

Convolution Neural Networks and Support Vector Machines for Automatic Segmentation of Intracoronary Optical Coherence Tomography

Caining Zhang¹, Huaguang Li², Xiaoya Guo³, David Molony⁴, Xiaopeng Guo², Habib Samady⁴, Don P. Giddens^{4,5}, Lambros Athanasiou⁶, Rencan Nie^{2,*}, Jinde Cao^{3,*} and Dalin Tang^{1,*7}

¹School of Biological Science & Medical Engineering, Southeast University, Nanjing, 210096, China.

²School of Information Science & Engineering, Yunnan University, Kunming, 650091, China.

³School of Mathematics, Southeast University, Nanjing, 210096, China.

⁴Department of Medicine, Emory University School of Medicine, Atlanta, GA, 30307, USA.

⁵The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology, Atlanta, GA, 30332 USA.

⁶Institute for Medical Engineering & Science, Massachusetts Institute of Technology, Cambridge, MA 02139 USA.

⁷Mathematical Sciences Department, Worcester Polytechnic Institute, Worcester, MA 01609 USA.

*Corresponding Authors: Dalin Tang, Southeast University, China. Email: dtang@wpi.edu; Rencan Nie. Email: renie@ynu.edu.cn; Jinde Cao. Email: jdcao@seu.edu.cn.

Abstract: Cardiovascular diseases are closely associated with deteriorating atherosclerotic plaques. Optical coherence tomography (OCT) is a recently developed intravascular imaging technique with high resolution approximately 10 microns and could provide accurate quantification of coronary plaque morphology. However, tissue segmentation of OCT images in clinic is still mainly performed manually by physicians which is time consuming and subjective. To overcome these limitations, two automatic segmentation methods for intracoronary OCT image based on support vector machine (SVM) and convolutional neural network (CNN) were performed to identify the plaque region and characterize plaque components. *In vivo* IVUS and OCT coronary plaque data from 5 patients were acquired at Emory University with patient's consent obtained. Seventy-seven matched IVUS and OCT slices with good image quality and lipid cores were selected for this study. Manual OCT segmentation was performed by experts using virtual histology IVUS as guidance, and used as gold standard in the automatic segmentations. The overall classification accuracy based on CNN method achieved 95.8%, and the accuracy based on SVM was 71.9%. The CNN-based segmentation method can better characterize plaque compositions on OCT images and greatly reduce the time spent by doctors in segmenting and identifying plaques.

Keywords: Atherosclerotic plaques; OCT; CNN; U-Net; SVM; segmentation

1 Introduction

Cardiovascular diseases are closely associated with atherosclerotic plaque development and rupture. Accurate and detailed identification of atherosclerotic plaques advances the understanding and diagnosis of cardiovascular diseases. However, the current common image modalities such as magnetic resonance imaging (MRI) and intravascular ultrasound (IVUS) were unable to identify vulnerable coronary plaques due to their limited imaging resolution [1-2]. Optical coherence tomography (OCT) is a recently emerged advanced intravascular imaging technique which has high resolution approximately 10 microns and could provide accurate morphology for coronary plaques [3-4]. Moreover, it has the unique ability to identify plaques with fibrous cap thickness < 65 μm , an accepted threshold value for vulnerable plaques [5]. However, the current segmentation of OCT images in clinic is still mainly performed manually by physicians, which is time consuming and can lead to inter-observer and intra-observer variability [6]. As such, automatic segmentation and recognition of vulnerable plaques through quantification of plaque

components have great clinical significance for cardiovascular research. To overcome time-consuming manual segmentation process, several methodologies have been proposed for automatic segmentation of OCT images. Van Soest et al. [7] presented a framework for automatic classification of atherosclerotic plaque constituents based on the optical attenuation coefficient of the tissue in IVOCT images. Athanasiou et al. [8] presented a fully automated methodology to detect the lumen border, identify the plaque region and detect four tissue types based on random forests (RF). Shalev et al. [9] and Guo et al. [10] applied support vector machine (SVM) to characterize plaques. Convolutional neural network (CNN) have become a popular method for image analysis and have recently been applied to OCT images. More recently, Abdolmanafi et al. [11] and Athanasiou et al. [12] investigated deep learning models based on CNN for tissue characterization. The above literature methods have demonstrated that both machine learning and deep learning can be applied to plaque segmentation with high accuracy. However, in all methods the annotations that were used were based on experts' experience and not on any form of *a priori* knowledge such as histology; histology/virtual-histology still remains the gold standard in intracoronary plaque segmentation. In addition, most of these methods were still limited by intricate image preprocessing and expensive computation.

