MCB, vol.15, no.2, pp.117-125, 2018

Intravascular Optical Coherence Tomography Image Segmentation Based on Support Vector Machine Algorithm

Yuxiang Huang[1], Chuliu He[1], Jiaqiu Wang[2], Yuehong Miao[1], Tongjin Zhu[1], Ping Zhou[1] and Zhiyong Li[1, 2, *]

Abstract: Intravascular optical coherence tomography (IVOCT) is becoming more and more popular in clinical diagnosis of coronary atherosclerotic. However, reading IVOCT images is of large amount of work. This article describes a method based on image feature extraction and support vector machine (SVM) to achieve semi-automatic segmentation of IVOCT images. The image features utilized in this work including light attenuation coefficients and image textures based on gray level co-occurrence matrix. Different sets of hyper-parameters and image features were tested. This method achieved an accuracy of 83% on the test images. Single class accuracy of 89% for fibrous, 79.3% for calcification and 86.5% lipid tissue. The results show that this method can be a considerable way for semi-automatic segmentation of atherosclerotic plaque components in clinical IVOCT images.

Keywords: IVOCT, image segmentation, support vector machine, attenuation coefficient, image texture features.

1 Introduction

In recent years, intravascular optical coherence tomography (IVOCT) [Bezerra, Costa, Guagliumi et al. (2009)] has been increasingly applied to the clinical diagnosis of coronary heart disease. Comparing to coronary angiography, IVOCT has the advantage on providing a tomographic image of the lesion vessel and obtaining the detailed information on the vessel wall. IVOCT also provides images with a resolution of around 10 μm, an order of magnitude higher than the 200 μm resolution of intravascular ultrasound (IVUS). This demonstrates that it has the ability to clearly characterize the endovascular tissues, providing detailed and accurate information for stent implantation and other interventional procedures.

During imaging diagnosis, the IVOCT imaging probe is pulled by the imaging guidewire and extends into the lesion vessels for imaging. A single imaging pullback process yields approximately 270 consecutive blood vessel cross-sectional images, which gives the clinician a large amount of reading work. Therefore, automated or assisted IVOCT image reading techniques, to distinguish between vascular tissue components, are in great need.

---

[1] School of Biological Science and Medical Engineering, Southeast University, Nanjing 210096, China.

[2] School of Chemistry, Physics and Mechanical Engineering, Queensland University of Technology, Brisbane, QLD 4000, Australia.

[*] Corresponding author: Zhiyong Li. Email: zylicam@gmail.com.

MCB. doi: 10.3970/mcb.2018.02478

Among the existing studies, He et al. [He and Chunliu (2016)] implements a fuzzy *c* means clustering with geometric constrains, followed by a linear regression to analysis the correlation between IVOCT image features and manual Fibrous-cap-thickness measurements. In His study, some OCT features were found to have good performance and could be utilized for automatic qualitative analysis and the identification of high-risk plaques. Tsantis et al. [Tsantis, Kagadis, Katsanos et al. (2012)] extracted several features of the stent based on wavelet transform and then used neural network model to detect it. Ughi et al. [Ughi, Adriaenssens, Onsea et al. (2012)] introduced a method using the attenuation coefficient, which is extracted from each *A*-scan line according to the optical properties of tissue, combined with other texture features to perform image segmentation on the plaque tissue area. Athanasiou et al. [Athanasiou, Bourantas, Rigas et al. (2014)] first utilized the *K*-means clustering algorithm to extract the calcified region, and then used the gray level co-occurrence matrix and extracted the LBP texture. The classification was implemented based on the comprehensive consideration of the random forest algorithm. Wang et al. [Wang, Kyono, Bezerra et al. (2010)] proposed a way for semi-automatic detection of calcified plaques. In this study, edge detection was used to roughly extract the boundaries of calcified plaques, and then the active contour method was used to further adjust them. This method provides a clear and delineated boarder of calcifications but with limitations as users must first artificially remove false-positive calcified areas.

