Automatic Classification of Calcification in the Coronary Vessel Tree

Rahil Shahzad1,2,3, Lucas J. van Vliet2, Wiro J. Niessen2,3, and Theo van Walsum3

1 Division of Image Processing, Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
R.Shahzad@lumc.nl
2 Quantitative Imaging Group, Faculty of Applied Science, Delft University of Technology, Delft, The Netherlands
3 Biomedical Imaging Group Rotterdam, Department of Radiology, Erasmus Medical Center, Rotterdam, The Netherlands

Abstract. This paper presents a pipeline for complete automatic detection and labelling of coronary artery calcium lesions. The pipeline not only distinguishes between coronary artery calcifications and non-calcifications but also labels the detected calcium lesion to one of the main coronary vessel trees. The pipeline presented uses a combination of atlas-based and machine learning approaches. The set of features supplied to the classifier are either directly obtained from the subject scans or are derived after using information from atlas scans. The classifier was trained using 76 CT datasets. The pipeline was tuned using the training set provided by the challenge evaluation framework. On the 32 test datasets we achieved a sensitivity of 67% and a PPV of 86%, for the volume of the detected calcifications it was 89% and 94% respectively. The ICC was 1.00 for LAD+LM, 0.98 for LCX and 0.92 for RCA. On the whole heart the ICC for Agatston and volume score were 0.99 and 0.98 respectively. An accuracy of 88% and a linear weighted kappa (κ) of 0.90 was achieved for the risk categorization accuracy.

Keywords: Coronary artery, calcium scoring, atlas, CT, CTA

1 Introduction

Coronary artery disease (CAD) is one of the leading causes of mortality worldwide [1]. Many clinical studies have shown that the amount of calcium in coronary artery plaques correlates with the risk of future cardiovascular events [2, 3]. It has also been suggested that vessel-specific calcium scoring is more informative compared to whole heart calcium scoring [4]. Manual calcium scoring is a time-consuming task because it consists of drawing contours to obtain the region of interest or clicking inside all calcium objects. Lately, automating the task of calcium detection/scoring on CT scans has received a lot of attention. Isgum et al [5] and Kurkure et al [6] have
demonstrated the feasibility of automating the task for whole heart calcium scoring. Shahzad et al [7] have demonstrated the feasibility of automatic vessel-specific calcium scoring. A few other studies also perform detection/scoring of calcium lesions on Computed Tomography angiography (CTA) scans [8–11].

In this paper we present a pipeline that enables us to automatically detect vessel-specific calcium lesions, and evaluate it using the orCaScore evaluation framework. The pipeline is an adaptation of our earlier work [7]. The main change is an improved initialization of the atlas registration.

2 Method

The method used is similar to [7], with some modifications. The initialization of the atlas registrations was changed to be able to work with multi-vendor datasets. The method consists of the following steps: 1) Perform a multi-atlas based registration, 2) detect candidate calcium lesions from the CT scans, 3) compute a set of image features for the detected candidate lesions, 4) classify the candidate lesions either as a calcification or noise, and 5) assign the calcified lesions to one of the coronary vessel trees viz. left anterior descending + left main (LAD+LM), left circumflex (LCX) and right coronary artery (RCA).

2.1 Atlas Registration

Our method uses a multi-atlas based registration. The atlas registration has two purposes: 1) to segment the heart from the CTA scans [12], and 2) to compute a set of location based features for the classifier.

We use a set of eight atlas CTA scans. All the atlas scans have been acquired on the same machine (Somatom Sensation, Siemens Medical Solutions, Forchheim, Germany). Each of the atlases consists of a segmentation of the heart [13], a location estimate for each individual artery [14], and a standardized coordinate space [7].

