Computers in Biology and Medicine 163 (2023) 107147

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/compbiomed

A CT-based deep learning system for automatic assessment of aortic root morphology for TAVI planning

Simone Saitta [a], Francesco Sturla [a,b,*], Riccardo Gorla [c], Omar A. Oliva [c], Emiliano Votta [a,b], Francesco Bedogni [c], Alberto Redaelli [a]

[a] Department of Electronics, Information and Bioengineering, Politecnico di Milano, Milan, Italy
[b] 3D and Computer Simulation Laboratory, IRCCS Policlinico San Donato, San Donato Milanese, Italy
[c] Department of Clinical and Interventional Cardiology, IRCCS Policlinico San Donato, San Donato Milanese, Italy

ARTICLE INFO

Keywords:
TAVI
Segmentation
Deep neural networks
Automatic planning
Aortic root

ABSTRACT

Accurate planning of transcatheter aortic valve implantation (TAVI) is important to minimize complications, and it requires anatomic evaluation of the aortic root (AR), commonly performed through 3D computed tomography (CT) image analysis. Currently, there is no standard automated solution for this process.

Two convolutional neural networks with 3D U-Net architectures (model 1 and model 2) were trained on 310 CT scans for AR analysis. Model 1 performs AR segmentation and model 2 identifies the aortic annulus and sinotubular junction (STJ) contours. After training, the two models were integrated into a fully automated pipeline for geometric analysis of the AR. Results were validated against manual measurements of 178 TAVI candidates.

The trained CNNs segmented the AR, annulus, and STJ effectively, resulting in mean Dice scores of 0.93 for the AR, and mean surface distances of 0.73 mm and 0.99 mm for the annulus and STJ, respectively. Automatic measurements were in good agreement with manual annotations, yielding annulus diameters that differed by 0.52 [-2.96, 4.00] mm (bias and 95% limits of agreement for manual minus algorithm). Evaluating the area-derived diameter, bias, and limits of agreement were 0.07 [-0.25, 0.39] mm. STJ and sinuses diameters computed by the automatic method yielded differences of 0.16 [-2.03, 2.34] and 0.1 [-2.93, 3.13] mm, respectively.

The proposed tool is a fully automatic solution to quantify morphological biomarkers for pre-TAVI planning. The method was validated against manual annotation from clinical experts and showed to be quick and effective in assessing AR anatomy, with potential for time and cost savings.

1. Introduction

Transcatheter aortic valve implantation (TAVI) has emerged as an alternative to traditional open-heart surgery to treat severe aortic valve stenosis, proving to be effective at reducing morbidity and mortality in high-risk patients [1,2]. Despite its benefits, TAVI still carries the risk of post-operative complications, including paravalvular leakage, device migration, annulus rupture, and conductive disturbances [3–5]. Meticulous preprocedural planning based on the analysis of clinical imaging is thus crucial for minimizing the risk of complications and should involve an accurate anatomic assessment of the aortic root (AR), i.e., the complex bulb-like anatomical structure that includes the diseased native valve and will host the prosthetic device. Such quantification is essential for selecting the optimal prosthetic device size [6], and it is focused mainly on the aortic annulus, whose diameters [7], angulation [8], and perimeter [9] are measured. Furthermore, a comprehensive analysis of the whole AR can provide anatomical measurements of the sinotubular junction (STJ) and sinuses of Valsalva [8,10,11]. To perform the mentioned anatomical quantifications, three-dimensional (3D) computed tomography (CT) angiography is the preferred imaging modality [12]. A comprehensive CT-based TAVI planning involves three main operations: segmentation of the anatomy, landmark detection, and measurement extraction [6,10,13–15]. However, there is currently no standardized fully automated solution, and taking the necessary measurements can be a time-consuming process that often involves several manual operations that introduce operator dependency and may limit

* Corresponding author. Department of Electronics, Information and Bioengineering, Politecnico di Milano, Via Camillo Golgi 39, 20133, Milan, Italy.
E-mail addresses: francesco.sturla@grupposandonato.it, francesco.sturla@polimi.it (F. Sturla).

https://doi.org/10.1016/j.compbiomed.2023.107147
Received 7 February 2023; Received in revised form 18 May 2023; Accepted 6 June 2023
Available online 7 June 2023

reproducibility. Automatic and semi-automatic systems have already been proposed to identify and quantitatively assess AR features from CT images. Lalys et al. [13] exploited semi-automatic segmentation tools relying on atlas-based methods to segment the AR, localize a wide range of anatomical landmarks (e.g., leaflet and coronary ostium positions) and obtain accurate quantification of annulus diameter. More recent approaches have taken advantage of convolutional neural networks (CNNs) to fully automate the detection of AR landmarks from 3D CT [6, 14]. To ensure accuracy when using deep neural networks, which are often opaque in nature, it is important to extensively validate their results against manual landmark tracing by experts. To date, the largest patient validation set has been reported by Astudillo et al. [6], who proved the feasibility of building fast and accurate CNN-based systems for detection of the three aortic cusp nadirs and coronary ostia, training their model on 444 CT scans and validating their landmark detection accuracy on 100 patients. Nonetheless, the proposed approach could only be applied to contrast-enhanced CT images and did not include automatic measurement extraction. Hence, full automation of the entire process, including segmentation, landmark detection and extraction of aortic features relevant for TAVI, is still lacking. Given the increasing adoption of TAVI also for intermediate and low-risk patients [1], a reliable, fast and efficient method for assessing AR anatomy and determining the appropriate device size could have an increasingly broad impact by making TAVI planning faster, reliable and fully repeatable.

The main contribution of the present work was the development of an automatic workflow that combines deep learning techniques and tools from differential geometry to segment the AR from 3D CT data, extract AR-specific anatomical landmarks and compute clinically relevant measurements for TAVI planning. The effectiveness of the proposed method was extensively evaluated on a group of 178 patients.

2. Methods

2.1. Data collection and manual annotation

CT scans of 512 subjects acquired between 2010 and 2022 were retrospectively collected. All images were acquired using a 64-slice multidetector system with retrospective ECG-gating (SOMATOM Definition, Siemens, Germany). Pixel spacing ranged from 0.26×0.26 mm^2 to 0.87×0.87 mm^2, while slice thickness ranged between 0.25 and 1 mm. We excluded twenty-four acquisitions because of reconstruction artifacts, presence of metal devices, or slice thickness greater than 1 mm. Accordingly, the dataset was split in two subsets: dataset A ($N = 310$), for which no AR measurement was available, and dataset B ($N = 178$), for which an expert operator manually took measurements of the aortic annulus, STJ and Valsalva sinuses plane using the commercial software 3mensio Structural Heart (v8.2, Pie Medical Imaging, Maastricht, Netherlands) (Fig. 1). After manual detection of the aortic annulus and STJ planes, 3mensio allowed to compute their maximum and minimum diameters, area, and perimeter. All patients in dataset B were TAVI candidates with severe aortic stenosis.

