Classification of calcified regions in atherosclerotic lesions of the carotid artery in computed tomography angiography images

Danilo Samuel Jodas · Aledir Silveira Pereira · João Manuel R. S. Tavares*

Abstract The identification of atherosclerotic plaque components, extraction and analysis of their morphology represent an important role towards the prediction of cardiovascular events. In this article, the classification of regions representing calcified components in Computed Tomography Angiography (CTA) images of the carotid artery is tackle. The proposed classification model has two main steps: the classification per pixel and the classification per region. Features extracted from each pixel inside the carotid artery are submitted to four classifiers in order to determine the correct class, i.e. calcification or non-calcification. Then, geometrical and intensity features extracted from each candidate region resulting from the pixel classification step are submitted to the classification per region in order to determine the correct regions of calcified components. In order to evaluate the classification accuracy, the results of the proposed classification model were compared against ground truths of calcifications obtained from micro Computed Tomography images of excised atherosclerotic plaques that were registered with *in vivo* CTA images. The average values of the Spearman correlation coefficient obtained by the Linear Discriminant Classifier were higher than 0.80 for the relative volume of the calcified components. Moreover, the average values of the absolute error between the relative volumes of the classified calcium regions and the ones calculated from the corresponding ground truths were lower than 3%. The new classification model seems to be adequate as an auxiliary diagnostic tool for identifying calcifications and allowing their morphology assessment.

1 Introduction

Cardiovascular diseases represent one of the main causes of the increasing number of deaths around the world. Therefore, the early diagnosis of pathological conditions is important to minimize clinical cases such as heart attacks, transient ischemic attacks and even the occurrence of strokes. In a broader research study, Mendis et al [1] revealed alarming numbers regarding cardiovascular disease prevention and control: according to World Health Organization, cardiovascular diseases represented 31% of deaths of people in the world in 2011; and in 2008, cardiovascular diseases were responsible for the death of more than 17 millions of people with less than 60-year-old [1].

Atherosclerosis is an underlying disease responsible for the occurrence of heart attacks and strokes. Atherosclerotic plaques are formed when fatty material and cholesterol are deposited inside the wall of the artery. In an advanced stage, the plaques are composed by lipid core, fibrous tissue, intraplaque hemorrhage and calcifications. Atherosclerosis reduces the blood flow through

Danilo Samuel Jodas
CAPES Foundation - Ministry of Education of Brazil, Brasilia - DF, 70040-020, Brazil
Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465, Porto, Portugal E-mail: danilojodas@gmail.com

Aledir Silveira Pereira
Universidade Estadual Paulista "Júlio de Mesquita Filho", Rua Cristóvão Colombo, 2265, 15054-000, S. J. do Rio Preto, Brazil E-mail: aledir@sjrp.unesp.br

*João Manuel R. S. Tavares (Corresponding author)
Instituto de Ciência e Inovação em Engenharia Mecânica e Engenharia Industrial, Faculdade de Engenharia, Universidade do Porto, Rua Dr. Roberto Frias, s/n, 4200-465, Porto, Portugal E-mail: tavares@fe.up.pt

the artery, leading to serious complications such as heart attacks and strokes. Hence, in order to prevent such risks, as appropriate treatment or rehabilitation plan should be designed based on imaging exams.

Technological advances in computational systems for imaging-based diagnosis allow the detection and assessment of atherosclerotic lesions. Computed Tomography (CT), Magnetic Resonance (MR) and ultrasound are examples of less invasive imaging modalities that have been widely used in evaluating the presence and assess the morphology of atherosclerotic plaques [2, 3, 4]. Although x-ray angiography allows the assessment of the lumen diameter, this imaging modality does not provide enough quality to identify the components of atherosclerotic plaques [5]. Currently, the correct diagnosis of atherosclerotic plaque components is performed based on images of the carotid artery acquired from ultrasound, CT and MR examinations in a noninvasive way, which allows the visualization of such plaques and the identification of the associated components.

Previous studies [6, 7, 8, 9, 10, 11] have confirmed the importance of the atherosclerotic plaque components in evaluating the risks of cerebrovascular diseases. The assessment of the atherosclerotic plaque composition is important to identify risks related to plaque rupture and embolization, as well as risks to transient ischemic attacks, amaurosis fugax and strokes. Although visual analysis has been proposed for quantifying the atherosclerotic plaque composition in images, the intra and intervariability between experts might impair the correct diagnosis. Therefore, the development of computational algorithms plays an important role to expedite the assessment of atherosclerotic plaques and avoid the intervariability between experts.

Computational algorithms have been proposed to segment atherosclerotic plaques and associated components in images [12], which allows the assessment of the plaques in order to predict the risk of cardiovascular and cerebrovascular diseases more quickly. Techniques of image processing, clustering and supervised classification are examples of computational approaches suggested in several studies for identifying the main components of atherosclerotic plaques in a semi-automatic or fully automatic way [12]. Moreover, advances in dynamic learning approaches have been made to adapt the classification of the input data in the context of dynamic scenarios. Multi-armed bandits are especial cases of algorithms where the learning of features of the input data is typically dynamic and based on the knowledge of the evaluated environment. It is important on the scope of recommendation systems such as advertising, videos and movies as reported in the studies of Gentile et al [13], Li et al [14] and Korda et al [15], but could

also be suitable for dynamic evaluation of medical images as proposed in Song et al [16] and Gutiérrez et al [17]. Concerning the performance measurements of classification tasks, several studies have also been carried out to address the relative frequency, i.e quantification of each class in the context of classification assignments. Quantification is important to accurately calculate the amount of samples that belong to each class regardless the imbalance of the data and the dynamic changes of the number of samples belonging to each class in the testing set. An online stochastic method was proposed by Kar et al [18] to tackle the balance between classification and quantification performance.

According to several studies, the presence of calcifications represents an advanced stage of atherosclerosis [19] [20]. Therefore, this article proposes the classification of calcified components in CTA images of the carotid artery. Briefly, the proposed classification model has two main steps: the classification per pixel and the classification per region. In the first step, intensity and distance features extracted from each pixel inside the carotid artery are submitted to four classifiers in order to obtain the candidate regions representing calcified components. Thereafter, geometrical and intensity features extracted from each candidate region are submitted to the classification per region step in order to determine the regions corresponding to true calcifications.

One of the main contributions of this study is the ability of the proposed model to overcome possible classification errors induced by misalignments of the registration between micro CT images of excised atherosclerotic plaques and related *in vivo* CTA images. Briefly, the used ground truths of the calcified components were obtained from micro CT images, including parts of the lumen and other regions with intensity similar to calcifications in CTA images that may affect the classification results, leading to erroneously determination of the calcium regions. The novelty of the proposed classification model is the inclusion of a second classification step that has the ability to handle the regions identified in the pixel classification step with intensity similar to calcifications of the atherosclerotic plaques. Geometrical features such as the area of the classified region and its relative percentage with respect to the carotid wall have particular aspects in calcified components when compared to non-calcium regions. In fact, the inclusion of the classification per region step provided results significantly better than the ones exclusively obtained by the classification per pixel step. In a practical application point-of-view, the proposed classification model is suitable as a computer aided diagnosis tool that might help physicians in expediting the identification of the

calcium regions for establishing a better treatment plan. In addition, since the presence of outliers and the distances of each pixel to the contours of the lumen and carotid artery under analysis can influence the final classification results, a study on these factors was also performed.

This article is organized as follow: Section 2 introduces previous studies related to the classification of atherosclerotic plaque components in images acquired by well-know imaging modalities. Section 3 presents a description of the proposed classification model. Section 4 presents the classification results obtained by the proposed classification model and the comparison between these results and the corresponding ground truths. The advantages and limitations of the proposed model are discussed in Section 5. Finally, the conclusions are drawn in Section 6.

2 Previous studies

Several studies have been proposed to identify atherosclerotic plaque components in MR and CTA images of carotid and coronary arteries. Vukadinovic et al [21] proposed the segmentation of calcium regions in atherosclerotic plaques of the carotid artery in CTA images based on the following steps: a level set approach is used to segment the lumen of the carotid artery in the CTA image under analysis; then, features of the candidate calcium regions are extracted and used to classify them as belonging to calcium or non-calcium components; afterwards, features are extracted to classify the pixels as inside or outside the carotid wall; and then, an ellipse fitting procedure is used to detect the carotid wall boundary. In this work, the lumen and calcium regions are combined to delineate the contour of the carotid wall since the calcium appears in the inner region of the carotid artery.

van Engelen et al [22] proposed the classification of atherosclerotic plaque components in CTA and MR images of the carotid artery. The main objective of this study was to handle the misalignments between *in vivo* and histological images of atherosclerosis by measuring the probability and Dice overlap of each voxel relatively to the corresponding ground truth. The combination of features extracted from each voxel of the CTA and MR images was also addressed. After the registration of the *in vivo* with histological and micro CT images, each component of the ground truth was binarized and blurred with a Gaussian filter in order to create soft labels that indicate the probability of each voxel to belonging to an atherosclerotic plaque component. Additionally, the rejection of outliers was also performed to address the misalignments of the plaque components

manually delineated in histological images with the corresponding CTA and MR images. For the classification of the plaque components into calcification, fibrous tissue and lipid-rich necrotic core, 24 features were extracted from each voxel and submitted to a Linear Discriminant Classifier, which provided better results than a Support Vector Machine (SVM) with a Radial Basis Function (RBF) kernel.

Wintermark et al [23] proposed the identification of atherosclerotic plaque components in CTA images of the carotid artery based on the analysis of the Hounsfield Unit (HU) values. The HU values obtained from a 2x2 mm square template centred at each pixel of the CTA image under analysis are used in a linear mixed model to obtain the appropriate mean HU values for each atherosclerotic plaque component. In terms of the classification results, the calcium regions classified based on the intensity obtained from the CTA images were in perfect accordance with the corresponding ground truths manually delineated in histological images.

de Graaf et al [24] proposed an automatic method to identify the atherosclerotic plaque components in CTA images of the coronary artery. The extraction of each component was performed by using two approaches: fixed threshold and dynamic threshold. The first approach is based on fixed HU values for extracting each plaque component, whereas the second one defines the cut-off values based on the luminal intensity. The dynamic threshold is based on the fact that the lower the luminal intensity is, the lower will be the HU value of the atherosclerotic plaque. Thus, the HU values of each plaque component are defined based on the luminal attenuation.

The main limitation of the above-mentioned studies regards the usage of fixed HU values for identifying the atherosclerotic plaque components, as in Vukadinovic et al [21], Wintermark et al [23] and de Graaf et al [24]. Additionally, the identification of calcified regions in atherosclerotic lesions can be more effective when features extracted from CTA images are used in the classification process. Hence, instead of using images acquired by different imaging modalities as suggested in van Engelen et al [22], the proposed approach only uses CTA images.

3 Materials and Methods

3.1 CTA images used

The images of the carotid artery selected for this study were previously used in a research by van Engelen et al [22] and kindly provided by the authors on request. The

proposed classification method was performed on images that are regions of interest surrounding the carotid arteries. A registration procedure was previously performed to align the original MR images with the corresponding histological images, which only contained the region of the arteries under study [22]. Once the alignment was completed, the MR images were cropped to only obtain the part that matched the histological images [22]. The original dataset was composed of five MRI scans acquired from thirteen patients: T1-weighted (T1W), Proton Density Weighted (PDW), Time-of-Flight (TOF) and two 3D-T1W scans. The first three MRI scans were acquired without administration of intravenous contrast media, whereas the 3D-T1W scans were acquired with and without contrast media. The post-contrast 3D-T1W scan was performed 4.6±3.4 minutes after the administration of the contrast media. Each MR imaging scan is composed of approximately 17.7±4.8 slices per patient; each slice has a pixel size of 0.25 mm x 0.25 mm. CTA images were also acquired to provide details for the registration of these images with the histological images and to facilitate the manual segmentations of the lumen, vessel wall and plaque components. The manual delineations of the lumen and carotid wall in the CTA and MR images were also provided. As already mentioned, the used images were acquired for a previous study that was approved by the Medical Committee of the Erasmus Medical Center; all the evaluated patients provided written consensus for the use of the images [22]. More details about the MR and CTA imaging datasets are available in van Engelen et al [22].

From the original image dataset, we used all CTA images with their corresponding ground truth of the atherosclerotic plaque components and manual delineations of the contours of the lumen and carotid wall made by one expert. In total, 230 CTA images are available with the dataset provided. The ground truths of the atherosclerotic plaque components are available in 177 CTA images and the manual delineations of the lumen and carotid wall are available in 184 CTA images. The manual delineations of the fibrous tissue and lipid-rich necrotic core were made on the histological images, whereas the ground truths of the calcified components were obtained by using a fixed threshold value on the micro CT images acquired from the excised plaques [22].

3.2 Proposed model

The proposed classification model has two main steps, as depicted in Fig. 1.

The classification per pixel represents the first step of the proposed model and it consists in classifying each pixel inside the carotid wall provided by the ground truth as belonging or not to calcifications. Features extracted from the following CTA images are used in this step: the original image, the original image after been smoothed by a Gaussian filter, the original image after been smoothed by a mean filter and the same input image after the application of a Sigmoid filter. The second step consists in classifying the regions resultant from the previous step in order to determine the ones that represent true calcifications of the atherosclerotic lesions. Hence, geometrical and intensity features of the regions obtained by the first step are extracted and submitted to the second step which performs the classification per region.

