



# Coronary Plaque Characterization From Optical Coherence Tomography Imaging With a Two-Pathway Cascade Convolutional Neural Network Architecture

Yifan Yin<sup>1</sup>, Chunliu He<sup>1</sup>, Biao Xu<sup>2</sup> and Zhiyong Li<sup>1,3\*</sup>

<sup>1</sup> School of Biological Science and Medical Engineering, Southeast University, Nanjing, China, <sup>2</sup> Department of Cardiology, Nanjing Drum Tower Hospital, Nanjing, China, <sup>3</sup> School of Mechanical, Medical, and Process Engineering, Queensland University of Technology, Brisbane, QLD, Australia

## OPEN ACCESS

### Edited by:

Antonios Karanasos,  
Hippokraton General Hospital, Greece

### Reviewed by:

Lampros Athanasiou,  
Canon, United States  
Kenichiro Otsuka,  
Massachusetts General Hospital,  
United States

### \*Correspondence:

Zhiyong Li  
zylicam@gmail.com

### Specialty section:

This article was submitted to  
Cardiovascular Imaging,  
a section of the journal  
Frontiers in Cardiovascular Medicine

**Received:** 21 February 2021

**Accepted:** 10 May 2021

**Published:** 16 June 2021

### Citation:

Yin Y, He C, Xu B and Li Z (2021)  
Coronary Plaque Characterization  
From Optical Coherence Tomography  
Imaging With a Two-Pathway  
Cascade Convolutional Neural  
Network Architecture.  
Front. Cardiovasc. Med. 8:670502.  
doi: 10.3389/fcvm.2021.670502

**Background:** The morphological structure and tissue composition of a coronary atherosclerotic plaque determine its stability, which can be assessed by intravascular optical coherence tomography (OCT) imaging. However, plaque characterization relies on the interpretation of large datasets by well-trained observers. This study aims to develop a convolutional neural network (CNN) method to automatically extract tissue features from OCT images to characterize the main components of a coronary atherosclerotic plaque (fibrous, lipid, and calcification). The method is based on a novel CNN architecture called TwopathCNN, which is utilized in a cascaded structure. According to the evaluation, this proposed method is effective and robust in the characterization of coronary plaque composition from *in vivo* OCT imaging. On average, the method achieves 0.86 in F1-score and 0.88 in accuracy. The TwopathCNN architecture and cascaded structure show significant improvement in performance ( $p < 0.05$ ). CNN with cascaded structure can greatly improve the performance of characterization compared to the conventional CNN methods and machine learning methods. This method has a higher efficiency, which may be proven to be a promising diagnostic tool in the detection of coronary plaques.

**Keywords:** optical coherence tomography, convolutional neural network, plaque characterization, cascaded structure, two-pathway architecture

## INTRODUCTION

Cardiovascular disease remains to be the leading cause of morbidity and mortality globally. Acute events, such as heart attack and stroke, are usually triggered by the development of plaque rupture/erosion and subsequent thrombus formation. It has been shown that plaque components and morphology are the main factors in the determination of plaque stability (1–3). A coronary plaque normally consists of three different tissue types: fibrous, calcified, and lipid-rich, with mixed forms appearing in many cases. Plaques that have a large lipid core and a thin fibrous cap are more prone to rupture, whereas plaques containing calcification tend to be stable (4). Therefore, to prevent the acute cardiovascular events, it is crucial to develop strategies to characterize plaque morphology and components so that the vulnerable plaques can be identified at an early stage to reduce the risk of rupture.

Based on the analysis of the reflected light, optical coherence tomography (OCT) has a high resolution. Its axial resolution can reach to 12–18  $\mu\text{m}$  (5), which is a promising diagnostic technique in interventional cardiovascular imaging. High-resolution cross-sectional images of coronary artery can be seen *via* OCT, which allow us to observe the plaque components and morphology directly and distinguish different plaque types. The expert consensus document indicates that the fibrous tissue appears as uniform high-signal regions, the calcified tissue appears as uneven, low-signal regions with clear boundaries, whereas the lipid tissue is the low-signal region with blurred boundaries in OCT (6). In current clinical application, vascular lumen segmentation and characterization of plaque components are performed manually, which is laborious and time-consuming. Moreover, the accuracy is dependent on the experience of the observers. Therefore, developing a fast, accurate, and reliable method for an automatic characterization of atherosclerotic plaque composition is crucial, which can reduce the time and cost of analyzing data and avoid human interference errors, and thus can improve risk assessment of patients with coronary artery diseases.

Levitz et al. (7) published a study demonstrating that atherosclerotic plaque characterization by OCT could be done by measuring the back-scattering and attenuation coefficients (8), which enhanced the differentiations between the fibrous, calcified, and lipid tissues. The preliminary data indicated that differences in scattering properties may exist between the normal and the atherosclerotic plaques and that optical scattering properties provided by OCT can contribute to plaque characterization. After that, more optical parameters were introduced to characterize components of plaques in OCT studies (9–13). Although optical properties are good representations of various intracoronary tissues, detailed information regarding plaque morphology is required for better recognition of various plaque types. To utilize more texture features, machine learning methods have been applied to plaque characterization. Wang et al. (14) proposed a method for semiautomatic segmentation of calcified plaques with the morphology operation. Ughi et al. (15) combined the attenuation coefficients and Gray-Level Co-occurrence Matrix (GLCM) (16) to obtain texture features and utilized the random forest (17) model to classify the test set and compared the result with the manually segmented ground truth. The overall classification accuracy was 81.5% with a feasibility of 76.5%. The accuracy for each type was 89.5% for fibrotic tissue, 72.1% for calcium, and 79.5% for lipid-rich tissue. Athanasiou et al. (18) used K-means (19) to divide calcified tissue and applied GLCM and local binary pattern (20) to characterize texture features. Rico-Jimenez et al. (21) applied least square optimization strategy to estimate the depth of plaque. Our group (22) also presented the characterization of atherosclerotic plaque components based on the machine learning methods. Although these methods have shown a better performance compared to the methods based on attenuation coefficients, the features extracted by algorithm cannot completely describe all the global morphological information, thus leading to classification noise.

Convolutional neural networks (CNNs) (23) have shown remarkable successes in image processing tasks, such as image classification, facial identification (24), and object detection (25). In terms of OCT, Gessert et al. (26) developed a CNN model to detect images with calcified plaques. Kolluru et al. (27) implemented a CNN model consisting of two convolutional and max-pooling layers and applied a fully connected conditional random field (CRF) as a post-processing step to improve classification sensitivity. Yong et al. (28) proposed a linear-regression CNN model to segment the lumen in OCT images automatically. Lee et al. (29) implemented SegNet and Deeplab v3+ for plaque characterization in terms of pixel-wise and A-line-based classifications. Our group also presented calcium classification on the OCT pullback with 3D deep neural networks (30). Gharaibeh et al. (31) used deep learning and transfer learning methods to analyze the calcification components in OCT images. Li et al. (32) proposed a deep residual U-Net network for the segmentation of vulnerable plaque components in coronary OCT. Athanasiou et al. (33) proposed a patch-based CNN network to characterize plaque components. Previous studies on deep learning-based plaque characterization mainly fall into two categories: one is utilizing segmentation networks to learn and evaluate OCT images directly, such as FCN, SegNet, etc. However, this method requires a huge training dataset. The other way is using a patch-based CNN network to achieve the segmentation goal by classifying small patches, but some position and boundary information may be lost, and the evaluating efficiency is low because the networks need to characterize each pixel by classifying a patch. Most patch-based CNN models realize segmentation function by predicting the  $M \times M$  patch centered on that pixel; the segmentation results are determined by the classification category of the  $M \times M$  patch. Therefore, when evaluating the whole OCT image by normal patch-based CNN models, the experts need to make patches for all the pixels in the region of interest and classify them with CNN models. This means the CNN models have to classify thousands of patches when evaluating an OCT image.