This work implements a method based on multiple image feature extraction and support vector machine classifier (SVM) to segment the IVOCT blood vessel cross-sectional images. Training labels and “precision standards” come from manual segmentation. This work aims to distinguish three major plaque components at the site of coronary atherosclerotic lesions, i.e. calcification, fibrous tissue, and lipid containing regions [Tearney, Regar, Akasaka et al. (2012)].

2 Methods

The segmentation process steps are described as follows: pre-processing, feature extraction, SVM training and acquiring classification results. The purpose of the preprocessing is to exclude various artifacts and interferences on the image for subsequent steps of feature extraction [Tearney, Regar, Akasaka et al. (2012)]. The extracted features are enumerated in Section 2.3. Specific settings of parameters and training procedures of the SVM classifier are described in Section 2.3.

2.1 Data acquisition

Intracoronary images employed in this study were collected from 11 patients (age was 61-73 years and four were women) at Nanjing Drum Tower Hospital in Nanjing with a Fourier-Domain OCT system (C7-XR, St. Jude Medical, Inc.) and the C7 Dragonfly intravascular OCT catheter (St. Jude). All data were de-identified before sending to Southeast University for analysis. Imaging specifications, according to the manufacturer, consist in an axial resolution of 15 μm and a lateral resolution of 19 μm. Scan parameters are 100 frames per second, pullback speed of 20 mm/s and scan depth of 1.0-2.0 mm in biological tissue.

The image datasets were acquired from 11 clinical patients and were non-identifiable. Each frame of raw image was of 504×976 size and stored in polar coordination. A total of 28 slices of images were chosen. Three well-trained IVOCT image reader implemented manual segmentation on each image as gold standard of this study. Average consistency between segmentation results from IVOCT readers is 91%.

2.2 Preprocessing

The original data were stored in polar coordinates as shown in Fig. 1(a). In order to eliminate the noise in the lumen, the pre-processing method uses Gaussian filtering on the image.

The steps to remove the guide wire artifacts refer to the method by Wang et al. [Wang, Kyono, Bezerra et al. (2010)]. Due to the consistency of the guide wire artifacts location in IVOCT frames within a single pullback, a darker trajectory representing the movement of guide wire position along the pullback, as presented in Fig. 1(b), can be generated by: 1) generate a single line containing means of every A-scan line in the origin frame; 2) arrange each average line together to acquire a map as shown in Fig. 1(b). Thus the aim to extract the high-lighted yellow boundary in Fig. 1(b) can be treated as a dynamic programming problem [Howard (1966)].

Fig. 1(c) was obtained by implementing an OSTU threshold filtering on a single frame image which has eliminated the guidewire artifacts with the position obtained. Lumen position can be easily acquired from Fig. 1(c). Fig. 1(d) is the Cartesian view obtained from Fig. 1(a) and Fig. 1(c), in which the red line in the figure is the contour of the vessel lumen [Ping, Zhu and Li (2015)]. Then, according to the lumen position, the region of interest (ROI) was set by extending a distance of 270 pixels along the axial direction.

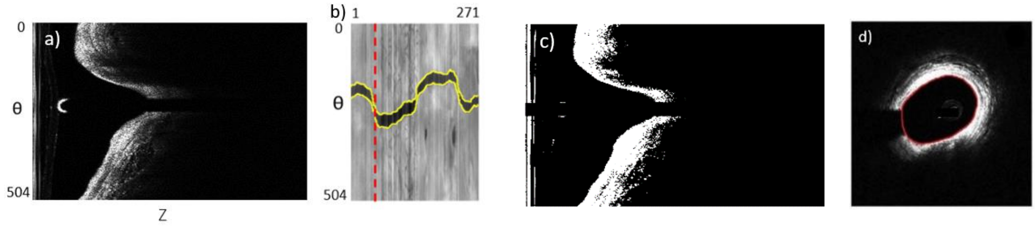


Figure 1: (a) An IVOCT frame in polar coordinate; (b) Average line map. Red dotted line refers to the frame position of Fig. 1(a); (c) OSTU threshold map; (d) Cartesian coordinate map of Fig. 1(a), in which the red line implies lumen position

