Machine learning for coronary artery calcification detection and labeling using only native computer tomography

Asmae Mama Zair[1] · Assia Bouzouad Cherfa[1] · Yazid Cherfa[1] · Noureddine Belkhamsa[1]

Abstract

In recent decades, the World Health Organization has found an increase in the death rate due to cardiovascular disease. Calcifications of the coronary arteries are the main sign of any cardiovascular event. Each individual's calcium score helps estimate the severity of the disease. However, the score for each artery is more significant. This study aims to research the segmentation, the labeling, and then the complete and partial quantification of calcium using only native coronary computed tomography with the help of machine-learning algorithms. Our semi-automatic system limited the region of interest by applying a defined preprocessing step. We then implemented two random forest classifiers; the first separated true coronary artery calcification (CAC) from the noise, and the second labeled CAC into the right coronary artery, left coronary artery, left anterior descending artery, and left circumflex artery using specific features. Agatston score and volume score of each CAC, each artery, and all of the arteries were calculated. This method gave promising results, comparable to those found in the literature, with the accuracy of 99.98% and 100% for CAC detection and labeling respectively.

Introduction

The heart is one of the main muscles in the body. It ensures blood flow by coordinating contractions. the coronary arteries supply the heart with oxygen and nutrients. According to the World Health Organization [1], coronary vascular disease (CVD) is considered to have the highest death rate in the world. Coronary artery calcification (CAC) is one of the lesions that affect the heart gradually, without symptoms. Preventive detection and the associated demand for stratification of individual cardiovascular disease risks depend largely on the accuracy and amount of information acquired. New coronary imaging techniques allow more accurate prediction of cardiovascular risk [2] and provide additional information on a patient's health.

Frequently, CAC is caused by the buildup of calcium, fats, cholesterol, and other substances in the inner wall of the arteries. This buildup hardens over time, eventually leading to blocked vessels [3]. Calcium scoring can be performed particularly in coronary computed tomography (CCT) without enhanced contrast. CCT is a non-invasive technology that detects and quantifies calcified lesions. CAC score is a powerful marker and an independent predictor of the cardiovascular event in a patient [4]. It was first reported by Agatston et al. [5]. The Agatston score (AS) is defined as a summary measure based on the total volume and density of the pericardial coronary calcification. It can be used to stratify patients into four risk categories, which range from 0 to 10 (minimal calcifications), 11 to 100 (moderate calcifications), 101 to 399 (modest calcifications), and 400 to several thousands (severe calcifications), indicating extensive coronary atherosclerosis.

Several recent studies aim to treat CAC and stop its progression. Doctors use numerous treatment strategies such as medications, balloon angioplasty, stenting, and bypass grafting in patients with advanced stages of CAC. The location and types of plaques (soft, mixed, and calcified calcification) contain key information about the stiffness of the injured area. Therefore, more meaningful information can help to improve treatment outcomes.

Doctors can benefit from visual information and other interesting data available in images (size, position, and shape), as well as a deeper understanding of the

✉ Asmae Mama Zair
asmaz.zair@gmail.com

1 LASICOM Laboratory, Department of Electronics, Faculty of Technology, University of Blida 1, 09000 Blida, Algeria

characteristics of lesions. Labeling the coronary arteries by their standard anatomical names provides more detail and speeds up the analysis process [6].

CAC exclusively resides in the walls of the coronary tree [7], which is a passage on the heart surface (epicardium). Exploiting this prior knowledge, researchers make efforts to crop a specific region of interest (ROI) focused on the heart and coronary artery region.

Several studies have used 2D segmentation to detect and quantify CAC [8–12]. These studies did not take into account the thickness between the volume sections, Thus losing key information. However, the volume and density of the coronary tree calcium, each separately, provide additional information regarding subsequent risk of clinical events.

The CAC segmentation is still a challenging task due to the distinction between true candidates and other dense objects (bone, calcified tissue, metal stents). Additionally, calcification produces blooming that may cause the erroneous enlargement of the appearance of calcification and prevent accurate evaluation [13]. Furthermore, the time required for manual segmentation and scoring of calcium makes the process prohibitive. In these cases, fully automatic or even semi-automatic methods are preferred to facilitate the workflow.

Various methods have been proposed for calcium scoring in non-contrast-enhanced CCT during the last decade. It started with rule-based approaches of mathematic morphology, threshold, and region and grew to the better-performing conventional machine learning (ML) approaches [14], finally arriving at the more sophisticated approaches of deep learning [2, 15, 16]. The first work that presented an automatic method of CAC detection and scoring to assign a risk category to a subject was performed by Išgum et al. [17]. Kurkure et al. [18] isolated the heart region based on anatomical markers; subsequently, the calcifications were segmented with a bone threshold and classified. The classifier's role is to eliminate false candidates.

Xie et al. [19] limited the heart region by removing the lungs, bones, aorta, and adipose tissue. Then, they applied $3 \times 3 \times 3$ medium filters and identified the CACs using the bone segmentation threshold. The CAC is considered a connected group of more than five voxels. Wolterink et al. [20] created an atlas to segment the midline of the coronaries and identified true calcifications by using random forest (RF). The work was finalized by a score calculation method proposed by Agatston. Wolterink et al. [21] used supervised learning by ConvNet to directly identify and quantify the CAC using the three planes: axial, sagittal, and coronal.

Yang et al. [10] generated a patient-specific ROI using the coronary computed tomography angiography (CCTA) contrast images. They extracted the heart and coronary vascular tree segmentation from the same patient. ROI

eliminated false calcifications localized in the surrounding tissues. The authors applied a support vector machine (SVM) classifier to discriminate the coronary calculations from image noise. Lessmann et al. [22] used two convolutional neural networks (CNN) without extracting ROI. The first network, which was a wide receptive field, allowed labeling the candidates according to their anatomical location; the second network identified the real calcifications among the candidates detected. Santini et al. [8] created the heart atlas. They applied a threshold of 130 *Hounsfield units* (*HU*) to extract CAC, valves, and aortic calcifications as well as noise. ConvNet was designed to detect real cardiac calcifications. Shadmi et al. [23] used the fully convolutional neural network (FCNN) approach to segment coronary calcium and predict the AS. Šprem et al. [24] used the expectation-maximization algorithm to separate true CAC from its background. Zreik et al. [25] segmented the midline of the coronary artery with the CNN method. Then, to address the task of analyzing plaque and stenosis of the coronary artery, they used 3D CNN associated with a recurrent neural network (RNN).

Zreik et al. [26] extracted image features from a volume centered around each centerline point of the arterial wall. Subsequently, they analyzed the features extracted from the sequence of analyzed volumes.

First, a 3D CNN detected coronary plaque and characterized the type of plaque detected (no plaque, uncalcified, mixed, calcified). Thus, it detected stenosis of the coronary artery and determined its anatomical significance (no stenosis, non-significant, significant). Second, an RNN was deployed to analyze and accumulate the characteristics extracted by the CNN. It arranged all the voxels in a single vector for classification. In a recent work [27], the authors created an automatic 3D deep learning regression network to perform direct regression of the AS by skipping the segmentation step.

