A comparative study of machine learning systems in abdominal aortic segmentation

IOANNIS THEOCHARAKIS[1], ELEFTHERIOS KONTOPODIS[1], DIMITRIS ARAMPATZIS[2], EMMANOUIL ATHANASIADIS[1], ILIAS THEODORAKOPOULOS[3], DIMITRIS GLOTSOS[1], PANTELIS ASVESTAS[1], ANASTASIOS RAPTIS[4], CHRISTOS MANOPOULOS[4], KONSTANTINOS MOULAKAKIS[5], JOHN KAKISIS[5], IOANNIS KALATZIS[1], SPIROS KOSTOPOULOS[1*]

[1]Medical Image and Signal Processing Laboratory, Department of Biomedical Engineering, University of West Attica, Egaleo, 12243 Athens, GREECE
[2]Intelligent Data Exploration and Analysis Laboratory, Department of Statistics and Actuarial – Financial Mathematics, University of the Aegean, Karlovasi, 83200, Samos, GREECE
[3]Department of Electrical and Computer Engineering, Democritus University of Thrace, Xanthi, 67100, GREECE
[4]Laboratory of Biofluid Mechanics & Biomedical Technology, School of Mechanical Engineering, National Technical University of Athens, 157 72 Zografos, GREECE
[5]Department of Vascular Surgery, Attikon University Hospital, National and Kapodistrian University of Athens, 106 79 Athens, GREECE

Abstract: - Abdominal aortic segmentation is a critical procedure for the diagnosis and monitoring of pathologies such as aneurysms and strictures. Accurate and efficient imaging of this anatomical region requires the use of advanced computational imaging technologies. In this work, we investigate the application of modern machine learning (ML) tools for automatic segmentation of the abdominal aorta from computed tomography (CT) data without retraining of available ML networks. We compared two ML image segmentation systems against annotations performed by skilled personnel. Nvidia's Medical Open Network for Artificial Intelligence (MONAI) and TotalSegmentator are two state-of-the-art tools that in recent years have become almost universally prevalent in the field of Medical Image Analysis as solutions that combine deep learning algorithms and ML to improve accuracy and efficiency in medical image segmentation. We used 19 CT datasets acquired in the framework of the SAFE-AORTA action, along with annotations of the lumen and intraluminal thrombus. We used the Dice Similarity coefficient (DSC) and the average Hausdorff Distance (HD) to quantitatively assess the performance of the two systems in segmenting the aorta region (lumen and intraluminal thrombus). Preliminary results indicate minimal differences between the two ML tools and their adaptation to expert segmentation remains within acceptable limits (DSC~0.83, HD~3 mm with 95%HD~10 mm). In a next phase, we will explore the possibility of improving the models through dataset enrichment and retraining, with the aim of increasing the accuracy of abdominal aortic segmentation.

Key-Words: - X-ray CT; aortic aneurysm; segmentation; machine learning; neural networks

Received: April 15, 2024. Revised: December 17, 2024. Accepted: January 19, 2025. Published: June 23, 2025.

1 Introduction

The abdominal aorta, as the main artery of the human body, is involved in a multitude of diseases, the most typical being abdominal aortic aneurysm (AAA). The AAA is defined as a permanent, localized dilatation of the aorta that exceeds more than 50% of its normal diameter (usually >3 cm) [1]. It is a serious vascular condition that is the 15th leading cause of death in the United States and the 10th leading cause of death in men over 55 years of age, with more than 15,000 deaths annually. The most serious complication of AAA is rupture, which is accompanied by high mortality rates (50-75%) [2] and is characterized by the absence of symptoms, often described as a ‘silent killer’.

The main imaging techniques used to capture aortic geometry are Computed Tomography (CT) and Magnetic Resonance Imaging (MRI) [3]. CT offers rapid image acquisition but exposes the patient to ionizing radiation and often requires the use of an iodinated contrast agent. In contrast, MRI provides high-resolution images of both the aorta and its wall, does not involve ionizing radiation or iodinated contrast agents, and can offer both functional and

biomechanical information about the aortic structure. However, MRI requires longer acquisition times and is not suitable for patients with metal implants.