Therefore, in this study, we designed a patch-based CNN model to analyze the OCT dataset of coronary arteries to characterize three plaque components automatically: fibrotic, calcified, and lipid-rich tissues. Three trained observers provided segmentations for each OCT image, and the selected regions were divided into patches with different sizes. All the patches were put into a CNN model to obtain an automated classifier of atherosclerotic plaques. For testing, the performance of the model was subsequently evaluated against the testing set manually segmented by a trained OCT observer, which includes 10 samples for three plaque components, respectively, 30 in total. To convince the reliability of testing results, these samples were selected from 10 different pullbacks and 3 different patients, which were not used in the training period. This is the first study that Two-Pathway Cascade CNN Architecture is applied into the field of coronary plaque characterization from OCT imaging. Compared with other patch-based CNN models, the method achieved a more accurate result than the methods published in the literature especially for the calcified plaques. The time complexity is also significantly reduced by using

convolutional output layer, which makes the method more clinically realistic. Several novel strategies on CNN architectures and training process were utilized to avoid the weakness of patch-based CNN model, and demonstrated the effectiveness with statistical methods.

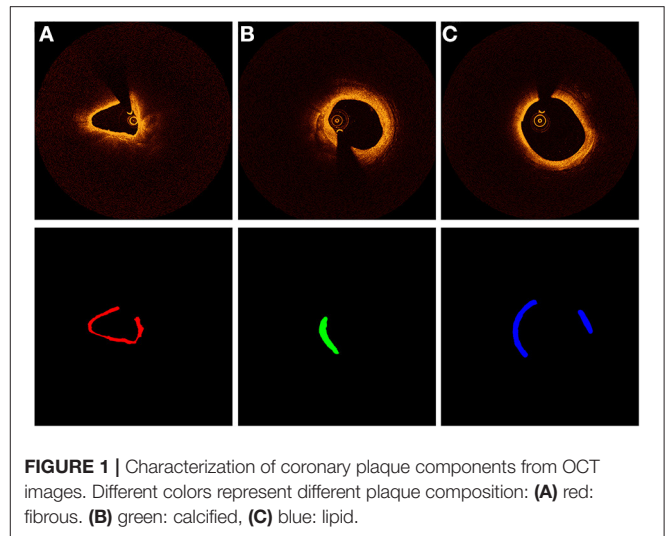
## METHODS

### Data Collection

All the coronary OCT imaging data were collected from Nanjing Drum Tower Hospital. The whole imaging procedure was performed using the C7XR OCT imaging system (St. Jude Medical, Minneapolis, MN, USA) connected to a C7 Dragonfly intravascular imaging catheter (St. Jude Medical). During imaging, the catheter was inserted into the patient's coronary artery. By rotating the probe inside the vessel and pulling it backwards, the images inside the vessel were acquired. The light source emits an optical signal and the detector receives the interference signal produced by the signal light backscattered by the vessel and the reference light reflected by the reflector, then the system further processes the signal to obtain a scan line signal (A-line). The C7XR OCT has 504 scan A-lines per blood vessel section. In one scan, the C7XR OCT can obtain a 271-frame blood vessel image in polar representation with the resolution of  $976 \times 504$  pixels in gray scale. For easier observation, all the images were transformed into Cartesian space with  $1,024 \times 1,024$  pixels, and the B-scan is color-mapped to RGB using a colormap for visualization contrast.

For ground-truth annotation, three experienced observers firstly excluded images with stenting cases. To avoid plaques with similar patterns containing in the dataset, we extracted one frame in each adjacent three frames for dataset making. All the selected images were segmented manually. To ensure the reliability of dataset, all images were separated randomly to three observers for manual segmentation. After the segmentation, they exchanged the segmented OCT images and double-checked the results. When disagreement occurred, the final segmentation results were confirmed by an interventional cardiologist. **Figure 1** shows the template of an example of the manually segmented plaque compositions.

In total, the dataset included 2,000 images from 31 patients, and 2,000 regions of interest were extracted from these images. Fibrous tissue and lipid tissue comprised 40% of the whole dataset, respectively, and calcified tissue occupied 20%. One sixth of the dataset were used to validate and the rest were used to train. The regions of interest in the training images (corresponding to color regions in the template) were sliced into the size of  $65 \times 65 \times 3$  and  $33 \times 33 \times 3$  pixels. Different patches (e.g.,  $17 \times 17 \times 3$  or  $83 \times 83 \times 3$ ) were also tested, but  $65 \times 65 \times 3$  and  $33 \times 33 \times 3$  achieved the best performance. A smaller size cannot obtain location features in the training period while a larger size introduces more noise that may affect the training results. All the patches were labeled according to their corresponding color in the template. Here, we used patches with RGB channels for CNN training to guarantee that the images for prediction are in the same condition with those for manual analysis.



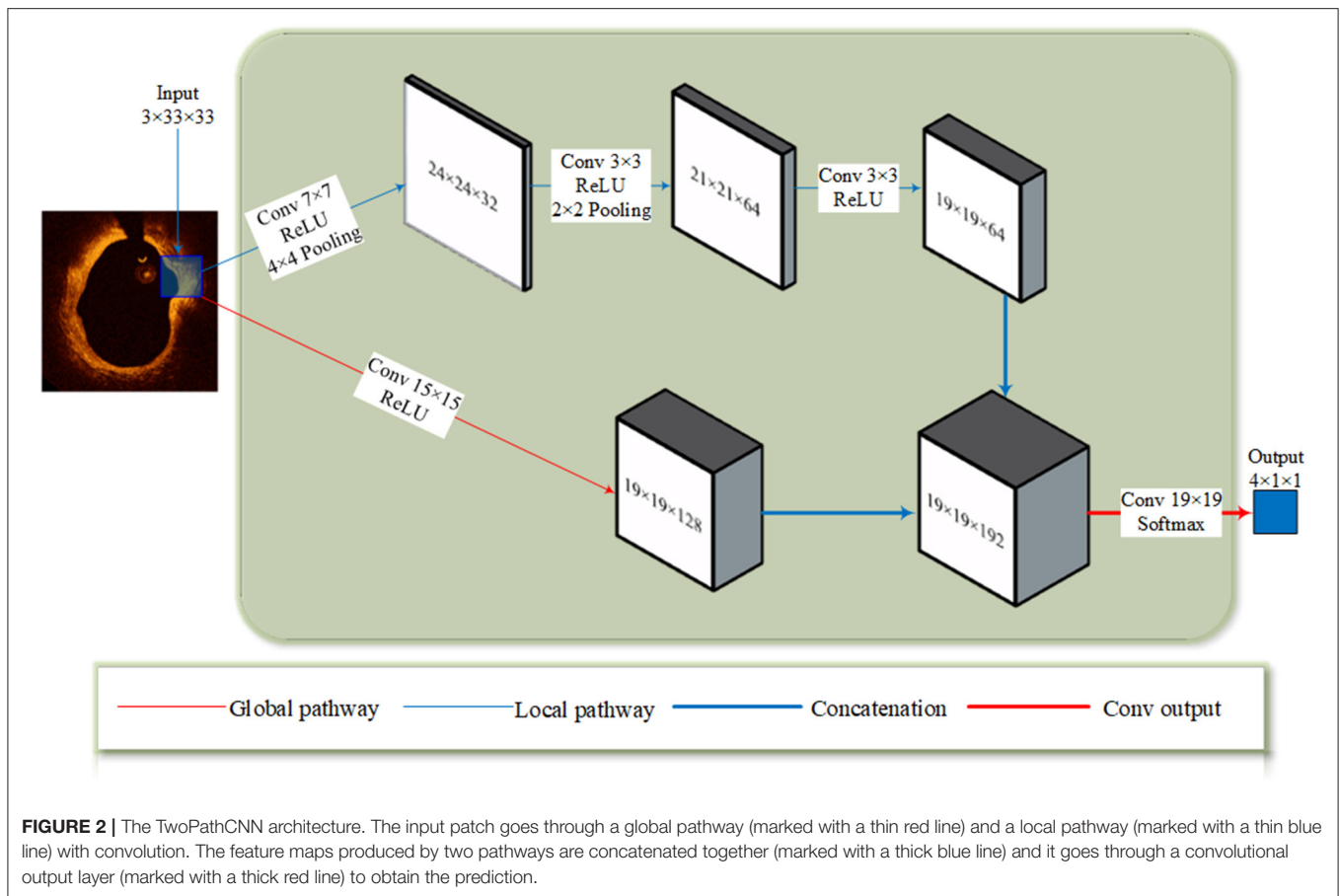
### Data Augmentation

From the number of patches belonging to each label, we discovered that patches belonging to calcified tissue were the least. Seriously, data imbalance can affect the training result of the CNN model. To deal with the problem of imbalanced data, the often-used approach is to modify loss function or use data augmentation. In our cases, modifying loss function did not perform efficiently because there was a considerable gap in the amount of data in each category. Thus, we used data augmentation to the calcified patches to enable the quantity of patches in each label to be strictly equal.

