AUTOMATED QUANTITATIVE ASSESSMENT OF CORONARY CALCIFICATION USING INTRAVASCULAR ULTRASOUND

SHENGNAN LIU,* TARA NELEMAN,* ELINE M.J. HARTMAN,* JURGEN M.R. LIGTHART,* KAREN T. WITBERG,* ANTONIUS F.W. VAN DER STEEN,*+ JOLANDA J. WENTZEL,* JOOST DAEMEN,* and GIJS VAN SOEST*

* Department of Cardiology, Erasmus University Medical Center, Rotterdam, The Netherlands; y Department of Imaging Physics, Faculty of Applied Sciences, Delft University of Technology, The Netherlands; and z Shenzhen Institutes of Advanced Technologies, Shenzhen, China

Abstract—Coronary calcification represents a challenge in the treatment of coronary artery disease by stent placement. It negatively affects stent expansion and has been related to future adverse cardiac events. Intravascular ultrasound (IVUS) is known for its high sensitivity in detecting coronary calcification. At present, automated quantification of calcium as detected by IVUS is not available. For this reason, we developed and validated an optimized framework for accurate automated detection and quantification of calcified plaque in coronary atherosclerosis as seen by IVUS. Calcified lesions were detected by training a supported vector classifier per IVUS A-line on manually annotated IVUS images, followed by post-processing using regional information. We applied our framework to 35 IVUS pullbacks from each of the three commonly used IVUS systems. Cross-validation accuracy for each system was >0.9, and the testing accuracy was 0.87, 0.89 and 0.89 for the three systems. Using the detection result, we propose an IVUS calcium score, based on the fraction of calcium-positive A-lines in a pullback segment, to quantify the extent of calcified plaque. The high accuracy of the proposed classifier suggests that it may provide a robust and accurate tool to assess the presence and amount of coronary calcification and, thus, may play a role in image-guided coronary interventions. (E-mail: g.vansoest@erasmusmc.nl) © 2020 The Author(s). Published by Elsevier Inc. on behalf of World Federation for Ultrasound in Medicine & Biology. This is an open access article under the CC BY license. (http://creativecommons.org/licenses/by/4.0/).

Key Words: Calcified plaque, Intravascular ultrasound, Automated quantification, Intravascular imaging, Coronary artery disease.

INTRODUCTION

Coronary artery disease, the most common heart disease, is caused by a long-term accumulation of atherosclerotic plaque in the intima of the arterial vessel wall (Falk 2006; Hong 2010). As plaques grow, they may impinge on the free coronary lumen, causing a stenosis that limits blood flow to the myocardial territory served by the coronary artery. Severe or acute coronary artery disease is frequently treated by stent implantation in a procedure called percutaneous coronary intervention (PCI). Atherosclerotic plaques are frequently heterogeneous in their composition, and typically consist of fibrous, lipid-rich and calcified tissue.

The presence of calcium, in particular, can hamper the feasibility of PCI (Hoffmann et al. 1998). Because of its rigidity, circumferential calcification may prevent full dilation of the stent, which may lead to stent underexpansion, associated with increased risk of target vessel failure (Witzenbichler et al. 2014). Specialized techniques, such as cutting balloons, rotational atherectomy or intracoronary lithotripsy (Sharma et al. 2019), can be used to prepare calcified plaques to enable complete stent expansion. The beneficial effects of these plaque modification techniques depend directly on the extent and severity of calcification, so detailed knowledge is needed to guide the choice of lesion preparation method (Wijns et al. 2015).

Large calcified plaques can often be unambiguously identified in intravascular ultrasound (IVUS) images because of their high reflection of and low penetration by ultrasound signals (Pu et al. 2014; Mintz and Guagliumi 2017). As illustrated in Figure 1, coronary calcium is characterized as a narrow band with high echo intensities, with a dark shadow behind it. Current IVUS software does not allow automated calcium quantification.
Because of the presence of inherent speckle, various system- and anatomy-dependent artifacts, limited resolution and image contrast, automated structure detection in IVUS remains a challenging task (Katouzian et al. 2012). Simple approaches, using thresholding (Kim et al. 2014) and adaptive thresholding, have been reported (Dos Santos Filho et al. 2007). A more complex approach to segmentation of the leading edge of calcified plaque combining Rayleigh mixture models, Markov random fields, graph searching and prior knowledge, has been reported (Gao et al. 2014). After detection of the intima and the media—adventitia borders, calcified regions were detected using a Bayesian classifier (Taki et al. 2008). Two recent studies have reported the detection of calcification per frame using a deep learning network (Balocco et al. 2018; Sofian et al. 2018).