The ML is used to solve segmentation and classification problems that are too difficult to distinguish. ML methods used in [9–12] provided high accuracy (Table 4). The model uses a maximum number of features to distinguish between true and false candidates. Qian et al. [28] extracted the morphological and geometric statistics of each calcifying lesion. They estimated the probability of risk caused by each injury by using a naïve Bayesian technique. Išgum et al. [29] created a multi-atlas probabilistic map to segment CAC. In addition, a statistical shape recognition system was designed to identify coronary calcifications by texture, size, and spatial characteristics. To achieve this, a classification of a mixture of classifiers (*k*-nearest neighbors [KNN] and SVM) was proposed in this work. Lessmann et al. [9] estimated the position of the heart using three deep CNNs in the axial, coronal, and sagittal sections of the image volume. A 130 HU threshold and an analysis of the connected

components were applied to identify the CACs. The authors considered that the components less than $1.5 \, mm^3$ or greater than $5000 \, mm^3$ are probably the noise or bone structures. de Vos et al. [12] proposed a calculation method that uses two ConvNets; the first performed a recording to align the fields of vision of the input CCTs, and the second performed a direct regression of the calcium score.

The automatic methods speed up the process and make the software easier to use for medical personnel, but their results are not always accurate [30]. Unlike manual methods, their results were more reliable but more difficult to manipulate. Therefore, semi-automatic methods brought enormous benefits and balance between execution time, result, and manual work [31, 32].

Using methods on large datasets is time consuming and requires a dedicated workstation. To reduce the calculation costs, we proposed a preprocessing step to eliminate the voxels that surround our volume of interest (VOI) but contain no relevant information.

Most of the mentioned approaches for the detection and labeling of CAC depend on CCTA and CCT scans as clinical data. Likewise, they identify the severity stage of the lesion using a total AS of the coronary tree. Performances of these approaches are limited and require more data for labeling. Moreover, they fail to adequately identify the precision of the diagnostic.

The main contribution of this work is the segmentation and labeling of CACs using only native computed tomography. For good segmentation and labeling, the choice of the features of the RF classification algorithm is essential. The scoring of each candidate, for each artery and then for the entire volume, gives more details on the lesion and guidance for a correct diagnosis. In this paper, CCT volumes were first pre-processed for removing unnecessary regions by using segmentation techniques, such as lungs and ribcage. The CAC segmentation was done followed by statistical analysis to detect the most-significant features that will characterize the true CAC. Two successive classification steps were devoted to detecting and labeling CAC lesions.

Materials and methods

Data and materials

The experiments were implemented using a PC with Intel Core $i7 - 5500U \, CPU$ (4×2.40 GHz), 8 *GByte* of RAM, and the computations were performed on MATLAB 9.6 (R2019a), under Windows 10 operating system. The orCaScore framework provided the data used for classifying the CAC and evaluating the classifier, which consisted both of CCT and CCTA data of 72 patients from four different hospitals and four different vendors. The CCTs were acquired

with a tube voltage of 120 *kV*, an in-plane resolution ranging from $0.4 \, mm \times 0.4 \, mm$ to $0.5 \, mm \times 0.5 \, mm$, and a thickness of 2.5 mm (GE), 3 mm (Philips, SIEMENS, and TOSHIBA). The images were reconstructed into a resolution of 512×512 matrices and a 16-bit Hounsfield scaling. The slices included the entire heart volume from the level of the tracheal bifurcation to the base of the heart (56 ± 9 slices per patient). The dataset contained volumetric images: There were 32 ground truth images, one per patient, in the form of segmented and artery-specific labeled CAC and a set of 40 patient's images without accompanying ground truth. All of the data was labeled manually by two experienced cardiologists (with 2 and 6 years of experience in cardiac CT imaging) using a special customized calculation software (CardIQ Xpress 2.0, AW workstation, GE Healthcare); the data were classified as belonging to the right coronary artery (RCA) calcification, left coronary artery (LCA) calcification, left anterior descending (LAD) artery calcification, or left circumflex (LCX) artery calcification.

The dataset was divided into 32 scans for training and 40 scans for testing and final evaluation. The training set contained 17 coronary calcifications and 15 non-coronary calcifications. The test set contained 26 coronary calcifications and 14 non-coronary calcifications. In this study, we used only the CCT sequences.

Methodology

In the first step, the pericardium was segmented by applying a series of treatments to the original volumes (pericardium segmentation, 3D heart segmentation and bounding box creation). This was used to reduce our CAC search space and avoid a maximum of false positives later in the classification step. The bone threshold was a standard way of identifying candidates for vascular calcification in CCT images. The minimum HU threshold to identify potentially calcified volumes was predefined at 130 HU. The 3D connected component analysis (26 connectivity) [33] was performed on the threshold image to obtain a set of candidate calcifications. According to Lessmann et al. [9], the candidates with a volume of less than $1.5 \, mm^3$ or greater than $1500 \, mm^3$ were considered as noise or metal implants, respectively; hence, they were rejected. However, some non-CAC lesions (valvular and aortic calcifications) in this VOI were included and still represented noisy objects. In this case, an ML approach plays the role of eliminating false positives and only represents the true CAC lesions. The resulting calcifications assigned different labels according to their corresponding coronary branches so that they are then subjected to multiclass classification. To distinguish among the candidates, it was imperative to study their different characteristics. For this, each candidate was described with a set of parameters. The input of the first RF classifier [34] contained the features

vectors of all the objects resulting from our segmentation. The feature vector used for classification comprised of the features listed in Table 1. The output of the classifier was the result of CAC detection. A second RF classifier was used for CAC labeling. For this, new feature extraction was based on the CAC anatomical location and the voxel information that was cited at the CAC labeling stage (see Fig. 1).

Preprocessing

The main objective of the preprocessing was to create volumes in a form suitable for further processing. This was accomplished by removing the unwanted parts in the background. Initially, the VOI was delimited, and the unnecessary structures were extracted. Next, the pericardium segmentation procedure itself, which used the data obtained in the previous step.

Pericardium segmentation The black space surrounding the body in the CT image and organs close to the heart do not transmit any diagnostic information. The idea was to isolate the pericardium and remove all the unnecessary

Fig. 1 Flowchart visualizing the proposed semi-automatic CAC detection and labeling method

structures (lung, bone, and anterior chest wall muscle). Therefore, the volumes were pretreated and cropped before being used in the experiments. This VOI cannot only highlight the coronary calcifications but also label the calcifications belonging to the different main coronary arteries.

1. *Volume of interest (VOI) limitation:* The operator manually selected one of the slices that preceded the bifurcation of the coronary arteries in such a way that all the arteries were included in the VOI. Anatomically, the length of a healthy patient's heart is equal to 120 mm [35]. From the first slice, the program calculated the number of slices that cover the whole heart using the thickness information of each vendor, using Eq. 1.

   $$Number\ of\ slices = 120\ \text{mm}/thickness \qquad (1)$$

2. *3D Ribcage segmentation:* After the VOI determination, we removed the ribcage and the lungs. The ribcage was segmented with a bone threshold (> 130 HU and greater than 1500 mm^3) followed by 3D connected component [33]. The 3D dilatation operation, in which the kernel size was $5 \times 5 \times 10$ and kernel $(1:3, :, :) = 1$, was applied to close the holes between the ribs. Here, the chest wall muscle was almost separated from the pericardium.

