# VASCULAR PLAQUE DETECTION FROM OPTICAL COHERENCE TOMOGRAPHY IMAGES

by

**Ammu Prakash**

A thesis submitted to the Faculty of Graduate Studies of

the University of Manitoba

in partial fulfillment of the requirements of the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering

University of Manitoba

Winnipeg, Manitoba, Canada

## ABSTRACT

It is difficult to detect atherosclerotic plaque from optical coherence tomography (**OCT**) images via visual inspection. In this work, we developed three algorithms to allow us to detect atherosclerotic plaque more effectively: (i) a statistical method that uses higher-order moments; (ii) a model-based method that enables vascular plaque to be automatically identified based on the textural features in **OCT** images; (iii) and a sparsity-based segmentation algorithm in the curvelet domain. All three algorithms do not rely on visual inspection at all.

The statistical method consists of three main components: extracting statistical image textural features using the Spatial Gray Level Dependence Matrix (**SGLDM**) method; applying an unsupervised *Fuzzy C-means* clustering algorithm to these features; and, finally, mapping specific clustered regions—namely, background, plaque, vascular tissue, and the deep-depth degraded signal in feature-space—back to the actual image. Since the use of the full set of 26 textural features is computationally expensive and may not be practical for real-time implementation, we identified a reduced set of 6 textural features, which were used to characterize vascular plaque via sparse principal component analysis. However, our clustering-based algorithm results had some limitations, most notably non-smooth and coarse segmentation results.

To overcome this low spatial resolution limitation, we developed a stochastic model to segment **OCT** images of vascular tissue into plaque and non-plaque (i.e., healthy tissue) regions, as well as background regions. Our stochastic model is based on a *maximum a posteriori-Markov Random Field* (**MRF-MAP)** framework wherein **OCT** images of vascular tissue were modeled as a Markov random field. This **MRF-MAP**-based algorithm yielded results with better spatial resolution, but it is not consistent and also computationally expensive, thereby impractical for real-time implementation.

Our third approach, using a sparsity-based segmentation algorithm in the curvelet domain, overcame the two limitations above by generating both fast and high-resolution vascular plaque detection from **OCT** images.

We verified the validity of the results of all three methods using both qualitative and quantitative methods. Specifically, all results were compared with 1) actual photographic images of vascular tissue samples, 2) histology results, and 3) ground truth obtained from manual segmentations performed by four cardiovascular surgeons from the Intervention Cardiology Group at St. Boniface Hospital, Winnipeg, Manitoba. These comparisons of results demonstrated that our three methods allow good plaque detection, thus making them potential clinical tools for the detection of vascular plaque from **OCT** images and for clinical studies involving **OCT** imaging of vascular plaque.

## ACKNOWLEDGMENTS

First and foremost, I wish to express my sincere gratitude to my advisor, Dr. Sherif Sherif, for his guidance, mentorship, and immense knowledge. Working with you has been my richest learning experience. Your advice on both my research and my career has been invaluable, and this dissertation would not have been possible without your support and motivation.

I would like to thank Dr. Michael Sowa and Mr. Mark Hewko from the Institute for Biodiagnostics, National Research Council Canada, for providing the vascular Optical Coherence Tomography (OCT) images that were used in my work. I would also like to thank Dr. David Allen, Dr. Kunal Minhas, Dr. Amir Ravandi, and Dr. Ashish Shah from the Intervention Cardiology Group at St. Boniface Hospital for their valuable help in identifying vascular plaque in these images.

My sincere thanks go to my Ph.D. advisory committee, Dr. Pradeepa Yahampath, and Dr. Jitendra Paliwal, for their valuable input on my research.

I would also like to thank my friends, Snehil, Shachi, Sumitha, Sujith, and Randupama, for their valuable advice, for our many enriching philosophical debates and exchanges of knowledge, and for their constant love and support.

Finally, I would like to thank my parents and my brother, Arvind, for showing faith in me and encouraging me to pursue my goals.

## TABLE OF CONTENTS

## LIST OF TABLES

## TABLE OF FIGURES

## THESIS CONTRIBUTION

The following publications are directly related to the contents of this thesis:

1. Prakash, A., Hewko, M., Sowa, M., & Sherif, S. (2012, October). Detection of atherosclerotic vascular tissue from optical coherence tomography images. In *Photonics North 2012* (Vol. 8412, p. 841204). International Society for Optics and Photonics. ***(published)***.

2. Prakash, A., Hewko, M., Sowa, M., & Sherif, S. (2013, May). Texture based segmentation method to detect atherosclerotic plaque from optical tomography images. In *European Conference on Biomedical Optics* (p. 88020S). Optical Society of America. ***(published)***.

3. Prakash, A., Hewko, M. D., Sowa, M., & Sherif, S. S. (2015). Detection of Atherosclerotic Plaque From Optical Coherence Tomography Images Using Texture-Based Segmentation," *Medical Technologies in Medicine*, vol. 7, no. 1 (***published***).

4. Prakash, A., Hewko, M. D., Sowa, M., & Sherif, S. S. (2016). Vascular plaque detection with a reduced textural feature set from optical coherence tomography images. *Photonics North (PN)* (***published)***.

5. Prakash, A., Macias, M. O., Hewko, M., Sowa, M., & Sherif, S. (2016, May). Vascular plaque detection with reduced textural feature set from optical coherence tomography images. In *2016 Photonics North (PN)* (pp. 1-1). IEEE ***(published)***.

6. Prakash, A., & Sherif, S. S. (2017). U.S. Patent application No. 20170251931A1. Washington, DC: U.S. Patent and Trademark Office (***Granted***).

7. Prakash, A., Hewko, M., Sowa, M., & Sherif, S. (2018, April). Detection of vascular plaque from optical coherence tomography images using hidden Markov random field based segmentation. In *Optics and the Brain* (pp. JTh3A-48). Optical Society of America. ***(published)***.

8. Prakash, A., Hewko, M., Sowa, M., Allen D., Minhas K., Ravandi A., Shah A., and Sherif, S. S., "Automated Identification of Vascular Plaque in Optical Coherence Tomography Images,"(submitted to Computer Methods and Programs in Biomedicine)

9. Prakash, A., Hewko, M., Sowa, M., Allen D., Minhas K., Ravandi A., Shah A., and Sherif, S. S. (2021, Feb). High-resolution segmentation of vascular plaque in optical coherence tomography images using sparsity-based s in the curvelet domain. In European Conference on Biomedical Optics (p. 88020S). Optical Society of America. (submitted).

# Chapter 1: INTRODUCTION

## 1.1 Thesis motivation

Cardiovascular diseases are one of the leading causes of mortality and morbidity around the world. As such, it is crucial for medical professionals to be able to detect the conditions that cause cardiovascular disease. The formation of vascular plaque is considered to be the primary underlying pathology of coronary heart disease, as it can accumulate to the point where it blocks arterial blood flow. Among the many medical imaging methods that have been utilized to detect vascular plaque, Intravascular Optical Coherence Tomography (**IVOCT**) has been proven to be equal to Intravascular High-Resolution Ultrasound (**IVUS**) in its ability to detect calcified plaque morphologies. In addition, **IVOCT** also possesses several features that make it highly suitable for intravascular imaging, including high imaging resolution, small-size fiber-based imaging probes, and the availability of advanced image-processing techniques, which allow physicians to extract diagnostic information from Optical Coherence Tomography (**OCT)** images. In this work, we present three algorithms that enable the automatic detection of vascular plaque from **OCT** images: (i) a statistical method that utilizes higher-order moments; (ii) a model-based method that enables vascular plaque to be automatically based on the textual features of **OCT** images; and (iii) a sparsity-based segmentation algorithm in the curvelet domain that does not rely on visual inspection. Our statistical method consists of 3 main steps: using the *Spatial Gray Level Dependence Matrix* (**SGLDM)** method to extract a full set of 26 statistical features; applying an unsupervised clustering algorithm method on these features; and mapping of the clustered regions: namely, background, plaque, vascular tissue, and the deep-depth degraded signal in the feature-space back to the actual image. Given that the use of the full set of 26 textural features is computationally expensive and may not be feasible for real-time implementation, we also identify

a reduced set of 6 textural features, which we use to characterize vascular plaque via sparse principal component analysis. However, our clustering-based algorithms are hampered by several limitations, most notably non-smooth and coarse segmentation results. To overcome this limitation, we developed a stochastic model based on a m*aximum a posteriori – Markov Random Field* (**MRF-MAP)** framework, wherein **OCT** images of vascular tissues are segmented into plaque and non-plaque. One of the main disadvantages of the **MRF-MAP** technique is that it has extremely high computational costs due to requiring numerically approximating the **MAP** estimate. To address this limitation, we utilize a sparsity-based method using a curvelets-based algorithm, wherein the image is sparsely represented in the curvelets transform domain. The results of these three algorithms were validated quantitatively by comparing them with ground truth, which was formed based on the consensus of four surgeons. In addition, we further validated our results visually by comparing them with histology and actual photographic images of vascular tissues. It is expected that our results will yield an efficient pre-clinical tool for detecting vascular plaque from **OCT** images.

## 1.2 Background on vascular plaque and a review of relevant imaging techniques

### 1.2.1 A general overview of vascular disease

Coronary heart disease (**CHD**) did not become a significant problem until the beginning of the 20th century. The later part of the 20th century, particularly between 1968 and 1980 [1], saw a dramatic increase in the prevalence of **CHD** and its entrenchment as a significant cause of mortality and morbidity in developed countries in Europe and North America [1,2], largely due to changing lifestyles within these countries. Comparative prevalence of the condition is currently witnessed in several countries across the globe. In the developed world, atherosclerosis is a leading

cause of death and disability, with vascular diseases being one of the leading causes of death in Canada.

Approximately 70,000 Canadians suffer strokes each year; another 16,000 die from heart attacks, which accounts for 29% of all deaths in Canada [3]. The decline is partially explained by improved survival rates after myocardial infarction (**MI**). Nine in ten Canadians, or 90% of the population, possess one of the major risk factors of heart attack or strokes, which include physical inactivity, smoking, obesity, high blood cholesterol, high blood pressure, and diabetes [3]. With the projected increase in the number of cases, there is a need for more research into novel methods that can be used to identify and characterize CHD efficiently.

It is difficult to track the development of vascular plaque on the artery wall during the early stages of **CHD** due to its asymptomatic and silent nature, and the inability to completely characterize vascular plaque lesions in individual patients is a key limitation currently faced by medical professionals [4]. Cardiovascular diseases are the clinical consequences of arteriosclerosis [5], which is a chronic lesion that affects segments of arteries. Arteriosclerosis is generally caused by a combination of hypertrophic (enlargement of cells) and fibrous changes in the entire thickness of the inner arterial wall [6].

Arteries are the blood vessels that carry blood from the heart to all parts of the body, and their thick wall structure is critical in managing the high pressure of ventricular heart contractions. Arteries are composed of 3 different tunica (layers): the endothelium (intima), the media, and the adventitia [5] (Fig.1.1). The endothelium, also referred to as the "intima," is the innermost layer of the artery, and it is composed of endothelium cells, which are responsible for lining the interior surface of blood vessels, fluid filtration, and blood clotting [7]. The media is the middle layer of the artery and is composed of a layer of elastic tissue and smooth muscle. The primary function of

this layer is to control vascular tonus [8]. The outermost layer, the adventitia, is composed of collagen and fibrous connective tissue, which cover and protect the artery structure in order to support the artery wall and prevent it from tearing [8, 9].

Plaque formation most commonly impacts the media and adventitia by obstructing blood flow, which can lead to angina, myocardial infarction, heart failure, sudden death, or stroke.

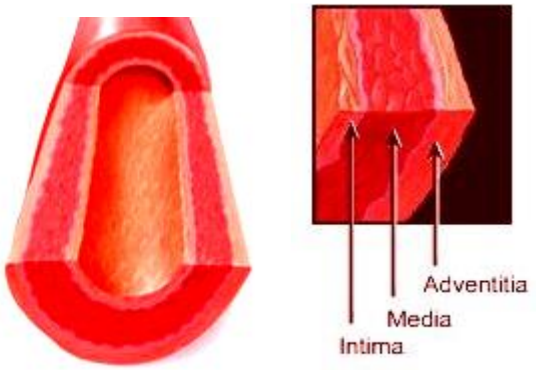



Fig. 1.1. Arterial anatomy.

### 1.2.2 Formation of vascular plaque

Vascular plaque is comprised of fatty substances such as cholesterol (low-density lipoprotein and very-low-density lipoprotein), cells, calcium, and fibrin [10]. The formation of vascular plaque occurs when cholesterol deposits on the arteries harden. In response, the body's immune system attempts to remedy the situation by sending white blood cells to trap the cholesterol; however, this causes the cholesterol to further turn into foamy cells that release more fat and cause more inflammation. The muscle cells of the arteries are subsequently triggered to multiply and form a cap over the affected area, which is problematic because the soft plaque trapped beneath the cap is often quite dangerous. Indeed, this soft plaque is the leading cause of heart attack. When the person's blood pressure spikes, the soft plaque exerts pressure on the plaque's thin walls, which consequently break open and cause a clot to form. This process causes the artery wall to thicken,

which reduces the artery's interior diameter, thus obstructing blood flow within [10]. This process occurs gradually, and larger plaques can completely block blood flow. When coronary arteries are affected, the obstruction of blood that supplies the heart muscle can lead to angina or myocardial infarction; conversely, when the carotid arteries that supply the brain are obstructed, a stroke may occur. Similarly, obstruction of the renal arteries may result in kidney disease, while obstruction of the peripheral arteries may lead to gangrene [10].

Unfortunately, the process of removing the cholesterol from the plaque is often difficult, as it involves adjusting the levels of blood cholesterol [11]. Low-density lipoprotein (**LDL**) deposits cholesterol into the blood vessels; however, high **LDL** removes deposited cholesterol from the bloodstream. The process of changing normal concentrations of **LDL** is not easy and wrought with complexities. Nonetheless, several processes exist for reducing increased **LDL** concentrations, including the use of statins, lovastatin, and pravastatin, which impede the liver's production of cholesterol-inducing enzymes. In addition, the use of another drug, ezetimibe, has also been shown to be effective at blocking cholesterol in the digestive tract. The process of shrinking plaques using statins has come to be considered desirable, especially when the **LDL** value gets below 70 mg/dl. There are other methods of preventing this condition as well. For instance, previous studies have demonstrated the importance of lifestyle changes in mitigating the condition. While post-mortem examinations provide invaluable information about the histopathologic tendencies of the vulnerable plaques, such studies are often limited by selection bias. Additionally, previous studies have only considered lesions at the most advanced stages. As such, these studies are considered a snapshot that cannot provide primary information relating to the natural history of the plaque-build-up process. This deficiency is partly due to the use of intracoronary imaging in patients [12].

This paucity in research clearly highlights the need for the development of intracoronary imaging methods that are able to quantify unstable and rapture-prone plaques.

### 1.2.3 Risk factors of vascular disease

Kakkos *et al*. (2007) [13] have described the causes of vascular plaque and its associated risk factors. High blood concentration of lipoprotein B-containing lipoproteins, of which **LDL** is the most dominant form, is one of the main determinants of **CHD**. **LDL** is also strongly associated with familial hypercholesterolemia (**FH**) and genetic hyperlipidemias (monogenic disease). Vascular plaque often develops under low **LDL** concentrations alongside a combination of other risk factors, including diabetes, smoking, hypertension, gender, and genetic susceptibility to the condition. The condition is normally prevented using statins and anti-hypersensitive drugs. Lifestyle modification is another necessary approach to preventing the condition. For instance, since smokers have an increased risk of suffering from the condition, it would be prudent for them to quit in order to reduce their chances of developing it. Lifestyle changes indicate the multifactorial basis of the condition.

Previous research has established that individuals with exceedingly low **LDL** fail to develop clinically relevant atherosclerosis, irrespective of other risk factors [12,13]. Mendelian randomization studies have demonstrated that **LDL** acts as a buffer against the pathogenesis of vascular plaque [14], which underscores its central higher order as a causal factor for the common, multifactorial form of the condition. Previous studies indicate that modifiable risk factors explain more than 90% of the causes of **MI** across the world [15]. Cardiovascular disease (**CVD)** is the main cause of mortality and disability in various countries across the world and initially manifests in the form of fatal or leaves irreversible sequelae [16]. While **CVD** must be stratified according to the patient's risk level, experts are exempted from using primary prevention risk equations when

assessing patients with genetic hypercholesterolemia (**GH**), as prior research has demonstrated that the use of such equations underestimates this group's risk level [17]. Indeed, **GH** patients are at high-risk for **CVD** due to their abnormal levels of atherogenic lipoproteins, which they normally possess from birth.

Apart from lifestyle changes, there are recommendations for food consumption. For instance, one study indicated that eating a Mediterranean diet—which is rich in olive oil, fruits, vegetables, fish, low-processed foods, and moderate wine consumption—reduces the chances of cholesterol development by approximately 30% [18]. Additionally, the role of a sedentary lifestyle in the pathophysiology of cardiovascular diseases has been well-established. Aerobic exercise has been shown to increase HDL levels, lower blood pressure, burn body fat, and lower blood sugar levels. Individuals with risk factors for **CVD** should consider engaging in moderate-intensity exercise for 150 minutes per week, as exercise regimens oriented towards weight loss have been shown to further lower **LDL** levels in the blood [19]. Smoking cessation is another notable step that one can take to reduce one's risk of developing **CVD**. Research indicates that quitting smoking can lower one's risk of developing **CVD** by approximately 20%, in addition to raising their **HDL** levels. In summary, the most common risk factors that contribute to the development of **CVD** are as follows:

- Hypertension
- Dyslipidemia (Referred to as High blood cholesterol)
- Cigarette Smoking
- Diabetes Mellitus
- Physical Inactivity
- Stressful Lifestyle

### 1.2.4 Types of vascular plaque

Type I (Adaptive Initial Thickening): This is the earliest vascular change that can be described microscopically and is found in at least 30% of neonates at birth. The intima of human arteries contains *resident smooth muscle cells* [20].

Type II (Fatty Streaks): This lesion starts when extracellular lipids are caused by risk factors such as high cholesterol, smoking, hypertension, obesity, insulin resistance, and the accumulation of adhesion molecules in the intima as a result of endothelium dysfunction. The inflammatory process begins with the adherence of leukocytes (immune system cells), especially monocytes (a larger version of leukocytes), in the intima. When the lipid accumulation is not significant enough, the leukocytes begin to accumulate lipids and transform them into foam cells. Fatty Streaks do not cause symptoms, can evolve into more complex lesions, or even disappear.

Type III (Pathologic Intimal Thickening): When the lipid accumulation (fatty streaks) does not disappear, it grows up into intermediate lesions due to the attraction of numerous leukocytes (monocytes). Known as an intermediate lesion, this lesion type is characterized by lipid pools near the medial wall in areas that generally lack smooth muscle cells (due to the apoptosis of foam cells produced by the down-regulated degeneration of the fatty streaks by the leukocytes).

The luminal surfaces, however, are most abundant in smooth muscle cells and are often accompanied by infiltrating foam cells. Another characteristic of these lesions is microcalcification.

Type IV Fibroatheromas (vulnerable plaques) are the first of the advanced lesions and are characterized by an acellular necrotic lipid core that attracts more macrophages from the bloodstream. The core is embedded in the depth of the plaque, surrounded by fibrous tissue.

Type V (Plaque Stability): This lesion is characterized by the migration of smooth muscle cells to the affected lesion. This results in the overgrowth of the foam cells, which leads to the formation of a layer of collagen called the fibrous cap. The fibrous cap (considered as stable plaque) plays a critical role in harboring the contents of the necrotic core, and its integrity is one of the defining influences on plaque stability.

Type VI (Thrombosis): This is the last lesion prior to atherosclerosis. In this stage, the artery wall thickens, resulting in reduced artery diameter and obstructed blood flow [20]. The endothelial cells (due to the degeneration of this layer) covering the fibrous cap become extremely thin, fragile, and susceptible to erosion. As a result, the endothelial layer becomes injured, and the released material exposed to the flowing blood initiates the sudden formation of a blood clot, also known as thrombosis, which blocks the lumen of the artery. Plaque rupture is the primary process responsible for myocardial infarction and stroke.

## 1.3 Imaging techniques

Vascular plaque formation is a chronic disease that progresses slowly and presents significant symptoms. Although it cannot be tracked during the early stages, it can be monitored via diagnostic imaging processes prior to clinical manifestations of **CVD**. New imaging modalities continue to emerge because of new developments, device improvements, and the application of new energy sources. A sample population from the Asymptomatic Polyvascular Abnormalities Community (**APAC**) study was examined via sonography with high-resolution beta-mode ultrasounds in order to investigate the epidemiology and presence of asymptomatic intracranial atherosclerosis stasis, carotid atherosclerosis, and peripheral artery disease [21]. The study examined the bilateral carotid, including the common carotid artery, carotid bifurcation, and the internal carotid artery (**ICA**). However, the researchers reported issues with the use of sonography, which emphasizes the need

for effective strategies for the detection and prevention of atherosclerosis and **IS**. The research findings further found that individuals with higher cardiovascular health (**CVH)** scores were at lower risk of developing carotid plaque.

Treating **CVD** typically involves targeting small and unstable plaques. For a 30% blockage, the goal is usually to reduce it to 15% by sucking the cholesterol from the inside of the artery. The common understanding of the pathology of unstable and vulnerable atherosclerotic plaques is mainly based on a limited number of studies [22], which have usually consisted of post-mortem examinations of the human coronary arteries. Other studies have used resected surgical specimens from patients who underwent carotid endarterectomy for either primary or secondary prevention of transient ischemic attack or stroke [22].

The most common imaging modalities used in vascular plaque imaging are:

Angiography: First performed in humans in 1958, this traditional method is used to visualize arteries, veins, and heart chambers [23]. Angiography is an invasive technique that provides *in-vivo* high-resolution images. In this imaging method, a contrast agent is applied to the area to enhance the medium, which is imaged using an X-ray-based technique (fluoroscopy) [23]. Angiography provides detailed anatomic imaging but is unable to provide a functional assessment.

Intravascular Ultrasound (**IVUS**): This technique provides extremely high-resolution cross-sectional images with direct arterial vascular wall imaging through high-frequency ultrasound. This kind of ultrasound is typically performed at 20-40 MHz, which allows resolutions of approximately 15-20 μm [23]. Compared with conventional angiography, **IVUS** imaging enables the tomographic assessment of the lumen area, as well as plaque size, distribution, and composition. However, this modality still relies on visual inspection, which can be problematic, as different tissue components may appear remarkably similar [24].

Computed Tomography (**CT**): This imaging procedure uses special x-ray equipment to create a series of scans of areas inside the body. A fast-gated helical **CT** (i.e., multidetector **CT**) can precisely detect the quantity of calcium in vascular plaque lesions.

Magnetic Resonance Imaging (**MRI**): **MRI** is a non-invasive medical imaging technique that employs magnetic radio waves to take anatomical images, and it is also capable of distinguishing soft-tissue contrast. The advantage of **MRI** is that it can detect plaque formation in its early stages, but it is unable to provide any information on the actual risk of plaque rupture [24].

Near-Infrared Spectroscopy: This technique is based on the principle that organic molecules absorb and scatter light to different degrees at various wavelengths (800-2500 nm). This approach works by sending light into a sample and then measuring the proportion that is returned. This method enables researchers to obtain the chemical characterization of coronary artery plaques (chemo grams, which display the probability of lipid core plaque). Although this technique has been demonstrated to be feasible, accurate, and safe for application in humans, it requires further extensive study before it can be applied in clinical practice [23].

Intravascular Optical Coherence Tomography (**IVOCT**): **IVOCT** is a minimally invasive light-based imaging modality based on low-coherence interferometry and developed explicitly for the identification of vascular plaque [25- 27]. **IVOCT** generates cross-sectional 2-dimensional images from emitted and reflected near-infrared light. It offers a superior resolution to that of **IVUS** (about 4-20 μm) but limited tissue penetration of 2-3 mm, which enables improved plaque characterization and a histology-grade definition of the coronary plaque microstructure. **OCT** can accurately detect thin fibrous caps, lipid pools, and macrophage infiltration.

Additionally**, OCT** uses a caterer to reach the medium that is to be imaged; however, since light emission is highly attenuated by blood, it should be displaced during **OCT** imaging by applying contrast flushes or an angioplasty balloon [23].

**OCT** has been employed in numerous applications, and it has the potential to be used in many medical imaging fields and applications. The two most promising areas of **OCT** application are heart disease and cancer detection. Furthermore, the use of **OCT** imaging can improve current cardiovascular therapies, such as stenting and balloon angioplasty, as it is able to provide vascular images in real-time.

**OCT** can differentiate between stable and unstable plaques by visually identifying the plaques in the bloodstream. These types of plaque are responsible for up to 70% of all heart attacks. The **OCT**'s optical fiber probe is easily adaptable to coronary catheters [28], which allows it to be inserted into arteries to acquire arterial pathology images. The first investigation of **IVOCT** demonstrated its ability to perform microscopic tomographic imaging of the internal microstructure of vascular plaques *in vitro* [29]. In addition, further developments in **IVOCT** technology have enabled intracoronary imaging in human patients [30-33]. **IVOCT** can differentiate the three layers of an artery wall. It depicts the intima (innermost layer) as a signal-rich layer, the media (middle layer) as a weak signal layer, and the adventitia as a signal-rich layer. Moreover, **OCT** is also able to identify three types of vascular plaques: lipid-rich, fibrous, and fibrocalcific.

Yabushita *et al.* (2002) developed the first steps to differentiating different vascular plaque components using **IVOCT** imaging, thereby validating and testing its accuracy. To perform this validation and accuracy test, they used 357 specimens and histology images to perform different vascular plaque characterizations.

Their validation results showed that sensitivity and specificity ranged from 71% to 79% and 97% to 98% for fibrous plaques, 90% to 94% and 90% to 92% for lipid-rich plaques, and 95% to 96% and 97% for fibrocalcific plaques, thus demonstrating **IVOCT**'s potential for differentiating lipid-rich plaques from other plaque types [28]. **IVOCT** is also capable of identifying additional plaque components that may be associated with coronary events. Some of these features [33] and calcic nodules are associated with plaque thrombosis in some cases [23, 24]. Studies have also shown that, compared to histopathology, **IVOCT** can diagnose calcific nodules with 96% sensitivity and 97% specificity [33]. Cholesterol crystals are another notable feature that can be imaged via **IVOCT**. Studies have shown that the presence of cholesterol crystals increases the stiffness of the lipid pool stiffness and may, therefore, decrease the likelihood of plaque rupture [34]. Cholesterol crystals appear in **OCT** images as oriented, linear, and highly reflecting structures within the plaques [35]. Multinucleated macrophages are an inflammatory response to a foreign body, such as cholesterol crystals, within the plaque. These cells can also be identified by **IVOCT,** appearing as large, highly reflecting regions [33]. The features of **IVOCT** that make it attractive and superior for intravascular imaging are its high resolution, the small size of its fiber-based imaging probes, and the availability of advanced image-processing techniques to extract and assess diagnostic information [36-45]. **IVUS** is capable of higher resolution compared to **IVOCT**, which enables it to provide more detailed visualizations of the anatomical features of the arterial wall and plaque. Furthermore, **IVOCT**'s ability to detect calcified plaque morphologies has also been shown to be equal to that of **IVUS**. Additionally, other studies have demonstrated the clinical application of **IVOCT** and its superiority to **IVUS** in detecting characteristics of vascular plaque [46]. Although **IVUS** is not capable of identifying microstructural features of vascular plaque, it can identify non-plaque vessels and arterial disruptions. A study was conducted to compare

**IVOCT-IVUS** image pairs obtained from different patients [47]; in all cases, it was found that the **IVOCT** observations were more consistent than the **IVUS** observations. These findings establish **IVOCT** as a promising imaging modality for extracting diagnostic information related to a vascular plaque. Finally, the use of texture analysis may make it possible to differentiate between visually uniform tissue types in **IVOCT** images [48]. Recently researchers also have also used deep learning techniques such as convolutional neural networks to characterize vascular plaques in OCT images [49].

Juhwan *et al*. developed a fully automated vascular plaque characterization using deep learning models to classify lipidous and calcification plaque regions. The algorithm was validated on 89 volumes of interest having calcification on 32 regions, lipidous on 36 regions, both calcification and lipidous on 12 regions, and 9 regions of both without calcification and lipidous regions. They found the sensitivities and specificities for pixel-wise classification to be 87.4% /85.1% and 85.1%/94.2%, respectively.

Abdolmanafi *et al.* [50] used a deep learning AlexNet model to extract features, and then the arterial borders and plaque region were classified using supervised classifiers, e.g., random forest and support vector machine classifiers. Addolmanafi *et al.* [51] also proposed another deep learning-based method using a convolutional neural network and a fully convolutional neural network to classify the normal arterial wall and diseased or affected arterial wall. Their dataset compromised of 45 OCT pullbacks with ~100 images per pullback. They found the accuracy to be 96% for a diseased arterial wall structure and 91 % for a normal arterial wall.

Yong et al. [52] proposed a fully automated algorithm based on linear regression convolutional neural network to segment coronary lumen in IVOCT. They achieved 98.5% in dice coefficient and 97% in the Jaccard similarity index.

He.et.al [53] proposed a convolutional neural network-based method to classify vascular plaque in OCT images automatically. Their dataset consisted of 269 OCT images, and they found the average prediction accuracy to be 86.6%.

In summary, **IVOCT** imaging could enable faster and more accurate diagnoses of vascular disease. As technology becomes increasingly advanced, it may be possible for physicians to use **IVOCT** as a clinical tool for diagnosis.

## 1.4 Dissertation outline

This thesis proposal consists of six chapters. In chapter 2, we present the dataset used in this thesis, we describe the main advantage of using a specific animal model, and we explain the specifics of the **OCT** imaging technique used in this work. Furthermore, we introduce our evaluation technique, which utilizes both qualitative and quantitative measurements. In Chapter 3, we describe our first plaque detection algorithm, and we also introduce the feature generation and selection methods used in this work. In Chapter 4, we present our model-based plaque detection algorithm, in which **OCT** images are modeled as a Markov random field. Chapter 5 presents our sparsity-based vascular plaque segmentation method in the curvelet domain. Finally, Chapter 6 presents the conclusions of our work and outlines directions for future work.

# Chapter 2: DATASETS

## 2.1 Texture analysis

**IVOCT**'s high spatial resolution (typically 10 to 15 μm), high contrast, and volumetric imaging make it a potentially useful tool for the detailed study of the morphological structures of plaques. **OCT**'s high resolution enables it to contrast structural proteins, which makes it invaluable in observing plaque risk stratification. While many medical imaging methods have been utilized to detect vascular plaque, **OCT** has proven to be particularly effective for high-resolution detection of calcified plaque morphologies. Furthermore, **OCT** possesses several additional features that make it highly suitable for intravascular imaging; however, the use of **OCT** images to visually quantify and detect atherosclerotic plaques is a difficult task. Some studies have attempted to analyze the texture of carotid plaques through ultrasound models using a statistical method based on the Gray Level Co-occurrence Matrix (**GLCM**). Textural analysis could be very useful in developing automated systems for characterizing and classifying carotid ultrasound images, and the use of **GSM** (gray-scale median) allows researchers to find the historical features of plaque. These studies quantified textural features such as contrast, energy, dissimilarities, difference variance, entropy, cluster variance, homogeneity, sum entropy, and the sum of square variance extracted from the areas of interest.

Other studies have demonstrated the potential of using textural analysis to determine the behavior and the interaction between anti-hypersensitive drugs and plaque [14]. Information from such changes can be useful in allowing medical practitioners to adjust patient medication and regimens based on the plaque texture.

The use of textural analysis to evaluate the prognosis of other diseases has also been demonstrated. Loizou *et al*. (2011) [54] used magnetic response images of the brain obtained via

multiscale amplitude moderation-frequency modulation (**AM-FM**) to study the textures of lesions in patients with multiple sclerosis. There has been significant clinical interest in identifying lesion texture and progression. In addition to showing that **AM-FM** is able to differentiate between various features of lesions, their results also showed that **AM-FM** could be used to quantify other features such as the gray-scale median, contrast, and coarseness, which can help to develop a more detailed understanding of how these features are linked to brain lesions. A similar **MRI** study [55] produced comparable results, demonstrating that the use of textural analysis is complementary to **MRI**. As these studies show, texture contains crucial information about particular diseases and can be used to identify various morphological features of a disease.

Doonan *et al*. (2013) [56] estimated the correlation between echo density and textural features by analyzing ultrasound and digital images of plaques taken from patients with bilateral carotid stenosis who were to undergo carotid endarterectomies. Their results showed that the textual and echo density characteristics of carotid plaques are similar between the two sides in patients suffering from bilateral stenosis. This observation supports previous findings that plaque instability is a confounding of systemic factors. Doonan *et al*.'s study was pioneering in that it showed the role that the textural analysis of images can play in determining the features and characteristics that define the histopathology of the conditions. Previous studies have further indicated that unstable carotid atherosclerotic plaques exhibit higher lipid and hemorrhage content and less fibrous tissue calcification. The textural analysis is based on previous findings that have shown that the tissue content of carotid plaques is often correlated to features obtained through ultrasound images, for example, **GSM**. More particularly, high-lipid-content plaques that ultimately hemorrhage appear echolucent in ultrasounds with low **GSM**. Similarly, fibrous plaques often appear echogenic with high **GSM,** irrespective of their calcification status. Furthermore,

symptomatic plaques are often more echo lucent than asymptomatic plaques. All of these variations in the features of plaques can be leveraged to quantify their activities and development.

Van Engelen *et al.*'s (2014) [57] study, in which carotid plaque volumes and plaque texture in 298 carotid atherosclerosis patients were monitored for one year, highlighted the need for fast, reliable, and cost-effective methods of monitoring patients with increased risk of **CVD**. After the initial phase of the study had concluded, they followed up with the patients for a period of up to five years in order to record instances of myocardial infarction, transient ischemic attack, and stroke. Using Kaplan-Meier analysis, the researchers realized that the use of plaque texture was invaluable in providing information about vascular events, thus confirming the usefulness of textural analysis in patients with increased risk.

The earlier detection of atherosclerosis is critical in determining the occurrence of stroke and other heart conditions and providing opportunities for their prevention [58]. However, the quantification of the early streaks of plaque is often nuanced in that interpretation is based on the competence of a medical professional. Predicting the progression of atherosclerosis requires novel methods to identify blood fluid and cardiovascular wall dynamics, which can be used to identify instances of vessel abnormalities. A recent analysis indicated that the inner surface of the arterial wall often becomes rough before increasing in thickness [59]. Therefore, the ability to quantify the roughness of the inner surface of the arterial walls is important for the early diagnosis of the condition. To this end, previous studies have used ultrasound as a non-destructive method for evaluating arterial wall roughness.

Various methods have been developed to measure arterial texture—for instance, the use of the angular spectrum-based formulation [19]. In the experiment, the roughness of the surface varied with the specular reflection and corresponding scattering intensities: the rougher the surface,

the higher the proportion of scattered ultrasound. Other researchers developed an ultrasonic spectroscopic technique that quantified the gross surface texture of non-medical materials using a coefficient of roughness. However, the use of these programs is limited by their error range, which is often high, sometimes in the millimeter range. Additionally, such methods have underscored the importance of using textural properties to determine carotid atherosclerotic plaque. The present study attempts to fill this research gap by assessing the usefulness of textural analysis methods in quantifying arterial roughness and by defining textural features that can provide an accurate analysis.

Voros *et al*. (2011) [60] indicated that the molecular and cellular events that underpin atherosclerosis—for example, the deposition of lipoprotein, inflammation, the proliferation of the smooth muscles, apoptosis, necrosis, calcification, and fibrosis—have a specific influence on the compositional and geometric changes in coronary vessels. The use of Computerized Tomography Angiography (**CTA)** has been imperative in evaluating these changes, specifically changes in positive remodeling, lipoprotein deposition, and calcification. When the specific focus is placed on plaques, information about the various features that underpin the histopathology of **CVD** can be quantified. For instance, one study that used **CTA** to determine plaque characteristics assessed plaque segments based on stenosis severity while classifying the plaques into three categories: calcified, non-calcified, and partially calcified. The novel **CTA** procedure showed minimal variation between interobserver and intra-observer agreements, with clinically reproducible results that can be used to guide further research in **CTA**. While textural analysis is still in its infancy, it could be a breakthrough in determining the characteristics of plaques. As recent studies have shown, textural analysis can be useful in segmenting tissue types with similar appearances based on their speckle features [61-65].

## 2.2 Datasets

This research in this thesis was carried out using **OCT** images of vascular tissues from Watanabe heritable hyperlipidemic rabbits, henceforth referred to as **WHHLMI** rabbits. Our repository of datasets also contained histological and photographic images of these vascular tissues. We used ten different tissue sections of arterial samples from 7 **WHHMI** rabbits aged 309, 316, 330, 342, 365, 456, and 577 days. The dataset used in this research is provided in Table 2.1.

