

Article

The Robust Vessel Segmentation and Centerline Extraction: One-Stage Deep Learning Approach

Rostislav Epifanov ^{1,*}, Yana Fedotova ^{1,†}, Savely Dyachuk ¹, Alexandr Gostev ², Andrei Karpenko ³
and Rustam Mullyadzhyanov ^{1,4}

¹ Department of Mathematics and Mechanics, Novosibirsk State University, Novosibirsk 630090, Russia; i.antonevich@g.nsu.ru (Y.F.); sawel.nsk@gmail.com (S.D.); rustammul@gmail.com (R.M.)

² Meshalkin National Medical Research Center, Novosibirsk 630055, Russia; dr.gostev@gmail.com

³ Scientific Research Institute of Physical-Chemical Medicine, Moscow 119435, Russia; andreikarpenko@rambler.ru

⁴ Institute of Thermophysics, Novosibirsk 630090, Russia

* Correspondence: rostepifanov@gmail.com

† These authors contributed equally to this work.

Abstract

The accurate segmentation of blood vessels and centerline extraction are critical in vascular imaging applications, ranging from preoperative planning to hemodynamic modeling. This study introduces a novel one-stage method for simultaneous vessel segmentation and centerline extraction using a multitask neural network. We designed a hybrid architecture that integrates convolutional and graph layers, along with a task-specific loss function, to effectively capture the topological relationships between segmentation and centerline extraction, leveraging their complementary features. The proposed end-to-end framework directly predicts the centerline as a polyline with real-valued coordinates, thereby eliminating the need for post-processing steps commonly required by previous methods that infer centerlines either implicitly or without ensuring point connectivity. We evaluated our approach on a combined dataset of 142 computed tomography angiography images of the thoracic and abdominal regions from LIDC-IDRI and AMOS datasets. The results demonstrate that our method achieves superior centerline extraction performance (Surface Dice with threshold of 3 mm: $97.65\% \pm 2.07\%$) compared to state-of-the-art techniques, and attains the highest subvoxel resolution (Surface Dice with threshold of 1 mm: $72.52\% \pm 8.96\%$). In addition, we conducted a robustness analysis to evaluate the model stability under small rigid and deformable transformations of the input data, and benchmarked its robustness against the widely used VMTK toolkit.

Keywords: vessel centerline extraction; one-stage centerline reconstruction; vessel segmentation; multitask neural network; computed tomography angiography images; vascular modeling toolkit



Academic Editor: Geoff Dougherty

Received: 7 May 2025

Revised: 18 June 2025

Accepted: 19 June 2025

Published: 26 June 2025

Citation: Epifanov, R.; Fedotova, Y.; Dyachuk, S.; Gostev, A.; Karpenko, A.; Mullyadzhyanov, R. The Robust Vessel Segmentation and Centerline Extraction: One-Stage Deep Learning Approach. *J. Imaging* **2025**, *11*, 209. <https://doi.org/10.3390/jimaging11070209>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Medical visualization techniques are instrumental in the diagnosis and management of vascular pathologies, including aortic aneurysms, femoral artery occlusions, and coronary artery disease. Morphological information obtained from vascular medical images provide critical support for enabling preoperative planning [1,2], image-guided surgical interventions [3,4], and postoperative assessment [4]. For instance, the precise measurements of aortic diameters and lengths at anchoring zones are pivotal for selecting optimal

endograft dimensions and configurations. Notably, inadequate preoperative sizing directly correlates with an elevated risk of type 1A endoleak due to compromised device-wall apposition [5]. Similarly, quantification of maximal aortic diameters, bifurcation angles, and vessel tortuosity serves as a cornerstone for clinical decision-making [6]. In image-guided aortic interventional surgery, centerline extraction is required for preoperative path planning, particularly in cases involving complex aortic anatomy. Furthermore, accurate reconstruction of vascular geometry is essential for patient-specific hemodynamic modeling. Two primary types of mathematical models are utilized: (1) three-dimensional (3D) fluid dynamics models [7], which depend on a precise 3D vascular reconstruction derived from segmentation, and (2) one-dimensional (1D) fluid flow models [8], which necessitate both segmentation and extraction of a centerline graph representing vascular tubular networks.

Computed tomography angiography (CTA) is the most widely adopted procedure for visualizing vessels in human organs *in vivo*. Segmentation and centerline extraction are essential steps in image-based vascular modeling and morphological analysis. In recent years, convolutional neural networks have been widely utilized in medical image processing due to their minimal manual intervention and high accuracy [9]. Li et al. [10] developed a 2D cascaded convolutional network for contour extraction on cross-sectional 2D images, followed by 3D reconstruction of the aortic wall and lumen using pre-extracted centerlines. For the segmentation of large vessels, Dou et al. [11] proposed a 3D supervised deep learning model capable of segmenting volumetric structures, including the heart, aorta, and liver. Cao et al. [12] segmented the entire aorta using a 3D U-Net and subsequently differentiated the true and false lumens via an additional U-Net. Epifanov et al. [13] introduced a single 3D U-Net architecture with a ResNeXt encoder for simultaneous segmentation of the lumen, an aortic wall with thrombotic masses, and calcifications, highlighting that calcifications constitute the most challenging class for segmentation. Tailored augmentation techniques implemented by authors help to mitigate the complexity of their detection. Cepero et al. [14] introduced SeqSeg (sequential segmentation), a deep learning-based algorithm for automatic tracing and segmentation of vascular structures from medical images, which effectively generalizes to unseen vascular structures not included in the training data. Although the aforementioned methods can accurately segment the vessel lumen with reduced reliance on empirical parameters, they do not directly provide centerlines. Vessel tracking is an effective strategy for extracting structural information.

Centerline provides a concise representation of the vessel topology. An ideal centerline extraction algorithm produces points that are closely aligned with the geometric centers of the object cross-sections, accurately capture all true branches of the structure, and avoid generating spurious or false-positive branches. Traditional centerline extraction methods can be broadly categorized into three main classes [15]: minimal cost path, thinning, and tracking methods.

Voronoi diagram-based methods extract centerlines by identifying paths within the Voronoi diagram that minimize the integral of the radii of maximal inscribed spheres along the trajectory. This approach has been specifically implemented in the Vascular Modeling Toolkit (VMTK) [16], a reference package for vascular segmentation and automatic centerline extraction [17–20]. Distance mapping methods [21] are commonly used to construct the shortest path between two points. The process typically begins with computing a distance-from-source map, which encodes the distance from a specified source point to each voxel within the 3D object. The shortest path to the source is then obtained by following the gradient descent of this map from the target point. However, this approach does not guarantee that the path remains centered within the object. A similar strategy can be applied to compute skeletons using distance-from-boundary (DFB) fields [22]. To improve centeredness, a penalty term can be added to the cost function at each node to

steer paths away from the object boundaries. This issue can be effectively addressed by using the DFB field as node weights and constructing a minimum-cost spanning tree build from the DFB field. Nevertheless, methods of this class rely on manually defined features, heuristic rules, and explicit specification of branch endpoints, thereby limiting their degree of automation and generalizability [21].

Topological thinning methods iteratively remove voxels located on the boundary of the shape while preserving both connectivity and overall topology [23]. To enforce the centeredness of the resulting skeleton, voxel removal is typically guided by the voxel distance to the boundary, with those farther from the center removed first. However, centerlines extracted by these methods often come along with spurious branches and usually need post-pruning [24].

Tracking-based techniques extract the medial axis by iteratively estimating and predicting the next optimal voxel to be included in the centerline path. A variety of tracking schemes have been proposed, including image intensity-based methods such as ridge tracking [25], inertia matrix-guided tracking [26], and local maxima approaches [27]. Additionally, model-based techniques were suggested to estimate the medial line, including the use of B-spline curves [28] and super-ellipsoids [29]. As an alternative, Sironi et al. [30] reformulated centerline extraction as a regression problem, while Schneider et al. [31] proposed multivariate Hough regression forests to simultaneously extract binary segmentations and centerlines. In traditional approaches, post-processing is typically employed following centerline extraction to remove spurious branches, reconnect discontinuous segments, and smooth or interpolate the resulting trajectories [32,33]. In summary, despite the wide range of preprocessing, segmentation, and skeletonization techniques proposed in the literature for vascular structures, no universally optimal pipeline exists.

Recent studies have proposed several algorithms based on deep learning methods. Tetteh et al. [34] introduced a neural network with an original convolutional architecture designed to perform vessel segmentation, centerline extraction, and bifurcation detection through three subsequent binary classification tasks. A conceptually similar approach is discussed in Kromm et al. [35], who also proposed a network for retinal vessel segmentation and centerline generation based on a Capsule Network inspired by the Inception architecture [36]. However, these methods usually do not take into account the direct constraint relationship between the two tasks.

The multitask learning (MTL) paradigm has garnered considerable attention in recent years [37]. Rather than developing separate models for individual tasks, MTL unifies them within a single architecture, thereby enhancing performance through knowledge transfer across tasks [38–43]. Most studies in this field have focused on coronary and retinal arteries [44–46]. Shit et al. [47] introduced a centerline Dice (clDice) loss function to realize the constraint of the centerline on blood vessels, but they only focused on the segmentation task, and the centerline is only an intermediate result. In contrast, Pan et al. [48] proposed MSC-Net, a multitask learning framework designed for simultaneous retinal vessel segmentation and centerline extraction, leveraging clDice loss for both tasks. Guo et al. [49] proposed a two-head multitask fully convolutional network (FCN) which simultaneously generates a locally normalized distance map and a list of branch endpoints for coronary arteries. Rouge et al. [50] presented a multitask cascaded network with a U-Net backbone for cerebrovascular segmentation, where the subsequent skeletonization task directly benefits from the segmentation output and may also leverage information from the input image. The application of MTL to diverse vascular structures demonstrates enhancement of performance in joined tasks, underscoring the methodological promise of the approach for this study.

This paper proposes a multitask network to combine the two tasks, blood vessel segmentation and centerline extraction. We specially designed the hybrid network with convolutional and graph layers and loss function to better discover the topological relationship between these two tasks and realize their complementary features. The network enables the end-to-end approach capable of directly predicting connected centerline points with real-valued coordinates, thereby eliminating the post-processing steps required by algorithms that generated centerlines either implicitly presented or without connectivity points. We evaluated the accuracy of our method against existing open-source solutions, including classical and neural network-based approaches, and demonstrated its superior performance. Additionally, we conducted a robustness analysis to assess the algorithm stability under small affine and elastic perturbations of the input image, and benchmarked its robustness against that of the widely used VMTK tool.

