Automatic Image Classification in Intravascular Optical Coherence Tomography Images

Mengdi Xu, Jun Cheng, and Damon Wing Kee Wong
Institute for Infocomm Research
A*STAR, Singapore
Email: xumd@i2r.a-star.edu.sg

Akira Taruya and Atsushi Tanaka
Cardiovascular Division
Wakayama Medical University, Wakayama, Japan

Jiang Liu
Ningbo Institute of Material Technology and Engineering
China

Nicolas Foin and Philip Wong
National Heart Centre
Singapore

Abstract—Vulnerable plaque detection to identify plaque is important in coronary heart disease diagnosis. Currently, it is conducted through manual reading of intravascular optical coherence tomography (IVOCT) images by an interventional cardiologist. However, human reading and understanding is highly subjective. An objective and automated assessment of plaque status is highly needed. This paper proposes a method for automatic image classification in IVOCT images based on different lesion types. In the proposed method, we first use detail-preserving anisotropic diffusion to remove speckle noise in IVOCT images. It removes the noise without losing details. Then, the IVOCT images are transformed to polar coordinates for feature extraction. In particular, Fisher vector and other texture features including local binary pattern and histogram of oriented gradients are studied. Finally, a support vector machine classifier is obtained to classify the IVOCT images into five groups: Normal (normal), FP (fibrous plaque), FA (fibroatheroma), PR (plaque rupture), and FC (fibrocalcific plaque). These five groups are obtained according to lesion characteristics. We evaluate the proposed method in a dataset of 1,000 images with five groups. Experimental results show that the proposed method achieves an average accuracy of 90% in image classification. The proposed automatic IVOCT image classification method can be used to save time and cost of cardiologist.

I. INTRODUCTION

As the most common type of heart disease, coronary heart disease is also the most cause of heart attacks. The plaques building up along the inner walls of the heart arteries, thus the heart arteries are narrowed and blood flow to the heart are reduced. This process causes heart disease. There is growing interest in the possibility that the progress made against coronary heart disease could be enhanced by identifying vulnerable plaques and vulnerable patients.

Intravascular ultrasound system (IVUS) has been used for intracoronary plaque assessment. However, the resolution of IVUS is relatively low and some details cannot be recognized. Recently, intravascular optical coherence tomography (IVOCT) [1] has been emerged as a new promising intravascular diagnostic tool. The acquisition process of IVOCT has already been proved to be safe, effective, and highly reproducible. The resolution of IVOCT is around 15um, while the resolution of IVUS is around 150 um. The higher resolution of IVOCT would help provide a level of detail never reached before. It has been shown that it would be possible for IVOCT to characterize the appearance of vulnerable plaques [1]. Currently, manual reading of IVOCT images is conducted by interventional cardiologist performing the cardiac IVOCT imaging. The limitation of manual assessment is that it is subjective and time-consuming. Therefore, an objective and automated assessment of plaque status is needed.

As IVOCT is a relative new modality, there exists only a few research works on plaque or lesion study in IVOCT images. Athanasious et al. [2] proposed a plaque characterization algorithm. They manually selected the plaque area (Region of Interest) and classified the plaque area into four plaque types. The proposed method is semi-automated. Later, they proposed an automatic method for plaque analysis [3]. But only calcified plaque region is detected in the proposed work. We introduced an automatic system which detects whether the given IVOCT...
The framework of the proposed automated system is shown in Fig. 2. The proposed classification system includes noise reduction, guide-wire artifact detection, feature extraction, and image classification to classify the input IVOCT image into five groups by utilizing linear SVM model.

**Fig. 2.** Framework of the proposed automated system. The proposed classification system includes noise reduction, guide-wire artifact detection, feature extraction, and image classification to classify the input IVOCT image into five groups by utilizing linear SVM model.

In this work, the IVOCT images are classified into five groups according to the lesion characteristics [6]. These five groups are: Normal (normal), FP (fibrous plaque), FA (fibroatheroma), PR (plaque rupture), and FC (fibrocalcific plaque). We only consider the images with single plaque type. The detection of TCFA (thin-cap fibroatheroma) is important for vulnerable plaque identification. Instead of identifying the TCFA only, we propose an automated method which accurately classifies the IVOCT images into five groups. The identification of TCFA from FA group will be studied in the future.

In this work, the IVOCT images are classified into five groups according to the lesion characteristics [6]. These five groups are: Normal (normal), FP (fibrous plaque), FA (fibroatheroma), PR (plaque rupture), and FC (fibrocalcific plaque). We only consider the images with single plaque type. The detection of TCFA (thin-cap fibroatheroma) is important for vulnerable plaque identification. Instead of identifying the TCFA only, we propose an automated method which accurately classifies the IVOCT images into five groups. The identification of TCFA from FA group will be studied in the future. Fisher vector and other texture features are extracted and compared in this paper. Our system includes four steps: noise reduction, guide-wire artifact detection to remove catheter and artifacts, feature extraction, and image classification to classify IVOCT images into five groups. In order to evaluated the system, we conduct four-fold cross-validation on a dataset with 1,000 annotated IVOCT images. There are 200 images for each group in the studied dataset. Some example images in our dataset are shown in Fig. 1. These IVOCT images belong to group Normal, FP, FA, PR, and FC. It can be seen that IVOCT could help in imaging the plaque details clearly. Quantitative evaluation is also studied and the results show that the proposed system is stable and accurate. To the best of our knowledge, this is the first work that automatically classifies the IVOCT images into five groups.

