

An Automated Deep Learning Software Tool for Assessing Aortic Remodeling in Aortic Dissection

Yu Nakano

`nyk3151@tokyo-med.ac.jp`

Tokyo Medical University

Ikki Kojima

Tokyo Medical University

Masaki Kano

Tokyo Medical University

Toshiki Fujiyoshi

Tokyo Medical University

Yusuke Shimahara

Tokyo Medical University

Article

Keywords: Aortic dissection, Deep learning, Image segmentation, Computed tomography angiography, Artificial intelligence

Posted Date: July 10th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10095993/v1>

License: © ⓘ This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Computed tomography (CT) is the gold standard for assessing aortic remodeling following aortic dissection treatment. Large-scale clinical studies utilizing CT data are needed to evaluate the effectiveness of interventions such as the frozen elephant trunk (FET) procedure. However, such studies are currently limited by the time-consuming and labor-intensive nature of manual image analysis. To address this, we developed an AI-based automated pipeline to assess morphological changes in aortic dissection, utilizing open datasets (TotalSegmentator, ImageTBAD, AortaSeg24). In this proof-of-concept study, we evaluated the efficacy of the pipeline using preoperative and postoperative CT scans from 14 patients who underwent FET repair. Two surgeons manually measured the areas of the whole aorta and true lumen at the Th8 and Th12 levels. Our automated pipeline analyzed the same metrics, and the results were compared with the manual measurements. The pipeline showed excellent agreement with the manual reference, with intraclass correlation coefficients (ICCs) comparable to interobserver reproducibility (aortic area, 0.913; aortic diameter, 0.911; true lumen [TL] area, 0.947; TL ratio, 0.901). Future work will require model refinement and validation with larger datasets.

Introduction

Aortic dissection is a life-threatening cardiovascular condition. Whereas timely surgical intervention is critical for survival in the acute phase, patients remain at significant risk for aneurysm formation and re-dissection in the late phase. Furthermore, the recent introduction of novel surgical techniques and devices, such as the frozen elephant trunk (FET) procedure¹ and the Ascyrus medical dissection stent (AMDS; Artivion®, Atlanta, GA, USA)², necessitates a comprehensive evaluation of postoperative aortic remodeling to determine the long-term clinical efficacy of these procedures.

Computed tomography (CT) is the established gold standard for assessing aortic remodeling in both routine clinical practice and clinical trials. However, evaluating morphological changes requires analyzing preoperative and postoperative CT scans across large patient cohorts. This process is currently heavily limited by the time-consuming, labor-intensive, and costly nature of manual image analysis, presenting a major bottleneck for large-scale clinical studies.

In recent years, artificial intelligence (AI), particularly deep learning, has made significant progress in medical image analysis, with numerous studies reporting automated segmentation of the true and false lumen in aortic dissection^{3–11}.

While recent technical trends in AI have increasingly explored three-dimensional (3D) volumetric analysis, clinical decision-making still fundamentally relies on established two-dimensional (2D) metrics, such as maximum diameter and cross-sectional area at specific anatomical landmarks. This continued reliance on 2D evaluation, along with the current limitation in user-friendly software interfaces for clinicians, remains a key factor to consider when integrating advanced AI models into routine clinical workflows.

To bridge this gap, we developed a deep learning-based automated software pipeline utilizing established open-source datasets (TotalSegmentator¹², ImageTBAD⁵, and AortaSeg24¹³). Our approach leverages the precision of advanced 3D AI segmentation to automatically extract the 2D metrics that are most clinically relevant to surgeons. In this study, we aimed to demonstrate the clinical applicability of our software. As proof of concept, we evaluated its efficacy by comparing the automated measurements of the true lumen and whole aorta areas against manual assessments by cardiovascular surgeons using preoperative and postoperative CT scans. We focused on FET repair because it produces clinically important postoperative aortic remodeling that currently lacks a standardized, automated method for quantitative assessment.

Materials and Methods

Software Implementation

The proposed pipeline was implemented as an extension module for 3D Slicer (version 5.8.1), an open-source medical image computing platform¹⁴. The source code is available at <https://github.com/nyk3151/SlicerAorticPipeline>. All processing was performed on a workstation equipped with an NVIDIA graphics processing unit (GPU) running compute unified device architecture (CUDA)-accelerated inference.

Deep Learning Framework

For the core segmentation tasks, we employed no new (nn)U-Net (version 2), a self-configuring framework for biomedical image segmentation¹⁵. This architecture was selected on the basis of its demonstrated efficacy and robustness in multiple prior studies of aortic dissection segmentation^{9–11}, ability to automatically adapt to dataset characteristics, and widespread adoption as a benchmark in medical image analysis. The 3D full-resolution configuration (3d_fullres) was used for both segmentation models. Inference was performed using an ensemble of all five cross-validation folds, as recommended by the nnU-Net framework, to maximize segmentation robustness.

Training Datasets and Label Preprocessing

Two independent open-source datasets were utilized to train the respective segmentation models.

The AortaSeg24 dataset¹³ was used to train the whole-aorta segmentation model, which comprises 100 cases including aortic dissections, originally annotated according to the Society of Thoracic Surgeons zone classification¹⁶ without true/false lumen differentiation. For our pipeline, the original multizone annotations were simplified into two anatomically meaningful classes: Class 1 (which represented the aorta from Zone 1 through to the bilateral common iliac arteries) and Class 2 (which represented the ascending aorta (Zone 0)). Annotations for the cervical and abdominal branches, as well as the internal and external iliac arteries, were excluded from the training labels. The model was trained using five-fold

cross-validation. The “whole-aorta” segmentation used for centerline extraction was reconstructed at inference time as the union of Class 1 and Class 2 predictions.

The ImageTBAD dataset⁵ was utilized for luminal segmentation, which contains 98 cases of Type B aortic dissection, originally annotated for the true lumen (TL), a patent false lumen (FL), and FL thrombosis. The model was trained using five-fold cross-validation. Because the reported baseline accuracy for FL thrombosis segmentation was as low as a Dice similarity coefficient score of 0.52 in the original study, our inference strategy used only the true lumen predictions for the final automated measurements to ensure clinical reliability. The FL area, when reported, was derived by subtracting the TL area from the total aortic cross-sectional area. Further details are provided in the Supplementary Materials.