For all CT scans in dataset A, an initial segmentation of the ascending aorta was obtained using a neural network previously trained by our group [16]. Segmentations that were erroneously inferred (e.g. incomplete filling) by the pre-trained model were edited by an experienced operator using a semi-automatic region growing algorithm [17]. All the resulting aortic segmentations included the left ventricle outflow tract (LVOT), the whole AR and the proximal portion of the ascending aorta according to the scan-specific field-of-view. Both aortic annulus and STJ were manually segmented for all CT scans in dataset A. To expedite the annotation process, a graphic-user-interface (GUI) was appositely developed using VTK [18]. Through the designed GUI, the user could easily position a plane in the 3D image space and, exploiting the previously obtained segmentation of the aorta, the annulus or STJ labels were automatically assigned to the image points within a Euclidean distance of 3 mm to the chosen plane and inside the aorta.

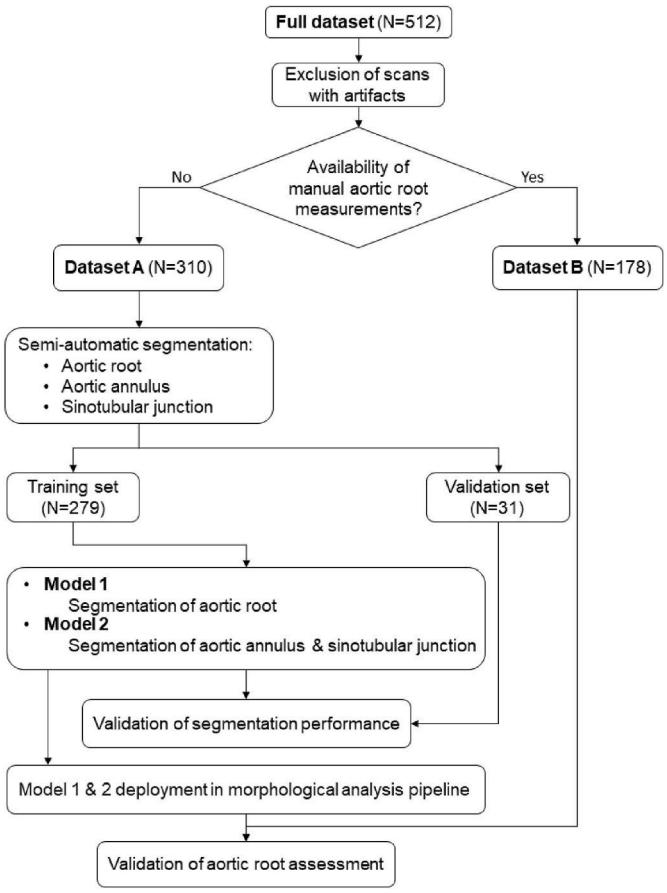

Fig. 1. Schematic representation of the adopted workflow. Dataset A included manual segmentations of the aortic root (AR), annulus and sinotubular junction (STJ); it was used to train and validate model 1 and model 2 for automatic segmentation. Dataset B included manual annotations of AR measurements, used to validate the developed morphological analysis pipeline.

All image data used in the present study were collected at IRCCS Policlinico San Donato (San Donato Milanese, Italy). The study was approved by the local ethics committee and informed consent was waived because of the retrospective nature of the study and the anonymized data analysis.

2.1.1. Analysis of inter-operator variability

To obtain an unbiased assessment of the inter-operator variability of aortic segmentation, a randomly selected subset of 20 CT scans belonging to dataset B was segmented by three experienced, completely independent operators (Op1, Op2 and Op3). Op1 and Op3 used the open-source segmentation software ITK-snap [17], exploiting a combination of region growing and paintbrush tools to annotate the images, while Op2 used the CE-marked commercial software Mimics Medical v25.0 (Materialise, Leuven).

2.2. Image pre-processing: aortic root region detection

For each 3D CT scan in our datasets, a region of interest (ROI) encompassing the AR was cropped out (Fig. 2). The detection of the ROI to crop was fully automatic and is achieved through a 3D template-matching approach [19]. An image X with its ground truth AR segmentation S was previously selected from dataset A. When processing a new case (target image, Y), the image X is roughly aligned to the target image using a rigid transformation ($\mathbf{R}$) followed by an affine transformation ($\mathbf{A}$) able to account for scaling and shear deformations. Defining the generic final transformation:

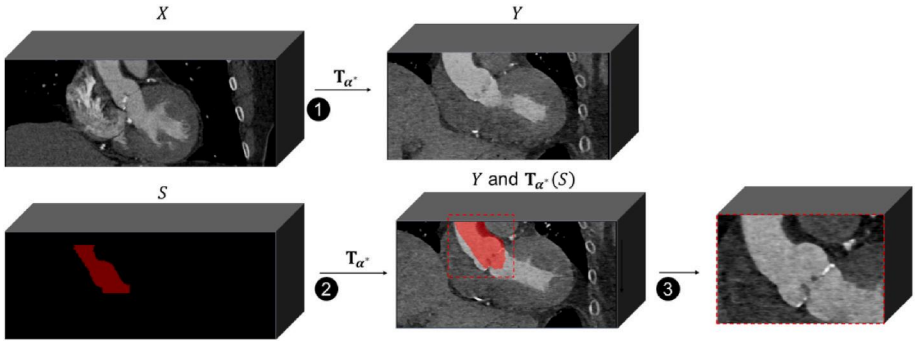

Fig. 2. Template-matching approach for automatic identification of the region of interest (ROI) from 3D computed tomography (CT). A template image (X) for which aortic segmentation (S) is known, is affinely registered to the processed image (Y). The transformation mapping X to Y ($\mathbf{T}_{\alpha^*}$) is applied to S and the bounding box of $\mathbf{T}_{\alpha^*}(S)$ is used to crop the ROI from Y.

$$\mathbf{T}_{\alpha^*}(X) = (\mathbf{A} \circ \mathbf{R})(X), \tag{1}$$

parametrized by $\boldsymbol{\alpha}^*$, the optimal mapping is found as:

$$\boldsymbol{\alpha}^* = \underset{\boldsymbol{\alpha}}{\operatorname{argmax}} I(\mathbf{T}_{\boldsymbol{\alpha}}(X), Y), \tag{2}$$

where $I(\mathbf{T}_{\boldsymbol{\alpha}}(X), Y)$ is the mutual information between the two images, which can be considered a nonlinear generalization of cross-correlation [20]. $\boldsymbol{\alpha}^*$ is found using a gradient-based optimization algorithm. Once found, $\mathbf{T}_{\alpha^*}$ is used to map the known ground truth AR segmentation into the target image domain, obtaining $\mathbf{T}_{\alpha^*}(S)$. The bounding box of $\mathbf{T}_{\alpha^*}(S)$ is dilated outward in each direction by 20 pixels and the resulting ROI is cropped out from the original image.

