Research on Intravascular Optical Coherence Tomography Image Analysis

March 2021

Graduate School of Systems Engineering
Wakayama University

Xinbo Ren

光干渉断層法による血管内画像の解析に関する研究

令和３年３月

和歌山大学大学院システム工学研究科

任 鑫博

Abstract

Coronary artery disease (CAD) as a well-known heart disease causes numerous deaths in countries, and the death number continuously increases every year. Lipid, fibrous and calcified plaques are the three common types of atherosclerosis plaque to be investigated the symptom reason by specialists through a new imaging modality, called intravascular optical coherence tomography (IVOCT). However, hundreds of IVOCT images would be produced for each treatment time to a patient. The biggest difficulty for CAD specialists is that they need a large time to manually analyze these IVOCT images. Therefore, automatic methods of IVOCT image analysis for quantitative measurement and lesion tissue classification are necessary and significant.

To overcome the problems, my research focuses on the automated methods on IVOCT image analysis. Chapter 1 introduces the CAD concept and the schema of the optical coherence tomography (OCT) system. Moreover, the characteristics of atherosclerosis plaques are presented and the relative work for the IVOCT image investigation is mentioned.

Chapter 2 discusses the pre-processing methods to eliminate noise, and to remove the catheter imaging region with a circle detector by consideration of the circle shape of the catheter. To segment the vessel lumen border and the outer border of the region of interest (ROI) in the following steps, I segmented a sector region containing the guide wire and its black shadow instead of detecting the guide wire spot directly.

Chapter 3 analyzes the intensity fluctuation degree of the A-line profile between the vessel tissue and the stent, and designs features based on the standard deviation to discriminate between the vessel tissue and the stent. At last, utilizing an adaptive region growing method to remove the guide wire spot.

Thinking that a 1-D A-line contains less information compared with a 2-D region composed of A-lines, and considering the light attenuation of lesion tissues, a local multi-layer model is built up to investigate the plaques feature extraction of the superficial layer in chapter 4. An A-line sub-region (ALSR) is defined as the basic unit of the local multi-layer to analyze features in a 2-D region.

In chapter 5, utilizing the outperformance of the deep learning method on image processing, I structured a VGG-like model, which contains 11 layers, to predict the plaque type of every ALSR. Three types of input with different channel amounts are designed for the recognition effect discussion on the texture information and multi-channel.

Chapter 6 researches the semantic segmentation of IVOCT images with the deep learning concept. The outer border of ROI is obtained with the level-set method to segment the superficial region containing useful information. The input data is cropped based on the ROI to feed into the built neural network (DB-SegNet) to solve the pixel-wise classification task. The classified prediction outputs are finally combined to compute the maximum possible category of the pixel.

Chapter 7 gives out a conclusion for each part of this thesis, and discuss the existing problems and future work.

概　要

冠動脈疾患（CAD）は有名な心臓病として各国で多くの死亡者を出しており、その数は年々増加の一途をたどっています。動脈硬化性プラークは、脂質性プラーク、繊維性プラーク、石灰化プラークの３つが一般的であり、血管内光コヒーレンス・トモグラフィー（IVOCT）と呼ばれる新しい画像モダリティを用いて、専門医が症状の原因を調査する必要があります。しかし、患者への治療時間ごとに数百枚のIVOCT画像が作成されることになる。ＣＡＤ専門医にとって最大の難点は、これらのＩＶＯＣＴ画像を手動で解析するために多くの時間を必要とすることである。そのため、定量的な計測や病変組織分類のためのIVOCT画像の自動解析手法が必要であり、意義がある。

これらの問題点を克服するために、私はIVOCT画像解析の自動化手法に焦点を当てて研究を行っている。第１章では、CADの概念と光コヒーレンス・トモグラフィー（OCT）システムのスキーマを紹介する。また、動脈硬化性プラークの特徴を紹介し、IVOCT画像解析のための相対的な作業についても言及している。

第２章では、ノイズを除去するための前処理方法と、カテーテルの円の形状を考慮して円検出器でカテーテルの撮像領域を除去する方法について述べる。以下のステップで血管内腔境界と関心領域（ROI）の外縁をセグメント化するために、ガイドワイヤスポットを直接検出するのではなく、ガイドワイヤとその黒い影を含むセクタ領域をセグメント化した。

第3章では、血管組織とステントの間のA線プロファイルの強度変動の程度を解析し、血管組織とステントを識別するための標準偏差に基づいた特徴量を設計する。最後に、適応領域成長法を活用してガイドワイヤースポットを除去する。

1次元A線はA線で構成される2次元領域に比べて情報量が少ないと考え、病変組織の光減衰を考慮して、第4章では表層のプラーク特徴抽出を検討するために局所多層モデルを構築した。2次元領域の特徴を解析するための局所多層モデルの基本単位として、A線サブ領域（ALSR）を定義した。

第5章では、ディープラーニング法の画像処理上の優れた性能を利用して、11層からなるVGGライクなモデルを構築し、ALSRごとのプラークタイプを予測した。テクスチャ情報とマルチチャンネルの認識効果を議論するために、チャンネル量の異なる3種類の入力を設計した。

第6章では、ディープラーニングの概念を用いてIVOCT画像の意味的セグメンテーションを研究している。レベルセット法を用いてROIの外縁を取得し，有用な情報を含む表層領域をセグメンテーションする．入力データはROIに基づいて切り取られ、構築されたニューラルネットワーク(DB-SegNet)に供給され、ピクセル単位の分類タスクを解決します。分類された予測出力は、最終的にピクセルの最大可能なカテゴリを計算するために結合されます。

第７章では、本論文の各部の結論を述べ、これまでの問題点と今後の課題について述べる。

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. I have duly acknowledged all the sources of information which have been used in the thesis.

Xinbo Ren
February 2021

Acknowledgements

Foremost, I would like to express my sincere gratitude to my supervisor Professor Toshiyuki Iami for the continuous support of my Ph.D. study and research. Simultaneously, I would like to thank Professor Haiyuan Wu and Associate Professor Qian Chen, they give me much more support to my investigation. Especially, Professor Haiyuan Wu supplies the important opportunity for me to improve my study ability and education level. Professor Chen also gives me a lot guidance on how to write a better thesis with English. I would like to thank Professor Takashi Kubo and Professor Takashi Akasaka with their kind help on the understanding of IVOCT images, the definition of vessel lesion plaques and review of my research papers. Also, thanks to the Department of Cardiovascular Medicine of Wakayama Medical University for supplying the experiment data (IVOCT images from patients) to support me to begin and complete the related research.

Thanks to my colleagues in Visual Design Laboratory to help me a lot in daily life. My sincere thanks also go to Doctor Peng Li, Doctor Yankun Lang, Doctor Yiqiang Qi, Doctor Yi Tian and Doctor Haoqi Gao, for giving me much more advice and contribution not only on my study, but also including other help. Especially for Doctor Peng Li, he helps me to familiarise the school environment and encourage me in my study. Doctor Yiqiang Qi goes with me to the hospital when my eyes feel uncomfortable. Thank Professor Yoshihiro Takuya invite me to his house for a nice summer holiday.

I would also like to express my gratitude to Mrs. Chunchun Yang from the international engagement office of Wakayama University, she helps me apply for the residency status. Sincere thanks to Mrs. Miyamoto, Mr. Takahashi and Professor Nagatomo bun-ko, they help me to improve my Japanese level in my spare time.

Many thankfulness to my fellow Mrs. Yan Li for help me to finish my class tasks when I study for a doctor's degree in Japan. Moreover, great gratefulness to my students Qing Tian, Jing Wang, ZhiJuan Li and Yidan Zhang, they finish the label work in my research.

Finally, I will give my deepest love to my family, my parents, my wife and my cute children for their support and encouragement for me as usual to finish my three years of study. Especially to my wife, without her understanding and help, I can't complete my research without any difficulty.

Publications

Journal papers

1. **Xinbo Ren**, Haiyuan Wu, Qian Chen, Takashi Kubo, Takashi Akasaka, *A Local Multi-Layer Model for Tissue Classification of in-vivo Atherosclerotic Plaques in Intravascular Optical Coherence Tomography*, IEICE Transactions on Information and Systems, Vol.E102-D, No.11, pp. 2238-2248, Nov. 2019.

2. **Xinbo Ren**, Jingbo Fan, Yi Tian, *Review of Vascular Stents and Atherrosclerotic Plaques Based on IVOCT Images*, Computer System Application, Institute of Software, Chinese Academy of Sciences, Vol.28, No.03, pp. 10-17, Mar. 2019.

3. **Xinbo Ren**, Yan Li, *Removal Method of OCT Image Catheter Based on Detection of Circumferential Intensity*, Computing Technology and Automation, Hunan University, Hunan Computer Society and Hunan Automation Society, Vol.28, No.03, pp. 107-111, Oct. 2019.

4. **Xinbo Ren**, *Automatic Segmentation of Nerve Dendrite Spines Based on Fluorescent Spot Detection*, Journal of Xi'an University of Post and Telecommunications, Vol.19, No.04, pp. 53-56, Aug. 2014.

International conference

1. **Xinbo Ren**, Haiyuan Wu, Qian Chen, Takashi Kubo, Takashi Akasaka, *A Tissue Classification Method of IVOCT Images Using Rectangle Region Cropped along the Circumferential Direction based on Deep Learning*, International Forum on Medical Imaging in Asia (IFMIA) 2019, Proceedings Volume 11050

2. **Xinbo Ren**, Qian Chen, Haiyuan Wu, Takashi Kubo, Takashi Akasaka, *A Novel Automatic Stent Detection Method in Intravascular OCT Images Based on Local Maximum*

of Standard Deviation, 24th International Workshop on Frontiers of Computer Vision (IW-FCV), 2018

3. Z. Sui, J. Li, J. Fan, **Xinbo Ren**, Y. Li and Y. Liu. *All-in-focus image reconstruction with depth sensing*, BioMedical Engineering and Informatics (CISP-BMEI), 9th International Congress on Image and Signal Processing, pp.178-183, 2016.

Table of contents

List of figures **xiii**

List of tables **xvii**

1 Introduction **1**

1.1 Cardiovascular disease 1

1.1.1 Cardiovascular disease 1

1.1.2 Healthy vessel and CAD plaques 3

1.2 Two main imaging modalities of vascular 5

1.2.1 Intravascular ultrasound 5

1.2.2 Optical coherence tomography 6

1.3 Atherosclerosis plaque types 9

1.4 Related work 11

1.5 Purpose of this thesis 19

1.6 Overview of this thesis 20

2 IVOCT pre-processing **23**

2.1 Image acquisition 23

2.2 IVOCT image conversion 24

2.3 Image noise eliminating 26

2.4 Catheter imaging removal 27

2.5 GW detection and black shadow segmentation 31

2.6 Experimental results 35

2.6.1 Parameter setting 35

2.6.2 Validation 38

2.6.3 Results 39

2.7 Discussion and conclusion 47

3 Lumen boundary segmentation and stents detection 49
3.1 Basic concept and model definition 49
3.2 Lumen boundary segmentation 54
3.3 Stents detection 55
  3.3.1 Stents candidate points detection 55
  3.3.2 Candidate points selection 58
  3.3.3 Stent area detection by adaptive region growing 60
3.4 Experimental results 61
  3.4.1 Parameter setting 61
  3.4.2 Results 61
3.5 Discussion and conclusion 62

4 Atherosclerosis plaque recognition 67
4.1 Feature analysis 67
4.2 Multi-layer model 70
4.3 Feature extraction 71
4.4 Experimental results 75
  4.4.1 Parameter setting 75
  4.4.2 Validation 75
  4.4.3 Results 76
4.5 Discussion and conclusion 76

5 Atherosclerosis plaque identification with deep learning 83
5.1 Deep learning basic concept 83
5.2 VGG-like model building 89
  5.2.1 VGG network architecture 89
  5.2.2 Input data 90
  5.2.3 VGG-like model 91
5.3 Experimental results 92
  5.3.1 Parameter setting 92
  5.3.2 Validation 93
  5.3.3 Results 94
5.4 Discussion and conclusion 94

6 Semantic segmentation for atherosclerosis plaques 99
6.1 ROI segmentation 100
6.2 Semantic segmentation architecture 102

6.3 Post-processing 103
6.4 Experimental results 104
6.4.1 Data acquisition and labeling 104
6.4.2 ROI segmentation algorithm parameters 104
6.4.3 Deep learning model and data augmentation 105
6.4.4 Validation 106
6.4.5 Results 107
6.5 Discussion and conclusion 109

7 Discussion and future work 121
7.1 Conclusion 121
7.2 Problems and future work 123

References 125

Appendix A Algorithms 133
A.1 Adaptive region growing algorithm 133

Appendix B Deep learning architectures 135
B.1 VGG-like architecture 135
B.2 DB-SegNet architecture 136

List of figures

1.1 Coronary artery disease 2
1.2 Artery with three layers structure 3
1.3 Atherosclerosis specimen 4
1.4 IVUS imaging modality 5
1.5 A-line illustration 6
1.6 FD-OCT Schematic illustration 7
1.7 Two imaging modalities to an artery 8
1.8 The components of an IVOCT image 9
1.9 Three samples of atherosclerosis plaques 11
1.10 Plaques appearance diversity 12
1.11 Other types of lesion plaque components 12

2.1 IVOCT image convert 25
2.2 Examples of different catheter shape types 28
2.3 Examples of circle detector in the catheter imaging 29
2.4 Illustration of catheter model 31
2.5 Demonstration of Wang's method[92] 32
2.6 GW region demonstration 33
2.7 CRDM description and application in IVOCT images 34
2.8 Illustration of catheter model 36
2.9 An example of error result for catheter removing 37
2.10 Examples of initial results of the catheter imaging removing 37
2.11 Examples of middle results with our GW segmentation 38
2.12 Linear regression for catheter area and GW shadow angle 40
2.13 Bland-Altman graphics for catheter area evaluation 41
2.14 Bland-Altman graphics for GW shadow angle evaluation 42
2.15 Examples of the final catheter imaging area detection results-1 43
2.16 Examples of the final catheter imaging area detection results-2 44

2.17 Examples of the GW shadow sector area segmentation results-1 45
2.18 Examples of the GW shadow sector area segmentation results-2 46

3.1 The complexity condition of the vascular 50
3.2 Coordinate transformation 51
3.3 A-line profile intensity 52
3.4 Example of smoothed A-line profile 53
3.5 Lumen border detection 54
3.6 Illustration of lumen boundary smoothness 56
3.7 Stents in the IVOCT image 57
3.8 Initial detected candidate points 58
3.9 Intensity analysis of the peak intensity and the region behind 59
3.10 Illustration of the stent detection 62
3.11 Results of lumen boundary detection 63
3.12 Results of stents detection 64
3.13 Illustration of local individual points 65
3.14 GW segmentation method-2 66

4.1 Illustration of the intensity distribution in the IVOCT image 69
4.2 Multi-layer model illustration 71
4.3 Results for the vessel tissue classification 78
4.4 Confusion matrix of lesion tissue 79
4.5 Misclassification examples 79

5.1 2D convolution illustration 85
5.2 Input images with different channel number 86
5.3 N layers of feature maps 88
5.4 Pooling operation for the downsampling 88
5.5 CNN architecture demonstration 89
5.6 Illustration of VGG-like model 93
5.7 Training presentation on our mode 95
5.8 Results of the tissue classification with our deep learning method 96

6.1 Illustration of ROI segmentation implemented by applying the level-set method 101
6.2 Cropped ROI illustration 106
6.3 Examples of ROI segmentation 108
6.4 ROI detection 110
6.5 An example of a dense block structure 111

6.6 Schematic of DB-SegNet architecture 112
6.7 3-D volume of cropped predictions 114
6.8 Semantic segmentation results of vessel lesion plaques 115
6.9 Illustrations of tow ROI segmentation methods 116
6.10 Demonstration of two lesion plaques classification methods 117
6.11 Prediction post-processing 118

List of tables

1.1 General characteristics of the main atherosclerosis plaque 10

2.1 Evaluation metrics for the catheter segmentation and GW region angle detection 39

4.1 Appearance of tissue thickness measurement in the previous studies 70
4.2 Lesion classification accuracy comparison between Ughi' method and our method 77
4.3 Evaluation metrics for the vessel tissue classification 81

5.1 Evaluation results of 4 classes in 3 different input channel types 96
5.2 Results of accuracy for 292 testing data in 3 different input channel types . 97

6.1 Evaluation metrics for the segmentation of ROI in IVOCT images. 107
6.2 Comparison of tissue classification between SegNet and DB-SegNet-3 (DB-SegNet-3 model presents the number of layers of a dense block is set 3.) . . 109
6.3 Evaluation metrics over 10 datasets for the vessel tissue classification . . . 113

Chapter 1

Introduction

1.1 Cardiovascular disease

1.1.1 Cardiovascular disease

Cardiovascular disease (CVD) is a general term for conditions affecting the function of heart or blood vessels. It remains the most frequent cause of death to people in the world, accounting for 17.3 million deaths per year[58], which makes a 31% proportion of the total, and the number of death is still continuously increasing. Generally, CVD includes the following common types: coronary artery diseases (CAD) (such as angina and heart attack), heart failure, stroke, aortic disease, and peripheral arterial disease[1, 2]. The reasons that contribute to the creation of CVD are various, including high blood pressure, smoking, high cholesterol, diabetes, lack of exercise, being overweight or obese, family history of CVD, and others, which all these causing as the "risk factors" would increase a person's risk of getting CVD.

Among the classes of CVDs, CAD as the most common type happens in the human blood tubes (in the body)[3] and is induced with the phenomenon that build-up plaques occasioning on the human vessel inner walls block the blood flow. This built-up plaque is medically called atherosclerosis, and it impacts the arteries become narrowed or hardened, which decreases the vessel lumen area and the flowing volume supplying blood to the heart. Simultaneously, the amount of oxygen contained in the human blood also be affected presenting cutting down correspondingly. In order to maintain the normal life activities of various tissues and organs of the human body, the heart needs to beat constantly to ensure blood transportation. Medically speaking, the heart, as a muscle-powered organ that pumps blood, also needs sufficient nutrients and energy itself which are supplied by the blood itself. Strikingly, coronary artery disease occurs when part of the smooth, elastic lining inside a coronary artery (the arteries that supply blood to the heart muscle) develops atherosclerosis[3]. When suffering atherosclerosis, the artery's lining becomes hardened, stiffened, and accumulates deposits of calcium, fatty lipids, and abnormal inflammatory cells – to form plaques around the human vessel wall[3]. As shown in Fig. 1.1, lesion plaques of atherosclerosis are produced in the inner vessel wall and the yellow substance depicts the shape, position, cover area of the lesion plaque in the blood vessel. The black arrow in the enlarged image section indicates the flowing direction of the blood. Three examples of cross-sections marked in different locations present 3 inner section situations of the coronary artery, including one healthy and other two unhealthy cross-sections containing plaques with diverse degrees. Observing in Fig. 1.1, clearly finding that the existence of atherosclerosis undoubtedly results in a condition of area decrease for the vessel cross-section, which in turns reduces the amount of blood flowing. If in the absence of any immediate and effective treatment and medical advice, the plaques will continue to increase in the size of the volume and finally block the blood flowing to the heart causing human death. Early coronary artery atherosclerosis has no obvious signs, neither significantly obstructing blood flow nor inducing acute symptoms, even in many cases plaque coronary artery plaque rupture will not have a significant impact on the normal physiological activities of people, so it is difficult to be aware of and found. However, as the patient ages and the size of the atherosclerotic plaque increases, the blood flow from the arteries to the heart decreases, resulting in symptoms such as coronary heart disease and angina, and in severe cases, the patient may face the threat of death. CAD can lead to various symptoms, such as chest pain or a heart attack, etc, especially, the heart attack could possibly cause sudden death to patients in a future time.

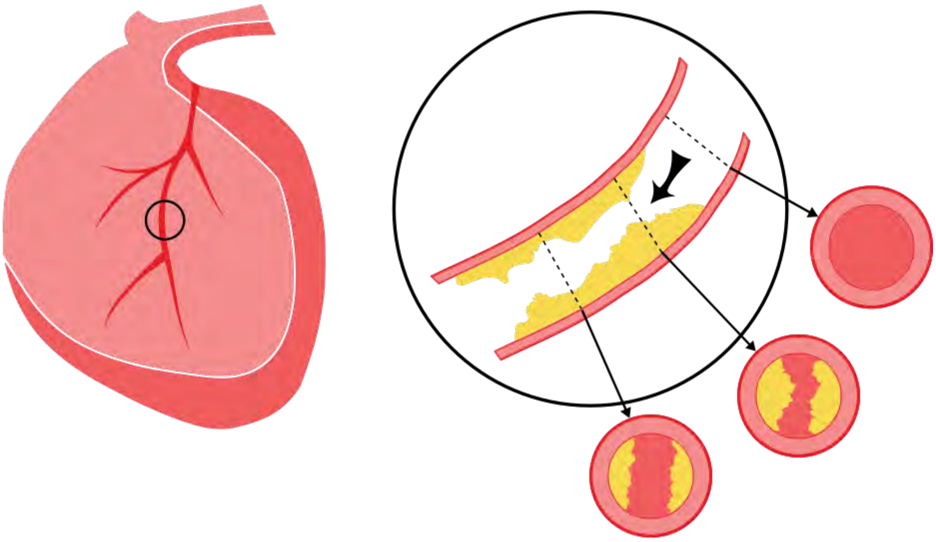


Fig. 1.1 Illustration depicting the human heart and the plaques of CAD in the blood vessel. Three cross-sections of the human artery with different positions is represented.

1.1.2 Healthy vessel and CAD plaques

A healthy artery wall is a three-layer morphology containing intima, media and adventitia respectively to construct a multi-layer vessel structure. The illustration of the blood vessel three layers of the structure is displayed in Fig. 1.2. The intima layer is near the vessel lumen and adventitia is the external layer of the artery wall, while media is the layer between intima and adventitia. The normal blood vessel can supply necessary oxygen and nutrients to the heart maintaining life. Atherosclerosis plaques regularly happen between the intima and media[72] to break the multi-layer structure and lead the complete blockage of blood flowing through the artery. Previous medical studies of CAD indicate that coronary atherosclerosis plaques are mainly divided into three categories: *lipid*, *calcified* and *fibrous* plaques. From clinical studies, these three types differ significantly in appearance and characteristics, histopathologically. Lipid plaque is possibly formed during the period of cholesterol oxidization in the vessel under the condition of insufficient high-density lipoprotein[4]. The fibrous plaque is also localized under the intima, within the wall of the artery resulting in thickening and expansion of the wall and, sometimes, spotty localized narrowing of the lumen with some atrophy of the muscular layer. The fibrous plaque contains collagen fibers (eosinophilic), precipitates of calcium (hematoxylinophilic) and, rarely, lipid-laden cells[4]. Calcification forms among vascular smooth muscle cells of the surrounding muscular layer, specifically in the muscle cells adjacent to atheromas and on the surface of atheroma plaques and tissue[59]. The lipid core plaque or fibroatheroma is a type of atherosclerotic lesion prone to develop unstable features under the influence of inflammatory processes and mechanical forces[28]. Besides, the form of mixed plaques composed of these three types is often observed during the treatment period, which contains the fibro-lipid

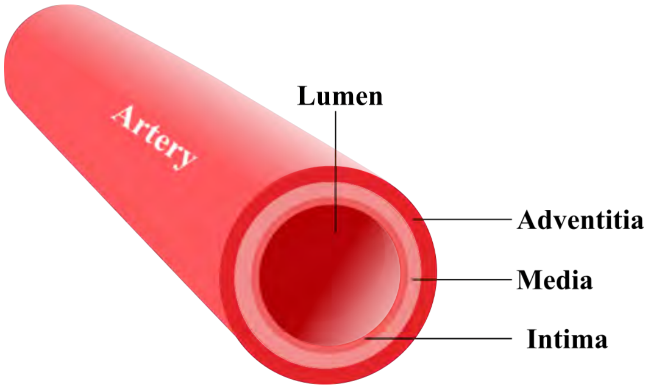


Fig. 1.2 A section example of the healthy human vessel. The artery wall is made up of three layers: intima, media and adventitia.

plaque, fibro-calcific plaque and thin cap fibroatheroma, etc. Figure 1.3[1] presents a real severe atherosclerosis specimen of a patient. Observing from Fig. 1.3, the luminal surface is studded with the proliferative lesion tissue recognized as the symptom of atherosclerosis.

A straightforward way to investigate the causes of the pathogenesis of CVD is to study the morphological changes and pathological manifestations of vascular lesions in the tissues firstly through the quantitative measurement of the appearance, lumen area, lumen thickness, etc. Human blood vessels as one of the exquisite organs, although the appearance of its size is dynamical (the large artery is $> 10\,mm$ diameter and the smaller ones are 0.1–$10\,mm$), it is still a micro-substance structure, which needs to be imaged with the professional technology of the coronary artery to assess the severity of coronary lesions.

Fig. 1.3 A specimen of the vessel with atherosclerosis tissues from one patient.

---

[1] This image is a work of the Centers for Disease Control and Prevention, part of the United States Department of Health and Human Services, taken or made as parhuman arteryt of an employee's official duties. As a work of the U.S. federal government, the image is in the public domain.

1.2 Two main imaging modalities of vascular

The technology for vascular imaging with high quantity can improve the understanding of the inner structure of the vessel and treat the vessel diseases helpfully to reduce the symptom of patients. In recent years, two significant technologies are generally performed as the main imaging modalities to the vascular to observe the realistic inside morphology of the vessel. One is called intravascular ultrasound (IVUS) and the other is intravascular optical coronary tomography (IVOCT).

1.2.1 Intravascular ultrasound

IVUS is a medical imaging methodology through the ultrasound equipment with an ultrasound probe placed in front of the catheter to receive the signal of the blood vessel and then produces the cross-section imaging by computer aid tools. Mechanically rotating the IVUS catheter around could gain an ultrasound signal reflected from the vessel inner wall. To a 20-40 MHz of IVUS, generally, the axial resolution is 80-100 μm with a pullback speed of $0.5\,mm/s$ and the lateral value is a range of 200 to 250 μm[19]. Figure 1.4 presents one transverse image of the human vessel. The left sub-figure is the original IVUS image and the right is the corresponding labelled figure for the denotation of an interesting area. The region with green indicates the area and thickness of atherosclerosis plaques of the *in-vivo* and the severity degree of this disease. IVUS can be applied to optimize the result of the angiographic assessment to give out a more definitive conclusion of the coronary lesion. Its

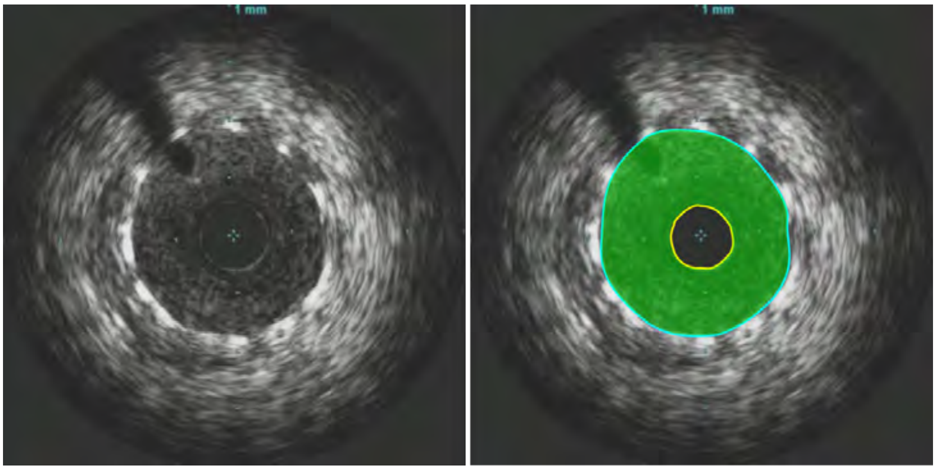


Fig. 1.4 An example of IVUS imaging, which shows the detailed information of a vessel through ultrasound. The interesting region labelled with green is an imaging area without more detailed information that can be observed (comparing with OCT).

clinical applications are to size the vessel and guide the implantation of coronary stents in percutaneous coronary intervention (PCI) procedures.

1.2.2 Optical coherence tomography

Optical coherence tomography (OCT) as a novel intravascular imaging technique, which is a light-based imaging modality, generates high-resolution cross-section images of tissue microstructure. The light of OCT is near-infrared based on optical 1-dimensional low-coherence reflectometry, which employs a Michelson interferometer and a broadband light source[13, 34, 43]. The principle of OCT is simply explained as that the OCT equipment measures the time delay of the light which is reflected or backscattered from vessel tissues, and which is collected by the catheter, by utilizing a technique known as interferometry.

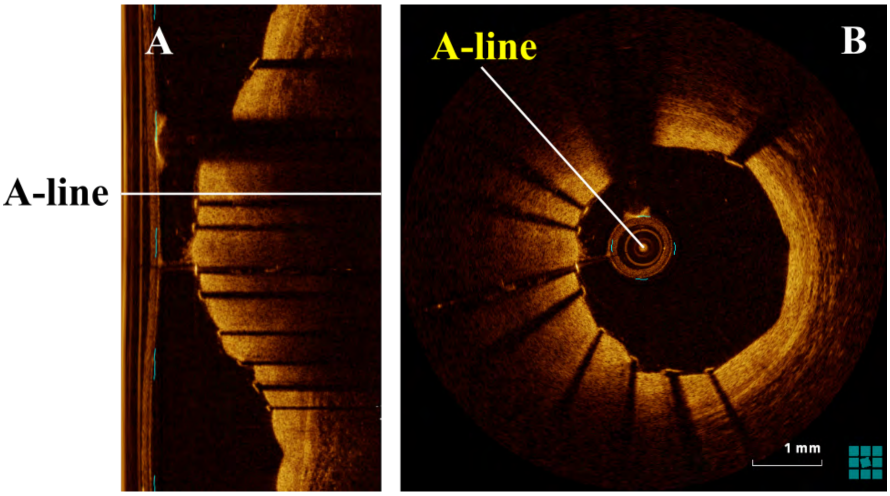


Fig. 1.5 The reflection light received by the catheter is stacked into the storage device to structure an IVOCT image slice. A (the polar domain) and B (Cartesian domain) are the two types of IVOCT image representation. The white lines in both example images denote A-lines generated with the OCT system.

Currently, two types of OCT systems, named respectively as time-domain OCT (TD-OCT) and frequency-domain OCT (FD-OCT), are applied for the CAD investigation and vessel disease treatment, which are mainly comprised of a low coherence and broad bandwidth light source, reference mirror, and photodetector. Compared with TD-OCT, FD-OCT acquires images significantly faster by using a fixed mirror with a variable frequency light source, which is more popular to be utilized in medical institutions[43]. Figure 1.6 shows a schematic representation of FD-OCT system and the relative significant components. Light is transmitted from the light source to a beamsplitter and then split into two portions: one goes to scan the tissue target of the patient through a catheter (*sample arm*) and the other

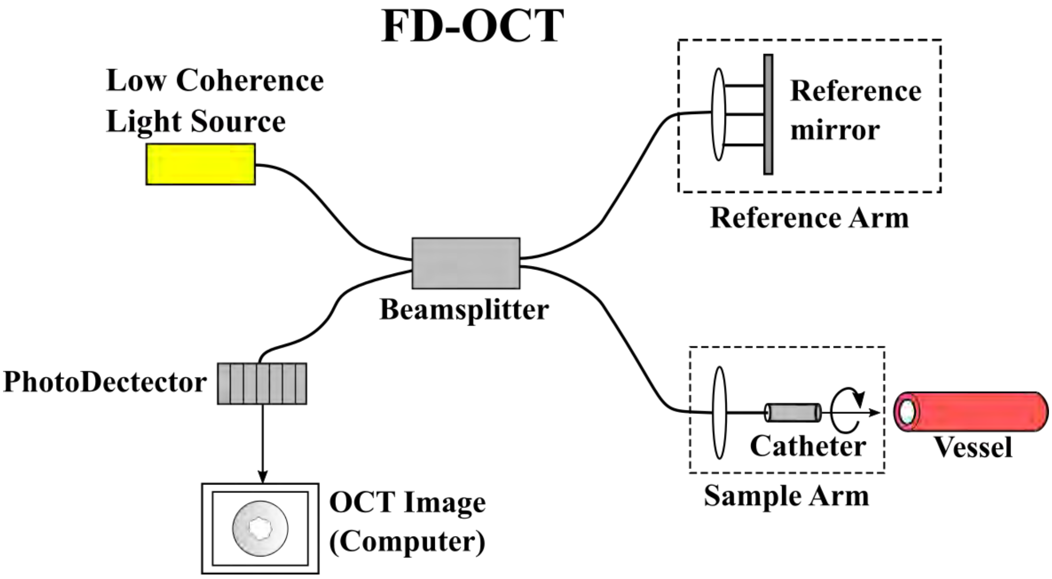


Fig. 1.6 Schematic of FD-OCT system.

portion is transmitted to the reference mirror (*reference arm*). Subsequently, the reflected light from the reference arm and the collected light from the sample arm are merged as a single one to be sent to the photodetector. Because of the distances that the collected light and the reflected light both travel are the same, therefore a pattern of high and low intensities can be detected, which is known as the interference method[78]. And then, the IVOCT image is determined through the interference pattern analysis considering the amount of backscattering as a function of delay time or depth within the tissues. Finally, a cross-section of an IVOCT image sample is produced through the OCT relevant software in the computer console. After the entirety "pull-back" procedure completing, a set of slices about the human vessel is collected to form the real situation of the treated vessel. The procedure of the OCT "pull-back" in the *in-vivo* vessel is that the intravascular OCT equipment uses the catheter to emit light to the artery inner wall and then records the reflection while simultaneously rotating the catheter through the guide-wire (GW) during a pull-back period. After collecting the backscattering of light from the artery inner wall through the catheter, the IVOCT equipment stacks the light called *A-line* (axial line) together to form a cross-section image as a slice of the original IVOCT image. Figure1.5 illustrates the A-line that can be conceptually thought of as illuminating tissue with pulses of light. This procedure is performed repeatedly by the OCT equipment until an entirety "pull-back" is completed. Although the light emitted by the catheter contains a center wavelength range from 1.25 μm to 1.35 μm, the value of 1.3 μm is usually chosen by medical institutions for the research and treatment[13].

The quality of IVOCT images depends on the spatial resolution which includes axial and lateral directions corresponding to parallel and perpendicular to the light beam respectively. Compared with IVUS, IVOCT resolution presents a range from 10 to 20 μm, which is almost 10 times of IVUS, consequently, it can obtain the description of the tissue structure with high resolution. Besides, under the wavelength of 1.3 μm, the tissue penetration of IVOCT is 1 to 3 *mm* while IVUS can achieve 4 to 8 *mm*[13], which IVOCT is lower than IUVS in the penetration capability of imaging. That means we observe a valid region with OCT technology in a smaller limited area than IVUS. Although IVUS provide a deeper through vessel tissues than IVOCT, IVOCT presents more detail for characterizing the superficial structure of the vessel wall and is considered as an important imaging technique applied to medical fields. Therefore, the main disadvantage of OCT is the poor penetration depth to tissues and makes the tissue visualization with a limitation of 2 to 3 *mm* in depth for a global scanning area of 7 *mm* in diameter. An example of two types of vascular imaging modalities to a cross-section of one vascular is given in Fig. 1.7. Left figure (Fig. 1.7 (A)) shows that IVUS provides information in anatomical severity of lesions, while right figure (Fig. 1.7 (B)) displays the same position imaging through the IVOCT technique. An obvious difference in appearance between these two representations is easily observed that utilizing OCT technique can reflect. For example, the position and the region of neointimal hyperplasia (NIH) and the intervention stents can patently be displayed in the IVOCT image (Fig. 1.7 (B)), but under the condition of IVUS observing the corresponding tissue and stent in the same location is not an easy work for observer if without any specialist knowledge.

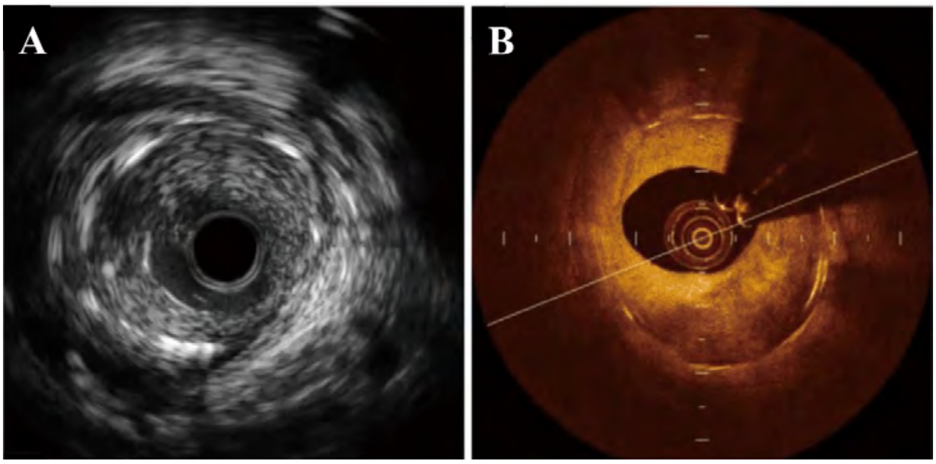


Fig. 1.7 A cross-section of an artery captured with two modalities: IVUS and IVOCT. (B) presents more useful lesion tissue information to represent the detailed inner structure of the vessel than (A).

With the high resolution, IVOCT presents the amount of content of the vessel inner structure in detail more than IVUS, although the reality is that the useful region in the IVOCT image is only the superficial area and the rest of the part of the IVOCT image supplies less beneficial information. Specialists in medical institutions apply the IVOCT for the measurements of the lumen, intimal elastic membrane (IEM), stent and the plaque characterization, including the angle, depth, thickness, area of lesion plaques, both in the step of pre-PCI lesion assessment and post-PCI lesion assessment[43, 78]. According to the representation of the vessel with OCT technique, a regular IVOCT image is comprised of an artery wall, a vessel lumen, a catheter imaging and a bright metal reflection of a guide-wire (GW) immediately followed by a black shadow region (Fig. 1.8), often also including residual blood artifacts inside the vessel lumen. As Sec. 1.1.1 mentioned, with the impact of the "risk factors", the atherosclerosis plaque, which is a mass lesion or focal thickening, are usually generated in different layer architectures (intima, media and adventitia) resulting in loss of a layered structure of the vessel wall.

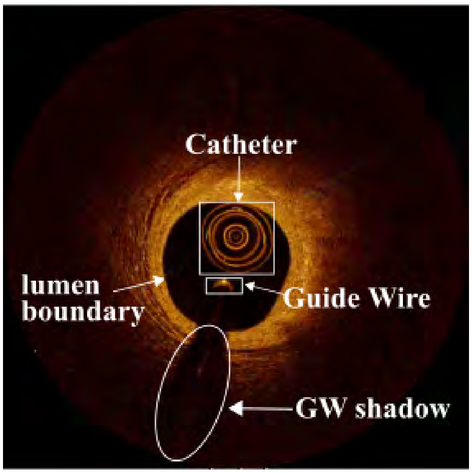


Fig. 1.8 The main components of an IVOCT image are comprised of a catheter (white big rectangle), a Guide-wire (white small rectangle), a GW sector shadow (white ellipse) and the vessel lumen with a black background.

1.3 Atherosclerosis plaque types

As atherosclerosis plaques explained in medical research, there are 3 main types of atherosclerosis plaque which can be discriminated by the OCT: *fibrous*, *calcific* and *lipid plaques*[94]. Previous studies [13, 28, 43, 93] described the features of atherosclerosis plaque characterizing through histological correlation. Table 1.1 summarizes the IVOCT features of the main types of atherosclerosis plaque in accordance with [13, 18, 93]. Fibrous plaques have a low attenuation coefficient and are a bright tissue characterized as a homogeneous signal-rich region. Lipid plaques associate with a high attenuation coefficient and appear in heterogeneous signal-poor regions with diffuse and irregular borders, which is usually depicted as a phenomenon of a superficial high signal region followed by a very low signal region. The calcified plaque is signal-poor regions but with sharp borders and presents a low attenuation coefficient. The above description of plaque characterizations can be observed in Fig. 1.9, respectively. Figure 1.9 (A) is the example of a fibrous plaque that presents a bright lesion tissue, and the thickness of the intima and media is larger than the condition in the normal vessel. In Fig. 1.9 (B), the contour of the vessel wall changes and the layer-structure of the normal vessel disappears obviously, simultaneously a low signal region appears (illustrated with white arrow). Different from Fig. 1.9 (A) and Fig. 1.9 (B), the lesion plaque in Fig. 1.9 (C) displays signal-poor regions with sharp texture borders, which is in accordance with the characterizations description of calcified plaque. As a matter of fact, it is known that the shape and representation of the plaques are diversity in representation. The representation for each major type of lesion plaques is variable and somewhat dissimilar during different periods for different patients, even with the condition of the same patient. As shown in Fig. 1.10, each row in the figure is the examples of the fibrous, lipid and calcified plaques with the representation of tissue diversity. Furthermore, other existing classes of plaque components are also observed in the IVOCT images with intricate morphology (or mixed formation) to demonstrate the histological diversity[13, 78]:

Table 1.1 General characteristics of the main atherosclerosis plaque

Plaque type	Backscattering	Attenuation	Characterization
Fibrotic	moderate	low	Signal-rich, homogeneous
Lipid	moderate	high	Signal-poor, homogeneous, irregular and diffuse borders
calcium	low	low	Signal-poor, heterogeneous with low signal, sharp borders

- **Fibrous cap** is a cap with thickness < 65 μm relating to plaque rupture (Fig. 1.11 (A)).
    - Thin cap fibroatheroma (TCFA), a general formation that usually is observed appearance paired with the lipid plaque.

- **Fibrocalcific plaque** that a kind of lesion plaque composes of fibrous and calcified two portions, where presents a circumferential bright lengthy and narrow region followed by a signal-poor heterogeneous region with sharp borders (Fig. 1.11 (B)).
- **Mixed plaques** (or heterogeneous plaques) containing calcific deposit with delineated borders and lipid-like region with unclear borders (Fig. 1.11 (C)).
- **Thrombus** described as a thick mass including red and white ones protruding into the lumen (Fig. 1.11 (D)).

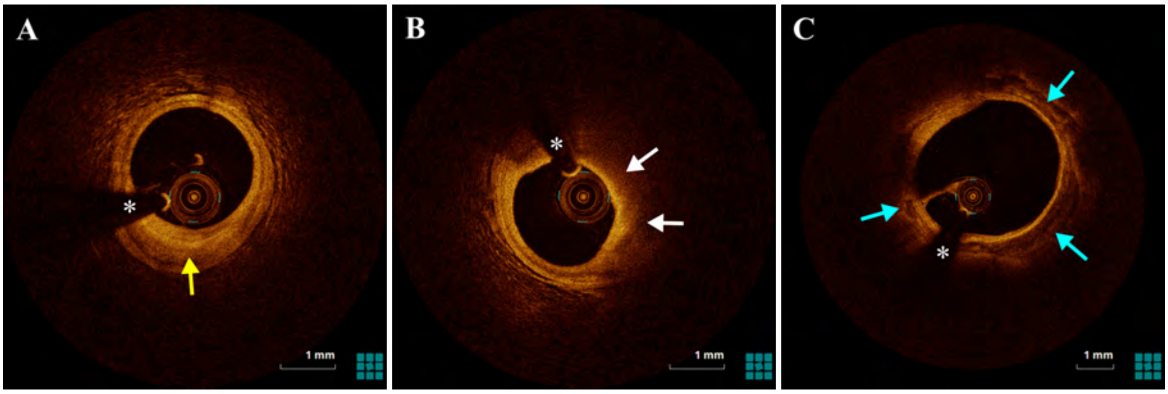


Fig. 1.9 Three IVOCT image examples of the atherosclerosis plaque corresponding to the 3 types, (A) fibrous plaque (yellow arrow), (B) lipid plaque (white arrow), (C) calcified plaque (blue arrow), respectively. * denotes GW artifact.

Overall, for different periods and different patients, the representation of lesion tissues is variance. As a consequence, the complex structure of the inner human vascular requires a suitable imaging modality on the high resolution to capture the detailed tissue information. IVOCT as the new imaging technique with high resolution gradually is applied in the clinical research and microstructure investigation instead of IVUS, although the penetration depth of IVOCT is less than IVUS that only superficial region characteristics are expressed. In brief, IVOCT is a powerful and significant imaging modality to be employed on the clinical research of human vessel diseases and the quantitative measurement of vessel features.

1.4 Related work

IVOCT provides a high-resolution imaging approach to capture the vessel inner structure for the purpose of CAD treatment with the disease diagnostic assessments, plaque recognition and characterization, PCI lesion assessment, guidance PCI, and eventually, improves the understanding of the vascular biology of atherothrombosis and the relevant clinical outcomes.