Study Cohort for Pipeline Evaluation

This study followed the guidelines of the Declaration of Helsinki and was approved by the Institutional Review Board of Tokyo Medical University Hospital (Approval No. T2025-0156). Informed consent was obtained from all patients using the opt-out method because of the retrospective nature of this study. We retrospectively reviewed 55 patients who underwent total arch replacement (TAR) + FET between January 1, 2024, and March 31, 2025. Patients were excluded based on the following criteria: (1) absence of a computed tomography angiography (CTA) scan; (2) a slice thickness exceeding 3.0 mm; (3) dissection not extending to the Th12 level; a history of aortic intervention; and (5) severe artifacts precluding accurate anatomical assessment. Consequently, the final dataset comprised 14 patients (28 CT scans: 14 preoperative and 14 postoperative), yielding 56 measurement points (14 patients × 2 timepoints × 2 anatomical levels). This cohort also enabled an exploratory assessment of postoperative aortic remodeling following FET repair. The patient selection flowchart is presented in Supplementary Figure S1.

Automated Pipeline Architecture

To automatically derive clinically relevant 2D measurements from 3D patient anatomy, we designed a sequential, multistep image processing pipeline (Figure 1). The complete workflow is fully automated and requires no manual intervention.

Step 1—Multi-structure segmentation. The pipeline performed comprehensive 3D segmentation as follows:

1. The aorta (Class 1 and Class 2), using the AortaSeg24-trained nnU-Net model
2. The aortic TL, using the ImageTBAD-trained nnU-Net model
3. Vertebral bodies—specifically, the 8th thoracic (Th8), 12th thoracic (Th12), and 4th lumbar (L4) vertebrae—were evaluated using TotalSegmentator, a publicly available tool with previously validated performance for vertebral segmentation.

Step 2—Whole-aorta reconstruction and endpoint definition. The whole-aorta segmentation was reconstructed by computing the voxel-wise union of the Class 1 and Class 2 predictions. The centerline endpoints were determined automatically as follows:

1. **Starting point:** the centroid (center of mass) of the ascending aorta segmentation (Class 2)
2. **Termination point:** the centroid of the aortic cross-sectional mask at the cranio-caudal (z-axis) level corresponding to the L4 vertebral centroid

Step 3—Centerline extraction. The aortic centerline was computed using the vascular modeling toolkit (VMTK)¹⁷ from the reconstructed whole-aorta segmentation. The extracted centerline was resampled to 200 control points to ensure a stable curved planar reformation while maintaining adequate longitudinal sampling along the vessel.

Step 4—Vertebral landmark projection. The centroids of the Th8 and Th12 vertebrae were computed from the TotalSegmentator output. The Th8 level was chosen because a previous nationwide Japanese multicenter study reported that the distal end of the FET was located at or above the Th8 level in nearly all patients¹⁸; Th12 was used to represent the lower level of the descending aorta. Each vertebral centroid was projected onto the aortic centerline by selecting the centerline point with the minimum cranio-caudal (z-axis) distance to the centroid.

Step 5—Image straightening (curved planar reformation). Using the extracted centerline, both the native CT volume and the segmentation label maps underwent curved planar reformation (CPR) to produce straightened multiplanar reconstructions. CPR was applied to the segmentation label maps in addition to the CT volume to ensure that cross-sectional area measurements were performed on geometrically aligned data.

Step 6—2D cross-sectional measurement. Strictly orthogonal 2D cross-sections were automatically extracted from the straightened volume at the projected Th8 and Th12 centerline points. The aorta and true lumen segmentations were resampled onto the CPR-straightened volume with 1×1×1 mm isotropic voxels using nearest-neighbor interpolation to preserve the original binary labels. At each level, the cross-sectional area was computed as the number of labeled voxels in the corresponding one-voxel-thick axial section multiplied by the in-plane voxel area (1 mm²).

Derived morphological parameters

The following parameters were calculated at each measurement level:

- Aortic cross-sectional area (mm²)
- Aortic diameter (mm), derived as $Diameter = 2\sqrt{\frac{Area}{\pi}}$
- TL area (mm²)
- TL ratio (%), defined as $TL\ ratio = \frac{TL\ area}{Aortic\ area} \times 100$
- FL area (mm²), derived by subtraction

The aortic diameter formula assumes a circular cross-section, which is a reasonable approximation for the overall aorta but is not appropriate for the often crescent-shaped TL. Therefore, TL measurements were reported as the area and TL ratio rather than as the diameter.

Manual Reference Measurements

Two independent observers (Observer A (a cardiovascular surgeon) and Observer B (a resident with 5 years of postgraduate training) performed manual measurements at the Th8 and Th12 levels on each CT scan. Each observer independently identified the corresponding axial slice using the multiplanar reconstruction (MPR) view of the workstation, with reference to the vertebral level. Manual segmentation of the aortic outer wall and the TL was then performed on the identified slice using Vincent (Fujifilm Corporation, Tokyo, Japan), a clinical 3D image analysis workstation. The cross-sectional areas were derived from the manual segmentations. The two observers were blinded to the automated pipeline results and each other's measurements throughout the process.

Statistical Analysis

Interobserver reproducibility (Observer A vs. Observer B) and pipeline accuracy (automated vs. mean of A and B) were assessed using the intraclass correlation coefficient (ICC(A,1)), two-way random effects, absolute agreement, and single rater. Bland–Altman analysis (mean bias and 95% limits of agreement (LoA = bias ± 1.96 × SD) ICC values were interpreted using the criteria proposed by TK Koo and Li MY¹⁹: poor (<0.50), moderate (0.50–0.75), good (0.75–0.90), and excellent (>0.90). All the statistical analyses were performed using the pingouin package in Python 3.12. Subgroup analyses were performed by the surgical phase (preoperative vs. postoperative) and anatomical level (Th8 vs. Th12) to identify potential factors affecting pipeline accuracy.

Results

Patient characteristics

The demographic and clinical characteristics of the 14 patients are summarized in Supplementary Table S1. The mean age of the study population was 59.1 ± 11.9 years, and 85.7% were male. Stanford Type A dissection was observed in 12 patients (85.7%) and Type B in two patients (14.3%). Notably, the dataset predominantly consisted of high-resolution images, with a mean CT slice thickness of 1.1 ± 0.5 mm. The median distal landing level of the FET was Th7 (IQR Th6–Th7; range Th5–Th8).