3. *3D lung segmentation:* The lung region segmentation was achieved through the following steps. We applied an optimal threshold as suggested by Sonka et al. [36] to extract the initial pulmonary mask. The lung area was an uninteresting region and had low density, which ranged from $-100\ HU\ to -500\ HU$. The interesting area contained the surroundings of lung lobes. Since the lungs were placed in an uninteresting area, we initially selected a threshold value of $T_0 = -500\ HU$ and then used the optimal thresholding technique to find the optimal threshold T_{op}. The optimal threshold was finally used to obtain the initial segmentation of the lungs.

$$f(x;y;z) = \begin{cases} 0 & f(x;y;z) \geq T_{op}, \\ 1 & f(x;y;z) \leq T_{op} \end{cases} \qquad (2)$$

Here, the *x* and *y* indices represent the slice coordinates and *z* indicates the slice number. The volume consisted of the total number of *z* slices, and each slice had a dimension of $x \times y$ pixels [36]. After that, the connected components labeling method was applied to eliminate all structures other than the pulmonary region. Morphological filling of the holes was used to fill and suppress the holes in the region obtained.

3D heart segmentation and bounding box creation Since the heart is positioned between the two lobes of the lungs, we constructed a convex hull of the pulmonary lobes. Then, we extracted the inner region, that is, the pericardium. A
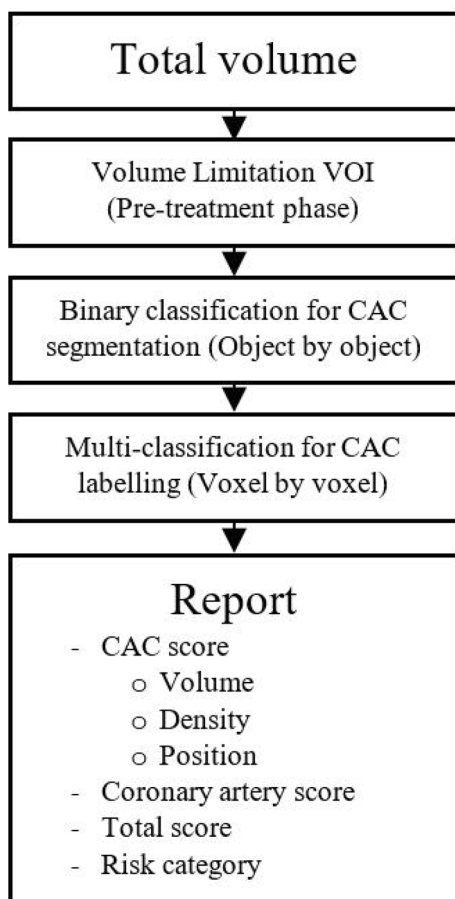
3D bounding box around the heart structure was detected. It was automatically determined using the algorithm described by de Vos et al. [37]. The volume around the bounding box was cropped out.

CAC segmentation

According to the orCaScore challenge, a bone threshold that included the voxels with an attenuation value > 130 HU threshold was applied to form CAC lesion candidates. Then, the 3D connected component tool (3D 26 connectivity) [33] was used to merge the voxels of the threshold image into singular lesions. The candidates with a volume between 1.5 mm^3 and 1500 mm^3 were considered as CAC. The resulting volume still included the non-CACs such as valvular and aortic calcifications. At this point, the role of ML classification was to eliminate non-candidates by calculating and learning their characteristics (see Fig. 2).

CAC classification

In this phase, we tried two different classifiers for comparison. RF [34] and AdaBoost [38] classifiers were trained with an ensemble of 100 trees to distinguish between true positive and false positive CACs.

RF and Adaboost were both ML classifiers. The difference between them is that RF uses parallel and independent decision trees, while Adaboost uses sequential trees. The motivation for using RF in this work is based on its advantages of the low chance of overfitting due to sufficient trees in the algorithm, good precision despite missing datasets, and fast convergence.

Feature extraction Feature extraction is a well-known method of giving a physician the ability to locate specific lesions. The goal was to find the most-discriminating and significant features that can differentiate between CAC and non-CAC objects. The classification, in this case, was done object by object. The following are the descriptions of the features used in this work:

– 2D features: For each object, the largest slice of the object volume was chosen to calculate features mentioned in Table 1.
– 3D features: In this case, we focused on the volumetric characteristics of objects. Features included size, position, shape, and appearance information of candidates and their surroundings. The statistical Haralick features
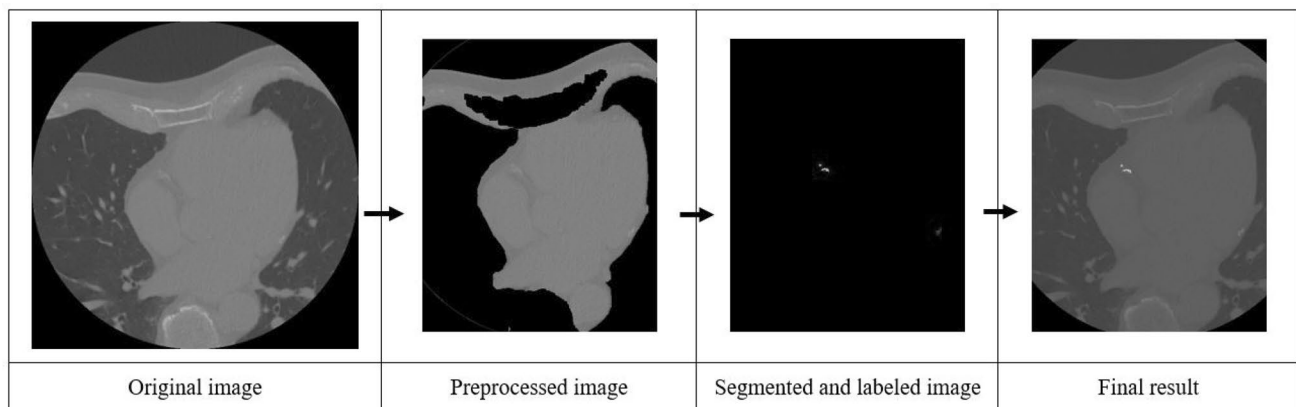

Fig. 2 The flowchart of CAC segmentation and labeling of the proposed method

Table 1 Important extracted 2D and 3D features used in the binary classification

2D/3D	Number of features	Features	Notes
2D	7	Perimeter, circularity, equivalent diameter, eccentricity, extrema, elongation, roundness	The largest slice of the object
	12	3D Haralick features: energy, entropy, correlation, contrast, variance, sum mean, inertia, cluster shade, cluster tendency, homogeneity, max probability, inverse variance	
3D	3	Invariant HU moment according to X/Y/Z	
	5	The maximum intensity of Gaussian kernel with sigma = 0.5,1,2,3,4.	
	7	The entropy of the 3D wavelet decomposition volumes of "Haar" level 2	

were calculated across the image/volume. It was based on the gray level co-occurrence matrix. These measures were used to describe the overall texture of the image/volume using entropy and the sum of the variance [39]. Other calculated and selected features are listed in Table 1.