Table 2.1. **OCT** dataset used in this research.

| Image # | OCT sample ID | Age (days) | Age (months) | OCT B-scan # |
|---|---|---|---|---|
| 1 | 2009_12_15_W21_092_01_511 | 309 | 10 | 165 |
| 2 | 2009-05-27_W21-087-511 | 316 | 10 | 310 |
| 3 | 2009-05-27_W21-087-611 | 316 | 10 | 330 |
| 4 | 2010_03_03_W20_084_02 | 330 | 11 | 190 |
| 5 | 2009-06-23_W20-087-01-6-1-2 | 342 | 11 | 220 |
| 6 | 2009-08-12_W21-085-01-611 | 365 | 12 | 350 |
| 7 | 2009-12-08_W20-084-04-6-1-1 | 456 | 15 | 170 |
| 8 | 2010_02_02_W20_086_02_512 | 577 | 19 | 450 |
| 9 | 2010-02-02-W20-086-02-612 | 577 | 19 | 130 |
| 10 | 2010-02-02-W20-086-511 | 577 | 19 | 75 |

### 2.2.1 Animal model

Samples of vascular tissue with atherosclerotic plaque were obtained from myocardial-infarction-prone **WHHLMI** rabbits [66-67]. We obtained the photographic, histology, and **OCT** dataset of arterial samples from the Institute for Biodiagnostics, National Research Council Canada. Arterial segments of tissue, starting from the ascending aorta to the external iliac artery, were excised from all specimens and subdivided into sections 20~30 mm in length. Arterial samples were harvested from seven **WHHLMI** rabbits aged 309 days, 316 days, 330 days, 342 days, 365 days, 456 days, 577 days. We used 10 **OCT** B-scan images from 10 different volume scans of tissue sections. A breakdown of a sample tissue section is shown in Fig. 2.1, and an actual photographic image of a sample vascular tissue section is shown in Fig. 2.2. This study was approved by the local animal care committee at the Institute for Biodiagnostics, National Research Council Canada (Winnipeg, Manitoba).

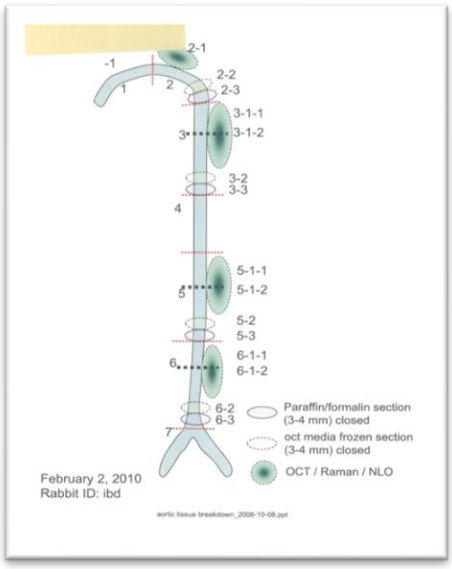



Fig. 2.1. Vascular tissue breakdown.

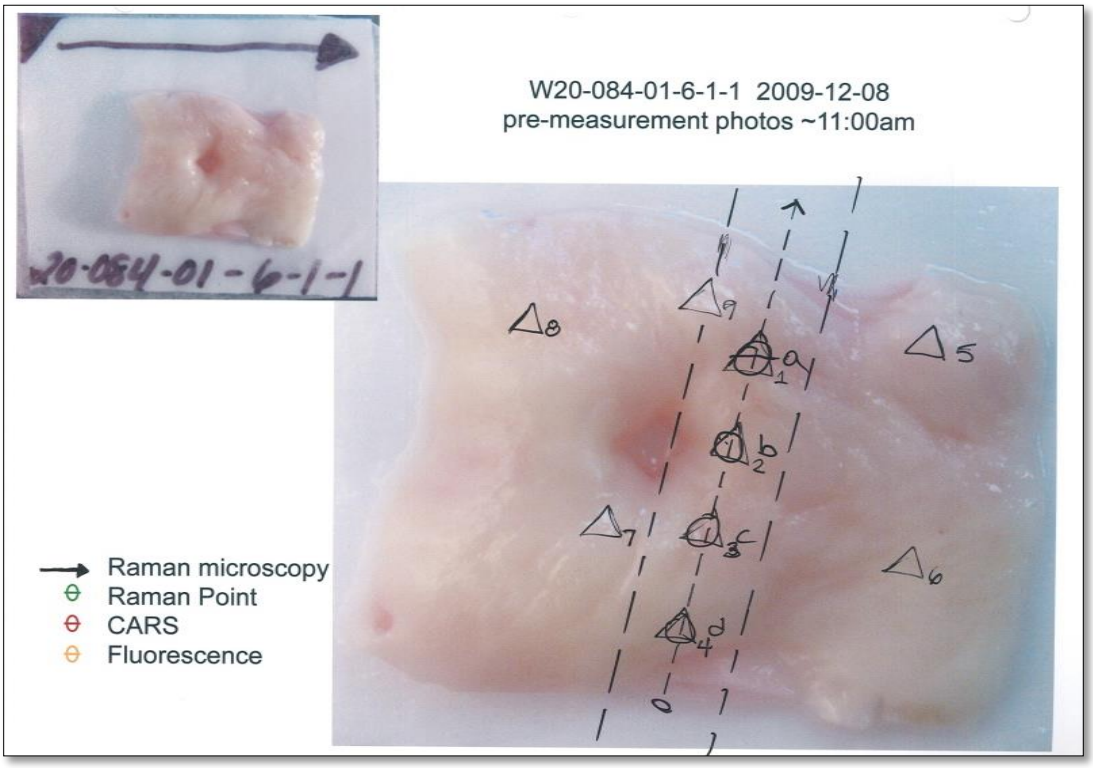



Fig. 2.2. Photographic image of vascular tissue of section 5-1-2 of a 456 days old **WHHML** rabbit.

### 2.2.2 OCT imaging modality used in this work

**IVOCT** is a catheter-based intravascular imaging technique that uses near-infrared light to create images [68-72]. **OCT** is very similar to ultrasound imaging; only it uses light waves instead of sound waves to create images. Because of this, **OCT** is able to produce images with resolutions ten times higher than those produced via ultrasound imaging. **OCT** uses light with wavelengths ranging from 1.25 to 1.350 μm, which minimizes light wave absorption in water, lipids, and hemoglobin. In **OCT**, the light from the source is split into two parts, with one part being directed toward the arterial wall and the other part being directed toward a mirror. The reflected signals interfere on the surface of a photodetector, which creates images based on the intensity of the

interference signal. The lateral resolution of the **OCT** system ranges between 20-90 µm, compared to 150-300 µm for **IVUS**. Similarly, **OCT**'s range of axial resolution is 12-18 microns, compared to 150-200 microns for **IVUS** [67]. However, the tissue penetration depth is limited to 1-3 mm in **OCT** as opposed to 4-8 µm for **IVUS**. The **IVOCT** system consists of a catheter, an imaging engine, and a computer (Fig. 2.3). In this work, we used a swept-source **OCT** (**SS-OCT**) [73] with a central wavelength of 1310 nm and a sweep rate and range of 30 kHz and 110 nm, respectively. Our **SS-OCT** unit was configured as a Mach-Zehnder interferometer with balanced optical detection.

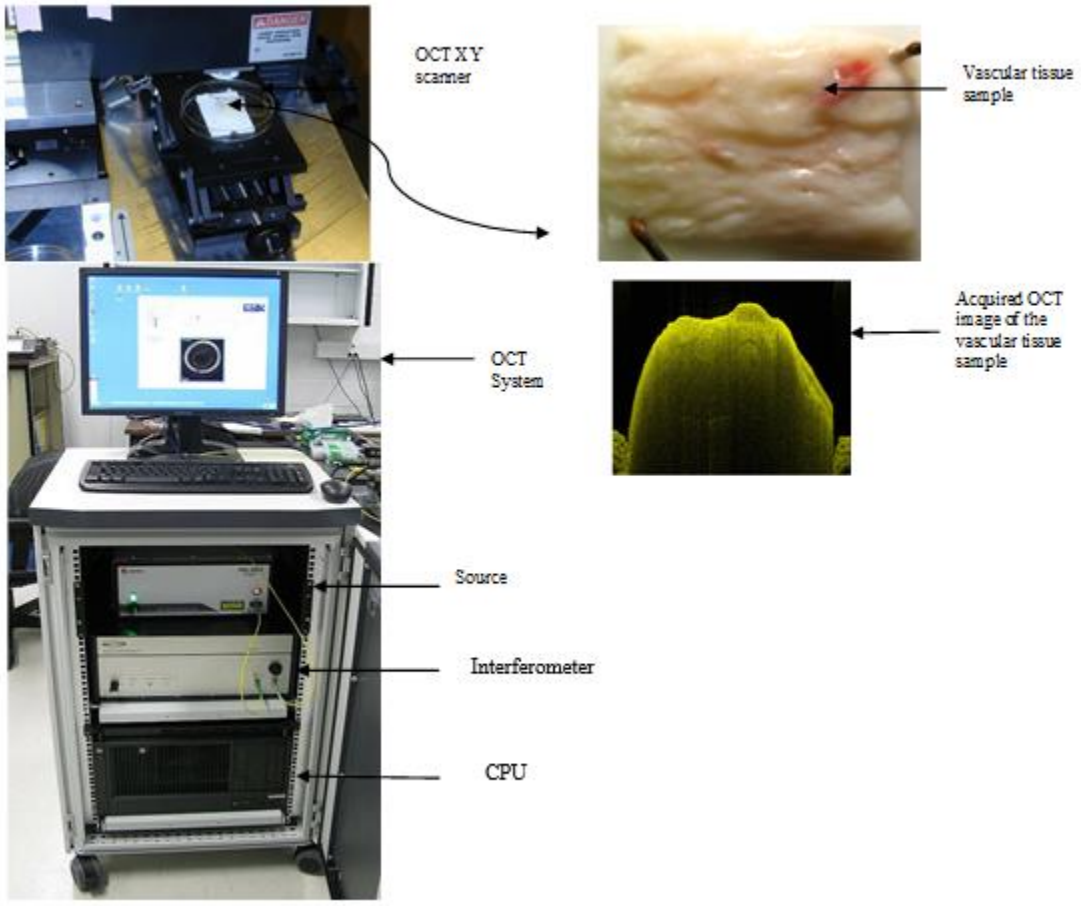



Fig. 2.3 (a). Intravascular OCT imaging setup.

The optical power at the sample was measured to be 8.45mW, and the collimator's lens focal distance was 25 mm. As shown in Fig 2.3 (b), the dimension of the generated image, in the X direction, is ~8 mm ×~20 mm.

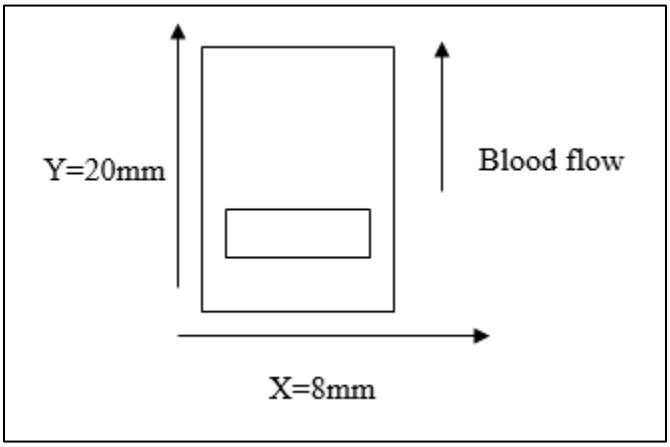



Fig. 2.3 (b). Intravascular **OCT** imaging setup.

## 2.3 Preprocessing of OCT data

Since our raw **OCT** vascular images were represented as floating point numbers, we performed segmentation using image normalization on each image file to achieve a uniform distribution of intensities on a standardized intensity range and to improve contrast. After image normalization, each pixel had a brightness value ranging from 0 to 255. Image normalization was performed via a Min-Max normalization operation, which preserves all relationships between the data values exactly [74] and compresses the normal range if extreme values or outliers exist. Min-Max normalization is carried using the following formula:

$$X' = \frac{\mathrm{x}-\min(\mathrm{x})}{\max(x)-\min(x)} \tag{2.1}$$

where x' is the processed image, and x is the raw **OCT** image.

After normalizing our image file, we improved the image quality by performing automatic image segmentation using a threshold. Examples of preprocessed images are shown in Fig. 2.4.

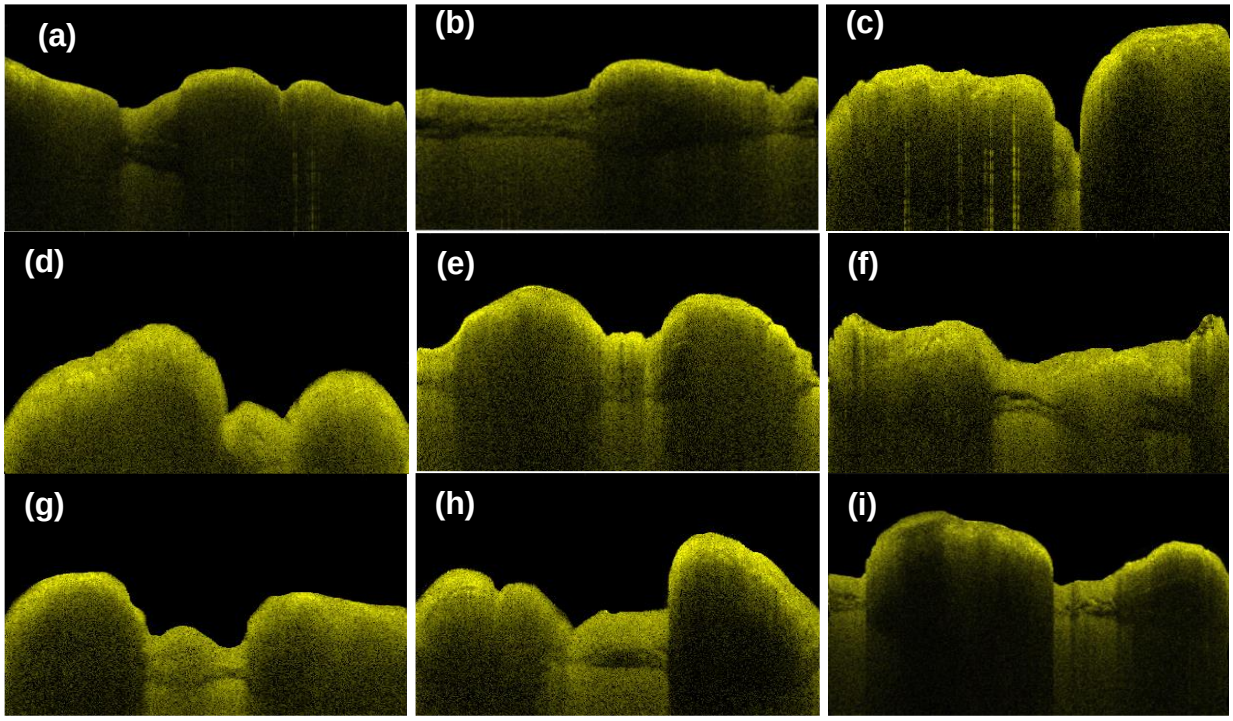



Fig. 2.4. Examples of preprocessed **OCT** images of tissue sections: (a) 6-1-1 from a 309 day-old rabbit; (b) 6-1-2 from a 316 day-old rabbit; (c) 5-1-1 from a 577 day-old rabbit; (d) 5-1-1 from a 309 day-old rabbit; (e) 6-1-1 from a 316 day-old rabbit; (f) 5-1-2 from a 577 day-old rabbit.

## 2.4 Evaluation of our segmentation results

We evaluated our segmentation results using both qualitative and quantitative approaches.

### 2.4.1 Qualitative evaluations

We evaluated our results using by comparing them with ground truth, histology images, and actual photographs of vascular tissues. Each surgeon independently and blindly outlined the plaque region in 10 selected **OCT** images. Ground truth for the segmented vascular plaque images was obtained based on the consensus of assessments of the images (Fig. 2.5) provided by four interventional cardiologists with the Intervention Cardiology group at St. Boniface Hospital (Winnipeg, Manitoba): Dr. David Allen, Dr. Kunal Minhas, Dr. Amir Ravandi, and Dr. Ashish Shah. We also compared our results with actual photographic images (Fig. 2.6) and histology images (Fig. 2.7).

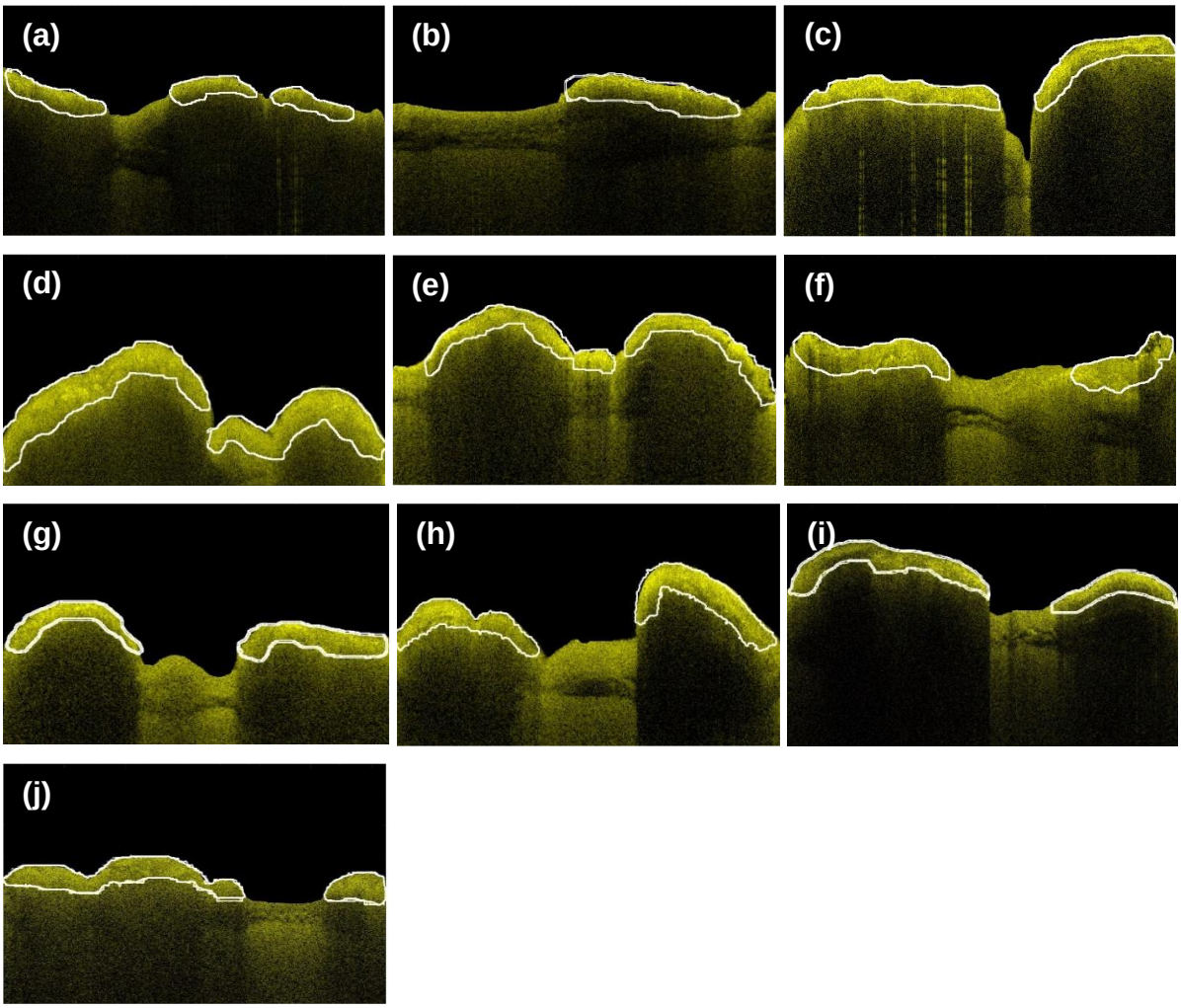



Fig. 2.5. **OCT** images of vascular tissues based on the consensus of 4 cardiac surgeons: (a) 5-1-2 from a 316 day-old rabbit; (b) 6-1-2 from a 316 day-old rabbit; (c) 6-1-2 from a 365 day-old rabbit; (d) 6-1-2 from a 456 day-old rabbit; (e) 6-1-2 from a 342 day-old rabbit; (f) 6-1-2 from a 330 day-old rabbit; (g) 6-1-2 from a 577 day-old rabbit; (h) 5-1-1 from a 577 day-old rabbit; (i) 5-1-2 from a 577 day-old rabbit; (j) 6-1-2 from a 309 day-old rabbit.

We used oil red histology images and actual photographic images to compare our plaque detection results visually. Oil red is a staining agent used for the qualitative measurement of lipid deposit formation. The red regions in the histology image highlight the plaque region, as shown in Fig 2.7.

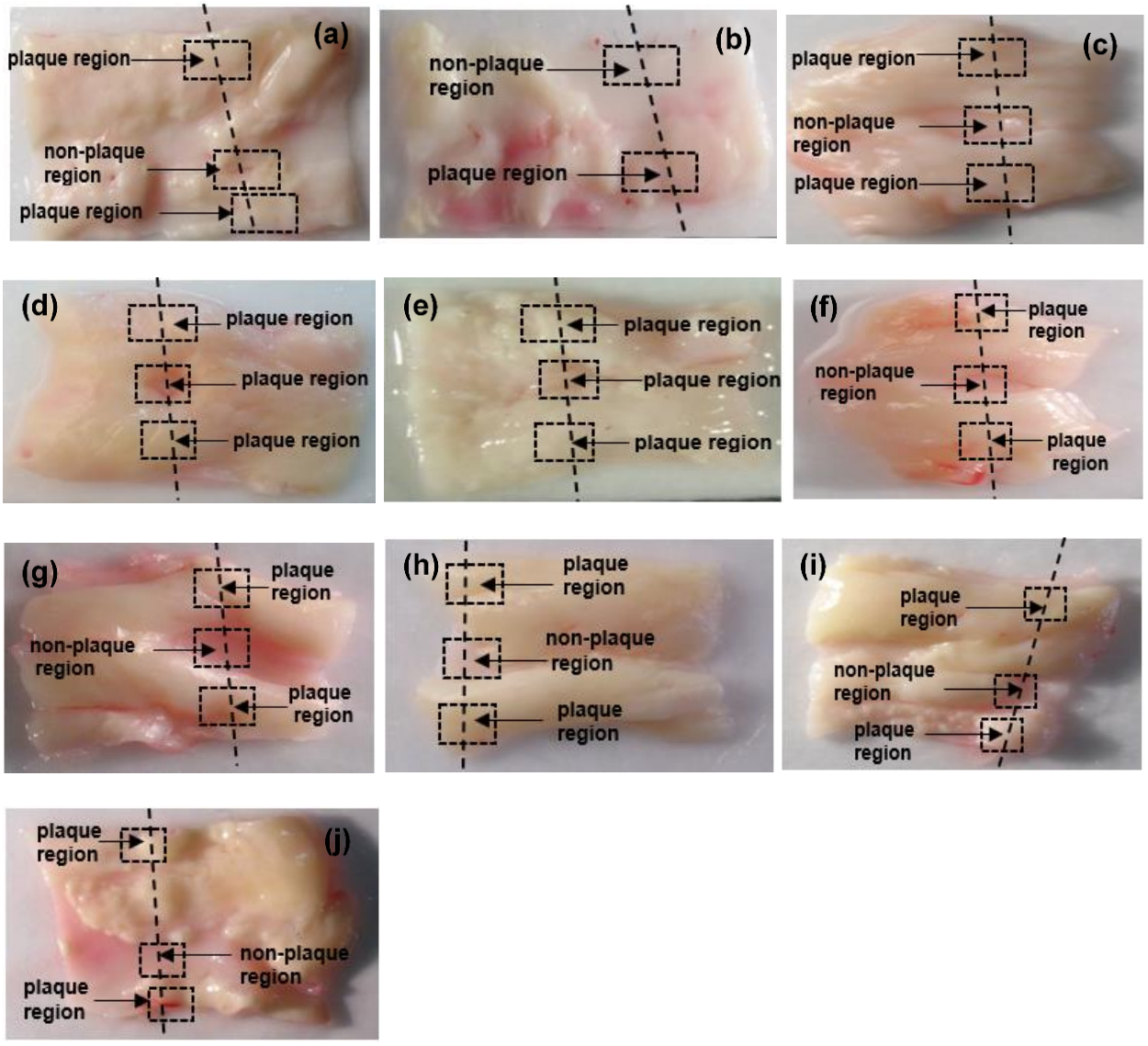



Fig. 2.6. Photographic images of vascular tissues: (a) 5-1-2 from a 316 day-old rabbit; (b) 6-1-2 from a 316 day-old rabbit; (c) 6-1-2 from a 365 day-old rabbit; (d) 6-1-2 from a 456 day-old rabbit; (e) 6-1-2 from a 342 day-old rabbit; (f) 6-1-2 from a 330 day-old rabbit; (g) 6-1-2 from a 577 day-old rabbit; (h) 5-1-1 from a 577 day-old rabbit; (i) 5-1-2 from a 577 day-old rabbit; (j) 6-1-2 from a 309 day-old rabbit.

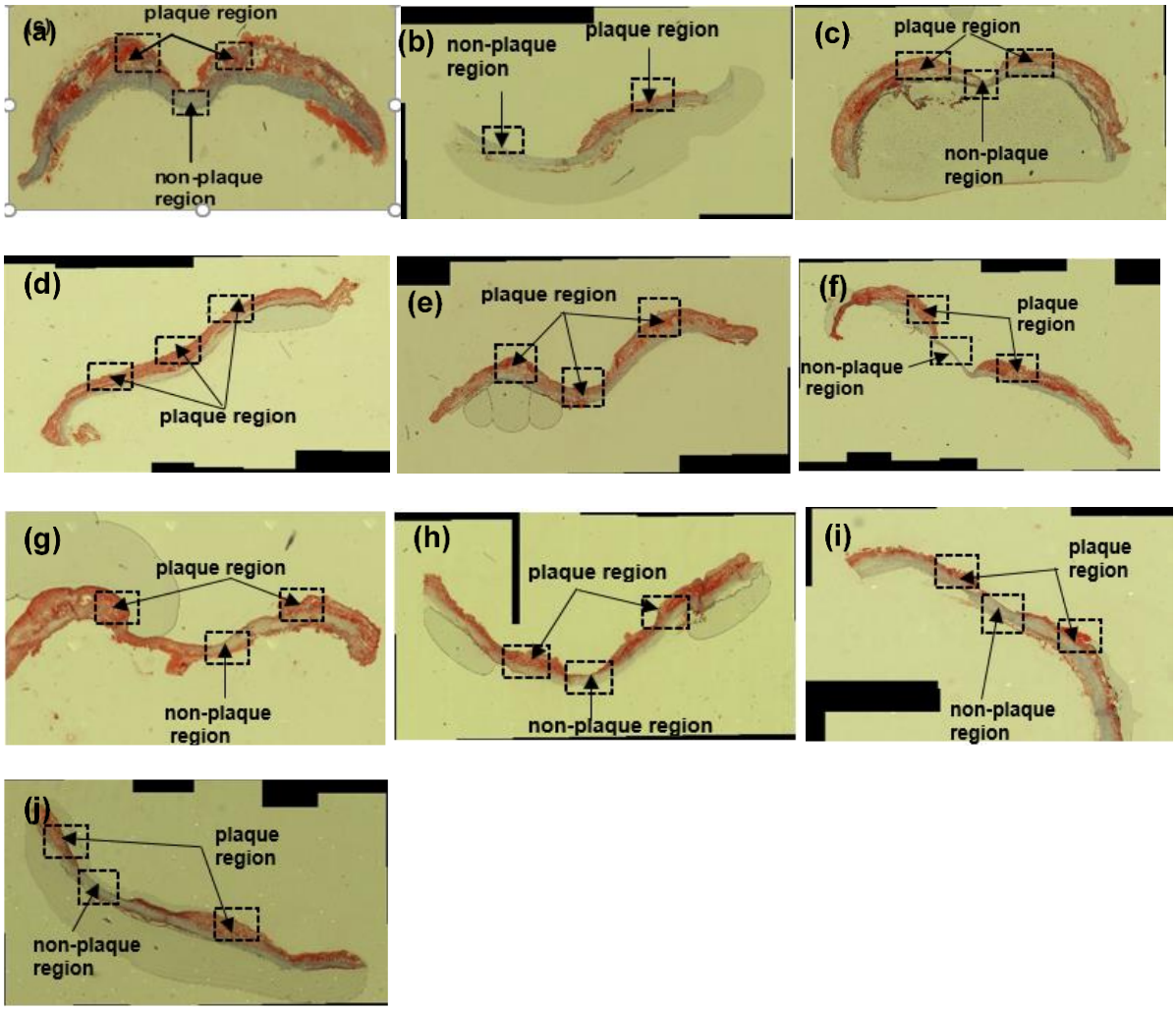



Fig. 2.7. Histology images of vascular tissues: (a) 5-1-2 from a 316 day-old rabbit; (b) 6-1-2 from a 316 day-old rabbit; (c) 6-1-2 from a 365 day-old rabbit; (d) 6-1-2 from a 456 day-old rabbit; (e) 6-1-2 from a 342 day-old rabbit; (f) 6-1-2 from a 330 day-old rabbit; (g) 6-1-2 from a 577 day-old rabbit; (h) 5-1-1 from a 577 day-old rabbit; (i) 5-1-2 from a 577 day-old rabbit; (j) 6-1-2 from a 309 day-old rabbit.

### 2.4.2 Quantitative evaluations

We evaluated the performance of our segmentation results using six well-known standard metrics: sensitivity (recall), specificity, accuracy, precision, Matthews Correlation Coefficient (**MCC**), and Dice similarity coefficient [70-73, 75-78]. The pixels in the segmented region and its corresponding ground truth could be categorized into four categories: true positive (**TP**), false positive (**FP**), true negative (**TN**), and false-negative (**FN**). An example illustrating the relationship between these four categories is shown in Fig. 2.8.

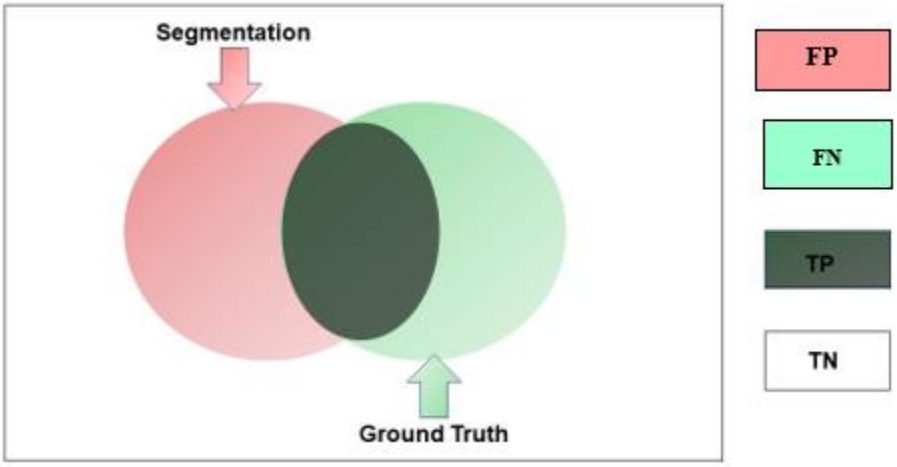



Fig. 2.8. Illustration of **TP**, **TN**, **FP**, and **FN**.

Sensitivity and specificity are the statistical quantitative error evaluation metric for classification problems. Sensitivity is also considered to be a true positive (TP) rate, as it measures the percentage of actual positives that are correctly labeled as such; in our case, the pixels labeled as plaque were correctly identified as such. Conversely, specificity is considered to be a measure of the true negative (TN) rate, as it gives the proportion of actual true negatives that are correctly detected as negatives. In our case, these pixels labeled as a non-plaque region are true negatives. Accuracy is the overall measure of the method's ability to classify positives and negatives correctly. To calculate the accuracy measure, the proportion of both true positives and true negatives must be considered. Also, precision is defined as the proportion of plaque, and non-plaque regions are correctly segmented as TP and TN. In other words, precision is the proportion of true positives out of all detected true positives.

Mathematically, sensitivity, specificity, precision, and accuracy can be stated as follows,

$$Sensitivity\ (Recall) = \frac{TP}{(TP+FN)} \tag{2.2}$$

$$Specificity = \frac{TN}{(TN+FP)} \tag{2.3}$$

$$Accuracy = \frac{TN}{(TN+TP+FN+FP)} \tag{2.4}$$

$$Precision = \frac{TP}{(TP+FP)} \tag{2.5}$$

Additionally, we also used the Matthews Correlation Coefficient (MCC), which gives an overall summary of the performance of the segmentation algorithm, and it takes into account all the four values: TP, TN, FP, and FN. It is generally regarded as a balanced measure that could be used even if the true negatives are very unbalanced compared with true positives. It returns a value between -1 and 1, where 1 represents a complete agreement, –1 represents a complete disagreement, and 0 is no better than a random prediction.

$$MCC = \frac{TP \times TN - FP \times FN}{(TP+FP)(TP+FN)(TN+FP)(TN+FN)} \tag{2.6}$$

Since the regions of interest are small, the specificity (true negatives), and even the accuracy, may appear artificially high relative to the sensitivity. Therefore, we also used an alternative error-measurement metric, called the Dice similarity coefficient. The Dice similarity coefficient evaluates the spatial overlap between the ground truth and the segmentation results. The value of the Dice similarity coefficient ranges from 0 to 1, with 0 representing no overlap between ground truth and the segmented result and 1 representing a perfect overlap between the two regions.

$$Dice\ similarity\ coefficient = 2\frac{A \cap B}{|A|+|B|} \tag{2.7}$$

In the above equation, A is the manual segmented volume of the ground truth, and B is the segmented volume extracted by our proposed automated segmentation algorithm. $|A|$ is the absolute value of the segmented region of the ground truth, $|B|$ is the absolute value of the automated segmented region by our proposed algorithms, and $A \cap B$ is the area of overlap between ground truth and automated segmentation.

For our plaque segmentation problem, the region of interest, which is the plaque only, is usually small, and therefore the three most relevant evaluation metrics are precision, **MCC**, and Dice coefficient metric.

## 2.5 Summary

In this chapter, we introduced the dataset used for this work. This chapter consists of five sections. In the first section, we briefly explained the texture analysis methods and their applications. In the second section, we introduced the details of the **OCT** dataset and the **OCT** imaging framework used for this work. In the third section, we explained the pre-processing process of raw **OCT** images. And the fourth section introduces images used for qualitative evaluation and explains the metrics used for quantitative evaluation of our segmentation results, and the final fifth section presents the conclusions.

# Chapter 3: DETECTION OF VASCULAR PLAQUE IN OPTICAL COHERENCE TOMOGRAPHY IMAGES USING TEXTURAL FEATURES

## 3.1 Introduction

In earlier work, we proposed a method of automatically detecting vascular plaque that uses a full set of 26 *Haralick* textural features and *K-means* clustering [79]. However, the computational complexity of our plaque detection method was limited by the dimensionality of the feature space. Therefore, we selected only the most important features in order to reduce the number of features while also retaining as much information as possible. This procedure is known as feature selection or reduction. If we select features with little discrimination power, the resultant segmentation algorithm will perform poorly. On the other hand, if information-rich features are selected, the algorithm design can be greatly simplified. As such, features should take distant values in the different classes, as well as closely located values within the same class. Thus, it is crucial to reduce the feature set by selecting only those features that characterize the vascular plaque texture. We used two techniques to perform feature selection: (i) genetic algorithm (**GA**) optimization and (ii) sparse principal component analysis (**PCA**).

Using **GA** enabled us to identify a reduced set of 3 textural features. Unfortunately, **GA** suffers from high computational complexity and is also prone to overfitting. In order to overcome these limitations, we used sparse **PCA**. This approach allowed us to identify a set of 6 textural features that characterize vascular plaque in **OCT** images, thereby making **OCT** a viable option for real-time applications. Finally, we also incorporated an advanced clustering technique (*Fuzzy C-means*) to detect plaque regions within this reduced feature space [80].

## 3.2 Methods of plaque detection

### 3.2.1 Textural feature generation

Texture can be defined as visual patterns composed of spatially repetitive organized structures. Although there is no precise mathematical approach to describing texture, it is sufficient to describe it using the specific qualitative properties of an image. For example, the texture of an image could be referred to as being fine, coarse, smooth, irregular, homogenous, or inhomogeneous. Textural features are those that can be used to quantify such properties in an image, and its histogram or statistical moments can be used to characterize its textural properties. The most commonly used second-order statistical features are calculated using a Spatial Gray Level Dependent Matrix (**SGLDM**) [81]; these features showed great potential for discriminating between different textures in biomedical images [82-83]. The **SGLDM**-based approach has also been shown to outperform transform-based methods, such as the use of wavelets in texture classification [84]. The main advantage of using **SGLDM** features is that they extract second-order statistics. Prior studies have compared visual-texture-differentiation [85], power-spectrum-based, and structural-based [86] methods and found that textures are discriminated only when they differ in their second-order statistics.