2.3 Extract features

The segmentation was based on individual pixels. The target was to determine which category it belongs to pixel by pixel. It should be noted that each feature for a single pixel was extracted from a neighborhood of a particular sized window where the pixel was located. The utilized features are described below:

(1) Attenuation coefficient: the method proposed by Van et al. [Van, Goderie, Regar et al. (2010); Soest, Goderie, Noorden et al. (2009)] was used to perform line-by-line attenuation coefficient extraction of *A*-scan line of IVOCT image. This method automatically extracts the attenuation coefficient μ_t by choosing a window to fit the *A*-scan line according to Eq. (1):

$$I_d(z) = I_0 \cdot \exp(-\mu_t z)(1) \tag{1}$$

(2) We extracted a series of gray level co-occurrence matrix based texture features. Let C_{ij} be the element of the *i*-th row and the *j*-column in the normalized matrix, Let P_{ij} be the element of the *i*-th row and *j*-column in the original GLCM, then: $C_{ij} = P_{ij}/\sum_{i,j}^{G} P_{ij}$ $C_{ij} = P_{ij}/\sum_{i,j}^{G} P_{ij}$, Then we can extract features such as:

Contrast:

$$f = \sum C_{ij}(i-j)^2 \tag{2}$$

Correlation:

$$f = \frac{\sum(ij)C(i,j)-\mu_i\mu_j}{\sigma_i\sigma_j} \tag{3}$$

Homogeneity:

$$f = \frac{C_{ij}}{(1+|i-j|)} \tag{4}$$

Entropy:

$$f = -\sum C_{ij}\log(C_{ij} + 2^{-30}) \tag{5}$$

Among them, in Eq. (3), $\mu_i = \sum iC_{ij}$, $\mu_j = \sum jC_{ij}$, $\sigma_i^2 = \sum(i-\mu_i)^2 C_{ij}$, $\sigma_j^2 = \sum\left(j-\mu_j\right)^2 C_{ij}$. The gray level co-occurrence matrix was calculated based on grayscale levels of 16, 32, and 64; distances between pixel pairs of 2, 4, and 6; angles of 0°, 45°, 90° and 135°; window widths of Grayscale co-occurrence matrix extraction at 5 pixels, 9 pixels, 15 pixels, 21 pixels.

We annotated the pixels in the ROI region in several images and extracted the features. A total of 3 types of labels were marked out: pixels in the calcification area, pixels in the lipid area, pixels in the fibrous tissue area. A total of 628400 252-dimensional vectors were obtained.

2.4 Feature processing and SVM training

First the missing or damaged part in the data acquired from Section 2.2 was wiped out. Since SVM classifier is scale-variant, a linear scale change was made for each dimension feature of the data so that the features of each dimension were roughly scaled to the range of (0, 1). The ratio of the linear scale change was calculated from the training data and used directly in the test data.

Then, the feature data were selected by sorting according to the Random forest feature selection and relevance scores between features and training labels. The filtered data was entered into the SVM class for training. Noted that SVM classifier is equipped with 'rbf' kernel function for which the parameter is $\gamma = 0.7$. We set up SVMs with different proportions of features as control groups to perform the experiments. In addition, the data set was divided into training sets and test sets in different proportions for subsequent analysis.

Classification accuracy was computed within the manually drawn ROI by comparing the ground-truth pixels to the output result of SVM classifier. Accuracy was computed as the number of pixels signed to the correct class label divided by the total number of pixels.

3 Results

Part of the statistics of texture features collected from the ROIs of the images is shown in Tab. 1.

