

Geometry Before Volume: A Review of Morphology-Aware Learning for Pulmonary Vessel Segmentation

Literature review and conceptual analysis

Abstract

Pulmonary vessel segmentation is often evaluated as a binary labeling problem, yet the anatomy to be recovered is a sparse, branching, multiscale network. A small break in a distal artery can have little influence on regional overlap while disrupting an entire downstream subtree; a slight boundary bias in a large vessel can dominate the voxel count while leaving clinically important connectivity unchanged. This mismatch has encouraged a shift from architecture-only innovation toward morphology-aware learning, in which centerlines, boundary distance, thickness, and continuity become explicit training signals. This review traces that shift from classical vessel enhancement and encoder-decoder segmentation to topology-aware losses, geometry-conditioned convolutions, and adapted foundation models. MorVess is examined as a recent synthesis of these ideas because it jointly predicts vessel masks, distance maps, and thickness maps while introducing limited three-dimensional context into a two-dimensional Segment Anything Model encoder. The analysis identifies three recurring design tensions: local detail versus global tree structure, full volumetric context versus computational economy, and general-purpose representation versus anatomy-specific supervision. Evidence suggests that geometric priors can improve small-branch recovery and diameter consistency, but their value depends on annotation quality, voxel spacing, skeleton stability, and evaluation beyond aggregate Dice. The review concludes with design principles for morphology-aware pulmonary vessel segmentation and with research needs for external validation, uncertainty-aware topology, and clinically grounded morphometry.

Pulmonary Vessels Are Networks, Not Foreground Regions

Segmenting pulmonary vessels from computed tomography is difficult for reasons that are anatomical, physical, and statistical at the same time. Vessels occupy a small portion of the thoracic volume, divide repeatedly, and taper from central trunks to structures near the resolution limit. Their appearance varies with contrast administration, reconstruction kernel, dose, slice thickness, disease, and cardiac phase. Peripheral branches may be only a few voxels wide and can resemble airway walls, fibrotic reticulation, or noise. A model can therefore achieve good average overlap by recovering central vessels while omitting precisely the fine structures needed for quantitative vascular analysis.

Traditional reviews already framed vessel extraction as a special class of image-analysis problem. Moccia et al. (2018) organized blood-vessel segmentation around model-based, tracking, filtering, and learning approaches, noting that datasets and metrics shape what counts as progress. Multiscale filters can enhance tubular patterns, and region-growing or front-propagation methods can exploit continuity, but their response depends on manually selected scales and intensity assumptions. Bifurcations, close artery-vein pairs, pathology, and partial-volume effects weaken those assumptions. Deep learning reduced dependence on handcrafted features, yet it did not make vessel geometry irrelevant.

The U-Net architecture established a practical encoder-decoder pattern in which contracting layers accumulate semantic context while skip connections restore spatial detail (Ronneberger et al., 2015). Its volumetric extension, 3D U-Net, replaced planar operations with three-dimensional counterparts and demonstrated dense prediction from sparse volumetric annotation (Cicek et al., 2016). These models remain influential because their inductive structure fits medical segmentation well. However, a conventional

voxel-wise loss does not know that two adjacent foreground segments should form one branch, that caliber should change smoothly, or that an isolated blob is less plausible than a connected terminal segment. Architecture supplies capacity; morphology-aware supervision tells the capacity what anatomical relationships matter.

Why Regional Overlap Is an Incomplete Objective

The ordinary Dice objective is effective for foreground-background imbalance because it rewards overlap relative to predicted and reference volume. For a tree, however, its weighting is dominated by large-caliber vessels. Removing a short central segment and deleting a long distal subtree may change fewer voxels than slightly shifting the wall of a major artery. Both are errors, but their effects on a reconstructed vascular graph are radically different. This is why visually obvious discontinuities can coexist with apparently competitive overlap scores.

Self-configuring systems such as nnU-Net show how far a carefully engineered baseline can go without a novel anatomy-specific architecture. Isensee et al. (2021) automate decisions about resampling, patch size, network configuration, training, and post-processing. That contribution is important for morphology-aware research because it raises the baseline: a proposed geometric module should be compared with a strong, appropriately configured segmentation pipeline, not an artificially weak U-Net. It also exposes an experimental principle. Improvements attributed to a new loss or prior may actually arise from better preprocessing, larger effective receptive field, or a more suitable patch size unless these factors are controlled.