2. Materials and Methods

2.1. General Pipeline

We have developed a multitask neural network that simultaneously addresses the interrelated tasks of segmentation and centerline extraction of blood vessels from input CTA images. The schematic diagram is presented in Figure 1. The network architecture is hybrid, incorporating convolutional layers to extract feature maps from CTA images and graph-based layers to transform the extracted features into centerline coordinates. Structurally, the architecture can be divided into three modules: an encoder (Figure 1a), a voxel decoder (Figure 1b), and a centerline decoder (Figure 1c). Each of these modules are split into five stages to make cross connections between modules. The neural network accepts an input CTA image of dimensions $H \times W \times D$, which is processed through the encoder to generate latent features of the CTA image with reduced dimension $H' \times W' \times D'$. To generate a vessel segmentation mask, the latent features are passed to a voxel decoder, which produces an output array of dimensions $H \times W \times D$. Each voxel in this array corresponds to the predicted probability of belonging to the vessel region. Simultaneously, these latent features are also input to the centerline decoder. Within the centerline decoder, sampling layers extract features corresponding to centerline points, using coordinates derived from the preceding stage. Subsequently, graph layers transform these features into three-dimensional coordinates of centerline points. In the first sampling layer, the three-dimensional polyline is randomly initialized in real coordinates. Thus, the proposed neural network enables simultaneous segmentation of lumen and reconstruction of the vessel centerline through a multitask architecture.

2.2. Neural Network Architecture

In selecting the architecture for the voxel encoder, we focused on convolutional architectures. On the one hand, they require less training data to achieve comparable performance than transformer-based models, which is particularly important given the difficulty of obtaining high-quality medical annotations. On the other hand, convolutional architectures are also less demanding in terms of computational resources, which is especially relevant for three-dimensional data. We chose the EfficientNetV2 architecture [51] as the backbone for our encoder because this architecture was selected through an automated architecture search process, optimally balancing performance quality and memory consumption. However, in Appendix A, we also provide performance measurements for other widely used convolutional architectures, such as ResNet [52] and DenseNet [53]. We adopted a layer composition similar to that used in the b0-sized model, as this variant is most suitable for processing data at a resolution of 224 pixels, which closely matches the dimensions of the used data in this study. Additionally, the encoder is divided into five

blocks to facilitate skip connections to the decoders. A detailed schematic of the encoder is illustrated in Figure 1a. Mobile inverted bottleneck convolution (MBCnv) following the original choice consists of convolution layer, squeeze and extraction block and convolution layer. The voxel decoder was constructed utilizing the same layers as those implemented in our prior study [13]. Each block incorporates an interpolation layer, followed by a twice-repeated sequence of a convolutional layer, a batch normalization layer, and a ReLU activation layer. After the interpolation layer, the upsampled feature map is concatenated with the feature map derived from the voxel encoder. This composition layer mirrors the U-Net architecture, which is widely recognized for its performance in segmentation tasks. We employ a voxel encoder–decoder framework to achieve segmentation of the vessel lumen.

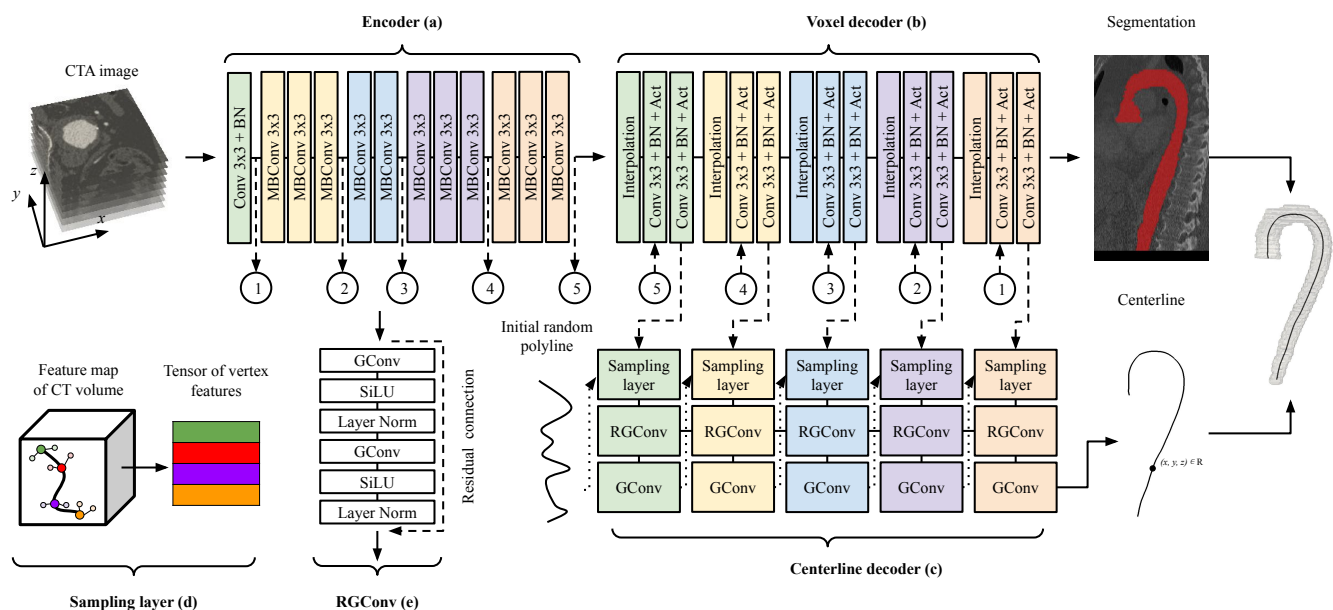


Figure 1. Overview of the proposed hybrid network.

The neural network-based methods for segmentation and centerline extraction discussed earlier rely on voxel-wise target predictions, such as label maps [34] and distance maps [54]. These target formats inherently limit the ability to reconstruct centerlines at subvoxel resolution without the aid of additional post-processing steps, including filtering, smoothing, and point interpolation. However, the use of such post-processing techniques to achieve subpixel accuracy introduces an additional source of error in the localization of centerline points, which is further compounded by potential inaccuracies in the neural network output. To address these limitations, we propose a graph-based architecture that circumvents voxel-level constraints by directly predicting centerline points with continuous real-valued coordinates. During model initialization, a randomly generated polyline serves as the starting structure, enabling the extraction of initial centerline point features via a sampling layer. Appendix B provides an analysis of the network performance depending on the initial polyline initialization. Based on the coordinates of candidate centerline points (pts^e), the sampling layer retrieves their corresponding feature representations (pts^f) from the latent features (x) of the CTA image produced by the voxel decoder. To incorporate contextual information surrounding each point, a convolutional layer predicts the neigh-

bors ($nbrs^c$) of the candidate points, and a unified representation of each point and its local neighborhood ($nbhd^f$) is subsequently constructed as follows:

$$\begin{aligned} pts^f &= \text{grid_sample}(x, pts^c), \\ nbrs^c &= \text{Conv}(pts^f), \\ nbrs^f &= \text{grid_sample}(x, nbrs^c), \\ nbhd^f &= \text{Conv}([pts^f, nbrs^f]). \end{aligned} \quad (1)$$

This neighborhood-aware representation is then processed by a RGConv block to refine the topological structure of the centerline. Each RGConv block consists of a main processing path and a residual connection. The architectural design of these blocks draws inspiration from the decoder structure of U-Net [55]. The main path comprises a repeated sequence of a graph convolution layer [56], an activation layer, and a normalization layer called LayerNorm [57]. The SiLU activation function is employed, in line with its use in the encoder architecture. After passing through the RGConv blocks, the final coordinates of the centerline points are predicted using a single graph convolution layer. These coordinates are subsequently utilized in the sampling layers of the following stages to iteratively refine the centerline geometry.

2.3. Dataset Preparation

In this study, we employ two publicly available datasets: LIDC-IDRI [58] and AMOS [59]. The LIDC-IDRI dataset comprises 1010 CT volumes, whereas the AMOS dataset contains 500 CT volumes with corresponding segmentation masks of the abdominal aorta. The centerline annotations for these datasets are provided in an external dataset [54], which includes centerline annotations for 101 CT volumes from the LIDC-IDRI dataset and 41 CT volumes from the AMOS dataset. Aortic centerlines were annotated by three independent, experienced radiologists, who placed points in the three-dimensional space by aligning their positions across axial, sagittal, and coronal CT projections [54]. Accordingly, a total of 142 CT volumes (101 from the LIDC-IDRI dataset and 41 from the AMOS dataset) were used in this study, as complete triplets of CT images, segmentation masks, and vessel centerline annotations are publicly available only for this subset. Pixel spacing was uniform across all axes at 2 mm. The $x - y$ plane dimensions ranged from 138×138 to 250×250 pixels, while the number of z -planes varied between 113 and 330. The mean number of centerline points per scan was 31.21 ± 7.88 , with an average inter-point distance of 6.03 ± 3.30 mm. The combined dataset of 142 CT images was randomly partitioned into five subsets for cross-validation. To enlarge the effective size of the train subsets in each fold, we augment data with flips, affine and distortion transformations, noising, and blurring [13].

2.4. Neural Network Training

The training of our network was supervised by optimizing a compound loss function incorporating ground truth segmentation masks and centerline points. The total training loss was formulated as a combination of voxel and line losses.

2.4.1. Voxel Loss

In the voxel part of the loss, we lean on a combination of focal and dice losses [60,61] as we have done previously during training the network only to make segmentation [13]:

$$L_{vox}(M^p, M^t) = - \sum_{i \in I} \sum_{k \in K} M_{ik}^t (1 - M_{ik}^p) \log M_{ik}^p - \frac{2}{|K|} \sum_{k \in K} \frac{\sum_{i \in I} M_{ik}^t M_{ik}^p}{\sum_{i \in I} M_{ik}^t + M_{ik}^p}, \quad (2)$$

where M_{ik}^p and M_{ik}^t stand for the predicted and ground truth segmentation probabilities for CT volumes with shape I . Indexes i and k iterate over I shape and K classes, respectively.