Segmentation performance

On five-fold cross-validation, the structures used for the downstream measurements were segmented accurately: the aorta achieved a mean Dice of 0.94 for Class and 0.91 for Class 2, and the true lumen achieved a mean Dice of 0.95. In contrast, the thrombosed false lumen was segmented far less reliably (mean Dice 0.31), which supported basing the automated measurements on true lumen predictions only (Supplementary Tables S2–S3). Building on these segmentations, the primary analysis assessed the agreement between the pipeline-derived 2D metrics and the manual reference, as described below.

Pipeline Performance

The fully automated pipeline successfully processed all 28 CT scans (14 patients $\times$ 2 timepoints) without manual intervention or pipeline failure. The mean processing time per scan was approximately 12.55 min on a workstation equipped with an NVIDIA Quadro RTX 6000 GPU (24GB), including all the segmentation, centerline extraction, CPR, and cross-sectional measurement steps.

Interobserver Reproducibility (Table 1)

Table 1 Interobserver Reproducibility						
Parameter	ICC(A,1)	95% CI	Bias	95% LoA	MAE	MAPE
Aortic diameter (mm)	0.963	0.930–0.980	+ 0.35	[-1.43, + 2.13]	0.72	2.34%
TL ratio (%)	0.983	0.970–0.990	+ 0.23	[-5.59, + 6.05]	2.38	6.86%
Aortic area (mm ²)	0.962	0.930–0.980	+ 16.57	[-71.12, + 104.26]	35.00	4.68%
TL area (mm ²)	0.989	0.980–0.990	+ 9.30	[-34.88, + 53.48]	16.73	6.00%
ICC (A,1), intraclass correlation coefficient (two-way random effects, absolute agreement, single rater); CI, confidence interval; LoA, limits of agreement; MAE, mean absolute error; MAPE, mean absolute percentage error.						

Interobserver reproducibility between Observer A and Observer B was excellent across all the measured parameters (n = 56 measurement pairs).

These results establish the reliability of the manual reference measurements, with ICC values exceeding 0.96 for all the parameters, supporting the use of the mean of A and B as the reference standard for assessing pipeline accuracy.

Pipeline Accuracy: Automated vs. Manual Measurements (Table 2)

Table 2						
Pipeline accuracy: automated pipeline vs. manual reference (mean of Observers A and B, n = 56).						
Parameter	ICC (A,1)	95% CI	Bias	95% LoA	MAE	MAPE
Aortic diameter (mm)	0.911	0.840–0.950	+ 0.47	[-2.31, + 3.25]	1.00	3.30%
TL ratio (%)	0.901	0.810–0.950	+ 3.09	[-10.71, + 16.89]	5.47	19.19%
Aortic area (mm ²)	0.913	0.850–0.950	+ 23.07	[-110.05, + 156.19]	48.21	6.58%
TL area (mm ²)	0.947	0.750–0.980	+ 35.78	[-45.72, + 117.28]	39.15	18.93%
Manual reference: arithmetic mean of Observer A and Observer B. ICC interpretation: poor (< 0.50), moderate (0.50–0.75), good (0.75–0.90), excellent (> 0.90).						

The agreement between the automated pipeline measurements and the manual reference (mean of Observers A and B) was excellent for all the parameters (n = 56).

All the parameters achieved an ICC > 0.90, which was classified as excellent according to the criteria of Koo and Li. The aortic diameter, the most clinically relevant parameter, demonstrated a bias of only + 0.47 mm, with a 95% LoA of – 2.31 to + 3.26 mm, which was well within the typical range considered clinically acceptable for aortic measurements. The TL ratio, the primary remodeling indicator, was slightly systematically overestimated by the pipeline (bias + 3.09%) but maintained excellent agreement (ICC = 0.901). Bland–Altman plots for all the parameters are shown in Fig. 2.

Subgroup Analysis by Surgical Phase (Table 3)

Table 3
Subgroup analysis by surgical phase (Auto vs. Manual reference).

Parameter	Preoperative (n = 28)			Postoperative (n = 28)		
	ICC	95% CI	MAE	ICC	95% CI	MAE
Aortic diameter (mm)	0.933	0.860–0.970	0.92	0.889	0.770–0.950	1.08
True lumen ratio (%)	0.923	0.800–0.970	4.67	0.881	0.750–0.940	6.27
Aortic area (mm ²)	0.938	0.860–0.970	43.30	0.890	0.770–0.950	53.12
TL area (mm ²)	0.956	0.740–0.990	38.61	0.936	0.710–0.980	39.70
Excellent agreement (ICC > 0.88) was maintained across both the pre- and postoperative phase.						

Pipeline accuracy was further analyzed separately for preoperative and postoperative scans. Excellent agreement (ICC > 0.88) was maintained in both preoperative and postoperative settings, with slightly higher ICC values in the preoperative phase.

Subgroup Analysis by Anatomical Level (Supplementary Table S4)

Excellent agreement (ICC > 0.85) was achieved at both the Th8 and Th12 levels for all the parameters, indicating that pipeline performance was not strongly dependent on the anatomical level of measurement.

Exploratory Analysis: Postoperative Aortic Remodeling (Supplementary Table S5-S6)

As an exploratory analysis, postoperative aortic remodeling was assessed using an automated pipeline. At the Th8 level, the automated pipeline detected significant postoperative remodeling, with a substantial increase in the TL area ($\Delta + 80.64 \text{ mm}^2$, $p = 0.017$) and TL ratio ($\Delta + 9.61\%$, $p = 0.030$). The aortic diameter increased slightly but nonsignificantly ($\Delta + 0.95 \text{ mm}$, $p = 0.173$), and the FL area decreased nonsignificantly ($\Delta - 34.57 \text{ mm}^2$, $p = 0.358$).

At the Th12 level, similar but attenuated trends were observed, with nonsignificant increases in the TL area ($\Delta + 41.36 \text{ mm}^2$, $p = 0.068$) and TL ratio ($\Delta + 5.42\%$, $p = 0.135$).

