Deep learning method for aortic root detection

Pablo G. Tahoces [a,*], Rafael Varela [b], Jose M. Carreira [b]

[a] *Department of Electronics and Computer Science, Universidad de Santiago de Compostela, Santiago de Compostela, Spain*
[b] *Complejo Hospitalario Universitario de Santiago (CHUS), Santiago de Compostela, Spain*

A B S T R A C T

Background: Computed tomography angiography (CTA) is a preferred imaging technique for a wide range of vascular diseases. However, extensive manual analysis is required to detect and identify several anatomical landmarks for clinical application. This study demonstrates the feasibility of a fully automatic method for detecting the aortic root, which is a key anatomical landmark in this type of procedure. The approach is based on the use of deep learning techniques that attempt to mimic expert behavior.
Methods: A total of 69 CTA scans (39 for training and 30 for validation) with different pathology types were selected to train the network. Furthermore, a total of 71 CTA scans were selected independently and applied as the test set to assess their performance.
Results: The accuracy was evaluated by comparing the locations marked by the method with benchmark locations (which were manually marked by two experts). The interobserver error was 4.6 ± 2.3 mm. On an average, the differences between the locations marked by the two experts and those detected by the computer were 6.6 ± 3.0 mm and 6.8 ± 3.3 mm, respectively, when calculated using the test set.
Conclusions: From an analysis of these results, we can conclude that the proposed method based on pre-trained CNN models can accurately detect the aortic root in CTA images without prior segmentation.

1. Introduction

Computed tomography (CT) is a well-established 3D imaging modality that provides a map of a patient's anatomy. Key advantages of the modality are the high spatial resolution and high relative contrast of the images obtained, the short acquisition time, the wide field-of-view, and the feasibility of obtaining high-quality three-dimensional multiplanar reconstructions. These features make this modality particularly effective for providing accurate information on diseases related to the morphology of most anatomical structures.

It is feasible to extract precise and reliable information on characteristics related to length, diameter, volume, and other anatomical parameters from the slices acquired by this modality [1]. These features are quantifiable and can help in evaluating normality or severity, degree of variation, or state of a disease or an injury [2]. In addition, this quantitative information has served as a basis for the publication of several clinical guidelines related to different types of diseases including cardiovascular diseases [3]. Therefore, the development of new algorithms that can provide accurate and reproducible values for extracting these quantitative characteristics has become an important issue.

With regard to the thoracic aorta, CTA images have been widely applied to establish the presence of aneurysms, dissections, mural thrombi, or elongations. These are considered as main anomalies to be treated in this part of the aorta [4]. However, it is difficult to develop automatic computer tools that use information extracted from CTA images, to address such diseases. It is necessary to develop several preceding tasks to 1) isolate the voxels of the aorta from the remainder of the volume that constitute the CTA (aorta segmentation) and 2) detect and tag several anatomical landmarks of regular use for extracting quantitative information [5].

Different algorithms have been proposed for automated aortic segmentation based on CTA volume. In general, the segmentation process starts with the automatic detection of the aorta in one of the CTA slices. The consideration of the first slice that is related to an anatomical landmark (such as the Carina [6] or the pulmonary trunk [7]) or is identifiable by its shape using the Hough transform [8,9], has been established to be a reliable solution for this task. Either a starting point or an initial contour is extracted from this first slice, and the segmentation process starts and continues until a pre-established criterion terminates it. However, the implementation of a termination criterion that

---

* Corresponding author.
E-mail address: pablo.tahoces@usc.es (P.G. Tahoces).

is suitable for all circumstances is challenging. More recently, deep learning solutions have also been proposed [10]. Nevertheless, in many cases, the segmentation process fails because leaks cause the process to advance inside the heart. Thus, an accurate detection of the aortic root before segmentation can improve the performance of segmentation algorithms.

Meanwhile, the detection of the aortic root has been the motivation for the development of different methods related to i) the characterization of the aorta to assess the relative location of a possible aneurysm [11,12] or analyze the stiffness of the aortic arch [13,14], ii) the detection and characterization of the coronary arteries to assess the presence of calcium deposits [15,16], and iii) the extraction of the dimensions of the aortic root prior to interventions [17]. Therefore, the detection of the aortic root is a fundamental issue to be considered while developing computer applications for aortic diseases either to improve segmentation or for tasks related to the extraction of measurements of clinical interest.

2. Related work

Manual and automatic detection schemes have been proposed in different studies for detecting aortic roots [18]. However, the implementation of automatic methods has been gaining increasing interest in the past few years [19]. The motivation for this is the elimination of the human factor, which is time-consuming and can yield results that introduce interobserver variations [20].

In general, ECG-gated CT angiography has been employed to detect the aortic root in preoperative scenarios related to planning for transcatheter aortic valve intervention (TAVI) [21]. This modality is highly recommended for use when precise measurements are required. This is because it reduces the presence of motion artifacts that increase the variability of the measurements obtained [22]. From this perspective, Elatar et al. [5] proposed a method for detecting the sinotubular junction and two coronary ostia in CTA images. Lalys et al. [23] proposed a method for automatically detecting aortic leaflets and coronary ostia locations to derive anatomical measurements from these landmarks.

Meanwhile, the application of the aortic root as a reference point has been proposed to extract quantitative information related to the shape and size of the aorta. Conventional CTA (non-ECG-gated) or MRI is used for this task. Kurugol et al. [24] and Tahoces et al. [25] proposed fully automated pipelines for calculating the aortic morphology in large cohorts of CTA scans. Herein, several landmarks were detected automatically for extracting features related to aortic morphology.

This study aimed to develop a method to automatically detect the aortic root from images acquired from CTA studies. Three main constraints were imposed. First, we considered cases involving only the thoracic aorta as well as those involving both thoracic and abdominal aorta in the same scan volume. Thus, ECG non-gated cases, constituted the main target of our approach. Second, we did not perform prior segmentation of the aorta. Third, the number of cases used for training was relatively small.

3. Dataset and methods

3.1. Dataset

CTA scans of 140 cases with approximately 79, 000 images were used to train and test the proposed scheme. For comparison with previous studies [25], only 39 cases (16, 353 images) were used for training, whereas 30 cases (18, 747 images) were used for validation. Furthermore, after the training was completed, the performance of the proposed algorithm was tested on the remaining 71 cases (43, 501 images).

All the CTA examinations were performed on patients from the Complejo Hospitalario Universitario de Santiago (CHUS) who were referred for CTA examination considering the clinical indications that they displayed. The indications were suspected thoracic aortic disease or

thoracic aortic disease with extension to the abdominal aorta, follow-up controls of patients diagnosed earlier, and controls of treatments performed on patients diagnosed and treated with endovascular or open surgery. In all the cases, the thoracic aorta was included in the examination. The abdominal aorta was also included when necessary. The dataset included 37 women and 103 men with an average age of 66 years (ranging from 27 to 89 years). The following conditions were imposed for the selection of cases: i) the aortic root should be visible and correctly opacified by the contrast agent in the CTA volume, and ii) the slice thickness should be 0.625 mm.

The number of slices per scan ranged from 280 (minimum) to 1, 128 (maximum). This implies that in certain cases, only the section corresponding to the thoracic aorta was scanned during the acquisition process, whereas in several other cases, both the thoracic and abdominal sections of the aorta were scanned. In all the case, the aortic sections included in a study depended on the study's clinical indications.

All the cases were annotated by two experts, who independently marked two points on the image: i) the center of the aortic root, which is located in the sinus of Valsalva (SOV) and ii) a point within the ascending aorta that provides the orientation of its centerline. The annotation process was performed using a tablet equipped with a touch screen and installed with the ITK-SNAP program, and a digital pen. Annotation was performed in the coronal plane, although the axial and sagittal views were simultaneously available to the expert. The SOV commissure was considered as a reference for the position of the center of the aorta. (Fig. 1).

3.2. Proposed detection scheme

Our scheme attempts to mimic the manner in which radiologists perform aortic root detection. First, the CTA axial slices were transformed into coronal slices (Fig. 2a). This step makes it more convenient to distinguish the location of the SOV. Then, the radiologist analyzed the region located within the ribcage and navigated through the coronal slices until the heart was visible (Fig. 2b). We call the *Heart Box (HB)* to this ribcage section. It includes the coronal slices where the heart is visible. Then, the slices that belonged to this volume were analyzed in detail to locate the center of the aortic root (Fig. 2c). Once located, the orientation of the ascending aorta was determined. We call the plane that includes the center of the aortic root as *SOV plane*. It is normal to the vector following the centerline of the ascending aorta.