2.4.2. Centerline Loss

The centerline loss function is formulated as a composite of three terms. The Chamfer Distance (CD) term minimizes Euclidean distance between predicted and ground truth centerline points:

$$L_{cd}(C^p, C^t) = \sum_{v^p \in C^p} \min_{v^t \in C^t} \|v^p - v^t\|_2 + \sum_{v^t \in C^t} \min_{v^p \in C^p} \|v^t - v^p\|_2, \quad (3)$$

where C^p and C^t denote the predicted and ground truth centerlines. The iteration over v of C corresponds to iterating over points of the corresponding centerline. CD is typically used to measure the similarity between two sets of points [62]. A regularization term on the edges encourages a more uniform edge length of the predicted centerline:

$$L_{elr}(C^p) = \sum_{e^p \in C^p} \|e^p\|_2, \quad (4)$$

where iteration over e^p of C^p means iteration over edges of centerline, e^p means the average edge length. In the third term of centerline loss, we explicitly leveraged the segmentation head to impose penalties on points situated outside the vessel lumen:

$$L_{plr}(C^p, M^p) = \sum_{v^p \in C^p} L_{vox}(\tilde{M}_p(v_p), 1), \quad (5)$$

where $\tilde{\cdot}$ means stop-gradient operator, and $M_p(v_p)$ denotes sampling with trilinear interpolation from M_p by v_p coordinates.

2.4.3. Loss Normalization

To normalize the loss between different network heads, we employed exponential averaging of the voxel and centerline losses:

$$\bar{L} = \alpha \bar{L} + (1 - \alpha)L, \quad (6)$$

where α is equal to 0.9. We selected the value of 0.9 based on our experiments. The final applied loss for training is equal to:

$$L = \bar{L}_{vox} + \bar{L}_{cd} + \bar{L}_{elr} + \bar{L}_{plr}. \quad (7)$$

2.4.4. Training Hyperparameters

We trained the model using the AdamW optimizer [63] with a learning rate of 10^{-3} . The training process comprised 80 epochs, each consisting of 1000 iterations, and was conducted on an NVIDIA (Santa Clara, CA, USA) RTX 3090 Ti GPU with 24 GB of VRAM. The total training time was under six hours, and the inference time per case was less than one minute on a CPU. Prior to input into the neural network, the CTA volumes were windowed and normalized to the range [0, 1]. The windowing process used a level of -200 Hounsfield Units (HU) and a width of 1400 HU. To ensure subpixel accuracy during centerline generation, we fixed the number of points at 96.

3. Results

3.1. Evaluation Metrics

In this section, we conducted a comparative evaluation of our proposed multitask neural network against both traditional classical methods and existing deep learning-based approaches. In alignment with Yaushev et al. [54], the performance of the centerline generation task was evaluated using one-dimensional adaptations of metrics for comparing sets of points: Surface Dice (SD), Hausdorff Distance (HD), and Average Symmetric Surface Distance (ASSD). Explicit definitions of these metrics are provided in Appendix C. For SD, tolerance thresholds of 1 mm and 3 mm were employed, while HD was calculated at the 95th percentile. We emphasized the SD-3 metric, as it most precisely characterizes the accuracy of centerline reconstruction at the voxel resolution level. This corresponds to the commonly accepted error tolerance of 1–2 voxels, or 2–4 mm, based on the 2 mm voxel spacing of the CT images used. To evaluate subvoxel-level reconstruction precision, we additionally utilized SD-1. The ASSD and HD metrics were applied to quantify the average and maximum distances, respectively, between the extracted and ground truth centerlines. Thus, these metrics enable comprehensive evaluation of both the proportion of accurately reconstructed centerline points at voxel and subvoxel resolutions using SD and the absolute geometric deviations using HD and ASSD, which quantify extreme and average spatial discrepancies, respectively. To evaluate the segmentation task performance quality, we employed the Volumetric Dice (VD), the most widely adopted metric for assessing performance in this task [64].

3.2. Comparative Results of Proposed Method with Other Baseline Methods

Comparative results are summarized in Table 1. Appendix D provides examples of ground truth segmentations alongside those produced by our algorithm. Figure 2 illustrates representative examples of the algorithms selected for comparison, applied to thoracic and abdominal aortic segments. For benchmarking against classical approaches, we selected the following centerline extraction methods: a mass centroid-based algorithm (CM1) [65], thinning algorithms implemented in scikit-image (CM2) [66], tracking-based algorithms from the Kimimaro library (CM3) [67], and the minimal cost path algorithm from VMTK (CM4) [68]. These algorithms were chosen due to their prevalence in vascular centerline extraction workflows and the accessibility of their reference implementations. In comparative evaluations, our method demonstrated superior performance across all metrics. The mass centroid algorithm achieved acceptable accuracy in extracting centerlines for unbifurcated abdominal aortas (Figure 2 CM1-A, CM1-B). However, it exhibited limitations in the thoracic region due to spurious branch generation during axial slice analysis near the aortic arch (Figure 2 CM1-C, CM1-D). Notably, the mass centroid algorithm yielded an SD-3 score of 89.17%, the lowest among classical evaluated methods. The thinning algorithm exhibited performance comparable to the mass centroid algorithm. While this method did not face fundamental limitations in reconstructing centerlines within the thoracic aorta, achieving an SD-3 score of 92.72%, it remained prone to generating short false branches under certain aortic configurations (Figure 2 CM2-B, CM2-D), which constrained its overall accuracy. Furthermore, as thinning method inherently produces integer-valued points for centerlines, it underperformed in subvoxel precision, yielding an SD-1 score of 58.63% compared to the mass centroid algorithm 64.89%. The Kimimaro algorithm and the Voronoi diagram-based method demonstrated similar performance, achieving SD-3 scores of approximately 95.8%. However, it is critical to note that the VMTK method failed to generate centerlines for 3% of cases (five cases). Thus, its performance metrics were computed only on successfully processed cases. Despite these failures, VMTK offers distinct advantages, such as reconstructing the centerline as a continuous polyline with explicit topological

connectivity, a feature absent in other classical methods. The connectivity ensures smooth transitions in the tangent vector to the centerline, whereas Kimimaro generates abrupt directional shifts in the tangential vector, necessitating post-processing procedures for points connecting and trajectory smoothing to enable clinical application in vessel geometric feature computation derived from the centerline. Conversely, while VMTK eliminates the need for post-processing, it requires extensive preprocessing, including surface mesh preparation and explicit endpoint specification, to initiate centerline extraction.

Table 1. A comparison of prediction accuracy across different methods.

Task Name	Segmentation		Centerline Extraction			
	VD (%) ↑	SD-1 (%) ↑	SD-3 (%) ↑	HD (mm) ↓	ASSD (mm) ↓	
Mass centroid	CM1	-	64.89 ± 11.18	89.17 ± 6.02	8.68 ± 3.72	8.68 ± 3.72
Thinning	CM2	-	58.63 ± 11.28	92.72 ± 8.86	4.88 ± 4.56	1.43 ± 0.68
Kimimaro	CM3	-	66.93 ± 11.92	95.77 ± 2.80	3.67 ± 1.69	1.17 ± 0.30
VMTK *	CM4	-	66.29 ± 10.73	95.99 ± 3.00	3.28 ± 1.62	1.09 ± 0.27
Tetteh et al. [34]	NM1	-	58.70 ± 7.02	95.24 ± 3.59	7.87 ± 17.09	1.51 ± 1.15
Yaushev et al. [54]	NM2	-	56.00 ± 20.00	97.00 ± 4.00	15.00 ± 16.00	1.4 ± 1.1
Proposed NN		91.09 ± 0.02	72.52 ± 8.96	97.65 ± 2.07	2.74 ± 0.81	0.93 ± 0.21
Proposed NN only centerline extraction		-	70.01 ± 10.37	96.71 ± 2.30	5.85 ± 9.12	1.15 ± 0.52
Proposed NN only mask segmentation		90.83 ± 0.02	-	-	-	-

* denotes that metrics are reported exclusively for cases successfully processed by VMTK method.

Among neural network-based algorithms, we selected the method proposed by Tetteh et al. [34] (NM1), which employs voxel-wise binary classification, and the approach introduced by Yaushev et al. [54] (NM2), which reconstructs centerlines through the prediction of attraction fields. As we were unable to reproduce the NM2 method, we utilized the quality metrics reported in the original study for comparison [54], which were obtained using the same dataset employed in our research. We attribute this failure to reproduce the method to differences in library versions used in the original studies. The NM1 method demonstrated comparable performance quality to the thinning algorithm while sharing the same fundamental limitations. The NM1 method achieved a quality score of 58.71% on the SD-1 metric and 95.24% on the SD-3 metric. Furthermore, when predicting points, the NM1 method exhibited a tendency to generate point clusters (Figure 2 (NM1-B, NM1-C)), likely attempting to improve the probability of accurate alignment with the true centerline. The NM2 method demonstrated the closest quality metrics to our approach, achieving comparable performance on the SD-3 metric but underperforming on the SD-1 metric. We attribute these outcomes to the fact that the NM2 method, unlike ours, requires complex post-processing steps. To obtain the centerline, the authors emphasize the necessity of employing a non-maximum suppression method to refine the predicted point cloud and using Isomap to establish connections within the refined point cloud. Suboptimal parameter selection for these steps may degrade the final quality of the generated centerline, particularly at subpixel resolution. In contrast, our method eliminates the need for any pre- or post-processing steps. This explains the superior SD-1 metric of 72% achieved by our approach, representing a 16% improvement over NM2. This improvement is further reflected in the ASSD metric, with our method attaining 0.93 mm compared to 1.4 mm for the NM2 method.