In this paper, two intracoronary OCT segmentation methods, based on SVM and CNN, were implemented to identify plaque region and characterize different plaque components. The current gold standard for training and testing the methods was a based on a hybrid plaque estimation approach: experts' knowledge and virtual histology were combined in an attempt to present a more robust plaque segmentation method. Therefore, the results of the two segmentation algorithms, were compare on the same dataset.

2 Data Acquisition and Segmentation Methods

2.1 Data Acquisition and Hybrid Gold Standard Formation

OCT and IVUS image data from 5 patients were acquired at Emory University with informed consent obtained. OCT were obtained from ILUMIEN OPTIS System (St. Jude, Minnesota, MN). The OCT catheter was traversed to the segment of interest and the catheter pullback was limited at a rate of 20 mm/sec. Following the OCT image acquisition, the IVUS catheter (Volcano Therapeutics, Rancho Cordova, CA) was traversed distally though the artery to the same coronary segment and the catheter pullback speed was at a standard rate of 0.5 mm/sec. Four different color were used to distinguish four plaque types in colored VH-IVUS (virtual histology IVUS) images: fibrous, fibro-fatty, lipid and dense calcified tissue (calcification, Ca in short) (see Fig. 1(b)). Dynamic angiography data were recorded which provided positions of both catheters for co-registration of OCT and IVUS images. Seventy-seven matched IVUS/VH-IVUS and OCT slices with good image quality and lipid cores were identified for our use in the training and validation of the segmentation methods. Manual OCT segmentation was performed by experts using VH-IVUS images as references and used as gold standard in the automatic segmentations [13]. Manual annotations were performed for the identification of lipid tissue (LT), fibrous tissue (FT) and background (BG). Fig. 1 gave 5 sample slices with VH-IVUS image slices, VH-IVUS contour plots, OCT image slices, OCT contour plots (expert-drawn), and registered OCT and IVUS slices with OCT and VH-IVUS overlapped. Tiny calcification was neglected in this study.

2.2 Machine Learning Segmentation Method

Two segmentation methods for intracoronary OCT image based on convolutional neural network (CNN) and support vector machine (SVM) were performed to detect the lumen borders and characterize the plaque component.