Powerful features selection for CAC classification: The features selection for classification is a tool that aims to reduce the dimensionality in datasets. It removes redundant information by performing information filtering. Thus, it can improve the efficiency of classification methods and reduce execution time. In this paper, we used the Relief-F function [40] for features selection. The Relief-F function ranks the importance of predictors. For each candidate, 35 features were calculated. A comparative study between the two classifiers is presented in "Results" section.

CAC labeling

The purpose of this part was to label CAC according to its location: RCA, LCA, LAD, and LCX. A complicated CAC was placed next to the separation of two main coronary arteries. We, therefore, could not consider it as a single CAC, but as two CACs glued together (bifurcation CAC; Fig. 3). Our system considered each voxel of a CAC to be an independent candidate. An RF classifier was trained to divide the calcifications based on their arterial bed.

Voxel-by-voxel CACs labeling Based on cardiovascular anatomy [6] and following the short-axis view of the heart, we created a map that allowed us to locate and label CACs without segmenting the coronary arteries. Some specific features were extracted and mentioned later to label CACs. To carry out this process, we followed these four steps:
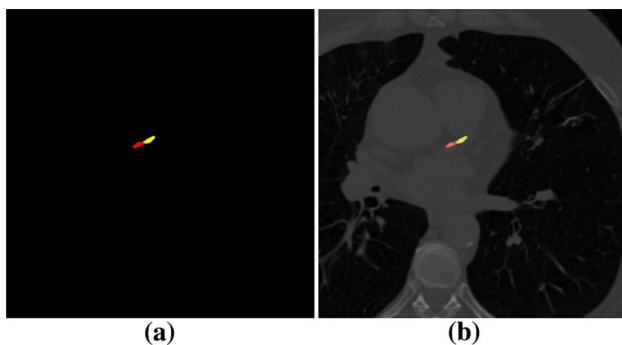

(a)   **(b)**

Fig. 3 Example of bifurcation CAC. **a**, **b** Represent the segmentation and labeling of a bifurcation CAC respectively. The red and the yellow colors represent the LCA and LCX calcifications respectively

– Aorta center was automatically defined: The aorta is an anatomical landmark found in the early sections of the heart (in the first slice manually chosen). The aorta, in this case, appeared circular. The Hough transform was applied to extract the ascending aorta [10].
– Volume Was divided into 32 fragments: Our idea Was inspired by the bull's eye plot [41]. We schematized a map to extract a location parameter that assists the classifier in correctly labeling CACs. The idea was to create two circles with the same center A and different diameters AB and AC. A was located in the center of the aorta, B was located at the point of bifurcation of the LCA into two coronaries; LCX and LAD. C was located outside of the heart with $AB < AC$. The circle was divided into 16 triangular fragments. We thus obtained 32 fragments numbered from 1 to 32. The schematic map was applied to the axial images of the volume (Fig. 4). Anatomically, $AB \simeq 30$ mm and the radius of the heart was almost equal to 60 mm in the axial view. Since the resolution of the image changes from one provider to another, we automatically calculated the distance AB according to the following equation:

$$AB = 30/EsX \tag{3}$$

and

$$AC = 60/EsX \tag{4}$$

where *EsX* is the element spacing of *X*.
– Voxel parameters were calculated: intensity, distance from the voxel-aorta center, the segment that contains the CAC, voxel position (*x*, *y*, *z*).
– Classification was done: CACs were labeled as RCA, LCA, LCX, and LAD.

The output of this classification was the result of labeled coronary calcifications. SVM and RF classifiers are presented in Table 3

Partial and total patient CAC scoring

In clinical practice, the AS [5] and volume score (VS) [42] are the most used indexes for coronary calcium quantification. In this study, the AS was calculated as the product of the lesion area A and a weighting factor W, which depends on peak lesion density. The total AS of each patient was calculated by summing the scores of every calcified focus across all volume bands (see Eq. 5).

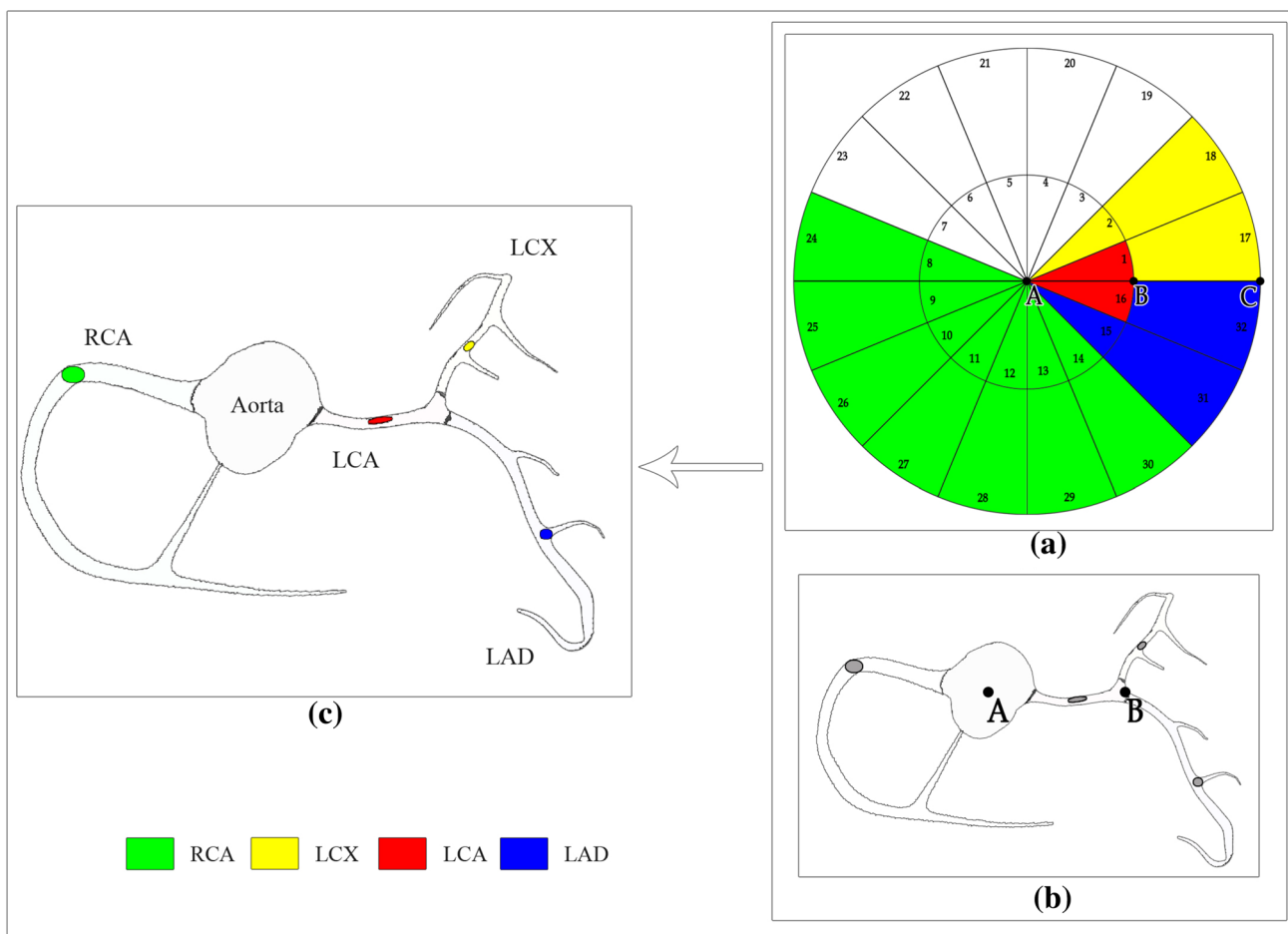$$Agatston\ score = \sum_{i=1}^{N} (A_i \times W_i) \tag{5}$$