3.2.1 Feature extraction

Feature extraction plays an important role in the accuracy of the classification result. Calcifications are characterized as regions having the highest intensity in CTA images. However, features extracted from the lumen region represent a challenge for classifying the calcified components due to the similarity of their intensity values. Moreover, small regions corresponding to image noise could also be classified as calcifications. Hence, the model proposed in this study takes into account the intensities of the calcium and lumen regions, as well as geometrical features extracted from the candidate regions previously classified in the pixel classification step.

The following filters are applied to highlight the calcium regions relatively to other structures present in the original CTA images: a Gaussian filter with a standard deviation σ; a mean filter with a NxN neighbourhood; and a Sigmoid filter that is applied to highlight a range of intensities and attenuate the intensities outside this range. The Gaussian and mean filters are used to smooth the original CTA images in order to remove noise artefacts, whereas the Sigmoid filter is employed to improve the contrast of regions having high intensities. The Sigmoid filter is based on a pixel-wise function defined as:

$$f(x) = min + \frac{max - min}{1 + e^{\frac{\beta - x}{\alpha}}}, \qquad (1)$$

where min and max are the minimum and maximum intensities of the resultant image $f(x)$, respectively, and α and β are enhancing parameters defined according to the intensities of the structure to be enhanced. Examples of CTA images obtained after the application of the above-mentioned imaging filters are illustrated in Fig. 2.
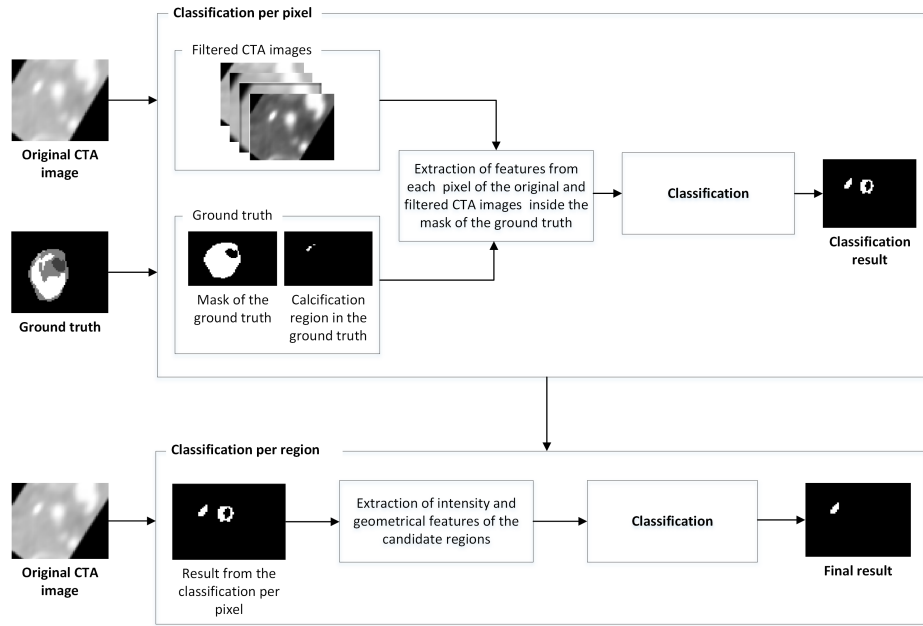
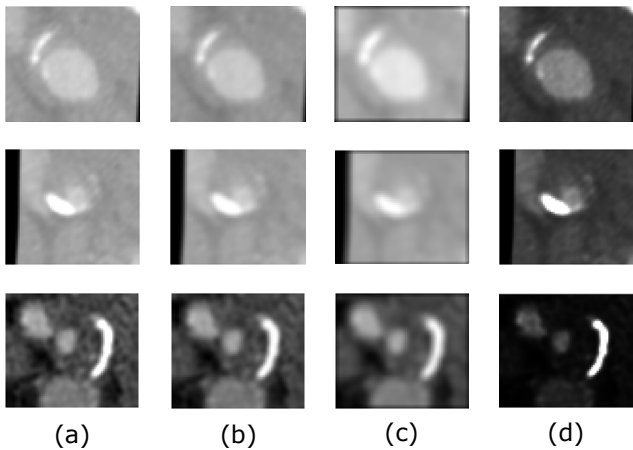**Fig. 1** Diagram of the proposed classification model.

Fig. 2 Examples of images resulting from the Gaussian, mean, and Sigmoid filters: a) Original CTA images; b) Images resulting from the original images after applying a Gaussian filter with $\sigma = 2$, c) a mean filter with $N = 3$, and d) a Sigmoid filter with $\alpha = 50$ and $\beta = 256$.

The brightest regions corresponding to possible calcifications are enhanced by the application of the Sigmoid filter, as one can realize in Fig. 2d. For the pixel classification step, the intensity features obtained from each pixel of the original and filtered CTA images are used as inputs for the classifiers. Table 1 indicates the features used in the two steps of the proposed classification model.

Two additional features are used in the pixel classification step: the distances of the pixel under analysis to the boundaries of the lumen and wall of the carotid artery. In addition, the average intensities of the lumen and carotid wall regions in the original CTA image are also used.

3.2.2 Outliers removal

The detection of outliers represents an important task in data analysis and one of the most important pre-processing steps for improving the robustness, performance and accuracy of a classification model. An outlier is characterized as an observation that is far from the remainder ones in a dataset. The presence of outliers often decreases the performance and accuracy of the used classifier due to the increase of the observations variance. Hence, the removal of outliers plays an important role to reduce the observations variance and improve the accuracy of the associated classification model.

Regarding the classification of calcified regions, the outliers are often related to pixels corresponding to misalignments of the histological images with the MR and CTA images resultant from the registration step described in van Engelen et al [22]. Therefore, the identification and removal of these pixels from the training and testing sets were tackled.

The boxplot analysis is an important statistical tool commonly used to evaluate the distribution and variability of the observations under study. Additionally, abnormal observations can also be identified by means of boxplots. An approach based on boxplots is proposed in this study to detect and remove outliers before the training phase of the classification process. Firstly, the examples of the training set are separated into calcifi-

Table 1 Features used in the proposed classification model.

Id.	Classification per pixel	Id.	Classification per region
F1	Intensity of the original image	F21	Percentage of the area relatively to the carotid wall
F2	Intensity of the image (Gaussian filter)	F22	Area of the region
F3	Intensity of the image (Mean filter)	F23	Average intensity of the lumen †
F4	Intensity of the image (Sigmoid filter)	F24	Average intensity of the carotid wall †
F5	Average intensity (Original image)*	F25	Distance of the centroid of the region to the lumen
F6	Average intensity (Gaussian filter)*	F26	Distance of the centroid of the region to the carotid wall
F7	Average intensity (Mean filter)*		
F8	Average intensity (Sigmoid filter)*		
F9	Minimum intensity (Original image)*		
F10	Minimum intensity (Gaussian filter)*		
F11	Minimum intensity (Mean filter)*		
F12	Minimum intensity (Sigmoid filter)*		
F13	Maximum intensity (Original image)*		
F14	Maximum intensity (Gaussian filter)*		
F15	Maximum intensity (Mean filter)*		
F16	Maximum intensity (Sigmoid filter)*		
F17	Average intensity of the lumen †		
F18	Average intensity of the carotid wall †		
F19	Distance of the pixel to the lumen contour		
F20	Distance of the pixel to the carotid wall contour		

*Features extracted from a $3x3$ neighbourhood centred at each pixel of the original and filtered CTA images.
†Features extracted from the original CTA image.

cations and non-calcifications. Then, boxplots of each intensity feature extracted from the pixels of the input CTA image that are inside the carotid wall are generated and combined in order to provide the distribution of the training examples for all features used. An illustration of the combined boxplots built is presented in Fig. 3.

From the boxplots in Fig. 3, it is possible to notice that the outliers of the features F4, F8, F12 and F16 (see Table 1) in the non-calcified components represented by the highest intensity values might belong to calcified regions (Fig. 3b). Likewise, the outliers in the calcified components represented by the lowest intensity values and the ones of the feature F12 represented by the highest intensity values (Fig. 3a) may belong to non-calcified regions. The outliers belonging to each class, i.e. calcification and non-calcification subsets, are evaluated in order to find the ones that are present in all features. Then, these outliers are removed from the corresponding subset. The boxplot analysis continues until all outliers have been removed. A new training set without the outliers found is generated and submitted to the classification model.

3.2.3 Training and testing of the classifiers

After selecting the features and removing the outliers, a set of classifiers is used for the classification model. Decision tree, Support Vector Machine (SVM), Naive Bayes and Linear Discriminant Classifier (LDC) were selected to perform the classification of the pixels and,

subsequently, the candidate regions representing possible calcified regions in atherosclerotic lesions.

Decision Trees are one of the most simple and effective models used in inductive inference. A decision tree is trained according to a training set and then, other examples are classified according to the same tree model. The graphical representation of the decision tree is composed of lines that are used to identify the decision to be made, for example, "yes" or "no", and nodes to identify the issues to be decided. Each branch formed by lines and nodes ends in a leaf node that identifies the most likely consequence of the sequence of decisions made. The choice of Decision Tree is due to its simple implementation, performance and effectiveness in classifying problems involving sequential decisions.

Support Vector Machines have been adopted for the classification of patterns into two classes separated by a decision hyperplane. The decision hyperplane is a surface that separates the input data into two classes. The goal of a SVM is to find an optimal decision hyperplane that keeps a maximal gap between the examples under analysis. Building an optimal hyperplane as a decision surface is a fundamental step for increasing the separation between the examples to be classified. However, problems that are not linearly separable are common in many classification problems. Hence, the mapping of the nonlinear training set into a linearly separable space is necessary in order to determine an optimal decision hyperplane. The linearly separable space is called feature space, which is generated by a kernel function that
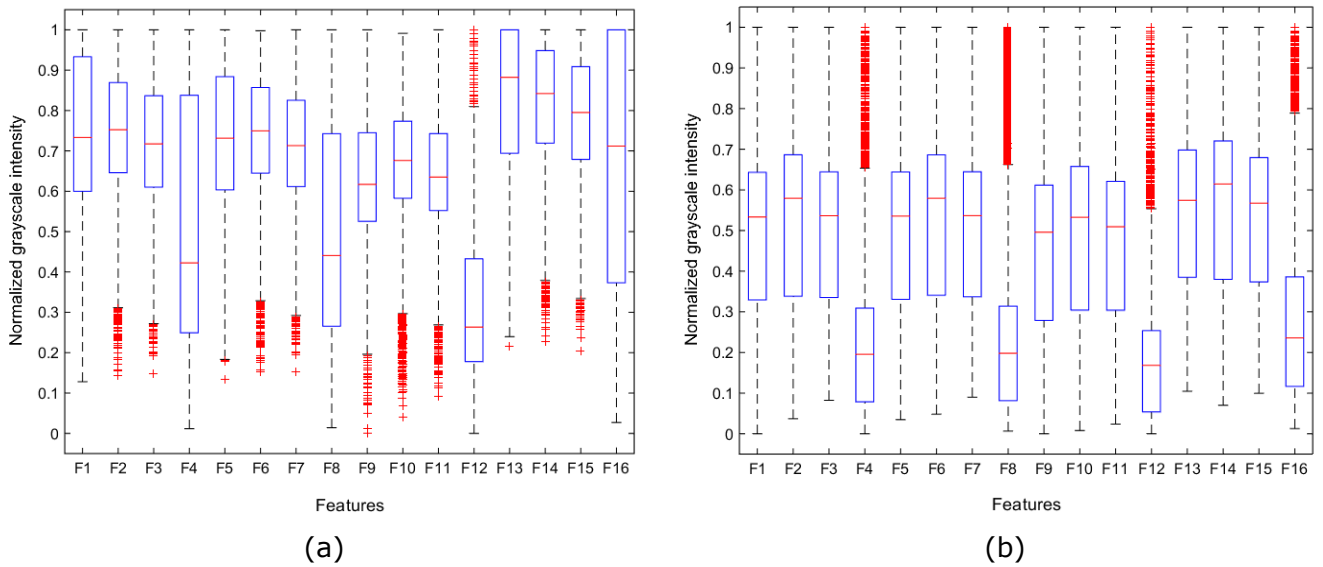
Fig. 3 Example of outliers detected by using boxplots built from all features and pixels of the images used in the classification model: a) Boxplot of the calcified examples with the outliers represented in red; b) Boxplot of the non-calcified examples with the outliers represented in red. (The identification of each feature is indicated in Table 1 in column "Classification per pixel".)

maps the input examples of the nonlinear space onto a space where the examples are linearly separable. Linear, polynomial, Radial Basis Function (RBF) and Hyperbolic tangent are examples of kernel functions used in several classification problems tackled by SVMs. SVMs are suitable for binary classification problems. In this study, a SVM was used with RBF, Sigmoid, Linear and Polynomial based kernels, which are defined as:

$$\mathbf{RBF} = k(x,y) = exp\left(-\frac{\|x-y\|^2}{2\sigma^2}\right) \qquad (2)$$

$$\mathbf{Sigmoid} = k(x,y) = tanh(\alpha x^T y + c) \qquad (3)$$

$$\mathbf{Linear} = k(x,y) = x^T y + c \qquad (4)$$

$$\mathbf{Polynomial} = k(x,y) = (\alpha x^T y + c)^d \qquad (5)$$

The Naive Bayes classifier is one of the most simple and effective probabilistic models used in many types of classification problems. Given an example x in the training phase, the Naive Bayes classifier calculates a distribution $Pr(x|c)$ for each class $c = \{-1, 1\}$, which represents the probability of example x to belong to class c. In the testing phase, the distribution with the highest probability generated from each example is calculated. The conditional probability used in the Naive Bayes classifier is defined as:

$$P(c_i|x) = \frac{p(c_i)p(x|c_i)}{p(x)}, \qquad (6)$$

where $P(c_i|x)$ is the posterior probability, i.e. the probability of example x belonging to class c_i, $p(c_i)$ is the probability of occurring class c_i, $p(x|c_i)$ is the probability of occurring example x given class c_i and $p(x)$ is the probability of occurring example x. Class c_i with the maximum posterior probability is assigned to the input example x. The Naive Bayes classifier can also be extended to examples having more than one feature; in such cases, example x is represented by a vector of features. Prior probabilities $p(x|c_i)$ of the features are calculated and combined to obtain posterior probability $P(c_i|x)$ of example x.

