

FULL ARTICLE

An automatic diagnostic system of coronary artery lesions in Kawasaki disease using intravascular optical coherence tomography imaging

Atefeh Abdolmanafi^{1,2*}  | Farida Cheriet^{1,2} | Luc Duong^{2,3} | Ragui Ibrahim⁴ | Nagib Dahdah²

¹Department of Computer Engineering, École Polytechnique de Montréal, Montréal, Canada

²Division of Pediatric Cardiology, Centre Hospitalier Universitaire Sainte-Justine, Montréal, Canada

³Department of Software and IT Engineering, École de technologie supérieure, Montréal, Canada

⁴Division of Cardiology, Hôpital Pierre Boucher, Longueuil, Canada

*Correspondence

Atefeh Abdolmanafi, Department of Computer Engineering, École Polytechnique de Montréal, Montréal, QC H3T 1J4, Canada.

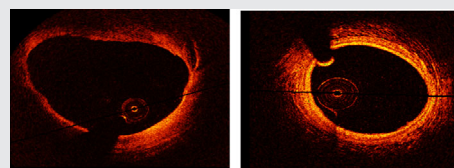
Email: atefeh.abdolmanafi@polymtl.ca

Funding information

NSERC Discovery grant, Grant/Award Number: RGPIN-2018-05283; JP Metal Children's fund, Montreal; and Fonds BoBeau Coeur—Fondation CHU Ste-Justine; Natural Sciences and Engineering Research Council of Canada (NSERC)

Abstract

Intravascular optical coherence tomography (IV-OCT) is a light-based imaging modality with high resolution, which employs near-infrared light to



provide tomographic intracoronary images. Morbidity caused by coronary heart disease is a substantial cause of acute coronary syndrome and sudden cardiac death. The most common intracoronary complications caused by coronary artery disease are intimal hyperplasia, calcification, fibrosis, neovascularization and macrophage accumulation, which require efficient prevention strategies. OCT can provide discriminative information of the intracoronary tissues, which can be used to train a robust fully automatic tissue characterization model based on deep learning. In this study, we aimed to design a diagnostic model of coronary artery lesions. Particularly, we trained a random forest using convolutional neural network features to distinguish between normal and diseased arterial wall structure. Then, based on the arterial wall structure, fully convolutional network is designed to extract the tissue layers in normal cases, and pathological tissues regardless of lesion type in pathological cases. Then, the type of the lesions can be characterized with high precision using our previous model. The results demonstrate the robustness of the model with the approximate overall accuracy up to 90%.

KEYWORDS

coronary artery, deep learning, optical coherence tomography, tissue characterization

Abbreviations: ACS, acute coronary syndrome; CAD, coronary artery disease; CHD, coronary heart disease; CNN, convolutional neural network; FCN, fully convolutional network; IV-OCT, intravascular optical coherence tomography; KD, Kawasaki disease; RF, random forest; ROI, region of interest.

1 | INTRODUCTION

Coronary artery disease leads to progression of pathological formations in arterial wall layers, which may be followed by acute coronary syndrome. Considering the significant role of coronary arteries in functionality of cardiac tissues by controlling the blood flow to myocardium, coronary artery disease is recognized as the main cause of myocardial

infarction and sudden death. The early mechanism, which leads to acute myocardial infarction is the formation of intracoronary pathological tissues and vulnerable coronary plaque rupture. This requires a high resolution imaging modality to be identified [1]. Catheter-based imaging modalities demonstrate higher resolution to visualize intracoronary structural information than noninvasive imaging techniques such as magnetic resonance and computerized tomography. Intravascular ultrasound is widely used in cardiology to evaluate coronary artery tissue layers and pathological formations, but the low pullback speed and limited axial resolution of Intravascular Ultrasound (IVUS) (100-150 μm) restricted its application to evaluate various cases with intimal hyperplasia, and pathological formations. Intracoronary optical coherence tomography (OCT) is recognized as a feasible and safe imaging technique with higher resolution of 10 to 15 μm than IVUS imaging, which can provide detailed structural tissue information [2]. Intravascular optical coherence tomography (IV-OCT) is a catheter-based invasive imaging modality, which employs a bandwidth in the near-infrared spectrum with central wavelength of approximately 1300 nm. Using such wavelength results in the tissue penetration of 1 to 3 mm. A single fiber-optic in OCT is responsible to emit the light and record the back-scattering of light from the arterial wall by simultaneous rotation, and pullback along the arterial wall. OCT works based on interferometry principle to measure the back-scattered signal since the direct measurements are impossible due to the high speed of light. OCT is significantly used in cardiology for diagnostic assessment of coronary atherosclerosis. As a limitation, light is strongly attenuated by blood as a result of light absorption by hemoglobin, and scattering by the red blood cells. Therefore, blood clearance is required during the imaging process [3, 4].

1.1 | Significance of the coronary artery lesion classification

Normal coronary artery has a three-layered structure. The outermost arterial wall, adventitia, is responsible to protect the arterial wall from over stretching and serves the mechanical connections with surrounding tissues. Adventitia is recognized as a signal rich pattern in OCT images. Media is the second arterial wall layer, which is composed of smooth muscle cells, elastic lamina and collagen. Media is the most significant mechanical layer, which is visualized as a signal poor pattern in OCT images. Intima is the innermost arterial wall layer in direct contact with blood flow. Intima is composed of endothelial cells and it is recognized as a signal rich pattern in OCT images [5]. Coronary arteries are responsible to deliver blood to the cardiac muscle, which supplies the required amount of oxygen and nutrients to the heart muscle.

Therefore, coronary artery disease can be followed by serious implications. This can lead to myocardial infarction and sudden death. In 95% of patients with symptomatic coronary artery disease and intracoronary pathology, there is a risk of atherosclerosis. In the remaining 5% of the patients, there is a huge risk of inflammatory, degenerative or congenital diseases, which are serious cardiac complications [5, 6]. Therefore, evaluation of intracoronary tissues in acute phase of the disease is important to prevent myocardial infarction. Manual segmentation of the tissues in coronary artery images is tedious, time-consuming and particularly error-prone from one observer to another and interpretation of the OCT images are highly challenging, even for a trained expert. Fully automatic method based on recent machine learning techniques, particularly deep learning, would have significant impact on efficient clinical diagnosis of coronary artery disease as a robust indicator of progression of pathological formations.

In this study, the experiments were performed on OCT pullbacks obtained from patients with Kawasaki disease (KD). 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 infusion reduces the risk of coronary artery complications, about 5% of treated children, and 15% to 25% of untreated children suffer a risk of experiencing coronary artery aneurysms or ectasia. Intimal thickening, media disappearance, lamellar calcification, fibrosis, macrophage and neovascularization are the most distinguished pathological features of late coronary artery lesions in KD. In severe cases, they can lead to myocardial infarction and sudden death.

1.2 | Related works

Optical coefficient approaches are used for intracoronary tissue characterization in some studies [7–12]. Atherosclerotic plaque characterization is performed using attenuation and back-scattering coefficient from intracoronary OCT images by Schmitt et al [8]. Three different plaque types (fibrosis, lipid and calcification) are recognized by considering their attenuation coefficients in the work of Xu et al [9]. Soest et al [10] classified the plaque into two groups with high and low attenuation coefficients. This method was not robust to measure the back-scattering coefficient in the cases with the lack of intensity calibration. A tissue characterization model based on quantification of the attenuation coefficients at different penetration depths of intracoronary OCT images is proposed by Veermeer et al [11]. However, the multi-scattered signal was not considered in this study. In addition, the results for uniform-layered phantom do not show a good consistency [11]. Evaluation of attenuation coefficient,

back-scattering coefficient and pixel-wise intensity in intracoronary OCT images is used to characterize various tissues in the work of Liu et al [12]. Moreover, various machine learning approaches are used for intracoronary tissue characterization [13–15]. Ughi et al proposed a model to characterize atherosclerotic plaques using random forest (RF) as the classifier. Combination of texture features and attenuation coefficients are used for plaque classification in intracoronary OCT images [13]. Athanasion et al proposed a tissue characterization model using RF as the classifier to discriminate between calcification, lipid and fibrous plaques [14]. A-line modeling method for plaque characterization is proposed by Rico-Jimenez et al [15]. However, in the study of Rico-Jimenez et al the effect of blood was not considered. Also, A-line analysis cannot be generalized to all challenging cases since it is intensity-based.

Although optical properties of tissues and texture features can provide a fair description of tissues, but considering the challenges of the OCT images, detailed tissue information is required for better representation and evaluation of various arterial wall tissues. Other than this limitation, none of the mentioned studies limited to few lesion types, and there is no complete tissue characterization system proposed in these studies, which can start by evaluating the arterial wall structure, and analyze all the tissues and lesions in details.