Image segmentation is a fundamental pre-processing step for many computational techniques, including artificial intelligence algorithms used in automated data analysis. Manual and semi-automatic segmentation methods are time-consuming, prone to human error, while they are inefficient when processing large volumes of data required for training automated computational workflows. As previously reported in [4], a significant degree of intra-observer variability was observed in the maximal infrarenal aortic diameter measurement on CT images of 2 mm or less. In contrast, a substantially higher interobserver variability was documented, reaching 82%. Automatic segmentation of medical images is a critical tool in the modern diagnostic and therapeutic process. Early and accurate detection and monitoring of AAA requires accurate imaging of the aortic morphology, making its segmentation critical for medical decision-making [5,6].

Previous studies have focused on computer-aided segmentation of the aorta using conventional image processing techniques, such as region-growing based models [7] and model-based approaches [8], usually combining several image processing steps to achieve the final result, which constitutes them susceptible to errors. In recent years, machine learning (ML) systems - and in particular deep learning (DL) techniques - have emerged as highly effective tools for segmenting anatomical structures from images such as CT scans [9,10]. These models are able to identify complex patterns in the data and provide reproducible segmentation results with increased accuracy. Previous review studies [11,12] have examined in detail the state-of-the-art Artificial Intelligence (AI) models for aortic diseases [9], highlighting their contribution in early diagnosis [13], treatment planning [14], and monitoring [15].

However, the variety of proposed methodologies highlights the need for a systematic benchmarking of their performance in terms of accuracy, speed and generalizability [16–18].

In this work, we investigate the application of modern ML tools for automatic segmentation of the abdominal aorta from CT data without training of existing networks. We compared two pre-trained ML image segmentation systems against segmentations performed by experts. Nvidia's Medical Open Network for Artificial Intelligence (MONAI) and TotalSegmentator are two state-of-the-art, freely available software, that in recent years have become almost universally prevalent in the field of Medical Image Analysis as solutions leveraging ML techniques to improve accuracy and efficiency in CT and MR image segmentation.

The remainder of this paper is structured as follows: Section 2 presents the materials and methods used in this study, including the datasets, segmentation frameworks and the evaluation metrics employed. In Section 3, we report and discuss the quantitative results of the segmentation experiments, comparing the performance of the two approaches and analyzing their clinical relevance. Finally, Section 4 summarizes the main conclusions, highlights the limitations of our work, and outlines potential directions for future research.

2 Material and Methods

2.1 Material

Anonymized CT imaging data from 19 abdominal aortic aneurysm (AAA) cases were provided by the Department of Vascular Surgery at the General University Hospital “Attikon”, Athens, Greece. The collection and secondary use of the data were approved by the Ethics Committee of GUH Attikon (Approval No. 168/19-02-2025) and by the Research Ethics Committee of the University of West Attica, Greece: Protocol ID: 17798/11-03-2024.

Subjects (mean age of 70y, mainly men) were scanned following routine protocol for AAA management. The mean maximum AAA diameter of the subjects was 44 mm.

The DICOM images collected for this study are 512×512 pixels while the number of slices per study varies from 59 to 1226 (average: 560). The median slice thickness is 1 mm.

Annotation was conducted by a team of experts, including physicians and medical engineers, using the 3D Slicer software [19]. All segmentations were subsequently reviewed and finalized by vascular surgeons. The segmented region extended along the abdominal aorta from the level of the celiac axis until the bifurcation of the common iliac arteries. The initial parts of the celiac axis, superior mesenteric artery, and renal arteries were also included. Separate masks were created for the blood lumen, the intraluminal thrombus (ILT), and the calcifications.

We then merged the lumen and ILT labels to a single annotation because the evaluated ML systems were trained (a priori) to segment the entire aorta. This merged annotation served as the gold standard for this study and was used to compare the results of the automatic ML segmentation methods. Thus, the segmentation target is the aorta region (lumen and intraluminal thrombus).

2.2 TotalSegmentator

The TotalSegmentator is a DL based tool built on the nnU-Net framework [20], optimized for automatic multi-organ and pathology segmentation from 3D CT images. In the context of AAA detection, it provides high-resolution segmentation of the abdominal aorta, enabling precise measurement of vessel diameter and aneurysmal bulges. The model supports over 100 anatomical structures and is trained on diverse datasets to enhance generalizability. It operates on volumetric data and supports both DICOM and NIfTI formats. With GPU acceleration, TotalSegmentator 12 delivers rapid inference times suitable for clinical use, making it a powerful asset in early AAA identification and risk stratification [21]. We employed the TotalSegmentator software to identify all available anatomical structures and retained only the label corresponding to the aorta.