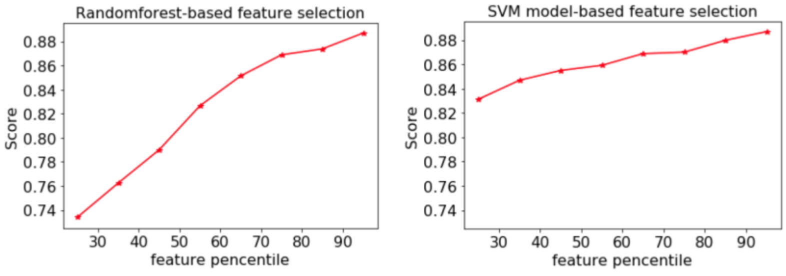
According to the random forest selecting score, a certain percentage, i.e. from 20% to 90%, of the top-scored features were taken as input of SVM classifier. Fig. 2(a) lists the test accuracies of SVM classifiers trained with different numbers of features. It can be seen that an increase in the number of features leads to an increase in the classification score. However, the score improved slowly when more than 70% features were added in, yet the calculation was time-consuming. Apart from this, to some extent, Fig. 2(a) implies that Random forest selection did not work well in estimating the correlation between the features and the target label. Relatively, the scores achieved with feature selection based on trained SVM weight is shown in Fig. 2(b).

Table 1: Feature naming rules: f*i*d*w*. i16 represents a greyscale level of 16; d2 indicates that the distance between pixel pairs is 2; w5 refers to a window size of 5×5; f* indicates the feature type: f1: Contrast, f2: Correlation, f3: Homogeneity, f4: Entropy, f5: Energy, f6: cluster shadow; Atten: attenuation coefficient. In this table, ‘f2i16d2w5’, ‘f3i16d2w5’, ‘f4i16d2w5’, ‘f5i16d2w5’, ‘f6i16d2w5’ implies ‘Correlation’, ‘Homogeneity’, ‘Entropy’, ‘Energy’, ‘Cluster shadow’ calculated from GLCM. The GLCM parameters are specified as: greyscale level=16, pixel distance offset=2, window size=5×5

	Atten	f2i16d2	f3i16d2	f4i16d2	f5i16d2	f6i16d2	…
count	628400	628400	628400	628400	628400	628400	…
mean	19.0368	0.22185	0.02882	0.35078	64.5151	3.67943	…
std	5.76380	0.21191	0.00809	0.04698	65.1912	0.16436	…
min	3.87244	-	0.02015	0.24229	-	1.70274	…
max	42.8254	0.90898	0.29154	0.75583	582.061	3.94259	…
	…						
				7 rows×504 columns			

According to Fig. 2(b), the first 151 dimensions feature were taken as input of SVM. This balanced the trade-off between 4% accuracy and the amount of calculations, reduced the test time from about 6 s to about 3 s per 100 pixels.

To exam how much we benefit from adding more training data and the capability of SVM classifier, we used different sizes of training set to train the SVM. With a constant number of training steps, the training and validation score increases with the gradual expansion of training data. The training score and validation score were close when taking 80% of dataset as the training set as Fig. 3 shows.



(a) (b)

Figure 2: The prediction results achieved by including different ratios of parameters that were selected based on Random forest estimation (a) and SVM model parameters (b) respectively

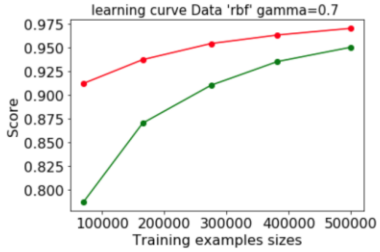


Figure 3: A cross-validation on the prediction results over splitting the training set and the validation set in different proportions

4 Discussion

In general, support vector machines can select a few training samples that are most effective for predicting tasks in high dimensional data, thus saving the required memory in the model learning process and improving the prediction performance of the model. A support vector machine with kernel function could obtain the optimal classifier in the determined data set. However, the support vector machine also consumes more computing resources and time when the input data dimension is very high. The computing devices utilized in the experiment: Feature extraction and classifier training were implemented on inter core7 CPU with Python multiprocessing. The features acquisition in a 977×977 image (ROI area only) costs about 10 min; the training process takes 25 min 40 s when counting in 252-dimensional feature entry; point-by-point prediction in a single-image (depends on the ROI) takes about 2-3 min.