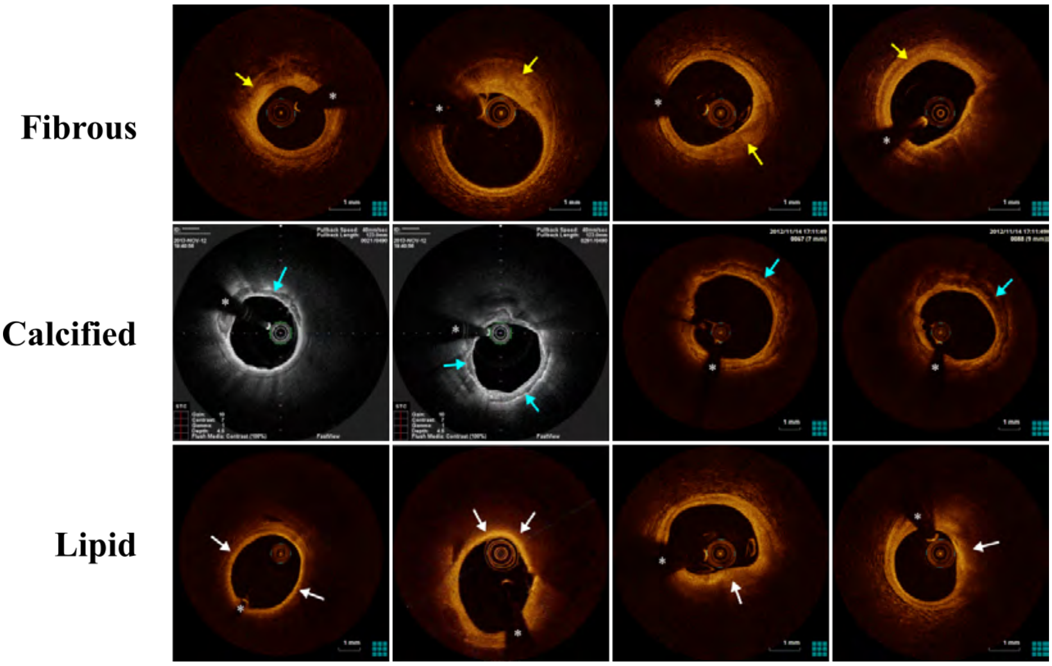


Fig. 1.10 Illustration of diversity appearance with different shape, area, angle, depth and thickness for 3 major types of atherosclerosis plaque (fibrous plaque:yellow arrow, lipid plaque:white arrow and calcified plaque:blue arrow). * denotes GW artifact.

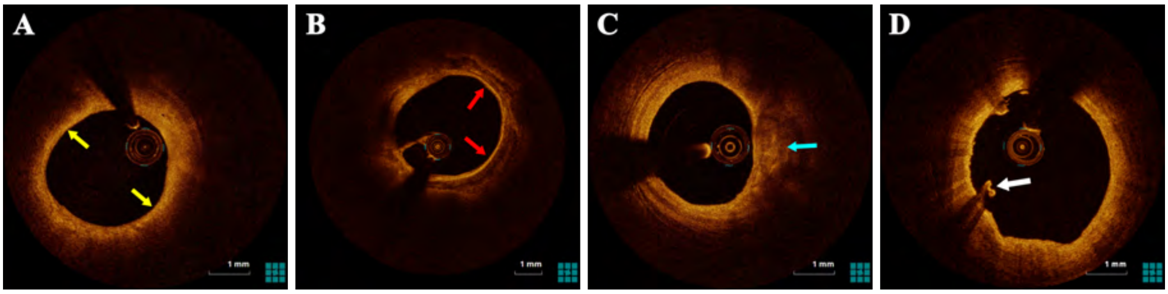


Fig. 1.11 Examples of other types of plaque components with different lesion morphology. (A) presents a fibrous cap region (yellow arrow) between the low signal region and the vessel lumen border. A heterogeneous region containing the fibrous plaque and calcified plaque is depicted as the fibrocalcific plaque in (B) indicated with red arrow. (C) is an example of mixed plaques combining calcific deposits with lipid-like region designated with blue arrow. (D) displays a lengthy and narrow protrusion tissue called thrombus (white arrow).

However, some challenges still exist to CAD specialists on the vessel diseases research and treatment by utilizing IVOCT as follows:

- Hundreds or thousands of image frames would be generated in an individual pullback scan for every treatment period of the patient, manual analysis becomes a heavy burden and time-consuming task to specialists, if without any reliable assist tool and automated image analysis methods
- Vessel lumen area measurement and morphology assessment based on lumen boundary segmentation with the condition of complexity and non-complexity
- Stent detection, contour evaluation and tissue coverage area measurement at every observation period after the stent implantation in IVOCT images
- Tissue characterization, plaque recognition and pixel-wise classification, the quantitative measurement of plaque in area, angle, thickness and depth
- Pre-processing of artifacts, including the elimination of the residual blood, as well as catheter and GW imaging removal, which truly impact the accuracy of the designed automatic methods

To overcome the above clinical problems and improve the effectiveness of CAD diagnostic and curing, previous studies have proposed various methodologies for semi-automatic or fully automatic *in-vivo* OCT images analysis and processing in artifacts removal, lumen segmentation, stent struts detection, lesion plaques identification and classification, and other relevant CAD clinical research tasks.

Artifacts Removal

As aforementioned in Sec. 1.2.2, a general IVOCT image without the stent implantation is a vessel inner structure imaging which is composed of the vessel wall, vessel lumen, the catheter imaging, a bright reflection of the GW and its dark shadow. Sometimes, the residual blood in the lumen is also captured and displayed in the IVOCT image if the blood in the checking segment of the vessel is not rinsed out completely before the catheter entrance. In the IVOCT image, the artifacts usually contains the catheter imaging, the metal reflection of GW and the residual blood. The elimination of the catheter, GW and the blood artifacts are the first facing challenges to researchers in their approaches when they complete the tasks of lumen boundary segmentation, stent detection, and both the above assessment examines. Too little work has been devoted to mainly developing methods for the artifacts elimination, most method utilized the prior-information to remove the catheter and GW reflection imaging.

For the catheter removal, several papers[22, 54, 91] utilized the known information of the catheter cross-section position or the maximum radius of the catheter rings to straightforwardly remove the catheter area. For example, [54] exploited the fact that the Dragonfly catheter diameter is $\sim 0.90\,mm$ to remove the catheter according to its center position. Observing that the position of the catheter is the same across all frames of IVOCT images, [22] computed the average intensity in the same position for catheter removal.

Ughi et al.[81] discovered that the catheter imaging is multiple bright concentric circular rings, and the internal structure of catheter always maintain a fixed appearance while the external plastic sheet with deformed expression. They transformed OCT images from Cartesian space to the polar space and discovered that the concentric circular rings become vertical lines in the polar image domain. A rapid algorithm based on the Hough transform[50] was used to define curves and transform them into vertical lines in the polar domain. Then, detecting the largest distance of the concentric ring as the outer border of the whole inner structure of the catheter. In other papers[82, 83], Ughi et al. converted the polar IVOCT image to a binary formation using the Otsu method[65] firstly, and then applied a morphological operation (closing) to eliminate small holes inside the binarized IVOCT images. Subsequently, applying an area constraint method to remove the individual pixel area because they considered catheter and GW imaging as unconnected regions containing pixels. If the area of these individual regions is smaller than a predefined threshold BWMA (black-white minimal area), it can be recognized as the catheter or GW.

Tsantis et al.[80] explained the reasons for the catheter distortion imaging that do not satisfy the circle parameterization due to the edge detection errors and the noisy pixels appearing nearby the catheter circle boundaries. They modelled two continuous concentric circles to limit the bright circles of catheter imaging. Then, through histogram of distances from the image center to the pixels inside the region defined by the two concentric circles, the pixels belong to the catheter would be detected.

The imaging GW presenting a bright metal reflection followed by a black shadow region impacts the lumen border segmentation and stent struts detection when it needs to assess the relationship between the vessel lumen morphology and CAD. Therefore, GW removing is also an important task for researchers. The diameter of GW in[54] is known as $0.3556\,mm^2$, and the bright region of GW imaging is verified as $0.0496\,mm^2$. Zhang et al.[92] converted each slice to a "accumulated intensity line" by adding all the pixels of each A-line to form as one intensity value. All the "accumulated intensity line" corresponding to the slices are compressed as one *en-face* image. Since the GW portion of the "accumulated intensity line" presents low energy, a long dark bar can be observed obviously and be segmented with applying dynamic programming twice to locate its contour. Clearly, the single position of GW can be obtained from the detected black bar contour. This GW segmentation method is also utilized in paper[16, 81]. Wang et al.[91] characterized GW as a gap in the bright superficial layer, then found out the brightest pixels along the A-line within the gap to detection GW position.

Although the above approaches overcome the catheter removal and GW segmentation on certain situations, the limitation still exists. Prior information only solves the same model OCT equipment by using the known catheter diameter and the GW size. Most examines did not mention the cases that the catheter location nearby the lumen boundary and the irregualr catheter with distorted shape. The GW segmentation method in [81, 92] needs to compute all the slices instead of analyzing the single OCT image, which can not be developed as a real-time assistant tool for specialists' clinical research. Other methods[60, 82, 83] employed morphological operations to segment GW pixel area with area constraint without considering the dynamic change. For the residual blood elimination, a general method is to apply the morphology operations (opening and closing, etc.) several times to remove the single pixel area[54, 81–83].

Lumen Boundary Segmentation

Morphologically and histologically, a healthy vascular lumen boundary is with a circle-like or ellipses-like shape presenting homogeneous attribute and smooth curve without any protuberance substance. Accordingly, the morphology of luminal boundary is normally used as the first step to judge the healthy condition of the vessel. Sihan et al.[74] firstly proposed a fully automatic lumen contour detection in OCT images. They employed the Canny filter[15] to detect the edges in the IVOCT images. However, due to the multi-layer structure of the vessel and the big difference between the OCT datasets, extra edge segments would be produced. Thus, unnecessary edges are removed by using the dot product between the gradient orientation and the catheter center, for the residual short lines, a threshold of line length is set for judgment.

Observing the IVOCT images, it is no doubt that the lumen border is a divided line between the dark lumen and the bright tissue. A significant gradient changing occurs along each A-line from the center of the IVOCT image to the vessel wall, which can be used to describe the intensity profile characterization of the A-line. According to the principle of light attenuation, there is a peak intensity existing in the A-line and soon occurring intensity falling phenomenon. Utilizing this attribute, Ughi et al.[81, 82] extracted four properties (peak intensity, shadow presence, length of a shadow and speed of the energy falling to a certain value) to distinguish the lumen border with other objects. The shallowest pixel $r_{Lsh}(\theta)$ with its intensity approximately equal to half of the maximum intensity is located. Then, a 2-dimensional cubic smoothing spline f[23] is employed to fit all the selected points to obtain the final lumen boundary. Similarly, Wang et al.[88, 92] also focused on this obvious attribute that an obvious intensity variance near the border of the intima closing to the lumen. They segmented the lumen boundary by searching the contour that maximizes the energy difference between the sum of gray values outside and inside the boundary[92]. Dynamic programming[10] method is selected to find the optimal solution to solve the path problem. After recursively computing all the possible paths that satisfy the condition, a contour with the maximum accumulated energy would be determined as the final luminal border.

Utilizing morphological operations to segment the lumen boundary is also a general method. Macedo et al.[53] used the Otsu method to separate the vessel wall with lumen area firstly. And then, applying significant gradient searching from the bottom to top of the in vivo OCT image and setting the value of the region below intima layer as zero, and making the region containing intima layer as well as lumen area as one. Subsequently, a subtraction was employed between the Otsu-processed results and the zero-one setting outcomes. The final segment results were gained after a sequence of five dilations and five erosions for eliminating holes and shadows. Besides, Macedo et al.[54] investigated the bifurcation of the lumen through defining 13 descriptors (such as distance centroid, circularity, bending energy) to produce 104 features. With the orthogonal least squares, feature selection operation was applied to search for the best features. After that, three state-of-the-art classifiers (support vector machine, random forest and adaboost) were implemented to classify the IVOCT images with bifurcation situations.

In [60], Moraes et al. utilized Discrete Wavelet Packet Frame to extract features and separate tissue information, which made an adequate data for the next step. Subsequently, the Otsu threshold was used to binarized the processed result for the lumen boundary segmentation in the polar domain. Gurmeric et al.[30] shot rays from the center point of the IVOCT image to each angle, subsequently, two Catmull-Rom splines were used to initialize the lumen boundary. At last, the desired boundaries were obtained via an edge-based active contour framework and the area of region of interesting (ROI). Tsantis et al.[80] denoted that the class probability of a pixel was depend on the membership of its neighbors. They combined the conditional and contextual information as the input of Markov random field (MRF) to determined the pixel class. The textural information in [80] was computed through continuous wavelet transform for each pixel. Roy et al.[72] built up a model that splitting the IVOCT image into two disjoint parts, I_{lumen} and I_{tunica}. Combining with the optical backscattering principle, its maximum was refined using a global gray-level statistic and was employed as the initial seeds of the random walks image segmentation to the lumen and tunica. Cao et al.[16] focused on the segmentation of the irregular lumen caused by the GW shadow, blood artifacts, bifurcation vessel. They proposed a divide-and conquer strategy to eliminate GW, then a gradient-base level set model utilizing edge information was established. To overcome the noise affection, paper[16] employs a Gaussian filter based on the kernel size of $N \times 1$ and $N \times N$ respectively to the bottom and top of the IVOCT image within the polar system.

Stent Detection

Stent implantation is an effective treatment for the patients to implement the coronary revascularization procedure. It can decrease the symptom caused by CAD and increase the life-time of patients through enlarging the area of the vessel lumen to let blood flow normally.

Wang et al.[89] synthesized an *en face* image that each line in this image was the "accumulated intensity line" by adding all the pixels along the A-line direction. A single case of the hundred IVOCT images can form the *en face*, where each line was derived from average intensity computation of the superficial pixels from the lumen border to a certain depth along the A-line direction. To detect stents in *en face* image which reveal the 3-dimension spatial information, Wang et al. utilized the minimum spanning tree to detect all the stent points. Similarly, in paper [90], to imporve the detection accuracy of stent and enhance utilizing the 3-dimension knowledge of stent structures, Wang et al. used a Bayesian network based on physical principles of OCT imaging to investigate the stent detection. They computed the probability of each A-line to roughly estimate the stent depth through 3-dimension information, subsequently, all struts' depth location in a pullback are obtained.

As the obvious features to the stent strut, shadows behind the small bright is an evident appearance to be utilized detecting the stent. Gurmeric et al.[30] analyzed the angular intensity energy distribution to find out the clues of dark shadows. They built up an energy map that transmitting rays from the image center to any angle to discover the trace of struts by investigation of falling and rising of energy on these rays. Strut position was determined through a second analysis over the detected shadow rays. Besides, NIH was discussed with the assessment of minimum NIH cases and mild to severe NIH cases. In paper [87], peak point detection, candidate pixel selection and shadow edge detection were investigated for the stent detection in NIH. Lu et al.[51, 52] introduced features of the candidate stent and the shadow region to detect strut locations. Totally, 17 intuitive characteristics were designed to depict the bright reflecting and shadow dark attributes. Thresholds and bagged decision trees were used to analyze the maximum possible stents in IVOCT images. As in [82], Ughi et al. described the stent with properties of high peak intensity, very fast rise and fall of energy and a significant drop in intensity based on the four attributes mentioned in Section of *Lumen Boundary Segmentation*. Other features, such as *mean*, *maximum* and *sum of values above mean*, were also utilized in the polar domain to discover the location of stents.

In [9, 24], clusters of malapposed and uncovered stent struts as a topic research were implemented. In [9], the mean and maximum malapposition distance within each cluster, the length of the cluster and the number of quadrants were introduced to characterize malapposed stent. Study [24] constructed a score map to reflect the overall apposition of a stent suing the interpolated distance between the stent and the lumen and supply as qualitative measurements of the stent position.

However, the bright reflection may also be caused by inner tissues of the vessel, and sometimes some stents do not show bright reflection. In these cases, it is hard to detect stents without considering the black shadow area behind them. Therefore, it is difficult to detect stents from the series of IVOCT images stably by using fixed threshold values. Moreover, the effect of luminal residual blood, image noise, guider-wire, and the catheter are also considerable reasons for the change of the intensity of the IVOCT image. Some reports analyzed a limited number and type of the cases in stent detection. Furthermore, few types of research investigated the detection of stents with neointima coverage.

Plaque Identification and Classification

Considering the IVOCT images formation that the catheter received a reflected signal of the vessel tissue and these sampled signals are constructed to an IVOCT image through the OCT equipment, A-line profile attributes are directly utilized for plaque detection and recognition analysis. Therefore, traditional methods based on machine learning principally focused on the feature extraction of A-lines. Rico-Jimenez et al.[71] modelled each A-line as a linear combination of N depth profiles $(\mathbf{p}_1, \dots, \mathbf{p}_N)$ and assessed the category of each A-line with a least-square optimization strategy. The divergence of optical attenuation among lesion plaques is regarded as a significant feature of A-line for the plaques recognition[66, 83, 86]. Athanasiou et al.[11] presented a method that extracting totally of 42 features for each pixel and then used a random forest classifier to classify four tissue types (calcium, lipid, fibrous and mixed tissues). Besides, with conventional approaches, segmentation and quantitative assessment of the fibrous cap and the border detection of the calcium plaque was investigated in [12, 18, 29, 88, 91, 92].

Other literature applied the CNN-based methods to recognize the images containing lesion tissues or as a feature extractor supplying features to the classifier. Kulluru et al.[41] used A-line as the fundamental unit and recognized its type with the deep learning method. Gessert et al.[25] discussed the recognition capability of two designed CNN models and applied their method both on Cartesian and polar images. Features extracted through pre-trained CNN were employed to three classifiers for the coronary layer identification by Abdolmanafi et al.[7]. Also, they evaluated three different pre-trained CNNs and obtained the final classification results through voting of the outcomes of these CNN models[6]. Meanwhile, other groups aimed to classify the detected lesion regions at pixel-level and proposed lesion tissue identification methodologies based on semantic segmentation technologies. With the CNN method, He et al.[32] determined the pixel category of the IVOCT image by analysis of a cropped tissue patch. Cheimariotis et al.[21] proposed two steps method to classify the plaque type by applying AlexNet. Utilizing the CNN as the feature extractor for each pixel could obtain a high accuracy of lesion plaques segmentation, but the fact is that an effective region of a single IVOCT image contains numerous pixels, which costs much more time to gain the recognition results. Fortunately, fully convoluational neural network (FCN) (e.g., U-Net and SegNet) were applied for the semantic segmentation of lesion tissues of IVOCT images to overcome the mentioned problem. Zhang et al.[96] compared two approaches (U-Net and SVM) for the identification of plaque types. Oliveira et al.[64] and Gharaibeh et al.[26] both utilized SegNet for the calcifications segmentation and assessment. Lee et al.[47] firstly compared the lesion tissue segmentation results through employing two deep learning models (SegNet and Deeplab V3+), respectively, and then assessed the clinically relevant measures based on the acquired results.

1.5 Purpose of this thesis

The research purpose of this thesis is comprised of the following parts:

1. Well understand the principle and pattern of the OCT imaging modality, the characterization of tissues, lumen wall, layer structure, stent structs, residual blood, atherosclerosis plaque of *in-vivo* OCT

2. Constructing a reliable pre-processing framework of IVOCT images: including speckle noise eliminating, catheter imaging region removal, GW and the section black shadow region detection, artifacts removing(residual blood, non-uniform rotational distortion, etc) algorithms, etc.

3. Researching a novel and rapid lumen segmentation method that processing the bifurcation and irregular morphology cases, which is employed in an individual IVOCT image with real-time when specialists face the treatment of CAD, and simultaneously as an approach to obtain the values of the qualitative measurements of the lumen

4. Developing a rapid stent detection method which not only focuses on the research of normal stent strut recognition but also including the condition of neointimal and malapposed

5. Investigation of feature extraction with considering the variance characteristic of tissue attenuation along the A-line direction on the region of interest (ROI) or available and analyzable regions belong to the superficial area in the IVOCT image

6. Investigation of developing an automatic method for the classification of lesion tissues with deep learning, and trying to build up a deep learning model for the research of semantic segmentation

1.6 Overview of this thesis

Chapter 2 discusses pre-processing methods for the catheter imaging circle region detection, GW dark shadow sector segmentation and residual blood elimination. For the catheter imaging removal, according to its circular shape and the structure consisting of several concentric rings, I constructed a circle detector for matching the circle pattern of the catheter imaging in the IVOCT image center region. Dynamically changing the radius of the circle detector to maximally cover the catheter region. For the GW bright spot detection, I detected the GW black shadow region instead of directly recognizing the GW bright spot by proposing a new automatic method building a circle-ring detection model. The possible portions as a part of the GW shadow can be obtained with our algorithm, then an entirety GW shadow region is segmented through the computed angle which can also be used in the lumen boundary detection to remove the GW borderline. Besides, residual blood is also eliminated through the Otsu method and morphological operations in this chapter.

Chapter 3 explores automated methods for lumen boundary segmentation and stent sturts detection. Analysing the intensity changing of the A-line profile, the intensity changing from the lumen area to the vessel wall presents a phenomenon of value mutation on the A-line. Simultaneously, after the peak intensity, fluctuating changes in light intensity behave differently. I chose the statistic variable (standard deviation) as the basic analysis factor and formed a new value with thresholds and coefficients for the lumen border and stents detection.

After acquiring the candidate points of bright reflection spots, adaptive region growing is applied for the GW removal to confirm the final stent spots. In our stent detection method, the stent covered with neointima could also be identified.

Chapter 4 focuses on the recognition and classification tasks of vessel lesion tissue by extracting tissue features with traditional methodologies. To determine the plaque distribution of the visible vessel wall, simultaneously, utilizing the 2-D A-lines region instead of 1-D single A-line profile for feature analysis, I proposed a local multi-layer model using the A-line sub-region (ALSR) as its basic element for feature extraction research. Different features of ALSR were extracted along the radial and circumferential dimensions to discuss the statistical intensity information of three types of plaques. Furthermore, I also employed the gray level co-occurrence matrix to gain the texture information of ALSR from the superficial layer. All these features were finally put into a machine learning classifier (random forest) to predict the classes of the ALSRs. This method can solve the angle distribution of lesion tissue in the circumferential dimension.

Chapter 5 explores using the deep learning methodologies to ALSRs for the deep features extraction. Given the advantage and architecture of VGG-Net, and considering the size of each ALSR, I constructed an 11-layer VGG-like deep neural network to complete this task. Additionally, using Local Binary Pattern (LBP) for the texture information analysis to process the input data, and examining the effect of channel amount on the ALSR classification, I used LBP to generate a single-channel input and merged it with RGB channels to produce a four-channel input data. Three kinds of the channel (LBP, RGB and LRGB) in the designed and learned 11-layer deep learning model were tested. The experiment showed that four channels performed better than the other two types.

Chapter 6 employs the pixel-wise classification of lesion plaques with the semantic segmentation technique of deep learning. To reduce the learning time and improve the pixel classification accuracy, considering the analyzable region of the IVOCT image simultaneously, I segmented the IVOCT image with a level-set method to create the region of interest (ROI) of the superficial layer of the human vessel wall. Each input data derived from the ROI is the patch that cropping the ROI with fixed width and height. Hence, an ROI of one IVOCT image can create a number of cropped patches containing some reduplicated pixels. Placing the prediction result of each patch to a 2-D space according to its original position, and using these 2-D regions to construct a 3-D volume of which the third domain is the numbers of input data (or cropped patches from one ROI). The class of every pixel with the same location in the 3-D volume is finally determined by the maximum number of the prediction results of each pixel.

Chapter 7 gives out a conclusion for each part of this thesis, and discuss the existing problems and future work.

Chapter 2

IVOCT pre-processing

2.1 Image acquisition

IVOCT image data were acquired by using a Fourier-Domain OCT imaging system (ILUMIEN$^{\mathrm{TM}}$ OPTIS$^{\mathrm{TM}}$, Abbott, Santa Clara, California, USA) and an intravascular OCT catheter (Dragonfly$^{\mathrm{TM}}$ OPTIS$^{\mathrm{TM}}$, Abbott, Santa Clara, California, USA) at Wakayama Medical University. The above system used in the Department of Cardiovascular Medicine at Wakayama Medical University is called OCT diagnostic imaging system with a JMDN code of “70031000”, and the general name of the catheter is the intravascular radiographic catheter with a JMDN code of “70275000”. The 2.7 Fr intravascular OCT catheter of the mono-rail rapid-exchange type was advanced to the distal coronary artery over a 0.014-inch conventional angioplasty guide wire through a 6 Fr guide catheter. As near-infrared light penetrates only a short distance through blood, temporary blood clearance is required for OCT imaging. In order to clear the blood in the human vessel, preheated contrast media at 37 degrees Celsius (Omnipaque 350 Injection, Daiichi Sankyo Co, Ltd, Tokyo, Japan) was flushed through the guiding catheter at a rate of 2-4ml/sec for approximately 3-6 seconds using an injector pump (Mark V; Medrad, Pennsylvania, USA). When a blood-free image was observed, the OCT imaging core was withdrawn at a rate of $18 mm/sec$ using the stand-alone electronic control of the pullback motor. For each signal transmitting and receiving, it is called the A-scan procedure to obtain each A-line data depicting the current tissue signal in the radial direction through the OCT technique. A complete A-scan containing the tissue signal is transmitted to the computer console to produce a full cross-section image presenting the inner situation of the vessel. During the pullback period of the catheter, which is called a B-scan procedure, the number of cross-section images (frames or slices) for one patient treatment were generated, and these OCT images were stored digitally for the subsequent analysis.

Two kinds of pullback speed exist in this OCT system, one is $18\,mm/sec$ and the other is $36\,mm/sec$. The catheter cost 3 seconds over a distance of $54\,mm$ with each frame interval of $0.1\,mm$ and the latter cost 2.1 seconds on the distance of $75\,mm$ with each frame interval setting as $0.2\,mm$. The data frame rate was 180 fps. This OCT system has both axial and lateral resolution $< 20\,\mu m$. In guidance materials, its range is 12–15 μm and currently, the spatial resolution in these laboratory image data is $9.9 \times 9.9\,\mu m$. The catheter radiates light along axis direction to scan the vessel inner wall tissue with the movement of pulling back. A single vessel OCT image is formed by A-lines that contain the information of reflected light from the vessel tissues[13]. With a rotating catheter, a cross section of the vessel is obtained to present the details of the *in-vivo* vessel. In the OCT system, the IVOCT images are modeled with the OCT signal $\langle I_d(r) \rangle$ which is denoted by fitting a single scattering model

$$\langle I_d(r) \rangle = T(r) \cdot S(r) \cdot I_0 \cdot exp(-\mu_t \cdot r), \tag{2.1}$$

where r denotes the penetration depth, $T(r)$ is the point spread function of the catheter[85], $S(r)$ indicates the signal roll-off with scan depth, parameter μ_t is the attenuation coefficient. Equation 2.1 is fitted every A-line starting from the lumen border in the polar domain to use a small window with variance length to obtain the tissue attenuation, which is also utilized to describe the signal with the homogeneous property. Generally, the vessel wall is a heterogeneous structure, that is the category of tissue along the radial is diverse. Hence, different tissues present different attenuation coefficient, $\langle I_d(r) \rangle$ of an A-line usually would fit more than one type. Furthermore, the intensity of each scanning A-line progressively decreases with depth increasing and each type of tissues has a different light attenuation coefficient. These characteristics are described in Lambert-Beer law[40, 86].

2.2 IVOCT image conversion

Notably, the coordinate of original IVOCT images stored in the OCT system is in the form of a polar domain, which is not convenient to analyze by using computer graphics and image technologies, sometimes. Therefore, it needs to convert the original IVOCT image to the Cartesian domain to obtain the IVOCT imaging as same as the morphology and structure of the real cross-section of the artery. Although this conversion can be performed through the OCT system if necessary, the interconversion that transforming of polar coordinate to Cartesian coordinate (P2C) or Cartesian domain to the polar domain (C2P) normally occurs in the subsequent IVOCT image processing procedures. As shown in Fig. 2.1, a convert operation is performed to the IVOCT image from the polar domain to the Cartesian domain for matching the normal morphology of vessel cross-section.

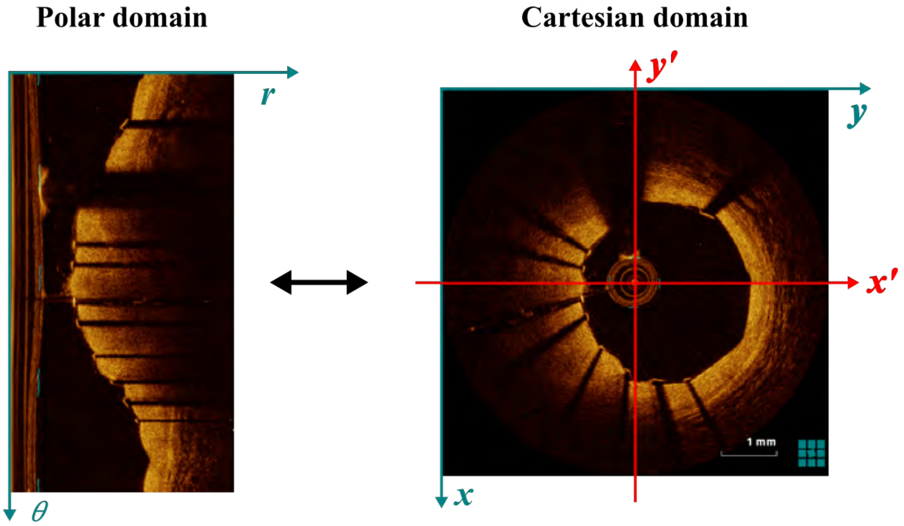


Fig. 2.1 Two different image representations within different coordinates. Left: shows the polar domain image, which is the original version of the IVOCT image by stacking each A-line. Right: is a transforming image in a Cartesian domain with converting each point from (r, θ) coordinate to (x, y) coordinate.

C2P method: generally, an IVOCT image associated with the Cartesian domain is a square size. Suppose that d is the width and height of (x, y) coordinate, simultaneously is a diameter in (x', y') coordinate of which the origin point is the center point of (x, y). The radius r in the Fig. 2.1 is defined as:

$$R = \frac{d}{2.0} \tag{2.2}$$

The coefficients for the θ and r associated with the rows and columns of polar image is denoted as:

$$c_r = \frac{R}{cols} \tag{2.3}$$
$$c_\theta = \frac{2\pi}{rows}. \tag{2.4}$$

where rows and cols is the height and width of the polar IVOCT image, respectively.

Lets (O_{x_0}, O_{y_0}) be the origin point of (x', y'), which also is the center point of (x, y), the coordinate relationship between (x', y') and (x, y) is

$$x' = y - O_{x_0} \tag{2.5}$$
$$y' = O_{y_0} - x, \tag{2.6}$$

Here, constructing a new polar domain (r', θ') based on the (x', y') coordinate, the r' and θ' in new coordinate is

$$r' = \sqrt{x'^2 + y'^2} \tag{2.7}$$

$$\theta' = \arctan(y', x') \tag{2.8}$$

For the cases of $r' > R-1$ and $\theta' < 0$, we use $r' = 0$ and $\theta' = 2\pi + \theta'$ instead.

Then, the (r, θ) in the polar domain as the left sub-figure of Fig. 2.1 illustrated is defined as

$$r = \frac{r'}{c_r} \tag{2.9}$$

$$\theta = \frac{\theta'}{c_\theta}. \tag{2.10}$$

P2C method: the P2C operation is to make up a relationship from (r, θ) to (x, y), which is defined as

$$x = R - c_r r \sin(c_\theta \theta) \tag{2.11}$$

$$y = R + c_r r \cos(c_\theta \theta) \tag{2.12}$$

where r is in a range from 0 to θ_{max}, usually, we set θ_{max} as 720, and r is in a range from 0 to R.

Overall, according formulates (2.2)-(2.12), a quick and reliable interconversion of IVOCT images between polar domain and the Cartesian domain can be performed to help complete task of the vessel image analysis. But a phenomenon that distorts happen simultaneously with among the conversion results should also pay attention and be processed. Observing the right sub-figure in Fig. 2.1, the shape of a bright reflection near the catheter imaging is distorted in the left sub-figure, obviously, including the catheter imaging too. This is because the relationship of points between the above two coordinates is not a one-to-one correspondence.

2.3 Image noise eliminating

The IVOCT image has an intrinsic problem of the noise speckle, which would impact the lumen boundary detection and the stent recognition. In general, speckle noise is the grainy salt-andpepper pattern present in radar imagery, also speckle noise can be understood as a granular “noise” that inherently exists in and degrades the quality of the medical ultrasound, synthetic aperture radar (SAR), active radar, and optical coherence tomography images. Images obtained from these surfaces by coherent imaging systems such as laser, SAR, and ultrasound suffer from a common interference phenomenon called speckle. Here, we utilized Gaussian filtering with a kernel size of 3×3 and a standard deviation $\sigma = 0.8$ to eliminate these speckle noise in IVOCT images. Of course, other filters can also be applied to smooth the image, such as bilateral filtering[79] used in [54].

2.4 Catheter imaging removal

As Sec. 1.2.2 described and Fig. 1.6 shown, the OCT catheter for *in vivo* vessel rotary imaging consisted of a rotating optical fiber with a microlens at its tip, which was placed inside a water-flushed sheath with an outer diameter of 0.9 mm. Observing the IVOCT images, the catheter imaging with multiple imaging shapes is captured presenting as a circle-like region that consisted of several bright concentric circular metal rings, of which the position locates in the center of the IVOCT image. According to publications[81, 92, 22, 80], the catheter's relative position to the vessel wall varies throughout the process of regression in the vasculature, and there is distortion in the concentric circles imaged, simultaneously. The representation of the catheter in the IVOCT image may be displayed as (1) the size of the catheter varies in different IVOCT image datasets, (2) circle rings of the catheter are in contact with each other, (3) the catheter against the wall of the lumen, (4) the width along the radius direction varies. These cases are illustrated in Fig. 2.2. Sub-figures (A)-(F) shows different shape types of catheter, including concentric circle distortion, morphology changing with dynamic radius, count variance of circle rings, touching to the lumen border, which makes it difficult for the catheter detection. The following might be the reasons causing the above phenomenons: (1) the wire receding motion, (2) the catheter in the human body, the location of blood vessel movement changes, (3) the complexity of the vascular internal tissue structure, the probe emits light waves are absorbed and scattered by these tissues, resulting in uneven reception of the signal and other reasons, making the catheter imaging characteristics of distortion, (4) distorted concentric circles are also accompanied by image noise in the vicinity of the circumference.

Utilizing the prior information of vessel imaging, [18, 22, 54, 91] removed the catheter region through the known position of the catheter or the determined max-radius of the catheter rings. In addition, methods of a constant mask employed to binary images[82, 83], dynamic radius detection using distance histogram[80], or Hough transform-based method contributed to the research of catheter removal. Although the above methods detected the catheter area successfully in some situations, the case that the catheter location

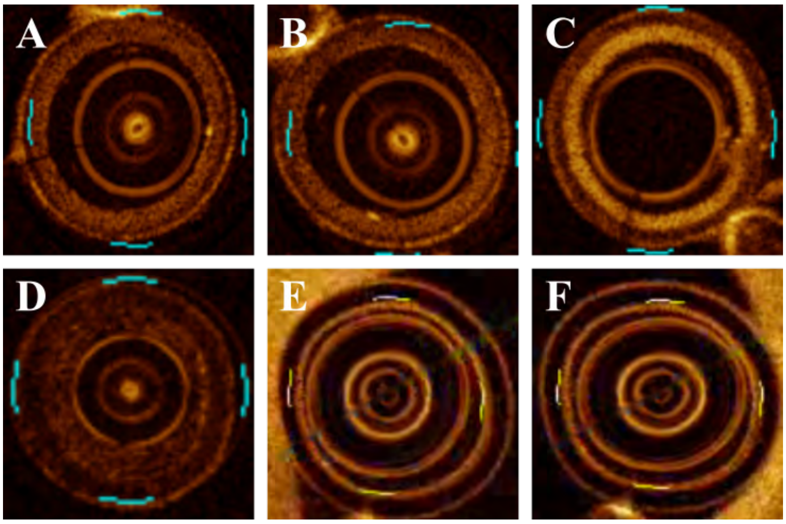


Fig. 2.2 Examples of different catheter shape types. (A)-(C) shows concentric circular rings with distortion; (D) represents a regular catheter but with a different thickness of rings. The circular rings of the irregular catheters are in contact with each other in (E) and (F), also in (A), (B) and (C). Observing (E) and (F), the outer circle rings touches to the lumen border.

touching to the lumen boundary was not considered. Secondly, these methods mostly were implemented to the IVOCT image containing a regular catheter area. Catheter removal is the first processing problem to face, which would impact the lumen border detection and qualitative measure, stent detection with complex conditions, or the further processing of the lesion tissue recognition and classification. To overcome the mentioned problems, we developed an automatic and rapid detection method of catheter area.

Ideally, under the condition that the vessel tissue absorption and reflection of the light ray transmitted by the catheter is the same, and the catheter imaging is regular and non-distortion concentric circle rings. The distribution of pixels in these concentric circular rings is uniform, that is the number and the intensity of the pixels in a unit portion of an individual circle is equal to each other. With a uniform circle, cutting the circle to two equal-half parts, beyond doubt, the mean of the intensity of the two half circles is equal. Therefore, the intensity variance between the two half circles is zero, ideally. From Fig. 2.2, the distortion part of each concentric circle in every catheter imaging is only a small portion of each circle while the shape of the most part of a circular ring is circle-like. Additionally, the thickness of the outer border along the radius direction contains a certain width. Therefore, an intensity variance of the average intensity of the two half circles can be calculated according to the above characterization description. With the true situation of nonuniform pixels distribution in the concentric circle and the noise impacting, we can build a circle detection model with a certain radius in a set of the concentric circular region, and then compute the variance results with these pair half-circles, to form a detection circle feature description with the corresponding the center coordinate and the radius. Transforming the circle center coordinate and varying the radius of the circle detector, a group of intensity variance values is gained with our method and the pixels satisfying the condition would be labelled as the catheter's pixels. As shown in Fig. 2.3, each circle (yellow line) each circle is depicted as a circle detector to match the corresponding pixels of the catheter area. Facing the shape and position alteration, the following issues exist:

- the scope of center coordinate of one circle detector
- the maximum radius value of the circle detector
- intensity variance value statistic

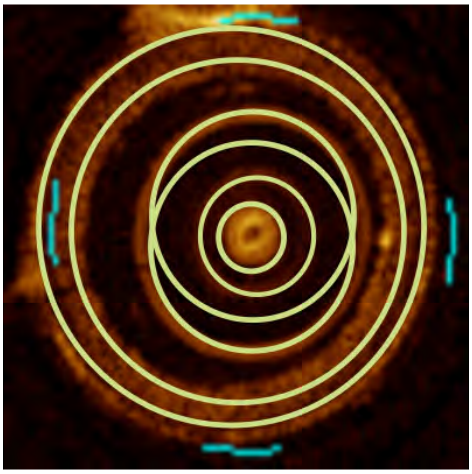


Fig. 2.3 Demonstration of circle detectors utilized as a model to acquire a circle pattern. Each circle line (yellow) indicates a circular detector.

To overcome the above problems, it need to (1) determine the range of the circle center coordinate; (2) the maximum possible radius of the circle detector should be cogitated to prohibit the radius value out of the IVOCT image range; (3) to statistically analyze the circle detector satisfying the judgment condition is the circle ring of the catheter region.

Catheter model definition

Supposing an imaging catheter composes of a set of regular continuous circular rings on which the pixel intensity distribution is uniform. Building a circumference detector to identify a region with the circle-shape characterization. Let G be the vessel OCT image, the circumference detector [$C(r_c, x_c, y_c)$] is modeled based on a radius r_c and a circle center (x_c, y_c), where $r_c \in G$. Theoretically, if a circumference detector is on one of the circular ring regions of the catheter, as shown in Fig. 2.4, the intensity of all points belonging to this circle detector should be same in our ideal model and the mean absolute difference of intensity (MADI) between the two individual half circumferences of the current circumference must be zero. That is, with the assumption that the pixel intensity distribution is uniform, the intensity of each pixel in both of the two half parts is equal in an ideal situation. In fact, considering the width of a circular ring along the radius direction and the case of distortion of the circular ring shape, the circumferential pixel intensities of one circular ring of a real catheter imaging are completely different. Hence, the MADI of a circumference obtained by the circumference detector from the real catheter is not zero. To solve this problem, the MADI threshold (TH_{MADI}) of the intensity is employed as the constraint condition of the circumference detector to control the intensity difference in a certain range, with which an approximate circle detector model is constructed for the catheter region detection. Moreover, we observed that almost all catheter imagings have the appearance of a distorted geometry in Fig. 2.2. Further speaking, our circle detector would meet an irregular circle that only part of it is on the circle detector while other parts are not. In this case, our circle detector with TH_{MADI} can also beat this issue. To the multi-concentric distortion circular rings, we shift the center of the circumference detector within a small range to detect the approximate circle rings of the catheter. On the other hand, if traversing all the points in G would increase the algorithm's computation and computation time, and it would be pointless to perform model detection outside the catheter region, defining the local region B_{center}, B_{center} denotes the local region defined with the center of the OCT image as its center point.

Let C_1 and C_2 be the two individual half circumferences of the circle detector in an IVOCT image G respectively. B_{center} with the size of $N \times N$ is defined as a local region for the dynamical movement of the detector center (x_c, y_c) accomplishing the detection of the distorted catheter. The relation among the above definitions are $(x_c, y_c) \in B_{center} \subset G$. The MAID of the ith circle detector is calculated as:

$$MADI^i = |MI^i_{C_1} - MI^i_{C_2}| \quad i = 1, 2, \ldots, N_{B_{center}}. \tag{2.13}$$

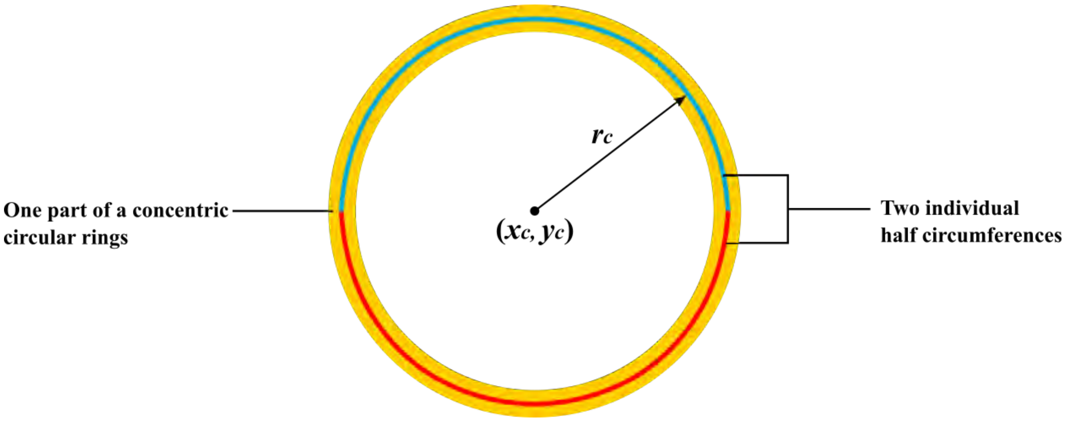


Fig. 2.4 A circle detector (blue and red) with certain values of radius r_c and circle center coordinate (x_c, y_c), which presents into two equal half parts, applies to a component (yellow) of a catheter imaging to explain the catheter model application on the catheter to obtain the related pixels.

Put the $MADI^i$ into the constraint condition for the circumferential shape judgment:

$$O = \begin{cases} a\ circle & \text{if } MADI^i \leq TH_{MADI}, \\ & MI^i_{C_1} \geq \gamma \text{ and } MI^i_{C_2} \geq \gamma. \\ not\ a\ circle & \text{otherwise.} \end{cases} \tag{2.14}$$

where $MI^i_{C_1}$ and $MI^i_{C_2}$ are respectively the mean intensity of two half circumferences of the i-th circle detector. $MI^i_{C_j}(r_c, x_c, y_c) = I(r_c, x_c, y_c)/N_c$, $(j = 1, 2)$. Circle detector radius $r_c \in [0, r_{min})$ and $r_{min} = \lambda min(x_c, y_c)$. N_c is the sampling number applied to both C_1 and C_2. $N_{B_{center}}$ indicates the number of the circle detector center coordinate (x_c, y_c) in B_{center}. O represents the current circumferential region defined by our model.