Therefore, with regard to the algorithm's design, there are three main tasks in the entire process: classification, detection, and
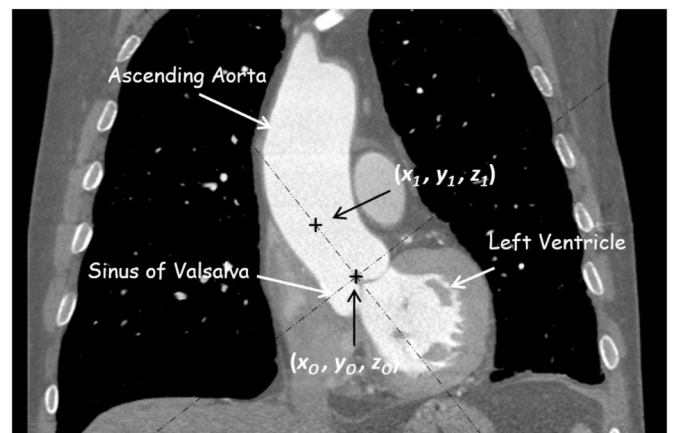

Fig. 1. Coronal view of a CT scan. The black crosses indicate the two marks inserted by the expert, i.e., the commissure of the SOV location (x_0, y_0, z_0) and a point on the aorta's centerline (x_1, y_1, z_1). The dashed black lines indicate the slopes of both the ascending aorta and the plane where the SOV is located that are calculated from the marks inserted by the expert.
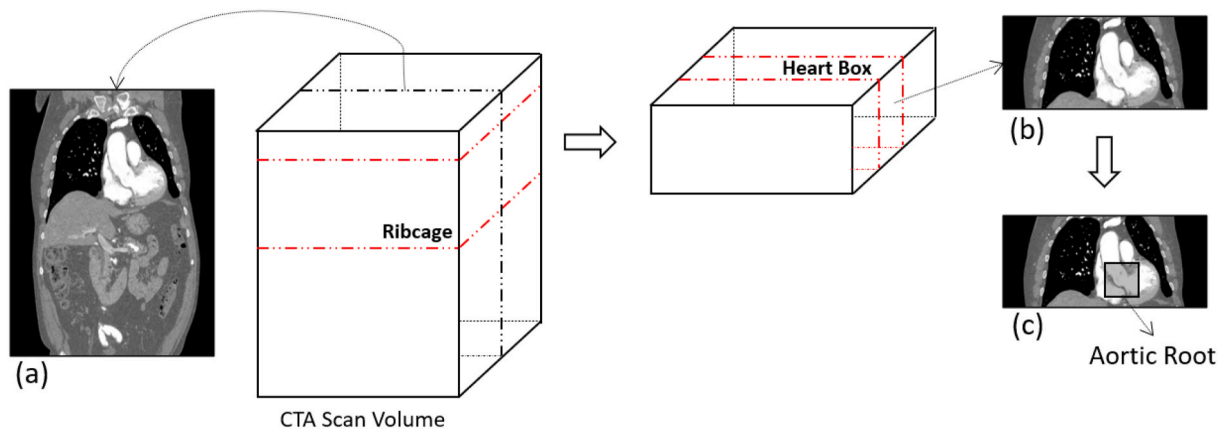
Fig. 2. Schematic representation of aortic root detection by radiologists: (a) Among the coronal views of the initial CT volume, the radiologist focuses on the ribcage area. (b) Within the ribcage volume, he selects coronal slices in which the heart is visible. (c) Finally, the aortic root is located between these slices.

characterization. i) A classification task is necessary to obtain axial sections in which the presence of the lungs is detected. Thereby, the volume containing the ribcage is isolated (ribcage bounding extraction). This eliminates the likelihood that the axial slices belonging to the remaining volume interfere with the calculation in the subsequent steps. ii) A classification task is also carried out to extract the coronal slices that include the HB from the ribcage volume (HB bounding extraction). iii) A detection process is performed to determine the aortic root's location within the HB (aortic root detection). iv) Finally, principal component analysis (PCA) is performed to determine the orientation of the ascending aorta. Thereby, the SOV plane can be calculated (aortic root characterization). Fig. 3 shows the overall scheme of the proposed method.

3.3. Classification: extraction of ribcage and HB

We trained models based on convolutional neural network (CNN) architectures to perform this task. We applied three architectures frequently used in medical imaging to construct these models: VGG, ResNet, and Inception. To summarize, each CNN is composed of two essential elements, namely, a convolutional base (where a succession of convolutional filters and pooling layers are defined) and a classifier block (where all the nodes are interconnected), which are generally constructed through dense layers. The convolutional filters extract the fundamental characteristics of the images in terms of shape and orientation. The pooling layers alter the size of the extracted features. An additional layer (flattened layer) functions as an interface to adapt the format of the data obtained from the convolutional base to the classifier. For a convolutional base, each model defines its topology. VGG consists of one or more blocks that are composed of a succession of convolutional filters and terminated by a pooling layer (Fig. 4a). ResNet is composed of blocks of residual connections, which provides the output of a layer as an input for a subsequent layer (Fig. 4b). The inception block consists of several filters placed into different parallel branches that converge into a concatenate layer (Fig. 4c).

We demonstrated three scenarios: i) training from scratch (VGG16, ResNet50, and InceptionV3), ii) use of pre-trained values while maintaining the convolutional base and retraining the classifier (VGG16

(pre), ResNet50 (pre), and InceptionV3 (pre)), and iii) use of pre-trained values while partially maintaining the convolutional base and retraining the classifier (VGG16 (pre, blk), ResNet50 (pre, blk), and InceptionV3 (pre, blk)). Models pre-trained on the ImageNet dataset were used for transfer learning in all the cases. We also applied three VGG architectures with different numbers of convolutional blocks (VGG1, VGG2, and VGG3) that were trained from scratch.

To homogenize the input of the proposed models, black and white images of 128 × 128 were used when the training process was performed from scratch. However, to effectively utilize pre-trained models, it is necessary to accommodate the input to each model's requirements. Therefore, we transformed the input images to 224 × 224 color images for VGG16 and ResNet50, and to 299 × 299 color images for InceptionV3. This was performed by adapting the input layers of the different models to the requirements of each case.

We used the general method to train the models. First, we trained on a training dataset using the validation set to fine-tune the parameters. Then, we used the test set to evaluate the performance of the trained model independently. The results obtained from the test set were used to compare and select the best models. The models were constructed in Python using the Keras library with TensorFlow (GPU version). Binary cross-entropy was used for the loss function in all the cases. A gradient descent (with momentum) optimizer was used to minimize the loss function using accuracy as a metric. All the intermediate activation layers used ReLU as an activation function. The output layer used a sigmoid function to deliver values in the range [0, 1]. A total of 20 epochs were performed for each model, and the model that achieved the highest accuracy in the validation set was saved for testing. The learning rate was maintained constant throughout the training process.

As mentioned above, classification was used for two tasks in the overall scheme: 1) Ribcage bounding extraction to select the slices that belong to the ribcage from the full set of axial slices of the CTA scan and 2) HB bounding extraction to select the coronal slices that include the heart, from the set of coronal slices. Therefore, two CNN networks were trained to perform these tasks.

To train and evaluate the first classification model (ribcage bounding extraction), the full set of axial slices was split into three datasets (training, validation, and test). Then, a sparse selection process was
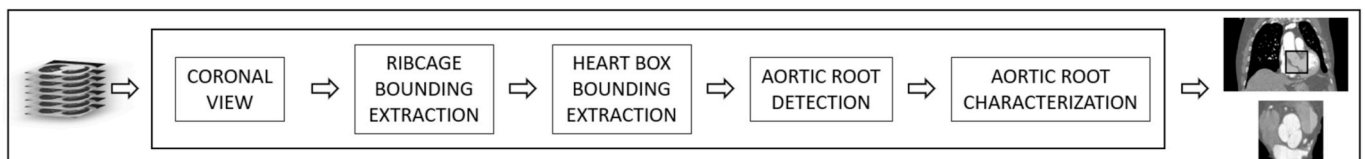

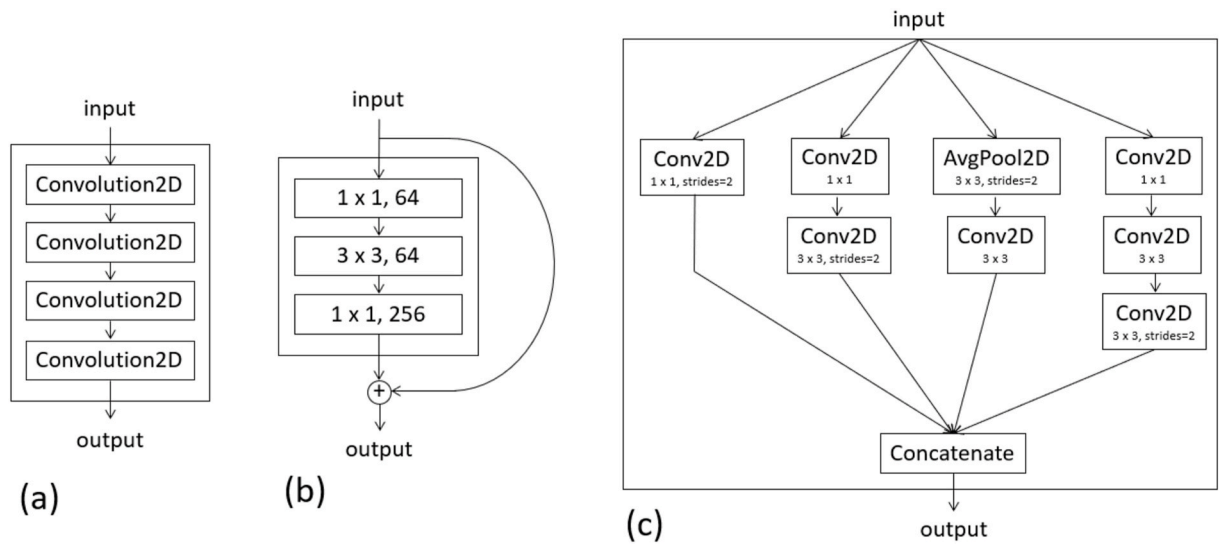
Fig. 3. Flowchart of the method.