Recently, deep learning was widely used in the field of medical imaging for various applications [16–19]. Convolutional neural networks (CNNs) are recognized as the robust neural network architectures for classification tasks, where their output deep feature is a feature vector per image with an associated single class label predicted by the network. Since training a network from scratch requires a lot of data and considering the limited available data in the field of medical image analysis, it is efficient to use pre-trained networks by fine-tuning and transfer learning. Therefore, using the same architecture of the pre-trained network, the weights at each layer are initialized by the weights of the pre-trained network to start the iterative weight update by layer-wised fine-tuning to find the optimal parameters for the new application [20]. A CNN-based plaque characterization method is proposed by He et al [21], which results do not show a good precision of the method to be used by clinicians. In another study, the performance of a CNN and an artificial neural network is compared to characterize calcification, and lipid vs other tissues [22]. However, all the pathological tissues and normal cases were not considered in this study. In these studies, pre-processing to remove the catheter was performed using Otsu's threshold in the work of He et al. Also, extraction of the region of interests was performed using OCT A-lines, which is mostly intensity-based. These approaches are not reliable to be generalized to all the cases, specifically challenging cases considering the artifacts of the

imaging system, noise, and challenging cases in terms of disease complexity. Also, the results show the low precision of the method to characterize the lesions.

Considering the strength of CNN features to describe various intracoronary tissues, in our previous studies, we designed a tissue characterization model to discriminate between arterial wall tissue layers, intima and media, as well as pathological formations, specifically calcification, neovascularization, fibrosis and macrophage accumulation. The final tissue characterization model was designed using CNNs as feature extractors from each tissue to train RF as a classifier. Majority voting approach was used for final classification decision. The model is highly precise to characterize various intracoronary tissues [23, 24]. In our previous studies, the main contribution was to find the features and classifier, which are reliable for intracoronary tissue characterization using OCT images. Therefore, the performance of various classifiers was assessed to design our model. In our previous study, pre-processing to remove the surrounding arterial wall tissues was performed on each frame of the OCT pullbacks by applying active contour and connected component approaches. In our previous studies, the OCT pullbacks with normal and diseased arterial wall structures were determined by expert cardiologist. We did not discriminate between the normal and diseased arterial wall structures automatically. To characterize the lesion types, the lesions were extracted manually using the ground-truth. Then, the type of lesions was characterized using our proposed model. In this study, we consider all the limitations of our previous studies since we aimed for a fully automatic and complete diagnostic system to analyze the coronary artery tissues.

1.3 | General limitations of the related works

1. A complete intracoronary tissue characterization model can be useful for clinicians for early detection of pathological tissues. To our knowledge, as the main limitation of all the current studies, there is no complete framework, which starts by evaluating the arterial wall structure to distinguish between normal and diseased arterial wall structure. Also, all the proposed tissue characterization models focused on characterizing a limited number or specific coronary artery lesions. There is no tissue characterization model, which can automatically detect all coronary artery tissue layers, and pathological lesions.
2. Pre-processing steps are additional computational steps, which applied in all the proposed methods in related works to prepare the images for classification task. Also, designing a pre-processing approach, which can be generalized to all the cases is very challenging.

3. In all the existing studies, the region of interests (lesions) were extracted either manually or during pre-processing steps in order to train the classification model. To have a fully automatic tissue characterization model, this step should be automatic as well.
4. Patch-based classification using CNNs has some limitations: (a) The network should be run for each patch separately, which results in redundant feature extraction process due to the overlapping patches. (b) The patch size selection is challenging since using small patches, the network considers a small context because of applying max-pooling. Using large patches may require more max-pooling steps, this can reduce the localization accuracy. (c) Considering the huge computational time, the pre-processing steps are required to remove the unnecessary tissues and reduce the processing time, which is an additional step.

Designing an automatic tissue characterization system for intracoronary OCT images is a challenging task considering the limited available coronary OCT data, artifacts of the system, residual blood and very small size of the artery. In addition, we have to consider a solution, which is precise with low computational burden to be useful for clinicians.

1.4 | Contributions

As the main contribution, in this study for the first-time, a deep learning-based computer-aided diagnostic model providing clinicians with an operator-independent diagnosis of coronary artery lesions was developed. This study contributes to:

1. Automatic evaluation of the arterial wall structure using deep features: To have a complete tissue characterization model, as the first step, it is very efficient if the model can automatically recognize between normal and pathological arterial wall structure.
2. Automatic intracoronary tissue analysis by developing a VGG-based fully convolutional network (FCN): According to the results of the previous step, the model can look for arterial wall layers in the coronary arteries with normal three-layered structure and lesions in pathological cases.
3. Avoiding the pre-processing steps: First, the pre-processing steps are additional computational steps and second, designing a pre-processing approach, which can be generalized to all the cases is very challenging. This can decrease the certainty level of the tissue characterization model, specifically in diseased coronary arteries. Considering the artifacts of the imaging system, and the small size of the arterial wall as well as the limitations of

the traditional approaches. It is more efficient if we can consider the original images without applying any pre-processing step to assure that all the details of the tissues are considered in our analyses.

This work is organized as follows: The proposed method is explained in Section 2. The results are discussed in Section 3, and the work is concluded in Section 4.

2 | MATERIALS AND METHODS

Different steps of our tissue characterization model are shown in the flowchart of Figure 1. The main focus of this study is to propose a fully automatic intracoronary tissue characterization model to evaluate coronary artery tissue layers as well as pathological formations. The proposed model is designed in the following steps:

1. Evaluating each OCT pullback frame by frame: The model starts by evaluating the overall structure of the arterial wall for each frame of the OCT pullback using a CNN-based approach (Section 2.2).
2. Characterization of different tissues based on the result of the first step for each frame of the OCT pullback (Section 2.3):

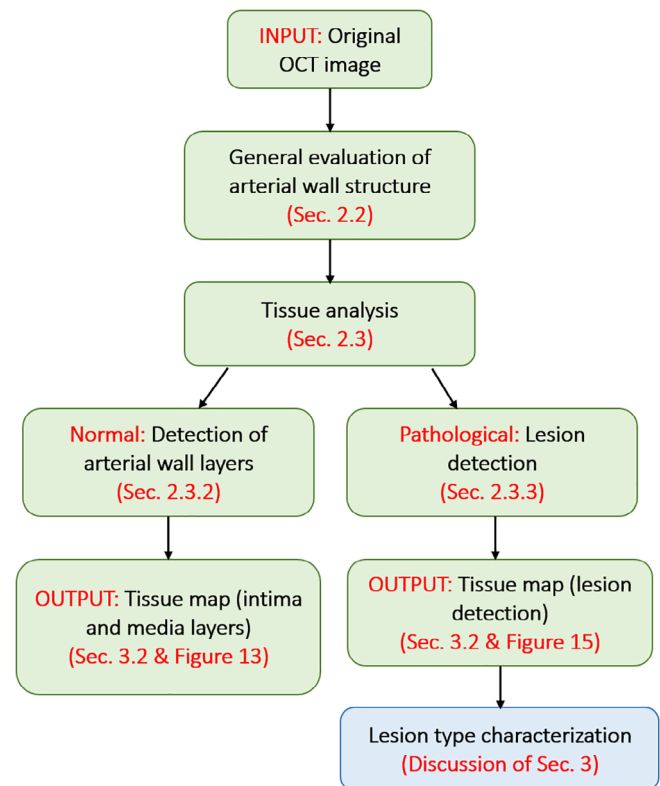


FIGURE 1 All steps of the proposed tissue characterization model

- If the arterial wall structure was recognized as normal in the first step, the model segments the arterial wall layers using FCN-based models. Tissue layer segmentation is important because it can assist clinicians in evaluating the arterial tissue by estimating the thickness of each layer (Section 2.3.2).
 - If the arterial wall structure was recognized as diseased, extraction of pathological formations (lesions) was performed. In pathological cases, the model looks for all the possible lesions developed in the arterial wall tissues due to the coronary artery disease. In this step, a FCN-based approach is used to extract all the lesions regardless of the lesion type (Section 2.3.3).
3. Lesion-type characterization: Having all the lesions extracted, we applied a CNN-based method with a final decision using majority voting to characterize the lesion types. This step was proposed in our previous work [24]. Although this step is not the main contribution of the current study, we discussed the final results of lesion-type characterization in Section 3 to demonstrate the final output of the complete intracoronary tissue characterization system. Each step of the work is explained in details in the following sections

2.1 | Data collection

The experiments are performed on 45 OCT pullbacks obtained from different patients with KD. Each OCT pullback consists of approximately 100 frames of DICOM images per patient. There are some cases with more than 100 frames per OCT pullback. The total number of frames, which are used for the experiments are 5040 frames. Twenty-six OCT pullbacks with the total of 2900 frames are considered as normal in this study since the three-layered structure of the arterial wall was preserved in all these frames. The other 19 OCT pullbacks with the total of 2140 frames are considered as pathological cases with neo-intimal development and developed lesions. Image acquisition is performed using FD-OCT (St. Jude Medical Inc., St. Paul, Minnesota) with the pullback speed of 20 mm/s. The axial and lateral resolutions of the OCT system are 12 to 15 and 20 to 40 μm , respectively. Permission to conduct this study on retrospective OCT studies was granted by the institutional review board.

OCT images were labeled by trained operator using custom software in Matlab. Each annotated image was reviewed by two cardiologists and if there was any disagreement, a consensus was reached by reviewing carefully each region of interest.