The organization of this paper is as follows. We introduce methodology in Section 2. Experimental results are shown in Section 3. Discussion and conclusions are presented in Section 4.

**II. METHODOLOGY**

In this section, we will introduce technical details of the proposed method. Fig. 2 demonstrates the framework of the proposed method. There are four main steps: a) noise reduction, b) guide-wire artifact detection, c) feature extraction, and d) image classification to classify the input unknown IVOCT image into five groups.

**A. Noise Reduction**

Speckle noise which occurred during IVOCT imaging process is difficult. It is generally serious, causing difficulties for image interpretation and feature extraction, thus should be removed first. Detail preserving anisotropic diffusion (DPAD) has shown to be promising as it removes the noise without losing the details [7]. In this paper, this algorithm is applied remove noise from IVOCT images.

**B. Guide-wire Artifact Detection**

To acquire IVOCT images, the catheter need to be inserted into the blood vessel. And the catheter is shown as bright circles in the middle of the image. Meanwhile, there is a guide-wire artifact which caused by guide-wire. Fig. 3 (a) shows the catheter and guide-wire artifact. The red circle in the middle of the IVOCT image shows the catheter, the red lines show the boundary of the guide-wire artifact. Because the guide-wire artifact appears at different locations, thus should be removed before feature extraction.

In the process, we first convert the image into polar coordinates, and then the catheter circle is transferred to line afterward. We detect and remove the catheter using Hough transform [8]. To remove the guide-wire artifact, we slide a box with fixed size over the polar image from left to right and...
compute the mean intensity within the box region (Fig. 3(c)). The location where the smallest mean intensity value appears is guide-wire artifact. The height of the sliding window is set as the height of the image in polar coordinates with the catheter removed and the width is set to be 30 empirically.

C. Feature Extraction

This part introduces the features used in the proposed system. The texture features are extracted on the polar coordinates images after preprocess (Fig 2 Guide-wire artifact detection). Since the appearance of the IVOCT images from five groups is different, we use texture features to represent the images. Fisher vector [9] is extracted for IVOCT image classification. It is applied because it adds additional information that describes the distribution of the descriptors. We also studied intensity, LBP, and HOG for feature comparison. The classification results of using Fisher vector, intensity, LBP, HOG, and the combination of these features are reported.

Fisher vector is an efficient feature used in image classification, which could generates high classification accuracy. It is shown to extend the bag-of-visual-words (BOV) by encoding additional information which describes the distribution of the descriptors. Fisher vector uses Gaussian Mixture Model (GMM) to construct visual word dictionary. In this work, we also compute the gradient of the Gaussian, $\gamma_t(i)$, to Gaussian $i$:

$$g_{u,i}^X = \frac{1}{T \sqrt{w_i}} \sum_{t=1}^{T} \gamma_t(i) \left( \frac{x_t - u_i}{\sigma_i} \right)$$

$$g_{\sigma,i}^X = \frac{1}{T \sqrt{2w_i}} \sum_{t=1}^{T} \gamma_t(i) \left[ \frac{(x_t - u_i)^2}{\sigma_i^2} - 1 \right]$$

where $\lambda$ is the soft assignment of $x_t$ to Gaussian $i$: $\gamma_t(i) = \frac{w_i u_i(x_t)}{\sum_{j=1}^{K} w_j u_j(x_t)}$. The concatenation of the $g_{u,i}^X$ and $g_{\sigma,i}^X$ is the final Fisher vector $g^X$. In this experiment, GMMs with $K = 256$ Gaussians is used to compute Fisher vectors. The GMMs are trained using Expectation-Maximization (EM) algorithm.

In [4], intensity features, Histograms of Oriented Gradients (HOG) [11], and Local Binary Patterns (LBP) [10] bag-of-visual-words (BOV) [12] features are extracted to represent the IVOCT images. In this work, we also extract these features and combine them with Fisher vector. These features are described below:

- Intensity: Intensity is calculated by resizing the IVOCT images to $32 \times 32$ pixels directly. In order to reduce computational cost, we apply principal component analysis (PCA) afterward to reduce the dimensionality to 64-D.
- LBP BOV: Bag-of-visual-words is to represent the image using the frequency of the words. First, we select IVOCT images, which are not used in the following evaluation experiment, as codebook images to construct codebook (words) of BOV. After that, we extract LBP features of each image, and apply K-means clustering method to get the codebook centres. In this experiment, we set K as 32 empirically.
- HOG BOV: The same as LBP BOV, we compute HOG features of each image and represent it using BOV. The HOG BOV centres are computed by using K-means clustering method. In this experiment, we set K as 32 empirically.