2.3 MONAI labeler

The MONAI, powered by NVIDIA, is an open-source, PyTorch-based framework designed to accelerate ML development in healthcare imaging. MONAI provides domain-optimized tools for training, deploying, and validating DL models across various modalities, such as CT, MRI, and ultrasound. It supports 3D medical imaging natively and includes advanced utilities for data augmentation, pre-processing, and multi-GPU training. With integration into NVIDIA Clara and support for standardized medical formats (i.e. DICOM, NIfTI), MONAI enables scalable, reproducible research and clinical translation [22]. We used the MONAI software to segment the aorta and retained the associated 'aorta' label.

2.4 Evaluation

We employed two well acceptable and widely used metrics in evaluating medical image segmentation performance; the Dice-Sørensen coefficient (DSC) [23], that quantifies the similarity/overlap between two sets X and Y, as calculated by (1) and the average Hausdorff distance (HD) [24], that considers the distance between segmented boundaries, as calculated by (2):

$$DSC(X,Y) = 2\frac{|X \cap Y|}{|X|+|Y|} \quad (1)$$

$$HD(X,Y) = \frac{1}{2}\left(\frac{1}{|X|}\sum_{x\in X}\min_{y\in Y} d(x,y) + \frac{1}{|Y|}\sum_{y\in Y}\min_{x\in X} d(x,y)\right) \quad (2)$$

where $|X|$ and $|Y|$ are the number of elements in each compared set, and $d(x,y)$ denotes the distance between points x and y. The DSC takes values in the range $[0,1]$, while values over 0.8 are considered acceptable. The HD takes values in the interval $[0,+\infty)$ with values less than 3 mm being considered a very good match.

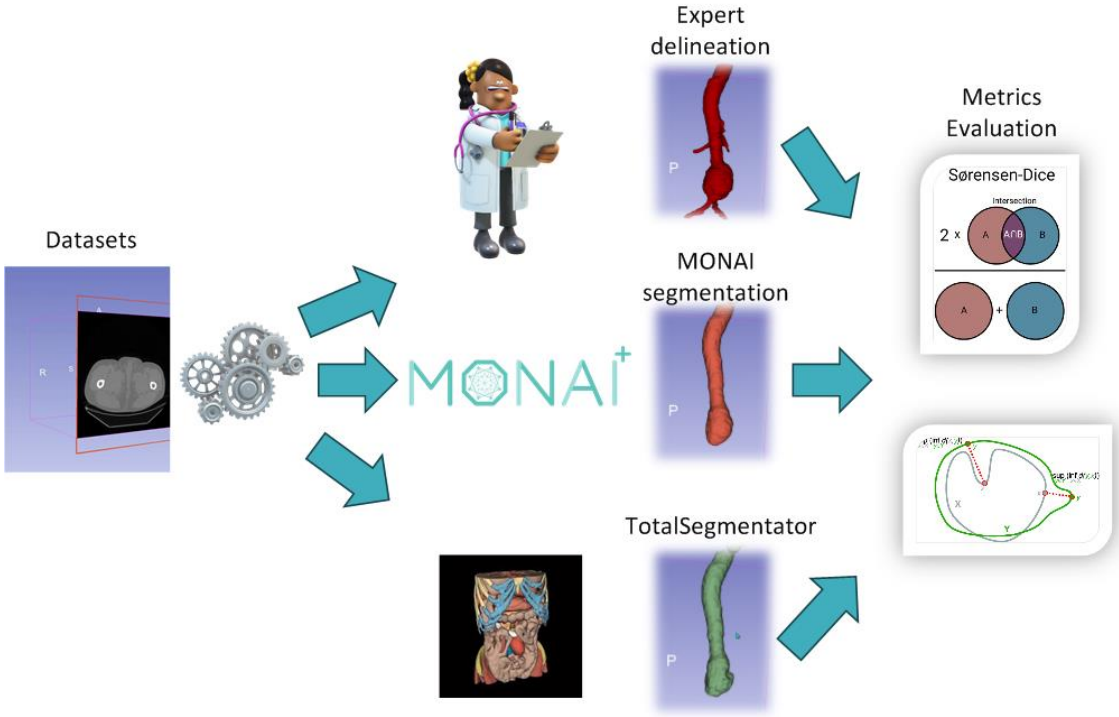

Fig. 1. The flow diagram of the present work

The DSC is simple to interpret, is automatically supported by the ML systems under study, and focuses on the internal coincidence of the partitioned regions. The HD highlights errors in boundaries that may be clinically important, is considered a more rigorous metric, useful in cases where precise contours matter, and is often used in conjunction with Dice for more comprehensive evaluation.