A certain number of calcified patches were randomly chosen for data augmentation. We applied flipping along the  $x$  and  $y$  directions on the chosen patches. In total, the final patches dataset consisted of 120,000 patches with equal quantity in each label. We split off an independent validation set of 20,000 patches to detect the learning condition of the CNN model during the training process and the others were considered as the training set. To further reduce the impact of imbalanced dataset, two-phased training strategy was utilized in the training process, which is introduced in section Segmentation Results.

### CNN Approach

Like most CNN-based segmentation models, we replaced the calculation of each pixel by predicting the  $M \times M$  patch centered on that pixel. Thus, the input of CNN model was  $M \times M$  patches with different types. By applying a series of convolutional layers, CNN had the ability to extract complex features. More detailed information of patches can be learned by treating the output feature maps of the previous convolutional layer as the input of the subsequent. To obtain non-linear features from the input, ReLU (34) was utilized as an activation function. A normalization process was also added after the activation function to bring mean value and variance close to zero and one, respectively. The normalization process can be summarized by



the following formula:

$$x_n = \frac{x - \mu}{\sqrt{\sigma^2 + t}}$$

Here,  $x_n$  and  $x$  represent the feature maps after and before the normalization process, respectively, and  $\mu$  and  $\sigma^2$  are the mean and variance of the feature maps.  $t$  is a small constant to avoid division by zero.

Due to the lack of training examples in our study, the model was vulnerable to over-fitting. Accordingly, we used Dropout (35) regularization method to overcome it. When computing the hidden layer of CNN, dropout masks the input or hidden unit with a certain probability. In this case, the prediction of CNN did not only rely on a small part of weights in the entire network.

In conventional CNN models, for classification purposes, the output layer is typically fully connected, which is inefficient in segmentation. To perform a prediction of the segmentation labels, we replaced the fully connected layer by a convolutional output layer. The number of kernels in the layer equaled to the number of labels (four in our case, the background was also included). The output of kernels was considered as the final estimation of CNN model, which can be normalized by SoftMax

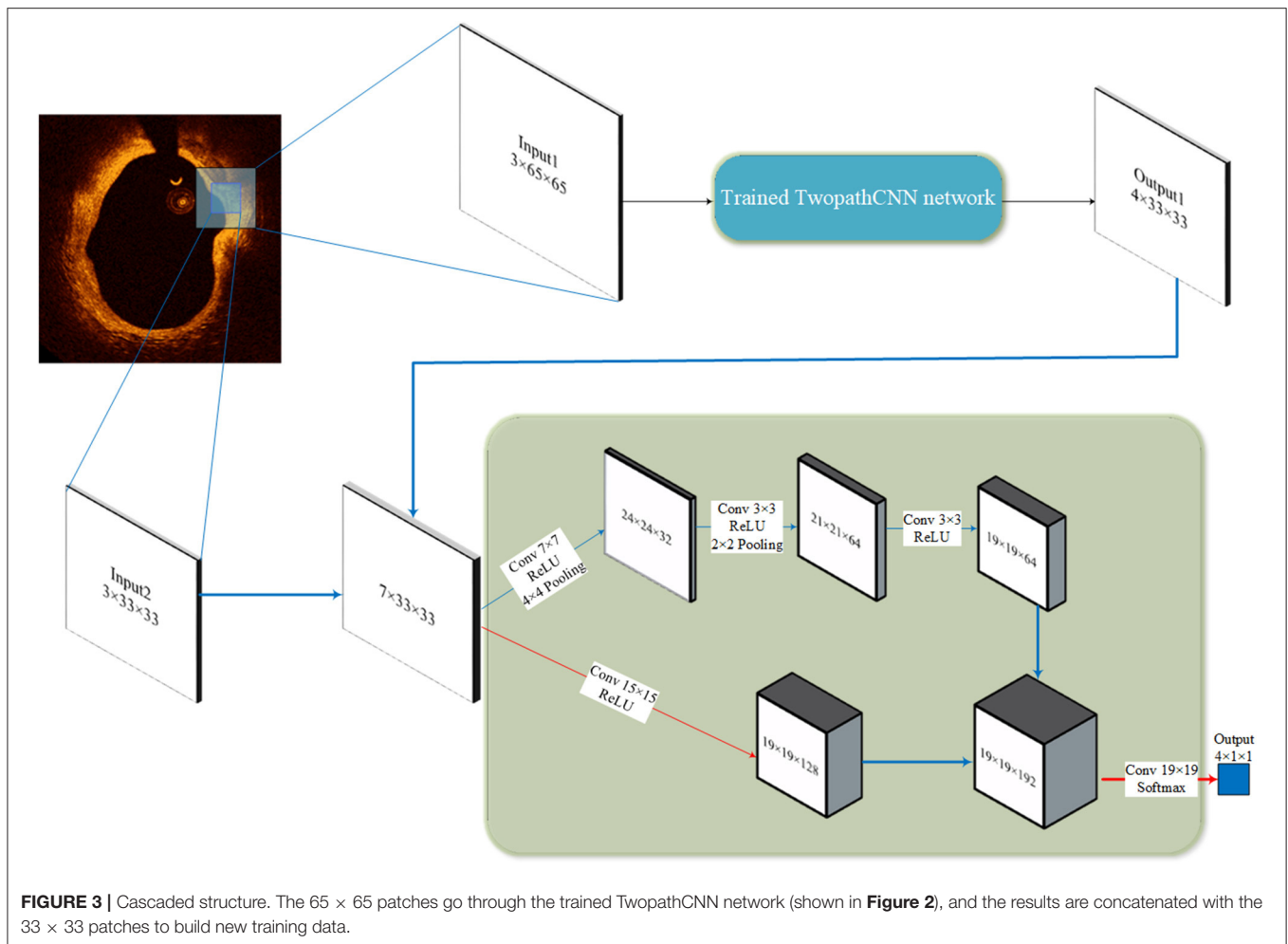
function as follows:

$$\text{Softmax}(a) = \frac{\exp(a)}{\sum_i \exp(a_i)}$$

Here,  $a$  is the output vector of convolutional output layer. Each element of  $w = \text{SoftMax}(a)$  is limited to the range of (0, 1) and the sum equals to 1. Thus,  $w$  is regarded as the probability distribution of all labels of the patches. The CNN approach performs the segmentation by assigning the label with largest probability to each pixel. By placing the convolutional output layer at the end, when doing segmentation tasks, the speed of testing was 40 times faster than the conventional CNN model. More details are shown in section Implementation Details.

### Two-Pathway CNN Architecture

To find the most appropriate CNN architecture for the segmentation task, we constructed a new CNN model instead of applying transfer learning because the state-of-the-art CNN architectures typically use large images as their training set (about  $270 \times 270$  in most cases). Two-pathway CNN (TwoPathCNN) (36) was introduced to extract features. Different from the straight structure of conventional CNN networks, TwoPathCNN is made of two streams: a pathway with several convolutional layers in smaller receptive fields (local pathway) and the other with a convolutional layer in a larger receptive field (global pathway). After the convolution operation, a concatenation layer



combined the feature maps from two streams together. In this architecture, the prediction of labels can be influenced by two aspects: the visual details of the region and the general features.

The full architecture is illustrated in **Figure 2**. To satisfy the concatenation, the size of the top hidden layers of both pathways must be the same. The local pathway contains three convolutional layers with  $7 \times 7$  and  $3 \times 3$  kernels. The global pathway contains one convolutional with  $15 \times 15$  kernels. In this network, more feature maps were extracted from the global pathway.