Fig. 4 Schematic diagram of the process to classify CACs. **a** The representation of the location of coronary artery in a circular map, **b** the representation of coronary artery tree containing different calcifications, where A and B are the aorta center and left main coronary artery bifurcation, respectively, an **c** labeled CAC in coronary artery tree

Calcium VS was calculated based on the volume and intensity of the voxel. This method is robust and applicable. The voxel volume was calculated using Eq. 6. Following the work of McCollough et al. [42], we calculated the calcification VS according to Eqs. 7–9.

$$Voxel_i = EsX_i \times EsY_i \times Thikness\,Z_i \tag{6}$$

$$CAC\ score = \sum_{i=1}^{M}(Voxel_i \times W_i) \tag{7}$$

$$Artery\ score = \sum_{i=1}^{M}(CAC\ score_i) \tag{8}$$

$$Total\ score = \sum_{i=1}^{N}(Artery\ score_i) \tag{9}$$

where EsX_i, EsY_i are the pixel sizes, and N and M are the number of pixels and voxels in the evaluated lesion, respectively.

Moreover, the weights (W) are represented as follows:

$$130 - 199\ HU \Rightarrow W = 1$$
$$200 - 299\ HU \Rightarrow W = 2$$
$$300 - 399\ HU \Rightarrow W = 3$$
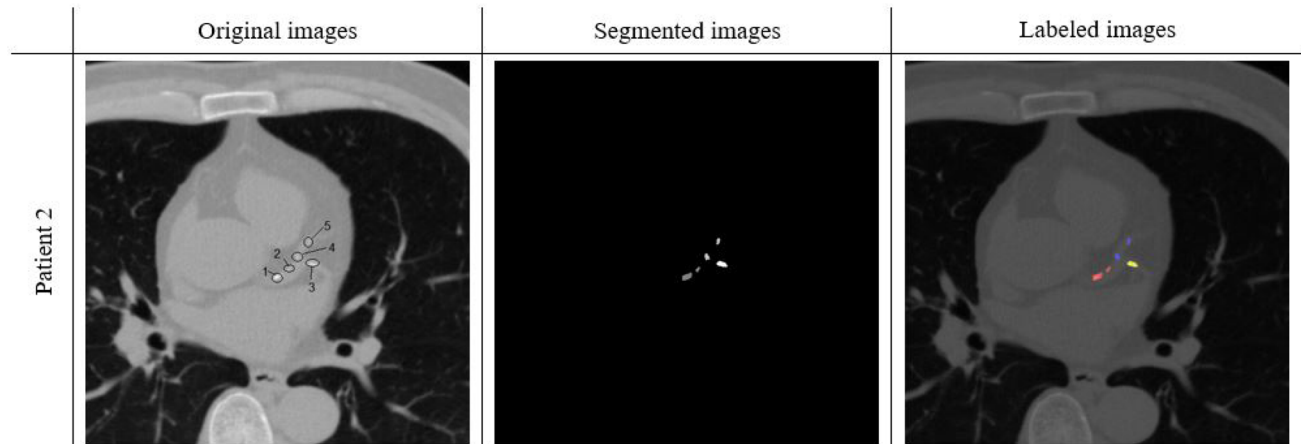$$\geq 400\ HU \Rightarrow W = 4$$

In previous works, the AS, total VS, and partial scores were obtained for the four main coronary arteries [20]. However, in this study, we were motivated to find out whether volume, density, and position of each CAC can explain the severity of the risk and give more clarification of the extent of the lesion. Figure 5 shows an example taken from two different patients.

Rumberger et al. [43] initially proposed the conventional categories for the CAC notation according to AS/VS value.
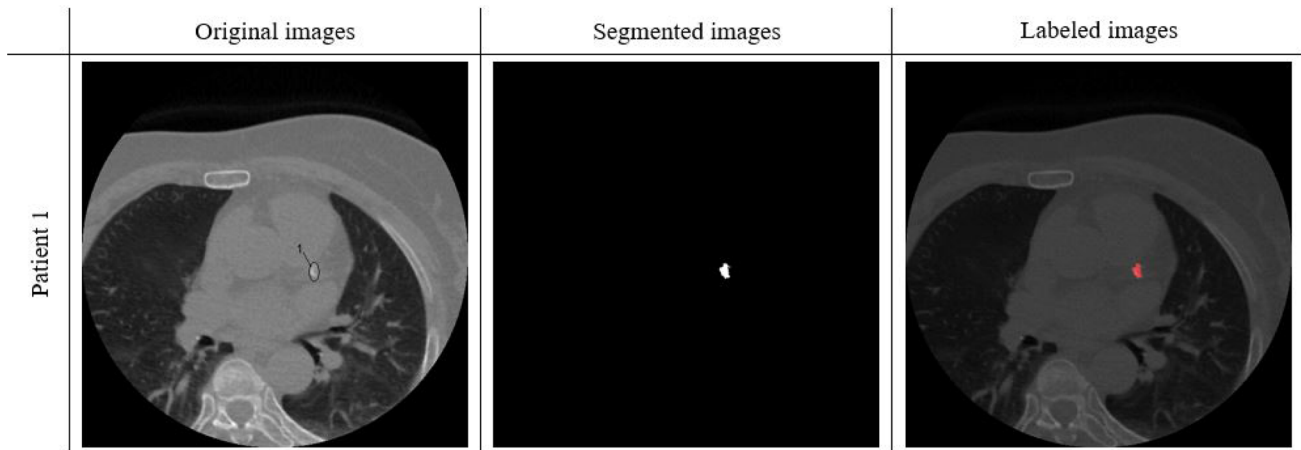
	Number of CAC	CAC Score				Risk category	Total score	Total risk category	Volume	Position	Density
		RCA	LCA	LAD	LCX						
Patient 1	1	-	185	-	-	Modest	185	Modest	354	[398; 310;20]	254.8941
TOTAL CAC Score per artery		-	185	-	-	-	-	-	-	-	-
Risk category per artery		-	Modest	-	-						

(a)

	Number of CAC	CAC Score				Risk category	Total score	Total risk category	Volume	Position	Density
		RCA	LCA	LAD	LCX						
Patient 2	1	-	97	-	-	Medium			184	[347; 320;18]	196.4149
	2	-	49	-	-	Medium			54	[362;310;18]	175.5926
	3	-	-	-	83	Medium	185	Modest	158	[394;303;18]	189.8354
	4	-	-	39	-	Medium			88	[374;284;18]	176.9318
	5	-	-	14	-	Medium			66	[389;273;18]	176.0909
TOTAL CAC Score per artery		-	146	53	83	-	-	-	Total volume = 850	-	Total density = 914.8656
Risk category per artery		-	Modest	Medium	Medium						

(b)

Fig. 5 Examples of original, segmented, and labeled images taken from two different patients with their reports: **a** the first patient with a single CAC at the LCA bifurcation and **b** the second patient with five CACs located in the LCA, LCX, and LAD

$AS/VS\ risk\ categories = 1 - 10 \Rightarrow Minim$

$\qquad = 11 - 100 \Rightarrow Medium$

$\qquad = 101 - 400 \Rightarrow Modest$

$\qquad > 400 \Rightarrow Sever$

Results

This section presents the findings we obtained through the evaluation of the CAC segmentation algorithm. The measurement between the segmentation results and the ground truth is presented in Table 2. The comparative analysis of the classification performance of the RF and Adaboost classifiers yielded an F1 score of 0.82 and 0.71, respectively, and a Dice coefficient of 0.74 and 0.61, respectively. Among these two ML algorithms, the RF classifier outperformed the AdaBoost classifiers, as indicated by its accuracy value of 99.98% compared to that of the Adaboost classifier (99%).