2.3. Automatic segmentation

Dataset A was divided into training and test sets: 279 (90%) scans with their corresponding ground truth segmentations of the aorta, annulus and STJ were randomly selected and used to train two different neural networks, respectively dedicated to the segmentation of the aorta (model 1) and to the multi-class segmentation of the aortic annulus and STJ (model 2). The remaining 31 (10%) scans were used for validation (Fig. 1). Both models were based on the 3D U-Net architecture proposed in Ref. [21], with encoding and decoding branches of 5 resolution levels each, defined using residual units as introduced by Ref. [22]. The number of filters used for the encoding branch was 32, 64, 128, 256, 512. Each encoding block consisted of a convolution layer with kernel size of 3x3x3 and stride of 2, followed by parametric rectified linear units (PReLUs) [23] and instance normalization. The encode and decode paths were connected with skip connections via concatenation. Model 1 and model 2 only differed by the number of output channels; given the different tasks, model 1 outputs a tensor with one channel, whereas model 2 returns a tensor with three channels (background, annulus and STJ). The number of trainable parameters for each of the two models was approximately 20 M.

For comparison, we considered two additional automatic approaches for semantic segmentation of the aorta, aortic annulus and STJ.

2.3.1. Atlas-based segmentation

As a baseline, an atlas-based segmentation method was implemented. Similar to the ROI detection solution previously described, this approach exploits the registration of a reference image (with its associated labels of the aorta annulus and STJ) with the image to segment. With respect to the approach described in the previous section, a free form deformation $\mathbf{D}$, based on cubic b-splines is added to the final transformation $\mathbf{F}_{\beta^*}$ parametrized by $\boldsymbol{\beta}^*$ such that:

$$\mathbf{F}_{\beta^*}(X) = (\mathbf{D} \circ \mathbf{A} \circ \mathbf{R})(X) \tag{3}$$

As in the formalization presented in equation (2), the $\boldsymbol{\beta}^*$ that maximizes $I(\mathbf{F}_{\beta^*}(X), Y)$ was found using a gradient-based optimization algorithm. More details of the algorithm can be found in Ref. [24]. Once found, $\mathbf{F}_{\beta^*}$ was used to map the known labels of the aorta, annulus and STJ into the target image domain.

2.3.2. Segmentation via 3D U-Net transformers

Similarly to what was described above, two additional models with a 3D U-Net transformer architecture were implemented; model 3 for segmentation of the aorta and model 4 for segmentation of the annulus and STJ. Models 3 and 4 were based on the transformer architecture proposed in Ref. [25]. During a forward pass of a U-Net transformer, the 3D tensor is divided into a series of non-overlapping patches of size 16x16x16. These patches are projected into an embedding space using a linear layer with 128 hidden channels. A position embedding is added to the resulting sequence, which is then used as input to a transformer model with 4 attention heads. The transformer model's various encoded representations are extracted from different layers, and they are combined with a decoder using skip connections. Finally, this process is used to predict the final segmentation.

During training of all neural networks, an online data augmentation routine including random Gaussian noise, cropping, mirroring and rotation was implemented in the MONAI framework [26]. A Dice loss was used to train model 1 and 3, while a combination of Dice and Focal loss [27] with equal weights was used to train model 2 and 4. Training was carried out on an NVIDIA A100 GPU over 1000 epochs, using an Adam optimizer with learning rate of 0.0001.

2.4. Pipeline implementation

After training, models 1 and 2 were embedded in a fully automated pipeline for AR analysis and TAVI pre-procedural planning. The different sequential steps of the implemented pipeline are described below and illustrated in Fig. 3.

2.4.1. Aortic surface processing

After ROI cropping, the CT scan is processed by model 1, generating a binary mask of the ascending aorta including the LVOT. A marching cubes algorithm extracts the corresponding contour as a triangulated surface; surface smoothing is applied with a windowed sinc function interpolation kernel with passband of 0.01, and adaptive remeshing is performed following the approach described in Ref. [28]. For each processed surface, the Laplace-Beltrami operator (LBO) [29] is computed. LBO eigenvectors form a set of bases for the definition of continuous functions on the surface manifold. For tubular structures with a dominant longitudinal direction, the second LBO eigenvector (E1) represents a scalar field approximating the curvilinear abscissa of the surface centerline; E1 isocontours on the aortic surface are shown in Fig. 3.