Fig. 4. Fundamental topologies of the convolutional base of the CNN architectures used. (a) VGG, (b) ResNet, and (c) Inception.

designed to maintain a balanced number of cases (with and without lungs). Thereby, a similar number of axial slices with and without lungs were selected for each CTA volume. However, the criteria for determining which slices include or does not include the lungs are not stringent. Thus, only the axial slices wherein the presence of the lungs was reflected in a high percentage of voxels were labeled as lung slices. The remaining samples were labeled as "without a lung."

To train and evaluate the second classifier (HB bounding extraction), the set of coronal slices obtained from the ribcage volume was also split into three datasets (training, validation, and test). Slices wherein both the aorta and heart were visible were tagged as *Heart*, those wherein neither the aorta nor the heart was visible were labeled as *Not Heart*, and the remaining ones (wherein the heart may have been visible whereas the aorta was not) were not labeled. As a result, a higher variability can be observed in coronal slices that do not contain the aorta. In contrast, the slices containing the aorta are highly similar. Therefore, we decided to construct unbalanced datasets because the *Not Heart* slice datasets were significantly larger than the *Heart* datasets.

3.4. Aortic root detection

The HB classifier's output is a number that represents the probability that a coronal view would include the aortic root in the image. Therefore, the classifier would assign a relatively high probability to all coronal views that include the left ventricle connected to the ascending aorta. However, to obtain the position of the aortic root, we need to calculate its coordinates ($x_{SOV}, y_{SOV}, z_{SOV}$) within the entire volume of the CTA. Therefore, y_{SOV} would be the result of selection from the set of coronal views, which is closest to the center of the SOV. To obtain this value, we first calculate $y_{SOV_{pre}}$ using the following expression:

$$y_{SOV_{pre}} = \frac{\sum_k k \cdot p_{TH}(k)}{p_{TH}(k)} \qquad (1)$$

where $p_{TH}(k)$ is the output of the HB classifier for coronal view k. It has a value above the threshold TH_1. Hence, only coronal slices with a high probability of including the aortic root ($TH_1 = 0.9$, in our case) were included in this computation.
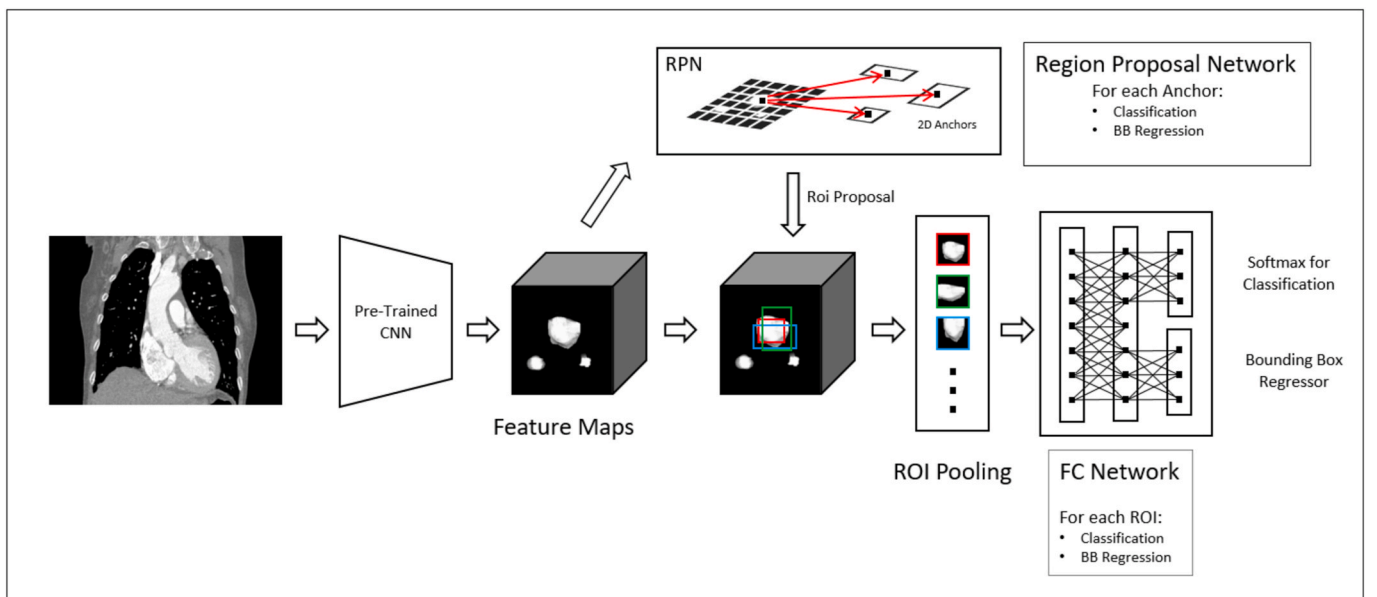

Fig. 5. Diagram of the architecture of a Faster R–CNN.

To calculate the final values of the aortic root coordinates ($x_{SOV}, y_{SOV}, z_{SOV}$), we trained a new model based on the faster region-based convolutional neural network (Faster R–CNN) [26]. This network consists of three main stages (Fig. 5). In the first stage, a CNN transforms the input image into feature maps. Pre-trained CNN models can be used for this task. In our case, a ResNet101 pre-trained on the MSCOCO object detection dataset was used as a pre-trained network. In the second stage, a region proposal network (RPN) proposes bounding boxes of candidate objects related to objects contained in the image. Several square regions of different sizes and aspect ratios (2D anchors) are used to perform this. The ROI pooling step accommodates the different sizes of the ROIs proposed in the second stage by matching these. The final step performs classification and bounding-box regression for each candidate ROI selected in the second stage by using fully connected layers (FC network). The features used by these FC layers originate from the feature maps obtained during the first stage. Finally, the output is forked into two branches: (i) one associated with a softmax layer that produces an estimate of the probability of belonging to a class and (ii) another associated with a linear regressor that generates four numbers. These numbers code the positions of the bounding boxes of the object detected in the input image.

In our case, the output of this network was a rectangular ROI bounded by four numbers (x_1, x_2, z_1, z_2) (see Fig. 6) calculated by the regressor of the network and the probability (q) that this ROI contains the aortic root. To calculate y_{SOV}, we applied the faster R–CNN model to the $2 \times W + 1$ coronal slices centered at the $y_{SOV_{pre}}$ computed previously ($W = cte$). Then, we calculated y_{SOV} as

$$y_{SOV} = \frac{\sum_k k \cdot q_{TH}(k)}{q_{TH}(k)} \tag{2}$$

where (again) $q_{TH}(k)$ is the probability delivered by the classifier of the Faster R–CNN network for coronal view k. It has a value above the threshold (TH_2). Thus, the final value for y_{SOV} was selected by the Faster R–CNN model. Moreover, x_{SOV} and z_{SOV} can be calculated as follows:

$$x_{SOV} = \frac{x_1 + x_2}{2} \tag{3}$$

$$z_{SOV} = \frac{z_1 + z_2}{2} \tag{4}$$

3.5. Aortic root characterization

The result of the detection step was a 2D ROI, which includes a portion of the ascending aorta, the SOV, and a portion of the left ventricle (Fig. 6a). To obtain an image of the aortic root displaying the SOV, we calculated the SOV plane as defined above. To obtain this plane and the direction of the ascending aorta's centerline, we used principal component analysis (PCA) on the 2D ROI. Thus, we obtained the covariance matrix of the coordinate vector of the pixels included in the ROI and calculated their eigenvectors and eigenvalues. The eigenvector with a higher eigenvalue corresponds to the direction of the ascending aorta in the coronal plane, at least in the vicinity of the heart. The following is a brief description of the procedure:

i) The 2D ROI obtained during the detection step is pre-processed by transforming the pixel values that belong to its upper right and the left bottom corners to zero (Fig. 6b). Thus, the contribution of these regions to the calculation of the covariance matrix is zero. This step is necessary to prevent the presence of artifacts in the computation of an optimal threshold (TH_3). It is performed as follows:

$$ROI_{pre}(x, z) = \begin{cases} 0 & \text{if } x < \frac{1}{4}x_{SOV} \text{ and } z < \frac{1}{4}z_{SOV} \\ 0 & \text{if } x > \frac{3}{4}x_{SOV} \text{ and } z > \frac{3}{4}z_{SOV} \\ ROI(x, z) & \text{any other case} \end{cases} \tag{5}$$

ii) To obtain the covariance matrix of pixel locations (x_i, z_i), we binarize the pixels of the ROI (Fig. 6c). Thus, only the positions of the pixels whose value is one become part of the covariance matrix. A thresholding algorithm is applied to the pixels of the ROI to perform this task. Thus, pixels below a threshold (TH_3) become zero and those above it become one:
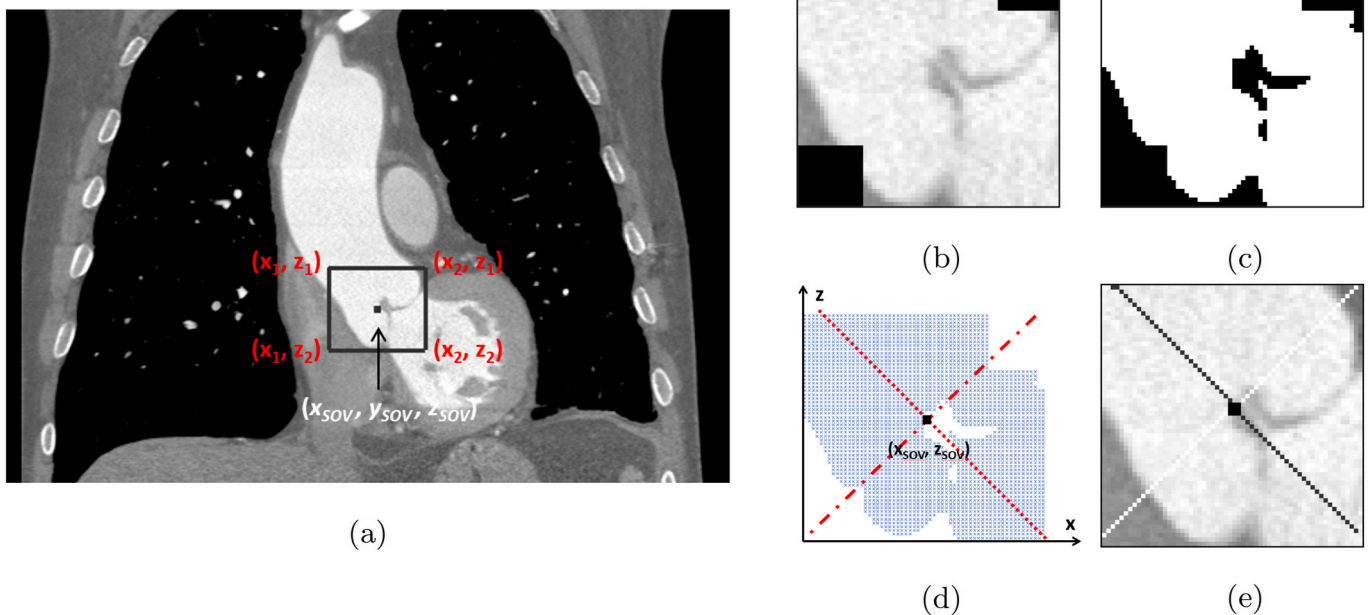

Fig. 6. (a) Coronal view superimposed with the aortic root detected. (b) and (c) preprocessing steps, (d) and (e) axis of the SOV plane and the aorta centerline computed using PCA.

$$ROI_{pre}(x,z) = \begin{cases} 0 & \text{if } ROI_{pre}(x,z) <= TH_3 \\ 1 & \text{if } ROI_{pre}(x,z) > TH_3 \end{cases} \tag{6}$$

The value of TH_3 was obtained from a smoothed version of the histogram of each ROI by identifying the first flat zone before the peak representing both the ascending aorta and left ventricle.

iii) We collect the coordinates of the pixels of ROI_{pre} that are non-zero and store these in the vectors x_{coord} and z_{coord} (Fig. 6d). Thus, the k-element of these vectors is

$$\left. \begin{array}{l} x_{coord}(k) = x_i \\ z_{coord}(k) = z_i \end{array} \right\} \text{if } ROI_{pre}(x_i, z_i) \neq 0, \; \forall i \in ROI_{pre} \tag{7}$$

where $0 < k < n$ and n is the number of non-zero pixels of ROI_{pre}.

iv) We compute the covariance matrix cov_{coord} between the two co-ordinate vectors x_{coord} and z_{coord}.

$$cov_{coord}(x_{coord}, y_{coord}) = \frac{\sum_{i=1}^{k}\left(x_{coord_i} - \overline{x_{coord}}\right)\left(z_{coord_i} - \overline{z_{coord}}\right)}{k-1} \tag{8}$$

v) We compute the eigenvectors (A) and eigenvalues (B) of the covariance matrix.

$$A = \begin{pmatrix} A_{00} & A_{01} \\ A_{10} & A_{11} \end{pmatrix} \tag{9}$$

$$B = \begin{pmatrix} B_0 \\ B_1 \end{pmatrix} \tag{10}$$

vi) We identify the highest eigenvalue (B_k) and select its associated eigenvector (A_{0k}, A_{1k}) that would be the main component of the 2D ROI. That is, it would follow the direction of the ascending aorta in the proximity of the heart (Fig. 6e). Its slope m is

$$m = \frac{A_{1k}}{A_{0k}} \tag{11}$$

Thus, the straight line

$$z = z_{SOV} + m(x - x_{SOV}) \tag{12}$$

corresponds to the computed centerline of the ascending aorta (the black line in Fig. 7a) and its corresponding orthogonal (white line in Fig. 7a).

$$z = z_{SOV} - \frac{1}{m}(x - x_{SOV}) \tag{13}$$

belongs to the SOV plane, i.e., the plane defined by the normal vector $\vec{n}(A_{0k}, A_{1k})$.

vii) Based on this normal vector $\vec{n}$ we can define the plane:

$$A_{0k}x + A_{1k}z + C = 0 \quad \text{with} \quad C = -A_{0k}x_{SOV} - A_{1k}z_{SOV} \tag{14}$$

where the SOVs are visible (Fig. 7b).

3.6. Performance evaluation

To determine the performance of the two-category classification tasks, we constructed the confusion matrix and then computed the metrics precision and recall:

$$Precision = \frac{TP}{TP + FP} \tag{15}$$

$$Recall = \frac{TP}{TP + FN} \tag{16}$$

In the last two expressions, True Positives (TP) are the slices that include the *ribcage/HB* and were classified as containing the *ribcage/HB*, TN (True Negatives) are the slices that do not include the *ribcage/HB* and were classified as not containing the *ribcage/HB*, FP (False Positives) are the slices that do not include the *ribcage/HB* and were classified as containing the *ribcage/HB* (False Positives), and FN (False Negatives) are the slices that include the *ribcage/HB* and were classified as not containing the *ribcage/HB*.

In addition, to obtain a single value related to the performance, we computed the accuracy by

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

and the F1 score:

$$F1 \; score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{18}$$

The accuracy is particularly effective for analyzing the performance of ribcage detection because the number of positives is balanced well with the number of negatives. In contrast, the F1 score is effective for analyzing the balance between precision and recall where there is a non-uniform distribution. This is the case for HB detection, as mentioned above.

The performance achieved by the Faster R–CNN network for aortic root detection was calculated in terms of the mean Euclidean distance between the positions calculated by the algorithm (x_{SOV}, y_{SOV}, z_{SOV}) and
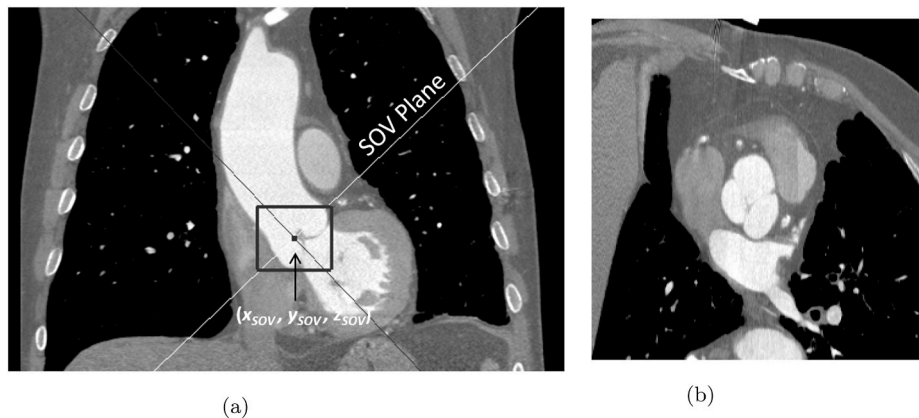

(a)

(b)

Fig. 7. Example of the result of the method. (a) Coronal view with the box of the aortic root calculated, including the ($x_{SOV}, y_{SOV}, z_{SOV}$) dot (black dot). The two lines representing the centerline of the ascending aorta (black straight line) and the SOV plane (white straight line) are superimposed. (b) Reformatted image following the SOV plane.

those marked by the two experts who participated in the study ($(x_{0_i}, y_{0_i}, z_{0_i})$, $i = 1, 2$). Finally, to analyze the performance of the characterization step, we calculated the difference between the tilt angles of the SOV plane as determined by the algorithm (θ_{SOV_c}) and that determined from the two marks entered manually by the experts ($\theta_{SOV_{0_i}}$, $i = 1, 2$).