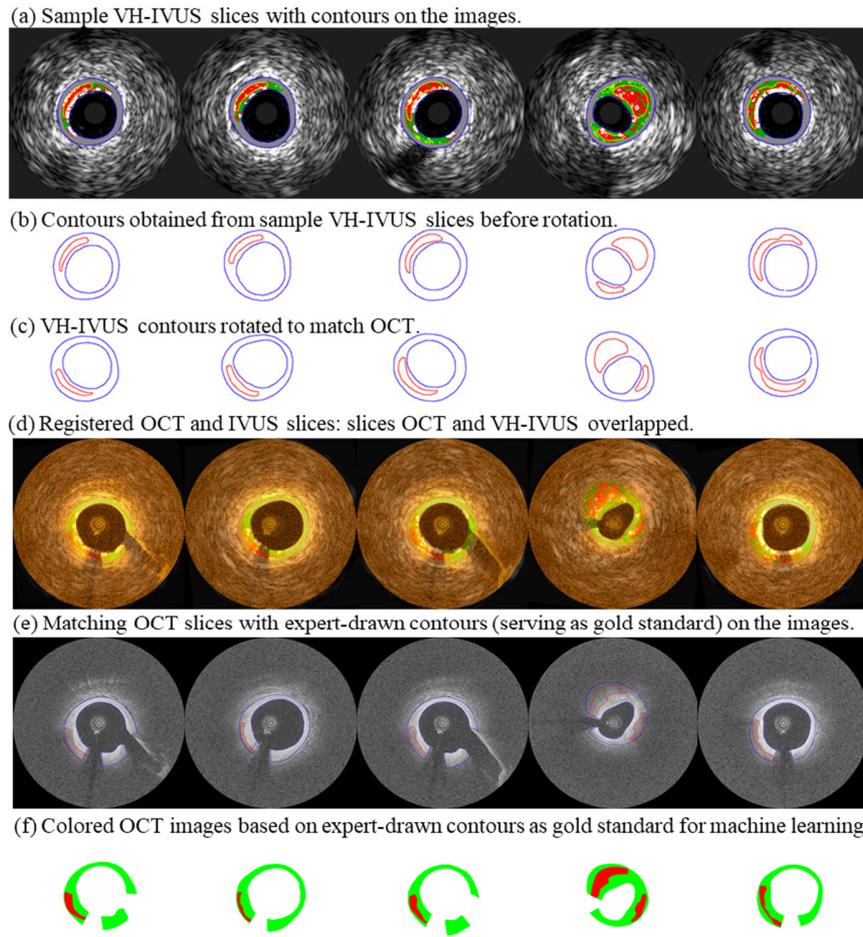


Figure 1: Sample VH-IVUS and OCT slices and their contour plots. Expert-drawn contours from OCT slices will serve as gold standard for machine learning segmentation method. Colors used in VH-IVUS: Red-necrotic core; White-dense calcium; Dark Green-Fibrous; Light Green-Fibro-Fatty. Colors used in gold standard: Red-lipid tissue, Green-fibrous tissue; White-background

2.2.1 Convolutional Neural Network (CNN) Method

Image processing and data augmentation

Due to limited penetration depth of OCT images, the input images of network was truncated from 1024 by 1024 pixels (original OCT data matrix) to 592 by 544 pixels (the truncated data matrix indicated by the marked rectangle) to remove invalid information (see Fig. 2) and save the time of model training. Furthermore, since the power of deep learning models highly relies on the size of the training data, we enhanced existing datasets *via* data augmentation to train a more robust deep convolutional neural network. In our study, the data augmentation operation contains rotation (the rotation angle $\theta = 90, 180, 270, 360, 450, 540$ degree) and brightness transformation (the parameter is 0.5, 0.7, 0.9, 1.1).

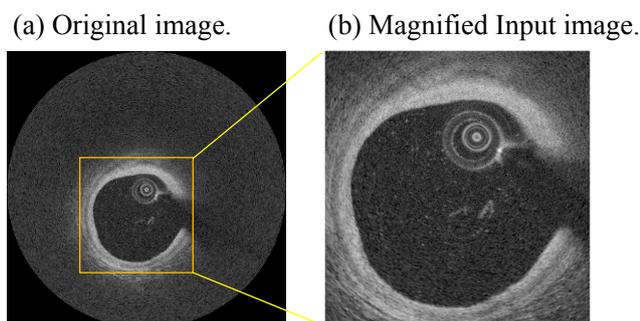


Figure 2: Image processing for CNN. (a) Original image (the yellow rectangular marked input image); (b) Magnified input image of network

Model architecture

Deeper layers of a convolutional neural network can extract detailed lower-level information from the original image, which is appealing for segmentation in OCT images. The neural network model chosen for this problem was based on the U-Net architecture, which has promising future in the field of segmentation, particularly for medical images [14]. The network can take a full image section as input and then, through a series of trainable weights, create the corresponding section segmentation mask. Focal loss function was used in U-Net model between the true segmentation and the output of our model due to the class imbalance of the proportion that lipid tissue make up compared with the other classes [15]. The original 77 OCT image datasets were divided into eleven test sets, and eleven-fold cross validation was employed, such that when a test set was selected, the rest of the dataset and the data augmentation based were served as the training set. Model selection was made by using the 1-standard-error rule on the validation data set [16]. Five times training and testing were done in this method. All U-Net models were implemented in Pytorch, version 3.6.

2.2.2 Support Vector Machine (SVM) Method

Image processing and ROI definition

With the same reason as we explained in CNN method section (limited penetration of OCT), it is not sensible to process the whole OCT image slice. To make the region of interest (ROI), an outline (called the out-boundary of the ROI) was generated by using the lumen border (the inner boundary of the ROI) in polar coordinate system and drawing it 1 mm outward from center of catheter. The region of interest (ROI) is the area bounded by lumen and the out-boundary with artery branches and artifacts removed (see Fig. 3, ROI-yellow).