### Cascaded Structure

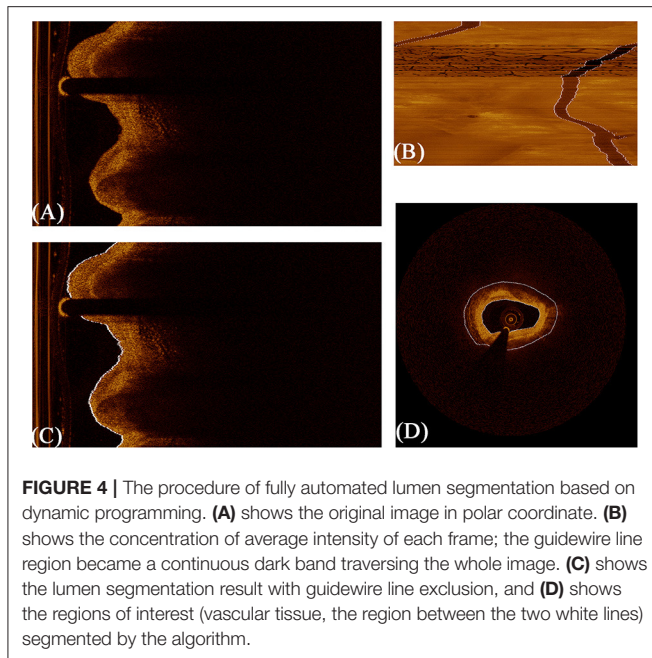
Single CNN structure predicts segmentation label according to each individual patch, lacking the information of position. In response to this shortcoming, the segmentation method in literatures often define a conditional random field (37, 38) to simulate the relationship between each label from a spatial perspective, and the final prediction is influenced by both the original prediction and segmentation labels in the vicinity.

In our study, the cascaded structure was applied to make a connection between CNN networks with different input sizes. As **Figure 3** illustrated, the output of the first CNN is concatenated

with the patches in a smaller size, the output and the patches both comprise the input to the second CNN. Thus, the input of the second CNN is the feature maps with seven channels. This kind of structure can overcome the ignorance of the relative position by regarding the probability distribution of labels of surrounding pixels as a prerequisite to the prediction. Compared to conditional random field, this method takes the advantage in the consuming of computational resource, which is important when dealing with large amounts of OCT data in clinical diagnosis.

### Lumen and Guidewire Line Segmentation

Due to the limitation of the tissue penetration of OCT (about 1–1.3 mm) (39), OCT may not be able to correctly display the region outside the region far away from the lumen border; therefore, the region of interest (vascular tissue) of an OCT image was considered as the domain that lumen expands 1 mm outward. Lumen and guidewire line segmentation can greatly improve the performance of characterization by implying an additional known condition during the analysis. This process is necessary because even experienced interventional cardiologists cannot always clarify the difference between the outer border of



lipid tissue and the background. If the region of interest is not strictly defined, the outer border of the segmentation result will be uneven.

The lumen and guidewire line segmentation was based on dynamic programming, which has been proven effective and robust in Wang et al.'s research (40). The lumen boundary was detected according to the gradient in original polar image and the contour with the highest pixel value difference across the vertical axis. Guidewire always shows a long dark shallow in OCT images. Therefore, in terms of guidewire line region, the average intensity of each A-line is significantly lower. When concentrating the average intensity of all images in the whole pullback together, due to the continuity of pullbacks, the guidewire appears as a dark strip throughout the vertical axis, as shown in **Figure 4B**. In this condition, using dynamic programming can directly extract the guidewire of all frames in the whole pullback. On the basis of resolution of intravascular C7-XR OCT (15  $\mu\text{m}$ ), a constant depth of 100 pixels was utilized in our study. **Figure 4** shows the whole segmentation process.

## Implementation Details

Our study is based on the deep learning toolbox in Matlab platform. Deep learning toolbox integrates multiple application and visualization tools to create and analyze deep learning architecture. It also supports the use of NVIDIA GPU, which can greatly accelerate the training process of CNN. Moreover, Matlab has excellent computing capability in matrix and is very effective in disposing data like images or feature maps.

The hyper-parameters of CNN architecture are shown in **Figures 2, 3**. The pooling layers uniformly applied max pooling with a stride of 1. As for optimizer selection, we chose RMSprop (41), which has been proven efficacious in many deep learning networks. The squared gradient decay factor was set to 0.9. These hyper-parameters and optimizer selection were determined

according to the cross-validation results. The chosen hyper-parameters and optimizer achieved the best performance on the validation sets.

For testing time, the specific structure of convolutional output layer mentioned in section CNN Approach can accelerate the proceeding by simplifying the computation. This was realized by feeding a full image instead of individual patches as the input. The trained models can automatically convert the image under test to all label probabilities of the entire image. This testing mode can be 40 times faster than extracting patches at each pixel and predicting them individually because convolution can better deal with large-sized input compared with multiple small-sized input data. Moreover, the extracting operation is spared. Finally, as post-processing, we removed small isolated regions in the predicted results, which may be classification noise. In total, the TwoPathCNN model can provide a segmentation for each image in 2 s with the NVIDIA 1080Ti card. When using cascaded structure for prediction, it takes <3 s on average.

Because our study was focused on the characterization of plaque composition, only the prediction on the region of vascular tissue was held. After going through the trained model, the region outside the border produced in section Implementation Details was set as background by default.

## EXPERIMENTS AND RESULTS

### Evaluation Standard

F1-score (42) is an evaluation standard of classification problems, which is the harmonic average of precision and recall. Many competitions about deep learning use it as the final evaluation method. F1-score divides the classification results of models into four types as follows:

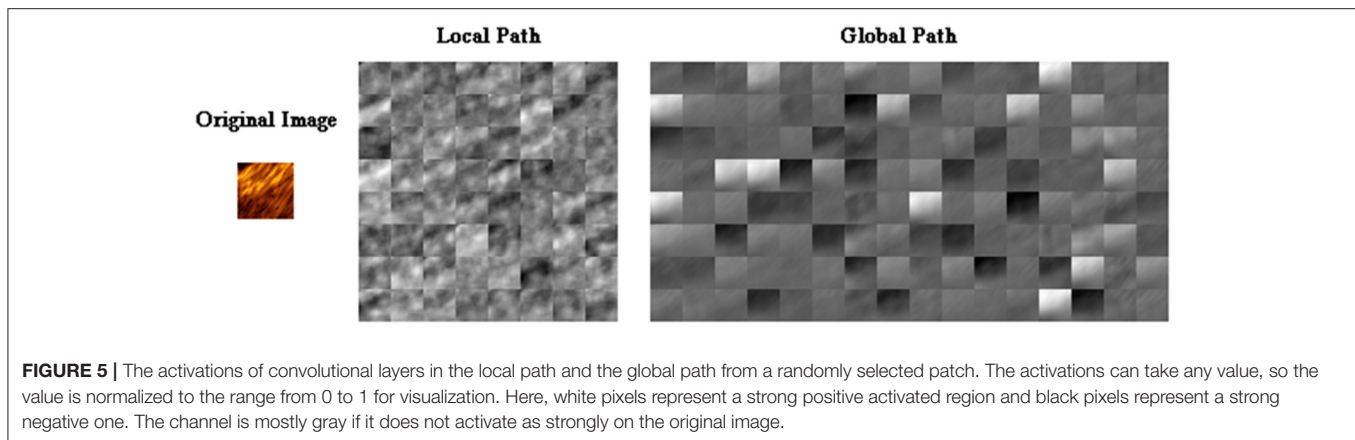
- True Positive (*TP*): Correctly predicting the target category.
- False Positive (*FP*): Incorrectly predicting other categories into the target category.
- True Negative (*TN*): Correctly predicting other categories.
- False Negative (*FN*): Incorrectly predicting the target category into other categories.

