CHARACTERIZATION OF CORONARY ARTERY TISSUES WITH KAWASAKI DISEASE FROM OCT IMAGES

by

Atefeh ABDOLMANAFI

MANUSCRIPT-BASED THESIS PRESENTED TO ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
IN PARTIAL FULFILLMENT FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
Ph.D.

MONTREAL, JUNE 27TH, 2018

ÉCOLE DE TECHNOLOGIE SUPÉRIEURE
UNIVERSITÉ DU QUÉBEC

BOARD OF EXAMINERS

THIS THESIS HAS BEEN EVALUATED

BY THE FOLLOWING BOARD OF EXAMINERS

Prof. Luc Duong, Thesis Supervisor
Department of Software and IT Engineering, École de technologie supérieure

Dr. Nagib Dahdah, Co-supervisor
Division of Pediatric Cardiology and Research Center, Centre Hospitalier Universitaire Sainte-Justine

Prof. Jean-Marc Lina, President of the Board of Examiners
Department of Electrical Engineering, École de technologie supérieure

Prof. Sylvie Ratté, Member of the jury
Department of Software and IT Engineering, École de technologie supérieure

Prof. Santiago Costantino, External Examiner
Research Center, Hôpital Maisonneuve-Rosemont

THIS THESIS WAS PRESENTED AND DEFENDED

IN THE PRESENCE OF A BOARD OF EXAMINERS AND THE PUBLIC

ON MAY 28TH,2018

AT ÉCOLE DE TECHNOLOGIE SUPÉRIEURE

ACKNOWLEDGEMENTS

Professor Luc Duong, I learned with you the real definition of education as a combination of science, respect, honesty, responsibility, and collaboration. I learned the importance of gaining a deep understanding of problems by touching them with observation and perception before thinking about the solutions based on algorithms, mathematical models, and physical concepts. You trusted me and gave me the courage to take strong steps forward, although you knew all of my weaknesses. Above all, I learned with you how to honor science by allowing the intentions for the work to be grounded in values. Deep in my heart, I would like to say that although human beings are forgetful by nature, I will never forget the fact that you have changed my life.

This research project is co-supervised by Dr.Nagib Dahdah, an experienced interventional cardiologist at CHU Sainte-Justine. Very special thanks to you for guiding me with great insight into the clinical aspects of the project, which were very important in understanding and stating the research problems and finding the appropriate solutions.

Professor Farida Cheriet, I appreciate your collaboration at each step of the work by providing me with valuable comments, constructive criticisms, and precious scientific suggestions.

I would like to appreciate Dr.Ibrahim Ragui Adib, an experienced cardiologist at Pierre Boucher hospital, for guiding me with the clinical aspects of the project.

I would like to thank the members of the jury for their time in evaluating this thesis and for their valuable comments.

I learned and practiced how to keep moving towards success by respecting and understanding values in a very positive environment. I am very thankful to all of my great lab-mates and amazing friends who have contributed to creating such a valuable achievement in my life.

Very special thanks to Samuel Docquier for his technical support, which is very important to keep the project in progress.

To my little angel, Abtin, I have had a great chance of having such an unbelievable love in my life. I've always been amazed by your huge amount of understanding and support. Thanks for all the paintings that you created on my assignments with a naughty smile that was reflecting the most beautiful childish passion. The many colorful twisted lines have reminded me that even if the path is upside down and complicated, it is still full of positivity, colors, and beauty.

My unique family, you have built an incredibly beautiful world inside of me with the foundation of respect, honesty, freedom, and love. I was responsible to start my journey in my world to discover it and strengthen its foundation. You gave me the courage to go on my journey even though you knew the way is long, hard, and cruel. But you never left me alone, since my travel bag is full of treasures of values that each of you created for me during all these years. When we linked our hands as a symbol of unity and solidarity and linked our hearts through a common world of light and brightness, we promised to be responsible for the universe, creation, and humanity with a common goal, which is creation of love. This motivates me to stand strong, and work hard for deep exploration of realities to discover more about the universe by knowing myself.

This research was funded by the Fonds de Recherche du Québec-Nature et technologies (FRQNT).

CARACTÉRISATION TISSULAIRE DES ARTÈRES CORONAIRES PAR TOMOGRAPHIE EN COHÉRENCE OPTIQUE POUR LA MALADIE DE KAWASAKI

Atefeh ABDOLMANAFI

RÉSUMÉ

La maladie de Kawasaki est une maladie infantile affectant les ganglions lymphatiques cutanéomuqueux et caractérisé par de la fièvre, des éruptions cutanées, une conjonctivite non exsudative bilatérale, des érythèmes des lèvres et de la muqueuse buccale et des mains et pieds érythémateux enflés. La maladie de Kawasaki conduit à l'inflammation dans les parois des artères de taille moyenne dans tout le corps. Bien qu'une forte dose de perfusion d'immunoglobulines intraveineuses (IVIG) réduise le risque de complications coronariennes, environ 15% à 25% des enfants non traités présentent un risque d'anévrisme de l'artère coronaire. L'épaississement de l'intima, la disparition de la media, les calcifications lamellaires, la fibrose, les macrophages et la néovascularisation sont les caractéristiques pathologiques les plus typiques des lésions coronaires tardives de la maladie de Kawasaki. Dans les cas graves, ils peuvent entraîner un infarctus du myocarde et engendrer une mort subite. Puisque la fonctionnalité des tissus cardiaques dépend du flux sanguin coronaire vers le myocarde, l'évaluation intravasculaire des tissus de l'artère coronaire est importante pour détecter les formations pathologiques provoquées par différentes complications de l'artère coronaire.

La visualisation in vivo des artères coronaires fournit une information très pertinente pour le suivi et le traitement des artères coronaires. Cependant, l'acquisition d'images est particulièrement difficile chez les patients pédiatriques en raison de la petite taille des vaisseaux et de la fréquence cardiaque élevée. La tomographie en cohérence optique (TCO) est une modalité d'imagerie interférométrique en infrarouge permettant d'acquérir des images en coupe transversale des tissus examinés à l'échelle micrométrique. La TCO a été développée pour le diagnostic et le traitement de la maladie coronarienne dans la population adulte. Il a récemment été utilisé en cardiologie pédiatrique pour l'imagerie des tissus de l'artère coronaire avec des résultats d'innocuité appropriés dans ce groupe d'âge. Il a une haute résolution allant de 10 à 20 μm pour caractériser la structure interne des tissus tels que les couches de paroi des vaisseaux et l'accumulation de plaque. La géométrie de la paroi interne des vaisseaux permet de détecter et d'évaluer les propriétés biophysiques et dynamiques de la paroi artérielle, l'épaisseur des couches de l'artère coronaire et diverses anomalies des artères coronaires.

L'objectif principal de cette thèse est la caractérisation des tissus coronariens à partir d'images TCO pour l'évaluation automatique des pathologies coronariennes. Les expériences sont réalisées sur des acquisitions intracoronaires de TCO chez des patients atteints de la maladie de Kawasaki. Les objectifs spécifiques se détaillent comme suit: 1. Classification des couches de l'artère coronaire pour reconnaître les caractéristiques de chaque couche, intima et media. 2. Correction de mouvement comme une étape de la reconstruction 3D pour l'évaluation longitudinale et transversale de différentes formations pathologiques et l'estimation de la rigidité de la

paroi artérielle. 3. Identification des lésions de l'artère coronaire causées par KD sur les tissus de l'artère coronaire pour évaluer la fonctionnalité des artères coronaires.

Premièrement, des caractéristiques qui décrivent les couches intima et média ont été extraites automatiquement en utilisant un réseau neuronal convolutionnel. Les activations de la dernière couche entièrement connectée sont utilisées pour entraîner un classifieur par forêts aléatoires pour la tâche de classification. Ce travail contribue à évaluer l'épaisseur des couches de l'artère coronaire à distinguer entre les segments normaux et pathologiques de l'artère coronaire. En second lieu, un modèle de correction de mouvement d'images TCO intracoronaires est proposé pour aligner les tranches 2D en 3D de la TCO. Notre algorithme est conçu pour la correction de mouvement intra-tranche dans les images intravasculaires TCO en utilisant des informations sur les tissus plutôt que sur le contour du lumen. Les caractéristiques sont extraites automatiquement en appliquant un réseau de neurones convolutif et la similitude cosinus entre les caractéristiques profondes est utilisée pour effectuer l'enregistrement. Pour la troisième contribution, un classifieur a été entrainé à reconnaitre les différentes formations pathologiques des artères coronaires causées par la maladie de Kawasaki. Spécifiquement, les complications de l'artère coronaire les plus typiques telles que la fibrose, les macrophages, la néovascularisation et les calcifications ainsi que les couches de l'artère coronaire (intima et média) sont détectées en utilisant des caractéristiques extraites d'un réseau de neurones convolutifs préentraîné et un vote par consensus pour la classification finale. Cette étude pourrait contribuer à mieux comprendre le mécanisme de formation des maladies coronariennes et pourrait prévenir de futures complications chez les enfants et les jeunes adultes souffrant de la maladie de Kawasaki.

Mots-clés: Caractérisation tissulaire, Tomographie en Cohérence Optique (TCO), artère coronaire, réseau neuronal convolutionnel, maladie de Kawasaki, Forêt aléatoire, compensation du mouvement

CHARACTERIZATION OF CORONARY ARTERY TISSUES WITH KAWASAKI DISEASE FROM OCT IMAGES

Atefeh ABDOLMANAFI

ABSTRACT

Kawasaki Disease (KD), mucocutaneous lymph node syndrome, is an acute childhood vasculitis syndrome, which is characterized by fever, rash, bilateral nonexudative conjunctivitis, erythema of the lips and oral mucosa, and swollen erythematous hands and feet. KD is an inflammatory disease, which leads to inflammation in the walls of medium-sized arteries throughout the body. Although a high dose of Intravenous Immune Globulin (IVIG) infusion reduces the risk of coronary artery complications, about 15% to 25% of untreated children suffer a risk of experiencing coronary artery aneurysms or ectasia. Intimal thickening, media disappearance, lamellar calcifications, fibrosis, macrophage, and neovascularization are the most distinguished pathological features of late coronary artery lesions in Kawasaki disease. In severe cases, they can lead to myocardial infarction and sudden death. Since the functionality of the cardiac tissues significantly depends on the coronary blood flow to the myocardium, intravascular assessment of coronary artery tissues is significant to detect the pathological formations caused by different coronary artery complications.

Although in vivo intravascular visualization of coronary arteries is significant to provide highly valuable progressive information, it is a challenging task, especially in pediatric patients because of the small size of the vessels and high heart rate. OCT is an interferometric imaging modality that maps the backscattered near-infrared (NIR) light to create cross-sectional images of the tissues under review in micrometer scale. OCT was developed for the diagnosis and treatment guidance of coronary artery disease in the adult population. It has been recently used in pediatric cardiology to image coronary artery tissues with appropriate safety results in this age group. It has high resolution ranging from 10 to 20 μm to characterize the internal structure of the tissues such as vessel wall layers and plaque accumulation. Inner vessel wall geometry allows detecting and evaluating biophysical and dynamic properties of arterial wall, the thickness of coronary artery layers, and various coronary artery abnormalities caused by the disease.

This thesis is focused on developing an intra-coronary tissue characterization model using OCT imaging to pave the way for evaluating the functionality of coronary artery tissues. The experiments are performed on intracoronary OCT acquisitions from patients affected by Kawasaki disease. Analysis of coronary artery tissues is a broad study field, which consists of three main steps: 1. Classification of coronary artery layers to recognize characteristic attributes of each layer, intima and, media. 2. Identification of coronary artery lesions caused by KD on coronary artery tissues to assess the functionality of coronary arteries. 3. Motion correction as the step of 3D reconstruction for longitudinal and transversal assessment of different pathological formations and estimation of arterial wall stiffness.

For the first contribution, we developed an automatic classification approach to characterize coronary artery layers in pediatric patients using the images obtained from OCT system. The goal of the study was to identify the features, which perfectly describe intima and media layers using a Convolutional Neural Network (CNN). The activations of the last fully connected layer are used to train Random Forest (RF) for the classification task. This work contributes to evaluating the thickness of coronary artery layers to distinguish between normal and diseased segments of the coronary artery.

A motion correction model of intracoronary OCT images is proposed for the second contribution. Our algorithm is designed for intra-slice motion correction in intravascular OCT images using tissue information rather than the lumen outline. Features are extracted automatically by applying a Convolutional Neural Network and the similarity between deep features is used to perform registration. For the first time, deep learning is applied on intracoronary OCT images for motion correction. This will contribute to evaluate the functionality of coronary arteries by analyzing the volume variation and considering the motion of the vessel. Also, it is a robust method to assess the pathological formations by finding the correlation between the tissues of adjacent frames.

For the third contribution, we focused on developing a tissue characterization approach to classifying various pathological formations of coronary arteries caused by KD. Specifically, the most distinguished coronary artery complications such as fibrosis, macrophage, neovascularization, and calcification as well as coronary artery layers (intima, and media) are detected using deep features extracted from pre-trained CNNs and majority voting from Random Forest classification. This study contributes to preventing future complications in children and young adults suffered from Kawasaki disease. Since mentioned pathological formations are recognized as the most common intracoronary complications caused by coronary artery disease (CAD), this work is not limited to intracoronary tissue characterization in KD patients.

Keywords: Tissue characterization, Optical Coherence Tomography (OCT), Coronary artery, Convolutional Neural Networks (CNNs), Kawasaki Disease (KD), Random Forest (RF), Motion correction

TABLE OF CONTENTS

Page

LIST OF TABLES

Page

LIST OF FIGURES

LIST OF ABREVIATIONS

KD	Kawasaki Disease
CA	Coronary Artery
CAD	Coronary Artery Disease
CAA	Coronary Artery Aneurysm
ACS	Acute Coronary Syndrome
MI	Myocardial Infarction
IVIG	Intravenous Immune Globulin
NO	Nitric Oxide
NIR	Near Infrared
DICOM	Digital Imaging and Communications in Medicine
OCT	Optical Coherence Tomography
3D-OCT	Three Dimensional Optical Coherence Tomography
FD-OCT	Frequency-Domain Optical Coherence Tomography
TD-OCT	Time-Domain Optical Coherence Tomography
A-Scan	Amplitude Scan
B-Scan	Brightness Scan
IVUS	Intra-Vascular Ultra-Sound
CT	Computed Tomography
CTPA	CT Pulmonary Angiography

CIMT	Carotid Intima-Media Thickness
ECG	Echocardiography
CMR	Cardiac Magnetic Resonance
CNNs	Convolutional Neural Networks
ReLU	Rectified Linear Unit
SGD	Stochastic Gradient Descent
FC	Fully Connected
RF	Random Forest
OOB	Out Of Bag
CART	Classification And Regression Trees
SVM	Support Vector Machine
RBF	Radial Basis Function
C-SVC	C-Support Vector Classifier
ROI	Regine Of Interest
CC	Connected Components
AC	Active Contour
RBC	Red Blood Cells
2D	Two Dimensional
3D	Three Dimensional
4D	Four Dimensional

SD	Standard Deviation
ED	Euclidean Distance
VSMC	Vascular Smooth Muscle Cell
STFT	Short-Time Fourier Transform
RVM	Relevance Vector Machine
LDA	Linear Discriminant Analysis
OA	Osteoarthritis
FMM	Fast Marching Method
PE	Pulmonary Embolism
EM sensor	Electromagnetic sensor

LISTE OF SYMBOLS AND UNITS OF MEASUREMENTS

L	Loss function in Convolutional Neural Networks (CNNs)
X	Size of the training set in CNNs
y	Feature labels
j	Number of images
i	Number of iterations
μ	Momentum (CNNs)
γ	Scheduling rate (CNNs)
α	Learning rate (CNNs)
W	Weights at each layer of CNN architectures
s	Strength of each tree (Random Forest)
ρ	Correlation between trees (Random Forest)
γ	Gaussian kernel bandwidth (SVM tuning parameter)
C	Regularization (SVM tuning parameter)
ρ	Correlation between trees (RF tuning parameter)
M_{try}	Number of randomly selected predictors (RF tuning parameter)
μm	Unit used for the measurements of tissue layer thickness
μm	Unit used for the spatial resolution of OCT system
nm	Unit used for the wavelength of the light source
mm	Unit used for penetration depth

m/sec	Unit used for the speed of light
mm/sec	Unit used for pullback speed

INTRODUCTION

Kawasaki disease (KD) is an acute childhood disease characterized by six principal symptoms such as fever that may last for five days, bilateral conjunctival congestion, changes of lips and oral cavity, which they specifically appear as redding of lips, strawberry tongue, diffuse injection of oral and pharyngeal mucosa; polymorphous exanthema, changes in the extremities, which they appear as redding of palms and soles and indurative edema in acute phase of the disease, and membranous desquamation from fingertips in convalescent phase; and cervical lymphadenopathy (see Figure 0.1). Patients with five of the principal symptoms are diagnosed as KD. If four of the principal symptoms are recognized in some patients and they are followed by the coronary aneurysm or dilatation using echocardiography or coronary angiography, these group of patients also are diagnosed as KD (Newburger *et al.*, 2004; Group, 2005; Group *et al.*, 2010). There are other clinical symptoms, which are important to be considered during the evaluation and follow up of the disease. These symptoms are categorized as cardiovascular changes, gastrointestinal tract, changes in blood factors, urine, skin, respiratory, and joints as well as neurological variations. The prevalence of the disease is about 180/100000 in children and infants. Around 80% to 85% of children, who are affected by KD is at the age of 0 to 4 years old. It is more prevalent in male than the female with the ratio of 1.3-1.5 boys:1 girl.

KD is an acute febrile disease with histopathological features of self-limited vasculitis (Hauser *et al.*, 2004; Newburger *et al.*, 2004). During the acute phase of the disease, as the initial therapy, patients receive Intravenous Immunoglobulin (IVIG). For patients who are not responding to initial IVIG therapy, additional IVIG treatment, steroid or combination of them is required. In 15-25% of untreated children, there is a risk of aneurysmal development, which is a serious complication of KD and may lead to coronary artery disease (CAD) and sudden death (Newburger *et al.*, 2004). Pathological analysis of coronary arteries in KD demonstrates arterial inflammation followed by intimal hyperplasia, media disappearance, neovascularization, fibrosis, calcification, and macrophage accumulation (Newburger *et al.*, 2004; Group *et al.*, 2010).

Figure 0.1 Symptoms of Kawasaki Disease (KD): rash, bilateral nonexudative conjunctivitis, erythema of the lips and oral mucosa, and swollen erythematous hands and feet.

Aneurysms can be filled in with the intimal fibrosis, which results in the development of arterial wall stiffness while the lumen diameter is normal. In some cases, although the aneurysms are treated and the lumen diameter is normal, but the coronary artery may represent less distensibility that can affect myocardial microcirculation (Hauser *et al.*, 2004; Newburger *et al.*, 2004). Cardiac function depends strongly on the coronary blood flow to the myocardium and subsequently the functionality of cardiac tissues.

In cross-sectional view, healthy coronary artery is characterized by three distinguished layers (see Figure 0.2). Normal layers of arterial wall, intima, and media, are characterized by OCT as signal-rich and signal-poor patterns respectively with the normal thickness of 61.7 ± 17.0 μm for intima and 61.4 ± 16.7 μm for media. Two types of intimal hyperplasia are visualized in KD. Intimal hyperplasia with preserved media occurs when the three-layered structure of the arterial wall is preserved, though intimal thickening is observed. In severe cases, intimal thickening is followed by media destruction, which is known as intimal hyperplasia with media disappearance. Neovascularization is also known as neointimal hyperplasia caused by the proliferation of the smooth muscle cells from the media. Neointimal hyperplasia is visualized as

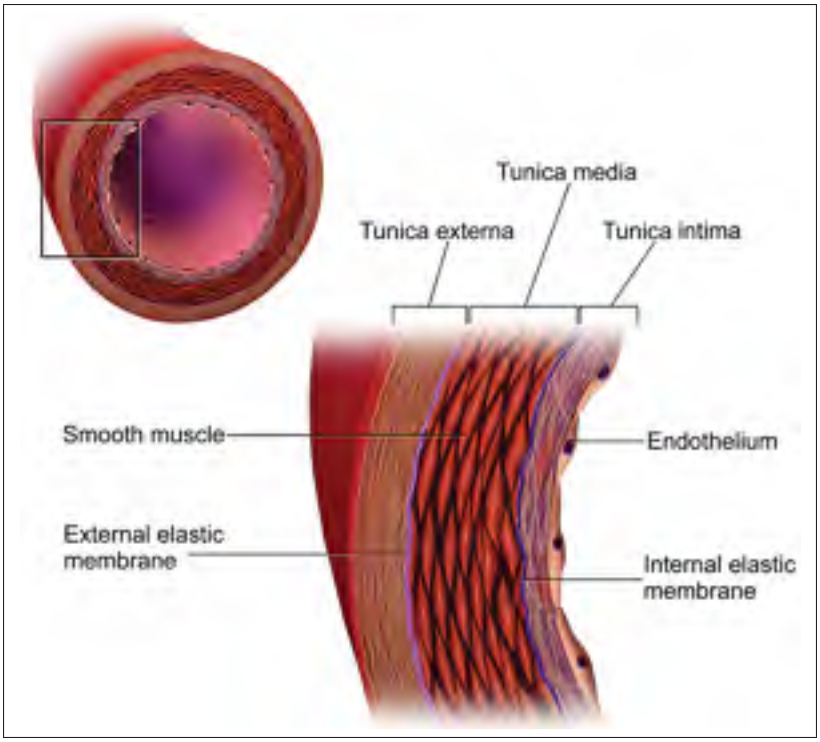


Figure 0.2 Coronary artery structure in cross-sectional view.

a signal-rich pattern followed by a dark appearance of the micro-vessels in OCT images. Calcifications are characterized as a well-delineated signal-poor region with clear borders in OCT images.The macrophage is the excessive absorption of glucose by inflammatory immune cells, which can reside in arterial plaques. It is characterized as an underlying signal-poor region in OCT images (Dionne *et al.*, 2015). In vivo intravascular visualization of coronary arteries is crucial as a part of follow up in patients with the history of coronary artery abnormalities to assist the physicians by increasing the efficiency of treatments (Orenstein *et al.*, 2012).

Various conventional cardiac imaging modalities such as echocardiography (ECG), computed tomography (CT), X-ray angiography, cardiac magnetic resonance imaging (CMR) are used to assess the functionality of coronary arteries. Although the mentioned imaging techniques are used to characterize the size of aneurysms and luminal diameter of the coronary artery segment affected by the disease, they are limited to provide information on the underlying

coronary tissue layers. Also, they are restricted to reflect the histological reality of the regressed aneurysmal coronary segments, which are inappropriately considered as normal coronary segments (Orenstein *et al.*, 2012; Newburger *et al.*, 2004; Group *et al.*, 2010; Dionne *et al.*, 2015). Catheter-based Intravascular Ultrasound (IVUS) has been used for many years in interventional cardiology to evaluate coronary artery tissues by providing information on coronary arterial wall and lumen (Rathod *et al.*, 2015). IVUS imaging is restricted to be used in pediatric cardiology due to its suboptimal spatial imaging resolution (100-150 μm), and low pullback speed. In contrast, Optical Coherence Tomography (OCT) is a new intravascular imaging modality with high spatial resolution, which is primarily used in adult cardiology for treatment of coronary artery disease (Tearney *et al.*, 2012; Yabushita *et al.*, 2002). Considering high spatial resolution of OCT ranging from 10 to 20 μm to characterize coronary artery tissues and demonstrating safety results in pediatric patients, it has been recently used in pediatric cardiology to evaluate structural variations in coronary arteries in patients suffering from Kawasaki disease and it has been demonstrated progressive diagnostic value among the conventional imaging systems in this population (Ulrich *et al.*, 2017; Dionne *et al.*, 2015; Mitani *et al.*, 2009; Harris *et al.*, 2014).

Since OCT is a new technology recently used in pediatric cardiology, and KD can be followed by severe coronary artery complications in children and young adults, quantitative and qualitative characterization of different coronary artery tissues should be supported by automated mathematical and technical algorithms based on the clinical knowledge reported by clinicians regarding the histological and pathological features of the disease. Consequently, this thesis aimed to develop an automated intracoronary tissue characterization model in patients with the history of KD using OCT imaging.

The thesis is structured in four chapters (see Figure 0.3). The importance of using OCT system to study and evaluate coronary arteries is explained in **Chapter 1**. The existing techniques in

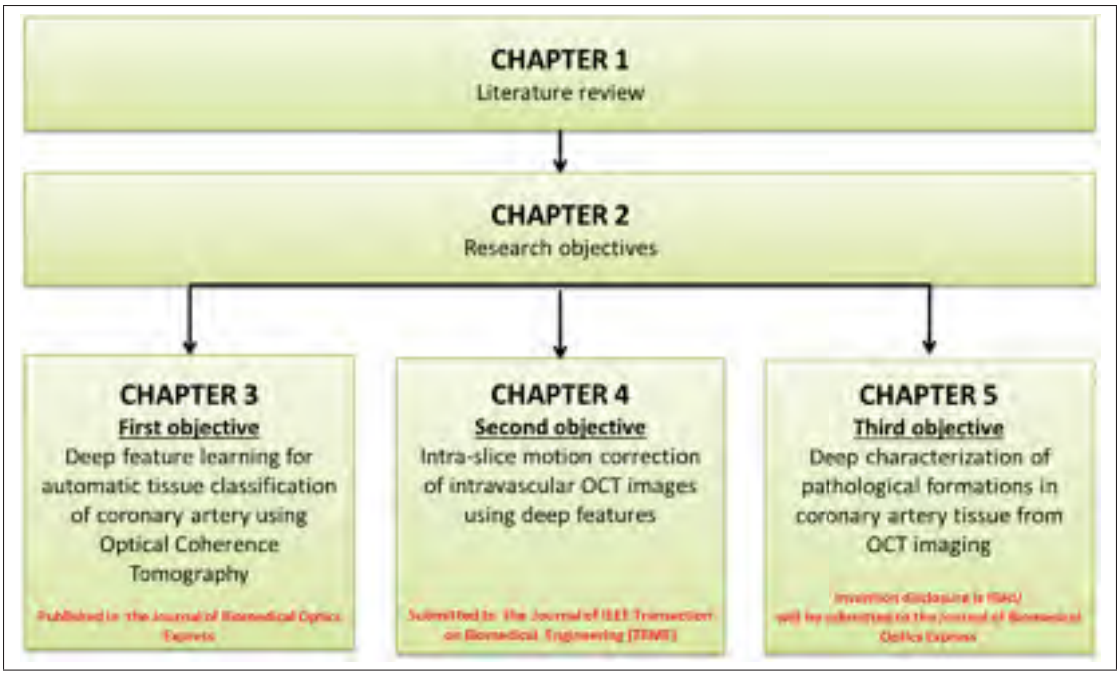


Figure 0.3 Thesis structure

the literature related to each objective are discussed in this chapter to demonstrate the significant contributions of our research project.

In **Chapter 2**, the objectives are defined by investigating, observing, and understanding the research problems as well as the clinical knowledge behind the tissues under review.

Chapter 3 is focused on the automatic classification of different coronary artery layers, namely intima and media to distinguish between normal and diseased coronary artery segments by characterizing the appropriate attributes of each layer of the arterial wall. Feature extraction is performed using a Convolutional Neural Network (CNN) and the performance of different classifiers are compared against each other to design an appropriate model of tissue classification, which is a step of characterization of different pathological formations and structural variations of diseased coronary artery segments.

Although 2D cross-sectional OCT images of coronary arteries play a crucial role to estimate tissue variations of coronary artery segments affected by KD, the accurate assessment of pathological formations may obtain by considering the adjacent frames and analyzing the volume variation of each tissue. Motion correction using 2D cross-sectional OCT images of coronary arteries is performed in **Chapter 4**. This work may add incremental steps to our tissue characterization model and also pave the way for evaluation of arterial wall stiffness caused by pathological formations on the diseased segments of the arterial wall.

In **Chapter 5**, an intracoronary tissue characterization model is proposed to classify the most significant intracoronary pathological formations caused by coronary artery disease as well as coronary artery layers (intima, and media). The possible solutions of applying the state-of-the-art pre-trained networks in the most efficient way for the application on OCT images is discussed in this work. Also, this study demonstrates the application of deep feature learning in pediatric cardiology to detect coronary artery complications caused by Kawasaki disease.

CHAPTER 1

LITERATURE REVIEW

Intravascular visualization of coronary arteries is promising to assess the functionality of coronary artery tissues and evaluate different pathological formations due to various coronary artery disease.

1.1 OCT imaging

Cardiovascular OCT is a catheter-based invasive imaging modality, which typically employs a near-infrared light to provide cross-sectional images of the coronary artery at depth of several millimeters relying on low-coherence interferometry. The unique characteristic of the OCT is its high axial resolution of 10-15 μm, which is measured by the light wavelength and is decoupled from the lens dependent lateral resolution ranging from 20-40 μm (Ferrante *et al.*, 2013). The OCT system is composed of an optical imaging engine, computer signal acquisition, and the imaging catheter as it is shown in Figure 1.1.

The probe is inserted into the coronary artery using an over-the-wire balloon catheter from patient's groin. A sequence of cross-sectional images of coronary artery segment is recorded using the backscattered light from the arterial wall through each pullback (see Figure 1.2). Considering the fact that light can be attenuated by blood before reaching the vessel wall, blood clearance is required before starting the image acquisition (Bezerra *et al.*, 2009; Boudoux, 2016; Drexler & Fujimoto, 2015).

1.1.1 Physical principles of the OCT system

A near-infrared spectrum of light with the wavelength between 1250 to 1350 nm is applied in OCT system. Using longer wavelength, we could have deeper penetration depth through the tissues under review; but in clinical purposes, the wavelength of the light source is selected considering the characterization of tissue absorption and other determining factors depending

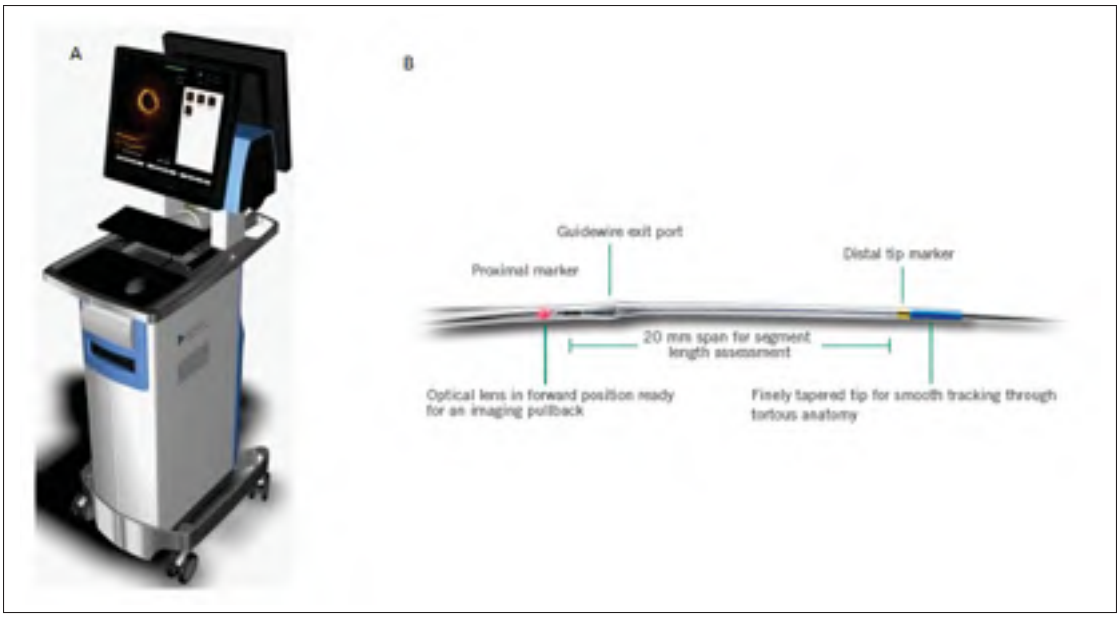


Figure 1.1 Optical imaging engine and computer signal acquisition, B) OCT imaging probe.