The **SGLDM** method is widely considered to be the most powerful textural feature extraction method [87]. Textural features are useful in many applications, including medical imaging. Image texture has been recognized as a significant feature in applications such as medical image analysis, image classification, and automatic image inspection [87-88]. Our method uses a statistical method to extract second-order statistical textural features where pixels are considered in pairs of plaque of images obtained via **OCT**. The use of first-order statistics is generally insufficient for measuring

an image's structural and textural characteristics, as they only provide information such as histograms, which capture pixel intensity distribution but do not provide information about the position or structure of these pixels within an image. To extract this information, we used second-order statistics, wherein pixels are considered in pairs. Prior studies have also shown that SGLDM based texture analysis outperformed other texture analysis methods such as the gray level run length method (GLRLM), the gray level difference method (GLDM), and the power spectral method (PSM) [89-91]. The **SGLDM** provides information on both the relative distance between the pixels and their relative orientation to one another. In our application of **SGLDM**, we used a distance ($d$) equal to 1 pixel (i.e., neighboring pixels), and two different orientations: one in a horizontal direction ($\theta$ =$0^0$), and one in a vertical direction ($\theta$ =$90^0$) (Fig. 3.1).

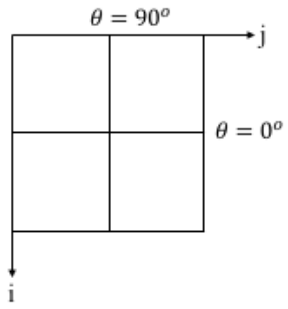



Fig. 3.1. The two ($0^o$ and $90^o$) orientations used to construct the **SGLDM** matrices in our algorithm.

For each combination of distance, $d$, and orientation, $\theta$, a two-dimensional histogram is defined as:

$$0^0 = P(I(i,j) = I_1, I(i \pm d, j) = I_2)$$
$$90^0 = P(I(i,j) = I_1, I(i, j \mp d) = I_2) \quad (3.1)$$

After using the probabilities of grey-level occurrence with respect to a pixel's position in order to form the **SGLDM** matrices, we then used them to calculate the *Haralick* textural features [92-94]. Some of these features can be directly interpreted with respect to texture; for example, the angular second-moment feature is the measure of the image's smoothness; contrast is the measure of the local gray level variation within the image, and entropy is the measure of randomness in an image,

and therefore produces low values for smooth images. While other features do not enable such direct interpretation, they can still convey texture-related information with high discriminatory power. Table 3.1 summarizes all the *Haralick* textural features extracted from **OCT** images of vascular tissues.

Table 3.1. *Haralick* textural features.

| **Feature Number** | **Formula** | **Feature Name** |
|---|---|---|
| $f_1$,$f_{14}$ | $\sum_i \sum_j (P(i,j)^2)$ | Angular Second Moment at Orientations (ASM) ($\theta = 0^0$, $\theta = 90^0$) |
| $f_2$,$f_{15}$ | $\sum_i \sum_j \frac{(i,j)P(i,j) - \mu_x \mu_y}{\sigma_x \mu \sigma_y}$ | Correlation ($\theta = 0^0$, $\theta = 90^0$) |
| $f_3$,$f_{16}$ | $\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_i \sum_j P(i,j) \right\}_{\lvert i-j \rvert = n}$ | Inertia at orientations ($\theta = 0^0$, $\theta = 90^0$) |
| $f_4$,$f_{17}$ | $\sum_i \sum_j (i-\mu)^2 P(i,j)$ | Variance at orientations ($\theta = 0^0$, $\theta = 90^0$) |
| $f_5$,$f_{18}$ | $\sum_i \sum_j \frac{P(i,j)}{1+(i-j)^2}$ | Inverse Difference Moment at orientations ($\theta = 0^0$, $\theta = 90^0$) |
| $f_6$,$f_{19}$ | $\sum_{i=0}^{2(N_g-1)} iP_{x+(-)y}(i)$ | Sum Average at angles ($\theta = 0^0$, $\theta = 90^0$) |
| $f_7$,$f_{20}$ | $\sum_{i=0}^{2N_g-2} (i-F_5)^2 P_{x+y}(i)$ | Sum Variance at orientations ($\theta = 0^0$, $\theta = 90^0$) |

| | | |
|---|---|---|
| $f_8$,$f_{21}$ | $-\sum_{i=0}^{2Ng-2} P_{x+y}(i)\, log\{P_{x+y}(i)\}$ | Sum Entropy at orientations ($\theta$ = $0^0$, $\theta$ = $90^0$) |
| $f_9$,$f_{22}$ | $-\sum_i \sum_j P(i,j)\, log\, P\,(i,j)$ | Entropy at orientations ($\theta$ = $0^0$, $\theta$ = $90^0$) |
| $f_{10}$,$f_{23}$ | $-\sum_{i=0}^{Ng-1} (i - F_5)^2 P_{x-y}(i)$ | Difference Variance at orientations ($\theta$ = $0^0$, $\theta$ =$90^0$) |
| $f_{11}$,$f_{24}$ | $-\sum_{i=0}^{Ng-1} P_{x-y}(i)\, log\, P_{x-y}\,(i)$ | Difference Entropy at orientations ($\theta$ = $0^0$, $\theta$ = $90^0$) |
| $f_{12}$,$f_{25}$ | $\frac{H_{xy} - H_{xy}^1}{max[\,H_x - H_y]}$ | Information Measure I of Correlation at orientations ($\theta$ = $0^0$, $\theta$ = $90^0$) |
| $f_{13}$,$f_{26}$ | $\sqrt{1 - exp(-2(H_{xy}^2 - H_{xy}))}$ | Information Measure II of Correlation at orientations ($\theta$ = $0^0$, $\theta$ = $90^0$) |

where, $H_{xy}^1 = -\sum_i \sum_j P(i,j)\, log(\,P_x(i)P_y(j))$, $H_{xy}^2 = -\sum_j \sum_i (P_x(i)P_y(j)\, log(\,P_x(i)P_y(j))$, *P(i, j)* is the $(i, j)^{th}$ entry in the SGLDM matrix, $P_x(i)$ is the marginal probability of the $i^{th}$ entry, $N_g$ is the number of gray levels in the image, and $P_y(i)$ is the marginal probability of the $j^{th}$ entry.

### 3.2.2 Feature normalization

Feature normalization is an important step in pre-processing features for clustering algorithms. Since the scale of our textural features had different dynamic ranges, we normalized the entire textural feature vector to ensure that all of the features had the same influence on our method's performance. Each textural feature vector was normalized as [95],

$$X' = \frac{x - \bar{x}}{\sigma} \tag{3.2}$$

where $X'$is the new rescaled feature vector, x is the raw feature vector before rescaling, $\bar{x}$ is the mean of all entries of *x*, and σ is the corresponding standard deviation.

### 3.2.3 Textural feature selection

Vascular plaque can be detected from Optical Coherence Tomography (**OCT**) images by using the full set of 26 *Haralick* textural features and the standard *K-means* clustering algorithm [96]. However, the use of the full set of 26 textural features is computationally expensive and may not be feasible for real-time implementation. Therefore, the use of a feature reduction step was critical for optimizing the performance and robustness of our method. Given a number of generated features, it is important to reduce the number of dimensions by selecting only the most informative features, as this will enable the number of dimensions to be reduced while retaining their class discriminatory information. This procedure, known as feature selection or feature reduction, is highly vital because the selection of features with little discrimination power may cause the segmentation algorithm to perform poorly. The two main disadvantages of irrelevant features are that they incur substantial computational costs, and they may also lead to overfitting. We aimed to reduce the number of features, selecting only those that are rich in information with respect to our plaque detection problem. Accordingly, we used two different feature selection algorithms—a genetic algorithm (**GA**) and a sparse principal component regression (**PCR**)—to attain the smallest number of textural features possible without sacrificing textural information.

*3.2.3.1 Feature selection using an evolutionary algorithm*

A **GA** is a computation model inspired by Darwin's Theory of Evolution [97-98]. Since any feature selection problem can be considered as a multi-criteria optimization problem, evolutionary algorithms offer an attractive approach to solving such problems [99-101].

Each individual in the algorithm represents a potential candidate to feature subset selection problem. The fitness function is an evaluation function, which is a combination of two criteria: the accuracy of the segmentation or classification and the costs involved in performing the segmentation. We chose a fitness function based on the max-relevance and min-redundancy principle. According to this principle, the optimal number of subset features will be selected, thus satisfying the following maximization problem,

$$maximize\,\{f(x)\}$$

where $f(x)$ the objective function. Subsequently,

A complete genetic algorithm consists of the steps below and those shown in Fig. 3.2.

1. Initialization: An initial population is entered into a loop, which runs until convergence or termination criteria are met.

2. Evaluation: The fitness or the quality-measure of the individuals is computed using the following entropy-based fitness function,

$$H(Y|x_i) = \sum_{k=1}^{N} -log(P(y_k\,|x_i)).P(y_k|x_i) \tag{3.5}$$

where *Y* is the target variable, and *x* is the input variable.

3. Selection: The fittest individuals from the current population are selected to form a mating pool for reproduction. Each individual, ***x,*** is selected and copied in the mating pool with probability being proportional to fitness (f(x) / Σf(x)). Selection is made via a roulette wheel selection,

$$\frac{exp(gu)-1}{exp(u)-1} \tag{3.6}$$

where $g$ is a positive constant value used to tune the selective pressure; the larger the value of $g$, the faster the algorithm will converge, and $u$ is a uniformly distributed random variable.

5. Crossover: The genes of two selected parents are merged to yield two new children. In this work, a single-point crossover is assumed. Two individual parents are selected from the mating pool. The crossover point is randomly chosen, and the strings are swapped with respect to the crossover point between the two parents.

6. Mutation: One or more of the elements of the genotype is spontaneously changed. The mutation operator is applied gene-wise; that is, each gene undergoes mutation with the probability, $pm$. When the mutation operation occurs in a gene, its value is flipped.

7. Termination: This evolution process is carried out until the termination criterion is met.

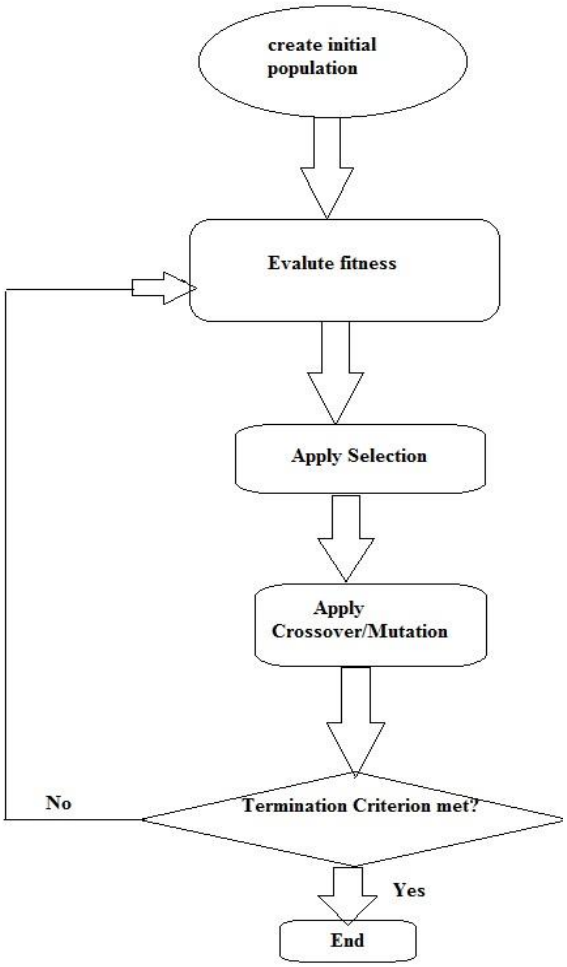



Fig. 3.2. Flowchart of the genetic algorithm optimization working principle [98].

We identified the best three features subset (Table 3.2).

Table 3.2. *Haralick* textural feature set in the $\Theta = 0^0$ and $\Theta = 90^0$ directions with d=1 selected using the **GA** feature selection algorithm.

| Selected feature set | Feature name |
|---|---|
| $f_1$ | Angular Second Moment at orientations(ASM) ($\Theta = 0^0$) |
| $f_3$ | Inertia at orientations ($\Theta = 0^0$) |
| $f_{14}$ | Angular Second Moment at orientations(ASM) ($\Theta = 90^0$) |

However, we noticed that one of the disadvantages of using a **GA** to select a subset of features is that **GA**s tend to converge towards the local optimum rather than the global optimum, and therefore

it does not converge to one unique solution, which makes **GAs** prone to overfitting. To improve the robustness of our feature selection problem, we employed another technique: sparse **PCA**.

*3.2.3.2 Feature selection using sparse principal component regression*

**PCA**, also known as the Karhunen-Loeve transformation, is one of the most straightforward and robust dimensionality reduction techniques. PCA is one of the widely used techniques in statistics and data analysis. It projects the high-dimensional input features into a low-dimensional subspace with only a few linear combinations of input features, known as principal components. If we have a set of features $X_1$, $X_2$,……$X_p$, PC is a linear combination of these variables,

$$Y_1 = \varphi_{11}X_1 + \varphi_{12}X_2 + \ldots + \varphi_{1p}X_p = V_1X$$

$$Y_2 = \varphi_{21}X_1 + \varphi_{22}X_2 + \ldots + \varphi_{2p}X_p = V_2X$$

.

.

$$Y_p = \varphi_{p1}X_1 + \varphi_{p2}X_2 + \ldots + \varphi_{pp}X_p = V_pX$$

which could rewritten as, $Y_{n\times p} = X_{n\times p}V_{p\times p}$

Each column of matrix V is called an eigenvector or loading vector, and coefficient(weights) are called the loadings for that vector. Each of the principal components $Y_1 \ldots Y_p$ is, therefore, a linear

combination of all the 26 original Haralick features. As most of the coefficients (or loadings) in the principal components are not zero, and, therefore, they are not suitable feature selection. However, it is possible to reduce the number of features using an ad-hoc **PCA** method, which involves threshold loadings. However, this approach can be misleading [102-104], and it has been hampered by the significant drawback of determining how to identify the threshold value and the best discrete-valued loading coefficients. Various methods based on imposing sparsity to **PCA** have been proposed to overcome this limitation, including sparse **PCA** [105-109], sparse factor analysis [114,116], sparse singular value decomposition [117,118], and sparse support vector machines [114-119]. The sparse **PCA** formulation is closely related to the dictionary learning problem with one main difference: sparsity is enforced on the dictionary's atoms (loadings), not on the coefficient matrix (principal components). In this work, we used the most common approach to inducing sparseness into **PCA:** *the least absolute shrinkage and selection operator* (**LASSO**) [108]. **LASSO** was originally used in linear regression models to select important variables by shrinking negligibly small estimates to zero via an $L_1$ penalty function. The same shrinkage concept can also be used in **PCA** by formulating **PCA** as a principal-component-regression-type problem and by adding a sparsity constraint on the loadings. In this work, we apply sparsity on the loadings in order to perform feature selection [116]. Sparse **PCA** methods adjust the **PCA** method to inject sparseness into the loading vectors, much like in regularization methods wherein sparseness is injected into the parameter estimates in the regression setting. Our method consists of the following three main steps. This motivated us to use a variant of the classical PCA problem called sparse principal component analysis (SPCA). The sparse algorithm is based on framing PCA as a least-squares type problem, and sparsity is enforced to coefficient (or loadings) by imposing Lasso (L1) type penalty.

The first step in SPCA involves computing PCA by singular value decomposition (SVD) of our feature matrix X. The **SVD** of X yields,

$$X = UDV^T \tag{3.7}$$

where, $Y_i = U_i D_{ii}$ gives the principal component of each observation, and $V_i$ is the corresponding loading vector of the PCA (principal direction). $X$ is our observation matrix of dimension $N \times p$ whose columns, $p$, are the centered input features. U is an $N \times p$ orthogonal matrix, whose columns, $u_j$, are referred to as the left singular vectors. V is also an orthogonal matrix with dimensions of $p \times p$ whose columns, $v_j$, are referred to as the right singular vectors. D is a $p \times p$ diagonal matrix with diagonal elements known as the singular values.

Since $Y_i$ can also be obtained by projecting X on the vector $V_i$ (i.e., $Y_i = XV_i$ ), one can view PCA as a regression type problem where $Y_i$ is the response vector and $V_i$ is the regression coefficient.

To improve sparsity, authors of SPCA propose to impose the lasso penalty [104], resulting in the following optimization problem,

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \|Y_i - X\beta\|_2^2 + \lambda\|\beta\|_1$$

where, $\|\beta\|_1 = \sum_{j=1}^{p} |\beta_j|$ is the $l_1$-norm of $\beta$. $\widehat{V_i} = \frac{\hat{\beta}}{\|\hat{\beta}\|}$ an approximation of $V_i$ and $XV_i$ is the $i^{\text{th}}$ principal component. $\lambda$ is the Lagrange multiplier.

The above formulation cannot be directly used to find sparse loadings; we must first perform **PCA** to find the principal components and their loadings or coefficients (loadings), and then we must solve a **LASSO** regression framework for the **PCA**. This allowed us to find the reduced four-feature set shown in Table 3.3.

Table 3.3. Selected *Haralick* textural feature set in $\Theta = 0^0$ and $\Theta = 90^0$ directions with d=1 obtained using the **PCR** feature selection method.

| Selected feature set | Feature name |
|---|---|
| $f_1$ | Angular Second Moment (Energy) at $0^o$, |
| $f_6$ | Sum average at $0^o$ |
| $f_{13}$ | Information measures of correlation II at $0^o$ |
| $f_{16}$ | Inertia at $90^0$ |
| $f_{23}$ | Sum Variance at $90^0$ |
| $f_{20}$ | Difference Variance at $90^0$ |

## 3.3 Application of Fuzzy C-means algorithm on reduced feature space

Various methodologies that use a clustering technique have also been proposed for the segmentation of vascular plaque [117-122]. Clustering analysis is an unsupervised technique wherein different grouping regions within an image are grouped into subsets of similar properties. Unsupervised methods do not require *a priori* knowledge of samples; that is, their class labels do not need to be known. Thus, unsupervised methods aim to organize a dataset into sensible clusters or groups by finding the similarities or differences within it. Since each region of vascular tissue is composed of different textural features, a clustering algorithm could be applied to the group using similarities among these textural features — clustering algorithms group feature vectors into their respective classes. In the present work, we performed clustering using the *Fuzzy C-means* algorithm [123]. One of the main advantages of *Fuzzy C-means* clustering over the standard *K-means* clustering is that it allows data points to belong to more than one cluster by assigning a membership value. Thus, this approach is useful for overlapping data sets, as *Fuzzy C–means* considers every data point to be a member of every cluster, with varying degrees of membership. The *Fuzzy C-means* algorithm tries to minimize the following objective function:

$$J = \sum_{i=1}^{C} \sum_{k=1}^{N} \mu_{i,k}^{m} d_{i,k}^{2} \tag{3.9}$$

where J is the objective function; $\mu_{ki}$ is the degree of membership, which is defined as the closeness of each feature vector to the cluster center; m is the weighting exponent, which determines the fuzziness of clusters; and $d_{i,j}^2$ is the Euclidean norm between the feature vector and the cluster center.

The first step in *Fuzzy C -means* is to initialize the cluster centroid randomly. The cluster centroid is computed as,

$$v_i = \frac{\sum_{k=1}^{N} \mu_{i,k}^m x_k}{\sum_{k=1}^{n} \mu_{i,k}^m} \tag{3.10}$$

where $x_k$ is the feature vector, and $\mu_{i,k}$ is the membership of a feature vector, $x_k$, to the i[th] cluster. Next, the Euclidean distance between cluster centroids and each data point is calculated, and the fuzzy membership matrix is subsequently determined based on each cluster centroid and each data point. The fuzzy membership criterion is determined by,

$$\mu_{i,k} = \frac{1}{\sum_{j=1}^{C} \left( \frac{\|x_k - v_i\|}{\|x_k - v_j\|} \right)^{\frac{2}{m-1}}} \tag{3.11}$$

The final step in this process is to find the centroid from the updated membership matrix and then repeat these procedures until the algorithm converges.

*Fuzzy C-means* clustering has three main parameters. The first parameter is the number of clusters (C), which is the only parameter that should be known *a priori*. In our vascular detection problem, there are 4 clusters in total (plaque region, healthy tissue region, **OCT** deep-depth degraded signal region, and background). The second parameter is the fuzziness parameter (m), which is also referred to as the weighting exponent. This parameter influences the fuzziness of the partitioning clustering; as m gets closer to 1, the partitioning clustering becomes hard or crisp, similar to conventional *K-means* clustering. As m $\rightarrow\infty$ (m>1), the partitioning clustering starts to become fuzzy, allowing for the overlapping of clusters. The standard value for the fuzziness parameter, m,

is 2. The selection of the fuzziness parameter is a complex process, and the accurate selection of the optimal parameter is subjective. The third parameter is the termination criterion: the *Fuzzy C-means* algorithm stops the iteration process once the distance between 2 successive iterations is smaller than the termination parameter ($\epsilon$=0.001), or once the algorithm has reached a certain number of iterations. In our problem, we used 100 iterations. Finally, we mapped the clustered regions (plaque region, healthy tissue region, **OCT** deep-depth degraded signal region, and background) from reduced feature space back to the original image.

## 3.4 Plaque detection results

### 3.4.1 Qualitative evaluation of our results by comparing with ground truth

Figure 3.3. shows different images of vascular tissue with plaque build-up taken from 10 and 19-month-old **WHHLMI** rabbits: Fig. 3.3(a) raw OCT image of vascular tissue Fig. 3.3(b) processed OCT image; Fig.3.3(c) shows a photographic **OCT** image at the marked B-scan location; Fig.3.3(d) shows the oil red histology image of vascular tissue where red region depicting both the plaque and the remaining as non-plaque regions; Fig. 3.3(d) shows the ground truth, which was established based on the consensus of all four surgeons; Fig. 3.3(e) shows the plaque detection results of analysis of the **OCT** image using the full set of 26 textural features, and Fig. 3.3(f) shows the plaque detection results of analysis of the **OCT** image using the reduced set of 6 textural features obtained from GA algorithm; Fig3.3(h) shows the plaque detection results of analysis of the **OCT** image using the reduced set of 6 textural features obtained from sparse PCA method. Similar results are shown in Fig. 3.4, Fig. 3.5, Fig. 3.6, Fig. 3.7, Fig. 3.8, Fig. 3.9, Fig. 3.10, Fig. 3.11, and Fig. 3.12.

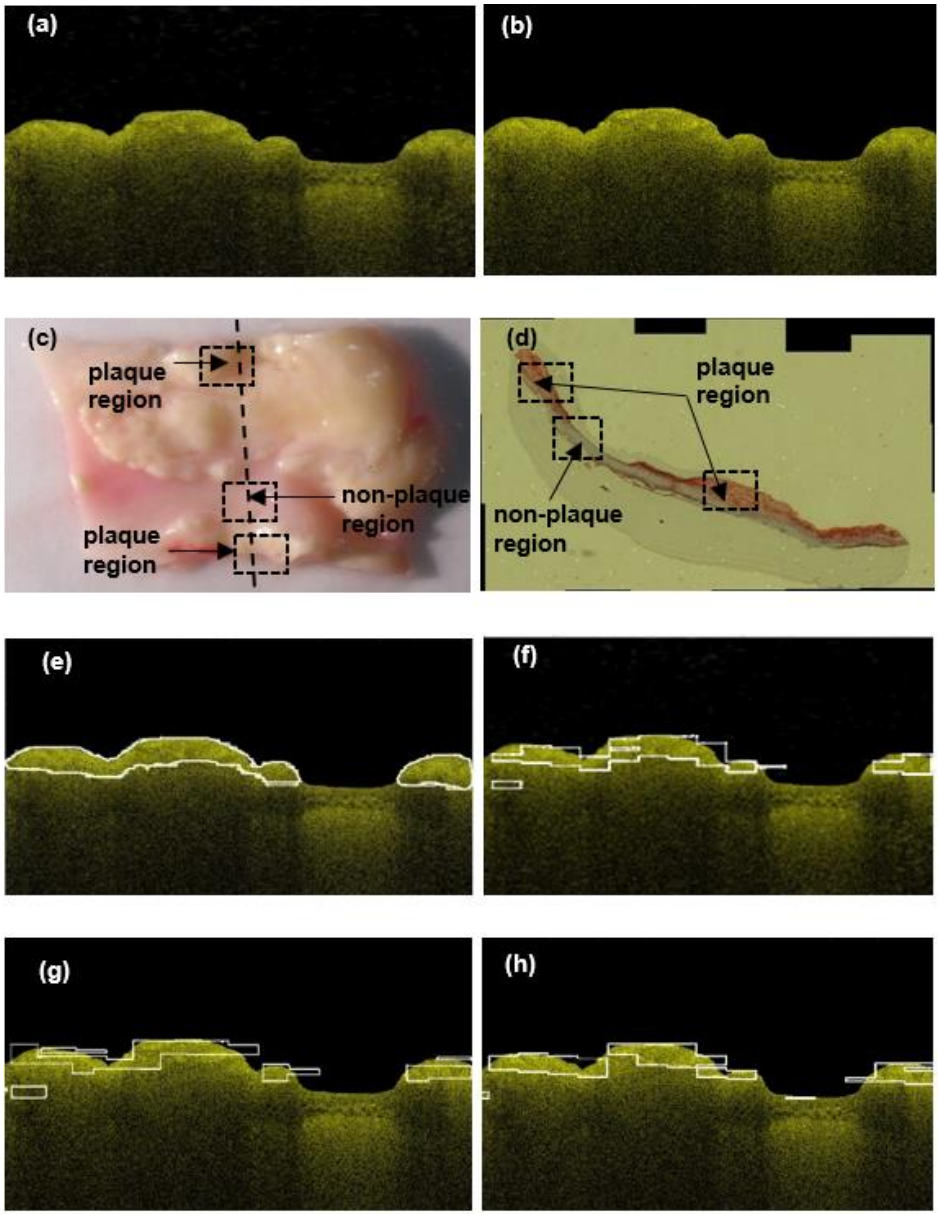



Fig. 3.3. (a) Raw **OCT** image of vascular tissue section 5-1-1 taken from a 309 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

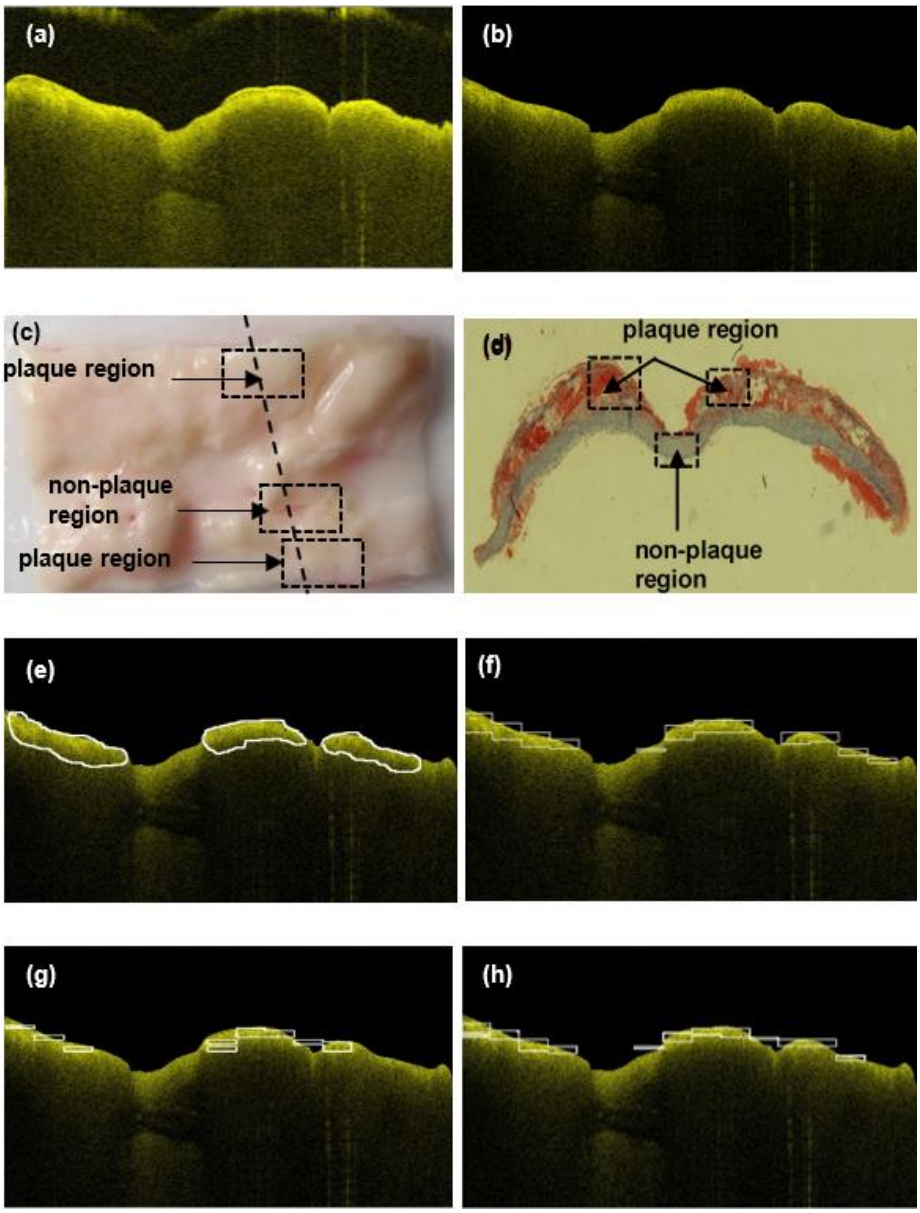



Fig. 3.4. (a) Raw **OCT** image of vascular tissue section 5-1-2 from a 316 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

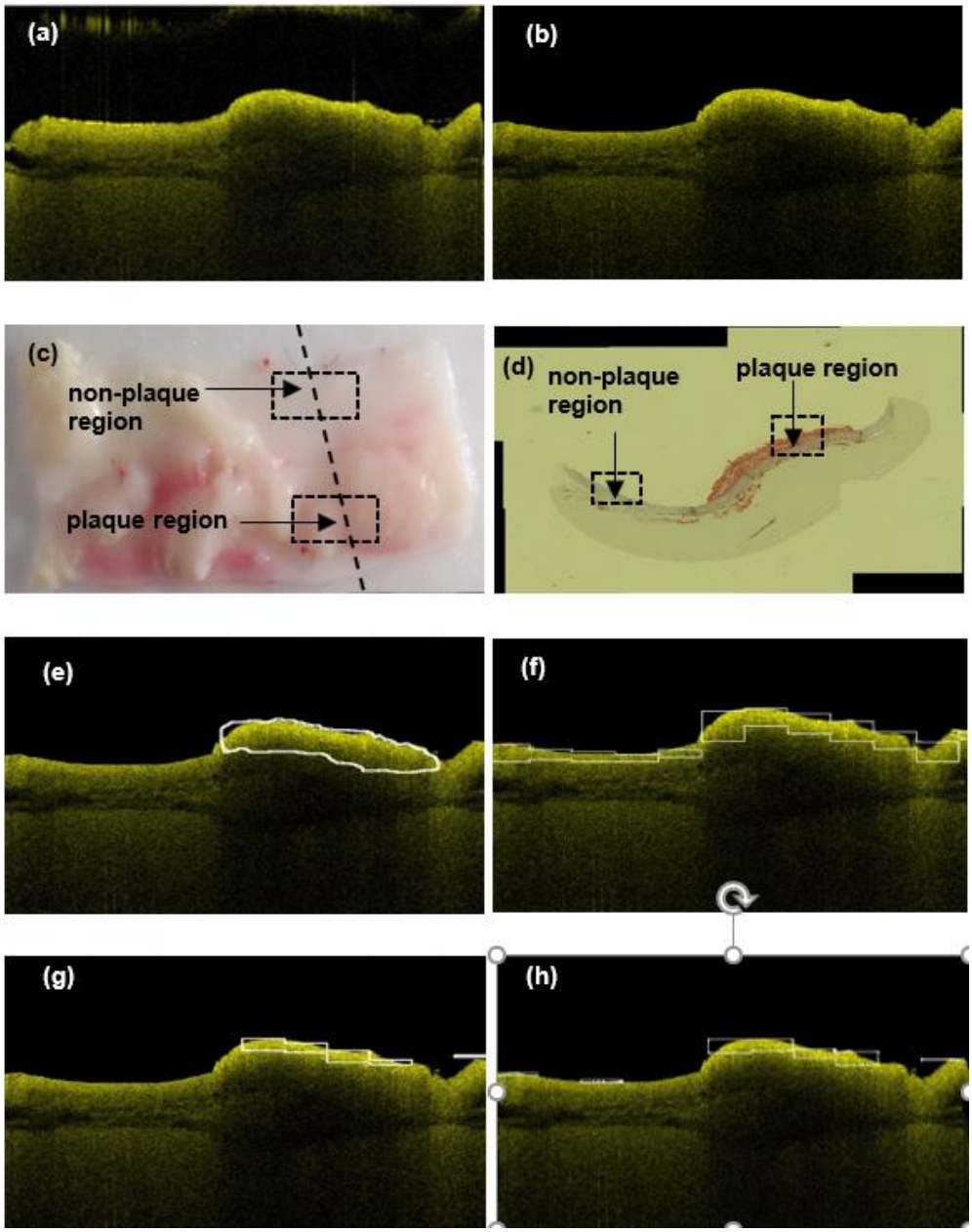



Fig. 3.5. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 316 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from GA algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

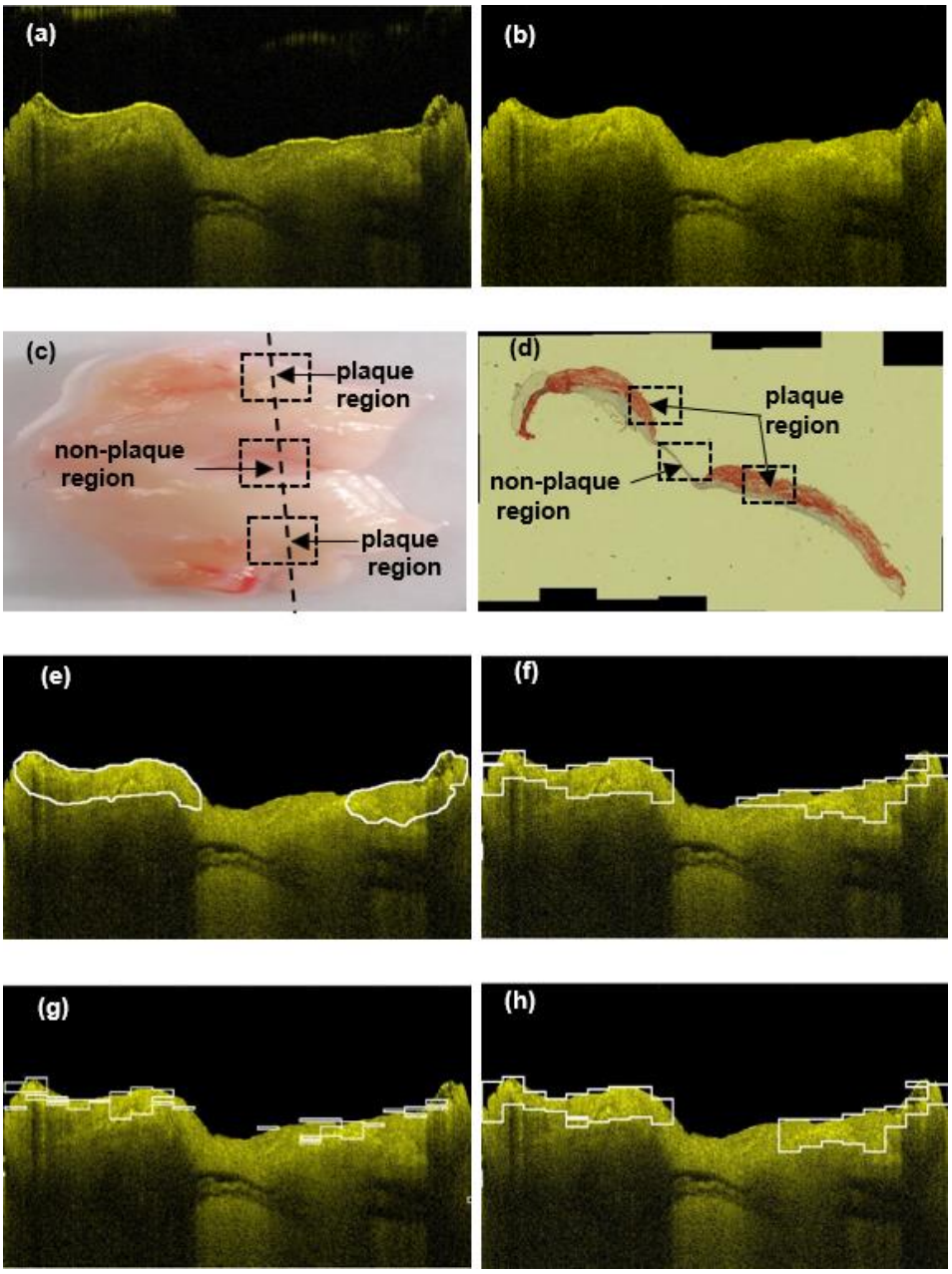



Fig. 3.6. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 330 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