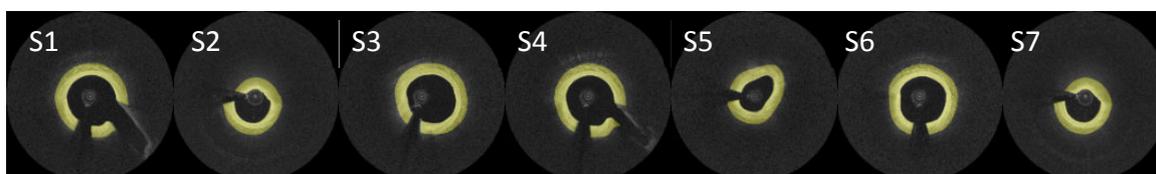


Figure 3: Seven OCT images with ROI and borders of ROI. ROI is marked by yellow

Feature extraction and classification

Local binary patterns (LBPs), gray level co-occurrence matrices (GLCMs) which contains contrast, correlation, energy and homogeneity, entropy and mean value were calculated as features. All features were computed in an 11 by 11 pixels' neighborhood window in ROI [9]. Ten features based on rotation invariant LBPs were calculated with $P = 8$, $R = 1$. Twenty-eight features were chosen and combined to improve the accuracy of our classification algorithm which contains 10 LBPs, 16 GLGMs (Four features at each angle

$\theta = 0^{\circ}, 45^{\circ}, 90^{\circ},$ and 135°), entropy and mean value. These selected features which extracted from all pixels in ROI were assembled into a data matrix to feed SVM. The dimension of data matrix is n by m , where n is equal to the number of pixels and m is 28, the length of the feature vector. According to the gold standard, three classes are needed to be characterized indicating three different types: lipid tissue, fibrous tissue and background. Multi-class support vector machine classifier was chosen to classify three types with Gaussian Radial Basis Function as the kernel function. In keeping with CNN, the 11-fold cross validation was employed in SVM method. For each group of training OCT images, 500 pixels were random selected from lipid class, fibrous class and background class each to give 105000 pixels. The testing set consists of all pixels in ROI of the rest testing OCT image. Five times training and testing were done in this method. All multi-class SVM models were implemented in Python, version 3.6.

3 Results

3.1 Gold Standard and Segmentation Accuracy

Figs. 4 & 5 gave seven sample OCT slices with our hybrid segmentation by expert (gold standard) and segmentation by CNN and SVM after processing, respectively. The results of segmentation methods were evaluated and compared using the OCT dataset and gold standard at pixel level. Accuracy (Acc), sensitivity (Sen) and specificity (Spe) are defined as follows:

$$Acc = \frac{TP+TN}{TP+FP+TN+FN} \tag{1}$$

$$Sen = \frac{TP}{TP+FN} \tag{2}$$

$$Spe = \frac{TN}{TN+FP} \tag{3}$$

where TP is the number of true positive outcomes, FP is the number false positive outcomes, TN is the number of true negative outcomes and FN is the false negative outcomes.

(a) Gold standard of CNN method



(b) Classification by CNN



Figure 4: Magnified image classification by CNN. Colors used in gold standard: Red lipid tissue, Green fibrous tissue, Black background

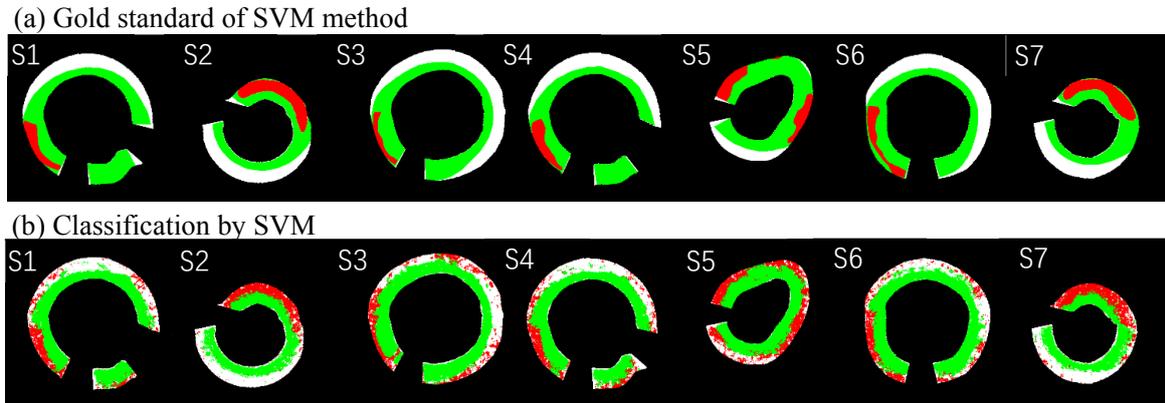


Figure 5: ROI classification by SVM. Colors used in gold standard: Red-lipid tissue; Green-fibrous tissue; White-background