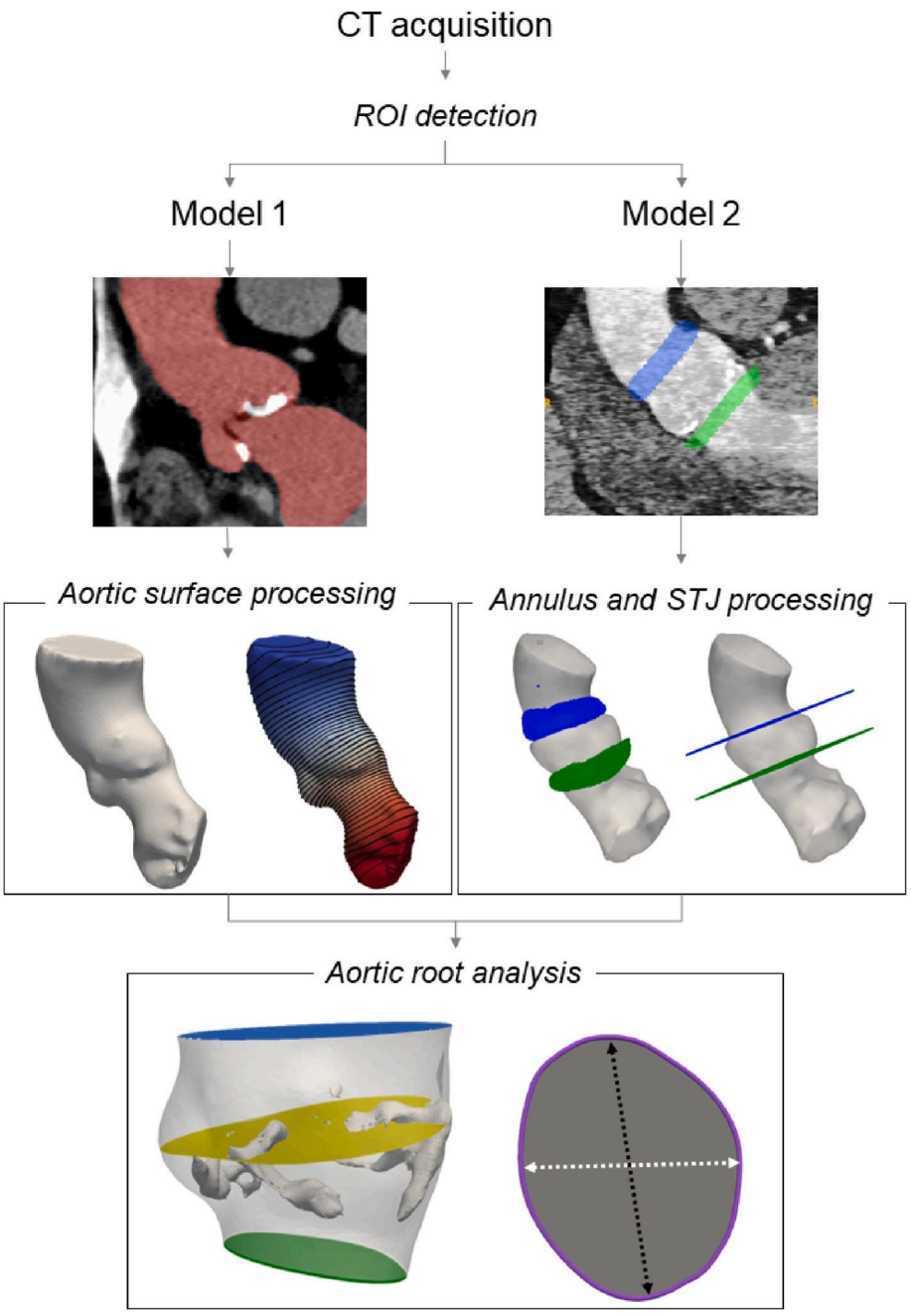

Fig. 3. Schematic representation of the implemented automatic pipeline. The ROI is detected from the input CT scan. Model 1 (left side) infers the segmentation of the AR with the left ventricle outflow tract (LVOT) (red label). The segmentation is the input of the aortic surface processing step, which computes the second eigenvector of the Laplace-Beltrami operator (LBO) and its contours. Model 2 (right side) infers the segmentation of the aortic annulus and STJ (in green and blue, respectively) (right side). This segmentation is the input to the annulus and STJ processing step, which refines it and identifies the annular and STJ planes. The results of the two parallel analyses are funneled in the AR analysis step (bottom): the AR is isolated by cropping the initial segmentation of AR and LVOT by means of the annulus and STJ planes, and the Sinuses of Valsalva are detected. Anatomical measurements are computed together with the volume of calcium deposits (shown in white, bottom panel).

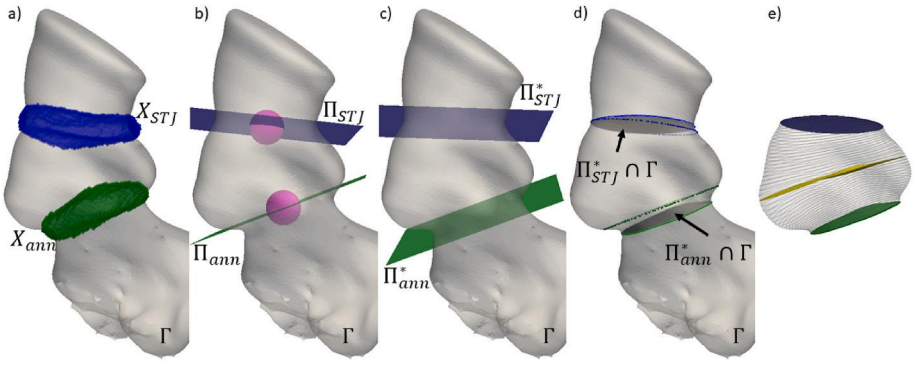

Fig. 4. Steps implemented for aortic annulus and STJ plane detection refinement. a) the point sets yielded by model 2 for the annulus (green) and STJ (blue), visualized with the aortic segmentation produced by model 1 (transparent gray). b) the best fitting planes for the two structures (Π_{ann} and Π_{STJ}), together with the region inside which their center of mass is allowed to move during the refinement procedure (magenta spheres). c) the resulting planes (Π^*_{ann} and Π^*_{STJ}) after the optimization procedure. d) visualization of the refined planes in dark gray and the initialization planes obtained by model 2. e) the isolated AR with the contours of the second eigenvalue of the Laplace-Beltrami operator (LBO) (in black) and the plane identifying the sinuses of Valsalva (yellow).

2.4.2. Annulus and STJ processing

The cropped CT scan is also processed by model 2, which infers the masks and the corresponding set of points X_{ann} and X_{STJ} for the annulus and STJ, respectively (Fig. 4). Using a random sample consensus (RANSAC) iterative algorithm with inlier threshold of 1.5 mm, two best-fitting planes are identified, namely Π_{ann} and Π_{STJ} from X_{ann} and X_{STJ}, respectively. Specifically, a generic plane Π is represented by a bounded rectangular mesh grid identified by its normal vector $\boldsymbol{n}$ and center of mass $\boldsymbol{c}$: $\Pi = \Pi(\boldsymbol{n}, \boldsymbol{c})$. The normal vector and center of mass of both Π_{ann} and Π_{STJ}, are refined (rotated and shifted) following a constrained optimization procedure, initialized using $\boldsymbol{n}$ and $\boldsymbol{c}$ computed by RANSAC. The optimal plane normal vector $\boldsymbol{n}^*$ and center of mass $\boldsymbol{c}+\boldsymbol{\delta}$ are defined as the ones that minimize the area of intersection with the aortic surface:

$$\Pi^* = \min_{\boldsymbol{n}^*, \boldsymbol{\delta}} area(\Pi(\boldsymbol{n}^*, \boldsymbol{c}+\boldsymbol{\delta}) \cap \Gamma) \quad (4)$$

subject to $\|\boldsymbol{\delta}\| < K$,

where Γ represents the AR and LVOT surface and $K = 5\,mm$. The described procedure is applied to both Π_{ann} and Π_{STJ} and the yielded planes are used to clip and isolate the AR (Fig. 4.d). Within the extracted AR surface, the E1 isocontour with maximum area is used to identify the plane of the sinuses of Valsalva (Fig. 4e).

2.4.3. Aortic root analysis

For the annulus, the STJ and the plane of Valsalva sinuses, the following metrics are automatically computed: i) area; ii) perimeter; iii) the maximum (D_{max}), iv) minimum (D_{min}) and v) mean ($D_{mean} = (D_{max} + D_{min}) / 2$) diameters. Maximum and minimum diameters are defined as the lengths of the largest and shortest segments connecting two opposite points of the perimeter while passing through the center of mass. In addition for the aortic annulus, the annulus angle, i.e., the angle formed by the plane normal to the foot-head image axis [8], is extracted. The AR calcium score is automatically quantified as the volume enclosing pixels with Hounsfield units (HUs) greater than 800 by a simple thresholding of the corresponding segmented AR [30].

2.5. Statistical analysis

Measurements based on expert manual annotations were used as reference values in assessing the proposed method's accuracy. Accuracy of areas, perimeters and diameters was evaluated using Bland-Altman analysis. Analyses were performed using python *scipy* 1.5.1 statistics library.