The LDC is a simple technique used for detecting the group or class having the average value closer to the one of the test examples. The average of each group is obtained from the features of all examples belonging to that group [25]. The LDC is defined as:

$$\delta_k(x) = x^T \sum{}^{-1} \mu_k - \frac{1}{2}\mu_k^T \sum{}^{-1} \mu_k + log(\pi_k), \qquad (7)$$

where $\delta_k(x)$ is the posterior probability, i.e. the probability of example x belonging to class k, k is the class, x is the vector of features, $\sum$ is the covariance matrix and μ_k is the mean of class k. Likewise the Naive Bayes

classifier, the example represented by feature vector x is assigned to class k that have the maximum posterior probability $\delta_k(x)$.

As already mentioned, the proposed classification approach has two main steps: the first step concerns the classification of each pixel of the original and filtered CTA images inside the mask of the ground truth, whereas in the second step the classification of the candidate regions resulting from the first step is performed to identify the ones that correspond to true calcified regions. Hence, two models were developed for each step of the classifier under analysis.

Leave-one-out and k-fold cross-validation were performed in the CTA images of the thirteen patients under study. The leave-one-out cross-validation was applied in repeatedly steps, with the images of 12 patients used in the training step and the images of the remaining 13th patient used for testing the classifiers. The k-fold cross-validation method consists in dividing the dataset into mutually exclusive subsets of the same size. A subset is used for testing the classifier and the $k - 1$ subsets are employed for building the model of the classifier. This process is performed k times by alternating the subset of observations used for testing the classifier.

In order to evaluate the influence of the distance features and of the outliers removal process, the approaches indicated in Table 2 were studied.

Two new sets of training and testing examples were generated from each approach indicated in Table 2 for each iteration of the leave-one-out and k-fold cross-validation techniques. Additionally, the outliers removal was performed in each iteration of the two cross-validation techniques. Hence, four models of each classifier were built in the training step of the pixel classification step.

For the training of the classifiers in the pixel classification step, a binary image of the carotid wall obtained from the ground truth corresponding to the atherosclerotic plaque components was built for selecting only the pixels of the CTA images that are inside the carotid artery. Then, the features extracted from each pixel were submitted to the classifier corresponding to each approach indicated in Table 2.

As to the region classification step, the region corresponding to the lipid-rich necrotic core was used as reference for the non-calcified class. The choice of the lipid-rich necrotic core was due to the intensity and geometrical features that differ from the calcium components in CTA images, making it appropriate for a binary classification in the region classification step. The features of the regions of the binary images resultant from each approach of the pixel classification step were extracted and submitted to the classifiers that perform the classification of each potential calcified region.

Regarding the testing of the classifiers under analysis, the pixels inside the manually delineated carotid wall were used for testing and validating the accuracy of the chosen classifiers. The extraction of the pixels inside the binary image representing the carotid wall is illustrated in Fig. 4.
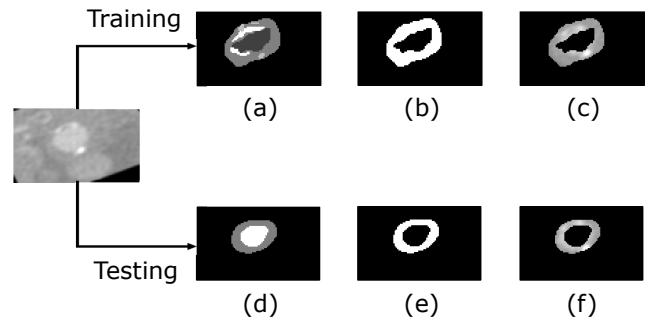

Fig. 4 Illustration of the procedure used to select the pixels inside the carotid wall for extracting the features to be used in the pixel classification step of the proposed classification model: a) The ground truth of the atherosclerotic plaque components; b) Binary image of the ground truth with the carotid wall represented in white; c) Part of the original CTA image inside the carotid wall; d) The manual delineations of the lumen and carotid wall; e) Binary image obtained from the manual delineations with the carotid wall represented in white; f) Part of the original CTA image inside the carotid wall.

For the pixel classification step, the intensity and distance features described in Table 1 were extracted from the original and filtered CTA images belonging to the testing set for the classification process. The intensities of each pixel of the original CTA image inside the previously identified regions along with geometrical features of each candidate region were then submitted to the classifiers that perform the classification per region.

4 Results

The accuracy of the classification model proposed in this study was evaluated by means of the absolute error and Spearman correlation between the areas and volumes of the classified calcium regions and those calculated from the corresponding ground truths. The relative and absolute areas and volumes of the calcium regions were calculated in order to evaluate the results generated by the classification model. The relative area and volume represent the percentage of the calcium region occupied in the total area and volume of the carotid artery. The absolute area and volume are the total area and volume of the related calcium region, respectively. In image processing and analysis, the correlation coefficient is an important index to assess the accu-

Table 2 Approaches adopted to verify the influence of the distance features and of the outliers removal process.

	Outlier removal	Usage of the distances to the lumen and carotid wall
Approach 1	No	Yes
Approach 2	No	No
Approach 3	Yes	Yes
Approach 4	Yes	No

racy of computational methods in identifying structures in images. By comparing the computationally identified structure with the corresponding manual delineation in the image under analysis, the correlation coefficient provides a value that represents the accuracy of the automatic or semi-automatic method. The higher is the value of the coefficient, more accurate is the computational method. Concerning the method proposed in this article, the Spearman correlation coefficient was used to compare our results with the ones reported in previous studies. The evaluation by means of the Area Under the Curve (AUC), sensitivity and specificity was not possible due to the existent misalignment between the ground truths obtained from the micro CT images and the corresponding *in vivo* CTA images.

The following parameters were defined to perform the classification process: the value of parameter σ of the Gaussian filter was defined as equal to 2; the values of α and β parameters of the Sigmoid filter were defined as equal to 50 and 256, respectively; neighbourhood NxN of the mean filter was set to 3x3; the number of iterations to remove outliers of the training and testing sets was defined as equal to 50; the value of k of the k-fold cross-validation approach was set to 10 since this is the most used value in several classification problems; the values of σ and of penalty term C of the RBF kernel were defined as equal to 1 (one) since these values have been commonly used in several studies; the best order d of the polynomial kernel was defined as equal to 3; tests using the Sigmoid kernel with the value of parameter α varying from 0.5 to 1 and parameter C equal to -1 were performed in each iteration of the k-fold and leave-one-out validation techniques in order to evaluate the classification performance. The choice of the SVM kernel depends on the linear separation of the data under analysis. The linear kernel is suitable for datasets where the groups to be classified are linearly separable for the input features. However, non-linear and sparse datasets are more difficult to separate using a linear plane. Fig. 5 illustrates the sparsity of the data corresponding to the calcium and non-calcium classes with respect to the intensity features indicated in Table 1.

The non-linear behaviour of the data depicted in Fig. 5b and Fig. 5c could make the SVM classification more prone to errors. Therefore, the use of the RBF, Sigmoid and polynomial kernels was an attempt to evaluate the classification results for the non-linear data.

The average absolute errors of the relative areas and volumes for each parameter α tested in the Sigmoid kernel by means of the leave-one-out validation technique are indicated in Tables 3 and 4, respectively.

As shown and emphasized in Tables 3 and 4, the lowest average absolute errors were obtained when parameter α was equal to 0.5 in the majority of the performed tests. Therefore, the remaining discussion of the Sigmoid kernel results is made with the value of parameter α equal to 0.5.

Examples of the classification results per pixel and per region obtained by each classifier under analysis are shown in Fig. 6.

As shown in Fig. 6a, the lumen region of the CTA image under study overestimated the size of the lumen of the related ground truth obtained from the manual delineation, causing that part of the lumen in the input image be used in the pixel classification step. Hence, the referred part of the lumen was classified as calcification due to the similarity of its grayscale intensity with the grayscale intensity of true calcified regions. In these cases, the region classification step plays an important role to evaluate the geometrical properties of the candidate regions previously classified by the pixel classification step in order to select those that represent true calcifications. After the classification per region, the mentioned part of the lumen was disregarded and only the region corresponding to a calcification remained in the final classification result. The results provided by the SVM with the RBF and polynomial kernels, as well as the results of the LDC classifier, are also in accordance with the ground truth as shown in Fig. 6a. The examples illustrated in Fig. 6b-c also show the good accordance of the results obtained by most of the classifiers under analysis with the corresponding ground truths.

Examples resultant from each approach indicated in Table 2 in the pixel classification step are shown in Fig. 7.

The classification results were improved after the outliers removal as can be perceived in Fig. 7(d-e), mainly in the case of the CTA image shown in the third row.
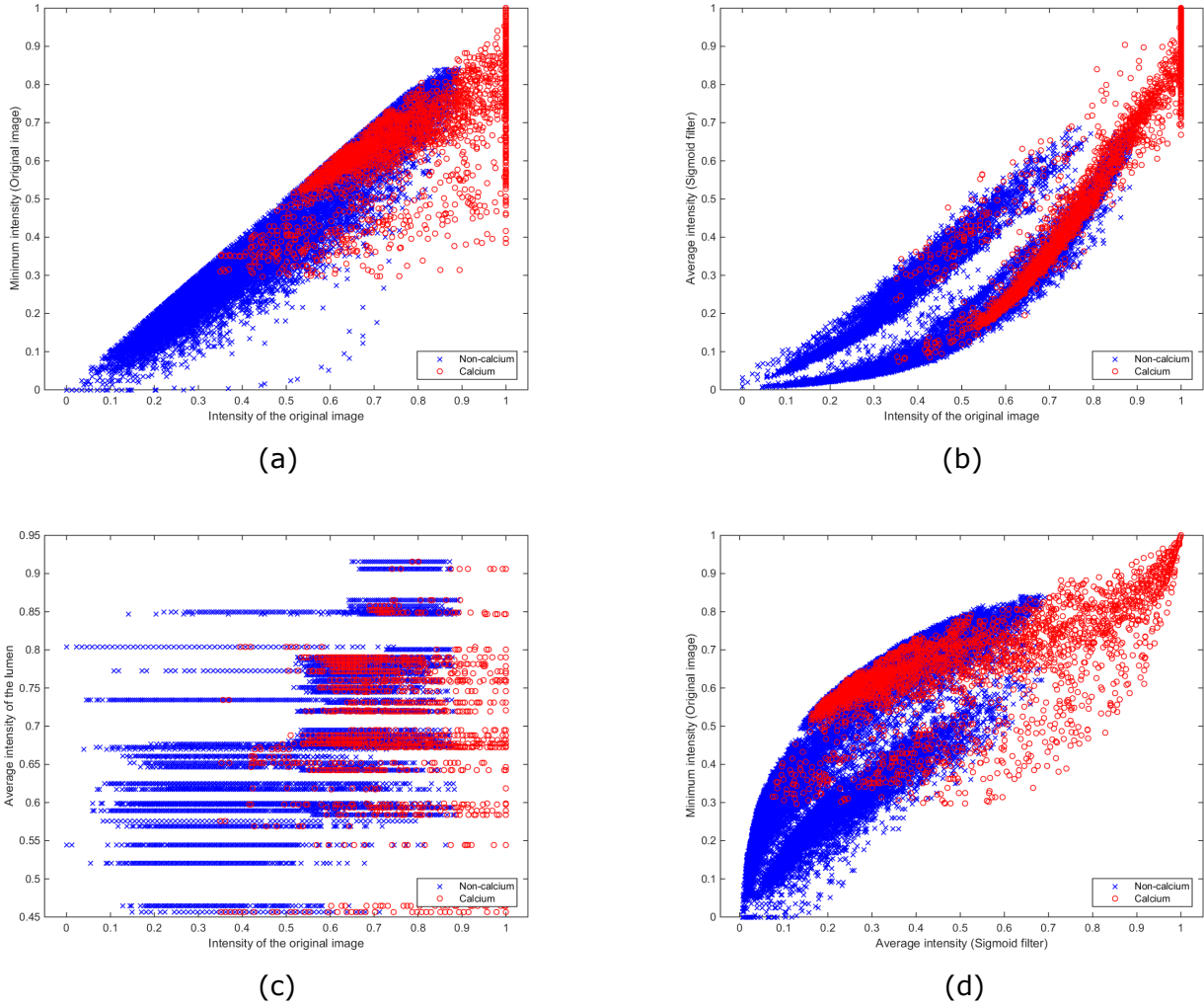
(a)

(b)

(c)

(d)

Fig. 5 Examples of the data distribution for the calcium and non-calcium pixels for some features described in Table 1: a) Distribution of the pixels' intensity of the original CTA images (feature F1) and the minimum intensity of the pixels of the original CTA images (feature F9); b) Distribution of the pixels' intensity of the original CTA images (feature F1) and the average intensities of the images filtered by the Sigmoid filter (feature F8); c) Distribution of the pixels' intensity of the original CTA images (feature F1) and the average intensities of the lumen of the original CTA images (feature F17); d) Distribution of the average intensities of the images filtered by the Sigmoid filter (feature F8) and the minimum intensities of the original CTA images (feature F9).