2.2 | General evaluation of arterial wall structure

In this step of the work, we investigated the features that a CNN learns to train RF as the classifier to discriminate between normal and diseased frames. Normal structure is referred to the three-layered structure of the arterial wall even if the artery is affected by disease, the three-layered structure is preserved. The diseased arterial wall structure is referred to the neo-intimal development [25] (Figure 2).

2.2.1 | Pre-trained VGG-19

CNNs are developed on convolutional layers. These layers are responsible to excite features from the local receptive field of the input image. Therefore, convolutional layers are composed of shared weights between the nodes to extract the similar local attributes in the input channels by sliding the filters through the input image with defined step size called stride. The extracted feature map from each convolutional layer is the input of the next layer. Using a non-saturating activation function, Rectified linear unit (ReLU), that replaces the negative values by zero in the feature map, which can accelerate the training process [26]. The pooling or subsampling is used for dimensionality reduction by keeping the most important information. To evaluate the arterial wall structure, the general information of the shape and borders and some details of the texture are enough at this step. Therefore, a CNN, which can provide us such information, can be quite useful.

VGG-19 is originally in the category of deep CNNs [27]. Deep networks consist of stacks of convolutional layers with very small receptive field kernels. The input size of VGG-19

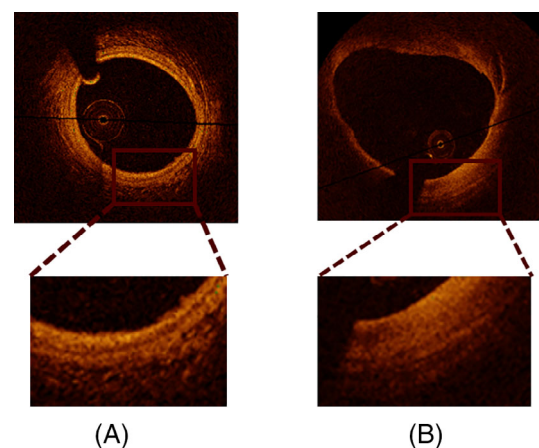


FIGURE 2 Normal structure of coronary artery is shown in (A). The arterial wall has three-layered structure with intima, media, adventitia and surrounding tissues. The diseased arterial wall with neo-intimal development and disappearance of the media layer is shown in (B)

is defined as $224 \times 224 \times 3$, which is fixed. Since OCT images are RGB images, the depth of our input is 3 and we needed to resize the OCT images with the original size of $352 \times 352 \times 3$ to $224 \times 224 \times 3$. The only pre-processing is that the mean RGB value was subtracted from each pixel for each input image. As it is shown in Figure 3, the VGG-19 architecture consists of five stacks of convolutional layers. Specific number of filters with small receptive field of 3×3 are applied in each convolutional layer with convolutional stride of 1 pixel. Also, 1-pixel padding preserves the spatial resolution through the convolutional layers. Considering that the output of each layer is the input of the next layer, for each convolutional layer the number of parameters is calculated as number of filters \times filter size \times input depth + number of filters (each filter has a bias). Also, the depth of each layer output is the number of filters at each convolutional layer. Every convolutional layer follows by a ReLU as the activation function to introduce nonlinearity. After each stack of convolutional layers, max pooling is applied for dimensionality reduction. Therefore, there is no learning parameter defined for pooling layers. Two-by-two-pixel window with stride of 2 is used for max pooling. Therefore, in pooling layers, the output is 1 pixel for every 2×2 -pixels and the windows do not have any overlap since the stride is defined as 2. As a result, the spatial resolution divides by 2 in each pooling layer while the depth remains

the same. The original VGG-19 architecture consists of two fully connected layers (fc6, fc7) with 4096 units and a classification layer (fc8). We removed the classification layer since we applied VGG-19 as feature extractor. We extracted the features from fc7, which is the last fully connected layer right before the classification layer.

To compare the activations excited by each layer of the network with the original image, all the activations are projected to the input pixel space [28]. Using the ReLU function, the positive activations are used to build the final feature map. Considering that the ReLU replaces negative values by zero, the white regions in Figure 4 show the positive activations. The channels in each layer learn various activations. The first layers learn and excite the abstract level information regarding the shape, corners and edges of the original image, which can effectively capture information regarding the borders of different layers. Complex invariance by evaluating the texture information is recognized in deeper layers. This may determine the robustness of deep features to describe and evaluate the general structural differences between images to discriminate between normal and diseased frames in various OCT pullbacks. Figure 4 is the representation of few activations extracted by VGG-19, which demonstrates the usefulness of this network as feature extractor for the first step of the work. The feature vectors for each frame of OCT pullbacks are extracted from

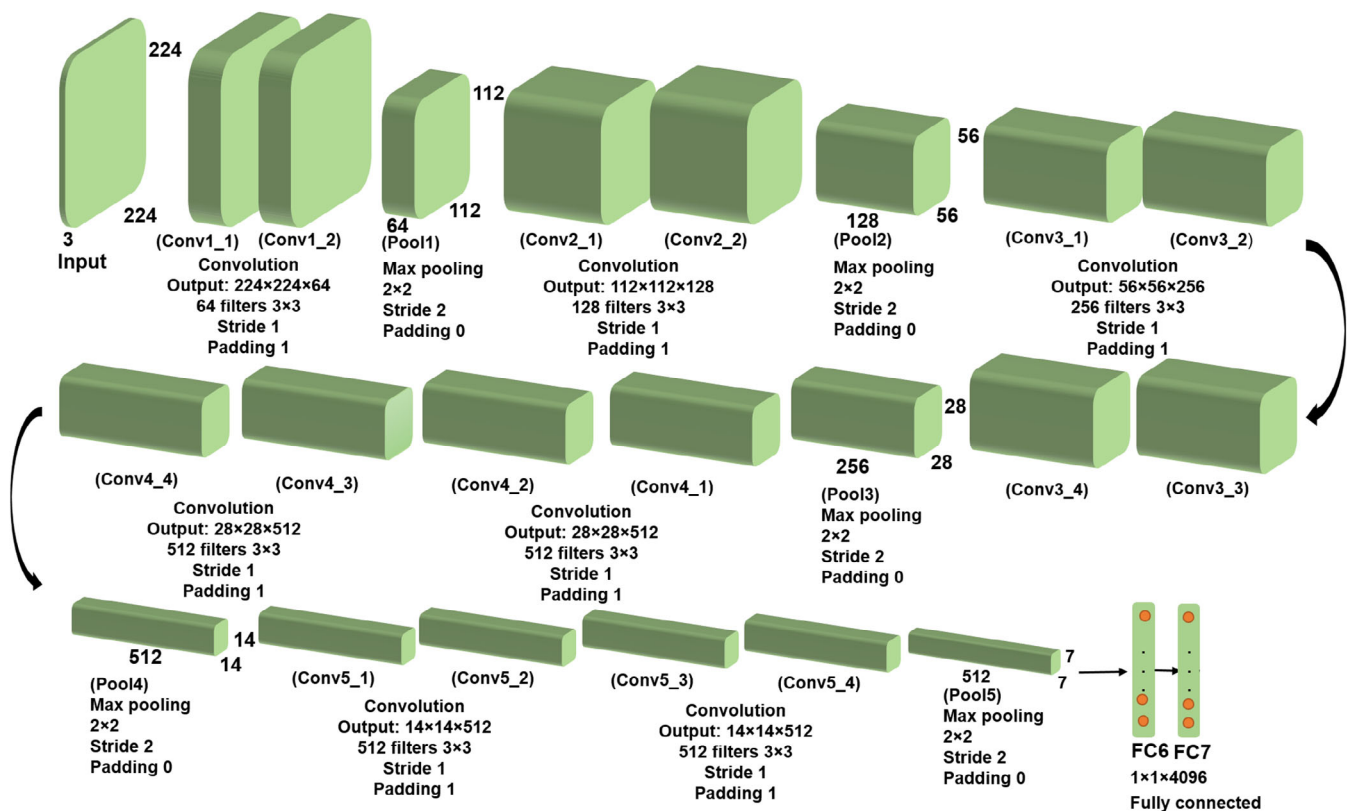


FIGURE 3 The architecture of VGG-19

layer fc7, which is the fully connected layer right before the classification layer. These feature vectors are used to train RF.

2.2.2 | Random forest

RF was developed based on generating an ensemble of trees [29, 30], which has considerable advantages compared against other existing classifiers:

1. The random vectors can control the growth of the trees in the ensemble and this increases the classification accuracy using RF.
2. RF can perform accurate classification on large datasets.
3. There is a very low risk of over-fitting in using RF compared against other classifiers.
4. In the case of medical images, which are noisy images, RF is robust to deal with noisy data.

The CART methodology is used to grow the trees to maximum size without pruning. The accuracy of RF depends on the strength of each tree (s), and the correlation between the trees (ρ). The smaller the ρ/s^2 ratio results in the better functioning of RF. In our experiments, the optimal number of trees is set to 429 by evaluating the performance of RF using out of bag error rate (Figure 5) for 500 of trees.