4. Experiments and results

Fig. 7a shows the result of applying our method to the example case. The straight lines corresponding to the SOV plane (Fig. 7b) and the ascending aorta's centerline calculated by our method are also included. Fig. 8 shows the results obtained applying our method to different cases included in our test set. The marks introduced by one of the experts (orange color) and the aortic root's location calculated by the computer
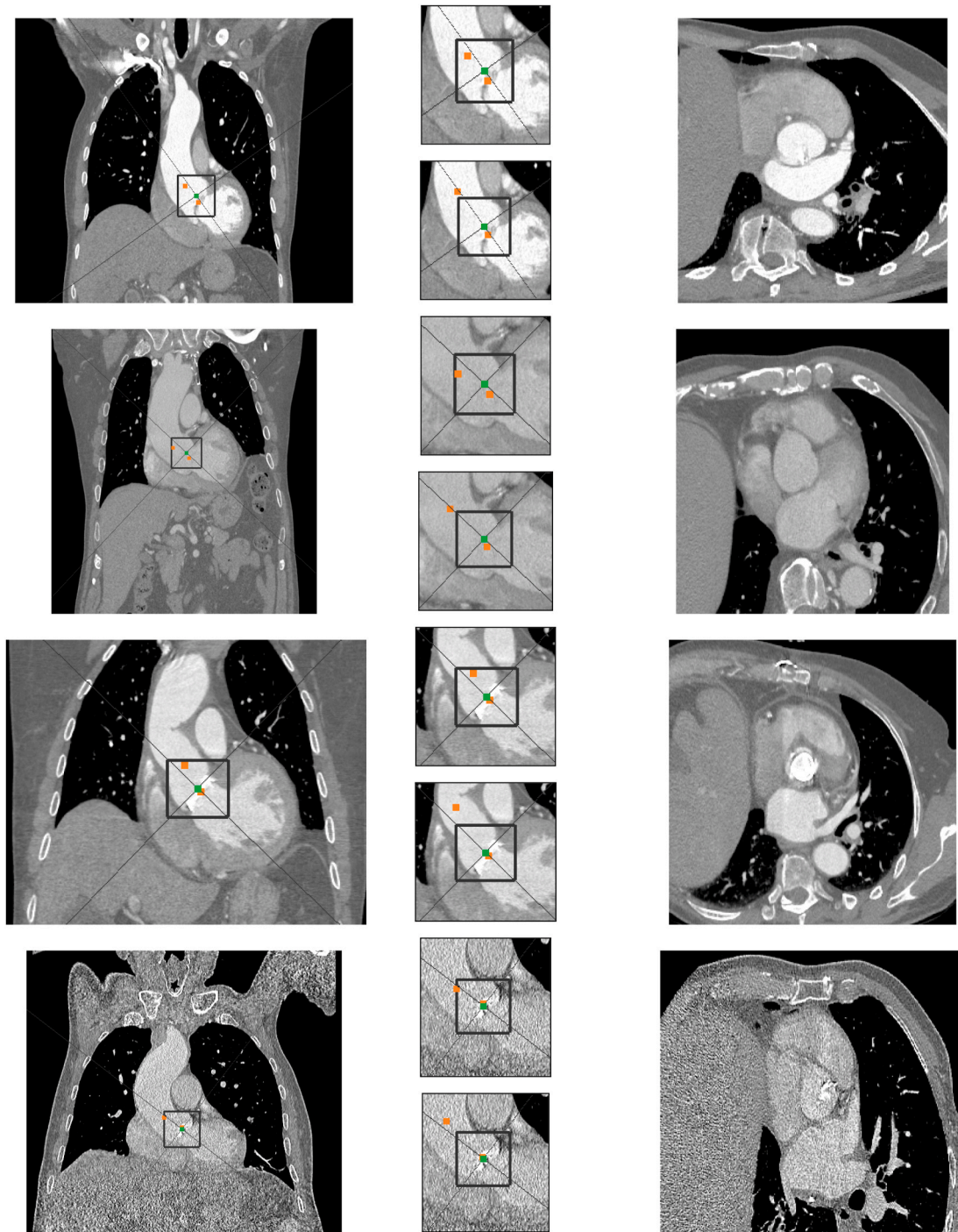

Fig. 8. Results obtained by our method under different assumptions. From upper to lower: (a) normal case. (b) low-contrast case. (c) case with a prosthetic heart valve inserted. (d) noisy case. The images depict a coronal slice with the aortic root (left), the SOV plane (right), and the ROIs with i) the location of the two points marked by the two experts (orange dots), ii) the root position computed by the algorithm (green dot), and iii) the axes representing both the central line of the aorta and the orientation of the SOV calculated by the computer (black lines) (center). The dots marked by *Expert*$_1$ are in the upper ROI, and those marked by *Expert*$_2$ are in the bottom ROI.

(green color) are superimposed. The cases selected reveal the behavior of our algorithm in four scenarios: i) normal case, ii) low contrast case, iii) case with prosthetic heart valve inserted, and iv) noisy case. The images include the location of the two points marked by the expert and the aortic root position calculated by our method.

4.1. Ribcage and HB bounding extraction

The ribcage and HB bounding extractions were implemented in terms of a binary classification task as explained above. Therefore, two models were specially trained for this purpose. Axial slices were used as inputs to the model for ribcage bounding extraction, whereas coronal slices were used for HB bounding extraction.

The performances and configurations of the three CNN architectures were computed. Tables 1 and 2 present a comparison of the results obtained from these approaches after their application to the cases included in the test set. We can observe that the complex architectures VGG16, ResNet50, and InceptionV3 achieved their best results for pre-trained models when both the final layers and the classification stage were retrained ((pre, blk) models, see Section 3.3). It is important to note that the accuracy achieved for the models VGG16, ResNet50, and InceptionV3 when trained from scratch was reduced to approximately 50%.

According to the results depicted in Table 1, we can conclude that the VGG16 (pre, blk) model achieved the highest classification accuracy (98.0%). Similar results were achieved for the HB bounding extraction (Table 2), where the f1 score was 0.72 for this configuration. Therefore, the models based on VGG16 (pre, blk) were finally employed to isolate the HB volume, from which the aortic root was extracted in the next step.

4.2. Aortic root detection

The differences between the positions marked by the experts and those obtained using the computer were calculated. The differences between the positions marked by the two experts were also calculated to estimate the interobserver error. The differences between the experts served as a reference for comparison with the results of the algorithm.

The QQ plot was analyzed to verify the normal distribution of the calculated distances of the test set. We had observed that most of the points are located close to the straight line corresponding to the diagonal of the plot. Similar results were obtained in the training and validation cases. Therefore, we can assume a normal distribution for these distances. Based on these results, the mean value and standard deviation of the distance were calculated to estimate the error of our method (Table 3).

Next, we analyzed the outliers in the upper right corner of the QQ-plots. These were considered the most relevant because these correspond to cases in which the difference between the position determined by the experts and that identified by the computer is more significant. We concluded that in general, these cases are related to the presence of artificial heart valves or an incorrect location of the root position because of the selection of an inaccurately determined coronal slice.

A detailed analysis of the values obtained for the test set and the calculation of the differences (separated by views) revealed that the worst results were obtained for the coronal view. This is true both in terms of the mean distance and dispersion of the values obtained (Table 4). These results are consistent with those obtained when the comparison was made with the experts' marks. In addition, in this case, the worst results were obtained with the selection of the coronal view, for both average value (2.8 for the coronal view vs. 2.1 and 1.9 for the sagittal and axial views, respectively) and dispersion (2.3, vs. 1.7 and 1.3, respectively).

An analysis of the number of cases in the test set revealed that the distance between the location marked by the expert and that calculated by the algorithm exceeds 10 mm for some patients (Table 5). We can observe that this scenario occurs only in the sagittal and coronal views. That is, the coronal view has the worst result in this case as well.

QQ-plot analysis was also performed to verify the normal distribution of the differences between the slope of the SOV plane computed from the locations marked by the experts and the slope calculated by the algorithm. Both mean value and standard deviation of these differences were calculated based on these results. The results for the different datasets are shown in Table 6.