Table 3 Average absolute errors of the relative area for each value of parameter α tested in the Sigmoid kernel by means of the leave-one-out validation technique. (Best values in bold.)

		α=0.5	α=0.6	α=0.7	α=0.8	α=0.9	α=1
Approach 1	**Per pixel**	**31.08 $\pm$ 23.51**	32.26 $\pm$ 24.84	34.67 $\pm$ 24.07	35.03 $\pm$ 24.81	33.67 $\pm$ 24.02	32.54 $\pm$ 22.60
	Per region	**6.16 $\pm$ 7.20**	6.18 $\pm$ 11.69	6.98 $\pm$ 10.20	6.74 $\pm$ 9.53	7.63 $\pm$ 11.45	6.40 $\pm$ 8.04
Approach 2	**Per pixel**	**20.74 $\pm$ 20.00**	22.26 $\pm$ 22.82	30.00 $\pm$ 26.84	29.83 $\pm$ 25.64	38.85 $\pm$ 23.20	33.11 $\pm$ 24.89
	Per region	**4.83 $\pm$ 5.50**	6.21 $\pm$ 10.07	8.73 $\pm$ 15.04	7.62 $\pm$ 8.97	8.77 $\pm$ 11.57	9.41 $\pm$ 16.08
Approach 3	**Per pixel**	21.93 $\pm$ 23.23	22.58 $\pm$ 22.91	23.85 $\pm$ 23.00	25.03 $\pm$ 22.61	25.77 $\pm$ 22.33	27.52 $\pm$ 21.86
	Per region	5.01 $\pm$ 6.48	6.81 $\pm$ 10.57	6.97 $\pm$ 10.98	**4.56 $\pm$ 6.26**	6.02 $\pm$ 7.69	5.43 $\pm$ 6.40
Approach 4	**Per pixel**	11.79 $\pm$ 14.95	11.55 $\pm$ 14.61	**11.43 $\pm$ 14.48**	13.27 $\pm$ 15.30	15.08 $\pm$ 16.10	17.92 $\pm$ 18.19
	Per region	**3.29 $\pm$ 4.56**	4.51 $\pm$ 9.38	4.28 $\pm$ 9.25	4.62 $\pm$ 6.37	4.77 $\pm$ 6.33	4.81 $\pm$ 6.02

*The values are expressed in percentage.

Table 4 Average absolute errors of the relative volume for each value of parameter α tested in the Sigmoid kernel by means of the leave-one-out validation technique. (Best values in bold.)

|		α=0.5	α=0.6	α=0.7	α=0.8	α=0.9	α=1
Approach 1	**Per pixel**	**24.31 ± 13.29**	25.16 ± 16.94	27.34 ± 15.93	27.89 ± 18.35	29.92 ± 18.40	27.82 ± 16.02
	Per region	**3.69 ± 2.92**	4.10 ± 3.40	5.27 ± 4.07	4.15 ± 3.37	4.34 ± 3.33	4.14 ± 3.63
Approach 2	**Per pixel**	**20.08 ± 15.91**	20.52 ± 15.57	29.01 ± 21.08	29.77 ± 21.56	37.15 ± 21.44	33.60 ± 21.83
	Per region	**3.35 ± 2.05**	4.34 ± 3.57	9.30 ± 14.39	4.57 ± 3.86	4.97 ± 3.65	6.01 ± 6.97
Approach 3	**Per pixel**	**16.03 ± 13.64**	16.91 ± 13.42	18.12 ± 13.16	19.40 ± 12.96	20.12 ± 12.90	21.83 ± 12.49
	Per region	3.42 ± 2.9	4.60 ± 3.92	4.97 ± 4.47	3.18 ± 2.83	3.72 ± 2.85	**3.17 ± 3.10**
Approach 4	**Per pixel**	**10.13 ± 9.51**	10.30 ± 9.41	10.33 ± 9.42	12.33 ± 9.46	13.78 ± 9.51	16.20 ± 12.16
	Per region	**2.74 ± 3.11**	3.43 ± 3.80	3.69 ± 4.29	3.36 ± 3.13	3.34 ± 2.81	2.90 ± 3.05

*The values are expressed in percentage.
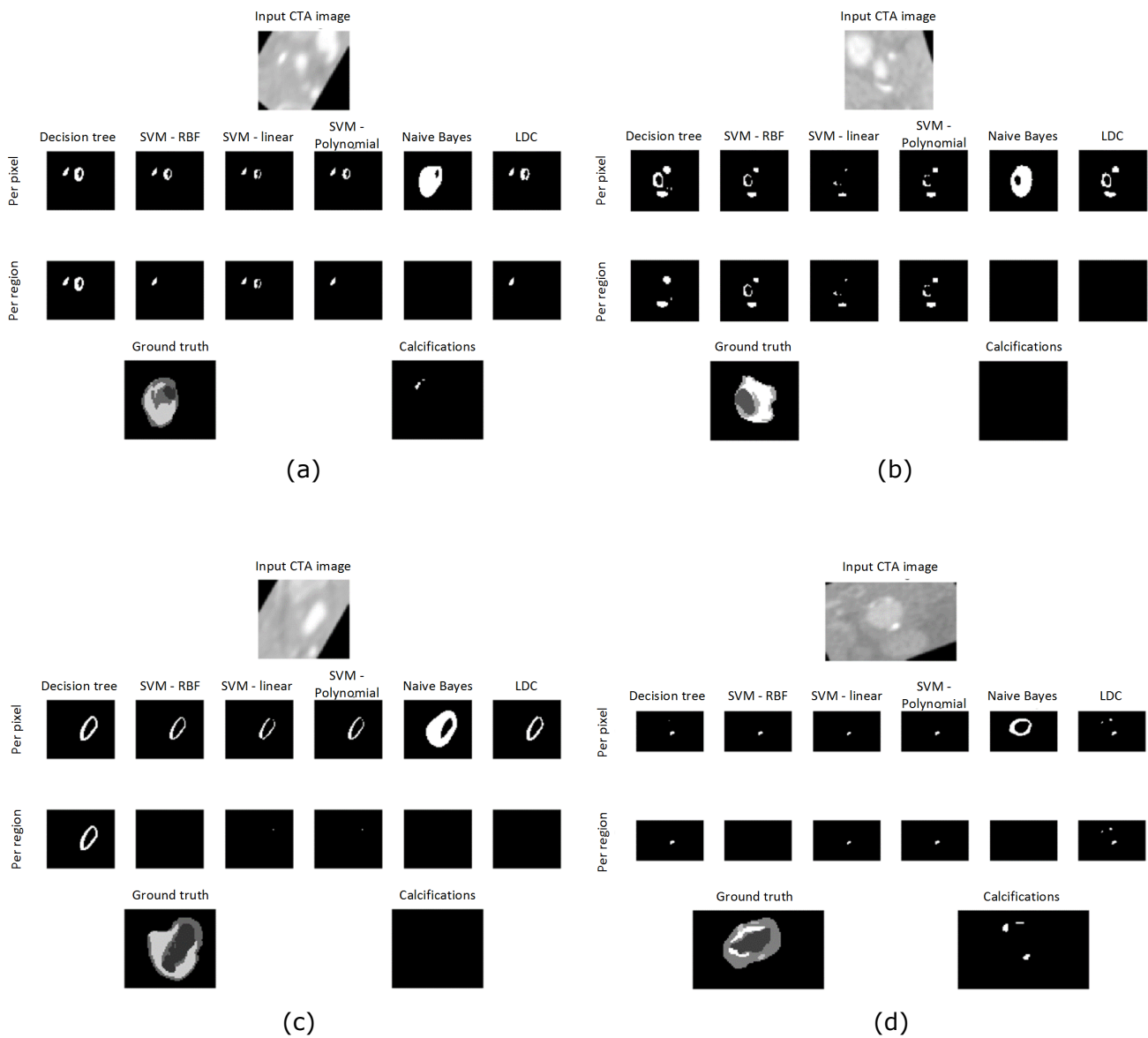
Fig. 6 Examples of classification results obtained by each classifier under comparison in the pixel and region classification steps.
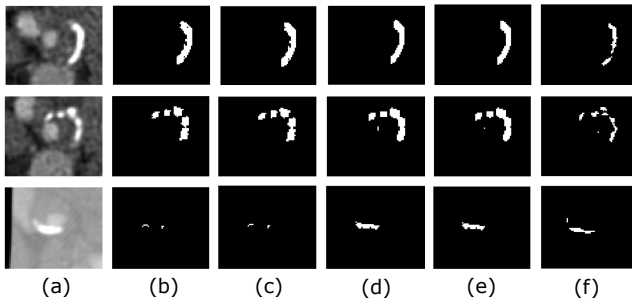
Fig. 7 Examples of classification results obtained by each approach indicated in Table 2 in the pixel classification step: a) Original CTA images; b) Classification results obtained by Approach 1; c) Classification results obtained by Approach 2; d) Classification results obtained by Approach 3; e) Classification results obtained by Approach 4; f) Corresponding ground truths of the calcified components.

Regarding the comparison of the computational results with the ground truths, the LDC obtained better results than the remainder classifiers as can be perceived in the examples of Fig. 6.

The average error and Spearman correlation between the computational results and the corresponding ground truths as to the k-fold cross-validation technique for each approach and classifier are shown in Fig. 8, Fig. 9 and Fig. 10. In addition, the average bias of the relative areas and volumes are indicated in Tables 5 and 6.

Regarding the relative percentage of the classified calcium regions with respect to the total area of the carotid wall, the average absolute errors decreased for almost of the classifiers after the employment of the classification per region, mainly for Approaches 3 and 4, as can be observed in Fig. 8. Additionally, the average absolute errors obtained from the classification performed by the LDC were lower than those obtained from the other classifiers, mainly for Approaches 1 and 2. In terms of the results obtained by Approach 3, the average absolute error of the relative area obtained from the LDC results decreased from 12.32% to 4.07% after the employment of the classification per region on the results obtained from the classification per pixel step. Similar improvements were also obtained for Approach 4, where the average absolute error of the relative area decreased from 12.22% to 4.06% after the classification per region. As shown in Fig. 8 (plots of the second row), a significant reduction of the absolute errors of the relative volume was also achieved after classifying each region resultant from the pixel classification step, leading the absolute errors of Approaches 3 and 4 to decrease from 10.94% and 10.84% to 2.03% and 2.14%, respectively. In terms of the absolute area and volume of the calcifications, similar improvements were also found with the classification per region as shown in Fig. 9. Ad-

ditionally, the average absolute errors of the absolute area and volume of the calcium regions were also lower for the LDC in all approaches (Fig. 9b).

Improvements on the Spearman correlation coefficients were also achieved from the classification of the candidate regions resultant from the pixel classification step. As shown in Fig. 10b, a significant increase of the correlation coefficients between the area and volume of the classification results and those calculated from the corresponding ground truths was obtained after the region classification step for most of the classifiers used, mainly for Approaches 3 and 4. Moreover, the results obtained by the LDC were also better than the ones obtained from the other classifiers. Regarding the results obtained by the LDC, the Spearman correlation coefficient between the relative area calculated from the classified calcium regions and the ones obtained from the corresponding ground truths was 0.14 for Approach 3 of the pixel classification step. The correlation coefficient for the same approach increased to 0.55 after the application of the classification per region. Regarding the relative volume obtained by using Approach 3, the Spearman correlation coefficient increased from 0 (zero) to 0.85 after the application of the classification per region. With respect to Approach 4, the Spearman correlation coefficients between the relative area and volume of the classification results and those calculated from the ground truths were 0.14 and -0.04, respectively, for the pixel classification step. The correlation coefficients increased to 0.56 and 0.83, respectively, after the classification of the candidate regions obtained from the pixel classification step. Similar improvements in the Spearman correlation coefficients of the absolute area and volume of the calcium regions were also achieved for all approaches and classifiers as shown in Fig. 10 (plots of the second row).

Similar results were also achieved by means of the leave-one-out cross-validation technique as illustrated in Fig. 11, Fig. 12 and Fig. 13. In addiction, the average bias of the relative area and volume obtained from the leave-one-out cross-validation technique are indicated in Tables 7 and 8.