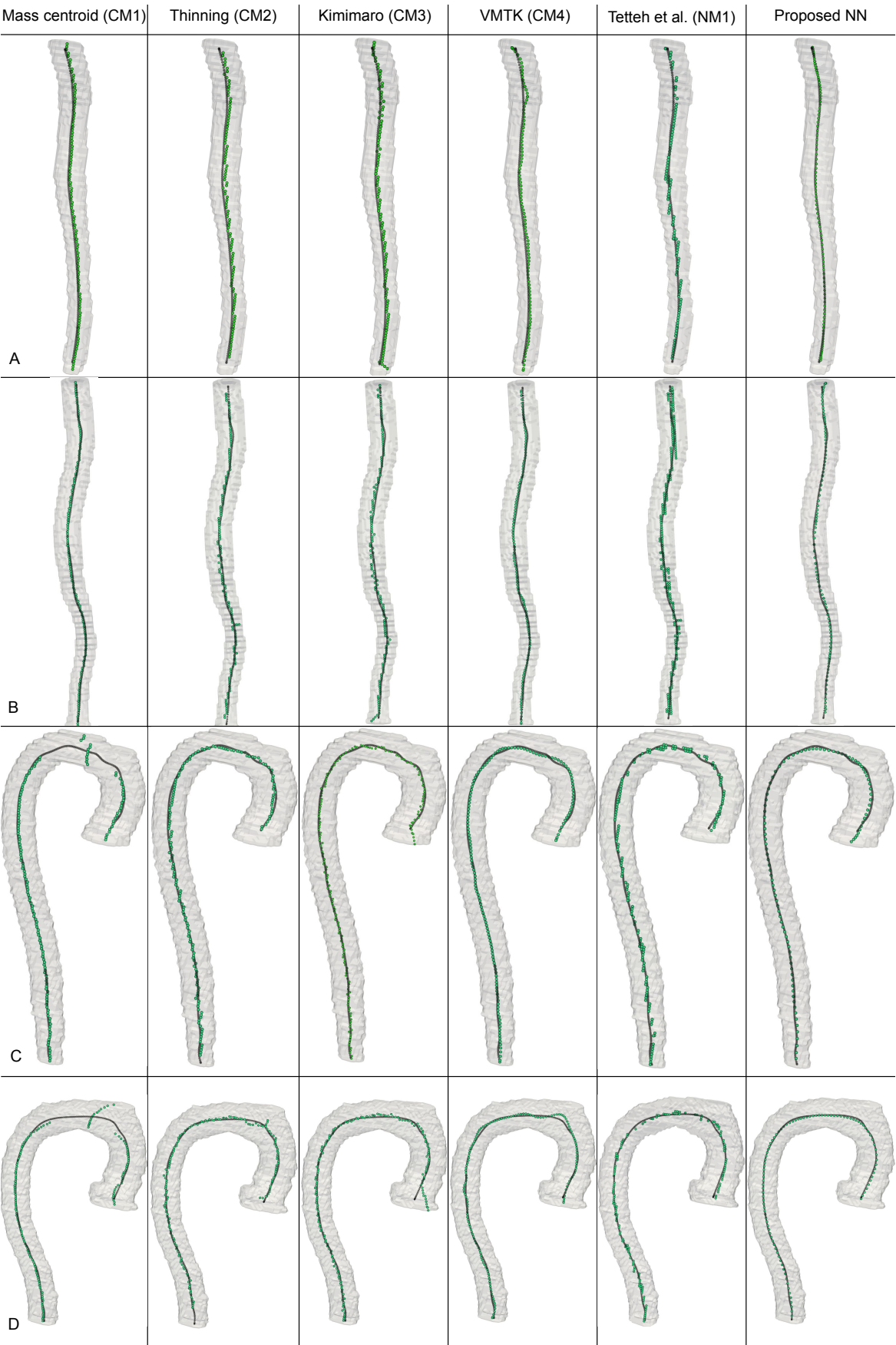


Figure 2. Example centerline reconstructions across different methods [34].

To evaluate the impact of MTL on individual task performance, we trained two single-task networks: one specialized exclusively in the segmentation and the other solely in the centerline extraction. In the segmentation task, despite the high performance of the single-task network, the multitask network outperformed its counterpart by 0.26% on the VD metric. In the centerline extraction task, the multitask network demonstrated improvements of 2.51% on the SD-1 metric and 0.94% on the SD-3 metric compared to the single-task network. In our opinion, these enhancements are driven by two key complementary factors. First, the shared layers between the voxel and centerline decoders implicitly provide the latter with access to spatial information about vessel localization. Second, since deviations of generated centerline points outside vessel boundaries are penalized during training, the usage of implicitly known vessel localization is encouraged during centerline generation. The accounting of vessel localization during the generation is further evidenced by a reduction in the HD from 5.85 mm to 2.74 mm. Thus, joint learning ensures knowledge sharing between tasks, thereby enhancing the performance of each individual task.

3.3. Robustness Comparison of the Proposed Algorithm and VMTK for Centerline Extraction

We compared the robustness of the proposed neural network against a commonly used Voronoi diagram-based centerline extraction method, implemented in the VMTK library. The method is integrated into medical imaging assessment toolkit, such as 3D Slicer [69] and CRIMSON [18]. To assess robustness, we performed a series of experiments involving data augmentations, such as scaling, rotation, and grid distortion, applied to CTA images and corresponding segmentation masks. During image augmentation, we used B-Spline and Gaussian smoothing interpolation techniques [70,71] for CTA scans and corresponding segmentation masks, respectively. Centerlines were generated using both the neural network and the VMTK method. The neural network successfully constructed the centerlines across all dataset, whereas the VMTK method failed to do so in 7% of cases (ten cases). It should be noted that before applying the image transformations, both methods were able to generate the centerlines successfully.

The results of centerline extraction for augmented images are presented in Figure 3. Unlike the neural network, which exhibits robustness to input augmentations, the VMTK method demonstrates sensitivity to deformations of segmentation mask induced by augmentation transformations. In 10% of cases produced by VMTK method, erroneous reconstructions were observed, with artifacts categorized as follows: (1) centerline discontinuities resulting in fragmented paths (Figure 3G); (2) topological inaccuracies, such as spurious branches or loops; and (3) significant deviations of the centerline toward the lumen boundary. These failures primarily originated from minor segmentation artifacts, such as spikes, which were reinforced by slight affine transformations of the segmentation masks used to centerline computation. In contrast, the neural network demonstrated robustness to such input image modifications.

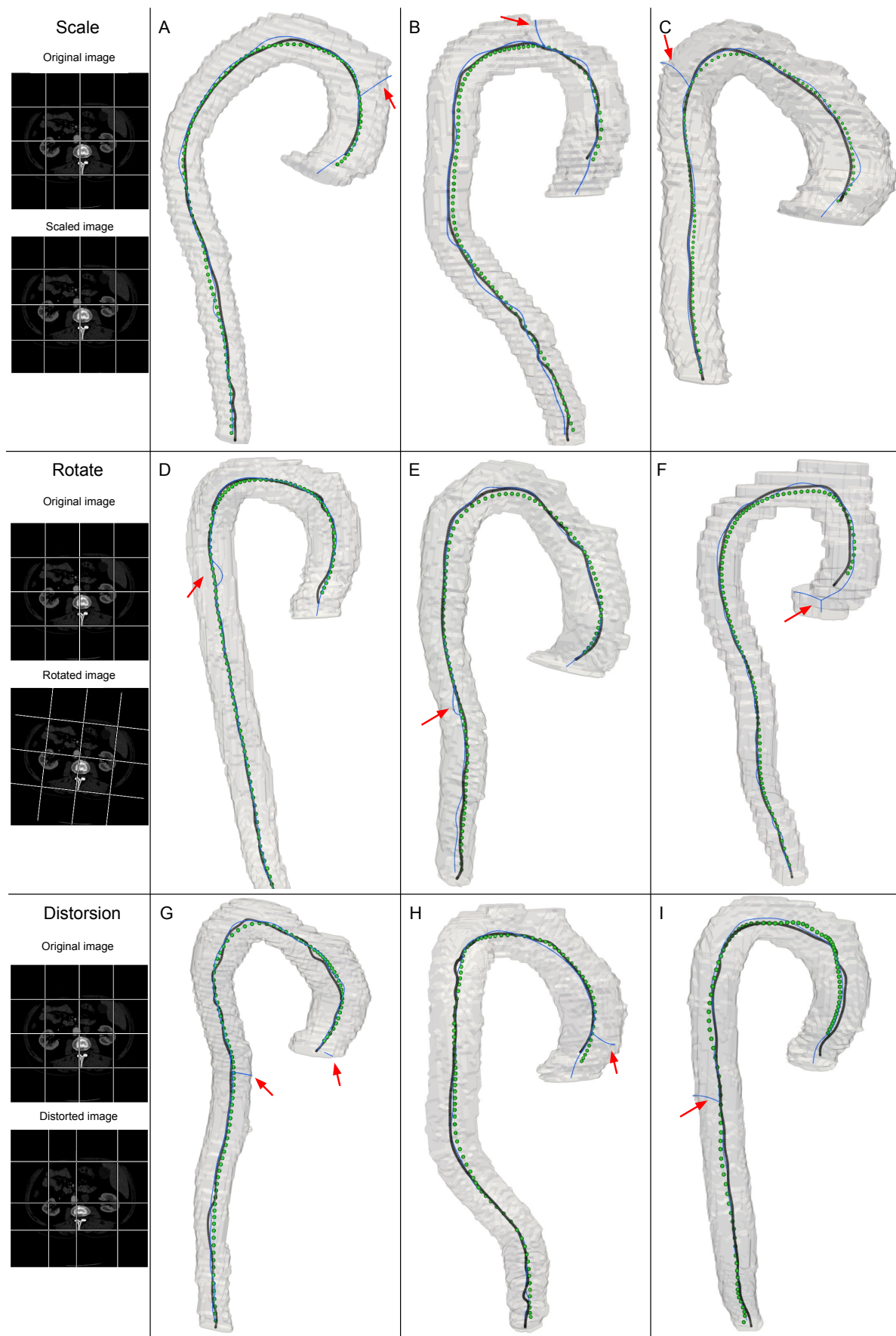


Figure 3. Examples of centerline reconstructions generated by the proposed method and the VMTK for input data subjected to small geometric deformations. White grid on CTA slices used as reference to visualize deformation. Green—proposed method, blue—VMTK, black—ground truth.

3.4. Robustness Evaluation of the Proposed Method Under CTA Image Artifacts

In this section, we evaluate the robustness of the proposed method for centerline extraction when applied to CTA images affected by common types of artifacts. We assess model performance under three artifact scenarios: (i) noise-induced artifacts, (ii) calibration-related artifacts, and (iii) motion-induced artifacts caused by patient movement during CT acquisition (Figure 4).