3. Results

3.1. Evaluation of segmentation performance

The generalization performance of the segmentation models was evaluated on the validation set (n = 31), for which manual ground truth labels of the aorta, annulus and STJ were available (dataset A). For all subjects in the validation set, both models 1 and 2 proved able to trace accurately the regions encompassing the AR, annulus and STJ. Segmentation performance was quantitatively assessed, in terms of Dice coefficient and MSD with respect to manual ground truth segmentations; for both indices, mean value and [min, max] range were computed. Model 1 and 2 outputs were compared against models 3 and 4, respectively, and on the atlas-based method on the same validation set. Results are reported in Table 1. For the AR, the U-Net architecture (model 1) gave 0.02 higher Dice scores and 0.08 mm lower MSDs than the U-Net transformer (model 3). For the aortic annulus and STJ, model 1 slightly outperformed model 3, yielding 0.03 higher Dice scores and 0.1 mm lower MSDs. In general, the atlas-based method gave the worst segmentation performance, reaching the lowest Dice score of 0.16 for the STJ.

The test cases corresponding to the best and worst performance of models 1 and 2 are shown in Fig. 5.

Table 1
Comparison between automatic segmentation approaches and ground truth manual segmentations for the AR, annulus and sino-tubular junction (STJ). Mean value and [min, max] ranges of Dice score and mean surface distance (MSD) are reported.

		Dice	MSD [mm]
AR	Model 1 (U-Net)	0.93 [0.80, 0.97]	0.39 [0.0, 4.7]
	Model 3 (UNETR)	0.91 [0.77, 0.97]	0.31 [0.0, 6.4]
	Atlas-based	0.53 [0.0, 0.91]	3.6 [0.1, 18]
Annulus	Model 2 (U-Net)	0.57 [0.34, 0.87]	0.73 [0.0, 3.9]
	Model 4 (UNETR)	0.54 [0.18, 0.71]	0.63 [0.0, 4.3]
	Atlas-based	0.25 [0.0, 0.78]	12 [0.2, 76]
STJ	Model 2 (U-Net)	0.66 [0.29, 0.90]	0.99 [0.0, 1.5]
	Model 4 (UNETR)	0.54 [0.18, 0.73]	0.85 [0.0, 4.3]
	Atlas-based	0.16 [0.0, 0.66]	14 [0.2, 91]

3.1.1. Inter-operator variability

For the 20 randomly selected subjects, the aortic segmentations obtained by the three independent operators were compared pairwise, in terms of Dice scores, MSDs and physical volume differences (ΔV). For a generic comparison OpX vs. OpY, ΔV is computed by subtracting the label mask created by OpY from the one of OpX. Comparison results are reported in Table 2. In general, Op1 and Op2 produced almost identical segmentations (mean Dice score of 0.96), whereas Op1 and Op3 reported the lowest agreement (0.89 mean Dice score).

3.2. Comparison with manual measurements

Through Bland-Altman plots (Fig. 6) we analyzed the agreement between the automatic pipeline output and manual measurements. Herein, we report comparisons as biases and 95% limits of agreement, i.e., *bias [lower limit, upper limit]*, where differences were computed as manual – algorithm. In general, automatic and manual anatomical measurements were in good agreement. A tendency of the automatic system to underestimate annulus diameters vs. manual measurements was observed, as evidenced by positive biases (solid horizontal lines in Fig. 6). For D_{max}, measurement differences were 0.51 [−2.79, 3.81] mm. A similar trend was found for D_{min}, resulting in 0.89 [−2.8, 4.62] mm. Significantly smaller discrepancies between the two measurement techniques were obtained for the annulus area. Evaluating the area-derived diameter, bias and limits of agreement were 0.07 [−0.24, 0.38] mm. For the annulus angle, an average difference $<3^\circ$ was found between measurements, while limits of agreement were [−17°, 11°]. D_{mean} computed at the STJ and Sinuses by the automatic method slightly underestimated manual measurements, yielding differences of 0.05 [−1.98, 2.07] and 0.17 [−2.63, 2.97] mm, respectively. As compared to diameter measurements, perimeter measurements showed larger differences: −1.8 [−8.06, 11.74] mm for the annulus perimeter, and 1.09 [−6.18, 8.37] mm for the STJ perimeter.

3.2.1. Critical cases

For a few cases, a large discrepancy was observed between the automatic and expert measurements. The largest differences in annulus diameters were observed for patient 057 and were equal to $\Delta D_{max} = -4.9\,mm$ and $\Delta D_{min} = -5.7\,mm$, where Δ is defined as (expert – algorithm). However, a detailed analysis of case 057 showed that AR segmentation was precise; thus, discrepancies are not due to errors in automated segmentation (Fig. 7, top row). Concerning STJ measurements, the largest discrepancies were obtained for patient 029 ($\Delta D_{mean} = 4.4\,mm$); these were likely due to the presence of calcifications around the aortic bulb (Fig. 7, mid row), possibly introducing larger uncertainty in STJ measurements. The largest difference in D_{mean} for the sinuses of Valsalva was obtained for patient 140 ($\Delta D_{mean} = -$

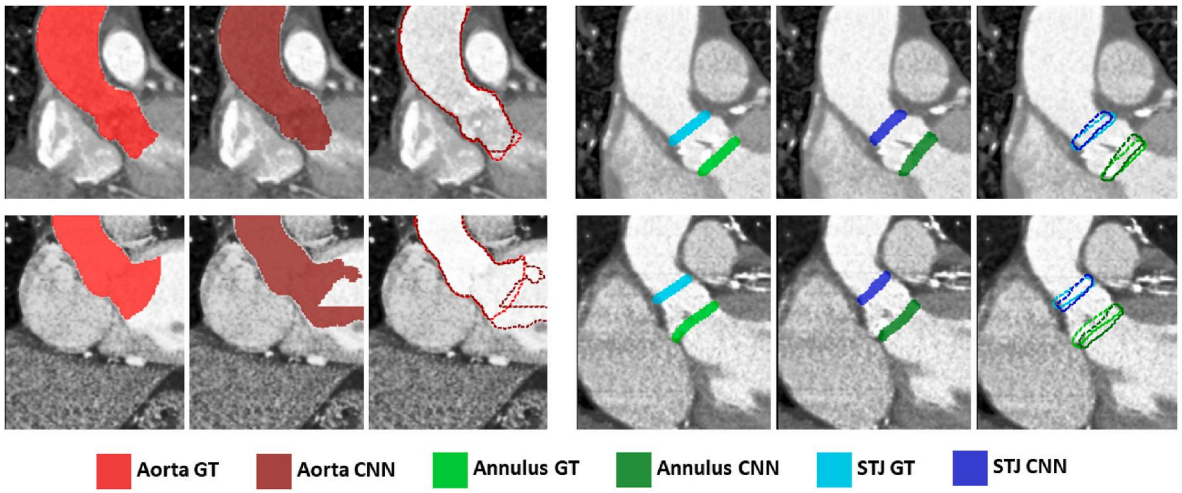

Fig. 5. Comparison of manual (ground truth, GT) segmentation in bright red (aorta), bright green (annulus) and light blue (STJ) vs. automatic segmentation obtained by the trained neural networks (CNN) in dark red (aorta), dark green (annulus) and dark blue (STJ) on four test cases: the best (top row) and the worst (bottom row) for the models 1 and 2. Coronal slices with color fill and with dashed contours are shown.