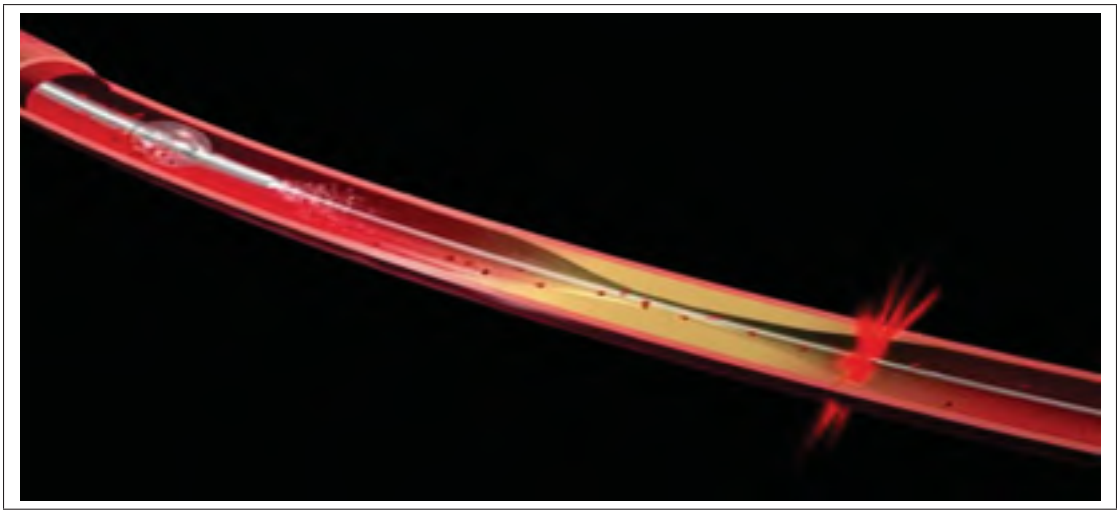


Figure 1.2 OCT pullback: A sequence of cross-sectional images of coronary artery segment is recorded using the backscattered light from the arterial wall through each pullback.

on the texture under review, such as the refractive index of the interface between the catheter and the vessel wall. In common OCT systems, a wavelength of 1300 nm is employed. Therefore, the penetration depth is restricted to 1 to 3 mm. Considering the high speed of light,

3×10^8 m/s, it is impossible to perform all the measurements directly and the backscattered signal should be measured using interferometry method in this short period of time (Bezerra *et al.*, 2009). A low-coherence light beam emitted from a light source is received by a beam splitter and is divided into two equal parts. One part passes through a reference mirror and reflects back to the splitter, the other part moves toward the sample. According to the optical characteristics of the sample, some portions of the incident light may be absorbed, deflected or reflected because light velocity may have variable values moving across various media. The reflected light from the sample reaches the splitter. Two reflected light beams with equal wavelengths and invariable phase difference interfere with each other and this is the cornerstone of creating an image by OCT. Constructive interference occurs if two reflected waves are in phase. For each particular area of the sample, one pixel is recorded by the detector using dark and bright patterns (Hamdan *et al.*, 2012). As a result of intracoronary pullbacks, raw data is obtained from the OCT scanner. The amplitude scan (A-Scan) is the signal obtained from a single axial scan. Brightness scan (B-Scan) is the cross-sectional image which is acquired as a result of moving the beam over the sample (see Figure 1.3).

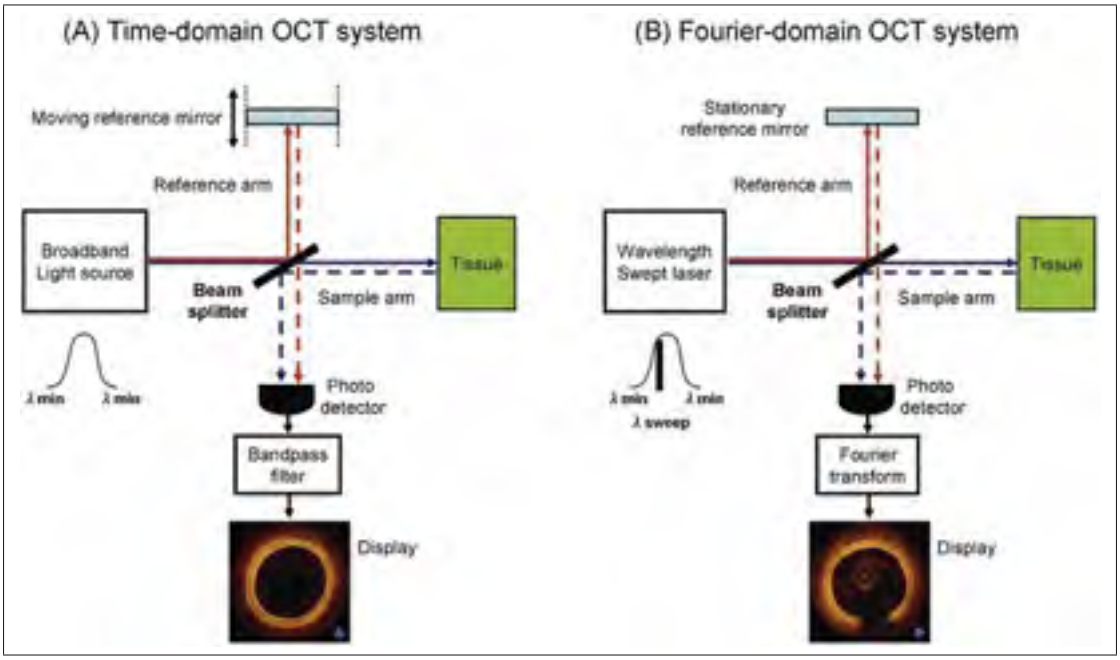


Figure 1.3 Image acquisition using OCT system.

Attenuation compensation along the OCT A-Scan is performed to improve the contrast of the intracoronary OCT images, enhance the visibility of the deep tissue structures, particularly tissue borders, and ameliorate the shadow artifacts. The enhanced cross-sectional image is acquired by polar reconstruction of the processed data (Foin *et al.*, 2013).

Theory of the Michelson interferometer

Interferometry calculates the field of the light rather than its intensity (Fercher *et al.*, 2003). Considering the Michelson interferometer, the incident light is a polychromatic plane wave with the following electric field:

$$E_i = S(k,\omega)e^{(kz-\omega t)} \tag{1.1}$$

Where s(k,ω) is the spectral distribution of the light source, k is wavenumber, and ω is the angular frequency. For different media, we have various indexes of refraction which are related to wavelength as below:

$$c/n(\lambda) = \lambda\nu \tag{1.2}$$

The beam splitter is assumed to have the splitting ratio of 0.5. The incident light is split into a reference beam $E_R(t)$ and the sample beam $E_S(t)$, they pass various distances between the two arms of the interferometer and they return to the beam splitter. The electric field reflected back from the reference mirror to the splitter is expressed by the following equation:

$$E_R = E_i/\sqrt{2}r_R e^{i2kz_R} \tag{1.3}$$

Where r_R is the electric field reflectivity of the reference mirror, and z_R is the position of reference mirror from the beam splitter. The factor 2 in the exponential term is entered because of the round-trip path length and $1/\sqrt{2}$ is the normalization term. The reference reflectivity power is defined as follows:

$$R_R = \|r_R\|^2 \tag{1.4}$$

Depending on the different penetration depth that light passes through the sample, we obtain various backscattered light from the sample which is characterized by their field reflectivity as follows:

$$E_s = E_i/\sqrt{2}[r_s(z_s) \circledast e^{(i2kz_s)}] \tag{1.5}$$

If the reflections are assumed as a series of N discrete real delta-function reflections, then we will have:

$$r_s(z_s) = \Sigma_{n=1}^{N} r_{s_n} \delta(z_s - z_{s_n}) \tag{1.6}$$

Equation (1.6) shows that each reflection is specified by its electric field reflectivity, r_{s_n} for n reflections, and the path from the splitter, z_{s_n} for n reflections. The power reflectivity is defined as follows:

$$R_{s_n} = \|r_{s_n}\|^2 \tag{1.7}$$

By replacing $r_s(z_s)$ from equation (1.6) in equation (1.5), we obtain the electric field of the sample as follows:

$$E_s = E_i/\sqrt{2}[\Sigma_{n=1}^{N} r_{s_n} e^{(i2kz_{s_n})}] \tag{1.8}$$

Finally, the reference beam and the sample beam interfere at the detector, which results in the following equation:

$$I(k,\omega) = \rho/2 < \|E_R + E_s\|^2 > = \rho/2 < (E_R + E_s)(E_R + E_s)^* > \tag{1.9}$$

Where ρ is the response factor of the detector. By replacing E_R and E_s from equations (1.3), and (1.5) and ignoring the temporal angular frequency, ω, in the above equation since the frequency oscillates much faster than the response time of the detector:

$$\begin{aligned} I(k) &= \rho/4[S(k)(R_R + R_{s_1} + R_{s_2} + ...)] \\ &+\rho/4[S(k)\Sigma_{n=1}^{N}\sqrt{R_R R_{s_n}}(e^{i2k(z_R - z_{s_n})} + e^{-i2k(z_R - z_{s_n})}] \\ &+\rho/4[S(k)\Sigma_{m=1}^{N}\sqrt{R_{s_n} R_{s_m}}(e^{i2k(z_{s_n} - z_{s_m})} + e^{-i2k(z_{s_n} - z_{s_m})}] \end{aligned} \tag{1.10}$$

Gaussian-shaped light spectrum is one of the best choices of the light source to model the OCT system. Using Gaussian light source with the central wavenumber of k_0 and the spectral bandwidth at half width of 1/e of its maximum (Δk), equation (1.10) can be simplified as follows:

$$\begin{aligned} I(k) &= \rho/4[S(k)(R_R + R_{s_1} + R_{s_2} + ...)] \\ &+\rho/2[S(k)\Sigma_{n=1}^{N}\sqrt{R_R R_{s_n}}cos(2k(z_R - z_{s_n}))] \\ &+\rho/4[S(k)\Sigma_{m=1}^{N}\sqrt{R_{s_n} R_{s_m}}cos(2k(z_{s_n} - z_{s_m}))] \end{aligned} \tag{1.11}$$

The first, second, and fourth terms of equation (1.11) are respectively called DC term, cross-correlation term, and auto-correlation term. DC term is a constant offset of the detector current and is independent of the path length. Cross-correlation term is the main component of the detector current which is desired for OCT imaging, and the last term is obtained as a result of the interference among various sample reflectors and leads to the artifacts of the OCT system. The main artifacts of the OCT system will be discussed later in this section. Obviously, it is understood from equation (1.11) that if a single reflector is considered, the auto-correlation term will not appear in the response of the detector (Drexler & Fujimoto, 2008; Fercher *et al.*, 2003).

Time-Domain OCT and Frequency-Domain OCT

In Time-Domain OCT (TD-OCT), a fiber-optic coupler is applied as the light splitter. In this system, an ambulant mirror is used as the reference arm. The reflected signals of the sample and the mirror have interfered and the electric field amplitude will be measured. Using TD-OCT in interventional cardiology has some limitations. The scanning process reduces the image acquisition rate. On the other hand, soft balloon inflation is used to prevent coronary blood circulation during the intervention process. Also, studying the proximal lesions and left main coronary artery is restricted by using the TD-OCT (Hamdan *et al.*, 2012). In Frequency Domain OCT (FD-OCT), the mirror is fixed. Therefore, the reference distance will be constant while the frequency of the light source varies during time. This new generation of OCT is

fast to produce the images considering the fact that detections at different penetration depths perform at the same time. There are two types of FD-OCT, Fourier-domain OCT, and swept-source OCT, which are different on their procedure to derive data from the interferometer. FD-OCT with the pullback speed of 20 mm/sec is 10-20 times faster than TD-OCT (Ferrante *et al.*, 2013).

OCT imaging artifacts

Image acquisition in realistic clinical imaging conditions leads to different OCT artifacts including residual blood, which can reduce the tissue brightness, non-uniform rotational distortion as a result of changes in optical fiber rotational speed, sew-up artifact caused by fast movements of imaging wire or artery, artifacts related to eccentric wire position, side branches and large vessels (Bezerra *et al.*, 2009; Herrero-Garibi *et al.*, 2010). Figure 1.4 illustrates ten common artifacts of the OCT imaging.

1.1.2 Clinical applications of OCT

At first, OCT was proved in retinal imaging and then it was developed for other medical and surgical proficiencies including gastroenterology, dermatology, cardiology, and oncology. Technological progress such as improvement of laser sources, development of beam transfer instruments and image detection have caused extensive clinical uses of OCT (Drexler & Fujimoto, 2008). Intravascular imaging modalities, such as OCT allow nowadays improving diagnosis, treatment, follow-up, and even prevention of coronary artery disease in the adult and children. At first, OCT is used in adult cardiology particularly to detect fibrosis, calcification, neovascularization, and intimal hyperplasia, which are the most significant histological lesions of coronary atherosclerotic disease in adults (Tearney *et al.*, 2012; Yabushita *et al.*, 2002). OCT is also used as stent positioning guidance. Also, it is accurate to detect the regressed aneurysmal coronary segments, which are considered as normal in X-ray angiography. Considering the high spatial resolution of OCT, high image acquisition rate, the feasibility and safety of OCT in pediatric population, it has been recently used in the pediatric cardiology to

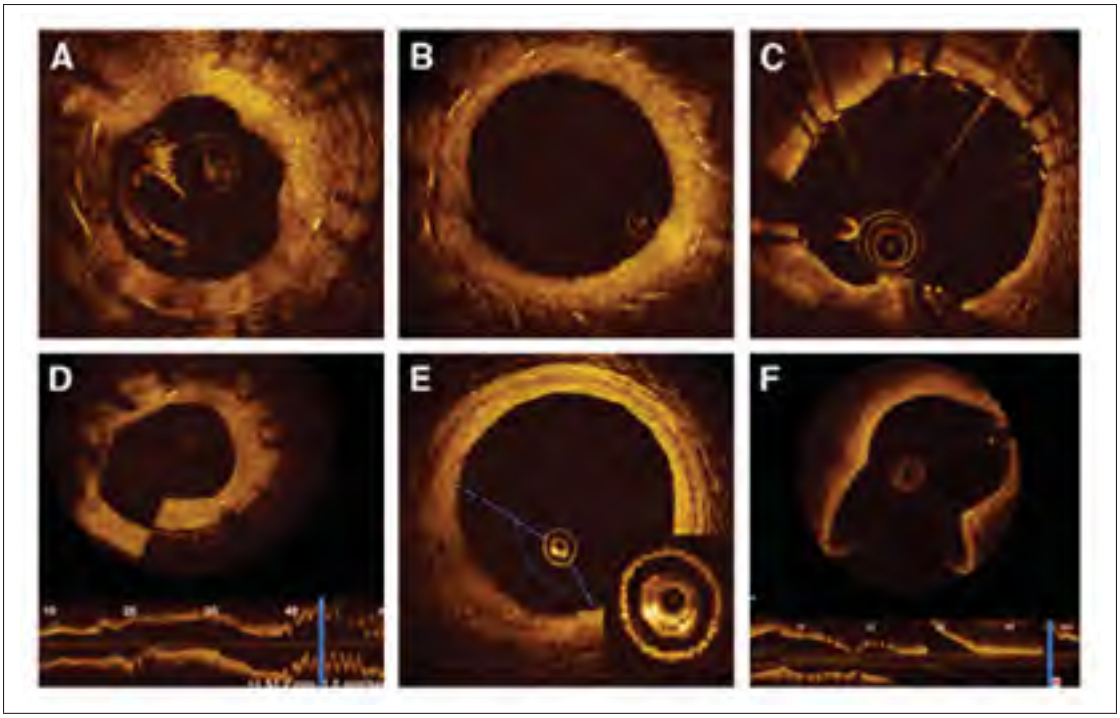


Figure 1.4 Common OCT artifacts have displayed in the cross-sectional image of the coronary artery: Residual blood, eccentric wire position, saturation, sew-up, bubble and fold-over artifacts are shown respectively from A to F.

evaluate structural variations of the coronary arterial wall in the segments of the artery, which are affected by disease (Fujino *et al.*, 2014; Kakimoto *et al.*, 2014; Dionne *et al.*, 2015).

Automated tissue characterization of diseased coronary artery segments is a challenging task considering the small size of coronary arteries in children and infants, and the complications caused by KD on coronary artery tissues. This chapter aims to review the recent studies related to our research project to demonstrate the scientific contribution of our research considering the strengths and limitations of the existing approaches.

1.2 Review of the tissue segmentation and classification approaches

1.2.1 CA segmentation methods using IVUS imaging

Intravascular Ultrasound (IVUS) has been used for many years in interventional cardiology to evaluate coronary artery tissues by providing information on coronary arterial wall and lumen (Rathod *et al.*, 2015). Many studies are focused on segmentation of coronary artery tissues using IVUS imaging. The automated technique for detection of lumen and media-adventitia borders for quantitative assessment of atherosclerosis in IVUS images was proposed by (Papadogiorgaki *et al.*, 2008) The technique is based on multilevel discrete wavelet using initialized contours. (Mendizabal-Ruiz *et al.*, 2013) proposed a computational method for detection of the luminal border using IVUS images. The method is based on the minimization of a probabilistic cost function considering the likelihood of the pixels belonging to the blood and non-blood regions. Support Vector Machine (SVM) is used to perform the classification. A segmentation approach based on the fast-marching method (FMM) in the context of intravascular ultrasound (IVUS) imaging is developed by (Destrempes *et al.*, 2014). Shape-driven approach for segmentation of the arterial wall from intravascular ultrasound images in the rectangular domain is proposed to detect luminal and media-adventitia borders (Unal *et al.*, 2008).

1.2.2 Classification of CA tissues using OCT imaging

IVUS imaging is restricted to be used in pediatric cardiology due to its suboptimal spatial imaging resolution (100-150 μm), and low pullback speed. OCT is recently used in pediatric cardiology. It has high resolution ranging from 10 to 20 μm to characterize the internal structure of tissues such as vessel wall layers and plaque accumulationFerrante *et al.* (2013). Automatic lumen segmentation to assess the stenosis grading and characterization of the plaque types in OCT images of coronary arteries have been performed by (Celi & Berti, 2014). (Yabushita *et al.*, 2002) have proposed a method of plaque characterization by correlating OCT images with histology. Other studies focused on atherosclerosis plaque characterization using optical properties of tissues (Levitz *et al.*, 2004). The actual backscattering and attenuation coeffi-

cients were measured by (Xu *et al.*, 2008). The automatic quantification of optical attenuation coefficients has been proposed by (Van Soest *et al.*, 2010). (Ughi *et al.*, 2013) have proposed a classification method using texture features and attenuation coefficient to characterize atherosclerosis tissues. Volumetric estimation of backscattered intensity and attenuation coefficient is performed by (Gargesha *et al.*, 2015). SVM is used to classify between fibrosis, calcification, and lipid. Identification and quantification of fibrous tissue based on Short-Time Fourier Transform (STFT) using OCT imaging are proposed by (Macedo *et al.*, 2016). (Gan *et al.*, 2016) developed a classification framework to detect normal myocardium, loose collagen, adipose tissue, fibrotic myocardium, and dense collagen. Graph searching method is applied to segment various tissue layers of the coronary artery. Combination of texture features and optical properties of tissues are used to train Relevance Vector Machine (RVM) to perform the classification task. A plaque tissue characterization technique based on intrinsic morphological characteristics of the A-lines using OCT imaging is proposed by (Rico-Jimenez *et al.*, 2016) to classify superficial-lipid, fibrotic-lipid, fibrosis, and intimal thickening by applying Linear Discriminant Analysis (LDA).

To the best of our knowledge, these studies did not address the challenging task of classification of coronary artery layers, particularly in pediatric patients. There is no study in the literature focused on characterization of all pathological formations caused by coronary artery disease, which is a challenging task since the model should be able to discriminate between various tissue formations as well as normal coronary tissue layers.

1.2.3 Application of CNNs in medical image analysis

The recent applications of CNNs in medical image analysis include pancreas segmentation using CT images of the abdomen (Roth *et al.*, 2015), classification of pulmonary peri-fussural nodules (Ciompi *et al.*, 2015), and brain tumor segmentation (Havaei *et al.*, 2016). The strength of CNNs originates from its deep structure, which allows feature extraction from various abstraction layers (Szegedy *et al.*, 2015; Simonyan & Zisserman, 2014; Zeiler & Fergus, 2014; Eigen *et al.*, 2013). Basically, all CNNs consist of a series of layers defined by a specific

number of filters or kernels that mainly have the role of feature detector from a set of input images. Sliding the filters on the input images and calculating the convolution of these filter matrices and the input image matrix creates a set of convolved features or simply feature maps. The exact meaning of learning a CNN is to train a CNN architecture by the values of these convolutional operations (Hochreiter & Schmidhuber, 1997).

The transferability of the information preserved on pre-trained CNNs which is one of the most significant characteristics of CNNs, has been demonstrated by the work of (Azizpour *et al.*, 2015). The recent studies show the significant applications of transfer learning in medical image analysis to extract features from a new dataset using pre-trained CNNs or to use CNNs as the classifier by fine-tuning a pre-trained network (van Ginneken *et al.*, 2015; Bar *et al.*, 2015; Arevalo *et al.*, 2015; Chen *et al.*, 2015; Carneiro *et al.*, 2015; Gao *et al.*, 2016; Margeta *et al.*, 2015). (van Ginneken *et al.*, 2015) have used the last layer of OverFeat pre-trained CNN as feature extractor for pulmonary nodules detection in CT acquisition. Then, extracted features have been used to train SVM for the classification task. Other studies used pre-trained CNNs as feature extractors on various medical image datasets such as chest radiograph data in the work of (Bar *et al.*, 2015) and mammography images in the work of (Arevalo *et al.*, 2015). Recently, pre-trained networks have been also used for classification tasks. For instance, (Chen *et al.*, 2015) applied pre-trained CNNs on ImageNet dataset to detect the fetal abdominal standard plane in ultrasound images by keeping the low-level representations extracted from natural images and modifying the parameters of the last layers based on the characteristics of ultrasound images. The application of transfer learning using pre-trained CNNs in medical image analysis has been shown in other studies (Carneiro *et al.*, 2015; Gao *et al.*, 2016; Margeta *et al.*, 2015). Also, it has been demonstrated in the work of (Tajbakhsh *et al.*, 2016) that using pre-trained CNNs with adequate fine-tuning works better than training a CNN from scratch in medical image analysis applications. In their experiments, they considered different categories of medical images in radiology, cardiology, and gastroenterology using different medical imaging systems such as colonoscopy images for polyp detection, CT pulmonary angiography (CTPA) for PE

diagnosis, and carotid intima-media thickness (CIMT) which is a noninvasive ultrasonography method in cardiology.

Nevertheless, most of the studies are focused on fine-tuning the available pre-trained networks to perform the classification task. How to use pre-trained networks in an efficient way to overcome the convergence issues, overfitting concerns, and long computational time is not mentioned in the literature. Choosing a classifier with robust classification performance, while it is trained on deep features is not reported in the literature. To the best of our knowledge, the application of deep feature learning on intracoronary tissue classification to detect pathological formations caused by coronary artery disease as well as tissue layers is not proposed in previous studies.

1.3 Review of motion correction and 3D reconstruction techniques

Since the OCT probe moves freely in coronary artery pathway, cross-sectional images might be misaligned. This is problematic to evaluate longitudinally each tissue and to design robust model for clinical measurements. Since OCT is recently used in cardiology, motion correction methods are mostly focused on intracoronary IVUS images and OCT retinal images.

Intravascular Ultrasound (IVUS) has been used for many years as an intracoronary imaging modality in cardiology. Therefore, many of the motion correction and 3D reconstruction methods are focused on IVUS images. (Wahle *et al.*, 1998) developed an image fusion technique to create the 3D reconstruction of intracoronary IVUS images. (Cothren *et al.*, 2000) proposed a 3D reconstruction technique by detecting the 3D trajectory of the IVUS transducer using the information of the angiographic images. It is demonstrated that the correct rotational orientation of IVUS images on the angiogram is based on the function of time by applying best fit angle function. To increase the accuracy of the 3D reconstruction, (Bourantas *et al.*, 2005) extracted the catheter path from biplane angiography. Also, the IVUS images and angiogram are visualized simultaneously using ECG wave. (Zheng, 2009) focused on 3D reconstruction method of intracoronary IVUS images by reconstructing the pullback path using snake algorithm. The

precision of the 3D reconstruction method is increased by focusing on 3D axial position, spatial orientation, and surface fitting. 3D artery is registered with intracoronary IVUS cross-sections using distance mapping algorithm in the work of (Tu *et al.*, 2011). Another study is focused on 3D reconstruction of IVUS images using the biplane angiography to detect the 3D centerline. The IVUS frames are distributed along the reconstructed centerline in equivalent time intervals (Ma *et al.*, 2013). (Athanasiou *et al.*, 2016) proposed a 3D reconstruction of intracoronary IVUS images by estimating the lumen and the outer vessel wall borders using the approximation of the arterial centerline to evaluate plaque formations. (Karlas & Lee, 2015) proposed a 3D reconstruction method based on the fusion of IVUS and preoperative CT data. The alignment of the detected lumen borders is performed using the ellipse fitting technique and CT data is used to detect the arterial shape. (Zhao *et al.*, 2016) developed a 3D reconstruction method using EM sensor to detect the catheter pose. Lumen contours are detected by applying a radial scan method.

A motion correction algorithm based on A-scans was proposed for retinal OCT images by (Kraus *et al.*, 2012). They have also extended their method to 3D-OCT motion correction using image registration and orthogonal raster scan patterns (Kraus *et al.*, 2014). (Braaf *et al.*, 2013) proposed a method using the combination of inter-B-scan phase-resolved OCT angiography with real-time eye tracking. (Lee *et al.*, 2011) have devised an algorithm relying on cross-correlation maximization for motion compensation of rodent cerebral cortex.(Wang *et al.*, 2015) are focused on motion tracking using speckle decorrelation of OCT signal by evaluating the sensitivity of speckle motion tracking using the derivative of cross-correlation coefficient.

Some other studies are focused on OCT images of coronary arteries. 3D reconstruction of coronary artery images is performed by (Ellwein *et al.*, 2011) using graph theory applied on computed tomography (CT) and OCT data of a single patient after stent placement. (Athanasiou *et al.*, 2012) proposed a semi-automated 3D reconstruction method using OCT images and biplane angiography.

Very few studies addressed the problem of non-rigid, intra-slice motion correction in intravascular OCT images using tissue information, rather than the lumen outline. Also, the application of deep features for registration problem is not mentioned in the literature. Considering the physical principals of the imaging system, the motion correction technique should be adapted to the artifacts, challenges, and the physical properties of the imaging system as well as the attributes of the acquired images.

Research problems are defined based on the clinical demands, which are mentioned by expert cardiologists. The main and specific research objectives are defined according to the research problems. For each specific objective, the models are designed by considering the physics of the imaging system, image properties, strengths and limitations of the related studies, and possible improvements of the proposed techniques in the literature.

CHAPTER 2

RESEARCH OBJECTIVES

2.1 Problem statement and Research objectives

Considering the significant role of coronary arteries in the functionality of cardiac tissues, coronary artery disease (CAD) is known as the main cause of myocardial infarction (MI) and sudden death. Progression of pathological formations caused by coronary artery disease can be followed by acute coronary syndrome (ACS). Therefore, the main goal of this research is **to design an intracoronary tissue characterization model using OCT imaging by finding the appropriate feature extraction, classification, and motion correction techniques.** To achieve the main goal of this research, we defined three specific objectives. Each specific objective is a complementary step of the main objective, which contributes to solving a set of problems involved in this project.

Characterization of the normal coronary artery layers is significant to discriminate between normal and diseased arterial segments. Therefore, for the first objective, we started with the intracoronary OCT images with three-layered structure, which can be considered as a normal coronary artery. In this objective, we aimed to find the features, which may describe arterial wall layers accurately, and a robust classifier to discriminate between features extracted from various coronary artery layers. Thus, the first objective is defined as: **Feature extraction and automatic classification of coronary artery layers (intima and media)**, which contributes to:

- Identifying the features, which perfectly describe coronary artery tissues.
- Choosing a classifier, which has a high performance compared against other classifiers to discriminate between different coronary artery layers.
- Analyzing the classification results at each step of the work to find the optimal classification model.

For each OCT pullback, intracoronary cross-sectional images might be misaligned, since the probe moves freely along the catheter during the imaging process. This is problematic for volumetric measurements of different tissues, which are significant for studying the progression and regression of various pathological formations particularly to evaluate the aneurysmal regions and stenotic segments of the arterial wall. Therefore, motion correction is a step of 3D reconstruction of intracoronary OCT images. Since, the optimal features, which can properly describe the arterial wall layers are determined in the first objective, they might be applied for motion correction of intracoronary OCT images by finding the maximum similarity between feature vectors extracted from the frames of an OCT pullback to perform registration. Thus, the second objective is defined as **Intra-slice motion correction using the similarity between extracted features**, which is the step of 3D reconstruction and contributes to:

- Applying deep learning on intracoronary OCT images for motion correction.
- Intra-slice motion correction in intravascular OCT images using tissue information rather than the lumen outline.
- Using deep features as similarity measures.

We define the third objective to design an intracoronary tissue characterization model to detect pathological formations caused by coronary artery disease. This is challenging considering the artifacts of the imaging system, the characteristics of each coronary artery tissue, and similar structures in different pathological formations. The classification foundation built in the first objective is employed in this step of the work to propose an appropriate tissue characterization model. Therefore, the third objective is defined as: **Automatic characterization of intracoronary pathological formations**, which contributes to:

- Characterization of complex pathological formations in KD from OCT imaging, namely intimal hyperplasia, media disappearance, neovascularization, fibrosis, calcification, and macrophage accumulation.

- Evaluation of different pre-trained CNN models for OCT image analysis with a limited labeled dataset.
- Assessment of the clinical usefulness of deep feature learning for OCT imaging in pediatric cardiology.

Flowchart of Figure 2.1 demonstrates the correlation between the main research objective and the three specific research objectives.

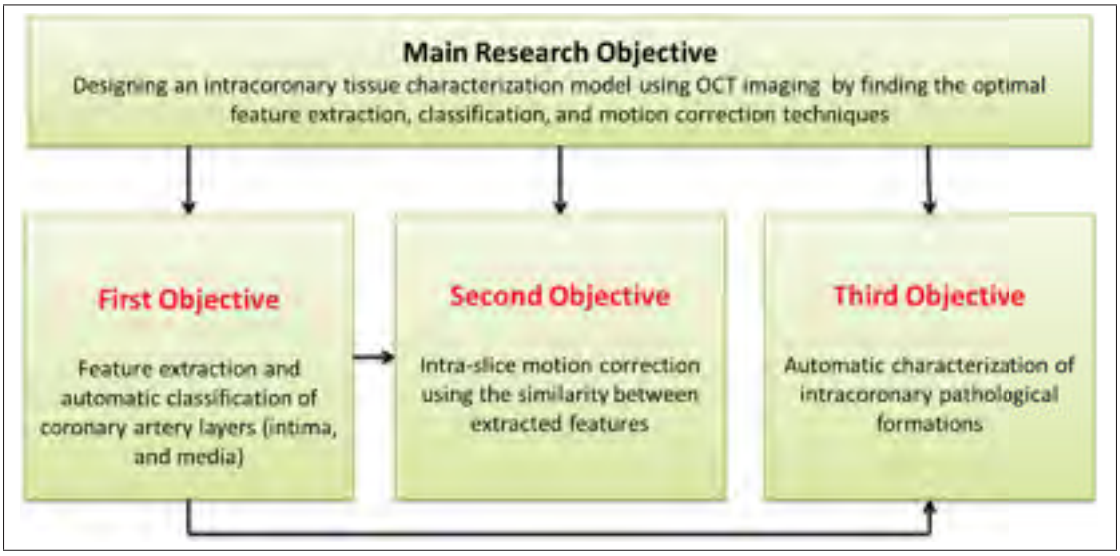


Figure 2.1 Main and specific research objectives

The detailed information of coronary artery structure and the OCT imaging system, as well as the clinical features of pathological formations caused by KD, should be considered to take into account all the possible aspects of the problems and build the optimal model related to each specific objective.

CHAPTER 3

DEEP FEATURE LEARNING FOR AUTOMATIC TISSUE CLASSIFICATION OF CORONARY ARTERY USING OPTICAL COHERENCE TOMOGRAPHY

Atefeh Abdolmanafi[1], Luc Duong[1], Nagib Dahdah[2], Farida Cheriet[3]

[1] Dept. of Software and IT Engineering, École de technologie supérieure, Montréal, Canada
[2] Div. of Pediatric Cardiology and Research Center, Centre Hospitalier Universitaire Sainte-Justine, Montréal, Canada
[3] Dept. of Computer Engineering, École Polytechnique de Montréal, Montréal, Canada

The article was published in Biomedical Optics Express on February 2017.

ABSTRACT

Kawasaki disease (KD) is an acute childhood disease complicated by coronary artery aneurysms, intima thickening, thrombi, stenosis, lamellar calcifications, and disappearance of the medial border. Automatic classification of the coronary artery layers (intima, media, and scar features) is important for analyzing Optical Coherence Tomography (OCT) images recorded in pediatric patients. OCT has been known as an intracoronary imaging modality using near-infrared light which has recently been used to image the inner coronary artery tissues of pediatric patients, providing high spatial resolution (ranging from 10 to 20 μm). This study aims to develop a robust and fully automated tissue classification method by using the Convolutional Neural Networks (CNNs) as feature extractor and comparing the predictions of three state-of-the-art classifiers, CNN, Random Forest (RF), and Support Vector Machine (SVM). The results show the robustness of CNN as the feature extractor and Random Forest as the classifier with classification rate up to 96%, especially to characterize the second layer of coronary arteries (media), which is a very thin layer and it is challenging to be recognized and specified from other tissues.

3.1 Introduction

3.1.1 Coronary artery structure

A normal arterial wall is composed of three layers. The first layer, the intima, is a transparent, achromatic, and extremely elastic structure comprised of endothelial cells in direct contact with circulating blood. It is characterized by a signal-rich pattern in OCT images, and a normal intima has a reported thickness of 61.7 $\pm$ 17.0 μm. The next layer of the arterial wall, the media, is homogeneous and composed of smooth muscle cells, lined by the inner and the outer elastica layers which are composed of elastic fibers. The media is specified by a signal-poor pattern in OCT images with a normal thickness of 61.4 $\pm$ 16.7 μm. Finally, the adventitia is the outermost layer of the artery, surrounding the media and characterized by a signal-rich layer in OCT images (Dionne *et al.*, 2015; Ligthart *et al.*, 2011).

One of the most important abnormalities resulting from Kawasaki disease is intimal hyperplasia, which can be eccentric or concentric and is followed by a mean intima thickening of 390.8 $\pm$ 166.0 μm. Mean media thickness in the case of the aneurysmal artery is 30.2 $\pm$ 56.9 μm. Changes to the normal structure of the vessel wall as a consequence of severe intimal hyperplasia lead to a partial disappearance of the media. The composition of calcified nodules, white thrombus, fibrosis, and macrophage accumulation are additional abnormalities (Dionne *et al.*, 2015).