By counting the number of patches or pixels of these four types, the precision, recall, and accuracy of each category were calculated, which can be summarized as follows:

$$\begin{aligned} \text{precision} &= \frac{TP}{TP + FP} \\ \text{recall} &= \frac{TP}{TP + FN} \\ \text{accuracy} &= \frac{TP + TN}{TP + FP + TN + FN} \end{aligned}$$

The F1-score of each category was obtained from its precision and recall, and the general evaluation score was the average value of all the F1-scores. The details are shown in the following formula:

$$F1_k = \frac{2\text{precision}_k \text{recall}_k}{\text{precision}_k + \text{recall}_k}$$



$$F1\text{-score} = \frac{1}{n} \sum_k F1_k$$

Compared to the conventional accuracy evaluation standard, F1-score has a more comprehensive estimation of the CNN model. In addition, F1-score provides different indicators, which can be selectively adopted according to actual classification tasks.

### TwoPathCNN Architecture

To clearly illustrate the different features extracted from different paths of TwoPathCNN architecture, the activations of convolutional layers in each path is shown in **Figure 5** to observe which areas activate on the patches in the convolutional layers. Although there are explicit explanations on the advanced features learned by CNN, different emphasis is indicated in these two paths. The global path mainly detects the edge information of plaques, whereas local path pays more attention on the localized texture features.

To study the effect of TwoPathCNN architecture compared to conventional linear CNN networks, we regarded the CNN model consisting of only the local path or the global path as the control group. The results of these variations are listed in **Table 1**, which contains the performance of TwoPathCNN, LocalPath, and GlobalPath in terms of precision and recall value of each class, the average F1-score, and the overall accuracy. According to the result of Wilcoxon signed-rank test, both F1-score and accuracy have a significant improvement compared to the control group. Moreover, the confidence interval of precision and recall on TwoPathCNN is evidently narrower than the other two CNN architectures. This indicates that TwoPathCNN is more stable and robust than the original CNN architecture because different OCT images can obtain similar performance on TwoPathCNN model. TwoPathCNN joins the local and global paths to co-adapt the features learnt from the training data, thus actualizing the promotion of capability and robustness of the CNN model.

### Cascaded Structure

To investigate the performance of cascaded structure, we applied the cascaded structure to TwoPathCNN architecture and compared it with the same architecture without using it. In order to maintain the consistency of the whole structure, when training

**TABLE 1 |** The confidence interval of performance on different CNN architectures, including precision, recall, F1-score, and accuracy.

Architecture	Class	Precision	Recall	F1-score	Accuracy
TwoPathCNN	Fibrous	0.96 ± 0.03	0.95 ± 0.03	0.86 ± 0.06*	0.88 ± 0.05*
	Calcified	0.68 ± 0.12	0.87 ± 0.06		
	Lipid	0.89 ± 0.13	0.81 ± 0.11		
LocalPath	Fibrous	0.94 ± 0.05	0.93 ± 0.05	0.81 ± 0.07	0.85 ± 0.06
	Calcified	0.63 ± 0.23	0.75 ± 0.15		
	Lipid	0.85 ± 0.12	0.77 ± 0.13		
GlobalPath	Fibrous	0.96 ± 0.04	0.87 ± 0.10	0.71 ± 0.09	0.77 ± 0.09
	Calcified	0.43 ± 0.18	0.67 ± 0.13		
	Lipid	0.80 ± 0.23	0.74 ± 0.16		

With TwoPathCNN, both F1-score and accuracy were significantly improved compared to GlobalPath and LocalPath (\* $p < 0.05$ ). For statistical analysis, the Wilcoxon signed-rank test was performed.

the newly built dataset by concatenation, we selected the same architecture as the network that generated output comprised in the concatenated dataset (Output 1 in **Figure 4**). At the same time, we also investigated the impact of two-phase training method in our dataset. Two-phase training method retrains the CNN network with the unbalanced training set and keeps all the layers fixed except the output layer in the second phase. Therefore, we divided the testing performance into four groups: cascaded structure with two-phase training, cascaded structure with original training, non-cascaded structure with two-phase training, and non-cascaded structure with original training. **Table 2** provides the corresponding performance results.

As shown in **Table 2**, we found that cascaded structure with two-phase training achieved best performance, which was shown by the Wilcoxon signed-rank test ( $p < 0.05$ ). Regardless of whether two-phase training was utilized or not, cascaded structure significantly improved the performance over the normal structure, with a statistically significant difference for all metrics ( $p < 0.05$ ). However, in terms of two-phase training, two-phase training with cascaded structure had shown significant improvements on F1-score and accuracy compared to one-phase training with cascaded structure ( $p < 0.05$ ). Without

**TABLE 2** | The confidence interval of performance with or without training strategies (cascaded structure and two-phase training).

Structure/training	Class	Precision	Recall	F1-score	Accuracy
Cascaded and two-phase training	Fibrous	0.96 ± 0.03	0.95 ± 0.03	0.86 ± 0.06*	0.88 ± 0.05*
	Calcified	0.68 ± 0.12	0.87 ± 0.06		
	Lipid	0.89 ± 0.13	0.81 ± 0.11		
Cascaded and one-phase training	Fibrous	0.97 ± 0.02	0.90 ± 0.06	0.82 ± 0.07	0.86 ± 0.08
	Calcified	0.58 ± 0.15	0.94 ± 0.03		
	Lipid	0.86 ± 0.14	0.78 ± 0.15		
Non-cascaded and two-phase training	Fibrous	0.95 ± 0.04	0.89 ± 0.04	0.75 ± 0.06	0.80 ± 0.06
	Calcified	0.51 ± 0.19	0.64 ± 0.08		
	Lipid	0.77 ± 0.17	0.81 ± 0.09		
Non-cascaded and one-phase training	Fibrous	0.93 ± 0.09	0.87 ± 0.08	0.74 ± 0.07	0.80 ± 0.07
	Calcified	0.50 ± 0.18	0.69 ± 0.14		
	Lipid	0.76 ± 0.18	0.79 ± 0.11		

With cascaded structure and two-phase training, both F1-score and accuracy had the best performance (\* $p < 0.05$ ). Cascaded structure significantly improved the performance of the segmentation ( $p < 0.05$ ). For statistical analysis, the Wilcoxon signed-rank-test was performed.

cascaded structure, the overall metrics of two-phase training were slightly increased but statistically insignificant ( $0.05 < p < 0.25$ ). Therefore, we consider that two-phase training may bring positive effects on the segmentation results, but is not as effective as cascaded structure.

## Segmentation Results

From the confidence interval of precision and recall value in each class shown in **Tables 1, 2**, we found that in all cases, CNN models had shown best performance in fibrous tissues, whereas the calcified tissue had a lower precision and recall coefficients compared to other tissues. From small patches, many texture features were learnt by CNN models, but location features and edge features were ignored at the same time. Fibrous tissues usually express high-intensity region in OCT images, which is significantly different from other tissues. Experts usually identify calcified tissues according to clear boundaries, but in terms of texture, the calcified tissues represent similar features with other tissue types in many cases. Although many strategies like cascaded structure were taken to add more location information into CNN models, some features were still lost in small patches. **Figures 6, 7** show the prediction of different types of plaques with different networks. Compared with the results of TwoPathCNN, the cascaded structure makes the boundary of characterized plaque smoother, which is more reasonable and accurate. Whereas, the LocalPath network mis-characterized the plaque in some cases.

**Table 3** shows a comparison between our CNN architectures and other published state-of-the-art methods in plaque characterization. We chose a recently published method using deep learning (29) and a machine learning method established by our group (22) for comparison. Compared with the deep learning method, although our method had a similar performance in calcified and lipid tissues, our method is able to achieve a more complicated task by segmenting the fibrous tissues. In terms of comparison with the machine learning method, our method achieved a better performance in all cases. Lambros et al. (33) used normal patch-based CNN to categorize four

**TABLE 3** | Comparison of our method with other state-of-the-art segmentation methods published previously in terms of sensitivity and specificity (— means the author did not segment the corresponding plaques).