Pearson's correlation and Bland–Altman comparison [44] were used to compare the total VS obtained by the algorithm to that obtained by expert manual scoring. The AS showed a correlation of R = 0.857, $p < 0.0001$ with the ground truth (Fig. 6). The Bland–Altman plot (Fig. 7) showed a positive bias of 53.31 and the 95% limits of agreement ranged from −417.05 to 525.69. In the second step, the classification for the labeling of calcium voxels in LAD, LCX, RCA was carried out. Table 3 presents a comparison between quantitative assessment of CAC labeling using SVM and RF classifiers. To assess the classification performance, we used the parameters of accuracy, sensitivity, and specificity. These parameters are widely used to quantify the quality and reliability of a classifier. While using the SVM classifier, the accuracy, sensitivity, and specificity ranged between 99.92% and 100%. However, these were all equal to 100% in the fourth coronaries when the RF classifier was applied. The RF classifier showed good accuracy and worked perfectly in our case.

Figure 8 shows four examples of different orCaScore data providers. The original images, ground truths, and our results are presented in the first, second, and third columns, respectively.

The artery score was calculated by adding the score volumes of all the calcifications located in that coronary artery. The CAC scores were calculated according to their volumes

Table 2 CAC detection

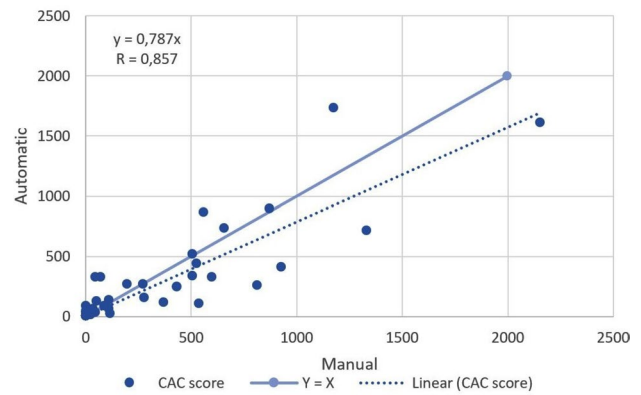|	Dice coefficient	Accuracy %	F1 score
RF	0.74	99.98	0.82
AdaBoost	0.61	99.00	0.71

Fig. 6 Pearson's correlation between automatic and expert manual CAC score

and intensities. The volume, intensity, and position of each CAC (the center of gravity of the CAC) were calculated and displayed in the volume report.

Discussion

The given work presented a semi-automatic method for CAC identification and AS by cardiac computed tomography. The method began with a preprocessing phase to make the input volumes identical and remove any unused structures.

The aim of the first slice selection, which contains the aorta, and the VOI limitation was delimited to keep only the heart and avoid the structures above and below it.

We selected a pack of slices according to the heart size and thickness of the slices from each vendor to cover the whole heart.

The diversity of equipment used to obtain CCT scans in the orCaScore database (GE, Philips, Siemens, and Toshiba), the different constants in each acquisition, and the morphological differences among individuals resulted
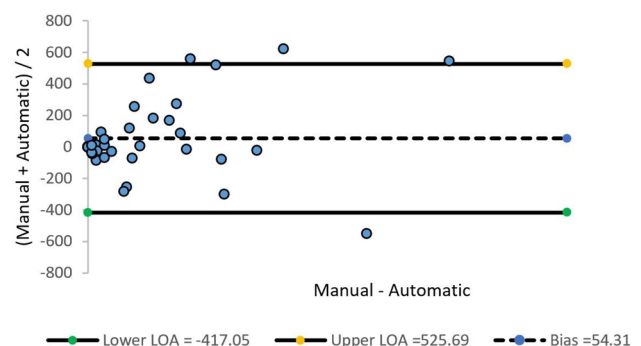

Fig. 7 Bland–Altman plot analysis between automatic and expert manual CAC score

Table 3 Comparison between quantitative assessment of CAC labeling using SVM and RF classifiers

Classifier method	SVM				RF			
Coronary artery	RCA	LCA	LAD	LCX	RCA	LCA	LAD	LCX
Accuracy (%)	100	99.92	99.92	100	100	100	100	100
Sensitivity (%)	100	99.34	100	100	100	100	100	100
Specificity (%)	100	100	99.89	100	100	100	100	100

in the limitation of using a fixed threshold for each scan in the database. These limitation led us to use an optimal threshold and 3D connectivity for 3D lung segmentation. The lungs appeared in the largest region. The bone threshold (130 HU) followed by 3D connectivity and dilatation with special structuring elements were done for ribcage segmentation.

In the bottom of the heart region, the diaphragm and liver still appeared, and their intensity was similar to the heart muscle. These will cause noise that will be corrected later in the classification step. Since the heart is positioned between the two lungs and above the diaphragm, the convex hull method of lungs contains the whole heart region. Ine this study the space on either side of the heart was cropped out using a bounding box algorithm. Compared to the original volumes, the process was easier and less time consuming $(19 \pm 5s)$.

Fig. 8 Example of sections of certain volumes of CCT. The original images, the ground truths, and the results of the proposed method are presented in the first, the middle, and the last column, respectively

Selected and pretreated input volumes were presented to a calcium detection classifier that directly predicts the CACs in these volumes.

In preliminary experiments, we found that the object-by-object classification could differentiate CAC from other types of calcifications, e.g., aorta calcification, pericardium calcification, and heart valve calcification. The voxel-by-voxel classification in this case was taking a lot of time during training. To simplify the problem, we extracted a bounding box around each object and calculated their characteristics.

By limiting the number of features (using the most-significant ones), we were able to use a future selection function Relief-F that ranks predictors according to their importance. After several attempts, we selected the important 2D and 3D predictors (Table 1).

RF is known for its efficiency and performance in complex cases, including the similarity between real and false candidates. Therefore, in this work, we chose RF for CAC detection.

Using the expert manual annotations as references, our classifier performed well with an accuracy of 99.98%.

Table 4 presents the comparison of the findings of previous studies presented in the literature review with the results obtained in our study on CAC segmentation. Our method surpassed most of the previous work. It is clear that the results of [11] were a little better than ours, but they required more information on the data for the realization (volumes CCTA and CCT). Conversely, our method used only CCT images. In addition, the latest research is directed towards the detection of CAC in three dimensions, and our work followed this direction. Some classification errors were inevitable due to the similarity between CAC and noise.