3.1.2 Optical Coherence Tomography (OCT)

Most of the traditional cardiac imaging modalities, such as X-ray angiography and computed tomography, are effective to visualize the outline of the lumen, where the contrast agent flows in the coronary artery. However, they do not characterize the internal structure of the tissues, such as the vessel wall layers and plaque accumulation (Preim & Bartz, 2007). The inner vessel wall geometry and visualization of the morphology of both plaques and arterial wall layers allow for detection and evaluation of the thickness of each layer of the coronary artery, various thrombi,

and calcifications, enabling improvements to the process of diagnosis, treatment, and follow-up the patients with coronary complications. Although intravascular ultrasound (IVUS) may be used to assess the inner part of the vessels, its application is restricted due to its suboptimal resolution of 100 - 150 μm (Ferrante *et al.*, 2013).

Optical Coherence Tomography (OCT), an intracoronary imaging modality that uses near-infrared light, has many clinical applications because of its high resolution (ranging from 10 to 20 μm). OCT is a promising method for quantifying all information about the inner parts of the vessels by producing a sequence of cross-sectional images of coronary arteries with the high resolution about 10 times higher than IVUS (Bezerra *et al.*, 2009). At first, OCT has been widely used in retinal imaging as one of the significant diagnostic technologies of retinal diseases and glaucoma (Costa *et al.*, 2006). Then, it has been developed for other medical applications, specifically in cardiology (Zysk *et al.*, 2007). The introduction of OCT for intravascular imaging was found to be an interesting alternative for intravascular ultrasound (IVUS) imaging.

3.1.3 Kawasaki Disease (KD)

Kawasaki disease is an acute childhood inflammatory disease characterized by fever, rash, bilateral nonexudative conjunctivitis, erythema of the lips and oral mucosa, changes in the extremities, and cervical lymphadenopathy. While a high dose of Intravenous Immune Globulin (IVIG) infusion decreases the occurrence of coronary abnormalities, about 15% to 25% of untreated children suffer a risk of experiencing coronary artery aneurysms or ectasia (Newburger *et al.*, 2004), which may be followed by intimal hyperplasia, thrombi, stenosis, lamellar calcifications, disappearance of the medial border, and significantly the stiffness of the arterial wall (Orenstein *et al.*, 2012).

In vivo intravascular visualization of coronary arteries and diagnostic assessment of coronary artery abnormalities are feasible in children and may provide highly valuable progressive information (Dionne *et al.*, 2015; Harris *et al.*, 2014). In this work, we focused on the segments

of coronary arteries of patients which are identified as normal by the cardiologist to extract all attributes describing each coronary artery layer. This information can be used to distinguish the normal versus abnormal coronary arteries to detect intima hyperplasia, fibrous, and calcification.

3.1.4 Related works

Considering intima hyperplasia, thrombi, stenosis, lamellar calcifications, and disappearance of the medial border, tissue classification and specifically classification of the coronary artery layers is very important to evaluate the thickness of the layers. Furthermore, intimal thickening and disappearance of the media complicate the classification task. Therefore, manual segmentation of the coronary artery layers is tedious, time-consuming, and particularly error-prone from one observer to another.

Automatic lumen segmentation to assess the stenosis grading and characterization of the plaque types in OCT images of coronary arteries have been performed by Celi et al. (Celi & Berti, 2014). Yabushita et al. have proposed a method of plaque characterization by correlating OCT images with histology (Yabushita *et al.*, 2002). Other studies focused on atherosclerosis plaque characterization using optical properties of tissues (Levitz *et al.*, 2004). The actual backscattering and attenuation coefficients were measured by Xu et al. (Xu *et al.*, 2008). The automatic quantification of optical attenuation coefficients has been proposed by Van Soest et al. (Van Soest *et al.*, 2010). Ughi et al. (Ughi *et al.*, 2013) have proposed their classification method using texture features and attenuation coefficients to characterize atherosclerosis tissues. However, these studies did not address the challenging task of classification of coronary artery layers.

Advances in machine learning and pattern classification have lead to significant advances in automatic image recognition. For instance, Convolutional Neural Networks (CNNs) have been demonstrated as very powerful techniques in the broad range of tasks and in various fields of studies such as computer vision, language processing, image processing, and medical image

analysis (Szegedy *et al.*, 2015; Simonyan & Zisserman, 2014; Zeiler & Fergus, 2014; Eigen *et al.*, 2013). A wide range of image detection problem using CNNs can be traced back to the 90's, including lung nodule detection, micro-calcification, mass tissue detection on mammography (Lo *et al.*, 1993; Chan *et al.*, 1995; Sahiner *et al.*, 1996). Also, unsupervised deep learning for multiple organ detection using 4D data is performed by (Shin *et al.*, 2013).

The recent applications of CNNs in medical image analysis include pancreas segmentation using CT images of the abdomen (Roth *et al.*, 2015), classification of pulmonary peri-fussural nodules (Ciompi *et al.*, 2015), and brain tumor segmentation (Havaei *et al.*, 2016). The strength of CNNs originates from its deep structure which permits to extract the features from various abstraction layers (Szegedy *et al.*, 2015; Simonyan & Zisserman, 2014; Zeiler & Fergus, 2014; Eigen *et al.*, 2013). Basically, all CNNs consist of a series of layers defined by a specific number of filters or kernels that mainly have the role of feature detectors from a set of input images. Sliding the filters on the input images and calculating the convolution of these filter matrices and input image matrix creates a set of convolved features or simply feature maps. The exact meaning of learning a CNN is to train the CNN architecture by the values of these convolutional operations (Hochreiter & Schmidhuber, 1997).

The transferability of the information preserved on pre-trained CNNs which is one of the most significant characteristics of CNNs, has been demonstrated by the work of Azizpour et al. (Azizpour *et al.*, 2015). The recent studies show the significant applications of transfer learning in medical imaging to extract features from a new dataset using pre-trained CNNs or to use CNNs as the classifier by fine-tuning a pre-trained CNN (van Ginneken *et al.*, 2015; Bar *et al.*, 2015; Arevalo *et al.*, 2015; Chen *et al.*, 2015; Carneiro *et al.*, 2015; Gao *et al.*, 2016; Margeta *et al.*, 2015). Van Ginneken et al. have used penultimate layer of OverFeat pre-trained CNN as feature extractor for pulmonary nodules detection in CT acquisition. Then, extracted features have been used in a SVM classifier (van Ginneken *et al.*, 2015). Other studies used pre-trained CNNs as feature extractors on various medical image datasets such as chest radiograph data in the work of Bar et al. (Bar *et al.*, 2015) and mammography images in the work of Arevalo et al. (Arevalo *et al.*, 2015). Recently, pre-trained CNNs have been also used for classification tasks.

For instance, Chen et al. have applied pre-trained CNNs on ImageNet dataset to detect the fetal abdominal standard plane in ultrasound images by keeping the low-level representations extracted from natural images and modifying the parameters of the last layers based on the characteristics of ultrasound images (Chen *et al.*, 2015). The application of transfer learning using pre-trained CNNs in medical image analysis has been shown in other studies (Carneiro *et al.*, 2015; Gao *et al.*, 2016; Margeta *et al.*, 2015). Also, it has been demonstrated in the work of Tajbakhsh et al. that using pre-trained CNNs with adequate fine-tuning works better than training a CNN from scratch in medical image analysis applications (Tajbakhsh *et al.*, 2016). In their experiments, they have considered different categories of medical images in radiology, cardiology, and gastroenterology using different medical imaging systems such as colonoscopy images for polyp detection, CT pulmonary angiography (CTPA) for PE diagnosis, and Carotid intima-media thickness (CIMT) which is a noninvasive ultrasonography method in cardiology (Tajbakhsh *et al.*, 2016).

In this study, our main contribution is the automatic classification of coronary artery layers in pediatric patients using the images obtained from OCT system. Our work contributes to identifying the features, which perfectly describe both intima and media layers in OCT images using a pre-trained CNN as feature extractor. We also determine if it is better to fine-tune a pre-trained network and use it as the classifier or applying pre-trained CNNs as feature extractor and using the activations of the last fully connected layer to train other classifiers in our application. Finally, we analyze the performance of the classifiers using CNN features and compare the results against tissue classification results of coronary arteries using texture analysis, which is recently done by our group (Abdolmanafi *et al.*, 2016).

3.2 Material and methods

3.2.1 Pre-processing

We started the pre-processing by automatic recognition and removal of the guide-wire from the images. This step is applied to all the sequences obtained from all the patients. The images are

subsequently converted to a planar representation by transformation from Cartesian coordinate to Polar coordinate, where the vertical and horizontal axes correspond to radial distance and polar angle, respectively. The approximate region of interest, which consists of the lumen, the arterial wall layers, and the catheter, is extracted by applying active contour. Finally, the catheter and unwanted red blood cells are removed from the images by finding the smallest connected components. All the pre-processing steps are shown in Figures 3.1 and 3.2.

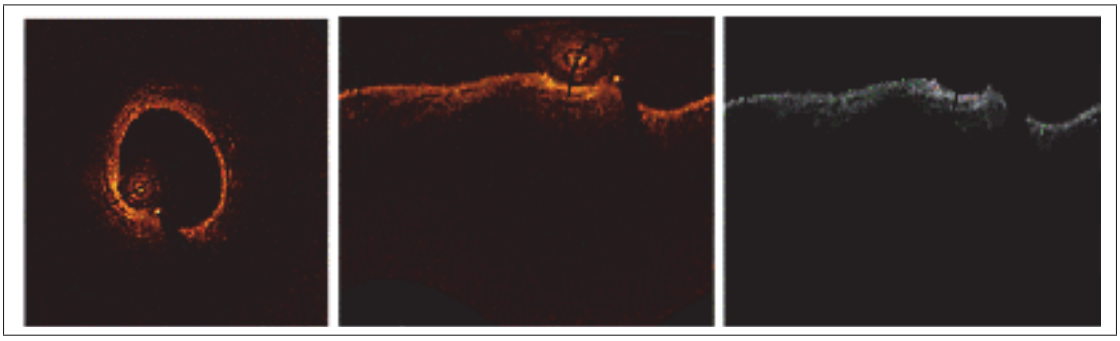


Figure 3.1 Pre-processing steps from left to right: original image, converting to planar representation, and extracting the region of interest by removing all the background.

3.2.2 Initial segmentation

Since OCT is a new imaging modality recently introduced in cardiology, to the best of our knowledge, there is no ground-truth available in the literature. Manual segmentation to create the ground-truth is tedious, time-consuming, and imprecise when the size of the tissues is very small. Moreover, the impacts of the disease on these small tissues make them more complicated to be recognized by trained operators. To improve the precision of our classification, we developed an automated approach based on the work of Azarnoush et al. on the intravascular images obtained from phantoms using OCT system (Azarnoush *et al.*, 2012). This initial segmentation is using peak information and image quantization to perform tissue segmentation to create ground-truth. The segmentation results were validated by an expert cardiologist.

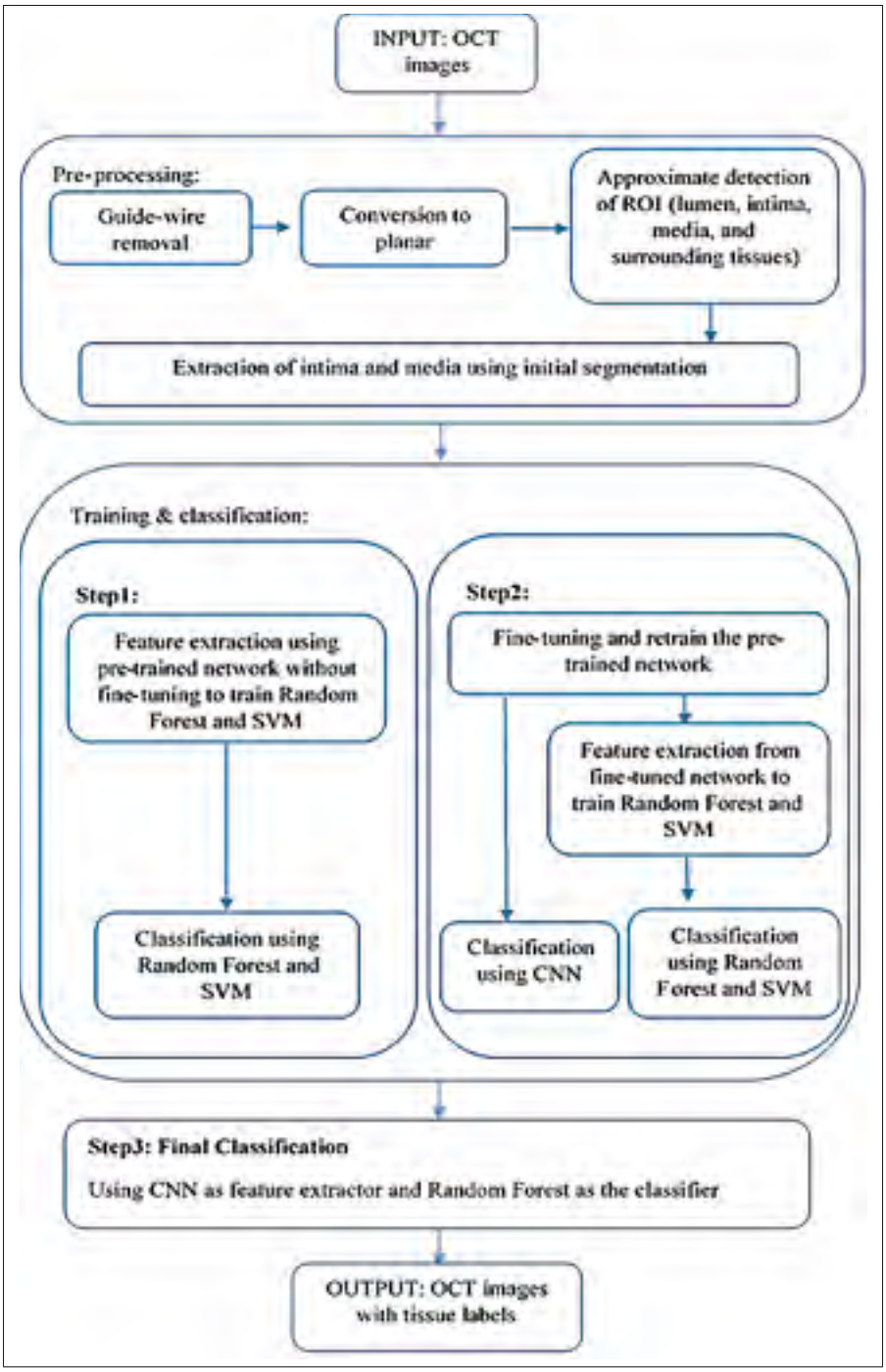


Figure 3.2 Flowchart of the tissue classification algorithm. The process of training, feature extraction, and classification using pre-trained CNN just as feature generator is shown in step 1 and fine-tuning the network to use it as the classifier as well as feature extractor to train Random Forest and SVM is demonstrated in step2. Step 3 show our final decision to select the optimal classification algorithm based on measured classification accuracy, sensitivity, and specificity at each step of the work and for each classifier.

For each frame of a sequence, we assess the image profile at different penetration depths by scanning the images from 0 to 359 degrees along the radius in planar representation. It is possible to recognize the number of layers present for each frame of a sequence by considering the fact that passing the light from one tissue of the sample to another creates a peak in the A-scan obtained from the backscattered light in the OCT system (Foin *et al.*, 2013).

Accordingly, a single peak signifies that just the intima is present; media disappearance is obvious in these parts of the images (see Figure 3.3a). Regardless of the level of noise, the presence of two peaks would correspond to two layers (intima and media) and correspondingly three borders (intima, intima-media, and media borders). In realistic laboratory conditions, however, three other possibilities must be considered:

- If the two peaks lie close to one another, then both peaks correspond to one layer, the only layer present in this case is intima (see Figure 3.3c).
- Considering the fact that the tissue surrounding the media (adventitia) is characterized by a signal-rich pattern in OCT images. Therefore, passing from the media results in a high peak value. A low value of the second peak also describes a case where just a single layer, the intima, is present (see Figure 3.3b).

When there are more than two peaks, they could all belong to one layer (e.g., if peak values are very close to each other), or they could belong to two different layers (see Figure 3.3d, Figure 3.3e, and Figure 3.3f).

To quantify what we mean by low peak values and how close peak values must be to one another to be considered as belonging to the same layer, we apply image quantization to segment each frame to 9 gray levels using 8 threshold values (see Figure 3.3), the first threshold for each frame determines the minimum difference between two peak values to consider them as belonging to one layer. Also, the peak values less than the first threshold value for each frame is considered as noise. By mapping points that are recognized as intima, intima-media, and media borders using peak values and image quantization to the real images, we can see the

precision of the method as illustrated in Figure 3.3a, a single peak corresponds to the presence of just intima. In Figure 3.3b, according to the thresholds obtained by image quantization, the second peak value is too low to be considered as corresponding to any of the tissues. Correspondingly, two borders are recognized in this part of the image (intima and intima-media). In Figure 3.3c, there are two peaks, but the two peaks are very close to each other (the difference between the peak values is less than the first threshold level). This is in agreement with the image, which shows no media border. In Figure 3.3d, there are three peaks, but the last two peak values are too low to be considered as corresponding to any of the tissues. In Figure 3.3e, the first two peaks are not close to each other to be considered as belonging to the same layer, and the second peak value is greater than the first threshold level. We have three borders of intima, intima-media, and media. Considering the width of the first peak, intima thickening is obvious in this part. In Figure 3.3f, an example of more than two peaks is depicted; the first two peaks belong to the intima, the last two peak values are very low to be considered as corresponding to any of the tissues.

3.2.3 Feature extraction and classification

Referring to the work of Tajbakhsh et al. (Tajbakhsh *et al.*, 2016), it is already determined that using pre-trained CNNs with proper fine-tuning works better in practice than the full training of a CNN on scarcely available medical image datasets. In our application, we use the pre-trained AlexNet model as feature generator by removing the top output layers (classification layers) and using the activations of the last fully connected layer as training input for Random Forest and Support Vector Machine (SVM). Also, we are interested to fine-tune the AlexNet model by finding the optimal learning rates for the weights at each layer and prepare the network for the classification task.

3.2.3.1 Convolutional Neural Networks (CNNs)

Generally, every Convolutional Neural Network architecture which is applicable in image processing builds on four main operations: convolution, nonlinearity (ReLu), pooling or sub-

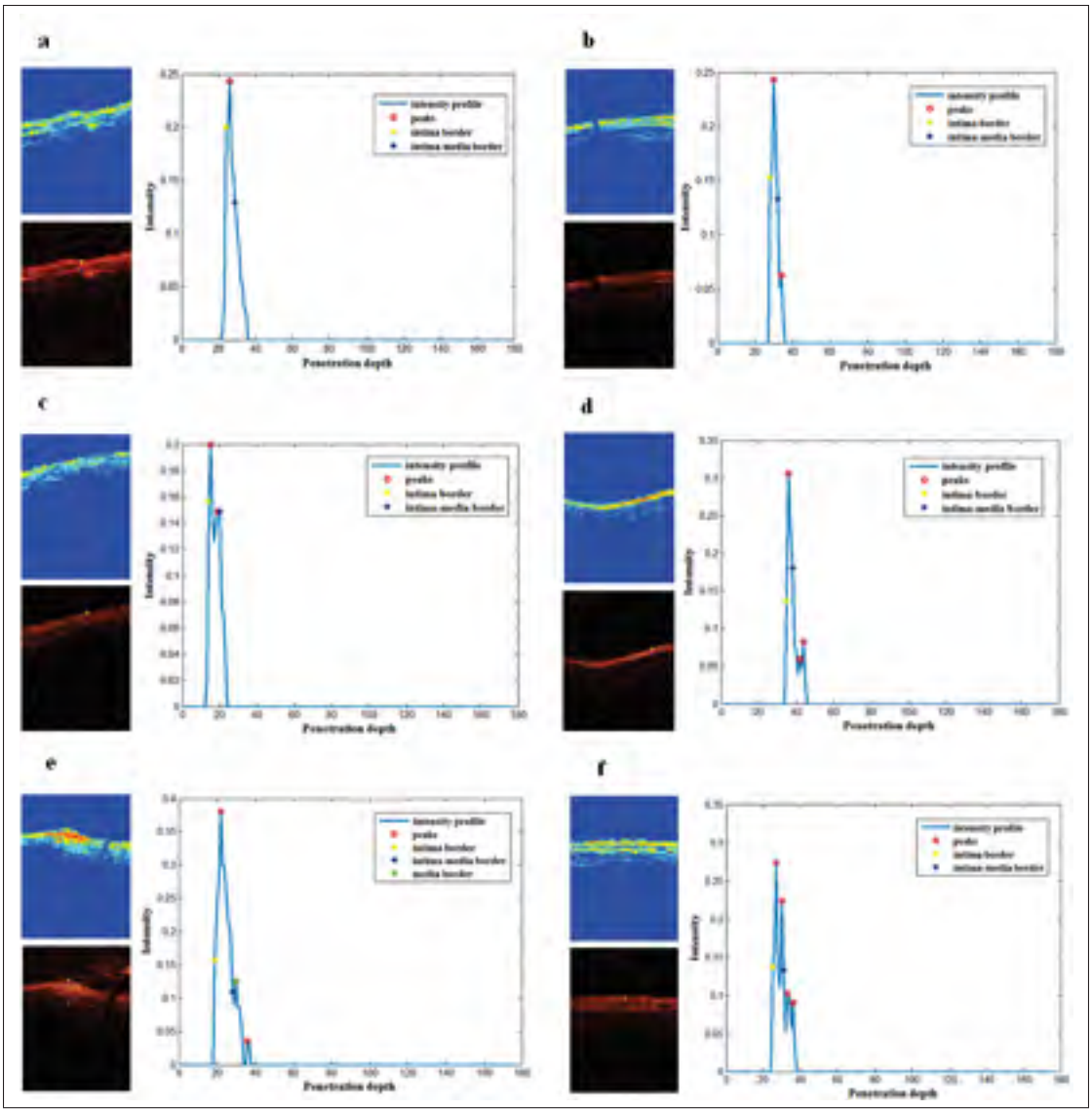


Figure 3.3 Peak detection and image quantization. Red circles show the peaks in the image profile; yellow, blue, and green are used to display intima, intima-media, and media borders, respectively.

sampling, and classification. In CNN, each convolutional filter creates one feature map when it moves through the whole image with a defined stride. Therefore, the size of the kernel determines the depth of the network. After every convolutional operation, a Rectified Linear Unit (ReLU) has been applied. Since convolution is a linear operator, it is required to introduce the non-linearity by storing non-negative values in the feature map and replacing the negative

Figure 3.4 Initial segmentation of three consecutive frames of four different patients. Yellow, blue, and green dots show intima, intima-media, and media borders, respectively.

values with zero. The pooling or sub-sampling is used for dimensionality reduction by keeping the most important information (Hochreiter & Schmidhuber, 1997; Hubel & Wiesel, 1959).

In detail, a CNN is trained by minimizing the cost function with respect to the weights at each layer using stochastic gradient descent. The cost function is defined as follows:

$$L = -(1/|X|)\Sigma_j^{|X|} ln(p(y^j|X^j)) \tag{3.1}$$

Where X is the size of the training set and $ln(p(y^j|X^j)$ denotes the probability of j^{th} image to be classified correctly with the corresponding label y. for each layer of the network, the weights are updated at each iteration i as follow:

$$V_{i+1} = \mu V_i - \gamma_i \alpha \partial L / \partial W \tag{3.2}$$

$$W_{i+1} = W_i + V_{i+1} \tag{3.3}$$

Where μ is the momentum, α is the learning rate, γ is the scheduling rate which reduces the learning rate at the end of iterations, and W is the weight at each iteration i for each layer (Krizhevsky *et al.*, 2012; Tajbakhsh *et al.*, 2016). The training process starts with initialized weights for each convolutional layer from a zero-mean Gaussian distribution and standard deviation.

Pre-trained AlexNet model

For both feature extraction and classification using CNN, the pre-trained AlexNet model (Krizhevsky *et al.*, 2012) is used in our experiments. AlexNet is trained on 1.2 million images from the ImageNet dataset, which are labeled with 1000 semantic classes. The network consists of 60 million parameters and 650000 neurons. It composed of eight learned layers, five convolutional layers, and three fully-connected layers with a final 1000-way softmax with GPU implementation of convolutional operation to speed up the training process of such a huge network. The architecture of the AlexNet used in our experiments is shown in Table 3.1. The model is trained using stochastic gradient descent with the batch size of 128, momentum of 0.9, and weight decay of 0.0005 to reduce the training error of the model (Krizhevsky *et al.*, 2012).

Transfer learning and fine-tuning

In transfer learning, we use the same architecture as the pre-trained CNN. The last fully connected layer (fc8 in this network), is designed based on the number of classes. Therefore, for the first step, the last 3 layers of the pre-trained network (fc8, prob, and classification layer) are

Table 3.1 AlexNet architecture consists of five convolutional layers, and three fully connected layers.

Layer	Layer type	Input of each layer	Output
conv1	convolution	227x227x3, stride [4 4], padding [0 0]	55x55x96
pool1	max pooling	55x55x96, stride [2 2], padding [0 0]	27x27x96
conv2	convolution	27x27x96, stride [1 1], padding [2 2]	27x27x256
pool2	max pooling	27x27x256, stride [2 2], padding [0 0]	13x13x256
conve3	convolution	13x13x256, stride [1 1], padding [1 1]	13x13x384
conv4	convolution	13x13x384, stride [1 1], padding [1 1]	13x13x384
conv5	convolution	13x13x384, stride [1 1], padding [1 1]	13x13x256
pooling5	max pooling	13x13x256, stride [2 2], padding [0 0]	6x6x256
fc6	fully connected	6x6x256	1x4096 feature vector
fc7	fully connected	1x4096 feature vector	1x4096 feature vector
fc8	fully connected	4096D feaure vector	1x1000

replaced by a set of layers, which are designed for multi-class classification to classify intima and media. Accordingly, the number of neurons in the last fully connected layer is set based on the number of classes in our dataset. The next step is fine-tuning; which means to initialize the weights of each layer in our network by transferring the weights from the pre-trained CNN and using the same structure of the pre-trained architecture.

Since low-level features are related to more general characteristics of the images such as edge orientation detectors, or color blob detectors that should be applicable to many tasks and over-fitting concerns of deep fine-tuning considering the small size of our dataset for each patient, we prefer just to fine-tune the weights of the last few layers of the network. But, from another perspective, our dataset is completely different from the original dataset of the pre-trained CNN. Therefore, it is more reliable if we fine-tune the pre-trained network by continuing the back-propagation and changing the network slightly. We started fine-tuning from the last fully connected layer (the new fc8 that we replaced based on our classification task). The weights of all other layers remain constant by forcing the learning rates to zero for those layers. The parameters are selected based on grid searching for an extensive interval of values. We keep μ and γ at 0.9 and 0.95 respectively and change the learning rate for the last fully connected layer by setting the learning rate to 0.1. For the next steps, we continue fine-tuning by changing the

learning rates of the last two layers, last three layers and so on, to reach the best performance of the network to stop fine-tuning. Table 3.2 shows the learning rates for each step. We started decreasing the learning rate to 0.01 from fc6 (first fully connected layer in the network). In this way, the weights of the last layers, which are more dataset specific are changing faster than the rest of the network.

Table 3.2 Learning rates at each step of fine-tuning the AlexNet model in our experiments. μ and γ are fixed at 0.9 and 0.95 respectively at all the steps of fine-tuning. We started fine-tuning from the last fully connected layer by setting the learning rate to 0.1 for this layer and zero for other layers. We continue changing the network slightly. We started decreasing the learning rates during fine-tuning from fc6. So, the weights of the last layers which are more dataset specific change faster than the rest of the network.

Layers	**Step 1**	**Step 2**	**Step 3**	**Step 4**	**Step 5**	**Step6**	**Step7**	**Step8**
fc8	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1
fc7	0.0	0.1	0.1	0.1	0.1	0.1	0.1	0.1
fc6	0.0	0.0	0.01	0.01	0.01	0.01	0.01	0.01
conv5	0.0	0.0	0.0	0.01	0.01	0.01	0.01	0.01
conv4	0.0	0.0	0.0	0.0	0.01	0.01	0.01	0.01
conv3	0.0	0.0	0.0	0.0	0.0	0.01	0.01	0.01
conv2	0.0	0.0	0.0	0.0	0.0	0.0	0.01	0.01
conv1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.01

3.2.3.2 Random Forest

Generating an ensemble of trees using random vectors, which control the growth of each tree in the ensemble significantly increases the classification accuracy. Random Forest works efficiently on large data sets, carries a very low risk of overfitting, and is a robust classifier for noisy data. The trees are grown based on the CART methodology to maximum size without pruning. Two important factors which affect the Random Forest accuracy are the strength, s, of each tree and the correlation, ρ, between them. Generalization error for Random Forest classifier is proportional to the ratio ρ/s^2. Hence, the smaller this ratio, the better functioning of Random Forest will be concluded. The correlation between trees is reduced by random selection of a

subset of features at each node to split on (Criminisi & Shotton, 2013; Kuhn & Johnson, 2013). To improve the performance of the classifier in our experiments, we started from 100 trees and increase the number of trees to 1000. The optimal number of trees is chosen by considering the Out Of Bag (OOB) error rate. By setting the number of trees to 241, the error rate is low, almost close to the minimum error rate, and fewer number of trees reduces the computational burden; so, classifier performance is faster. The number of randomly selected predictors (another tuning parameter in Random Forest) is set to 7. Random Forest training and validation is described in section 2.3.4.

3.2.3.3 Support Vector Machine (SVM)

SVM is robust against a large number of variables, large data sets, and noisy data, which are the principal challenges of medical images (Wang & Xue, 2014). Non-linear decision boundary is obtained using SVM by means of a kernel function. We employed two-class SVM classifier with Gaussian Radial Basis Function (RBF) as the kernel using C-Support Vector Classifier (C-SVC) available in the LIBSVM library (Chang & Lin, 2011). Using grid searching, we found the optimal values of regularization and Gaussian kernel bandwidth parameters, C and γ, which are set to 10 and 0.5 respectively. γ is related to the inverse of the RBF kernel extent. Therefore, the smaller γ, the wider kernel will be resulted. The trade-off between the SVM complexity and the number of non-separable samples is controlled by C. The larger C, the higher training accuracy is obtained (Wang & Xue, 2014). SVM training and validation is described in section 2.3.4.

3.2.4 Training and validation

A total of 26 sequences of intracoronary cross-sectional OCT images are obtained from patients with the history of KD using the ILUMIEN OCT system (St. Jude Medical Inc., St. Paul, Minnesota, USA) with the axial and lateral resolutions of 12-15 μm and 20-40 μm respectively.

In the experiments, for each patient, the ROIs (intima and media) are extracted from each frame of the sequence using the initial segmentation and they are labeled as one and two for intima and media respectively. We have a total of 4800 ROIs adapted to the pre-trained network for all the 26 patients (with an average of 180 ROIs per patient).

The experiments are performed separately for each patient and the classifiers retrained for each sequence of images per patient. For each experiment, the ROIs are divided randomly into three equal parts to create the training, validation, and test sets. To ensure that there is no correlation between these three sets, 2/3 of the ROIs are randomly selected as the training set and the remaining 1/3 is considered as the test set. To create the validation set, half of the training set is randomly selected as the validation set and the remaining built the final training set. Then, the accuracy is calculated on the validation set and the training process is stopped when the highest accuracy on the validation set is obtained. By terminating the training process, the features are extracted from the last fully connected layer just before the classification layer (fc7) of the fine-tuned network. Extracted features are used to train Random Forest and SVM. Classification is performed on the test set using CNN, Random Forest, and SVM (see Figure 3.2). We also apply the pre-trained network as feature extractor without fine-tuning. The extracted features from the layer fc7 of the network are used to train Random Forest and SVM. Then, the classification of the layers is performed on the test set using Random Forest and SVM Figure 3.2.

In the first part of this work, we performed experiments for each patient, separately, as it is reported in the literature because of the variety of textures in the coronary artery tissues from patient to patient. We randomly selected different frames of the sequence for each patient to ensure that consecutive frames are not chosen and our system does not become biased. To show the generalization of our method, another experiment is performed with the same configuration as the previous steps but with different selections of training, validation, and test sets. We train our algorithm using the images obtained from various patients. To create the training set, we select the OCT sequences from ten different patients. The remaining eighteen patients are split into two equal sets to create the validation and test sets.

Classification accuracy at each step of the work and for each ROI is calculated by comparing the predicted labels and the ground-truth for both intima and media. By considering the intima as the positive class, and media as the negative class, the sensitivity is measured as the true positive rate for intima and specificity is calculated as the true negative rate for media.

3.3 Results and discussion

The results of fine-tuning at each step are shown in Table 3.3 as measured values of accuracy, sensitivity, and specificity. Also, Figure 3.5 shows the accuracy of tissue classification for both classes, intima, and media, for all 26 patients at each step of fine-tuning. The classification rate up to 94% shows that the learning rates listed in Table 3.2 are selected properly. From the sixth step of fine-tuning (fc8 to third convolutional layer), the results are almost the same compared to the steps seven and eight. Figure 3.5 also shows very close results for the last three steps of fine-tuning. Therefore, it is reasonable to stop fine-tuning at step six (fc8 to the third convolutional layer).