Similar to the k-fold cross-validation technique, the results obtained by the LDC were also better than the ones generated by the remainder classifiers. In terms of the relative area, the average absolute errors obtained by the LDC in the pixel classification step were 2.49%, 2.46%, 12.48% and 12.27% for Approaches 1, 2, 3 and 4, respectively. After the application of the classification per region, the average absolute errors were 2.56%, 2.38%, 4.10% and 4.07% for the same Approaches, respectively. It is possible to notice a significant reduction of the average absolute errors of the relative area
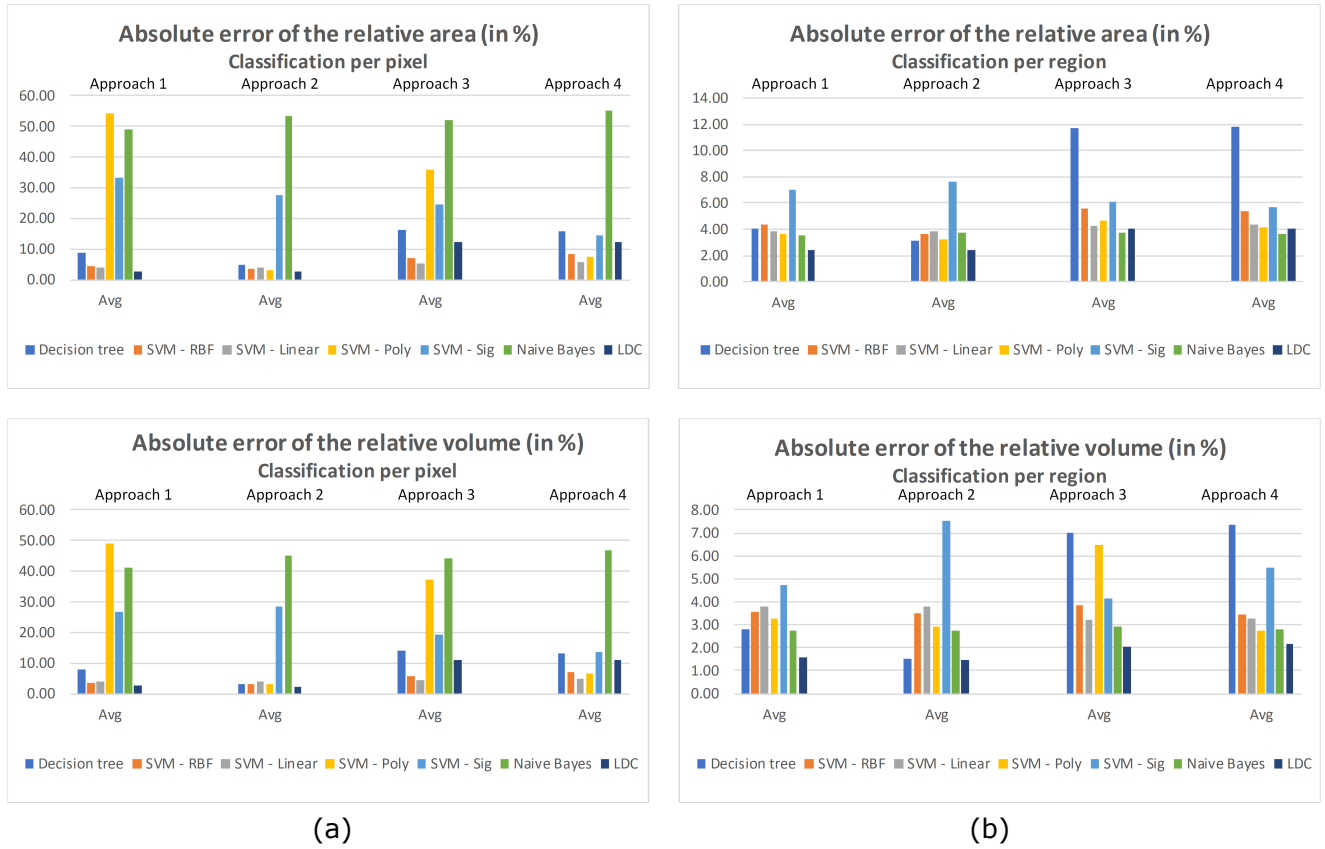
Fig. 8 Average absolute errors of the relative areas and volumes obtained from the k-fold cross-validation technique for the classifications (a) per pixel and (b) per region. (The first row represents the average absolute errors for the relative area, and the second row the average absolute errors for the relative volume of the classified calcium regions.)

Table 5 Average bias obtained as to the relative areas of the classified calcium regions in each approach and classifier of the k-fold cross validation technique. (Best values in bold.)

		Decision tree	SVM - RBF	SVM - Linear	SVM - Poly	SVM - Sig	Naive Bayes	LDC
A1	Per pixel	7.19 ± 12.07	-2.08 ± 5.98	-3.82 ± 5.23	54.19 ± 28.29	33.15 ± 23.27	48.75 ± 42.00	**-0.06 ± 4.74**
	Per region	1.23 ± 6.43	-2.59 ± 5.81	-3.82 ± 5.23	-2.80 ± 5.06	**0.91 ± 10.84**	-2.47 ± 5.56	-1.20 ± 3.96
A2	Per pixel	3.55 ± 7.27	-3.29 ± 5.12	-3.82 ± 5.23	-3.15 ± 4.85	27.16 ± 25.13	53.36 ± 41.70	**0.29 ± 4.60**
	Per region	**0.46 ± 4.90**	-3.42 ± 5.17	-3.82 ± 5.23	-3.16 ± 4.86	2.20 ± 15.19	-2.49 ± 5.73	-0.70 ± 4.05
A3	Per pixel	16.18 ± 16.16	2.94 ± 9.87	**1.77 ± 7.85**	34.79 ± 26.47	24.19 ± 24.04	52.07 ± 42.17	11.67 ± 11.94
	Per region	8.95 ± 17.19	**0.18 ± 7.54**	-0.31 ± 6.02	-2.46 ± 9.23	0.30 ± 10.69	-2.63 ± 5.64	1.89 ± 5.67
A4	Per pixel	15.49 ± 15.87	6.92 ± 10.97	**3.09 ± 8.16**	6.04 ± 9.97	13.50 ± 19.03	54.99 ± 41.49	11.60 ± 12.18
	Per region	9.55 ± 17.11	0.76 ± 8.05	**0.07 ± 6.17**	1.59 ± 5.99	0.38 ± 14.16	-2.58 ± 5.62	1.92 ± 5.55

*The values are expressed in percentage.
*A1, A2, A3 and A4 represent Approaches 1 to 4.

obtained from Approaches 3 and 4 after the region classification step. For the absolute error of the relative volume, the average values obtained from Approaches 1, 2, 3 and 4 were 2.13%, 1.89%, 11.00% and 10.80%, respectively, for the pixel classification step. After the classification per region, the average absolute errors decreased to 1.91%, 1.56%, 2.15% and 2.19%, respectively, for the same Approaches. A significant reduction of the absolute error was also achieved for Approaches 3 and 4. As shown in Fig. 12, regarding the average absolute

errors for the absolute area and volume of the calcium regions, a similar decreasing behaviour was found after the classification per region step.

The Spearman correlation coefficients obtained from the leave-one-out cross-validation technique were similar to the ones calculated for the k-fold cross-validation technique, showing the stability of the suggested classification model. The LDC classifier also showed better results than the other classifiers. The Spearman correlation coefficients between the relative areas of the
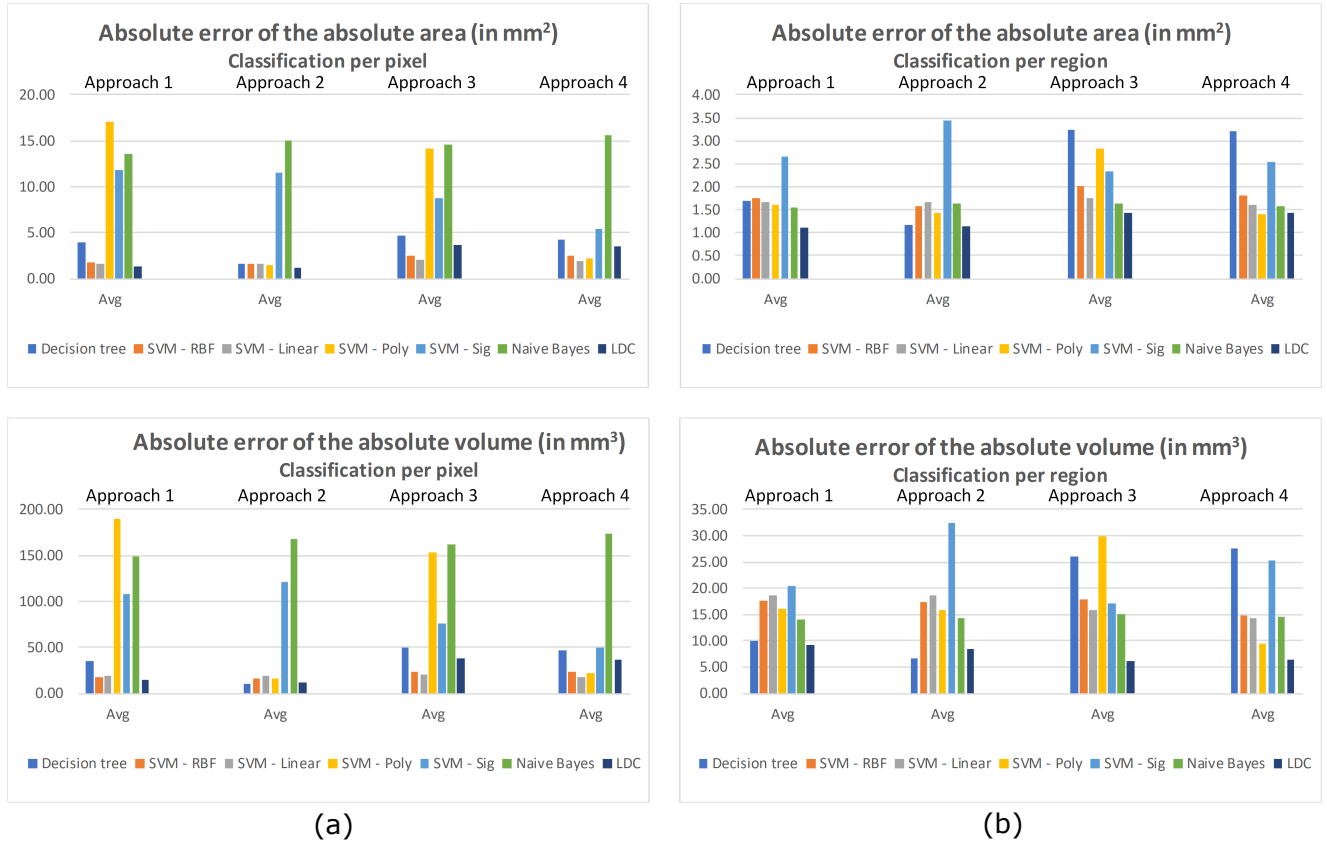
(a)   (b)

Fig. 9 Average absolute errors of the absolute areas and volumes obtained by the k-fold cross-validation technique for the classifications (a) per pixel and (b) per region. (The first row represents the average absolute errors for the absolute area, and the second row the average absolute errors for the absolute volume of the classified calcium regions.)

Table 6 Average bias obtained as to the relative volumes of the classified calcium regions in each approach and classifier of the k-fold cross validation technique. (Best values in bold.))

		Decision tree	SVM - RBF	SVM - Linear	SVM - Poly	SVM - Sig	Naive Bayes	LDC
A1	Per voxel	7.81 ± 10.79	-2.08 ± 4.14	-3.79 ± 3.38	49.13 ± 23.88	24.84 ± 16.90	41.23 ± 35.88	**0.42 ± 3.95**
	Per region	1.17 ± 3.55	-2.70 ± 3.92	-3.79 ± 3.39	-2.90 ± 3.65	**0.73 ± 6.53**	-2.47 ± 2.87	-1.22 ± 2.34
A2	Per voxel	3.10 ± 2.12	-3.08 ± 3.39	-3.79 ± 3.39	-2.89 ± 3.14	26.53 ± 25.31	45.26 ± 37.04	**0.65 ± 3.25**
	Per region	**0.35 ± 1.72**	-3.43 ± 3.31	-3.79 ± 3.39	-2.90 ± 3.15	3.88 ± 15.93	-2.47 ± 3.02	-0.82 ± 1.87
A3	Per voxel	14.06 ± 10.92	2.32 ± 7.54	**1.81 ± 6.07**	37.39 ± 19.33	17.34 ± 16.76	44.32 ± 36.55	10.71 ± 8.96
	Per region	6.46 ± 10.19	-0.54 ± 5.10	-0.72 ± 4.31	0.22 ± 9.82	**0.19 ± 5.58**	-2.57 ± 3.29	1.74 ± 1.94
A4	Per voxel	13.29 ± 10.60	6.08 ± 7.89	**2.96 ± 5.85**	5.40 ± 7.10	11.39 ± 15.36	46.76 ± 36.77	10.63 ± 9.10
	Per region	6.92 ± 10.73	**-0.22 ± 4.66**	-0.62 ± 4.05	1.55 ± 3.30	1.08 ± 11.15	-2.51 ± 3.06	1.81 ± 1.77

*The values are expressed in percentage.
*A1, A2, A3 and A4 represent Approaches 1 to 4.