D. Image Classification

Considering the efficiency of classification scheme, we employ Linear Support Vector Machine (SVM) [13] to classify these five groups. In this experiment, LIBLINEAR package [14] is used for classification.

III. RESULTS

A. Dataset Construction

In this experiment, we collect 1,000 IVOCT frames from 47 patients to construct our dataset. There are 200 images from each group in our dataset. The images are annotated by a cardiologist manually. These images were acquired using
TABLE I
MEAN ACCURACY AND STANDARD DEVIATION OF THE EXPERIMENT. HOG AND LBP ARE BAG-OF-VISUAL-WORDS FEATURES.

<table>
<thead>
<tr>
<th></th>
<th>Mean Accuracy</th>
<th>Standard Deviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intensity</td>
<td>77%</td>
<td>0.01</td>
</tr>
<tr>
<td>LBP</td>
<td>55%</td>
<td>0.04</td>
</tr>
<tr>
<td>HOG</td>
<td>56%</td>
<td>0.02</td>
</tr>
<tr>
<td>FP</td>
<td>83%</td>
<td>0.04</td>
</tr>
<tr>
<td>FA</td>
<td>82%</td>
<td>0.01</td>
</tr>
<tr>
<td>Intensity+LBP+HOG [4]</td>
<td>89.0%</td>
<td>0.03</td>
</tr>
<tr>
<td>FP</td>
<td>95.0%</td>
<td>0.01</td>
</tr>
<tr>
<td>FA</td>
<td>84.0%</td>
<td>0.03</td>
</tr>
<tr>
<td>FP</td>
<td>82%</td>
<td>0.01</td>
</tr>
<tr>
<td>Intensity+LBP+HOG+FV</td>
<td>90%</td>
<td>0.02</td>
</tr>
</tbody>
</table>

TERUMO LUNAWAVE, which is an optical frequency domain imaging (OFDI) equipment. The original image size is 800 × 800 pixels. In this experiment, the images are resized to 256 × 256 pixels to reduce computational cost.

B. Experimental Results

In this work, we propose a four-fold cross-validation method to evaluate our proposed system. We divide the dataset into four subsets. These subsets are divided randomly and have equal size. Considering that there are 200 images for each group, therefore each subset includes 50 IVOCT images from each group. We run the classification process four times and get the mean accuracy value and standard deviation of the four accuracy values. For each time, three subsets are used as training sets, and the rest subset is used as testing set.

Table I shows the mean accuracy value and standard deviation. Here HOG and LBP are bag-of-visual-words features (BOV). It can be seen from Table I that, Fisher vector outperforms other single features, and has comparable performance with the feature combinations using in [4]. Although the mean accuracy of Fisher vector is slightly lower than [4], the standard deviation is also lower, which indicates that Fisher vector is more stable. By integrating Fisher vector with intensity, LBP, and HOG features, the system achieves the highest accuracy value of 90%. The small standard deviations of the studied features indicating that our system is stable.

Table II shows the mean accuracy of the five groups respectively. As shown in Table II, the classification accuracy for Normal and FP are higher than the other three groups, it may be explained by the fact that Normal and FP share the similar texture, while for other three groups, the texture variation within each group is relatively larger.

Table II
MEAN ACCURACY ON EACH GROUP.

<table>
<thead>
<tr>
<th></th>
<th>Normal</th>
<th>FP</th>
<th>FA</th>
<th>PR</th>
<th>FC</th>
</tr>
</thead>
<tbody>
<tr>
<td>I+LBP+HOG [4]</td>
<td>84.0%</td>
<td>91.0%</td>
<td>83.5%</td>
<td>82.5%</td>
<td>73.0%</td>
</tr>
<tr>
<td>FV</td>
<td>90.5%</td>
<td>89.5%</td>
<td>80.0%</td>
<td>80.0%</td>
<td>71.0%</td>
</tr>
<tr>
<td>I+LBP+HOG+FV</td>
<td>93.5%</td>
<td>95.0%</td>
<td>85.5%</td>
<td>89.0%</td>
<td>84.5%</td>
</tr>
</tbody>
</table>

IV. DISCUSSIONS AND CONCLUSIONS

In this work, we propose an automatic image classification system to classify the IVOCT images into five groups. These five groups are classified according to lesion characteristics. Fisher vector, intensity feature, LBP, and HOG features are investigated in this work. We conduct a four-fold cross-validation process to evaluate the system, and achieves mean accuracy of 90%. The low standard deviation value shows that the proposed system is stable. To the best of our knowledge, the proposed system is the first automated system which classifies the five groups in IVOCT images. In the future, our method would be used to characterize plaque pathology during intervention and provide important information to the assessment of coronary artery disease. Algorithms for TCFA detection will be studied in the future.

REFERENCES