In this work, we present a framework for accurate automated detection and quantification of the extent of calcification in coronary atherosclerosis as seen by IVUS. Without multistep pre-processing, extraction of complex features or design of a deep learning approach to perform the task, we found that accurate classification can be achieved by applying a kernel-based support vector classifier on simple statistical features, originating from the imaging physics, extracted directly from un-processed images. Using data from three commonly used IVUS systems, we trained the classifier for recognition of calcified plaques per A-line. Using the detection result, we propose an IVUS calcium score (ICS) to evaluate the calcified plaque load and compare it with the calcium score based on manual labels.

**METHODS**

**IVUS data**

IVUS pullback data sets, acquired in native coronary arteries, were extracted from the clinical database of the Department of Cardiology, Erasmus MC. The data sets were anonymized and contained no identifying information; all selected patients consented to the use of their data in retrospective studies. All data in this study were collected as part of routine clinical care. Consequently, institutional review board approval and individual patient consent are not required under Dutch law. We selected 105 pullbacks, 35 each from three commonly used systems: Infraredx (40 MHz, TVC NIRS Catheter System, Infraredx Inc., Burlington, MA, USA), Volcano (20 MHz, Eagle Eye Platinum Rx, and ST Rx, Digital IVUS Catheters, Volcano Corp., Rancho Cordova, CA, USA) and Boston Scientific (40 MHz, Atlatis SR Pro and OptiCross Coronary Imaging Catheter, Boston Scientific Corp., Natick, MA, USA). Example images from the three systems are provided in Figure 1.

An overview is given in Table 1. Pullbacks, stored in DICOM format, were manually annotated by trained experts (E.M.J.H. and T.N.) at 1-mm intervals. Centering at the imaging catheter, we divided and labeled each frame as a calcified or non-calcified pie sector (Fig. 2).

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Population</th>
<th>Pullback</th>
<th>Frame rate (fs/s)</th>
<th>Pullback speed (mm/s)</th>
<th>Training</th>
<th>Test</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infraredx</td>
<td>34</td>
<td>35</td>
<td>16</td>
<td>0.5</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>Volcano</td>
<td>35</td>
<td>35</td>
<td>30</td>
<td>0.5</td>
<td>31</td>
<td>4</td>
</tr>
<tr>
<td>Boston Scientific</td>
<td>35</td>
<td>35</td>
<td>30</td>
<td>0.5</td>
<td>31</td>
<td>4</td>
</tr>
</tbody>
</table>
Frames (in Cartesian coordinates) and their corresponding labels were converted into polar coordinates, such that each A-line was binary labeled as 1 (calcified) or 0 (non-calcified).

**Identifying features**

Calcified plaque can be confidently classified by human observers. The A-line profile itself is quite variable though, within each (calcified or non-calcified) category. By comparing the A-line amplitude statistics, we observed that the typical combination of a thin reflection and dorsal shadow of calcium in IVUS may be used as a discriminating characteristic, whereas an A-line scan of soft tissue contains many gray values and most amplitude values in a line sampling calcium are low (lumen and shadow), with a few very high ones (calcium border). To build an unbiased classifier, we decided to train it upon a rich feature set, including 10 distribution densities, 10 distribution quantiles and the mean value (21 in total; examples for calcified and non-calcified A-lines are illustrated in Fig. 3).

**Training the detection model**

We trained a radial basis function (RBF) support vector classification (SVC) model to classify IVUS A-lines. An SVC is a flexible structure used to classify high-dimensional data. A basic trained SVC is a linear hyperplane, which can separate only linearly separable clusters. To deal with data that are not linearly separable, a non-linear kernel needs to be introduced. When little is known about the structure of data, a Gaussian RBF kernel is a robust choice, assuming only general smoothness (Smola and Schölkopf 2004).

Each A-line is characterized by a set of statistical features $x_i$, and has a binary label $y_i$ identifying it as calcified or not. For $M$ labeled A-line distributions, \{($x_i$, $y_i$) | $y_i \in \{0, 1\}$, $i \in \{0, \ldots, M\}$\}, a classifier

$$f(x) = \text{sgn}\left[\sum_{i=1}^{n} \alpha_i y_i K(x_i, x) + b\right]$$

was trained to optimize the problem (Guyon et al. 1993)

$$\min_{\alpha} \frac{1}{2} \alpha^T Q \alpha - e^T \alpha \quad \text{such that } y^T \alpha = 0, \quad 0 \leq \alpha_i \leq C, \quad i = 1, \ldots, n$$

(1)