Bland—Altman plots (Figs. 9 and 10) were plotted to analyze the agreement between the two procedures (manual and automatic) proposed to quantify the SOV slope. As is evident, biases exist between the two methods. If we compare these biases ($-7.2°$ for $Expert_1$ vs. *Computer*, and $-10.4°$ for $Expert_2$ vs. *Computer*), the values are negative and close in both cases. Moreover, the results obtained for most cases were within the 95% confidence interval. This indicates the feasibility of introducing a correction factor to correct for bias.

5. Discussion

The automatic extraction of quantitative information from medical images is an important challenge that has been addressed with a certain degree of success in recent years. However, to achieve optimal results, it is necessary to obtain effective segmentation of the organ/structure to be analyzed and the precise location of several reference points used to extract such information.

In this study, we focused on the automatic detection of the aortic root. The tilt of the SOV plane was also obtained after the detection. The developed method consists of three main steps: classification, detection, and slope calculation. For classification and detection, we used models based on different CNN architectures. A PCA-based method was proposed for the calculation of the slope.

For classification, models based on networks with complex architectures (namely, as VGG16, ResNet50, or InceptionV3) were demonstrated to be invalid when trained from scratch, owing to overtraining. In this case, accuracies close to 50% were obtained, which implies that they could not distinguish between the groups. The large number of parameters to be tuned (over 20 million) and the limited size of the datasets hindered the obtainment of optimal solutions for such complex networks. However, the results improved substantially when pre-trained models were used on these architectures. We obtained reliable classifiers

Table 1
Summary of results of different CNN architectures used for ribcage extraction.

CNN	NPTT	Acc (%)	Prec	Rec	f1	TP	TN	FP	FN
VGG1	16787617	97.3	0.96	0.99	0.97	21019	21805	963	203
VGG2	8454433	97.4	0.97	0.97	0.97	20680	22146	622	542
VGG3	4629153	97.4	0.96	0.98	0.97	20864	21999	769	358
VGG16 (pre)	1048833	96.5	0.95	0.97	0.96	20672	21784	984	550
VGG16 (pre, blk)	3408641	98.0	0.97	0.98	0.98	20886	22209	559	336
ResNet50 (pre)	1312257	95.8	0.92	0.99	0.96	21064	21060	1708	158
ResNet50 (pre, blk)	5777921	97.5	0.97	0.98	0.97	20696	21194	574	526
InceptionV3 (pre)	1312257	50.9	0.49	0.39	0.43	8255	14149	8619	12967
InceptionV3 (pre, blk)	7385793	95.9	0.97	0.94	0.96	20016	22162	606	1196

Table 2

Summary of results of different CNN architectures used for HB extraction.

CNN	NPTT	Acc (%)	Prec	Rec	f1	TP	TN	FP	FN
VGG1	16787617	92.0	0.42	0.97	0.58	2049	31877	2877	61
VGG2	8454433	89.5	0.35	1.00	0.52	2110	30888	3866	0
VGG3	4629153	92.9	0.44	0.97	0.61	2039	32200	2554	71
VGG16 (pre)	1048833	92.6	0.44	0.98	0.60	2065	32089	2665	45
VGG16 (pre, blk)	3408641	95.9	0.59	0.95	0.72	2001	33335	1419	109
ResNet50 (pre)	1312257	76.6	0.19	0.96	0.32	2036	26204	8550	74
ResNet50 (pre, blk)	5777921	94.1	0.49	0.86	0.62	1819	32862	1892	291
InceptionV3 (pre)	1312257	76.6	0.17	0.80	0.28	1681	26558	8196	429
InceptionV3 (pre, blk)	7385793	87.7	0.31	0.94	0.47	1977	30341	4413	133

Table 3

Distance between the location of the aortic root calculated by our method and that determined by two experts.

Data Set	Cases	Expert 1 vs Computer	Expert 2 vs Computer	Expert 1 vs Expert 2
		(mm)	(mm)	(mm)
Training set	39	4.9 ± 2.5	6.8 ± 3.3	5.5 ± 2.8
Validating Set	30	7.5 ± 2.7	7.1 ± 3.1	5.7 ± 3.1
Test Set	71	6.6 ± 3.0	6.8 ± 3.3	4.6 ± 2.3

Table 4

Distance between the location of the aortic root as calculated by our method and that determined by the two experts for the test set, separated by view.

View	Expert 1 vs Computer	Expert 2 vs Computer	Expert 1 vs Expert 2
	(mm)	(mm)	(mm)
Sagittal	3.6 ± 2.5	3.6 ± 2.7	2.1 ± 1.7
Coronal	3.9 ± 3.0	3.9 ± 3.2	2.8 ± 2.3
Axial	2.5 ± 1.8	2.7 ± 1.8	1.9 ± 1.3

Table 5

Detailed analysis of the values obtained for the test set. The maximum distance between the aortic root locations determined by our method and that marked by each of the two experts is depicted in the *Max Diff* column. The number of cases where the distance between the value calculated by the algorithm and that by each of the experts exceeds 10 mm is depicted in the *NC* column.

View	Expert 1 vs Computer		Expert 2 vs Computer	
	Max Diff	NC ≥ 10	Max Diff	NC ≥ 10
Sagittal	11.0	1	13.8	2
Coronal	15.6	2	17.6	2
Axial	7.5	0	6.9	0

Table 6

Differences between the SOV plane slope computed from the locations marked by the experts and that computed by the algorithm.

Data Set	Expert 1 vs Computer	Expert 2 vs Computer	Expert 1 vs Expert 2
	(degrees)	(degrees)	(degrees)
Training set	−10.9 ± 11.4	−10.4 ± 9.2	−0.5 ± 5.3
Validating Set	−5.5 ± 12.9	−10.7 ± 9.9	5.2 ± 13.3
Test Set	−7.2 ± 11.4	−10.4 ± 9.3	3.1 ± 6.5

cases where the classifier and specific layers of the convolutional part were retrained ((pre, blk) models). This indicates that the second classification is more complex, which is consistent with the fact that it was necessary to construct an unbalanced dataset for training. It is likely that a higher number of cases in the training set would improve these numbers. However, the model trained with VGG16 (pre, blk) attained accuracies close to 95.9% (0.72 for F1 score) and recall and precision of 0.95 and 0.59, respectively.

With regard to the detection of the aortic root, we verified that for the test set, the differences between the results obtained using the algorithm and the location marked by the experts follow a normal distribution. Accordingly, we can conclude that in 95% of the cases, the difference between the location marked by $Expert_1$ and the position calculated by the computer (see Table 3) is less than 12.6 mm (13.4 mm for $Expert_2$). Following this argument, the differences between the positions marked by the experts ($Expert_1$ vs. $Expert_2$) are less than 9.2 mm in 95% of the test set cases. Thus, the interobserver agreement was relatively similar regardless of whether the observers were humans or computers. It is important to note that the cases used in our study were acquired with non ECG-gated CT scanners. Therefore, an implicit error occurs while attempting to mark the exact location of the aortic root. Meanwhile, we know that the mean ascending aorta diameter is approximately 35 mm for men and 33 mm for women [27]. Thus, the error can be regarded as reasonable considering both interobserver agreement error and size of the aorta.

There are few papers in the literature on the automatic detection of aortic root in CTA images. Most of these are related to the detection of different landmarks during TAVI procedure planning. Thus, Elatar et al. [5] reported that automatic landmark detection had a mean error of 2.66 ± 1.63 mm and 2.96 ± 2.52 mm when compared with Observers I and II, respectively. The mean distance of the matched observers was 2.38 mm. Astudillo et al. [28] used Euclidean distance and reported a median difference of 1.5 (1.1, 2.0) mm for all landmarks combined. This was smaller than the difference of 2.0 (1.3, 2.8) mm achieved for the manually detected landmarks by the first and second observers. For Al et al. [29], the overall landmark error for patients without and with TAVI was 1.94 ± 0.93 mm and 2.74 ± 1.78 mm, respectively. Finally, in Lalys et al. [23], the results for the detection of the two ostia revealed distance errors of 1.80 ± 0.74 mm and 1.96 ± 0.87 mm.

To summarize, all the previous results show differences of approximately 2.0 mm between computer-annotated and hand-annotated landmarks, which is similar to those obtained while comparing observers. However, in all these cases, the problem to be solved is related to preoperative TAVI procedures. Consequently, the images used were ECG-gated CTAs. Therefore, a comparison with our method would be biased. In our method, the mean distance between the computer and the human observer was 6.7 mm, which is close to the interobserver error of 4.6 mm (see Table 3). In our previous work [25], the results obtained for aortic root detection on non-ECG-gated CTA are similar (5.7 ± 7.3 mm) to those reported here. However, in that case, the aorta was segmented previously and its centerline calculated. Considering the two facts (the use of non-ECG gated cases and no previous segmentation of the aorta), the results obtained by our new method are significant.