Statistical differences were tested based on the Wilcoxon signed-rank test, the non-parametric alternative of the two sample Student's t-test.

Fig. 1 shows the flow diagram of the present work. The dataset was distributed in three directions; to expert staff/physicians where they manual delineated the lumen, to TotalSegmentator and MONAI systems.

3 Results and Discussion

Fig. 2a presents the overlap of the three segmentations. The experts' delineation (the gold standard) is in red, and the TotalSegmentator and MONAI segmentations are in orange and green respectively. Fig. 2a clearly shows under-segmented regions in the thrombus area, primarily located around the aortic bifurcation along the Z-axis from the legs to the head. Physicians tend to delineate the aortic bifurcation low and not high, while both ML systems ignore the bifurcation (they are not trained to do so) and move the entire width of the aorta upward. This discrepancy distorts our comparison with the region of interest.

Fig. 3 shows a dual graph of the DSC metric and the volume difference for the MONAI system versus the expert's segmentation. As we may observe, the substantial discrepancies in DSC are attributable to the considerable variations in volume (i.e instances of T1-P4, T1-P6, T1-P9, and T2-P15).

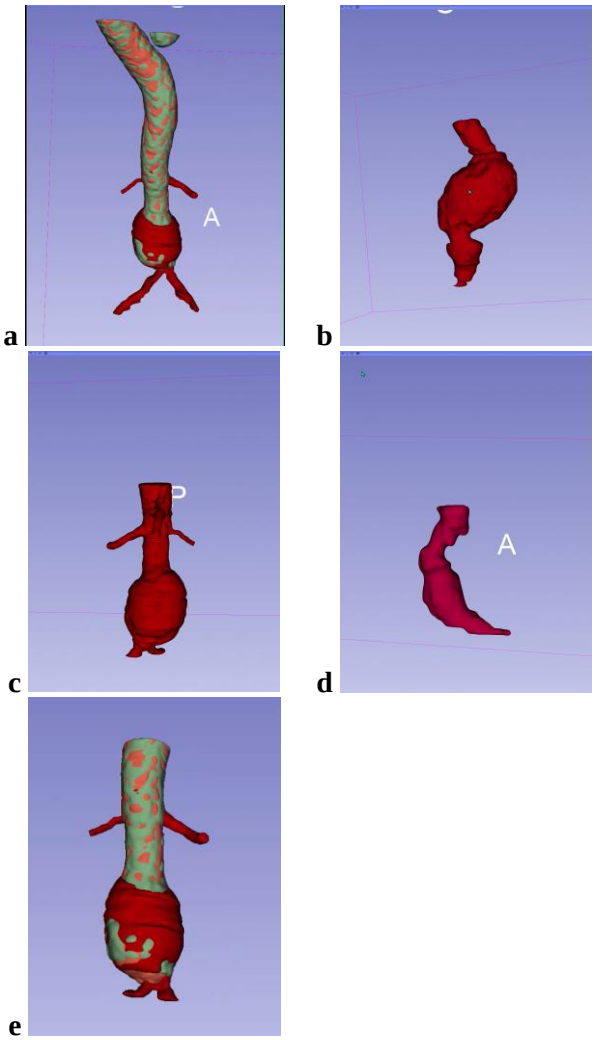

Fig. 2. a. Segmented regions of entire aorta volume, b-d segmentation of cropped aorta volumes, e. Segmented regions of the cropped volume (red: the experts' delineation, orange: and green: TotalSegmentator and MONAI ML systems respectively)