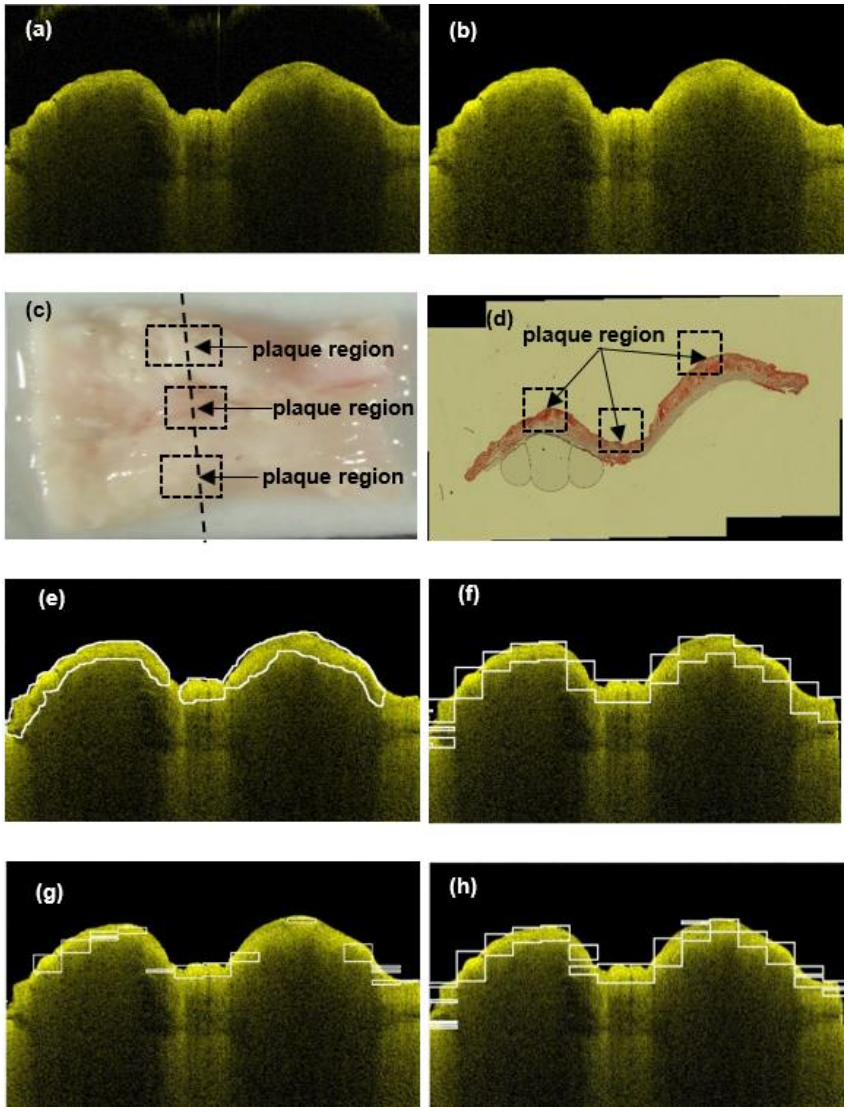



Fig. 3.7. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 342 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

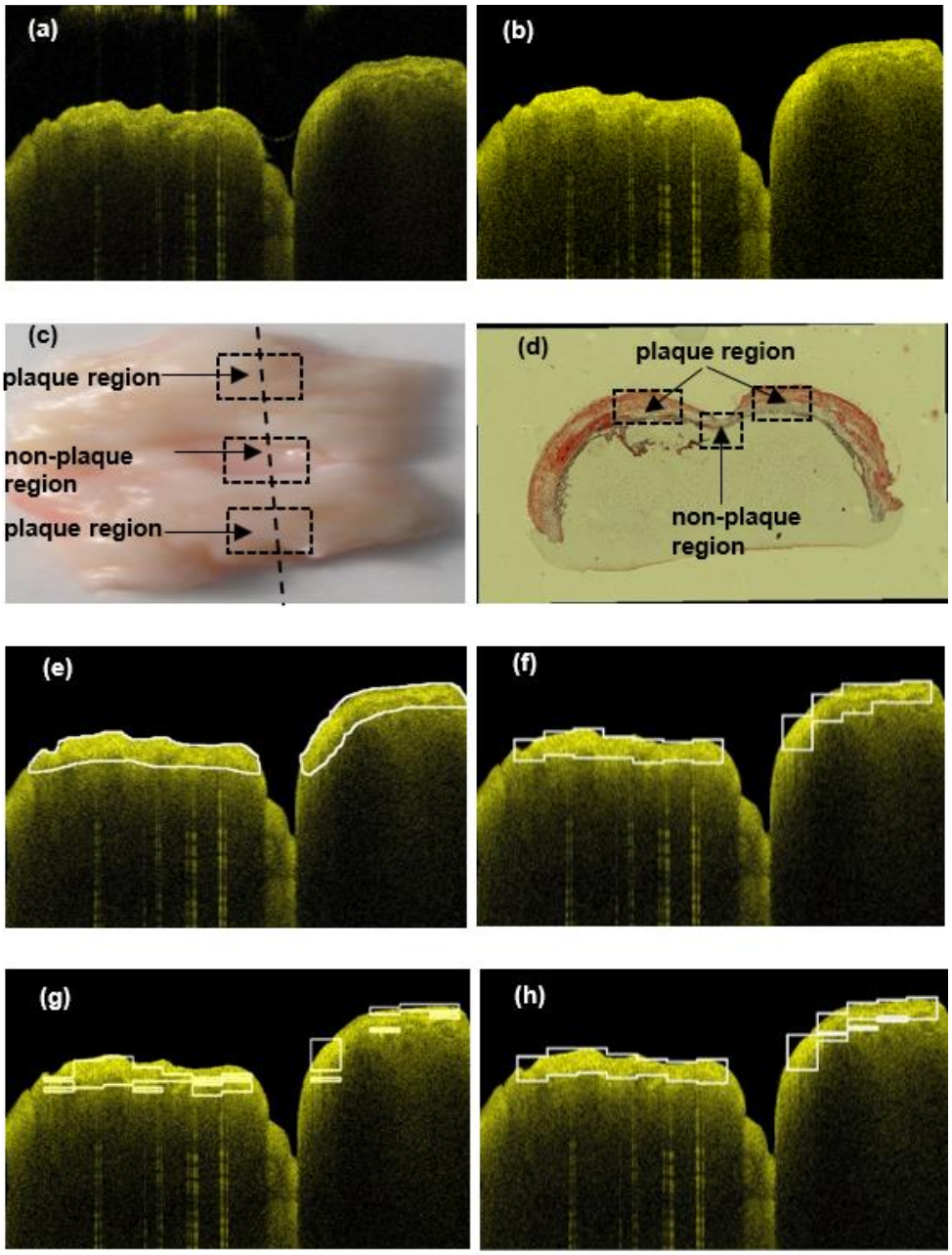



Fig. 3.8. (a) Raw OCT image of vascular tissue section 6-1-2 taken from a 365 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

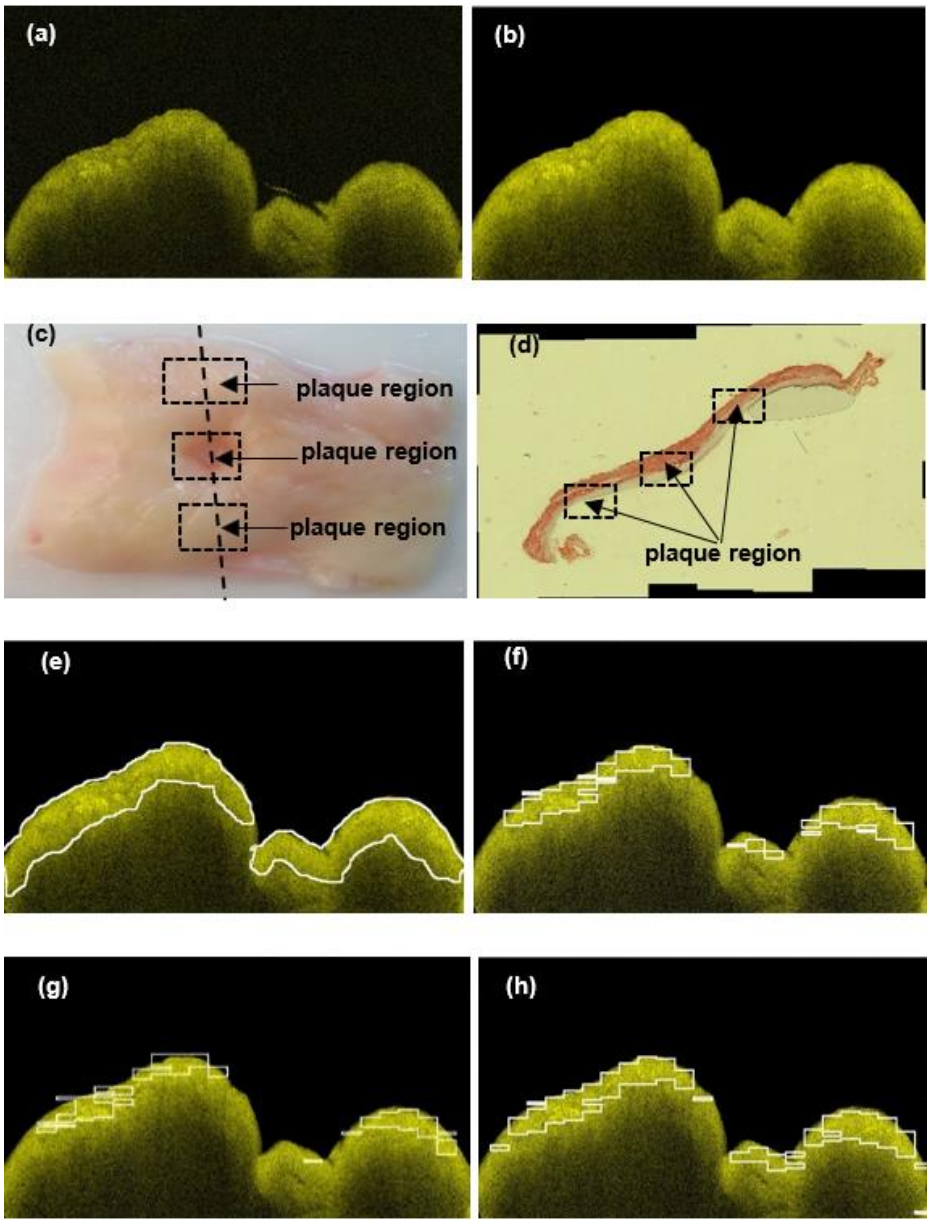



Fig. 3.9. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 456 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

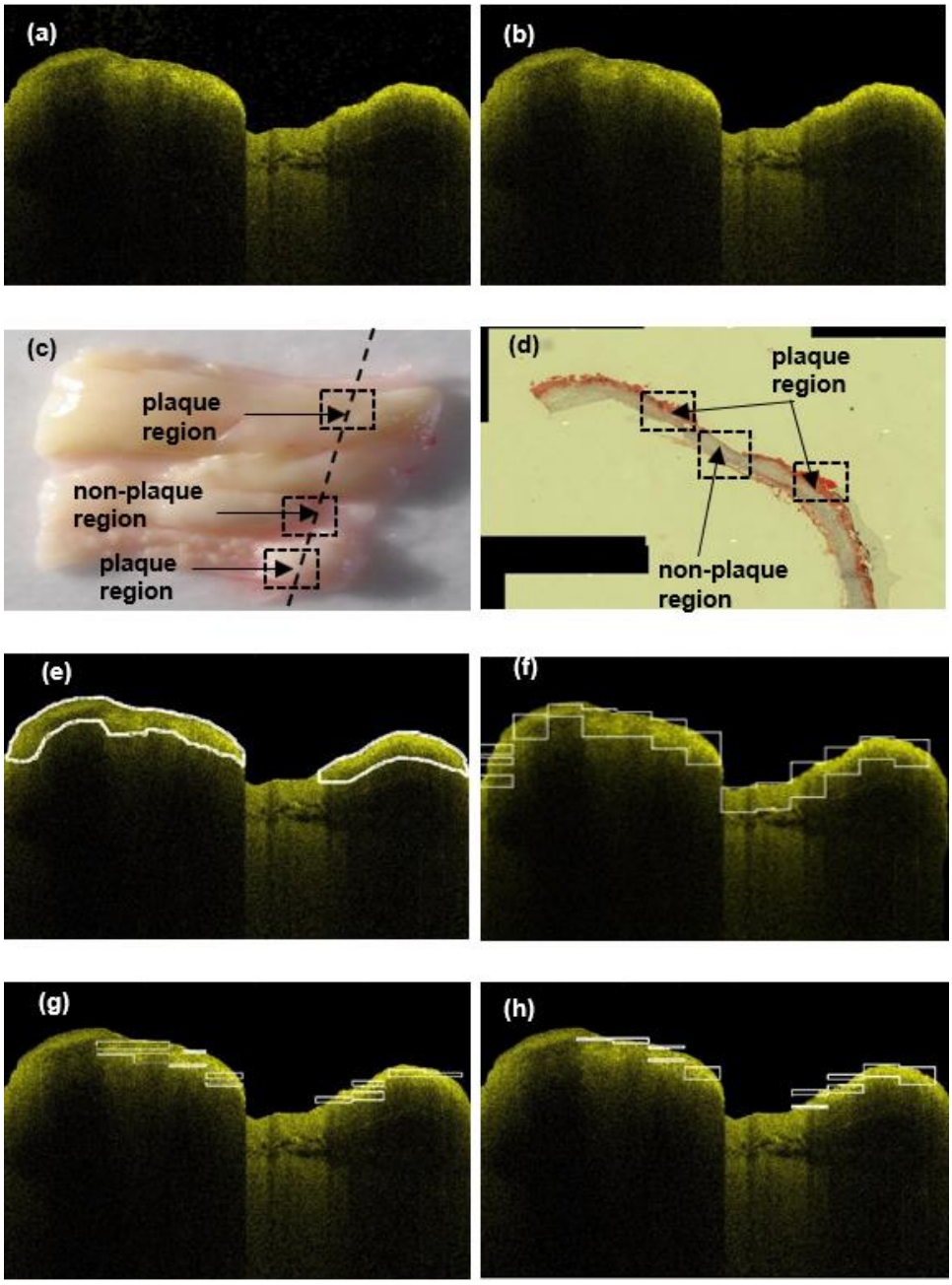



Fig. 3.10. (a) Raw **OCT** image of vascular tissue section 5-1-2 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

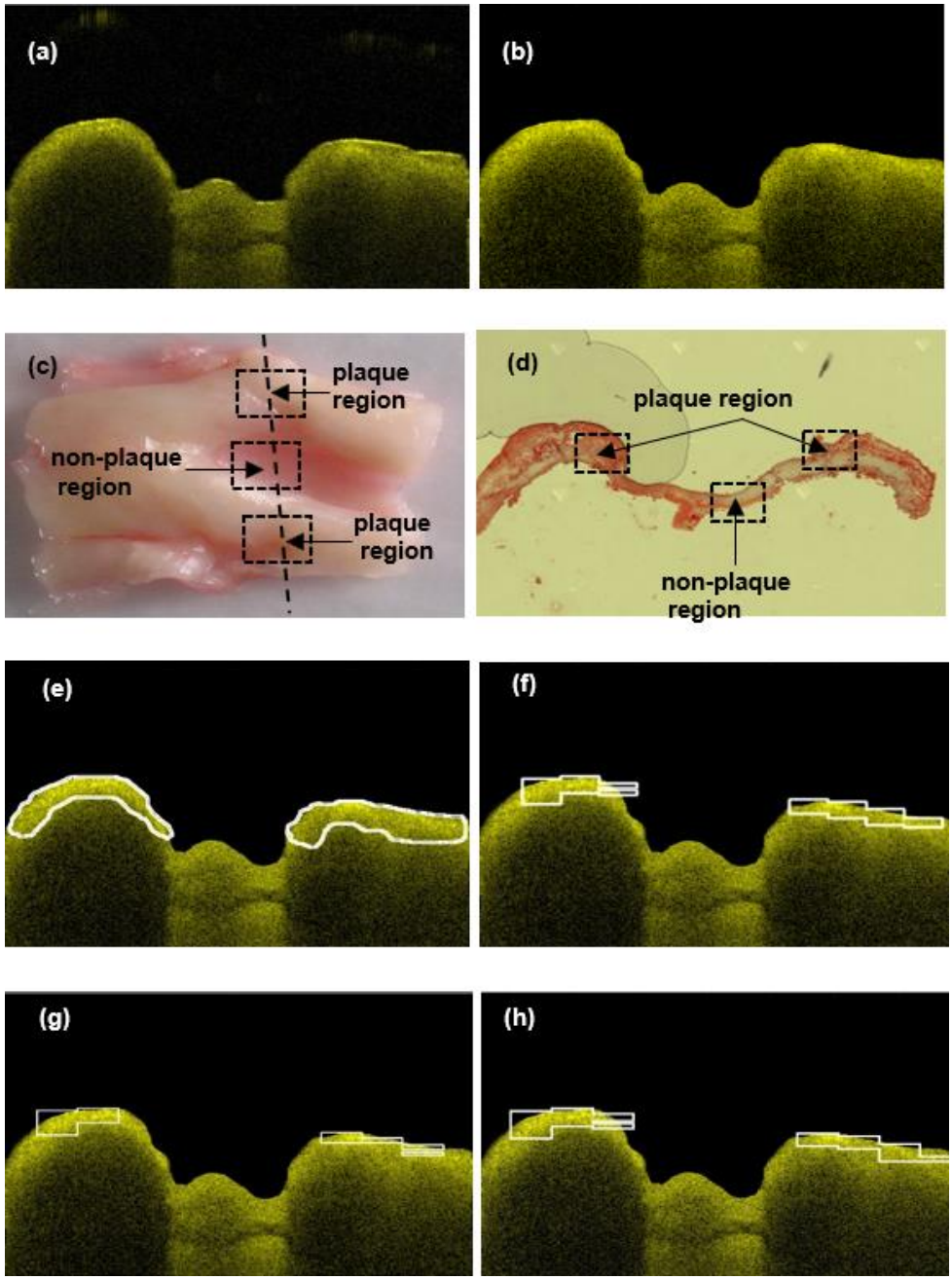



Fig. 3.11. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

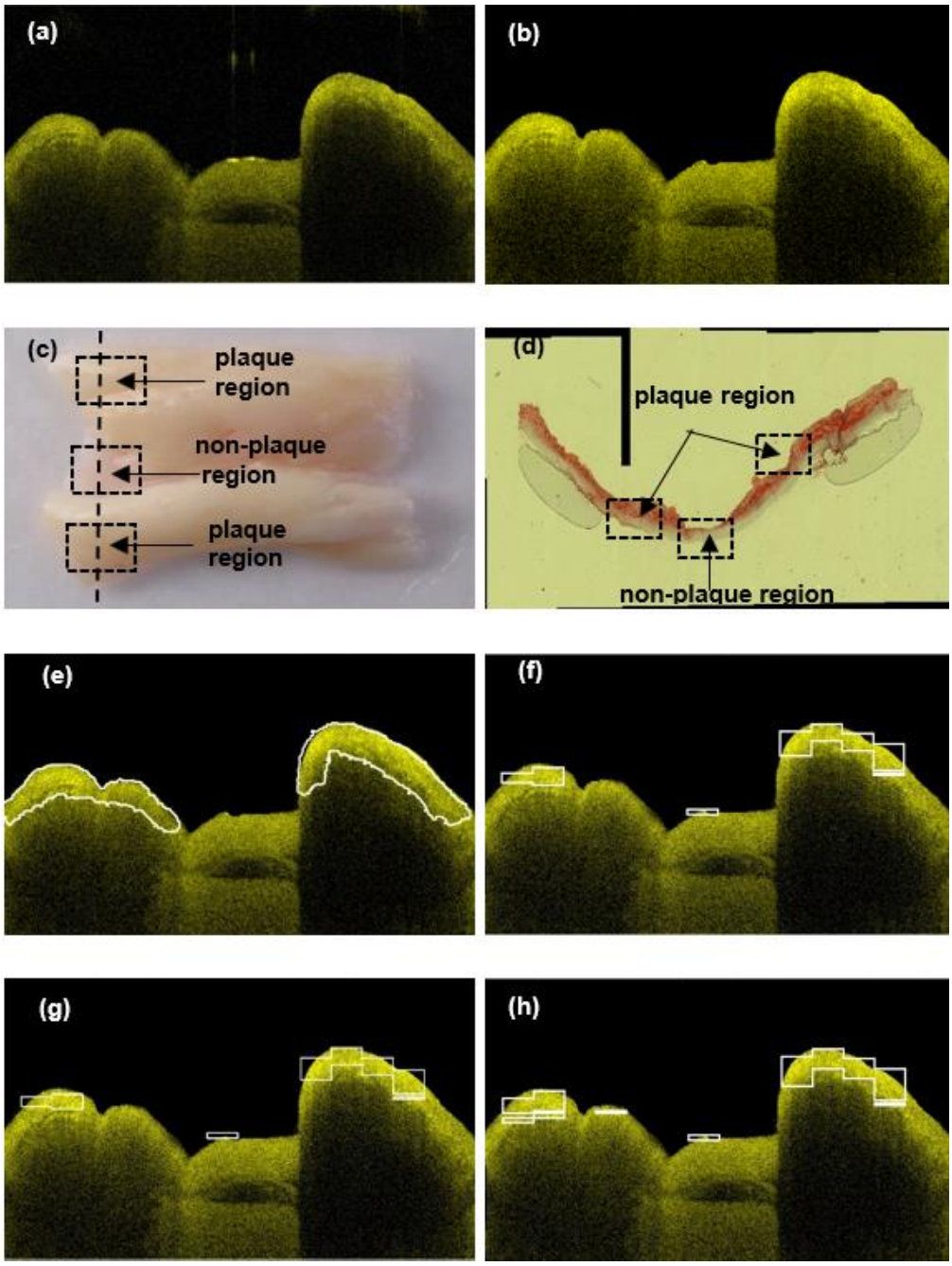



Fig. 3.12. (a) Raw **OCT** image of vascular tissue section 5-1-1 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue section; (b) oil red histology of vascular tissue; (c) ground truth; (d) plaque detection results of analysis of an **OCT** image of vascular tissue showing plaque and no-plaque regions using the full set of 26 textural features; (e) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from **GA** algorithm; (f) plaque detection results of analysis of an **OCT** image of vascular tissue using the reduced set of 6 textural features from sparse **PCA** algorithm.

### 3.4.2 Quantitative evaluation of our results by using standard metrics

We evaluated our plaque segmentation results quantitatively by comparing the pixel accuracy, sensitivity, precision, **MCC**, Dice coefficient, and specificity of our results with the ground truth. Table 3.4 shows the segmentation results for all ten **OCT** images with the full feature set, Table 3.5 shows the results for all ten **OCT** images with reduced features set using **GA**, and Table 3.6 shows the results for all ten **OCT** images with the reduced features using sparse **PCA**.

Table 3.4. Quantitative evaluation of Fuzzy C-means plaque segmentation algorithm using the full 26 feature set

| Image # | Image ID | Sample rabbit age (days) | Sensitivity (%) | Specificity (%) | Accuracy (%) | Precision (%) | MCC (%) | Dice Coefficient (%) |
|---|---|---|---|---|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 65.947 | 96.557 | 94.513 | 57.808 | 58.817 | 61.610 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 38.362 | 98.542 | 96.013 | 53.586 | 43.347 | 44.713 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 35.088 | 97.868 | 94.316 | 49.665 | 38.868 | 41.123 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 65.522 | 97.026 | 94.978 | 60.496 | 60.275 | 62.909 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 78.567 | 94.472 | 93.453 | 49.338 | 59.092 | 60.613 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 80.410 | 98.608 | 97.469 | 79.402 | 78.557 | 79.906 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 36.647 | 99.317 | 90.656 | 89.583 | 53.740 | 52.015 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 25.303 | 99.305 | 94.021 | 73.689 | 40.973 | 37.671 |
| 9 | 2010-02-02-086-02-612 | 577 | 36.981 | 99.268 | 95.248 | 77.709 | 51.621 | 50.113 |
| 10 | 2010-02-02-086-511 | 577 | 34.893 | 98.299 | 92.953 | 65.390 | 44.493 | 45.504 |

Table 3.5. Quantitative evaluation of Fuzzy C-means plaque segmentation algorithm using reduced feature set obtained from the **GA** algorithm.

| **Image #** | **Image ID** | **Sample rabbit age (days)** | **Sensitivity (%)** | **Specificity (%)** | **Accuracy (%)** | **Precision (%)** | **MCC (%)** | **Dice Coefficient (%)** |
|---|---|---|---|---|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 50.481 | 97.689 | 94.537 | 60.973 | 52.617 | 55.233 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 22.985 | 99.413 | 96.201 | 63.221% | 36.639 | 33.713 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 19.972 | 99.720 | 95.209 | 81.062 | 38.806 | 32.048 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 30.189 | 98.747 | 94.291 | 62.615 | 40.942 | 40.737 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 32.130 | 98.807 | 94.531 | 64.851 | 43.212 | 42.971 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 40.812 | 99.303 | 95.643 | 79.631 | 55.144 | 53.966 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 22.628 | 99.702 | 89.050 | 92.404 | 42.617 | 36.354 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 20.932 | 99.757 | 94.128 | 86.897% | 40.975 | 33.737 |
| 9 | 2010-02-02-086-02-612 | 577 | 25.683 | 97.985 | 93.318 | 46.789 | 31.460 | 33.163 |
| 10 | 2010-02-02-086-511 | 577 | 25.098 | 98.882 | 92.661 | 67.403 | 38.210 | 36.577 |

Table 3.6. Quantitative evaluation of Fuzzy C-means plaque segmentation algorithm using the reduced feature set obtained from the sparse **PCA** algorithm.

| **Image #** | **Image ID** | **Sample rabbit age (days)** | **Sensitivity (%)** | **Specificity (%)** | **Accuracy (%)** | **Precision (%)** | **MCC (%)** | **Dice Coefficient (%)** |
|---|---|---|---|---|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 66.308 | 97.006 | 94.957 | 61.307 | 61.058 | 63.710 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 37.319 | 99.378 | 96.770 | 72.483 | 50.607 | 49.270 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 29.347 | 99.199 | 95.248 | 68.729 | 42.954 | 41.131 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 74.833 | 96.370 | 94.970 | 58.903 | 63.774 | 65.920 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 69.710 | 96.678 | 94.949 | 58.977 | 61.448 | 63.896 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 79.111 | 99.033 | 97.786 | 84.519 | 80.598 | 81.725 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 47.690 | 99.363 | 92.222 | 92.315 | 63.067 | 62.891 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 32.473 | 99.670 | 94.004 | 90.056 | 52.018 | 47.733 |
| 9 | 2010-02-02-086-02-612 | 577 | 24.323 | 99.678 | 94.297 | 85.308 | 43.761 | 37.853 |
| 10 | 2010-02-02-086-511 | 577 | 34.292 | 99.702 | 95.480 | 88.810 | 53.583 | 49.479 |

We also compared the quantitative segmentation results using a full feature set, the reduced feature set obtained from **GA**, and the reduced feature set obtained from the sparse **PCA** method for precision, **MCC**, and Dice similarity metric, as these three metrics are the most important and relevant metrics for plaque segmentation problem.

Fig. 3.13 shows the comparison of segmentation results of **OCT** vascular tissue images obtained using full features, reduced feature set using **GA**, and the reduced feature set using sparse **PCA** method. Our reduced feature set using sparse **PCA** generated notably higher results in comparison to the full feature set and **GA**.

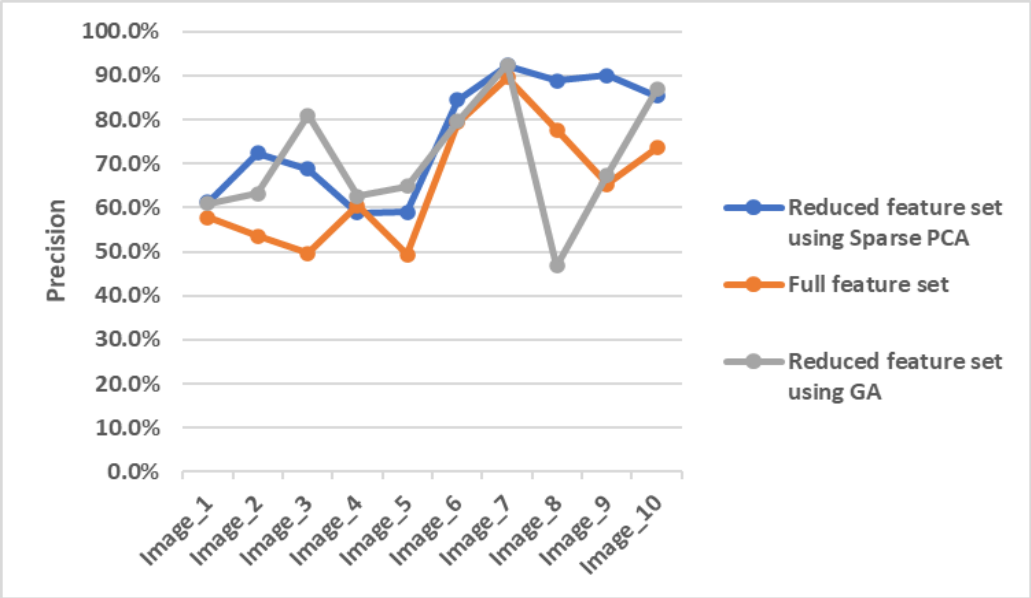



Fig. 3.13. Comparison of precision metric of segmentation results obtained using reduced feature sets from **GA** and sparse **PCA** method, and a full feature set.

Fig. 3.14 and Fig.3.15 show the comparison of segmentation results obtained using full features, reduced feature set using **GA**, and the reduced feature set using sparse **PCA** method for **MCC** and Dice similarity coefficient metric, respectively. It is clear from both graphs that our reduced feature set using sparse **PCA** generated better results, followed by the full feature set and then **GA**.

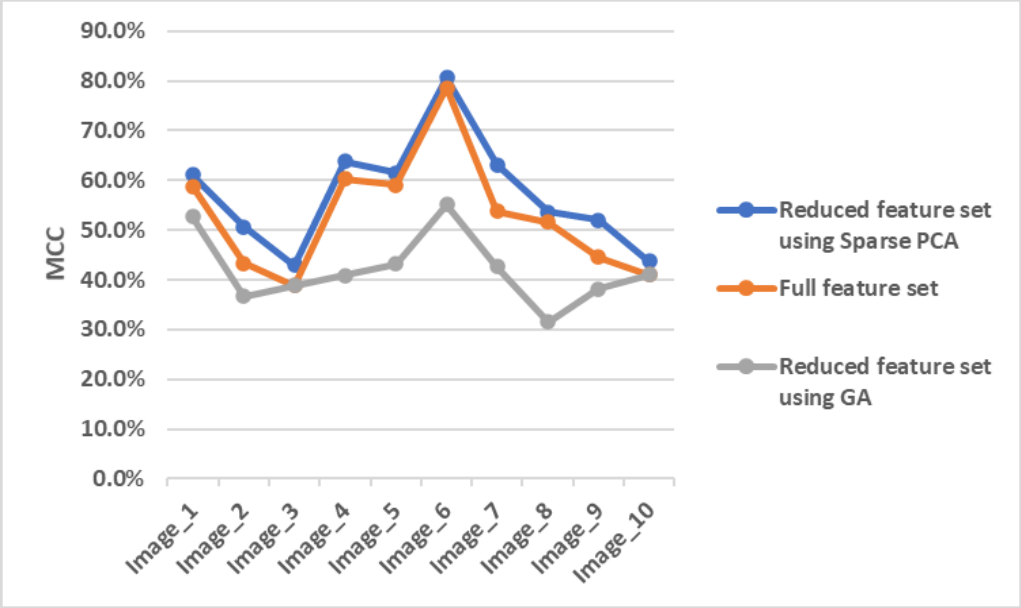



Fig. 3.14. Comparison of **MCC** metric of segmentation results obtained using reduced feature sets from **GA** and sparse **PCA** method, and a full feature set

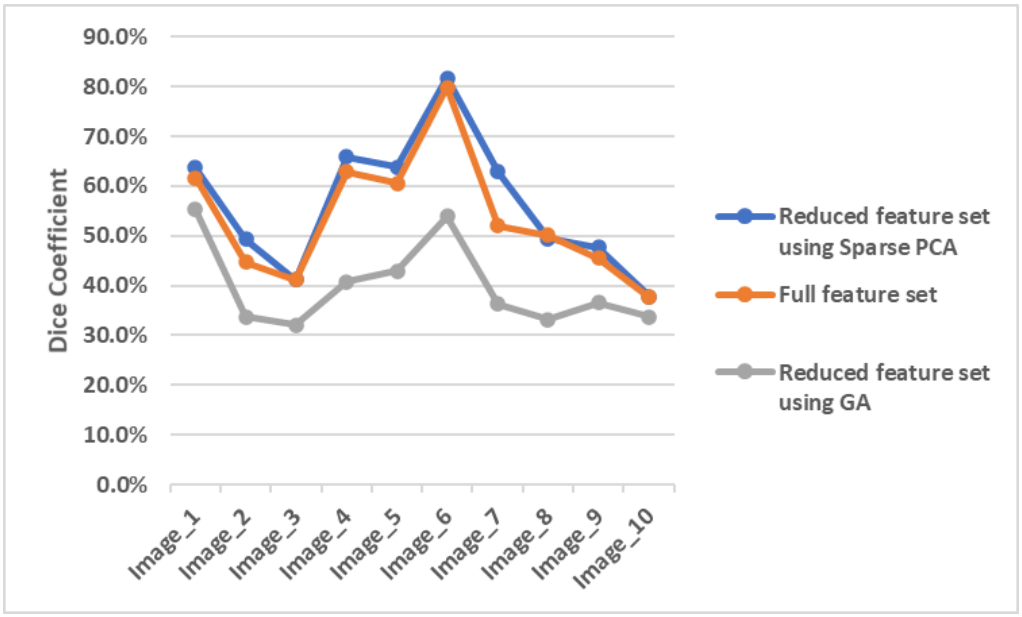



Fig. 3.15. Comparison of Dice coefficient metric of segmentation results obtained using reduced feature sets from **GA** and sparse **PCA** method, and a full feature set

## 3.5 Discussions

### 3.5.1 Discussion on quantitative results

Fig. 3.16 shows the average values of all the evaluation metrics over all the 10 images. We compared the segmentation results obtained using reduced feature sets from **GA** and the sparse **PCA** method, and a full feature set. In general, reduced features obtained from the sparse **PCA** method performed the best. We now discuss results obtained for individual error metrics.

For accuracy and specificity, although all the three approaches performed well, the best result of 95% accuracy and 99% specificity obtained for the reduced feature set using sparse **PCA**. Since the count of true negatives, which is the non-plaque region, is far greater than true positive counts, i.e., plaque region, accuracy, and specificity may inflate the performance, which could be misleading.

For sensitivity, segmentation results obtained from the reduced feature set method of 50% performed far better than using the full feature set. This could be attributed to the fact that full features could increase model overfitting. This is also clear from Fig. 3.5 and Fig. 3.10 panel (e), where the results obtained from using the full feature set show a non-plaque region segmented as a plaque region.

For precision, **MCC**, Dice similarity coefficient, the segmentation result obtained from the reduced set using sparse **PCA** method performed better than the **GA** method and using full feature set. This is because **GA** is not quite robust, and the optimal solution (i.e., the best combination of features) difficult to be found as the algorithm could get stuck in local minima. With a full feature set (i.e., without any feature selection), the results are prone to overfitting and thereby segmenting the nonplaque region as plaque region, which in turn increases the false positive count.

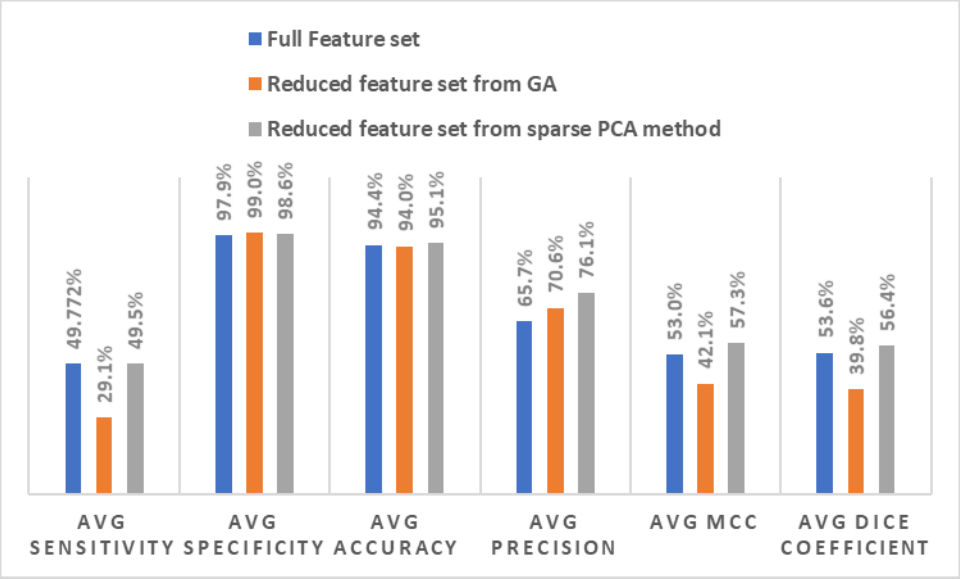



Fig. 3.16. Comparison of average values of evaluation metrics of segmentation obtained using reduced feature sets from **GA** and sparse **PCA** method, and a full feature set

### 3.5.2 Discussion on computation time

We also profiled the computation time required in case both the full feature set and a reduced feature set, as shown in table 3.5.