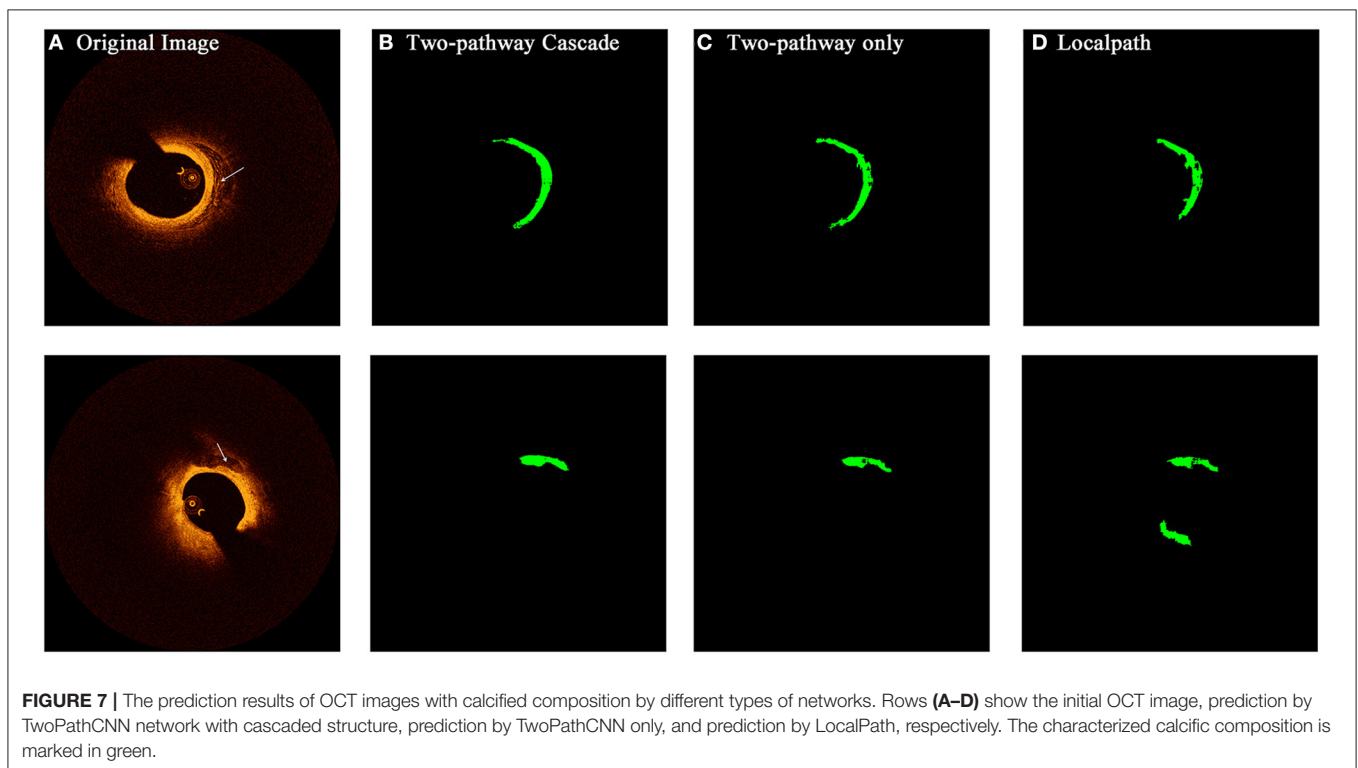
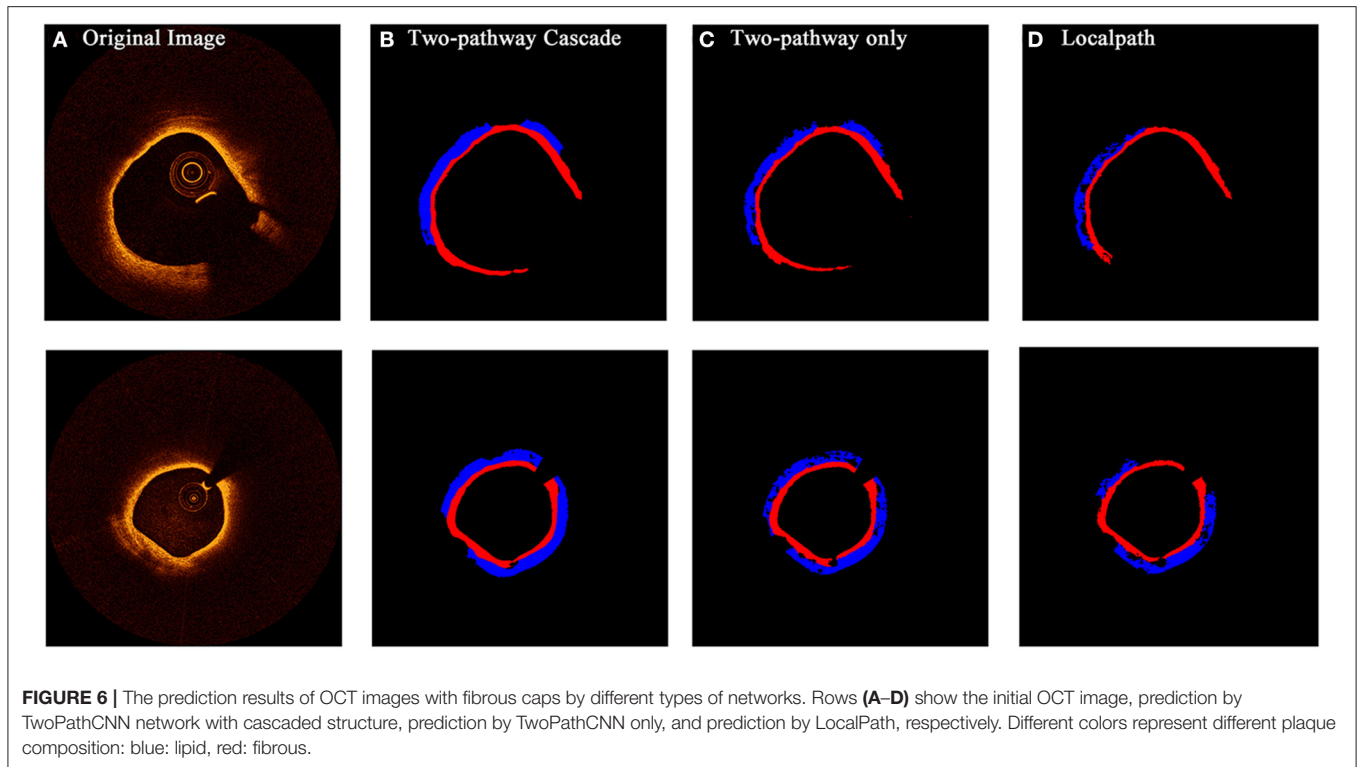
Methods		Fibrous	Calcified	Lipid
Our method	Sensitivity	0.95 ± 0.03	0.87 ± 0.06	0.81 ± 0.11
	Specificity	0.97 ± 0.02	0.92 ± 0.06	0.96 ± 0.03
Deep learning method	Sensitivity	—	0.85 ± 0.07	0.87 ± 0.07
	Specificity	—	0.94 ± 0.02	0.90 ± 0.05
Machine learning method	Sensitivity	0.81	0.65	0.80
	Specificity	0.91	0.91	0.88
Other patch-based method	Sensitivity	0.89	0.34	0.88
	Specificity	0.98	0.97	0.98

plaques and achieved good results in overall accuracy. However, the sensitivity is relatively low especially for calcified plaques. According to the consensus, calcified plaque contains IVOCT evidence of fibrous tissue, along with calcium that appears as a signal-poor or heterogeneous region with a sharply delineated border (43). Therefore, it is hard to characterize calcified plaques by only using texture features; thus, the normal patch-based CNN models have a poor performance. In this study, by applying cascaded structure, more information of location and border is learnt by CNN models, which can greatly improve the characterizing performance on calcified plaques. In addition, by using convolutional output layer, the computation time can be reduced.

## DISCUSSION

Currently, plaque characterization from OCT imaging requires trained observers, which is impractical when analyzing hundreds of pullbacks in the clinical setting. Therefore, developing an automatic algorithm for plaque characterization can effectively assist clinician in analyzing OCT images and facilitate the clinical diagnosis. We proposed a CNN-based algorithm for





the automated characterization of coronary plaque compositions from OCT imaging and reformed the conventional CNN architecture according to the features of plaque composition. The results show that our CNN-based method can effectively segment

different plaque compositions using trained CNN models. For each OCT image, all pixels in the region of vascular tissue are classified to one of the tissue types of atherosclerotic plaque components.

