

Anatomy-Informed Deep Learning for Abdominal Aortic Aneurysm Segmentation

Osamah Sufyan¹, Martin Brückmann¹, Ralph Wickenhöfer², Babette Dellen¹,
and Uwe Jaekel¹

¹ Faculty of Mathematics, Informatics and Technology, University of Applied Sciences
Koblenz, 53424 Remagen, Germany

² Department of Radiology, Herz-Jesu-Krankenhaus, 56428 Dernbach, Germany
sufyan@hs-koblenz.de

Abstract. In CT angiography, the accurate segmentation of abdominal aortic aneurysms (AAAs) is difficult due to large anatomical variability, low-contrast vessel boundaries, and the close proximity of organs whose intensities resemble vascular structures, often leading to false positives. To address these challenges, we propose an anatomy-aware segmentation framework that integrates organ exclusion masks derived from TotalSegmentator into the training process. These masks encode explicit anatomical priors by identifying non-vascular organs and penalizing aneurysm predictions within these regions, thereby guiding the U-Net to focus on the aorta and its pathological dilation while suppressing anatomically implausible predictions. Despite being trained on a relatively small dataset, the anatomy-aware model achieves high accuracy, substantially reduces false positives, and improves boundary consistency compared to a standard U-Net baseline. The results demonstrate that incorporating anatomical knowledge through exclusion masks provides an efficient mechanism to enhance robustness and generalization, enabling reliable AAA segmentation even with limited training data.

Keywords: Image Segmentation · CT angiography · U-Net · Abdominal aneurysms

1 Introduction

An abdominal aortic aneurysm (AAA) is a life-threatening dilation of the aorta associated with a rupture mortality rate exceeding 80% [1]. Accurate and early assessment is therefore critical for clinical decision-making. Computed tomography angiography (CTA) is the clinical reference standard due to its high spatial resolution and ability to visualize the lumen, thrombus, and surrounding anatomy [2]. However, manual segmentation—particularly of the outer aneurysm wall—is time-consuming and prone to inter-observer variability [2].

Deep learning, particularly U-Net-based encoder-decoder architectures, has significantly advanced medical image segmentation [3] and has been successfully applied to AAA analysis [4]. However, these methods typically require large annotated datasets, which are scarce in AAA due to the difficulty of labeling the aneurysm wall and challenges such as low contrast, thrombus heterogeneity, and anatomical variability. Segmentation performance is further degraded near

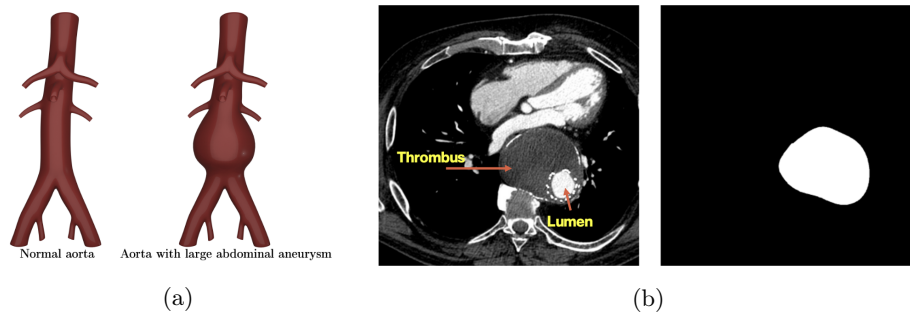


Fig. 1: (a) Comparison between a normal aorta and an aorta with a large aneurysm. (b) Example axial CT slice with lumen, thrombus, and corresponding ground-truth segmentation.

anatomically similar structures (e.g., vertebrae, bowel, kidneys), leading to false positives [5]. While prior work has incorporated anatomical context through multi-organ segmentation, shape priors, or attention mechanisms, these approaches often increase annotation effort, model complexity, or computational cost [5, 6, 7].

TotalSegmentator enables automatic segmentation of 104 anatomical structures from CT images without manual labeling [8], offering a practical source of anatomical context. In this work, we propose an anatomy-aware AAA segmentation framework that integrates organ exclusion masks derived from TotalSegmentator into U-Net training. These masks identify non-vascular regions and penalize anatomically implausible predictions, reducing false positives and improving boundary consistency without requiring architectural changes or additional annotations. Despite being trained on a relatively small dataset, the proposed method achieves high accuracy and improved robustness compared to a baseline U-Net. To our knowledge, this is the first work to leverage automatically generated multi-organ masks as exclusion priors for AAA segmentation.