Table 3.4. Computation-time comparison between the 26-feature set and the 6-feature set.

| Image # | Image ID | Sample rabbit age (days) | Computation time with 26 feature set (s) | Computation time with 6 feature set (s) |
|---|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 27.438 | 13.92 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 15.56 | 9.97 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 14.232 | 10.71 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 18.539 | 16.26 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 15.616 | 9.34 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 22.298 | 11.00 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 18.970 | 11.75 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 22.44 | 12.27 |
| 9 | 2010-02-02-086-02-612 | 577 | 13.042 | 8.28 |
| 10 | 2010-02-02-086-511 | 577 | 14.404 | 8.86 |

As our time comparison results show, the speed of our algorithm more than doubles when the set of features is reduced, as shown in Fig. 3.17.

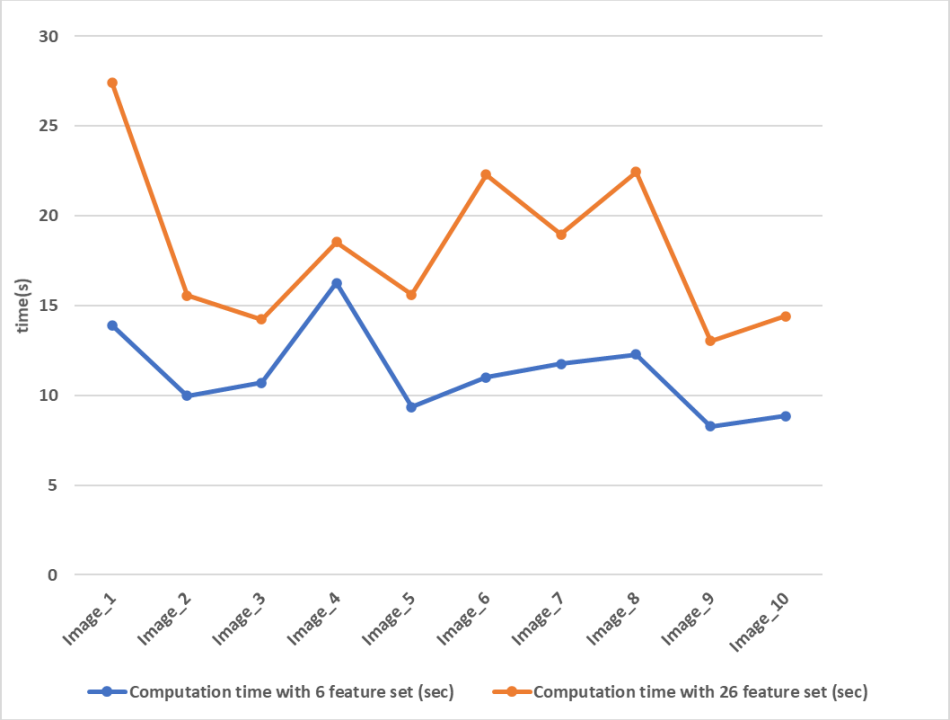



Fig. 3.17. computation time required using the full feature set and the reduced feature set

All experiments were performed on a machine with an Intel Core (i5) CPU, 8 GB RAM, and a Windows 7 OS using an interpreted programming language, MATLAB. All execution times exclude the time required to read the images and calculate the metrics.

### 3.5.3 Discussion on the effect of image window size for feature generation

Texture segmentation consists of dividing an image into different regions of similar textural properties. In the first stage of texture segmentation, we extract textural features and perform feature selection, while the second stage entails using previously selected features to segment similar regions via clustering-based algorithms. However, the quality of any texture-based segmentation dramatically depends upon the window size over which features are calculated. We analyzed various window sizes over a region of interest from the plaque and non-plaque regions, as shown in Fig. 3.18. Small windows may not have enough pixels to accurately capture the texture

of underlying tissue, while windows that are too large may contain tissues with significantly different textures, which may result in coarse segmentation results. We chose the window size experimentally. We tried several different window sizes, with our findings showing that a window size of 32x32 pixels yielded the best results for plaque segmentation from other regions.

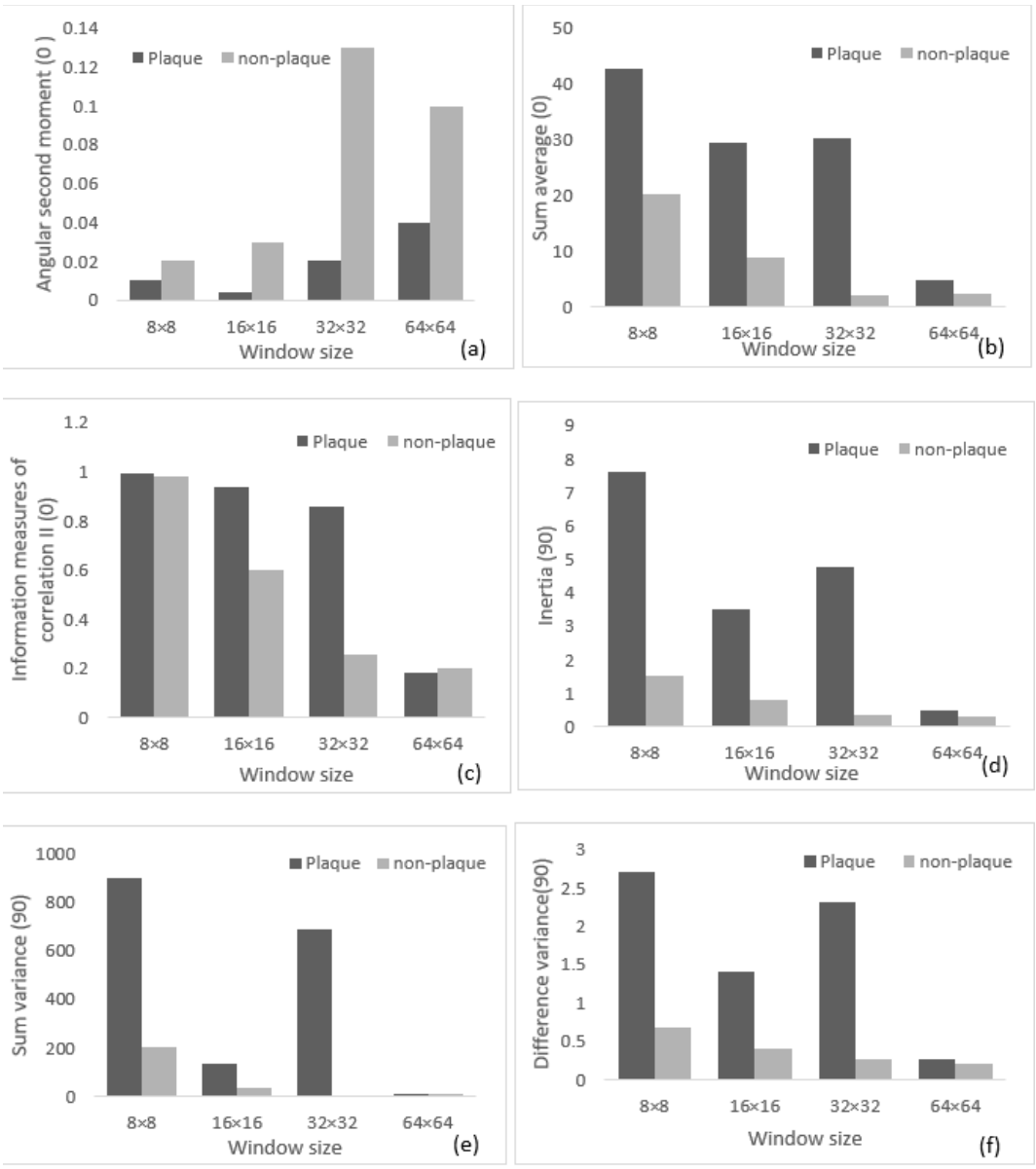



Fig. 3.18. Comparison of second-order parameters extracted from the **SGLDM** of plaque vs non -plaque regions on different windows size: (a) angular second moment (**ASM** ); (b) sum average; (c) information measures of correlation II; (d) inertia; (e) sum variance; (f) difference variance. Features were evaluated with various window sizes — Gray bars plaque region; White bars: healthy regions.

## 3.6 Conclusions

In this chapter, we presented two different feature selection methods (**GA** and sparse **PCA**) and an advanced clustering technique (*Fuzzy C-means*) to enable the detection of vascular tissue texture using **OCT** images.

We were able to successfully reduce the full set of 26 textural features down to a set of 6 textural features, and we quantitatively evaluated the accuracy of our method by comparing our plaque detection results with both the full texture feature set and the reduced feature sets obtained from **GA** and sparse **PCA** method (see Table 3.4, 3.5 and 3.6) using the ground truth images. Our results show that although all the three approaches (full features set, reduced feature set obtained using **GA**, reduced feature set obtained using sparse **PCA**) were successful in identifying plaque region, our reduced feature set obtained using sparse **PCA** based method provided the most satisfactory results both visually and quantitatively. We found the average sensitivity, specificity, accuracy, precision, MCC and Dice similarity coefficient for our three approaches (full features set, reduced feature set obtained using **GA**, reduced feature set obtained using sparse **PCA**) as (49.541%, 97.926%, 94.362%, 65.667%, 52.978%, 53.618), (29.091%, 99.001%, 93.957%, 70.585%, 42.062%, 39.850%) and (49.541%, 98.608%, 95.068%, 76.141%, 57.287%, 56.361%), respectively.

Our new reduced feature sets from sparse **PCA** method, which are $f_1$, $f_6$, $f_{13}$, $f_{16}$, $f_{20}$, and $f_{23}$ (Angular Second Moment (**ASM**) at $0^o$, Sum average at $0^o$, Information measures of correlation II at $0^o$, Inertia at $90^0$, Sum Variance at $90^0$, and Difference Variance at $90^0$), along with *Fuzzy C-means* clustering, and with its average processing time of ~ 11 secs could help to detect vascular plaque using **OCT** images.

# Chapter 4: Detection of Vascular Plaque from Optical Coherence Tomography Images Using Markov Random Field Model-Based Segmentation

## 4.1 Introduction

In any image segmentation problem, the objective is to divide an image into finite subregions; therefore, image segmentation is inherently a discrete problem. Typically, image segmentation approaches can be classified into three groups: (*i*) region-based, (*ii*) contour-based, and (*iii*) clustering. The contour-based algorithm is based on boundary features and partitions the image based on the boundaries of each object in the image. It does not involve any stochastic technique; it starts with a spline curve and optimizes it based on the energy function. One of the disadvantages of such methods is that they tend to get stuck in the local minima easily. In addition, these methods are not completely automatic and require manual assistance to initialize the curves. Region-based segmentation methods are based on partitioning different homogeneous regions. One of the traditional region-based algorithms is watershed segmentation, which segments the image into different regions based on pixel intensity. Usually, this algorithm is used to segment the foreground and background in an image. However, this approach often produces unsatisfactory results for images with many different regions. Two of the most common non-parametric statistical-based clustering methods are *K-means* and *Fuzzy C-means.* Some of the major limitations of the *K-means* clustering method are that it does not work very well with global clusters, and the algorithm has a difficult time converging when data has outliers. Thus, non-parametric methods may not result in robust image segmentation. Many machine-learning algorithms have also been proposed for detecting vascular plaque using medical imaging modalities. These modalities include computed

tomography using a regression approach, ***K-NN*** classifier, graph cuts, and kernel regression approaches. Ultrasounds using contour-based segmentation, *Fuzzy C-means* and ensemble clustering, backpropagation, support vector machines. In **OCT** using contour-based, clustering-based method. These clustering-based algorithms face several limitations, including non-smooth and coarse plaque detection results and less robust outcomes when high noise levels are present.

Therefore, to overcome these limitations, we propose a more robust stochastic algorithm that uses Bayesian segmentation based on an **MRF-MAP** approach. Our proposed algorithm considers images as random objects and pixels as random variables. In any image texture labeling problem, we observe the pixels that form a digitized picture, but not the type of texture (texture labels). Thus, the purpose of texture image segmentation is to estimate the texture type labels from the observed image.

Pixels that are close to each other or neighboring pixels tend to have similar textural properties. This information is known as contextual information. **MRF** is a probabilistic model that captures such contextual constraints with respect to neighboring pixels.

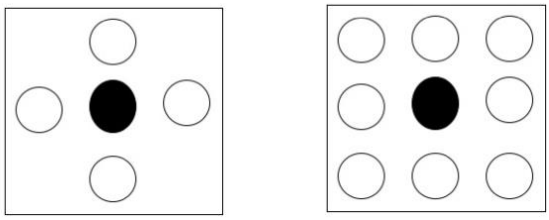
Consequently, we developed a stochastic model generated by **MRF** in order to address our plaque segmentation problem. We considered our **OCT** images as random objects and the constituent pixels as random variables, and we aimed to extract vascular plaque regions, which were label type and cannot be observed directly.

Our second approach is an **MRF**-based model. There are also models based on fractals, which have shown potential for use in modeling natural textural images [124-126]; however, these methods are not suitable for characterizing the local structures of images. Our model uses an **MRF**, which is a stochastic process wherein all interactions within a pixel are local; that is, the probability

of a pixel is solely determined by its neighboring states [127-128], with interactions between pixels always being direct with neighboring pixels.

## 4.2 Definition of Markov Random Field (MRF)

One of the basic building blocks of **MRF** is a neighborhood system, wherein each pixel's neighborhood is defined based on the pixels immediately surrounding it (Fig. 4.1).



(a) (b)

Fig. 4.1. (a) First-order neighborhood system, where the conditional probability of Xs (black pixel) depends only on four neighboring random variables represented as white pixels; (b) second neighborhood system, where the conditional probability of Xs (black pixel) depends only on eight neighboring random variables.

Let us assume an image to be a random field, X, and divided into $N_x \times N_y$ non-overlapping blocks. In this scenario, $X = \{X_{i,j} | (i,j) \in \Omega\}$, where $\Omega = \{(i,j) | 1 \leq i \leq N_x, 1 \leq j \leq N_y\}$ is an index set on $N_x \times N_y$ blocks (Fig. 4.2).

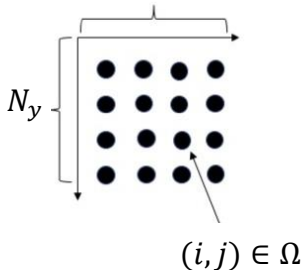



Fig. 4.2. 2D rectangular image block.

$X_{i,j}$ is an **MRF** if it is independent of other pixels outside the neighborhood, given all of the pixels in the neighborhood (i.e., $X_{i,j}$ is conditionally independent of all the pixels outside its neighborhood). Since pixel $X_{i,j}$ is completely characterized by its local conditional probability, there are certain limitations of **MRF** in terms of its conditional pdfs, such as:

(i) Computing the joint probability distribution function (**JPDF**) from its local conditional pdf is cumbersome and not straightforward.

(ii) The relationship between local spatial characteristics and the local conditional probability distribution function form is not very obvious.

To overcome these limitations, **MRF** can be characterized by a Gibbs Random Field (**GRF**) according to the *Hammersley-Clifford theorem*, which relates the **GRF-MRF** [129], to construct the **JPDF** of the random variable $X_{i,j}$ from its local conditional probabilities. This theorem bridges **MRF**s and **GRF**s and makes them equivalent.

The *Gibbs* distribution, which is the **JPDF** of all of the random variables in the random field, takes the following form:

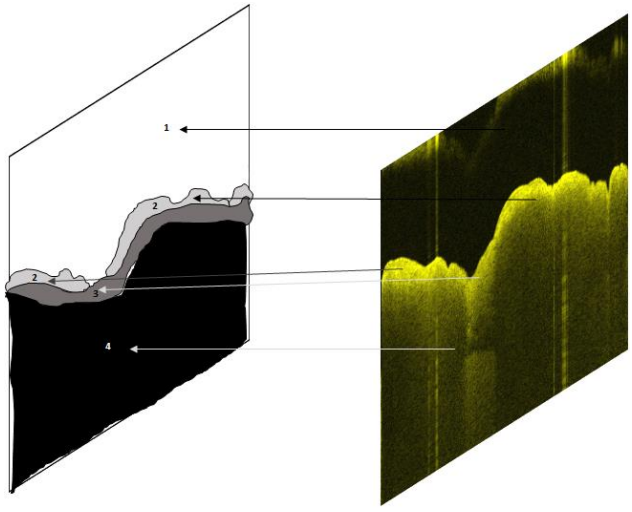
$$P(X = x) = \frac{1}{Z} exp\left[\frac{-u(x)}{T}\right] \tag{4.1}$$

where Z is the normalizing constant (also known as the partition function), and T is the temperature parameter. $U(x)$ is the energy function of the form,

$$U(x) = \sum_{c \in C} V_c\,(x) \tag{4.2}$$

where C is a set of cliques. A clique is a subset if every pair of pixels in the subset are neighbors. For each clique, c $\in C$, $V_c(x)$ is a clique potential.

Any observed two-dimensional image can be represented as a random field, $Y = \{Y_s\}$, on the lattice, S, where $Y_s$ is the random variable in the observed space representing gray-level value (0…255). In an image segmentation problem, there exists a second different random field, *X*, that represents variables that are not directly observable, taking values in the space, $X = \{L1, L2, .. Lx\}$, where $L_x$ is the number of segmented regions that can only assume integer values, which, in our problem, is 4 (Fig. 4.3). These hidden variables, known as a Hidden Markov Random Field (**HMRF)**, represent the labels of different image regions.




***X*: Hidden OCT image random field with four hidden classes (labels)** ***Y*: Observed OCT image**

Fig. 4.3. Y is the observed **MRF**, X is the **HMRF** whose labels are estimated using the observed random field.

## 4.3 Representing OCT images as MRFs

We formulated our **OCT** image segmentation problem as an **MRF** estimation problem. The advantage of using **MRF**-model-based segmentation is that it allows spatial information to be incorporated into an image as *a priori* information [129]. We then used a maximum *a posteriori* (**MAP**) approach to obtain optimal values for labels corresponding to different image regions

(**MRF-MAP**). **MAP** criterion is the most commonly used in Bayesian segmentation problems due to the binary nature of its cost function.

$$C(x,\hat{x}) = 1 - \delta(x - \hat{x}) = \begin{cases} 1, if\ x \neq \hat{x} \\ 0, if\ x = \hat{x} \end{cases} \tag{4.3}$$

This cost function has a value of 1 if there is an error in labeling. Since the image segmentation problem is discrete by nature, this cost function will assign equal error to single mislabeled pixels or all mislabelled pixels.

Given the observed random field (i.e., the raw **OCT** image, Y), the **MAP** formulation finds the class label, which maximizes the conditional probability,

$$\hat{x} = \underset{x}{\operatorname{argmax}}\, P(x|y) \tag{4.4}$$

which could be rewritten using Bayes' formula as,

$$\hat{x} = \underset{x}{\operatorname{argmax}}\, P(y|x)\, p(x) \tag{4.5}$$

The term, $P(x)$, is the prior probability. One of the major advantages of using **MRF**-based models is that they allow priors to be represented using a *Gibbs* distribution, which is characterized by energy functions. These energy functions are more intuitive for modeling than working directly with probabilities. The prior in our texture segmentation problem is hidden texture label types, which we assumed to follow a *Gibbs* distribution.

As an *a priori* probability distribution, Gibb*s* Random Field (**GRF)** takes the form,

$$P(X = x) = \frac{1}{Z} exp\{-U(x)\} \tag{4.6}$$

with the energy function, $U(x)$, taking the form,

$$U(x) = -\beta \sum_{j \in \mathcal{N}_i} \delta(x_i - x_j) \tag{4.7}$$

where the summation is over all neighboring pairs, $\mathcal{N}_i$ is a neighborhood of $j$, and $x_i$ is the spatial label distribution that can take values of 1,2, 3, 4, $\delta(.)$ is the Kronecker delta function, and $\beta \geq 0$ is a scalar model parameter.

In our implementation, we used the four nearest neighbors. This type of **MRF** model is referred to as the *Potts* model, which is a generalized form of the *Ising* model.

$P(y|x)$ is the likelihood and represents the conditional probability density function of the observed image data given its segmentation labels. We assumed that the pixel intensity in the image would follow a *Gaussian* distribution given the segmentation labels.

$$P(y_i|x_i) = \prod_{i \in S} \frac{1}{\sqrt{2\pi}} \exp\left( \frac{-\left(y_i - \mu_{x_i}\right)^2}{2\sigma_{x_i}^2} - \log\left(\sigma_{x_i}\right) \right) \tag{4.8}$$

Using *Gaussian* distribution as the likelihood and *Potts* model as the MRF gives the posterior distribution as,

$$P(y_i|x_i) = \frac{1}{\sqrt{2\pi}\sigma_{x_i}^2} \exp\left( \frac{-\left(y_i - \mu_{x_i}\right)^2}{2\sigma_{x_i}^2} - \beta \sum_{j \in \mathcal{N}_i} \delta(x_i - x_j) \right) \tag{4.9}$$

where *y* is the observed data, and the summation is over the four neighboring pixels of $i$.

Given the form of the prior and likelihood functions, our next step in formulating the complete statistical model is to estimate all the unknown parameters, which is known as the parameter estimation step. We assumed a *Gaussian* distribution for the observed image and then used the Expectation-Maximization (**EM**) algorithm to estimate its parameters, including their mean and variance. The algorithm begins by estimating the initial values for the parameters, i.e., mean and standard deviation. Since the **EM** algorithm is sensitive to the initial estimates of these unknown parameters, we also used a *K-means* clustering algorithm to obtain appropriate initial estimates.

This was followed by **EM** steps, which iteratively updates our class labels and the parameters. In each iteration of our **MRF-MAP** approach, the iterated conditional modes **(ICM**) based optimization method [124] were used to estimate class. **ICM** is a deterministic optimization algorithm that uses a greedy strategy to maximize local conditional probability in a sequential manner. The computation time required in **ICM** is linear to the number of labels.

## 4.4 MRF-MAP model-based segmentation results

### 4.4.1 Qualitative evaluation of our results by comparing with ground truth

We validated our **OCT** image segmentation results by comparing them with both histology images and photographic images of vascular plaque samples. The validation results indicated that our proposed algorithm had a good ability with respect to detecting vascular plaque from OCT images, which means that it could be of significant help in using such images to diagnose cardiovascular diseases

Our **OCT** image-segmentation-based vascular plaque detection results are shown in Fig 4.4. To validate our results, we compared our segmentation results with actual photographic images (Fig. 4.4(b)) and oil red histology images (Fig. 4.4(c)). The red regions in the histology images highlight the plaque region, which accurately matches our segmented plaque region. Similar results are presented in Fig. 4.5, Fig. 4.6, Fig. 4.7, Fig. 4.8, Fig. 4.9, Fig. 4.10, Fig. 4.11, Fig. 4.12, and Fig. 4.13.

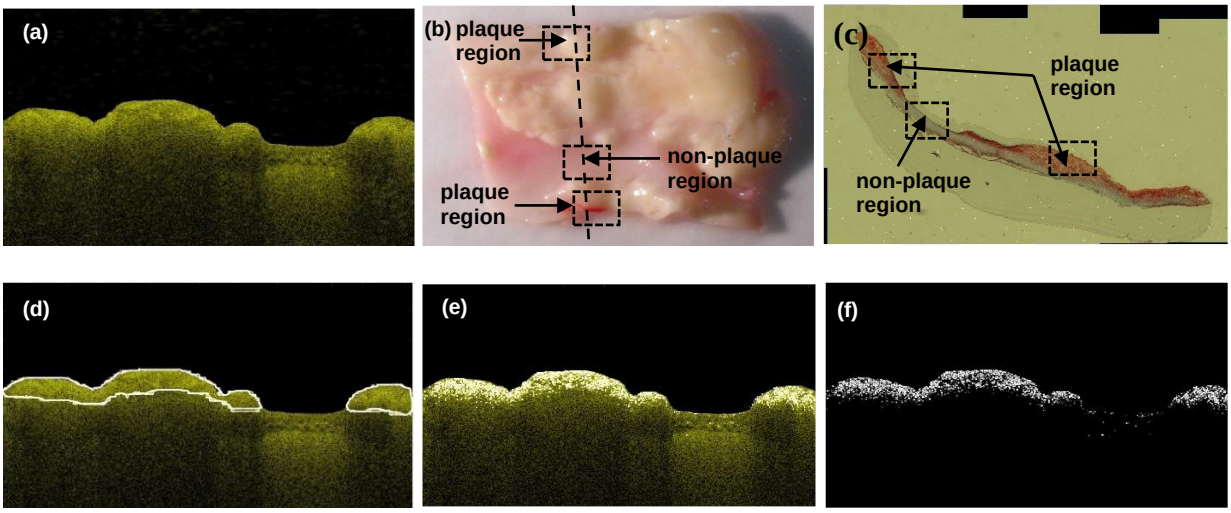



Fig. 4.4. (a) Raw **OCT** image of vascular tissue section 5-1-1 taken from a 309 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

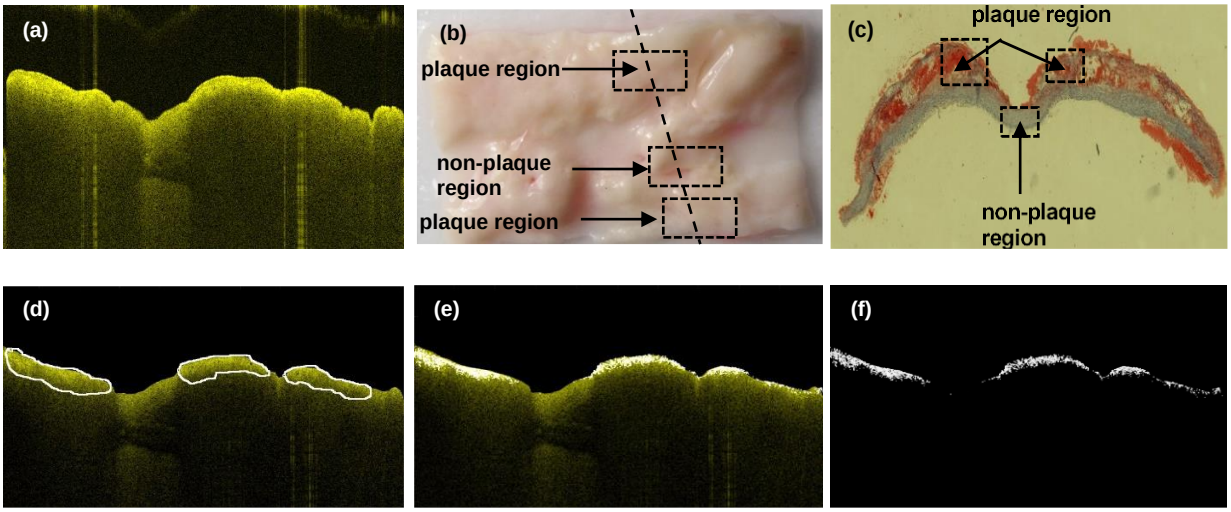



Fig. 4.5. (a) Raw **OCT** image of vascular tissue section 5-1-2 taken from a 316 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

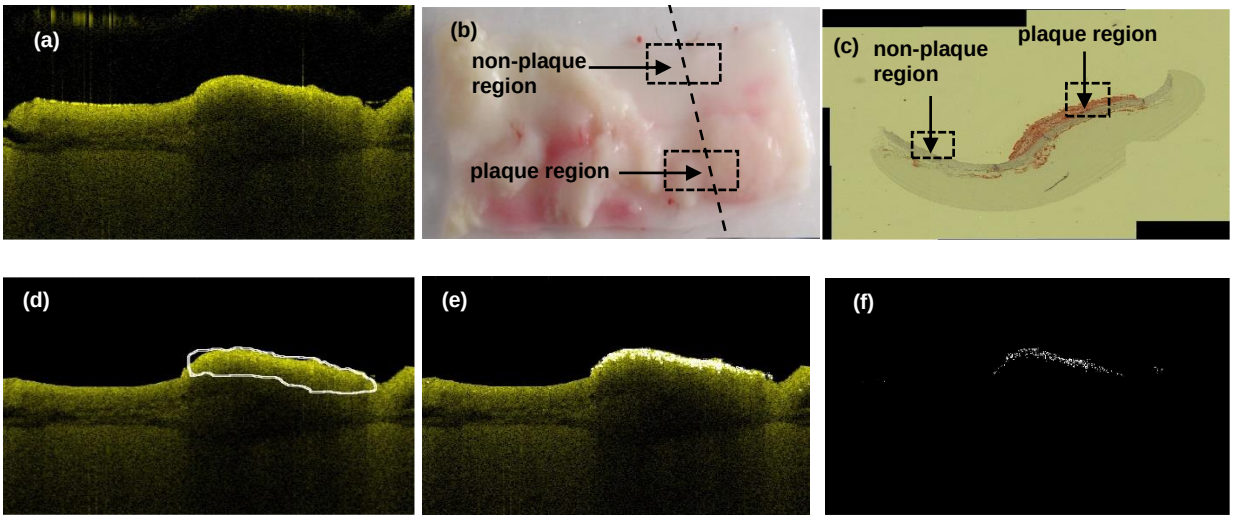



Fig. 4.6. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 316 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

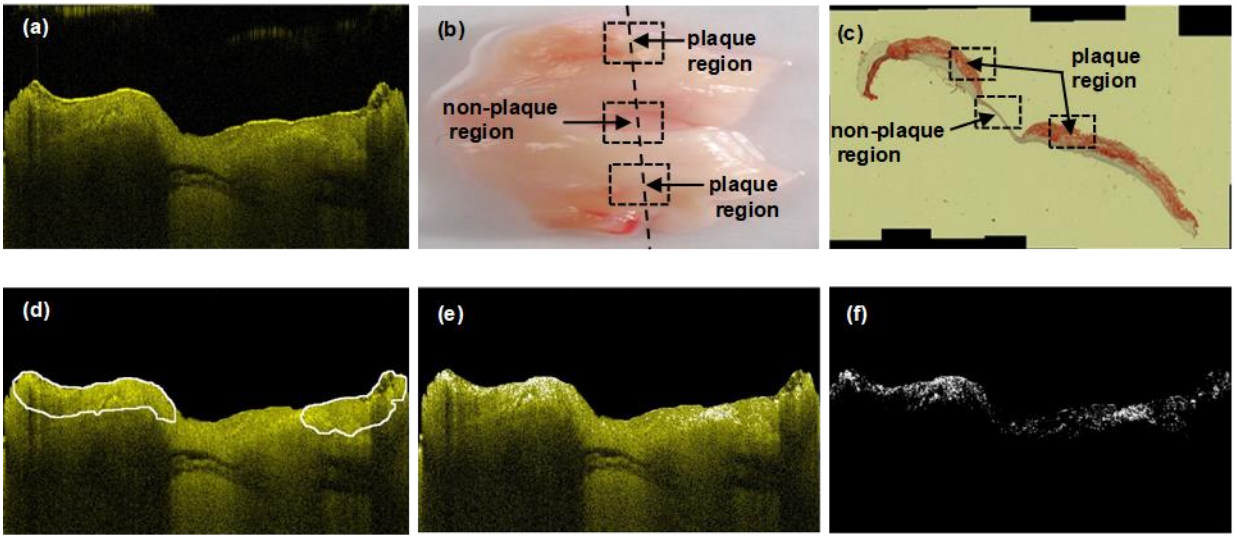



Fig. 4.7. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 330 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

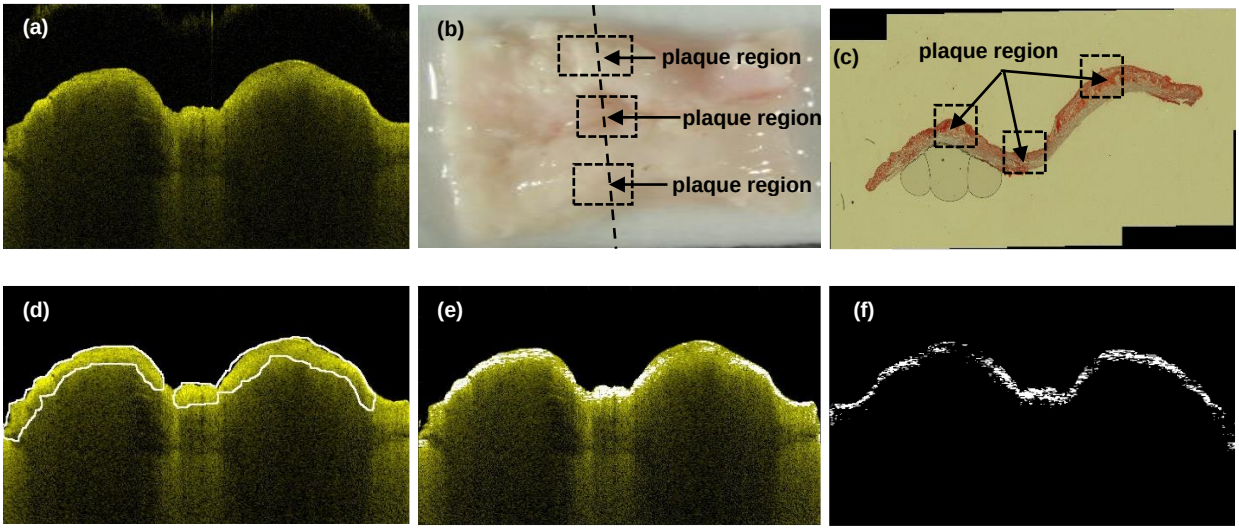



Fig. 4.8. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 342 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

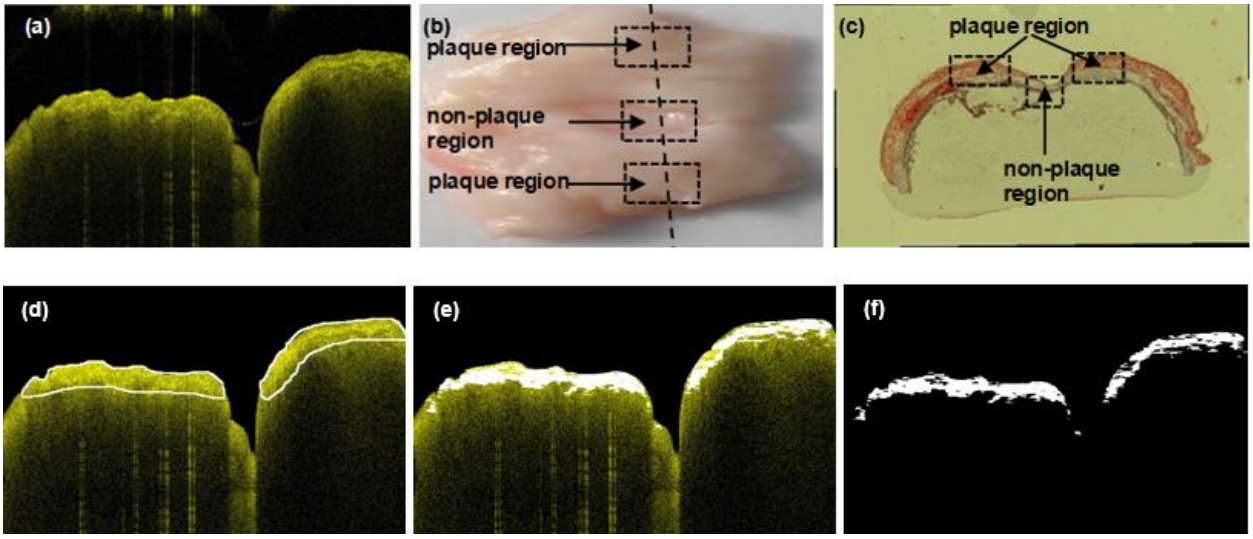



Fig. 4.9. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 365 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

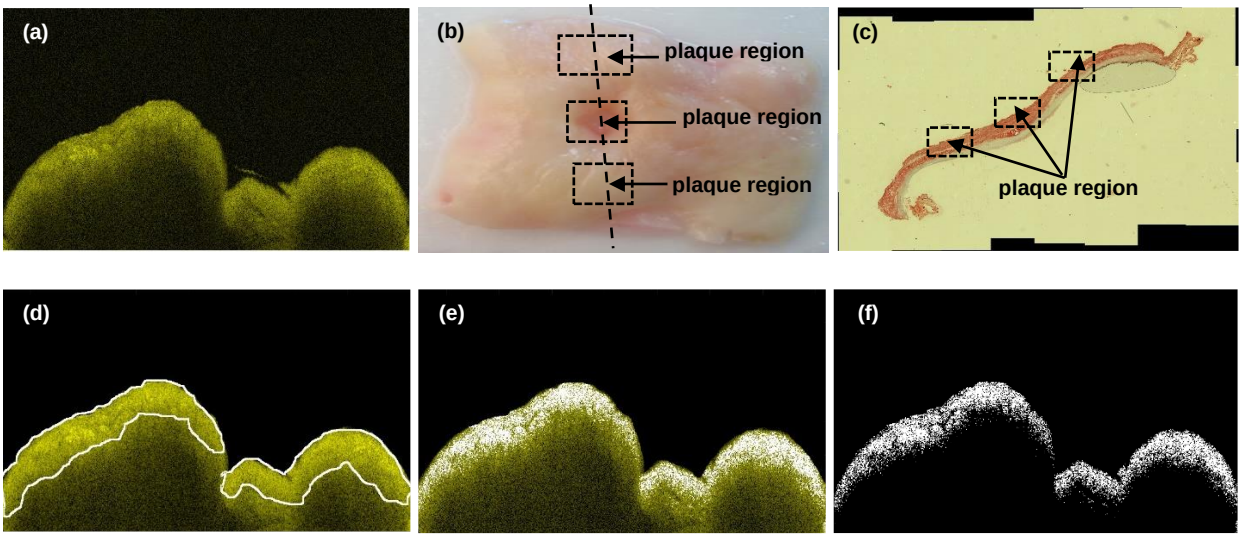



Fig. 4.10. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 456 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