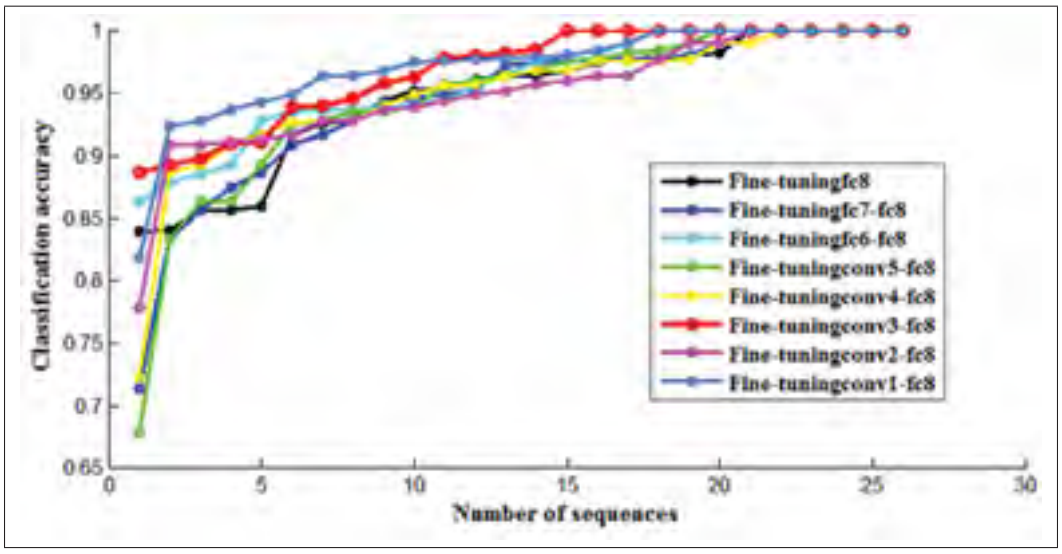


Figure 3.5 Tissue classification accuracy for all 26 sequences of intravascular OCT images at each step of fine-tuning the network from fc8 to the first convolutional layer to find the optimal depth of fine-tuning.

Table 3.3 Measured values of accuracy, sensitivity, and specificity to find the optimal depth of fine-tuning based on the performance of the network to classify intima and media at each step of fine-tuning. Values are reported as means ± standard deviation for 26 sequences.

Depth of fine-tuning	Accuracy	Sensitivity	Specificity
fc8	0.947±0.054	0.927±0.089	0.968±0.056
fc7-fc8	0.944±0.067	0.929±0.117	0.958±0.070
fc6-fc8	0.959±0.043	0.947±0.060	0.971±0.043
conv5-fc8	0.945±0.072	0.915±0.105	0.974±0.049
conv4-fc8	0.950±0.056	0.957±0.050	0.943±0.095
conv3-fc8	0.968±0.040	0.967±0.053	0.969±0.050
conv2-fc8	0.951±0.048	0.946±0.050	0.955±0.074
conv1-fc8	0.971±0.039	0.957±0.068	0.985±0.039

By fine-tuning the pre-trained network from the classification layer (fc8) to the third convolutional layer, the classification of the coronary artery layers are performed using the deep fine-tuned network, Random Forest, and SVM by applying the features extracted from fc7 (the last fully connected layer just before the classification layer) for each patient separately. The reported results in Table 3.4 and Figure 3.6 show the good performance of Random Forest and CNN than SVM for classification of the layers.

Table 3.4 Measured values of accuracy, sensitivity, and specificity to evaluate the performance of CNN, Random Forest, and SVM to classify intima and media.Values are reported as mean ± standard deviation for 26 sequences. In this experiment, fine-tuning is performed from fc8 to the third convolutional layer for CNN. Features are extracted from fc7 (the last fully connected layer just before the classification layer) to train Random Forest and SVM.

Classifier	Accuracy	Sensitivity	Specificity
CNN	0.97±0.04	0.97±0.05	0.97±0.05
Random Forest	0.96±0.04	0.97±0.05	0.96±0.07
SVM	0.94±0.07	0.94±0.07	0.95±0.11

For the next step of the work, CNN is used as the feature extractor for each sequence of images per patient. Then, Random Forest and SVM are trained using the activations of the last fully

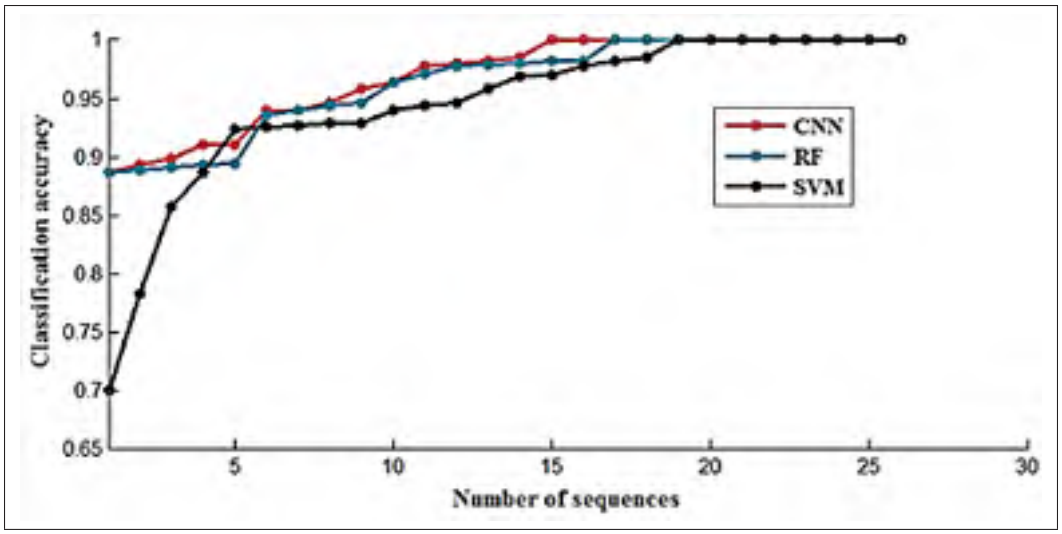


Figure 3.6 Performance of CNN, Random Forest, and SVM based on classification accuracy for each patient. Fine-tuning is performed from fc8 to the third convolutional layer for CNN. Features are extracted from fc7 (the last fully connected layer just before the classification layer) to train Random Forest and SVM.

connected layer just before the classification layer. The results demonstrate that using the pre-trained CNN as a feature generator and employing the extracted CNN features to train Random Forest compete against using CNN as the classifier even with deep fine-tuning the network (see Table 3.5 and Figure 3.7).

Table 3.5 Measured values of accuracy, sensitivity, and specificity.Values are reported as mean ± standard deviation for 26 sequences. In this experiment, CNN is used as feature extractor for our dataset. Features are extracted from fc7 (the last fully connected layer just before the classification layer) to train Random Forest and SVM. The performances of Random Forest and SVM are compared against the best performance of the CNN as classifier in our experiments when the network is fine-tuned from fc8 to the third convolutional layer.

Classifier	**Accuracy**	**Sensitivity**	**Specificity**
CNN	0.97±0.04	0.97±0.05	0.97±0.05
Random Forest	0.96±0.06	0.95±0.08	0.95±0.06
SVM	0.90±0.10	0.87±0.13	0.93±0.11

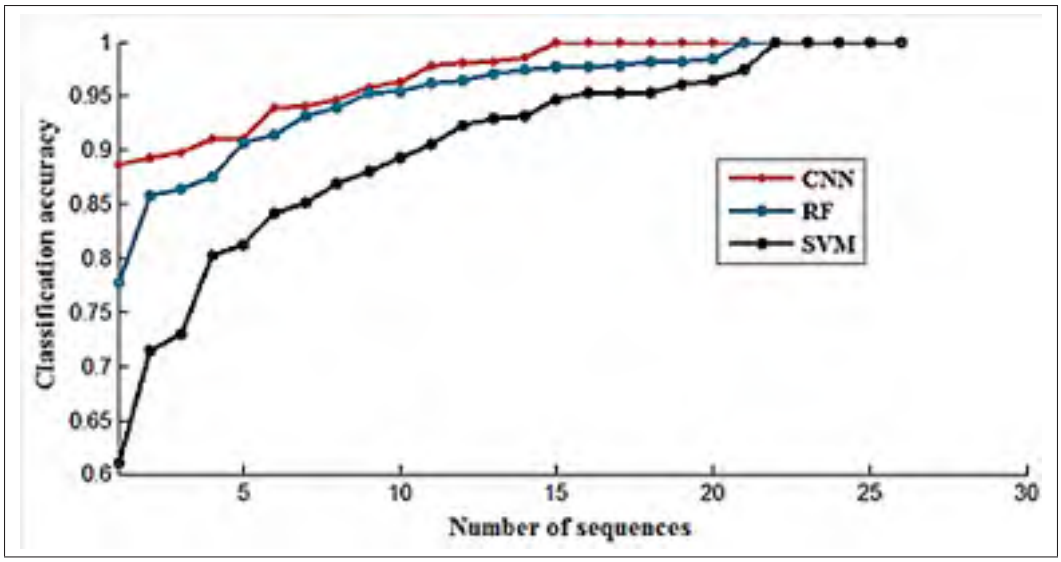


Figure 3.7 Performance of CNN, Random Forest, and SVM based on classification accuracy for each patient. CNN is used as feature extractor for our dataset. Features are extracted from fc7 (the last fully connected layer just before the classification layer) to train Random Forest and SVM. The performance of RF and SVM compared against the best performance of the CNN as the classifier in our experiments when the network is fine-tuned from fc8 to the third convolutional layer.

To show the generalization of our method, the training process is performed using different OCT images obtained from various patients. Measured accuracy, sensitivity, and specificity for this step of the work are shown in Table 3.6 and Table 3.7. Table 3.6 shows the tissue classification results of fine-tuned CNN, Random Forest, and SVM. Table 3.7 represents the tissue classification results of Random forest and SVM using the features generated from pre-trained CNN without fine-tuning. The results in Table 3.6, and Table 3.7 show the capability of our method to be generalized to other future cases.

In this study, to validate our results, the experiments are performed on 26 different sequences of intracoronary OCT images obtained from various patients. Our findings show that although CNNs are robust to be used as the classifier and using pre-trained CNNs significantly decreases the computational burden, training time, and convergence issues. But, retraining the network during fine-tuning requires a considerable amount of time. Overfitting concerns in deep fine-tuning the network and finding proper learning rates for each layer are other issues of using

Table 3.6 Measured values of accuracy, sensitivity, and specificity to evaluate the performance of CNN, Random Forest, and SVM to classify intima and media for the next step of the work when our algorithm is trained on different patients. In this experiment, fine-tuning is performed from fc8 to the third convolutional layer for CNN. Features are extracted from fc7 (the last fully connected layer just before the classification layer) of the fine-tuned network to train Random Forest and SVM.

Classifier	**Accuracy**	**Sensitivity**	**Specificity**
CNN	0.92	0.85	0.99
Random Forest	0.92	0.90	0.94
SVM	0.92	0.88	0.96

Table 3.7 Measured values of accuracy, sensitivity, and specificity for the next step of the work when our algorithm is trained on different patients. In this experiment, CNN is used as feature extractor for our dataset. Features are extracted from fc7 (the last fully connected layer just before the classification layer) to train Random Forest and SVM. The performances of Random Forest and SVM are compared against the best performance of the CNN as the classifier in our experiments when the network is fine-tuned from fc8 to the third convolutional layer.

Classifier	**Accuracy**	**Sensitivity**	**Specificity**
CNN	0.92	0.85	0.99
Random Forest	0.88	0.98	0.78
SVM	0.83	0.94	0.71

CNNs as the classifier. Our results show that it is more efficient to use pre-trained CNNs as feature generators for our application by removing the classification layer and using the activations of the last fully connected layer to train Random Forest. By comparing the results of tissue classification using CNN features against our previous work (Abdolmanafi *et al.*, 2016), CNN features are substantially robust to describe the characteristics of objects of interest than textural features. Also, comparing the tissue classification accuracy measured in our experiments with the work of Ughi et al. (Ughi *et al.*, 2013) show that using the same classifier

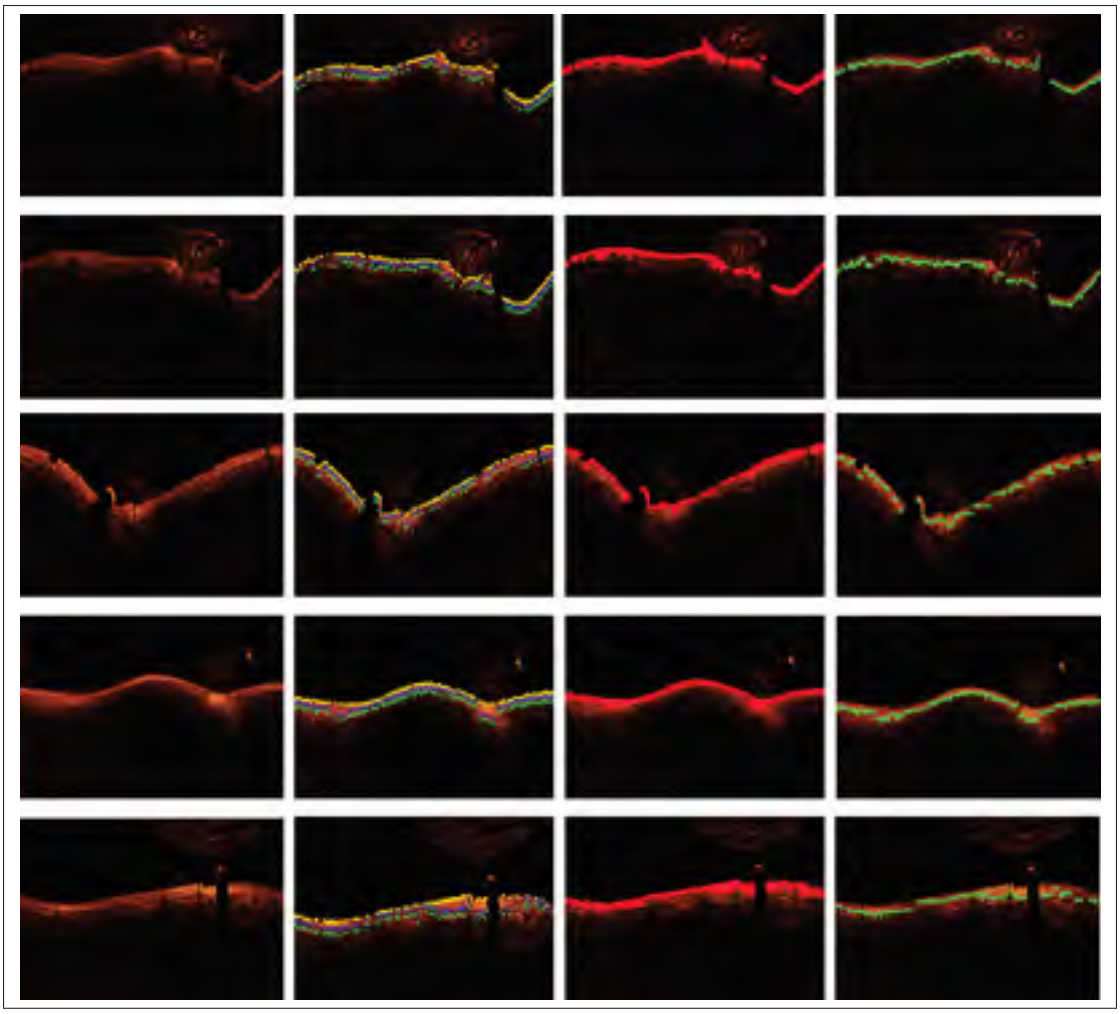


Figure 3.8 Classification results for one frame of five different patients. From left to right for each patient: original image converted to planar representation, initial segmentation, intima (red), and media (green).

(Random Forest) CNN features are more discriminant than optical properties and texture features for tissue classification.

As the classifier, Random Forest works efficiently on large data sets, carries a very low risk of overfitting, and training the model using Random Forest is considerably fast compared to CNN. In our experiments, it takes 5 hours to fine-tune and retrain the pre-trained network. Feature extraction using CNN takes 45 minutes, and training Random Forest takes 4 minutes. Therefore, training the model using Random Forest for all 26 patients was 75 times faster

than fine-tuning the pre-trained network exclusive of the time which has been spent to find the optimal values for learning rates and depth of fine-tuning. The measured values of accuracy, sensitivity, and specificity using Random Forest as the classifier also compete against the same values obtained by deep fine-tuning the CNN and using that as the classifier.

3.4 Conclusion

The main contribution of this study is the classification of the coronary artery layers using OCT imaging in pediatric patients. Fully automated tissue classification method is proposed in this work by using a pre-trained CNN as feature extractor by removing the classification layers and using the activations of the last fully connected layer to train Random Forest and SVM. The results confirm the robustness of CNN features to describe the tissue characteristics and Random Forest as the classifier considering the small size of the arteries in children and infants, which is followed by very thin layers in the structure of coronary arteries, and OCT artifacts. Considering the results obtained from different steps of this work, two major points can be noted: 1. Training the algorithm on a specific patient and classify the layers. 2. Training the algorithm on a set of patients and generalize it for future cases.

This will contribute to estimating intima-media thickening to evaluate the functionality of coronary arteries in patients suffering from Kawasaki disease. The stiffness of the coronary artery tissues (distensibility) as a result of calcium deposits and fibrous scarring, while the layers have normal thickness, is another significant abnormality caused by KD. Future works will be focused on detecting other abnormalities, evaluating distensibility, dynamic, and geometry of the vessels using stationary OCT imaging.

CHAPTER 4

INTRA-SLICE MOTION CORRECTION OF INTRAVASCULAR OCT IMAGES USING DEEP FEATURES

Atefeh Abdolmanafi[1], Luc Duong[1], Nagib Dahdah[2], Farida Cheriet[3]

[1] Dept. of Software and IT Engineering, École de technologie supérieure, Montréal, Canada
[2] Div. of Pediatric Cardiology and Research Center, Centre Hospitalier Universitaire Sainte-Justine, Montréal, Canada
[3] Dept. of Computer Engineering, École Polytechnique de Montréal, Montréal, Canada

The article was submitted to the IEEE Journal of Biomedical and Health Informatics (JBHI) on May 2018.

ABSTRACT

Intra-slice motion correction is an important step for analyzing volume variations and pathological formations from intravascular imaging. Optical Coherence Tomography (OCT) has been recently introduced for intravascular imaging and assessment of coronary artery disease. 2D cross-sectional OCT images of coronary arteries play a crucial role to characterize the internal structure of the tissues. Adjacent images could be compounded, however, they might not fully match due to motion, which is a major hurdle for analyzing longitudinally each tissue in 3D. The aim of this study is to develop a robust tissue matching based motion correction approach from a sequence of 2D intracoronary OCT images. Our motion correction technique is based on the correlation between deep features obtained from Convolutional Neural Network (CNN) for each frame of a sequence. The optimal transformation of each frame is obtained by maximizing the similarity between the tissues of reference and moving frames. The results show a good alignment of the tissues after applying CNN features and determining the transformation parameters.

4.1 Introduction

Coronary arteries, which are responsible to deliver oxygenated blood to the heart muscles can be affected by arterial stenosis and lead to myocardial infarction (Newburger *et al.*, 2004; Hauser *et al.*, 2004). The functionality of the cardiac tissues significantly depends on the coronary blood flow to the myocardium. Therefore, coronary artery disease (CAD) is the main leading cause of myocardial infarction and sudden death (Newburger *et al.*, 2004; Hauser *et al.*, 2004). Angiographic images allow to visualize the trajectory of the contrast agent but they cannot provide any information on the underlying coronary tissue layers. Considering the limitations of coronary angiography to evaluate coronary artery disease, catheter-based Intravascular Ultrasound (IVUS) has been used for many years to evaluate coronary artery tissue layers and pathological formations on different coronary artery layers (Rathod *et al.*, 2015). IVUS imaging is restricted by limited spatial image resolution (100-150 μm) to detect the thickness of various pathological formations and low pullback speed. In contrast, Optical Coherence Tomography (OCT) is another catheter-based imaging system, which plays a significant role in the development of medical imaging modalities with interesting advantages over IVUS imaging modality (Rathod *et al.*, 2015). OCT is an interferometric imaging modality that maps the backscattered near-infrared (NIR) light to create cross-sectional images of the tissues under review in micrometer scale (Zysk *et al.*, 2007). The image-wire is inserted into the coronary artery using an over-the-wire balloon catheter from patient's groin. A sequence of cross-sectional images of a coronary artery segment is recorded using the backscattered light from the arterial wall through each pullback. Considering the fact that light can be attenuated by blood before reaching the vessel wall, blood clearance is required before starting the image acquisition (Bezerra *et al.*, 2009; Drexler & Fujimoto, 2015). OCT has been developed for the diagnosis and treatment guidance of coronary artery disease. It has the high resolution ranging from 10 to 20 μm to characterize the internal structure of tissues such as vessel wall layers and plaque accumulation (Ferrante *et al.*, 2013). In the cross-sectional view, the normal coronary artery has a three-layered structure. Intima is the first layer, which is composed of endothelial collagen and is connected to lumen by a single layer of endothelial cells. The second layer, me-

dia, consists of muscle cells and is determined by internal and external elastic lamina. Media is enclosed by the outermost layer, adventitia (Regar *et al.*, 2011).

Intravascular assessment of coronary artery tissues is a challenging task considering the pathological formations due to various coronary artery complications, limitations and the possible artifacts of the imaging system. 2D cross-sectional OCT images of coronary arteries play a crucial role to estimate the thickness variations of arterial wall layers and evaluate the severity of the disease by detecting the various scarring remodeling features (Dionne *et al.*, 2015). But, the accurate assessment of pathological formations is obtained by considering the information of adjacent frames and analyzing the volume variation of each tissue using 3D reconstruction. This can be useful to evaluate the dynamics of coronary artery motion and distensibility variation as a result of calcium deposits in a specific coronary artery segment. Volumetric measurements of different tissues are significant for studying the progression and regression of various pathological formations particularly to evaluate the aneurysmal regions and stenotic segments of the arterial wall (Harris *et al.*, 2014; Orenstein *et al.*, 2012; Tearney *et al.*, 2008).

3D assessment of the morphological tissues is problematic because they are highly affected by motion artifacts resulted in the rotation and translation of the imaging catheter along the artery during image acquisition. Generally, heart beating and respiration during OCT acquisition are the main sources of both axial and longitudinal motions (Balocco *et al.*, 2016). Furthermore, when the vessel dimensions are changed during a cardiac cycle, it results in larger lumen area in cross-sectional images (Balocco *et al.*, 2016). The frequency-domain OCT with the pullback speed of 20 mm/sec is 10 - 20 times faster than the previous generation of OCT (time-domain OCT). Therefore, accelerated image acquisition within fewer numbers of cardiac cycles reduces motion artifacts but they cannot be entirely eliminated (Takarada *et al.*, 2010).

4.1.1 Related works

4.1.1.1 Motion correction

Since the OCT probe moves freely in coronary artery pathway, cross-sectional images might be misaligned. This is problematic to evaluate longitudinally each tissue and to design robust clinical measurements. Since OCT is recently used in cardiology, motion correction methods are mostly focused on intracoronary IVUS images and OCT retinal images.

Intravascular Ultrasound (IVUS) has been used for many years as an intracoronary imaging modality in cardiology. Therefore, many of the motion correction and 3D reconstruction methods are focused on IVUS images. (Wahle *et al.*, 1998) developed an image fusion technique to create the 3D reconstruction of intracoronary IVUS images. Cothren et al. proposed a 3D reconstruction technique by detecting the 3D trajectory of the IVUS transducer using the information of the angiographic images. It is demonstrated that the correct rotational orientation of IVUS images on the angiogram is based on the function of time by applying best-fit angle function (Cothren *et al.*, 2000). To increase the accuracy of the 3D reconstruction, Bourantas et al. extracted the catheter path from biplane angiography. Also, the IVUS images and angiogram are visualized simultaneously using ECG wave (Bourantas *et al.*, 2005). Zheng et al. focused on a 3D reconstruction method of intracoronary IVUS images by reconstructing the pullback path using snake algorithm. The precision of the 3D reconstruction method is increased by focusing on 3D axial position, spatial orientation, and surface fitting (Zheng, 2009). 3D artery is registered with intracoronary IVUS cross-sections using distance mapping algorithm in the work of (Tu *et al.*, 2011). The other study is focused on a 3D reconstruction of IVUS images using biplane angiography to detect the 3D centerline. The IVUS frames are distributed along the reconstructed centerline in equivalent time intervals (Ma *et al.*, 2013). Athanasiou et al. proposed a 3D reconstruction of intracoronary IVUS images by estimating the lumen and the outer vessel wall borders using the approximation of the arterial centerline to evaluate plaque formations (Athanasiou *et al.*, 2016). Karlas et al. proposed a 3D reconstruction method based on the fusion of IVUS and preoperative CT data. The alignment of the detected lumen borders

is performed using the ellipse fitting technique and CT data is used to detect the arterial shape (Karlas & Lee, 2015). Zhao et al. developed a 3D reconstruction method using EM sensor to detect the catheter pose. Lumen contours are detected by applying a radial scan method (Zhao *et al.*, 2016). Some studies are performed on OCT images of coronary arteries. 3D reconstruction of coronary artery images is performed by Ellwein et al. using graph theory applied on computed tomography (CT) and OCT data of a single patient after stent placement (Ellwein *et al.*, 2011). (Athanasiou *et al.*, 2012) proposed a semi-automated 3D reconstruction method using OCT images and biplane angiography.

Nevertheless, very few studies addressed the problem of non-rigid, intra-slice motion correction in intravascular OCT images using tissue information, rather than the lumen outline. Also, the physics of the imaging modality is important to be considered to find the proper solution of motion correction. We looked at the problem from the perspective of tissue matching. As much as we can align the tissues, we have the more accurate assessment of deformation and pathological changes from one tissue to another. To solve the problem of motion correction using OCT imaging, various considerations have been taken into account:

- **Physical principals of the imaging system:** Although there are some motion correction methods, which are performed on IVUS images, considering the fact that the problem of motion correction is defined differently regarding the functionality of the imaging systems, we thought about the solution that might tackle our problem more accurately. High resolution of OCT images is one of the characteristics of the OCT images that contributes to have a better interpretation of different tissues and extract useful features compared against IVUS images. Also, the maximum IVUS pullback speed is four times slower than the OCT pullback speed. This can completely change the definition of the problem of motion correction for IVUS images and OCT images since the accelerated image acquisition within fewer numbers of cardiac cycles reduces the motion artifacts. Therefore, we need to have more accurate motion correction technique based on the advantages of the imaging modality (the high resolution in this problem) to deal with very small variations in the rotation and translation of the catheter form one frame to another due to the accelerated image acquisition.

- **Transformation parameters:** Translation and rotation should be considered in the problem of motion correction since the position of the catheter is generally out of the center due to the blood flow and high heart beat, particularly in infants and children. Also, movements of the interventionist hand can cause a very small rotation of the catheter. Arterial wall as a soft tissue has some deformations due to the cardiac motion, which should be considered in the problem of motion correction. Therefore, the motion correction technique should be sensitive to very small variations.

- **Challenges:** In challenging cases, there are some pullbacks with pathological tissues; in some cases, the tissue borders are not clear and the shape of the arterial wall is not specific and visible. Considering the fact that the application of the proposed motion correction technique should be generalized to all different OCT pullbacks, the solution for the problem of motion correction of the intracoronary OCT images should be based on tissue information rather than the lumen outline.

4.1.1.2 Convolutional Neural Network (CNN)

Convolutional Neural Networks (CNNs) have been demonstrated very effective in different study fields (Szegedy *et al.*, 2015; Simonyan & Zisserman, 2014; Zeiler & Fergus, 2014; Eigen *et al.*, 2013). Recently, CNNs are widely used in the field of medical image analysis to perform segmentation and classification tasks (Roth *et al.*, 2015; Ciompi *et al.*, 2015; Havaei *et al.*, 2017). Considering the fact that it is rare to access sufficient data to train a network from scratch, specifically in the field of medical image analysis, it is wise to transfer the preserved knowledge of a pre-trained network to a new application. Transfer learning is one of the most efficient characteristics of CNNs, which is demonstrated by the work of (Azizpour *et al.*, 2015) in the field of medical image analysis. Pre-trained networks have been used as feature generators (van Ginneken *et al.*, 2015; Bar *et al.*, 2015; Arevalo *et al.*, 2015; Abdolmanafi *et al.*, 2017). Also, fine-tuned pre-trained networks have been used recently for classification tasks for different applications in the field of medical image analysis (Chen *et al.*, 2015). A tissue classification model is introduced in our previous work using deep feature learning. Fine-tuned

network can be used to generate features from each coronary artery layer to train random forest for classification of various arterial wall layers (Abdolmanafi *et al.*, 2017).

In this study, we demonstrate the application of deep features to solve the problem of motion correction for intracoronary OCT images considering the physics of the imaging system, the movement of the catheter during the image acquisition, and the challenges of the OCT imaging. Our motion correction algorithm is designed to correct the rotation and translation of the catheter as well as the arterial wall deformation through image acquisition by finding the best transformation parameters that will maximize the similarity between deep features extracted from the reference and moving frames at each transformation.

The originality of this approach is as follows:

- Application of deep learning on intracoronary OCT images for motion correction for the first time.
- Intra-slice motion correction in intravascular OCT images using tissue information rather than the lumen outline.
- Automatic feature extraction by applying a Convolutional Neural Network and using the similarity between deep features to perform registration.
- This study is not limited to our application and it can be a key to evaluate different plaque formations, and tissue deformations.

This study is organized as follows: We start with a brief explanation of data and pre-processing steps in section 4.2.1. Then, we present the CNN feature learning process and introduce our motion correction approach in section 4.2.2. The results using patient dataset are shown and discussed in section 4.3. Finally, we conclude with the possible extension of this work in section 4.4.

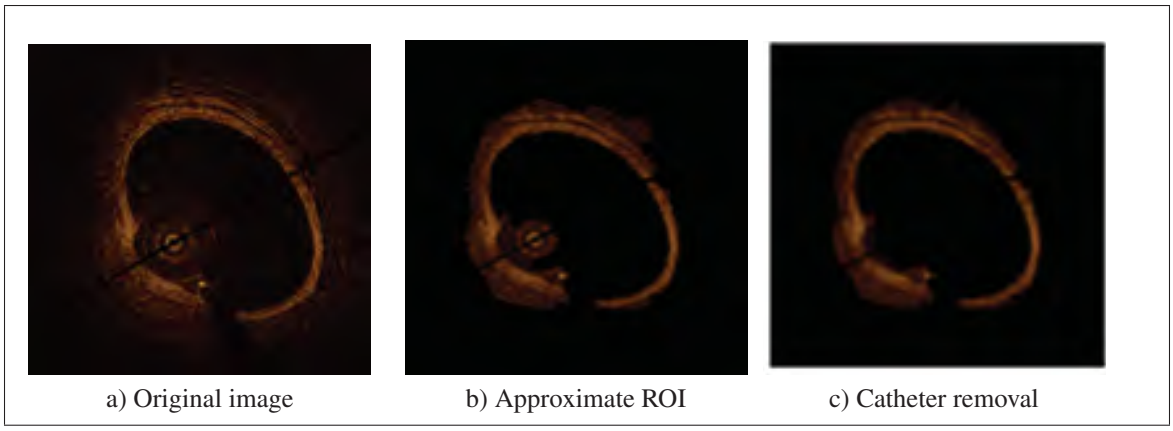
4.2 Material and method

4.2.1 Data collection and pre-processing

Kawasaki Disease (KD), mucocutaneous lymph node syndrome, is an acute vasculitis syndrome in infants and young children, which is characterized by fever, rash, conjunctivitis, and swollen erythematous hands and feet and leads to inflammation in the walls of medium-sized arteries throughout the body. Coronary arteries are affected by arterial inflammation. Intimal thickening is the most distinguished pathological feature of late coronary artery lesions in Kawasaki disease. In severe cases, it can lead to localized stenosis, extensive intimal hyperplasia, and consequently disappearance of media (Newburger *et al.*, 2004; Hauser *et al.*, 2004).

The experiments are performed on 26 retrospective cases comprising of pullbacks of intracoronary cross-sectional images obtained from different pediatric patients with KD using ILUMIEN OCT system (St. Jude Medical Inc., St. Paul, Minnesota, USA). The axial and lateral resolutions of the OCT system are 12-15 μm and 20-40 μm respectively. FD-OCT is used for image acquisition with the pullback speed of 20 mm/sec and frame rate of 100 frames/sec. Each pullback consists of 270 frames of DICOM images per patient. Permission to conduct this study on retrospective OCT studies was granted by the institutional review board.

The pre-processing is started by recognizing and removing the guide-wire automatically from all the images per pullback. Then, the approximate region of interest (ROI) including the lumen, arterial wall layers, and the catheter are recognized and extracted using active contour. Active contour does not perform properly to remove the catheter and unwanted blood cells, specifically when the catheter is located very close to the arterial wall. Therefore, for the last step of the pre-processing, we removed the catheter and unwanted blood cells by applying the smallest connected components approach (Figure 4.1).



a) Original image b) Approximate ROI c) Catheter removal

Figure 4.1 Pre-processing steps: Original image is shown in (a), using active contour the approximate ROI including lumen, catheter, intima, media, and surrounding tissues are detected in (b). Then, catheter and unwanted blood cells are removed using connected components approach as shown in (c).

4.2.2 Motion correction model

4.2.2.1 Deep feature extraction

As it is mentioned in the related works, pre-trained CNNs work efficiently as fixed feature generators for OCT images by removing the uppermost fully connected layer (classification layer) from the network architecture and using the activations of the last fully connected layer, right before the classification layer, as deep features of the new images. The process of feature extraction using CNNs is briefly explained in the diagram of Figure 4.2.

Figure 4.2 Feature Extraction Using Convolutional Neural Networks (CNNs): Convolutional operation is performed by moving the filters with fixed stride through the input image and computing the convolution between each filter matrix and input image matrix. Non-linearity is introduced after each convolutional operation by applying a Rectified Linear Unit (ReLU) since convolution is a linear operator. Features are extracted from the fully connected layer right before the classification layer.