Table 2
Inter-operator variability analysis results. Dice scores, mean surface distance (MSD) and physical volume differences (ΔV) are computed for each pairwise comparison. Median [minimum, maximum] values are reported.

	Op1 vs. Op2	Op1 vs. Op3	Op2 vs. Op3
Dice	0.96 [0.93, 0.99]	0.89 [0.78, 0.95]	0.91 [0.78, 0.96]
MSD [mm]	0.11 [-0.58, 0.67]	1.0 [-0.02, 5.3]	1.2 [-0.21, 6.1]
ΔV [cm^3]	1.6 [-10, 4.9]	7.7 [-4.1, 37]	7.7 [-7.7, 38]

4.3 *mm*). In this case, the high curvature of the aortic bulb surface could cause larger discrepancies between the two methods (Fig. 7, bottom row).

4. Discussion

In this work, we presented a fully automatic pipeline for AR morphological analysis and TAVI pre-procedural planning support from 3D CT. Our system provides 3D segmentations of the anatomical structures of interest without requiring any human intervention and extracts quantitative morphological parameters of the AR with good accuracy as compared to manual annotations performed by a clinical expert.

To obtain a quantitative morphometric assessment of the AR, our approach relies on two CNNs trained to perform automatic 3D segmentation of the ascending aorta including the LVOT (model 1), and of the aortic annulus and STJ (model 2). The CNN trained to segment the aorta (model 1) proved able to delineate accurately the anatomical region of interest in all 31 validation cases from dataset A and all 178 test cases in dataset B. For the validation cases (dataset A), the average Dice score and MSDs were comparable to state-of-the-art approaches focusing on similar anatomical districts, with values of 0.93 [0.80, 0.97] and 0.39 [0.0, 4.7] mm, respectively. In a previous paper by our group dealing with automatic segmentation of the thoracic aorta from CT, a mean Dice score of 0.954 was reported [16]. Elattar et al. [31] adopted an AR segmentation approach based on normalized cuts and achieved Dice scores of 0.95 [0.85, 0.98] and MSDs of 0.74 ± 0.39 mm. In our study, model 1 was trained on ground truth segmentations encompassing

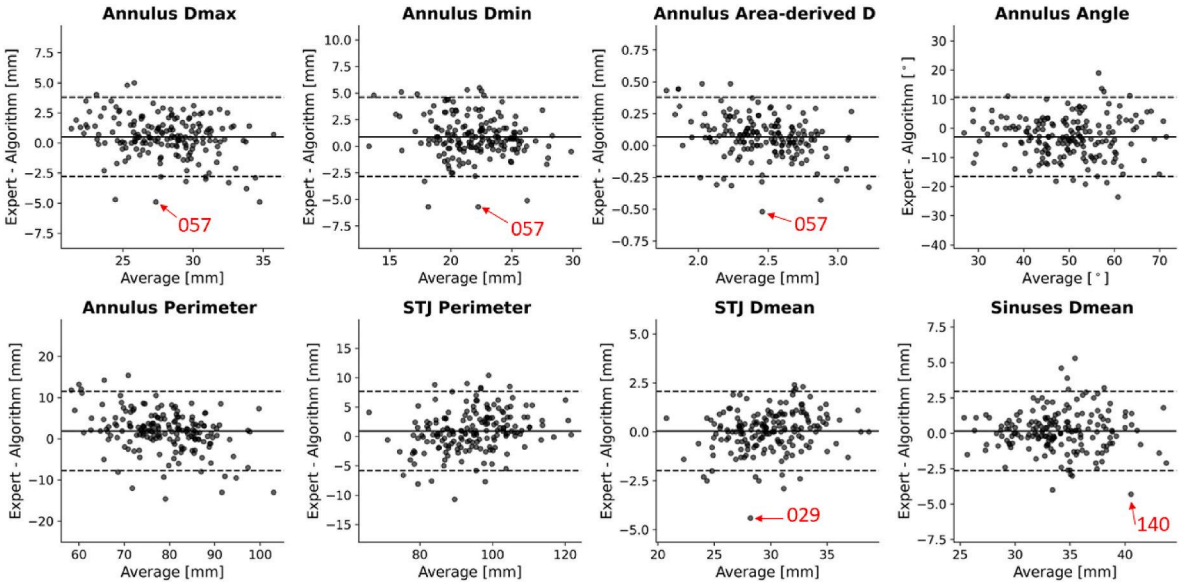

Fig. 6. Bland–Altman plots of the proposed algorithm versus expert manual measurements. Mean differences are shown as continuous horizontal lines, while 95% limits of agreement are shown as dashed horizontal lines. Critical cases are highlighted with a red arrow and identification number.