Interestingly, the method gave good results despite the limited size of the database. Although the cardiac tomographic images were synchronized with the ECG, a few volumes with artifacts still existed. The misclassification of CAC occurred primarily where the right and left coronary arteries originated (Ostia). Even manual classification of calcifications can be very difficult in this part. Moreover, noise in the liver was not fully corrected.

Multi-classification SVM and RF classifiers were tested with CAC labeling. Table 3 presents the evaluation of the four principal coronaries with each classifier. The RF trained with our features (intensity, distance from the voxel-aorta
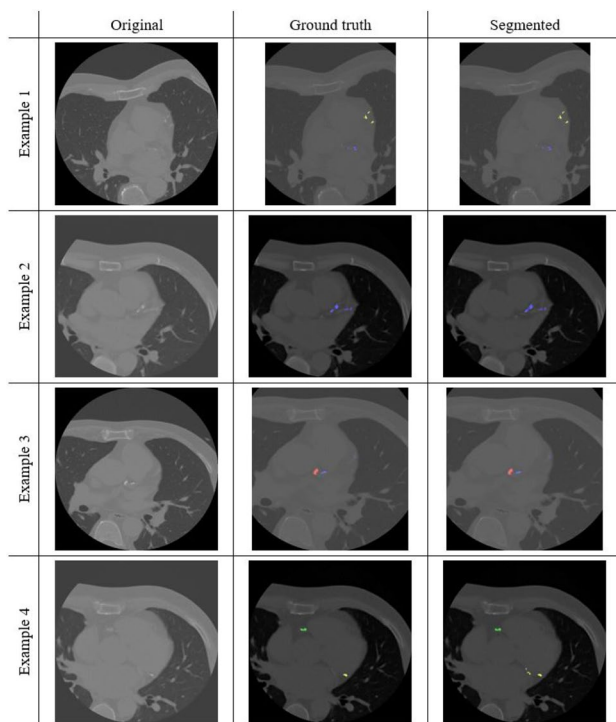
Table 4 Results comparison of CAC segmentation studies presented in the literature

Authors	Method	2D/ 3D	CCT/ CCTA	DATA (volume)	PPV	F1 score	Accuracy %
Lessmann et al. [9]	CNN	2D	CCT	1028	–	–	84.4
Yang et al. [10]	SVM	2D	CCT/ CCTA	32	0.97	–	–
Durlak et al. [11]	Atlas-based feature approach in combination with a RF classifie	2D	CCT/ CCTA	72	0.97	0.95	100
De Vos et al. [12]	ANN	2D	CCT	903 cardiac	–	–	93
				1687 chest	–	–	99
Our method	Two RF classifiers	3D	CCT	72	0.83	0.82	99.98

center, the segment that contains the CAC, voxel position) provided high classification accuracy (100) in fourth coronaries (RCA, LCA, LCX, and LAD). As a result, the proposed algorithm outperformed the previous work studied in state-of-the-art.

To achieve this perfect result, we tried to choose specific features, which are mentioned in "Methodology" section. These features have not been utilized in any work till date. They included the location, shape and parameters of the voxel, without using an atlas and CCTA.

The calcium scores per-artery presented the risk of each artery. This is mostly used in clinical research. In this work, we calculated AS after CAC labeling for a precise view of the severity of each coronary artery. Consequently, we achieved great reliability.

Figure 5 shows two examples of partial and total AS/VS calculation. The first patient had a VS of 185, and the second had a VS of 282. The total risk category in both patients was the same (modest), but the partial risk was modest and medium in the first and second patient, respectively. The risk of the first patient was more severe than that of the second because of the location (bifurcation), density, and volume of the CAC. In this case, we can highlight the importance of a partial risk study to explain the severity of the risk and show the extent of the lesion.

The contribution of the current work lies in

- the speed of program execution (19 ± 5 s),
- the use of only native scans,
- new and simple features for CAC labeling that give a perfect result at different vendors, and
- calculation of total and partial AS/VS to properly diagnose stenosis.

Limitations and future works

The limitations of this study were the size of the CACs and the lack of data. To fill these gaps, we plan to increase the input volumes from different datasets and to add chest CCT volumes in our future work. The liver can be cropped before CAC segmentation for more precision.

Conclusion

Our approach enriches the state of the art in medical imaging by improving the speed of calculation and the simplicity of implementation. The coronary artery calcium score is an essential tool in measuring cardiovascular disease. Until now, physicians manually distinguished CAC with increased radio-density of slice-by-slice CCT images. In this study, we presented a 3D CAC detection and labeling strategy. For 3D calcium scoring, the preprocessing step of the original data improved classification results. We implemented two RF classifiers applied in the heart region. The first RF was used for true CAC identification and the second for labeling CACs using some special features. We obtained perfect results without extracting the coronary artery tree from a CCTA or using the atlas registration. Despite the low contrast and high noise existence in the images, the proposed algorithms introduced promising results in detecting the CAC in most of the patients in comparison with the ground truth and give good results in labeling CAC in CCT scan (accuracy of 99.98% and 100% for CAC detection and CAC labeling in the four coronary branches, respectively).