These findings are consistent with the expected pattern of postoperative aortic remodeling following FET repair, characterized by TL expansion at the distal end of the FET stent (Th8), with a more attenuated effect at the more distal Th12 level^{20–22}. The greater magnitude of remodeling at Th8 suggests that more favorable aortic remodeling occurred in the segment near the deployed FET device. When the same exploratory remodeling analysis was performed using manual reference measurements, the directional changes and effect sizes were highly concordant with those derived from the automated pipeline (e.g., TL ratio at Th8: manual $\Delta + 10.43\%$ vs. automated $\Delta + 9.61\%$). Manual analysis revealed

statistically significant changes in additional parameters (aortic diameter and area at Th8), reflecting the smaller measurement variability of manual segmentation. Nevertheless, the consistency in the magnitude and direction of remodeling demonstrates that the automated pipeline reliably captures the clinically meaningful patterns of post-FET aortic remodeling. Representative automated segmentation results at the Th8 and Th12 levels are shown in Fig. 3.

Discussion

In this pilot study, we developed and validated a fully automated pipeline for the quantitative assessment of aortic remodeling. The pipeline integrates deep learning-based segmentation, automated anatomical landmark identification, and curved planar reformation (CPR) to produce reproducible cross-sectional measurements at standardized vertebral levels. Excellent agreement with the manual reference measurements was achieved across all parameters ($ICC > 0.90$), and the pipeline reproduced the expected pattern of postoperative true lumen (TL) expansion, particularly at the Th8 level adjacent to the distal end of the FET stent graft. In one case, however, the model failed to delineate the true lumen within the stented segment at Th8: while the aorta was correctly segmented, the automated true lumen was confined to a small region coincident with the high-attenuation stent struts, markedly underestimating the TL area (Fig. 4). This likely reflects the out-of-distribution nature of stent-graft anatomy, as the TL segmentation model was trained on ImageTBAD, which does not contain stented cases.

Comparison with Prior Studies

Several prior studies have applied deep learning to aortic dissection segmentation. Recently, clinically oriented automated analysis pipelines have begun to emerge. Zhuang et al.¹⁰ developed a deep learning pipeline focused on the diagnostic classification of aortic dissection, including Stanford type categorization, and achieved promising accuracy on a large multicenter cohort. Song et al.¹⁰ reported a comprehensive analysis pipeline for Type A aortic dissection, in which clinically relevant features such as the location of the primary entry tear, branch vessel involvement, and FL area ratio were extracted, providing a multidimensional automated characterization of dissection morphology. These contributions represent significant advances in the automated analysis of aortic dissection. However, the scope of these studies is primarily focused on diagnostic characterization at a single time point rather than on the longitudinal assessment of postoperative remodeling. Furthermore, although these pipelines extract sophisticated 3D features of aortic dissection, to our knowledge no clinically actionable framework has been reported that translates automated segmentations into the simple 2D cross-sectional metrics used by cardiovascular surgeons and radiologists in routine practice—namely aortic diameter, TL area, and TL ratio at standardized vertebral levels.

Our study addressed this gap by focusing on three distinguishing aspects:

(1) Clinical scope — postoperative remodeling assessment: Whereas prior pipelines target diagnostic classification or initial morphological characterization, our pipeline is designed specifically for the

longitudinal evaluation of aortic remodeling following FET repair, a clinical question with direct prognostic implications.

(2) Output framed in clinically familiar 2D metrics: The pipeline outputs cross-sectional area, diameter, and TL ratio at standardized vertebral levels (Th8 and Th12) that align directly with how clinicians describe and discuss aortic morphology in routine clinical communication and surgical decision-making.

(3) Standardized landmark-based reproducibility: By projecting vertebral centroids onto the aortic centerline, the pipeline ensures that preoperative and postoperative measurements—and measurements across different patients—are obtained at anatomically equivalent positions, eliminating one of the main sources of variability in manual longitudinal assessment.

Advantages and Extensibility of the Proposed Pipeline

End-to-end automation: The pipeline runs from a raw CT volume to final measurements without any manual landmark placement, slice selection, or contour editing. This is a key practical differentiator from existing tools that typically require operator interaction at multiple steps.

Modular and extensible architecture. The pipeline is implemented as a 3D Slicer extension¹ and is openly distributed on GitHub. Each processing step (segmentation, centerline extraction, CPR, measurement) is independently swappable, enabling future enhancement without architectural overhaul. For example:

- The analysis is not restricted to Th8 and Th12; any vertebral level (e.g., Th6, L1) can be specified through configuration, allowing the framework to be applied to other aortic segments..
- Alternative segmentation models—including those trained on different cohorts or imaging protocols—can be substituted without modifying the rest of the pipeline.
- Volumetric analysis can be extended naturally from the current 2D measurements by leveraging the already-segmented 3D structures, providing a path toward future research in 3D remodeling quantification.
- Batch processing is already supported, allowing efficient large-cohort analysis once additional cases become available.

Clinical Workflow Potential

The mean processing time per scan was approximately 12.5 min, including all segmentation and measurement steps, making the pipeline practical for retrospective cohort studies. This performance was obtained on a workstation equipped with an NVIDIA RTX 6000 GPU; broader or near-real-time clinical use would therefore require adequate GPU resources and/or further model optimization (e.g., lightweight model variants) to achieve comparable processing times. Integration into clinical review platforms, although beyond the scope of this study, is a feasible future direction given the open-source 3D Slicer foundation.

Limitations

Sample Size and Single-center Design

The most fundamental limitation of this study is the small sample size ($n = 14$ patients; 56 measurement points), which reflects the strict imaging quality criteria applied during patient selection. Although this rigorous selection enhances data quality, it limits the generalizability of the findings. A larger multicenter prospective study is needed to confirm the performance of the pipeline across different scanner platforms, imaging protocols, and patient populations.

TL-only Luminal Segmentation

Because the original ImageTBAD study reported relatively low accuracy for FL thrombosis segmentation, our pipeline restricted luminal output to the TL only. The FL area was therefore derived by subtraction, without distinguishing between patent and thrombosed components. This simplification was deliberate to avoid introducing unreliable data into the pipeline, but it limits the analysis of FL status, which has prognostic relevance after FET repair. The incorporation of more accurate FL and thrombosis segmentation models, when available, would enable a more nuanced characterization of postoperative remodeling.