classification results and the ones calculated from the corresponding ground truths were 0.70, 0.73, 0.12 and 0.14, for Approaches 1, 2, 3 and 4, respectively, in the pixel classification step. The Spearman correlation coefficients obtained from the region classification step were 0.65, 0.70, 0.54 and 0.56 for the same Approaches, leading to a significant increase of the correlation coefficients for Approaches 3 and 4. Similar improvements were also obtained as to the relative volumes. The correlation coefficients between the relative volumes of the

classification results and the ones calculated from the corresponding ground truths were 0.65, 0.70, -0.04 and -0.07 for Approaches 1, 2, 3 and 4, respectively, in the pixel classification step. On the other hand, the correlation coefficients obtained after the classification per region step were 0.69, 0.79, 0.83 and 0.81 for the same Approaches. Likewise the relative area, significant improvements in the absolute areas of the classified calcium regions were also achieved for Approaches 3 and 4
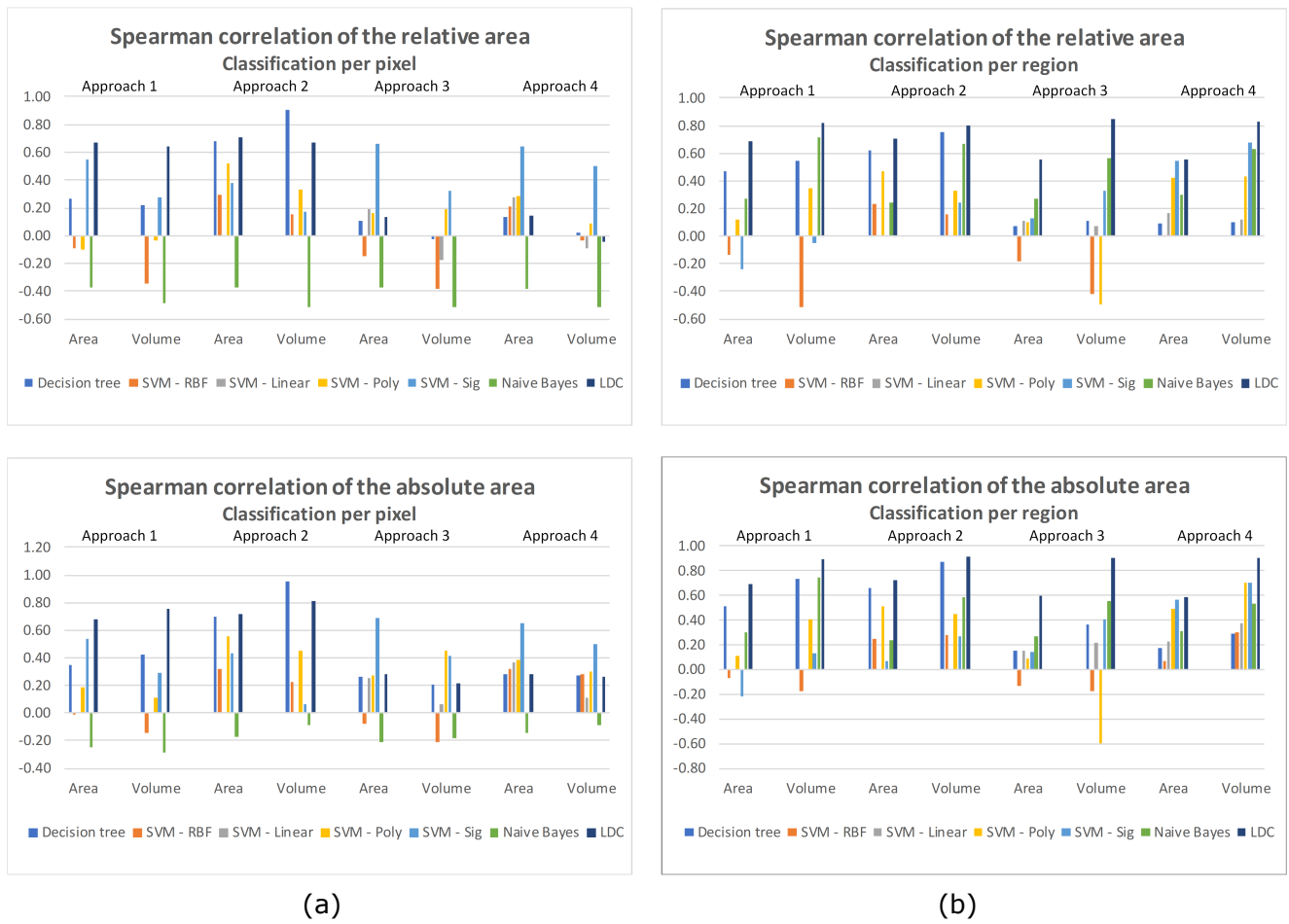
(a)   (b)

Fig. 10 Spearman correlation coefficients of the relative and absolute areas and volumes obtained by the k-fold cross-validation technique for the classifications (a) per pixel and (b) per region. (The first row represents the Spearman correlation coefficients for the relative area and volume of the classified calcium regions, whereas the second row the Spearman correlation coefficient of the absolute area and volume of the same regions.)

Table 7 Average bias obtained as to the relative areas of the classified calcium regions in each approach and classifier of the leave-one-out cross validation technique. (Best values in bold.)

		Decision tree	SVM - RBF	SVM - Linear	SVM - Poly	SVM - Sig	Naive Bayes	LDC
$A1$	Per pixel	6.06 ± 10.00	-2.05 ± 5.95	-3.82 ± 5.23	53.32 ± 31.08	30.96 ± 23.67	48.94 ± 41.87	$\mathbf{-0.44 \pm 4.08}$
	Per region	1.31 ± 6.76	-2.63 ± 5.70	-3.82 ± 5.23	-3.04 ± 4.96	$\mathbf{-0.23 \pm 9.48}$	-2.41 ± 5.64	-0.99 ± 4.18
$A2$	Per pixel	3.32 ± 7.38	-3.25 ± 5.09	-3.82 ± 5.23	-3.08 ± 4.95	20.45 ± 20.30	53.41 ± 41.62	$\mathbf{-0.08 \pm 4.03}$
	Per region	$\mathbf{0.33 \pm 4.98}$	-3.38 ± 5.14	-3.82 ± 5.23	-3.09 ± 4.96	1.08 ± 7.25	-2.43 ± 5.79	-0.53 ± 3.88
$A3$	Per pixel	16.56 ± 16.62	3.30 ± 10.27	$\mathbf{2.26 \pm 8.35}$	32.92 ± 26.41	21.48 ± 23.65	52.31 ± 42.02	11.84 ± 12.44
	Per region	9.73 ± 17.42	0.39 ± 8.04	$\mathbf{0.09 \pm 6.76}$	-2.13 ± 9.42	-1.12 ± 8.12	-2.56 ± 5.76	2.03 ± 5.63
$A4$	Per pixel	15.83 ± 16.56	7.24 ± 11.24	$\mathbf{3.20 \pm 8.32}$	6.20 ± 10.08	10.82 ± 15.66	55.16 ± 41.36	11.67 ± 12.50
	Per region	10.38 ± 17.40	1.01 ± 8.50	$\mathbf{0.24 \pm 6.39}$	1.74 ± 6.11	-2.68 ± 4.95	-2.51 ± 5.72	2.03 ± 5.49

*The values are expressed in percentage.
*A1, A2, A3 and A4 represent Approaches 1 to 4.

after the region classification step as can be perceived in Fig. 13 (plots of the second row).

5 Discussion

The characterization of atherosclerotic plaque components plays an important role towards the evaluation of the disease progression. The composition of atherosclerotic plaques has been seen as an important factor for evaluating the risks of plaque rupture, as well as risks
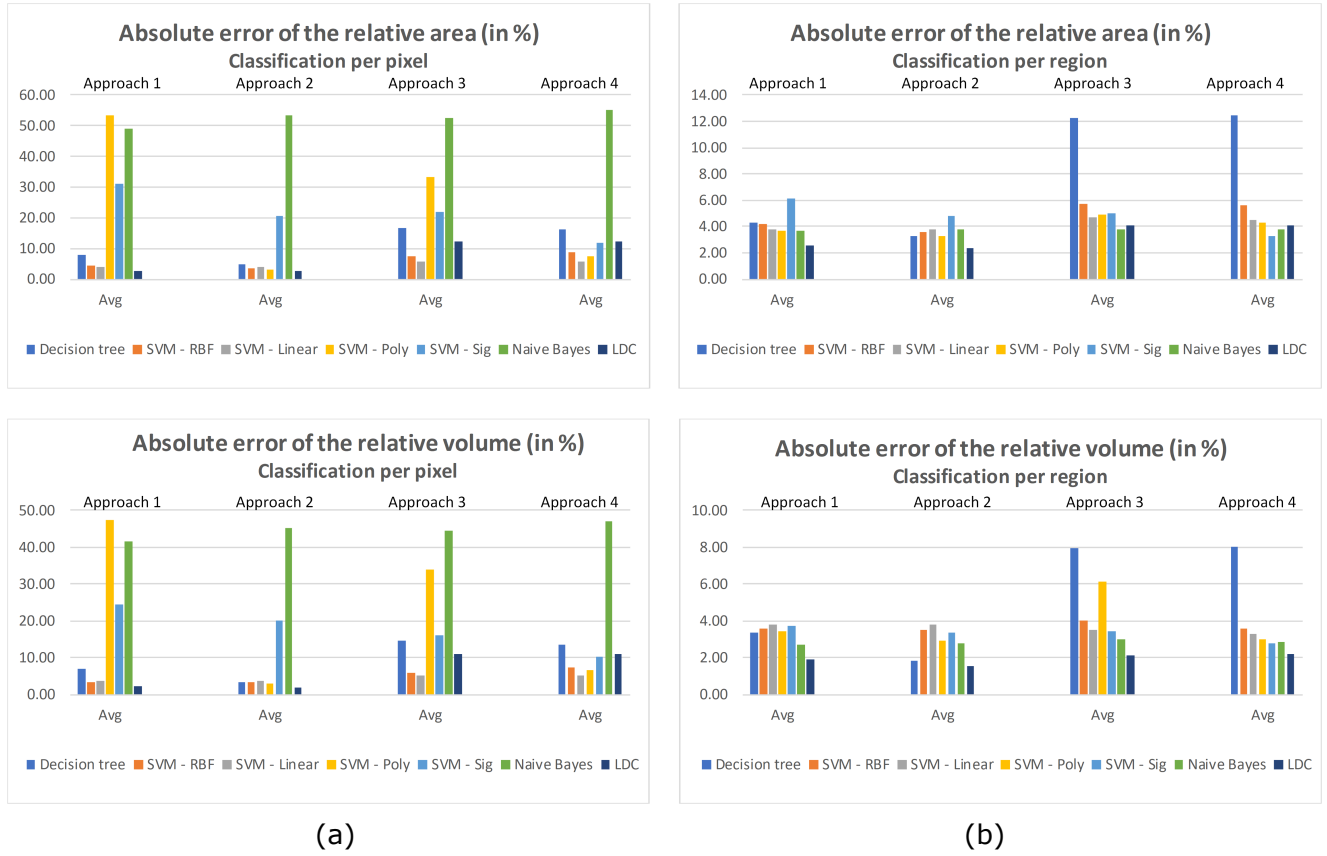
Fig. 11 Average absolute errors of the relative areas and volumes obtained from the leave-one-out cross-validation technique for the classifications (a) per pixel and (b) per region. (The first row represents the average absolute errors for the relative area, and the second row the average absolute errors for the relative volume of the classified calcium regions.)

Table 8 Average bias obtained as to the relative volumes of the classified calcium regions in each approach and classifier of the leave-one-out cross validation technique. (Best values in bold.)

		Decision tree	SVM - RBF	SVM - Linear	SVM - Poly	SVM - Sig	Nave Bayes	LDC
A1	Per voxel	6.53 ± 9.37	-2.02 ± 4.19	-3.79 ± 3.39	47.41 ± 27.00	22.35 ± 16.61	41.47 ± 35.81	**-0.03 ± 3.32**
	Per region	1.18 ± 4.30	-2.69 ± 3.95	-3.79 ± 3.39	-3.19 ± 3.48	**-0.66 ± 4.77**	-2.39 ± 2.95	-1.02 ± 2.66
A2	Per voxel	3.16 ± 2.39	-3.04 ± 3.39	-3.79 ± 3.39	2.81 ± 3.29	18.12 ± 18.28	45.33 ± 36.95	**0.19 ± 2.70**
	Per region	**0.27 ± 2.19**	-3.40 ± 3.32	-3.79 ± 3.39	-2.82 ± 3.29	-1.14 ± 3.87	-2.40 ± 3.08	-0.61 ± 2.11
A3	Per voxel	14.43 ± 11.36	2.69 ± 7.71	**2.19 ± 6.15**	34.00 ± 20.27	14.07 ± 15.81	44.60 ± 36.45	10.80 ± 9.48
	Per region	7.25 ± 9.96	**-0.41 ± 5.23**	-0.49 ± 4.48	0.54 ± 9.69	-1.52 ± 4.37	-2.51 ± 3.38	1.87 ± 1.91
A4	Per voxel	13.57 ± 11.34	6.37 ± 8.04	**3.00 ± 5.88**	5.53 ± 7.14	8.02 ± 11.49	46.95 ± 36.64	10.63 ± 9.46
	Per region	7.71 ± 10.43	**0.01 ± 4.91**	-0.48 ± 4.06	1.67 ± 3.38	-2.63 ± 3.21	-2.47 ± 3.14	1.91 ± 1.77

*The values are expressed in percentage.
*A1, A2, A3 and A4 represent Approaches 1 to 4.

for embolization and neurological events. Previous studies suggested that the presence of calcified components in atherosclerotic plaques represents an advanced stage of the disease.