2.5 GW detection and black shadow segmentation

Tiny GWs are designed to navigate the catheter to reach a lesion or vessel segment. As shown in Fig. 1.8, GW presents a bright reflection characteristic due to its metallic property, immediately followed by a big black shadow region with both blurred borders, which shelters a segment of the vascular wall. Similarly, GW and its black shadow impact the lumen border segmentation and stents detection and related qualitative measures. Yet, GW segmentation is always a challenge in a single vessel OCT image because of lacking the priori information on the size and position of GW and the angle range of the GW shadow region. Normally, the size and position of GW would vary and be not fixed in a set of IVOCT images of one patient in a treatment period because of the pull-back motion of GW. Studies[81, 92] accumulated pixel intensities as a function of A-line depth to construct a 2-D *en face* image with a set of IVOCT images firstly, and then segmented GW through dynamic programming and Otsu's method[65] respectively. But this method needs to utilize all the IVOCT images to obtain a *en face* view image, it can't be used in real-time analysis of IVOCT image during the treatment period for a patient, and it processed a set of IVOCT images instead of a single one (Fig. 2.5[1]). Additionally, an adaptive region growing algorithm was applied to the obtained candidate seeds to recognize the maximum area which was denoted as GW[69]. The above methods are all employed to segment GW in the analysis processing of IVOCT images after they are fully acquired.

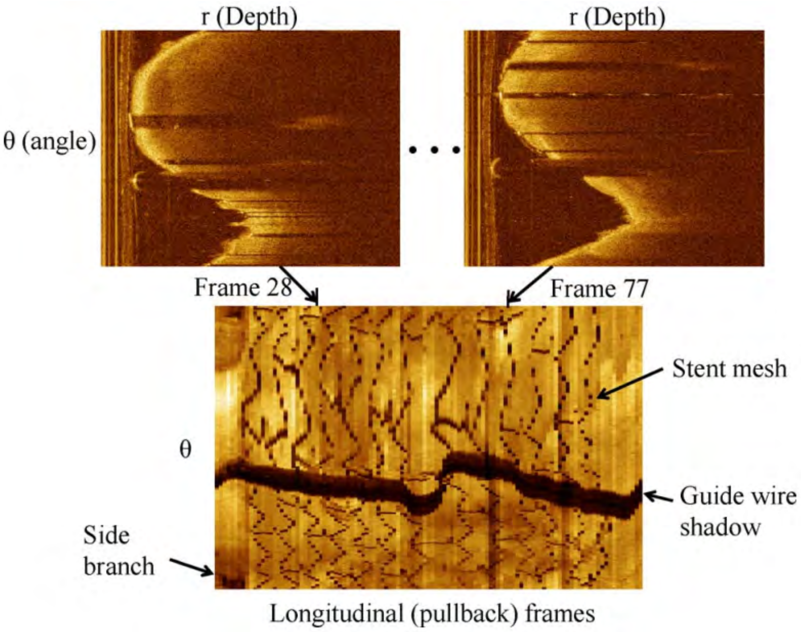


Fig. 2.5 Illustration of Wang's method on the GW segmentation method[92] that accumulating the A-line intensities of individual slices to synthesize along the longitudinal direction to produce *en face* projection image. The final longitudinal frames image composed of a set of IVOCT images.

To overcome the above problems and develop a real-time GW segmentation method during the patient examinations, in this section work, we propose a novel automatic algorithm using a circle-ring detection model (CRDM) to segment GW and its black shadow region. As we know, the segmenting object consists of two components: (1) a bright reflection of

---

[1] This figure is the original source and derived from the publication[92] to illustrate the creation procedure of *en face* image.

GW and (2) its black shadow which presents a sector shape (Fig. 2.6). Our main idea is to recognize the borders of both sides of the black shadow and then obtain the corresponding sector region, including the sector angle and area. With the investigation of the IVOCT images, the obvious fact is that the intensity of GW shadow is lower than other tissues' as same as the condition of the stents, and presents a larger area mostly comparing with other black shadow regions. Additionally, the black shadow region covers the entire depth of a portion of the vessel wall along the radius direction.

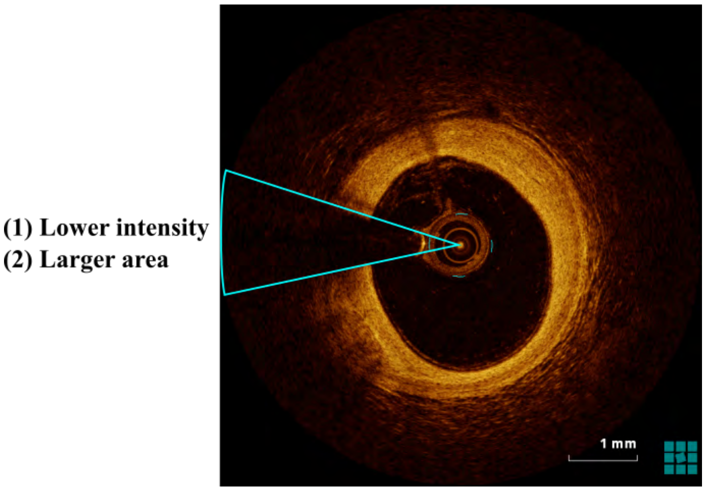


Fig. 2.6 An example of a GW black shadow region denoted with the cyan sector line. Comparing with the tissues around, GW shadow is characterized as a visually lower intensity and its area generally presents a larger measurement.

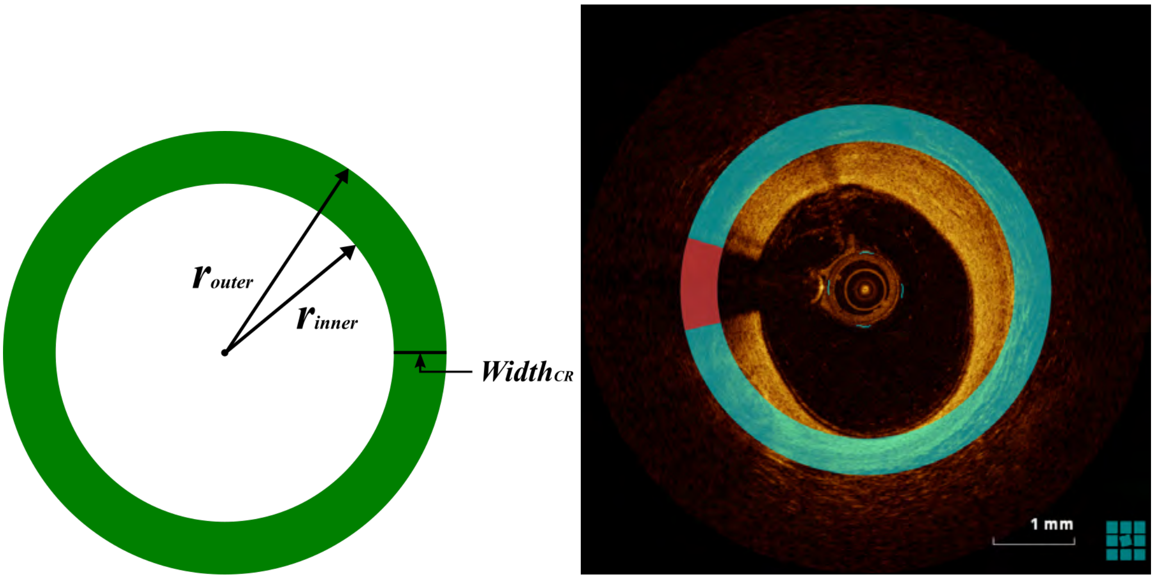
CRDM definition

The CRDM contains three parameters: the width of CR ($width_{CR}$), the inner radius (r_{inner}) and the outer radius (r_{outer}) (Fig. 2.7a). Let $C_{r_{inner}}$ be a circular region with the radius r_{inner} while circle region $C_{r_{outer}}$ is defined based on the radius parameter r_{outer}, the described CR mask region is obtained by the following formula:

$$CR = C_{r_{outer}} \cap \overline{C_{r_{inner}}} \tag{2.15}$$

where $r_{min} \leq r_{inner} < r_{outer} \leq r_{max}$, r_{min} and r_{max} are the minimum and maximum radius defined in the IVOCT image, respectively. The width of CR $width_{CR} = r_{outer} - r_{inner}$.

Here, we describe the GW black shadow region detection method by using CRDM as follows. The first step is to use the modelled CR to mask a gray-scale IVOCT image for gaining the corresponding CR region, as illustrated in Fig. 2.7b, which demonstrates an



(a) CRDM concept drawing defined with the width of CR ($width_{CR}$), the inner radius (r_{inner}) and the outer radius (r_{outer}).

(b) An example of CRDM (cyan CR) applying to an IVOCT image to detect the GW black shadow region (red segment).

Fig. 2.7 CRDM illustration and application in an IVOCT image.

example of one of CRs which obtained the CR region of an IVOCT image with CRDM at a certain time. With the r_{inner} and r_{outer} simultaneously increasing and iterably altering its values within a range, a number of CRs is produced in one IVOCT image, as represented in the first row of Fig. 2.8, totally 7 CRs generated with start and end radius to define the value boundary. Using these generated CRs to mask a original gray IVOCT image, the corresponding original CR regions (CR_{oriImg}) can be obtained from the current gray IVOCT image. For each $CR_{origImg}$, morphology operations containing dilate and erode with kernel 3×3 and iteration of 5 times are implemented to $CR_{origImg}$, immediately followed by a opening operation with a 3×3 kernel. Before applying a series of morphology operations to the CR_{oriImg}, a threshold method with setting th_{min} and th_{max} respectively is employed to CR_{oriImg} to remove the noise and unrelated object. The object region (like tissue and black shadow) of each CR_{oriImg} is divided into two classes (white and black) after morphology operations, which produces a corresponding binarized CR ($CR_{binarized}$). The middle row of Fig. 2.8 are the results from the CR_{oriImg} processed with the our method. Seven CRs with different radius yielding seven matching ($CR_{binarized}$), which the black areas present the shadow regions of GW and stents or a portion of the lumen and the portions of the white display of the vessel tissue in one CR. In the middle row of Fig. 2.8, obviously, not all the CRs hold a portion of GW shadow for the reason of the diversity and complexity of the morphology of the inner lumen of on of GW shadow. Here, the $CR_{binarized}$ containing the maximum number of white area is firstly considered as the candidate CR ($CR_{candidate}$) for the sector generation in the following step, as shown in the right of the third row of Fig. 2.8. The angle of every CR segment in $CR_{candidate}$ can be gained, then. Based on the white areas of the $CR_{candidate}$, the corresponding sector regions are obtained respectively, which is shown at the bottom of Fig. 2.8. Subsequently, a reverse operation is employed to the binarized sectors to convert the original black areas to white areas and the original white areas to black areas. As aforementioned, the types of sectors of the reverse image are GW shadow, the mixed region including an inner lumen and the tissue, and stents shadow. Compared with the stent shadow, the area of GW shadow region is usually lager than the stents', which is an obvious characteristic to GW shadow. To distinguish the mixed region with GW shadow, features of the candidate sectors should be taken into account. Moreover, the appearance of the bifurcation and irregular shape of the vessel lumen shape also causes difficulties for the GW shadow recognization. Observing from the IVOCT image, we know that the GW shadow is usually the biggest black sector region with a highlight block (the reflection of GW) inside if the shape of the vessel lumen is a regular circle shape. Therefore, the following characteristics: (1) the average intensity of the sector (AIS); (2) the sector area (SA); (3) the number of highlight block (NHB); (4) the area ratio between the highlight block and the black region (AR); (5) the intensity ratio between the highlight block and the black region (IR); (6) the maximum intensity of highlight block (MIHB); and (7) the distance between the highlight block and the image center (DHB), are analysed from each sector of an IVOCT image for features extraction of the GW shadow region. To the candidate sector of the stent shadow region, characteristics (1), (2), (4), (5) and (7) can be utilized to recognition of GW black shadow region. For the vessel bifurcation or irregular shape condition, characteristics (1), (3), (4), (5), (6) and (7) can be used for a judgement of GW shadow region.

2.6 Experimental results

2.6.1 Parameter setting

In Sec. 2.4, $TH_{MADI} = 1.0 \pm 0.1$ and finally was set as 1.0. For circle detector radius r_c, if the coefficient λ of r_{min} was set as 1.0, superfluous detection circles satisfying Eq. 2.14 would be produced in the experiment results (Fig. 2.9). To overcome this problem, two approaches were determined: (1) Statistical analysis was performed to select the appropriate value for λ. (2) It was observed that the distance (white arrow) between the two regions in the figure was significant, so the data were sorted using statistical methods to remove excess detection circles and retain the detection region of the catheter at the central area. Through experiments

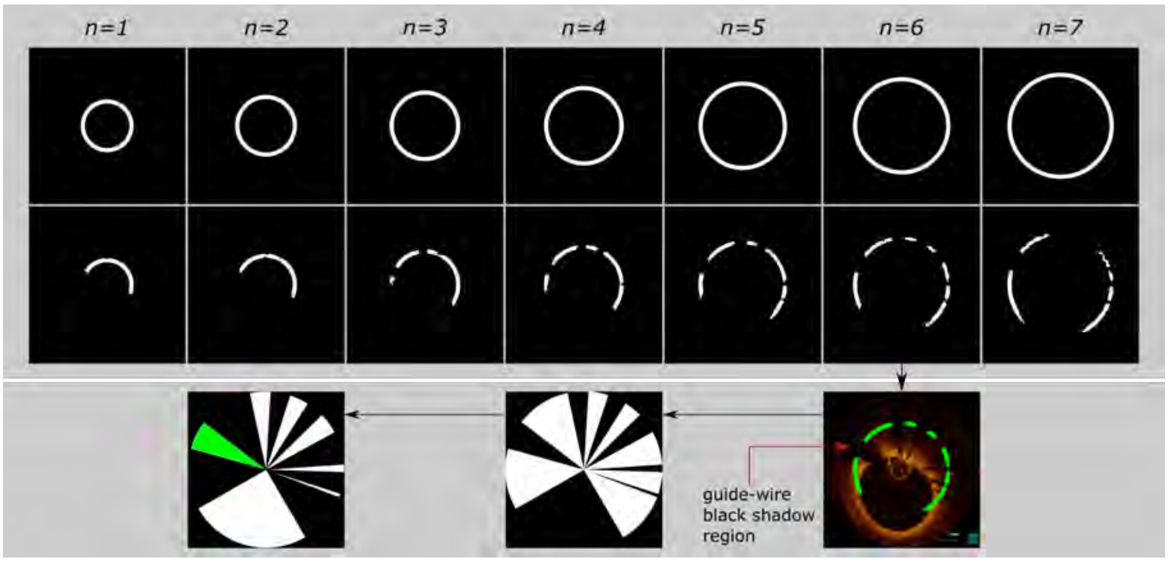


Fig. 2.8 Top: Illustration of the CRs with different radius. Middle: A sequence of $CR_{binarized}$ by applying morphology operations to the CR_{oriImg}. Bottom: A simple flowchat from the obtained $CR_{binarized}$ to the final candidate sectors. The middle binarized sector image is developed from the corresponding $CR_{binarized}$. The red arrow points to the black shadow region of GW. The green sector in the first image presents the GW shadow area.

to test the above two schemes respectively, scheme (1) can solve the problem of catheter area detection in all kinds of cases, but in the processing of OCT images for different classes, a small amount of manual involvement in setting and adjusting parameters is required. Scheme (2) can automate the algorithm, but for the detection of the catheter area in some special cases, the results have errors. Therefore, in this paper, scheme (1) was chosen for the final testing of the algorithm. Radius coefficient $\lambda = 0.25 \pm 0.125$ to fit the multiple cases of catheter imaging of IVOCT images, the threshold value γ for $MI^i_{C_1}$ and $MI^i_{C_2}$ was set to 20. N_c was the total number of pixels in $C_j, (j = 1,2)$ in this paper. B_{center} is a symmetry region of 12×12 size as the location transforming region of the circle center coordinate of the circle detector. A set of initial results of catheter removal by our method is displayed in Fig. 2.10.

For GW shadow segmentation, the range of CRDM radius is important to control the CRDM obtaining the detection circles that contain agreeable targets (tissue and shadow). If it too small, the extra object would impact the segmentation, while with a bigger value, some useless information would be included making the unnecessary computation and value boundary problem. In this paper, we tested and determined that CRDM radius range, $[r_{min}, r_{max}]$, was set to [100, 240] and the CR width $width_{CR} = 20$, finally. In our single detection circle algorithm, th_{min} and th_{max} are used as the threshold values to remove the unnecessary object, and th_{min} and th_{max} was set to 12 and 255, respectively. After that, an

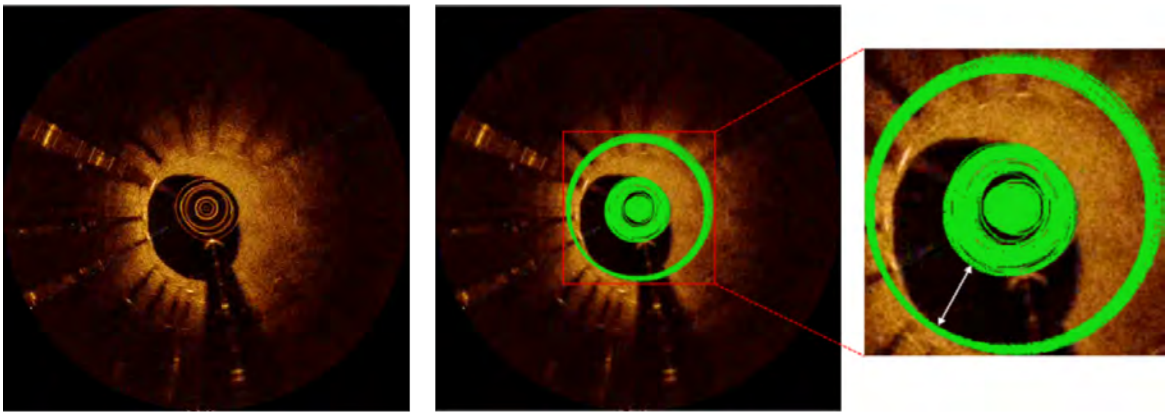


Fig. 2.9 Illustration of catheter imaging region detection. Superfluous detection circles were also generated simultaneously with the correct results if the value of λ larger than 0.375.

area constraint method was utilized to take off the extra and small blocks, here we used e_{area} to illustrate this value and set a threshold value E_{area} with 500 to implement this procedure. Fig. 2.11 displays the middle results in one detection iteration period, where it can obviously notice that GW shadow region (asterisk) is not covered. Subsequently, we can gain the portion of the detection circle with respect to GW shadow after a reverse operation is applied.

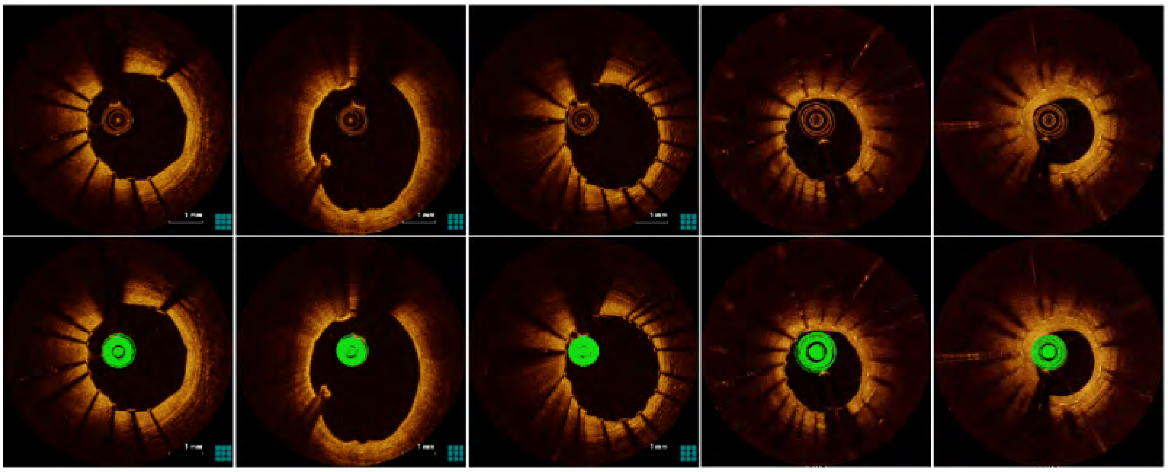


Fig. 2.10 Catheter removal results by using our catheter detection method. Top: The catheter with different position in Original IVOCT images. Bottom: Catheter detection results. The last three results show that our method could also detect the catheter even it adjacent to the lumen boundary or stents.

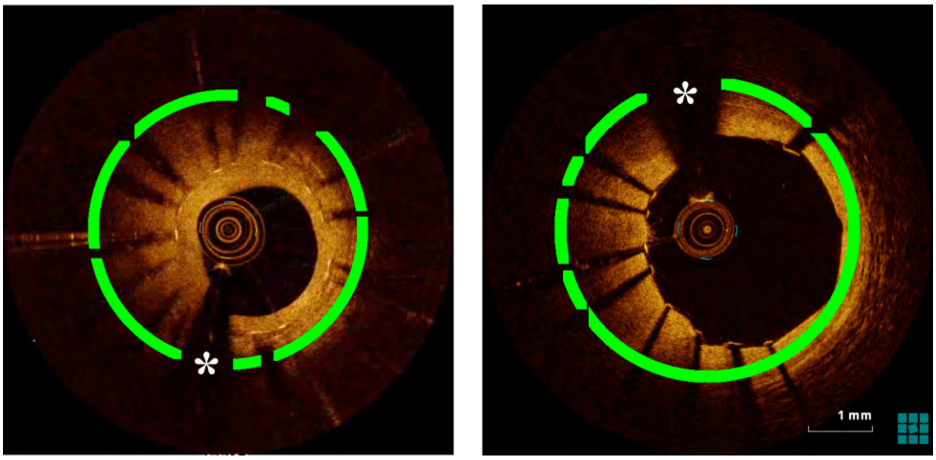


Fig. 2.11 Examples of middle results (green) for GW shadow region (asterisk) detection with our algorithm.

2.6.2 Validation

The proposed catheter removing and GW shadow region segmentation were tested on 7 human IVOCT datasets (total 490 IVOCT images). We retrospectively investigated OCT datasets obtained from 7 coronary arteries in 7 patients with stable coronary artery disease. The ground truth of catheter areas and the central angles of GW shadow region were all measured and labelled by an expert observer.

Mean absolute difference of area (MAD_{area}) and center angle (MAD_{angle}) are used to validate the difference between the manual and automated methods on the catheter segmentation and GW center angle respectively:

$$MAD_{area} = \frac{\sum_{n=1}^{N} |A_{manual}^{n} - A_{auto}^{n}|}{N} \tag{2.16}$$

$$MAD_{angle} = \frac{\sum_{n=1}^{N} |CA_{manual}^{n} - CA_{auto}^{n}|}{N} \tag{2.17}$$

where A_{manual}^{n} and A_{auto}^{n} are the area results through the manual and automatic segmentation methods. CA_{manual}^{n} and CA_{auto}^{n} are the GW shadow center angles derived by manual work and our algorithm respectively. N is the amount of the data sample. Table 2.1 shows the mean absolution difference evaluation metrics for the catheter area and the angle of GW shadow region obtained by our methods.

Table 2.1 Evaluation metrics for the catheter segmentation and GW region angle detection

Dataset	$\mathbf{MAD_{area}}$ ($\mathbf{mm^2}$)	$\mathbf{MAD_{angle}}$ (**rad.**)
set 0	0.088±0.009	0.052±0.027
set 1	0.051±0.007	0.063±0.022
set 2	0.053±0.006	0.115±0.057
set 3	0.033±0.004	0.022±0.014
set 4	0.064±0.007	0.117±0.029
set 5	0.048±0.009	0.078±0.030
set 6	0.052±0.005	0.088±0.036

2.6.3 Results

Comparing the manual measurements determined by the expert with automatically derived catheter areas and GW angles respectively, the linear regression method was applied to measure the correlation of manual results and automated results. Figure 2.12 show the agreement between automated and manual results of catheter segmentation and GW angle detection, respectively. The correlation coefficients in both correlation plots presented well, $r^2 = 0.99$. To evaluate the results on the 7 datasets, we compared automated catheter imaging detection against manual assessments and GW shadow segmentation with experts' labeling both in Bland-Altman graphics (Fig. 2.13 and Fig. 2.14). Figure 2.15, 2.16, 2.17 and 2.18 illustrate the results (green for catheter and cyan for GW) by using our automated catheter imaging region detection and GW black shadow sector segmentation. Obviously, Fig. 2.12 indicates that the size of the catheter area and GW angle is fully different for various datasets, which also can be observed in the results presented in Fig. 2.15, 2.16, 2.17 and 2.18. By using our methods, catheter region, GW and its shadow sector can be detected in many cases.

All these experiments in this chapter were implemented with the version of Python 2.7, Numpy 1.14, and Scikit-image 0.14 based on a system of Windows 10 64-bit, 64 GB RAM. The configuration of our experiment computer contained a graphics card of NVIDIA GeForce GTX 1080 Ti and a processor of Intel(R) Core(TM) i7-3930K CPU with 3.20 GHz.

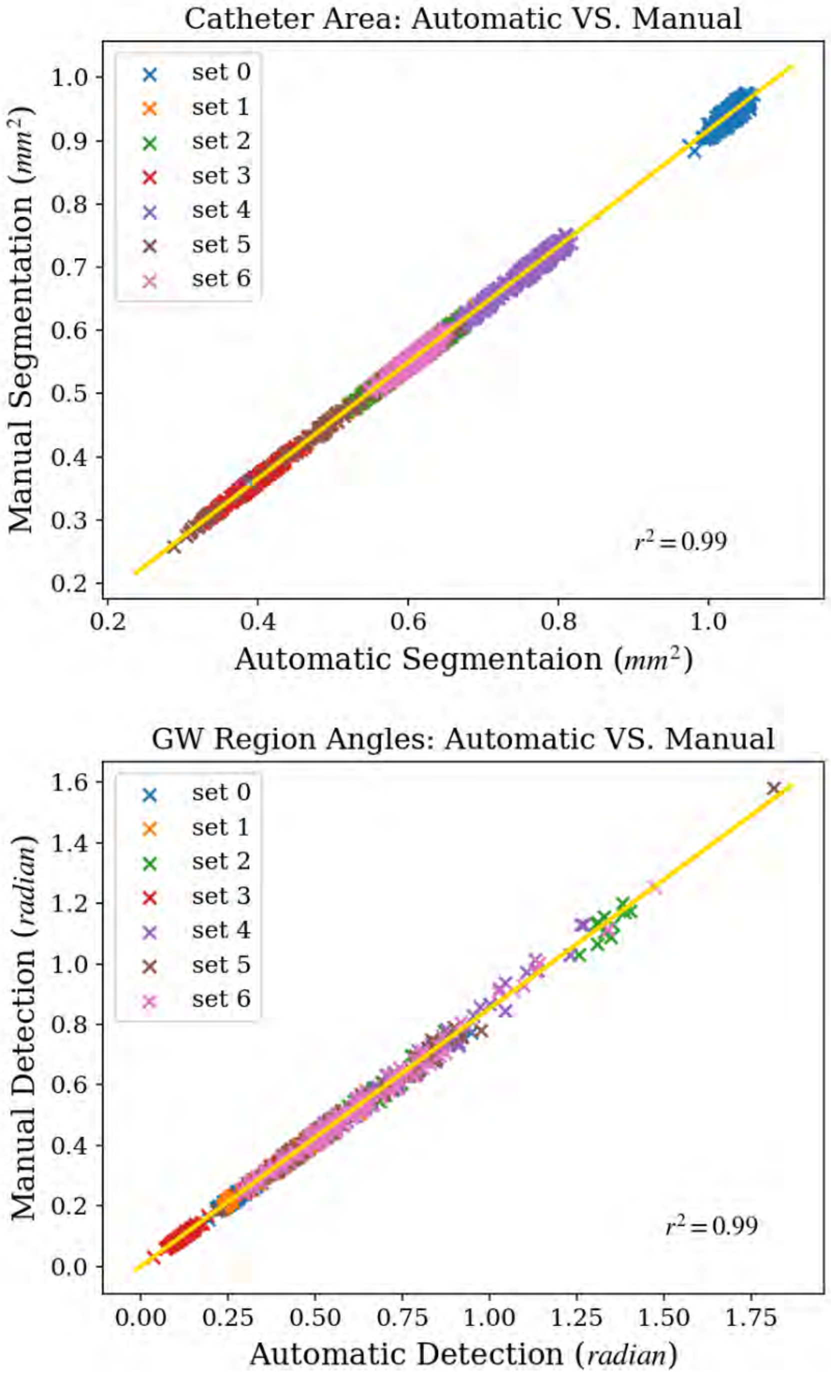


Fig. 2.12 Comparison of automatic and manual catheter and GW segmentation, which is presented through liner regression method. Top: catheter areas segmentation. Bottom: GW shadow center angle detection.

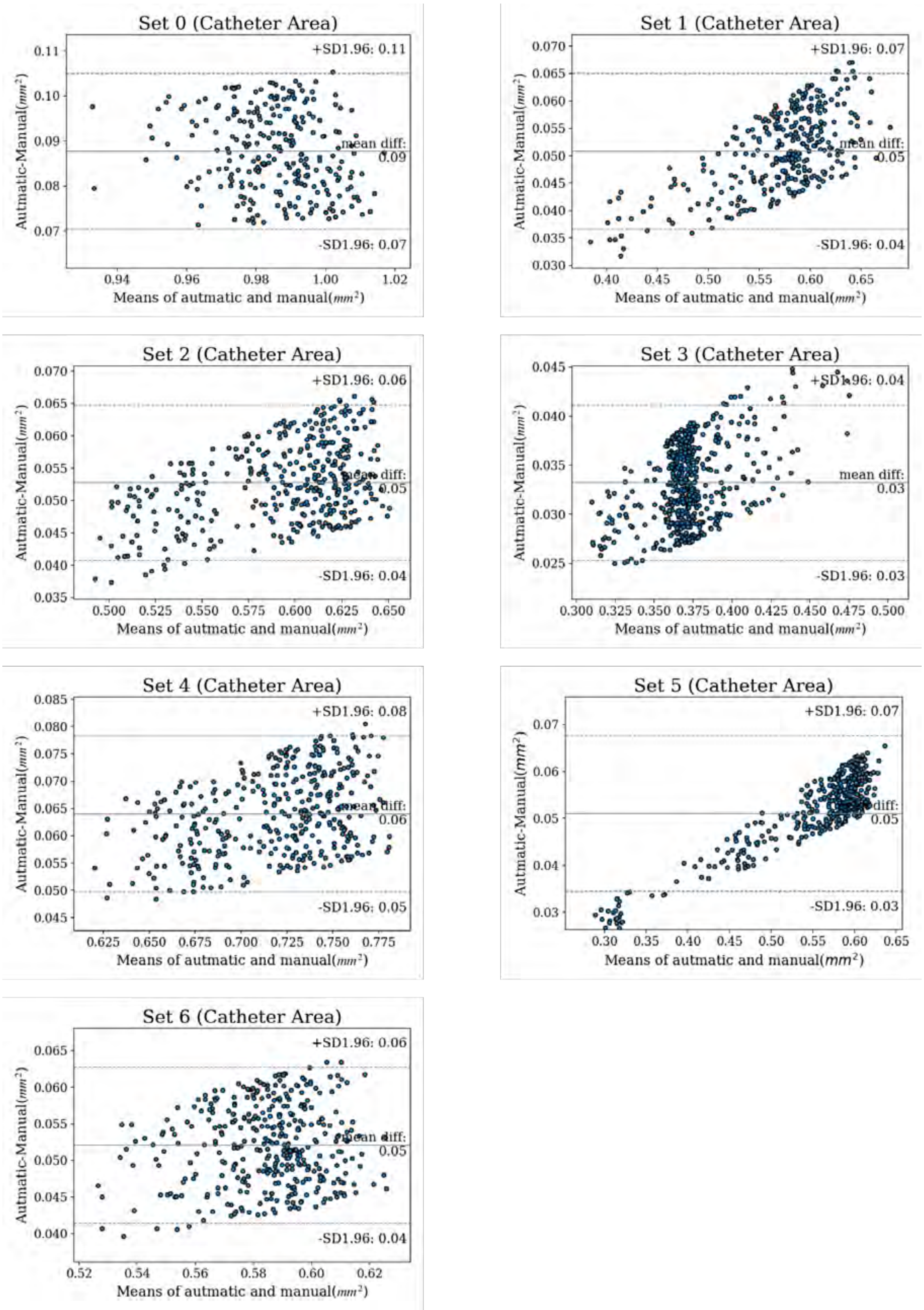


Fig. 2.13 Seven Bland-Altman plots corresponding to seven groups of assessment of catheter area detection by manual against by automated method.

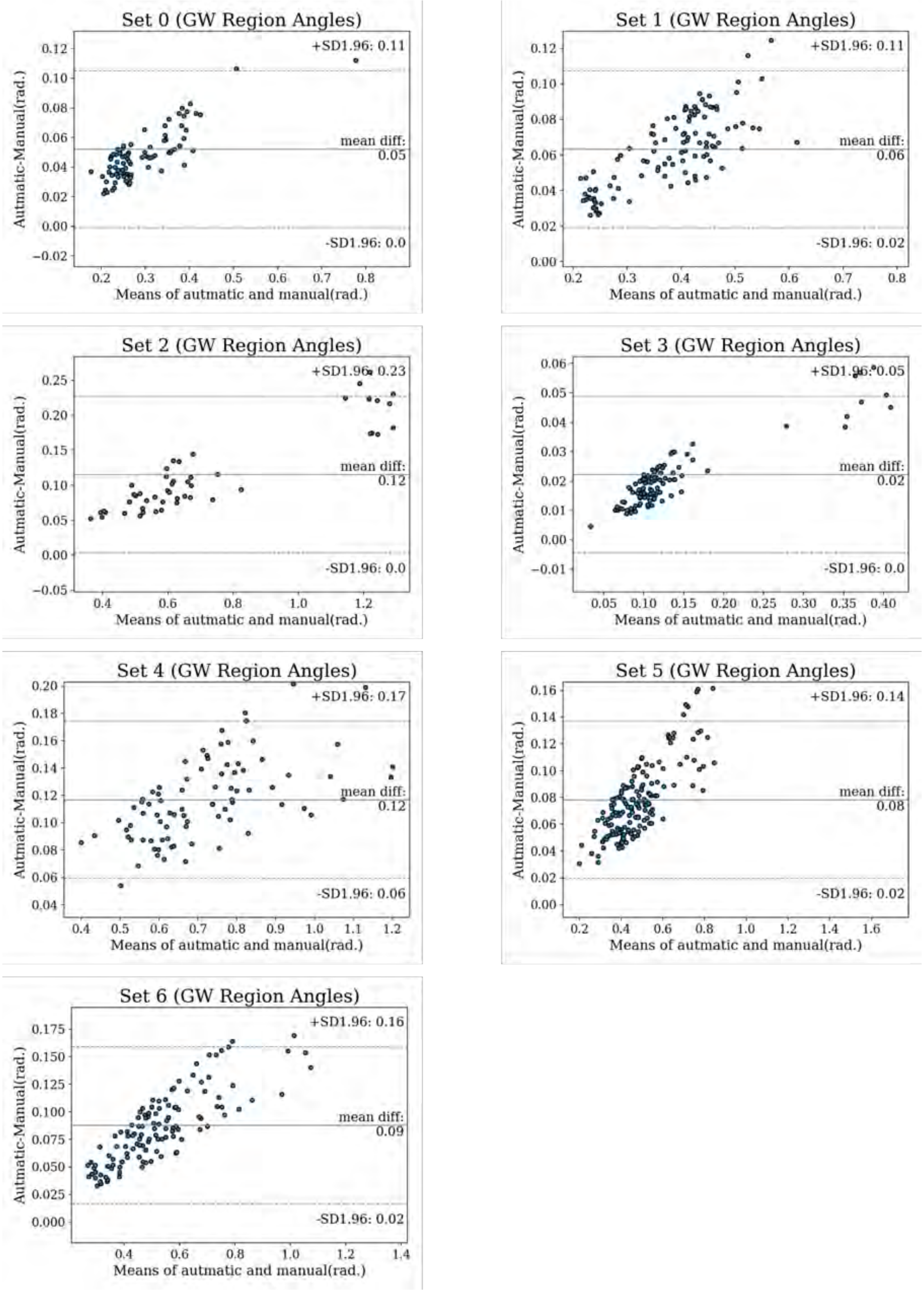


Fig. 2.14 Seven Bland-Altman plots corresponding to seven groups of assessment of GW shadow region angle by manual against by automated method.

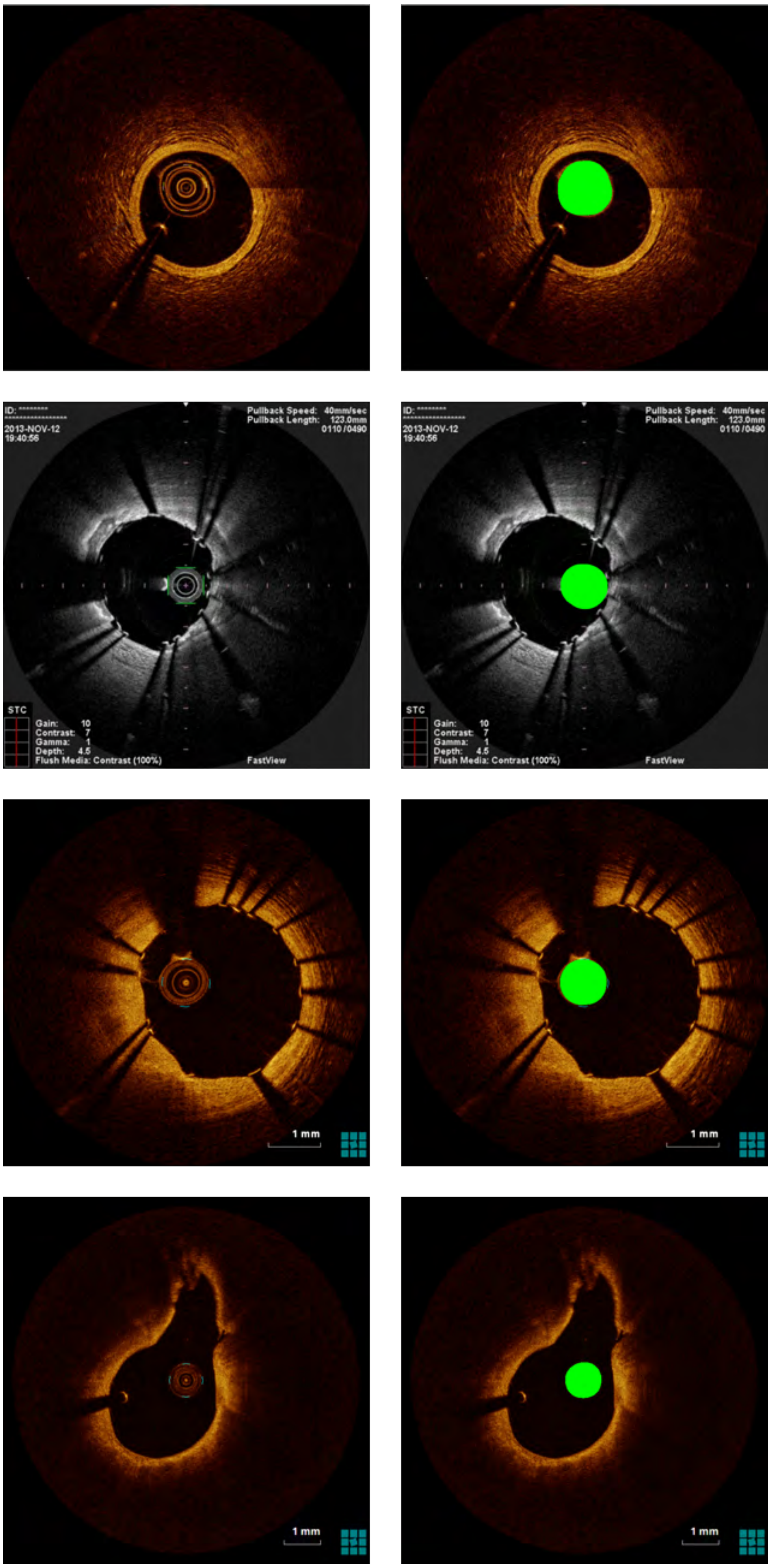


Fig. 2.15 Results of catheter region (green) recognition derived from dataset 0, 1, 2, 3.

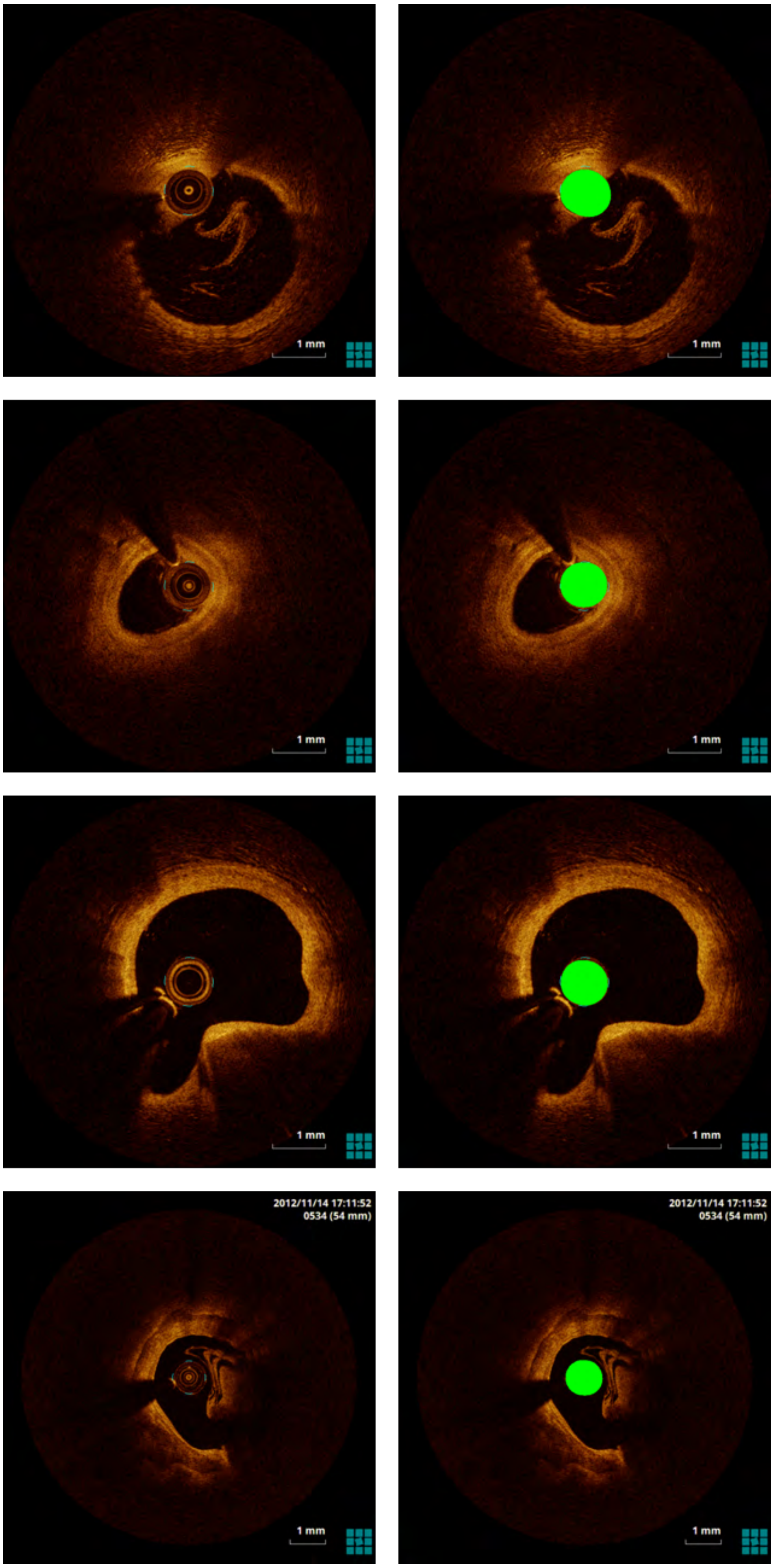


Fig. 2.16 Results of catheter region (green) recognition derived from dataset 4, 5, 6.