Here, $e$ is a vector of ones, and $C$ is a parameter balancing the complexity and training error to be tolerated. $Q$ is a $n \times n$ positive semidefinite matrix, and $Q_{ij} = y_i y_j K(x_i, x_j)$ and $K(x_i, x_j) = \exp(\|x_i - x_j\|^2 / 2\sigma^2)$ is the commonly used Gaussian RBF kernel. $\sigma$ governs the linearity of the classifier: larger $\sigma$ values allow greater non-linearity. The hyperparameters $C$ and $\sigma$ were determined by an exhaustive grid search optimizing overall classification performance.
Fig. 3. Features extracted from image intensities, depicting an example of a calcified A-line (blue) and a non-calcified A-line (red) from the manually annotated set. The mean value and the distribution densities are given in the left panel, and the quantiles, in the right panel.

Fig. 4. Flowchart for training the support vector machine. The numbered steps are explained in the text under Training the Detection Model.

Table 2. Overview of performance of support vector machine

<table>
<thead>
<tr>
<th>Vendor</th>
<th>Experiment</th>
<th>Measurement</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Accuracy*</td>
</tr>
<tr>
<td>Infraredx</td>
<td>Training</td>
<td>0.9170 ± 0.0003</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>0.87</td>
</tr>
<tr>
<td>Volcano</td>
<td>Training</td>
<td>0.9113 ± 0.0002</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>0.89</td>
</tr>
<tr>
<td>Boston Scientific</td>
<td>Training</td>
<td>0.9084 ± 0.0004</td>
</tr>
<tr>
<td></td>
<td>Testing</td>
<td>0.89</td>
</tr>
</tbody>
</table>

* For training we report the cross-validation accuracy.
Figure 4 is the flowchart for model training and testing; the numbered steps are as follows:

Step 1. The pullbacks in polar coordinates were randomly split 9:1; 90% were used for training and 10% were used for testing. The splitting numbers of pullbacks for each system are given in Table 2.

Step 2. After generating the training and testing data set, we extracted the image gray values for each A-line in the labeled frames.

Step 3. The distribution densities were calculated by generating a histogram of the image gray values in [0; 255] in 10 bins. They were normalized such that the sum is 1. The 10 quantiles were linearly sampled in the range 0.5%–99.5%.

Step 4. Because there were more non-calcified than calcified A-lines, we balanced the training categories by randomly downsampling the non-calcified A-lines such that the total number of non-calcified A-lines was equal to that of calcified A-lines.

Step 5. The input was further normalized using Z-score normalization. The estimated mean (μ) and the standard deviation (σ) were derived from the training set and later applied to Z-score normalize the testing set.

Step 6. Hyperparameters C and σ were selected. Grid points were evenly chosen on a double log scale, \( \{C_i = 10^b \mid u_i \in [-3, 3], i = 1, \ldots, M \} \times \{ \sigma_j = 10^g \mid \eta_j \in [-5, 3], j = 1, \ldots, N \} \); \( M = N = 21 \). The A-lines were randomized in the ratio 3:7 to training and validation sets for threefold cross-validation. The parameters with the highest accuracy were chosen for use in the final model. We find that there is a large range with nearly optimal performance for \( \sigma \in [10^{-4}, 10^{-1}] \), approximately, with minimal effect of the value of \( C \), indicating limited sensitivity to algorithm or data specifics.

Step 7. The final trained model was applied to the testing data. Precision, recall and the F1 score, given in eqn (2), were computed on balanced data where the negative examples were downsampled to be equal to the amount of positive examples.

Post-processing

After performing the classification, we applied the dense fully connected conditional random field (CRF) as the post-processing for noise removal (Krähenbühl and Koltun 2011). The method applies a Gaussian penalty when two pixels in the defined neighborhood have different labels. In combination of prior probability, the a posteriori probability was maximized with an optimized labeling solution. Here the prior probability was estimated for each pullback using the SVC-detected label, and the neighborhood was empirically chosen to be 10 frames and 21 A-lines. Followed by an in-frame morphological closing \((1 \times 21)\) and opening \((1 \times 51)\), the post-processing steps remove isolated positive and negative classifications and integrate labels with small gaps.