In this study, fine-tuned AlexNet is applied to generate the features for our dataset (Krizhevsky *et al.*, 2012). The AlexNet model is built on 60 million parameters and 650000 neurons and is trained on 1.2 million images from ImageNet dataset. It consists of five convolutional layers (conv1 to conv5) , three max pooling, and three fully connected layers (fc6, fc7, and fc8) (Krizhevsky *et al.*, 2012). The network architecture is shown in Figure 4.3. Each convolutional filter or kernel has the role of feature detector of the images to create the feature maps by sliding through the whole image with defined stride and computing the convolution of the filter matrices and the input image matrix. The depth of the network is determined by the number of kernels. It is also important to mention that each layer extracts the features from the output of the previous layer (Hochreiter & Schmidhuber, 1997). The process of updating the weights at each layer and for each iteration, i, is as follows:

$$V_{i+1} = \mu V_i - \gamma_i \alpha \partial L / \partial W \tag{4.1}$$

$$W_{i+1} = W_i + V_{i+1} \tag{4.2}$$

Where μ is the momentum, α is the learning rate, γ is the scheduling rate which reduces the learning rate at the end of the iterations (Krizhevsky *et al.*, 2012; Abdolmanafi *et al.*, 2017). L is the cost function which is aimed to be minimized with respect to the weights, W, at each layer and during the training process. In this work, we modified pre-trained AlexNet for our application based on our previous study (Abdolmanafi *et al.*, 2017). Since the pre-trained network is applied as a fixed feature generator, we removed the last three layers (fc8, prob, and classification layer), which are designed for the classification task. To start the inductive transfer, we used the same architecture as the pre-trained network and the weights of each layer are initialized by transferring the weights from the pre-trained network. To start fine-tuning, two facts are considered: 1. There is a risk of overfitting by deep fine-tuning the network, 2.Deep features extracted from the upper layers are more dataset specific, which characterize distinctive features of the images. Therefore, fine-tuning is started from the last fully connected layer (fc7) using grid searching for an extensive interval of learning rates. After finding the optimal learning rates, fine-tuning is performed by applying the learning rate of 0.1 for fc7

and the learning rate values are decreased to 0.01 from the layer fc6. Fine-tuning is stopped at the third convolutional layer since continuing the fine-tuning does not improve the network performance. The learning rates of the first and the second convolutional layers are fixed at zero to keep the weights constant. Also, μ and γ are fixed at 0.9 and 0.95 respectively for the whole process of fine-tuning. Table 4.1 shows the learning rates of fine-tuned network. The features are extracted from the last fully connected layer (fc7).

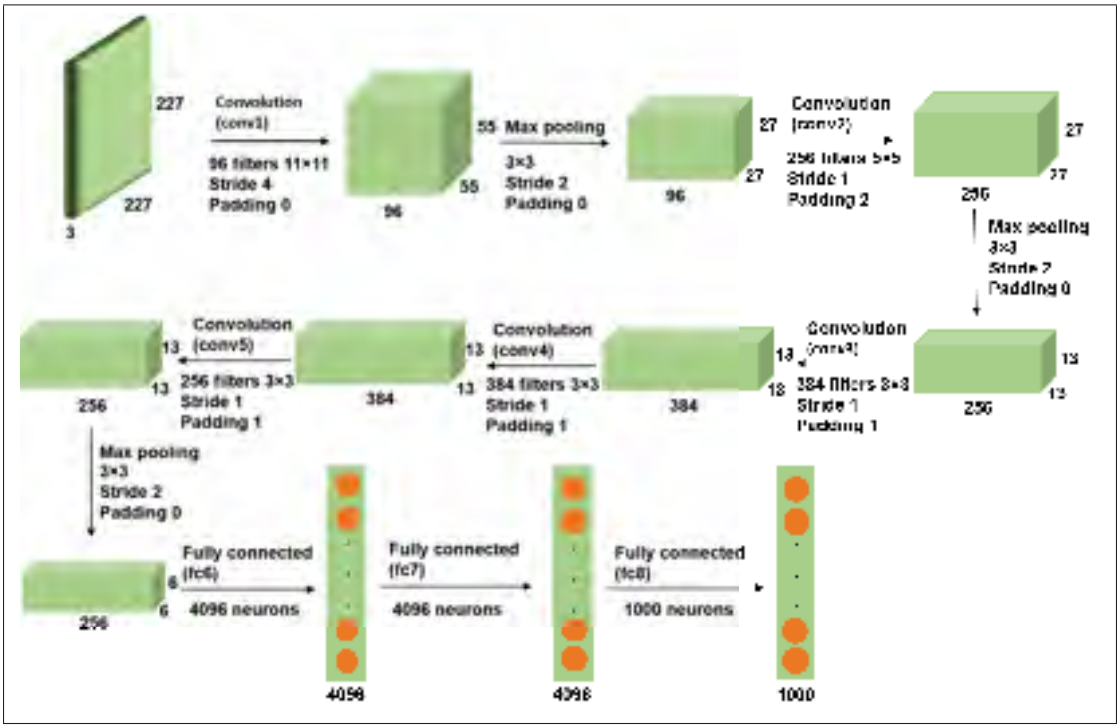


Figure 4.3 Architecture of AlexNet with five convolutional layers and three fully connected layers.

4.2.2.2 Motion correction

The motion correction is formulated as a non-rigid registration problem when the first frame is considered as the reference frame. The deep feature vector for the reference frame is extracted using the fine-tuned CNN as described in the previous section. In this step of the work, the original images before applying the pre-processing are used to ensure that all the tissue information is considered during feature extraction. To perform the motion correction, other

Table 4.1 Learning Rates at Each Step of Fine-tuning the AlexNet Model In Our Experiments:μ and γ are fixed at 0.9 and 0.95 respectively at all the steps of fine-tuning. Learning rates are modified from fc7 to the third convolutional layer (conv3)

Layers	**Learning Rates**
fc7	0.10
fc6	0.01
Conv5	0.01
Conv4	0.01
Conv3	0.01
Conv2	0.00
Conv1	0.00

frames are considered as the moving frames. For the first step, by translating the images at each value of ρ, $\rho = \sqrt{x^2+y^2}$, deep feature vectors are extracted using CNN model and the cosine similarity is computed as a dot product of the feature vectors of the fixed and moving frames at each translation. To capture the rotation, at each rotational angle, the CNN features are extracted and the cosine similarity between the extracted feature vectors from reference and moving frames is computed. The same process is performed for scaling and shearing in both x and y directions to find the optimal parameters to correct vessel wall deformation. The final transformation matrix to map a pair of input coordinates P = (x,y) to a pair of output coordinates P$\prime$ = (x$\prime$, y$\prime$) can be summarized in the following equation,

$$P\prime = \boldsymbol{M}P \tag{4.3}$$

Where **M** is the final transformation matrix, which is defined as follows,

$$\begin{aligned}\mathbf{M} &= M_{shear} * M_{scale} * M_{rigid} \\ &= \begin{bmatrix} 1 & \lambda_y & 0 \\ \lambda_x & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_x & 0 & 0 \\ 0 & v_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} cos\theta & -sin\theta & t_x \\ sin\theta & cos\theta & t_y \\ 0 & 0 & 1 \end{bmatrix}\end{aligned} \tag{4.4}$$

Where M_{rigid}, M_{scale}, and M_{shear} are rigid transformation, scaling, and shearing matrices respectively. t_x and t_y are translation parameters along x and y axes respectively, θ is the angle

of rotation in 2D, v_x and v_y are scaling parameters, λ_x and λ_y are shearing parameters along x and y directions.

The objective function is defined based on the cosine similarity between the feature vectors of reference and moving frames as follows,

$$argmin_{\psi} f(\psi) = argmin_{\psi}(1 - cos(\psi)) \tag{4.5}$$

$$cos(\psi) = (\mathbf{a}.\mathbf{b}) / \mid \mathbf{a} \mid\mid \mathbf{b} \mid \tag{4.6}$$

Where ψ is the angle between the extracted feature vectors from reference and moving frames (**a** and **b** respectively) at each transformation. Unlike other similarity measures, cosine similarity is a measure of the direction-length correspondence between vectors, which is mostly used in high dimensional space and large-scale studies.

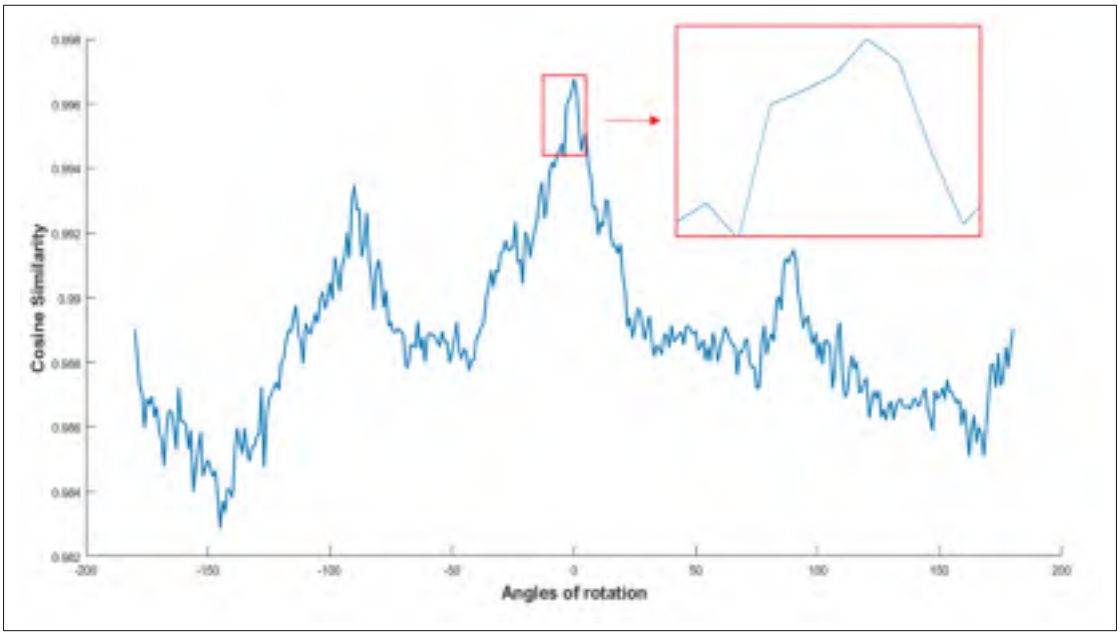


Figure 4.4 Cosine similarity between the feature vectors of fixed and moving frames: The calculations are preformed in a small interval of angles considering that the probe can not have a huge movement during image acquisition using OCT system.

Since we aimed to investigate the robustness of CNN features to be applied for motion correction, the optimal transformation parameters are solved using exhaustive search for all the possible transformations. At each translation, rotation, and deformation, we look for the angle which can maximize the cosine similarity of two feature vectors. Since using the OCT system, large rotation of the probe is not possible physically, we bound our exhaustive search with the small range of angles to make the process more accurate and also faster (see Figures 4.4 and 4.5). After obtaining the optimal transformation parameters, all the pre-processing steps are performed on the moving frame. The 3D matrix is built based on the information of the reference frame and the moving frames after applying the optimal transformations. The number of frames in each pullback is considered as the length of the z-axis.

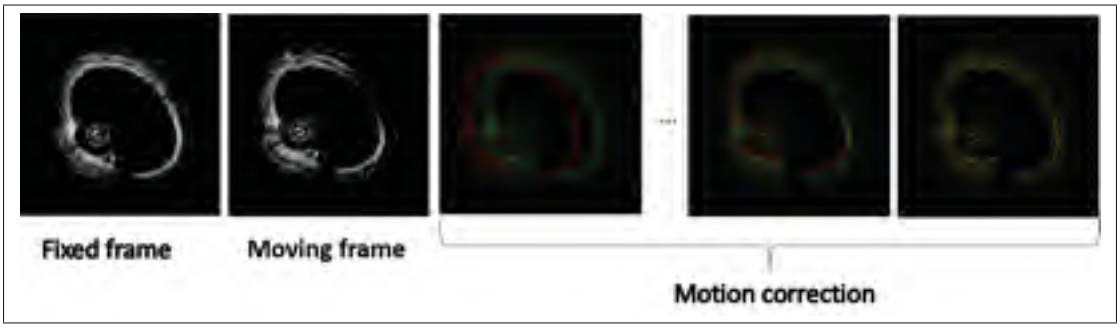


Figure 4.5 Motion correction process for one frame of a sequence: The first frame is considered as the reference frame, the second one is moving based on ρ and θ. At each transformation the cosine similarity between the deep feature vectors of both fixed and moving frames are computed. The process is terminated when the maximum similarity between the feature vectors is reached.

4.2.2.3 Validation

Since OCT is a new imaging system which is recently used in cardiology, to the best of our knowledge, there is no ground-truth available for OCT images. The motion correction of intracoronary OCT images is visually validated by the expert cardiologist. Also, the quantitative validations are performed in two steps:

- **The centerlines of the 3D models are compared against each other before and after motion correction:** For each sequence of images, the approximate center of the lumen for each frame is calculated before and after motion correction. Then, the Euclidean distance between the lumen center and the corresponding point on the fixed straight line centered on the lumen of the reference frame (the first frame of each pullback) is calculated using equation 4.7. The results are reported in Figure 4.6.

$$D(x,y) = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2} \tag{4.7}$$

  Where D is the Euclidean distance between the points (x_1, y_1) and (x_2, y_2), which are respectively the lumen center of each frame and the corresponding point on the fixed straight line centered on the lumen of the reference frame.

- **Intra-slice tissue alignment for each 3D model before and after motion correction:** To validate the alignment of tissues, we evaluated the alignment between the tissues of each frame and the previous one by estimating the joint entropy of two consecutive frames before and after motion correction (equation 4.8). The results are shown in Figure 4.7 for all 26 pullbacks.

$$H(f_1, f_2) = -\Sigma_j \Sigma_k P(j,k) log_2[P(j,k)] \tag{4.8}$$

  Where f_1 and f_2 are two consecutive frames with pixel values of j and k. P(j,k) is the joint probability of appearing the pixel values j and k at corresponding pixels in the two consecutive frames. Joint entropy is calculated for every two consecutive frames. Then, the result is reported as mean$\pm$std of all the calculations for each volume before and after motion correction.

4.3 Results and discussion

Motion correction is performed on 26 OCT pullbacks obtained from 26 different patients with KD. The results that are shown in Figure 4.6 demonstrate the measured Euclidean distance between all the points along the centerline of the 3D model and the corresponding points on the

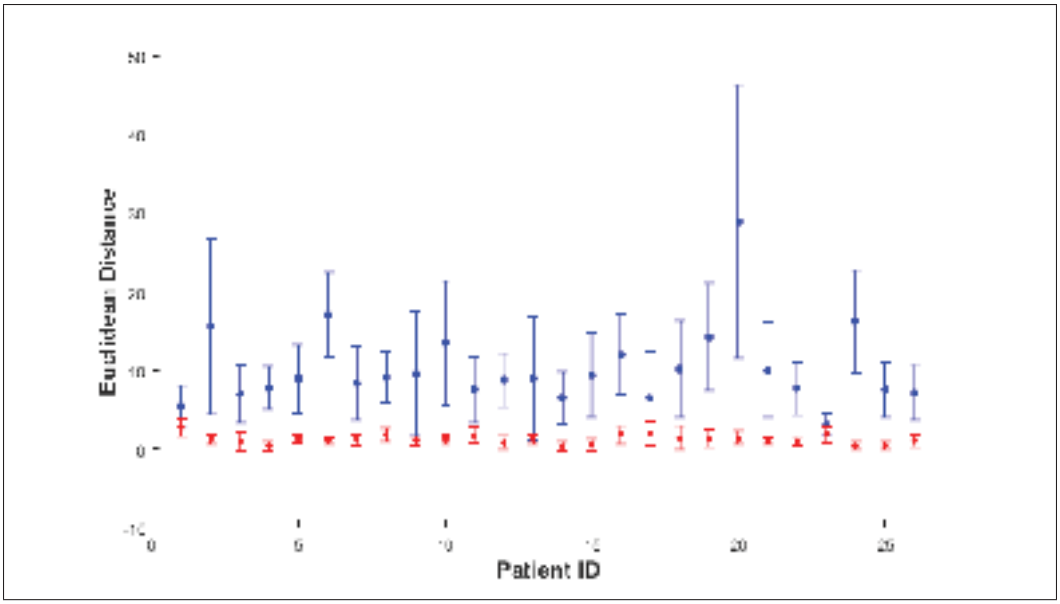


Figure 4.6 Comparison of Euclidean distances between the centerline of the 3D model and the reference centerline before and after motion correction which are shown in blue and red respectively. Error bars show the standard deviation. Euclidean distance is reported as mean$\pm$std for each patient.

fixed straight line centered on the lumen of the first frame before and after motion correction. For each patient, the mean value of all the measured Euclidean distances and standard deviations are calculated. As it is illustrated in Figure 4.6, there is a considerable improvement of measured values of Euclidean distance after motion correction. To ensure that all the aspects of quantitative validation of motion correction are considered, the alignment of the corresponding tissues are evaluated by estimating the joint entropy between the region of interests of every two consecutive frames for each volume before and after motion correction. As it is illustrated in Figure 4.7, the lower joint entropy after motion correction shows the higher dependency and good alignment of the tissues between the frames for each patient.

The results of motion correction of intracoronary OCT images are visually illustrated in Figures 4.8, 4.9, and 4.10 for three different patients. The centerlines of the 3D models show a good alignment of the frames since the centers of the frames are well aligned. Joint entropy is estimated for every two consecutive frames, which demonstrates the robustness of the proposed motion correction approach to align the tissues. The longitudinal cross-section of the 3D

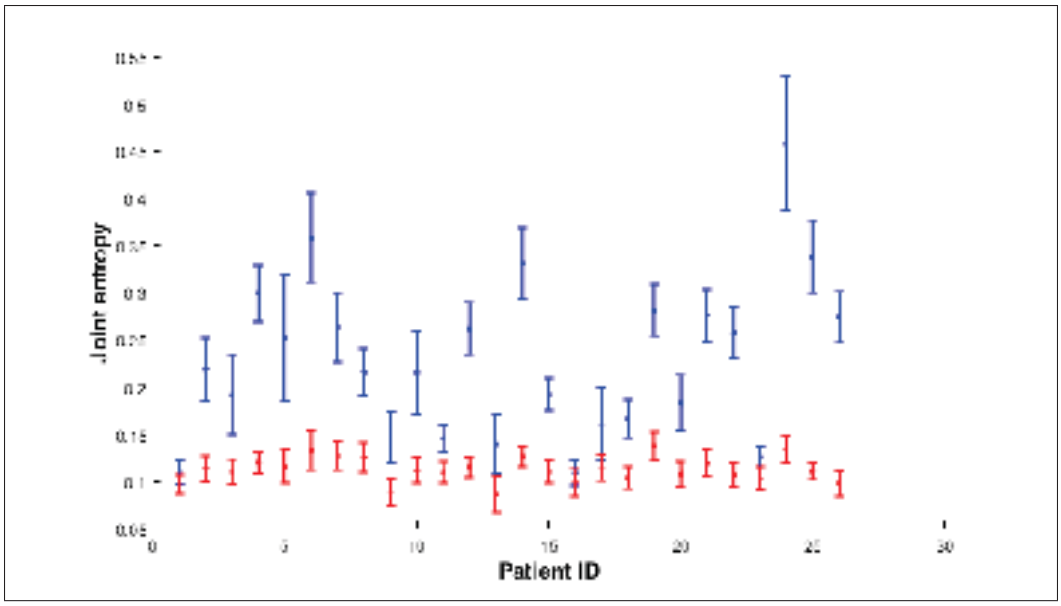


Figure 4.7 Intra-slice joint entropy to evaluate tissue matching before and after motion correction which are shown in blue and red respectively for all 26 pullbacks. Error bars show the standard deviation. Joint entropy is reported as mean±std for each patient.

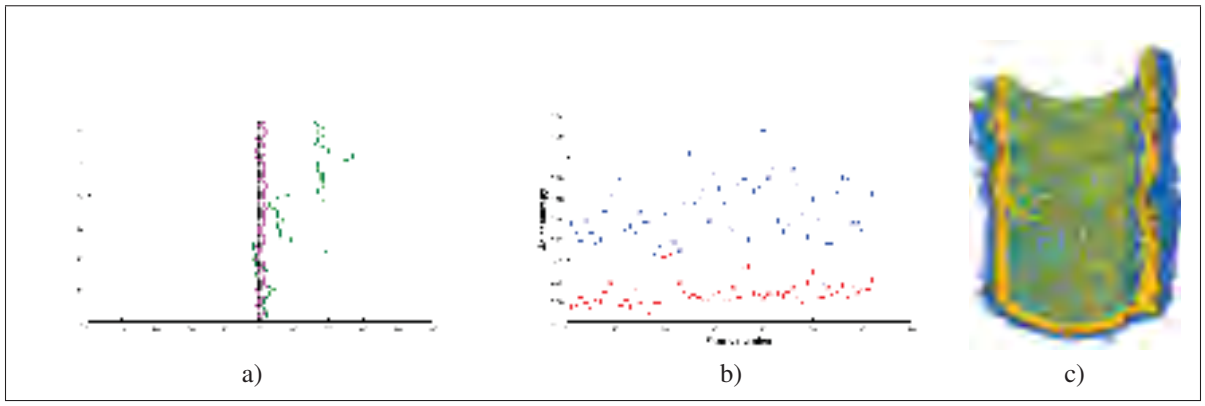


Figure 4.8 Patient 1: Motion correction of intracoronary OCT images: (a) shows the x-z projection of centerlines in green and purple before and after motion correction respectively compared against the fixed straight line centered on the lumen of the first frame (black). Joint entropy between each frame and the previous one is calculated in (b) for the whole sequence of frames before and after motion correction (blue and red respectively). The Longitudinal Cross-section is shown in (c) to illustrate tissue matching after motion correction.

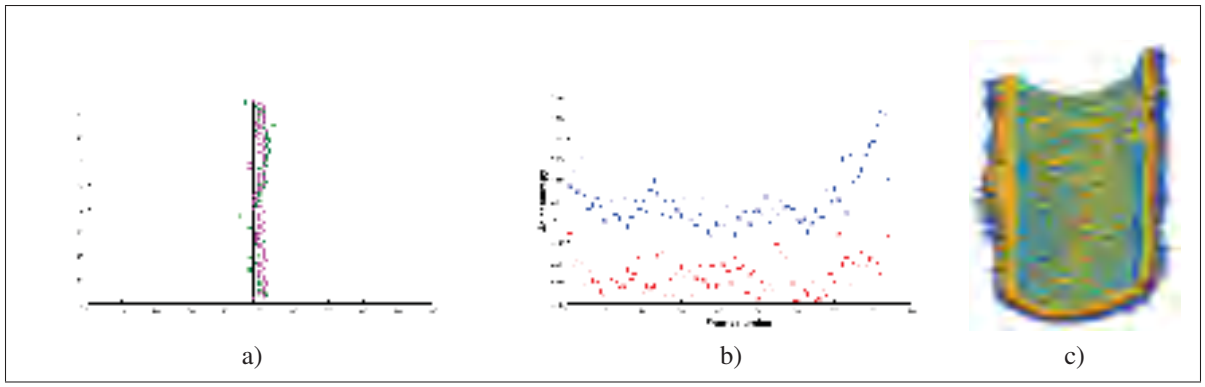


Figure 4.9 Patient 2: Motion correction of intracoronary OCT images: (a) shows the x-z projection of centerlines in green and purple before and after motion correction respectively compared against the fixed straight line centered on the lumen of the first frame (black). Joint entropy between each frame and the previous one is calculated in (b) for the whole sequence of frames before and after motion correction (blue and red respectively). The Longitudinal Cross-section is shown in (c) to illustrate tissue matching after motion correction.

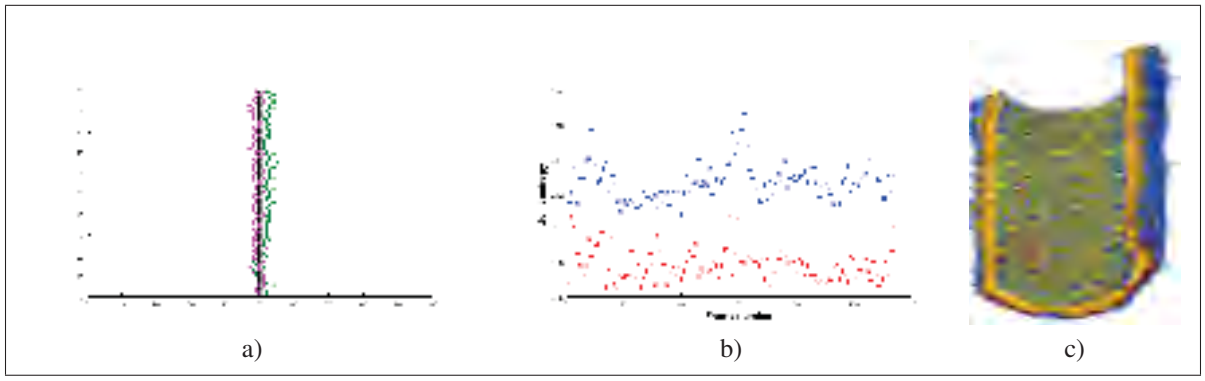


Figure 4.10 Patient 3: Motion correction of intracoronary OCT images: (a) shows the x-z projection of centerlines in green and purple before and after motion correction respectively compared against the fixed straight line centered on the lumen of the first frame (black). Joint entropy between each frame and the previous one is calculated in (b) for the whole sequence of frames before and after motion correction (blue and red respectively). The Longitudinal Cross-section is shown in (c) to illustrate tissue matching after motion correction.

model after motion correction shows a good alignment of different coronary artery layers. The parameters of transformation are changed slightly from one frame to the other as it is expected based on very small movements of the catheter and probe along the artery.

The proposed motion correction method is based on deep features extracted from fine-tuned CNN and is aimed to find the best possible correspondence between tissues to facilitate the 3D assessment of tissues and pathological formations. The results show the susceptibility of deep features to describe and characterize different coronary artery tissues while the images obtained from young children with the very small size of blood vessels, high heart rate, and lack of collaboration of the patient during the imaging process. Considering different artifacts caused by the imaging system, which definitely leave some deteriorating effects on the image quality, intravascular assessment of coronary artery tissues is a challenging task. In some cases, the cross-section is positioned at a side branch that can cause fold-over artifact and can complicate the tissue alignment process.

Our method overcomes the significant limitations of the proposed methods in the literature. Also, considering the physics of the imaging modality and the possible artifacts, we proposed a motion correction method, which is applicable to OCT images. First of all, tissue information is more reliable than lumen and media-adventitia borders to perform the registration, since it can be generalized to all challenging cases when the coronary artery is affected by the disease. As an example, intimal hyperplasia with media disappearance is one of the significant coronary artery complications caused by Kawasaki disease on the coronary arterial wall. Therefore, we do not have any information regarding the media-adventitia border. Also, lumen shape can be completely deformed by disease. In addition, the shadow of the catheter on arterial wall causes loss of information in OCT imaging. As a result, even in the least challenging cases, the lumen border does not appear in a complete circular shape. When the registration is based on the geometry of the lumen border, the algorithm tries to find transformation parameters that can match the lost part of the lumen border along each other. This sometimes causes large rotational angles, which is not possible in the reality of OCT imaging, since the probe has very small rotations during image acquisition. Also, the Gaussian filter that is used in some techniques to detect the lumen border is not reliable enough, since it is not edge preserving. In some studies, EM sensors are attached to the tip of the catheter to facilitate the accessibility of spatial position of the catheter. But, using EM sensors can reduce the manipulability of the

catheter. Using mutual information and joint entropy as objective functions for registration has some limitations. Both measures have a low performance with increasing noise. In monomodal image registration, both mutual information and joint entropy can be insensitive to intensity variations between the images. Therefore, they are used for multimodal image registration. From another perspective, deep features give us detailed information of the tissues, which are more reliable than the distribution of pixel values to be considered for motion correction and registration problem. The limitations of mentioned studies demonstrate the advantages of using deep features and tissue information for the problem of motion correction in OCT imaging.

Analysis of coronary artery tissues is a broad study field, which consists of four main steps: 1. Classification of coronary artery layers to recognize characteristic attributes of each layer. 2. Characterizing the abnormalities caused by the disease on coronary artery tissues. 3. Measuring the dimensions of the abnormal segments of the artery, which requires both longitudinal and transversal assessment of different pathological formations. 4. Evaluating the functionality of coronary arteries by estimating the stiffness of coronary artery tissues, which affects vascular elasticity and can reduce the vascular distensibility. This work is another complimentary study of our framework for tissue analysis of coronary artery in pediatric cardiology, which is open for any improvement.

4.4 Conclusion

In this study, a framework for motion correction of the OCT images is proposed. The main contribution of this work is the application of deep features in solving the problem of motion correction of intracoronary OCT images. This will contribute to evaluate the functionality of coronary arteries by analyzing the volume variation and considering the motion of the vessel. Also, it is a robust method to assess the pathological formations by finding the correlation between the tissues of adjacent frames, which are recognized using deep features. Since there is no technical method proposed in the literature for longitudinal assessment of the coronary artery abnormalities caused by KD in pediatric patients, this study paves the way for identifying

and precisely evaluating vascular wall abnormalities to prevent future complications in young adults who suffered from KD.

CHAPTER 5

DEEP CHARACTERIZATION OF CORONARY ARTERY TISSUES FROM OCT IMAGING

Atefeh Abdolmanafi[1], Luc Duong[1], Nagib Dahdah[2], Ibrahim Ragui Adib[3], Farida Cheriet[4]

[1] Dept. of Software and IT Engineering, École de technologie supérieure, Montréal, Canada
[2] Div. of Pediatric Cardiology and Research Center, Centre Hospitalier Universitaire Sainte-Justine, Montréal, Canada
[3] Div. of Cardiology, Hôpital Pierre Boucher, Longueuil, Canada
[4] Dept. of Computer Engineering, École Polytechnique de Montréal, Montréal, Canada

An invention disclosure is currently under review by Aligo for this study. This article will be submitted to Biomedical Optics Express.

ABSTRACT

Kawasaki Disease (KD) or mucocutaneous lymph node syndrome is an acute childhood vasculitis syndrome. Intimal hyperplasia, media disappearance, lamellar calcifications, fibrosis, macrophage, and neovascularization are the most distinguished pathological formations in coronary artery lesions. In severe cases, they can lead to myocardial infarction and sudden death. In-vivo intravascular imaging has been promising to assess the functionality of coronary artery tissues. Optical Coherence Tomography (OCT) is an interferometric imaging modality, which has been recently used in cardiology to characterize coronary artery tissues providing safety results and high resolution ranging from 10 to 20 μm. In this study, we investigate different deep learning models for robust tissue characterization to learn the various intracoronary pathological formations caused by KD from OCT imaging. Our approach evaluates deep features computed from three different pre-trained convolutional networks. Then, a majority voting approach is applied to provide the final classification result. The results demonstrate high values of accuracy, sensitivity, and specificity for each tissue (up to 0.99 ± 0.01). Hence, deep learning models and especially, majority voting method are promising for automatic interpretation of OCT images.

5.1 Introduction

Kawasaki Disease (KD), mucocutaneous lymph node syndrome, is an acute childhood vasculitis syndrome, which is the leading cause of coronary artery sequelae, complicated by coronary artery aneurysms with subsequent intimal hyperplasia, media disappearance, neovascularization, fibrosis, calcification, and macrophage accumulation (Newburger *et al.*, 2004; Group *et al.*, 2010). In the normal three-layered structure of coronary artery using OCT imaging, intima is characterized as a signal rich well-delineated layer and media appears as a homogeneous signal poor pattern specified by the internal and external elastic lamina. The outermost layer is adventitia, which is characterized as a signal rich layer (Newburger *et al.*, 2004; Orenstein *et al.*, 2012; Dionne *et al.*, 2015; Baim & Grossman, 1996). Intimal hyperplasia is the most frequent complication caused by KD, which is thickening of the intima and can be followed by media destruction since media becomes thinner and finally disappeared as a result of plaque accumulation and vessel remodeling. Intimal thickening can disturb oxygen diffusion and cause proliferation of vasa vasorum in inner layers of arterial wall, which is called neovascularization. Presence of neovascularization may be a sign of plaque instability and rupture and is characterized in OCT images as signal poor voids (Kitabata & Akasaka, 2013). Fibrosis is scarring of the connective tissues, which may occur as a result of arterial inflammation and is characterized as signal rich areas in OCT imaging. The macrophage may be accumulated within a fibrous cap as a result of monocytes differentiation in confronting with arterial wall inflammation. Macrophage is visualized as a confluent signal rich focal area in OCT imaging (Baim & Grossman, 1996; Jang *et al.*, 2002; Fujii *et al.*, 2010; Taguchi *et al.*, 2017; Swirski *et al.*, 2016). Vascular smooth muscle cells (VSMCs) regulate mineralization in intima and media. Rising lipid content within arterial lesions and inflammatory mediators may transform vascular smooth muscle cells into an osteoblast phenotype, resulting in intimal calcification. Calcification may be extended within a fibrous cap, which is visualized as a signal poor area with sharply delineated borders in OCT imaging (Baim & Grossman, 1996; Liu *et al.*, 2015; Madhavan *et al.*, 2014).