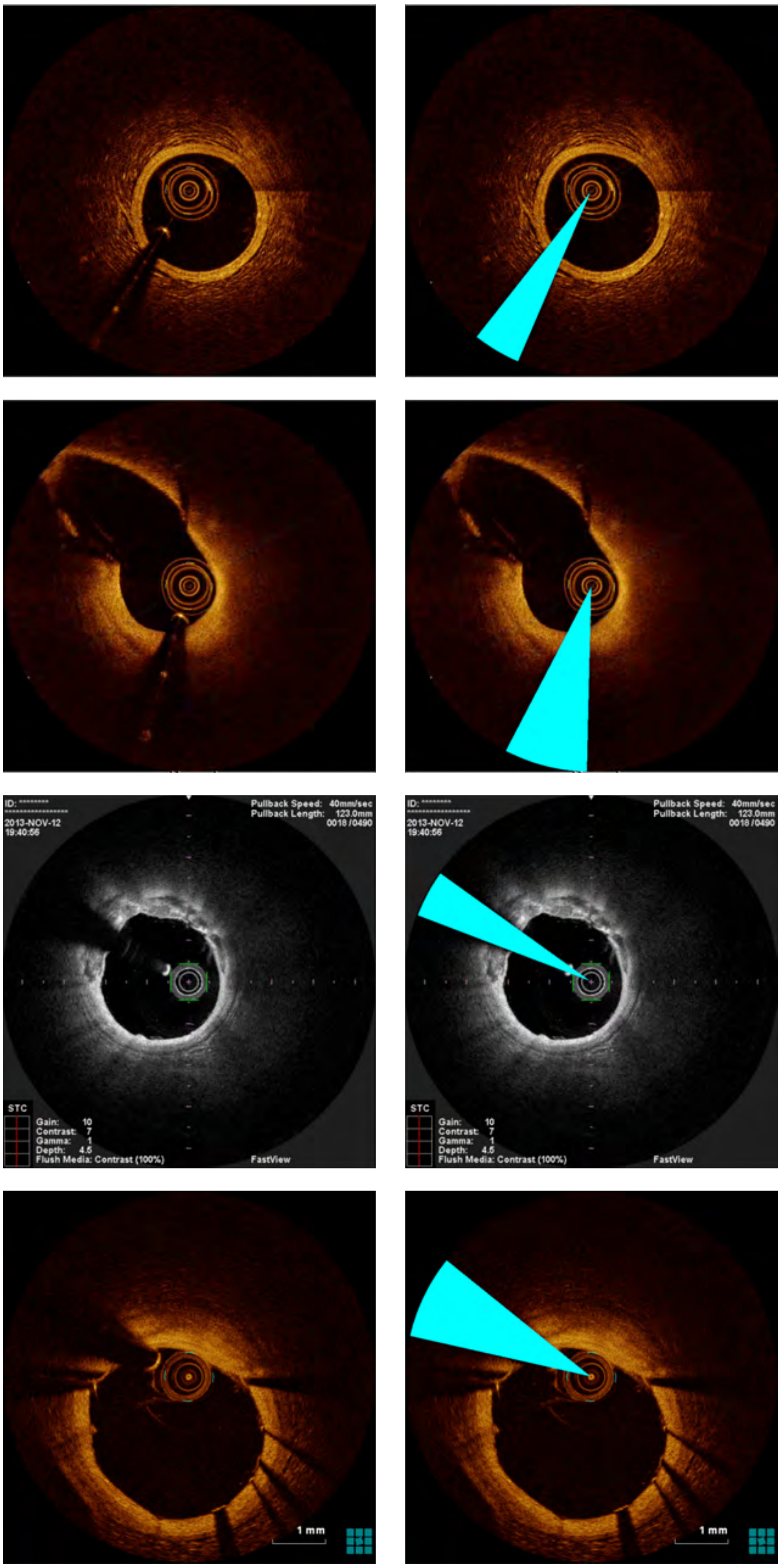


Fig. 2.17 Results of GW black shadow sector area (cyan) derived from dataset 0, 1, 2.

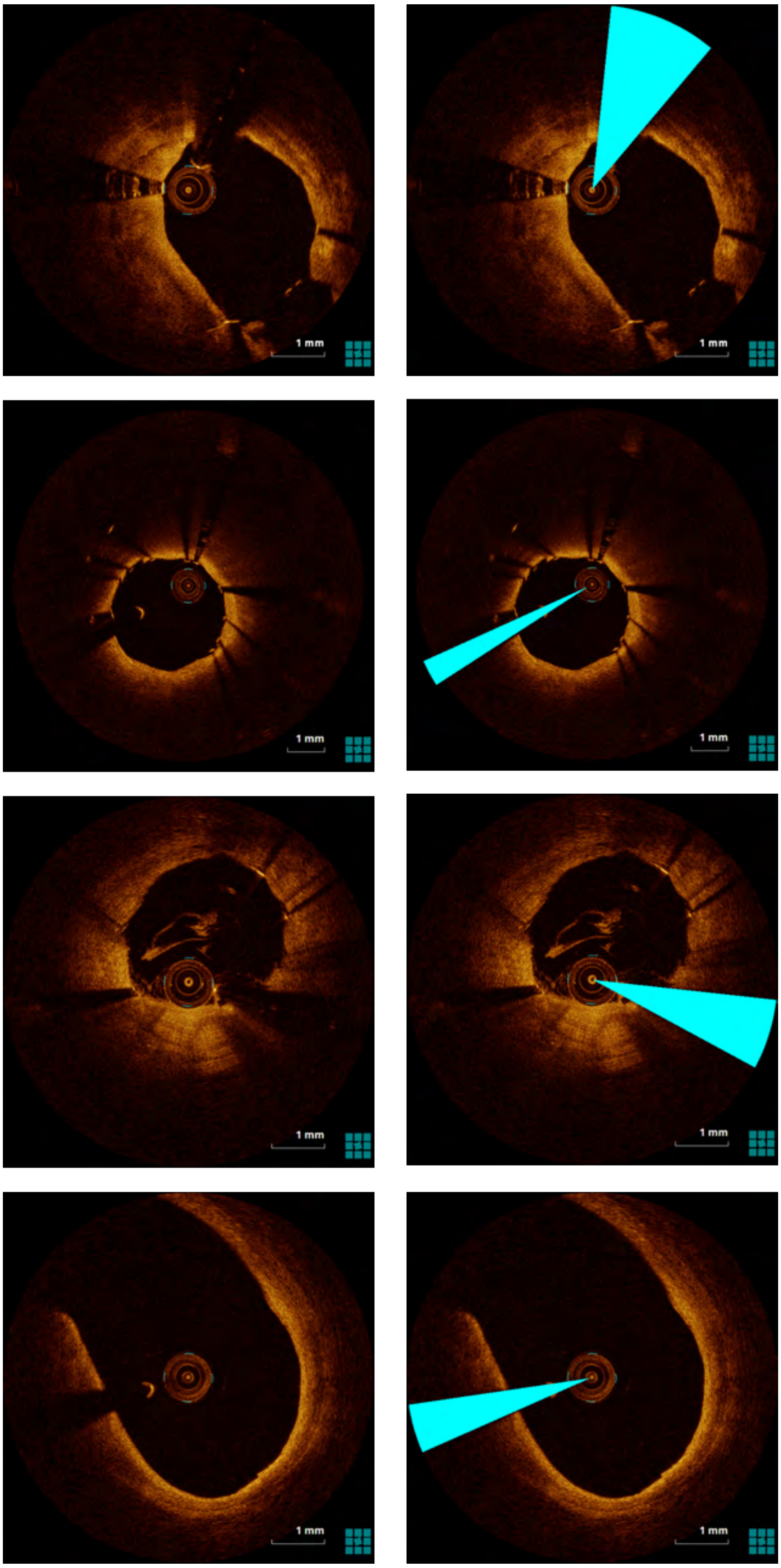


Fig. 2.18 Results of GW black shadow sector area (cyan) derived from dataset 3, 4, 5, 6.

2.7 Discussion and conclusion

In this chapter, a fully automatic catheter segmentation, GW shadow location model are defined according to the imaging properties of catheter and GW. Seven datasets were used for the evaluation of the detection and segmentation results, respectively. The mentioned methods could be considered as a medical-aid tool to assist the CAD doctor treatment in the pre-processing step.

Catheter impacts the post-processing steps using the morphology operations to detect the vessel lumen boundary[54] or extracting features of the stent strut along the A-line direction in the stent quantitative analysis. From Fig. 2.10, we see that our method could depict the concentric circular shape of the catheter area with different size very well, even in the case of the catheter with the position adjoining the lumen border. In order to accelerate the detection speed, we define a local region B_{center} for the circle detector to convert its center coordinate, which it costs 0.7 ± 0.1sec. for each IVOCT image.

Guide wire also causes an effect to the lumen boundary detection and stent detection. The previous literatures[81, 92] detected the GW in the synthesize *en face* projection view image. But this method must compress all OCT images firstly to an accumulated intensity image. Our GW shadow region detection method could be applied to a single OCT image. With the regression methods of machine learning, the location or the angle of the GW and its shadow region of the unprocessed OCT images containing a complicated inner structure can be predicted.

Overall, in this chapter, we mainly proposed two novel automatic catheter and guide wire segmentation methods to address the existing difficulties:

1. We defined a circle detector to detect an approximately circular shape in the catheter imaging area by computing the average intensity variance of both two half-circle, and all the condition-satisfied circles composed a detected area identified as the catheter region.

2. We developed an algorithm that utilizing a CRDM to obtain the corresponding circle ring portion of an IVOCT image to compute the angle of the GW black shadow to gain the corresponding sector region for the GW segmentation.

Experimental results for the catheter removal and GW dark shadow region segmentation shows that our method can perform satisfying outcomes comparing with the manual work. However, the parameter setting for the catheter removing method is not fully automated and the angle value of the sector gained with our GW approach is sometimes bigger than the real value causing a result of segmenting an extra area of the vessel tissue. Therefore, the future work is to continue to investigate the fully automatic catheter imaging removal and improve the segmentation precision of the GW black shadow region.

Chapter 3

Lumen boundary segmentation and stents detection

During CAD treatment, the lumen characteristic assessment performs a significant effect on the condition of diagnosis work by specialists, such as the morphology (circularity or bifurcation), area variance, and a conclusion of vessel health degree initial examination. The quantitative measurement and diagnosis assessment of the vessel lumen boundary is an indispensable job before the step of treatment with stent struts and is significant to indicate the formation of stenosis. As the relative examinations showed, the area of the vessel lumen in different periods that including initial stent implantation and middle or last times that generating neointima tissue, would manifest exceedingly discrepant. It is extremely helpful and necessary to stent implantation with the evaluation of lumen boundary in PCI. In this chapter, I discussed the characteristics of the vessel lumen border and the stents in IVOCT images, and developed methods to complete the tasks of segmentation the vessel lumen boundary and the detection of stents.

3.1 Basic concept and model definition

The obvious appearance of the lumen border is that it's a dividing line to distinguish from the vessel tissue side to the lumen. The significant characteristic of it is that the intensity changing occurs strongly and presents rapidly up or down. Hence, the A-line attribute is usually utilized and characterized as the intensity profile in [81, 82], which was used to detected the borders on each A-line and then obtained a smooth lumen boundary. However, some detected border points were assessed as fake compared to the ground truth. Moraes et al.[60] and Macedo et al.[53, 54, 56] both utilized the Otsu method[65] and morphology operations to separate the IVOCT image into two parts: vessel tissue and lumen, subsequently, the histogram for each A-line to distinguish the tissue and non-tissue and the first non-zero points in the binary IVOCT image with a subtraction operation were employed respectively. Moreover, Canny filter [15], Markov random field (MRF)[48] and level-set model [17, 73] mentioned in Sec. 1.4 were also investigated for the lumen boundary segmentation.

Generally, three challenges depicted in [16] are existing as the common phenomenon when facing the lumen segmentation task. Firstly, the irregularity and complexity of the lumen boundary[57], also including complex bifurcation vessel frame, bring difficulties to the segmentation work. Second, GW dark shadow region makes the lumen border incomplete with the portion information of the vessel lost. And the last one is luminal blood artifacts presenting in the lumen increase the difficulty of lumen border recognition. Figure 3.1 shows real inner conditions of the vascular, irregularity, bifurcation and the blood artifacts can be observed in the IVOCT image. The method combining the Ostu approach with morphology operations can solve the problem of irregular lumen but the GW and the blood artifacts may cause challenges. Other approaches utilizing Computer vision, like the Canny filter, MRF and level-set-based method, can complete the task of lumen border segmentation but need extra pre-processing or post-processing steps.

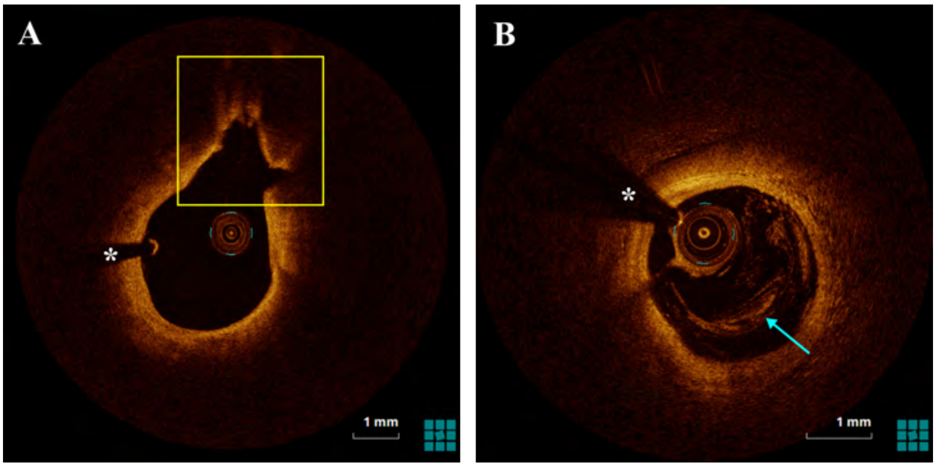


Fig. 3.1 An irregular region denoted with a yellow rectangle in (A) is described as a bifurcation of the IVOCT image, which illustrates the complexity of the lumen boundary. (B) presents the blood artifacts (cyan arrow) captured in the lumen. * denotes GW artifact.

The challenge (2) can be solved through our automated method described in Sec. 2.5, and the blood artifacts can be eliminated with the artifact removing approach depicted in this section. So the important issue is to detect the real boundary of the vessel lumen including the condition of irregularity and complexity. In our lumen border detection method, we converted the Cartesian system to polar domain with the approach denoted in Sec. 2.2 (Fig. 3.2), firstly.

Quantitative measuring the intensity value of each A-line, the obvious characteristic of it is the intensity variance sharply, especially from the lumen to the vessel tissue along the A-line direction. Figure 3.3 shows an example of the intensity distribution of A-line profile referring to the stent and the vessel wall, where the A-line intensity of stent (blue) presents a high value and fall down rapidly followed with a low energy area as the presence of the shadow. The drop in the vascular tissue intensity profile is signified as a continuous descending line (orange) on a larger scale while the intensity profile of the stent's A-line immediately falling down to an approximately stable line after the peak intensity point. Study [82] characterizes each A-line with 4 features: (1) peak intensity, (2) presence of a shadow, (3) length of a shadow and (4) speed at which the energy rises and falls as a function of depth. Obviously, the intensity of the lumen region is much lower than the intensity distribution of the vessel tissue or the stent. Although the energy of the portion of the vessel tissue falls when it far away from the center, the alteration in light intensity is a gradual decrease process without any sudden change if we smooth the A-line intensity (Fig. 3.4). As observed, the fluctuate of each A-line region behind the lumen border point is not strong, while the change of energy rapidly and strongly near the lumen boundary. That is, the intensity fluctuation behind the peak intensity point shows a considerable difference between the tissue and stent strut. To detect the lumen boundary with this feature, we used the standard deviation as the basic statistical value to describe this phenomenon. Of course, lumen border segmentation also utilized the obvious characterization of intensity fluctuation to investigate the intensity variance in both sides of the lumen border.

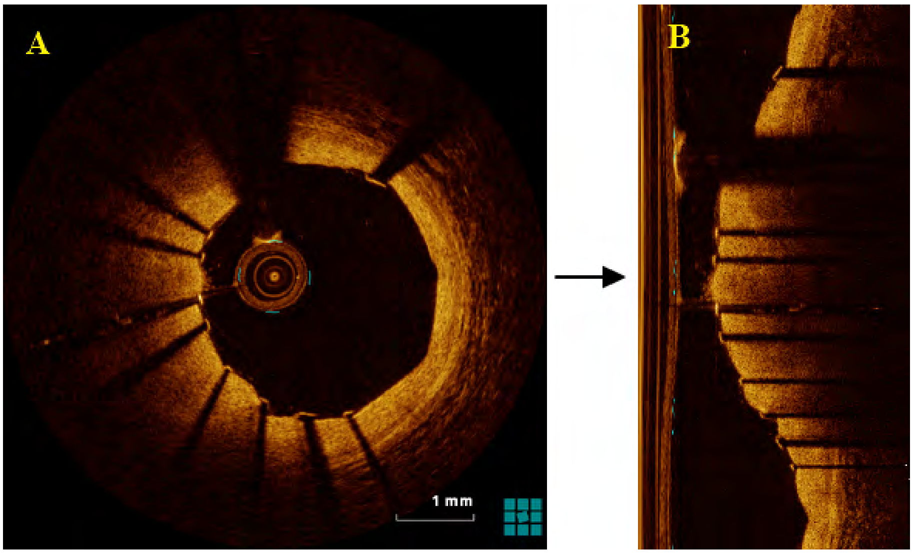


Fig. 3.2 (A) Original IVOCT image, (B) the transformed IVOCT image presented in the polar coordinate system. Each row in (B) is called A-line which is obtained by the catheter detecting reflected light from tissue.

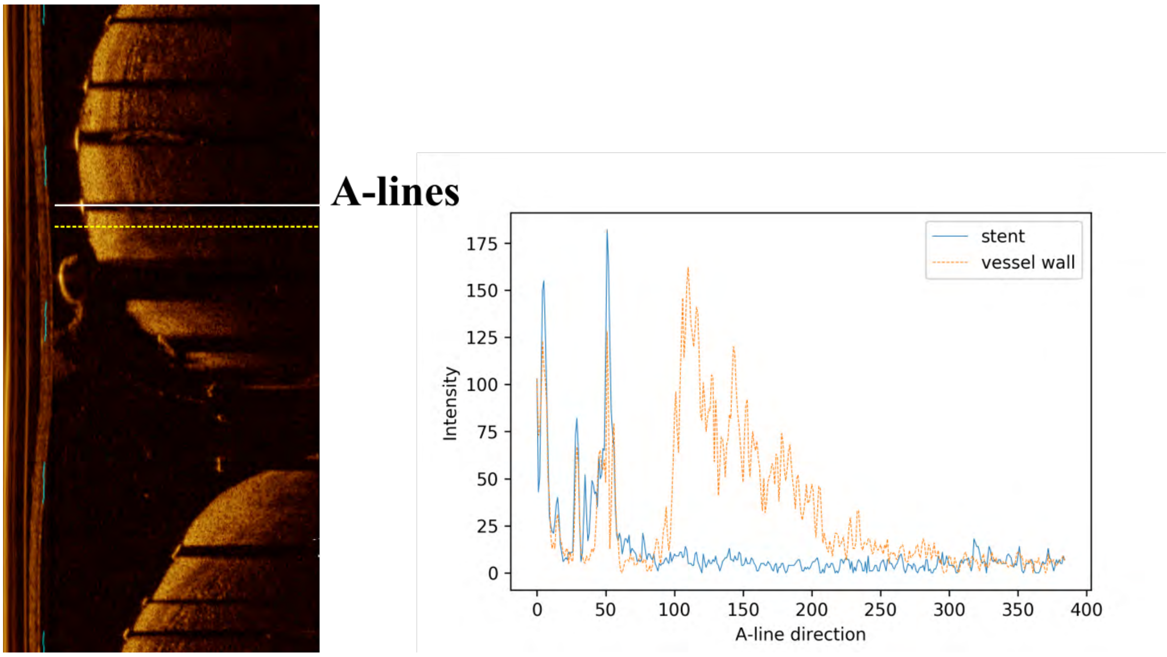


Fig. 3.3 White line and yellow dotted are A-lines (blue and orange) corresponding to the stent and the vessel wall respectively.

Definition: let $F_{st}(i,j)$ be a set of each pixel of the IVOCT image in the polar domain, which is the function of the standard deviation and pixel intensity

$$\begin{aligned} STDI(i,j) &= SD(i,j-n,j)+SD(i,j,j+n) && (3.1)\\ F_{st}(i,j) &= STDI(i,j)+kI(i,j) && (3.2)\end{aligned}$$

where $SD(i,j-n,j)$ and $SD(i,j,j+n)$ are the function of the standard deviation for each pixel at the position of (i,j), the computation range for $SD(i,j_1,j_2)$ is from the j_1-th column to the j_2-th in the i-th row, n is the parameter to define the size of the area around pixel (i,j). $I(i,j)$ is the intensity value of each pixel, and k is a weight parameter to adjust the balance between the standard deviation value and the intensity value.

Let C be the set of the point with the maximum F_{st} of each row (A-line) in the transformed IVOCT image,

$$C=\{C_i\}, \quad i=1,\ldots,R \tag{3.3}$$

where R is the index of rows of the image and

$$C_i=\begin{bmatrix} i \\ \arg\max_j(F_{st}(i,j)) \end{bmatrix}. \tag{3.4}$$

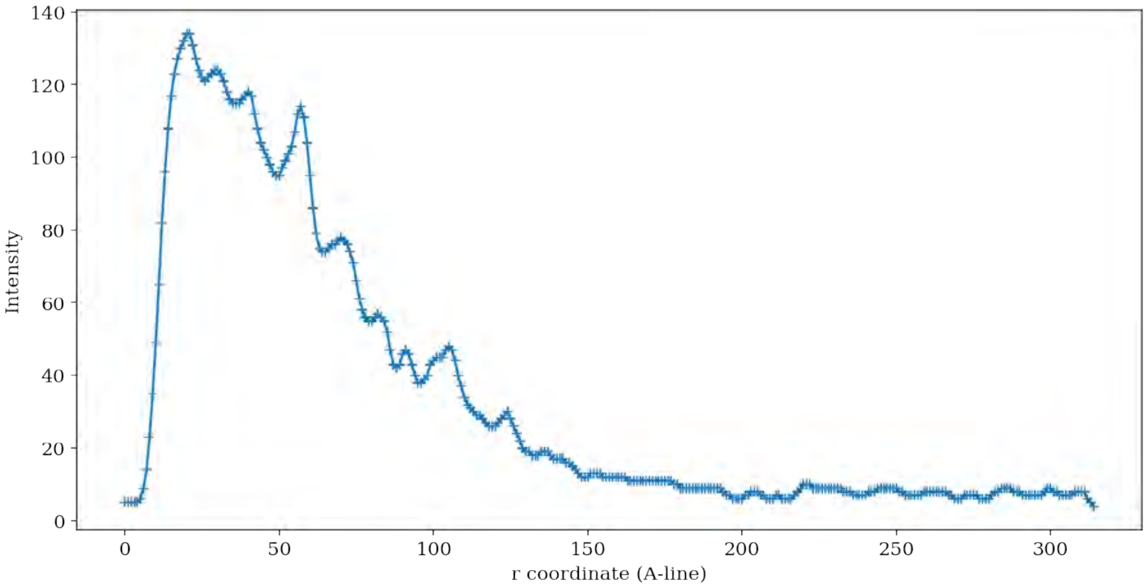


Fig. 3.4 An example of the A-line intensity profile after the smooth filter operation. Each intensity value corresponding to the r (here is the depth of the vessel wall) is denoted with "+".

In statistics, the standard deviation is a measure of the amount of variation or dispersion of a set of values[14]. Our purpose is to use the standard deviation to depict the fluctuation degree of the intensity curve. For each A-line, $STDI(i, j)$ is used to describe the variation of a set of values containing total $2n$ pixels of which the center point is (i, j), that is we obtained $STDI(i, j)$ of each pixel by calculating the standard deviation value of 2n pixels. On one A-line, if the sum of the standard deviation gained from both side of a certain center point (i, j) is larger than the same calculation from other center points, this center point (i, j) indicates a maximum degree of the intensity fluctuation. The pixels on both sides of it may show the lower intensity value and its energy changes drastically. If simultaneously considering the intensity of the current point, we can construct a variable F_{st} to describe this characterization. Actually, F_{st} of some blood artifacts sometimes present high value, which impacts the stents or lumen boundary detection. To eliminate the effect of the blood artifact, the intensity value of point (i, j) multiplied with a coefficient k is added to the $STDI(i, j)$ to remove the possible artifact object. Though Eq. 3.2, the candidate point of each A-line is computed by finding out the (i, j) which can obtain the maximum F_{st} on every row. These candidate points C_i finally form a set C which is used for the next step of lumen border pixels or stent pixels detection.

3.2 Lumen boundary segmentation

As analyzed in the Sec. 3.1, $STDI(i,\, j) = SD(i,\, j-n,\, j) + SD(i,\, j,\, j+n)$ is defined to calculate every pixel in each A-line, and then using Eq. 3.4 detect the candidate pixels containing the maximum F_{st}. In Eq. 3.2, the parameter n is explained as a fixed range for the computation of the standard deviation. Range $[j-n,\, j]$ and $(j,\, j+n]$ are set for the left and right side of the center pixel $(i,\, j)$ to obtain $STDI$. However, to recognize the lumen boundary, we need only to utilize the $STDI(i,\, j)_{left} = SD(i,\, j-n,\, j)$ computed in the left side $[j-n,\, j]$ to get candidate pixels of the lumen boundary according to the feature of big intensity variance from the lumen to the vessel tissue. Before applying the $STDI(i,\, j)_{left}$, we applied the methods described in Chapter 2 for the segmentation catheter imaging and the GW including its shadow region, and then implementing the residual blood removal algorithm to eliminate the blood artifacts with 4 times morphological erosion and 4 times dilation. Figure 3.5 respectively displays the converting result in the polar domain and the border detection result of one original IVOCT image.

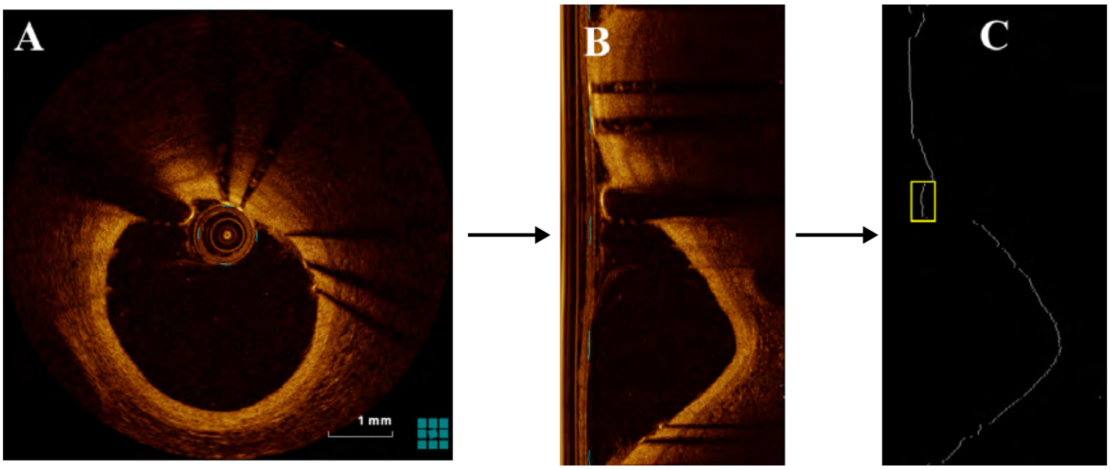


Fig. 3.5 An example of lumen boundary detection by using Eq 3.2 and Eq 3.4 to obtain the initial lumen boundary. Yellow rectangle denotes the border of GW that we did not remove GW in this case firstly.

In order to obtain the best appearance of the lumen border, we chose Savitzky-Golay filter[68] as the smooth method to a set of points for the purpose of getting a smoothed curve. Suppose that the filter window width is $n = 2m+1$, each point $x = (-m,\, -m+1,\, \dots,\, 0,\, 1,\, \dots,\, m-1,\, m)$. Fitting of data points within a window using a $k-1th$ polynomial

$$y = a_0 + a_1 x + a_2 x^2 + \cdots + a_{k-1} x^{k-1}. \tag{3.5}$$

Therefore, n equations with k-element linear group are constructed

$$Y_{(2m+1)\times 1} = X_{(2m+1)\times k} \cdot A_{k\times 1} + E_{(2m+1)\times 1}. \tag{3.6}$$

Then $\widehat{A}$ obtained with least squares as the solution of A is

$$\widehat{A} = (X^T \cdot X)^{-1} \cdot X^T \cdot Y, \tag{3.7}$$

where X is a Vandermonde matrix. The model prediction value of Y can be computed as follows

$$\widehat{Y} = X \cdot A = X \cdot (X^T \cdot X)^{-1} \cdot X^T \cdot Y = B \cdot Y. \tag{3.8}$$

The above formulas give estimates of the smoothed data at the central point of each sub-set. Then, after fitting all the points with moving polynomial, a smoothed curve reflecting the original data is obtained. In this chapter, we used Savitzky-Golay filter to process all the points belong to the lumen boundary to get a smoothed curve, simultaneously, the missing points in the lumen border is also filled with Savitzky-Golay method. Figure 3.6 illustrates an example result (green line) by using the Savitzky-Golay filter to the detected lumen points presented in (r, θ) coordinate. Comparing with the initial segmentation result in Fig. 3.5 (C), the computed result in Fig. 3.6 is obviously smoother than before and the missing portion in GW shadow region is automatically supplemented.

3.3 Stents detection

3.3.1 Stents candidate points detection

A stent is a tiny reticulate tube made of either metal or plastic, which is used to insert a hollow structure (like the artery) of the human body to help keep a blocked artery from closing. Stent implantation as the most common coronary revascularization procedure is implemented for the patients to improve the disease condition or CAD treatment. Approximately 2 million people in the world receive this implantation each year. For the metallic stents, the outstanding characteristic of it in the IVOCT image is strongly reflected light presenting bright reflection blocks followed with black shadows. Figure 3.7 illustrates an unhealthy human vessel with stent struts implemented inside to enlarge the area of the vessel lumen. Due to the metal reflection attribute, every stent strut is followed with a narrow and long shadow correspondingly. Previous studies have discussed many semi-automatic or fully automatic method for stent struts detection utilizing the features of each A-line of stents. In the study[52], the notable features for the classification of struts and shadows include

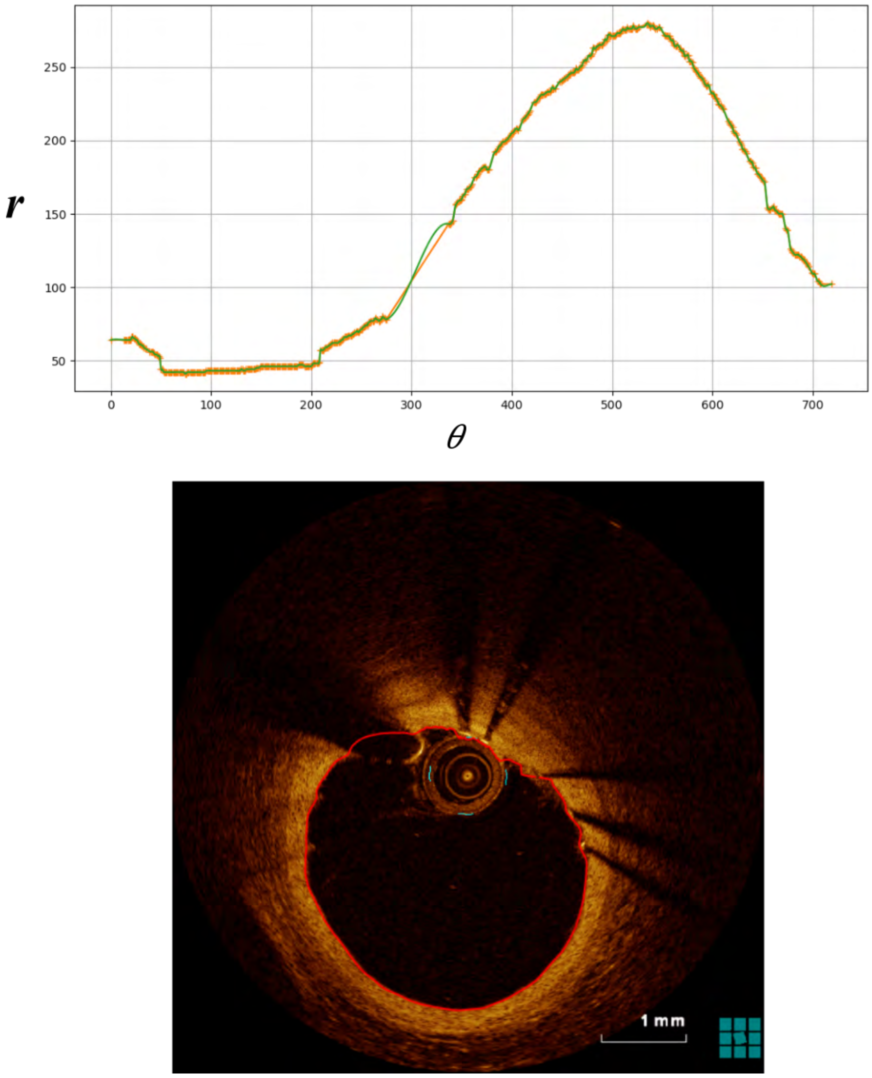


Fig. 3.6 Top: the initial lumen boundary(orange +) is presented in (r, θ) coordinate. The smoothed line (green) overlapping with the initial lumen border was obtained through Savitzky-Golay filter. The prediction position of the missing portion belonging to the GW shadow was computed through this method. Bottom: is the illustration of the smoothed lumen border line (red) in the IVOCT image.

maximum intensity, mean intensity, median intensity, solidity, area, intensity variance, percentage of dark area and mean of dark area. All the analysis features extracted are based on the imaging characteristic of the stent strut. As we described in Sec. 3.1, C_i $(i = 1, \ldots, r)$ is defined to calculate the highest F_{st} for each A-line to obtain the candidate pixels in the IVOCT image. However, in cases of IVOCT images containing bifurcation and residual blood, applying Eq. 3.2 and 3.4 would produce pseudo-candidate points. Simultaneously, the vessel tissue wall also generates extra candidate points. Observing Fig. 3.8, except the reality stents labelled with the yellow rectangle, candidate points in the red circle present the fake points that there are residual blood artifacts satisfying the initial condition designed for the candidate stent points. So, the next task is to remove the fake points from the candidate points obtained with the Eq. 3.2 and Eq. 3.4. The types of fake points contains residual blood pixels (marked with red circle in Fig. 3.8), GW pixels and lumen border pixels. For the residual blood pixels, it usually presents several individual pixels or a small segment continuous line composed with detected pixels. Here, we proposed a method to remove this kind of fake point by judging the neighborhood intensity distribution around the current point. The approach of recognition of lumen border pixels and GW pixels would be respectively described in the following Sec. 3.3.2 and Sec. 3.3.3.

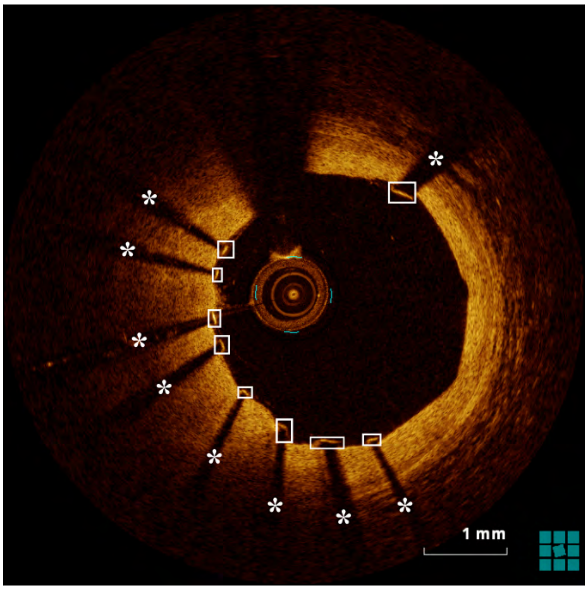


Fig. 3.7 An example of IVOCT image containing implemented stents labelled with white rectangle. Stents shows high bright reflection due to the metallic attribute. Each stent is followed with a dark shadow (*).

3.3.2 Candidate points selection

From Fig. 3.8, we know that C (Eq. 3.3) contains not only the stent points but also other points including vessel tissue, residual blood, etc. Therefore, in this section, our task is to select the real stent points from all C points. The variance of the intensity behind the stent and non-stent along the A-line shows big difference (Fig. 3.9). For a stent spot, there is a dark shadow area adjacent to it (Fig. 3.7, labelled with stars), thus the alteration of intensity belongs to the stent shadow areas should be very small (the part marked with “B” in top of Fig. 3.9). And for any non-stent point, such as the vessel tissue region, the variance for the region adjacent to it (the part marked with “B” in the bottom of Fig. 3.9) is obviously big. The entirety slope tend of region B presents slow down compared with the corresponding region in the top sub-figure of Fig. 3.9, simultaneously, local peak intensity points exist and its intensity profile is an up-and-down motion. The phenomenon described above can be utilized as a feature to distinguish the stent and non-stent regions. Hence, in order to select the stent points from the detected candidate points C, we define two features. The first one is the standard deviation value (V_{ij}) of the whole right part adjacent to each candidate point

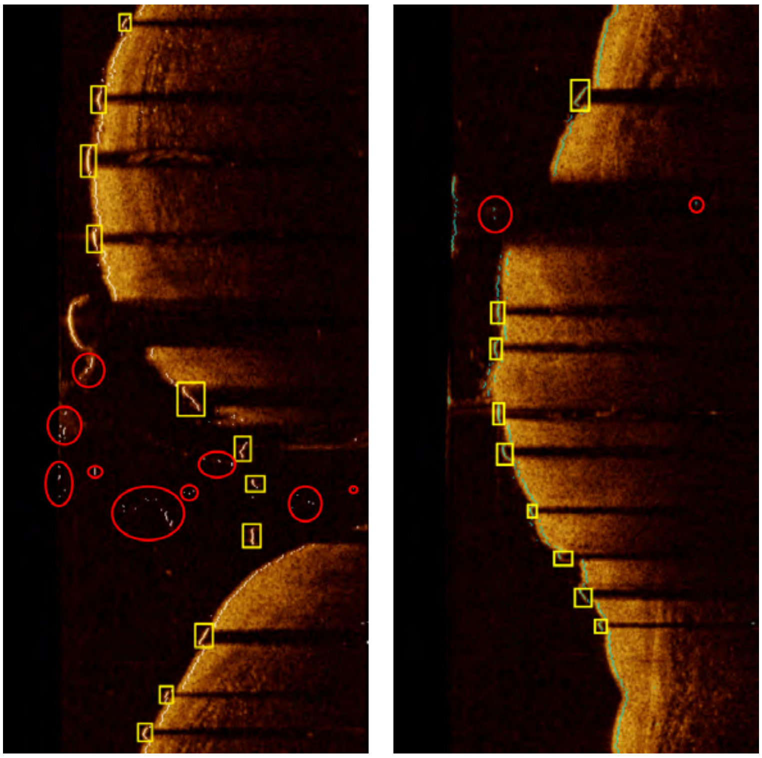


Fig. 3.8 Detecting the candidate points C in IVOCT images. Two types of C are illustrated in the examples, one is the real stent strut pixels (yellow rectangle) and the other is the pseudo-points (red circle).

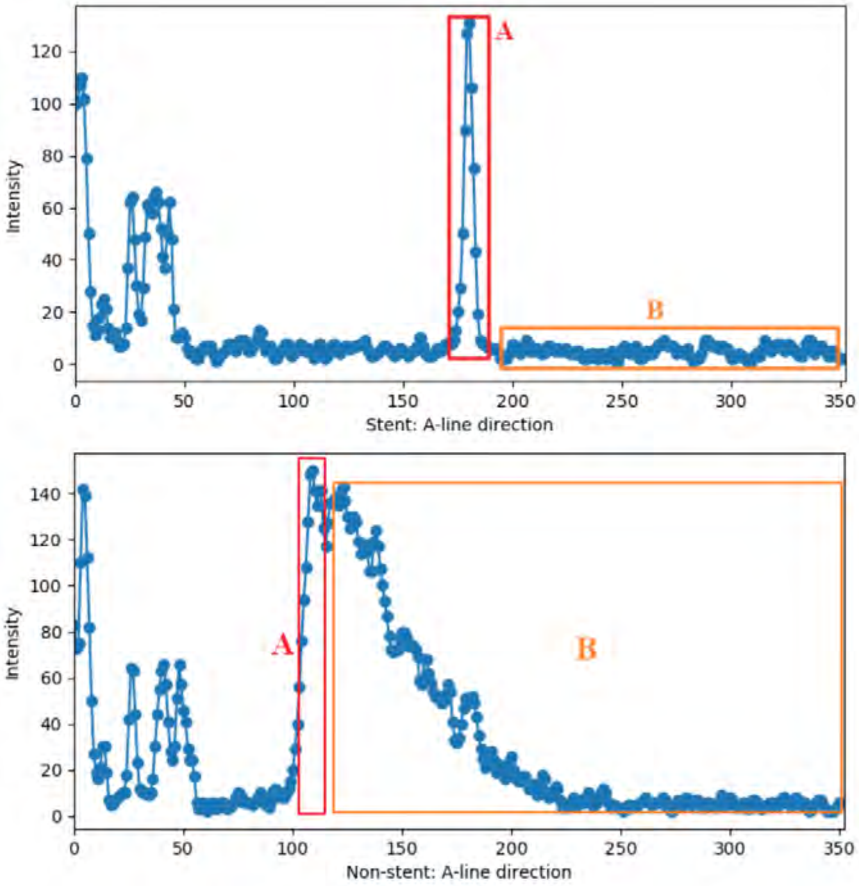


Fig. 3.9 Top: presents an A-line intensity of stent, bottom: illustrates the A-line of the vessel tissue region. A: Brightest pixel (or peak intensity). B: Interested region. The variance of the interested region adjacent to the brightest pixel is significant different for stent and non-stent cases.

C_i, where the calculation range for V_{ij} is from the current point $(i,\, j)$ of C_i to the end of the current A-line.

With the analysis, there is a dark shadow area existing immediately adjacent to a real stent spot, the intensity of the pixels on the right side of a stent pixel in an A-line would drop very rapidly. Based on this consideration, we design the second feature S that measures the relative rapidness of the intensity dropping of the pixels on the right side of each candidate point in position $(i,\, j)$

$$S_{ij} = \frac{W_i}{I_{ij}} \tag{3.9}$$

where I_{ij} is the intensity value of a candidate point C_i, $W_i = j_{stable} - j$ and j_{stable} as the column index is obtained by computing the most appearance of points containing a stable standard deviation value.

According to our pre-experiments, V_{ij} just presents the intensity fluctuation degree in a range from the candidate point to the end boundary of IVOCT image in the polar domain, and S_{ij} measures the relative rapidness of the intensity dropping of the candidate point, it is difficult to select stent points effectively by only using the V_{ij} feature or the S_{ij} feature to select the real stent pixels from the candidate points. These two features are easily affected by the noise, the imaging condition and quality. The V_{ij} of stents often smaller than the tissue's, and the S_{ij} can be used to avoid stent artifacts affecting the stent detection, we choose to merge the V_{ij} and S_{ij}, simultaneously utilizing *log* function to increase the recognition between the stent and tissue. For a better selection of stent points, we combined V_{ij} and S_{ij} to construct a new formula to increase the accuracy of stent detection

$$M_{ij} = -\log\left(\alpha V_{ij} + (1-\alpha)S_{ij} + \varepsilon\right), \tag{3.10}$$

where α is a parameter to adjust the relative weights of V_{ij} and S_{ij}, and ε is a tiny positive number to avoid M becomes an infinity. A threshold, defined as λM_{th}, classifies the candidate points into stent points and non-stent points, where M_{th} is a mean value of the total M values and λ is a coefficient to control M_{th}. For the convenience of data representation and processing, we normalize the value of $\alpha V_{ij} + (1-\alpha)S_{ij}$ for an image to a range of $[0,1]$ before we calculate the M_{ij}.

3.3.3 Stent area detection by adaptive region growing

The stent points detected with the methods described in the previous sub-sections still contain the GW points because the judgement condition for the realistic stent struts is also satisfied by the detected GW pixels. In order to detect stent reliably, we need the information about the shape and the area of stent parts. Here we use the adaptive region growing algorithm (Appendix A.1) to obtain the area generated by the remaining points and compare these areas statistically to discover the GW with the maximum area. Region growing algorithm[5, 8, 38, 67] supposes that all the points in the same region have a similarity of properties. Finding the starting points (initial seeds) is the first step for the region growing. The neighbour pixels around the seeds are then compared with them and if the difference between the neighbours and the seeds are small enough, the neighbour points are classified as the one belonging to the same region and added in Q as the new seeds. This operation is repeatedly applied by using the new seeds and will terminate when no more new seeds can be detected.