Fewer number of trees reduces the computational complexity in training the RF model. The number of randomly selected predictors (mtry) is set to 7 using grid searching.

2.2.3 | Training and validation

We consider all the frames of the OCT pullbacks obtained from the total of 45 patients. We resize the images to $224 \times 224 \times 3$, which is the input size of VGG-19. To generalize the model to all the cases, we extract the features from each frame of the OCT pullbacks and consider all the features with associated labels as a single feature matrix. Then, we split the features into training, validation and test sets. To ensure that consecutive frames are not selected and our system is not biased. First, we split randomly 75% of the feature vectors as the training features, and the remaining 25% of the features as the test features. Then, to create the validation set, we split the training set into the validation set by randomly selecting the 25% of the training set and the remaining 50% built the final training set. RF is trained using deep features of the training set and it is validated and tested using validation and test sets, respectively to classify between normal and diseased arterial wall structures.

To reduce the computational burden, we preformed the experiments by one random selection of training, validation and test sets in this step of the work. However, we

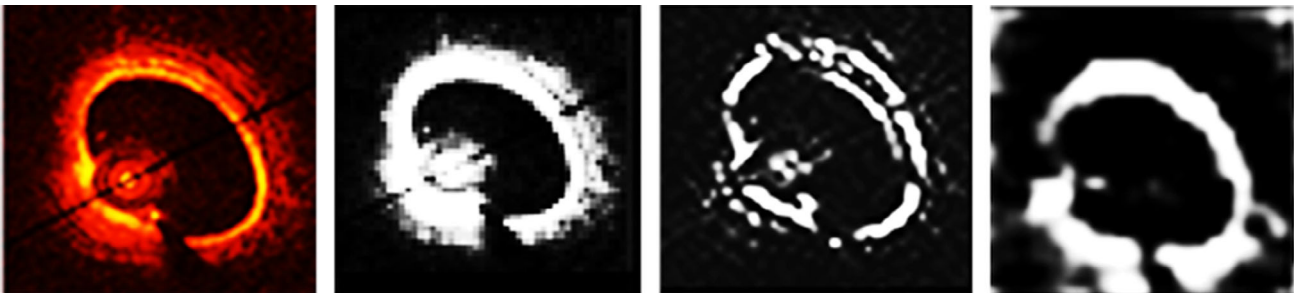
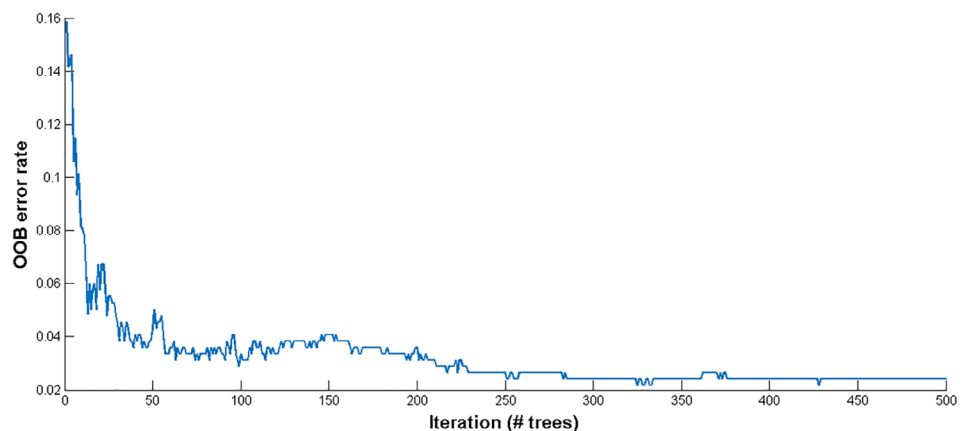


FIGURE 4 Visualization of few activations learned by VGG-19 architecture in different layers. Projection of features in pixel space demonstrate the usefulness of the network to determine the general structure of the arterial wall

FIGURE 5 Out of bag error rate to find the optimal number of trees of Random Forest. The error rate remains almost constant from the tree number 429



performed leave-one-out cross-validation to assure that there is no over-fitting, also to evaluate the performance of the model in different selections of training, and validation sets.

Therefore, in this step, VGG-19 is used as feature generator to train RF as the classifier to generally discriminate between normal and diseased images in each pullback.

2.3 | Tissue analysis

In this step, if the pullback frame was recognized as normal in the first step, the system looks for segmenting different arterial wall layers (intima and media), and if the pullback frame was recognized as diseased in the previous step, in this pathological cases, the system looks for possible developed lesions in arterial wall tissues. We used FCN-based approach in this step for the following reasons: we do not have the region of interests extracted (the arterial wall layers in normal cases and the lesions in pathological cases) to feed them to a CNN and obtain the features to distinguish between various tissues of coronary artery. Therefore, if we want to use CNN, we should use it as patch-based classification. However, using CNN for patch-based classification has some important limitations: (a) we should perform patch-based classification and run the network for each patch in the image separately. For each patch, we have a single feature vector and this is computationally very expensive because in patch-based approaches, the patches overlap and this results in redundant feature extraction. Pre-processing of the images to remove unnecessary information from the images can help the processing time using CNNs, but pre-processing is an additional step, and there is a risk of losing important information during the pre-processing as well. On the other hand, patch size selection is very challenging because the patch sizes should fit the size of the filters defined for each layer of the network. For all these reasons, CNNs are not very useful at this step of the work. To solve this problem, we found that FCNs are very efficient in this step because they do not have the limitations of the patch-based classification using CNNs. The network is trained end-to-end, pixels-to-pixels to exceed the training process and it avoids the problem of redundant feature extraction of multiple overlapped patches. Therefore, there is no need for pre-processing because the training is fast enough to consider the whole image and make sure that all the information is considered in our analyses.

2.3.1 | Fully convolutional networks

Generally, the network architecture used for semantic segmentation is composed of an encoder network followed by a decoder network. The encoder can be a pre-trained network. The decoder is responsible to project the learning features by

the encoder from the feature space to the pixel space to get the dense classification. FCNs perform semantic segmentation by considering the context as well as each pixel localization in the images. Compared to CNNs, FCNs can take the image with an arbitrary size as an input, since there is no fully connected layer involved to restrict the input size. FCNs are built on locally connected convolution, pooling and up-sampling layers. The network does not have any fully connected layer, which considerably reduce the number of parameters and training time. Considering the local connected layers in the network architecture, the network works independently from the original image size [31]. In our experiments, the input size of the network is defined as $177 \times 360 \times 3$ based on our application. The main parts of the FCNs are down-sampling path to extract the contextual information and the up-sampling path to recover the pixel localization.

VGG-based FCN

One of the standard networks, which is used as the basis of semantic segmentation is VGG architecture. VGG-19 is used in this study. Using VGG-based FCN, the knowledge is transferred from VGG-19 to perform semantic segmentation. The VGG-19 is used as the encoder of the FCN model. Fully connected layers are converted to fully convolutional layers using 1×1 convolution, which produce the feature map. Then, the up-sampling is started to convert the feature map from feature space to pixel space using transposed convolutions. Besides the deconvolutional layers, up-pooling is required as well. Considering that the max-pooling operation is noninvertible, the max location switches are recorded during max-pooling to approximately reconstruct the data from the above layer using the recorded positions.

Network configuration

The learning parameters are set by grid searching to find the optimal parameters at each layer. For transfer learning and fine-tuning, we initialized the weights of each layer of the VGG-based FCN by transferring the weights from the pre-trained VGG-19. The first layers of the network provide the abstract level information regarding the shapes, borders and edges, which are general attributes. Therefore, we started fine-tuning from the last layer of the network and we continued 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 on validation set. An extensive interval of learning rate values is chosen for grid searching to find the optimal learning parameters at each layer of the network. The interval of learning rate values is chosen from zero to two with the step of 0.001. We kept the momentum and the scheduling rate at 0.9 and 0.95, respectively at each step of fine-tuning. The learning rates for all the

convolutional layers are set to 0.01 based on the best performance of the network. Other significant factors, which determine the performance precision of FCNs are the choices of loss function and optimizer. In both normal and pathological cases, the arterial wall layers (intima and media) and the lesions represent a very small fraction of the cross-sectional coronary artery images, which causes the occurrence of the class imbalance problem. This results in suboptimal performance of the network. To deal with the problem of class imbalance, we chose Generalized Dice Loss (GDL) as our objective function [32], which is defined as follows,

$$GDL = 1 - [2(\sum_l w_l \sum_n r_{ln} p_{ln}) / (\sum_l w_l \sum_n r_{ln} + p_{ln})] \quad (1)$$

where w is the weight assigned to each class with label l , n is the number of image from the total of N images, r demonstrates the pixel values of the ground-truth assigned to each label for the image n , and p is the probabilistic decision map for each class with label l . The weights for each class label is calculated using the following equation,

$$w_l = 1 / (\sum_{n=1}^N r_{ln})^2 \quad (2)$$

The weights demonstrate the contribution of each label in minimizing the loss function, which is defined as the inverse of the region size for foreground (the region of interest), and background (all the other tissues). This makes the model suitable to deal with the class imbalance problem. We selected adaptive momentum estimation (Adam) for stochastic optimization, which uses the first order gradients with little memory requirement and fast convergence. Figure 6 shows the architecture of VGG-based FCN used in this study.