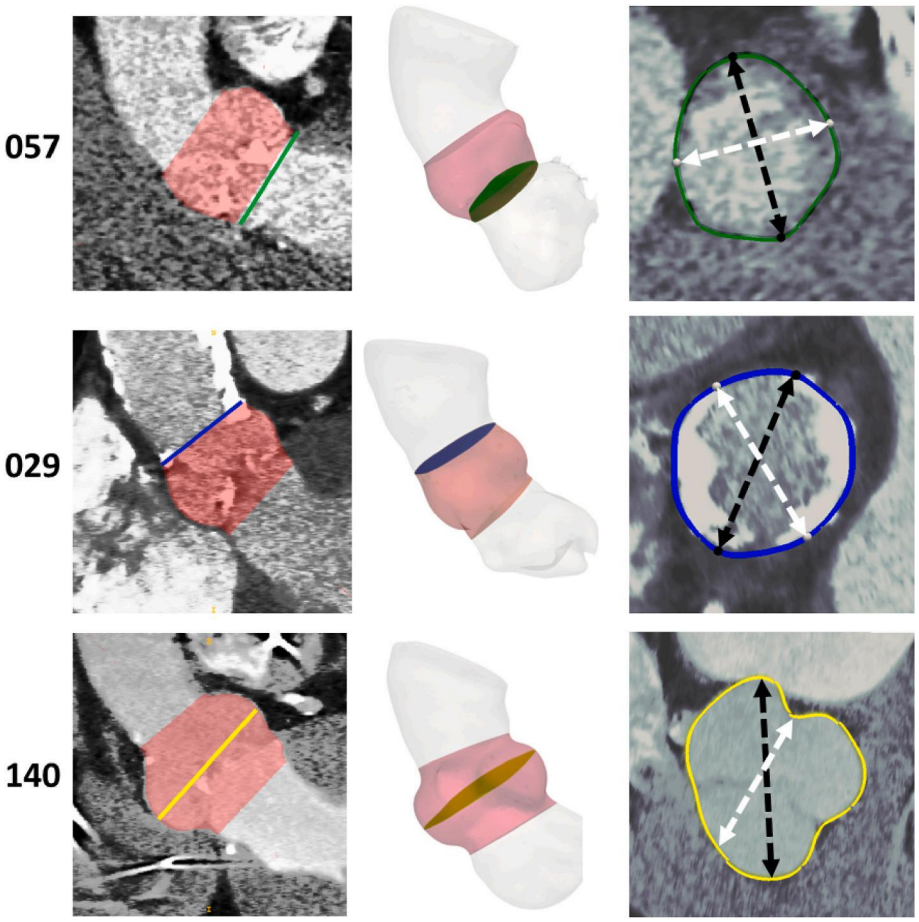

Fig. 7. Results of the implemented automatic pipeline for 3 cases for which the discrepancies between expert and automatic measurements were largest. Left column: segmentation inferred by model 1 and bounded to the AR region (in red). Central column: 3D anatomy reconstructions together with either the annulus (green), STJ (blue) or sinuses plane (yellow). Right column: interpolation of the image onto the plane identified by either the annulus, STJ or sinuses planes. Segmentation contours are show in green (annulus), blue (STJ) and yellow (sinuses), together with D_{max} (dashed arrows, black) and D_{min} (dashed arrows, white).

different portions of the left ventricle due to CT scan field of view or inter-operator variability. This led to label map predictions with differing extents (as it is visible in Fig. 5), entailing slightly worse accuracy with respect to other similar studies. This inconsistency was irrelevant to the downstream tasks of our pipeline and did not invalidate the overall effectiveness of our approach, which only requires the LVOT to be segmented correctly. In addition, when considering only the AR region, mean values of Dice score and MSD of 0.96 and 0.14 mm were found. These errors are in the range of twice the average pixel spacing of 0.56 mm of our image dataset. In our experiments, models with 3D U-Net architectures (models 1 and 2) performed better than transformer-based architectures (models 3 and 4) and outperformed the implemented atlas-based approach. In some cases, the b-spline registration method underlying the atlas-based approach proved to be unsuitable for the task at hand, yielding segmentations with no overlap with the ground truth. The realm of neural networks for image segmentation is vast, and further studies are necessary to investigate whether different architectures such as UNet++ [32] or models based on atrous convolutions [33] may provide better performance.

Concerning the identification of the aortic annulus and STJ, our approach (model 2) differs from previously published techniques since it derives anatomical landmarks from image segmentation and differential geometry. Elattar et al. [10] based their methodology on extracting the AR centerline, computing the Gaussian curvature and an harmonic decomposition of the aortic shape. Their automatic measurements were in excellent agreement with manual ones, with differences in annulus radii of 0.24 ± 0.70 mm and 0.37 ± 0.82 mm for two observers, respectively, and differences in annulus angles of 6.86 ± 5.39° and 6.34 ± 4.00°. We obtained slightly larger differences for annulus mean radii (0.34 ± 0.78 mm), lower differences for annulus area-derived radii (0.006 ± 0.16 mm), and similar discrepancies for annulus angles (−3.07 ± 7.15°). However, [10] still represents a proof of concept, lacking generalization, since automatic measurements are validated against a limited set of 40 patients. In our experience, Gaussian curvature-based criteria for landmark detection would perform sub-optimally or fail for anatomies with less pronounced bulb curvature. Additionally, we showed how atlas-based methods are limited in their capability to generalize to very heterogeneous anatomies [13]. The significantly larger number of test patients included in our study allowed us to devise robust solutions that proved able to deal with great anatomical variability among subjects. To date, with 178 patients, our study presents the most comprehensive validation of an automatic AR assessment system.

Another type of landmark detection approach implies training a neural network to directly localize key anatomical points. Astudillo et al. [6] adopt this kind of strategy using a DenseVNet architecture to detect coronary ostia position and height with respect to the annulus, validating their accuracy against manual annotations on 100 patients. Compared to manual annotations, they report mean differences of 0.54 mm and −0.16 mm in left and right coronary heights, and of 1.4 mm in 3D Euclidean distances of coronary ostia. However, when utilizing a landmark-based method for anatomical measurement quantification, an error in identifying even one landmark can result in significant errors in more complex anatomical metrics. In contrast, in our approach we did not rely on direct neural network-based landmark detection, but rather on a series of geometric computations for robust identification of the aortic annulus and STJ. In particular, the annulus and STJ plane refinement step enabled our system to cope with segmentation errors, requiring only a rough localization of the two anatomical regions. This makes the results of our pipeline more dependent on the accuracy of AR

and LVOT segmentation, which are more clearly defined and visible from CT scans, and thus pose an easier challenge to deep learning segmentation systems.

Given the large number of cases on which we tested our automatic measurement method, it is reasonable to expect some large discrepancies with respect to manual measurements. In the critical cases reported in Fig. 7, the automatic segmentation algorithms were able to segment the anatomical structures with good precision. However, some anatomical features such as high curvature of the aortic bulb, and presence of an abundant of calcium deposits could cause differences in how diameters are computed by the two approaches. Our system has the potential to automatically segment calcium deposits, giving the end-user the choice to include or exclude them in the computed measurements.

An additional limitation of the proposed approach consists in not having tested our pipeline on severely anomalous anatomies, at those that may be associated with congenital heart diseases. In these cases, an accurate detection of the relevant anatomical structures may not be feasible with the current models. Moreover, we did not include an uncertainty quantification phase in our workflow; future efforts could focus on the development of Bayesian models for detecting anatomical landmarks with their associated uncertainty to enhance the models' explainability. Other potential sources of error that may affect the tool's accuracy include variations in image quality or the presence of artifacts (e.g., due to patient motion), limitations of the deep learning algorithms themselves (e.g., overfitting to training data or generalization to new data), and limited identification of the STJ and aortic annulus in some anatomies.

Upon solving the mentioned limitations, our system will have the potential to be used in clinical routine as a support to TAVI planning, as it will combine reliability and repeatability of AR measurements, which are critical for selecting the appropriate TAVI device size and type (e.g., balloon vs. self-expandable) and to reduce post-procedural complications such as paravalvular leakage or device migration, with unprecedented time-efficiency. Indeed, our system requires less than 45 s to run from CT to measurements on a GPU-accelerated workstation, while, even when leveraging commercially available semi-automated tools [12], the full AR anatomical assessment currently requires around 30 min on average. This improvement in time efficiency would allow for accurate TAVI planning for every patient and by any clinician, improving procedural success rates and patient outcomes, thus reducing hospital stays and healthcare costs. Also, owing to its time efficiency, our system may be well suited to large population studies.