Supposing that Q is defined as an initial region for the region growing. Let q be a neighbor point of Q, that is $q \in \overline{Q}$ and $|q-p| \leq 3.0$ stands for all $p \in Q$. Defining a region

$W = \{w; |w-q| < d\}$ where d is a distance parameter. q will be added into Q as a new seed if the following inequality is satisfied.

$$|I(q) - \mu| \leq 2\sigma \tag{3.11}$$

Where $I(q)$ is the intensity of point q, μ is the average value, $\mu = \sum I(q)/N$, and σ is the standard deviation of the set $Q \cap W$, $\sigma = \sqrt{\sum(I(q)-\mu)^2/N}$, N is the pixel number of the set $Q \cap W$. Applying Eq. 3.11 to the detecting points will generate corresponding regions with a certain area. All these regions present homogeneous and similarity with intensity.

3.4 Experimental results

3.4.1 Parameter setting

Three different values (6, 8, 10) of parameter n was tested, and $n = 10$ presented better results and was finally selected for the computation of F_{st}, the coefficient $k = 0.25$. At the pre-processing step of the stent detection, we used a *threshold minimum function*[84] to obtain the minimum threshold at a global level firstly and multiplied it with a coefficient of 1.4 to eliminate small artifacts. Let V_{thv} be the threshold for one row of an IVOCT image for an initial judgement of the tissue points and stent points, simultaneously, W_i is also computed with it. The parameter α and ε were set as 0.3 and 0.01 respectively. The λ in the threshold λM_{th} was set to 1.2 ± 0.2. To gain the best result with our adaptive region growing method, the distance coefficient d defined in Sec. 3.3.3 was tested as 4.

3.4.2 Results

Figure 3.10 shows the procedure of stent detection with our designed features. Figure 3.10 (A) shows $V_{i,j}$ and $S_{i,j}$ of C set gained from an IVOCT image respectively, and Fig. 3.10 (B) is the $M_{i,j}$ value profile for every candidate point. Stents can be identified with the bright spot area constraint after applying the adaptive region growing method (Fig. 3.10 (C)). Figure 3.11 displays example results by applying the automatic lumen border detection method described in Sec. 3.2. The size and morphology of these lumen borders in Fig. 3.11 are different. The last three IVOCT images contain residual blood in the vessel lumen, which might impact the lumen border detection in other studies when they used their detection methods, but our results (inner green lines) can depict the real luminal border and gives out a better performance. From Fig. 3.12, the best situation for stent detection is that the stent meshes the vascular wall closely. three different situations corresponding to different periods of stent position can be observed in the graphics. Especially, in the later stages of stenting, neointima would be generated to cover the implanted stents sometimes. The detection results present that our method can give out a nice performance on this condition.

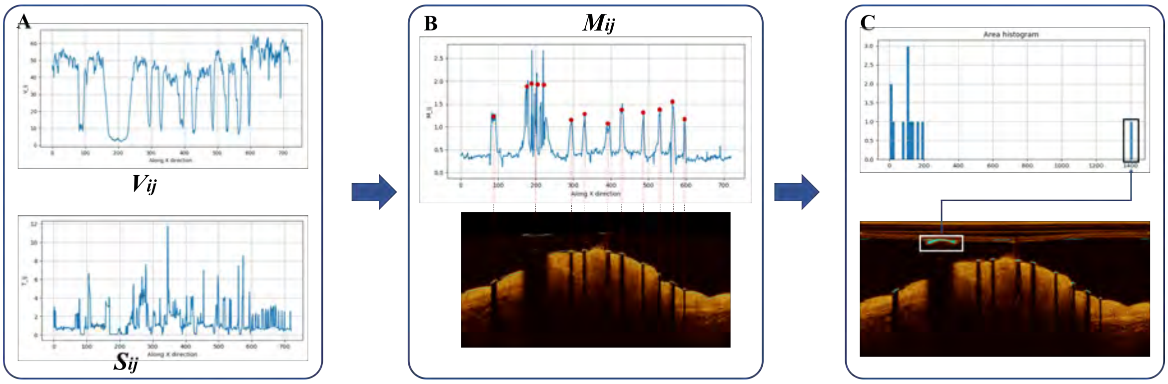


Fig. 3.10 A flowchart shows the procedure of stent detection. The red points (top of B) are the detected stent results with M_{ij} calculated through Eq. 3.10. The bottom subfigure of (B) displays the points (stents and GW) with larger M_{ij} values corresponding to the red points. (A) presents the illustration of V_{ij} and S_{ij}. The final stent detection step can be recognized by employing the adaptive region growing method, and also can remove the GW (Fig. 3.10 (C)).

3.5 Discussion and conclusion

In this chapter, we developed two automated methods based on the local maximum standard deviation for the lumen boundary segmentation and stents detection, which reduces human interaction and time-consuming. We computed the left side of $STDI(i, j)$ ($STDI(i, j)_{left}$) for the lumen boundary segmentation after the catheter and GW removing. Notably, after applying the catheter removal method, the original catheter region would be filled with zero values, therefore, in the polar domain, the start point for F_{st} computation is not the first column of the polar image. Two approaches can be employed to solve this issue: (1) we set an initial start position equal to 45 for each F_{st} calculation in the individual row, (2) make the last zero value belong to the catheter region in each A-line as the start point to compute $F_{st}s$. Additionally, in some complexity condition of the vascular lumen, including deformity morphology lumen and cases containing tiny or small bulge, some portion of the vessel lumen boundary can not be recognized correctly in all with our method. Here, we combined the Ostu method and the morphology operations (erosion and dilation four times, respectively) to obtain an optimized binary result of the vessel tissue presenting high intensity. Sometimes, in the polar system (r, θ), the first point and the last point may be both or respectively lost after employing a candidate luminal points selection method for removing the single points

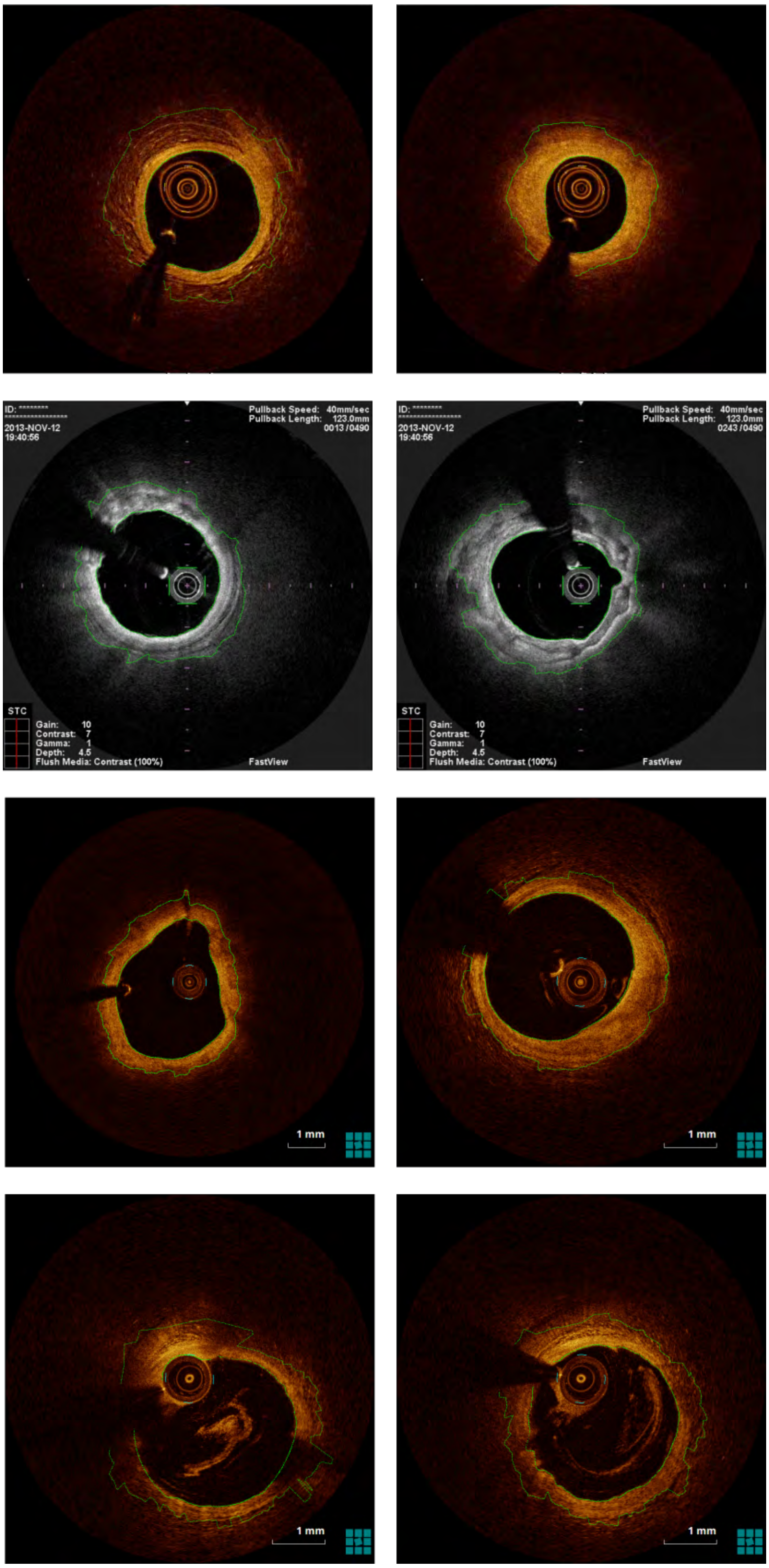


Fig. 3.11 Application examples of lumen boundary detection with our automatic method. The lumen boundaries in the eight IVOCT images with different morphology of lumen are indicated with inner green lines, respectively.

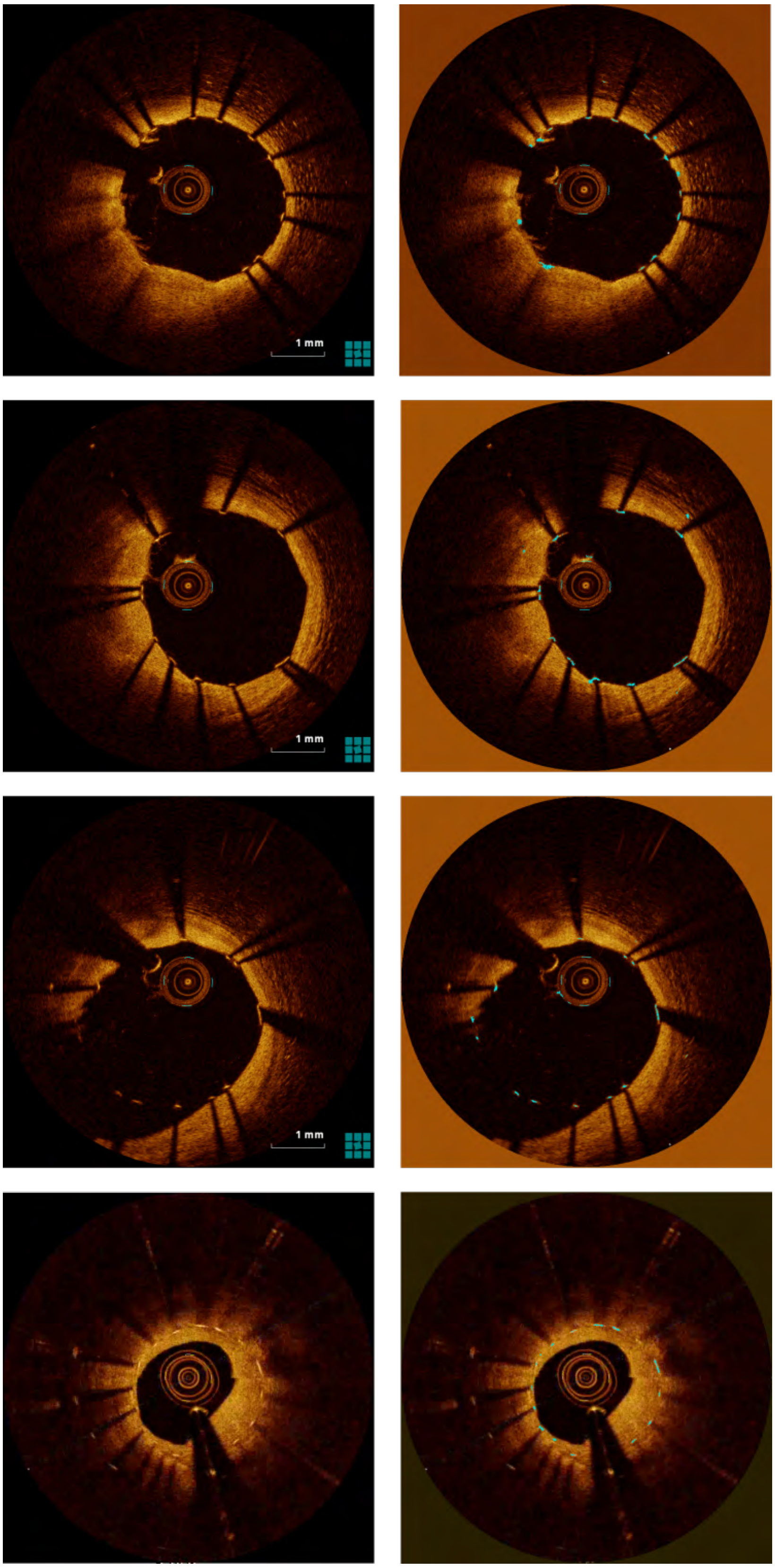


Fig. 3.12 Example IVOCT images that the detected stents (cyan spots) in the second column are marked. Each stent in every IVOCT image can be detected, even with the conditions of the stent malapposition and the neointima (in the third and fourth rows).

which may be the speckle noise or residual blood. Our strategy is to use the point value adjacent to the first or last point instead, or an average value computed with the sum of the first non-zero and the last non-zero of the detected lumen border instead.

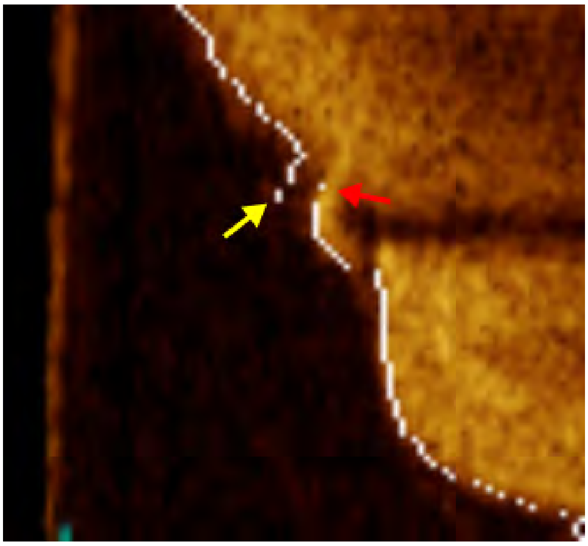


Fig. 3.13 Illustration for the candidate points extraction for the stent detection. A segment of residual pixels (yellow arrow) and an single pixel (red arrow) are observed.

For the stent detection, we apply our method to recognize stents from candidate seeds with M value that combines two defined stent features, and then use an adaptive region growing algorithm to segment the stent in IVOCT images. The experiment presents that our method can also give a good result for the stent detection with neointimal hyperplasia coverage in the IVOCT image. There are still existing a lot of problems with stent detection in IVOCT image, currently. Luminal residual blood, catheter close to the lumen wall in some images, the spot noise, indistinct stent and shadow, extra stents existing in the shadow region are considerable reasons that complicate the stent detection. In addition, stent detection in the case of a black shadow without a stent nearby still takes a big challenge for researchers especially in some IVOCT images with NIH cases.

Finally, utilizing the adaptive region growing in the stent detection can obtain the area histogram of all the final bright spots. Generally, the GW area much larger than other bright reflection spots (stents). Therefore, our detection method of the bright refection spot combined with the adaptive region growing can be indirectly employed to segment the GW as well, as shown in Fig. 3.14. The rest bright spots of which area is less than 200 are recognized as the stents except the individual one (marked with the red rectangle) of which area is significantly larger than the others.

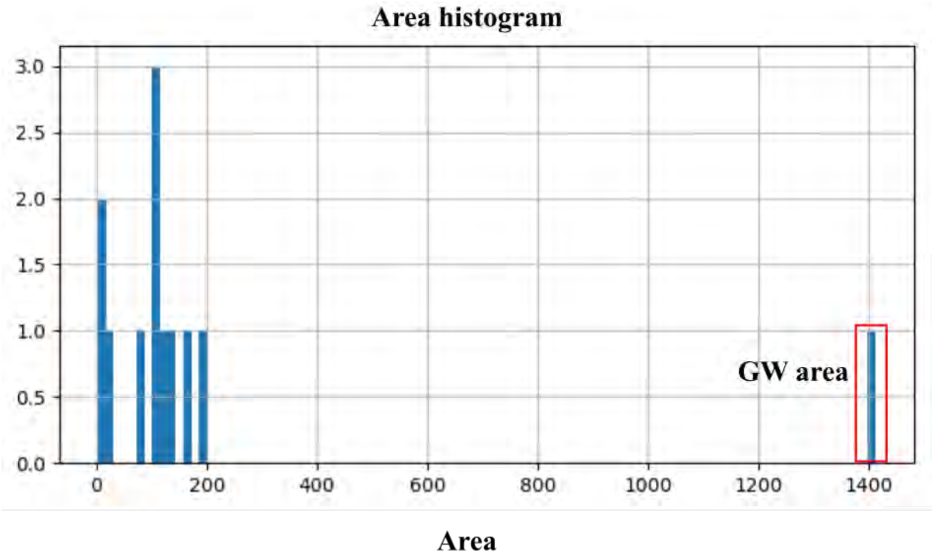


Fig. 3.14 Applying the adaptive region growing method to the candidate points, GW is measured with the biggest area, generally. The data in the red rectangle illustrates this feature which can be used as another method to segment the bright reflection region of GW.

Overall, in this chapter, we proposed an automatic method for the lumen boundary segmentation and stent struct detection. Our method can rapidly finish the above task with high accuracy. Observing the stent detection, our method can recognize certain stents in NIH cases, but for a certain situation, the detection of stents covered with NIH may be hard. The next step for stent detection is to improve the accuracy under the condition of covering NIH and increase the capability of lumen boundary segmentation in complexity cases (diffusive lumen border).

Chapter 4

Atherosclerosis plaque recognition

PCI with the stent strut medication as a progressive treatment indeed can relieve or cure the symptom of patients who catch CAD. Recognizing and measuring the location, angle, categories of the atherosclerosis plaque is the elementary and significant procedure before applying PCI, which provides beneficial information to specialists during the following treatment step, such as vessel inner examination and stent implantation. However, the manual qualitative analysis job of classification for the vessel lesion tissue and tracing of plaque components is time-consuming to doctors because a single-time IVOCT data set of a patient usually contains hundreds of *in-vivo* vessel images with respect to one segment vascular. Identifying and classifying rapidly and accurately atherosclerosis plaques and obtaining its relative measurement data instead of manual work are challenge tasks. As such, automated methods for IVOCT tissue characterization is necessary for the cardiovascular research and clinical practice. In this chapter, I focus on the need for developing an automated tissue identification method, and we will discuss the tissue feature definition and extraction of the vessel lesion.

4.1 Feature analysis

IVOCT images are generated in the computer terminal by receiving the OCT signal $\langle I_d(r) \rangle$ from the vessel tissue and present sorts of tissue imaging according to its different attenuation coefficients. In a rotary-style OCT scanning procedure, due to the vessel tissue scatter and attenuation, the ability of light-absorbing of the vessel wall is increasing as the tissue depth increment[94]. The relationship between the OCT signal $\langle I_d(r) \rangle$ and the depth r is simply illustrated in Eq. 2.1 of Sec. 2.1. Considering the movement pattern of GW (pull-back) and the imaging modality of rotary scanning of OCT light transmitted by the catheter, the IVOCT image presents the appearance of the vessel tissues is continuous in the distribution along the radial and circumferential directions, respectively. As such, an obvious phenomenon can be observed in IVOCT images. First, the vessel tissue (including the lesion plaques) appears significantly different due to the existence of the tissue attenuation coefficient μ presenting difference. Different types of vessel tissue can be displayed clearly in the OCT image with the optical coherence technique. Secondly, the ability of light penetration gradually declines as the depth increases in the radial direction. Only the superficial layer of the IVOCT image can be clearly observed, while the most rest of the region manifest dark area containing useless information. Third, mixed plaques increase the difficulty of lesion plaques recognition. Tissue stratification would occur in the radial direction, that is the different tissues would appear in one radial direction. Except for the center high-light region of the IVOCT image, the rest surrounding areas of the vessel imaging present poor signal and low contrast. The superficial tissue near the lumen boundary maybe consist of different types of tissue but its distribution is continuous in a limited region. According to the human vessel image acquisition procedure described in Sec. 2.1, the catheter scans the vessel with continuous angles to form a vessel OCT image. That is the tissue imaging presents continuity in the circumferential dimension and the contrast among different tissues is obvious. Figure 4.1 displays an example of the similarity and contrast comparison of the intensity along the radial and circumferential direction by using different colors in the superficial layer of the vessel wall. Based on the statistical tissue thickness data indicated in Tab. 4.1, we used 8 pixels as the basic depth for each layer and 24 pixels as the unit step length to partition regions in the circumferential direction to define an interesting region to compare the variance of intensity and texture of adjacent areas. Histologically, the intensity characteristic of tissues on the same layer having a fixed thickness presents homogeneous and coherent. Additionally, tissue stratification would occur in the radial direction, that is the different tissues would appear in one radial direction. That is the intensity distribution has a gradient property in the radial direction and certain continuous in the circumferential direction because of the heterogeneous of the vessel lesion tissue. Notably, Except for the center high-light region of the IVOCT image, the rest surrounding areas of the vessel imaging present poor signal and low contrast, which is worthless to the lesion plaque research.

Furthermore, several researches[13, 28, 43, 93] have proved that lipid plaques associate with a high attenuation coefficient or appear in signal-poor regions, the fibrous plaque has a low attenuation coefficient or is in signal-rich regions, and the calcified plaque presents low attenuation coefficient but with sharp borders. Soest et al.[86] investigated that the attenuation coefficient (μ_t) of the healthy vessel wall and fibrous plaque were 2–5 mm^{-1}, calcified plaque $\mu_t \approx 6 \pm 1.0 mm^{-1}$, lipid tissue $\mu_t \geqslant 10 mm^{-1}$. Observing the Tab. 4.1, the thickness of the healthy vessel wall and the calcified plaque are in a limited range while the lipid plaque

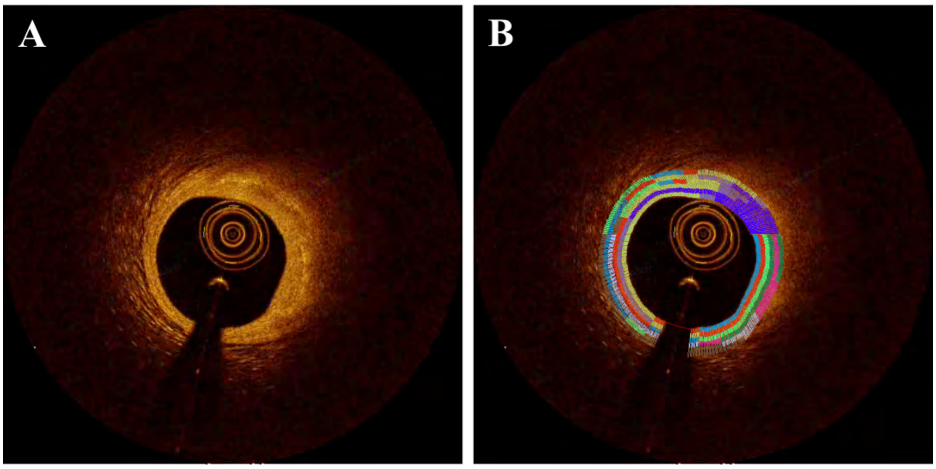


Fig. 4.1 (A) is the original IVOCT image, and (B) presents the characterization of the intensity similarity and variation in the superficial tissue nearby the lumen boundary.

occurs in a dynamic range due to its diffuse attribute. The fibrous plaque is classed into three categories based on its thickness measurement. Several studies have investigated to utilize the intensity profile of A-line for the tissue or lesion plaque classification analysis. The intensity change in the A-line profile indirectly or directly reflects the absorption and scattering attributes of different tissues. However, analyzing the vessel lesion plaque identification only with the intensity profile can not effectively and completely solve the tasks of atherosclerosis plaque identification, classification and quantitative measurement. Some papers combined the A-line profile with light attenuation for the feature analysis of the vessel tissue, such as Ughi et al.[83] used a rectangular window to iteratively fit the OCT A-lines for different k values. At every k, all the possible fits were calculated, and the best fitting curve was selected as the indicator μ_t of the corresponding tissue. Although the single A-line as the direct and important analysis object to extract significant features for vessel tissue quantitative measurement, note that a single A-line only contains 1-D information to present the vessel tissue appearance with OCT technology, e.g. intensity profile or attenuation degree, can't provide extra information for the lesion plaque identification and classification. Additionally, according to the A-line characteristic analysis of per type of the vessel tissue and quantitative thickness measurement researches of different tissues (illustrated in Tab. 4.1) in the earlier literatures, the proximity luminal boundary region (PLBR) of a vessel contains more useful morphology information about the vessel tissue (health vessel wall or lesion plaques). The relationship between adjacent A-lines will not be indicated and the local region con-texture information of the vessel tissue will also be lost simultaneously if only considering using methods based on the single A-line along the radius direction for the tissue classification analysis. Undoubtedly, the information of the 2-D region containing multiple adjacent A-lines would give out more benefits for tissue identification and recognition. Obviously, the texture and intensity relativity in a set of adjoining A-lines can be easily detected with the existing graph image analysis method. Meanwhile, extra features could be discovered and analyzed to help the lesion plaque examination in clinical research.

Table 4.1 Appearance of tissue thickness measurement in the previous studies

Tissue type	**Thickness measurement**
Healthy vessel wall	$< 0.3mm$[86]
Fibrous plaque	Three categories[88]:
	$< 0.065mm$;
	0.065–$0.15mm$;
	$> 0.15mm$
Lipid plaque	$3.9 \pm 2.1mm$[28]
Calcified plaque	$< 0.7mm$[91]

4.2 Multi-layer model

Thus, in order to utilize the local A-lines texture information in radial and circumferential dimensions, we build a multi-layer model which divide the analyzable area near the inner lumen boundary into several layers with different depths for the tissue classification investigation. Based on the definition of local multi-layer shown in Fig. 4.2, an A-line sub-region (ASLR) is developed as a basic 2-D region unit to extract the regionally statistical features of multiple adjacent A-lines. Comparing with the 1-D information of a single A-line, an ASLR with 2-D contains more local region information than the former, such as tissue texture and homogeneous characteristic, etc. An ASLR composed of multiple adjacent A-lines can satisfy the depiction of the A-line relative relationship and 2-D local information continuity. In our local multi-layer model, an individual ASLR of each layer is remarked with a superscript for distinguishing and the ASLR in each layer is denoted as $ALSR^1$, $ALSR^2$ and $ALSR^3$ for describing the three ASLRs in different layers respectively, suppose we use three layers in our model, which is shown in Fig. 4.2. The size of the three mentioned ALSRs corresponding to the three layers is $[l_w \times l_h^1]$, $[l_w \times l_h^2]$ and $[l_{wh} \times l_h^3]$ respectively, where l_w indicates the height of ALSR and l_h^j, $(j = 1, 2, 3)$ is the width of ALSR. A joint region consisting of $ALSR_i^1$, $ALSR_i^2$ and $ALSR_i^3$ is presented as $ALSR_i^{123}$ ($[l_w \times (l_h^1 + l_h^2 + l_h^3)]$) to denote the entirety ALSR with the index of i, where i is the ALSRs number depending on the value of $length_{LB}/l_w$, where $length_{LB}$ is the length of the vessel lumen boundary. With the multi-layer model definition, a PLBR in an individual IVOCT image can be divided into several $ALSR^{123}$s in the circumferential dimension to characterize the circumferential continuity and difference of tissues, and each $ALSR^{123}s$ composed of three ALSRs to describe the hierarchical properties of tissues. Figure 4.2 illustrates the schematic of ALSR applying to PLBR and the hierarchical structures for ALSRs with the defined parameters l_w, l_h^1, l_h^2 and l_h^3. In our model, I define three ALSR with three different sizes of l_h^i by according to the property that the different tissues present different thicknesses in PLBR. (Tab. 4.1).

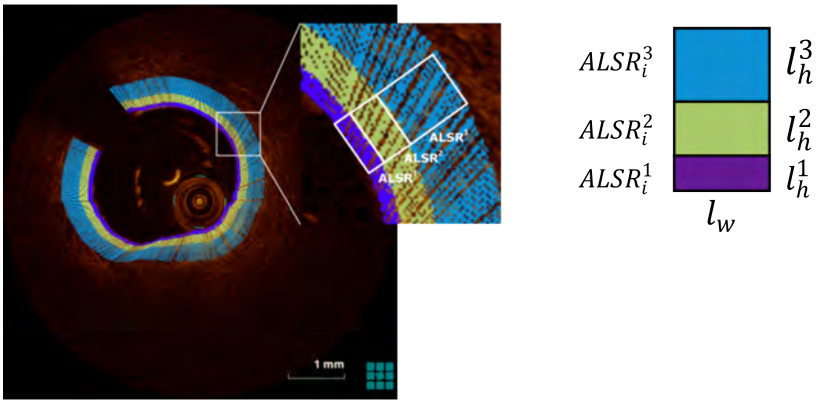


Fig. 4.2 The PLBR is divided into 3 layers along the radial direction by considering the light attenuation property of different vessel tissues. A individual $ALSR^m$ is defined in each layer for the 2-D local features extraction.

4.3 Feature extraction

Through the definition of the local multi-layer model, the following characteristic descriptors are investigated and applied to extract features from an ALSR of each layer of PLBR. We move counterclockwise the ALSR with a specific stride in the circumferential dimension to obtain a local A-line cluster, and then extract features based on this local multi-layer. Let $ALSR_i^m$ denotes the i-th ALSR of the m-th layer.

Kullback-Leibler Divergence

Formally, the Kullback-Leibler divergence (KL divergence)[44] is a measure of the asymmetry of the difference between two probability distributions P and Q. For discrete probability distribution P and Q, its KL divergence is defined as

$$D_{KL}(P \parallel Q) = -\sum_i P(i) ln \frac{Q(i)}{P(i)}, \tag{4.1}$$

or

$$D_{KL}(P \parallel Q) = \sum_i P(i) ln \frac{P(i)}{Q(i)}, \tag{4.2}$$

where the KL divergence is valid only if for all x, $Q(x) > 0$ and $P(x) > 0$, simultaneously the sum of probability P and Q respectively equal to 1.

Here, we use the histogram of an ALSR to denote its intensity probability distribution. The similarity of two regions (A and B, which are selected from $ALSRs$ belong to one $ALSR_i^m$) is calculated with the symmetric form of the KL divergence presented as follows:

$$D_s = \frac{D_{KL}(P_A, P_B) + D_{KL}(P_B, P_A)}{2} \tag{4.3}$$

where $D_{KL}(P_A, P_B) = \sum P_A log(P_A/P_B)$, P_A and P_B both are the discrete probability distributions of A and B. Therefore, KL divergences are calculated among the three ALSRs ($ASLR_i^1$, $ASLR_i^2$ and $ASLR_i^3$) which have the same polar angle along the radial direction. To measure the relative attribute of any pairwise from the $ASLR_i^{123}$, three KL values [D_s^{12}, D_s^{13}, D_s^{23}] are gained through pairwise calculation.

Radial direction intensity difference

The intensities among ALSRs corresponding to regions of vessel lesions tissues show obvious intensity difference in the radial dimension because of the variance of light attenuation of each tissue Visible and valid region with a certain thickness demonstrate the above characterization in Fig. 4.1 (B). I construct an assessment factor to statistically describe the radial dimension intensity difference. This measurement expresses the sum of intensity difference of ALSRs in the same polar angle direction. It is defined as:

$$RDID_i = \sum_{m=1}^{M} \frac{|e_i^{mp} - e_i^{mq}|}{e_i^{mp} + e_i^{mq}} \tag{4.4}$$

where $e_i^{mp} = \|V_i^m - \mu_i^p\|$, $e_i^{mq} = \|V_i^m - \mu_i^q\|$, V_i^m is the 2-D pixel intensity matrix of $ALSR_i^m$ that is the *i-th* ALSR in the *m-th* layer, μ_i^p and μ_i^q respectively indicate the average intensity of $ALSR_i^p$ and $ALSR_i^q$. The m, p, $q \in \{1, 2, \ldots, M\}$ are used to denote the layer index , M is the maximum layer number and $m \neq p \neq q$. Radial direction intensity difference (RDID) can measure the statistic information of the intensity variance of ALSRs in the same angle direction to characterize the local adjacent A-lines intensity alteration.

Accumulated circumference difference

Observing from Fig. 4.1, at the same layer, ALSRs with the same tissue class have similar features and appearance while the texture characteristics between different tissues exist in obvious contrast. That is, if measuring the distribution of vascular tissues in the circumferential domain, the angle range used for the measurement of tissues is a significant quantitative characterization. In a homogeneous tissue, the intensity variation circumferentially of ALSRs presents less difference, but for the different tissues, the variance of these ALSRs appear outstanding. To express this statistical character, an accumulated intensity difference of $ALSR_i^m$ in the mth layer is designed:

$$ACD_i^m = \sum_{j=0}^{K-1} \frac{|\mu_i^m - \mu_{i,j}^m|}{\sigma_{i,j}^m} \tag{4.5}$$

where μ_i^m and $\mu_{i,j}^m$ are the average intensity of the $ALSR_i^m$ and $ALSR_{i,j}^m$ respectively. $ALSR_{i,j}^m$ is a joint region composing of $ALSR_i^m$ and $ALSR_j^m$ in the mth layer. $\sigma_{i,j}^m$ is the standard deviation of $ALSR_{i,j}^m$. Parameter K is the total number of ALSRs of a single layer.

Depth of region of interest (ROI)

The aforementioned content[11, 18, 28, 86, 88, 91, 96] discussed that lipid, fibrous and calcified plaques contain different penetration depths and attenuation coefficients when utilizing the catheter to capture the vessel inner structure situation. Therefore, the ROI defined in an IVOCT image should contain almost all useful information about the vessel tissues. In this paper, we applied the Chan-Vese-level-set method[20] to obtain an energy dividing line as the outer boundary of the ROI. Corresponding to the number of the ALSRs of a single layer, ROI is split into K parts, computing the average distance between the points of lumen boundary and the corresponding points on the outer boundary for each ALSR part as the distance feature of the ROI. The set of distance D is defined:

$$D_{ROI} = \{\overline{d}_k | k = 1, 2, \ldots, K\} \tag{4.6}$$

where $\overline{d}_k$ is the average distance of kth part of the ROI.

Gray level co-occurrence matrix

The texture filter functions provide a statistical view of texture based on the image histogram. These functions can provide useful information about the texture of an image but cannot provide information about shape, i.e., the spatial relationships of pixels in an image. Gray level co-occurrence matrix (GLCM)[31] calculates how often a pixel with the gray value i occurs in a specific spatial relationship (based on the angle θ and distance d) to a pixel with the value j. To a $m \times n$ image I that its gray level is supposed as G (the gray range of $[1, G]$), let C be a matrix $N \times N$, and $C_{\Delta x,\Delta y}(i, j)$ denotes that the number of times in I of the adjacent pixels i and j, the adjacent relation is given by Δx and Δy, $C_{\Delta x,\Delta y}(i, j)$ is defined as the following formula:

$$C_{\Delta x,\Delta y}(i, j) = \sum_{p=1}^{n}\sum_{q=1}^{m} \begin{cases} 1, & if\, I(p,q) = i\, and\, I(p+\Delta x, q+\Delta y) = j, \\ 0, & otherwise. \end{cases} \tag{4.7}$$

For an individual ALSR unit, statistical measures of GLCM could be calculated as the features of ALSR. Here, we utilized 4 statistical features considered for the analysis of image texture information. The properties are computed as follows:

Dissimilarity:

$$f_{dissimilarity} = \sum_{i,j=0}^{N-1} C_{i,j}|i-j| \tag{4.8}$$

Homogeneity:

$$f_{homogeneity} = \sum_{i,j=0}^{N-1} \frac{C_{i,j}}{1+(i-j)^2} \tag{4.9}$$

Energy:

$$f_{energy} = \sum_{i,j=0}^{N-1} C_{i,j}^2 \tag{4.10}$$

Correlation:

$$f_{correlation} = \sum_{i,j=0}^{N-1} C_{i,j}\left[\frac{(i-\mu_i)(j-\mu_j)}{\sqrt{\sigma_i^2\sigma_j^2}}\right] \tag{4.11}$$

where $C_{i,j}$ is the co-occurrence matrix, i and j are the labels of the columns and rows of the GLCM. μ is the mean and σ is the standard deviation. In this paper, we tested and chose the distance $d = 3$ and angle θ = [0, 90, 180, 270 deg] of GLCM, 16 texture features for each $ALSR_i^m$ are selected as its features.

Overall, 22 features totally generating 56 data are calculated and utilized to describe the texture, intensity contrast and homogeneous information of $ALSR^1$, $ALSR^2$, $ALSR^3$ and a joint region $ALSR^{123}$.

4.4 Experimental results

4.4.1 Parameter setting

The layers number of the local multi-layer model was set as $M = 3$ by considering the thickness of each tissue mentioned in Tab. 4.1. According to the previous studies and the description in Tab. 4.1, $[l_w \times l_h^1]$, $[l_w \times l_h^2]$ and $[l_w \times l_h^3]$are the spatial size of $ALSR^1$, $ALSR^2$ and $ALSR^3$ respectively, where $l_h^1 = 8$ pixels ($\approx 79\mu m$), $l_h^2 =$16 pixels ($\approx 158\mu m$) and $l_h^3 =$32 pixels ($\approx 317\mu m$) and $l_w = 24$ pixels ($\approx 238\mu m$). The sum of three layer depth is $l_h^1 + l_h^2 + l_h^3 = 554\mu m$, which can satisfy the thickness range from $150\mu m$ to $700\mu m$ for the four types of vessel tissue (the thickness of calcified plaque $< 700\mu m$). In the depth of ROI D_{ROI} computation, due to the thicknesses of the vessel wall in one individual IVOCT image or among IVOCT images are inhomogeneous, directly calculating D_{ROI} may cause errors with out of range, therefore, we multiplied the original depth with a coefficient $\alpha = 0.8$ for a dynamical alteration. The detail parameter setting for the outer border of ROI is explained in the Sec. 6.4.2 of Chapter 6. In the features of GLCM calculating, the gray level G was set to 255.

4.4.2 Validation

The vessel datasets containing 490 images were processed to totally generate 14399 samples based on our local multi-layer model and ASLR definition. In this chapter, an entire sample in our tissue classification method was the combined region consisting of $ALSR^1$, $ALSR^2$ and $ALSR^3$. We used 5 numbers (0, 1, 2, 3 and 4) to denote 5 classification categories with respect to the healthy vessel wall, lipid plaque, fibrous plaque, calcified plaque and other non-tissue objects (residual GW artifact or small bifurcation region). Each sample was labelled as one of the 5 categories to manifest the tissue type of the current $ALSR^{123}$ in PLBR. To assess the generalization of our method, the dataset was split into 2 parts: training set (67%) and testing set (33%).

According to the literatures[11, 83], we chose Random Forests (RF) as the tissue estimator to accomplish the tissue classification task. RF consists of a number of single decision trees to estimate the final classification result by voting the class prediction obtained from each individual tree in RF. We tested the number of trees of RF from 30 to 100 respectively and finally selected the number trees of RF as 100 for a higher accuracy score (1–2% improvement). To evaluate the efficiency of the classifier, a cross validation strategy of machine learning named KFold was implemented to the training data. The number of KFold was tested from 5 to 20 and finally set as 10. Classification accuracy was calculated by comparing the prediction results with the expert manual analysis. Here, we used the following formulas, statistically analyzing the relation of TP (true positives), TN (true negatives), FP (false positives) and FN (false negatives), to illustrate the accuracy of classification results through our method. Recall, true positive rate, is an assessment of the percentage of correct results in all true results. Precision indicates the percentage of correct results in predict results. ACC denotes the overall accuracy of the detecting objects.

$$Recall = \frac{TP}{TP+FN} \tag{4.12}$$

$$Precision = \frac{TP}{TP+FP} \tag{4.13}$$

$$ACC = \frac{TP+TN}{TP+FP+FN+TN} \tag{4.14}$$

4.4.3 Results

Table 4.3 presents the results of the classification statistic for 7 datasets measured through recall, precision and accuracy metrics. Five classes, including healthy vessel wall, lipid plaque, fibrous plaque, calcified plaque and non-tissue object, were used to as the tissue categories for the tissue classification with our local multi-layer model. Notably, not all the datasets in our experiments contain all types of tissue, we used "-" to denote the non-existing classes, as shown in Tab. 4.3. The tissue classification results are illustrated in Fig. 4.3, four colors (dark turquoise, yellow, green and white) are used to code the corresponding tissue type for presenting the tissue occurring range in the circumference dimension. Comparing the predictions by using our method with the ground truth by manual analysis, the results present that our approach can be developed to a considerable assist tool for lesion tissue recognition and classification.

4.5 Discussion and conclusion

Local multi-layer model defined in our method is applied to investigate the tissue characteristics of the superficial layer region (in our paper, depth = $554\,\mu m$) of the inner vessel wall, namely, PLBR. We chose the layer number as 3 by according to the light attenuation shown in Tab. 4.1 and intensity variance in the A-line profile. Furthermore, to gain more

Table 4.2 Lesion classification accuracy comparison between Ughi' method and our method

Method	**Tissue Categories**		
	Fibrous	Lipid	Calcified
Ughi's method	89.50%	79.5%	72.1 %
Our method	89.53%	91.78%	93.81%

tissue information and compare the characteristics difference of various tissues in the radial dimension, 3 dissimilar depth for the 3 layers were determined. ALSR was employed to extract the 2-D region features with a specific size instead of a single A-line method. We used the ALSR as a unit to extract features from the PLBR in radial and circumferential dimension for the tissue classification of the superficial layer region. Along the circumferential direction, we obtained a distribution range of a certain type of one tissue. To our knowledge, this is the first time to define the local multi-layer for quantitative analysis of the human vessel lesion tissue classification. Comparing with Ughi's[83] method based on the single A-line feature research, our method improves the accuracy of the lipid and calcified classification, which is shown in Tab 4.2. Athanasiou et al.[11] implemented the lesion tissue classification within the pixel-level and their method presented a precision of 70% with the depth of $0.625\,mm$, while the mean precision of our method in the 3-layer model was 83.45%. Despite our model achieves a better result, the limitations still remain. Validation shown in Tab. 4.3 presents a low score results for the recall metric of lipid plaque (in set 0, 1, 2, 6) and fibrous plaque (in set 1, 4). Some examples of tissue misclassification are illustrated in Fig. 4.5, and the confusion matrix about the classification results from one data set, which is shown in Fig. 4.4, demonstrates the possible error classification. The factors causing the misclassification of the lesion tissue as follows: (1) Helicoidal image data acquisition and the pull-back imaging form sometimes generate a set of blur IVOCT image frames. (2) Another factor affecting the classification results is the lipid plaque contains a diffuse border. As discussed in the literature[88], the true boundary of the fibrous plaque is not always clear and presented obviously. (3) Additionally, the location of GW in the human vessel, the noise generated by PSF (point spread function), the OCT images with blurring and low resolution are also the reasons for the error classification. The ROI depth of the lipid plaque, which extracted with the Chan-Vese-level-set method, might present a great difference in some continuous IVOCT frames, even existing in one IVOCT image. In our classification strategy, the principal component of a region will be considered as the basis for categorization labelling, e.g., the lipid region containing a thin fibrous cap was still determined as a lipid type. In this chapter,

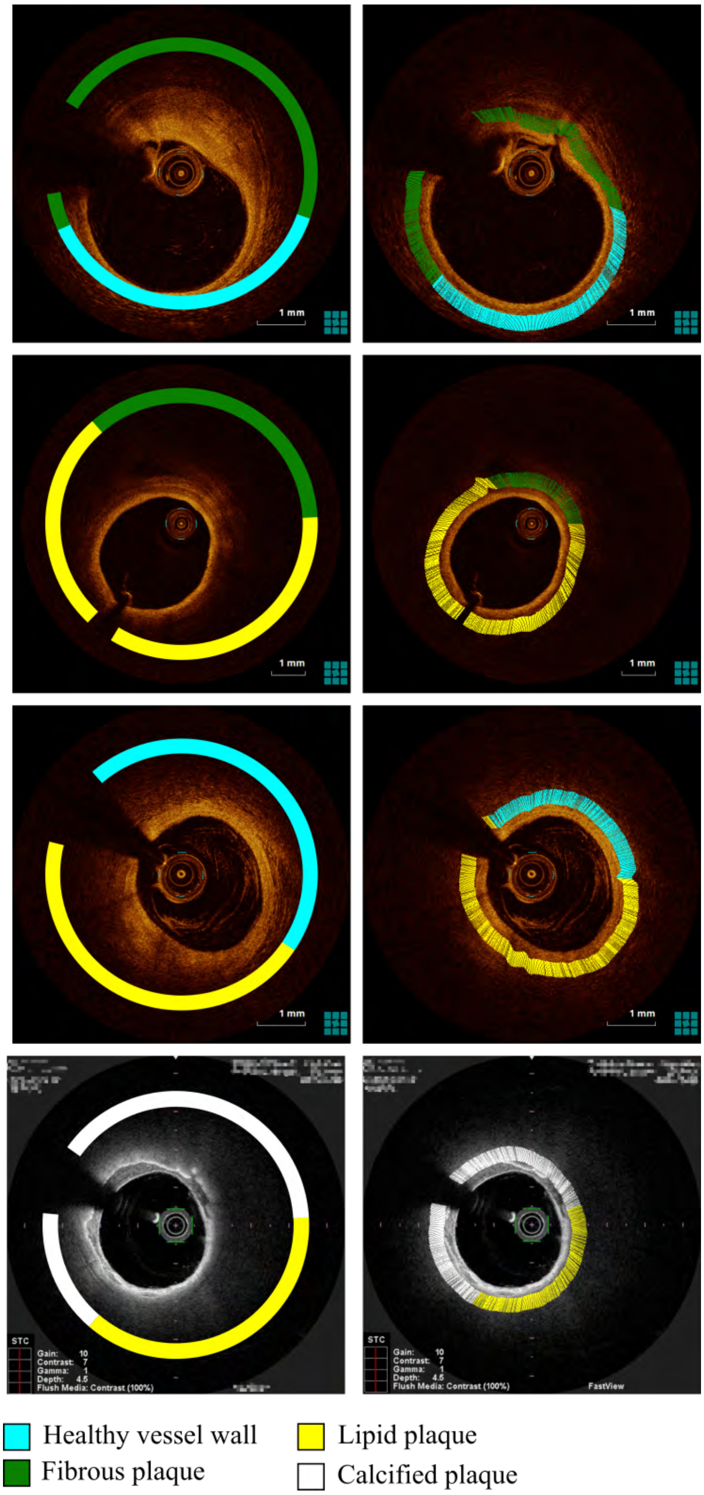


Fig. 4.3 Left: 4 examples of lesion vessel (including lipid plaque, fibrous plaque and calcified plaque), which were labelled with the ground truth determined by the expert. Right: The prediction results obtained by our tissue classification method. Each tissue type was illustrated with the color code. The range of the polar angle presented the region where the tissue occurred along the circumferential direction.