2.3.2 | Intima-media detection in normal cases

The experiments are performed on 26 various OCT pullbacks. All the frames at each pullback represent the coronary arterial wall with three-layered structure. Two

VGG-based FCNs with the same structure explained in the previous section are trained in parallel to extract intima and media layers, respectively. The first VGG-based FCN performs two class segmentation to extract intima layer vs all other tissues, and the second FCN takes the same frame of the pullback and simultaneously performs segmentation of the second layer, media, vs all other tissues. The segmentation result is combined as the final decision to extract intima and media layers. The steps of tissue layer detection are visually shown in Figure 7. The images are categorized in three sets of training, validation and test sets. The total of 70% of the images are selected to build the training set, 15% of the images are used for the validation set, and the remaining 15% of the images are used to build the test set.

2.3.3 | Lesion extraction in pathological cases

The experiments are performed on 19 different OCT pullbacks with pathological formations, such as calcification, neovascularization, fibrosis and macrophage accumulation. The VGG-based FCN with the same architecture shown in Figure 6 is trained for two class segmentation of pathological tissues vs all other arterial wall tissues. Therefore, the output of the network is the detection of lesions regardless of the lesion type. The steps of lesion detection are visually shown in Figure 8. The total of 70% of the images are selected to build the training set, 15% of the images are used for the validation set and the remaining 15% of the images are used to build the test set.

We performed the experiments by one random selection of training, validation and test sets at each step. But, leave-one-out cross-validation was performed by leaving out the OCT images of one patient for validation and training the model on the OCT images of the remaining patients at each step of the experiment. The mean \pm SD of the overall accuracy for all the experiments are reported at each step. This is performed to assure that there is no over-fitting concern and also to evaluate the performance of the model in different selections of training, and validation sets.

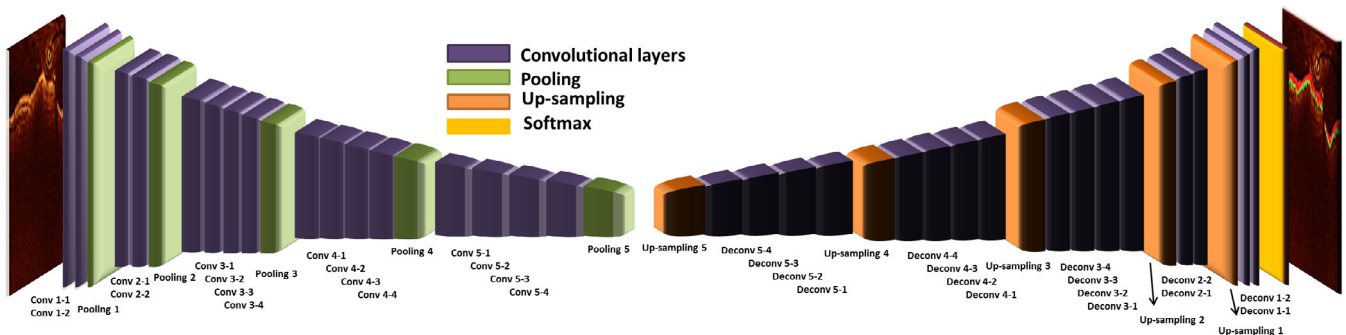


FIGURE 6 The architecture of VGG-based fully convolutional network

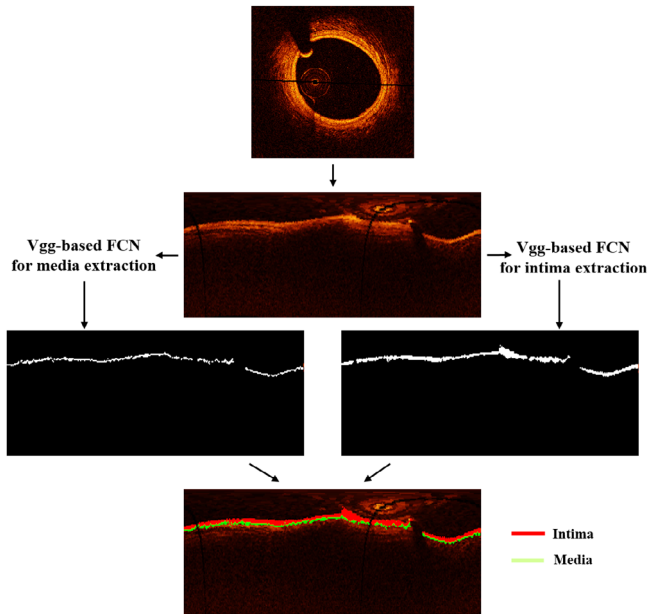


FIGURE 7 Visual representation of the VGG-based fully convolutional network (FCN) process to segment intima and media layers

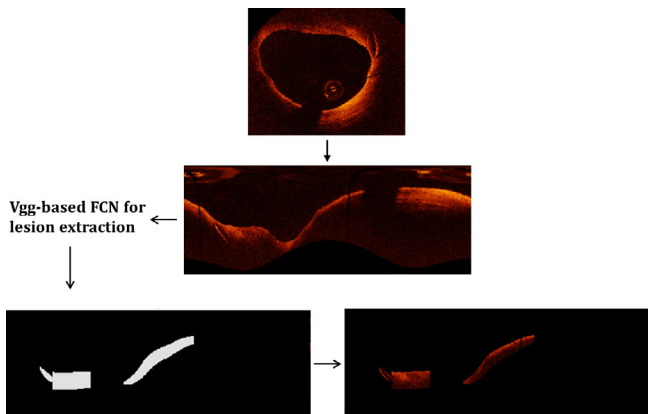


FIGURE 8 Visual representation of the VGG-based fully convolutional network (FCN) process to detect intracoronary lesions regardless of the lesion type

3 | RESULTS AND DISCUSSION

The experiments are performed on the total of 45 intracoronary OCT pullbacks obtained from patients with KD. The cross-sectional images of the 26 OCT pullbacks are recognized as the coronary artery segments with three-layered structure of the arterial wall that we called them normal structure. The cross-sectional images of the remaining 19 OCT pullbacks are recognized as diseased coronary artery segments with neo-intimal development and formation of lesions.

	Predicted Class	
	Positive	Negative
Positive	True Positive (TP)	False Positive (FP)
Negative	False Negative (FN)	True Negative (TN)

FIGURE 9 Confusion matrix structure used to evaluate the results. Positive refers to the class of interest (the tissue that we wanted to segment or characterize) and the Negative refers to the rest of the tissues including the image background

To evaluate the results, at each step of the work, we calculate the confusion matrix as it is shown in Figure 9. Having the confusion matrix, we measure the per class accuracy, sensitivity and specificity as follows,

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (4)$$

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

where TP, FP, FN and TN are true positive, false positive, false negative, and true negative, respectively. Using FCNs, since it is also efficient to measure the Boundary F1 score (BF-score) to validate the segmentation results, we calculated the BF-score as follows,

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

$$Sensitivity = \frac{TP}{TP + FN} \quad (7)$$

$$BFscore = \frac{2 \times Precision \times Sensitivity}{Precision + Sensitivity} \quad (8)$$

3.1 | General evaluation of the arterial wall structure

For the first step, RF is trained to evaluate the general structure of the arterial wall. The classification result is reported as measured per class accuracy, sensitivity and specificity in Table 1. The result shows the robustness of CNN features to detect the general structure of the arterial wall in normal and affected coronary arteries. The result of leave-one-out cross-validation to evaluate the general structure of the arterial wall is shown in details in Figure 10 for each patient that was left as validation set. The mean \pm SD of all the experiments were calculated as the overall accuracy of 0.94

TABLE 1 Measured accuracy, sensitivity and specificity to evaluate general arterial wall structure

	Arterial wall structure		
	Accuracy	Specificity	Sensitivity
Normal structure	0.95	0.97	0.94
Diseased structure	0.97	0.97	0.97

± 0.05 . Using CNN as feature extractor, we have one feature vector for each frame. We did not consider each OCT pullback separately. We considered different frames of different OCT pullbacks in training, validation and test sets since tissue texture can be different from one patient to another. This way, RF was trained on the feature vectors extracted from different frames of various patients. Therefore, the model can be generalized to all the cases since the training and test sets are not restricted to a single patient with specific tissue attributes.