Pulmonary vessel studies therefore need objectives that distinguish regional, boundary, and network errors. Shit et al. (2021) introduced centerline Dice, or clDice, by measuring the relationship between a predicted skeleton and the opposite mask. Its differentiable soft-skeleton form can be used as a training loss, making continuity more visible than it is under regional overlap alone. clDice does not encode every anatomical rule: a connected false shortcut can still score favorably, and skeletonization can be unstable for thick or noisy structures. Nevertheless, it changed the methodological vocabulary by making topology a trainable and reportable target.

Geometry as an Auxiliary Prediction Task

Another strategy asks the network to predict continuous geometric fields in addition to a binary mask. Wang et al. (2020) proposed Deep Distance Transform for tubular structures in CT. A distance map assigns foreground voxels values related to their distance from the surface, implicitly representing centerline location and local radius without requiring a brittle binary skeleton label at every training step. The predicted distance information then guides mask refinement. This approach illustrates why auxiliary regression can be easier to optimize than direct classification of a one-voxel-wide centerline: neighboring voxels receive smoothly varying targets rather than opposite labels.

MorVess extends this logic by using two complementary fields. Mao et al. (2026) generate a vessel distance map that emphasizes distance to the boundary and a vessel thickness map obtained by propagating centerline radius through the vessel volume. The network jointly predicts these fields and the semantic mask. The distance field supplies local boundary guidance, while the thickness field encodes a more global expectation that caliber should vary coherently along the vascular tree. A composite loss combines cross-entropy, Dice, clDice, distance regression, and thickness regression. In the reported ablation sequence, adding distance supervision reduced the 95th-percentile Hausdorff distance, while thickness supervision improved centerline overlap. The result supports a division of labor between local edge geometry and branch-scale caliber structure.

The conceptual contribution is stronger than simply adding two output heads. The derived targets transform an annotation into several views of the same anatomy. The mask says which voxels belong to a vessel; the distance field locates the wall; the thickness field relates lumen voxels to a medial representation; and clDice

tests centerline continuity. Agreement among these tasks can regularize feature learning. It can also expose contradictions. A predicted mask with an implausibly abrupt diameter transition may have acceptable overlap but disagree with thickness supervision.

There are limitations. All derived priors inherit annotation errors. If a distal branch is absent from the reference, its distance and thickness fields will also be absent, encouraging the model to suppress a potentially real structure. Skeleton-based thickness can be unstable near bifurcations or in irregular lumens. Resampling must respect physical voxel spacing, because a distance measured in voxel units changes meaning across scanners. Morphology-aware targets should therefore be generated in physical coordinates where possible, visually audited, and tested for sensitivity to annotation perturbations.

Learning Features That Follow Curves

Geometric supervision changes the objective, while specialized operators change the representation. Qi et al. (2023) proposed dynamic snake convolution, whose sampling pattern adapts to slender and tortuous structures. Combined with multi-view fusion and a continuity constraint, the operator seeks features along a curvilinear path rather than a fixed rectangular grid. This reflects an important insight: the informative neighborhood of a vessel voxel often lies along the branch direction, not isotropically around it.

Edge-focused networks pursue a related goal. Xia et al. (2022) developed an edge-reinforced approach for three-dimensional vessel-like structures, using boundary attention to improve fine detail. Boundary supervision is especially relevant where partial-volume effects blur the lumen wall. Yet emphasizing edges alone can create paired contours without a filled, connected lumen. Morphology-aware designs work best when boundary, interior, and centerline constraints are coordinated rather than optimized as separate visual cues.

Pulmonary anatomy also demands long-range reasoning. Arteries and veins can run close together and have similar intensities, while branch identity depends on how a local segment connects to central structures. Wu et al. (2023) combined a transformer-based 3D U-Net with contextual modeling for pulmonary vessel segmentation and artery-vein separation. Such approaches increase receptive field, but full-resolution three-dimensional attention is expensive. Patch-wise processing can fit memory limits while severing global coordinates and splitting branches at crop boundaries. The central design problem is therefore not simply choosing convolution or attention; it is allocating limited computation to the spatial relationships most relevant to connectivity.

Foundation Models Meet Volumetric Anatomy

The Segment Anything Model demonstrated that a large promptable model trained on extensive natural-image masks could transfer broadly across two-dimensional segmentation tasks (Kirillov et al., 2023). Medical images differ from natural photographs in contrast, texture, semantics, and dimensionality. MedSAM addressed part of this gap by fine-tuning a SAM-derived model on more than one million medical image-mask pairs spanning multiple modalities and targets (Ma et al., 2024). These models offer rich representations and reusable prompt interfaces, but pulmonary vessel segmentation tests their weakest conditions: low contrast, very small structures, dense branching, and a need for continuous three-dimensional topology.