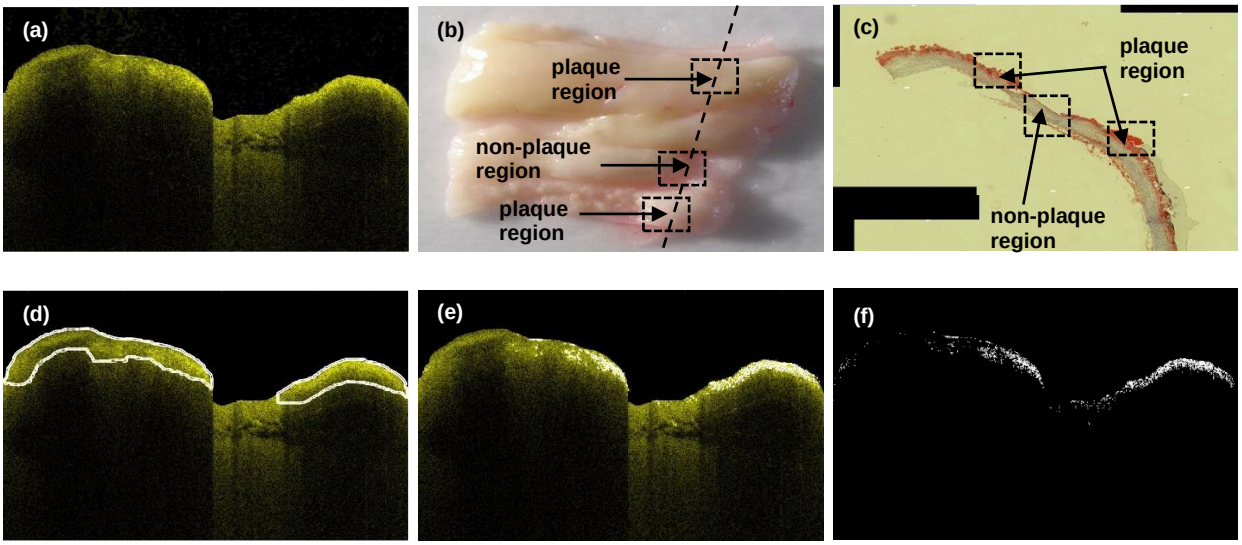



Fig. 4.11. (a) Raw **OCT** image of vascular tissue section 5-1-2 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

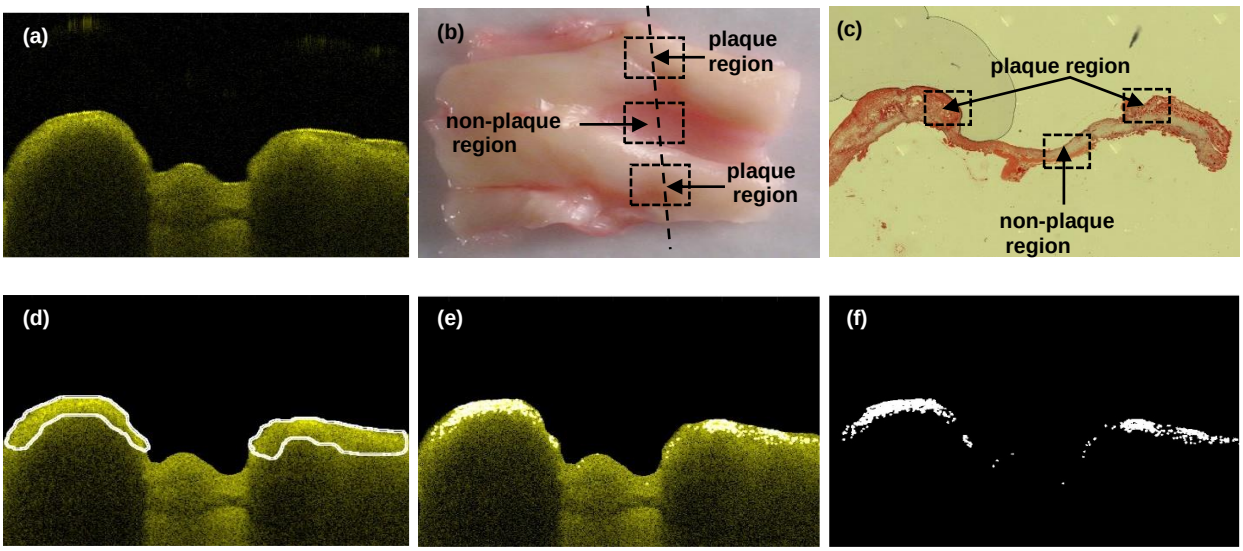



Fig. 4.12. (a) Raw **OCT** image of vascular tissue section 6-1-2 from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

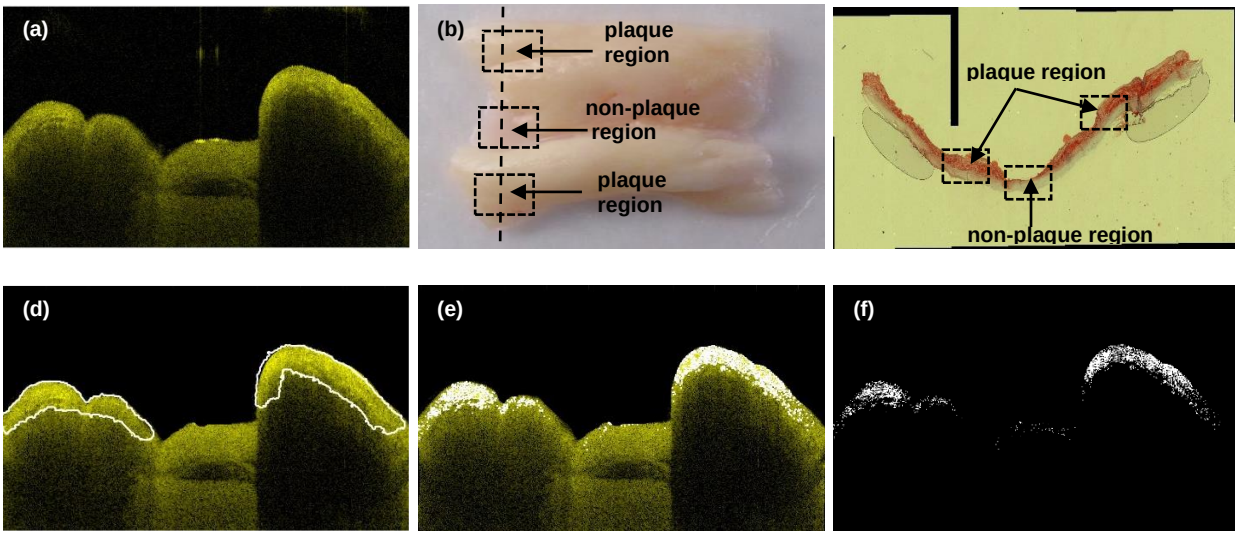



Fig. 4.13. (a) Raw **OCT** image of vascular tissue section 5-1-1 from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using **MRF-MAP** based segmentation; (f) extracted plaque region.

### 4.4.2 Quantitative evaluation of MRF-MAP based segmentation results

We also evaluated the performance of our **MRF-MAP** algorithm quantitatively by forming a ground truth based on the consensus of four surgeons, as shown in Table 4.1.

Table 4.1. Quantitative evaluation of **MRF-MAP**-based plaque segmentation algorithm.

| **Image #** | **Image ID** | **Sample rabbit age (days)** | **Sensitivity (%)** | **Specificity (%)** | **Accuracy (%)** | **Precision (%)** | **MCC (%)** | **Dice Coefficient (%)** |
|---|---|---|---|---|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 58.597 | 98.550 | 95.882 | 74.292 | 63.865 | 65.518 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 41.566 | 99.039 | 96.624 | 65.497 | 50.567 | 50.857 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 12.308 | 99.979 | 95.020 | 97.288 | 33.675 | 21.852 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 46.086 | 97.750 | 94.392 | 58.749 | 49.127 | 51.653 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 31.446 | 99.192 | 94.848 | 72.718 | 45.710 | 43.905 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 35.801 | 99.773 | 95.770 | 91.333 | 55.701 | 51.438 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 48.230 | 99.674 | 92.564 | 95.951 | 65.023 | 64.193 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 26.292 | 99.843 | 94.590 | 92.774 | 47.793 | 40.972 |
| 9 | 2010-02-02-086-02-612 | 577 | 29.861 | 99.359 | 94.874 | 76.284 | 45.753 | 42.921 |
| 10 | 2010-02-02-086-511 | 577 | 37.994 | 99.900 | 94.680 | 97.218 | 58.984 | 54.636 |

## 4.5 Discussions

### 4.5.1 Discussion on quantitative results

Fig. 4.14 shows the average values of all the evaluation metrics over all the 10 images. When compared to the clustering method, we notice that **MRF-MAP** performed poorly on some of the error evaluation metrics. This is attributed to the fact that we used a deterministic algorithm, **ICM,** to compute the posterior probability and, the major disadvantage of the deterministic **ICM** method is that it uses a greedy strategy in the iterative local minimization which makes it easy to get stuck in a local optimum. Also, the **MRF-MAP** method is overly sensitive to the model parameters; therefore, it performed poorly.

We now discuss results obtained for individual error metrics.

For accuracy and specificity, **MRF-MAP** based method performed slightly better with the result of 95.1% accuracy when compared to the clustering method of 94.9% accuracy and specificity of 99.3% for the **MRF-MAP** method and 98.6% for the clustering method. Both accuracy and specificity are sensitive to true negative counts, i.e., plaque region pixels segmented as non-plaque. The **MRF-MAP**-based method detected most of the non-plaque region pixels and missed plaque regions. Since non-plaque region pixels are far greater in number than plaque region pixels, accuracy and specificity metrics may appear to inflate the performance for the **MRF-MAP** method. For sensitivity, **MCC**, Dice similarity coefficient, the segmentation result of the clustering method performed better than the **MRF-MAP** method. For precision, **MRF-MAP** with 82.2% performed better than the clustering method of 76.1%. The main advantage of **MRF-MAP** based method is that prior information could be imposed on the pixel labels as a random field model. In real-world images, regions are often homogenous; neighboring pixels usually have similar properties, and the **MRF** model captures such contextual constraints through the clique potential (i.e., by adding a continuous windowing effect). To make sure we obtain a smooth segmentation,

we chose the smallest possible window of a first-order neighborhood with a distance of 1 pixel. And therefore, similar to the clustering method, the **MRF-MAP** method also has an inherent windowing effect, which would result in a low-resolution segmentation. Another disadvantage of the **MRF-MAP** method is its sensitivity to the model parameters. Also, by using a greedy optimization method, **ICM, MRF-MAP** did not give very satisfactory results.

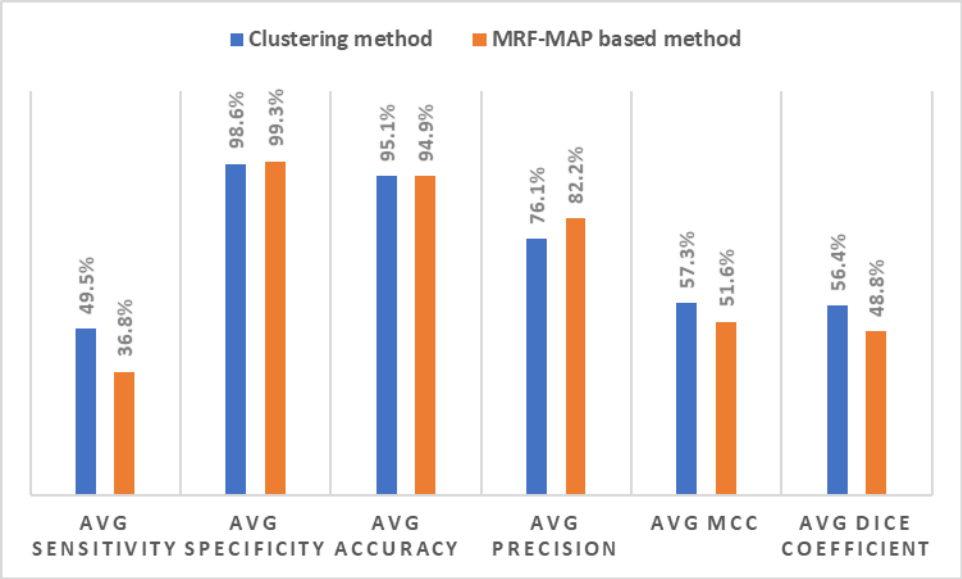



Fig. 4.14. Comparison of average values of quantitative results of all the metrics of **MRF-MAP** and clustering algorithm results

The performance of **MRF-MAP** could be improved further by using by applying stochastic algorithms, such as the *Gibbs* sampler or Simulated Annealing, instead of **ICM** to estimate the class labels. However, such an improvement will necessarily come at the cost of a further significant increase in computation complexity.

### 4.5.2 Discussion on computation time

With the current implementation using **ICM,** we already see an extremely high computation cost of the **MRF-MAP** method, as shown in table 4.2. On average, the **MRF-MAP**-based segmentation

method is about 9 times slower than the curvelet method and about 13 times slower than the clustering method. All experiments were performed on a machine with an Intel Core (i5) CPU, 8 GB RAM, and a Windows 7 OS using an interpreted programming language, MATLAB. All execution times exclude the time required to read the images and calculate the metrics.

Table 4.2. Computation time required to implement the **MRF-MAP** method.

| **Image #** | **Image ID** | **Age (days)** | **Computation Time(s)** |
|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 300 | 121.18 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 75.75 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 118.76 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 114.71 |
| 5 | 2009-06-23-W21-087-04-612 | 34**2** | 85.68 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 90.06 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 122.33 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 91.15 |
| 9 | 2010-02-02-086-02-612 | 577 | 74.59 |
| 10 | 2010-02-02-086-511 | 577 | 78.83 |

## 4.5 Conclusions

We implemented a fully automatic algorithm that employs a stochastic approach to conduct vascular plaque segmentation using **OCT** images. Our method is based on the **MRF-MAP** framework. We assumed a *Gaussian* distribution for likelihood, and we modeled prior, which is a latent random field, as the *Potts* model. We used the **EM** algorithm to estimate our parameter values and the deterministic **ICM** algorithm to estimate class labels. We showed that our **MRF-MAP**-based model is capable of providing high-resolution segmentation results as opposed to coarse segmentation results, which were obtained using the clustering algorithm. However, our findings also showed that, compared to the clustering results, the **MRF-MAP**-based model did not improve quantitative results for all of the images, and it also requires a very high computation time.

# Chapter 5: HIGH-RESOLUTION CLASSIFICATION OF VASCULAR PLAQUE IN OPTICAL COHERENCE TOMOGRAPHY IMAGES USING SPARSITY-BASED SEGMENTATION IN THE CURVELET DOMAIN

## 5.1 Introduction

**OCT** generates cross-sectional images of tissue with spatial resolutions of approximately 10 $\mu$m, which is higher than other clinical intravascular imaging methods, such as intravascular ultrasound. Such high spatial resolution, in addition to the small size of **OCT**'s fiber-based imaging probes, make it a highly suitable option for intravascular imaging.

Despite these advantages, the ability to detect vascular plaque from **OCT** images by visual inspection is often limited due to a lack of contrast. To overcome this problem, we developed an image segmentation algorithm based on *Haralick* textural features. However, the disadvantage of this image segmentation method is that it uses local image windows to calculate textural features. While the use of large local image windows makes it possible to compute these textural features accurately, it often also results in non-smooth low-resolution segmentation. On the other hand, using small local image windows results in image segmentation with an acceptable resolution but reduced accuracy. Therefore, there is always an inherent trade-off between resolution and accuracy when using texture-based image segmentation methods.

We also developed an **MRF**-based model that is capable of providing high-resolution segmentation results, as opposed to the coarse segmentation results obtained with the clustering algorithm. However, one of the main disadvantages of the **MRF-MAP** technique is that it requires

numerically approximating the **MAP** estimate, which results in high computational costs. To overcome this problem, we develop a method that utilizes sparsity-based segmentation in the curvelet domain to enable the high-resolution classification of vascular plaque in **OCT** images. Our classification results show an excellent match with histology and photographic images of vascular plaque samples, in addition to having significantly higher spatial resolution than both MRF and clustering methods.

## 5.2 Sparsity-based image segmentation

Our objective is to classify vascular plaque from **OCT** images by identifying different image regions represented by directional lines and edges. We note that such image discontinuities correspond to the high-magnitude non-zero coefficients in an appropriate image representation domain (e.g., wavelet domain). Therefore, image segmentation can be viewed as the process of obtaining a sparse representation of an image, typically in the wavelet domain [130].

Instead of using a sparse wavelet representation to classify vascular plaque regions from **OCT** images in this work, we segmented the image using the curvelet transform [131] due to its superior directional sensitivity. Therefore, the curvelet transform offers better performance than the wavelet transform with respect to representing curves and segmenting curvilinear regions. The curvelet transform is an extension of the traditional wavelet transform that overcomes its poor directionality by representing two-dimensional signals (e.g., images) using wavelets with different translations, scales, and angular orientations.

### 5.2.1 Curvelet transform

The first step in computing the curvelet transform of an image is to apply the 2D fast Fourier transform (**FFT**). The Fourier plane is constructed so that the center frequency (i.e., the D.C. value)

is located at the center of the plane. Next, the **FFT** plane is divided into tiles. The coarsest level (scale 0, innermost level) is not directional, and the curvelet becomes more sensitive to curved edges as the resolution level increases. Thus, curvelets are capable of effectively capturing curved edges at finer scales, making it possible to approximate curved singularities with few curvelet coefficients accurately. At a higher scale (scale 4), the curvelet waveform is very fine, with a needle-like structure.

When we combine the frequency responses of curvelets at all of the different scales and orientations, we obtain a frequency tiling wherein periodic extension is used in the outer scales. In order to achieve high computational efficiency, curvelet transforms are usually implemented in the frequency domain. In other words, both the curvelet and the image are transformed and then multiplied in the frequency domain. Next, the inverse Fourier transform of the product is used to obtain the curvelet coefficients; however, the product forms a trapezoidal wedge in the spectral domain, which is not suitable for inverse Fourier transform (Fig. 5.1).

Fortunately, Candes *et al*. (2006) have developed a wedge-wrapping operation that enables the inverse Fourier transform of these non-rectangular wedges to be acquired. In this operation, every wedge is localized inside a parallelogram with sides, $2^{j}$ and $2^{j/2}$, to support the wedge. The wrapping operation is performed via the periodic tiling of the spectrum inside the wedge data, with the curvelet coefficients being obtained by taking the inverted Fourier transform of this wrapped wedge data.

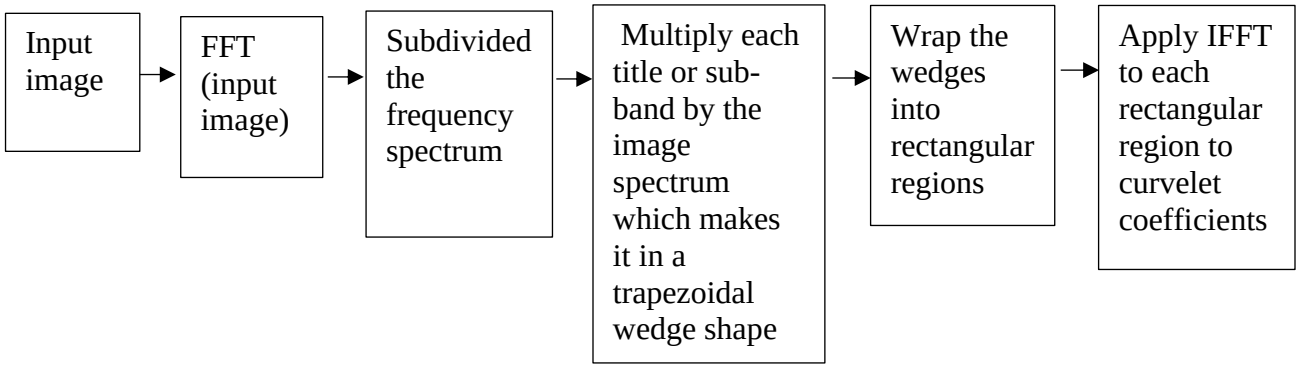



Fig. 5.1. Illustration of the curvelet algorithm.

## 5.3 Implementation and results of our vascular plaque classification method

We used Meyer wavelets with four scales and eight angular orientations to decompose each of our **OCT** images (3700 × 650 pixels) into cells of curvelet coefficients ranging from low-frequency coefficients to high-frequency coefficients. The Discrete Curvelet Transform (**DCT**) is a decomposition of the **OCT** image, *I*, into curvelet coefficients, $C_{ldk_1,k_2}$

$$C_{idk_1,k_2} = \sum_{0\le n_2<M}^{0\le n_1<N} I[n_1, n_2]\, \varphi_{idk_1,k_2}(n_1, n_2) \qquad (5.1)$$

where $(n_1, n_2)$ are pixel indices, $\varphi_{idk}$ are discrete curvelets with scales *i*, directions *d*, and spatial shifts $k_1$, $k_2$. At scale $i \geq 2$, there are *N* orientations,

$$N = N_\theta(2^{\left\lceil \frac{i-2}{2} \right\rceil}) \qquad (5.2)$$

where $N_\theta = 2^a$ and $a \geq 3$. Thus, the minimum number of orientations or angles must be 8.

The discrete curvelet transform (**DCT**) decomposes the image into dyadic rectangular coronae in different scales and orientations (angles). These rectangular coronae are divided into wedges, which double every second level. To extract the curvelet coefficients, each **OCT** image is divided into different subbands of different frequencies. In this work, we used i=4 scales and set the second coarsest curvelet transform to 16 angles. The next two higher frequency levels contain 32 angles in each scale, for a total of 80 angular wedges. Table 5.1 shows the scale and orientation of curvelet subbands,

Table 5.1. Scale and orientation of subbands.

| **Scale** | **Number of orientations** |
|---|---|
| 1 | 1 |
| 2 | 8 |
| 3 | 16 |
| 4 | 32 |

On the finest level (i.e., i = 4), we used wavelets to exclude the noise, which helped to decrease the computer's memory requirements and total execution time. At the coarsest level (i.e., when *i* = 1), the curvelets behave like Meyer wavelets and are nondirectional. To remove the background information, we discarded the coarsest level, which makes our method insensitive to variations in background intensity. All the parameters used in the curvelet domain are shown in Table 5.2.

To enhance computational efficiency, we used the simpler and faster discrete fast curvelet transform followed by a simple hard thresholding step to retain only the large magnitude

coefficients [132] from each sub-band. Hard thresholding sets any coefficients less than or equal to our threshold to zero.

Table 5.2. Parameters used in image transformation in the curvelet domain.

| | |
|---|---|
| Curvelab version | CurveLab 2.1.3 |
| Open-source site | www.curvelab.org |
| Matlab file | Fdct_wrapping.org |
| Input image | OCT image |
| is_real | 0 (complex curvelet coefficients) |
| finest | 2 (wavelet at scale 0, finest scale) |
| Nbscales | 4 (number of levels of decomposition) |
| Nbscales_coarse | 8 (number of orientations at the second coarsest level) |

### 5.3.1 Qualitative evaluation of our results by comparing with ground truth

To demonstrate the validity of our plaque classification method, we compared our results—which were obtained using **OCT** images of vascular tissues from **WHHLMI** rabbits—with actual photographic images, oil red histology images, and ground truth images.

Fig. 5.2 shows a raw **OCT** image of vascular tissue from a 309 days old **WHHLMI** rabbit (panel a), the ground truth, which was formed based on the consensus of all four surgeons (panel d), and the sparsity-based segmented images (panel e). A comparison of the photographic (panel b) and oil red histology images (panel c) with the plaque region extracted using our sparsity-based

segmentation method (panel f) revealed an excellent match between the plaque regions in all of these images. We also note that our method enabled significantly improved resolution in plaque classification compared to the textural-feature-based classification method. This improvement in resolution and reduction in computational complexity with respect to plaque classification is the main contribution of this work. To validate our results further, we show other successful plaque detection results for 316, 330, 342, 365, 456, and 577 days old **WHHLMI** rabbits in Fig. 5.3, Fig 5.4, Fig. 5.5, Fig. 5.6, Fig. 5.7, Fig. 5.8, Fig. 5.9, Fig. 5.10, and Fig. 5.11.

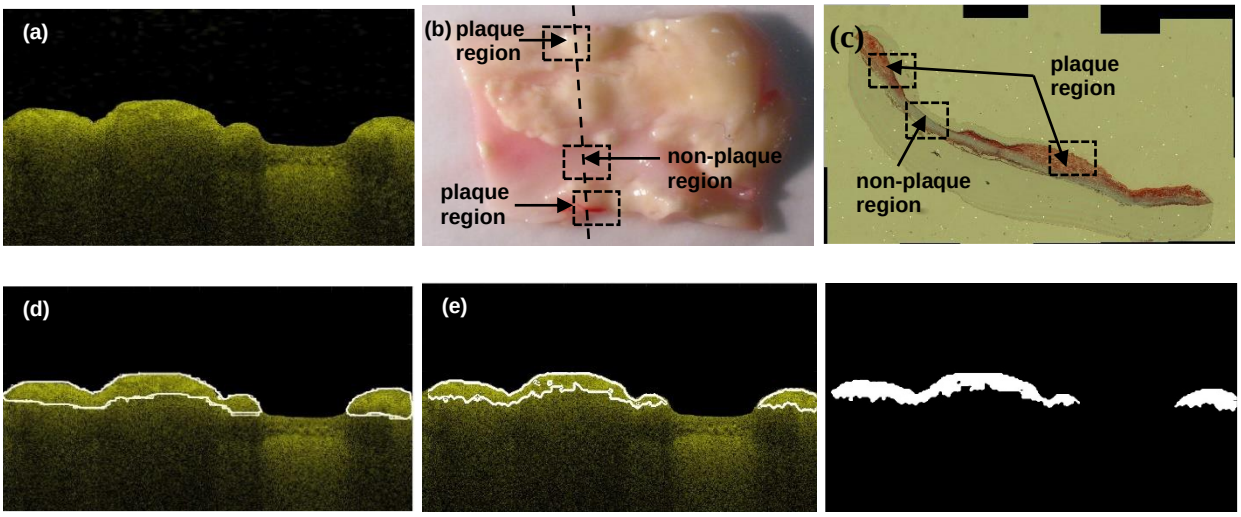



Fig. 5.2. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 309 day-old **WHHLMI** rabbit; (b) photographic image of vascular; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

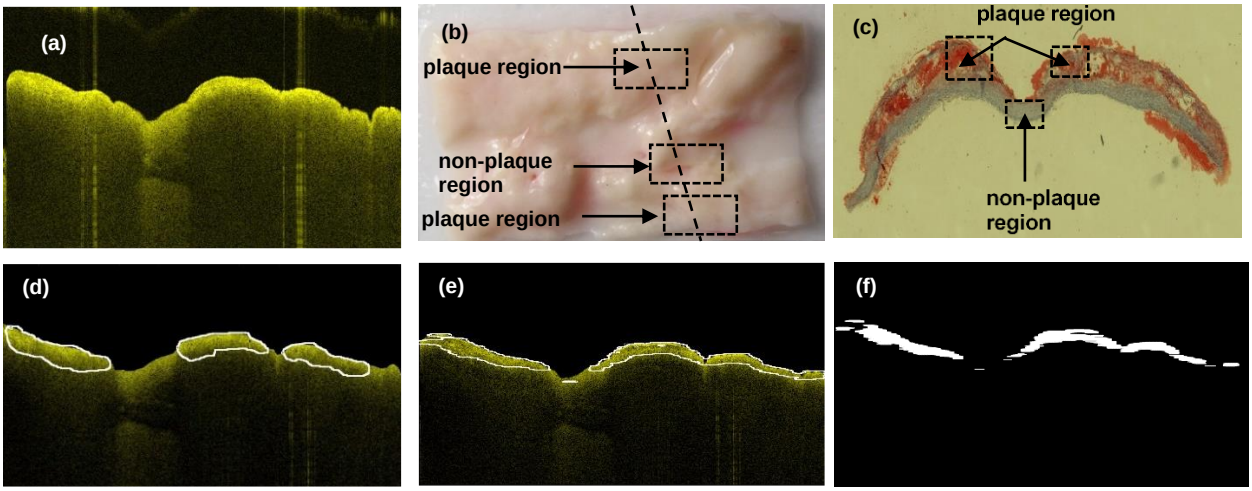



Fig. 5.3. (a) Raw **OCT** image of vascular tissue section of 5-1-2 taken from a 316 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

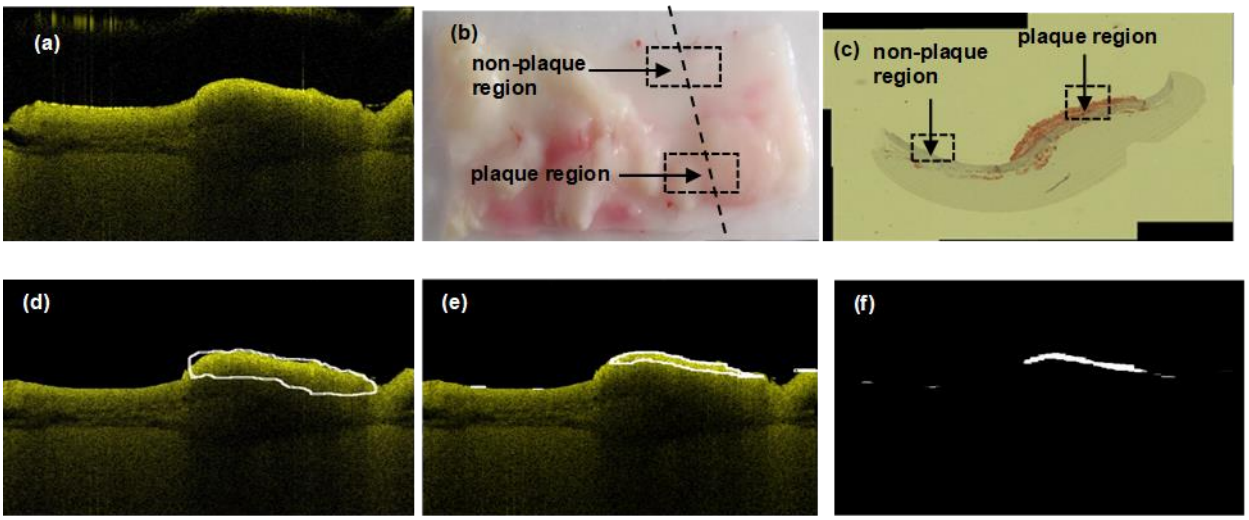



Fig. 5.4. (a) Raw **OCT** image of vascular tissue section of 6-1-2 taken from a 316 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

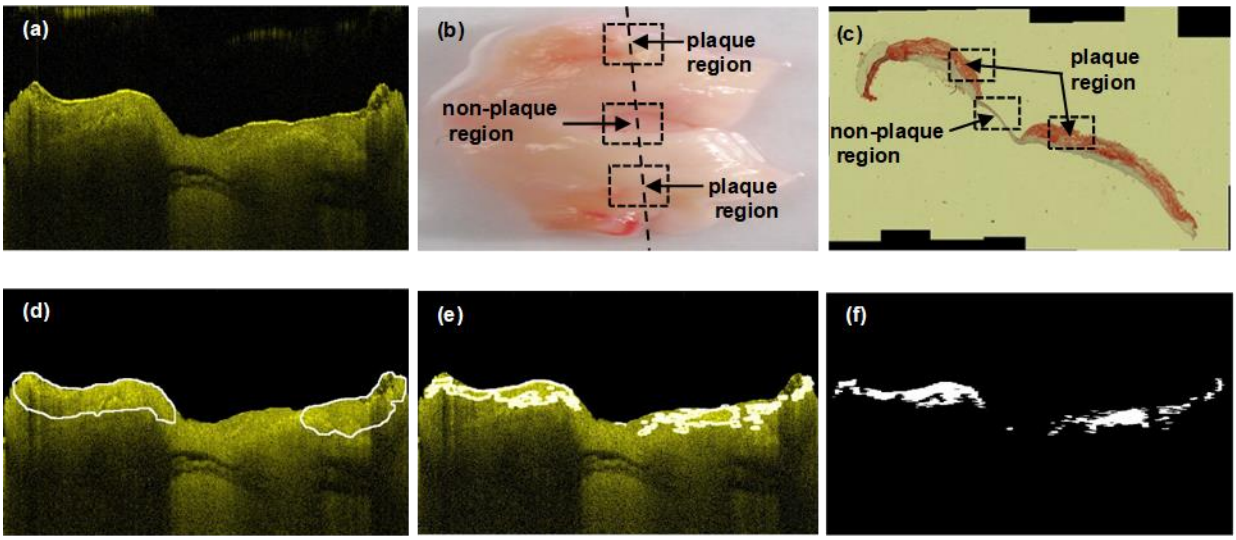



Fig. 5.5. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 330 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

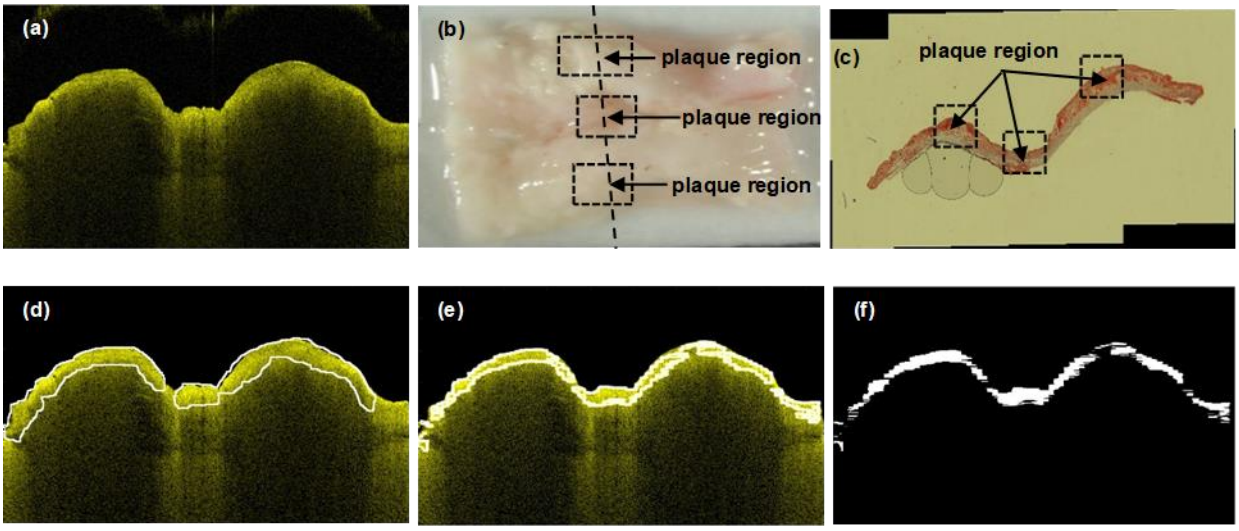



Fig. 5.6. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 342 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

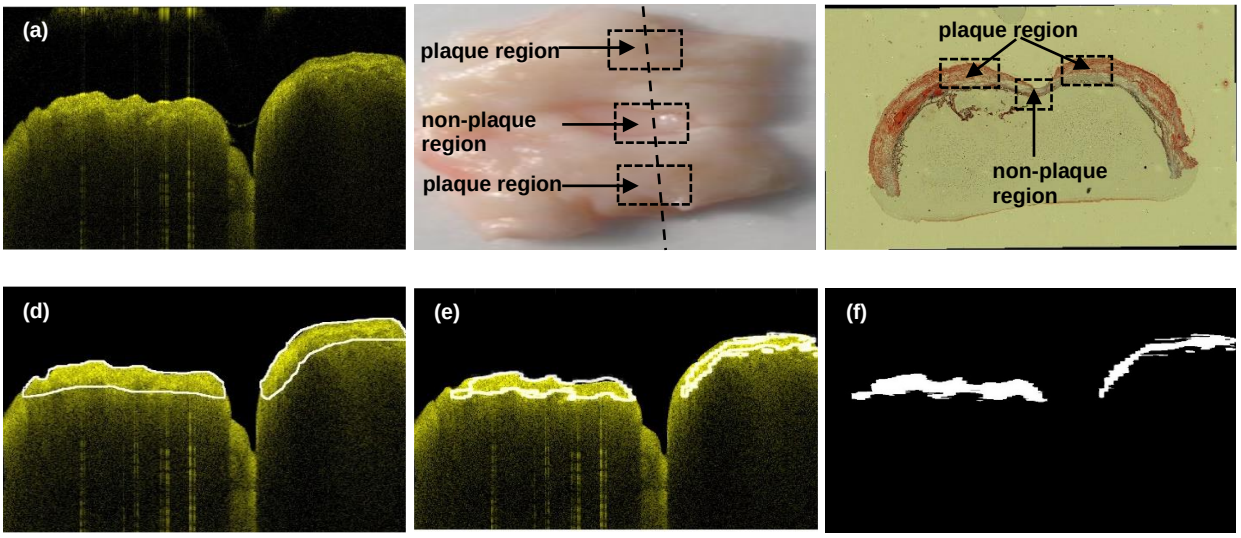



Fig. 5.7. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 365 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