3.2 Accuracy of Tissue Region Detection

In order to evaluate the accuracy of tissue region detection, lipid tissue and fibrous tissue were considered as positive patterns (prediction target) while background was regarded as negative pattern. The accuracy of tissue region using two methods are shown in Tab. 1. The overall average prediction accuracy from the 11 sets was 96.20% and 79.54% for CNN and SVM, respectively. The best accuracy was 97.04% for CNN and 82.65% for SVM. Even for the worst case, the prediction accuracy was 95.08% by CNN and 76.58% by SVM. Obviously, CNN provided more accurate results than SVM, especially in specificity. SVM will always underestimate the tissue region as it uses as input the ROI.

Table 1: Accuracy of tissue region detection by using CNN and SVM

| | | CNN | | | SVM | | |
|-----------------|--------|-------|-------|-------|-------|-------|-------|
| | | Sen | Spe | Acc | Sen | Spe | Acc |
| Average | Set-1 | 92.36 | 96.42 | 95.67 | 91.02 | 56.44 | 77.87 |
| | Set-2 | 93.02 | 97.49 | 96.50 | 84.79 | 66.78 | 78.47 |
| | Set-3 | 90.63 | 98.24 | 96.56 | 84.92 | 70.09 | 78.68 |
| | Set-4 | 91.20 | 97.78 | 96.49 | 86.52 | 65.93 | 79.42 |
| | Set-5 | 95.13 | 97.53 | 97.04 | 87.94 | 67.37 | 80.52 |
| | Set-6 | 87.36 | 98.36 | 96.05 | 88.21 | 61.46 | 82.61 |
| | Set-7 | 90.08 | 97.58 | 95.89 | 87.64 | 65.73 | 81.52 |
| | Set-8 | 91.65 | 98.13 | 96.80 | 87.61 | 60.54 | 78.18 |
| | Set-9 | 87.16 | 97.37 | 95.08 | 87.40 | 53.26 | 76.58 |
| | Set-10 | 90.00 | 97.73 | 96.12 | 88.78 | 69.40 | 82.65 |
| | Set-11 | 89.63 | 97.69 | 96.00 | 86.08 | 62.54 | 78.44 |
| Overall average | | 90.75 | 97.67 | 96.20 | 87.36 | 63.59 | 79.54 |

3.3 Accuracy of Objects Classification

The suggested methods can extract lipid (LT), fibrous tissue (FT), and background (BG) directly from OCT images, and categorize each pixel into its corresponding class. Tab. 2 indicated that prediction accuracy based on CNN was 81.92%, 74.14%, 97.67% and 94.29% for FT, LT, BG and overall, respectively. Prediction accuracy using SVM was 74.46%, 65.33%, 63.59% and 69.46% for FT, LT, BG and overall, respectively. CNN produced

much better scores in all indexes, and the overall accuracy rate of CNN was 35.75% higher than that from SVM. The reason that LT accuracy scores for CNN were lower than scores for other objects is that the percentage of LT pixels in OCT images was lower compared to the other 2 objects (FT and BG), causing training imbalance.

Table 2: Classification accuracy of objects using CNN and SVM

| | CNN | | | | SVM | | | |
|-----------------|-------|-------|-------|---------|-------|-------|-------|---------|
| | FT | LT | BG | Overall | FT | LT | BG | Overall |
| Set-1 | 82.73 | 79.18 | 96.42 | 94.02 | 83.50 | 72.67 | 56.44 | 72.59 |
| Set-2 | 84.32 | 83.80 | 97.49 | 94.91 | 74.07 | 62.40 | 66.78 | 69.87 |
| Set-3 | 83.36 | 79.62 | 98.24 | 95.15 | 70.47 | 60.16 | 70.09 | 67.02 |
| Set-4 | 81.66 | 67.11 | 97.78 | 94.28 | 68.13 | 64.58 | 65.93 | 66.42 |
| Set-5 | 90.62 | 76.34 | 97.53 | 95.98 | 81.72 | 58.87 | 67.37 | 74.73 |
| Average Set-6 | 77.94 | 73.61 | 98.36 | 93.78 | 69.31 | 71.69 | 61.46 | 67.99 |
| Set-7 | 80.58 | 65.31 | 97.58 | 93.28 | 71.71 | 69.70 | 65.73 | 69.43 |
| Set-8 | 81.98 | 82.84 | 98.13 | 95.14 | 74.30 | 65.28 | 60.54 | 68.29 |
| Set-9 | 78.57 | 65.68 | 97.37 | 92.90 | 74.77 | 62.93 | 53.26 | 66.56 |
| Set-10 | 77.51 | 69.06 | 97.73 | 93.44 | 76.57 | 63.76 | 69.40 | 72.11 |
| Set-11 | 81.80 | 72.99 | 97.69 | 94.31 | 74.49 | 66.58 | 62.54 | 69.09 |
| Overall average | 81.92 | 74.14 | 97.67 | 94.29 | 74.46 | 65.33 | 63.59 | 69.46 |

