Learning Hybrid Representations for Automatic 3D Vessel Centerline Extraction

Jiafa He[1], Chengwei Pan[2], Can Yang[1], Ming Zhang[2], Yang Wang[1],
Xiaowei Zhou[3(✉)], and Yizhou Yu[4(✉)]

[1] Department of Mathematics,
The Hong Kong University of Science and Technology, Hong Kong, China
[2] Department of Computer Science, School of EECS, Peking University,
Beijing, China
[3] The State Key Lab of CAD&CG, Zhejiang University, Hangzhou, China
xzhou@cad.zju.edu.cn
[4] Deepwise AI Lab, Beijing, China
yizhouy@acm.org

Abstract. Automatic blood vessel extraction from 3D medical images is crucial for vascular disease diagnoses. Existing methods based on convolutional neural networks (CNNs) may suffer from discontinuities of extracted vessels when segmenting such thin tubular structures from 3D images. We argue that preserving the continuity of extracted vessels requires to take into account the global geometry. However, 3D convolutions are computationally inefficient, which prohibits the 3D CNNs from sufficiently large receptive fields to capture the global cues in the entire image. In this work, we propose a hybrid representation learning approach to address this challenge. The main idea is to use CNNs to learn local appearances of vessels in image crops while using another point-cloud network to learn the global geometry of vessels in the entire image. In inference, the proposed approach extracts local segments of vessels using CNNs, classifies each segment based on global geometry using the point-cloud network, and finally connects all the segments that belong to the same vessel using the shortest-path algorithm. This combination results in an efficient, fully-automatic and template-free approach to centerline extraction from 3D images. We validate the proposed approach on CTA datasets and demonstrate its superior performance compared to both traditional and CNN-based baselines.

Keywords: Centerline extraction · Vessel segmentation · Hybrid representations

---

J. He and C. Pan—These authors contributed equally to this work.

---

1   Introduction

Extracting tubular objects, e.g., blood vessels, has become a crucial task in computer-assisted diagnosis (CAD) of many diseases. For example, vessel lumen segmentation and centerline extraction are prerequisites for vessel curved-planar reconstruction (CPR) [6] from computed tomography angiography (CTA) images, which further facilitates stenosis detection and plaque identification in clinical diagnosis. However, it is usually time-consuming to segment vessel and extract centerline from various medical images. Instead, automatic vessel segmentation and centerline detection play a more and more important role for quantitative analysis of vascular diseases. [1,7,26].

Recently, convolutional neural networks (CNNs) have been widely applied in 3D medical image segmentation. However, segmenting vessels in 3D medical images is still very challenging. The blood vessels have delicate tubular structures with a large variety in long-range topology, which cannot be captured by slice-wise or patch-wise convolutional operations in most deep learning based segmentation methods [5,9,24]. Moreover, at the presence of imaging artifacts which often exist in medical images, CNN-based segmentation algorithms are prone to missing some segments of vessels, resulting discontinuities in the extracted vessel centerline [18,23,28]. To preserve the correct topology of the extracted vessels, previous methods may rely on user input that annotates start and end points of each vessel [17,22]. Then, a more complete vessel centerline can be found by a minimal-cost path-based algorithm [2,4,8,10]. To avoid manual input, another approach to ensure correct topology is to build an atlas or template of the target vessels from training samples and register the template to the test image [27,28]. However, this approach is not very generalizable as the vascular structure of the test sample might be very different from the template.

In this paper, we present a novel approach for vessel centerline extraction, which is able to ensure the connectivity of extracted vessels without the need of any manual input or vessel template. The key idea is to use a patch-wise 3D CNNs to segment vessel mask and regress vessel centerline heatmap from the input image, and meanwhile use another point-cloud network to label the extracted vessel segments, such that segments belonging to the same vessel can be connected in a post-processing step. This hybrid approach makes the best use of both worlds: patch-wise CNNs for local appearance learning and point-cloud networks for global geometry learning, resulting in a robust and efficient algorithm for automatic centerline extraction. We also propose a geometry-aware grouping strategy to improve the performance of point-cloud network for vessel labeling. The effectiveness of the proposed framework is validated on two datasets: a public dataset of coronary artery CTA scans and an in-house dataset of head and neck artery CTA scans. Experimental results show that our approach outperforms existing baseline methods in terms of both accuracy and completeness of extracted centerlines.