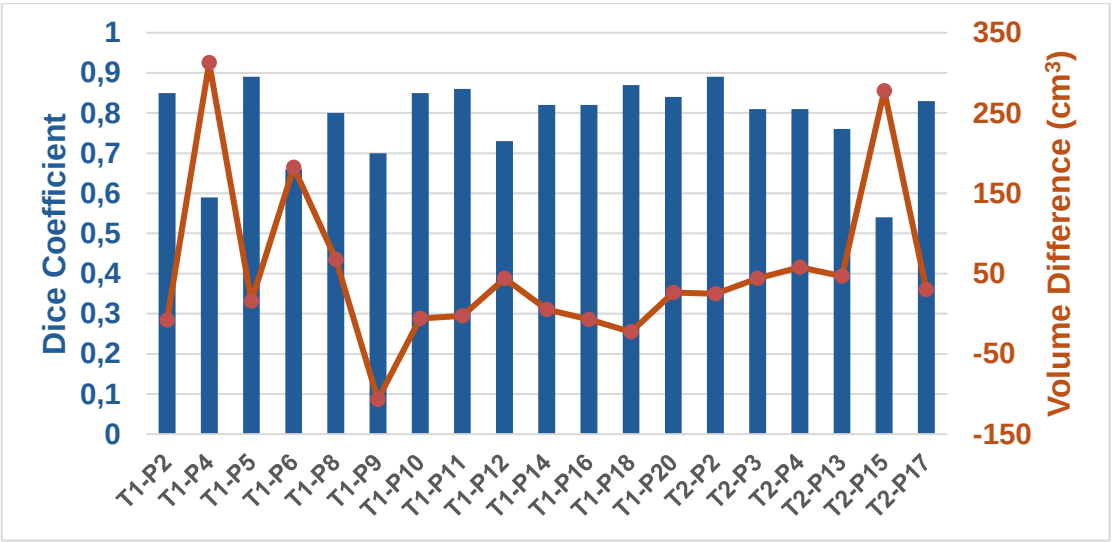

Fig. 3. Dice Coefficient and volume difference for MONAI versus expert segmentation (pre-cropped condition)

Table 1. The systems' performance in terms of Dice-Sørensen coefficient and Hausdorff distance metrics

	Dice-Sørensen coefficient		**Hausdorff Distance in mm (95% HD)**	
	MONAI	**TotalSegmentator**	**MONAI**	**TotalSegmentator**
T1-P2	0.83	0.87	2.02 (6.99)	1.71 (6.15)
T1-P4	0.92	0.92	2.11 (9.96)	2.12 (9.87)
T1-P5	0.91	0.88	0.89 (3.06)	1.06 (2.64)
T1-P6	0.69	0.65	14.56 (49.35)	15.64 (53.57)
T1-P8	0.88	0.89	2.09 (8.96)	1.92 (8.47)
T1-P9	0.64	0.86	4.55 (13.83)	2.02 (8.99)
T1-P10	0.87	0.89	2.01 (7.12)	1.82 (6.80)
T1-P11	0.86	0.83	2.28 (8.40)	2.32 (8.58)
T1-P12	0.91	0.92	1.64 (5.77)	1.50 (5.20)
T1-P14	0.85	0.86	2.05 (7.18)	1.81 (6.05)
T1-P16	0.89	0.92	2.15 (11.10)	1.55 (7.64)
T1-P18	0.88	0.89	2.19 (7.66)	1.90 (7.27)
T1-P20	0.85	0.90	2.10 (6.60)	1.55 (6.57)
T2-P2	0.91	0.89	1.32 (3.61)	1.59 (4.23)
T2-P3	0.84	0.83	1.79 (4.68)	1.78 (4.68)
T2-P4	0.74	0.71	3.93 (15.80)	4.46 (16.92)
T2-P13	0.71	0.71	3.93 (15.80)	3.80 (13.72)
T2-P15	0.78	0.77	3.56 (13.70)	3.80 (13.72)
T2-P17	0.90	0.85	2.29 (12.81)	2.92 (14.85)
Median	**0.86**	**0.87**	**2.11 (8.4)**	**1.90 (7.6)**
Range	**0.28**	**0.27**	**13.67 (46.29)**	**14.58 (50.93)**

For this reason, we cropped the volumes in the area around the aneurysm region (see Fig 2b-d). Fig. 2e presents the segmentations overlap in the cropped volume.

Table 1 presents the systems' performance in terms of DSC and HD metrics for the cropped dataset. The median of DSC metric is similar between the methods tested; 0.86 (Confidence Interval (CI): $[0.79, 0.87]$) for MONAI and 0.87 (CI: $[0.80, 0.88]$) for TotalSegmentator respectively ($p > 0.05$). Consequently, it was determined that the HD metric exhibited negligible variation ($p > 0.05$) between MONAI and TotalSegmentator with median values of 2.11 and 1.9 respectively. In relation to the T1-P6 case, it was determined that both systems' metrics lay outside the confidence intervals, so it may be considered as potential outlier. This is primarily attributable to the presence of two distinct thrombi along the aorta and the under-segmented aortic arch in the gold standard. Moreover, the 95th percentile of the HD distance was calculated (see Table 1 in parenthesis) as a most robust measure of similarity and was found 8.4 for MONAI and 7.6 for TotalSegmentator respectively ($p > 0.05$).