Vertebral Centroid Landmark

The vertebral masks generated by TotalSegmentator include the spinous and transverse processes, so the resulting centroid may not coincide with the geometric center of the vertebral body. We therefore used only the cranio-caudal (z-axis) coordinate of the centroid as the reference, because the posterior elements shift the centroid mainly within the axial plane and have a comparatively small effect on its cranio-caudal position. Moreover, any residual offset is expected to be systematic across patients and vertebral levels, and therefore does not bias the relative, level-to-level and pre- versus postoperative comparisons on which our analysis is based. Vertebral segmentation could also be affected by compression fractures or prior spinal interventions; however, no such cases were included in our dataset.

2D Cross-sectional Analysis

In this study, aortic morphology was quantified using 2D cross-sectional measurements at two predefined vertebral levels, which serve as surrogates for full 3D volumetric assessment. Although 2D measurements are clinically intuitive and align with standard radiological reporting, they cannot fully characterize complex 3D remodeling patterns. The current pipeline forms a foundation for future extension to volumetric analysis, leveraging the same segmentation and centerline outputs.

Future Directions

Enlargement and Expansion of Training Data

Pipeline accuracy was fundamentally constrained by the existing training datasets available—AortaSeg24 (100 cases) and ImageTBAD (98 cases). FET-specific imaging characteristics, including

stent metal artifacts and altered postoperative geometry, were not well represented in these datasets. Construction of a curated FET-specific dataset, ideally through multicenter collaboration, is the most direct path toward improving accuracy in this clinical setting.

Longitudinal Remodeling Tracking

Beyond pre/post comparisons, the same pipeline can be applied serially over multiple postoperative time points to characterize the kinetics of aortic remodeling, including the differential remodeling at different aortic levels, temporal pattern of false lumen thrombosis, and relationship between remodeling and clinical outcomes. The reproducibility of standardized landmarks and deterministic nature of the pipeline make it especially well suited for such longitudinal analysis.

Clinical Decision Support Integration

The combination of automated measurement and standardized reporting provided potential pathways toward clinical decision support—for example, automated flagging of cases with insufficient remodeling or progressive FL enlargement. The realization of such applications will require regulatory consideration, prospective validation, and integration with hospital information systems.

Conclusion

We developed and validated a fully automated pipeline for quantitative cross-sectional measurement of the aorta from CT and demonstrated its application to the assessment of aortic remodeling after FET repair. The pipeline achieved excellent agreement with manual reference measurements (ICC > 0.90 for all parameters) and successfully detected the expected pattern of postoperative TL expansion at the distal level of the FET device. Although the present pilot study has clear limitations—most notably, its small sample size and reliance on existing public-training datasets— the proposed framework provides a methodologically sound and clinically extensible foundation for future larger-scale studies of FET-related aortic remodeling.

Declarations

Acknowledgments

Generative AI tools were used for coding support, translation, and spelling correction of the manuscript. The authors reviewed and edited the output and take full responsibility for the final manuscript.

Author Contributions

Yu Nakano: Conceptualization, Data curation, Formal analysis, Methodology, Software, Visualization, Writing—original draft.

Ikki Kojima: Methodology, Data curation, Writing—review & editing.

Masaki Kano: Writing—review & editing.

Toshiki Fujiyoshi: Writing—review & editing.

Yusuke Shimahara: Supervision, Writing—review & editing.

Data Availability statement

The source code is available at <https://github.com/nyk3151/SlicerAorticPipeline>. The datasets used and/or analyzed during the current study are available from the corresponding author upon reasonable request except original CT scans.

Competing Interests

The authors declare that they have no competing interests.

Funding

This study was supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI (grant number JP25K15998).

References

1. Kato, M. *et al.* New graft-implanting method for thoracic aortic aneurysm or dissection with a stented graft. *Circulation* **94**, 11188-93 (1996).
2. Bozso, S. J. *et al.* Dissected Aorta Repair Through Stent Implantation trial: Canadian results. *J. Thorac. Cardiovasc. Surg.* **157**, 1763–1771 (2019).
3. Li, Z. *et al.* Lumen segmentation of aortic dissection with cascaded convolutional network. in *Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges* 122–130 (Springer International Publishing, Cham, 2019). doi:10.1007/978-3-030-12029-0_14.
4. Hahn, L. D. *et al.* CT-based true- and false-lumen segmentation in type B aortic dissection using machine learning. *Radiol. Cardiothorac. Imaging* **2**, e190179 (2020).
5. Yao, Z. *et al.* ImageTBAD: A 3D computed tomography angiography image dataset for automatic segmentation of type-B Aortic Dissection. *Front. Physiol.* **12**, 732711 (2021).
6. Chen, D. *et al.* Multi-stage learning for segmentation of aortic dissections using a prior aortic anatomy simplification. *Med. Image Anal.* **69**, 101931 (2021).