Validation and reporting

The model was trained for each system separately. Performance in the testing set was evaluated in the number of true/false positive/negative (TP, TN, FP, FN) A-line classifications, and reported in precision, recall and F1 score (Fawcett 2006):

\[
\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}
\]
\[
\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}
\]
\[
\text{F1 score} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}
\]

For further validation, we introduced the ICS, which is defined as the fraction of detected calcified A-lines in the total acquired number. Two ICSs were calculated using labeled frames; one was estimated using the manual labels (denoted as \( S \)) and the other was estimated using the detection results (denoted as \( \hat{S} \)):

\[
\hat{S} = \frac{\#(\text{detected calcified A-lines})}{\#(\text{all A-lines})}
\]

\[ S = \frac{\#(\text{labeled calcified A-lines})}{\#(\text{all A-lines})} \]

During the comparison we observed that \( S \) and \( \hat{S} \) are linearly related to each other. Therefore, we applied the random sample consensus (RANSAC) regression to highlight outliers and to fit a linear function \( \hat{S} = kS + b \) with inliers. Outliers were removed before computing Pearson’s correlation coefficient:

\[
r_{xy} = \frac{\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{\sum x_i^2 (\sum x_i)^2} \sqrt{\sum y_i^2 (\sum y_i)^2}}, \quad \text{for } i \in \{0, 1, \ldots, n\}
\]

We further adjusted the detected values by applying the linear transform \( f(\hat{S}) = (\hat{S} - b)/k \).

Using the estimation result on a whole pullback, we then calculated a total ICS (denoted as \( S_{\text{total}} \)), which can give an overall indication of the amount of calcified plaque in the whole pullback.

\[
S_{\text{total}} = \frac{\#(\text{detected calcified A-lines})}{\#(\text{all A-lines})}
\]
Furthermore, similar to intravascular near-infrared spectroscopy (Gardner et al. 2008), we present a local ICS (denoted as $\hat{s}_M$) in short segments, for instance, 2 mm, which can intuitively highlight artery sections with a heavy calcium burden:

$$\hat{s}_M = \frac{\# \text{(detected calcified A-lines)}}{\# \text{(all A-lines)}} \text{ in } M \text{ neighboring frames}$$

Again inspired by previous work on IVUS palpography (Schaar et al. 2004) and parametric intravascular optical coherence tomography (Gnanadesigan et al. 2016), we represent the detection result in a so-called “carpet view”, depicting the classification result in a display with dimensions of circumferential angle and pullback length.

**RESULTS**

Table 2 summarizes the overall performance of A-line-based calcium detection by the support vector machine (SVM) trained on the data. We observe that an average accuracy $>0.9$ was achieved, with small variations across validation experiments and similarly high in testing sets. The manual ICS $s$ and the detected ICS $\hat{s}$ are compared in scatterplots in Figure 5. Despite a moderate overestimation, the two numbers are highly correlated (Infraredx: $r = 0.94$, Volcano: $r = 0.88$, Boston Scientific: $r = 0.97$). The Wilcoxon tests suggest that the manual and automated measurements are sampled from the same distribution (Infraredx: $p = 0.6138$, Volcano: $p = 0.9426$, Boston Scientific: $p = 0.8370$).

A carpet view representation of the calcium detection results is provided in Figure 6. White areas indicate calcification. Vertical lines represent manually labeled frames, where red and blue designate calcium-positive and calcium-negative A-lines, respectively. The color bar above the carpet view displays the local ICS calculated every 2 mm. The local ICS provides an intuitive overview of the distribution of calcified plaques, with circumferential calcium ($s_{2\text{mm}} = 1000$) occurring in Figure 6a and b.

**DISCUSSION**

In the present study, we developed a pipeline for automated detection of calcified plaque on IVUS images. Using an SVC classifier and CRF post-processing, we attempted to use simple statistical features of image intensities for an A-line-based identification of calcium in the arterial wall. Results indicate that the proposed framework can be used for a robust estimation of an IVUS-based calcium score, which can be used as an objective evaluation of the presence and amount of calcified plaque in the vessel, overall and locally.

ICS overestimates calcium in specific situations

In total, 16 pullbacks were detected as outliers (Infraredx: 9, Volcano: 6, Boston Scientific: 1), 15 of which are overestimations comparing with the manual scores. All outliers were part of the training set. If the outliers had arisen in testing data, they could be indicative of overfitting. Rather, post hoc examination suggested that the overestimation occurred mainly in three scenarios.

First, some non-calcium image features cannot be distinguished from calcium features, based on A-line intensity statistics only. When the pericardial cavity is visible in the IVUS image, it appears as a thin bright band (visceral pericardium) followed by an abrupt dark cavity. This structure usually appears in large series of consecutive frames and can lead to massive overestimation. This was observed in three arteries and, in one case, was observed in 4264 of 5281 frames in one pullback. Occasionally, when large arteries are imaged with an eccentric catheter, leaving only a bright band of signals on the far wall, the statistical features can be similar to those from calcified regions. This was observed in five arteries.