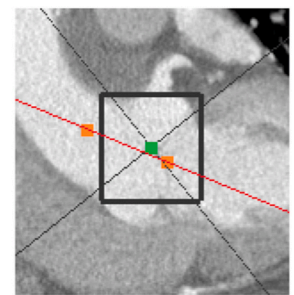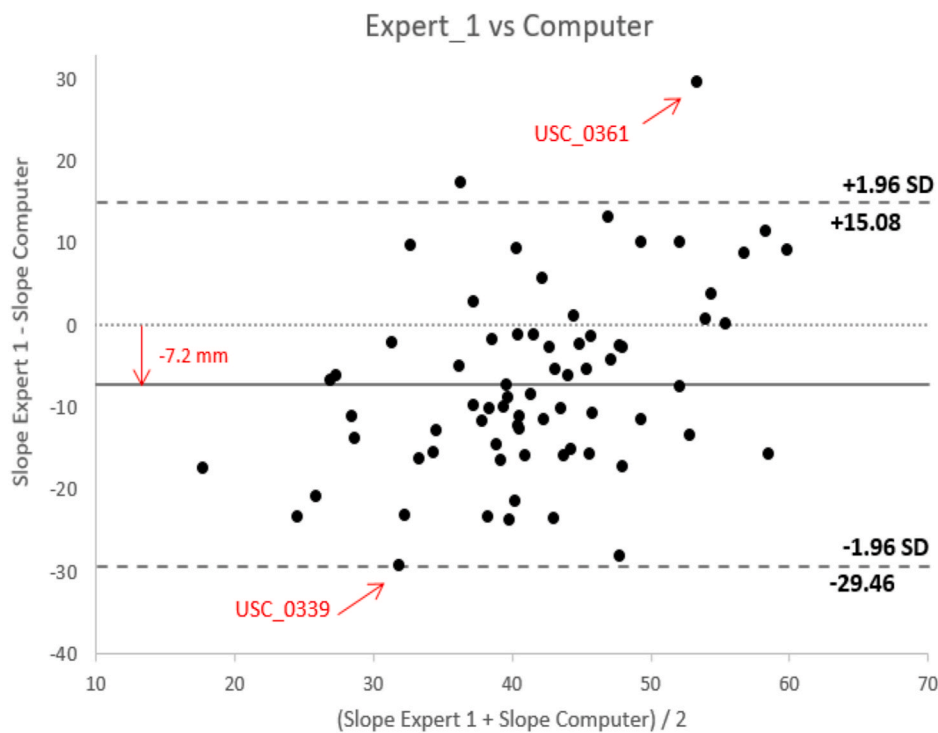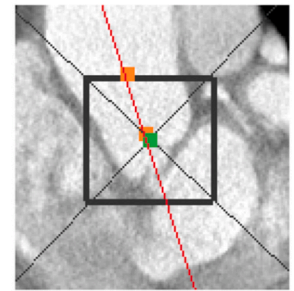even when datasets of limited size were employed for training.

Comparing the results achieved for both the classification tasks (see Table 1 and 2) and focusing on the results achieved for the different models, we observe that in general, values above 90% were obtained only for a few of the pre-trained model configurations and only in those

Fig. 9. Bland–Almant plot of the differences between the tilt angle of the SOV plane calculated from the marks inserted by *Expert*₁ and that calculated by the *Computer*, for cases in the test set. Two outliers are shown in (a) and (b).
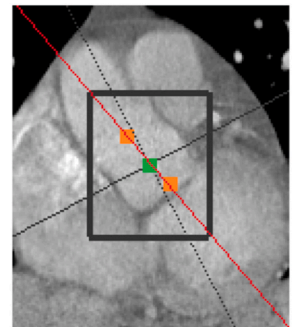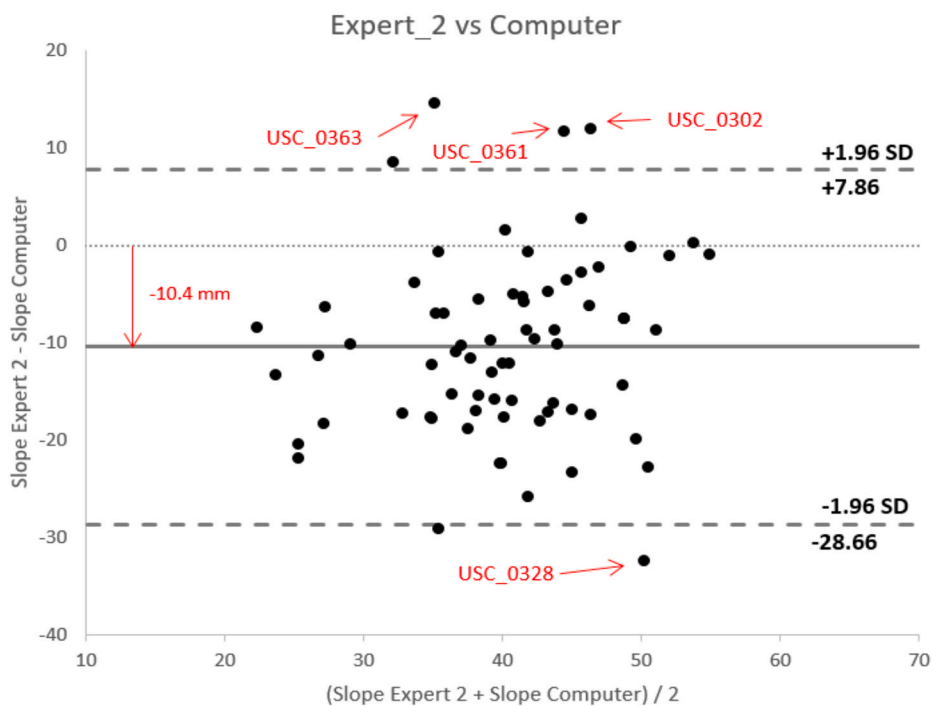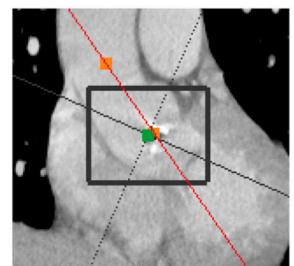

Fig. 10. Bland–Almant plot of the differences between the tilt angle of the SOV plane calculated from the marks inserted by *Expert*₂ and that calculated by the *Computer*, for cases in the test set. Two outliers are shown in (a) and (b).

The computation of the SOV plane slope is the second task provided by the proposed algorithm. A method based on PCA was used in this study. As demonstrated by the QQ-plot analysis, the differences between the angles computed by the experts and those computed by the algorithm follow a normal distribution. An analysis of the Bland–Almant plots (Figs. 9 and 10) reveal a certain bias between the slope proposed by the *Experts* and *Computer*. In both cases, this quantity is negative and close in terms of absolute value. This demonstrates the presence of a

systematic error that could be adjusted. However, there is a relatively high dispersion of the data relative to these differences. If we consider $1.96SD$ to include 95% of the cases, these values increase to $\pm 22.3°$ and $\pm 18.2°$ for $Expert_1$ vs. *Computer* and $Expert_2$ vs. *Computer*, respectively. Thus, these values appear to be exceptionally high.

The determination of the exact SOV plane slope is challenging even for human experts. The dispersion between the values calculated from the marks inserted by the experts shows relatively high differences (approximately $12.7°$). Furthermore, discrepancies between experts and computers are not always evident. Fig. 9a shows a case where the difference between the slope calculated by the algorithm and that calculated from the expert's marks is large. It appears evident that the computations made based on these marks are significantly better than those by the algorithm. This is a case in which the computation performed by the PCA-based algorithm fails owing to the large curvature of the aorta. However, there are several cases where the results achieved by the computer appear to be better than those obtained by humans (Figs. 9b and 10b). At this point, we should indicate that our method is aimed at obtaining a reasonable approximation of the location of the SOV plane rather than an accurate value of its slope. The slope of the aorta calculated by our method enables us to obtain the first approximation of the aortic root inclination. This can be useful for visualizing the SOV plane and obtaining a reference slice at the aortic root level. This reference slice can be used in subsequent processes such as segmentation or detection of other anatomical landmarks.

6. Conclusion

This paper presents an automated method for detecting the aortic root in CTA scans without the need for prior segmentation. The proposed method attempts to mimic the general procedure adopted by an expert. The developed algorithm uses models based on CNNs to perform this task. The models are trained on datasets with limited cases. The results obtained appear to be optimal and demonstrate how transfer learning enables the obtainment of reasonable results in this type of procedure.

Declaration of competing interest

The authors have no competing interest to declare.

Acknowledgment

References

[1] T.G. Flohr, S. Schaller, K. Stierstorfer, et al., Multi–detector row CT systems and image-reconstruction techniques, Radiology 235 (2005) 756–773.