Cardiovascular Optical Coherence Tomography (OCT) is a catheter-based invasive imaging modality, which typically employs a near-infrared light to provide cross-sectional images of the coronary artery at depth of several millimeters relying on low-coherence interferometry. The unique characteristic of OCT is its high axial resolution of 10-15 μm, which is measured by the light wavelength and is decoupled from the lens dependent lateral resolution ranging from 20-40 μm. The image-wire is inserted into the coronary artery using an over-the-wire balloon catheter from patient's groin. A sequence of cross-sectional images of coronary artery segment is recorded using the backscattered light from the arterial wall through each pullback. Considering the fact that light can be attenuated by blood before reaching the vessel wall, blood clearance is required before starting the image acquisition (Bezerra *et al.*, 2009; Boudoux, 2016; Drexler & Fujimoto, 2015).

5.1.1 Related works

5.1.1.1 Tissue characterization

Progression of pathological formations caused by coronary artery disease can be followed by acute coronary syndrome (ACS). Therefore, it is significant to develop robust coronary artery tissue characterization techniques to evaluate pathological formations (Kawasaki *et al.*, 2002). While conventional imaging techniques such as CT and MRI may be used for clinical assessment of the coronary arteries, they are limited to providing useful information about the underlying coronary artery tissue layers. Also, they are restricted to reflect the histological reality of the regressed aneurysmal coronary segments, which are inappropriately considered as normal coronary segments (Orenstein *et al.*, 2012; Newburger *et al.*, 2004; Group *et al.*, 2010; Dionne *et al.*, 2015). Catheter-based Intravascular Ultrasound (IVUS) has been used for many years in interventional cardiology to evaluate coronary artery tissues by providing information on coronary arterial wall and lumen (Rathod *et al.*, 2015). IVUS imaging is restricted to be used in pediatric cardiology due to its suboptimal spatial imaging resolution (100-150 μm),

and low pullback speed. Arterial plaque formations are structural abnormalities, which require a high-resolution imaging modality to be detected (Orenstein *et al.*, 2012; Jang *et al.*, 2002).

Automated tissue analysis and plaque detection were focused on 2D intracoronary OCT images in adult patients to visualize plaque formations. Combination of light backscattering and attenuation coefficients have been estimated from intracoronary time domain OCT for three different atherosclerosis tissues, namely calcification, lipid pool, and fibrosis (Xu *et al.*, 2008). Van et al. detected fibrosis and calcification in coronary atherosclerosis by estimating the optical attenuation coefficient. The estimated values were compared with histopathological features of each tissue to determine the corresponding optical properties (Van Soest *et al.*, 2010). (Ughi *et al.*, 2013) proposed a tissue classification method using support vector machine (SVM) with the combination of texture features and optical attenuation coefficient extracted form atherosclerotic tissues. Volumetric estimation of backscattered intensity and attenuation coefficient was performed by (Gargesha *et al.*, 2015). Classification approach using SVM was used to discriminate between fibrosis, calcification, and lipid. Identification and quantification of fibrous tissue based on Short-Time Fourier Transform (STFT) using OCT imaging were proposed by (Macedo *et al.*, 2016). Gan et al. developed a classification framework to detect normal myocardium, loose collagen, adipose tissue, fibrotic myocardium, and dense collagen. Graph searching method is applied to segment various tissue layers of the coronary artery. Combination of texture features and optical properties of tissues is used to train a relevance vector machine (RVM) to perform the classification task (Gan *et al.*, 2016). A plaque tissue characterization technique based on intrinsic morphological characteristics of the A-lines using OCT imaging is proposed by (Rico-Jimenez *et al.*, 2016). to classify superficial-lipid, fibrotic-lipid, fibrosis, and intimal thickening by applying Linear Discriminant Analysis (LDA).

5.1.1.2 Convolutional Neural Network (CNN) and transfer learning

Convolutional Neural Networks (CNNs) have gained a wide popularity in medical image analysis. Application of CNNs in medical image analysis was first demonstrated in the work of (Lo *et al.*, 1995) for lung nodule detection. This idea was extended to various applications in the

field of medical imaging (Hochreiter & Schmidhuber, 1997; Eigen *et al.*, 2013; Zeiler & Fergus, 2014; Simonyan & Zisserman, 2014; Szegedy *et al.*, 2015; Roth *et al.*, 2015; Ciompi *et al.*, 2015; Havaei *et al.*, 2016). More recently, CNNs have been introduced in coronary OCT imaging to classify tissue layers between media and intima using SVM and Random Forest classifiers (Abdolmanafi *et al.*, 2017).

Transferability is defined as transferring the knowledge embedded in the pre-trained CNNs for other applications, which is performed in two different ways: Using a pre-trained network as feature generator and fine-tuning a pre-trained network to be used for classification of medical images. Common networks, which are used as pre-trained models with applications in medical image analysis are categorized into three groups. Simple networks with few convolutional layers use kernels with large receptive fields in upper layers close to the input and smaller kernels in deeper layers. The popular network in this group, which has a broad application in medical image analysis is AlexNet and is introduced by (Krizhevsky *et al.*, 2012; Litjens *et al.*, 2017). The second group of architectures is deep networks such as Vgg models. They have the same configuration as simple networks with more convolutional layers and kernels with smaller receptive fields (Simonyan & Zisserman, 2014; Litjens *et al.*, 2017). The third group of networks is categorized as complex building blocks with higher efficiency of the training process compared to other groups of networks. GoogleNet was the first network in this category, which is introduced by Szegedy et al. ResNet and Inception models are other networks of this group. An improved version of GoogleNet, which is used recently in the field of medical image analysis is Inception-v3 (Szegedy *et al.*, 2016; Szegedy *et al.*; Litjens *et al.*, 2017). Vgg-16, VGG-M-128, and BVLC reference CaffeNet are used as feature extractors to classify the knee osteoarthritis (OA) images by training SVM using deep features (Antony *et al.*, 2016). Kim et al. used the pre-trained network for the application of cytopathology classification. It is demonstrated in their work that using a pre-trained network as feature extractor outperforms the fine-tuning of the network (Kim *et al.*, 2016). Esteva et al. fine-tuned a version of Inception-v3 model to classify skin lesions. Gulshan et al. used the fine-tuned network to evaluate the retinal fundus photographs from adults by detecting referable diabetic

retinopathy. In these studies, it is demonstrated that the results of classification using fine-tuned network competes against the human expert performance (Esteva *et al.*, 2017; Gulshan *et al.*, 2016). Nevertheless, most of the studies are focused on fine-tuning the networks and comparison of the results of fine-tuned networks with the results of other classifiers, which are trained on deep features.

In this study, we focused on designing a tissue characterization model to detect pathological formations in coronary artery tissues using OCT imaging. To find the optimal model to classify calcification, fibrosis, neovascularization, macrophage, and coronary artery layers, intima, and media, we performed four different experiments. Then, the optimal model is chosen by applying three different pre-trained networks as feature extractors. Random Forest is trained on each set of deep features separately to perform classification. classification results of Random Forest using three various set of deep features are combined using majority voting approach to provide the final classification result. Hence, the contributions of this study are:

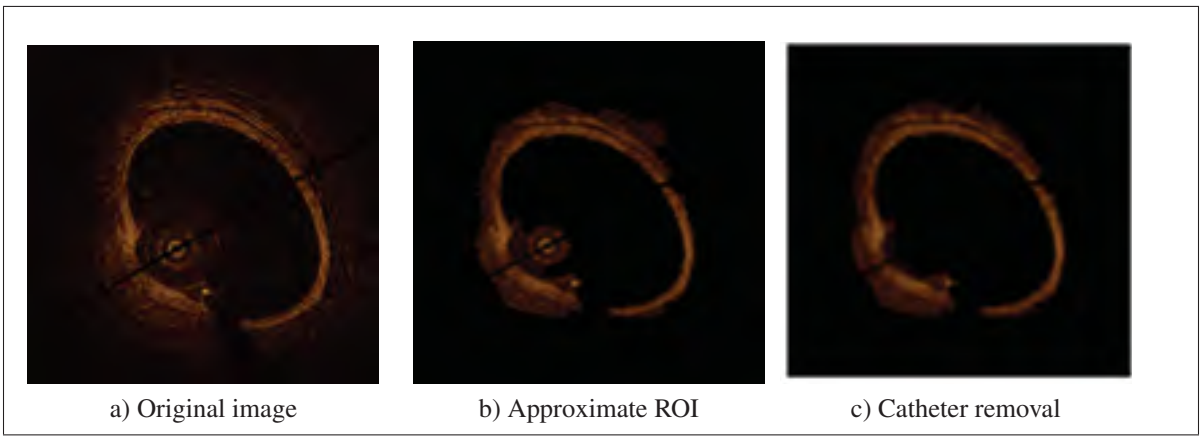
- Characterization of complex pathological formations in KD from OCT imaging, namely intimal hyperplasia, media disappearance, neovascularization, fibrosis, calcification, and macrophage accumulation.
- Evaluation of different pre-trained CNN models for OCT image analysis with a limited labeled dataset.
- Assessment of the clinical usefulness of deep feature learning for OCT imaging in pediatric cardiology.

This work is organized as follows. First, data collection and pre-processing are explained in section 5.2.1. Second, Convolutional Neural Networks (CNNs) and pre-trained network architectures are developed in section 5.2.2. The training of CNN and validation are presented in Section 5.2.3. The results of the experiments are reported and they are discussed in section 5.3. Finally, this study is concluded in section 5.4.

5.2 Material and method

5.2.1 Data collection and pre-processing

The experiments are performed on 33 pullbacks of intracoronary cross-sectional OCT images of patients affected by KD. The images are acquired using the ILUMIEN OCT system (St. Jude Medical Inc., St. Paul, Minnesota, USA) with the axial and lateral resolutions of 12-15 μm and 20-40 μm respectively. Image acquisition is performed using FD-OCT with pullback speed of 20 mm/sec and frame rate of 100 frames/sec. This study was approved by our institutional review board. For the first step, the pre-processing is performed on all the frames of each sequence by automatic recognition and removal of the guide-wire. Then, the approximate region of interests including the lumen, intima, media, calcification, neovascularization, macrophage, fibrosis and surrounding tissues are detected for each pullback frame using active contour. The catheter and unwanted blood cells are removed by applying the smallest connected components approach (Figure 5.1). The images were converted to planar by transferring all the points from Cartesian coordinates to planar representation in Polar coordinates to simplify the calculations.



a) Original image b) Approximate ROI c) Catheter removal

Figure 5.1 Pre-processing steps: (a) Original image, (b) ROI detection using active contour, (c) Applying smallest connected components approach to remove catheter and unwanted blood cells.

5.2.2 Convolutional Neural Network (CNN)

CNNs are built on convolutional layers, which are responsible to extract features from the local receptive field of the input image. Each convolutional layer consists of n sets of shared weights between the nodes to find similar local features in the input channels, which are called convolutional kernels. Each kernel creates a feature map when it slides through the whole input image with a defined stride. Feature maps extracted from one convolutional layer will be the input of the next layer (Litjens *et al.*, 2017). It is standard to calculate the output of a neuron by applying a hyperbolic tangent or logistic regression, which are both saturating activation functions. Saturating nonlinearities are slower than non-saturating non-linearities while stochastic gradient descent is used to minimize the cost function with respect to the weights at each convolutional layer. Therefore, a non-saturating activation function, which is called Rectified Linear Unit (ReLU) can accelerate the training process by keeping non-negative values and replacing negative values by zero in the feature map. ReLU is defined as follows:

$$f(x) = max(0,x) \tag{5.1}$$

Where f is the output of each neuron as a function of its input x (Krizhevsky *et al.*, 2012). CNNs alternate between the convolutional and pooling layers to achieve computational efficiency, since pooling layers are used for dimensionality reduction by aggregating the outputs of neurons at one convolutional layer and reducing the size of the feature maps. Pooling layers can keep the network invariant to small transformations, distortions, and translations in the input image as well as control overfitting by reducing the number of parameters and computations. Max pooling is used in most of the CNN architectures, which choose the superior invariant features in specified neighborhoods of a feature map (Krizhevsky *et al.*, 2012).

CNNs are trained using back-propagation algorithm and stochastic gradient descent is commonly used to minimize the following cost function:

$$L = -(1/|X|)\Sigma_j^{|X|} ln(p(y^j|X^j)) \tag{5.2}$$

Where X is the size of the training set and $ln(p(y^j|X^j)$ denotes the probability of j^{th} image to be classified correctly with the corresponding label y. for each layer of the network, the weights are updated at each iteration i as follows:

$$V_{i+1} = \mu V_i - \gamma_i \alpha \partial L / \partial W \tag{5.3}$$

$$W_{i+1} = W_i + V_{i+1} \tag{5.4}$$

Where μ is the momentum, α is the learning rate, γ is the scheduling rate, which reduces the learning rate at the end of iterations and W is the weight at each iteration i for each layer (Krizhevsky *et al.*, 2012; Abdolmanafi *et al.*, 2017).

5.2.2.1 Pre-trained networks

Pre-trained networks are widely used as both feature extractor and classifier for different tasks. Among the most common architectures, we selected three pre-trained networks with different architectures. AlexNet is a simple and shallow network, which is popular for clinical applications. The network consists of five convolutional layers, and three fully connected layers, which are followed by a final softmax with GPU implementation of the convolutional operation. The model is trained on 1.2 million images from the ImageNet dataset, which are annotated and categorized into 1000 semantic classes. The model uses 60 million parameters and consists of 650000 neurons, which is trained using stochastic gradient descent with the batch size of 128, momentum of 0.9, and weight decay of 0.0005 to reduce the training error of the model (Krizhevsky *et al.*, 2012). The network architecture is shown in Figure 5.2.

Deeper models were designed by stacking convolutional layers to increase the depth of the network. Instead of using a large receptive field, kernels with very small receptive field and fixed size were applied in each convolutional layer. Every set of convolutional layers is followed by a max pooling to reduce dimensionality, and every convolutional layer is followed by a ReLU to introduce non-linearity. Vgg networks are trained on 1.2 million images of 1000 classes from ImageNet. The batch size and momentum are set to 256, and 0.9 respectively. The learning rate

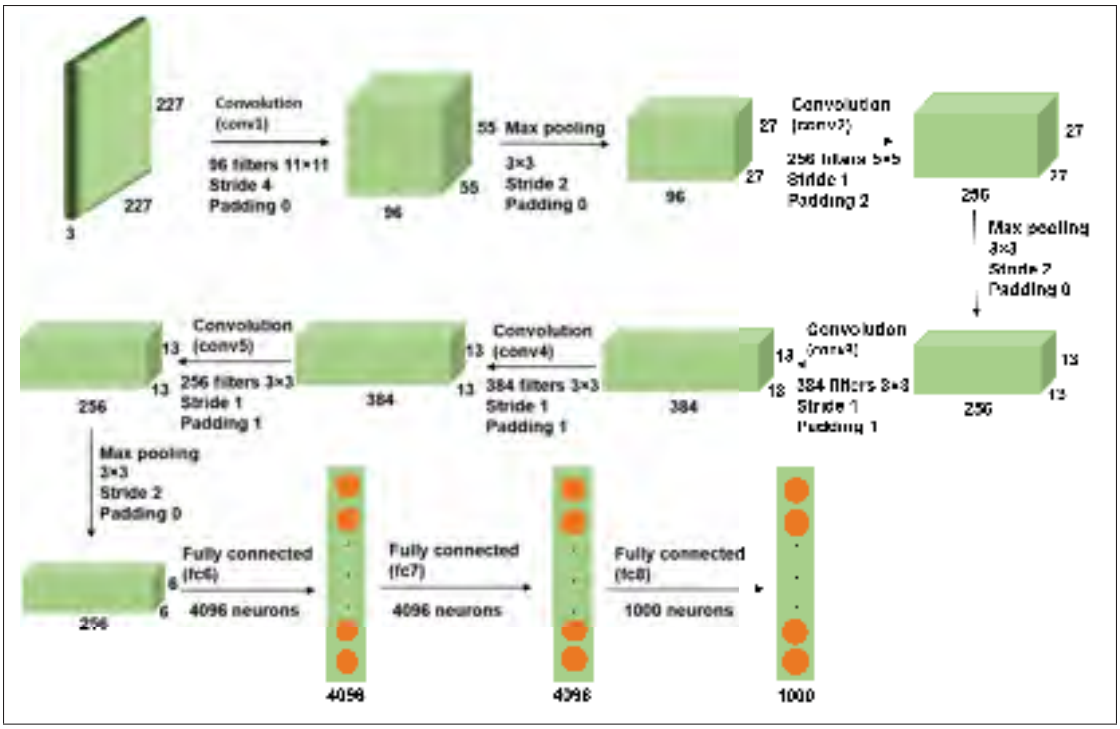


Figure 5.2 AlexNet architecture consists of five convolutional layers, and three fully connected layers.

was initialized to 0.01 and was decreased by the factor of 10 when the accuracy on validation set stopped improving (Simonyan & Zisserman, 2014). Among deep network architectures of Vgg, we selected Vgg-19 with 144 millions of parameters and deeper network architecture consists of 16 convolutional layers, and three fully connected layers, which is shown in details in Figure 5.3.

Complex building blocks (inception blocks) are introduced as models with the fewer numbers of parameters and higher efficiency of the training process by replacing the fully connected architectures with sparsely connected architectures. The network has been built from convolutional building blocks called inception modules, which are stacked on top of each other. Each inception module consists of combination of convolutional layers with kernel sizes of 1×1, 3×3, and 5×5, which their output filter banks concatenated into a single output vector that will be the input of the next stage. 1×1 convolutions in each inception module is used for dimensionality reduction before applying computationally expensive 3×3 and 5×5 convolu-

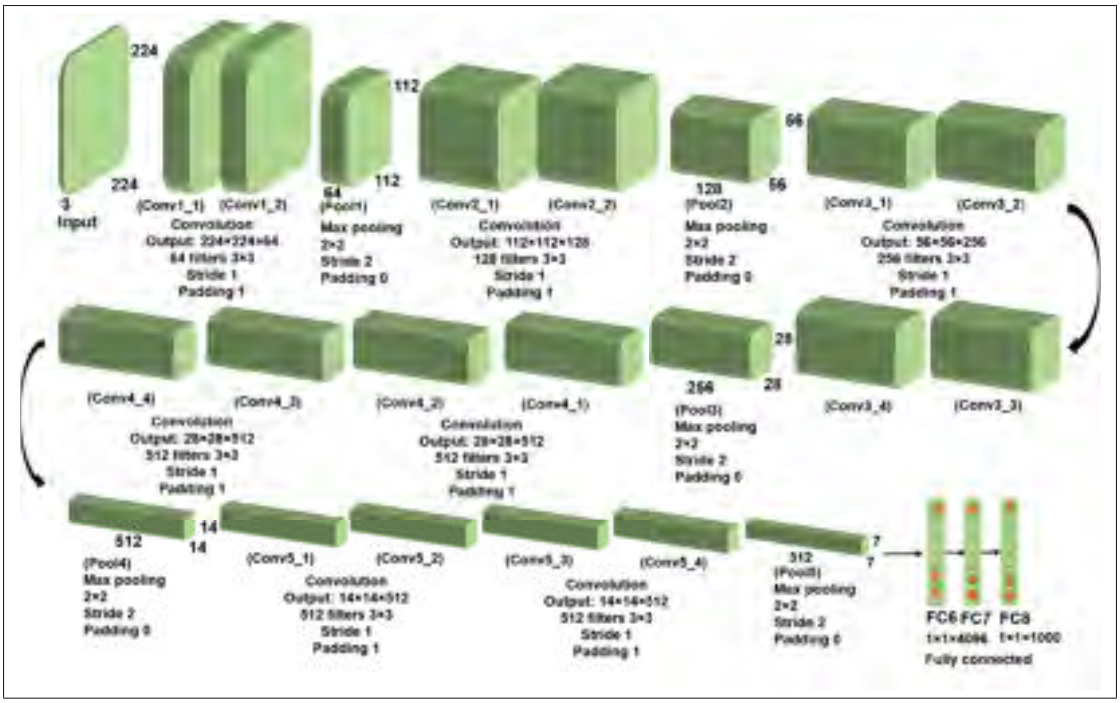


Figure 5.3 Vgg-19 architecture consists of sixteen convolutional layers, and three fully connected layers.

tions. Factorization of convolutions into smaller convolutions results in aggressive dimension reduction inside the network, which leads to the fewer numbers of parameters and low computational cost. Inception models are trained using stochastic gradient descent. Batch size is chosen as 32 for 100 epochs and momentum with the decay of 0.9. Learning rate is initialized by 0.045 and decayed every second epoch by the exponential rate of 0.94 (Szegedy *et al.*, 2016; Szegedy *et al.*). Pre-trained Inception-v3 is used in our experiments. The inception models are updated in this version of the network to further boost ImageNet classification accuracy. The last part of the network, which is used for fine-tuning in our experiments is shown in Figure 5.4.

5.2.3 Training and validation

In our experiments, the total of ∼5100 different tissues are extracted from OCT pullback images and are manually labeled as calcification, fibrosis, intima, macrophage, media, and neo-

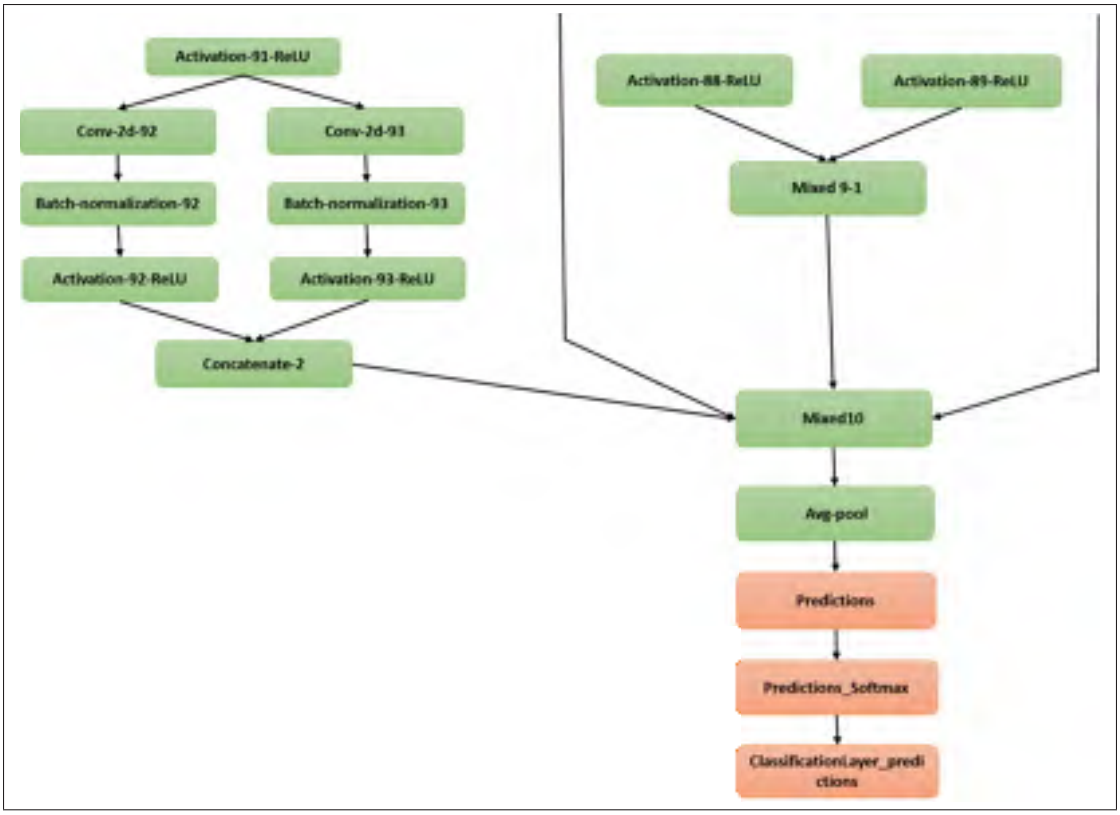


Figure 5.4 Last layers of Inception-v3 architecture

vascularization. To start the experiments, 66% of the ROIs are selected randomly as the training set. To avoid any correlation between training, test, and validation sets, 50% of the remaining ROIs are randomly selected as the validation set and test set is built on the last residual ROIs. The experiments are performed in four different steps to find the optimal tissue characterization framework.

5.2.3.1 Classification using fine-tuned networks

For each convolutional neural network, before starting the training process, the iterative weight update is performed by random weight initialization at each layer of the network. Since the number of labeled data is limited in our experiments, weight initialization can be performed using the weights of the pre-trained networks. Therefore, iterative weight updates of equations 5.3 and 5.4 lead to a fast convergence to find the desirable local minimum for the cost function

(equation 5.2). Considering the fact that number of nodes in the last fully connected layer depends on the number of classes in each dataset, the first step before starting the fine-tuning is to keep the same architecture as the pre-trained network architecture, remove the classification layers and replace them by the layers, which are designed appropriately for our classification task. The next step is to initialize the weights at each layer of the network with the weights of the pre-trained network, which is called transfer learning. the iterative weight update can be started using layer-wise fine-tuning by finding the optimal learning parameters at each convolutional and fully connected layer. Fine-tuning AlexNet for classification of coronary artery layers (intima and media) is preformed by our group Abdolmanafi *et al.* (2017). Since, the goal of this study is to develop a tissue characterization model to detect pathological formations (calcification, fibrosis, macrophage, and neovascularization) as well as the arterial wall layers, the process of fine-tuning the pre-trained AlexNet is improved based on our new dataset. the last three layers of the pre-trained network (fc8, prob, and classification layer) are replaced by a set of layers which are designed for multi-class classification task to classify calcification, fibrosis, macrophage, neovascularization, intima, and media. The depth of fine-tuning is increased compared against our previous work since AlexNet is not a very deep model, low-level features extracted from the first convolutional layers can improve the classification precision while we are dealing with various tissue labels to perform the classification task. the values of μ and γ are kept at 0.9 and 0.95 respectively and the learning rate for the last fully connected layers (fc6, fc7, and fc8) is set to 0.1 to learn faster in the last layers and we started decreasing the learning rates to 0.01 from the last convolutional layer (Conv3).

Since by adding convolutional layers with corresponding kernels, we will have access to detailed image information, increasing the depth and width of the network can improve the quality of the network architecture. To have a fair comparison among the performance of pre-trained networks, we selected Vgg-19 from the category of very deep CNN architectures. As it is explained in the previous section, Vgg-19 has almost the same configuration of the AlexNet with more convolutional layers. Therefore, fine-tuning the Vgg-19 is performed using the same strategy that is applied to AlexNet. We started fine-tuning by removing the classification

layers (fc8, prob, and output) and replaced them by a set of layers, which are appropriate for multi-class classification of various coronary artery tissues (calcification, fibrosis, macrophage, neovascularization, intima, and media). We started fine-tuning from the last fully connected layer (fc8) and increase the depth of fine-tuning gradually by evaluating the network performance at each fine-tuning level. To find the optimal parameters at each level of fine-tuning, an interval of values close to the optimal values of fine-tuned AlexNet is chosen. Then the optimal parameters are determined by grid searching for the defined interval of values and evaluating the performance of the network at each step. The best performance of the network obtained by assigning fixed values of 0.8, and 0.85 to μ and γ respectively. The learning rate is determined as 0.2 for the last fully connected layers (fc6, fc7, and fc8) and is decreased to 0.01 from the last convolutional layer (Conv5-4).

complex building blocks are very deep network architectures, which uses the particular configuration of inception modules to reduce the number of parameters and consequently improve the efficiency of the training procedure. We selected Inception-v3 from the category of complex network architectures to perform our experiments. Considering the complexity of the Inception architectures, changing the network can interfere with computational gains. Therefore, it is more difficult to adapt these types of networks to a new classification task. To fine-tune the network, we removed the last layers of the network (predictions, predictions-softmax, and ClassificationLayer-predictions), which aggregating the extracted features from the network for classification task and added a new set of classification layers adapted to our data set to the network graph. The new layers are connected to the transferred network graph and the learning rate for the fully connected layer is set to 0.1.

At each step of fine-tuning for all networks, the accuracy is calculated on the validation set and the training process is stopped when the highest accuracy on the validation set is obtained. By terminating the training process, classification is performed on the test set using each fine-tuned network separately.

5.2.3.2 Training Random Forest using deep features generated by pre-trained networks

In this experiment, pre-trained networks are used as feature generators. The activations extracted from the last layer before classification layer is used to train Random Forest to classify various coronary artery tissues. Using AlexNet, and Vgg-19, features are extracted from the last fully connected layer right before the classification layer (fc7). Each feature vector represents 4096 attributes of the labeled tissue. Using Inception-v3, features are extracted from the last depth concatenation layer (mixed10). Each feature vector represents 131072 attributes of the labeled tissue. It is demonstrated in our previous work that Random Forest is a robust classifier with quick training process and low risk of overfitting. It works based on generating an ensemble of trees. The trees are grown based on the CART methodology to maximum size without pruning. Generalization error for Random Forest classifier is proportional to the ratio ρ/s^2, which (s) and (ρ) are respectively defined as the strength of the trees and correlation between them. the smaller this ratio results in the better performance of Random Forest Criminisi & Shotton (2013); Kuhn & Johnson (2013). To find the optimal number of trees, The performance of Random Forest is evaluated for 1000 of trees while it is trained on each set of features extracted from each network separately. The OOB error rate is stopped decreasing when the tree number is assigned to 250 using the features extracted from Inception-v3, and Vgg-19, and 300 using the features extracted from AlexNet (see Figure 5.5). The fewer number of trees accelerates the training process by reducing the computational complexity. The number of randomly selected predictors (m_{try}) is set to 7.

Training features extracted from each pre-trained network and used separately to train Random Forest. Classification is performed on the test set using the test features extracted by each pre-trained network.

5.2.3.3 Classification using majority voting

Inspired by the ensemble learning approaches, we applied weighted majority voting on the classification results obtained by the second experiment. Classification is performed by Ran-

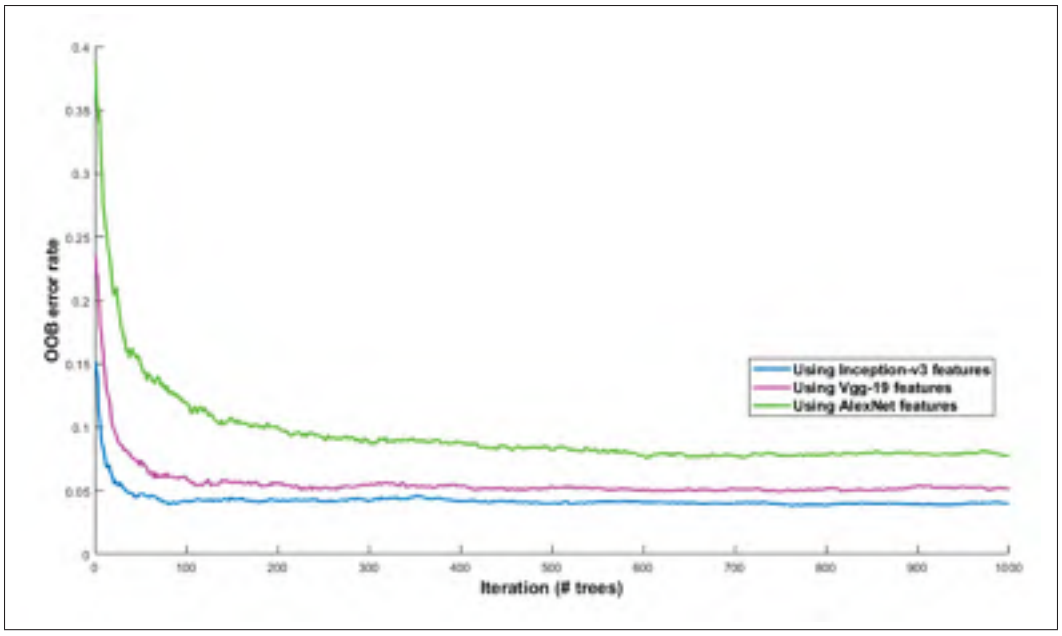


Figure 5.5 OOB error rate is calculated to find the obtimal number of trees to train Random Forest model. The performance of Random Forest is evaluated by calculating OOB errors while it is trained on each set of features extracted from each network separately. The OOB error rate is calculated for 1000 of trees

dom Forest using the features extracted from AlexNet, Vgg-19, and Inception-v3. Using the following equation, for weighted majority voting, weights are set to 1/3 for all the three sets of classification results except those labels with three different tissue labels.

$$C(x) = argmax_i \sum_j w_j I(C_j(x) = i) \tag{5.5}$$

Where C(x) is the classification label with the majority vote, i is the class label (it can be varying from 1 to 6 for calcification, fibrosis, intima, macrophage, media, and neovascularization), w_j is the weight of j^{th} tissue label and I is the indicator function. Thus, majority voting is applied to search in all the classification labels for the most frequent label assigned to each tissue using equation (5.6).

$$C(x) = mode\{C_1(x), C_2(x), C_3(x)\} \tag{5.6}$$

Where $C_1(x)$, $C_2(x)$, and $C_3(x)$ are Random Forest classification results using the features extracted from AlexNet, Vgg-19, and Inception-v3 respectively. Since the mode of $C_1(x)$,

$C_2(x)$, and $C_3(x)$ when $C_1(x) \neq C_2(x) \neq C_3(x)$ gives us the smallest tissue label as the majority vote, we put more weight on the third group of predicted labels if $C_1(x) \neq C_2(x) \neq C_3(x)$ considering the strength of deep Inception-v3 features. Therefore, the majority vote will be on the class label with the highest probability of belonging to the true class label.

5.2.3.4 RF classification using deep feature fusion

To consider all possible ways to find the optimal tissue characterization framework, we combined the features obtained from AlexNet, and Vgg-19 to train Random Forest. Classification is performed on the test set and the results are compared against the previous experiments. The features extracted from Inception-v3 is not used in this experiment since the size of the feature matrix is huge to be combined with other feature matrices.