We noticed that the pre-trained models we tested were not always able to accurately identify the lumen and the intraluminal thrombus. This is probably because these ML systems have not trained to discriminate sub-regions. The above-mentioned findings highlight both the potential and limitations of current pre-trained ML segmentation systems when used without additional training or customization for complex vascular structures such as the aorta.

The overall agreement of automated segmentation results with clinical expert annotations was moderate to high, a finding that highlights the importance of domain-specific adaptation and potential fine-tuning. The results also reinforce the need for more granular training datasets that reflect the variability encountered in real-world imaging.

Although the DSC and HD metrics did not show statistically significant differences between the two ML systems, the choice of MONAI as a platform for future use is based on its ability to adapt and train on specialized structures that are absent from general pre-trained models, such as intraluminal thrombus and calcifications. These structures have direct clinical relevance, as their identification and accurate

segmentation can influence decisions on the choice of invasive treatment, rupture risk assessment, or even the design of intimal grafts [25]. To improve accuracy our future work will focus on: Transfer learning from existing pre-trained networks, with further training (fine-tuning) on local data with specific annotations for ILT and calcifications [26]. The flexibility of MONAI allows the integration of the above strategies into a single pipeline, providing a tool that can evolve alongside the clinical needs and specificities of each institution.

We continue our research along with the employment of MONAI as a more comprehensive image processing platform. The initial findings are promising in terms of the lumen and intraluminal thrombus, yet the method is encountering challenges with the calcifications.

4 Conclusion

In the present work two machine learning driven methods compared for their ability to segment the aortic lumen and possible intraluminal thrombus. Both systems are comparable in terms of classical metrics and both encountered difficulties in the clot area due to incomplete training on the specific labels. Although, MONAI system seems more promising for the particular task since it provides the possibility of training from scratch with new labels (like intraluminal thrombus), making refinements in network setup and developing robust solutions for tasks like abdominal aortic aneurysm detection.

The future direction of this research will include two primary aspects. Firstly, the augmentation of the data currently held in the repository will be pursued, accounting for the range of aneurysm sizes, presence of thrombus/calcification and sex distribution. Secondly, a DL network will be trained using these expanded data to ascertain whether there is potential for enhancement in terms of segmentation. In addition to these objectives, the scope of the research will be expanded to encompass the automatic segmentation of multiple aortic structures, including the lumen, calcifications, and intraluminal thrombus.

Contribution of individual authors to the creation of a scientific article (ghostwriting policy)

Conceptualization: I.T., E.K., S.K. methodology: I.T., E.K., E.A., S.K., resources and data collection: A.R., C.M., K.M., J.K., software: I.T., data curation: D.A., I.T., validation: D.G., P.A., writing—review and editing: I,T, E.K., S.K., supervision: I.K. All authors have read and agreed to the published version of the manuscript.

Sources of funding for research presented in a scientific article or scientific article itself

This research is carried out within the framework of the Action “Flagship actions in interdisciplinary scientific fields with a special focus on the productive fabric”, which is implemented through the National Recovery and Resilience Plan Greece 2.0, funded by the European Union –NextGenerationEU (Project ID: TAEDR-0535983).

Creative Commons Attribution License 4.0 (Attribution 4.0 International, CC BY 4.0)

This article is published under the terms of the Creative Commons Attribution License 4.0
https://creativecommons.org/licenses/by/4.0/deed.en_US

References:

[1] Bobadilla JL, Kent KC. Screening for Abdominal Aortic Aneurysms. Adv Surg 2012;46:101–9. https://doi.org/10.1016/j.yasu.2012.03.006.

[2] Palazzuoli A, Gallotta M, Guerrieri G, Quatrini I, Franci B, Campagna MS, Neri E, Benvenuti A, Sassi C, Nuti R. Prevalence of risk factors, coronary and systemic atherosclerosis in abdominal aortic aneurysm: Comparison with high cardiovascular risk population. Vasc Health Risk Manag 2008;4:877–83. https://doi.org/10.2147/VHRM.S1866.

[3] Nienaber CA, Clough RE, Sakalihasan N, Suzuki T, Gibbs R, Mussa F, Jenkins MP, Thompson MM, Evangelista A, Yeh JSM, Cheshire N, Rosendahl U, Pepper J. Aortic dissection. Nat Rev Dis Primer 2016;2:1–18. https://doi.org/10.1038/nrdp.2016.53.