As Fig. 4 implies, in most of the images acquired, there are much more pixels belong to fibrous class than lipid or calcification, which means there are more samples of fibrous class than the rest of two. In this case, we managed to keep the amount of the samples belonging to each class in balance during training. Apart from this, Fig. 4(d) shows that some pixels belonging to calcification class were misclassified to fibrous class, which is probably because of the inhomogeneity and relatively low brightness display of calcified plaque area. Fig. 4(e) shows an example of lipid pool with bright fibrous cap. Lipid plaques are characterized by their low brightness and low contrast, which makes it easy to be distinguished from other plaques once the ROI is determined. Most of the misclassified pixels of lipid class are located on the border of lipid and fibrous tissues. Similar to this, the accuracy appears low at the location of different tissue boundaries or complex plaque components area. This relatively poor performance may be caused by the lack of data diversity, which is the limitation of this study. However, the advantage of this method is based on per pixel classification, which means some part of the image, e.g. the border of the fibrous and the lipid tissue, could be fixed utilizing morphological based image processing method and other image features.

To apply the SVM classifier trained with 151 features in test set's image ROI we achieved an accuracy of 83%. Single class accuracy of fibrous, calcification and lipid tissue are 89%, 79.3% and 86.5% respectively. As an example, Fig. 4 shows the predict result of two individual frames. Figs. 4(a), 4(e) are original image in Cartesian coordination. Figs. 4(b) and 4(f) are the images with segmentation results overlays of Figs. 4(a) and 4(e), respectively. Figs. 4(c), 4(d) are the actual label results and predictions of 4(a). Figs. 4(g), 4(h) are the true label results and predictions of 4(e). Figs. 5(a) and 5(b) refer to the confusion matrix and the normalized confusion matrix [Powers (2011)] for the prediction.

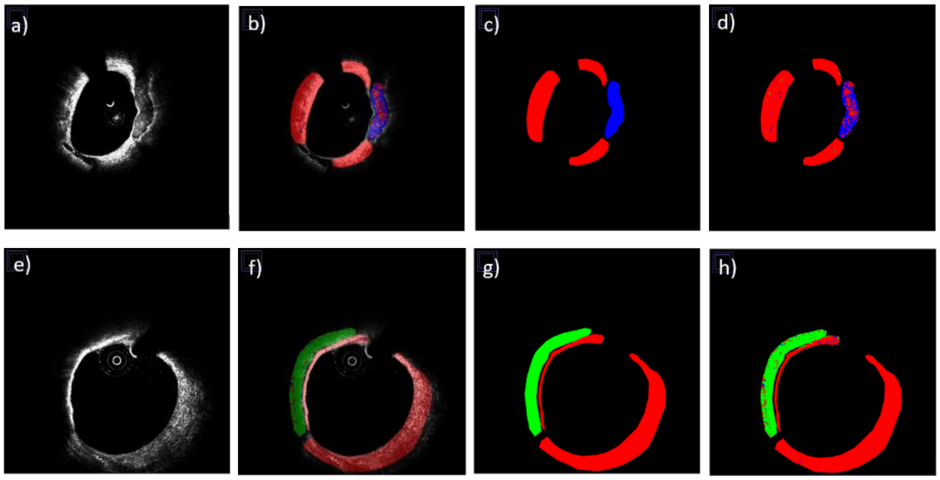
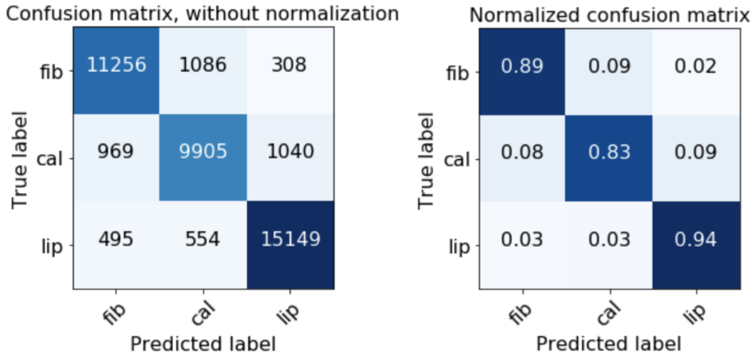


Figure 4: (a), (e) are two different original IVOCT frames; (b) is the image with segmentation result overlay of (a); (c) and (d) are the manual segmentation result and automatic segmentation result of (a); similarly, (f), (g), (h) are the overlay image, manual result and automatic segmentation result of (e). Red, blue, green parts indicate fibrous, calcification and lipid tissues respectively