The classification of calcium regions in CTA images of the carotid artery was successfully tackled in this study. Firstly, the proposed classification model performs the classification of each pixel inside the carotid artery in order to obtain the candidate regions representing the calcifications of the atherosclerosis. Then,

the regions previously identified are submitted to the region classification step in order to identify the ones that represent true calcifications based on geometrical and intensity features extracted from each region. Misalignments between the histological images and the corresponding *in vivo* CTA images resulting from the used registration procedure might affect the classification results due to the selection of parts of the lumen and other regions having intensity similar to the one of calcified components. Hence, pixels belonging to regions
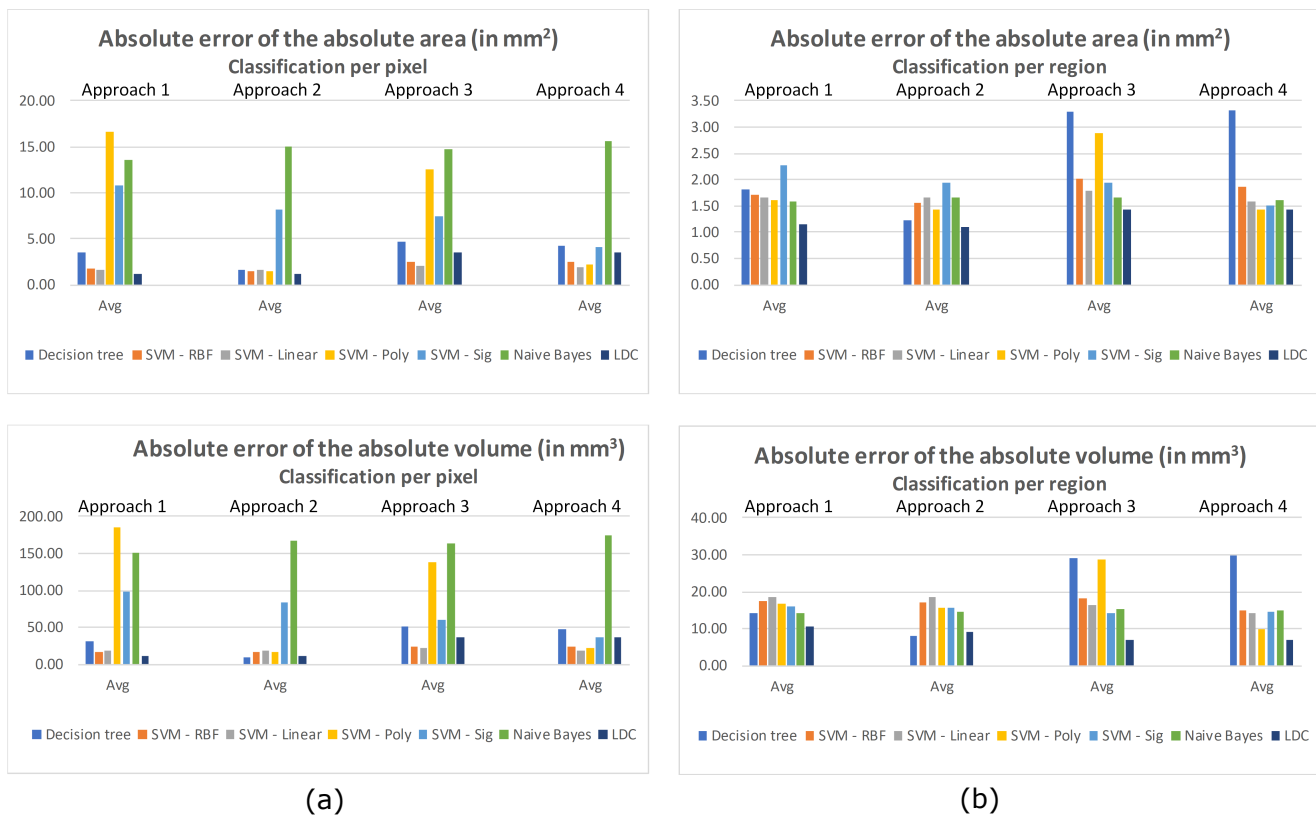
Fig. 12 Average absolute errors of the absolute areas and volumes obtained from the leave-one-out cross-validation technique for the classifications (a) per pixel and (b) per region. (The first row represents the average absolute errors for the absolute area, and the second row the average absolute errors for the absolute volume of the classified calcium regions.)

corresponding to other structures of the carotid artery may affect the accuracy of the results. Hence, a second step to identify and maintain only the regions correctly representing calcifications was included in the proposed classification model, leading to results significantly better than the ones exclusively obtained by the classification per pixel step.

Besides the assessment of regions that do not correspond to true calcifications of the atherosclerotic lesions, the removal of outliers represents an important approach to increase the performance of the classifiers and provide a better separation of the pixels belonging to each class, i.e. calcium and non-calcium regions. The presence of outliers decreases the ability of the classifiers in separating the classes and determining the correct class of the input pixel, leading to errors in the classification results and, consequently, the decreasing of the accuracy of the classification model. Hence, an approach to remove outliers of the pixels selected for training and testing the classifiers in each cross-validation iteration was proposed in this study. The boxplots of the pixels belonging to each class were generated and evaluated to determine the presence of outliers in the experimental dataset based on all intensity features ex-

tracted from each pixel to be classified. Hence, common pixels representing outliers in all features were removed before the training and testing processes. The outliers removal approach proposed here proved to be effective in determining the correct class of each pixel and improving the shapes of the calcified regions to make them as close as possible to the ones of the corresponding ground truths.

The areas and volumes of the calcium regions in the CTA images used in this study are often larger than the ones of the corresponding calcifications in the ground truth. The ground truth of the calcified components was obtained using micro CT images of the excised plaques as described in van Engelen et al [22]. According to van Engelen et al [22], the blooming artefacts often present in CTA images may cause the overestimation of the calcium components. The bias of the areas and volumes resulting from Approaches 3 and 4 showed a positive value when compared to Approaches 1 and 2. The average bias of the relative area obtained from the classification per region in Approaches 1, 2, 3 and 4 were -1.20 ± 3.96%, -0.70 ± 4.05%, 1.89 ± 5.67% and 1.92 ± 5.55%, respectively, for the classification results of the LDC in the k-fold cross-validation technique. In terms
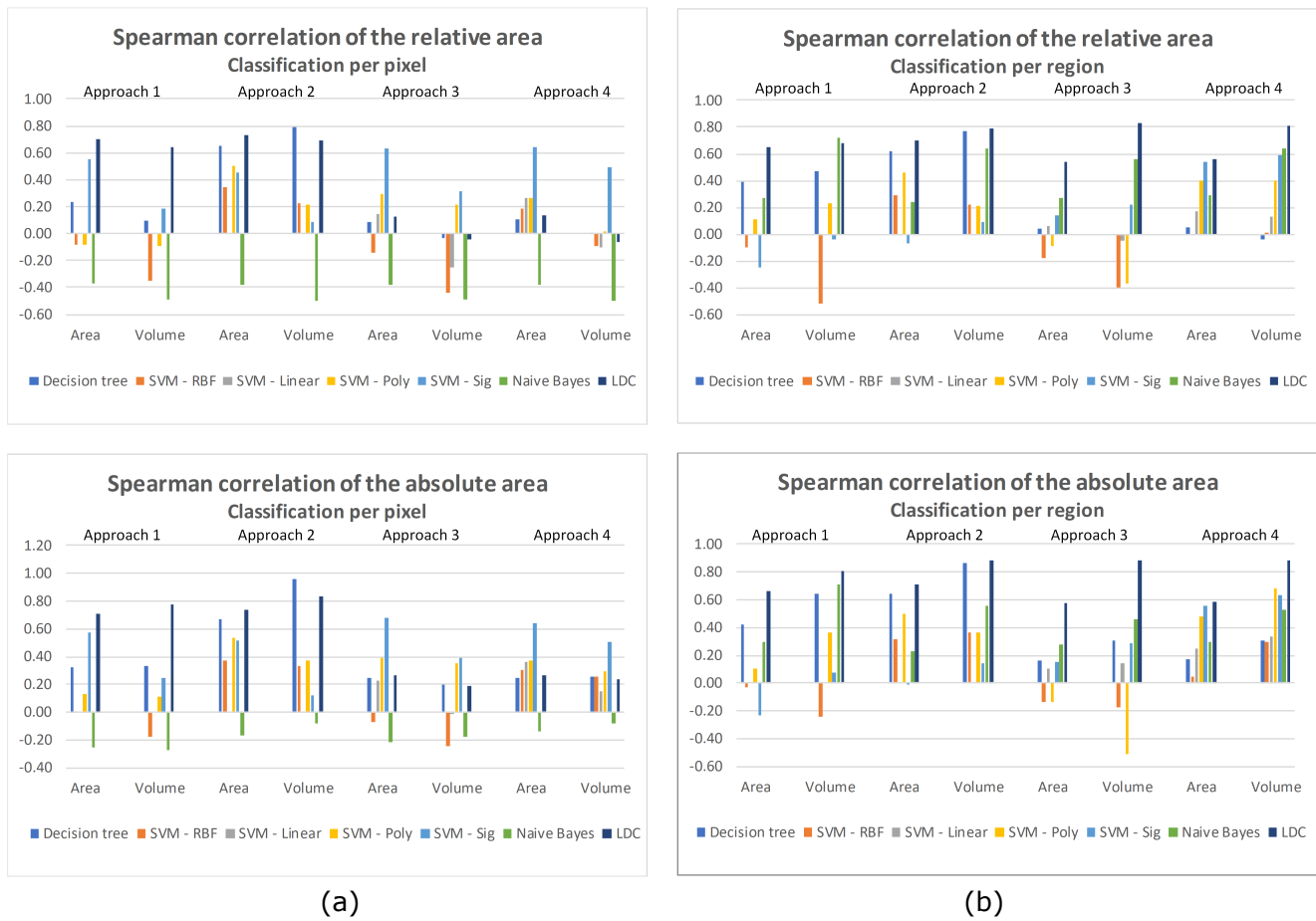
Fig. 13 Spearman correlation coefficients of the relative and absolute areas and volumes obtained from the leave-one-out cross-validation technique for the classifications (a) per pixel and (b) per region. (The first row represents the Spearman correlation coefficients for the relative area and volume of the classified calcium regions, whereas the second row the Spearman correlation coefficients of the absolute area and volume.)

of the relative volume, the average bias obtained from the LDC in the k-fold cross-validation technique in the region classification step were -1.22 ± 2.34%, -0.82 ± 1.87%, 1.74 ± 1.94% and 1.81 ± 1.77% for Approaches 1, 2, 3 and 4, respectively. Regarding the leave-one-out cross validation technique, the average bias of the relative areas obtained by the LDC for Approaches 1, 2, 3 and 4 were -0.99 ± 4.18%, -0.53 ± 3.88%, 2.03 ± 5.63% and 2.03 ± 5.49%, respectively, for the region classification step. Similarly, the average bias of the relative volumes obtained from the same Approaches were -1.02 ± 2.66%, -0.61 ± 2.11%, 1.87 ± 1.91% and 1.91 ± 1.77%, respectively, concerning the results generated by the LDC in the region classification step. Similarly to the k-fold cross validation technique, the average bias of Approaches 3 and 4 overestimated the ground truth relatively to Approaches 1 and 2. The outliers removal is responsible for the overestimation of the calcium regions classified by Approaches 3 and 4. After removing the outliers from the training and testing sets in each

iteration of the cross-validation technique, the good separation of the examples belonging to each class induces the classifier to assign the correct class to the pixels under analysis. Hence, the calcified components tend to overestimate the corresponding ground truths since all pixels of the calcium regions are correctly classified.

The Spearman correlation of the relative and absolute areas of the classified calcium regions decreased in Approaches 3 and 4 in comparison to the results obtained from Approaches 1 and 2 in the region classification step. As previously discussed, the overestimation of the calcifications after the outliers removal is responsible for affecting the size of the classified components, leading the area of the calcified regions to increase significantly in some cases. Since the areas of the classified calcium regions increase in comparison to the correspondent ground truths, the Spearman correlation coefficients also decreased. As to the distance features, no significant differences were found when the distances of each pixel to the lumen and carotid wall contours were

removed from the classification process. Hence, the classification model could be designed with only 18 intensity features in the pixel classification step.

The classification results obtained by the LDC were better than those obtained by the other classifiers that are often used in many classification problems. The results obtained by van Engelen et al [22] also indicated the better performance of the LDC in comparison to a SVM with a RBF kernel. Other studies have also indicated the superior performance of the LDC in classifying atherosclerotic plaque components in images [26] [27] [28]. Although decision trees have been widely used in several studies dealing with different classification problems, the large number of decisions makes the tree more complex and prone to overfitting. Additionally, small changes in the input values may still cause significant changes in the model of the decision tree. As shown in Fig. 10 and Fig. 13, the Spearman correlation coefficients obtained from the decision tree were better for Approaches 1 and 2 of the pixel and region classification steps, achieving results equivalent to the ones of the LDC. However, the correlation coefficients obtained from the decision tree decreased significantly in Approaches 3 and 4. Improvements in the results obtained by the Naive Bayes classifier were also achieved after applying the region classification step. However, the results were still of lower quality than those of the LDC. Although the Naive Bayes classifier is simple and independent of irrelevant features, an important condition for better results is that the features might be conditionally independent for the given class. The results of the LDC were stable in all approaches, which indicates that this classifier is more effective to classify atherosclerotic plaque components, particularly the calcified regions as proposed in this study.