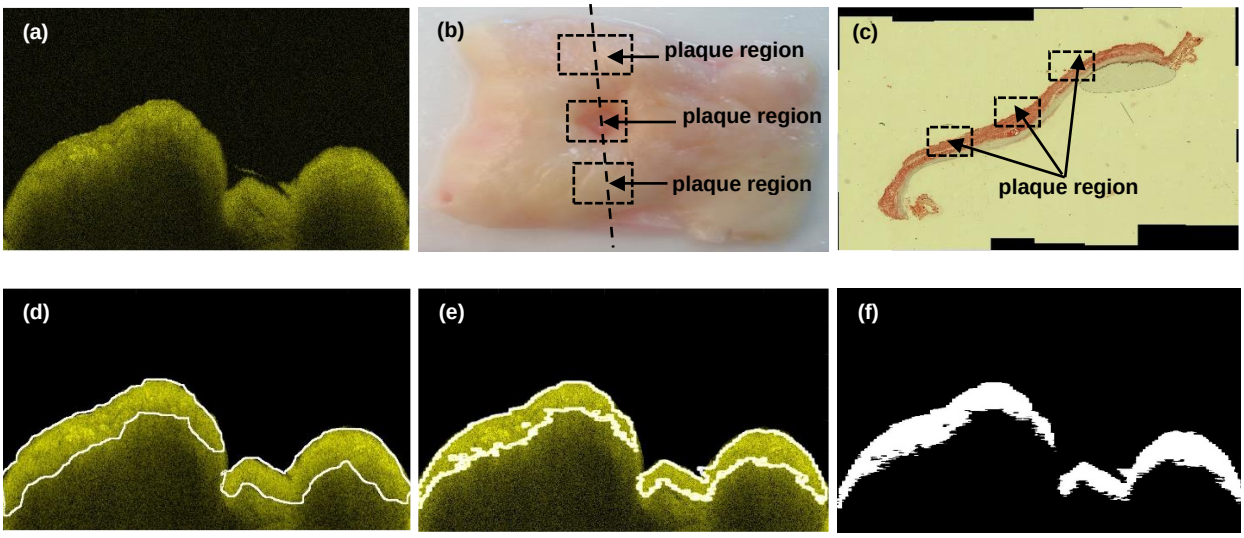



Fig. 5.8. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 456 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

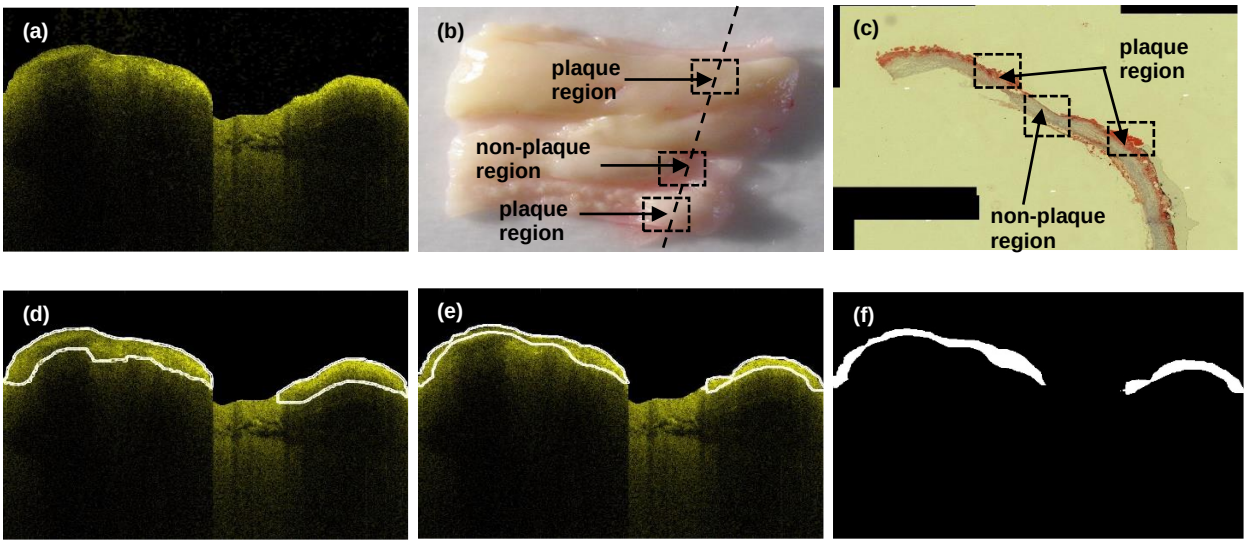



Fig. 5.9. (a) Raw **OCT** image of vascular tissue section 5-1-2 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

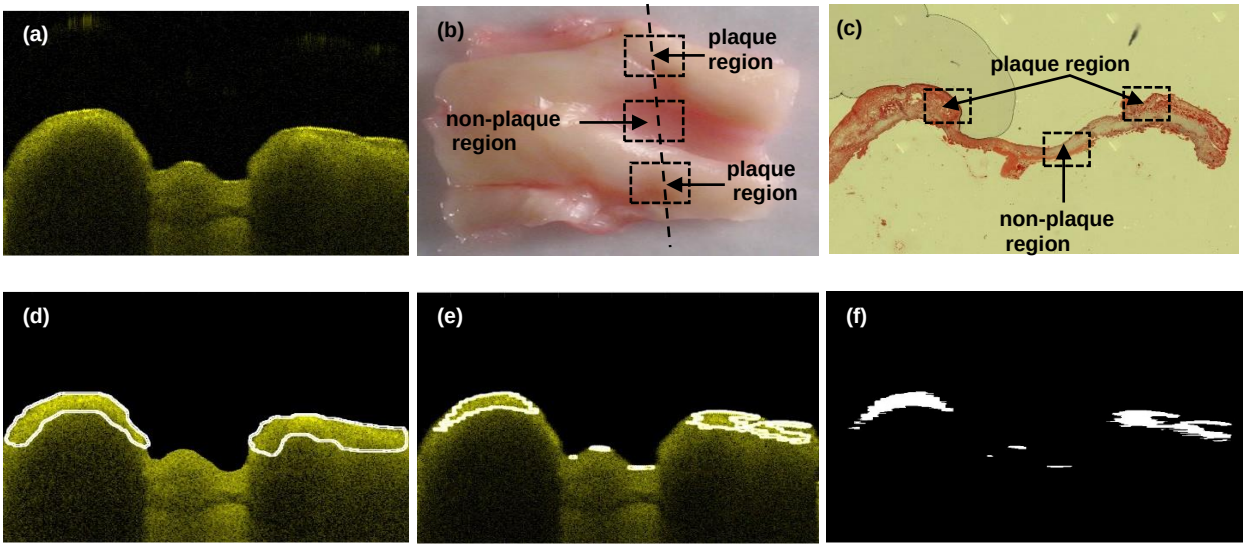



Fig. 5.10. (a) Raw **OCT** image of vascular tissue section 6-1-2 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

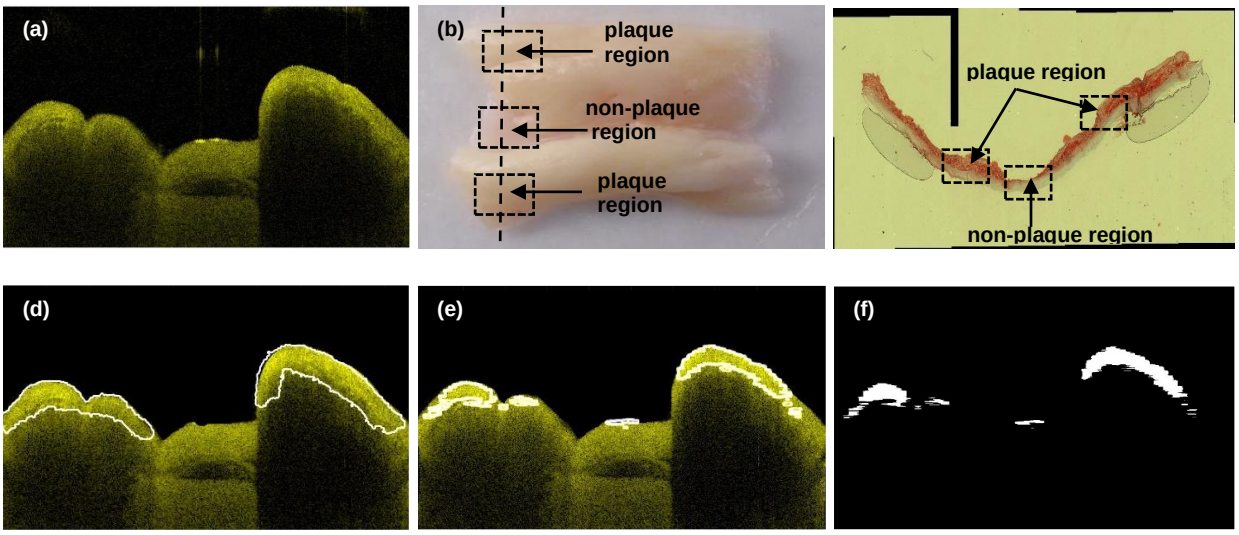



Fig. 5.11. (a) Raw **OCT** image of vascular tissue section 5-1-1 taken from a 577 day-old **WHHLMI** rabbit; (b) photographic image of vascular tissue ; (c) corresponding oil red-stained histology image; (d) ground truth; (e) plaque segmentation using curvelet-based segmentation; (f) extracted plaque region.

### 5.3.2 Quantitative evaluation of curvelet based segmentation results

We also evaluated the performance of our segmentation results quantitatively by computing the evaluation metric described in chapter 2. Table 5.3 shows the quantitative results of segmentation using our curvelet-based algorithm.

Table 5.3. Quantitative evaluation of curvelet-based plaque segmentation algorithm.

| **Image #** | **Image ID** | **Sample rabbit age (days)** | **Sensitivity (%)** | **Specificity (%)** | **Accuracy (%)** | **Precision (%)** | **MCC (%)** | **Dice Coefficient (%)** |
|---|---|---|---|---|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 70.493 | 99.595 | 97.653 | 92.573 | 79.642 | 80.038 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 90.664 | 99.027 | 98.676 | 80.353 | 84.676 | 85.198 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 26.770 | 99.999 | 95.856 | 99.927 | 50.620 | 42.228 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 78.789 | 97.738 | 96.507 | 70.777 | 72.819 | 74.568 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 84.782 | 98.361 | 97.491 | 77.997 | 79.986 | 81.248 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 84.762 | 99.511 | 98.589 | 92.052 | 87.591 | 88.256 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 88.740 | 99.449 | 97.969 | 96.275 | 91.286 | 92.354 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 44.290 | 99.653 | 95.699 | 90.750 | 61.698 | 59.528 |
| 9 | 2010-02-02-086-02-612 | 577 | 73.862 | 99.340 | 97.696 | 88.542 | 79.693 | 80.538 |
| 10 | 2010-02-02-086-511 | 577 | 59.373 | 99.847 | 96.434 | 97.274 | 74.476 | 73.739 |

## 5.4 Discussions

### 5.4.1 Discussion on quantitative evaluations and computation time

Fig. 5.12 shows the average values of different evaluation metrics over all the 10 images. Our curvelet-based method generated notably higher results in comparison to the **MRF-MAP** and clustering method.

As mentioned in chapter 3, our clustering-based method suffered from a poor resolution because of its windowing effect, which resulted in high false positives. And therefore, our clustering method performed worst on precision and specificity with an average score of 76.1% and 98.6%.

**MRF-MAP** resulted in high false negatives and performed worst on sensitivity, **MCC,** and Dice similarity coefficient metric with the average score of 36.8%, 57.3%, and 48.8%, respectively.

On all the evaluation metrics, our curvelet-based segmentation performed the best. While we are at 88.7% for precision, 76.2% for **MCC,** and 75.8% for Dice similarity coefficient, it is not perfect. This could because the manual segmentation which formed the ground truth is not perfect. There was some variance in the ground truth, as the manual segmentation formed by different cardiologists had some differences. We, therefore, feel confident in the ability of our curvelet-based algorithm to detect plaque effectively.

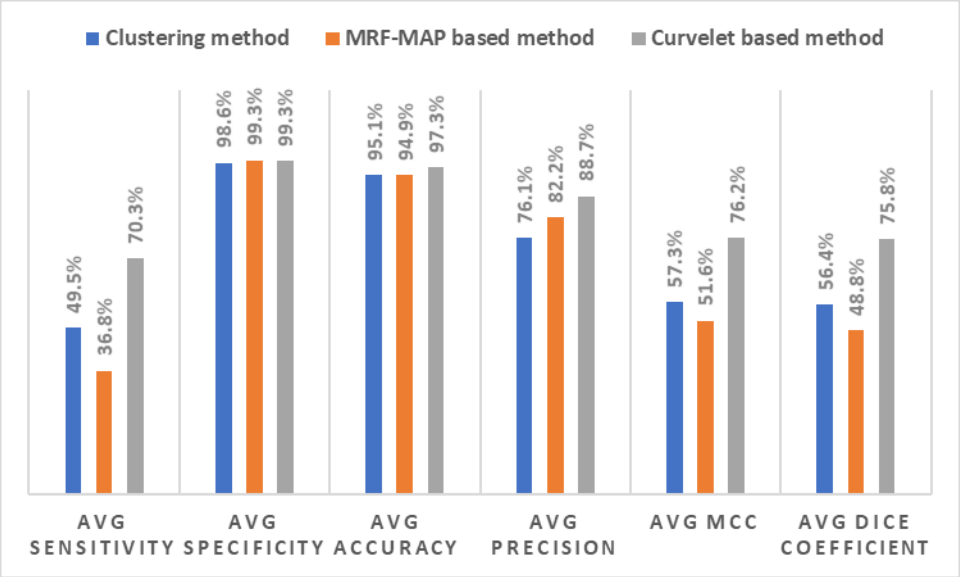



Fig. 5.12. Comparison of average values of quantitative results of all the error metrics of Curvelet based, **MRF-MAP** based, and clustering methods.

The computation time that was required to implement our curvelet method on all of the image samples is shown in Table 5.4. Our curvelet-based method shows a considerable improvement in both segmentation quality and computation efficiency when compared to clustering-based and **MRF**-based algorithms. All experiments were performed on a machine with an Intel Core (i5) CPU, 8 GB RAM, and a Windows 7 OS using an interpreted programming language, MATLAB. All execution times exclude the time required to read the images and calculate the metrics.

Table 5.4. Computation time required for the implementation of our curvelet method.

| **Image #** | **Image ID** | **Sample rabbit Age (days)** | **Computation Time(s)** |
|---|---|---|---|
| 1 | 2009_12_15_21_092_01_511 | 309 | 8.089 |
| 2 | 2009_05_27_21_087_01_5L2 | 316 | 6.662 |
| 3 | 2009_05_27_21_087_01_6L2 | 316 | 7.961 |
| 4 | 2010_03_03_20_084_02_612 | 330 | 9.362 |
| 5 | 2009-06-23-W21-087-04-612 | 342 | 6.591 |
| 6 | 2009_08_12_W21_085_01_611 | 365 | 7.742 |
| 7 | 2009-12-08-20-08-04-01-611 | 456 | 9.872 |
| 8 | 2010_02_02_20_086_02_512 | 577 | 7.646 |
| 9 | 2010-02-02-086-02-612 | 577 | 6.483 |
| 10 | 2010-02-02-086-511 | 577 | 6.932 |

## 5.5 Conclusions

This chapter presented a method that uses sparsity-based segmentation in the curvelet domain to achieve a high-resolution detection of vascular plaque region in **OCT** images. Our results visually showed an excellent match with the histology, photographic and ground truth image. To confirm the validity of our method, we also evaluated our plaque detection results quantitatively, attaining average sensitivity, specificity, accuracy, precision, **MCC**, Dice similarity coefficient values of 70.253%, 99.252%, 97.257%, 88.652%, 76.249%, 75.770% respectively.

Thus, our automated curvelet-based method of identifying vascular plaque from **OCT** images produced very promising results and could be adopted as practical clinical tools for automated real-time identification of vascular plaque in **OCT** images.

# Chapter 6: CONCLUSIONS AND FUTURE WORK

## 6.1 Summary

In this dissertation, we have shown three methods that are capable of differentiating tissue types in **OCT** images without relying on visual structures. Although **OCT** is an important modality for intravascular imaging, it is usually quite difficult to accurately identify atherosclerotic plaque from **OCT** images through visual inspection. We have described three novel fully automated algorithms that are capable of identifying vascular plaque in **OCT** images without relying on visual inspection: (i) a statistical-based method, (ii) a model-based method, and (iii) a sparsity-based model that uses curvelets to detect vascular plaque regions in **OCT** images automatically. While our first clustering-based algorithm, which was based on reduced feature-sets, provided successful visual results, however, because of its underlying windowing effect, its segmentation had a low resolution.

Our second algorithm, which was based on **MRF**, was able to improve the resolution of plaque identification regions in some images, though it was not very consistent. However, the main issue with the **MRF**-based method was that it required a high amount of computation time.

Finally, our third algorithm, which used a curvelet domain sparsity-based image segmentation method, provided the most satisfactory results with respect to both computation speed and resolution.

We qualitatively compared our algorithms' plaque identification results to photographic images and histology slides, as well as to the ground truth, which was based on the consensus assessments of the images by four interventional cardiologists.

We also quantitatively evaluated our results for all the error evaluation metrics presented in chapter 2. However, the three most relevant evaluation metrics for our plaque detection problem are

precision, **MCC,** and Dice similarity coefficient as the region of interest, i.e., the plaque-only region is significantly small when compared to the non-plaque region. An unbalanced plaque and non-plaque region may inflate the values of accuracy and specificity, which could be a misleading result. Where in the case of sensitivity, it may appear to be very low.

The average precision, **MCC**, and Dice similarity coefficient values obtained for our three algorithms (curvelet, **MRF**, and clustering method) were (88.7%,76.2%, and 75.8%), (82.2%, 51.6%, and 48.8%) and (76.1%, 57.1%, and 56.4%). Comparisons of our three algorithms with respect to each of the above evaluation metrics are shown in Fig. 6.1, Fig. 6.2, Fig. 6.3

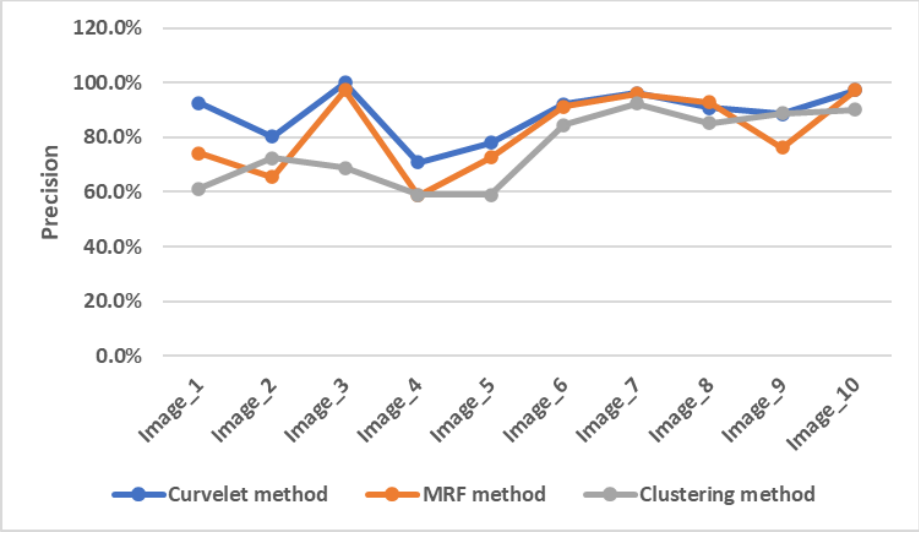



Fig. 6.1. Comparison of the precision metric for all three methods

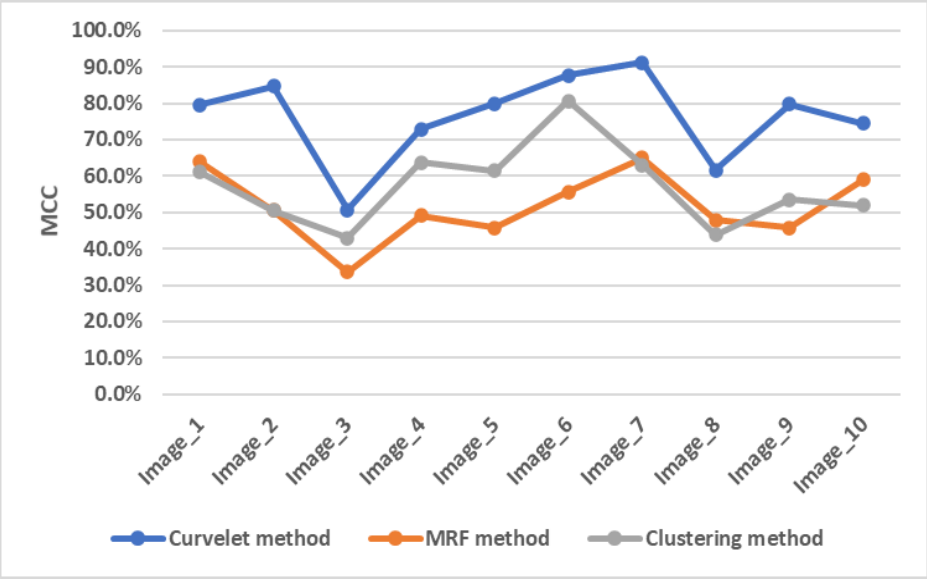



Fig. 6.2. Comparison of the **MCC** metric for all three methods

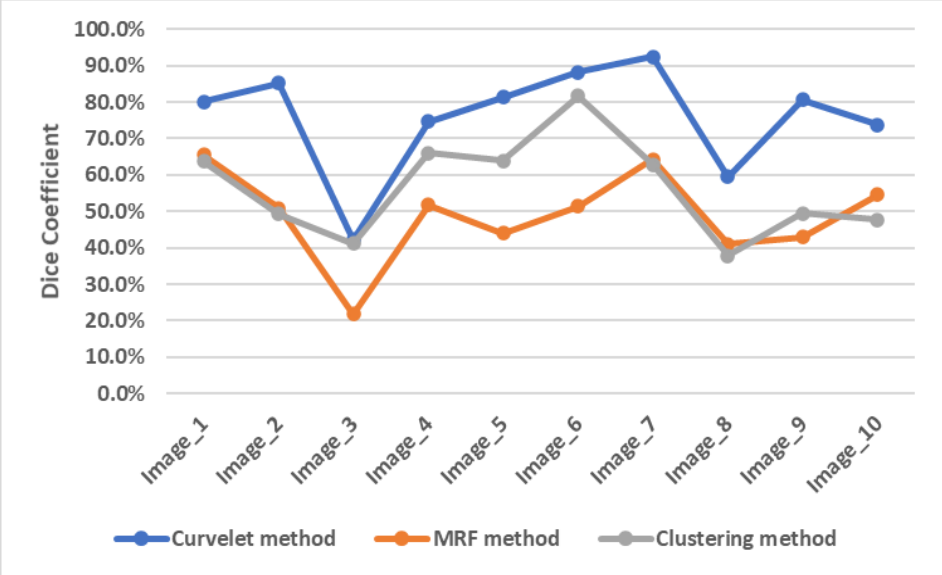



Fig. 6.3. Comparison of the Dice coefficient metric for all three methods

We also compared the computation time required for each of the three methods, with the curvelet method providing the best performance. All experiments were performed on a machine with an Intel Core (i5) CPU, 8 GB RAM, and a Windows 7 OS using an interpreted programming language, MATLAB. All execution times exclude the time required to read the images and calculate the metrics. The average computation times obtained for the curvelet-based, clustering-based, and **MRF-MAP**-based methods were 7.735secs, 11.235secs, 97.304secs, respectively. The imaging parameters of a commercial extreme resolution **IVOCT** of St. Jude Medical Inc. for maximum frame rate is 100 fps, and nominal pullback speed is 20 mm/sec. The average frame rate of **MRF-MAP** based is 0.010 fps which might be too slow for real-time application. The clustering method is fast enough with its average frame rate of 0.089 fps, but it suffers from poor resolution. The average rate for our curvelet-based method is 0.128 fps which is suitable for the commercial system and real-time applications.

A comparison of the computation time required for the three algorithms is provided in Fig. 6.4.

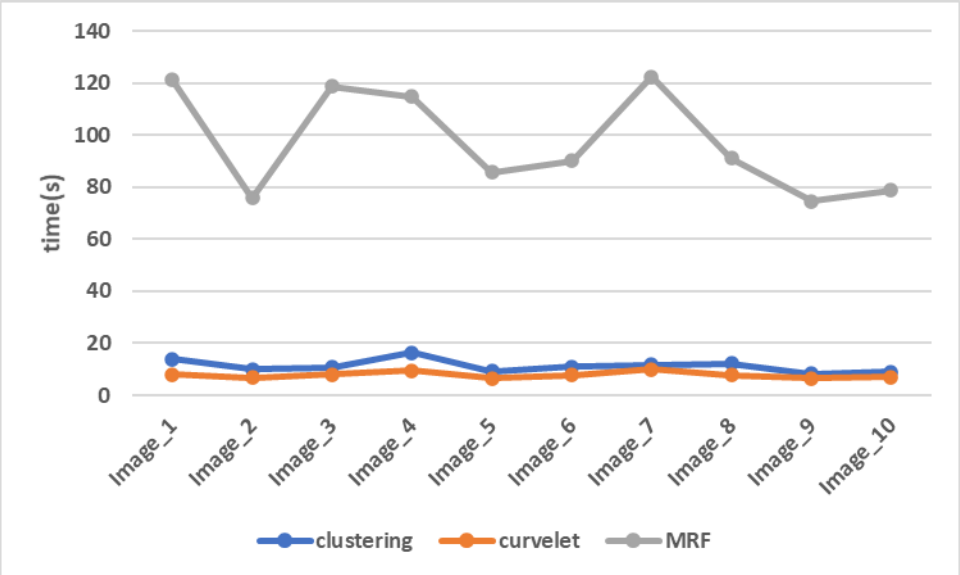



Fig. 6.4. Comparison of computation time for all three methods.

Our results show that the novel algorithms presented herein could be adapted as clinical tools for the automated real-time identification of vascular plaque in **OCT** images.

The main contributions made by this thesis can be summarized as follows:

- Our statistical approach is based on extracting second-order moments as textural features derived from **SGLDM** matrices and applying a clustering algorithm. As the use of this full set of 26 texture features is computationally expensive and could also lead to overfitting, it may not be practically fast for real-time clinical implementation. So, we reduced the computational complexity of our earlier method by using a reduced set of only 6 texture features, along with a *Fuzzy C-means* clustering algorithm. This additional step of feature selection is important to reduce computational complexity, as only features resulting in significant vascular plaque discrimination should be included in this reduced set, and other features should be ignored. Our goal then is to select those features that are rich in information related to our plaque identification problem. Sparse **PCA** expands the well-known **PCA** data analysis method by requiring its loading vectors to be sparse, i.e., to have as many zeros as possible. These sparse loading vectors are typically computed using the *least absolute shrinkage and selection operator* (***LASSO***) method [29]. To select only the information-rich features from the full set of 26 *Haralick* texture features, we computed the sparse **PCA** for all the vectors containing the full set of 26 *Haralick* image texture features and used the sparsity of its loading vectors for such selection.
- To overcome the limitation of redundancy and to reduce computation, we further improved our statistical method using a reduced set of 6 textural features and the *Fuzzy C-means* clustering algorithm. Our feature-reduction work comprised three main steps: implementing an **SGLDM** method to extract the full set of 26 textural features, identifying

a reduced set of 6 textural features from the full set, and applying the clustering algorithm on the reduced feature space. Our feature-reduction method combines sparse **PCA** with an advanced clustering technique (*Fuzzy C-means*) to detect vascular plaque in **OCT** images. We evaluated the accuracy of our method by comparing our plaque detection results with both the 26-feature set and the 6-feature set, with promising results being obtained for the reduced-feature set. We also validated our plaque detection results using histology, photographic and ground truth images. Our method, which combines the use of a reduced set of 6 textural features and **OCT** imaging, may offer an efficient tool for the real-time detection of vascular plaque in a clinical setting, thereby enhancing clinicians' ability to diagnose vascular disease earlier. Although our proposed method yielded promising results, it is hampered by some limitations, such as non-smooth and coarse plaque detection results. Therefore, to further improve our results and to overcome these limitations, we implemented an **MRF-MAP-**based approach.

- Image segmentation using the **MAP-MRF** approach has become extremely useful in the area of image segmentation. Since in the **MRF-MAP**-based approach**,** prior information could be imposed on pixel labels as a random field model, which makes it mathematically more feasible. In real-world images, regions are often homogenous; neighboring pixels usually have similar properties, and the **MRF** model captures such contextual constraints through the clique potential (i.e., by adding a continuous windowing effect). We assumed the pixel segmentation labels have a *Gibbs* probability distribution, characterized by energy functions that promote similarity of neighboring pixel labels. To make sure a smooth segmentation, we chose the smallest possible window of a first-order neighborhood with a distance of 1 pixel. We also assumed a conditional Gaussian distribution, with unknown

mean and variance, for each observed image pixel given its segmentation label. Then we used the **EM** algorithm to estimate both pixel segmentation labels and unknown Gaussian distribution parameters. Because the **EM** algorithm is sensitive to the initial values of unknowns, we used a *K*-means clustering algorithm to obtain good initial estimates for the pixel segmentation labels. This initialization is then followed by **EM** steps, which iteratively update pixel labels and the parameters using the **ICM** optimization method. Similar to the clustering method **MRF-MAP** method also has an inherent windowing effect, which resulted in low resolution. Another disadvantage of the **MRF-MAP** method is sensitivity to the model parameters, and by using a greedy optimization method, **ICM** we, therefore, did not achieve a satisfactory result for all the images.

The performance of **MRF-MAP** could be improved further by applying stochastic algorithms, such as the *Gibbs* sampler or Simulated Annealing, instead of **ICM** to estimate the class labels. However, such an improvement will necessarily come at the cost of a further significant increase in computation complexity.

- To overcome the low spatial resolution of results from our first method and the impractical long computation time required for our second method, we developed a third plaque identification method based on **OCT** image segmentation in the curvelet domain wherein the image is sparsely represented. To achieve significantly better directional sensitivities in representing curvilinear region boundaries, our method obtains a sparse representation of the image in an extension of the wavelet domain, i.e., the curvelet domain. Images are represented in the curvelet domain using wavelets that have different angular orientations (angles), in addition to their original spatial translations and scales.

A 2-D discrete curvelet transform could be easily implemented with Meyer wavelets having four scales, eight angular orientations, and by using the discrete Fourier transform. Each **OCT** image is divided into subbands of different frequencies before decomposing it into dyadic rectangular coronae that are further divided into wedges having different angular orientations and scales. Afterward, we applied a hard thresholding step to all subband coefficients, where any coefficient less than our chosen threshold is set to zero.

Our results show that although all three methods were successful in identifying plaque regions, our curvelet-based method provided the most satisfactory results both visually and quantitatively. We believe our curvelet-based algorithm, with its average processing times of ~7 secs and with high segmentation performance, and with high segmentation performance, could be adopted as an efficient pre-clinical tool for automated real-time identification of vascular plaque in **OCT** images. To the best of our knowledge, this research marks the first attempt to perform the automatic segmentation of vascular plaque regions from structureless **OCT** images using **MRF-MAP** and sparsity-based algorithms. Our novel algorithms could be adopted as practical clinical tools for automated real-time identification of vascular plaque in **OCT** images.

As future work, we anticipate the application of our intravascular methods to human *in-vivo* **OCT** 3-D images. Other unsupervised plaque segmentation methods based on sparse signal processing and dictionary learning could be promising. In addition, supervised plaque segmentation methods based on Deep learning architectures and convolutional neural networks could be promising as well.

## REFERENCES

[1] Bentzon, J. F., Otsuka, F., Virmani, R., & Falk, E. (2014). Mechanisms of plaque formation and rupture. *Circulation Research*, *114*(12), 1852-1866.

[2] Muraki, M., Mikami, T., Yoshimoto, T., Fujimoto, S., Kitaguchi, M., Kaga, S., ... & Kashiwaba, T. (2016). Sonographic detection of abnormal plaque motion of the carotid artery: its usefulness in diagnosing high-risk lesions ranging from plaque rupture to ulcer formation. *Ultrasound in medicine & biology*, *42*(2), 358-364.

[3] Benjamin, Emelia J., Michael J. Blaha, Stephanie E. Chiuve, Mary Cushman, Sandeep R. Das, Rajat Deo, Sarah D. De Ferranti *et al*. "Heart disease and stroke statistics—2017 update." (2017).

[4] Taylor, A. J., & Villines, T. C. (Eds.). (2012). *Atherosclerosis: clinical perspectives through imaging*. Springer Science & Business Media.

[5] Deldicque, L., & Francaux, M. (2012). Encyclopedia of exercise medicine in health and disease.

[6] Khan, M. G. (2005). *Encyclopedia of heart diseases*. Elsevier.

[7] Simionescu, M., & Sima, A. V. (2012). Morphology of atherosclerotic lesions. In *Inflammation and atherosclerosis* (pp. 19-37). Springer, Vienna.

[8] Gellman, M. D., & Turner, J. R. (Eds.). (2013). *Encyclopedia of behavioral medicine*. Springer New York.

[9] Carmeliet, P. (2003). Blood vessels and nerves: common signals, pathways and diseases. *Nature Reviews Genetics*, *4*(9), 710-720.

[10] Caplan, B., Kreutzer, J. S., & DeLuca, J. (2011). *Encyclopedia of Clinical Neuropsychology; With 199 Figures and 139 Tables*. Springer.

[11] Bungart, B. L., Lan, L., Wang, P., Li, R., Koch, M. O., Cheng, L., ... & Cheng, J. X. (2018). Photoacoustic tomography of intact human prostates and vascular texture analysis identify prostate cancer biopsy targets. *Photoacoustics*, *11*, 46-55.

[12] Chen, Y. C., Huang, A. L., Kyaw, T. S., Bobik, A., & Peter, K. (2016). Atherosclerotic plaque rupture: identifying the straw that breaks the camel's back, A*rteriosclerosis, thrombosis, and vascular biology*, *36*(8), e63-e72.

[13] Kakkos, S. K., Stevens, J. M., Nicolaides, A. N., Kyriacou, E., Pattichis, C. S., Geroulakos, G., & Thomas, D. (2007). Texture analysis of ultrasonic images of symptomatic carotid plaques can identify those plaques associated with ipsilateral embolic brain infarction. *European Journal of Vascular and Endovascular Surgery*, *33*(4), 422-429.

[14] Banchhor, S. K., Londhe, N. D., Araki, T., Saba, L., Radeva, P., Laird, J. R., & Suri, J. S. (2017). Wall-based measurement features provides an improved IVUS coronary artery risk assessment when fused with plaque texture-based features during machine learning paradigm. *Computers in biology and medicine*, *91*, 198-212.

[15] Johri, A. M., Herr, J. E., Li, T. Y., Yau, O., & Nambi, V. (2017). Novel ultrasound methods to investigate carotid artery plaque vulnerability. *Journal of the American Society of Echocardiography*, *30*(2), 139-148.

[16] Bea, A. M., Civeira, F., Jarauta, E., Lamiquiz-Moneo, I., Pérez-Calahorra, S., Marco-Benedí, V., & Mateo-Gallego, R. (2017). Association between the Presence of Carotid Artery Plaque and Cardiovascular Events in Patients with Genetic Hypercholesterolemia. *Revista Española de Cardiología (English Edition)*, *70*(7), 551-558.

[17] Mitchell, C., Korcarz, C. E., Gepner, A. D., Kaufman, J. D., Post, W., Tracy, R., ... & Stein, J. H. (2018). Ultrasound carotid plaque features, cardiovascular disease risk factors and events: the multi-ethnic study of atherosclerosis. *Atherosclerosis*, *276*, 195-202.

[18] Celletti, F. L., Waugh, J. M., Amabile, P. G., Brendolan, A., Hilfiker, P. R., & Dake, M. D. (2001). Vascular endothelial growth factor enhances atherosclerotic plaque progression. *Nature medicine*, *7*(4), 425-429.

[19] Niu, L., Qian, M., Yang, W., Meng, L., Xiao, Y., Wong, K. K., & Zheng, H. (2013). Surface roughness detection of arteries via texture analysis of ultrasound images for early diagnosis of atherosclerosis. *PloS one*, *8*(10), e76880.