7. Yu, Y. *et al.* A three-dimensional deep convolutional neural network for automatic segmentation and diameter measurement of type B aortic dissection. *Korean J. Radiol.* **22**, 168–178 (2021).
8. Zhang, X. *et al.* Deep learning-based multi-stage postoperative type-b aortic dissection segmentation using global-local fusion learning. *Phys. Med. Biol.* **68**, (2023).
9. Zhuang, C. *et al.* A fully automatic pipeline of identification, segmentation, and subtyping of aortic dissection from CT angiography. *Cardiovasc. Eng. Technol.* **16**, 465–480 (2025).
10. Song, S. *et al.* Domain knowledge based comprehensive segmentation of Type-A aortic dissection with clinically-oriented evaluation. *Med. Image Anal.* **102**, 103512 (2025).
11. Fournel, J. *et al.* Deep learning-driven false-lumen volumes predict adverse remodeling better than diameter in patients with residual aortic dissection on CT. *Eur. Radiol.* 1–14 (2025)
doi:10.1007/s00330-025-12116-9.
12. Wasserthal, J. *et al.* TotalSegmentator: Robust segmentation of 104 anatomic structures in CT images. *Radiol. Artif. Intell.* **5**, e230024 (2023).
13. Imran, M. *et al.* Multi-class segmentation of aortic branches and zones in computed tomography angiography: The AortaSeg24 challenge. *arXiv [eess.IV]* (2025).
14. Fedorov, A. *et al.* 3D Slicer as an image computing platform for the Quantitative Imaging Network. *Magn. Reson. Imaging* **30**, 1323–1341 (2012).
15. Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J. & Maier-Hein, K. H. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**, 203–211 (2021).
16. Lombardi, J. V. *et al.* Society for Vascular Surgery (SVS) and Society of Thoracic Surgeons (STS) reporting standards for type B aortic dissections. *J. Vasc. Surg.* **71**, 723–747 (2020).
17. Antiga, L. *et al.* An image-based modeling framework for patient-specific computational hemodynamics. *Med. Biol. Eng. Comput.* **46**, 1097–1112 (2008).
18. Ogino, H. *et al.* Comparative study of Japanese frozen elephant trunk device for open aortic arch repairs. *J. Thorac. Cardiovasc. Surg.* **164**, 1681-1692.e2 (2022).
19. Koo, T. K. & Li, M. Y. A guideline of selecting and reporting intraclass correlation coefficients for reliability research. *J. Chiropr. Med.* **15**, 155–163 (2016).
20. Inoue, Y. *et al.* Comparative study of the frozen elephant trunk and classical elephant trunk techniques to supplement total arch replacement for acute type A aortic dissection. *Eur. J. Cardiothorac. Surg.* **56**, 579–586 (2019).
21. Yamamoto, H. *et al.* Total arch repair with frozen elephant trunk using the “zone 0 arch repair” strategy for type A acute aortic dissection. *J. Thorac. Cardiovasc. Surg.* **159**, 36–45 (2020).
22. Tochii, M. *et al.* Aortic remodeling with frozen elephant trunk technique for Stanford type A aortic dissection using Japanese J-graft open stent graft. *Heart Vessels* **34**, 307–315 (2019).

Figures

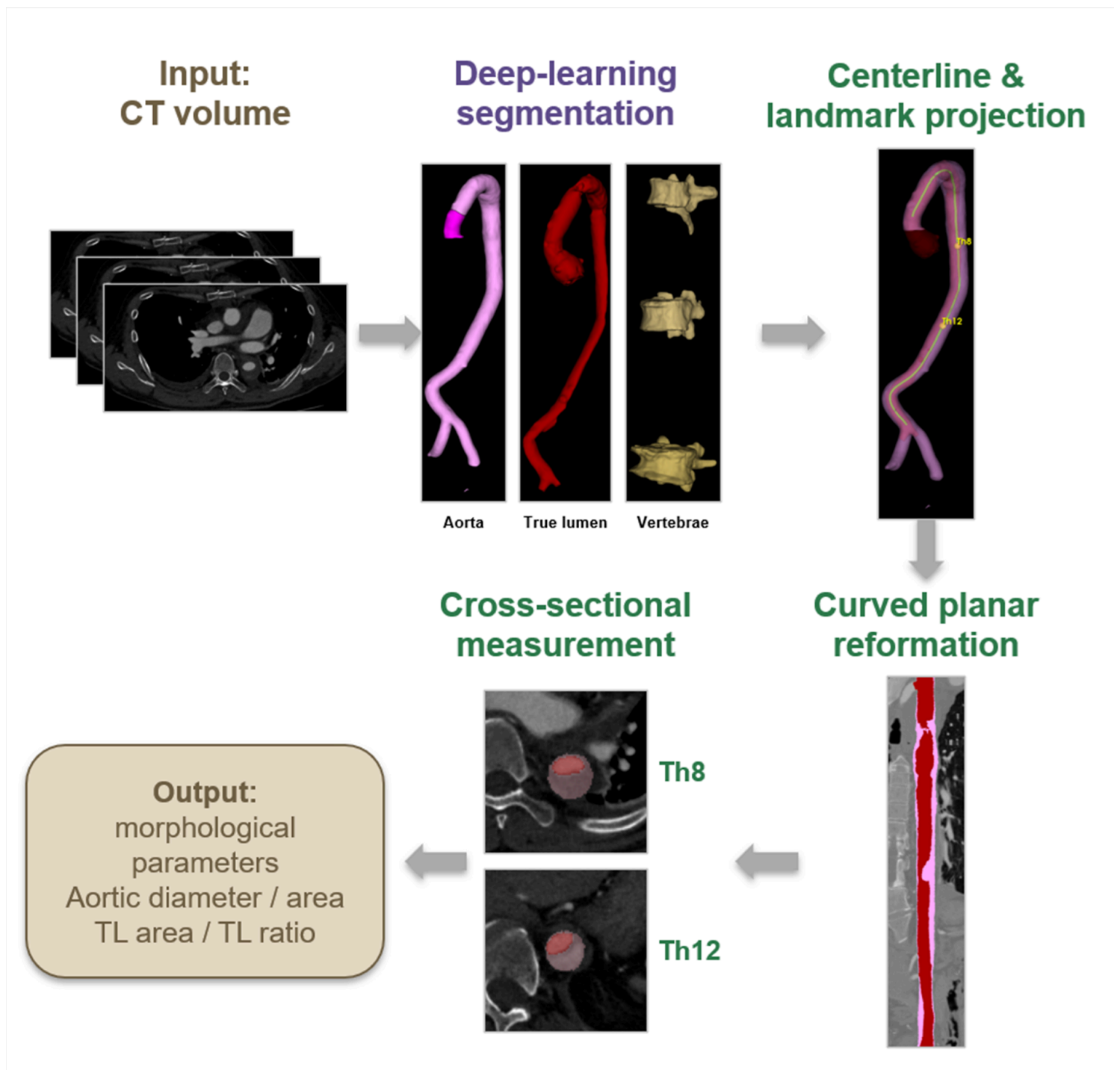


Figure 1

Overview of the fully automated analysis pipeline. From the input CT volume (pre- or postoperative), the aorta, true lumen (TL), and vertebrae are segmented using deep-learning models. The aortic centerline is then extracted, and the Th8 and Th12 levels are defined by projecting the corresponding vertebral centroids onto the centerline. The aortic segmentation is straightened along the centerline by curved planar reformation (CPR), and cross-sectional measurements are obtained at the Th8 and Th12 levels. The pipeline outputs the morphological parameters: aortic diameter, aortic area, TL area, and TL ratio.