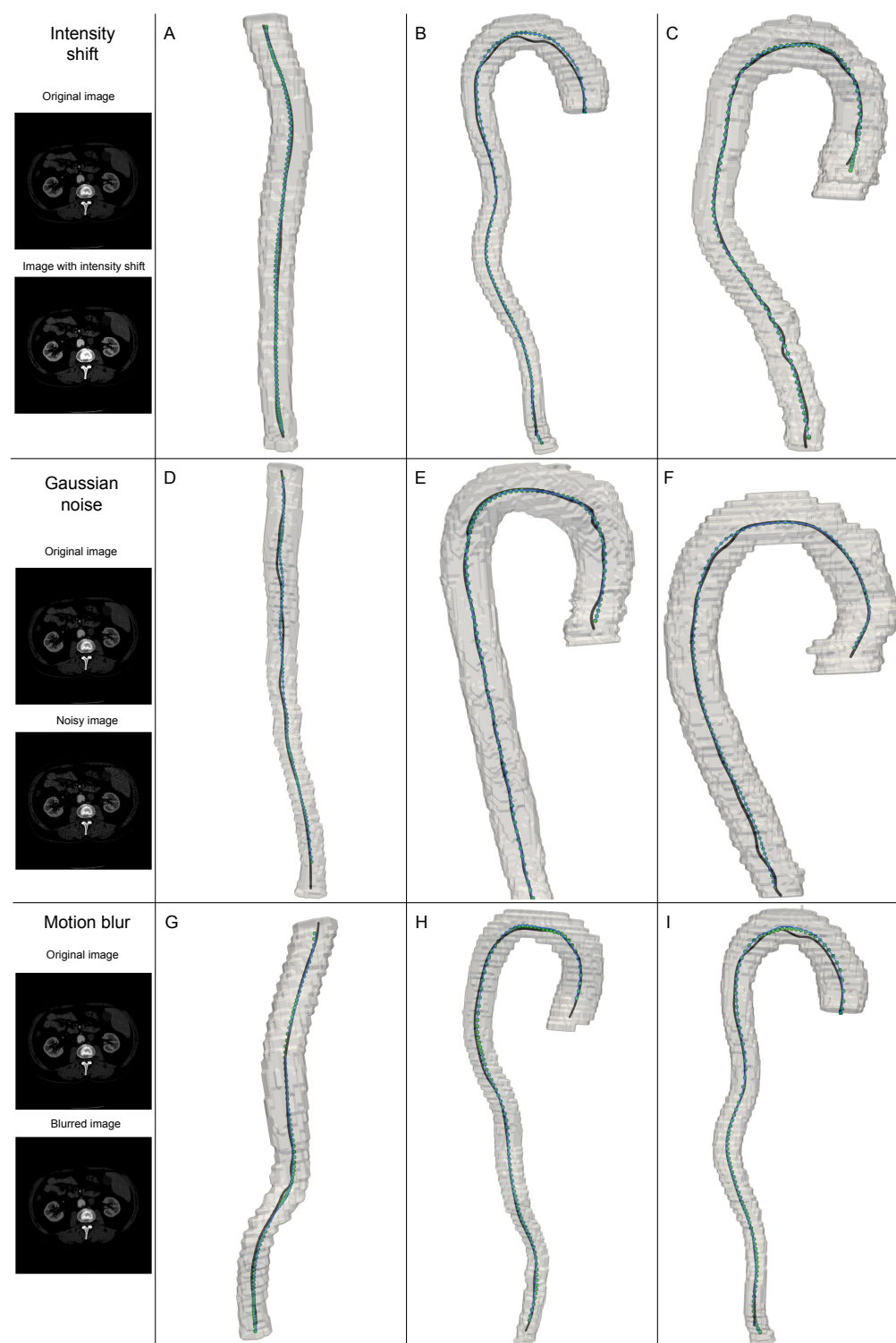


Figure 4. Comparison of centerline reconstructions generated by the proposed method: green—CTA with artifacts, blue—CTA without artifacts, black—ground truth.

To simulate image noise, additive Gaussian noise with zero mean and a variance of 15 HU was applied. Calibration artifacts were modeled by applying a global linear intensity shift sampled uniformly in the range of $[-10, +10]$ HU, simulating scanner calibration errors. Motion artifacts were emulated using linear motion blur with a displacement amplitude ranging from 1 to 3 voxels.

Quantitative results are summarized in Table 2, which compares the model performance under each artifact type to that on unaugmented (original) data. The model exhibited highly stable performance under noise and calibration artifacts: the SD-3 metric remained above 97.6%, and other metrics, including HD and ASSD, showed negligible variation compared to the baseline. The most pronounced impact was observed in the presence of motion blur, which led to a minor decrease of approximately 1.77% in SD-3 accuracy and slightly increased surface distance metrics. Overall, these results demonstrate that the proposed method maintains strong performance across a range of challenging acquisition conditions, exhibiting robustness to typical artifacts frequently encountered in clinical CTA imaging.

Table 2. A comparison of prediction accuracy under CT image artifacts.

Task Name	Segmentation		Centerline Extraction		
	VD (%) $\uparrow$	SD-1 (%) $\uparrow$	SD-3 (%) $\uparrow$	HD (mm) $\downarrow$	ASSD (mm) $\downarrow$
Noised data	91.09 ± 0.02	72.37 ± 8.21	97.61 ± 1.89	2.74 ± 0.94	0.93 ± 0.23
Data with calibration artefacts	91.09 ± 0.03	72.69 ± 7.58	97.67 ± 1.75	2.72 ± 0.89	0.92 ± 0.21
Motion blurred data	91.09 ± 0.03	67.56 ± 11.31	95.88 ± 4.78	3.4 ± 2.33	1.08 ± 0.37
Original data	91.09 ± 0.02	72.52 ± 8.96	97.65 ± 2.07	2.74 ± 0.81	0.93 ± 0.21

4. Discussion and Conclusions

This study presents a novel one-stage multitask network for simultaneous vessel segmentation and centerline extraction, eliminating the need for any pre- and post-processing steps. The proposed method achieved superior accuracy on non-bifurcating vessels compared to both classical and deep learning-based commonly used methods with publicly available implementations. The combination of convolution and graph layers in the decoders enabled direct centerline generation with subvoxel precision, distinguishing our approach from previously developed techniques. Furthermore, we demonstrated the advantages of the multitask learning strategy in terms of both performance and robustness. Our neural network was significantly more robust to image transformations than the classical VMTK method, highlighting its strong potential for fully automated and reliable centerline path tracking.

Previously proposed methods [34,48,49,54] for segmentation and centerline extraction, including multitask neural networks, predominantly rely on intermediate representations or hybrid pipelines rather than end-to-end solutions. Tetteh et al. [34] reduced the centerline detection task to segmentation problem of the input image pixels. In contrast, Yaushev et al. [54] formulated centerline detection as an implicit task. One head of their network generates vessel segmentation masks, while the second head predicts displacement fields from any pixel of the input image to the closest centerline point. Using these predicted displacement fields and segmentation masks, the authors first generate candidate centerline points and then apply a non-maximum suppression technique to refine the point cloud. Although the authors claim that their approach outperforms segmentation-based methods in aortic centerline reconstruction, it requires extensive post-processing steps to obtain final centerline points. Furthermore, none of the aforementioned methods directly generate continuous centerlines. Instead, they produce discrete centerline points that

necessitate subsequent connectivity determination. Thus, the proposed method not only effectively overcomes the identified issue but also demonstrates superior performance compared to existing approaches.

The limitations of our method include its reliance on constructing the centerline for a vascular network with a predefined topology, which reduces its flexibility in cases where the topology of the target vessel exhibits significant variability. However, this issue could potentially be addressed by modifying the method, for instance, by abandoning the assumption of predefined connectivity between centerline segments and instead generating only its points. The graph layers utilized in our network, which are based on polyline edges, can be substituted with graph layers that dynamically define connectivity between points. Several k-nearest neighbor-based approaches have been proposed in the literature, for example, to analyze point clouds [72] and particle detector data [73]. Connectivity between these points could then be established using a minimal spanning tree algorithm [74]. Moreover, this might involve employing a differentiable variant of the minimal spanning tree algorithm [75] alongside regularization terms in the loss that penalize discrepancies between the generated and the ground truth centerline topologies. Unlike existing methods that also extract unconnected centerline points, the described modification of our method explicitly incorporates a connectivity method during training and proposes regularization to ensure topological consistency.

Nevertheless, despite the aforementioned limitations, the proposed method can be effectively applied in tasks where the centerline topology is known a priori. For example, in preoperative path planning [76], where the target trajectory is assumed to be unbranched, or in spinal column centerline extraction [77], where the topology is fixed and exhibits minimal variability, our approach is guaranteed to reconstruct the correct structure without requiring post-processing to remove spurious branches or loops. Given that the method demonstrated high accuracy in vessel segmentation and centerline extraction tasks, as well as robustness to input image perturbations, it holds potential for integration into automated frameworks for vascular morphological analysis in clinical practice.

Author Contributions: Conceptualization, R.E. and Y.F.; methodology, R.E. and Y.F.; software, R.E.; validation, R.E. and Y.F.; formal analysis, R.E. and Y.F.; investigation, R.E.; resources, R.M.; data curation, R.E.; writing—original draft preparation, R.E., Y.F., S.D. and A.G.; writing—review and editing, R.E., Y.F., S.D. and R.M.; visualization, R.E. and Y.F.; supervision, R.M.; project administration, A.G., A.K. and R.M.; funding acquisition, A.G. and R.M. All authors have read and agreed to the published version of the manuscript.

Funding: The work of Epifanov R., Fedotova Y., and Gostev A. is supported by the Russian Science Foundation grant No. 23-75-10047. The work of Dyachuk S. and Mullyadzhyanov R. was supported by the Mathematical Center in Akademgorodok under the Agreement No. 075-15-2025-349 with the Ministry of Science and Higher Education of the Russian Federation.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The datasets are available in their original repository. For AMOS: <https://amos22.grand-challenge.org>, accessed on 6 May 2025, LIDC-IDRI: <https://www.cancerimagingarchive.net/collection/lidc-idri>, accessed on 6 May 2025, and additional annotations: <https://github.com/neuro-ml/curve-detection>, accessed on 6 May 2025. The source code of the proposed method is available in <https://github.com/rostepifanov/paper-vessel-centerline>, accessed on 6 June 2025.

Conflicts of Interest: The authors declare no conflicts of interest.