Confusion matrix, without normalization

True label \ Predicted label	fib	cal	lip
fib	11256	1086	308
cal	969	9905	1040
lip	495	554	15149

Normalized confusion matrix

True label \ Predicted label	fib	cal	lip
fib	0.89	0.09	0.02
cal	0.08	0.83	0.09
lip	0.03	0.03	0.94

(a) (b)

Figure 5: (a) confusion matrix; (b) normalized confusion matrix

With an increasing amount of training set to be added into the developed algorithm, the features can be improved significantly in the future. This method may be applied as a useful tool for plaque segmentation based on OCT imaging.

Acknowledgement: This work was partially supported by the Natural Science Foundation of China (NSFC) (Nos. 11422222, 11772093), the National 973 Basic Research Program of China (No. 2013CB733800) and ARC (FT140101152).

References

Athanasiou, L. S.; Bourantas, C. V.; Rigas, G.; Sakellarios, A. I.; Exarchos, T. P. et al. (2014): Methodology for fully automated segmentation and plaque characterization in intracoronary optical coherence tomography images. *Journal of Biomedical Optics*, vol. 19, no. 2.

Bezerra, H. G.; Costa, M. A.; Guagliumi, G.; Rollins, A. M.; Simon, D. I. (2009): Intracoronary optical coherence tomography: A comprehensive review clinical and research applications. *JACC Cardiovascular Interventions*, vol. 2, no. 11, pp. 1035.

He, C.; Wang, J.; Zhu, T.; Li, Z.; Miao, Y. (2016): The correlation between texture features and fibrous cap thickness of lipid-rich atheroma based on optical coherence tomography imaging. *Molecular & Cellular Biomechanics*, vol. 13, no. 1, pp. 23-36.

Howard, R. A. (1966): Dynamic programming. *Management Science*, vol. 12, no. 5, pp. 317-348.

Powers, D. M. W. (2011): Evaluation: From precision, recall and *f*-factor to roc, informedness, markedness & correlation. *Journal of Machine Learning Technologies*, vol. 2, pp. 2229-3981.

Ping, Z.; Zhu, T.; Li, Z. (2015): Segmentation and classification method in IVOCT images. *Joint International Mechanical, Electronic and Information Technology Conference*, pp. 327-330.

Soest, G. V.; Goderie, T. P. M.; Noorden, S. V.; Steen, A. F. W. V. D. (2009): Algorithm optimization for quantitative analysis of intravascular optical coherence tomography data. *Proceedings of SPIE 7161, Photonic Therapeutics and Diagnostics V*.

Tearney, G. J.; Regar, E.; Akasaka, T.; Adriaenssens, T.; Barlis, P. et al. (2012): 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*, vol. 59, no. 12, pp. 1058-1072.

Tsantis, S.; Kagadis, G. C.; Katsanos, K.; Karnabatidis, D.; Bourantas, G. et al. (2012): Automatic vessel lumen segmentation and stent strut detection in intravascular optical coherence tomography. *Medical Physics*, vol. 39, no. 1, pp. 503.

Ughi, G. J.; Adriaenssens, T.; Onsea, K.; Kayaert, P.; Dubois, C. et al. (2012): Automatic segmentation of *in-vivo* intra-coronary optical coherence tomography images to assess stent strut apposition and coverage. *International Journal of Cardiovascular Imaging*, vol. 28, no. 2, pp. 229-241.

Van, S. G.; Goderie, T.; Regar, E.; Koljenović, S.; van Leenders, G. L. et al. (2010): Atherosclerotic tissue characterization *in vivo* by optical coherence tomography attenuation imaging. *Journal of Biomedical Optics*, vol. 15, no. 1.

Wang, Z.; Kyono, H.; Bezerra, H. G.; Wang, H.; Gargesha, M. et al. (2010): Semiautomatic segmentation and quantification of calcified plaques in intracoronary optical coherence tomography images. *Journal of Biomedical Optics*, vol. 15, no. 6.