2 Methods

2.1 Dataset

The dataset consists of abdominal CT angiography (CTA) scans from 20 patients with confirmed abdominal aortic aneurysms, annotated under expert supervision by a senior radiologist. Each scan covers the abdominal aorta from the diaphragmatic hiatus to the iliac bifurcation, with slice thickness between 0.6 and 1.0 mm. All data were anonymized and stored in DICOM format.

Scans were converted to NIfTI using `dcm2niix` and resampled to an isotropic resolution of $1.0 \times 1.0 \times 1.0 \text{ mm}^3$. Axial slices were extracted and resized to 512×512 , with intensities normalized to $[0, 255]$ after applying a soft-tissue window (-150 to 600 HU). Slices not containing the abdominal aorta were automatically excluded using TotalSegmentator-based anatomical filtering.

Manual ground-truth masks delineating the outer aneurysm wall, including thrombus, were generated for all relevant slices. The dataset was partitioned at the patient level to prevent slice-level data leakage. Specifically, 10 patients were used for training, while 2 patients were reserved for validation. An additional 3 patients were allocated for testing. Each patient contributed approximately 150–300 slices. In scenarios where 16 patients were available for training, 4 patients were instead assigned to the test set.

To improve generalization given the limited dataset size, extensive on-the-fly data augmentation was applied, including rotations ($\pm 10^\circ$), flips, translations (± 10 pixels), scaling ($0.9\text{--}1.1\times$), intensity variations, Gaussian noise, and elastic deformations. All preprocessing and augmentation steps were implemented in Python using PyTorch-based pipelines.

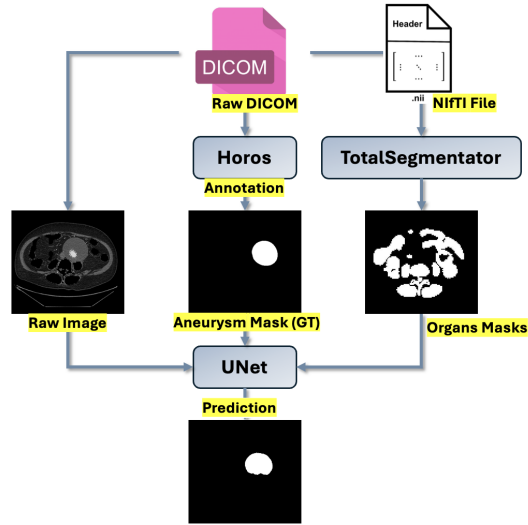


Fig. 2: Pipeline of the proposed method from preprocessing to training.

2.2 Anatomy-Aware Organ Exclusion and Masked Loss

To incorporate anatomical priors, we applied *TotalSegmentator* [8], a 3D U-Net-based framework that automatically segments 104 anatomical structures from CT volumes. From these outputs, non-vascular abdominal structures were merged into a binary *organ exclusion mask*, while the aorta and iliac arteries were excluded to preserve relevant vascular regions. For each axial slice X , the exclusion mask was resampled and aligned with the corresponding aneurysm ground truth Y .

This exclusion mask is converted into a binary allow mask A , where $A_i = 1$ denotes anatomically plausible regions and $A_i = 0$ excluded regions. Let $P, Y, A \in \mathbb{R}^N$ denote flattened prediction, ground truth, and mask ($N = H \cdot W$). Loss computation is restricted to pixels with $A_i = 1$.

The masked Dice and binary cross-entropy (BCE) losses are defined as

$$\mathcal{L}_{\text{Dice}} = 1 - \frac{2(A \cdot P \cdot Y) + \epsilon}{(A \cdot P) + (A \cdot Y) + \epsilon}, \quad (1)$$

$$\mathcal{L}_{\text{BCE}} = \frac{A \cdot \text{BCE}(P, Y)}{A \cdot \mathbf{1} + \epsilon}, \quad (2)$$

where \cdot denotes element-wise multiplication followed by summation. The total loss is

$$\mathcal{L} = w \mathcal{L}_{\text{BCE}} + (1 - w) \mathcal{L}_{\text{Dice}}, \quad w = 0.5. \quad (3)$$

This formulation removes anatomically implausible regions from optimization rather than penalizing them, reducing noisy gradients and improving generalization without requiring architectural changes or additional annotations.