Appendix A

In this section, we investigate how the performance of our network is affected when the selected EfficientNetV2 b0 encoder is replaced with other commonly used convolutional architectures, such as ResNet and DenseNet. All models were trained using the maximum batch size that could fit into GPU memory, without modifying any other training hyperparameters. The results are presented in Table A1. We observe that segmentation accuracy is relatively consistent across different encoder architectures. However, the performance on the centerline reconstruction task varies significantly between encoders. Although DenseNet-121 was the most memory-intensive encoder, it demonstrated inferior performance compared to ResNet-50, despite being trained with the same batch size. On the SD-3 metric, DenseNet-121 achieved a score of 49.25%, substantially lower than the 66.39% attained by ResNet-50. Smaller ResNet variants allowed for larger batch sizes during training. Specifically, ResNet-18 and ResNet-34 supported maximum batch sizes of four and two, respectively. However, ResNet-34 showed slightly lower SD-3 performance than ResNet-50. However, on absolute metrics such as HD and ASSD, ResNet-34 performed noticeably better, indicating that larger batch sizes reduce the variance of centerline point localization errors. While the performance of ResNet-18 was generally comparable to that of EfficientNetV2 b0, the latter consistently achieved slightly better results across all evaluated metrics.

Table A1. A comparison of prediction accuracy across different encoder architectures.

Task Name Architecture	Max BS ↑	Segmentation		Centerline Extraction		
		VD (%) ↑	SD-1 (%) ↑	SD-3 (%) ↑	HD (mm) ↓	ASSD (mm) ↓
Resnet 18	4	89.36 ± 2.01	69.74 ± 8.67	97.48 ± 2.41	3.01 ± 1.88	1.01 ± 0.24
Resnet 34	2	87.78 ± 4.56	25.81 ± 13.49	63.98 ± 19.01	8.07 ± 3.44	3.13 ± 1.31
Resnet 50	1	86.41 ± 5.18	29.78 ± 15.89	66.39 ± 25.59	21.27 ± 19.21	4.85 ± 4.22
Densenet 121	1	85.18 ± 6.17	19.01 ± 11.08	49.25 ± 23.76	19.64 ± 11.57	5.51 ± 3.74
EffitientnetV2 b0	4	91.09 ± 0.02	72.52 ± 8.96	97.65 ± 2.07	2.74 ± 0.81	0.93 ± 0.21

Appendix B

Based on our observations, random initialization does not significantly affect the final performance of the neural network. To quantify this effect, we trained three instances of the network using the same data fold but with different initializations of the centerline points. The results are presented in Table A2, which reports the mean and variance of the averaged performance metrics of the neural networks. The observed variance is relatively low, indicating that the initialization of the centerline points has a minimal impact on the overall network performance.

Table A2. The mean and variance of the averaged performance metrics of the neural networks with different polyline initialization in the centerline extraction task.

SD-1 (%) ↑	SD-3 (%) ↑	HD (mm) ↓	ASSD (mm) ↓
72.08 ± 0.38	97.46 ± 0.21	2.69 ± 0.04	0.93 ± 0.04

Appendix C

According to [54], we employed one-dimensional versions of the following metrics to evaluate the quality of centerline generation: $SD - \tau$, HD, and ASSD. $SD - \tau$ was defined as follows:

$$SD - \tau(C_p, C_t) = \frac{|\{v_t \in C_t \mid d(v_t, C_p) \leq \tau\}| + |\{v_p \in C_p \mid d(v_p, C_t) \leq \tau\}|}{|C_t| + |C_p|},$$

where $v \in C$ denotes iteration over the points of the centerline, $d(v, C)$ is the Euclidean distance between the point v and the centerline C , and τ is a threshold expressed in millimeters. ASSD was defined as follows:

$$ASSD(C_p, C_t) = \frac{1}{|C_p|} \sum_{v_t \in C_t} d(v_t, C_p) + \frac{1}{|C_t|} \sum_{v_p \in C_p} d(v_p, C_t).$$

HD was defined as follows:

$$HD(C_p, C_t) = \max\left\{\text{perc}_{v_t \in C_t}^{95} d(v_t, C_p), \text{perc}_{v_p \in C_p}^{95} d(v_p, C_t)\right\},$$

where max denotes the maximum value, and perc^{95} refers to the 95th percentile.

Appendix D

We compared the segmentation quality of the proposed method with the segmentation quality achieved by the nnU-Net method [78]. The nnU-Net is a strong baseline method that automatically configures an optimal convolutional neural network architecture for a given segmentation task based on data analysis. For training, nnU-Net employs the Adam optimizer with an initial learning rate of 3×10^{-4} , and automatically reduces the learning rate every 30 epochs twice if no improvement in performance is observed. One epoch is defined as an iteration over 250 training batches, and training is performed for up to a maximum of 1000 epochs. To augment the training dataset, nnU-Net applies various techniques such as random rotations, random scaling, random elastic deformations, gamma correction augmentation, and mirroring. The results of the nnU-Net method and our proposed method are presented in Table A3. Examples of segmentation results are shown in Figure A1. Based on the experimental results, we conclude that the segmentation performance of both methods is comparable.

Table A3. A comparison of prediction accuracy of segmentation across different methods.

	VD (%) $\uparrow$
nnU-net	91.93 $\pm$ 0.02
Proposed NN	91.09 $\pm$ 0.02
Proposed NN only mask segmentation	90.83 $\pm$ 0.02

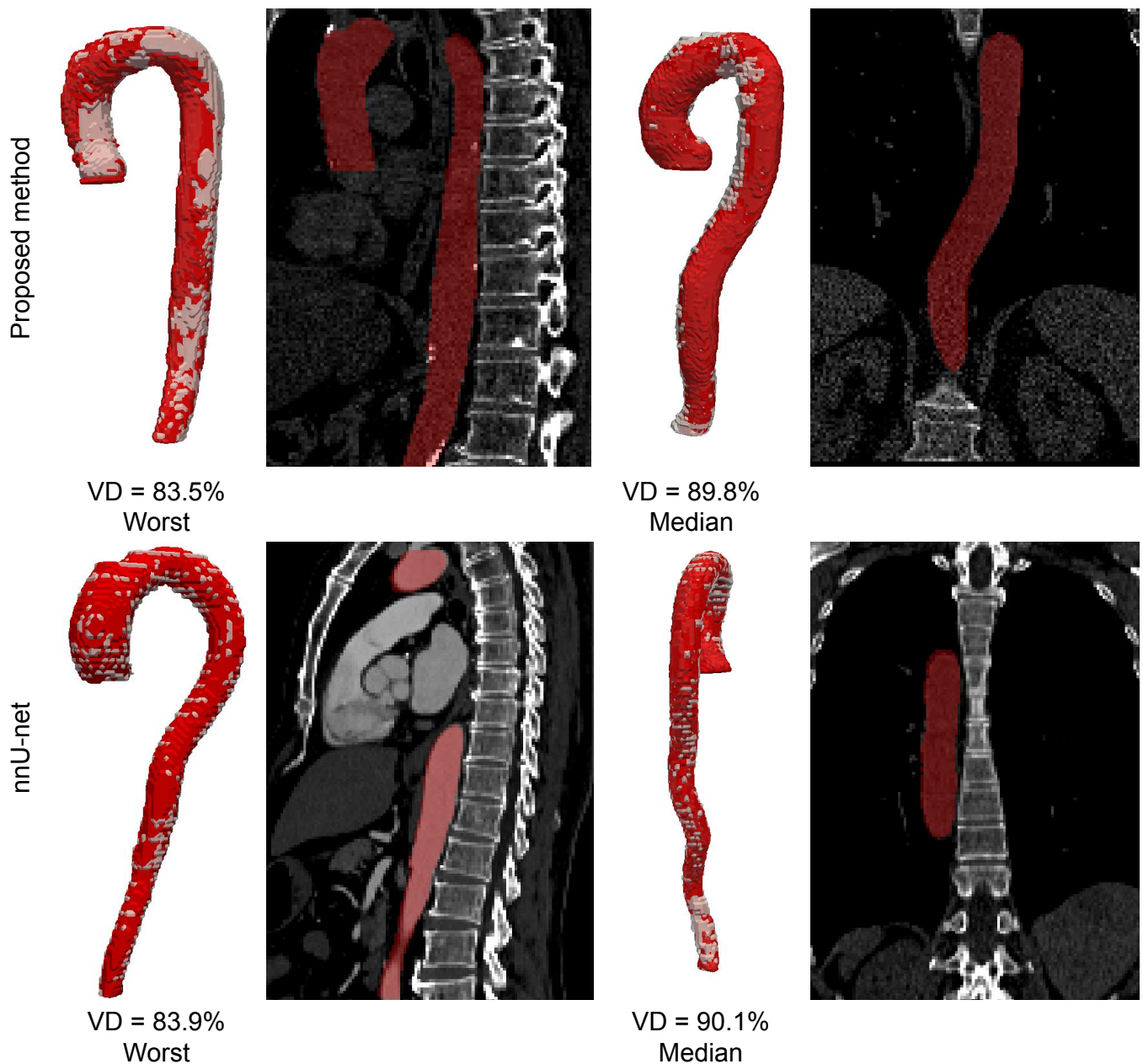


Figure A1. Examples of ground truth segmentation masks (shown in white) and segmentation results (shown in red) for the proposed method and nnU-net.

References

- Boix-Garibo, R.; Uzzaman, M.M.; Bapat, V.N. Review of minimally invasive aortic valve surgery. *Interv. Cardiol. Rev.* **2015**, *10*, 144. [\[CrossRef\]](#)
- Luo, Z.; Cai, J.; Peters, T.M.; Gu, L. Intra-operative 2-D ultrasound and dynamic 3-D aortic model registration for magnetic navigation of transcatheter aortic valve implantation. *IEEE Trans. Med. Imaging* **2013**, *32*, 2152–2165.
- Shimura, S.; Odagiri, S.; Furuya, H.; Okada, K.; Ozawa, K.; Nagase, H.; Yamaguchi, M.; Cho, Y. Echocardiography-guided aortic cannulation by the Seldinger technique for type A dissection with cerebral malperfusion. *J. Thorac. Cardiovasc. Surg.* **2020**, *159*, 784–793. [\[CrossRef\]](#)
- Kumar, P.; Bhatia, M. Role of computed tomography in postoperative follow-up of arterial switch operation. *J. Cardiovasc. Imaging* **2020**, *29*, 1. [\[CrossRef\]](#)
- Buijs, R.V.; Zeebregts, C.J.; Willems, T.P.; Vainas, T.; Tiellu, I.F. Endograft sizing for endovascular aortic repair and incidence of endoleak type 1A. *PLoS ONE* **2016**, *11*, e0158042. [\[CrossRef\]](#)
- Chiesa, R.; Melissano, G.; Zangrillo, A.; Coselli, J.S. *Thoraco-Abdominal Aorta: Surgical and Anesthetic Management*; Springer: Berlin/Heidelberg, Germany, 2011; Volume 783.