MorVess treats this mismatch as an adaptation problem rather than discarding the foundation model. Mao et al. (2026) freeze the two-dimensional SAM vision encoder and introduce a lightweight 2.5D adapter. Five adjacent CT slices form the input context, and a small three-dimensional convolution along depth communicates inter-slice information. A global-local fusion block then combines multi-level semantic features with the predicted geometric cues. The reported design uses about one million trainable parameters and substantially less memory than the full volumetric baselines in its comparison.

This compromise is attractive because the highest in-plane resolution can be retained without the cost of a full three-dimensional transformer. It also has a clearly stated boundary. The depth receptive field covers only two neighboring slices on either side, so it cannot directly model a vessel trajectory spanning many millimeters. The authors also note reduced robustness to anisotropic voxels. The limitation is not incidental; it defines when a 2.5D solution is plausible. Thin-section CT with stable spacing may support local trajectory inference, whereas thick or irregular slice spacing may make adjacent images poor geometric neighbors.

Reading the MorVess Evidence

MorVess is evaluated primarily on the Parse2022 pulmonary artery dataset and the AIIB2023 chest dataset. The first contains high-resolution CT with multi-level pulmonary artery labels, while the second emphasizes structurally difficult scans associated with pulmonary fibrosis. The reported comparisons include Dice, cDice, HD95, missing-rate and branch-oriented measures, plus geometric analyses of vessel volume, diameter distribution, tortuosity, and small-vessel fraction. This breadth is important because it tests whether a gain reflects more foreground, a better surface, a more connected tree, or more faithful morphometry.

The Parse challenge itself was designed to compare both central and branch pulmonary artery segmentation and to include efficiency in the evaluation (Luo et al., 2023). On Parse2022, Mao et al. (2026) report a mean Dice of 86.84, cDice of 83.22, and HD95 of 4.53 mm. On AIIB2023, the corresponding reported values are 94.31, 89.34, and 3.24 mm. Bootstrap comparisons are reported against the second-best method, although not every metric reaches the same significance threshold. The study also presents cross-domain tests: a Parse-trained model is applied to pulmonary veins in HiPaS, and an AIIB-trained model is applied to peripheral airways in ATM2022 without fine-tuning. Performance drops relative to in-domain evaluation but retains substantial centerline agreement.

These experiments support the claim that distance and thickness fields capture transferable tubular properties. They do not yet prove clinical reliability. The benchmark annotations define what counts as a vessel, and their conventions may differ across centers. Cross-anatomy testing on airways demonstrates structural transfer but is not a substitute for multi-site pulmonary vessel validation. The reported geometric correlations are closer to downstream use than Dice alone, yet associations with diagnosis, treatment planning, or hemodynamic modeling require separate prospective evidence.

From Better Masks to Better Vascular Measurements

The clinical value of morphology-aware segmentation lies in derived measurements: vascular volume, branch count, caliber distribution, pruning, tortuosity, and spatial relationships between arteries, veins, airways, and lesions. A recent large-scale study by Chu et al. (2025) illustrates this transition. Their HiPaS framework segments pulmonary arteries and veins from contrast and non-contrast CT and applies the outputs to population-scale anatomical analysis. The study shows that abundance and branch recovery can matter as much as aggregate overlap when the scientific objective is vascular phenotyping.

This creates a stricter standard for segmentation. An error that has little effect on Dice may bias a biomarker systematically. Missing peripheral branches lowers small-vessel fraction; a false connection changes graph hierarchy; over-smoothed caliber reduces apparent heterogeneity; and artery-vein confusion can reverse interpretation. Morphology-aware networks should therefore be validated by measurement agreement and repeatability, not merely by mask similarity. A useful study would compare derived biomarkers under repeat scans, reconstruction changes, and expert annotation variability.

Uncertainty must also be localized. A single confidence score for the volume cannot show whether uncertainty lies at peripheral terminals, bifurcations, or vessel-wall interfaces. Geometry-aware models already produce multiple fields that could support structured uncertainty. Disagreement between mask, distance, and thickness heads may flag a questionable branch. Ensembles or perturbation tests could

estimate the stability of centerlines and caliber. The aim is not to hide uncertainty through post-processing, but to expose where the reconstructed network is least dependable.

Design Principles for the Next Generation

First, morphology should be specified in physical units. Distance, thickness, and surface errors need voxel-spacing-aware computation. Second, supervision should combine region, boundary, centerline, and graph information without allowing one abundant signal to dominate. Third, training patches should preserve or encode global location so that local predictions remain consistent with the tree. Fourth, comparisons should include a strong self-configuring baseline and ablations that hold preprocessing and data splits constant. Fifth, evaluation should stratify vessels by caliber and branch generation, because averaged results conceal distal failure.