5. Conclusions

We have presented a pipeline for automating quantification of complex morphological biomarkers that are relevant for preprocedural planning of TAVI from CT images, providing an unprecedented validation of our approach against 178 patients, the largest to date. Our proposed method demonstrates a quick, accurate, and consistent assessment of AR anatomy from CT data, yielding 3D segmentations of the AR (mean Dice score 0.96). Incorporating this deep learning-based tool into the preoperative planning routine in TAVI environments could potentially lead to time and cost savings, as well as improved accuracy.

Declaration of competing interest

The authors declare no competing interests.

Acknowledgements

This work was supported by IRCCS Policlinico San Donato, a clinical research hospital partially funded by the Italian Ministry of Health.

References

[1] M.B. Leon, et al., Transcatheter or surgical aortic-valve replacement in intermediate-risk patients, N. Engl. J. Med. 374 (17) (2016) 1609–1620.
[2] M.J. Mack, et al., 5-year outcomes of transcatheter aortic valve replacement or surgical aortic valve replacement for high surgical risk patients with aortic stenosis (PARTNER 1): a randomised controlled trial, Lancet 385 (9986) (2015) 2477–2484.
[3] A. Khosravi, O. Wendler, TAVI 2018: from guidelines to practice, Eur J Cardiol Pract 15 (2018) 29.
[4] P. Breitbart, et al., Implantation depth and its influence on complications after TAVI with self-expanding valves, Int. J. Cardiovasc. Imag. 37 (10) (2021) 3081–3092.
[5] S. Bhushan, et al., Paravalvular leak after transcatheter aortic valve implantation its incidence, diagnosis, clinical implications, prevention, management, and future perspectives: a review article, Curr. Probl. Cardiol. 47 (10) (2022), 100957.
[6] P. Astudillo, et al., Automatic detection of the aortic annular plane and coronary ostia from multidetector computed tomography, J. Intervent. Cardiol. (2020) 2020.
[7] A.G. Cerillo, et al., Sizing the aortic annulus, Ann. Cardiothorac. Surg. 1 (2) (2012) 245.
[8] R. Gorla, et al., Impact of aortic angle on transcatheter aortic valve implantation outcome with Evolut-R, Portico, and Acurate-NEO, Cathet. Cardiovasc. Interv. 97 (1) (2021) E135–E145.
[9] C.J. Schultz, et al., Three dimensional evaluation of the aortic annulus using multislice computer tomography: are manufacturer's guidelines for sizing for percutaneous aortic valve replacement helpful? Eur. Heart J. 31 (7) (2010) 849–856.
[10] M. Elattar, et al., Automatic aortic root landmark detection in CTA images for preprocedural planning of transcatheter aortic valve implantation, Int. J. Cardiovasc. Imag. 32 (3) (2016) 501–511.
[11] S. Queirós, et al., Automatic 3D aortic annulus sizing by computed tomography in the planning of transcatheter aortic valve implantation, Journal of cardiovascular computed tomography 11 (1) (2017) 25–32.
[12] V. Delgado, et al., Automated assessment of the aortic root dimensions with multidetector row computed tomography, Ann. Thorac. Surg. 91 (3) (2011) 716–723.
[13] F. Lalys, et al., Automatic aortic root segmentation and anatomical landmarks detection for TAVI procedure planning, Minim Invasive Ther. Allied Technol. 28 (3) (2019) 157–164.
[14] P.G. Tahoces, R. Varela, J.M. Carreira, Deep learning method for aortic root detection, Comput. Biol. Med. 135 (2021), 104533.
[15] Y. Zheng, et al., Automatic aorta segmentation and valve landmark detection in C-arm CT: application to aortic valve implantation, in: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, 2010.
[16] S. Saitta, et al., A deep learning-based and fully automated pipeline for thoracic aorta geometric analysis and planning for endovascular repair from computed tomography, J. Digit. Imag. 35 (2) (2022) 226–239.
[17] P.A. Yushkevich, Y. Gao, G. Gerig, ITK-SNAP: an interactive tool for semi-automatic segmentation of multi-modality biomedical images, in: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), IEEE, 2016.
[18] W.J. Schroeder, L.S. Avila, W. Hoffman, Visualizing with VTK: a tutorial, IEEE Computer graphics and applications 20 (5) (2000) 20–27.
[19] K. Marstal, et al., SimpleElastix: a user-friendly, multi-lingual library for medical image registration, in: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2016.
[20] F. Maes, D. Vandermeulen, P. Suetens, Medical image registration using mutual information, Proc. IEEE 91 (10) (2003) 1699–1722.
[21] F. Isensee, et al., nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation, Nat. Methods 18 (2) (2021) 203–211.
[22] Z. Zhang, Q. Liu, Y. Wang, Road extraction by deep residual u-net, Geosci. Rem. Sens. Lett. IEEE 15 (5) (2018) 749–753.
[23] K. He, et al., Delving deep into rectifiers: surpassing human-level performance on imagenet classification, in: Proceedings of the IEEE International Conference on Computer Vision, 2015.
[24] S. Klein, et al., Elastix: a toolbox for intensity-based medical image registration, IEEE Trans. Med. Imag. 29 (1) (2009) 196–205.
[25] A. Hatamizadeh, et al., Unetr: transformers for 3d medical image segmentation, in: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, 2022.
[26] M.J. Cardoso, et al., MONAI: an Open-Source Framework for Deep Learning in Healthcare, 2022 arXiv preprint arXiv:2211.02701.
[27] M. Yeung, et al., Unified Focal loss: generalising Dice and cross entropy-based losses to handle class imbalanced medical image segmentation, Comput. Med. Imag. Graph. 95 (2022), 102026.
[28] S. Valette, J.M. Chassery, R. Prost, Generic remeshing of 3D triangular meshes with metric-dependent discrete Voronoi diagrams, IEEE Trans. Visual. Comput. Graph. 14 (2) (2008) 369–381.
[29] M. Reuter, F.-E. Wolter, N. Peinecke, Laplace–Beltrami spectra as 'Shape-DNA' of surfaces and solids, Comput. Aided Des. 38 (4) (2006) 342–366.
[30] N. Bettinger, et al., Practical determination of aortic valve calcium volume score on contrast-enhanced computed tomography prior to transcatheter aortic valve replacement and impact on paravalvular regurgitation: elucidating optimal threshold cutoffs, Journal of Cardiovascular Computed Tomography 11 (4) (2017) 302–308.

[31] M.A. Elattar, et al., Automatic segmentation of the aortic root in CT angiography of candidate patients for transcatheter aortic valve implantation, Med. Biol. Eng. Comput. 52 (7) (2014) 611–618.
[32] Z. Zhou, et al., Unet++: a nested u-net architecture for medical image segmentation, in: Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, *and 8th International Workshop, ML-CDS 2018*, Held in Conjunction with MICCAI 2018, Granada, Spain, September 20, 2018, Proceedings 4, Springer, 2018.
[33] L.-C. Chen, et al., Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs, IEEE Trans. Pattern Anal. Mach. Intell. 40 (4) (2017) 834–848.