7. Fedotova, Y.; Kalachev, I.; Epifanov, R.; Totmina, E.; Borisova, K.; Lysikov, D.; Karpenko, A.; Mullyadzhyanov, R. Association of hemodynamics and morphology with local surface growth of abdominal aortic aneurysm using spatial pattern analysis. *Phys. Fluids* **2025**, *37*, 021919. [[CrossRef](#)]
8. Simakov, S.S. Modern methods of mathematical modeling of blood flow using reduced order methods. *Comput. Res. Model.* **2018**, *10*, 581–604. [[CrossRef](#)]
9. Yagis, E.; Aslani, S.; Jain, Y.; Zhou, Y.; Rahmani, S.; Brunet, J.; Bellier, A.; Werlein, C.; Ackermann, M.; Jonigk, D.; et al. Deep Learning for Vascular Segmentation and Applications in Phase Contrast Tomography Imaging. *arXiv* **2023**, arXiv:2311.13319.
10. Li, Z.; Feng, J.; Feng, Z.; An, Y.; Gao, Y.; Lu, B.; Zhou, J. Lumen segmentation of aortic dissection with cascaded convolutional network. In Proceedings of the Statistical Atlases and Computational Models of the Heart. Atrial Segmentation and LV Quantification Challenges: 9th International Workshop, STACOM 2018, Held in Conjunction with MICCAI 2018, Granada, Spain, 16 September 2018; Revised Selected Papers 9; Springer: Berlin/Heidelberg, Germany, 2019; pp. 122–130.
11. Dou, Q.; Yu, L.; Chen, H.; Jin, Y.; Yang, X.; Qin, J.; Heng, P.A. 3D deeply supervised network for automated segmentation of volumetric medical images. *Med. Image Anal.* **2017**, *41*, 40–54. [[CrossRef](#)] [[PubMed](#)]
12. Cao, L.; Shi, R.; Ge, Y.; Xing, L.; Zuo, P.; Jia, Y.; Liu, J.; He, Y.; Wang, X.; Luan, S.; et al. Fully automatic segmentation of type B aortic dissection from CTA images enabled by deep learning. *Eur. J. Radiol.* **2019**, *121*, 108713. [[CrossRef](#)] [[PubMed](#)]
13. Epifanov, R.U.; Nikitin, N.A.; Rabtsun, A.A.; Kurdyukov, L.N.; Karpenko, A.A.; Mullyadzhyanov, R.I. Adjusting U-Net for the aortic abdominal aneurysm CT segmentation case. *Computer Optics* **2024**, *48*, 418–424. [[CrossRef](#)]
14. Sveinsson Cepero, N.; Shadden, S.C. SeqSeg: Learning Local Segments for Automatic Vascular Model Construction. *Ann. Biomed. Eng.* **2024**, *53*, 158–179. [[CrossRef](#)]
15. Sobiecki, A.; Yasan, H.C.; Jalba, A.C.; Telea, A.C. Qualitative comparison of contraction-based curve skeletonization methods. In Proceedings of the Mathematical Morphology and Its Applications to Signal and Image Processing: 11th International Symposium, ISMM 2013, Uppsala, Sweden, 27–29 May 2013; Proceedings 11; Springer: Berlin/Heidelberg, Germany, 2013; pp. 425–439.
16. Antiga, L.; Piccinelli, M.; Botti, L.; Ene-Iordache, B.; Remuzzi, A.; Steinman, D.A. An image-based modeling framework for patient-specific computational hemodynamics. *Med. Biol. Eng. Comput.* **2008**, *46*, 1097–1112. [[CrossRef](#)]
17. Wang, Z.; Chi, Y.; Huang, W.; Venkatesh, S.K.; Tian, Q.; Oo, T.; Zhou, J.; Xiong, W.; Liu, J. Comparisons of centerline extraction methods for liver blood vessels in ImageJ and 3D slicer. In Proceedings of the 2nd Asia-Pacific Signal and Information Processing Association Annual Summit and Conference, APSIPA ASC 2010, Biopolis, Singapore, 14–17 December 2010; pp. 276–279.
18. Arthurs, C.J.; Khlebnikov, R.; Melville, A.; Marčan, M.; Gomez, A.; Dillon-Murphy, D.; Cuomo, F.; Silva Vieira, M.; Schollenberger, J.; Lynch, S.R.; et al. CRIMSON: An open-source software framework for cardiovascular integrated modelling and simulation. *PLoS Comput. Biol.* **2021**, *17*, e1008881. [[CrossRef](#)]
19. Piccinelli, M.; Veneziani, A.; Steinman, D.A.; Remuzzi, A.; Antiga, L. A framework for geometric analysis of vascular structures: Application to cerebral aneurysms. *IEEE Trans. Med. Imaging* **2009**, *28*, 1141–1155. [[CrossRef](#)]
20. Rezaeitalshmahalleh, M.; Mu, N.; Lyu, Z.; Gemmete, J.; Pandey, A.; Jiang, J. Developing a nearly automated open-source pipeline for conducting computational fluid dynamics simulations in anterior brain vasculature: A feasibility study. *Sci. Rep.* **2024**, *14*, 30181. [[CrossRef](#)]
21. Vigneshwaran, V.; Sands, G.B.; LeGrice, I.J.; Smaill, B.H.; Smith, N.P. Reconstruction of coronary circulation networks: A review of methods. *Microcirculation* **2019**, *26*, e12542. [[CrossRef](#)]
22. Au, O.K.C.; Tai, C.L.; Chu, H.K.; Cohen-Or, D.; Lee, T.Y. Skeleton extraction by mesh contraction. *ACM Trans. Graph. (TOG)* **2008**, *27*, 1–10. [[CrossRef](#)]
23. Bai, X.; Latecki, L.J.; Liu, W.Y. Skeleton pruning by contour partitioning with discrete curve evolution. *IEEE Trans. Pattern Anal. Mach. Intell.* **2007**, *29*, 449–462. [[CrossRef](#)]
24. Danilov, A.; Ivanov, Y.; Pryamonosov, R.; Vassilevski, Y. Methods of graph network reconstruction in personalized medicine. *Int. J. Numer. Methods Biomed. Eng.* **2016**, *32*, e02754. [[CrossRef](#)]
25. Aylward, S.R.; Bullitt, E. Initialization, noise, singularities, and scale in height ridge traversal for tubular object centerline extraction. *IEEE Trans. Med. Imaging* **2002**, *21*, 61–75. [[CrossRef](#)]
26. Hoyos, M.H.; Orłowski, P.; Piątkowska-Janko, E.; Bogorodzki, P.; Orkisz, M. Vascular centerline extraction in 3D MR angiograms for phase contrast MRI blood flow measurement. *Int. J. Comput. Assist. Radiol. Surg.* **2006**, *1*, 51–61. [[CrossRef](#)]
27. Krissian, K.; Malandain, G.; Ayache, N.; Vaillant, R.; Troussset, Y. Model-based detection of tubular structures in 3D images. *Comput. Vis. Image Underst.* **2000**, *80*, 130–171. [[CrossRef](#)]
28. Flasque, N.; Desvignes, M.; Constans, J.M.; Revenu, M. Acquisition, segmentation and tracking of the cerebral vascular tree on 3D magnetic resonance angiography images. *Med. Image Anal.* **2001**, *5*, 173–183. [[CrossRef](#)]
29. Tyrrell, J.A.; di Tomaso, E.; Fuja, D.; Tong, R.; Kozak, K.; Jain, R.K.; Roysam, B. Robust 3-D modeling of vasculature imagery using superellipsoids. *IEEE Trans. Med. Imaging* **2007**, *26*, 223–237. [[CrossRef](#)]
30. Sironi, A.; Türetken, E.; Lepetit, V.; Fua, P. Multiscale centerline detection. *IEEE Trans. Pattern Anal. Mach. Intell.* **2015**, *38*, 1327–1341. [[CrossRef](#)]