Data design is equally important. Complete labeling of peripheral vessels is difficult and reader-dependent. Consensus annotation, explicit stopping rules, and uncertainty labels are preferable to presenting one mask as indisputable truth. Training on several acquisition protocols can reveal whether a geometric prior is genuinely invariant or merely compatible with one resampling pipeline. External tests should preserve native variation rather than normalizing every domain until the challenge disappears.

Finally, structural plausibility should not be confused with anatomical correctness. A network can produce a beautifully connected tree that joins the wrong branches. Topology-aware loss must be paired with image evidence and, where the task requires it, artery-vein or lobe-level anatomical constraints. Human review remains necessary for high-stakes applications. The best role for morphology-aware learning is to make likely errors rarer, more measurable, and easier to inspect.

Conclusion

Pulmonary vessel segmentation is moving from foreground recovery toward explicit reconstruction of a multiscale anatomical network. U-Net and its volumetric descendants established powerful representations, while self-configuring training made strong baselines accessible. Distance transforms, centerline losses, thickness fields, curvilinear convolutions, and foundation-model adapters now encode properties that regional overlap does not see. MorVess brings these strands together through joint mask-distance-thickness prediction, 2.5D adaptation of SAM, global-local fusion, and topology-oriented evaluation. Its benchmark and cross-domain results provide substantive evidence for geometry-aware supervision, while its limited depth context, dependence on derived annotation priors, and need for multi-site clinical validation define the next research problems. Progress should ultimately be judged not only by more overlapping voxels, but by whether the recovered tree supports stable and clinically meaningful vascular measurements.

References

- Mao, F., Chen, Y., Wu, B., Lin, L., Dai, J., Li, Z., ... & Qin, F. (2026). MorVess: Morphology-Aware Pulmonary Vessel Segmentation Network. *arXiv preprint arXiv:2606.24214*.
- Chu, Y., Luo, G., Zhou, L., Cao, S., Ma, G., Meng, X., ... & Gao, X. (2025). Deep learning-driven pulmonary artery and vein segmentation reveals demography-associated vasculature anatomical differences. *Nature Communications*, 16, 2262.
- Cicek, O., Abdulkadir, A., Lienkamp, S. S., Brox, T., & Ronneberger, O. (2016). 3D U-Net: Learning dense volumetric segmentation from sparse annotation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016* (pp. 424-432). Springer.
- Isensee, F., Jaeger, P. F., Kohl, S. A. A., Petersen, J., & Maier-Hein, K. H. (2021). nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nature Methods*, 18(2), 203-211.
- Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., ... & Girshick, R. (2023). Segment Anything. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 4015-4026).
- Luo, G., Wang, K., Liu, J., Li, S., Liang, X., Li, X., ... & Gao, X. (2023). Efficient automatic segmentation for multi-level pulmonary arteries: The PARSE challenge. *arXiv preprint arXiv:2304.03708*.

- Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B., ... & Wang, B. (2024). Segment anything in medical images. *Nature Communications*, 15, 654.
- Moccia, S., De Momi, E., El Hadji, S., & Mattos, L. S. (2018). Blood vessel segmentation algorithms - review of methods, datasets and evaluation metrics. *Computer Methods and Programs in Biomedicine*, 158, 71-91.
- Qi, Y., He, Y., Qi, X., Zhang, Y., & Yang, G. (2023). Dynamic Snake Convolution Based on Topological Geometric Constraints for Tubular Structure Segmentation. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6070-6079).
- Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015* (pp. 234-241). Springer.
- Shit, S., Paetzold, J. C., Sekuboyina, A., Ezhov, I., Unger, A., Zhylka, A., ... & Menze, B. H. (2021). clDice - A novel topology-preserving loss function for tubular structure segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 16560-16569).
- Wang, Y., Wei, X., Liu, F., Chen, J., Zhou, Y., Shen, W., Fishman, E. K., & Yuille, A. L. (2020). Deep Distance Transform for tubular structure segmentation in CT scans. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 3833-3842).
- Wu, Y., Qi, S., Wang, M., Zhao, S., Pang, H., Xu, J., Bai, L., & Ren, H. (2023). Transformer-based 3D U-Net for pulmonary vessel segmentation and artery-vein separation from CT images. *Medical & Biological Engineering & Computing*, 61(10), 2649-2663.
- Xia, L., Zhang, H., Wu, Y., Song, R., Ma, Y., Mou, L., ... & Zhao, Y. (2022). 3D vessel-like structure segmentation in medical images by an edge-reinforced network. *Medical Image Analysis*, 82, 102581.