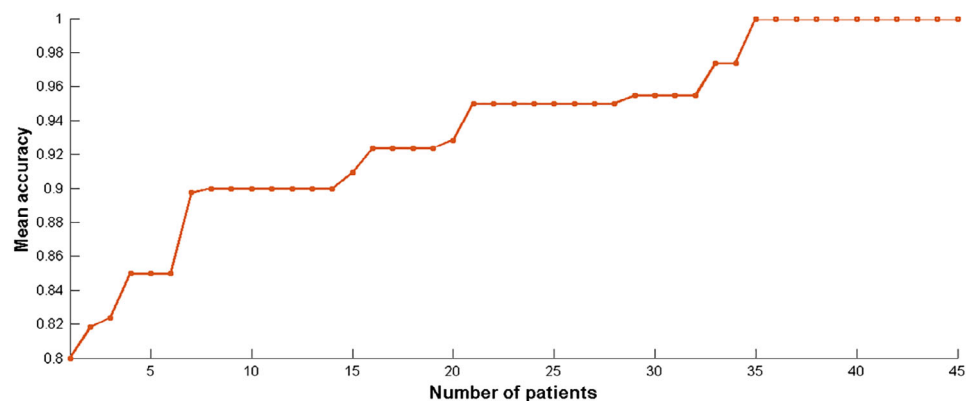
The model starts by evaluating each OCT pullback frame by frame. The reasons why pre-trained CNN and RF are used in this step of the work are as follows: (a) It is efficient to use pre-trained networks considering the fact that training a network from scratch requires a lot of data and we deal with the limited available data in the field of medical image analysis specifically for infants and children. (b) It is more efficient and computationally less expensive to use CNNs as feature extractors to train another classifier, when we deal with the problem of characterization between the region of interests (normal vs diseased arterial wall structure, where

the whole frame is our region of interest). Using CNNs as feature extractor avoids retraining the network during fine-tuning, which requires a considerable amount of time. (c) Over-fitting concerns in deep fine-tuning the network and finding proper learning rates for each layer are other issues of using CNNs as the classifier. There are many pre-trained CNNs that can be used as feature extractor, but in this step of the work, we wanted to evaluate each frame generally to discriminate between normal and pathological arterial wall structure. Since, VGG-19 is a strong network to perform feature extraction, and we used the architecture of VGG-19 to build the FCNs in the next steps of the work, to keep the model consistent in terms of choosing the networks, VGG-19 is used in this study. Using CNN as feature extractor, the processing time to extract the features for each frame is less than 10 seconds, and the training process using RF in this step of the work takes about 3 minutes, which is considerably fast.

3.2 | Tissue analysis

3.2.1 | Intima-media detection in normal cases

As the next step, in normal cases, we detect arterial wall layers (intima and media). Sometimes, the artery is affected by disease, which results in thickening the arterial wall layers although the three-layered structure of the arterial wall is preserved. We reported the mean \pm SD of the measured per class accuracy, specificity and sensitivity as well as the BF-score of all the test set images in Table 2 for intima, and

FIGURE 10 Leave-one-out cross-validation to evaluate the general arterial wall structure. The mean accuracy of the classification is measured at each iteration to evaluate the performance of the model**TABLE 2** Measured accuracy, sensitivity, specificity and Boundary F1 score (BF-score) for intima and media detection using FCN model

	Arterial wall layers			
	Accuracy	Specificity	Sensitivity	BF-score
Intima	0.90 ± 0.04	0.86 ± 0.06	0.93 ± 0.03	0.99 ± 0.01
Media	0.87 ± 0.04	0.82 ± 0.05	0.91 ± 0.02	0.99 ± 0.01

Abbreviation: FCN, fully convolutional network.

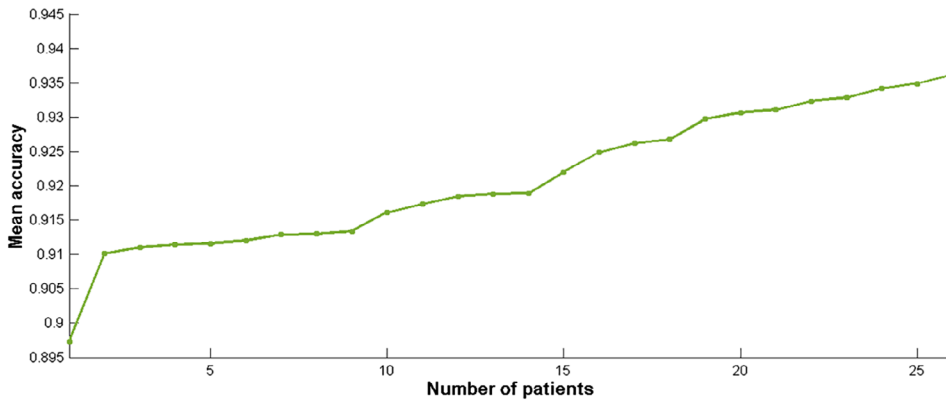


FIGURE 11 Leave-one-out cross-validation for detection of the first layer of the arterial wall (intima) in normal cases using VGG-based fully convolutional network (FCN). The mean accuracy of the classification is measured for each patient to evaluate the performance of the model

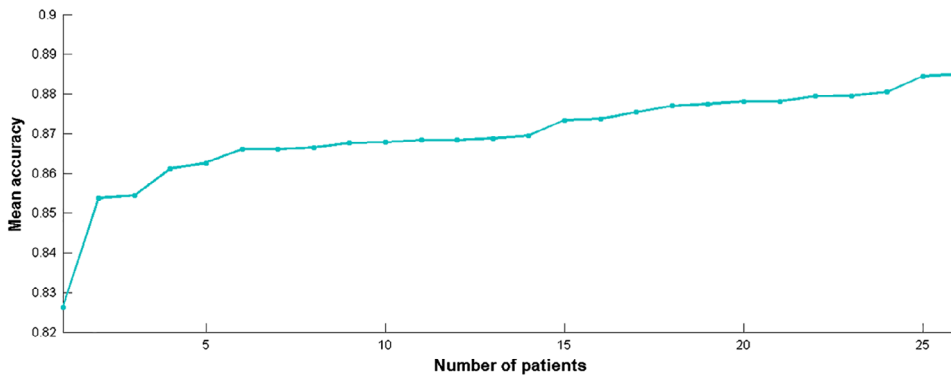


FIGURE 12 Leave-one-out cross-validation for detection of the second layer of the arterial wall (media) in normal cases using VGG-based fully convolutional network (FCN). The mean accuracy of the classification is measured for each patient to evaluate the performance of the model

media detection. The results of leave-one-out cross-validation to detect intima and media are shown in details in Figures 11 and 12, respectively for each patient that was left as validation set. The mean \pm SD of the results of all the experiments were calculated as the overall accuracy of 0.92 ± 0.01 and 0.87 ± 0.01 for intima and media, respectively. The results demonstrate the good performance of the model using different selections of training and validation sets. This result can also overcome the over-fitting concern. Figure 13 is the visual representation of the intima, and media detection for the frames of three different OCT pullbacks of various patients.

3.2.2 | Lesion extraction in pathological cases

In pathological cases, we extract the pathological tissues regardless of tissue type. In this step, we aimed to extract all the developed lesions automatically. We reported the mean \pm SD of the measured per class accuracy, specificity and sensitivity as well as the BF-score in Table 3 for extraction of the lesions. The result of leave-one-out cross-validation to extract the lesions is shown in Figure 14 for each patient that was left as validation set. The mean \pm SD of the results of all the experiments were calculated as the overall accuracy of 0.95 ± 0.02 for all the experiments. The results demonstrate the good performance of the model using different selections of training and validation sets as well as

overcoming the over-fitting concern. Figure 15 is the visual representation of the pathological tissue extraction for the frames of four different OCT pullbacks of various patients. The results show a high precision of the model to extract the lesions, which is the most challenging and significant problem in coronary arteries affected by disease.

All the lesions, which are extracted in this step of the work, are fed to the model developed in our previous study to characterize the type of each lesion and show the results for our complete tissue characterization framework of intracoronary OCT images. To characterize the lesion type, we extracted features from three different CNN networks (AlexNet, VGG-19 and Inceptionv3). For each set of features, we trained RF as the classifier. Then, we used majority voting for final classification result using all the RF decisions. This approach is explained in details in our previous study [24]. The final results of lesion type characterization are reported in Table 4.