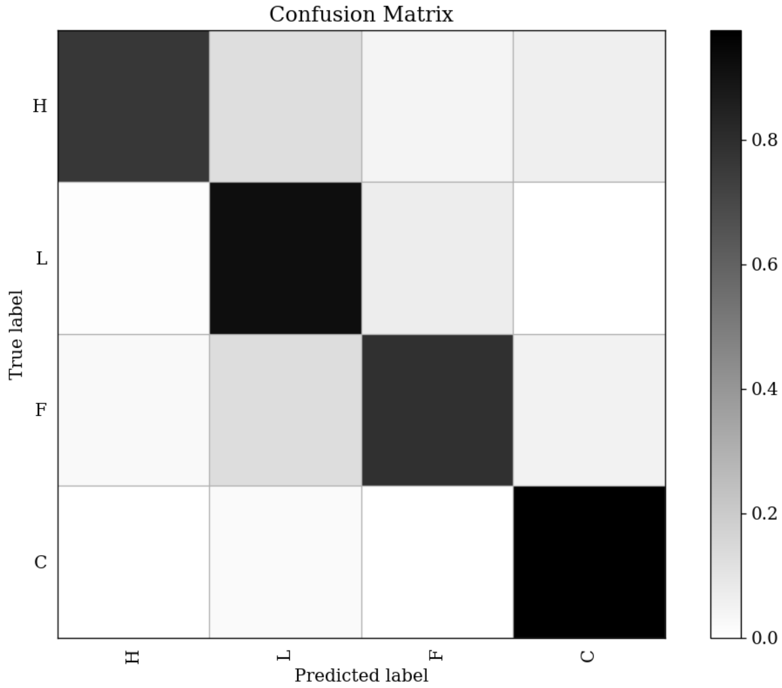


Fig. 4.4 A confusion matrix for the data set, which the classification results were obtained with our local multi-layer model.

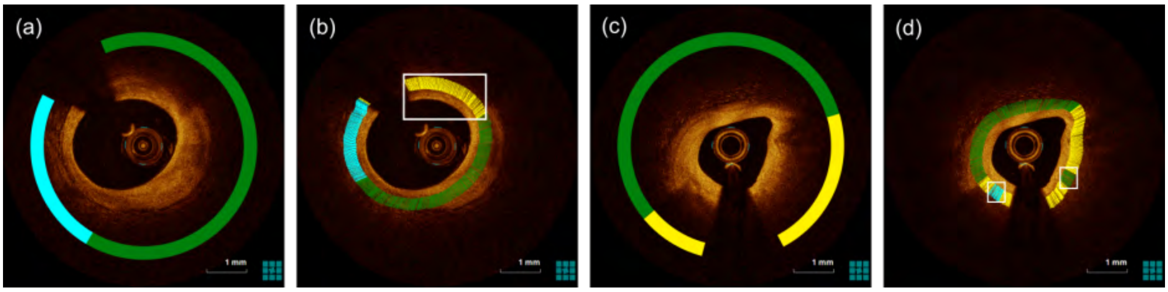


Fig. 4.5 (a) and (c) are the ground truth corresponding to (b) and (d). The category labeled with a white box in image (b) presents an error type (lipid plaque) for the possible reason is the texture of the original plaque is not clear. Some ALSRs might be recognized as wrong tissue types, in (d) the error categories are healthy and fibrous types, which is marked with the white boxes.

we did not discuss the mixed tissue case. Furthermore, the location of GW in the human vessel, the noise generated by point spread function (PSF), the OCT images with blurring and low resolution can also cause a decrease in Recall, Precision and Accuracy.

In summary, we proposed a local multi-layer model for the tissue characteristic analysis and extraction based on the pre-processing result obtained with the methods depicted in chapter 2 and chapter 3. Our method discussed the possible features extracted from the PLBR of IVOCT images and utilized GLCM to calculate four features (dissimilarity, homogeneity, energy and correlation) to produce a total of 56 features for RF classifier.

Table 4.3 Evaluation metrics for the vessel tissue classification

Dataset	Validation Metrics														
	Recall (%)					**Precision (%)**					**ACC. (%)**				
	H[1]	L[1]	F[1]	C[1]	N[1]	H	L	F	C	N	H	L	F	C	N
set 0	88.00	59.67	84.03	-	-	92.31	76.06	69.06	-	-	92.25	85.70	81.15	-	-
set 1	80.85	46.09	55.26	97.60	-	92.68	68.60	89.36	86.91	-	98.63	89.05	95.55	87.80	-
set 2	83.19	14.28	95.86	49.98	-	94.15	99.90	85.80	99.90	-	90.40	98.87	89.08	99.81	-
set 3	-	99.88	83.30	-	-	-	98.65	49.98	-	-	-	98.53	98.53	-	-
set 4	70.69	94.58	47.37	-	-	70.69	91.29	99.99	-	-	90.40	88.70	91.71	-	-
set 5	93.68	78.33	83.33	-	-	92.71	81.03	81.39	-	-	94.58	90.00	87.50	-	-
set 6	86.03	50.65	89.14	-	24.99	87.12	76.47	80.87	-	99.90	91.00	91.59	83.18	-	99.75

[1] H: *Healthy vessel wall*; L: *Lipid plaque;* F: *Fibrous plaque*; C: *Calcified plaque*; N: *Non-tissue object*

Chapter 5

Atherosclerosis plaque identification with deep learning

Over the last years, publications have employed automatic methods with machine learning technologies for the detection and classification task of vessel lesion tissues. Based on the extracted features of A-lines or each pixel in IVOCT images, utilizing support vector machine (SVM), RF or decision trees, etc. with proper parameters setting to identify and classify lesion plaques. Extracting A-line features combining with optical properties characteristics, such as light attenuation, to produce n features of each A-line or each pixel. Traditional machine learning methods are always applied to investigate the feature extracted from the image data with some special feature engineering technologies, but for deeper abstract features, it's still difficult to draw out. Especially for the type distinguishing work of fibro-lipid and lipid, the diffuse border property of the lipid plaque makes the recognition difficult for people, even sometimes including specialists, which brings challenges to this task. Therefore, developing a new technology to improve the accuracy of plaque recognition significantly impact the treatment effect for patients with this symptom and increase the life-time of patients. Recently, deep learning technology performs an important breakthrough in image classification and target detection, which progresses and is applied in more areas containing medicine image processing, auto-driving and voice recognition.

5.1 Deep learning basic concept

Deep learning is a specific subfield of machine learning: a new take on learning representations from data that puts an emphasis on learning successive layers of increasingly meaningful representations. Deep learning in the early years was depicted as simple linear models with a view of neural science. These models were designed to use a group of n variables $x_1, \ldots, x_n$ with respect to to an output y. Computing a set of weights $w_1, \ldots, w_n$ to make $y = f(\mathbf{x}, \mathbf{w}) = x_1 w_1 + \cdots + x_n w_n$. Linear models have a number of limitations that most famously, they cannot learn the asymptotic (XOR) function. Until 2006, in Geoffrey Hinton research group, there's been a major breakthrough in neural networks[33]. Simultaneously, with the appearance of the big data and high performance of hardware (faster CPU, generality GPU and more powerful algorithm and framework) as the fundamental conditions, it's easy to construct a deep learning neural network and train data with relative models. Since the 1980s, the ability of deep learning to provide accurate identification and prediction has been improving. Moreover, deep learning continues to be successfully applied to a growing range of practical problems. In 2012, deep learning with the convolutional neural network (CNN) first won the ImageNet large scale visual recognition competition (ILSVRC) and decreased the errors of the top 5 from 26.1% to 15.3%[42], which produces a positive and significant impact on the image recognition field.

Convolution

CNN[95, 46, 61], as one of the important deep learning neural framework, is a neural network specifically designed to process data with grid-like structures. Its name comes from the convolution operation applied in this neural network. Generally, convolution is a mathematical operation on two real variable functions $f(x)$ and $g(x)$. Let $s(x)$ be

$$s(t) = \int f(x)g(t-x)dx, \tag{5.1}$$

where $f(x)$ and $g(x)$ are integrable function in $\mathbf{R}$, respectively. $s(t)$ is a multiplied result of these two functions and denoted with respect to variable t.

Actually, time is defined as a discrete variable when we process the data on the computer. Therefore, the discrete formation of Eq. 5.1 is shown as follows:

$$s(t) = (f * g)[t] = \sum_{x=-\infty}^{\infty} f(x)g(t-x). \tag{5.2}$$

In the terminology of convolutional networks, the first parameter $f(x)$ usually called *input*, and the $g(x)$ is called *kernel function*, the output sometimes is defined as a *feature map*. In fact, the *input* in machine learning and deep learning is generally a data with a multidimensional array. If making a convolution operation to a two-dimension image I with a two-dimension kernel K, the convolutional results with this kernel K can be calculated through the following formula:

$$S(i,\, j) = (I * K)(i,\, j) = \sum_{m}\sum_{n} I(m,\, n)K(i-m,\, j-n), \tag{5.3}$$

where K is $S \times T$ and I is $M \times N$. With the commutative attribute of the convolution, Eq. 5.4 can be written as

$$S(i,\, j) = (I * K)(i,\, j) = \sum_{m}\sum_{n} I(i-m,\, j-n)K(m,\, n). \tag{5.4}$$

Figure 5.1 simply shows the schematic of a convolution operation in a 2-dimensional image through a 2 × 2 kernel. Each time, a 2 × 2 region (red rectangle) corresponding to the kernel size is acquired and employed convolution with the kernel K to generate a new related output. That is, sliding the kernel filter over the image spatially, and computing the dot products of the corresponding local 2-dimension region. When sliding over all spatial locations, to a 3 × 4 image and a 2 × 2 kernel (Fig. 5.1), a 2 × 3 output (*feature map*) finally would be obtained through the convolution operation. The above description just demonstrates the general principle of the convolution procedure, notably, the 2-D image of the introduced example contains only one single channel for convolution operation. As we know, most 2-D images that are used for the image analysis and understanding work is a structure of three

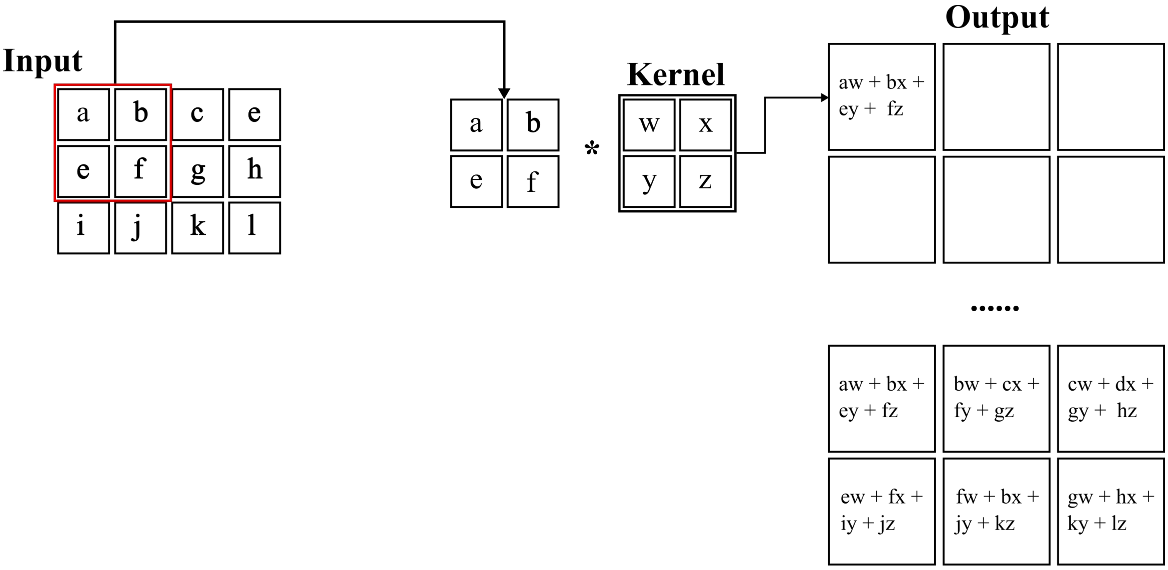


Fig. 5.1 An example of a 2D convolutional operation with a kernel K (2 × 2) applied to a 2D image. The same size region (red rectangle) acquired from the input iteratively is used to participate in the convolutional operation with the convolution kernel. Finally, (2 × 3) output would be produced after the implementation for each convolution to the *input*.

channels (RGB), therefore, the convolution operation to the image with RGB is different from the single-channel image. Figure.5.2 is the illustration of the input images with two kinds of channel number. For the convolution of images with RGB channels, each channel respectively performs the convolution with the *kernel* to produce a corresponding feature map, and then an accumulation among these feature maps occurs to generate a new feature map that contains only one channel.

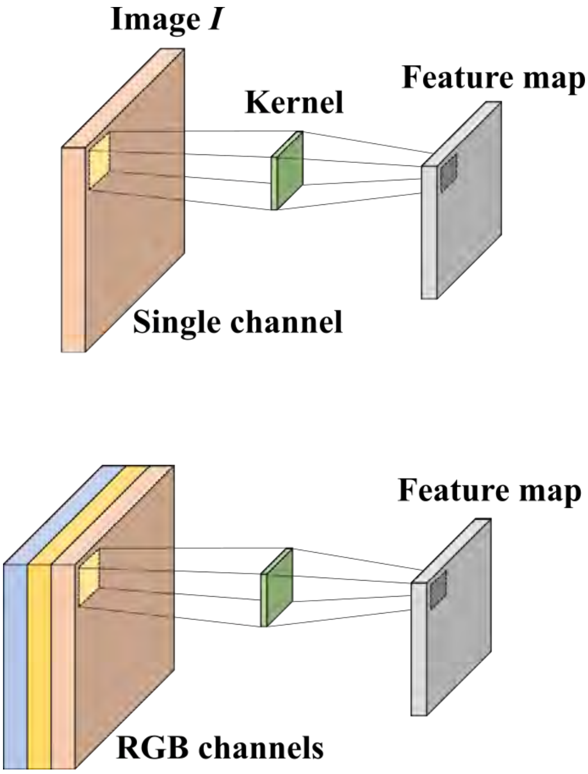


Fig. 5.2 Two types of input images containing a single channel and RGB channels, respectively. The number of created feature map after convolution operation with one kernel is one channel.

In the above case, there is only one channel kernel filter implemented the image convolution, however, the number of the kernel filter is more than one, actually, for the extraction of more spatial information. Suppose there are n kernel filters with the same size, each kernel performs the same convolutional operation to the gray images or RGB images. As known from the above content, one kernel produces one corresponding feature map, therefore, if utilizing n kernel filters to convolve the input data, n layers of feature maps would be created, which are also the output of the input image obtained with these kernel filters. The number of kernels in Fig. 5.3 increases from 1 (Fig. 5.2: bottom) to n ($n > 1$). Subsequently, sliding the kernel over the spatial location of the input data and repeating this operation until all the filters finished. As a result, the feature map containing n layers corresponding to each filter is created. Each filter is regarded as an "eye" to "watch or fee" the local region (receptive field) to generate a volume of the feature maps. And the result of each filter on the current sliding position is denoted as

$$\mathbf{y} = \mathbf{w}^{\mathbf{T}}\mathbf{x} + \mathbf{b}, \tag{5.5}$$

where **y** is the output result with applying the filter to the input data, **x** is the input. **w** is the weight parameter and **b** is the bias.

The convolution operation to the image data can compress the spatial size of the original image and simultaneously, improve the feature extraction ability and the "thickness" of the feature volume to obtain more useful spatial and abstract information from the input data. The size of the feature map is depended on the size of the input data and kernel filter. Suppose that the size of the input data is $W_1 \times H_1 \times D_1$, where W_1 is the width of the input, H_1 is the height, and D_1 is the channels or the depth to describe the input data in the third dimension. Four hyperparameters are respectively included: the number of filters is K and its size of $F \times F$ with respect to width and height, the stride variable S, and the zero padding P. Producing a volume of feature map with the size of $W_2 \times H_2 \times D_2$, it can be calculated as follow formulas:

$$\begin{aligned} W_2 &= \frac{(W_1 - F + 2P)}{S} + 1, && (5.6) \\ H_2 &= \frac{(H_1 - F + 2P)}{S} + 1, && (5.7) \\ D_2 &= K, && (5.8) \end{aligned}$$

where W_2 and H_2 are the width and height of the feature map, respectively, D_2 is the depth of the feature map. In the output volume, the d-th depth slice (of size $W_2 \times H_2$) is the result of performing a valid convolution of the d-th filter over the input volume with a stride of S, and then offset by d-th bias.

A strategy to decrease the number of parameters is to use the parameter sharing principle. Each slice layers in the input data share a set of parameters with the same settings. Of course, parameter sharing also can be helpful to speed up the derivative computation of the backpropagation[45]. Hence, it introduces $F \times F \times D_1$ weights pre filter, and for a total of $(F \times F \times D_1) \times K$ weights and K biases.

Pooling

Although applying parameter sharing is the purpose of the parameters reducing, the number of parameters is still large that it causes a huge computation time for the cost. Implementing the pooling operation to decrease the parameter number and the neural network computation

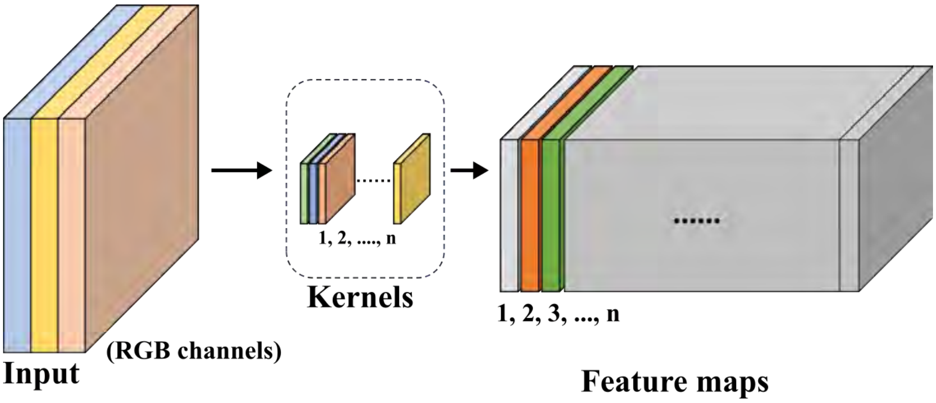
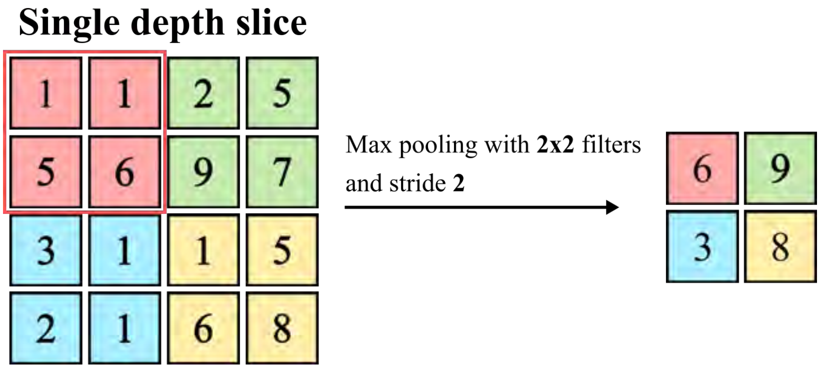


Fig. 5.3 Demonstration of the n layers feature maps generating with using n kernel filters after convolving all the spatial locations on the input data.

volume, and preventing overfitting results from occurring. Therefore, it is common to periodically insert a pooling layer after the convolutional operation for each input data. Pooling layer operates on each slice of the input and makes the representations smaller and more manageable. Suppose the input with the size of $W_1 \times H_1 \times D_1$, both the width and height of the pooling layer is F. The stride is S. A pooling output from the input $[W_1 \times H_1 \times D_1]$ is $W_2 = (W_1 - F)/S + 1$, $H_2 = (H_1 - F)/S + 1$, $D_2 = D_1$. The depth of the pooling output is the same as the input. The most common form is a pooling layer with filters of size 2×2 applied with a stride of 2 downsamples every depth slice in the input by 2 along both width and height, discarding 75% of the activations. As shown in Fig. 5.4, the pooling layer is implemented to an individual slice to generate the max results for every operation, and finally, a 2×2 output is acquired with 2×2 filter and stride 2.



1	1	2	5
5	6	9	7
3	1	1	5
2	1	6	8

6	9
3	8

Fig. 5.4 Schematic of the pooling operation applied to an independent slice. Red rectangle indicates the max pooling size of 2×2. Pooling operating slides iteratively for a stride of 2 to produce compressing outputs.

Convolution neural network architecture

The typical CNN architecture mainly consists of 3 parts: convolution, pooling, and Fully-connected (FC), additionally with an extra function of ReLU activation[42] which applies element-wise non-linearity. The most common CNN architecture can be presented as the following pattern: INPUT→[[CONV→RELU]⋆N]→POOL]⋆M→[FC→RELU]⋆K→FC, where INPUT is the input dataset, [CONV→RELU] is the convolution layer combines convolution and activation function *ReLU*, FC is the fully-connected layer which has full connections to all activations in the previous layer. N, M and K is the layer number. With properly designing for the specific image classify tasks, the layer number of the CNN architecture can be different. Many deep learning models derive and are developed from the architecture proposed by Krizhevsky et al. An example of CNN architecture is illustrated in Fig. 5.5.

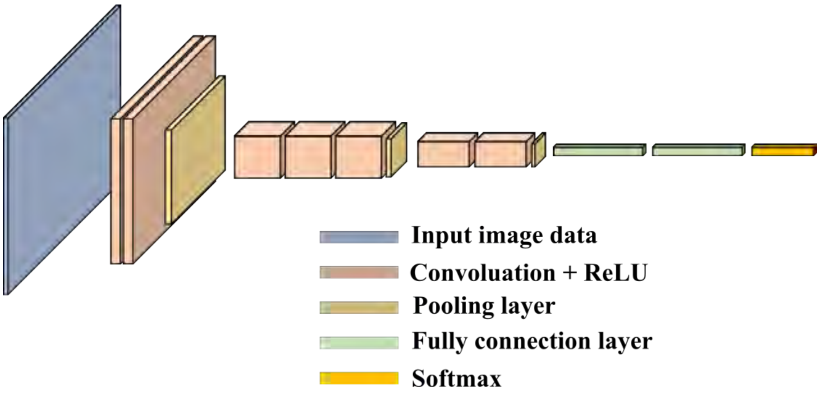


Fig. 5.5 Schematic of CNN structure containing input data layer, convolutional layers, pooling layers, FC layers and a final layer called *softmax* applied for the final prediction. This CNN architecture is usually used to classify the topic of the input image data.

5.2 VGG-like model building

5.2.1 VGG network architecture

In 2014, a new deep convolutional neural network: the VGGNet[75] was developed by the Computer Vision Group (Visual Geometry Group) at the University of Oxford, together with researchers from Google DeepMind, and took second place in the classification project of the ILSVRC 2014 competition and first place in the localization project. VGGNet explored the relationship between the depth of a convolutional neural network and its performance, and successfully constructed a convolutional neural network 16-19 layers deep, demonstrating that increasing the depth of the network can affect the final performance of the network to some extent, resulting in a significant reduction in the error rate, while at the same time being highly scalable and generalizable for migration to other image data. VGGNet can be seen as a deepened version of AlexNet[42], both consisting of two major parts: the convolutional layer and the full connection layer. VGG consists of 5 convolutional layers, 3 fully connected layers, and softmax output layers, separated from each other using max-pooling, and the ReLU function is used for all hidden layer activation units. VGG uses multiple convolutional layers with smaller convolutional cores 3×3) instead of one convolutional layer with larger convolutional cores, which on one hand can reduce parameters and on the other hand corresponds to more non-linear mapping, which can increase the fit/expression of the network. To date, VGG is still being used to extract image features.

5.2.2 Input data

OCT image cropping

The vessel OCT image is produced by the sensor of the OCT equipment by receiving the backscattering of light reflected from the vessel wall when the catheter runs with the form of "pull back". Soest et al.[86] investigated the attenuation coefficient (μ_t) for different types of the vessel tissues: healthy vessel wall and fibrous plaque are both $2-5mm^{-1}$, lipid tissue is $\mu_t \geq 10mm^{-1}$.

Cheimariotis[21] divided an IVOCT image into square patches with each of 8×8 pixels as the input data. From the IVOCT image, we could know that most areas of the vessel image consist of the background and part of the vessel wall containing useless information. Furthermore, the size of 8×8 pixels is too small to have more information about the vessel tissue. With taking account of the OCT imaging mechanism, we crop the patch with a rectangle shape along the circumferential direction based on the detected lumen boundary, rather than the traditional square patch cropping. Here, the patch size is set to 24×56 (a size of $ALSR^{123}$), which the width value setting (equal to 56) is considering the fact that the thickness of the healthy tissue layer is less than $300\mu m$, the fibrous thickness is less than $65\mu m$, while the lipid thickness is more than $500\mu m$ (described in Tab. 4.1).

Channel and texture

To analyze the effect of the channel number and the texture information to the accuracy of prediction of tissue category, three different channel types of input data, which are a single channel based on Local Binary Pattern (LBP)[62, 63], original channels of RGB and 4 channels with merging channels LRGB (LBP + RGB), were experimented. LBP is an operator used to describe the local texture features of an image; it has significant advantages such as rotation invariance and grayscale invariance. It compares each pixel with its neighbors and saves the result as a binary number. E.g., in a 3×3 window, the intensity of the adjoining 8 pixels is compared with the center pixel in the window, if the value of neighbor pixels is bigger than the center intensity value, the corresponding position in the neighbor is labelled with 1, else, marked with 0. As a result, an 8-bit binary number is produced as the LBP code of the center pixel, which is used to present the texture feature. The most important properties of LBP are its robustness to grayscale changes such as those caused by changes in lighting and its computational simplicity.

In our procedure, the LBP principle is used to present the intensity variance between the center pixel and the N-neighbor pixels around it is considered to be the joint feature to present the relationship of the center pixel with the around pixels. This principle is defined as follows:

$$V_i = \frac{1}{N} \sum_{i=1}^{N} (I_i - \mu)^2, \tag{5.9}$$

where V_i indicates the joint feature of pixel i, N value is the number of neighbors, I_i is the intensity of the i-th neighbor pixel around the center pixel expressed with the V_i value, μ is the average intensity of the N-neighbor pixels.

I explored the texture information affecting to the plaque classification result of ALSR in the PLBR of IVOCT images. And designing three types of input with changing the channel numbers of input data, simultaneously. The LBP channel contains a single channel and the RGB type has 3 channels. In the third one, we merge the single LBP-channel with RGB channels to construct a 4-channel input data type. Different kind of data with or without texture information was fed into the deep learning model to watch the prediction.

5.2.3 VGG-like model

VGG-Net explores the relationship between the depth of CNN and its performance. The contribution of VGG-Net is using a very small receptive field: 3×3 instead of the 5×5 or 7×7 and increasing the depth of the CNN layers. VGG-16 and VGG-19, which are discussed in the paper[75] , have a fixed-size 244×244 as the model input size. Considering the fixed-size (24×56) of a patch we defined as well as the receptive field size and max-pooling window size which are proposed in VGG-Net, the layers of our model should not be deeper, which is determined by the size of the input data fed into the model. Therefore, a CNN model with a depth of 11 layers based on the VGG-Net is constructed for the vessel tissue classification task. The input data is passed to convolution (conv.) layers which combine the conv. operation and ReLU function. The size of the conv. filter is selected with the size of 3×3 as same as employing in VGG-Net, and the filter stride is set as 1. Simultaneously, we increased the size of the input data before every conv. layer to make the corresponding output in the same layer keeping the same size. This is called the "Padding" operation which is usually employed in deep learning models to retain or process the boundary information of the input data. In our paper, the padding is 1 pixel for 3×3 conv. layers, which can maintain the output size as same as the input one. After a stack of conv. layers, spatial pooling is performed with a filter size of 2×2 and stride 2. The structure of [Conv.+Pooling] is implemented three times and the depth of the feature map for each group is larger than the former group. The first group contains two conv. layers and one pooling layer, and the conv.filter applied to the input data is a size of $64 \times 3 \times 3$, which 64 is the filter number and 3 denotes the width and height of the filter. The filter depth in the last two groups increases gradually to make sure that the designed model can extract deeper abstract feature information. We also use the *Drop out* technology[76] to avoid overfitting and time-consuming, where the *Drop out* can effectively relieve the overfitting and achieve regularization to a certain degree by ignoring a portion of the feature detection (let the hidden layers be shut down) to reduce the mutual effect of hidden layers. Two FC layers with 128 channels subsequently followed by a 4 channels FC layer are applied to fully connect every element of the former feature map. In the last FC layer, the *softmax function* is combined to obtain the final prediction results of the classes of the input data. Four categories (healthy vessel wall, fibrous plaque, lipid plaque and residual guide wire region) are predicted with classification scores, in which the highest score indicates the most probability of a class the sample belongs to. The detailed information of the model architecture is presented in Fig. 5.6 and the network configuration is in Appendix B.1.

5.3 Experimental results

5.3.1 Parameter setting

Three models (LBP, RGB and LRGB patterns) were tested for the tissue classification of IVOCT images. We implemented our method and constructed the VGG-like model with Keras framework. The dataset was first split into two parts: the training set (80%) and the test set (20%), and to the validation set, we acquired it from the 20% of the training set. The random state in our approach was set 42 for the shuffling operation. The size of all the conv. kernel was 3×3 and the pooling filter was 2×2 with a stride of 2. Padding operation in conv. layers was employed to keep the boundary information of the input image data. With Keras

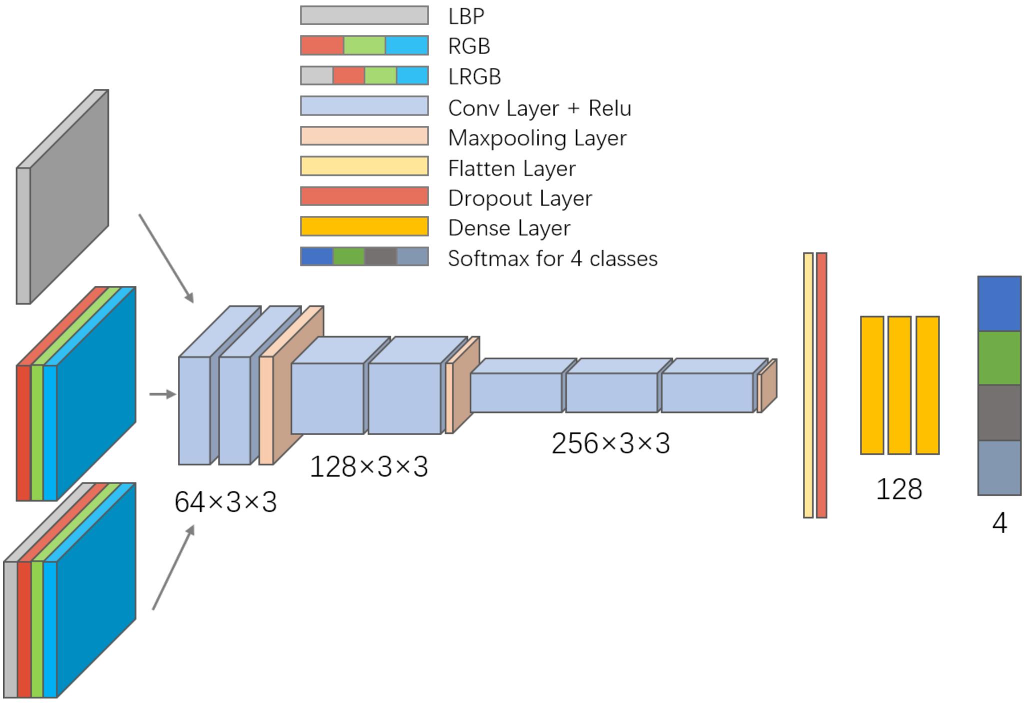


Fig. 5.6 This architecture is constructed based on the VGG-Net model. Three kinds of input type, including single-channel LBP-based input, the original image and the input of merging LBP channel and RGB channels, are respectively fed to our model for the tissue classification. At the end of this model, 4 scores corresponding to 4 classes (healthy wall, fibrous plaque, lipid plaque and residual guide-wire region) are obtained through the *softmax* function, which the highest score illustrates the most likely class that the input data is.

framework, a flatten layer was utilized to flat a 2-D feature map to 1-D data. The dropout rate was set as 0.5 to control the ratio of the active neural cells. In our model, the batch size for the training set and validation set were both set as 32 for the one-time input data feeding. For the gradient descent optimization of the model, we selected Adaptive Moment Estimation (Adam)[39] with weight decay of 10^{-6} as the optimizer to compute adaptive learning rates for each parameter. The learning rate was 0.0001.

5.3.2 Validation

For evaluation of multi-classification, we statistic the scores of the 4 classes with measurements of Sensitivity (Sens.), Specificity (Spec.), Precision (Prec.) and Accuracy (Acc.) in different input channel types. Sens. denotes the percentage of correct results in all true samples. In contrast, Spec. indicates the incorrect percentage in all false values. Prec. describes a measure of the percentage of correct predicted results in all predictions. Acc. is the percentage of all the correct prediction results (true or false) in the entirety predictions

Defining TP, TN, FP and FN as the basic indicators to depict the relationship between the predication results and the true sample values, where TP is the true positive, TN is true negative, FP is false positive and FN is false negative.

$$Sens. = \frac{TP}{TP+FN}, \tag{5.10}$$

$$Spec. = \frac{TN}{TN+FP}, \tag{5.11}$$

$$Prec. = \frac{TP}{TP+FP}, \tag{5.12}$$

$$Acc. = \frac{TP+TN}{TP+FP+FN+TN}. \tag{5.13}$$

5.3.3 Results

In this chapter, we applied our procedure to 57 IVOCT image frames of 3 patients and produced 1458 rectangle patches with cropping along the circumferential direction as the input samples. The tags of 4 tissue classes, including the healthy wall, lipid plaque, fibrous plaque and residual guide-wire-region, are labelled to these samples by an expert through manual job. The extra class of residual guide-wire-region, partially existing in some IVOCT images, is due to a threshold setting in our guide wire removal algorithm depicted in Sec. 2.5 of chapter 2. Totally, 933 samples with the rectangle shape were used as the training data set to train our model while 233 samples for the validation data set. The accuracy and loss values of our model during the training period are presented in Fig. 5.7, which we used 3 color lines to display the results of 3 channel types in epochs, respectively. From these two figures, we could find that the Acc. result with 4 channels is better than the other two types, and the loss value with 4 channels is also lower than the rest types.

The evaluation score results are shown in Tab. 5.1, and the prediction results of the tissues, which are labelled with 3 different colors, are shown in Fig. 5.7 comparing with the ground-truth that was marked by an expert. It's clearly noticed that tissue classification with our method could compete for the lesion tissue classification task.

5.4 Discussion and conclusion

We tested for 1458 input samples cropping from the IVOCT image along the circumferential direction based on the lumen boundary. Data augmentation technology, via a number of random transformations of images, was employed to improve the generalization of our model and reduce the overfitting. The operators of the horizontal flip, height shift, width shift and

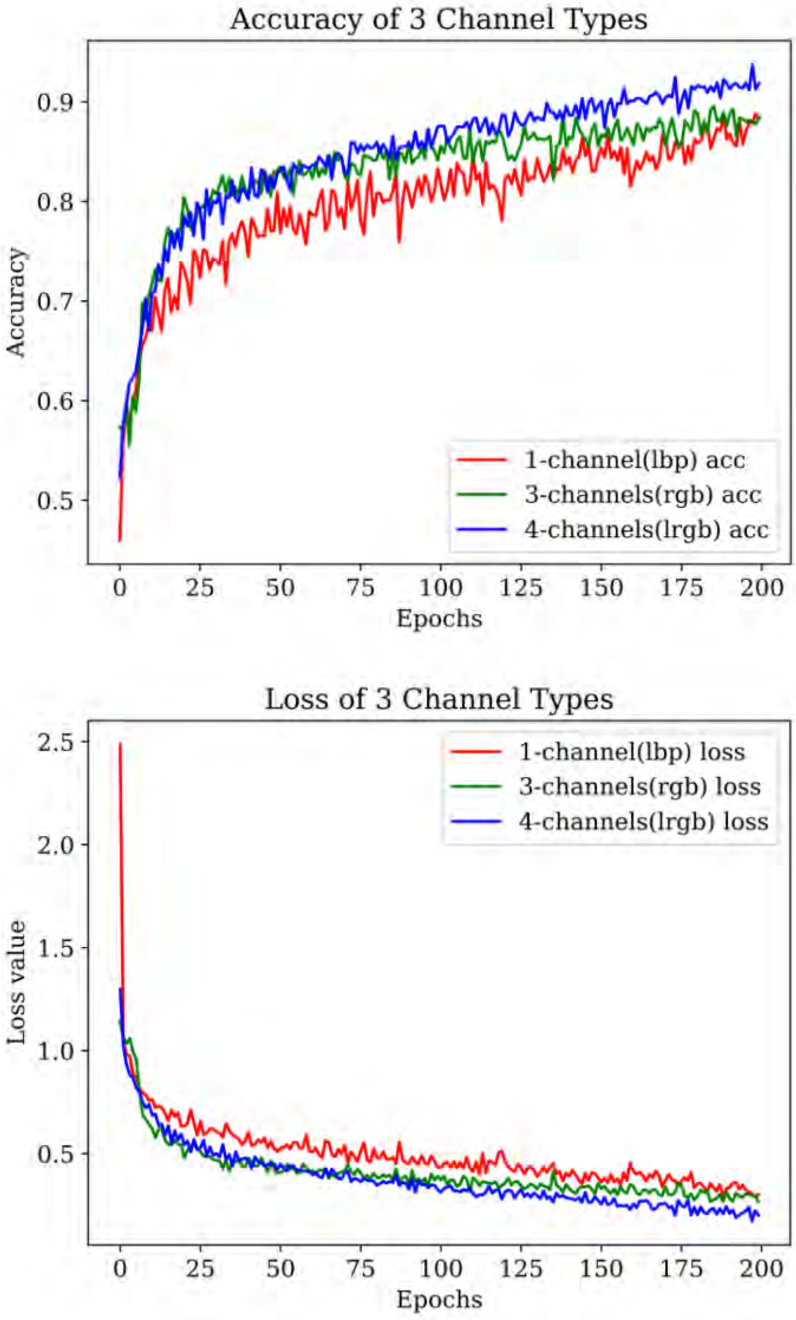


Fig. 5.7 Left: respectively shows the accuracy results of the 3 input channel types which are presented in different colors. Right: the loss values for each kind of input. Both of the epoch values are set to 200.

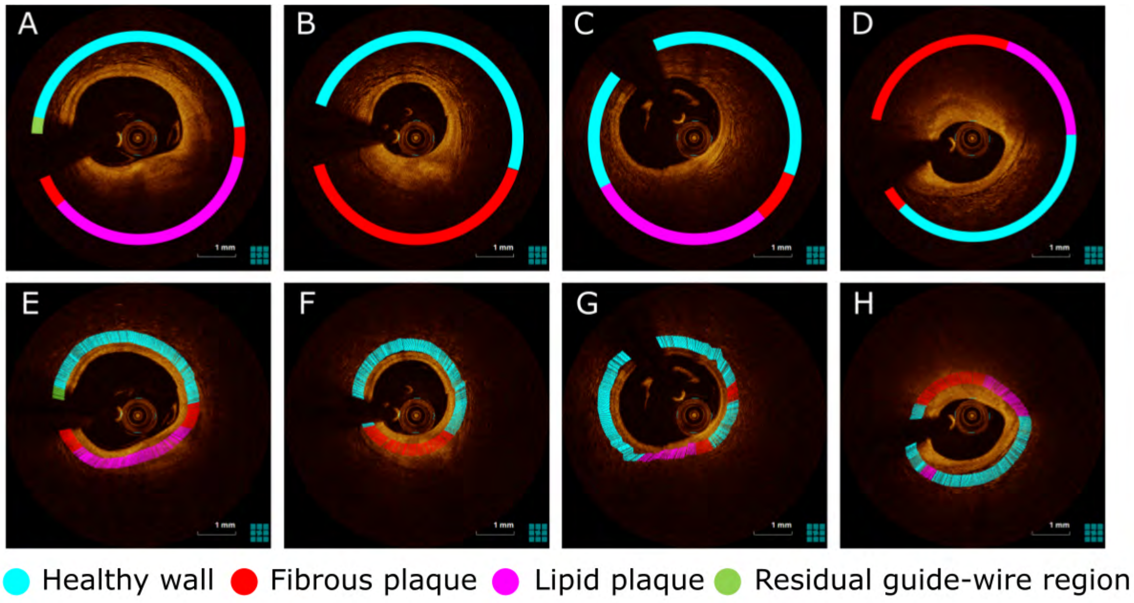


Fig. 5.8 Results for the healthy wall, lipid plaque and fibrous plaque classification. Top: is the ground-truth of the experiment data handled by an expert. Bottom: presents the tissue classification prediction results by using our model.