2.3 Network Architecture

We employ a two-dimensional U-Net for aneurysm segmentation, motivated by the limited dataset size, where 2D models are less prone to overfitting and allow larger batch sizes and stronger data augmentation.

The network follows a standard encoder–decoder design with four resolution levels. Each encoder block consists of two 3×3 convolutions with ReLU and batch normalization, followed by 2×2 max-pooling. Feature channels increase from 32 to 256. The decoder mirrors this structure using transposed convolutions for upsampling and incorporates skip connections to preserve fine anatomical details.

A final 1×1 convolution with sigmoid activation produces a single-channel probability map. No architectural modifications were introduced, allowing isolation of the effect of anatomical priors.

Training was performed end-to-end using the Adam optimizer (10^{-4}), a batch size of 3, and 200–300 epochs with early stopping based on validation Dice. Implementation was done in PyTorch on an NVIDIA L40S GPU.

2.4 Baseline Models

To evaluate the proposed anatomy-aware approach, we compare it to a standard U-Net trained without anatomical priors. The baseline uses the same preprocessing, training schedule, optimizer, and data augmentation, ensuring that performance differences arise solely from the incorporation of organ exclusion masks.

The model is trained with a conventional Dice loss without spatial masking, such that all pixels contribute equally to the optimization. As a result, aneurysm segmentation must be learned purely from image intensity and contextual information.

For a fair comparison, both models are initialized consistently and trained under identical conditions, with early stopping based on the validation Dice score. This setup allows improvements to be attributed directly to the anatomy-aware formulation.

3 Results

3.1 Quantitative Evaluation

The proposed anatomy-aware model demonstrated consistently higher segmentation accuracy than the standard U-Net baseline for all test patients. Table 1 summarizes the slice-wise Dice coefficients for both models. For each patient, we report the mean Dice score, the standard deviation for all slices, and the number of slices in which aneurysm tissue was present.

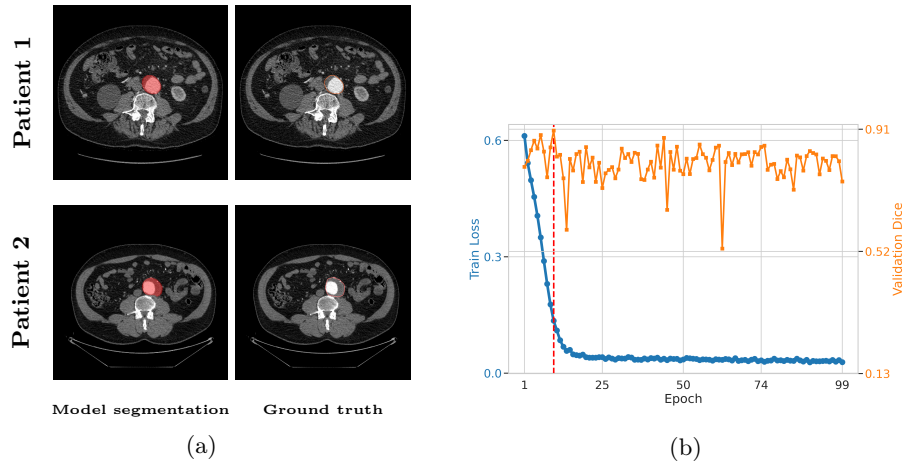


Fig. 3: (a) Qualitative comparison showing two representative slices from different patients. Each row corresponds to a different patient case. The model correctly segments the aneurysm with thrombus. (b) Loss curve and validation DICE vs epoch with the epoch with the best validation result marked with vertical red line

Table 1: Model comparison on the test set.

| Model | Train Patients | Mean Dice | Precision | Recall |
|---------------------|----------------|-------------------------------------|--------------|--------------|
| Baseline U-Net | 10 | 0.637 ± 0.442 | 0.796 | 0.762 |
| Anatomy-Aware U-Net | 10 | 0.914 ± 0.033 | 0.965 | 0.882 |
| Anatomy-Aware U-Net | 16 | 0.906 ± 0.179 | 0.961 | 0.942 |

Overall, our method achieved higher Dice scores on nearly all patients, with the largest improvements observed in anatomically challenging cases where the aneurysm was adjacent to bowel loops, vertebrae, or renal structures. These gains reflect the utility of organ exclusion masks in reducing false activations in regions with intensity profiles similar to aneurysm thrombus.