Using the FCN-based model, we could avoid the pre-processing steps that were performed in our previous study. This can help the model to be less expensive computationally and more accurate since we make sure that there is no tissue information, which was missed during the pre-processing steps. In addition, technically, it is very difficult to train the networks from scratch since we have limited number of available datasets specifically in infants and children population. To overcome this problem, we fine-tuned

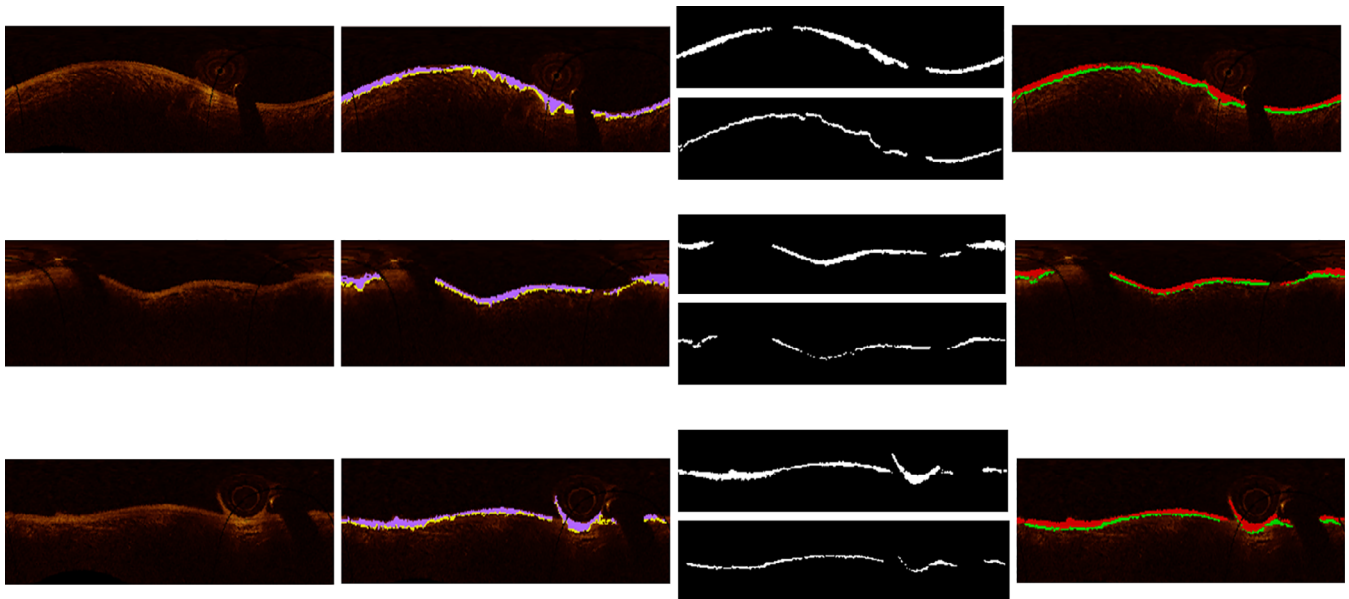


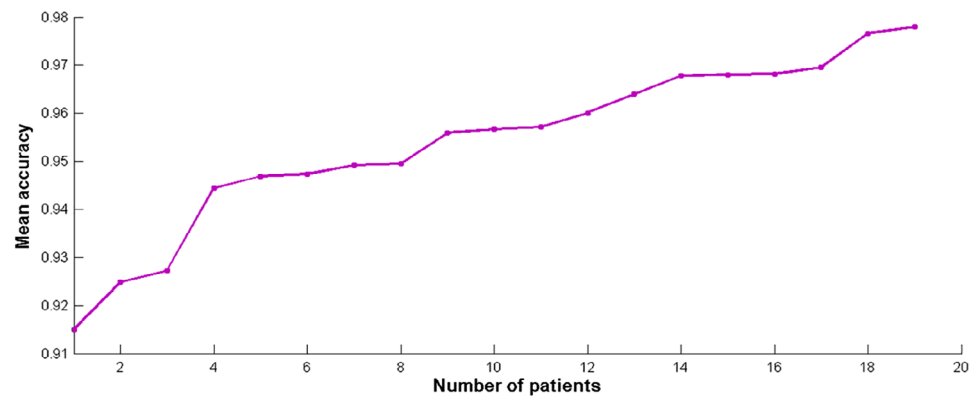
FIGURE 13 Visual representation of the VGG-based fully convolutional network (FCN) results to detect intima and media layers for three frames of three different patients. From left: the first image is the planar representation of the original optical coherence tomography image, the second image is the ground-truth, which shows intima with purple label and media with yellow label, the third image is the network result to extract the tissues (the upper region is intima, and the other one is media), and the fourth image is the overlap of the network results on the original image

TABLE 3 Measured accuracy, sensitivity, specificity and Boundary F1 score (BF-score) of lesion detection using FCN model

	Pathological cases			
	Accuracy	Specificity	Sensitivity	BF-score
Pathological tissues	0.96 ± 0.04	0.95 ± 0.05	0.97 ± 0.03	0.96 ± 0.04

Abbreviation: FCN, fully convolutional network.

FIGURE 14 Leave-one-out cross-validation for lesion extraction in pathological cases using VGG-based fully convolutional network (FCN). The mean accuracy of the classification is measured for each patient to evaluate the performance of the model



pre-trained networks. Therefore, instead of starting from scratch to initialize the weights of each layer, we initialized the weights of our network with the weights of the pre-trained network (VGG-19) and fine-tuned the parameters in each layer to make sure that the model performs accurately in our application. Then, we investigated the proper loss function to train our model since we were dealing with the problem of imbalanced classes during the training considering that the arterial wall layers and lesions are very small

areas compared against the surrounding tissues and the background of the image.

In normal cases, two VGG-based FCNs with the same architecture are trained separately to segment intima vs other tissues and media vs other tissues, respectively. Therefore, if the frame was recognized as normal in the first step of the model, the two VGG-based FCNs work in parallel to detect intima and media layers in the image. Training the network is considerably fast, which was

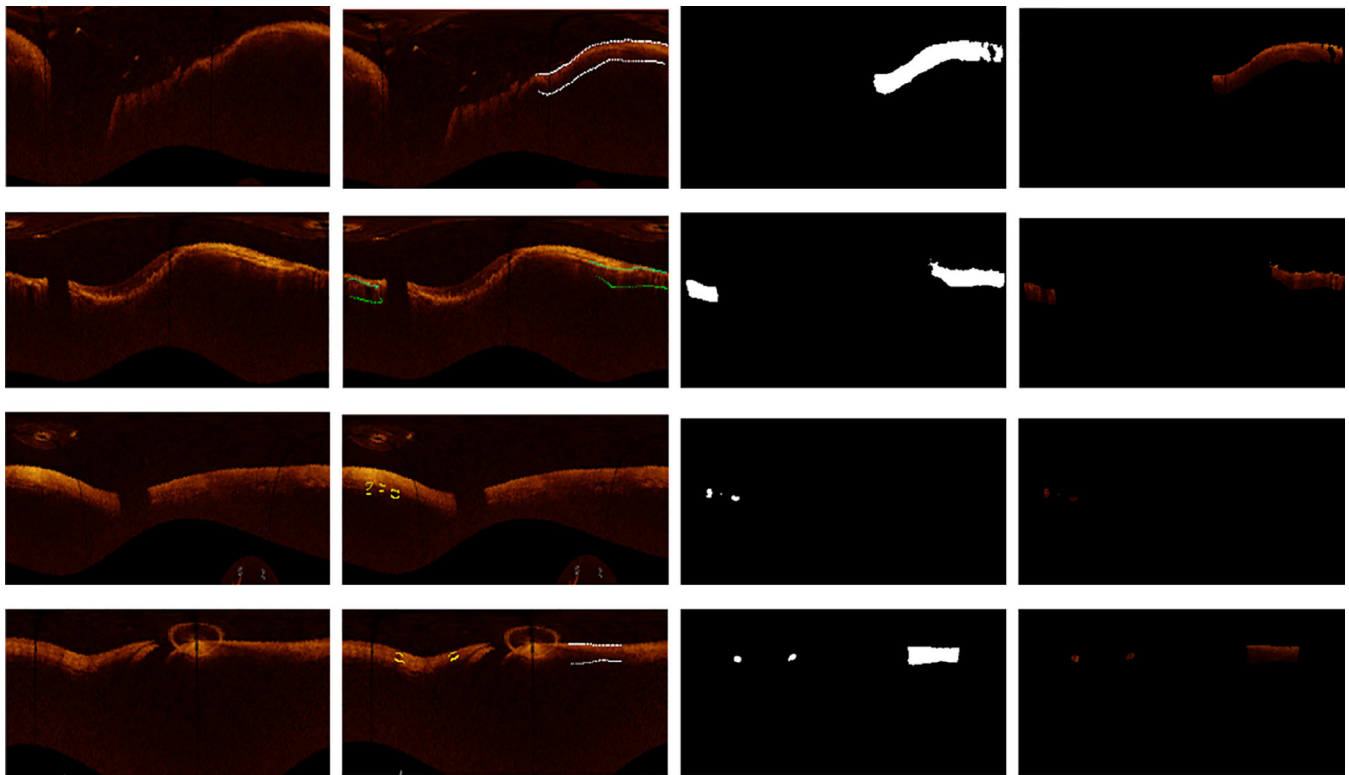


FIGURE 15 Visual representation of the VGG-based fully convolutional network (FCN) results to detect pathological tissues for one frame of four different patients. From left: the first image is the planar representation of the original optical coherence tomography (OCT) image, the second image is the ground-truth, which is manual segmentation of the pathological tissues in OCT images, the third image is the network result to extract the lesions regardless of the tissue type, and the fourth image is the extraction of all the regions, which is detected as pathological tissues from the original image

TABLE 4 Measured accuracy, sensitivity and specificity to characterize lesion types

	Lesion type		
	Accuracy	Specificity	Sensitivity
Calcification	0.90	0.95	0.84
Fibrosis	0.94	0.96	0.94
Macrophage	0.92	0.97	0.89
Neovascularization	0.95	0.97	0.90

approximately 59 minutes with 354 iteration per epoch and the total of 35 400 epoch.