3.4 Evaluation of the Discrimination Between LT and FT in Correct Tissue Region Detected

To further evaluate the discrimination between LT and FT in correctly detecting the tissue region, LT and FT were considered as positive pattern and negative pattern, respectively. The results of differentiating LT and FT are shown in Tab. 3. The overall average accuracy (*Acc*) were 88.76% and 83.36% for CNN and SVM, respectively. The prediction based on CNN presented the best and worst average *Acc*: 92.82% and 87.01% on Set 5 and Set 4, respectively. While SVM depicted better *Sen* for LT, it had worse *Spe* for FT. Overall, both CNN and SVM provided good discrimination between LT and FT.

Table 3: Evaluation of the discrimination between LT and FT in correctly detecting the tissue region

| | CNN | | | SVM | | |
|-----------------|-------|-------|-------|-------|-------|-------|
| | Sen | Spe | Acc | Sen | Spe | Acc |
| Set-1 | 80.66 | 90.32 | 89.02 | 87.49 | 90.94 | 90.63 |
| Set-2 | 84.39 | 92.33 | 90.80 | 87.13 | 85.19 | 84.61 |
| Set-3 | 79.97 | 93.38 | 90.76 | 79.17 | 81.81 | 80.00 |
| Set-4 | 68.61 | 90.96 | 87.01 | 88.99 | 77.77 | 78.53 |
| Set-5 | 76.73 | 96.63 | 92.82 | 81.24 | 91.07 | 89.33 |
| Average Set-6 | 73.85 | 93.43 | 87.87 | 91.62 | 77.20 | 78.28 |
| Set-7 | 68.38 | 90.71 | 86.13 | 87.77 | 80.14 | 80.86 |
| Set-8 | 83.57 | 90.91 | 90.64 | 90.97 | 82.93 | 83.42 |
| Set-9 | 68.20 | 92.29 | 87.80 | 87.78 | 83.56 | 83.34 |
| Set-10 | 71.31 | 87.90 | 84.34 | 83.16 | 84.08 | 83.05 |
| Set-11 | 76.23 | 91.82 | 89.14 | 85.67 | 85.67 | 84.94 |
| Overall average | 75.63 | 91.88 | 88.76 | 86.45 | 83.67 | 83.36 |

4 Discussion

4.1 Signification

Identification of plaque region and segmentation of plaque components in OCT images could provide accurate plaque morphological information which is the basis for clinical diagnosis, image analysis and image-based computation modeling applications. Image-based biomechanical plaque models may be used for researching and quantifying plaque mechanical stress and strain conditions, identifying possible mechanical factors associated with plaque rupture, and assessing plaque rupture risk. Hence, the accuracy of segmentation for OCT images play an important role in model construction and prediction of plaque rupture.

The proposed methods are capable of detecting the plaque region and characterizing the plaque component automatically with high accuracy. In addition, CNN method using U-Net model has a specific advantage over previous segmentation methods in that it is an “end-to-end” method. The automatic segmentation methods proposed by Athanasiou et al. [8,12] needs to detect both the lumen border and outer border before segmentation. Another method proposed by Shalev et al. [9] requires five steps in preprocessing. The proposed CNN method has no production line that requires large image preprocessing and post-processing, resulting in prominent improvements in performance time. Moreover, it is trained and tested on more realistic data which are produced by combining experts’ knowledge with virtual histology, the only commercially available plaque segmentation method which is used in clinical practice.

4.2 Limitations

Although the proposed study uses a more robust gold standard than the methods presented in the literature, the lack of histology is a drawback. Histological data remain the current gold standard in tissue detection. Histology could provide more effective and additional information which could be used to improve and optimize the efficacy of our methodology. In addition, given the limitation of a small dataset, a large amount of data would be required to improve the accuracy and reliability of our study.