The proposed anatomy-aware U-Net consistently outperformed the baseline U-Net across different training-set sizes. Table 1 summarizes the quantitative results. With a batch size of 3 and 10 training patients, the baseline U-Net achieved a

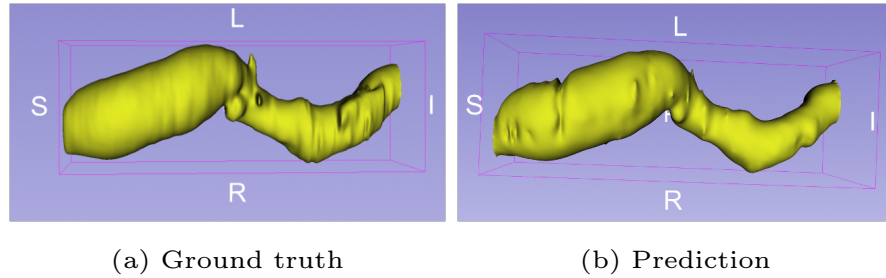


Fig. 4: 3D reconstructed aneurysm from predicted segmentation (right) and corresponding ground truth (left).

mean Dice score of 0.637, while the anatomy-aware U-Net achieved a score of 0.914. Increasing the batch size and training patients further kept the boosted performance, reaching a mean Dice of 0.906 for the anatomy-aware U-Net with 16 training patients and a better recall of 0.942. These results demonstrate that incorporating anatomical priors significantly enhances segmentation accuracy, particularly when training data is limited.

Together, the quantitative and visual results demonstrate that the anatomy-aware training strategy improves segmentation quality, reduces anatomically implausible predictions, and enhances robustness, particularly in scenarios with limited training data.

3.2 Three-Dimensional Reconstruction and Morphological Analysis

We present a pipeline for reconstructing patient-specific 3D vascular geometries and extracting centerline-based morphological descriptors from segmentation masks. Three-dimensional (3D) models were reconstructed from two-dimensional (2D) binary segmentation masks of the aneurysm lumen by aligning slices using imaging metadata and stacking them into a volumetric grid. To reduce voxelization artifacts, the volume was smoothed and converted into a watertight triangular mesh using the marching cubes algorithm, followed by Laplacian and Taubin smoothing to preserve anatomical fidelity (Fig. 4).

Table 2: Predicted and ground-truth geometric measurements for each patient.

| Measurement | Test Patient 1 | Test Patient 2 |
|-------------------------|---------------------|---------------------|
| Maximal Diameter (pred) | 5.3 cm | 4.7 cm |
| Maximal Diameter (true) | 5.8 cm | 5.0 cm |
| Surface Area (pred) | 616 cm ² | 253 cm ² |
| Surface Area (true) | 647 cm ² | 191 cm ² |
| Volume (pred) | 492 cm ³ | 203 cm ³ |
| Volume (true) | 538 cm ³ | 163 cm ³ |

Vascular centerlines were extracted using the Vascular Modelling Toolkit [9]. This method employs a distance transform to assign each voxel its minimum

distance to the vessel wall and defines a traversal cost inversely proportional to this distance. A fast-marching minimal-cost path between inlet and outlet points yields a smooth centerline approximating the medial axis. Local tangent vectors were computed along the centerline, and orthogonal planes were defined at each point to align cross-sections with the local flow direction. Intersecting these planes with the surface mesh produced closed lumen contours, from which anatomical radii and diameters were derived (e.g., maximal inscribed radius, equivalent-area radius, and maximum chord diameter), enabling accurate characterization of asymmetric aneurysm geometries.

These descriptors support computation of clinically relevant metrics such as maximum diameter, surface area, volume, curvature, and shape indices (e.g., aspect and size ratios). Local diameter variations relate to wall stress via Laplace’s law, with enlarged regions experiencing higher mechanical load, and prior studies have linked geometric features—including diameter, curvature, and tortuosity—to aneurysm rupture risk [10, 11].