---

Fig. 5. Scatterplot of ground-truth intravascular ultrasound calcium score (ICS) calculated using manual labels ($s$) and the ICS calculated using the detected labels ($\hat{s}$), acquired with the three different intravascular ultrasound systems: (a) Infraredx, (b) Volcano, (c) Boston Scientific. RANSAC = random sample consensus.
Second, post-processing is designed to remove small positive regions (which are likely to be false). False positives that are adjacent to a lesion are difficult to rule out, however. This may cause structures such as the guidewire to be classified as calcium.

Third, calcified regions appearing as a bright band, with a dorsal shadow, are relatively easy to identify for the experts. However, some calcified plaques exhibit normal brightness with dark shadows (in the case of a directional reflection from a non-normal surface not received by the transducer). The correct classification of this appearance of calcified plaque requires the observation of several neighboring frames, which were not available to the experts in this study. Compared with the experts’ labels, the detection framework performs more consistently for the detection of “dark” calcified lesions. However, in our reporting this is counted as an overestimation when compared with the experts’ labels.

Comparison with previous work

The present approach differs in a number of ways from recent work, which employed deep learning frameworks (Balocco et al. 2018; Sofian et al. 2018). First, we employed data from three different IVUS systems and developed a universally applicable analysis that differs only in the weights of the SVM. Second, our analysis did not require pre-processing (motion correction, gating, conversion to polar coordinates). This formulation of the calcium detection problem, which respects the independence of A-lines and relies on statistical features in the data that result directly from the imaging physics, outperforms the convolutional neural network-based classifier described by Balocco et al. (2018) as measured by the F1 score. Sofian et al. (2018) analyzed only isolated, selected frames, which are not necessarily representative of clinical data.

Outlook on application: Clinical research

For analysis of large intravascular imaging data sets, algorithmic quantification of plaque features can accelerate quantification studies by eliminating the time-consuming manual annotation of thousands of images, while simultaneously improving reproducibility and reducing inter-observer variation. Future work, including prospective studies, will be needed to evaluate the value of the ICS for stratification of the risk of follow-up events after the index PCI.

For asymptomatic populations, the relation between calcified atherosclerosis and cardiovascular events has been quantified in the coronary artery calcium score, an
important non-invasive diagnostic metric used to predict cardiac risk with computed tomography (CT) (Shah and Coulter 2012). CT detects calcified plaque with high accuracy, with a systematic overestimation of the plaque volume (Leber et al. 2006; Sun et al. 2008; Voros et al. 2011). CT lacks sensitivity to small calcifications compared with IVUS (Van Der Giessen et al. 2011). In our study, no such independent assessment of coronary calcium was available. As a result, the ICS has not been validated as a measure of the true amount of calcification in a vessel.

**Outlook on application: Intervention guidance**

Operator experience plays an important role in the practical use and utility of IVUS in PCI. Automated image analysis can be helpful for the development of objective and fast IVUS-guided PCI strategies. Many studies have found that imaging techniques can effectively reduce periprocedural (Witzenbichler et al. 2014; Zhang et al. 2018) and late ischemic events (Di Mario et al. 2018). Robust, real-time automated detection of plaque features can make the technology easier to use and thus more accessible.

Observational studies reported that calcified lesions can be associated with post-PCI adverse events, including restenosis and stent thrombosis (Witzenbichler et al. 2014). Moderate and severe calcifications were associated with major adverse events and revascularization in the target artery observed in the 3 years after the implantation of a stent (Shioide et al. 2018). Untreated calcified lesions may also trigger later adverse cardiac events (Mintz 2015). Exposed calcified nodules have been identified as a substrate for thrombus formation, leading to acute coronary syndrome (Virmani et al. 2000; Jia et al. 2013).

We introduced here a local measure of the calcium burden, by computing the ICS in short (2-mm) segments. In the future, the ICS may serve to indicate heavily calcified regions that may be likely to benefit from lesion preparation during PCI. A threshold value of the ICS for intervention guidance remains to be determined in follow-up research.

**Conclusions**

We presented an optimized framework for accurate automated detection and quantification of the presence and extent of calcification in coronary atherosclerosis as seen by IVUS. Calcified lesions were detected by training a support vector classifier per IVUS A-line on manually annotated pullback data, followed by post-processing using regional information.

With manual annotations as a standard, an overall accuracy of ~0:9 was achieved. Based on this classifier, we proposed an ICS that comprehensively characterizes the extent of coronary calcification in a vessel examined by IVUS.

**Conflict of interest disclosure**—The authors have no conflicts of interest to declare.

**References**