[2] Quantitative Imaging Biomarkers Alliance, 2021 url: https://www.rsna.org/res earch/quantitative-imaging-biomarkers-alliance. (Accessed 27 April 2021).

[3] L.F. Hiratzka, G.L. Bakris, J.A. Beckman, et al., ACCF/AHA/AATS/ACR/ASA/SCA/SCAI/SIR/STS/SVM guidelines for the diagnosis and management of patients with thoracic aortic disease, J. Am. Coll. Cardiol. 55 (2010) (2010) e27–e129.

[4] Authors/Task Force members, R. Erbel, V. Aboyans, et al., ESC Guidelines on the diagnosis and treatment of aortic diseases: document covering acute and chronic aortic diseases of the thoracic and abdominal aorta of the adult. The Task Force for the Diagnosis and Treatment of Aortic Diseases of the European Society of Cardiology (ESC), Eur. Heart J. 35 (2014) (2014) 2873–2926.

[5] M. Elattar, E. Wiegerinck, F. van Kesteren, et al., Automatic aortic root landmark detection in CTA images for preprocedural planning of transcatheter aortic valve implantation, Int. J. Cardiovasc. Imag. 32 (2016) 501–511.

[6] Y. Xie, J. Padgett, A.M. Biancardi, A.P. Reeves, Automated aorta segmentation in low-dose chest CT images, Int J Comput Assist Radiol Surg 9 (2014) 211–219.

[7] G. Sedghi, L. Bons, M. Giordano, D. Bos, R. Budde, K. Kofoed, J. Pedersen, J. Roos-Hesselink, M. de Bruijne, Automated 3d segmentation and diameter measurement of the thoracic aorta on non-contrast enhanced ct, Eur. Radiol. 29 (2019) 4613–4623.

[8] A. Dasgupta, S. Mukhopadhyay, S.A. Mehre, P. Bhattacharyya, Morphological geodesic active contour based automatic aorta segmentation in thoracic ct images,

[9] in: B. Raman, S. Kumar, P. Roy, S. D (Eds.), Proceedings of International Conference on Computer Vision and Image Processing. Advances in Intelligent Systems and Computing, 2017, pp. 187–195.

[9] P.G. Tahoces, L. Alvarez, E. González, C. Cuenca, A. Trujillo, D. Santana-Cedrés, J. Esclarín, L. Gomez, L. Mazorra, M. Alemán-Flores, J.M. Carreira, Automatic estimation of the aortic lumen geometry by ellipse tracking, International Journal of Computer Assisted Radiology and Surgery 14 (2019) 345–355.

[10] Y. Yu, Y. Gao, J. Wei, F. Liao, Q. Xiao, J. Zhang, W. Yin, B. Lu, A three-dimensional deep convolutional neural network for automatic segmentation and diameter measurement of type B aortic dissection, Korean J. Radiol. 22 (2021) 168–178.

[11] A. Hager, H. Kaemmerer, U. Rapp-Bernhardt, S. Blücher, K. Rapp, T. Bernhardt, M. Galanski, J. Hess, Diameters of the thoracic aorta throughout life as measured with helical computed tomography, J. Thorac. Cardiovasc. Surg. 123 (2013) 1060–1066.

[12] L. Quint, P. Liu, A. Booher, K. Watcharotone, J. Myles, Proximal thoracic aortic diameter measurements at ct: repeatability and reproducibility according to measurement method, Int. J. Cardiovasc. Imag. 29 (2013) 479–488.

[13] A. Redheuil, W.-C. Yu, E. Mousseaux, H. Ahmed A, K. Nadjia, W. Colin O, B. David, L. Joao A C, Age related changes in aortic arch geometry, J. Am. Coll. Cardiol. 58 (2011) 1262–1270.

[14] J. Sugawara, K. Hayashi, T. Yokoi, H. Tanaka, Age-associated elongation of the ascending aorta in adults, Cardivascular Imaging 1 (2008) 739–748.

[15] J. Shemesh, Coronary artery calcification in clinical practice: what we have learned and why should it routinely be reported on chest ct, Ann. Transl. Med. 4 (2016).

[16] G. Brunner, D.R. Chittajallu, U. Kurkure, I.A. Kakadiaris, Toward the automatic detection of coronary artery calcification in non-contrast computed tomography data, Int. J. Cardiovasc. Imag. 26 (2010) 829–838.

[17] R. Prabhakar, P. Schoenhagen, The role of computed tomography in pre-procedural planning of cardiovascular surgery and intervention, Insights Imaging 4 (2013) 671–689.

[18] A.V. Linden, J. Kempfert, J. Blumenstein, H. Möllmann, W.-K. Kim, S. Alkaya, C. Hamm, T. Walther, Manual versus automatic detection of aortic annulus plane in a computed tomography scan for transcatheter aortic valve implantation screening, Eur. J. Cardio. Thorac. Surg. 46 (2014) 207–212.

[19] M. Hardy, H. Harvey, Artificial intelligence in diagnostic imaging: impact on the radiography profession, Br. J. Radiol. 93 (2020) 20190840.

[20] S. Queirós, C. Dubois, P. Morais, T. Adriaenssens, J. Fonseca, J. Vilaça, J. D'hooge, Automatic 3d aortic annulus sizing by computed tomography in the planning of transcatheter aortic valve implantation, Journal of Cardiovascular Computed Tomography 11 (2017) 25–32.

[21] R. Koos, E. Altiok, A. Mahnken, M. Neizel, G. Dohmen, N. Marx, H. Kühl, R. Hoffmann, Evaluation of aortic root for definition of prosthesis size by magnetic resonance imaging and cardiac computed tomography: implications for transcatheter aortic valve implantation, J. Thorac. Cardiovasc. Surg. 158 (2012) 353–358.

[22] T.-L.C. Lu, E. Rizzo, P.M. Marques-Vidal, L.K. von Segesser, J. Dehmeshki, S. D. Qanadli, Variability of ascending aorta diameter measurements as assessed with electrocardiography-gated multidetector computerized tomography and computer assisted diagnosis software, Interact. Cardiovasc. Thorac. Surg. 10 (2010) 217–221.

[23] F. Lalys, S. Esneault, M. Castro, L. Royer, P. Haigron, V. Auffret, J. Tomasi, Automatic aortic root segmentation and anatomical landmarks detection for tavi procedure planning, Minim Invasive Ther. Allied Technol. 28 (2019) 157–164.

[24] S. Kurugol, C.E. Come, A.A. Diaz, et al., Automated quantitative 3D analysis of aorta size, morphology, and mural calcification distributions, Med. Phys. 42 (2015) 5467–5478.

[25] P.G. Tahoces, D. Santana-Cedrés, L. Alvarez, M. Alemán-Flores, A. Trujillo, C. Cuenca, J.M. Carreira, Automatic detection of anatomical landmarks of the aorta in CTA images, Med. Biol. Eng. Comput. 58 (2020) 903–919.

[26] S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, IEEE Trans. Pattern Anal. Mach. Intell. 39 (2017) 1137–1149.

[27] B. McComb, R. Munden, F. Duan, A. Jain, C. Tuite, C. C, Normative reference values of thoracic aortic diameter in american college of radiology imaging network (acrin 6654) arm of national lung screening trial, Clin. Imag. 40 (2016) 756–773.

[28] P. Astudillo, P. Mortier, J. Bosmans, O. De Backer, P. de Jaegere, F. Iannaccone, M. De Beule, J. Dambre, Automatic detection of the aortic annular plane and coronary ostia from multidetector computed tomography, J. Intervent. Cardiol. 2020 (2020).

[29] W. Al, H. Jung, I. Yun, Y. Jang, H. Park, C. Hj, Automatic aortic valve landmark localization in coronary CT angiography using colonial walk, PloS One 13 (2018) 1–23.

Pablo G. Tahoces has received an M.Sc. in Applied Physics in 1989 and a doctorate in Physics in 1992, both from the University of Santiago de Compostela (Spain). During the 90's he worked at the Kurt Rossmann Laboratories of the University of Chicago (USA) as Visiting Research Associate for different periods. Since 2007 he is a professor at the University of Santiago de Compostela. His main areas of research interest are medical imaging and computer vision.

Rafael Varela is Medical Doctor since 2007 from the University of Navarra (Spain). Since 2012 he has worked as a vascular and interventional radiologist at the University Hospital of Santiago de Compostela (Spain). He received a Ph.D degree in 2020 from the University of Santiago de Compostela. His main research interests are focused on peripheral vascular imaging analysis, vascular intervention, and tumor detection.

Jose M. Carreira is Medical Doctor since 1983 and received a Ph.D. degree in 1991 from the University of Santiago de Compostela (Spain). Between 1993 and 1998, he worked as a vascular and interventional radiologist with Prof. Manuel Maynar at the University Hospital of Las Palmas de Gran Canaria (Spain). Since 1998 he is a professor at the University of Santiago de Compostela, and his main research interests are focused on peripheral vascular imaging analysis and vascular intervention.