[20] Offermanns, S., & Rosenthal, W. (Eds.). (2008). *Encyclopedia of molecular pharmacology*. Springer Science & Business Media.

[21] Guo, L., Cheng, L., He, W., Ju, Y., & Zhao, X. (2018). Ideal cardiovascular health and incidence of carotid plaque among middle-aged and elderly adults. *Journal of Stroke and Cerebrovascular Diseases*, *27*(2), 391-396.

[22] Chen, Y. C., Huang, A. L., Kyaw, T. S., Bobik, A., & Peter, K. (2016). Atherosclerotic plaque rupture: identifying the straw that breaks the camel's back: a*rteriosclerosis, thrombosis, and vascular biology*, *36*(8), e63-e72.

[23] Taylor, A. J., & Villines, T. C. (Eds.). (2012). *Atherosclerosis: clinical perspectives through imaging*. Springer Science & Business Media.

[24] Wick, Georg, Nicole Buhr, Gustav Fraedrich, and Cecilia Grundtman. "A Darwinian-evolutionary concept for atherogenesis: the role of immunity to HSP60." In *Inflammation and Atherosclerosis*, pp. 171-196. Springer, Vienna, 2012.

[25] Patel, N. A., Stamper, D. L., & Brezinski, M. E. (2005). Review of the ability of optical coherence tomography to characterize plaque, including a comparison with intravascular ultrasound. *Cardiovascular and interventional radiology*, *28*(1), 1-9.

[26] Yabushita, H., Bouma, B. E., Houser, S. L., Aretz, H. T., Jang, I. K., Schlendorf, K. H., ... & Tearney, G. J. (2002). Characterization of human atherosclerosis by optical coherence tomography. *Circulation*, *106*(13), 1640-1645.

[27] Drexler, W., & Fujimoto, J. G. (2008). State-of-the-art retinal optical coherence tomography. *Progress in retinal and eye research*, *27*(1), 45-88.

[28] Tearney, G. J., Boppart, S. A., Bouma, B. E., Brezinski, M. E., Weissman, N. J., Southern, J. F., & Fujimoto, J. G. (1996). Scanning single-mode fiber optic catheter–endoscope for optical coherence tomography. *Optics letters*, *21*(7), 543-545.

[29] Brezinski, M. E., Tearney, G. J., Bouma, B. E., Izatt, J. A., Hee, M. R., Swanson, E. A., ... & Fujimoto, J. G. (1996). Optical coherence tomography for optical biopsy: properties and demonstration of vascular pathology. *Circulation*, *93*(6), 1206-1213.

[30] Prati, F., Arbustini, E., Labellarte, A., Dal Bello, B., Sommariva, L., Mallus, M. T., ... & Boccanelli, A. (2001). Correlation between high frequency intravascular ultrasound and histomorphology in human coronary arteries. *Heart*, *85*(5), 567-570.

[31] Bouma, B. E., & Tearney, G. J. (1999). Power-efficient nonreciprocal interferometer and linear-scanning fiber-optic catheter for optical coherence tomography. *Optics letters*, *24*(8), 531-533.

[32] Jang, I. K., Bouma, B. E., Kang, D. H., Park, S. J., Park, S. W., Seung, K. B., ... & Houser, S. L. (2002). Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. *Journal of the American College of Cardiology*, *39*(4), 604-609.

[33] Kume, T., Akasaka, T., Kawamoto, T., Watanabe, N., Toyota, E., Neishi, Y., ... & Yoshida, K. (2005). Assessment of coronary intima-media thickness by optical coherence tomography. *Circulation Journal*, *69*(8), 903-907.

[34] Tearney, G. J., Jang, I. K., & Bouma, B. E. (2006). Optical coherence tomography for imaging the vulnerable plaque. *Journal of biomedical optics*, *11*(2), 021002.

[35] Tearney, G. J., & Bouma, B. E. (2001, May). Atherosclerotic plaque characterization by spatial and temporal speckle pattern analysis. In *Technical Digest. Summaries of papers presented at the Conference on Lasers and Electro-Optics. Postconference Technical Digest (IEEE Cat. No. 01CH37170)* (pp. 307-308). IEEE.

[36] Martin, A. J., Ryan, L. K., Gotlieb, A. I., Henkelman, R. M., & Foster, F. S. (1997). Arterial imaging: comparison of high-resolution US and MR imaging with histologic correlation. *Radiographics*, *17*(1), 189-202.

[37] Tobis, J. M., Mallery, J., Mahon, D., Lehmann, K., Zalesky, P., Griffith, J., ... & Dwyer, M. L. (1991). Intravascular ultrasound imaging of human coronary arteries in vivo. Analysis of tissue characterizations with comparison to in vitro histological specimens. *Circulation*, *83*(3), 913-926.

[38] Prati, F., Arbustini, E., Labellarte, A., Dal Bello, B., Sommariva, L., Mallus, M. T., ... & Boccanelli, A. (2001). Correlation between high frequency intravascular ultrasound and histomorphology in human coronary arteries. *Heart*, *85*(5), 567-570.

[39] Rumberger, J. A., Behrenbeck, T., Breen, J. F., & Sheedy, P. F. (1999). Coronary calcification by electron beam computed tomography and obstructive coronary artery disease: a model for costs and effectiveness of diagnosis as compared with conventional cardiac testing methods. *Journal of the American College of Cardiology*, *33*(2), 453-462.

[40] Wong, N. D., Kouwabunpat, D., Vo, A. N., Detrano, R. C., Eisenberg, H., Goel, M., & Tobis, J. M. (1994). Coronary calcium and atherosclerosis by ultrafast computed tomography in asymptomatic men and women: relation to age and risk factors. *American heart journal*, *127*(2), 422-430.

[41] Budoff, M. J., & Brundage, B. H. (1999). Electron beam computed tomography: screening for coronary artery disease. *Clinical cardiology*, *22*(9), 554-558.

[42] Budoff, M. J., Cohen, M. C., Garcia, M. J., Hodgson, J. M., Hundley, W. G., ... & Rodgers, G. P. (2005). ACCF/AHA clinical competence statement on cardiac imaging with computed tomography and magnetic resonance: a report of the American College of Cardiology Foundation/American Heart Association/American College of Physicians Task Force on clinical competence and training: developed in collaboration with the American Society of Echocardiography, American Society of Nuclear Cardiology, Society of Atherosclerosis Imaging, and the Society for Cardiovascular Angiography & Interventions: endorsed by the Society .... *Circulation*, *112*(4), 598-617.

[43] Baer, F. M., Theissen, P., Schneider, C. A., Kettering, K., Voth, E., Sechtem, U., & Schicha, H. (1999). MRI assessment of myocardial viability: comparison with other imaging techniques. *Rays*, *24*(1), 96.

[44] MacNeill, B. D., Lowe, H. C., Takano, M., Fuster, V., & Jang, I. K. (2003). Intravascular modalities for detection of vulnerable plaque: current status. *Arteriosclerosis, thrombosis, and vascular biology*, *23*(8), 1333-1342.

[45] Jang, I. K., Bouma, B. E., Kang, D. H., Park, S. J., Park, S. W., Seung, K. B., ... & Houser, S. L. (2002). Visualization of coronary atherosclerotic plaques in patients using optical coherence tomography: comparison with intravascular ultrasound. *Journal of the American College of Cardiology*, *39*(4), 604-609.

[46] Gossage, K. W., Tkaczyk, T. S., Rodriguez, J. J., & Barton, J. K. (2003). Texture analysis of optical coherence tomography images: feasibility for tissue classification. *Journal of biomedical optics*, *8*(3), 570-576.

[47] Christodoulou, C. I., Pattichis, C. S., Pantziaris, M., & Nicolaides, A. (2003). Texture-based classification of atherosclerotic carotid plaques. *IEEE transactions on medical imaging*, *22*(7), 902-912.

[48] Dhawan, A. P., Chitre, Y., & Kaiser-Bonasso, C. (1996). Analysis of mammographic microcalcifications using gray-level image structure features. *IEEE Transactions on medical imaging*, *15*(3), 246-259.

[49] Lee, J., Prabhu, D., Kolluru, C., Gharaibeh, Y., Zimin, V. N., Dallan, L. A., ... & Wilson, D. L. (2020). Fully automated plaque characterization in intravascular OCT images using hybrid convolutional and lumen morphology features. *Scientific Reports*, *10*(1), 1-13.

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

[51] Abdolmanafi, A., Dahdah, N., Duong, L., Adib, R. I., & Cheriet, F. (2019). Fully automatic artificial intelligence diagnostic model of coronary artery lesions using OCT. *Canadian Journal of Cardiology*, *35*(10), S61-S62.

[52] Yong, Y. L., Tan, L. K., McLaughlin, R. A., Chee, K. H., & Liew, Y. M. (2017). Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography. *Journal of biomedical optics*, *22*(12), 126005.

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

[54] Loizou, C. P., Murray, V., Pattichis, M. S., Seimenis, I., Pantziaris, M., & Pattichis, C. S. (2010). Multiscale amplitude-modulation frequency-modulation (AM–FM) texture analysis of multiple sclerosis in brain MRI images. *IEEE Transactions on Information Technology in Biomedicine*, *15*(1), 119-129.

[55] Zhang, Y. (2012). MRI texture analysis in multiple sclerosis. *International Journal of Biomedical Imaging*, *2012*.

[56] Doonan, R. J., Dawson, A. J., Kyriacou, E., Nicolaides, A. N., Corriveau, M. M., Steinmetz, O. K., ... & Daskalopoulou, S. S. (2013). Association of ultrasonic texture and echodensity features between sides in patients with bilateral carotid atherosclerosis. *European Journal of Vascular and Endovascular Surgery*, *46*(3), 299-305.

[57] van Engelen, A., Wannarong, T., Parraga, G., Niessen, W. J., Fenster, A., Spence, J. D., & de Bruijne, M. (2014). Three-dimensional carotid ultrasound plaque texture predicts vascular events. *Stroke*, *45*(9), 2695-2701.

[58] Niu, L., Qian, M., Yang, W., Meng, L., Xiao, Y., Wong, K. K., & Zheng, H. (2013). Surface roughness detection of arteries via texture analysis of ultrasound images for early diagnosis of atherosclerosis. *PloS one*, *8*(10), e76880.

[59] Christodoulou, C. I., Pattichis, C. S., Kyriacou, E., & Nicolaides, A. (2010). Image retrieval and classification of carotid plaque ultrasound images. *The Open Cardiovascular Imaging Journal*, *2*(1).

[60] Voros, S., Rinehart, S., Qian, Z., Joshi, P., Vazquez, G., Fischer, C., ... & Villines, T. C. (2011). Coronary atherosclerosis imaging by coronary CT angiography: current status, correlation with intravascular interrogation and meta-analysis. *JACC: Cardiovascular Imaging*, *4*(5), 537-548.

[61] Tearney, G. J., & Bouma, B. E. (2002). Atherosclerotic plaque characterization by spatial and temporal speckle pattern analysis. *Optics letters*, *27*(7), 533-535.

[62] Lindenmaier, A. A., Conroy, L., Farhat, G., DaCosta, R. S., Flueraru, C., & Vitkin, I. A. (2013). Texture analysis of optical coherence tomography speckle for characterizing biological tissues in vivo. *Optics letters*, *38*(8), 1280-1282.

[63] Gossage, K. W., Smith, C. M., Kanter, E. M., Hariri, L. P., Stone, A. L., Rodriguez, J. J., ... & Barton, J. K. (2006). Texture analysis of speckle in optical coherence tomography images of tissue phantoms. *Physics in Medicine & Biology*, *51*(6), 1563.

[64] Narayan, N. S., Marziliano, P., Kanagalingam, J., & Hobbs, C. G. (2015). Speckle patch similarity for echogenicity-based multiorgan segmentation in ultrasound images of the thyroid gland. *IEEE journal of biomedical and health informatics*, *21*(1), 172-183.

[65] Moon, W. K., Lo, C. M., Huang, C. S., Chen, J. H., & Chang, R. F. (2012). Computer-aided diagnosis based on speckle patterns in ultrasound images. *Ultrasound in medicine & biology*, *38*(7), 1251-1261.

[66] Shiomi, M., Ito, T., Yamada, S., Kawashima, S., & Fan, J. (2003). Development of an animal model for spontaneous myocardial infarction (WHHLMI rabbit). *Arteriosclerosis, thrombosis, and vascular biology*, *23*(7), 1239-1244.

[67] Kobayashi, T., Ito, T., & Shiomi, M. (2011). Roles of the WHHL rabbit in translational research on hypercholesterolemia and cardiovascular diseases. *BioMed Research International*, *2011*.

[68] Arevalo, J. F. (Ed.). (2008). *Retinal angiography and optical coherence tomography*. Springer Science & Business Media.

[69] Yamaguchi, T., Terashima, M., Akasaka, T., Hayashi, T., Mizuno, K., Muramatsu, T., ... & Takayama, T. (2008). Safety and feasibility of an intravascular optical coherence tomography image wire system in the clinical setting. *The American journal of cardiology*, *101*(5), 562-567.

[70] Bouma, B. E., Villiger, M., Otsuka, K., & Oh, W. Y. (2017). Intravascular optical coherence tomography. *Biomedical optics express*, *8*(5), 2660-2686.

[71] Yamaguchi, T., Terashima, M., Akasaka, T., Hayashi, T., Mizuno, K., Muramatsu, T., ... & Takayama, T. (2008). Safety and feasibility of an intravascular optical coherence tomography image wire system in the clinical setting. *The American journal of cardiology*, *101*(5), 562-567.

[72] Li, Y., Jing, J., Heidari, E., Zhu, J., Qu, Y., & Chen, Z. (2017). Intravascular optical coherence tomography for characterization of atherosclerosis with a 1.7 micron swept-source laser. *Scientific reports*, *7*(1), 1-6.

[73] Kishi, S. (2016). Impact of swept source optical coherence tomography on ophthalmology. *Taiwan journal of ophthalmology*, *6*(2), 58-68.

[74] Jain, A., Nandakumar, K., & Ross, A. (2005). Score normalization in multimodal biometric systems. *Pattern recognition*, *38*(12), 2270-2285.

[75] Boughorbel, S., Jarray, F., & El-Anbari, M. (2017). Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. *PloS one*, *12*(6), e0177678.

[76] Moccia, S., De Momi, E., El Hadji, S., & Mattos, L. S. (2018). Blood vessel segmentation algorithms—review of methods, datasets and evaluation metrics. *Computer methods and programs in biomedicine*, *158*, 71-91.

[77] Hossin, M., & Sulaiman, M. N. (2015). A review on evaluation metrics for data classification evaluations. *International Journal of Data Mining & Knowledge Management Process*, *5*(2), 1.

[78] Popovic, A., De la Fuente, M., Engelhardt, M., & Radermacher, K. (2007). Statistical validation metric for accuracy assessment in medical image segmentation. *International Journal of Computer Assisted Radiology and Surgery*, *2*(3-4), 169-181.

[79,] Prakash, A., Hewko, M. D., Sowa, M., & Sherif, S. S. (2015). Detection of atherosclerotic plaque from optical coherence tomography images using texture-based segmentation. *Современные технологии в медицине*, *7*(1 (eng)).

[80] Prakash, A., Macias, M. O., Hewko, M., Sowa, M., & Sherif, S. (2016, May). Vascular plaque detection with reduced textural feature set from optical coherence tomography images. In *2016 Photonics North (PN)* (pp. 1-1). IEEE.

[81] Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE*, *67*(5), 786-804.

[83] Lerski, R. A., Straughan, K., Schad, L. R., Boyce, D., Blüml, S., & Zuna, I. (1993). VIII. MR image texture analysis—an approach to tissue characterization. *Magnetic resonance imaging*, *11*(6), 873-887.

[84] Materka, A., & Strzelecki, M. (1998). Texture analysis methods–a review. *Technical university of lodz, institute of electronics, COST B11 report, Brussels*, *10*(1.97), 4968.

[85] Valkealahti, K., & Oja, E. (1998). Reduced multidimensional co-occurrence histograms in texture classification. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(1), 90-94.

[86] Julesz, B. (1975). Experiments in the visual perception of texture. *Scientific American*, *232*(4), 34-43.

[81] Galloway, M. M. (1974). Texture analysis using grey level run lengths. *STIN*, *75*, 18555.

[87] Weszka, J. S., Dyer, C. R., & Rosenfeld, A. (1976). A comparative study of texture measures for terrain classification. *IEEE transactions on Systems, Man, and Cybernetics*, (4), 269-285.

[88] Lendaris, G. G., & Stanley, G. L. (1970). Diffraction-pattern sampling for automatic pattern recognition. *Proceedings of the IEEE*, *58*(2), 198-216.

[89] Conners, R. W., & Harlow, C. A. (1980). A theoretical comparison of texture algorithms. *IEEE transactions on pattern analysis and machine intelligence*, (3), 204-222.

[90] Nguyen, N. G., Poulsen, R. S., & Louis, C. (1983). Some new color features and their application to cervical cell classification. *Pattern recognition*, *16*(4), 401-411.

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

[92] Haralick, R. M., Shanmugam, K., & Dinstein, I. H. (1973). Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6), 610-621.

[93] Argenti, F., Alparone, L., & Benelli, G. (1990, December). Fast algorithms for texture analysis using co-occurrence matrices. In *IEE Proceedings F (Radar and Signal Processing)* (Vol. 137, No. 6, pp. 443-448). IET Digital Library.

[94] Theodoridis, S. (2003). Konstantinos koutroumbas. *Pattern recognition*.

[95] Joel, G. (2015). Data Science from Scratch.

[96] Prakash, A., Hewko, M., Sowa, M., & Sherif, S. (2013, May). Texture based segmentation method to detect atherosclerotic plaque from optical tomography images. In *European Conference on Biomedical Optics* (p. 88020S). Optical Society of America.

[97] Koza, J. R., Andre, D., Keane, M. A., & Bennett III, F. H. (1999). *Genetic programming III: Darwinian invention and problem solving* (Vol. 3). Morgan Kaufmann.

[98] Whitley, D. (1994). A genetic algorithm tutorial. *Statistics and computing*, *4*(2), 65-85.

[99] Tan, F., Fu, X., Zhang, Y., & Bourgeois, A. G. (2008). A genetic algorithm-based method for feature subset selection. *Soft Computing*, *12*(2), 111-120.

[100] Kim, Y., Street, W. N., & Menczer, F. (2000, August). Feature selection in unsupervised learning via evolutionary search. In *Proceedings of the sixth ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 365-369).

[101] Prakash, A., Macias, M. O., Hewko, M., Sowa, M., & Sherif, S. (2016, May). Vascular plaque detection with reduced textural feature set from optical coherence tomography images. In *2016 Photonics North (PN)* (pp. 1-1). IEEE.

[102] Silvera, S. A. N., Mayne, S. T., Risch, H. A., Gammon, M. D., Vaughan, T., Chow, W. H., ... & West, A. B. (2011). Principal component analysis of dietary and lifestyle patterns in relation to risk of subtypes of esophageal and gastric cancer. *Annals of epidemiology*, *21*(7), 543-550.

[103] Al-Kandari, N. M., & Jolliffe, I. T. (2005). Variable selection and interpretation in correlation principal components. *Environmetrics: The official journal of the International Environmetrics Society*, *16*(6), 659-672.

[104] Zou, H., Hastie, T., & Tibshirani, R. (2006). Sparse principal component analysis. *Journal of computational and graphical statistics*, *15*(2), 265-286.

[105] Jolliffe, I. T., Trendafilov, N. T., & Uddin, M. (2003). A modified principal component technique based on the LASSO. *Journal of computational and Graphical Statistics*, *12*(3), 531-547.

[106] Shen, H., & Huang, J. Z. (2008). Sparse principal component analysis via regularized low rank matrix approximation. *Journal of multivariate analysis*, *99*(6), 1015-1034.

[107] d'Aspremont, A., Ghaoui, L., Jordan, M., & Lanckriet, G. (2004). A direct formulation for sparse PCA using semidefinite programming. *Advances in neural information processing systems*, *17*, 41-48.

[108] Witten, D. M., Tibshirani, R., & Hastie, T. (2009). A penalized matrix decomposition, with applications to sparse principal components and canonical correlation analysis. *Biostatistics*, *10*(3), 515-534.

[109] Guan, Y., & Dy, J. (2009, April). Sparse probabilistic principal component analysis. In *Artificial Intelligence and Statistics* (pp. 185-192).

[110] Bernardo, J. M., Bayarri, M. J., Berger, J. O., Dawid, A. P., Heckerman, D., Smith, A., & West, M. (2003). Bayesian factor regression models in the "large p, small n" paradigm. *Bayesian statistics*, *7*, 733-742.

[111] Carvalho, C. M., Chang, J., Lucas, J. E., Nevins, J. R., Wang, Q., & West, M. (2008). High-dimensional sparse factor modeling: applications in gene expression genomics. *Journal of the American Statistical Association*, *103*(484), 1438-1456.

[112] Lee, M., Shen, H., Huang, J. Z., & Marron, J. S. (2010). Biclustering via sparse singular value decomposition. *Biometrics*, *66*(4), 1087-1095.

[113] Yang, D., Ma, Z., & Buja, A. (2014). A sparse singular value decomposition method for high-dimensional data. *Journal of Computational and Graphical Statistics*, *23*(4), 923-942.

[114] Sill, M., Kaiser, S., Benner, A., & Kopp-Schneider, A. (2011). Robust biclustering by sparse singular value decomposition incorporating stability selection. *Bioinformatics*, *27*(15), 2089-2097.

[115] Bi, J., Bennett, K., Embrechts, M., Breneman, C., & Song, M. (2003). Dimensionality reduction via sparse support vector machines. *Journal of Machine Learning Research*, *3*(Mar), 1229-1243.

[116] Cadima, J., & Jolliffe, I. T. (1995). Loading and correlations in the interpretation of principle compenents. *Journal of applied Statistics*, *22*(2), 203-214.

[117] Rezaei, Z., Selamat, A., Taki, A., Rahim, M. S. M., & Kadir, M. R. A. (2017). Automatic plaque segmentation based on hybrid fuzzy clustering and k nearest neighborhood using virtual histology intravascular ultrasound images. *Applied Soft Computing*, *53*, 380-395.

[118] Abdel-Dayem, A. R., & El-Sakka, M. R. (2007, August). Fuzzy c-means clustering for segmenting carotid artery ultrasound images. In *International Conference Image Analysis and Recognition* (pp. 935-948). Springer, Berlin, Heidelberg.

[119] Hassan, M., Chaudhry, A., Khan, A., & Kim, J. Y. (2012). Carotid artery image segmentation using modified spatial fuzzy c-means and ensemble clustering. *Computer methods and programs in biomedicine*, *108*(3), 1261-1276.

[120] Hassan, M., Chaudhry, A., Khan, A., & Iftikhar, M. A. (2014). Robust information gain based fuzzy c-means clustering and classification of carotid artery ultrasound images. *Computer methods and programs in biomedicine*, *113*(2), 593-609.

[121] Adame, I. M., van der Geest, R. J., Wasserman, B. A., Mohamed, M., Reiber, J. H. C., & Lelieveldt, B. P. (2004, May). Automatic plaque characterization and vessel wall segmentation in magnetic resonance images of atherosclerotic carotid arteries. In *Medical Imaging 2004: Image Processing* (Vol. 5370, pp. 265-273). International Society for Optics and Photonics.

[122] Abdel-Dayem, A. R., & El-Sakka, M. R. (2007, August). Fuzzy c-means clustering for segmenting carotid artery ultrasound images. In *International Conference Image Analysis and Recognition* (pp. 935-948). Springer, Berlin, Heidelberg.

[123] Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers & Geosciences*, *10*(2-3), 191-203.

[124] Pentland, A. P. (1984). Fractal-based description of natural scenes. *IEEE transactions on pattern analysis and machine intelligence*, (6), 661-674.

[125] Chaudhuri, B. B., & Sarkar, N. (1995). Texture segmentation using fractal dimension. *IEEE Transactions on pattern analysis and machine intelligence*, *17*(1), 72-77.

[126] Kaplan, L. M., & Kuo, C. C. (1995). Texture roughness analysis and synthesis via extended self-similar (ESS) model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *17*(11), 1043-1056.

[127] Owen, Tony. "Visual Reconstruction by Andrew Blake and Andrew Zisserman, The MIT Press, Massachusetts, USA, 1987 ." *Robotica* 6, no. 2 (1988): 166-166.

[128] Besag, J. (1974). Spatial interaction and the statistical analysis of lattice systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, *36*(2), 192-225.

[129] Won, C. S., & Gray, R. M. (2004). *Stochastic image processing*. Springer Science & Business Media.

[130] Unser, M. (1995). Texture classification and segmentation using wavelet frames. *IEEE Transactions on image processing*, *4*(11), 1549-1560.

[131] Ma, J., & Plonka, G. (2010). The curvelet transform. *IEEE signal processing magazine*, *27*(2), 118-133.

[132] Candes, E., Demanet, L., Donoho, D., & Ying, L. (2006). Fast discrete curvelet transforms. *Multiscale Modeling & Simulation*, *5*(3), 861-899.

## APPENDIX A: SPATIAL GRAY LEVEL DEPENDENCE MATRICES (SGLDM)

This appendix details the **SGLDM** matrices used in this dissertation.

The features resulting from first-order statistics, such as central moments, provide information about the image's gray-level distribution, but they do not provide any information about the relative positions of the various gray levels, e.g., if all low-value gray-level pixels are placed together or if are they are interchangeable. Such information can be generated from the second-order histograms, where the pixels are considered in pairs. The two new parameters are the relative distance and their relative orientations among the pixels. Let *d* be the relative distance measured in pixel numbers (*d* equals1 for neighboring pixels). The orientation, **θ**, is measured in four directions: horizontal, diagonal, vertical, and anti-diagonal (0°, 45°, 90°, 135°), as shown in Fig. A-1.

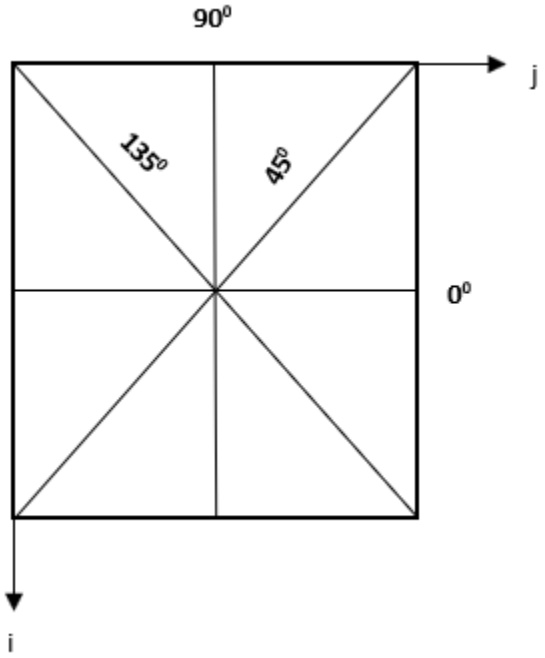



Fig. A-1. The four orientations used to construct SGLDM matrices.

For each combination of *d* and **θ**, a two-dimensional histogram is defined as,

$$0^0 = P(I(i,j) = I_1, I(i \pm d, j) = I_2)$$
$$90^0 = P(I(i,j) = I_1, I(i, j \mp d) = I_2)$$

These probability density functions, $p(i,j;d,\theta)$, measure the probability that two pixels located with an inter-sample distance, *d*, and direction, $\theta$, will have gray levels *i* and *j*.

This is known as the *Spatial Gray Level Dependence Matrix* (**SGLDM**) method, which is a known method for extracting second-order statistical texture features. Rosenfeld and Troy [A1] and Haralick *et al.* [A2] first proposed **SGLDM** matrices for arbitrary spatial distances and angular directions. The **SGLDM** method determines the probability of the occurrence of gray levels with respect to relative spatial pixel positions in an image. **SGLDM** matrices are based on an estimate of the second-order joint-conditional probability density functions, $p(i,j;d,\theta)$.

The following examples illustrate the construction of **SGLDM** matrices:

Let $I(i,j)$ be a 4 ×4 image

$$I = \begin{bmatrix} 1 & 2 & 1 & 2 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 2 & 1 \end{bmatrix}$$

then 4×4 image with three gray levels $N_g$= 0, 1, and 2.

$$A = \frac{1}{R}\begin{bmatrix} \eta(0,0) & \eta(0,1) & \eta(0,2) \\ \eta(1,0) & \eta(1,1) & \eta(1,2) \\ \eta(2,0) & \eta(2,1) & \eta(2,2) \end{bmatrix}$$

An **SGLDM** matrix for a pair ($d$, $\boldsymbol{\theta}$), where R is the total number of pixel pairs.

$$A^0(d=1) = \frac{1}{23}\begin{bmatrix} 0 & 2 & 2 \\ 2 & 0 & 6 \\ 2 & 7 & 2 \end{bmatrix}\begin{matrix} 0 \\ 1 \\ 2 \end{matrix}$$
$$0 \quad 1 \quad 2$$

The **SGLDM** matrix for the above image with d = 1, **θ**=0.

For each of the intensity pairs, such as (0, 0), we count the number of pixel pairs at a relative distance *d* =1 and orientation ***θ***=0◦ that take these values.

After using the probabilities of gray-level occurrence with respect to the pixel position in order to form the SGLDM matrices, we use them to calculate the corresponding *Haralick* features. Some of these features have a direct interpretation with respect to texture; for example, the angular second-moment feature is the measure of the image's smoothness; contrast is the measure of local gray-level variation within the image, and entropy is the measure of randomness in an image, and therefore produces low values for smooth images. However, there are other features that do not possess such a direct interpretation but that can still convey texture-related information with high discriminatory power. Even though these textural features contain information about the textural characteristics of an image, it is difficult to identify which specific textural characteristic is represented by each of these features.

## References

A1. Haralick, R. M. (1979). Statistical and structural approaches to texture. *Proceedings of the IEEE, 67*(5), 786-804.

## APPENDIX B: LEAST ABSOLUTE SHRINKAGE AND SELECTION OPERATOR

Least Absolute Shrinkage and Selection Operator (**LASSO**) was originally formulated in the context of least squares models [B1]. In signal processing, the **LASSO** method is also known as Basis Pursuit [B2]. The **LASSO** method is a coefficient-shrunken version of the ordinary least square estimated. As such, it minimizes the squared sum of residual subjection to constrain the sum of the absolute value of the coefficients, which should be no greater than a constant. Let us assume,

$$Y = f(x) + error \tag{B.1}$$

and $E(\varepsilon) = 0$, $Var(\varepsilon)=\sigma^2$, then the prediction error of the estimate $\hat{f}(x)$ is

$$Error(x) = E\left[\left(y - \hat{f}(x)\right)\right]$$

$$Error(x) = \sigma^2 + \left[E\hat{f}(x) - f(x)\right]^2 + E\left[\hat{f}(x) - E\hat{f}(x)\right]^2$$

$$Error(x) = \sigma^2 + Bias^2\left(\hat{f}(x)\right) + var(\hat{f}(x)) \tag{B.2}$$

Ordinary least square estimates often have low bias but high variance; however, the **LASSO** method can improve overall prediction accuracy by sacrificing a little bias to reduce the variance of the predicted value.

**LASSO** takes the following form:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{\sum_{i=1}^{N}\left(y_i - \alpha - \sum_j \beta_j - x_{ij}\right)^2\right\} \text{ subject to } \left|\beta_j\right| \le t \tag{B.3}$$

The criterion, $\sum_{i=1}^{N}\left(y_i - \sum_j \beta_j - x_{ij}\right)^2$, equals the quadratic function as

$$\left(\beta - \hat{\beta}^0\right)^T X'X\left(\beta - \hat{\beta}^0\right) \tag{B.4}$$

This function is represented as elliptical contours.

The $L_1$ norm constraint in the **LASSO** method is represented as a square centered at the origin, and the **LASSO** solution is the first place where the contour touches the square.

Another equivalent of **LASSO** exists in a *Lagrangian form*:

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \left\{ \frac{1}{2} \sum_{i=1}^{N} \left( y_i - \sum_j \beta_j - x_{ij} \right)^2 + \lambda \sum_j |\beta_j| \right\}$$

$$= \underset{\beta}{\operatorname{argmin}} \|Y - X\beta\|_2^2 + \lambda \|\beta\|_1 \qquad \text{(B.5)}$$

where $\|\beta\|_1$ is the $L_1$ **LASSO** penalty. This constraint makes the solution nonlinear in Y and lacking a closed-form, making it solvable using LARS.

## References

B1. Hastie, T., Tibshirani, R., & Wainwright, M. (2015). *Statistical learning with sparsity: the lasso and generalizations*. CRC press.

B2. Chen, S., & Donoho, D. (1994, October). Basis pursuit. In *Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers* (Vol. 1, pp. 41-44). IEEE.

## APPENDIX C: LEAST ANGLE REGRESSION (LARS)

Least Angle Regression (**LARS**) is a very efficient algorithm that solves for the entire path of the solution as Lambda is varied [C1]. **LARS** requires the same computation as least square fit with *p* variables and takes p steps to reach a full least-squares fit. **LARS** follows a similar concept to forward stepwise regression, which builds the model sequentially by adding one variable at a time. At each sequential step, forward stepwise regression finds the best set of variables and keeps updating the least square fit to include all the best variables. **LARS** operates similarly: in the first step, **LARS** identifies the variable most correlated with the response variable and keeps it in the active set until the second variable enters. Unlike forward stepwise regression, **LARS** does not take many small steps to select variables; instead, the steps are determined algebraically. The computation cost of **LARS** is $O(p^3 + np^2)$ computations, where p is the number of variables. This computational cost is the same as least squares for p variables. A more detailed description of **LARS** follows below:

Step 1: Standardize all variables to have a mean of zero and a unit norm. Start with all coefficients, $\beta_1, \beta_2, \ldots, \beta_p$, equal to zero

Step 2: find the variable, $x_j$, most correlated with the residual, $r$, by finding the largest value of $\langle x_j, r_j \rangle$, and define the active set, $\mathcal{A} = \{x_j\}$, and the active matrix, $X_{\mathcal{A}_j}$, containing active variables as its columns.

Step 3: Move $\beta_j$ from 0 in the direction of its least-squares coefficients $\langle x_j, r_j \rangle$. The direction is defined as $\Delta_j = \left(X_{\mathcal{A}_j}^T X_{\mathcal{A}_j}\right)^{-1} X_{\mathcal{A}_j}^T r_j$, and the coefficient becomes $\beta_{\mathcal{A}_j}(\eta) = \beta_{\mathcal{A}_j} + \eta . \Delta_j$, where $\eta$

is the step size defined as $\eta = \min_{j \in X_{\mathcal{A}}} \left\{ \frac{\hat{C} - \hat{c}_j}{X_{\mathcal{A}_j} - r_j}, \frac{\hat{C} + \hat{c}_j}{X_{\mathcal{A}_j} + r_j} \right\}$ , $c_j$ is the current correlation, and $\hat{C} = \min_j \{\hat{c}_j\}$.

Step 4: Move $\beta_j$ until some other variable, $x_k$, catches up and has as much as correlation with the residual as $x_j$ and enters in the active set.

Step 5: Continue until all of the variables have been entered into the active set. At this point, we will have arrived at the full least-squares solution.

## References:

C1. Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression. *The Annals of statistics*, *32*(2), 407-499.

## LIST OF ACRONYMS

**AIC**: Akaike Information Criterion

**AM-FM**: Amplitude Modulation-Frequency Modulation

**APAC**: Asymptomatic Polyvascular Abnormalities Community

**ASM**: Angular Second Moment

**CHD**: Coronary Heart Disease

**CT**: Computed Tomography

**CTA**: Computerized Tomography Angiography

**CVD**: CardioVascular Disease

**CVH**: CardioVascular Health score

**FFT**: Fast Fourier transform

**FH**: Familial Hypercholesterolemia

**FP**: False Positive

**FN**: False Negative

**GLCM**: Gray Level Co-occurrence Matrix

**GA**: Genetic Algorithm

**GRF**: Gibbs Random Function

**GSM**: Gray-Scale Median

**HC**: HyperCholesterolemia

**HMRF**: Hidden Markov Random Field

**ICA**: Internal Carotid Artery

**ICM**: Iterated Conditional Modes

**IVOCT**: Intravascular Optical Coherence Tomography

**IVUS**: Intravascular Ultrasound

**JPDF**: Joint Probability Distribution Function

**MCC**: Matthews' Correlation Coefficient

**MRF**: Markov Random Field

**MRF-MAP**: Markov Random Field-Maximum A Posteriori

**MI**: Myocardial Infarction

**MRI**: Magnetic Resonance Imaging

**LARS**: Least Angle Regression

**LASSO**: Least Absolute Shrinkage and Selection Operator

**LDL**: Low-Density lipoprotein

**OCT**: Optical Coherence Tomography

**PC**: Principal Component

**PCA**: Principal Component analysis

**PCR**: Principal Component Regression

**SGLDM**: Spatial Gray Level-Dependent Matrix

**SS-OCT**: Swept-Source Optical Coherence Tomography

**SVD**: Singular Value Decomposition

**TP**: True Positive

**TN**: True Negative

**VLDL**: Very Low-Density Lipoprotein

**WHHLMI**: Watanabe Heritable Hyperlipidemic Rabbits