[4] Singh K, Jacobsen BK, Solberg S, Bønaa KH, Kumar S, Bajic R, Arnesen E. Intra- and interobserver variability in the measurements of abdominal aortic and common iliac artery diameter with computed tomography. The Tromsø study. Eur J Vasc Endovasc Surg Off J

Eur Soc Vasc Surg 2003;25:399–407. https://doi.org/10.1053/ejvs.2002.1856.

[5] Raut SS, Chandra S, Shum J, Finol EA. The Role of Geometric and Biomechanical Factors in Abdominal Aortic Aneurysm Rupture Risk Assessment. Ann Biomed Eng 2013;41:1459–77. https://doi.org/10.1007/s10439-013-0786-6.

[6] Philip NT, Patnaik BSV, Sudhir BJ. Hemodynamic simulation of abdominal aortic aneurysm on idealised models: Investigation of stress parameters during disease progression. Comput Methods Programs Biomed 2022;213:106508. https://doi.org/10.1016/j.cmpb.2021.106508.

[7] Seada NA, Hamad S, Mostafa MGM. Automatically Seeded Region Growing Approach for Automatic Segmentation of Ascending Aorta. Proc 10th Int Conf Inform Syst 2016:127–32.

[8] Seada NA, Hamad S, Mostafa MGM. Model-based automatic segmentation of ascending aorta from multimodality medical data. Int J Electr Comput Eng 2016;6:3161–73. https://doi.org/10.11591/ijece.v6i6.12598.

[9] Hahn LD, Baeumler K, Hsiao A. Artificial intelligence and machine learning in aortic disease. Curr Opin Cardiol 2021;36:695–703. https://doi.org/10.1097/HCO.0000000000000903.

[10] Mavridis C, Vagenas TP, Economopoulos TL, Vezakis I, Petropoulou O, Kakkos I, Matsopoulos GK. Automatic Segmentation in 3D CT Images: A Comparative Study of Deep Learning Architectures for the Automatic Segmentation of the Abdominal Aorta. Electronics 2024;13:4919. https://doi.org/10.3390/electronics13244919.

[11] Pepe A, Li J, Rolf-Pissarczyk M, Gsaxner C, Chen X, Holzapfel GA, Egger J. Detection, segmentation, simulation and visualization of aortic dissections: A review. Med Image Anal 2020;65:101773. https://doi.org/10.1016/j.media.2020.101773.

[12] Wang T-W, Tzeng Y-H, Hong J-S, Liu H-R, Wu K-T, Fu H-N, Lee Y-T, Yin W-H, Wu Y-T. Deep Learning Models for Aorta Segmentation in Computed Tomography Images: A Systematic Review And Meta-Analysis. J Med Biol Eng 2024;44:489–98. https://doi.org/10.1007/s40846-024-00881-9.

[13] Spinella G, Fantazzini A, Finotello A, Vincenzi E, Boschetti GA, Brutti F, Magliocco M, Pane B, Basso C, Conti M. Artificial Intelligence Application to Screen Abdominal Aortic Aneurysm Using Computed tomography Angiography. J Digit Imaging 2023;36:2125–37. https://doi.org/10.1007/s10278-023-00866-1.

[14] Xiang D, Qi J, Wen Y, Zhao H, Zhang X, Qin J, Ma X, Ren Y, Hu H, Liu W, Yang F, Zhao H, Wang X, Zheng C. ADSeg: A flap-attention-based deep learning approach for aortic dissection segmentation. Patterns 2023;4:100727. https://doi.org/10.1016/j.patter.2023.100727.

[15] Caradu C, Pouncey A-L, Lakhlifi E, Brunet C, Bérard X, Ducasse E. Fully automatic volume segmentation using deep learning approaches to assess aneurysmal sac evolution after infrarenal endovascular aortic repair. J Vasc Surg 2022;76:620-630.e3. https://doi.org/10.1016/j.jvs.2022.03.891.

[16] Memon AA, Zarrouk M, Ågren-Witteschus S, Sundquist J, Gottsäter A, Sundquist K. Identification of novel diagnostic and prognostic biomarkers for abdominal aortic aneurysm. Eur J Prev Cardiol 2020;27:132–42. https://doi.org/10.1177/2047487319873062.