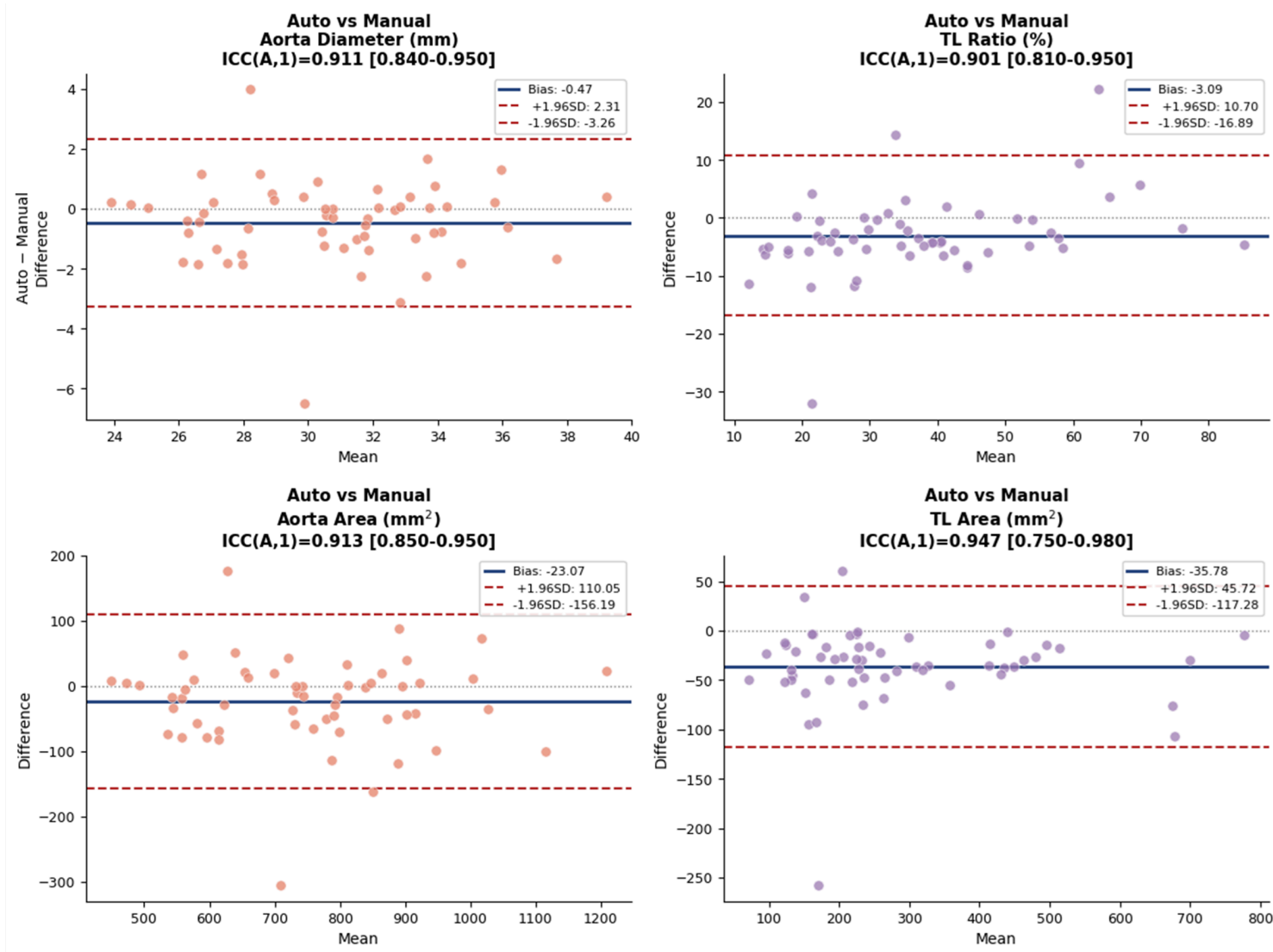


Figure 2

Bland–Altman plots assessing pipeline accuracy (Auto vs. mean of A and B) for four parameters (n = 56). Solid line, mean bias; dashed lines, 95% limits of agreement. $ICC(A,1)$ values with 95% CI are shown.