Compared to the accuracy of conventional methods using machine learning (around 81%), most architectures we mentioned in this paper can achieve this accuracy under the premise of considering a more complicated situation that an additional category besides fibrous, calcified, and lipid is taken into consideration. Methods based on texture and attenuation coefficients analyze each individual pixel with relevant intensity and texture features. This often leads to scattered noise in final prediction when the texture of tissue is uneven. Thanks to the cascaded structure, this problem has been greatly improved by providing more information on local tissue.

The algorithm was programmed in Matlab R2019b and tested in the computer with the following configuration: Windows 64-bits, CPU Intel i7 with 6 cores at 3.4 GHz, 32 GB RAM, and two NVIDIA 1080Ti cards. In machine learning methods, computational time of analysis is around 80 s per image. However, it takes only 3 s on average when using cascaded structure for prediction, which includes the lumen segmentation operation and plaque characterization. CNN methods and GPU equipment greatly accelerate the characterization. This gap in efficiency is huge especially when dealing with large datasets consisting of hundreds of OCT images.

When processing the dataset consisting of OCT imaging data, we had to exclude some images, which were hard to identify their plaque composition to prevent controversial ground truth from misleading the CNN models. This may result in error when testing on similar OCT images. Therefore, referring to segmentation results of neighboring images in the pullback and using interpolation to adjust the prediction is sometimes necessary.

OCT images *in vivo* do not have their corresponding histology slices, so the ground truth datasets were produced by trained observers *via* manual analysis. Although a series of actions were taken to reduce bias, inevitably, manual analysis of OCT images is prone to emerge error resulting in a weaker ground truth. This may affect the testing results and mislead the CNN network during the training process.

To deal with these limitations, a platform providing large amounts of OCT images and agreed validation standard can greatly improve the performance of characterization. The training sets of plaque composition require trained observers to segment substantial images manually, which is time- and labor-consuming. In addition, more observers may be required to reduce possible errors due to inter-operator variability. Lack of dataset has become a huge barrier in the promotion of automatic characterization of plaque composition. For example, accumulating evidence has shown that plaque erosion and calcified nodule are the underlying pathological causes for acute coronary events. However, due to the limitation of dataset, only a small amount of OCT images indicates the features of plaque erosion and calcified nodules, which cannot meet the requirement of CNN training. Increasing classification categories also affect the performance of CNN models especially when the training dataset is small. Given that our computational resource is limited, we selected a simple structure of CNN with only 20 layers and the input size of  $33 \times 33 \times 8$ . Under the precondition

of enough memory, more complex network structures can be developed to improve the performance of characterization.

In addition, although conventional OCT can provide morphological detail because of the high resolution, it also has some limitations in diagnosis. Lipid plaque by OCT is a signal-poor region with poorly delineated borders. However, the signal-poor region may represent other tissues like macrophage or intraluminal debris. Therefore, conventional OCT cannot provide a complete description of coronary artery, which may be one of the reasons why current methods to characterize plaque components has been reported with limited success. The next-generation OCT (44–46) can detect signal like unique human coronary autofluorescence signature during the imaging procedure, which may provide complementary information to that obtained by conventional OCT. In further research, auto-characterizing algorithms can be combined with these features provided by the next-generation OCT, which can improve the capacity of plaque characterization and risk assessment.

## CONCLUSION

We developed a novel automated algorithm based on CNN to characterize coronary atherosclerotic plaque composition from *in vivo* OCT imaging. F1-score and accuracy index were used to evaluate its performance. The experimental results show that CNN with cascaded structure can greatly improve the performance of characterization compared to the conventional machine learning methods. This method has a higher efficiency, which may be proven to be a promising practical method for future clinical practice. With further development and validation, it may be developed to be an automatic OCT-based diagnostic tool for coronary plaque characterization.

## DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

## AUTHOR CONTRIBUTIONS

YY and CH contributed to the design of the study and performed the statistical analysis. BX provided the clinical OCT data and instructed the manual segmentation. YY wrote the first draft of the manuscript. ZL was in charge of the whole research and polished the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

## FUNDING

This research was funded by the National Natural Science Foundation of China (NSFC) (Grant Nos. 11972118, 11772093, and 61821002) and Australian Research Council (DP200103492 and DP200101970).