[17] Lee R, Jarchi D, Perera R, Jones A, Cassimjee I, Handa A, Clifton DA, Bellamkonda K, Woodgate F, Killough N, Maistry N, Chandrashekar A, Darby CR, Halliday A, Hands LJ, Lintott P, Magee TR, Northeast A, Perkins J, Sideso E. Applied Machine Learning for the Prediction of Growth of Abdominal Aortic Aneurysm in Humans. EJVES Short Rep 2018;39:24–8. https://doi.org/10.1016/j.ejvssr.2018.03.004.

[18] Li B, Feridooni T, Cuen-Ojeda C, Kishibe T, de Mestral C, Mamdani M, Al-Omran M. Machine learning in vascular surgery: a systematic review and critical appraisal. NPJ Digit Med 2022;5:7. https://doi.org/10.1038/s41746-021-00552-y.

[19] Fedorov A, Beichel R, Kalpathy-Cramer J, Finet J, Fillion-Robin J-C, Pujol S, Bauer C, Jennings D, Fennessy F, Sonka M, Buatti J, Aylward S, Miller JV, Pieper S, Kikinis R. 3D Slicer as an image computing platform for the Quantitative Imaging Network. Quant Imaging Cancer 2012;30:1323–41. https://doi.org/10.1016/j.mri.2012.05.001.

[20] Isensee F, Jaeger PF, Kohl SAA, Petersen J, Maier-Hein KH. nnU-Net: a self-configuring method for deep learning-based biomedical image segmentation. Nat Methods 2021;18:203–11. https://doi.org/10.1038/s41592-020-01008-z.

[21] Wasserthal J, Breit H-C, Meyer MT, Pradella M, Hinck D, Sauter AW, Heye T, Boll DT, Cyriac J, Yang S, Bach M, Segeroth M. TotalSegmentator: Robust Segmentation of 104 Anatomic Structures in CT Images. Radiol Artif Intell 2023;5:e230024. https://doi.org/10.1148/ryai.230024.

[22] Cardoso MJ, Li W, Brown R, Ma N, Kerfoot E, Wang Y, Murrey B, Myronenko A, Zhao C, Yang D, Nath V, He Y, Xu Z, Hatamizadeh A, Myronenko A, Zhu W, Liu Y, Zheng M, Tang Y, Yang I, Zephyr M, Hashemian B, Alle S, Darestani MZ, Budd C, Modat M, Vercauteren T, Wang G, Li Y, Hu Y, Fu Y, Gorman B, Johnson H, Genereaux B, Erdal BS, Gupta V, Diaz-Pinto A, Dourson A, Maier-Hein L, Jaeger PF, Baumgartner M, Kalpathy-Cramer J, Flores M, Kirby J, Cooper LAD, Roth HR, Xu D, Bericat D, Floca R, Zhou SK, Shuaib H, Farahani K, Maier-Hein KH, Aylward S, Dogra P, Ourselin S, Feng A. MONAI: An open-source framework for deep learning in healthcare 2022.

[23] Zou KH, Warfield SK, Bharatha A, Tempany CMC, Kaus MR, Haker SJ, Wells WM, Jolesz FA, Kikinis R. Statistical validation of image segmentation quality based on a spatial overlap index1: scientific reports. Acad Radiol 2004;11:178–89. https://doi.org/10.1016/S1076-6332(03)00671-8.

[24] Aydin OU, Taha AA, Hilbert A, Khalil AA, Galinovic I, Fiebach JB, Frey D, Madai VI. On the usage of average Hausdorff distance for segmentation performance assessment: hidden error when used for ranking. Eur Radiol Exp 2021;5:4. https://doi.org/10.1186/s41747-020-00200-2.

[25] Schoenhagen P, Tuzcu EM, Apperson-Hansen C. Determinants of arterial wall remodeling during lipid-lowering therapy: Serial intravascular ultrasound observations from the reversal of atherosclerosis with aggressive lipid Lowering therapy (REVERSAL) trial. J Vasc Surg 2007;45:861. https://doi.org/10.1016/j.jvs.2007.02.018.

[26] López-Linares K, Aranjuelo N, Kabongo L, Maclair G, Lete N, Ceresa M, García-Familiar A, Macía I, González Ballester MA. Fully automatic detection and segmentation of abdominal aortic thrombus in post-operative CTA images using Deep Convolutional Neural Networks. Med Image Anal 2018;46:202–14. https://doi.org/10.1016/j.media.2018.03.010.