Table 5.1 Evaluation results of 4 classes in 3 different input channel types

Category	**Input Channel Type**	**Sens.**	**Spec.**	**Prec.**	**Acc.**
Healthy vessel wall	Single channel (LBP)	0.8846	0.8636	0.9148	0.8767
	Three channels (RGB)	0.8477	**0.9053**	**0.9489**	0.8664
	Four channels (LRGB)	**0.9039**	0.8609	0.9091	**0.887**
Lipid plauqe	Single channel (LBP)	0.9	0.9127	0.6207	0.911
	Three channels (RGB)	**0.9998**	0.8182	0.1034	0.8219
	Four channels (LRGB)	0.8267	**0.9419**	**0.7586**	**0.9281**
Fibrous plaque	Single channel (LBP)	0.5574	0.9394	0.7083	0.8596
	Three channels (RGB)	0.4045	**0.9409**	**0.75**	0.7774
	Four channels (LRGB)	**0.6111**	0.937	0.6875	**0.8767**
Residual GW Region	Single channel (LBP)	**0.9999**	**0.9965**	**0.8999**	**0.9966**
	Three channels (RGB)	0.0	0.9658	0.0	0.9658
	Four channels (LRGB)	0.8999	**0.9965**	**0.8999**	0.9931

shear intensity were tested for increasing the number of data. We used the testing samples to validate the capability of tissue classification and the prediction results for 3 channel type are shown in Tab. 5.2. In Tab. 5.1 and Tab. 5.2, we see that 4-channel (LRGB) performs a better result in vessel tissue classification. We directly add the LBP channel to the RGB channels to increase the texture information of the vessel. In other words, the healthy vessel wall and the fibrous plaque contain layer information, such as intima, media and adventitia layers, while the intensity alteration of lipid plaque is slow within A-line depth, which presents a diffuse characteristic.

Table 5.2 Results of accuracy for 292 testing data in 3 different input channel types

Input Channel Type	**Single channel (LBP)**	**Three channels (RGB)**	**Four channels (LRGB)**
Acc.	0.8219	0.7158	0.8425

Although we applied the data augmentation technology, one limitation is that the overfitting will be generated if the epoch takes a big value, and the classification accuracy would not improve again. For this issue, a large number of original IVOCT image set should be supplied for a better model generalization. Additionally, data calculated based on the LBP principle from the original IVOCT image lead to a large value range and it is different from the value range of RGB. Therefore, in the third input type, we didn't process the merging data by regularization. It should positively discuss the impact of the regularization in the 4-channel input in the future time. We took account of the light attenuation in vessel tissue and set a fixed size to crop for samples, which could quickly classify the single type tissue in a sample but not for a mixed tissue situation. Thus, an experiment should be employed to investigate the patch size setting in detail. Based on this experiment, which could be used as a baseline, our next step is to progressively apply a fully convolutional neural network (FCN) on IVOCT tissue region detection and classification.

We constructed a VGG-like model to classify the vessel tissues and discuss the results of 3 different types of the input channel. With taking account of the light attenuation, we cropped the original IVOCT image with defined size patches along a circumferential direction instead of the traditional method[21, 32] which feeds a whole vessel image into a CNN model or crops the vessel image along the horizontal and vertical direction. As we know, this is the first time to merge a single channel LBP-based with RGB channels and crop the sample patches along the circumferential direction in a Cartesian coordinate. The results show that our method has the potential to attack the tissue classification problem of the IVOCT image.

Chapter 6

Semantic segmentation for atherosclerosis plaques

Indeed, the aforementioned studies have accomplished promising results on the research of detection and identification of vessel lesion types with machine learning or deep learning methods, but some limitations exist. (1) Almost all relevant studies calculated the outer borders by defining a fixed value (e.g. $1\,mm$ or $1.5\,mm$) of depth from the lumen boundary along the radial direction. (2) The size of input data for pixel-wise segmentation with deep learning was the same as the dimension of the original IVOCT images. Observing from the IVOCT image (Fig. 1.8), obviously, the superficial region of the vessel contains the most useful information about the vessel inner tissue while useless data of IVOCT images would increase the time-costing for computation. (3) Some studies just investigated the classification with limited types of lesion tissue. An automatic pixel-wise method to segment simultaneously is necessary and important for multi-types of the lesion plaque. (4) Most of the existing approaches recognized plaques based on machine learning or CNN as the feature extractor. A deep learning model designed for the semantic segmentation of IVOCT images is not discussed.

In this chapter, we use a novel method to segment ROI of the IVOCT image and apply DB-SegNet of which the basic unit is comprised of the dense block, downsampling layer and upsampling layer to these ROIs for vessel tissue classification. At the pre-processing stage, automated methods[70] proposed by our group are applied to remove the catheter area, segment the GW shadow region (a district with sector shape) and detect the lumen boundary, respectively. Subsequently, ROI of the individual IVOCT image is segmented to decrease the useless information and accelerate the learning speed in the neural network. The ROI is the district between the lumen boundary and the outer border, where the outer border is obtained with a level-set model in the vessel images based on the detected luminal borders.

Utilizing the pixels of ROI as the center points of the cropped square patches to produce 320×320 input image data. Totally, 490 IVOCT images derived from 7 patients including various types of plaques are used to build a data set for training and testing in our experiment. Evaluation is implemented with ten-fold cross-validation for observing the measure metrics of semantic segmentation of vessel lesion plaques. Besides, we compare the segmentation results of lesion tissues between SegNet and our model.

6.1 ROI segmentation

In rotary-style OCT scanning, due to the vessel tissue scatter and attenuation, the ability of light-absorbing of the vessel wall is increasing as the tissue depth increment[94]. As shown in Fig. 1.8, except for the center high-light region of the IVOCT image, the rest surrounding areas of the vessel imaging present poor signal and low contrast. The light penetrates into only the superficial layer of the vessel wall adjacent to the lumen to make the superficial plaques be observed clearly. All the studies paid attention to the benefit of the visible region of the IVOCT image. The thickness with a range of $1\text{–}1.5\,mm$ from the lumen boundary to the outer border was determined as the depth of ROI for the research of plaques recognition[11, 18, 70, 96]. However, the thickness of the healthy vessel wall and calcified plaque are in a limited range with $< 0.3\,mm$[86] or $< 0.7\,mm$[91]. Lipid plaques usually demonstrate $3.9 \pm 2.1\,mm$[28]. Moreover, Wang et al.[88] divided the fibrous plaque into 3 categories: $< 0.065\,mm$, $0.065\text{–}0.15\,mm$ and $> 0.15\,mm$. Obviously, various tissue components exist a divergence of thickness. Even if in the same type of plaque condition, the lesion tissue thickness also generally presents differences.

According to the attenuation coefficient of various tissues and the definition of the valid analyzable vessel tissue region, we proposed an automatic ROI segmentation method based on a level-set model[20]. To obtain the outer border of ROI, the level-set method was applied to an extended ALSR (defined in Sec. 4.2 of chapter 4) which is illustrated in Fig. 6.1(a), a dotted rectangle annotated with cyan, and its transformation image is displayed as Fig. 6.1(b). Our aim is to discover a divided line in the extended ALSR with the level-set method, subsequently, all the extended ALSRs in one individual IVOCT image compose a line as the outer border of ROI of the current vessel image.

According to the level-set definition and concept, we define a zero level contour $C = \{(x,y)|\phi(x,y) = 0\}$, where $\phi(x,y)$ is a signed distance function (SDF). The region inside C presents $\phi(x,y) > 0$ while the region outside C illustrates $\phi(x,y) < 0$. Let's denote the C_0 as the initial contour. The desired segmenting border C with keeping a minimum length is to partition the ALSR into two regions, each of the regions presents a minimal intensity variance. That is minimizing the following energy term:

$$\begin{aligned} E_{CV} = \mu \int_{\Omega} \delta(\phi(x,y)) |\bigtriangledown \phi(x,y)| dxdy \\ + \nu \int_{\Omega} H(\phi(x,y)) dxdy \\ + \lambda_1 \int_{\Omega} |I_0(x,y) - c_1|^2 H(\phi(x,y)) dxdy \\ + \lambda_2 \int_{\Omega} |I_0(x,y) - c_2|^2 H(1 - \phi(x,y)) dxdy \end{aligned} \tag{6.1}$$

where $H(\phi)$ is the Heaviside function, δ is a smoothed Dirac function. I_0 denotes ALSR, and Ω is the image domain. c_1 is the average intensity inside C, and c_2 is the average intensity outside C. μ, ν, λ_1 and λ_2 are parameters. The first term in Eq. 6.1 is the length of C. The second term refers to the area information inside the contour C. The third and fourth terms denote intensity information of the two disjoint regions (inside and outside contour C), respectively.

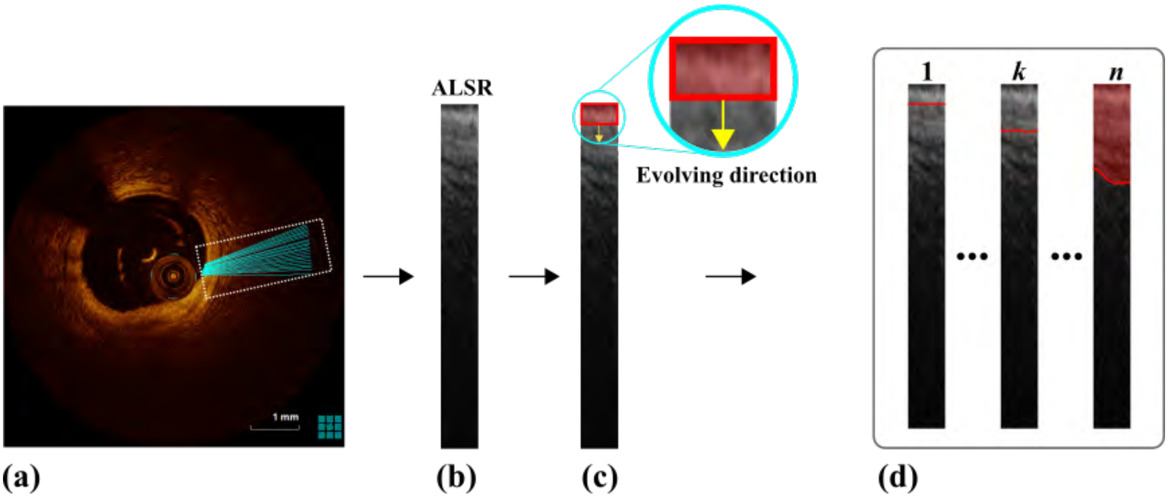


Fig. 6.1 An ALSR, which is labeled with the white dotted line in (a), obtained through our method[70] from an IVOCT image. (b) is the segmented ALSR after transformation. We defined an initial 2-D region (marked with a red rectangle contour) shown in (c), and the evolving direction is denoted with a yellow arrow. (d) illustrates a position alteration of the red line when the iterator of the level-set is changed from iterator = 1 to iterator = n. Then, a promising red segmented region of ALSR was gained and the red solid line is the final dividing line.

According to the definition of level-set, shortly, its principle is to evolve the size and shape of a contour iteratively to obtain a minimize the energy in a 2-D domain. Utilizing the feature of level-set, finding a stable status for contour C to segment a 2-D domain into two subregions. Similarly, in our extended ALSR images (Fig. 6.1(b)), our purpose is to use a divided line to partition extended ALSR into two subregions (Fig. 6.1(d)), where one of these subregions contains a visible lesion plaque region as many areas as possible. Here, our strategy is that we initialize the shape of the contour to a rectangle (Fig. 6.1(c)), where 3 borders (left, top and right) of the contour are coincident with the corresponding extended ASLR borders. This approach leaves one border (bottom direction in Fig. 6.1(c)) to complete the segmentation task. To keep the single divided line (bottom) moving iteratively from the initial position along the evolving direction demonstrated in Fig. 6.1(c) (yellow arrow), we set the SDF values of the initial area as 1, not only containing the contour border but also including the region inside this level-set contour. Finally, the dividing line would stop after a certain number of iterations. As we know in the level-set method, there are two part divided by the zero line, we called these two areas that one is the inner area and the other is outer area respectively. If we fill all the inner area with 1 and simultaneously the other three borders is the border of the extend ALSR, the only moving border is the bottom border. We described the width and height of an ALSR with l_w and l_h and the number of ALSRs of an individual IVOCT image depends on the value of $length_{LB}/l_w$, where LB indicates the lumen boundary. Afterwards, linear interpolation (Savitzky-Golay filter) was implemented to smooth the initial outer border of ROI.

6.2 Semantic segmentation architecture

In this step, we developed a deep learning model constructed of dense blocks[35] as its elements to investigate the semantic segmentation of the tissue classification in IVOCT images. The DenseNets model designer considered reusing the features of previous layers to the current layer to increase the learning ability of the network, where the basic component is the dense block consisting of several convolutional layers.

Let l depicts the index of the l^{th} layer, and x_l presents the output of the layer l. $H_l(\cdot)$ is a non-linear transformation, which is applied on the output of the $(l-1)^{th}$ layer. In a dense block, the input of the l^{th} layer is consist of x_0 and $l-1$ feature maps derived from the previous convolutional layers. That is, all the feature maps distributed in the preceding layers are connected to the subsequent layers. Now the input variable is consists of $x_0, x_1, \ldots$, and x_{l-1}. x_l is calculated as

$$x_l = H_l([x_0, x_1, \ldots, x_{l-1}]), \tag{6.2}$$

where $H_l(\cdot)$ is a composite function of operations including batch normalization (BN)[36], followed by a ReLU[27] and a 3×3 convolution layer. A schematic of a dense block structure surrounded in a dotted line is illustrated in Fig. 6.5. From Fig. 6.5, x_0 is an initial input data for the current dense block and four layers in the dense block produce four corresponding outputs, $[x_1, x_2, x_3, x_4]$, computed through Eq. 6.2. As shown in Fig. 6.5, “cat.” indicates a *concatenation* operation to join all the previous feature maps. For each $H_i(i = 1, 2, 3, 4)$, the channel number of input grows linearly with a *growth rate*. The final connective feature maps are transmitted to the transition layer for a down-sampling operation to reduce the spatial dimensionality of the feature maps.

In this research, the architecture of our semantic segmentation model (DB-SegNet, appendix B.2) contains two main portions: one is the down-sampling path and the other is the up-sampling path (Fig. 6.6), which construct our semantic segmentation model. Before entering the dense blocks in the down sampling path, a 3×3 convolution is firstly applied to generate feature-maps with a size of $m = n \times k$, where n is a parameter and k denotes the *growth rate*. Immediately, the output determined by the first convolutional layer is feed into the subsequent dense block for feature learning. In our model, an individual dense block of DB-SegNet contains l layers and each layer consists of [*bottleneck*, BN, ReLU, convolution:3×3], where the *bottleneck*[35, 37, 77] is 3 consecutive operations of [BN, ReLU, convolution:1×1] to reduce the dimensions of input feature maps. In the transition layer of the down-sampling path, [*bottleneck*, dropout, max pooling:2×2] following each dense block are combined to reduce the size of feature-maps and curtail the spatial resolution. In the up-sampling path, the reverse sampling operations are implemented to recover the size of the feature maps to the original input dimension gradually. Skip-connections are employed to concatenate the output of each up-sampling layer with the corresponding feature-maps in the down-sampling layer for reusing the previous superficial position information. Here, we use the skip-connection as *outer cat. operation* to join the features in the downsampling path to the decoded output in the upsampling path to keep the spatial information and improve the accuracy in the following upsampling layer. The joined feature maps are used as the input of the following dense block. It is worth noting that the concatenated results in the up-sampling path combine the corresponding feature-maps of both paths to produce a deeper feature-maps.

6.3 Post-processing

Firstly, to determine the final type of each pixel in an IVOCT image. As mentioned above, one IVOCT image produce a number of input data, and the corresponding predictions through DB-SegNet are obtained. Notably, all the input data with our cropped approach contain partly repetitive area, namely, the pixels in same position are included in different input data. Considering a pixel within different input data may be predicted as different type through DB-SegNet, hence, we construct a 3-D volume to place the classification result in the right position of each slice. The final type of each pixel can be got by computing the maximum number of classes along the *depth* axis. This is illustrated in Fig. 6.7. To reduce the classification errors, in the following step, we use morphological operations to smooth region, remove isolated spots and fill small holes.

6.4 Experimental results

6.4.1 Data acquisition and labeling

The IVOCT datasets in our experiments contained 490 IVOCT images from 7 patients, which were supplied by the Department of Cardiovascular Medicine, Wakayama Medicine University. The vessel tissue type of our experiment materials contained four different categories: the health vessel wall, fibrous, lipid and calcified plaques. We retrospectively investigated OCT datasets obtained from 7 coronary arteries in 7 patients with stable coronary artery disease. Each IVOCT images contains several tissue types among the four classes. Furthermore, all the unlabeled objects not belonging to the four categories were treated as the “background" type. The labeling job was manually accomplished by two experts through an annotation tool named “*Labelme*”. Annotation areas were both used for the assessment of the ROI segmentation results and the lesion tissue classification. Additionly, to evaluate the classification capability of our model and prepare for another research about the lesion plaque, the fibrous cap, a long and narrow area between the lumen boundary and the lipid necrotic core, was also annotated in one dataset.

6.4.2 ROI segmentation algorithm parameters

In our ROI segmentation strategy, as mentioned in Sec. 6.1, the out border of ROI in the IVOCT image was obtained based on the detected lumen boundary. We divided the lumen boundary into several segments and constructed ALSRs in one IVOCT image. The width of each ALSR, l_w, was set to 24 pixels (as same as in Chapter 5). Due to the alteration of the angle and position of GW in an entirety pullback, the distance between the lumen border to the outer edge in each vessel imaging is different. Hence, in our ROI segmentation algorithm, we multiplied the original depth with a coefficient $\beta = 0.8$ to dynamically altered the l_h to overcome the issue of depth variance. The coefficients for Eq. 6.1 was set as follows: $\mu = 1.0$, $v = 0$ and $\lambda_1 = \lambda_2 = 1.0$ by only considering the contour length, intensity information of subregion inside and outside the contour for a simple and fast result. To control the number of iteration ($iteration_{max}$) for achieving an outperformance of segmentation, we tested several numbers (600, 700, 800, 900, 1000 and 1200) respectively and finally set $iteration_{max} = 1000$. We finally used a digital filter, Savitzky-Golay filter, to smooth the initial outer border.

6.4.3 Deep learning model and data augmentation

In our research, pixels in ROI were selected as the center points and were utilized to cropped an individual IVOCT image into square patches that the areas containing pixels belong to ROI was at least more than 20% of the ROI to avoid the significant information about vessel tissue too less (Fig. 6.2). Using cropped patches is benefit to (1) improve the learning speed, (2) reduce the useless region analysis and (3) increase the generalization of data set. Consequently, 490 IVOCT images generated a total of 22, 210 square patches as the input data and each patch presented with a size of 320×320. Besides, a zero-padding method was designed to keep all patches produced with a consistent size. In our model, we also chose ADAM as the optimizer to gradually update the weight parameters in every backward for gaining the lowest loss value. ADAM optimization utilizes the power of adaptive learning rates methods to find individual learning rates for each parameter of the neural network.

Observing from the raw IVOCT images, statistically, the instance number of each tissue type in the vessel datasets (including healthy and unhealthy) presents unbalanced, which causes a negative impact on the accuracy of class recognition of some lesion plaques. Hence, we used *Focal Loss*[49] as the loss function to address the vessel tissue category imbalance during the training procedure. The focal loss is defined as follows:

$$FL = \begin{cases} -\alpha(1-p)^{\gamma} log(p), & if\, y = 1 \\ \\ -(1-\alpha)p^{\gamma} log(1-p), & if\, y = 0 \end{cases} \tag{6.3}$$

where $y \in \{0, 1\}$ denotes the ground-truth class. $p \in [0,\ 1]$, calculated with softmax function, specifies the prediction probability corresponding to the class with label y. $\alpha \in [0,\ 1]$ is a weight factor that balances the importance of positive/negative examples. $\gamma \geq 0$ addresses the imbalance contribution of the easy/hard samples to the loss.

We set the weight decay of ADAM optimizer with an initial value of 0.001. The coefficients in Eq. 6.3 were set as $\alpha = 0.75$ and $\gamma = 2$. As we know, a robust deep learning model highly relies on the volume and diversity of the training data, we enhanced the datasets with data augmentation technology. Random horizontal flip, random vertical flip and random rotation (the angle of 90, 180 and 270 degree) were performed to the original datasets for the training size enlarging. All the programmes were implemented with the version of Python

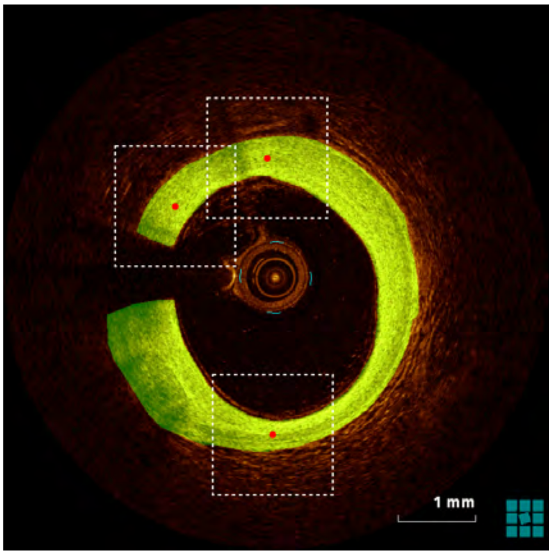


Fig. 6.2 White dotted square denotes the cropped region for the deep learning model. Red points indicate the center pixel belonged to ROI. Hundreds of cropped square patches as the input data are generated from ROI of one single IVOCT image.

3.6, Pytorch 1.0.0, CUDA 9.0.176 and the image processing libraries were OpenCV 2.0, Pillow 7.1.1 and Scikit-image 0.16.2.

6.4.4 Validation

Mean absolute difference of ROI area (MAD_{area}) and the Dice coefficient were respectively applied to validate the difference and the similarity between the manual and automatic methods on the ROI segmentation of IVOCT imageS. Both formulas are shown as follows:

$$MAD_{area} = \frac{1}{N}\sum_{n=1}^{N} |A^n - B^n|, \tag{6.4}$$

$$Dice(A,B) = \frac{2|A \cap B|}{|A|+|B|}, \tag{6.5}$$

where A and B denote segmentation results of manual approach and our proposed method respectively, N is the amount of IVOCT image in one data set, $Dice(A,B)$ indicates the Dice coefficient calculating the overlap of A and B.

We applied ten-fold cross-validation to the training data to assess the quality of pixel-wise classification of vessel inner tissues. Each fold was split into three parts: training set (80%), validation set (10%) and test set (10%) for evaluation of the plaques recognition. The test set

Table 6.1 Evaluation metrics for the segmentation of ROI in IVOCT images.

Dataset	**Dice coefficient**	$\mathbf{MAD_{area}}$ ($\mathbf{mm^2}$)
set 0	0.77±0.11	2.72±1.55
set 1	0.68±0.16	1.55±0.95
set 2	0.78±0.14	0.56±0.48
set 3	0.78±0.06	0.98±0.46
set 4	0.61±0.22	2.26±1.14
set 5	0.78±0.05	0.50±0.29
set 6	0.63±0.28	1.70±1.03

was the fully new data that never be used in the training set or validation set. Evaluation of the segmentation results obtained for test datasets were measured with sensitivity, specificity, pixel accuracy (PA) and mean intersection over union (MIoU) to demonstrate the effect of our deep learning neural network. Moreover, a comparison of the semantic segmentation results of IVOCT images between our model and SegNet model was also employed.

$$PA = \frac{\sum_{i=0}^{k} p_{ii}}{\sum_{i=0}^{k}\sum_{j=0}^{k} p_{ij}}, \tag{6.6}$$

$$MIoU = \frac{1}{k+1}\sum_{i=0}^{k}\frac{p_{ii}}{\sum_{j=0}^{k} p_{ij} + \sum_{j=0}^{k} p_{ji} - p_{ii}}, \tag{6.7}$$

where k is the total number of the categories, p_{ij} is the number of pixels predicted with class j while i is the true amount of it. p_{ii} denotes that the total number of pixels for which the true pixel category i is predicted to be category$i$.

6.4.5 Results

Figure 6.4 shows the initial ROI segmentation results with our method. The outer border of ROI is obtained based on the extended ALSR through the level-set methodology. We utilized the labeling region as the ground-truth of ROI directly to assess the automatic ROI segmentation results. Fig. 6.3 gives out the examples of our ROI segmentation method and the corresponding results manually labeled by experts. Table 6.1 demonstrates the Dice and

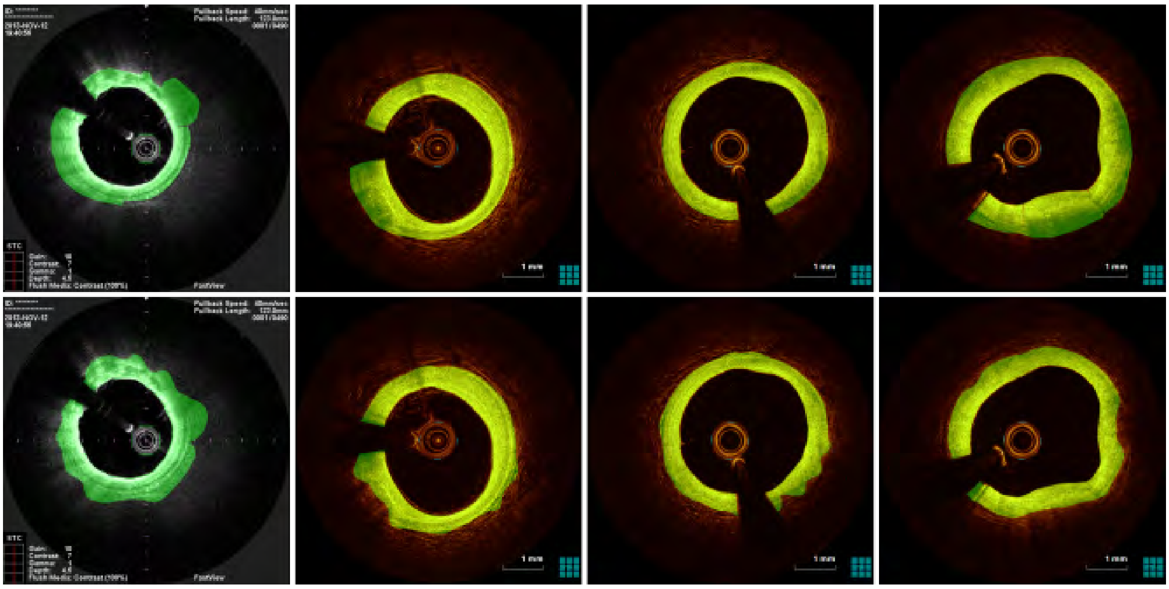


Fig. 6.3 Top: Ground-truth, which was labeled by specialists. Bottom: Segmentation results of ROI with our automatic method based on the level-set model.

MAD_{area} metrics between our proposed method and the manual work, that were applied to the 7 original datasets. The scale of the IVOCT image utilized in our experiments was 100 $pixels/mm$.

A 3-D volume with its depth as same as the number of cropped input is structured for the prediction of each pixel by computing the maximum class amount of the pixel in its current location (Fig. 6.7). Each prediction result is put in the right position corresponding to the position of input, which the size of every slice is the same as the original IVOCT image. Here, we construct regions that its scale is as same as the original IVOCT image, in which the cropped patches are put into the right position of each defined region. All these regions are combined to structure a 3-D volume with each slice contains the trimmed square prediction patch. The depth of the defined 3-D volume is the number of cropped square patches.

Three groups of the semantic segmentation instance with our method applying to the test datasets are presented in Fig. 6.8. Various types containing fibrous, calcific, lipidic plaques and the health vessel wall are displayed and the corresponding ground truth and the results of vessel tissue classification are demonstrated in Fig. 6.8(b) and Fig. 6.8(c). Four color maps corresponding to four types of vessel tissue are used to illustrate situations of tissue distribution. Ten individual test datasets were used for the model testing to evaluate the pixel-wise segmentation capability of our deep learning model with the segmentation metrics. Observing from Tab. 6.3 denoting the classifications of vessel tissue computed for each measure, results present that our methodology achieves an attractive effect on the lesion plaques classification. We used SegNet architecture as the baseline to compare the segmentation results with our deep learning model (Tab. 6.2). Here, we used the testing model by setting the number of layers in all the dense blocks as 3. MeanPA and MeanIoU in Tab. 6.2 respectively indicate the mean pixel accuracy and mean intersection over union. The comparison shows that MeanPA and MeanIoU of our model improve 6.13% and 4.67% than the related results of SegNet.

Table 6.2 Comparison of tissue classification between SegNet and DB-SegNet-3 (DB-SegNet-3 model presents the number of layers of a dense block is set 3.)

Model	**MeanPA(%)**	**MeanIoU(%)**
SegNet	81.06	72.87
DB-SegNet-3	87.19	77.54

6.5 Discussion and conclusion

As mentioned, traditional methods have won a progressive performance on tissue identification or the lesion plaques recognition. Many approaches[55, 66, 71, 88] focused on the investigation of A-line characteristics, such as A-line peaks alteration, because of each A-line presenting an intensity profile that the intensity amplitude varies as the depth increasing. With this characteristic, the A-line research approach also performed a significant role in the clinical research of the lumen boundary segmentation and stent struts detection, etc. Besides, considering the vessel tissue with attribution of absorption and scattering and assuming that the tissue is homogeneous, the attenuation coefficient (μ_t) is used as one feature representation for every pixel in an IVOCT image[66, 83]. In spite of the fact that utilizing the A-line investigation method could extract features along the radius-direction in Cartesian domain, the information along the circumferential-course is not considered[70]. Clearly, only 1-D information (intensity profile or μ_t) is taken into consideration. Related information including the circumferential distribution of tissue is lost. Other literature [11, 32, 83] utilized machine learning methods or CNN-like modes as the feature extractor for pixel-wise segmentation of lesion tissue, and obtained agreeable outcomes. However, it is actually a wasting-time procedure and low efficiency when executing a training action of tissue identification pixel-by-pixel, particularly facing massive IVOCT images.

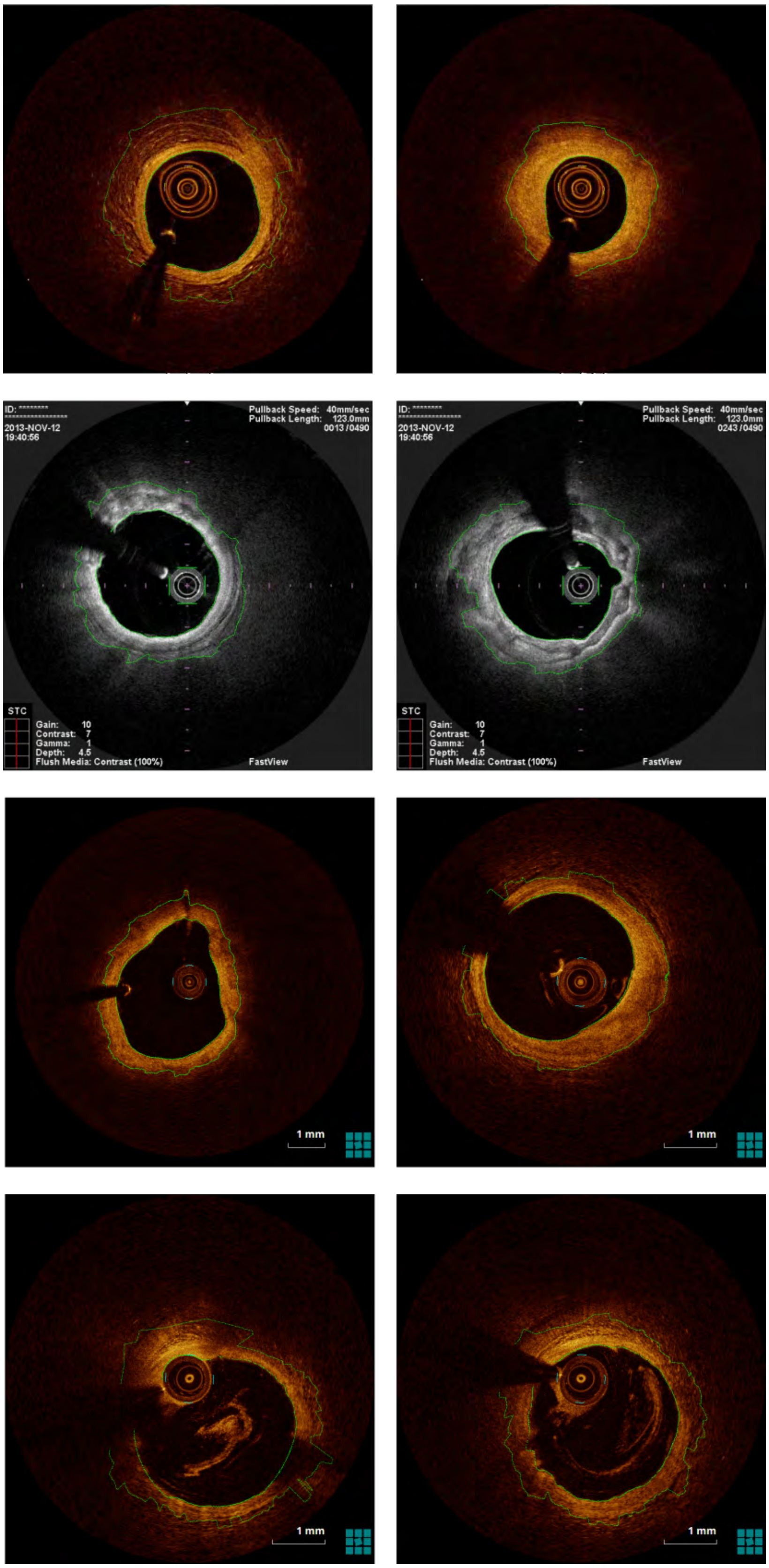


Fig. 6.4 Initial outer border (green line) detected with level-set method combines with the lumen border to construct ROI.

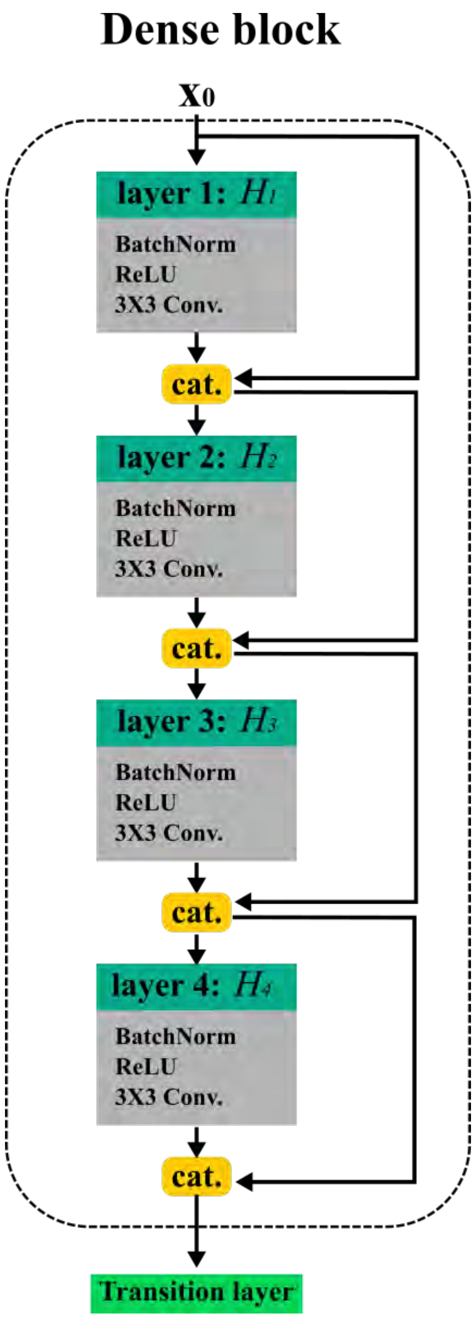


Fig. 6.5 In the dotted rectangle, X_0 is the initial input data. The example of a dense block is composed of four layers that each layer consists of continuous operations of BN, ReLU and a 3×3 Conv. A concatenation operation marked with yellow is performed to join previous feature maps with the output of the current layer. Subsequently, the concatenated feature maps are used as a new input for the next layer. The feature dimension of output for each layer is compressed to a fixed size of k that is described as the *growth rate* of the network. With the concatenation from previous layers to subsequent layers, the output dimension of the dense block grows linearly. Transition layer here is connected to the dense block for reducing the feature maps dimension and the spatial size of input data.

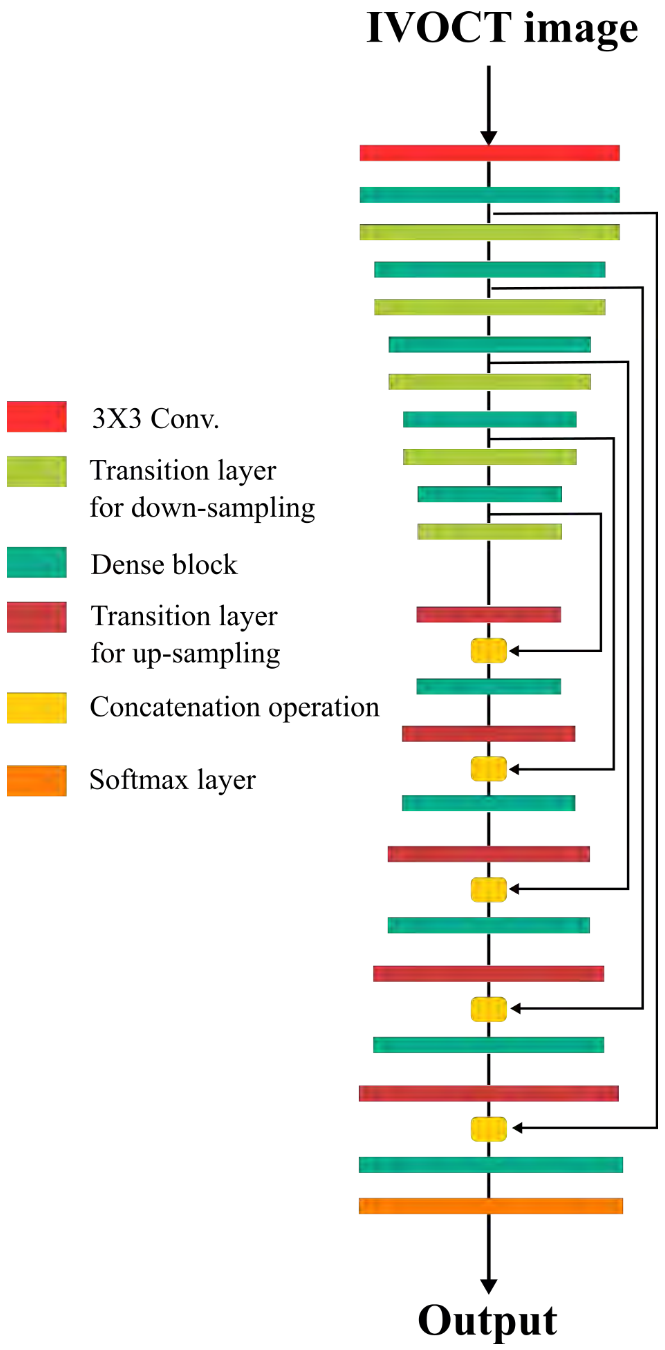


Fig. 6.6 The two-part architecture of DB-SegNet is the downsampling path and the upsampling path respectively. The main foundational units of DB-SegNet are dense block (jungle green), downsampling layer (yellow-green) and upsampling layer (brown). First, a 3×3 convolutional layer (blush) is applied to produce feature maps. Afterwards, 5 couples of dense block and downsampling layers are used for feature extraction and spatial reduction. In the latter part of our model, a sequence of upsampling layers and dense blocks followed by a *softmax* layer is combined to produce a desired prediction of the vessel tissue classification. In order to utilize the available information generated from the previous dense blocks and gain deeper feature maps, a concatenation is applied to join the output of the upsampling layer and the corresponding feature maps from the preceding dense blocks.

Table 6.3 Evaluation metrics over 10 datasets for the vessel tissue classification

Dataset	**Validation Metrics**															
	Sensitivity (%)					**Specificity (%)**					**PA (%)**					**MIoU (%)**
	O[1]	C[1]	F[1]	L[1]	H[1]	O	C	F	L	H	O	C	F	L	H	
set 0	98.57	83.19	90.56	91.64	91.36	94.54	90.15	88.46	92.94	93.70	97.80	94.91	90.95	90.28	87.41	80.83
set 1	93.68	93.79	92.10	89.90	92.11	94.46	89.94	91.30	93.29	96.56	97.98	93.90	93.78	94.54	89.35	75.17
set 2	90.66	93.75	89.89	92.81	93.30	95.03	84.91	92.62	94.16	89.61	98.07	88.85	89.19	93.54	90.37	78.69
set 3	95.10	95.69	90.80	92.79	91.43	95.40	93.94	95.46	96.91	86.72	97.93	82.90	91.85	92.32	93.67	81.32
set 4	97.52	90.52	92.95	93.13	93.84	90.92	96.94	91.49	93.11	93.61	98.04	90.88	92.08	90.60	94.40	83.10
set 5	98.58	95.06	89.84	91.27	93.76	91.23	95.93	90.40	90.95	97.63	97.89	93.18	98.63	89.30	92.35	71.94
set 6	96.37	92.56	91.41	91.74	91.59	94.34	90.95	89.60	99.31	92.66	98.07	95.30	94.13	92.68	91.41	75.31
set 7	92.27	92.56	93.23	90.27	92.57	95.43	96.95	94.19	98.31	99.68	97.75	94.89	93.12	93.11	96.43	80.46
set 8	90.66	90.95	91.69	90.75	91.80	94.58	93.93	90.22	94.15	92.31	97.90	90.30	95.20	91.28	89.70	79.01
set 9	98.71	94.73	90.03	90.78	89.59	94.61	93.15	90.66	90.73	97.91	97.92	93.77	89.13	90.15	93.69	81.22

[1] O: *Other unlabeled object*; C: *Calcified plaque*; F: *Fibrous plaque*; L: *Lipid plaque*; H: *Healthy vessel wall*

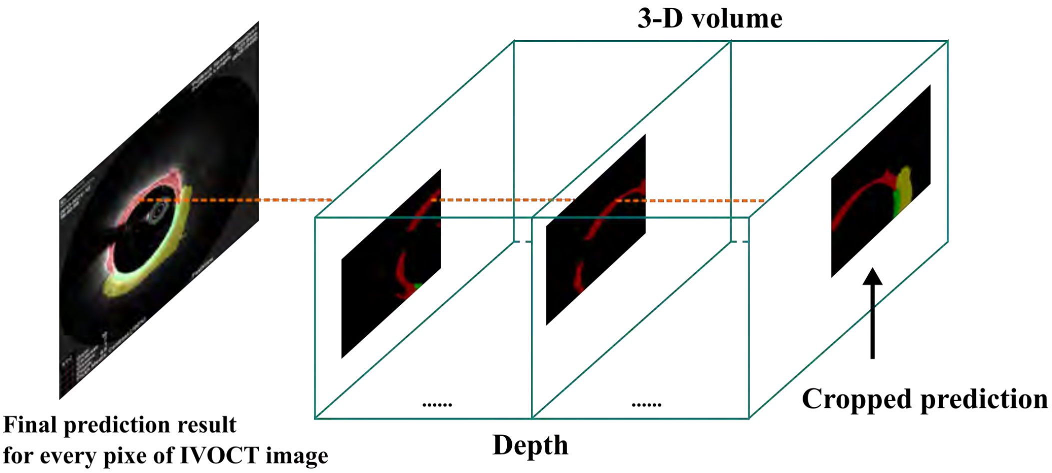


Fig. 6.7 3-D volume for the cropped prediction results is constructed in order to assess the pixel's final class by computing the maximum number of the pixels which are in the same location.

As such, semantic segmentation combining CNN and ROI is a worthy topic for the vessel lesion tissue. Semantic segmentation on the IVOCT images always is a challenging task for the classification of vessel lesion tissues at a pixel-level.

Due to the light attenuation existing, the region outward from the image center or the vessel lumen contains a great of useless information, and only superficial tissue nearby the lumen border could be easily identified. Therefore, the ROI of vessel tissue, an area between the lumen boundary and an outer border, was characterized and used for tissue type classification and to reduce the task of tissue recognization at the same time. In some studies, the thickness of ROI was set with a fixed value[11, 18, 96].