4 Discussion

The results demonstrate that incorporating anatomical priors via organ exclusion masks improves both the robustness and accuracy of abdominal aortic aneurysm segmentation. While the baseline U-Net performed well in high-contrast regions, it frequently produced false positives in adjacent structures such as bowel, vertebrae, and paraspinal muscles. The anatomy-aware model mitigated these errors by penalizing anatomically implausible predictions, resulting in more consistent and realistic segmentations. It also showed strong performance under limited training data, where the baseline degraded, indicating that anatomical priors act as an effective regularization mechanism that enhances generalization in challenging clinical scenarios.

The proposed method is simple and easily integrable, requiring no architectural modifications or additional annotations beyond aneurysm labels. However, its slice-wise design limits the use of full 3D spatial context, which could be addressed in future work through volumetric models or soft anatomical constraints. Overall, this anatomy-aware approach offers a practical and clinically relevant improvement for AAA segmentation.

Acknowledgments This research has received funding from the Ministry of Science and Health of Rhineland-Palatinate, Germany, and the Debeka Krankenhausesversicherungsverein a.G. through the Forschungskolleg Data2Health.

Informed Consent All patients provided informed consent for the clinically indicated CT examination in accordance with standard clinical practice. Institutional Review Board (IRB) approval was obtained.

References

- [1] Nicola Leone et al. “Systematic Review and Meta-Analysis of the Incidence of Rupture, Repair, and Death of Small and Large Abdominal Aortic Aneurysms under Surveillance”. In: *Journal of Clinical Medicine* **12.21** (2023), p. 6837. DOI: 10.3390/jcm12216837.

- [2] Taehun Kim et al. “Computed Tomography-Based Automated Measurement of Abdominal Aortic Aneurysm Using Semantic Segmentation with Active Learning”. In: *Scientific Reports* **14** (2024), p. 8924. DOI: 10.1038/s41598-024-59735-8.
- [3] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. 2015, pp. 234–241.
- [4] Shubhashree Sahoo et al. “A Deep Learning System for the Automated Detection and Segmentation of Abdominal Aortic Aneurysm from Computed Tomography Angiography”. In: *Vascular and Endovascular Review* **8.16s** (2025), pp. 251–259.
- [5] Yabo Fu et al. “A Review of Deep Learning Based Methods for Medical Image Multi-Organ Segmentation”. In: *Physics in Medicine & Biology* **85** (2021), pp. 107–122. DOI: 10.1016/j.ejmp.2021.05.003.
- [6] Pengfei Zhang, Yuanzhi Cheng, and Shinichi Tamura. “Shape Prior-Constrained Deep Learning Network for Medical Image Segmentation”. In: *Computerized Medical Imaging and Graphics* (2024). DOI: 10.1016/j.compmedimag.2024.102356.
- [7] Jo Schlemper et al. “Attention Gated Networks: Learning to Leverage Salient Regions in Medical Images”. In: *Medical Image Analysis* **53** (2019), pp. 197–207. DOI: 10.1016/j.media.2019.01.012.
- [8] Jakob Wasserthal et al. “TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images”. In: *Radiology: Artificial Intelligence* **5.5** (2023), e230024. DOI: 10.1148/ryai.230024.
- [9] Richard Izzo et al. “The Vascular Modeling Toolkit: A Python Library for the Analysis of Tubular Structures in Medical Images”. In: *Journal of Open Source Software* **3.25** (2018), p. 745. DOI: 10.21105/joss.00745. URL: <https://doi.org/10.21105/joss.00745>.
- [10] Samarth S. Raut et al. “The Role of Geometric and Biomechanical Factors in Abdominal Aortic Aneurysm Rupture Risk Assessment”. In: *Annals of Biomedical Engineering* **41.7** (2013), pp. 1459–1477. DOI: 10.1007/s10439-013-0786-6. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3679219/>.
- [11] Biyun Teng et al. “Combined Curvature and Wall Shear Stress Analysis of Abdominal Aortic Aneurysm: An Analysis of Rupture Risk Factors”. In: *Cardiovascular and Interventional Radiology* **45.6** (2022), pp. 752–760. DOI: 10.1007/s00270-022-03140-z.