5.3 Results and discussion

For each experiment, the classification is performed to characterize six different coronary artery tissues (calcification, fibrosis, macrophage, neovascularization, intima, and media). The results obtained for each experiment as follows,

5.3.1 Classification using fine-tuned networks

For the first experiment, fine-tuning is performed on AlexNet, Vgg-19, and Inception-v3 from different categories of simple architectures, very deep architectures, and complex networks respectively. The optimal fine-tuning parameters are estimated and the networks are trained by assigning the new learning parameters. Classification is performed by each network separately and accuracy, sensitivity, and specificity are measured using the corresponding confusion matrix for each network. The results are shown in Figures 5.6-5.8, and Tables 5.1-5.3.

The results of the experiments demonstrate the higher performance of Vgg-19 and Inception-v3 compared against AlexNet, which was expected considering the deep structure of Vgg-19, and Inception-v3 architectures. Although using pre-trained networks reduce the computational

burden, which results in reducing the training time and convergence issues, but a considerable amount of time is still required to find the optimal learning parameters and retrain the fine-tuned networks (approximately two hours for each network). Also, there is a risk of overfitting in deep fine-tuning a network. The following steps are proposed to find the optimal tissue characterization model, which can overcome the mentioned issues in an efficient way.

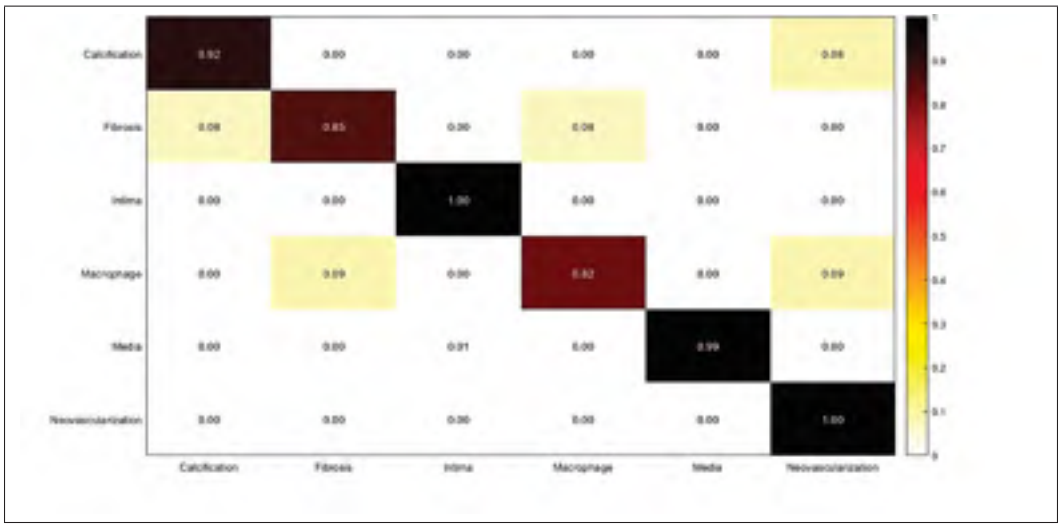


Figure 5.6 Confusion matrix of intracoronary tissue classification using fine-tuned AlexNet

Table 5.1 Measured sensitivity, specificity, and accuracy of tissue classification using fine-tuned AlexNet

Tissue	Accuracy	Sensitivity	Specificity
Calcification	0.95	0.92	0.99
Fibrosis	0.92	0.85	0.99
Intima	1.00	1.00	1.00
Macrophage	0.89	0.82	0.97
Media	1.00	0.99	1.00
Neovascularization	0.98	1.00	0.97

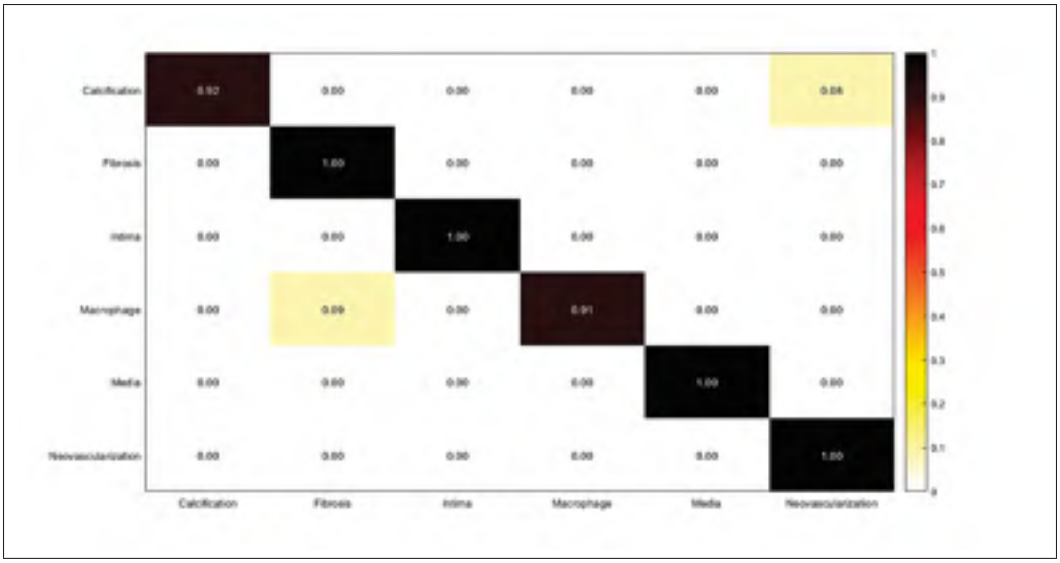


Figure 5.7 Confusion matrix of intracoronary tissue classification using fine-tuned Vgg-19

Table 5.2 Measured sensitivity, specificity, and accuracy of tissue classification using fine-tuned Vgg-19

Tissue	Accuracy	Sensitivity	Specificity
Calcification	0.95	0.92	0.99
Fibrosis	1.00	1.00	1.00
Intima	1.00	1.00	1.00
Macrophage	0.95	0.91	1.00
Media	1.00	1.00	1.00
Neovascularization	0.99	1.00	0.99

Table 5.3 Measured sensitivity, specificity, and accuracy of tissue classification using fine-tuned Inception-v3

Tissue	Accuracy	Sensitivity	Specificity
Calcification	1.00	1.00	1.00
Fibrosis	0.96	0.92	0.99
Intima	1.00	1.00	1.00
Macrophage	0.95	0.91	0.99
Media	1.00	1.00	1.00
Neovascularization	0.97	0.95	1.00

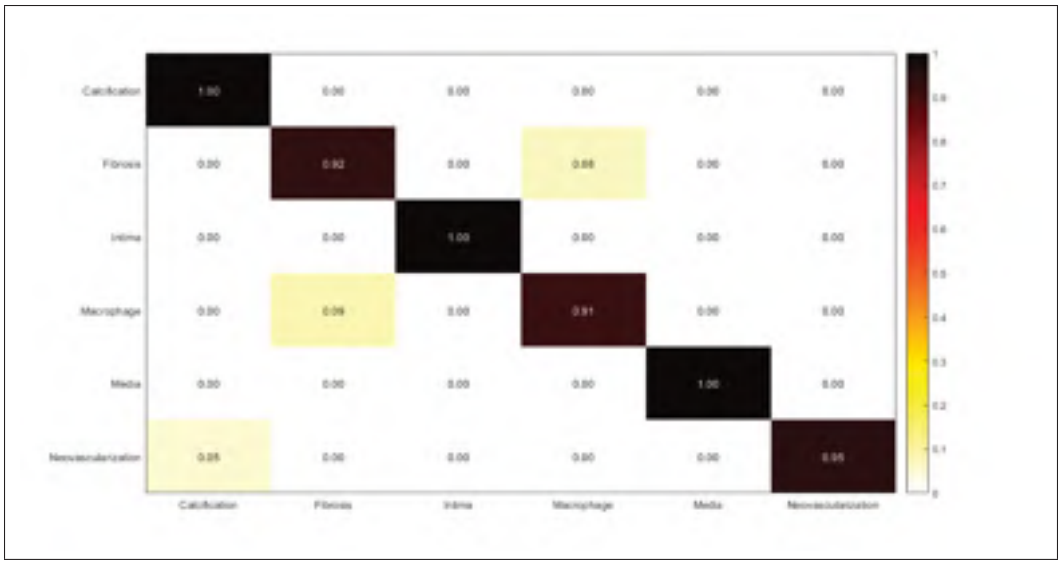


Figure 5.8 Confusion matrix of intracoronary tissue classification using fine-tuned Inception-v3

5.3.2 Training Random Forest using deep features generated by pre-trained networks

In this experiment, deep features are extracted from AlexNet, Vgg-19, and Inception-v3. By applying each network separately as feature generator, the training features are extracted to train Random Forest and the classification is performed on the test set. Features are extracted from the last fully connected layer before the classification layer (fc7) in AlexNet, and Vgg-19 architectures, and the last depth concatenation layer (mixed10) in Inception-v3 architecture. Accuracy, sensitivity, and specificity are measured using the corresponding confusion matrix for each classification result, which are shown in Figures 5.9-5.11, and Tables 5.4-5.6.

Regardless of the time, which is spent to find the optimal learning parameters, the process of feature extraction from all the three networks, and training the Random Forest using each set of features takes approximately twice less time than retraining a network. Using pre-trained networks as feature extractor overcomes the problems of fine-tuning, training time, and over-fitting concerns. But, the classification performance is not as high as using CNNs as classifiers (Figures 5.14-5.16). To solve this problem, the following two experiments are performed and the results of all experiments compared with each other.

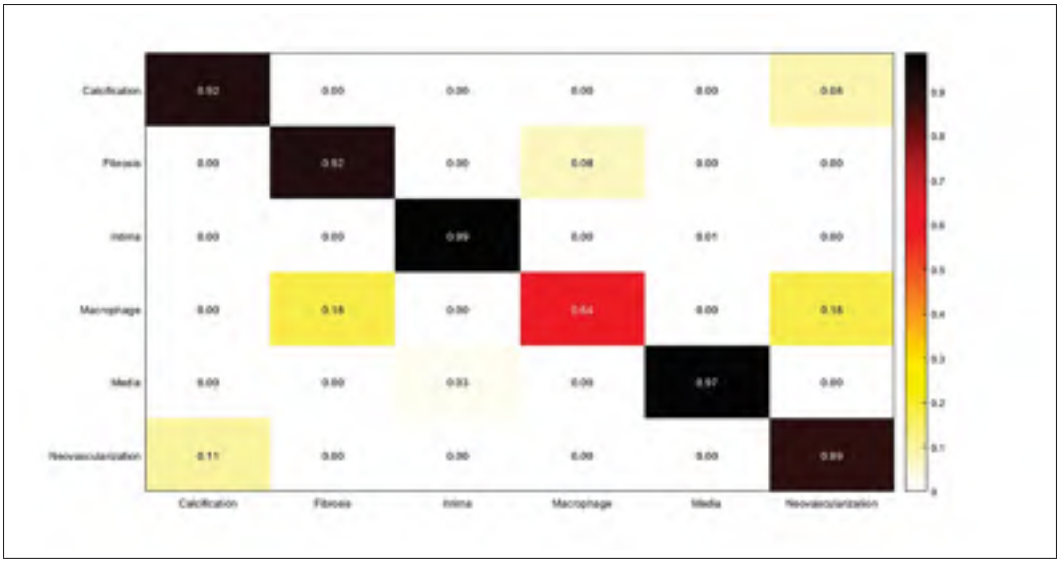


Figure 5.9 Confusion matrix of intracoronary tissue classification: Random Forest is trained using the deep features extracted from AlexNet

Table 5.4 Measured sensitivity, specificity, and accuracy of tissue classification using RF. Features are extracted from AlexNet.

Tissue	Accuracy	Sensitivity	Specificity
Calcification	0.95	0.92	0.98
Fibrosis	0.95	0.92	0.99
Intima	0.99	1.00	1.00
Macrophage	0.80	0.64	0.95
Media	0.99	1.97	1.00
Neovascularization	0.92	1.89	0.95

Table 5.5 Measured sensitivity, specificity, and accuracy of tissue classification using RF. Features are extracted from Vgg-19

Tissue	Accuracy	Sensitivity	Specificity
Calcification	1.00	1.00	1.00
Fibrosis	0.92	0.85	0.99
Intima	0.98	0.97	0.99
Macrophage	0.91	0.82	1.00
Media	0.98	0.97	0.98
Neovascularization	0.97	0.95	1.00

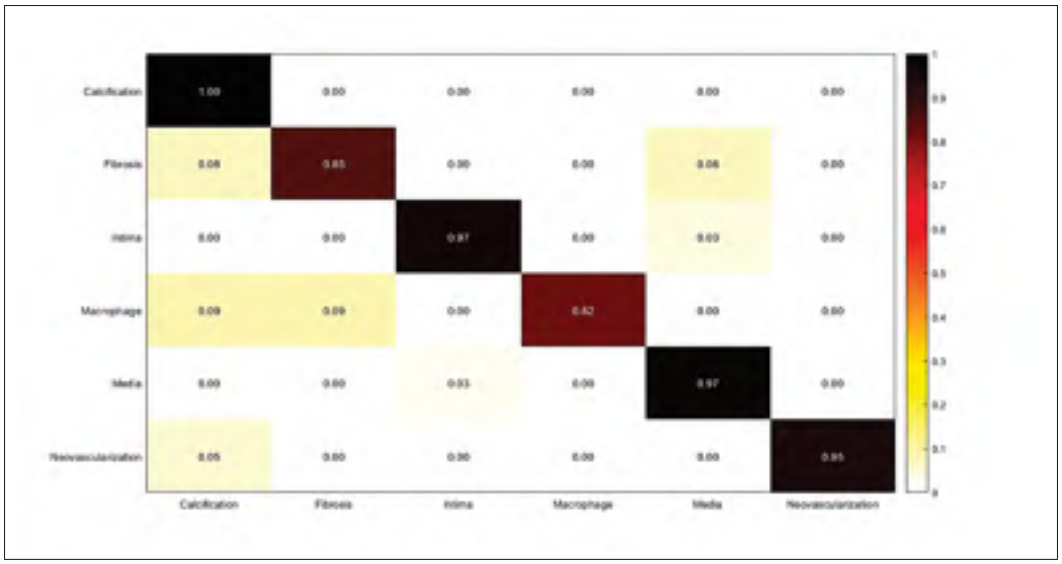


Figure 5.10 Confusion matrix of intracoronary tissue classification: Random Forest is trained using the deep features extracted from Vgg-19

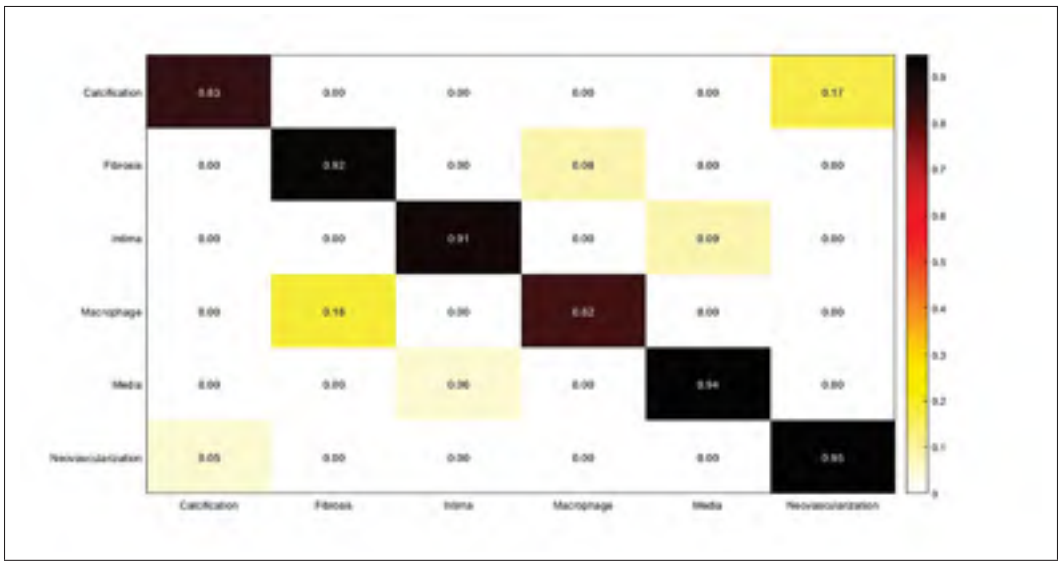


Figure 5.11 Confusion matrix of intracoronary tissue classification: Random Forest is trained using the deep features extracted from Inception-v3

5.3.3 Majority voting

In this experiment, weighted majority voting is applied on Random Forest classification results using each set of features extracted from the three mentioned networks. The results are

Table 5.6 Measured sensitivity, specificity, and accuracy of tissue classification using RF. Features are extracted from Inception-v3

Tissue	Accuracy	Sensitivity	Specificity
Calcification	0.90	0.83	0.97
Fibrosis	0.95	0.92	0.99
Intima	0.95	0.91	0.98
Macrophage	0.90	0.82	0.99
Media	0.96	0.94	0.98
Neovascularization	0.96	0.95	0.97

illustrated in Figure 5.12, and Table 5.7. The results show a good improvement of accuracy, sensitivity, and specificity, which are calculated for the final classification using majority voting.

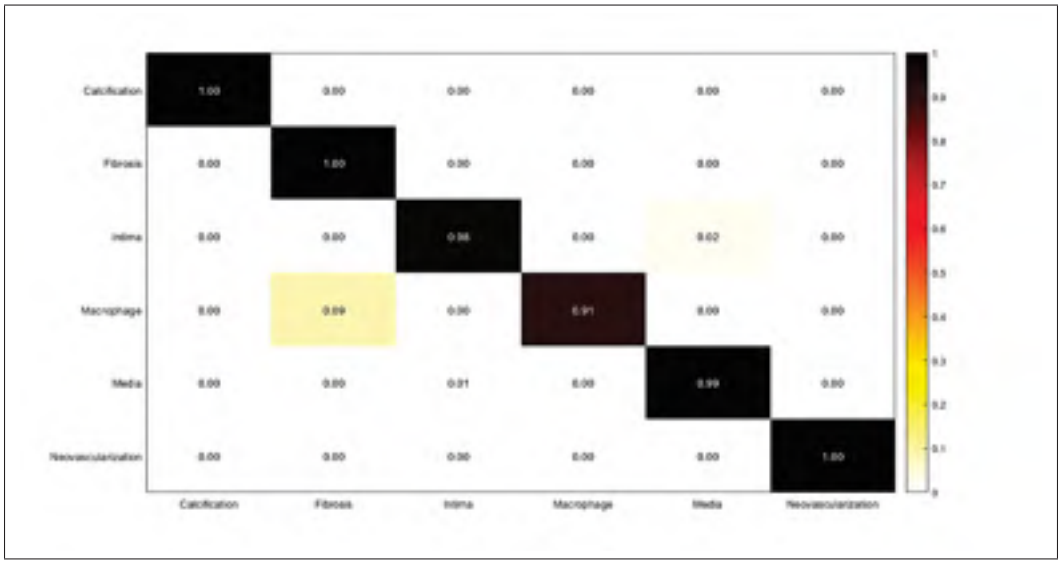


Figure 5.12 Confusion matrix of intracoronary tissue classification using majority voting approach.

5.3.4 RF classification using deep feature fusion

In this experiment, we combined deep features extracted from AlexNet, and Vgg-19 to train Random Forest. The results are shown in Figure 5.13 and Table 5.8. The results of the last two

Table 5.7 Measured sensitivity, specificity, and accuracy of tissue classification using majority voting approach.

Tissue	Accuracy	Sensitivity	Specificity
Calcification	1.00	1.00	1.00
Fibrosis	1.00	1.00	1.00
Intima	0.99	0.98	1.00
Macrophage	0.95	0.91	1.00
Media	0.99	0.99	1.00
Neovascularization	1.00	1.00	1.00

experiments show that majority voting approach performs better than Random Forest classification result while it is trained on the combination of features.

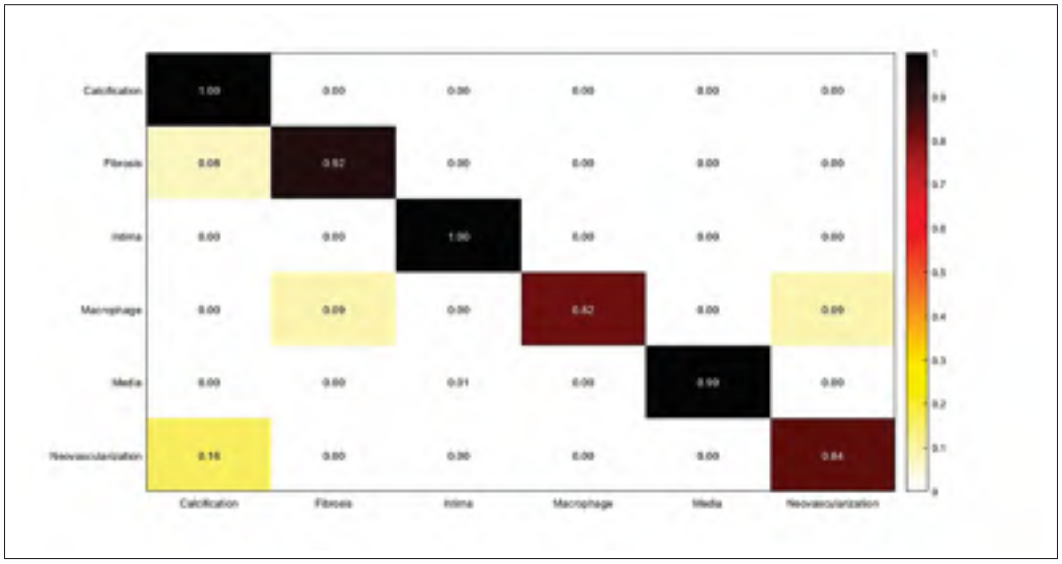


Figure 5.13 Confusion matrix of intracoronary tissue classification using RF: Combination of features extracted from pre-trained AlexNet, and Vgg-19 are used to train Random Forest.

To choose the optimal tissue characterization model considering all the experiments, and to compare the results of the experiments against each other, the mean $\pm$ standard deviation of the values of accuracy, sensitivity, and specificity obtained for all tissues performing each experiment are calculated and the results are shown in Figures 5.14-5.16 and Table 5.9. Although the combination of features can improve the classification results compared against using each

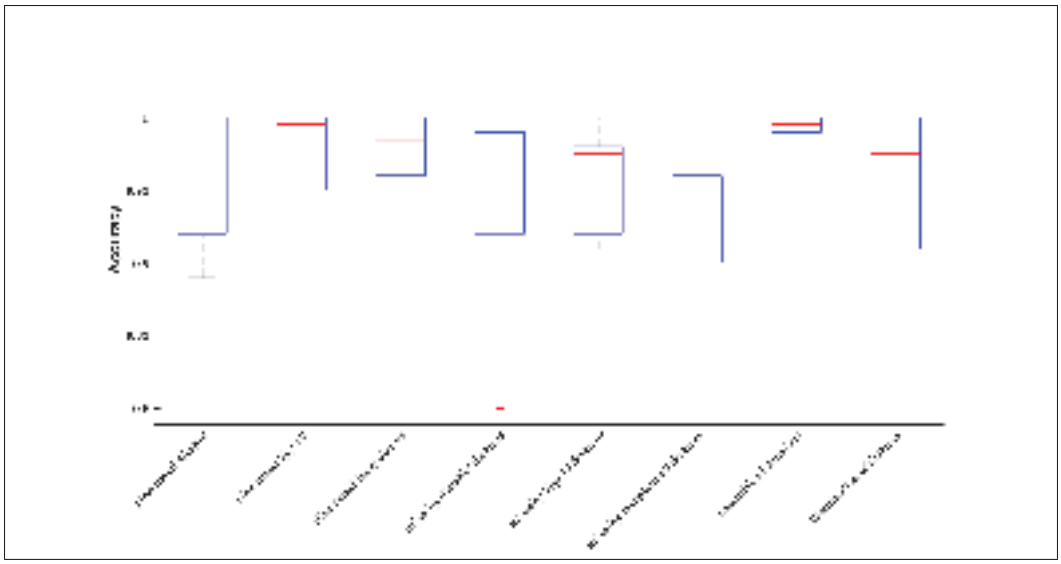


Figure 5.14 Accuracy is reported as the mean ± standard deviation of the measured accuracies for all tissues in each experiment.

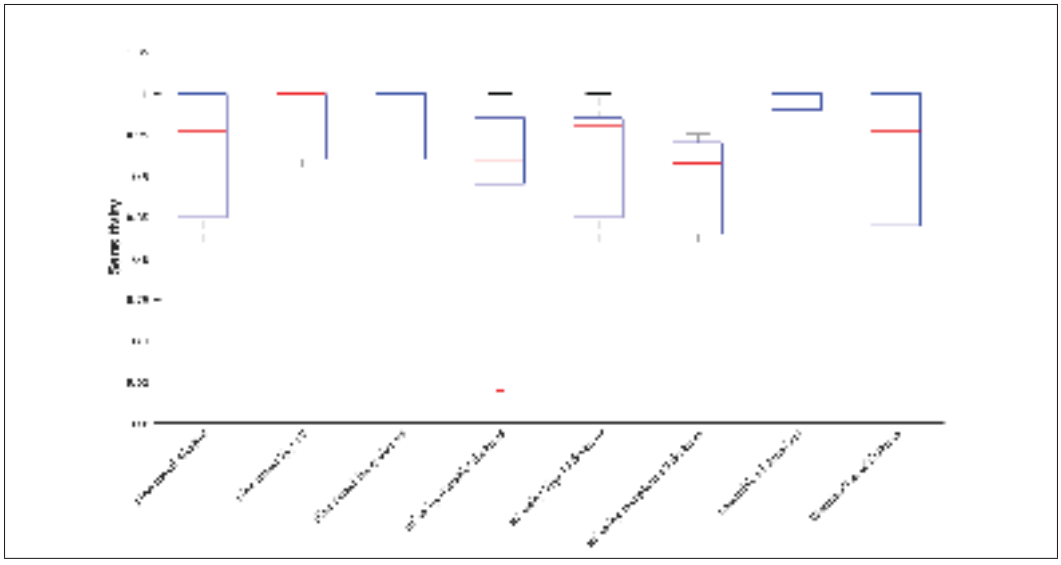


Figure 5.15 Sensitivity is reported as the mean ± standard deviation of the measured sensitivities for all tissues in each experiment.

network separately as feature extractor, the results of majority voting approach are considerably higher than the classification results using the combination of features (Figures 5.14-5.16).

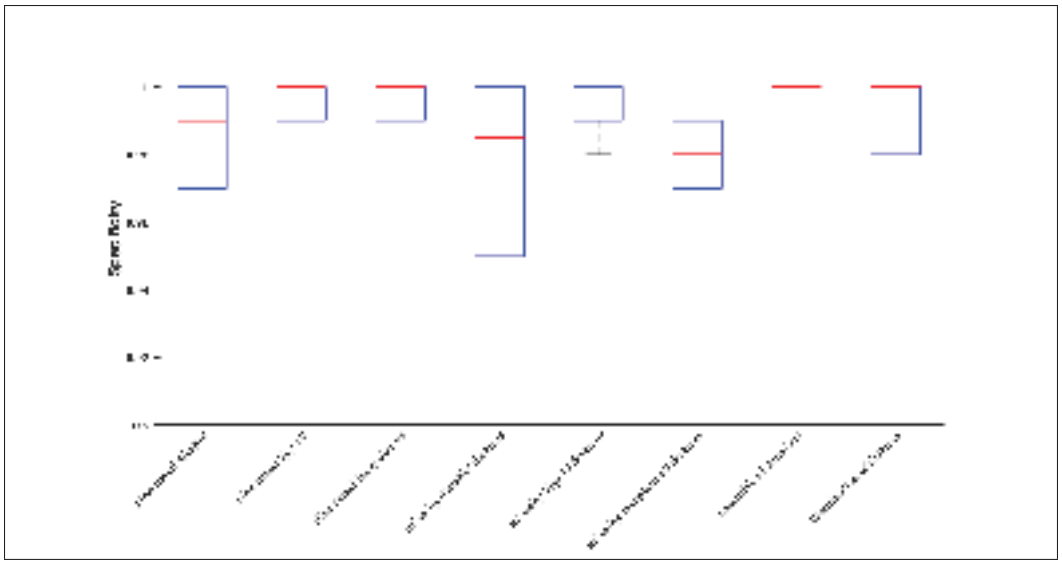


Figure 5.16 Specificity is reported as the mean ± standard deviation of the measured specificities for all tissues in each experiment.

Table 5.8 Measured sensitivity, specificity, and accuracy of tissue classification: Combination of features extracted from pre-trained AlexNet, and Vgg-19 are used to train Random Forest.

Tissue	Accuracy	Sensitivity	Specificity
Calcification	1.00	1.00	1.00
Fibrosis	0.96	0.92	1.00
Intima	1.00	1.00	1.00
Macrophage	0.90	0.82	0.98
Media	0.99	0.99	1.00
Neovascularization	0.91	0.84	0.98

In this study, the performance of pre-trained networks is discussed. Three different state-of-the-art networks (AlexNet, Vgg-19, and Inception-v3) are used in four different experiments. The experiments started with fine-tuning the networks and using them for tissue classification of six different tissue labels (calcification, fibrosis, neovascularization, macrophage, intima, and media). We started with fine-tuning the networks, which is the most common way of applying pre-trained networks for various applications in the field of medical image analysis. Each experiment is designed based on the limitations of the previous experiment to achieve the main goal of this study, which defined as designing an accurate intracoronary tissue classification

Table 5.9 Accuracy, sensitivity, and specificity obtained from each experiment. The accuracy, sensitivity, and specificity are reported as the mean ± standard deviation of the values of accuracy, sensitivity, and specificity obtained for all tissues performing each experiment.

Tissue	**Accuracy**	**Sensitivity**	**Specificity**
Fine-tuned AlexNet	0.96±0.04	0.92±0.08	0.99±0.01
Fine-tuned Vgg-19	0.98±0.02	0.97±0.03	1.00±0.00
Fine-tuned Inception-v3	0.98±0.02	0.96±0.04	1.00±0.00
RF(AlexNet features)	0.93±0.07	0.89±0.13	0.98±0.02
RF(VGG-19 features)	0.96±0.04	0.92±0.07	0.99±0.01
RF(Inception-v3 features)	0.94±0.03	0.90±0.06	0.98±0.01
Majority voting RF	**0.99±0.01**	**0.98±0.02**	**1.00±0.00**
RF(combination of features)	0.94±0.06	0.90±0.10	0.99±0.01

model using deep feature learning in an efficient procedure. The second experiment is performed to avoid convergence issues in fine-tuning the networks, overfitting by deep fine-tuning the networks, and training time. Deep features are very strong to describe arterial tissues and Random Forest works efficiently on large datasets with a very low risk of overfitting. Also, the training process is considerably fast using Random Forest. But, when pre-trained networks are used as feature generators without fine-tuning, the classification results show lower accuracy, sensitivity, and specificity compared against using fine-tuned networks as classifiers. Majority voting on classification results of Random Forest classifiers can considerably improve the results of the second experiment without adding a huge computational burden. The accuracy, sensitivity, and specificity obtained from the third experiment (majority voting from Random Forest classification) can compete against the classification performance of the fine-tuned networks.

By evaluating the results of all the experiments, it is more efficient if we use pre-trained networks as feature extractors and train Random Forest for each set of generated features to perform the classification. Then, majority voting method provides the final tissue classification result. Figure 5.17 shows classification results for each coronary artery tissue.

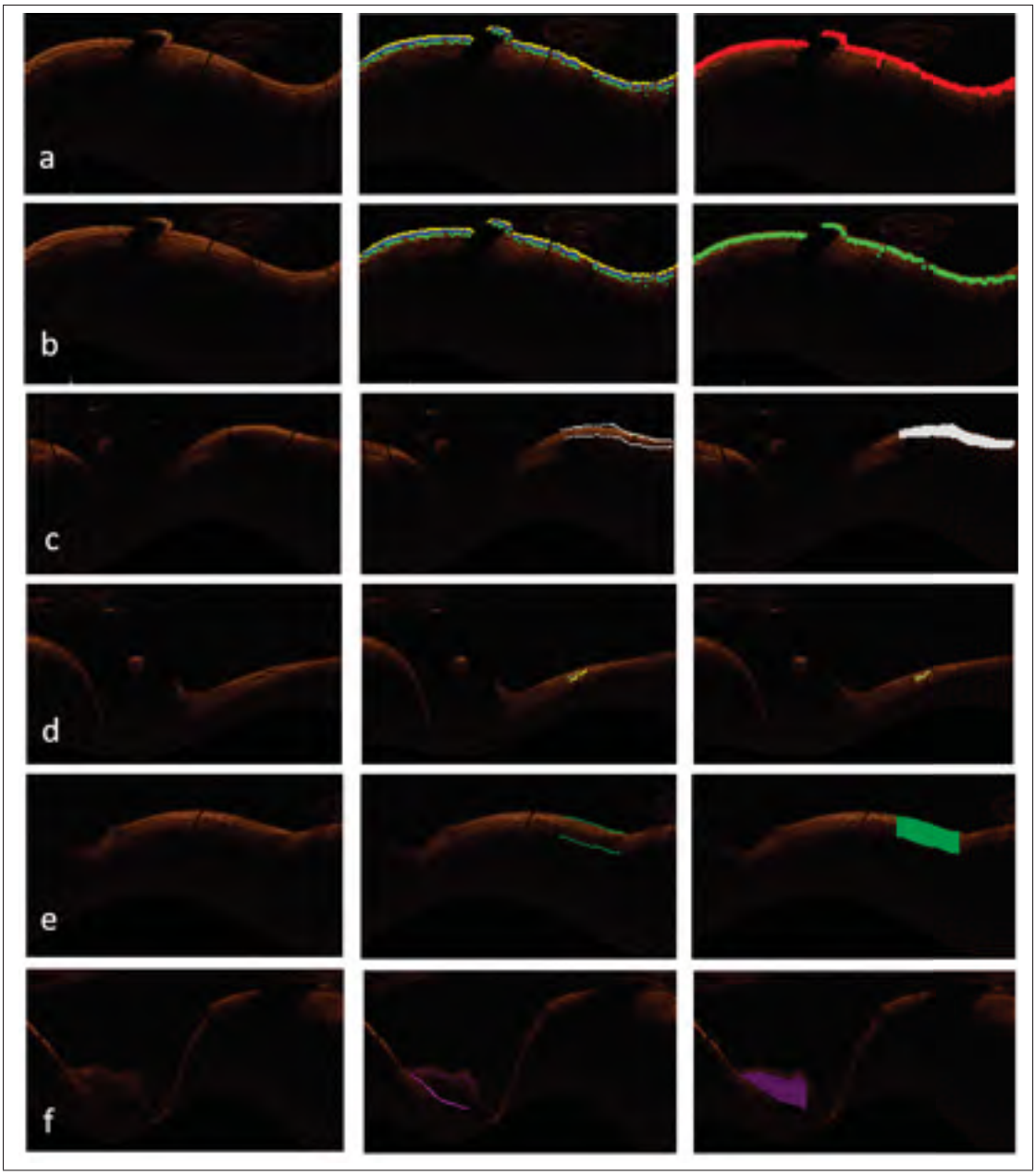


Figure 5.17 From left to right: The first image shows the original OCT image in planar representation, manual segmentation for each tissue is illustrated in the second image, and the third image is the classification result, which is shown for intima in (a), media in (b), fibrosis in (c), neovascularization in (d), macrophage in (e), and calcification in (f).