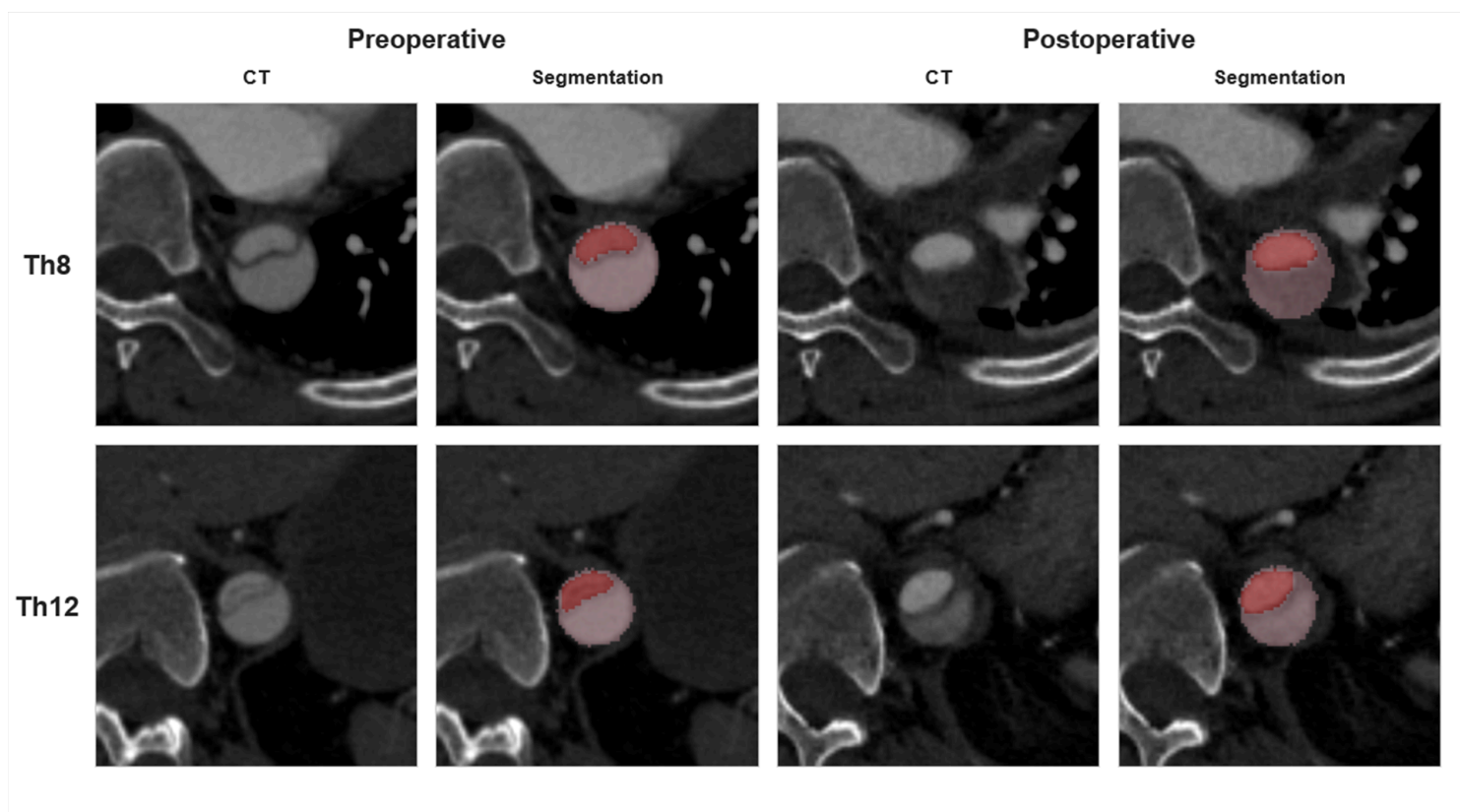


Figure 3

Representative case demonstrating accurate automated segmentation. Curved planar reformation (CPR) cross-sections at the Th8 (top) and Th12 (bottom) levels are shown before (left) and after (right) FET repair. For each level and time point, the CT image is displayed alongside the corresponding automated segmentation. The automated true lumen (red) and aorta (pink) closely matched the underlying anatomy at both levels and time points, before and after surgery. Red, true lumen; pink, aorta.

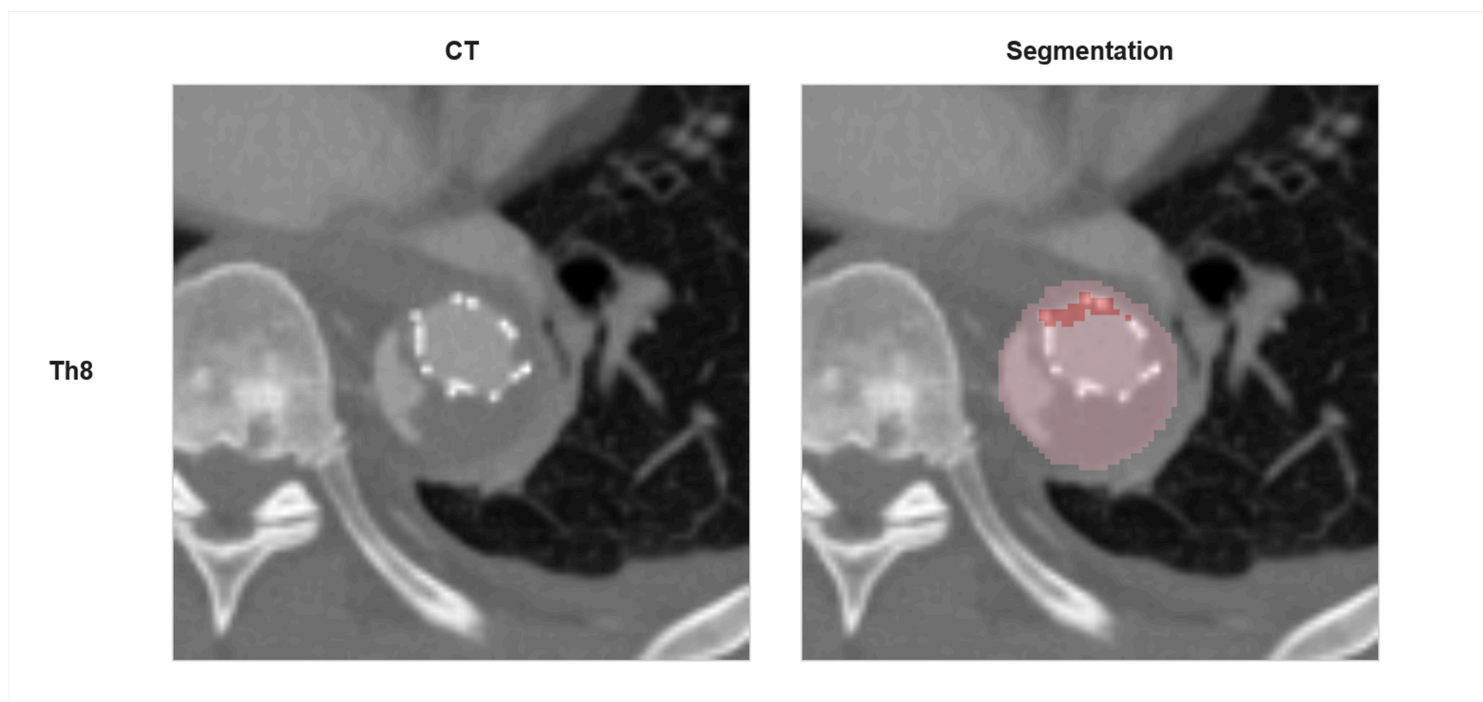


Figure 4

Stent-related segmentation failure in the single case with the largest true lumen discrepancy.

Postoperative CPR cross-section at the Th8 level, where the FET stent graft was located. Left, CT image showing the metallic stent struts as high-attenuation dots; right, the corresponding automated segmentation. The aorta (pink) was correctly segmented, whereas the automated true lumen (red) collapsed to a small region overlapping the stent struts, markedly underestimating the true lumen area. Red, true lumen; pink, aorta.