In pathological cases, the same FCN architecture was used, but it was trained to extract the lesions regardless of lesion type in each frame of the OCT pullback. The network training time was 41 minutes and 22 seconds. The network converged in 100 epoch with the total of 4900 iterations (49 iteration per epoch).

In this study, different factors were considered. First, we do not have all types of the pathological tissues in all the

frames of each pullback. In some cases, intima is thickened without the development of pathological tissues. In other cases, we may have one, two or more pathological tissues developed in the arterial wall layer as a result of the disease. Second, the number of images, particularly in pathological cases, are very limited. Therefore, training a FCN to segment all the tissue types is not possible since we aimed to propose a model, which is not limited to a specific type of pathological formations and can be extended to all pathological lesions. Although we considered the four most significant coronary artery complications caused by CAD in this work, but we may extend the model to other possible lesions. Therefore, it is not wise to train a single FCN model for each pathological tissue type separately since it is computationally very expensive and requires a huge memory. For this reason, we decided to train a FCN, which can extract all the pathological tissues without considering the lesion type. Then, using our proposed tissue characterization model [24], extracting the CNN features, and training a RF to distinguish between the tissue types demonstrated a high precision.

4 | CONCLUSION

In this study, we aimed to propose a fully automatic tissue characterization model, which can assist clinicians for better diagnosis of the coronary artery complications caused by coronary artery disease using OCT images. Our complete tissue characterization model starts by evaluating the arterial wall tissue structure for each frame of the pullback to recognize between the normal three-layered structure of the arterial wall and neo-intimal development. Then, in normal cases, the model can detect the arterial wall layers, and in pathological cases, all the existing lesions can be extracted regardless of the lesion type using a FCN model. The extracted lesions can be categorized based on the lesion type using CNN features and majority voting on RF decisions. Our future work will be concentrated on evaluating the distensibility variations of the arterial wall tissues to assess the mechanical properties of the arterial wall using stationary OCT.

ACKNOWLEDGMENTS

This study is supported by Natural Sciences and Engineering Research Council of Canada (NSERC), and Fonds BoBeau Coeur—Fondation CHU Ste-Justine. OCT equipment was acquired thanks to funding from JP Metal Children's fund, Montreal.

CONFLICT OF INTEREST

The authors declare no potential conflict of interests.

AUTHOR BIOGRAPHIES

Please see Supporting Information online.

ORCID

Atefeh Abdolmanafi  <https://orcid.org/0000-0001-7707-2158>

REFERENCES

- [1] E. Falk, M. Nakano, J. F. Bentzon, A. V. Finn, R. Virmani, *Eur. Heart J.* **2012**, *34*(10), 719.
- [2] I.-K. Jang, B. E. Bouma, D.-H. Kang, S.-J. Park, S.-W. Park, K.-B. Seung, K.-B. Choi, M. Shishkov, K. Schlendorf, E. Pomerantsev, S. L. Houser, H. T. Aretz, G. J. Tearney, *J. Am. Coll. Cardiol.* **2002**, *39*(4), 604.
- [3] H. G. Bezerra, M. A. Costa, G. Guagliumi, A. M. Rollins, D. I. Simon, *J. Am. Coll. Cardiol. Interv.* **2009**, *2*(11), 1035.
- [4] M. Terashima, H. Kaneda, T. Suzuki, *Korean J. Intern. Med.* **2012**, *27*(1), 1.
- [5] H. Chen, G. S. Kassab, *J. Biomech.* **2016**, *49*(12), 2548.
- [6] A. C. van der Wal, *Heart* **2007**, *93*(11), 1484.
- [7] J. M. Schmitt, A. Knüttel, R. F. Bonner, *Appl. Opt.* **1993**, *32*(30), 6032.
- [8] M. Joseph, A. Schmitt, Knüttel, M. Yadlowsky, M. A. Eckhaus, *Phys. Med. Biol.* **1994**, *39*(10), 1705.
- [9] C. Xu, J. M. Schmitt, S. G. Carlier, R. Virmani, *J. Biomed. Opt.* **2008**, *13*(3), 034003.
- [10] G. Van Soest, T. P. M. Goderie, E. Regar, S. Koljenovic, A. G. J. L. H. van Leenders, N. Gonzalo, S. van Noorden, T. Okamura, B. E. Bouma, G. J. Tearney, et al., *J. Biomed. Opt.* **2010**, *15*(1), 011105.
- [11] K. A. Vermeer, J. Mo, J. J. A. Weda, H. G. Lemij, J. F. De Boer, *Biomed. Opt. Express* **2014**, *5*(1), 322.
- [12] S. Liu, Y. Sotomi, J. Eggermont, G. Nakazawa, S. Torii, T. Ijichi, Y. Onuma, P. W. Serruys, B. P. F. Lelieveldt, J. Dijkstra, *J. Biomed. Opt.* **2017**, *22*(9), 096004.
- [13] G. J. Ughi, T. Adriaenssens, P. Sinnaeve, W. Desmet, J. D'hooge, *Biomed. Opt. Express* **2013**, *4*(7), 1014.
- [14] L. S. Athanasiou, C. V. Bourantas, G. Rigas, A. I. Sakellarios, T. P. Exarchos, P. K. Siogkas, A. Ricciardi, K. K. Naka, M. I. Papafaklis, L. K. Michalis, F. Prati, D. I. Fotiadis, *J. Biomed. Opt.* **2014**, *19*(2), 026009.
- [15] J. J. Rico-Jimenez, D. U. Campos-Delgado, M. Villiger, K. Otsuka, B. E. Bouma, J. A. Jo, *Biomed. Opt. Express* **2016**, *7*(10), 4069.
- [16] M. Biswas, V. Kuppili, D. R. Edla, H. S. Suri, L. Saba, R. T. Marinho, J. Miguel Sanches, J. S. Suri, *Comput. Methods Prog. Biomed.* **2018**, *155*, 165.
- [17] Y. Dong, Y. Pan, X. Zhao, R. Li, C. Yuan, W. Xu, in Smart Computing (SMARTCOMP), 2017 IEEE International Conf. on, IEEE **2017**, pp. 1–8.
- [18] K. Lekadir, A. Galimzianova, À. Betriu, M. del Mar Vila, L. Igual, D. L. Rubin, E. Fernández, P. Radeva, S. Napel, *IEEE J. Biomed. Health Inform.* **2017**, *21*(1), 48.
- [19] R.-M. Menchón-Lara, J.-L. Sancho-Gómez, A. Bueno-Crespo, *Appl. Soft Comput.* **2016**, *49*, 616.
- [20] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. A. Setio, F. Ciompi, M. Ghafoorian, J. A. Van Der Laak, B. Van Ginneken, C. I. Sánchez, *Med. Image Anal.* **2017**, *42*, 60.
- [21] S. He, J. Zheng, A. Maehara, G. Mintz, D. Tang, M. Anastasio, H. Li, In *Medical Imaging 2018: Image Processing*, International Society for Optics and Photonics, **2018**, p. 1057432.
- [22] C. Kolluru, D. Prabhu, Y. Gharaibeh, H. Bezerra, G. Guagliumi, D. Wilson, *J. Med. Imaging* **2018**, *5*(4), 044504.
- [23] A. Abdolmanafi, L. Duong, N. Dahdah, F. Cheriet, *Biomed. Opt. Express* **2017**, *8*(2), 1203.
- [24] A. Abdolmanafi, L. Duong, N. Dahdah, I. R. Adib, F. Cheriet, *Biomed. Opt. Express* **2018**, *9*(10), 4936.
- [25] A. Dionne, R. Ibrahim, C. Gebhard, M. Bakloul, J.-B. Selly, M. Leye, J. Déry, C. Lapierre, P. Girard, A. Fournier, et al., *J. Am. Heart Assoc.* **2015**, *4*(5), e001939.
- [26] A. Krizhevsky, I. Sutskever, G. E. Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems* **2012**, p. 1097.
- [27] K. Simonyan, A. Zisserman, arXiv preprint arXiv:1409.1556 **2014**.
- [28] M. D. Zeiler, R. Fergus. Visualizing and understanding convolutional networks. In *European Conference on Computer Vision*, Springer, Cham, **2014**, p. 818.

- [29] A. Criminisi, J. Shotton, *Decision Forests for Computer Vision and Medical Image Analysis*, Springer Science & Business Media, **2013**.
- [30] M. Kuhn, K. Johnson, *Applied Predictive Modeling*, Vol. 26, Springer, New York, **2013**.
- [31] E. Shelhamer, J. Long, T. Darrell, arXiv preprint arXiv: 1605.06211 **2016**.
- [32] C. H. Sudre, W. Li, T. Vercauteren, S. Ourselin, M. J. Cardoso. Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*, Springer, Cham, **2017**, p. 240.

How to cite this article: Abdolmanafi A, Cheriet F, Duong L, Ibrahim R, Dahdah N. An automatic diagnostic system of coronary artery lesions in Kawasaki disease using intravascular optical coherence tomography imaging. *J. Biophotonics*. 2020;13:e201900112. <https://doi.org/10.1002/jbio.201900112>