5.4 Conclusion

The goal of this study was to propose a new approach for OCT imaging using deep feature learning from different CNN models and to evaluate their performance on a complex multi-class classification problem such as pathological formations in coronary artery tissues. The most significant outcome is to be able to automatically differentiate between intracoronary pathological formations observed from OCT imaging. This might be highly relevant for the automatic assessment of coronary artery disease in KD. With the proper dataset and manual annotation, this might be adapted for adult coronary artery diseases to fully assess the structural information of the coronary artery. Majority voting from Random Forest classification using deep features have been successful in classifying coronary artery tissues. The final tissue labels were obtained with high accuracy, sensitivity, and specificity, which confirm the robustness of our proposed technique considering the high variability of pathological formations, OCT artifacts, and the small size of the arteries in pediatric patients, which is followed by very thin layers in coronary artery structure. In this work, we have outlined the relevance of deep features obtained using transfer learning for OCT imaging and the practical aspect of using RF classification to obtain the final decision in a clinically acceptable computational time. For future works, we will focus on detecting intimal hyperplasia by measuring the thickness of intima, and severity of pathological formations by evaluating distensibility variations as a result of calcification, and fibrous scarring.

CONCLUSION AND RECOMMENDATIONS

Progression of pathological formations caused by coronary artery disease can be followed by the acute coronary syndrome. Therefore, it is significant to develop robust coronary artery tissue characterization techniques to evaluate plaque formation. Interpretation of OCT images is challenging and requires a strong expertise from the cardiologist. Moreover, manual segmentation of coronary artery tissues is tedious, time-consuming, and particularly error-prone from one observer to another. The artifacts of the imaging system, the small size of the arteries in infants and children, the small available population with coronary artery disease (CAD) in infants and children, and similar structure of pathological formations are the challenges that result in imprecise manual segmentation of tissues. Since detailed tissue information is required to accurately discriminate among pathological formations, the main goal of this research was to design an automatic intracoronary tissue characterization model to evaluate the functionality of coronary artery tissues by detecting coronary artery layers, and pathological formations caused by coronary artery disease.

The first contribution of this thesis illustrates the robustness of CNN features to describe tissue characteristics, and Random Forest (RF) to classify between intima and media layers. Random Forest (RF) works efficiently on large datasets, carries a very low risk of overfitting, and training process is considerably fast compared against retraining a fine-tuned network for the classification task.

The accurate assessment of pathological formations may obtain by considering the adjacent frames and analyzing the volume variation of each tissue. For each OCT pullback, intracoronary cross-sectional images might be misaligned, since the probe moves freely along the catheter during the imaging process. This is problematic for volumetric measurements of different tissues, which are significant for studying the progression and regression of various pathological formations particularly to evaluate the aneurysmal regions and stenotic segments

of the arterial wall. Since, the optimal features, which can properly describe the arterial wall layers are determined in the first step of the work, they might be applied for motion correction of intracoronary OCT images. Considering the physics of the imaging system, possible artifacts of the OCT imaging, and limitations of the proposed motion correction methods in the literature, we designed a motion correction model by finding the maximum similarity between deep features extracted from OCT pullback frames to perform registration. This will contribute to evaluate the functionality of coronary artery tissues, and distensibility variations of the arterial wall as well as the assessment of pathological formations by finding the correlation between tissues of adjacent frames.

For the final step of this study, we proposed a new approach for OCT imaging using deep feature learning from various CNN models. Majority voting from Random Forest classification using deep features provided the final classification result to detect intima, media, fibrosis, calcification, macrophage, and neovascularization. This work contributes to the automatic characterization of intracoronary pathological formations observed form OCT imaging in KD patients that can be adapted to assess the structural information of coronary artery in adult patients.

OCT imaging is a new modality for pediatric cardiology, especially for KD patients. With proper registration, OCT may complement coronary assessment by adding tissue layers not visible on X-ray angiography and CT acquisitions. Since interpretation of OCT images is highly challenging, even for a trained expert, an automatic tissue characterization model not only have a significant impact on efficient clinical diagnosis in KD patients but also in adult cardiology, where the coronary arteries are larger and more cases with the coronary artery disease are available in this population. On the other hand, deep learning using CNN gained a huge popularity in image recognition tasks but requires a large amount of annotated images and they require a lot of computing resources (GPU). Therefore, transfer learning and fine-tuning of

an existing network were found to be successful for OCT imaging. Specifically, if the networks are used as feature generators to train a classifier such as Random Forest in our application. The designed models are robust indicators of progression of pathological formations, which can result in an efficient clinical diagnosis with a higher level of certainty than current methods and early detection of intracoronary complications to enhance patient outcomes. Also, clinical indices from OCT can now be computed in 3D (area, volume) to fully characterize coronary vessel, tissue-by-tissue.

For future works, we will focus on detecting the intimal hyperplasia and media disappearance, which are the most frequent finding in pediatric patients with KD. Since we designed the model to detect coronary artery layers, intimal hyperplasia can be recognized by measuring the thickness of coronary artery layers. The absence of media may be reported as media destruction as a result of thickening the first coronary artery layer (media).

Volumetric measurements of different tissues are significant for studying the progression and regression of various pathological formations particularly to evaluate the aneurysmal regions and stenotic segments of the arterial wall. 3D assessment of the morphological tissues is problematic because they are highly affected by motion artifacts resulted in the rotation and translation of the imaging probe along the artery during image acquisition. Generating a 3D-reconstruction from 2D cross-sectional OCT pullbacks including motion compensation may be an incremental step.

Pathological analysis of coronary arteries in KD demonstrates that aneurysms can be filled in with the intimal fibrosis, which results in the development of arterial wall stiffness. In some cases, although the aneurysms are treated and the lumen diameter is normal, but the coronary artery may represent less distensibility, that can affect myocardial microcirculation. To evaluate the functionality of arterial wall, decreased arterial distensibility should be estimated by considering the characteristics of arterial wall elasticity.

Moreover, this work may introduce a safe navigation system for coronary arteries. The first step in any navigation system is to locate the target and to provide accurate information about its surroundings. Stereoradiographic 3D-reconstruction from biplane angiography can provide a global representation of vascular structures and more importantly, vessel pulsatility in real-time. Such information is important to estimate the stiffness of the artery. However, the 3D-reconstruction is limited and does not provide any insight about the inner tissue layers. Consequently, intravascular imaging from coronary OCT images is appealing for characterization of the biophysical properties of the inner tissue layers. Stationary-OCT is used to outline the elasticity properties and estimate the stiffness of coronary artery tissues. In stationary-OCT, using the same configuration of the OCT system, the probe is stabilized and a sequence of time-varying cross-sectional images are recorded from the arterial wall motion during each cardiac cycle. In stationary OCT imaging, the pullback path reveals both the axial change of the arterial structure and the local variations of the vessel morphology as a result of cardiac motion. Measuring the strain in response to the changes of the luminal pressure and estimating the local differential motion of the vessel wall tissues can be helpful to evaluate the degree of stiffness of the arterial wall and severity of the disease. Finally, a proper characterization of the tissues in 3D may lead to a fully personalized model for the evaluation of coronary artery diseases. This personalized model can be inputted to a biomechanical simulation of the arterial wall geometry for a better understanding of pathologic formations in coronary arteries.

BIBLIOGRAPHY

Abdolmanafi, A., Prasad, A. S., Duong, L. & Dahdah, N. (2016). Classification of coronary artery tissues using optical coherence tomography imaging in Kawasaki disease. *SPIE Medical Imaging*, pp. 97862U–97862U.

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.

Antony, J., McGuinness, K., O'Connor, N. E. & Moran, K. (2016). Quantifying radiographic knee osteoarthritis severity using deep convolutional neural networks. *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pp. 1195–1200.

Arevalo, J., González, F. A., Ramos-Pollán, R., Oliveira, J. L. & Lopez, M. A. G. (2015). Convolutional neural networks for mammography mass lesion classification. *2015 37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 797–800.

Athanasiou, L., Rigas, G., Sakellarios, A. I., Exarchos, T. P., Siogkas, P. K., Bourantas, C. V., Garcia-Garcia, H. M., Lemos, P. A., Falcao, B. A., Michalis, L. K. et al. (2016). Three-dimensional reconstruction of coronary arteries and plaque morphology using CT angiography–comparison and registration with IVUS. *BMC medical imaging*, 16(1), 9.

Athanasiou, L., Bourantas, C., Siogkas, P., Sakellarios, A., Exarchos, T., Naka, K., Papafaklis, M., Michalis, L., Prati, F. & Fotiadis, D. (2012). 3D reconstruction of coronary arteries using frequency domain optical coherence tomography images and biplane angiography. *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, pp. 2647–2650.

Azarnoush, H., Vergnole, S., Pazos, V., Bisaillon, C.-É., Boulet, B. & Lamouche, G. (2012). Intravascular optical coherence tomography to characterize tissue deformation during angioplasty: preliminary experiments with artery phantoms. *Journal of biomedical optics*, 17(9), 096015–096015.

Azizpour, H., Sharif Razavian, A., Sullivan, J., Maki, A. & Carlsson, S. (2015). From generic to specific deep representations for visual recognition. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 36–45.

Baim, D. S. & Grossman, W. (1996). *Cardiac catheterization, angiography, and intervention*. Lippincott Williams & Wilkins.

Balocco, S., Zuluaga, M., Zahnd, G., Demirci, S. & Lee, S. (2016). *Computing and Visualization for Intravascular Imaging and Computer-Assisted Stenting*. Elsevier Science & Technology. Consulted at https://books.google.ca/books?id=AAE0vgAACAAJ.

Bar, Y., Diamant, I., Wolf, L. & Greenspan, H. (2015). Deep learning with non-medical training used for chest pathology identification. *SPIE Medical Imaging*, pp. 94140V–94140V.

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

Boudoux, C. (2016). Fundamentals of Biomedical Optics. Pollux.

Bourantas, C. V., Kourtis, I. C., Plissiti, M. E., Fotiadis, D. I., Katsouras, C. S., Papafaklis, M. I. & Michalis, L. K. (2005). A method for 3D reconstruction of coronary arteries using biplane angiography and intravascular ultrasound images. *Computerized Medical Imaging and Graphics*, 29(8), 597–606.

Braaf, B., Vienola, K. V., Sheehy, C. K., Yang, Q., Vermeer, K. A., Tiruveedhula, P., Arathorn, D. W., Roorda, A. & de Boer, J. F. (2013). Real-time eye motion correction in phase-resolved OCT angiography with tracking SLO. *Biomedical optics express*, 4(1), 51–65.

Carneiro, G., Nascimento, J. & Bradley, A. P. (2015). Unregistered multiview mammogram analysis with pre-trained deep learning models. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 652–660.

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

Chan, H.-P., Lo, S.-C. B., Sahiner, B., Lam, K. L. & Helvie, M. A. (1995). Computer-aided detection of mammographic microcalcifications: Pattern recognition with an artificial neural network. *Medical Physics*, 22(10), 1555–1567.

Chang, C.-C. & Lin, C.-J. (2011). LIBSVM: a library for support vector machines. *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2(3), 27.

Chen, H., Ni, D., Qin, J., Li, S., Yang, X., Wang, T. & Heng, P. A. (2015). Standard plane localization in fetal ultrasound via domain transferred deep neural networks. *IEEE journal of biomedical and health informatics*, 19(5), 1627–1636.

Ciompi, F., de Hoop, B., van Riel, S. J., Chung, K., Scholten, E. T., Oudkerk, M., de Jong, P. A., Prokop, M. & van Ginneken, B. (2015). Automatic classification of pulmonary peri-fissural nodules in computed tomography using an ensemble of 2D views and a convolutional neural network out-of-the-box. *Medical image analysis*, 26(1), 195–202.

Costa, R. A., Skaf, M., Melo, L. A., Calucci, D., Cardillo, J. A., Castro, J. C., Huang, D. & Wojtkowski, M. (2006). Retinal assessment using optical coherence tomography. *Progress in retinal and eye research*, 25(3), 325–353.

Cothren, R. M., Shekhar, R., Tuzcu, E. M., Nissen, S. E., Cornhill, J. F. & Vince, D. G. (2000). Three-dimensional reconstruction of the coronary artery wall by image fusion of intravascular ultrasound and bi-plane angiography. *The International Journal of Cardiac Imaging*, 16(2), 69–85.

Criminisi, A. & Shotton, J. (2013). *Decision forests for computer vision and medical image analysis*. Springer Science & Business Media.

Destrempes, F., Cardinal, M.-H. R., Allard, L., Tardif, J.-C. & Cloutier, G. (2014). Segmentation method of intravascular ultrasound images of human coronary arteries. *Computerized Medical Imaging and Graphics*, 38(2), 91–103.

Dionne, A., Ibrahim, R., Gebhard, C., Bakloul, M., Selly, J.-B., Leye, M., Déry, J., Lapierre, C., Girard, P., Fournier, A. et al. (2015). Coronary wall structural changes in patients with Kawasaki disease: new insights from optical coherence tomography (OCT). *Journal of the American Heart Association*, 4(5), e001939.

Drexler, W. & Fujimoto, J. G. (2008). *Optical coherence tomography: technology and applications*. Springer Science & Business Media.

Drexler, W. & Fujimoto, J. G. (2015). *Optical coherence tomography: technology and applications*. Springer.

Eigen, D., Rolfe, J., Fergus, R. & LeCun, Y. (2013). Understanding deep architectures using a recursive convolutional network. *arXiv preprint arXiv:1312.1847*.

Ellwein, L. M., Otake, H., Gundert, T. J., Koo, B.-K., Shinke, T., Honda, Y., Shite, J. & LaDisa Jr, J. F. (2011). Optical coherence tomography for patient-specific 3D artery reconstruction and evaluation of wall shear stress in a left circumflex coronary artery. *Cardiovascular Engineering and Technology*, 2(3), 212–227.

Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M. & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639), 115.

Fercher, A. F., Drexler, W., Hitzenberger, C. K. & Lasser, T. (2003). Optical coherence tomography-principles and applications. *Reports on progress in physics*, 66(2), 239.

Ferrante, G., Presbitero, P., Whitbourn, R. & Barlis, P. (2013). Current applications of optical coherence tomography for coronary intervention. *International journal of cardiology*, 165(1), 7–16.

Foin, N., Mari, J. M., Nijjer, S., Sen, S., Petraco, R., Ghione, M., Di Mario, C., Davies, J. E. & Girard, M. J. (2013). Intracoronary imaging using attenuation-compensated optical coherence tomography allows better visualisation of coronary artery diseases. *Cardiovascular Revascularization Medicine*, 14(3), 139–143.

Fortier, A., Gullapalli, V. & Mirshams, R. A. (2014). Review of biomechanical studies of arteries and their effect on stent performance. *IJC Heart & Vessels*, 4, 12–18.

Fujii, K., Kawasaki, D., Masutani, M., Okumura, T., Akagami, T., Sakoda, T., Tsujino, T., Ohyanagi, M. & Masuyama, T. (2010). OCT assessment of thin-cap fibroatheroma distribution in native coronary arteries. *JACC: Cardiovascular Imaging*, 3(2), 168–175.

Fujino, Y., Attizzani, G. F., Tahara, S., Takagi, K., Bezerra, H. G., Nakamura, S. & Costa, M. A. (2014). Frequency-domain optical coherence tomography evaluation of a patient with Kawasaki disease and severely calcified plaque. *International journal of cardiology*, 171(2), 281–283.

Gan, Y., Tsay, D., Amir, S. B., Marboe, C. C. & Hendon, C. P. (2016). Automated classification of optical coherence tomography images of human atrial tissue. *Journal of biomedical optics*, 21(10), 101407.

Gao, M., Bagci, U., Lu, L., Wu, A., Buty, M., Shin, H.-C., Roth, H., Papadakis, G. Z., Depeursinge, A., Summers, R. M. et al. (2016). Holistic classification of CT attenuation patterns for interstitial lung diseases via deep convolutional neural networks. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1–6.

Gargesha, M., Shalev, R., Prabhu, D., Tanaka, K., Rollins, A. M., Costa, M., Bezerra, H. G. & Wilson, D. L. (2015). Parameter estimation of atherosclerotic tissue optical properties from three-dimensional intravascular optical coherence tomography. *Journal of medical imaging*, 2(1), 016001.

Group, J. C. S. J. R. (2005). Guidelines for diagnosis and management of cardiovascular sequelae in Kawasaki disease. *Pediatrics International*, 47(6), 711–732.

Group, J. J. W. et al. (2010). Guidelines for diagnosis and management of cardiovascular sequelae in Kawasaki disease (JCS 2008). *Circulation Journal*, 74(9), 1989–2020.

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J. et al. (2016). Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama*, 316(22), 2402–2410.

Hamdan, R., Gonzalez, R. G., Ghostine, S. & Caussin, C. (2012). Optical coherence tomography: from physical principles to clinical applications. *Archives of cardiovascular diseases*, 105(10), 529–534.

Harris, K. C., Manouzi, A., Fung, A. Y., De Souza, A., Bezerra, H. G., Potts, J. E. & Hosking, M. C. (2014). Feasibility of optical coherence tomography in children with Kawasaki disease and pediatric heart transplant recipients. *Circulation: Cardiovascular Imaging*, 7(4), 671–678.

Hauser, M., Bengel, F., Kuehn, A., Nekolla, S., Kaemmerer, H., Schwaiger, M. & Hess, J. (2004). Myocardial blood flow and coronary flow reserve in children with "normal" epicardial coronary arteries after the onset of Kawasaki disease assessed by positron emission tomography. *Pediatric cardiology*, 25(2), 108–112.

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M. & Larochelle, H. (2016). Brain tumor segmentation with deep neural networks. *Medical Image Analysis*.

Havaei, M., Davy, A., Warde-Farley, D., Biard, A., Courville, A., Bengio, Y., Pal, C., Jodoin, P.-M. & Larochelle, H. (2017). Brain tumor segmentation with deep neural networks. *Medical image analysis*, 35, 18–31.

Herrero-Garibi, J., Cruz-González, I., Parejo-Díaz, P. & Jang, I.-K. (2010). Optical coherence tomography: its value in intravascular diagnosis today. *Revista Española de Cardiología (English Edition)*, 63(8), 951–962.

Hochreiter, S. & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, 9(8), 1735–1780.

Hubel, D. H. & Wiesel, T. N. (1959). Receptive fields of single neurones in the cat's striate cortex. *The Journal of physiology*, 148(3), 574–591.

Jang, I.-K., Bouma, B. E., Kang, D.-H., Park, S.-J., Park, S.-W., Seung, K.-B., Choi, K.-B., Shishkov, M., Schlendorf, K., Pomerantsev, E. et al. (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.

Kakimoto, N., Suzuki, H., Kubo, T., Suenaga, T., Takeuchi, T., Shibuta, S., Ino, Y., Akasaka, T. & Yoshikawa, N. (2014). Evaluation of coronary arterial lesions due to Kawasaki disease using optical coherence tomography. *Canadian Journal of Cardiology*, 30(8), 956–e7.

Karlas, A. & Lee, S.-L. (2015). Towards an IVUS-driven system for endovascular navigation. *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, pp. 1324–1327.

Kawasaki, M., Takatsu, H., Noda, T., Sano, K., Ito, Y., Hayakawa, K., Tsuchiya, K., Arai, M., Nishigaki, K., Takemura, G. et al. (2002). In vivo quantitative tissue characterization of human coronary arterial plaques by use of integrated backscatter intravascular ultrasound and comparison with angioscopic findings. *Circulation*, 105(21), 2487–2492.

Kim, E., Corte-Real, M. & Baloch, Z. (2016). A deep semantic mobile application for thyroid cytopathology. *Medical Imaging 2016: PACS and Imaging Informatics: Next Generation and Innovations*, 9789, 97890A.

Kitabata, H. & Akasaka, T. (2013). Visualization of Plaque Neovascularization by OCT. In *Optical Coherence Tomography*. InTech.

Klingensmith, J. D., Shekhar, R. & Vince, D. G. (2000). Evaluation of three-dimensional segmentation algorithms for the identification of luminal and medial-adventitial borders in intravascular ultrasound images. *IEEE Transactions on Medical Imaging*, 19(10), 996–1011.

Kraus, M. F., Potsaid, B., Mayer, M. A., Bock, R., Baumann, B., Liu, J. J., Hornegger, J. & Fujimoto, J. G. (2012). Motion correction in optical coherence tomography volumes on a per A-scan basis using orthogonal scan patterns. *Biomedical optics express*, 3(6), 1182–1199.

Kraus, M. F., Liu, J. J., Schottenhamml, J., Chen, C.-L., Budai, A., Branchini, L., Ko, T., Ishikawa, H., Wollstein, G., Schuman, J. et al. (2014). Quantitative 3D-OCT motion correction with tilt and illumination correction, robust similarity measure and regularization. *Biomedical optics express*, 5(8), 2591–2613.

Krizhevsky, A., Sutskever, I. & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, pp. 1097–1105.

Kuhn, M. & Johnson, K. (2013). *Applied predictive modeling*. Springer.

Lang, A., Carass, A., Hauser, M., Sotirchos, E. S., Calabresi, P. A., Ying, H. S. & Prince, J. L. (2013). Retinal layer segmentation of macular OCT images using boundary classification. *Biomedical optics express*, 4(7), 1133–1152.

Lee, J., Srinivasan, V., Radhakrishnan, H. & Boas, D. A. (2011). Motion correction for phase-resolved dynamic optical coherence tomography imaging of rodent cerebral cortex. *Optics express*, 19(22), 21258–21270.

Levitz, D., Thrane, L., Frosz, M., Andersen, P., Andersen, C., Andersson-Engels, S., Valanciunaite, J., Swartling, J. & Hansen, P. (2004). Determination of optical scattering properties of highly-scattering media in optical coherence tomography images. *Optics express*, 12(2), 249–259.

Ligthart, J., Bruining, N., van Soest, G. et al. (2011). The diagnostic value of intracoronary optical coherence tomography. *Herz*, 36(5), 417–429.

Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A., van Ginneken, B. & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical image analysis*, 42, 60–88.

Liu, W., Zhang, Y., Yu, C.-M., Ji, Q.-W., Cai, M., Zhao, Y.-X. & Zhou, Y.-J. (2015). Current understanding of coronary artery calcification. *Journal of geriatric cardiology: JGC*, 12(6), 668.

Lo, S.-C., Lou, S.-L., Lin, J.-S., Freedman, M. T., Chien, M. V. & Mun, S. K. (1995). Artificial convolution neural network techniques and applications for lung nodule detection. *IEEE Transactions on Medical Imaging*, 14(4), 711–718.

Lo, S.-C. B., Lin, J.-S., Freedman, M. T. & Mun, S. K. (1993). Computer-assisted diagnosis of lung nodule detection using artificial convoultion neural network. *Medical Imaging 1993*, pp. 859–869.

Ma, H. T., Wang, H., Wang, C. & Hau, W. K. (2013). 3D reconstruction of coronary arteries using intravascular ultrasound (IVUS) and angiography. *TENCON 2013-2013 IEEE Region 10 Conference (31194)*, pp. 1–4.

Macedo, M. M., Nicz, P. F., Campos, C. M., Lemos, P. A. & Gutierrez, M. A. (2016). Spatial-frequency approach to fibrous tissue classification in intracoronary optical images. *Computing in Cardiology Conference (CinC), 2016*, pp. 477–480.

Madhavan, M. V., Tarigopula, M., Mintz, G. S., Maehara, A., Stone, G. W. & Généreux, P. (2014). Coronary artery calcification: pathogenesis and prognostic implications. *Journal of the American College of Cardiology*, 63(17), 1703–1714.

Margeta, J., Criminisi, A., Cabrera Lozoya, R., Lee, D. C. & Ayache, N. (2015). Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition. *Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization*, 1–11.

Mattace-Raso, F. U., van der Cammen, T. J., Hofman, A., van Popele, N. M., Bos, M. L., Schalekamp, M. A., Asmar, R., Reneman, R. S., Hoeks, A. P., Breteler, M. M. et al. (2006). Arterial stiffness and risk of coronary heart disease and stroke. *Circulation*, 113(5), 657–663.

Mendizabal-Ruiz, E. G., Rivera, M. & Kakadiaris, I. A. (2013). Segmentation of the luminal border in intravascular ultrasound B-mode images using a probabilistic approach. *Medical image analysis*, 17(6), 649–670.

Mitani, Y., Ohashi, H., Sawada, H., Ikeyama, Y., Hayakawa, H., Takabayashi, S., Maruyama, K., Shimpo, H. & Komada, Y. (2009). In vivo plaque composition and morphology in coronary artery lesions in adolescents and young adults long after Kawasaki disease. *Circulation*, 119(21), 2829–2836.

Newburger, J. W., Takahashi, M., Gerber, M. A., Gewitz, M. H., Tani, L. Y., Burns, J. C., Shulman, S. T., Bolger, A. F., Ferrieri, P., Baltimore, R. S. et al. (2004). Diagnosis, treatment, and long-term management of Kawasaki disease. *Circulation*, 110(17), 2747–2771.

Orenstein, J. M., Shulman, S. T., Fox, L. M., Baker, S. C., Takahashi, M., Bhatti, T. R., Russo, P. A., Mierau, G. W., de Chadarévian, J. P., Perlman, E. J. et al. (2012). Three linked vasculopathic processes characterize Kawasaki disease: a light and transmission electron microscopic study. *PloS one*, 7(6), e38998.

Papadogiorgaki, M., Mezaris, V., Chatzizisis, Y. S., Giannoglou, G. D. & Kompatsiaris, I. (2008). Image analysis techniques for automated IVUS contour detection. *Ultrasound in medicine & biology*, 34(9), 1482–1498.

Preim, B. & Bartz, D. (2007). *Visualization in medicine: theory, algorithms, and applications*. Morgan Kaufmann.

Rathod, K. S., Hamshere, S. M., Jones, D. A. & Mathur, A. (2015). Intravascular ultrasound versus optical coherence tomography for coronary artery imaging–Apples and oranges. *Interventional Cardiology Review*, 10(1), 8–15.

Regar, E., Ligthart, J., Bruining, N. & van Soest, G. (2011). The diagnostic value of intracoronary optical coherence tomography. *Herz*, 36(5), 417.

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

Roth, H. R., Farag, A., Lu, L., Turkbey, E. B. & Summers, R. M. (2015). Deep convolutional networks for pancreas segmentation in CT imaging. *SPIE Medical Imaging*, pp. 94131G–94131G.

Sahiner, B., Chan, H.-P., Petrick, N., Wei, D., Helvie, M. A., Adler, D. D. & Goodsitt, M. M. (1996). Classification of mass and normal breast tissue: a convolution neural network classifier with spatial domain and texture images. *IEEE transactions on Medical Imaging*, 15(5), 598–610.

Shin, H.-C., Orton, M. R., Collins, D. J., Doran, S. J. & Leach, M. O. (2013). Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4D patient data. *IEEE transactions on pattern analysis and machine intelligence*, 35(8), 1930–1943.

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

Swirski, F. K., Robbins, C. S. & Nahrendorf, M. (2016). Development and function of arterial and cardiac macrophages. *Trends in immunology*, 37(1), 32–40.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S. & Anguelov, D. & Rabinovich, A.(2015). Going deeper with convolutions. *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 1–9.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. & Rabinovich, A. (2015). Going deeper with convolutions. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1–9.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818–2826.

Taguchi, Y., Itoh, T., Oda, H., Uchimura, Y., Kaneko, K., Sakamoto, T., Goto, I., Sakuma, M., Ishida, M., Terashita, D. et al. (2017). Coronary risk factors associated with OCT macrophage images and their response after CoCr everolimus-eluting stent implantation in patients with stable coronary artery disease. *Atherosclerosis*, 265, 117–123.

Tajbakhsh, N., Shin, J. Y., Gurudu, S. R., Hurst, R. T., Kendall, C. B., Gotway, M. B. & Liang, J. (2016). Convolutional Neural Networks for Medical Image Analysis: Full Training or Fine Tuning? *IEEE transactions on medical imaging*, 35(5), 1299–1312.

Takarada, S., Imanishi, T., Liu, Y., Ikejima, H., Tsujioka, H., Kuroi, A., Ishibashi, K., Komukai, K., Tanimoto, T., Ino, Y. et al. (2010). Advantage of next-generation frequency-domain optical coherence tomography compared with conventional time-domain system in the assessment of coronary lesion. *Catheterization and Cardiovascular Interventions*, 75(2), 202–206.

Tearney, G. J., Waxman, S., Shishkov, M., Vakoc, B. J., Suter, M. J., Freilich, M. I., Desjardins, A. E., Oh, W.-Y., Bartlett, L. A., Rosenberg, M. et al. (2008). Three-dimensional coronary artery microscopy by intracoronary optical frequency domain imaging. *JACC: Cardiovascular imaging*, 1(6), 752–761.

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

Tu, S., Holm, N. R., Koning, G., Huang, Z. & Reiber, J. H. (2011). Fusion of 3d qca and ivus/oct. *The international journal of cardiovascular imaging*, 27(2), 197–207.

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

Ulrich, S. M., Lehner, A., Birnbaum, J., Heckel, S., Haas, N. A., Hakami, L., Schramm, R., Dalla Pozza, R., Fischer, M. & Kozlik-Feldmann, R. (2017). Safety of optical coherence tomography in pediatric heart transplant patients. *International journal of cardiology*, 228, 205–208.

Unal, G., Bucher, S., Carlier, S., Slabaugh, G., Fang, T. & Tanaka, K. (2008). Shape-driven segmentation of the arterial wall in intravascular ultrasound images. *IEEE Transactions on Information Technology in Biomedicine*, 12(3), 335–347.

van Ginneken, B., Setio, A. A., Jacobs, C. & Ciompi, F. (2015). Off-the-shelf convolutional neural network features for pulmonary nodule detection in computed tomography scans. *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*, pp. 286–289.

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

Vermeer, K., Van der Schoot, J., Lemij, H. & De Boer, J. (2011). Automated segmentation by pixel classification of retinal layers in ophthalmic OCT images. *Biomedical optics express*, 2(6), 1743–1756.

Wahle, A., Prause, G., DeJong, S. & Sonka, M. (1998). Determination of the absolute axial orientation of intracoronary ultrasound images in fusion with biplane angiography. *Computers in Cardiology 1998*, pp. 153–156.

Wang, Y., Wang, Y., Akansu, A., Belfield, K. D., Hubbi, B. & Liu, X. (2015). Robust motion tracking based on adaptive speckle decorrelation analysis of OCT signal. *Biomedical optics express*, 6(11), 4302–4316.

Wang, Z. & Xue, X. (2014). Multi-class support vector machine. In *Support Vector Machines Applications* (pp. 23–48). Springer.

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

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

Yang, C., Bach, R. G., Zheng, J., Naqa, I. E., Woodard, P. K., Teng, Z., Billiar, K. & Tang, D. (2009). In vivo IVUS-based 3-D fluid–structure interaction models with cyclic bending and anisotropic vessel properties for human atherosclerotic coronary plaque mechanical analysis. *IEEE Transactions on biomedical engineering*, 56(10), 2420–2428.

Zeiler, M. D. & Fergus, R. (2014). Visualizing and understanding convolutional networks. *European Conference on Computer Vision*, pp. 818–833.

Zhao, L., Giannarou, S., Lee, S.-L. & Yang, G.-Z. (2016). SCEM+: real-time robust simultaneous catheter and environment modeling for endovascular navigation. *IEEE Robotics and Automation Letters*, 1(2), 961–968.

Zheng, S. (2009). Three-dimensional reconstruction of vessels from intravascular ultrasound sequence and X-ray angiograms. *Biomedical Engineering and Informatics, 2009. BMEI'09. 2nd International Conference on*, pp. 1–5.

Zysk, A. M., Nguyen, F. T., Oldenburg, A. L., Marks, D. L. & Boppart, S. A. (2007). Optical coherence tomography: a review of clinical development from bench to bedside. *Journal of biomedical optics*, 12(5), 051403–051403.