31. Schneider, M.; Hirsch, S.; Weber, B.; Székely, G.; Menze, B.H. Joint 3-D vessel segmentation and centerline extraction using oblique Hough forests with steerable filters. *Med. Image Anal.* **2015**, *19*, 220–249. [\[CrossRef\]](#)
32. Chen, Z.; Molloy, S. Automatic 3D vascular tree construction in CT angiography. *Comput. Med. Imaging Graph.* **2003**, *27*, 469–479. [\[CrossRef\]](#)
33. Soltanian-Zadeh, H.; Shahrokni, A.; Khalighi, M.M.; Zhang, Z.G.; Zoroofi, R.A.; Maddah, M.; Chopp, M. 3-D quantification and visualization of vascular structures from confocal microscopic images using skeletonization and voxel-coding. *Comput. Biol. Med.* **2005**, *35*, 791–813. [\[CrossRef\]](#)
34. Tetteh, G.; Efremov, V.; Forkert, N.D.; Schneider, M.; Kirschke, J.; Weber, B.; Zimmer, C.; Piraud, M.; Menze, B.H. Deepvesselnet: Vessel segmentation, centerline prediction, and bifurcation detection in 3-d angiographic volumes. *Front. Neurosci.* **2020**, *14*, 592352. [\[CrossRef\]](#)
35. Kromm, C.; Rohr, K. Inception capsule network for retinal blood vessel segmentation and centerline extraction. In Proceedings of the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), Iowa City, IA, USA, 3–7 April 2020; IEEE: Piscataway, NJ, USA, 2020; pp. 1223–1226.
36. Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; Rabinovich, A. Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Boston, MA, USA, 7–12 June 2015; pp. 1–9.
37. Zhao, Y.; Wang, X.; Che, T.; Bao, G.; Li, S. Multi-task deep learning for medical image computing and analysis: A review. *Comput. Biol. Med.* **2023**, *153*, 106496. [\[CrossRef\]](#) [\[PubMed\]](#)
38. Myronenko, A. 3D MRI brain tumor segmentation using autoencoder regularization. In Proceedings of the International MICCAI Brainlesion Workshop, Granada, Spain, 16 September 2018; Springer: Berlin/Heidelberg, Germany, 2018; pp. 311–320.
39. Chen, S.; Bortsova, G.; García-Uceda Juárez, A.; Van Tulder, G.; De Bruijne, M. Multi-task attention-based semi-supervised learning for medical image segmentation. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part III 22; Springer: Berlin/Heidelberg, Germany, 2019; pp. 457–465.
40. Wang, S.; He, K.; Nie, D.; Zhou, S.; Gao, Y.; Shen, D. CT male pelvic organ segmentation using fully convolutional networks with boundary sensitive representation. *Med. Image Anal.* **2019**, *54*, 168–178. [\[CrossRef\]](#) [\[PubMed\]](#)
41. Xue, Z.; Xin, B.; Wang, D.; Wang, X. Radiomics-enhanced multi-task neural network for non-invasive glioma subtyping and segmentation. In Proceedings of the International Workshop on Radiomics and Radiogenomics in Neuro-Oncology, Shenzhen, China, 13 October 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 81–90.
42. Cao, L.; Li, L.; Zheng, J.; Fan, X.; Yin, F.; Shen, H.; Zhang, J. Multi-task neural networks for joint hippocampus segmentation and clinical score regression. *Multimed. Tools Appl.* **2018**, *77*, 29669–29686. [\[CrossRef\]](#)
43. Ruan, Y.; Li, D.; Marshall, H.; Miao, T.; Cossetto, T.; Chan, I.; Daher, O.; Accorsi, F.; Goela, A.; Li, S. MB-FSGAN: Joint segmentation and quantification of kidney tumor on CT by the multi-branch feature sharing generative adversarial network. *Med. Image Anal.* **2020**, *64*, 101721. [\[CrossRef\]](#)
44. Salahuddin, Z.; Lenga, M.; Nickisch, H. Multi-resolution 3d convolutional neural networks for automatic coronary centerline extraction in cardiac CT angiography scans. In Proceedings of the 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), Nice, France, 13–16 April 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 91–95.
45. Dorobanțiu, A.; Ogorean, V.; Brad, R. Coronary centerline extraction from ccta using 3d-unet. *Future Internet* **2021**, *13*, 101. [\[CrossRef\]](#)
46. Xu, R.; Liu, T.; Ye, X.; Liu, F.; Lin, L.; Li, L.; Tanaka, S.; Chen, Y.W. Joint extraction of retinal vessels and centerlines based on deep semantics and multi-scaled cross-task aggregation. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 2722–2732. [\[CrossRef\]](#) [\[PubMed\]](#)
47. Shit, S.; Paetzold, J.C.; Sekuboyina, A.; Ezhov, I.; Unger, A.; Zhylka, A.; Pluim, J.P.; Bauer, U.; Menze, B.H. clDice-a novel topology-preserving loss function for tubular structure segmentation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 16560–16569.
48. Pan, L.; Zhang, Z.; Zheng, S.; Huang, L. MSC-net: Multitask learning network for retinal vessel segmentation and centerline extraction. *Appl. Sci.* **2021**, *12*, 403. [\[CrossRef\]](#)
49. Guo, Z.; Bai, J.; Lu, Y.; Wang, X.; Cao, K.; Song, Q.; Sonka, M.; Yin, Y. Deepcenterline: A multi-task fully convolutional network for centerline extraction. In Proceedings of the International Conference on Information Processing in Medical Imaging, Hong Kong, China, 2–7 June 2019; Springer: Berlin/Heidelberg, Germany, 2019; pp. 441–453.
50. Rougé, P.; Passat, N.; Merveille, O. Topology aware multitask cascaded U-Net for cerebrovascular segmentation. *PLoS ONE* **2024**, *19*, e0311439. [\[CrossRef\]](#)

51. Tan, M.; Le, Q. Efficientnetv2: Smaller models and faster training. In Proceedings of the International Conference on Machine Learning, Online, 18–24 July 2021; PMLR, 2021; pp. 10096–10106.
52. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
53. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely connected convolutional networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
54. Yaushev, F.; Nogina, D.; Samokhin, V.; Dugova, M.; Petrash, E.; Sevryukov, D.; Belyaev, M.; Pisov, M. Robust Curve Detection in Volumetric Medical Imaging via Attraction Field. In Proceedings of the International Workshop on Shape in Medical Imaging, Marrakesh, Morocco, 7 October 2024; Springer: Berlin/Heidelberg, Germany, 2024; pp. 84–96.
55. Ronneberger, O.; Fischer, P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; proceedings, part III 18; Springer: Berlin/Heidelberg, Germany, 2015; pp. 234–241.
56. Hamilton, W.; Ying, Z.; Leskovec, J. Inductive representation learning on large graphs. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 1025–1035.
57. Ba, J.L.; Kiros, J.R.; Hinton, G.E. Layer normalization. *arXiv* **2016**, arXiv:1607.06450.
58. Armato III, S.G.; McLennan, G.; Bidaut, L.; McNitt-Gray, M.F.; Meyer, C.R.; Reeves, A.P.; Zhao, B.; Aberle, D.R.; Henschke, C.I.; Hoffman, E.A.; et al. The lung image database consortium (LIDC) and image database resource initiative (IDRI): A completed reference database of lung nodules on CT scans. *Med. Phys.* **2011**, *38*, 915–931. [[CrossRef](#)] [[PubMed](#)]
59. Ji, Y.; Bai, H.; Ge, C.; Yang, J.; Zhu, Y.; Zhang, R.; Li, Z.; Zhanng, L.; Ma, W.; Wan, X.; et al. Amos: A large-scale abdominal multi-organ benchmark for versatile medical image segmentation. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 36722–36732.
60. Lin, T.Y.; Goyal, P.; Girshick, R.; He, K.; Dollár, P. Focal loss for dense object detection. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 2980–2988.
61. Cardoso, M.J.; Arbel, T.; Carneiro, G.; Syeda-Mahmood, T.; Tavares, J.M.R.; Moradi, M.; Bradley, A.; Greenspan, H.; Papa, J.P.; Madabhushi, A.; et al. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: Third International Workshop, DLMIA 2017, and 7th International Workshop, ML-CDS 2017, Held in Conjunction with MICCAI 2017, Québec City, QC, Canada, September 14, Proceedings*; Springer: Berlin/Heidelberg, Germany, 2017; Volume 10553.
62. Fan, H.; Su, H.; Guibas, L.J. A point set generation network for 3d object reconstruction from a single image. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 26 July 2017; pp. 605–613.
63. Loshchilov, I.; Hutter, F. Fixing weight decay regularization in adam. *arXiv* **2017**, arXiv:1711.05101.
64. Taha, A.A.; Hanbury, A. Metrics for evaluating 3D medical image segmentation: Analysis, selection, and tool. *BMC Med. Imaging* **2015**, *15*, 29. [[CrossRef](#)]
65. Raut, S.S.; Chandra, S.; Shum, J.; Finol, E.A. The role of geometric and biomechanical factors in abdominal aortic aneurysm rupture risk assessment. *Ann. Biomed. Eng.* **2013**, *41*, 1459–1477. [[CrossRef](#)]
66. Zhang, T.Y.; Suen, C.Y. A fast parallel algorithm for thinning digital patterns. *Commun. ACM* **1984**, *27*, 236–239. [[CrossRef](#)]
67. Silversmith, W.; Bae, J.; Li, P.; Wilson, A. Kimimaro: Skeletonize densely labeled 3D image segmentations. *Zenodo* **2021**, 5539913. [[CrossRef](#)]
68. Izzo, R.; Steinman, D.; Manini, S.; Antiga, L. The vascular modeling toolkit: A python library for the analysis of tubular structures in medical images. *J. Open Source Softw.* **2018**, *3*, 745. [[CrossRef](#)]
69. Pieper, S.; Halle, M.; Kikinis, R. 3D Slicer. In Proceedings of the 2004 2nd IEEE International Symposium on Biomedical Imaging: Nano to Macro (IEEE Cat No. 04EX821), Arlington, VA, USA, 15–18 April 2004; IEEE: Piscataway, NJ, USA, 2004; pp. 632–635.
70. Lehmann, T.M.; Gonner, C.; Spitzer, K. Survey: Interpolation methods in medical image processing. *IEEE Trans. Med. Imaging* **1999**, *18*, 1049–1075. [[CrossRef](#)]
71. Schaerer, J.; Roche, F.; Belaroussi, B. A generic interpolator for multi-label images. *Insight J.* **2014**, *950*, 1–4. [[CrossRef](#)]
72. Wang, Y.; Sun, Y.; Liu, Z.; Sarma, S.E.; Bronstein, M.M.; Solomon, J.M. Dynamic graph cnn for learning on point clouds. *ACM Trans. Graph. (TOG)* **2019**, *38*, 1–12. [[CrossRef](#)]
73. Qasim, S.R.; Kieseler, J.; Iiyama, Y.; Pierini, M. Learning representations of irregular particle-detector geometry with distance-weighted graph networks. *Eur. Phys. J. C* **2019**, *79*, 608. [[CrossRef](#)]
74. Graham, R.L.; Hell, P. On the history of the minimum spanning tree problem. *Ann. Hist. Comput.* **1985**, *7*, 43–57. [[CrossRef](#)]
75. Wang, Z.J.; Zhan, Z.H.; Zhang, J. Distributed minimum spanning tree differential evolution for multimodal optimization problems. *Soft Comput.* **2019**, *23*, 13339–13349. [[CrossRef](#)]
76. Li, Z.; Dankelman, J.; De Momi, E. Path planning for endovascular catheterization under curvature constraints via two-phase searching approach. *Int. J. Comput. Assist. Radiol. Surg.* **2021**, *16*, 619–627. [[CrossRef](#)]

77. Glocker, B.; Zikic, D.; Konukoglu, E.; Haynor, D.R.; Criminisi, A. Vertebrae localization in pathological spine CT via dense classification from sparse annotations. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013: 16th International Conference, Nagoya, Japan, 22–26 September 2013; Proceedings, Part II 16; Springer: Berlin/Heidelberg, Germany, 2013; pp. 262–270.
78. Isensee, F.; Jaeger, P.F.; Kohl, S.A.; Petersen, J.; Maier-Hein, K.H. nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **2021**, *18*, 203–211. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.