## REFERENCES

- Shah PK. Mechanisms of plaque vulnerability and rupture. *J Am Coll Cardiol.* (2003) 41:15–22. doi: 10.1016/S0735-1097(02)02834-6
- Li Z, Howarth SPS, Tang T, Gillard JH. How critical is fibrous cap thickness to carotid plaque stability? A flow-plaque interaction model. *Stroke.* (2006) 37:1195–9. doi: 10.1161/01.STR.0000217331.61083.3b
- Li Z, Howarth SPS, Trivedi RA, Ukingim JM, Graves MJ, Brown AP, et al. Stress analysis of carotid plaque rupture based on *in vivo* high resolution MRI. *J Biomech.* (2006) 39:2611–22. doi: 10.1016/j.jbiomech.2005.08.022
- Arbab-Zadeh A, Fuster V. The myth of the “vulnerable plaque”: transitioning from a focus on individual lesions to atherosclerotic disease burden for coronary artery disease risk assessment. *J Am Coll Cardiol.* (2015) 65:846–55. doi: 10.1016/j.jacc.2014.11.041
- Bezerra HG, Costa MA, Guagliumi G, Rollins AM, Simon DI. Intracoronary optical coherence tomography: a comprehensive review: clinical and research applications. *JACC Cardiovasc Interv.* (2009) 2:1035–46. doi: 10.1016/j.jcin.2009.06.019
- Yabushita H, Bouma BE, Houser SL, Aretz HT, Jang I, Schlenker KH, et al. Characterization of human atherosclerosis by optical coherence tomography. *Circulation.* (2002) 106:1640–5. doi: 10.1161/01.CIR.0000029927.92825.F6
- Levitz D, Thrane L, Frosz MH, Andersen PE, Andersen CB, Valanciunaite J, et al. Determination of optical scattering properties of highly-scattering media in optical coherence tomography images. *Optics Exp.* (2004) 12:249–59. doi: 10.1364/OPEX.12.000249
- Thrane L, Yura HT, Andersen PE. Analysis of optical coherence tomography systems based on the extended Huygens-Fresnel principle. *JOSA A.* (2000) 17:484–90. doi: 10.1364/JOSAA.17.000484
- van der Meer FJ, Faber DJ, Perrée J, Pasterkamp G, Sasso DB, van Leeuwen, TG. Quantitative optical coherence tomography of arterial wall components. *Lasers Med Sci.* (2005) 20:45–51. doi: 10.1007/s10103-005-0336-z
- van der Meer FJ, Faber DJ, Sasso DB, Aalders MC, Pasterkamp G, van Leeuwen, T. G. Localized measurement of optical attenuation coefficients of atherosclerotic plaque constituents by quantitative optical coherence tomography. *IEEE Trans Med Imag.* (2005) 24:1369–76. doi: 10.1109/TMI.2005.854297
- Xu C, Schmitt JM, Carlier SG, Virmani R. Characterization of atherosclerosis plaques by measuring both backscattering and attenuation coefficients in optical coherence tomography. *J Biomed Optics.* (2008) 13:034003. doi: 10.1117/1.2927464
- Popescu DP, Fluieraru C, Mao Y, Chang S, Sowa MG. Signal attenuation and box-counting fractal analysis of optical coherence tomography images of arterial tissue. *Biomed Optics Exp.* (2010) 1:268–277. doi: 10.1364/BOE.1.000268
- Van Soest G, Goderie TP, Regar E, Koljenovic S, van Leenders AGJ, Gonzalo N, et al. Atherosclerotic tissue characterization *in vivo* by optical coherence tomography attenuation imaging. *J Biomed Optics.* (2010) 15:011105. doi: 10.1117/1.3280271
- Wang Z, Kyono H, Bezerra HG, Wang H, Garghesha M, Alraies C, et al. Semiautomatic segmentation and quantification of calcified plaques in intracoronary optical coherence tomography images. *J Biomed Optics.* (2010) 15:061711. doi: 10.1117/1.3506212
- Ughi GJ, Adriaenssens T, Sinnaeve P, Desmet W, D’hooge J. Automated tissue characterization of *in vivo* atherosclerotic plaques by intravascular optical coherence tomography images. *Biomed Optics Exp.* (2013) 4:1014–30. doi: 10.1364/BOE.4.001014
- Chen C-H. *Handbook of Pattern Recognition and Computer Vision.* Singapore: World Scientific (2015).
- Breiman L. Random forests. *Mach Learn.* (2001) 45:5–32. doi: 10.1023/A:1010933404324
- Athanasiou LS, Bourantas CV, Rigas G, Sakellarios AI, Exarchos TP, Siogkas PK, et al. Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *J Biomed Optics.* (2014) 19:026009. doi: 10.1117/1.JBO.19.2.026009
- Jain AK. Data clustering: 50 years beyond K-means. *Pattern Recognit Lett.* (2010) 31:651–66. doi: 10.1016/j.patrec.2009.09.011
- Nanni L, Lumini A, Brahmam S. Local binary patterns variants as texture descriptors for medical image analysis. *Artif Intell Med.* (2010) 49:117–25. doi: 10.1016/j.artmed.2010.02.006
- Rico-Jimenez JJ, Campos-Delgado DU, Villiger M, Otsuka K, Bouma BE, Jo JA. Automatic classification of atherosclerotic plaques imaged with intravascular OCT. *Biomed Optics Exp.* (2016) 7:4069–85. doi: 10.1364/BOE.7.004069
- He C, Li Z, Wang J, Huang Y, Li Z. Atherosclerotic plaque tissue characterization: an OCT-based machine learning algorithm with *ex vivo* validation. *Front Bioeng Biotechnol.* (2020) 8:749. doi: 10.3389/fbioe.2020.00749
- Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. *Adv Neural Inform Process Syst.* (2012) 25:1097–105. doi: 10.1145/3065386
- Sun Y, Wang X, Tang X. Deeply learned face representations are sparse, selective, and robust. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* Boston, MA (2015). p. 2892–900. doi: 10.1109/CVPR.2015.7298907
- Ren S, He K, Girshick R, Sun J. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Trans Pattern Anal Mach Intell.* (2016) 39:1137–49. doi: 10.1109/TPAMI.2016.2577031
- Gessert N, Lutz M, Heyder M, Latus S, Leistner DM, Abdelwahed YS, et al. Automatic plaque detection in IVOCT pullbacks using convolutional neural networks. *IEEE Trans Med Imaging.* (2018) 38:426–34. doi: 10.1109/TMI.2018.2865659
- Kolluru C, Prabhu D, Gharaibeh Y, Bezerra HG, Guagliumi G, Wilson DL. Deep neural networks for A-line-based plaque classification in coronary intravascular optical coherence tomography images. *J Med Imaging.* (2018) 5:044504. doi: 10.1117/1.JMI.5.4.044504
- Yong YL, Tan LK, McLaughlin RA, Chee KH, Liew YM. Linear-regression convolutional neural network for fully automated coronary lumen segmentation in intravascular optical coherence tomography. *J Biomed Optics.* (2017) 22:126005. doi: 10.1117/1.JBO.22.12.126005
- Lee J, Prabhu D, Kolluru C, Gharaibeh Y, Zimin V, Bezerra HG, et al. Automated plaque characterization using deep learning on coronary intravascular optical coherence tomographic images. *Biomed Optics Exp.* (2019) 10:6497–515. doi: 10.1364/BOE.10.006497
- He C, Wang J, Yin Y, Li Z. Automated classification of coronary plaque calcification in OCT pullbacks with 3D deep neural networks. *J Biomed Optics.* (2020) 25:095003. doi: 10.1117/1.JBO.25.9.095003
- Gharaibeh Y, Prabhu D, Kolluru C, Lee J, Zimin V, Bezerra H, et al. Coronary calcification segmentation in intravascular OCT images using deep learning: application to calcification scoring. *J Med Imaging (Bellingham, Wash).* (2019) 6:45002. doi: 10.1117/1.JMI.6.4.045002
- Li L, Jia T. Optical coherence tomography vulnerable plaque segmentation based on deep residual U-net. *Rev Cardiovasc Med.* (2019) 20:171–7. doi: 10.31083/j.rcm.2019.03.5201
- Athanasiou LS, Olender ML, de la Torre Hernandez JM, Ben-Assa E, Edelman ER. A deep learning approach to classify atherosclerosis using intracoronary optical coherence tomography. In: *Medical Imaging 2019: Computer-Aided Diagnosis.* Vol. 10950. San Diego, CA: International Society for Optics and Photonics (2019). doi: 10.1117/12.2513078
- Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th International Conference on Machine Learning (ICML-10).* Haifa (2020). p. 807–14.
- Srivastava N, Hinton G, Krizhevsky A, Sutskever I, Salakhutdinov R. Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res.* (2014) 15:1929–58. doi: 10.5555/2627435.2670313
- Havaei M, Davy A, Warde-Farley D, Biard A, Courville A, Bengio Y, et al. Brain tumor segmentation with deep neural networks. *Med Image Anal.* (2017) 35:18–31. doi: 10.1016/j.media.2016.05.004
- He X, Zemel RS, Carreira-Perpiñán MÁ. Multiscale conditional random fields for image labeling. In: *Proceedings of the 2004 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2004. CVPR 2004. IEEE II-II.* Washington, DC (2004).
- Fu H, Xu Y, Lin S, Wong DWK, Liu J. Deepvessel: retinal vessel segmentation via deep learning and conditional random field. In: *International Conference*

- on *Medical Image Computing and Computer-Assisted Intervention*: Springer. Athens (2016). p. 132–9.
39. Prati F, Regar E, Mintz GS, Arbustini E, Di Mario C, Jang K, et al. Expert review document on methodology, terminology, and clinical applications of optical coherence tomography: physical principles, methodology of image acquisition, and clinical application for assessment of coronary arteries and atherosclerosis. *Eur Heart J*. (2010) 31:401–15. doi: 10.1093/eurheartj/ehp433
  40. Wang Z, Chamie D, Bezerra HG, Yamamoto H, Kanovsky J, Wilson DL, et al. Volumetric quantification of fibrous caps using intravascular optical coherence tomography. *Biomed Optics Exp*. (2012) 3:1413–26. doi: 10.1364/BOE.3.001413
  41. Tieleman T, Hinton G. Lecture 6.5-rmsprop: divide the gradient by a running average of its recent magnitude. *COURSERA Neural Netw Mach Learn*. (2012) 4:26–31.
  42. Sokolova M, Japkowicz N, Szpakowicz S. Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In: *Australasian Joint Conference on Artificial Intelligence*. Hobart: Springer (2006). p. 1015–21.
  43. Tearney GJ, Regar E, Akasaka T, Adriaenssens T, Barlis P, Bezerra HG. Consensus standards for acquisition, measurement, and reporting of intravascular optical coherence tomography studies: a report from the International Working Group for Intravascular Optical Coherence Tomography Standardization and Validation. *J Am Coll Cardiol*. (2012) 59:1058–72. doi: 10.1016/j.jacc.2011.09.079
  44. Ughi GJ, Wang H, Gerbaud E, Gardecki JA, Fard AM, Hamidi E, et al. Clinical characterization of coronary atherosclerosis with dual-modality OCT and near-infrared autofluorescence imaging. *JACC Cardiovasc Imag*. (2016) 9:1304–14. doi: 10.1016/j.jcmg.2015.11.020
  45. Nishimiya K, Yin B, Piao Z, Ryu J, Osman H, Leung HM, et al. Micro-optical coherence tomography for endothelial cell visualization in the coronary arteries. *JACC Cardiovasc Imag*. (2019) 12:1878–80. doi: 10.1016/j.jcmg.2019.01.021
  46. Otsuka K, Villiger M, Karanasos A, Zandvoort L, Bouma BE. Intravascular polarimetry in patients with coronary artery disease. *JACC Cardiovasc Imag*. (2019) 13:790–801. doi: 10.1016/S0735-1097(19)30681-3

**Conflict of Interest:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2021 Yin, He, Xu and Li. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.