5 Conclusions

Currently, we presented two machine learning approaches to automatically identify plaque region and characterize plaque compositions in OCT images. The current gold standard is based on a hybrid annotation approach where VH-IVUS was used to guide experts’ annotations. Both methods can potentially reduce the time spent by doctors in segmenting and evaluating coronary plaque OCT images. However, it is apparent that the CNN based method can provide better segmentation accuracies compared to those achieved by SVM and is more likely to be used in the clinical and research arena.

Acknowledgement: This research was supported in part by NIH grant R01 EB004759, National Sciences Foundation of China grant 11672001 and a Jiangsu Province Science and Technology Agency grant BE2016785.

References

1. Schaar JA, Mastik F, Regar E, den Uil CA, Gijssen FJ et al. Current diagnostic modalities for vulnerable plaque detection. *Current Pharmaceutical Design* **2007**, 13(10): 995-1001.
2. Nair A, Kuban BD, Tuzcu EM, Schoenhagen P, Nissen SE et al. Coronary plaque classification with intravascular ultrasound radiofrequency data analysis. *Circulation* **2002**, 106(17): 2200-2206.
3. Bezerra HG, Costa MA, Guagliumi G, Rollins AM, Simon DI. Intracoronary optical coherence tomography: a comprehensive review clinical and research applications. *JACC: Cardiovascular Interventions* **2009**, 2(11): 1035-1046.
4. Tearney GJ, Regar E, Akasaka T, Adriaenssens T, Barlis P et al. International Working Group for Intravascular Optical Coherence, T. Consensus standards for acquisition, measurement, and reporting of intravascular optical coherence tomography studies: a report from the International Working Group for Intravascular Optical Coherence Tomography Standardization and Validation. *Journal of the American College of Cardiology* **2012**,

- 59(12): 1058-1072.
5. Virmani R, Kolodgie FD, Burke AP, Farb A, Schwartz SM. Lessons from sudden coronary death: a comprehensive morphological classification scheme for atherosclerotic lesions. *Arteriosclerosis, Thrombosis, and Vascular Biology* **2000**, 20(5): 1262-1275.
 6. Kume T, Akasaka T, Kawamoto T, Watanabe N, Toyota E et al. Assessment of coronary arterial plaque by optical coherence tomography. *American Journal of Cardiology* **2006**, 97(8): 1172-1175.
 7. Van Soest G, Goderie TP, Regar E, Koljenovic S, van Leenders AGJ et al. Atherosclerotic tissue characterization *in vivo* by optical coherence tomography attenuation imaging. *Journal of Biomedical Optics* **2010**, 15(1): 011105.
 8. Athanasiou LS, Bourantas CV, Rigas G, Sakellarios AI, Exarchos TP et al. Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *Journal of Biomedical Optics* **2014**, 19(2): 026009.
 9. Shalev R, Prabhu D, Tanaka K, Rollins AM, Costa M et al. Intravascular optical coherence tomography image analysis method. *41st Annual Northeast Biomedical Engineering Conference* **2015**: 1-2.
 10. Guo X, Tang D, Molony D, Yang C, Samady H et al. A segmentation method for intracoronary oct image based on least squares support vector machine: vulnerable coronary plaque cap thickness quantification. *Proceedings of the International Conference on Computational Methods* **2017**: 346-359.
 11. Abdolmanafi A, Duong L, Dahdah N, Adib IR, Cheriet F. Characterization of coronary artery pathological formations from OCT imaging using deep learning. *Biomedical Optics Express* **2018**, 9(10): 4936-4960.
 12. Athanasiou LS, Olender ML, José M, Ben-Assa E, Edelman ER. A deep learning approach to classify atherosclerosis using intracoronary optical coherence tomography. *SPIE Medical Imaging* **2019**, 10950.
 13. Raber L, Heo JH, Radu MD, Garcia-Garcia HM, Stefanini GG et al. Offline fusion of co-registered intravascular ultrasound and frequency domain optical coherence tomography images for the analysis of human atherosclerotic plaques. *EuroIntervention* **2012**, 8(1): 98-108.
 14. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention* **2015**: 234-241.
 15. Lin TY, Goyal P, Girshick R, He K, Dollár P. Focal loss for dense object detection. *Proceedings of the IEEE International Conference on Computer Vision* **2017**: 2980-2988.
 16. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning. New York, NY: Springer. **2009**.