The results obtained by the proposed classification model are in accordance with the majority of the studies proposed in the literature as to calcium identification in CTA images of atherosclerotic plaques. De de Graaf et al [24] reported a correlation of 0.73 between the automatic and manual delineations of dense calcium regions in CTA images of coronary arteries. In the study carried out by Vukadinovic et al [29], the Pearson correlation coefficient between the automatic and manual delineations of the calcified regions was equal to 0.94 considering the volume of the identified components. The identification of atherosclerotic plaque components proposed by Wintermark et al [23] shown a perfect agreement between the automatic and manual delineations of the calcified components. Although the automatic identification of the calcium regions was well-correlated with the ground truths in the mentioned studies, the proposed methods were based on HU values of the cal-

cified components. However, the HU values are defined based on the used CTA image dataset, which can lead to different results when the same HU values are used in images belonging to distinct datasets. van Engelen et al [22] reported a Spearman correlation coefficient of 0.91 between the automatic and manual delineations of calcium regions with respect to the relative volume of the identified components. The average bias of the relative volume between the automatic and manual delineations reported in their study was equal to -0.2 $\pm$ 1.4. On the other hand, the classification model proposed here obtained a Spearman correlation coefficient of 0.83 and 0.81 for the relative volumes in Approaches 3 and 4, respectively, relatively to the manual delineations of the calcium regions. The average bias of the relative volumes obtained from the same approaches were 1.87 $\pm$ 1.90 and 1.91 $\pm$ 1.77, respectively. The classification model proposed by van Engelen et al [22] is based on 24 features extracted from CTA and MR images of carotid arteries. However, the classification model proposed here is based on the identification of calcium regions solely in CTA images. Hence, on a clinical point-of-view, the proposed process is simpler and less expensive since only one image modality is needed in order to obtain similar results to the ones reported in van Engelen et al [22].

The main limitation of the proposed classification model is the number of features used to perform the identification of calcified and non-calcified regions. Dimensionality reduction is often used to decrease the computational cost of classifiers and represents an important step towards the selection of the relevant features to be used in classification models. The outliers removal is performed to avoid the incorrect assignment of the correct class to each pixel due to the misalignments between the ground truths manually delineated in histological images and the input CTA images. However, the number of pixels of the calcified components is significantly lower than the ones belonging to the non-calcified regions. The relative area of the calcified components is low relatively to the total area of the carotid wall and the outliers removal might decrease the number of examples of the calcified components. Hence, manual delineations made directly in the CTA images represent an effective approach to avoid the removal of pixels with important information for the classification process. Although the number of images used in this study could be higher to make the analysis more comprehensive and self-contained, the proposed classification model was designed to eliminate the requirement of using MR Images as proposed by the previous study that used the same experimental dataset van Engelen et al [22]. Therefore, the same images were used

to make the comparison more trustworthy and closer to the referred study. In addition, the proposed classification model is scalable as the number of images increase, and adaptable to other datasets similar to the one used here.

Despite the above-mentioned limitation, the classification in two steps of calcium regions in CTA images of the carotid artery seems to be effective in eliminating regions not belonging to the true calcifications of the atherosclerotic lesions.

6 Conclusions

The characterization of atherosclerosis is an intensive focus of research and represents an important step in evaluating the progression of the disease. The classification of calcified components in CTA images of carotid arteries was proposed. In the proposed classification model, the original CTA images are submitted to an initial step that processes the classification of each pixel inside the carotid artery wall. Then, the candidate regions are submitted to the next step that performs the classification per region in order to identify the ones corresponding to true calcified components. Additionally, the proposed outliers removal approach proved to be effective in improving the separation of the pixels belonging to each class, i.e. calcified and non-calcified regions, leading the shapes of the classified calcium regions close to the ones of the corresponding ground truths.

In this study, the region classification step was proposed to effectively handle the incorrect classification of regions resultant from misalignments of the ground truths with the corresponding *in vivo* CTA images. The method proved to be effective in eliminating regions that do not correspond to true calcified components, leading to improvements of the classification results and, consequently, to a more efficient, reliable and accurate classification model.

Conflict of interest statement

The authors report no conflict of interest.

References

1. Mendis S, Puska P, Norrving B (2011) Global atlas on cardiovascular disease prevention and control. World Health Organization
2. Widder B, Paulat K, Hackspacher J, Hamann H, Hutschenreiter S, Kreutzer C, Ott F, Vollmar J (1990) Morphological characterization of carotid artery stenoses by ultrasound duplex scanning. Ultrasound in Medicine & Biology 16(4):349–354, DOI http:10.1016/0301-5629(90)90064-J
3. de Weert TT, Ouhlous M, Meijering E, Zondervan PE, Hendriks JM, van Sambeek MR, Dippel DW, van der Lugt A (2006) In vivo characterization and quantification of atherosclerotic carotid plaque components with multidetector computed tomography and histopathological correlation. Arteriosclerosis, Thrombosis, and Vascular Biology 26(10):2366–2372, DOI 10.1161/01.ATV.0000240518.90124.57
4. Serfaty JM, Chaabane L, Tabib A, Chevallier JM, Briguet A, Douek PC (2001) Atherosclerotic plaques: classification and characterization with T2-weighted high-spatial-resolution MR imaging– an in vitro study. Radiology 219(2):403–410, DOI 10.1148/radiology.219.2.r01ma15403
5. Fayad ZA, Fuster V (2000) Characterization of atherosclerotic plaques by magnetic resonance imaging. Annals of the New York Academy of Sciences 902:173–186, DOI 10.1111/j.1749-6632.2000.tb06312.x
6. Seeger JM, Barratt E, Lawson GA, Klingman N (1995) The relationship between carotid plaque composition, plaque morphology, and neurologic symptoms. The Journal of Surgical Research 58(3):330–336, DOI 10.1006/jsre.1995.1051
7. Yuan C (2002) Identification of Fibrous Cap Rupture With Magnetic Resonance Imaging Is Highly Associated With Recent Transient Ischemic Attack or Stroke. Circulation 105(2):181–185, DOI 10.1161/hc0202.102121
8. Redgrave JNE, Lovett JK, Gallagher PJ, Rothwell PM (2006) Histological assessment of 526 symptomatic carotid plaques in relation to the nature and timing of ischemic symptoms: the Oxford plaque study. Circulation 113(19):2320–2328, DOI 10.1161/CIRCULATIONAHA.105.589044
9. Takaya N, Yuan C, Chu B, Saam T, Underbill H, Cai J, Tran N, Polissar NL, Isaac C, Fergu-

son MS, Garden GA, Cramer SC, Maravilla KR, Hashimoto B, Hatsukami TS (2006) Association between carotid plaque characteristics and subsequent ischemic cerebrovascular events: A prospective assessment with MRI - Initial results. Stroke 37(3):818–823, DOI 10.1161/01.STR.0000204638. 91099.91

10. Kwee RM, van Oostenbrugge RJ, Mess WH, Prins MH, van der Geest RJ, ter Berg JW, Franke CL, Korten AG, Meems BJ, van Engelshoven JM, Wildberger JE, Kooi ME (2013) MRI of carotid atherosclerosis to identify TIA and stroke patients who are at risk of a recurrence. Journal of Magnetic Resonance Imaging 37(5):1189–1194, DOI 10.1002/jmri.23918

11. Salem MK, Bown MJ, Sayers RD, West K, Moore D, Nicolaides A, Robinson TG, Naylor AR (2014) Identification of patients with a histologically unstable carotid plaque using ultrasonic plaque image analysis. European Journal of Vascular and Endovascular Surgery 48(2):118–125, DOI 10.1016/j. ejvs.2014.05.015

12. Jodas DS, Pereira AS, Tavares JMRS (2016) A review of computational methods applied for identification and quantification of atherosclerotic plaques in images. Expert Systems with Applications 46:1–14

13. Gentile C, Li S, Kar P, Karatzoglou A, Zappella G, Etrue E (2017) On context-dependent clustering of bandits. In: Proceedings of the 34th International Conference on Machine Learning - Volume 70, JMLR.org, ICML'17, pp 1253–1262

14. Li S, Karatzoglou A, Gentile C (2016) Collaborative filtering bandits. In: Proceedings of the 39th International ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, New York, NY, USA, SIGIR '16, pp 539–548, DOI 10.1145/2911451.2911548

15. Korda N, Szörényi B, Li S (2016) Distributed clustering of linear bandits in peer to peer networks. In: Proceedings of the 33rd International Conference on International Conference on Machine Learning - Volume 48, JMLR.org, ICML'16, pp 1301–1309

16. Song L, Hsu W, Xu J, van der Schaar M (2016) Using contextual learning to improve diagnostic accuracy: Application in breast cancer screening. IEEE Journal of Biomedical and Health Informatics 20(3):902–914, DOI 10.1109/JBHI.2015.2414934

17. Gutiérrez B, Peter L, Klein T, Wachinger C (2017) A multi-armed bandit to smartly select a training set from big medical data. In: Descoteaux M, Maier-Hein L, Franz A, Jannin P, Collins DL, Duchesne S (eds) Medical Image Computing and Computer-

Assisted Intervention MICCAI 2017, Springer International Publishing, Cham, pp 38–45

18. Kar P, Li S, Narasimhan H, Chawla S, Sebastiani F (2016) Online optimization methods for the quantification problem. In: Proceedings of the 22nd International Conference on Knowledge Discovery and Data Mining, ACM, New York, NY, USA, KDD '16, pp 1625–1634, DOI 10.1145/2939672. 2939832

19. Herbert C, Chandler A, Dinsmore R (1995) A Definition of Advanced Types of Atherosclerotic Lesions and a Histological Classification of Atherosclerosis. Circulation 92:1355–1374, DOI 10.1161/01. CIR.92.5.1355

20. Huang H, Virmani R, Younis H, Burke AP, Kamm RD, Lee RT (2001) The impact of calcification on the biomechanical stability of atherosclerotic plaques. Circulation 103(8):1051–1056, DOI 10. 1161/01.CIR.103.8.1051

21. Vukadinovic D, van Walsum T, Manniesing R, Rozie S, Hameeteman R, de Weert TT, van der Lugt A, Niessen WJ (2010) Segmentation of the outer vessel wall of the common carotid artery in CTA. IEEE Transactions on Medical Imaging 29(1):65–76, DOI 10.1109/TMI.2009.2025702

22. van Engelen A, Niessen WJ, Klein S, Groen HC, Verhagen HJM, Wentzel JJ, van der Lugt A, de Bruijne M (2014) Atherosclerotic plaque component segmentation in combined carotid MRI and CTA data incorporating class label uncertainty. PLoS ONE 9(4):1–14, DOI 10.1371/journal.pone. 0094840

23. Wintermark M, Jawadi SS, Rapp JH, Tihan T, Tong E, Glidden DV, Abedin S, Schaeffer S, Acevedo-Bolton G, Boudignon B, Orwoll B, Pan X, Saloner D (2008) High-resolution CT imaging of carotid artery atherosclerotic plaques. American Journal of Neuroradiology 29(5):875–882, DOI 10.3174/ajnr.A0950

24. de Graaf MA, Broersen A, Kitslaar PH, Roos CJ, Dijkstra J, Lelieveldt BPF, Jukema JW, Schalij MJ, Delgado V, Bax JJ, Reiber JHC, Scholte AJ (2013) Automatic quantification and characterization of coronary atherosclerosis with computed tomography coronary angiography: cross-correlation with intravascular ultrasound virtual histology. The International Journal of Cardiovascular Imaging 29(5):1177–1190, DOI 10.1007/s10554-013-0194-x

25. Yan H, Dai Y (2011) The comparison of five discriminant methods. In: International Conference on Management and Service Science, pp 1–4, DOI 10.1109/ICMSS.2011.5999201

26. Van 't Klooster R, Naggara O, Marsico R, Reiber JHC, Meder JF, Van Der Geest RJ, Touzé E, Oppenheim C (2012) Automated versus manual in vivo segmentation of carotid plaque MRI. American Journal of Neuroradiology 33:1621–1627, DOI 10.3174/ajnr.A3028

27. van Engelen A, Niessen WJ, Klein S, Groen HC, Verhagen HJM, Wentzel JJ, van der Lugt A, de Bruijne M (2012) Multi-feature-based plaque characterization in ex vivo MRI trained by registration to 3D histology. Physics in Medicine and Biology 57(1):241–256, DOI 10.1088/0031-9155/57/1/241

28. van Engelen A, Niessen W, Klein S, Groen H, van Gaalen K, Verhagen H, Wentzel J, van der Lugt A, de Bruijne M (2013) Automated segmentation of atherosclerotic histology based on pattern classification. Journal of Pathology Informatics 4(2):1–7, DOI 10.4103/2153-3539.109844

29. Vukadinovic D, Rozie S, Van Gils M, Van Walsum T, Manniesing R, Van Der Lugt A, Niessen WJ (2012) Automated versus manual segmentation of atherosclerotic carotid plaque volume and components in CTA: Associations with cardiovascular risk factors. International Journal of Cardiovascular Imaging 28:877–887, DOI 10.1007/s10554-011-9890-6