In this paper, a fully automatic ROI segmentation method and a deep learning neural network architecture for semantic segmentation of vessel tissue were proposed. We did not directly utilize the binarization method to segment the ROI, e.g., the Otsu method. Undoubtedly, obtaining ROI with Otsu is a speedy and effective approach, yet the lumen boundary or the outer boundary might be non-continuous and unsmooth, as shown in Fig. 6.9 (a). Some edge information that is significant to the pixel-wise segmentation may be lost. Part area of lipidic plaque might be also missed because of its character of the diffuse border. Meanwhile, we did not determine an ROI with a fixed thickness, e.g. $1mm$[11, 18, 96]. As mentioned in Sec. 6.1, different tissues exist dissimilarity of thickness. Based on the ALSR defined in [70], we used the level-set model to divide each extended ALSR into 2 parts: ROI and the outer worthless area. The level-set method gradually iterates the contour to minimize the energy of the inside area and outside area. Hence, a stable contour could be gained to include the maximum useful region as ROI (Fig. 6.9(b)). For the 7 datasets, most ROIs

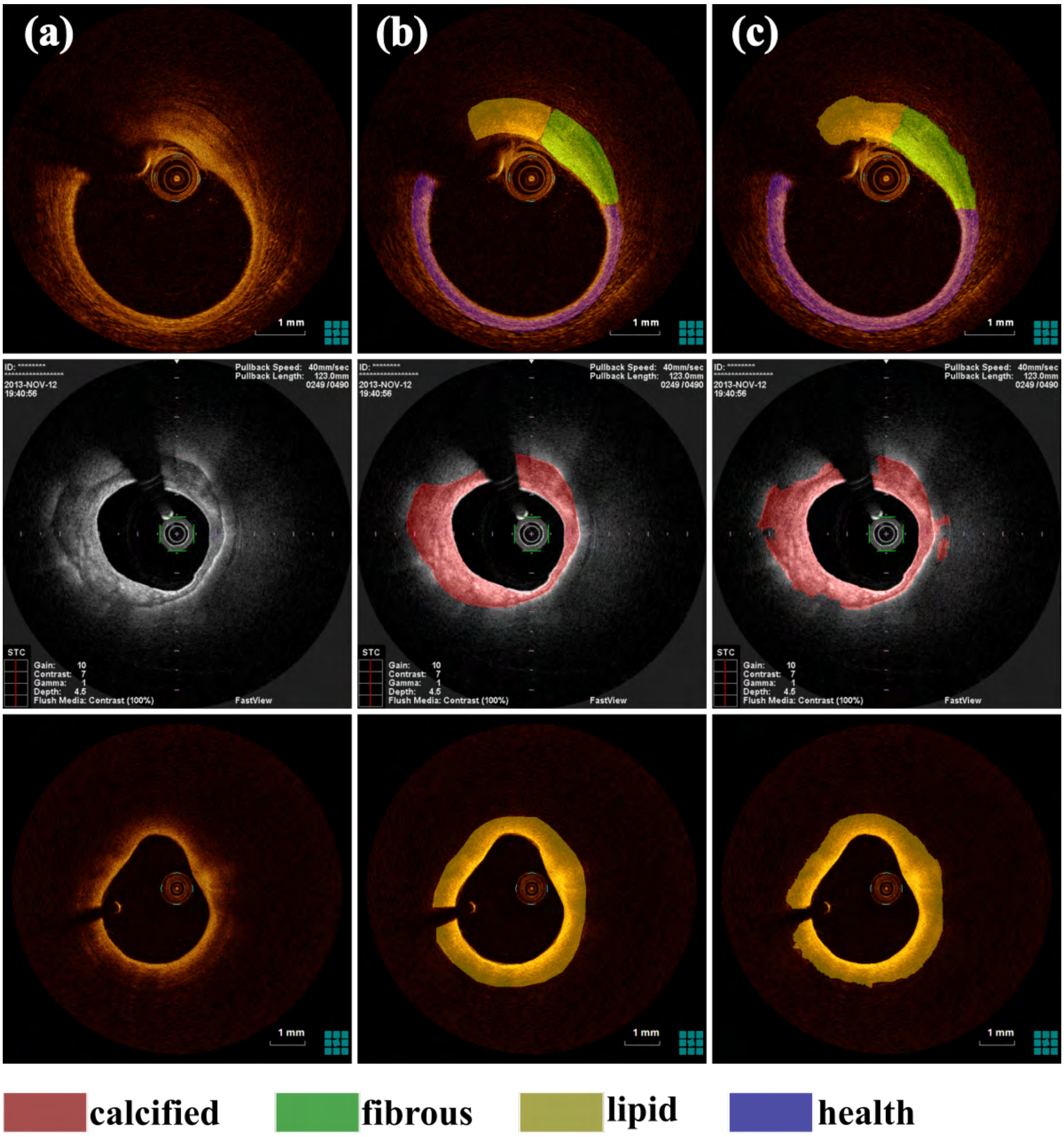


Fig. 6.8 Successful pixel-wise classification examples from certain datasets. column-(a) IVOCT image, column-(b) ground-truth, and column-(c) illustrates the segmentation results obtained with our proposed deep learning neural network. The annotation colors for each tissue is denoted at the bottom of the resulting plane.

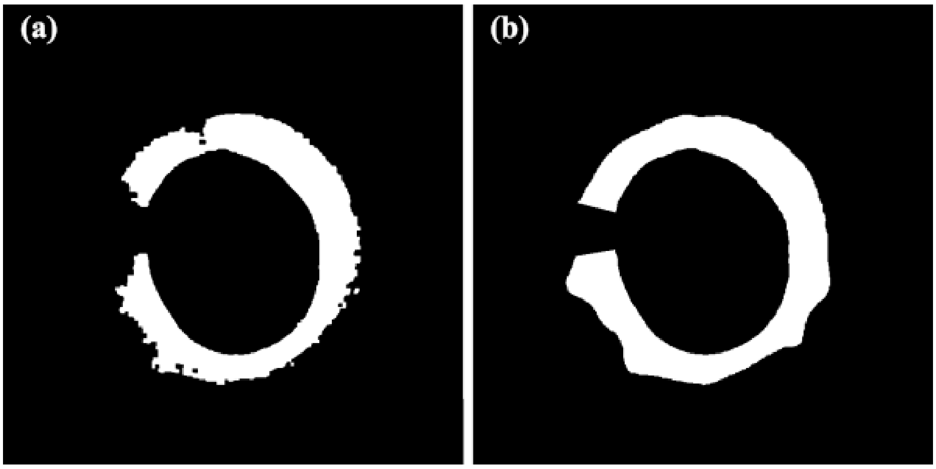


Fig. 6.9 After using the Otsu method, morphology operations, including opening and closing, are applied to the IVOCT image to get the final segmentation result displayed in (a). (b) is the ROI segmentation with our method.

present agreeable results containing useful information of the vessel tissue. From Tab. 2.1, the uncertain determination of lipid core with a diffuse border can cause an impact on the score of the Dice coefficient. This automated ROI segmentation strategy can be an aid tool to help cardiovascular specialists procuring attention areas in IVOCT images.

Although applying machine learning technologies (e.g., SVM and Random forest) combining texture information have accomplished an outperformance, the imaging conditions, such as the location of the catheter and the speed of pullback, really have effects on the image quality, which cause a difficult problem to vessel image analysis. Besides, the machine learning-based method for the pixel-wise segmentation could be impacted by the noise, small human artifacts, etc, and a lesion plaque with the same type in different patients would present various appearances, which is also a significant factor. Zhang et al.[96] contrasted pixel classification of tissue both with CNN and SVM, and they proved that CNN-based outcomes performed better than those achieved from SVM. As shown in Fig. 6.10, the left plot presents the lesion tissue classification by using Athanasiou's method[11] based on the feature extraction and Random forest classifier, and the right illustration is our classification result. However, the cropped approach of input data in Zhang's paper lost certain significant information of the visible district, which decreased the scope of recognizable areas. Furthermore, it is convenient that feeding a complete IVOCT image into the deep learning model directly[47, 66], but it would greatly expand the learning time and reduce efficiency. In our proposed method, we acquired the input image by cropping a square from the ROI obtained in Sec. 6.1 to beat the above issues. Our cropping strategy is to produce hundreds

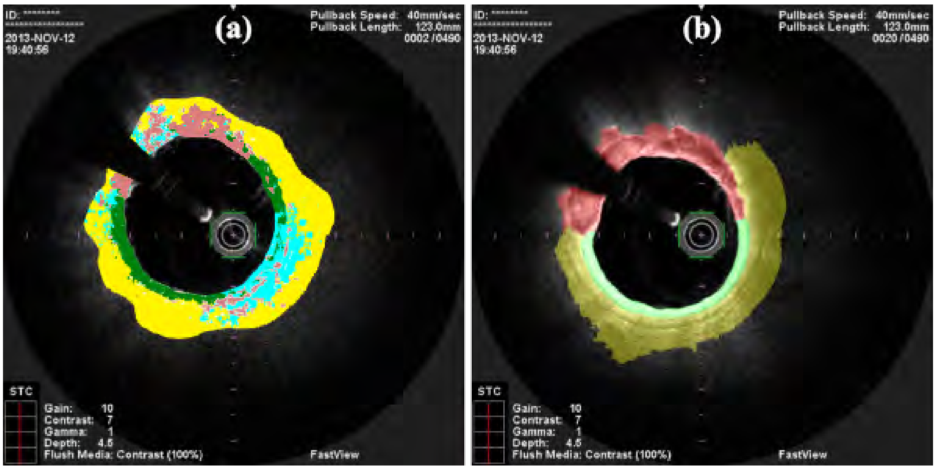


Fig. 6.10 (a) is implemented with Athanasiou's method, and (b) presents our classification results of lesion tissues.

of sub-images (or patches) from one single IVOCT image as the input image data to the DB-SegNet and each input image is 320×320 pixels. The content of trimmed input images from ROI contains a great deal of useful data about superficial tissues. Therefore, these trimmed patches can be used as the input data of our model. Besides, our cropping method could extend the volume of the training data-set and increased the generalization ability of our model.

To reuse the previous feature maps to obtain the information of the position and edge of vessel tissues, we utilized the dense block as the basic unit to construct a deep learning model with "end-to-end" architecture. This model is consists of 2 parts corresponding to the downsampling path and upsampling path, as shown in Fig. 6.6. Skip connections from the downsampling path to the upsampling path were executed in our model to supply high-resolution information to the recovering operation of the upsampling path. Concatenating operation occurred not only in the dense block but also between downsampling layers and upsampling layers. Thus, spatial information was fully joined with abstract information to help the vessel tissue semantic segmentation. Gharaibeh et al.[26] and Lee et al.[47] both applied SegNet to segment lesion plaques and achieved a sensitivity of $85.0 \pm 4.0\%$ and $85.1 \pm 7.2\%$ respectively after post-processing using the fully connected conditional random field (CRF) to the initial segmentation calcification while our sensitivity score is $92.28 \pm 3.43\%$ without CRF. In papers [26, 47, 66], CRF is employed for refinement of initial prediction results by smoothing region and preventing isolated spots. In our approach, we respectively refined the semantic segmentation results with CRF, traditional morphological methods (removing small holes and object) and the combination approach (morphology +

CRF). As shown in Fig. 6.11, the appearance of the best post-processing result is obtained with the methodology of conventional morphological methods, which is displayed in the fourth column. Here, we selected the conventional morphology approaches or the post-processing of IVOCT image predictions to fill the small hole and remove small unnecessary spots. The sensitivity of lipid plaque segmentation in [47] is $87.74 \pm 7.2\%$ after CRF processing, while our validation outcome is $91.51 \pm 1.06\%$ without CRF. From Tab. 4.2, it can be seen that both MeanPA and MeanIoU are improved comparing with SegNet.

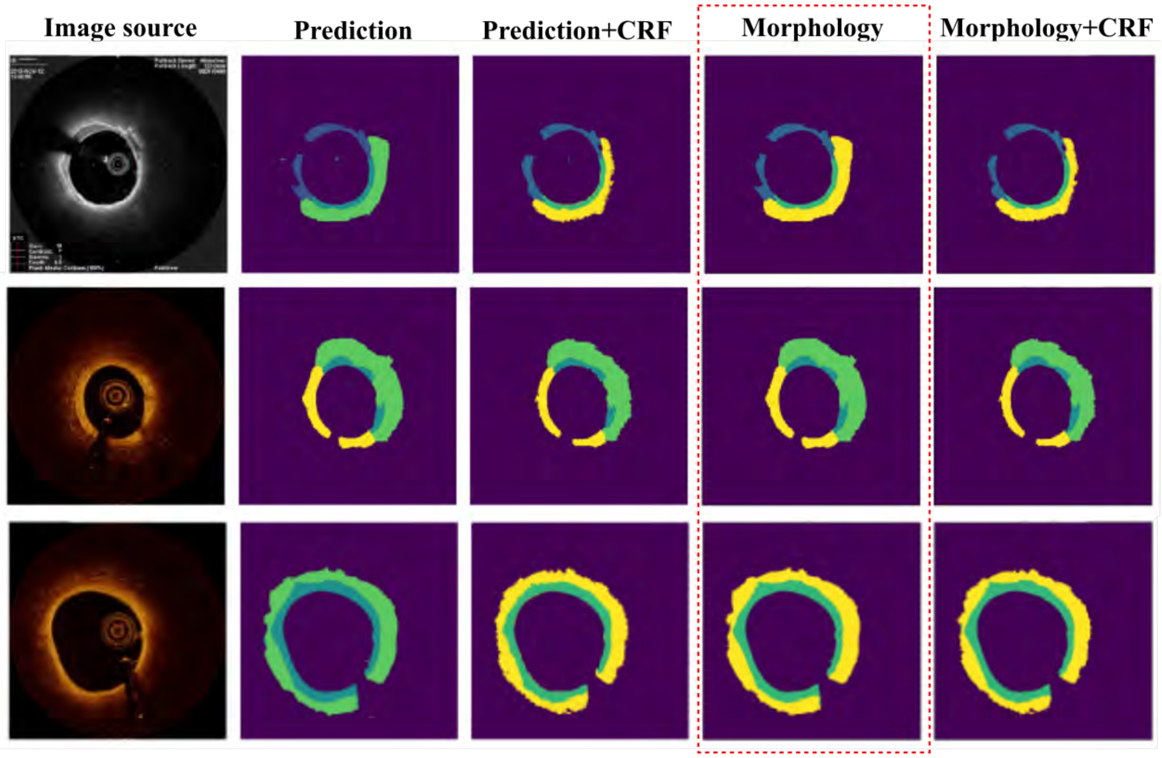


Fig. 6.11 Comparison of original prediction results, CRF processed results, morphological operations (only) and morphology processing combined with CRF. We find that the prediction results (red dotted) with morphological operation perform better in our method.

Note that, in view of our cropping method, certain pixels are shared between trimmed patches that are belonged to one IVOCT image, namely, the same pixel may appear in different locations of many cropped patches. We recorded the location information of the selected pixel, including the original center point coordinates and the four related vertex coordinates of each cropped patch. At the post-processing stage, we constructed a 3-D volume (Fig. 6.7) with the shape of $[H \times W \times C]$ for the purpose of combining all the predictions belonging to a single IVOCT image to a final result, where H and W are the height and width of the original IVOCT image, C is the number of patches, which denotes the depth of the 3-D volume. Each prediction of the trimmed patches was filled to the corresponding position of every slice. As a result, slices containing trimmed prediction patches construct a 3-D volume where the pixels with the same positions appear in different patches. For each IVOCT image, the final type of each pixel is determined by computing its statistical data of the maximum number along the C axis.

In this chapter, we present an automatic semantic segmentation method to classify pixels of the vessel inner tissue in IVOCT images based on the deep learning neural network (DB-SegNet). A cropping strategy based on the pixels of ROI was employed to yield input image data with a size of 320×320, where the trimmed image patches contained enough effective and worth information of the vessel tissue. We constructed an “end-to-end” deep learning neural network using the dense block as the basic element. This architecture is comprised of two paths for image downsampling and upsampling. Each path contains 5 dense blocks and corresponding transition layers. Validation results show that our proposed method achieves a promising exhibition on the pixel-wise segmentation of tissue in IVOCT images and could be used as a clinical assist analysis tool for specialists researching atherosclerotic plaques.

Chapter 7

Discussion and future work

7.1 Conclusion

In this thesis, I focus on the investigation of intravascular optical coherence tomography images analysis, including the schema and characterization of OCT image modality. I propose algorithms for the pre-processing of catheter removal and GW region segmentation, develop automatic methods for lumen boundary and stent struts detection, present a local multi-layer region model to analyze and extract the features from the superficial layers in IVOCT images for lesion plaques identification and classification, build up two deep learning models (VGG-like model and DB-SegNet) to discuss lesion tissue recognition with CNN.

Chapter 2 mainly discusses methodologies applied in the pre-processing step for the catheter imaging region removing, the GW shadow segmentation and human artifacts elimination. The contributions in this chapter include:

1. A circle detector (in Sec. 2.4) is built up to extract pixels in the catheter imaging region matching the condition of circumferential difference. I divide a single circle into two half ones which are used to compute the intensity difference of these two half-circles, which is used as the statistic information characterizing the catheter imaging circle shape. Parameters are employed to fit the condition of the distorted catheters.

2. I propose an algorithm that using a CRDM to detect the angle of the GW black shadow through the sector features defined Sec. 2.5. Our method iteratively increases the radius of CRDM and finds out a circle ring containing a maximum number of black sub-segment as the candidate ring. Utilizing the statistic method to obtain the GW angle.

3. For speckle noise elimination and the residual blood removing, I apply a Gaussian filter, and a series of morphological operations implemented to the binarized IVOCT images.

Chapter 3 investigates an automatic method for the lumen boundary segmentation and stent strut detection.

1. I propose a method to compute the sum of local standard deviation distributing both sides of the center point by considering the changing trend presented by the intensity of A-line profile.
2. I compute the left side of the STDI to obtain the candidate points of vessel lumen boundary and eliminate the fake points by our algorithm before smooth the luminal border.
3. The candidate points belong to stent struts are judged through the feature that the intensity variance after the peak point of each A-line, statistic variables are designed for the final result.
4. In this chapter, I remove the GW, which presents a metallic bright spot, by detecting the maximum area gained with the adaptive region growing method to the candidate points.

Chapter 4 researches on the atherosclerosis plaque recognition with a local multi-layer model defined in the PLBR (superficial layer of the vessel wall).

1. I examine and experiment with the light attenuation and the thickness for every vessel tissue.
2. To overcome the problem of the single A-line only with 1-D appearance and information, I combine the adjacent A-lines into a defined ALSR to construct a local area composed with A-lines.
3. I produce three layers of ALSR with fixed depths, and each ALSR features are extracted to investigate the intensity statistic information and tissue texture, RF classifier is utilized for these features for the ALSR types recognition.

Chapter 5 builds up a deep learning model based on VGG-Net to extract the deep and abstract features that can not easily be detected, simultaneously, three types of image channel are designed for the texture and mixed multi-feature effect to the prediction result.

1. I directly use ALSRs as the input data of the VGG-like model in three types of channels.

2. I use LBP to process the IVOCT images to generate LBP pattern images with one channel, and merge this LBP image with RGB channels IVOCT image to produce four channels combined input data, while other two kinds of input data is one channel LBP input and RGB channels.
3. An eleven-layer VGG-like deep learning model is created for the deep learning method on the research of tissue classification of the superficial layer of the human vessel wall.

Chapter 6 continues to focus on the deep learning method for the semantic segmentation of atherosclerosis plaques in IVOCT images.

1. I divide the vessel wall into several regions based on the definition of the ALSR width, and employ level-set method on each region to obtain the outer boundary of ROI.
2. Pixels in ROI are selected as the center points to crop a square region with a size of 320×320 as the input data.
3. These cropped data is fed into a designed deep learning semantic segmentation model named DB-SegNet for the pixel-wise classification.
4. To predict the final class of each pixel, I construct a 3-D volume structured with all the segmented results from one IVOCT image, where the location of these segmented results are as same as the original cropped input of the IVOCT image. To determine the final class of each pixel, the maximum class possibility of each pixel is calculated as the prediction result.

7.2 Problems and future work

Although our research performs better outcomes on the IVOCT image analysis, some issues still exist and need to be overcome in the future. In the GW region segmentation, the divided line for the GW shadow and vessel wall is blurred in the original IVOCT images, sometimes, due to the imaging modality conditions. It causes us to mislead our GW segmentation method to obtain a GW angle that may larger than the ground-truth. In the lumen boundary detection, the fake points may not be eliminated completely, which causes a mistake in the border segmentation. Additionally, I didn’t discuss the complex situation with rupture in the human vessel. Robustness methods for the GW and lumen boundary segmentation should be experimented with considering more CAD cases for clinical research. The ALSR as an entirety part of the local multi-layer model is utilized to distinguish tissue type, but in the view of morphology and histology of vessel, mixed tissue indeed appearances in the ALSR.

The method with ALSR predicts the type of ALSR containing the most possible plaques. Our aim is to recognize the angle range of the lesion plaque through the local multi-layer method, then consider a new method to identify the plaque type contained in the ALSR, including the mixed plaque case. Combining the level-set with the binarized IVOCT image acquired with morphological operations is the next step for the research of ROI, in order to improve the ROI segmentation results to reduce the difference compared with the manual labelling. I only discuss the cases that applying semantic segmentation technique of deep learning to the vessel without any stent implantation, however, analyzing the vessel inner tissue situation after stent implantation is also an very important work to specialists who execute the task of periodic examination of patients, therefore, the aim of the future work is to apply deep learning method to the stent-implantated vessel.

References

[1] (2018). Cardiovascular disease. https://www.nhs.uk/conditions/cardiovascular-disease/.

[2] (2019). Cardiovascular disease. https://en.wikipedia.org/wiki/Cardiovascular_disease/cite_note-WHO2011-2/.

[3] (2019). Coronary artery disease. https://medlineplus.gov/coronaryarterydisease.html/.

[4] (2020). Atherosclerosis. https://en.wikipedia.org/wiki/Atherosclerosis/cite_note-:1-53/.

[5] Abdelsamea, M. M. (2014). An automatic seeded region growing for 2d biomedical image segmentation. *arXiv preprint arXiv:1412.3958*.

[6] Abdolmanafi, A., Duong, L., Dahdah, N., Adib, I. R., and Cheriet, F. (2018). Characterization of coronary artery pathological formations from oct imaging using deep learning. *Biomedical Optics Express*, 9(10):4936–4960.

[7] Abdolmanafi, A., Duong, L., Dahdah, N., and Cheriet, F. (2017). Deep feature learning for automatic tissue classification of coronary artery using optical coherence tomography. *Biomedical optics express*, 8(2):1203–1220.

[8] Adams, R. and Bischof, L. (1994). Seeded region growing. *IEEE Transactions on pattern analysis and machine intelligence*, 16(6):641–647.

[9] Adriaenssens, T., Ughi, G. J., Dubois, C., Onsea, K., De Cock, D., Bennett, J., Wiyono, S., Vanhaverbeke, M., Sinnaeve, P., Belmans, A., et al. (2014). Automated detection and quantification of clusters of malapposed and uncovered intracoronary stent struts assessed with optical coherence tomography. *The international journal of cardiovascular imaging*, 30(5):839–848.

[10] Amini, A. A., Weymouth, T. E., and Jain, R. C. (1990). Using dynamic programming for solving variational problems in vision. *IEEE Transactions on pattern analysis and machine intelligence*, 12(9):855–867.

[11] Athanasiou, L. S., Bourantas, C. V., Rigas, G., Sakellarios, A. I., Exarchos, T. P., Siogkas, P. K., Ricciardi, A., Naka, K. K., Papafaklis, M. I., Michalis, L. K., et al. (2014). Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *Journal of biomedical optics*, 19(2):026009.

[12] Athanasiou, L. S., Exarchos, T. P., Naka, K., Michalis, L. K., Prati, F., and Fotiadis, D. I. (2011). Atherosclerotic plaque characterization in optical coherence tomography images. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 4485–4488. IEEE.

[13] Bezerra, H. G., Costa, M. A., Guagliumi, G., Rollins, A. M., and Simon, D. I. (2009). Intracoronary optical coherence tomography: a comprehensive review: clinical and research applications. *JACC: Cardiovascular Interventions*, 2(11):1035–1046.

[14] Bland, J. M. and Altman, D. G. (1996). Statistics notes: measurement error. *Bmj*, 312(7047):1654.

[15] Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on pattern analysis and machine intelligence*, (6):679–698.

[16] Cao, Y., Cheng, K., Qin, X., Yin, Q., Li, J., Zhu, R., and Zhao, W. (2017). Automatic lumen segmentation in intravascular optical coherence tomography images using level set. *Computational and mathematical methods in medicine*, 2017.

[17] Caselles, V., Kimmel, R., and Sapiro, G. (1997). Geodesic active contours. *International journal of computer vision*, 22(1):61–79.

[18] Celi, S. and Berti, S. (2014). In-vivo segmentation and quantification of coronary lesions by optical coherence tomography images for a lesion type definition and stenosis grading. *Medical image analysis*, 18(7):1157–1168.

[19] Chan, P. H., Alegria-Barrero, E., and Di Mario, C. (2012). Intravascular ultrasound and optical coherence tomography.

[20] Chan, T. F. and Vese, L. A. (2001). Active contours without edges. *IEEE Transactions on image processing*, 10(2):266–277.

[21] Cheimariotis, G., Riga, M., Toutouzas, K., Tousoulis, D., Katsaggelos, A., and Maglaveras, N. (2019). Automatic characterization of plaques and tissue in ivoct images using a multi-step convolutional neural network framework. In *World Congress on Medical Physics and Biomedical Engineering 2018*, pages 261–265. Springer.

[22] Chiastra, C., Montin, E., Bologna, M., Migliori, S., Aurigemma, C., Burzotta, F., Celi, S., Dubini, G., Migliavacca, F., and Mainardi, L. (2017). Reconstruction of stented coronary arteries from optical coherence tomography images: feasibility, validation, and repeatability of a segmentation method. *PloS one*, 12(6):e0177495.

[23] Dierckx, P. (1995). *Curve and surface fitting with splines*. Oxford University Press.

[24] Dubuisson, F., Péry, E., Ouchchane, L., Combaret, N., Kauffmann, C., Souteyrand, G., Motreff, P., and Sarry, L. (2015). Automated peroperative assessment of stents apposition from oct pullbacks. *Computers in Biology and Medicine*, 59:98–105.

[25] Gessert, N., Lutz, M., Heyder, M., Latus, S., Leistner, D. M., Abdelwahed, Y. S., and Schlaefer, A. (2018). Automatic plaque detection in ivoct pullbacks using convolutional neural networks. *IEEE transactions on medical imaging*, 38(2):426–434.

[26] Gharaibeh, Y., Prabhu, D., Kolluru, C., Lee, J., Zimin, V., Bezerra, H., and Wilson, D. (2019). Coronary calcification segmentation in intravascular oct images using deep learning: application to calcification scoring. *Journal of Medical Imaging*, 6(4):045002.

[27] Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323.

[28] Gnanadesigan, M., Hussain, A. S., White, S., Scoltock, S., Baumbach, A., van der Steen, A. F., Regar, E., Johnson, T. W., and van Soest, G. (2017). Optical coherence tomography attenuation imaging for lipid core detection: an ex-vivo validation study. *The international journal of cardiovascular imaging*, 33(1):5–11.

[29] Guo, X., Tang, D., Molony, D., Yang, C., Samady, H., Zheng, J., Mintz, G. S., Maehara, A., Wang, L., Pei, X., et al. (2018). A machine learning-based method for intracoronary oct segmentation and vulnerable coronary plaque cap thickness quantification. *International Journal of Computational Methods*, page 1842008.

[30] Gurmeric, S., Isguder, G. G., Carlier, S., and Unal, G. (2009). A new 3-d automated computational method to evaluate in-stent neointimal hyperplasia in in-vivo intravascular optical coherence tomography pullbacks. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 776–785. Springer.

[31] Hau, C. C. (2015). *Handbook of pattern recognition and computer vision*. World Scientific.

[32] He, S., Zheng, J., Maehara, A., Mintz, G., Tang, D., Anastasio, M., and Li, H. (2018). Convolutional neural network based automatic plaque characterization for intracoronary optical coherence tomography images. In *Medical Imaging 2018: Image Processing*, volume 10574, page 1057432. International Society for Optics and Photonics.

[33] Hinton, G. E., Osindero, S., and Teh, Y.-W. (2006). A fast learning algorithm for deep belief nets. *Neural computation*, 18(7):1527–1554.

[34] Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., Hee, M. R., Flotte, T., Gregory, K., Puliafito, C. A., et al. (1991). Optical coherence tomography. *science*, 254(5035):1178–1181.

[35] Huang, G., Liu, Z., Van Der Maaten, L., and Weinberger, K. Q. (2017). Densely connected convolutional networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4700–4708.

[36] Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167*.

[37] Jegou, S., Drozdzal, M., Vazquez, D., Romero, A., and Bengio, Y. (2017). The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*.

[38] Jiang, H., He, B., Fang, D., Ma, Z., Yang, B., and Zhang, L. (2013). A region growing vessel segmentation algorithm based on spectrum information. *Computational and mathematical methods in medicine*, 2013.

[39] Kingma, D. P. and Ba, J. (2014). Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*.

[40] Kodach, V., Faber, D., Van Marle, J., van Leeuwen, T., and Kalkman, J. (2011). Determination of the scattering anisotropy with optical coherence tomography. *Optics express*, 19(7):6131–6140.

[41] Kolluru, C., Prabhu, D., Gharaibeh, Y., Bezerra, H., Guagliumi, G., and Wilson, D. (2018). Deep neural networks for a-line-based plaque classification in coronary intravascular optical coherence tomography images. *Journal of Medical Imaging*, 5(4):044504.

[42] Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105.

[43] Kubo, T., Tanaka, A., Kitabata, H., Ino, Y., Tanimoto, T., and Akasaka, T. (2012). Application of optical coherence tomography in percutaneous coronary intervention. *Circulation Journal*, pages CJ–12.

[44] Kullback, S. and Leibler, R. A. (1951). On information and sufficiency. *The annals of mathematical statistics*, 22(1):79–86.

[45] LeCun, Y., Bottou, L., Orr, G. B., and Müller, K.-R. (1998). Efficient backprop. In *Neural Networks: Tricks of the Trade, This Book is an Outgrowth of a 1996 NIPS Workshop*, page 9–50, Berlin, Heidelberg. Springer-Verlag.

[46] LeCun, Y., Huang, F. J., and Bottou, L. (2004). Learning methods for generic object recognition with invariance to pose and lighting. In *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004.*, volume 2, pages II–104. IEEE.

[47] Lee, J., Prabhu, D., Kolluru, C., Gharaibeh, Y., Zimin, V. N., Bezerra, H. G., and Wilson, D. L. (2019). Automated plaque characterization using deep learning on coronary intravascular optical coherence tomographic images. *Biomedical Optics Express*, 10(12):6497–6515.

[48] Li, S. Z. (2012). *Markov random field modeling in computer vision*. Springer Science & Business Media.

[49] Lin, T.-Y., Goyal, P., Girshick, R., He, K., and Dollár, P. (2017). Focal loss for dense object detection. In *Proceedings of the IEEE international conference on computer vision*, pages 2980–2988.

[50] Linda G. Shapiro, G. C. S. (2001). *Computer Vision*. Prentice-Hall Inc.

[51] Lu, H., Gargesha, M., Wang, Z., Chamie, D., Attizani, G. F., Kanaya, T., Ray, S., Costa, M. A., Rollins, A. M., Bezerra, H. G., et al. (2013). Automatic stent strut detection in intravascular oct images using image processing and classification technique. In *Medical Imaging 2013: Computer-Aided Diagnosis*, volume 8670, page 867015. International Society for Optics and Photonics.

[52] Lu, H., Gargesha, M., Wang, Z., Chamie, D., Attizzani, G. F., Kanaya, T., Ray, S., Costa, M. A., Rollins, A. M., Bezerra, H. G., et al. (2012). Automatic stent detection in intravascular oct images using bagged decision trees. *Biomedical optics express*, 3(11):2809–2824.

[53] Macedo, M., Takimura, C., Lemos, P., and Gutierrez, M. (2014). An automatic segmentation method applied to images of iv-oct in human arteries with atherosclerotic plaques. pages 737–740.

[54] Macedo, M. M., Guimarães, W. V., Galon, M. Z., Takimura, C. K., Lemos, P. A., and Gutierrez, M. A. (2015). A bifurcation identifier for iv-oct using orthogonal least squares and supervised machine learning. *Computerized Medical Imaging and Graphics*, 46:237–248.

[55] Macedo, M. M., Nicz, P. F., Campos, C. M., Lemos, P. A., and Gutierrez, M. A. (2016a). Spatial-frequency approach to fibrous tissue classification in intracoronary optical images. In *Computing in Cardiology Conference (CinC), 2016*, pages 477–480. IEEE.

[56] Macedo, M. M. G. d., Takimura, C. K., Lemos, P. A., and Gutierrez, M. A. (2016b). A robust fully automatic lumen segmentation method for in vivo intracoronary optical coherence tomography. *Research on Biomedical Engineering*, 32(1):35–43.

[57] McCarthy, M. J., Loftus, I. M., Thompson, M. M., Jones, L., London, N. J., Bell, P. R., Naylor, A. R., and Brindle, N. P. (1999). Angiogenesis and the atherosclerotic carotid plaque: an association between symptomatology and plaque morphology. *Journal of vascular surgery*, 30(2):261–268.

[58] Mendis, S., Puska, P., Norrving, B., Organization, W. H., et al. (2011). *Global atlas on cardiovascular disease prevention and control*. Geneva: World Health Organization.

[59] Miller, J. D. (2013). Cardiovascular calcification: orbicular origins. *Nature materials*, 12(6):476–478.

[60] Moraes, M. C., Cardenas, D. A. C., and Furuie, S. S. (2013). Automatic lumen segmentation in ivoct images using binary morphological reconstruction. *Biomedical engineering online*, 12(1):78.

[61] Nielsen, M. A. (2015). *Neural networks and deep learning*, volume 2018. Determination press San Francisco, CA.

[62] Ojala, T., Pietikäinen, M., and Harwood, D. (1994). Performance evaluation of texture measures with classification based on kullback discrimination of distributions. *Proceedings of 12th International Conference on Pattern Recognition*, 1:582–585 vol.1.

[63] Ojala, T., Pietikäinen, M., and Harwood, D. (1996). A comparative study of texture measures with classification based on featured distributions. *Pattern recognition*, 29(1):51–59.

[64] Oliveira, D. A., Macedo, M. M., Nicz, P., Campos, C., Lemos, P., and Gutierrez, M. A. (2018). Coronary calcification identification in optical coherence tomography using convolutional neural networks. In *Medical Imaging 2018: Biomedical Applications in Molecular,*

Structural, and Functional Imaging, volume 10578, page 105781Y. International Society for Optics and Photonics.

[65] Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, and cybernetics*, 9(1):62–66.

[66] Prabhu, D., Bezerra, H., Kolluru, C., Gharaibeh, Y., Mehanna, E., Wu, H., and Wilson, D. (2019). Automated a-line coronary plaque classification of intravascular optical coherence tomography images using handcrafted features and large datasets. *Journal of Biomedical Optics*, 24:1.

[67] Preetha, M. M. S. J., Suresh, L. P., and Bosco, M. J. (2012). Image segmentation using seeded region growing. In *2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET)*, pages 576–583. IEEE.

[68] Press, W. H. and Teukolsky, S. A. (1990). Savitzky-golay smoothing filters. *Computers in Physics*, 4(6):669–672.

[69] Ren, X., Chen, Q., Wu, H., Takashi, K., and Takashi, A. (2018). A novel automatic stent detection method in intravascular oct images based on local maximum of standard deviation. In *Proceedings of The International Workshop on Frontiers of Computer Vision*.

[70] Ren, X., Wu, H., Chen, Q., Imai, T., Kubo, T., and Akasaka, T. (2019). A local multi-layer model for tissue classification of in-vivo atherosclerotic plaques in intravascular optical coherence tomography. *IEICE Transactions on Information and Systems*, 102(11):2238–2248.

[71] Rico-Jimenez, J. J., Campos-Delgado, D. U., Villiger, M., Otsuka, K., Bouma, B. E., and Jo, J. A. (2016). Automatic classification of atherosclerotic plaques imaged with intravascular oct. *Biomedical optics express*, 7(10):4069–4085.

[72] Roy, A. G., Conjeti, S., Carlier, S. G., Dutta, P. K., Kastrati, A., Laine, A. F., Navab, N., Katouzian, A., and Sheet, D. (2015). Lumen segmentation in intravascular optical coherence tomography using backscattering tracked and initialized random walks. *IEEE journal of biomedical and health informatics*, 20(2):606–614.

[73] Sethian, J. A. (1999). *Level set methods and fast marching methods: evolving interfaces in computational geometry, fluid mechanics, computer vision, and materials science*, volume 3. Cambridge university press.

[74] Sihan, K., Botha, C., Post, F., de Winter, S., Regar, E., Hamers, R., and Bruining, N. (2008). A novel approach to quantitative analysis of intravascular optical coherence tomography imaging. In *2008 Computers in Cardiology*, pages 1089–1092. IEEE.

[75] Simonyan, K. and Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

[76] Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: A simple way to prevent neural networks from overfitting. *Journal of Machine Learning Research*, 15(56):1929–1958.

[77] Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., and Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*.

[78] Tearney, G. J., Regar, E., Akasaka, T., Adriaenssens, T., Barlis, P., Bezerra, H. G., Bouma, B., Bruining, N., Cho, J.-m., Chowdhary, S., et al. (2012). Consensus standards for acquisition, measurement, and reporting of intravascular optical coherence tomography studies: a report from the international working group for intravascular optical coherence tomography standardization and validation. *Journal of the American College of Cardiology*, 59(12):1058–1072.

[79] Tomasi, C. and Manduchi, R. (1998). Bilateral filtering for gray and color images. In *Sixth international conference on computer vision (IEEE Cat. No. 98CH36271)*, pages 839–846. IEEE.

[80] Tsantis, S., Kagadis, G. C., Katsanos, K., Karnabatidis, D., Bourantas, G., and Nikiforidis, G. C. (2012). Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography. *Medical physics*, 39(1):503–513.

[81] Ughi, G. J., Adriaenssens, T., Desmet, W., and D'hooge, J. (2012a). Fully automatic three-dimensional visualization of intravascular optical coherence tomography images: methods and feasibility in vivo. *Biomedical optics express*, 3(12):3291–3303.

[82] Ughi, G. J., Adriaenssens, T., Onsea, K., Kayaert, P., Dubois, C., Sinnaeve, P., Coosemans, M., Desmet, W., and D'hooge, J. (2012b). Automatic segmentation of in-vivo intra-coronary optical coherence tomography images to assess stent strut apposition and coverage. *The international journal of cardiovascular imaging*, 28(2):229–241.

[83] Ughi, G. J., Adriaenssens, T., Sinnaeve, P., Desmet, W., and D'hooge, J. (2013). Automated tissue characterization of in vivo atherosclerotic plaques by intravascular optical coherence tomography images. *Biomedical optics express*, 4(7):1014–1030.

[84] van der Walt, S., Schönberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., Gouillart, E., Yu, T., and the scikit-image contributors (2014). scikit-image: image processing in Python. *PeerJ*, 2:e453.

[85] van Leeuwen, T. G., Faber, D. J., and Aalders, M. C. (2003). Measurement of the axial point spread function in scattering media using single-mode fiber-based optical coherence tomography. *IEEE Journal of Selected Topics in Quantum Electronics*, 9(2):227–233.

[86] Van Soest, G., Goderie, T. P., Regar, E., Koljenovic, S., van Leenders, A. G. J., Gonzalo, N., van Noorden, S., Okamura, T., Bouma, B. E., Tearney, G. J., et al. (2010). Atherosclerotic tissue characterization in vivo by optical coherence tomography attenuation imaging. *Journal of biomedical optics*, 15(1):011105.

[87] Wang, A., Eggermont, J., Dekker, N., Garcia-Garcia, H. M., Pawar, R., Reiber, J. H., and Dijkstra, J. (2014). Automatic metallic stent strut detection in intravascular optical coherence tomographic pullback runs. *Automatic Quantification of Intravascular Optical Coherence Tomography*, 29(1):19.

[88] Wang, Z., Chamie, D., Bezerra, H. G., Yamamoto, H., Kanovsky, J., Wilson, D. L., Costa, M. A., and Rollins, A. M. (2012a). Volumetric quantification of fibrous caps using intravascular optical coherence tomography. *Biomedical optics express*, 3(6):1413–1426.

[89] Wang, Z., Jenkins, M. W., Bezerra, H., Costa, M., Wilson, D., and Rollins, A. (2012b). Single-shot stent segmentation in intravascular oct pullbacks. pages BTu4B–5.

[90] Wang, Z., Jenkins, M. W., Linderman, G. C., Bezerra, H. G., Fujino, Y., Costa, M. A., Wilson, D. L., and Rollins, A. M. (2015). 3-d stent detection in intravascular oct using a bayesian network and graph search. *IEEE transactions on medical imaging*, 34(7):1549–1561.

[91] Wang, Z., Kyono, H., Bezerra, H. G., Wang, H., Gargesha, M., Alraies, C., Xu, C., Schmitt, J. M., Wilson, D. L., Costa, M. A., et al. (2010). Semiautomatic segmentation and quantification of calcified plaques in intracoronary optical coherence tomography images. *Journal of biomedical optics*, 15(6):061711.

[92] Wang, Z., Kyono, H., Bezerra, H. G., Wilson, D. L., Costa, M. A., and Rollins, A. M. (2011). Automatic segmentation of intravascular optical coherence tomography images for facilitating quantitative diagnosis of atherosclerosis. In *Optical Coherence Tomography and Coherence Domain Optical Methods in Biomedicine XV*, volume 7889, page 78890N. International Society for Optics and Photonics.

[93] Xu, C., Schmitt, J. M., Carlier, S. G., and Virmani, R. (2008). Characterization of atherosclerosis plaques by measuring both backscattering and attenuation coefficients in optical coherence tomography. *Journal of biomedical optics*, 13(3):034003.

[94] Yabushita, H., Bouma, B. E., Houser, S. L., Aretz, H. T., Jang, I.-K., Schlendorf, K. H., Kauffman, C. R., Shishkov, M., Kang, D.-H., Halpern, E. F., et al. (2002). Characterization of human atherosclerosis by optical coherence tomography. *Circulation*, 106(13):1640–1645.

[95] Zeiler, M. D. and Fergus, R. (2014). Visualizing and understanding convolutional networks. In *European conference on computer vision*, pages 818–833. Springer.

[96] Zhang, C., Li, H., Guo, X., Molony, D., Guo, X., Samady, H., Giddens, D., Athanasiou, L., Nie, R., Cao, J., et al. (2019). Convolution neural networks and support vector machines for automatic segmentation of intracoronary optical coherence tomography. *Molecular & Cellular Biomechanics*, 16(2):153–161.

Appendix A

Algorithms

A.1 Adaptive region growing algorithm

Adaptive region growing algorithm

```
Initialize: $Q \leftarrow SeedList, V \leftarrow empty, W \leftarrow empty,$
    $T \leftarrow 0, \mu \leftarrow 0, \sigma \leftarrow 0$
while $Q$ is not empty
    extract $p$ from $Q$
    for each $q$ adjacent to $p$
        if $Dist(p,q) \leq d$
            append $q$ to $T$
    for each $q$ in $T$
        $V = W \cap Q$
            if $V \neq empty$
                calc $\mu, \sigma$
                if $|I(q) - \mu| \leq 2\sigma$
                    append $q$ to $Q$
    set $W = 0$
```

Line 4-6 is to calculate the q adjacent to p with the $Dist(q, p)$(the distance between q and p) less than a parameter d. T is a list to store q that fits for the above condition in each loop. W represents 8-neighbors of each q, and V is the set of $Q \cap W$.

Appendix B

Deep learning architectures

B.1 VGG-like architecture

Input, m=1, 3, 4
Conv.: 3×3, m=64
Conv.: 3×3, m=64
Maxpooling: 2×2, $stride$=2
Conv.: 3×3, m=128
Conv.: 3×3, m=128
Maxpooling: 2×2, $stride$=2
Conv.: 3×3, m=256
Conv.: 3×3, m=256
Conv.: 3×3, m=256
Maxpooling: 2×2, $stride$=2
Dropout(0.5)
FC, m=128
FC, m=128
FC, m=c
Softmax

Architecture details of VGG-like model introduced in Sec. 5.2.3 of chapter 5. m indicates the channel number of each layer. In the input layer, m is set as 1, 3, 4 corresponding to the three types of input channel (LBP, RGB and LRGB). c stands for the number of classes.

B.2 DB-SegNet architecture

DenseLayer
Batch Normalization
ReLU
Conv.: 1×1
Drop
Batch Normalization
Conv.: 3×3
Drop

TransitionDown
Batch Normalization
ReLU
Conv.: 1×1
Drop
Average Pooling: 2×2

TransitionUp
Conv.(Transpose): 3×3, *stride*=2

Architecture
Input, m=3
Conv.: 3×3, m=24
DenseBlock1, TransitionDown
DenseBlock2, TransitionDown
DenseBlock3, TransitionDown
DenseBlock4, TransitionDown
DenseBlock5, TransitionDown
TransitionUp + DenseBlock5
DenseBlockUp5
TransitionUp + DenseBlock4
DenseBlockUp4
TransitionUp + DenseBlock3
DenseBlockUp3

TransitionUp + DenseBlock2
DenseBlockUp2
TransitionUp + DenseBlock1
DenseBlockUp1
Conv.: 1×1, m=4

*Transitionup + DenseBlock** indicates the concatenation operation between the upsampling result and the corresponding previous feature map from the preceding dense block, where * denotes the dense block number.