ElastiX, a publicly available registration package [15] is used to register the atlas CTA scans to the subject’s CTA scan. Registration is performed in two stages using a multi-resolution strategy. An initial rigid registration is used to approximately align the two scans. This rigid registration is followed by a non-rigid registration. The rigid registration is performed on masked distance transform scans using normalized cross correlation as similarity measure. The non-rigid registration is performed on the original CTA scans in which a B-spline transformation model was employed that was initialized by the result of the rigid transformation, and mutual information is used as similarity measure. The resulting transformations from the registration steps are used to map the atlas heart segmentation, coronary artery location estimates and the standardized coordinate space onto the subject’s scan. See Figure 1 for an overview of the registration stage. Finally, the information from all eight atlases are combined (majority voting for the heart segmentation, averaging for coronary artery location estimates and standardized coordinate space) and mapped onto the subject’s CT scan using an additional intra-subject (CTA-CT) registration step.
Fig. 1. An overview of the registration stage, this stage is performed for each of the eight atlas scans.

2.2 Candidate Detection

Candidate calcium lesions are obtained from the CT scans by thresholding the segmented heart at 130 HU (See Figure 2). All objects with a size smaller than 1.5 mm\(^3\) are discarded, as these are assumed to be noise. The detected candidate lesions consist of true coronary calcifications, aortic calcifications, calcifications in the valves, and false objects due to noise and imaging artefacts.

2.3 Feature Computation

For each of the detected candidate lesions a set of 21 features are calculated (see Table 1 for the detailed list). The features are divided into subcategories:

*Object based features* - Three object based features are computed. 1) volume, 2) maximum intensity and, 3) average intensity of the candidate lesion.

*Multi-scale image derivatives* - The value of the Gaussian derivative upto the second order and at five different scales were computed at the point of maximum intensity per object. (Gaussian standard deviation in-slice between 0.3 mm to 4.8 mm with one sample per octave; between slices from 1.5 mm to 24 mm with one sample per octave) up to the second order. A total of 15 multi-scale image derivative features are computed.
Fig. 2. A random axial slice from one of the CT training dataset, the blue overlay shows the segmented heart, the red overlay represents the coronary artery density estimates (obtained after the atlas-based segmentations) and the green objects represents the detected candidate lesions.

Location estimates - The z-positions of the candidates in the actual image space and standardized coordinate are computed. The coronary artery location feature is also computed.

2.4 Classifier

A 9-nearest neighbour (NN) classifier is used to classify the candidate objects into coronary calcifications and the rest. The classifier was trained on 76 CT datasets that were obtained from a Siemens scanner (Definition, Siemens Medical Solutions, Forchheim, Germany). More details about the scans and the classifier training can be obtained from [7].

3 Results

Our method was applied to both the training and the test datasets which were provided by the orCaScore challenge organizers. Table 2 shows the results of our pipeline on the training datasets for each of the vendors.

On the 32 testing datasets we achieved an overall sensitivity of 67% and a positive predictive value (PPV) of 86%, with respect to the volume of the identified calcifications we were able to detect 89% and 94% respectively. The intraclass correlation coefficient (ICC) for Agatston score and volume score was
Table 1. Set of features used.

<table>
<thead>
<tr>
<th>Feature Description</th>
<th>Number of features</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volume of object in mm$^3$</td>
<td>1</td>
</tr>
<tr>
<td>Intensity of object: maximum and average</td>
<td>2</td>
</tr>
<tr>
<td>Position: z coordinate in the image space and mean space</td>
<td>2</td>
</tr>
<tr>
<td>Coronary artery location estimate</td>
<td>1</td>
</tr>
<tr>
<td>Intensity of voxel after Gaussian filtering at scale 1, 2, 4, 8</td>
<td>4</td>
</tr>
<tr>
<td>Intensity of voxel after 1$^{st}$ order Gaussian derivative in z direction at scale 1, 8, 16</td>
<td>3</td>
</tr>
<tr>
<td>Intensity of voxel after 2$^{nd}$ order Gaussian derivatives in xx, yy direction at scale 1, 2, 4</td>
<td>6</td>
</tr>
<tr>
<td>Intensity of voxel after 2$^{nd}$ order Gaussian derivatives in zz direction at scale 1, 2</td>
<td>2</td>
</tr>
</tbody>
</table>

Table 2. Performance of the pipeline on the training datasets.