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

1. Organization WH, et al (2019) World health statistics 2019: monitoring health for the SDGs, sustainable development goals
2. Litjens G, Ciompi F, Wolterink JM, de Vos BD, Leiner T, Teuwen J et al (2019) State-of-the-art deep learning in cardiovascular image analysis. JACC Cardiovasc Imaging 12(8):1549–1565
3. Pravina P, Sayaji D, Avinash M et al (2013) Calcium and its role in human body. Int J Res Pharm Biomed Sci 4(2):659–668
4. Neves PO, Andrade J, Monção H (2017) Coronary artery calcium score: current status. Radiol Bras 50(3):182–189
5. Agatston AS, Janowitz WR, Hildner FJ, Zusmer NR, Viamonte M, Detrano R (1990) Quantification of coronary artery calcium using ultrafast computed tomography. J Am Coll Cardiol 15(4):827–832
6. Drake R, Vogl AW, Mitchell AW, Tibbitts R, Richardson P (2020) Gray's atlas of anatomy e-book. Elsevier, Amsterdam
7. McEvoy JW, Blaha MJ, DeFilippis AP, Budoff MJ, Nasir K, Blumenthal RS et al (2010) Coronary artery calcium progression: an important clinical measurement? A review of published reports. J Am Coll Cardiol 56(20):1613–1622
8. Santini G, Della Latta D, Martini N, Valvano G, Gori A, Ripoli A, et al (2017) An automatic deep learning approach for coronary artery calcium segmentation. In: EMBEC & NBC 2017, Springer, pp 374–377
9. Lessmann N, Išgum I, Setio AA, de Vos BD, Ciompi F, de Jong PA, et al (2016) Deep convolutional neural networks for automatic coronary calcium scoring in a screening study with low-dose chest CT. In: Medical imaging 2016: computer-aided diagnosis, vol 9785. International Society for Optics and Photonics. p 978511
10. Yang G, Chen Y, Sun Q, Ning X, Shu H, Coatrieux JL (2016) Fully automatic coronary calcification detection in non-contrast CT images. Med Phys 43(5):2174–186
11. Durlak F, Wels M, Schwemmer C, Sühling M, Steidl S, Maier A (2017) Growing a random forest with fuzzy spatial features for fully automatic artery-specific coronary calcium scoring. In: International workshop on machine learning in medical imaging, Springer, pp 27–35
12. de Vos BD, Wolterink JM, Leiner T, de Jong PA, Lessmann N, Išgum I (2019) Direct automatic coronary calcium scoring in cardiac and chest CT. IEEE Trans Med Imaging 38(9):2127–2138
13. Li P, Xu L, Yang L, Wang R, Hsieh J, Sun Z et al (2018) Blooming artifact reduction in coronary artery calcification by a new de-blooming algorithm: initial study. Sci Rep 8(1):1–8
14. Singh G, Al'Aref SJ, Van Assen M, Kim TS, van Rosendael A, Kolli KK et al (2018) Machine learning in cardiac CT: basic concepts and contemporary data. J Cardiovasc Comput Tomogr 12(3):192–201
15. Henglin M, Stein G, Hushcha PV, Snoek J, Wiltschko AB, Cheng S (2017) Machine learning approaches in cardiovascular imaging. Circulation 10(10):e005614
16. Siegersma K, Leiner T, Chew D, Appelman Y, Hofstra L, Verjans J (2019) Artificial intelligence in cardiovascular imaging: state of the art and implications for the imaging cardiologist. Netherlands Heart J 1–11
17. Išgum I, Rutten A, Prokop M, van Ginneken B (2007) Detection of coronary calcifications from computed tomography scans for automated risk assessment of coronary artery disease. Med Phys 34(4):1450–1461
18. Kurkure U, Chittajallu DR, Brunner G, Le YH, Kakadiaris IA (2010) A supervised classification-based method for coronary calcium detection in non-contrast CT. Int J Cardiovasc Imaging 26(7):817–828
19. Xie Y, Cham MD, Henschke C, Yankelevitz D, Reeves AP (2014) Automated coronary artery calcification detection on low-dose chest CT images. In: Medical imaging 2014: computer-aided diagnosis, vol 9035. International Society for Optics and Photonics, p 90350F
20. Wolterink JM, Leiner T, Takx RA, Viergever MA, Išgum I (2014) An automatic machine learning system for coronary calcium scoring in clinical non-contrast enhanced, ECG-triggered cardiac CT. In: Medical imaging 2014: computer-aided diagnosis, vol 9035. International Society for Optics and Photonics, p 90350E
21. Wolterink JM, Leiner T, de Vos BD, van Hamersvelt RW, Viergever MA, Išgum I (2016) Automatic coronary artery calcium scoring in cardiac CT angiography using paired convolutional neural networks. Med Image Anal 34:123–136
22. Lessmann N, van Ginneken B, Zreik M, de Jong PA, de Vos BD, Viergever MA et al (2017) Automatic calcium scoring in low-dose chest CT using deep neural networks with dilated convolutions. IEEE Trans Med Imaging 37(2):615–625
23. Shadmi R, Mazo V, Bregman-Amitai O, Elnekave E (2018) Fully-convolutional deep-learning based system for coronary calcium score prediction from non-contrast chest CT. In: IEEE 15th international symposium on biomedical imaging (ISBI 2018), IEEE, pp 24–28
24. Šprem J, De Vos BD, Lessmann N, Van Hamersvelt RW, Greuter MJ, De Jong PA et al (2018) Coronary calcium scoring with partial volume correction in anthropomorphic thorax phantom and screening chest CT images. PLoS ONE 13(12):e0209318
25. Zreik M, Van Hamersvelt RW, Wolterink JM, Leiner T, Viergever MA, Išgum I (2018) A recurrent CNN for automatic detection and classification of coronary artery plaque and stenosis in coronary CT angiography. IEEE Trans Med Imaging 38(7):1588–1598
26. Zreik M, van Hamersvelt RW, Wolterink JM, Leiner T, Viergever MA, Isgum I (2018) Automatic detection and characterization of coronary artery plaque and stenosis using a recurrent convolutional neural network in coronary CT angiography
27. Cano-Espinosa C, González G, Washko GR, Cazorla M, Estépar RSJ (2018) On the relevance of the loss function in the Agatston score regression from non-ECG gated CT scans. In: Image analysis for moving organ, breast, and thoracic images, Springer, pp 326–334
28. Qian Z, Marvasty I, Rinehart S, Voros S (2011) A lesion-specific coronary artery calcium quantification framework for the prediction of cardiac events. IEEE Trans Inf Technol Biomed 15(5):673–680
29. Isgum I, Prokop M, Niemeijer M, Viergever MA, Van Ginneken B (2012) Automatic coronary calcium scoring in low-dose chest computed tomography. IEEE Trans Med Imaging 31(12):2322–2334
30. Six O, Quantib B (2018) The ultimate guide to AI in radiology
31. Christler A, Felföldi E, Mosor M, Sauer D, Walch N, Dürauer A et al (2020) Semi-automation of process analytics reduces operator effect. Bioprocess Biosyst Eng 43(5):753–764
32. Buckee GK, Hickman E (1975) A review of automated and semi-automated analysis in brewing. J Inst Brew 81(5):399–407
33. Brahim W, Mestiri M, Betrouni N, Hamrouni K (2016) Semi-automated rib cage segmentation in CT images for mesothelioma detection. In: International image processing, applications and systems (IPAS). IEEE, pp 1–6

34. Breiman L (2001) Random forests. Mach Learn 45(1):5–32
35. Mohammadi S, Hedjazi A, Sajjadian M, Ghoroubi N, Mohammadi M, Erfani S (2016) Study of the normal heart size in Northwest part of Iranian population: a cadaveric study. J Cardiovasc Thorac Res 8(3):119
36. Sonka M, Hlavac V, Boyle R (2014) Image processing, analysis, and machine vision. Cengage Learning
37. De Vos BD, Wolterink JM, De Jong PA, Viergever MA, Išgum I (2016) 2D image classification for 3D anatomy localization: employing deep convolutional neural networks. In: Medical imaging 2016: image processing, vol 9784, International Society for Optics and Photonics, p 97841Y
38. Freund Y, Schapire R, Abe N (1999) A short introduction to boosting. J Jpn Soc Artif Intell 14(771–780):1612
39. Zayed N, Elnemr HA (2015) Statistical analysis of haralick texture features to discriminate lung abnormalities. Int J Biomed Imaging
40. Roffo G (2016) Feature selection library (MATLAB toolbox). arXiv:160701327
41. Termeer M, Bescós JO, Breeuwer M, Vilanova A, Gerritsen F, Gröller E (2008) 1104 The volumetric bull's eye plot. J Cardiovasc Magn Reson 10:1–3
42. McCollough CH, Ulzheimer S, Halliburton SS, Shanneik K, White RD, Kalender WA (2007) Coronary artery calcium: a multi-institutional, multimanufacturer international standard for quantification at cardiac CT. Radiology 243(2):527–538
43. Rumberger JA, Brundage BH, Rader DJ, Kondos G (1999) Electron beam computed tomographic coronary calcium scanning: a review and guidelines for use in asymptomatic persons. In: Mayo Clinic Proceedings, vol 74, Elsevier, pp 243–252
44. Giavarina D (2015) Understanding Bland Altman analysis. Biochem Med 25(2):141–151