<table>
<thead>
<tr>
<th>Vendor Id</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity (%)</td>
<td>83</td>
<td>80</td>
<td>68</td>
<td>68</td>
<td>75</td>
</tr>
<tr>
<td>Specificity (%)</td>
<td>99</td>
<td>99</td>
<td>98</td>
<td>99</td>
<td>99</td>
</tr>
<tr>
<td>PVV (%)</td>
<td>78</td>
<td>77</td>
<td>50</td>
<td>78</td>
<td>71</td>
</tr>
<tr>
<td>Volume TP Detected (%)</td>
<td>74</td>
<td>92</td>
<td>84</td>
<td>92</td>
<td>86</td>
</tr>
<tr>
<td>Avg no. FN</td>
<td>0.87</td>
<td>1.12</td>
<td>0.87</td>
<td>2.65</td>
<td>1.38</td>
</tr>
<tr>
<td>Avg no. FP</td>
<td>1.87</td>
<td>3.25</td>
<td>2.50</td>
<td>2.65</td>
<td>2.56</td>
</tr>
<tr>
<td>Avg FP volume mm$^3$</td>
<td>11.7</td>
<td>129.4</td>
<td>41.8</td>
<td>53.1</td>
<td>46.4</td>
</tr>
</tbody>
</table>

0.99 and 0.98 respectively. For vessel specific volume scores we obtained an ICC of 1.00 for LAD+LM, 0.98 for LCX and 0.92 for RCA. Comparing the risk category accuracy of our automatic method to the manual Agatston we obtain a correlation of 0.88 and a linearly weighted kappa ($\kappa$) value of 0.90.

4 Discussion and Conclusions

The results on the test datasets show a sensitivity of 67% and PPV of 86%, whereas 89% of the total volume of the calcium lesions has been detected. This shows that the lesions that were missed were small calcium lesions that could be easily mistaken as noise. The risk categorization accuracy is in very good agreement with the manually obtained ground truth. It can also be appreciated that our atlas-based method can be easily applied to data acquired by scanners from different vendors or with different scanning protocols without loosing performance. It can be noted from the training datasets that the performance
Fig. 3. The figures show a few axial slices from one of the test datasets after applying our pipeline. Pink overlay represents the structures with intensities above 130 HU, red objects are labelled as lesions belonging to LAD+LM, yellow to LCX and green to RCA on scans acquired on different scanners is similar. Also, from the training data it was observed that our method is sometimes prone to registration related errors, especially when the field of view (FOV) of the subject’s scan was very different from the atlases used. In the present pipeline we do not use an extensive atlas combination/selection strategy, but we simply average the labels together. This sometimes causes FP or FN detections. The method described was also evaluated on another dataset consisting of 56 datasets from a Siemens scanner [7]. On this dataset we achieved a sensitivity of 87%, and 93% of the subjects were assigned to the correct CVD risk category. We obtained vessel-specific correlations of 0.99, 0.93 and 0.93 for the LAD+LM, LCX and RCA volume scores respectively.

In the current method both CTA and CT scans were used. In case only CT scans are available for analysis, our method can easily be adapted to work only on the CT scans after some small modifications to the registration parameters.
The registrations and transformations are the most computationally expensive processes in our pipeline, they take approximately 10 minutes for each of the atlases, feature set computation and classification takes a couple of seconds on an Intel Xeon CPU E5-1620@3.60GHz with 16GB RAM. This process can be parallelized further by running the proposed pipeline on a computational cluster.

In conclusion, we presented an adapted version of our earlier calcium scoring method, which was based on a combination of atlas-based and machine learning approaches. The method was evaluated on the orCaScore framework, and obtained a PPV of 86%, an ICC of 0.99 and 0.98 for Agatston and volume score respectively, vessel-specific correlation of 1.00, 0.98 and 0.92 for LAD+LM, LCX and RCA respectively. With respect to risk category classification we obtained a correlation of 0.88 and a $\kappa$ value of 0.90.

Acknowledgements. We would like to thank Dr. Hortense Kirisli and Dr. Nora Baka for their suggestions and fruitful discussions.

References