In summary, we make the following contributions: (1) A novel hybrid representation learning approach for fully-automatic and template-free vessel centerline extraction; (2) A geometry-aware grouping method that utilizes the skele-
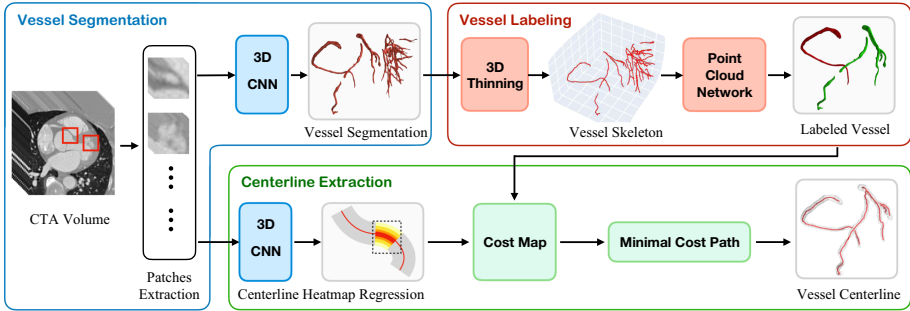
Fig. 1. Overview of the proposed approach. The framework consists of three components: **Vessel segmentation (blue block)**: the input CTA images are divided into 3D patches with overlap and then a 3D CNNs are utilized to efficiently learn local features of vascular objects. **Vessel labeling (red block)**: vascular skeletons produced by thinning the segmentation results are fed into a point-cloud network to learn the global geometry of vessels and realize vascular branch labeling. **Centerline extraction (green block)**: based on a cost map, which is constructed from the centerline heatmap and the labeled skeleton, a minimal cost path algorithm is finally utilized to extract the complete vessel centerline. (Color figure online)

ton's connection property to improve the performance of vessel labeling; (3) The state-of-the-art performance on the public benchmark.

2   Methods

Given a 3D CTA image consisted of a sequence of 2D slices, the objective is to segment the arteries and delineate their centerlines. The state-of-the-art segmentation methods are mostly based on CNNs. Due to the heavy computation of 3D convolutions, an input 3D image needs to be divided into overlapped patches and fed into the segmentation network separately. This leads to restricted local receptive fields, which may not provide sufficient information to distinguish between arteries and veins, resulting false detection. Moreover, there is no guarantee on the connectivity of extracted vessel segments from patch-based CNNs. A solution is to connect the segments that belong to the same vascular branch in post-processing. But to achieve this we need to label all the segments, which is also difficult for a patch-based CNN as vessel labeling requires considering the global geometry of the vessels.

To address these issues, we propose a hybrid approach that consists of a patch-based CNNs for local **vessel segmentation**, a point-cloud network for global **vessel labeling** and a path-finding algorithm for final **centerline extraction**. Figure 1 provides an overview to our approach.

2.1   Vessel Segmentation

At first, we use a 3D CNNs to learn vascular local appearance features from 3D patches of the original CTA images and produce a coarse vessel segmentation. An
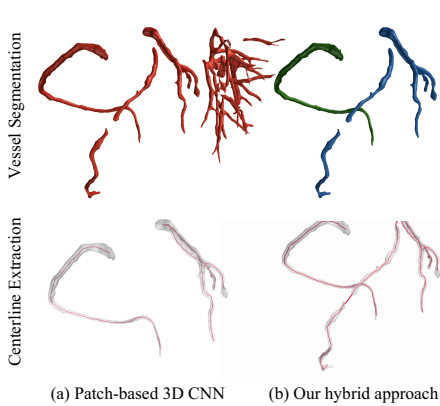
Vessel Segmentation

Centerline Extraction

(a) Patch-based 3D CNN   (b) Our hybrid approach

Neck and Head Vessel

Coronary Artery

Original Vessel Skeleton   Labeled Vessel Skeleton

Fig. 2. Comparison. Patch-based 3D CNNs extract extra non-vascular tissues and miss some true segments (left). Our hybrid approach is able to remove non-vascular tissues and connect the disjointed segments (right).
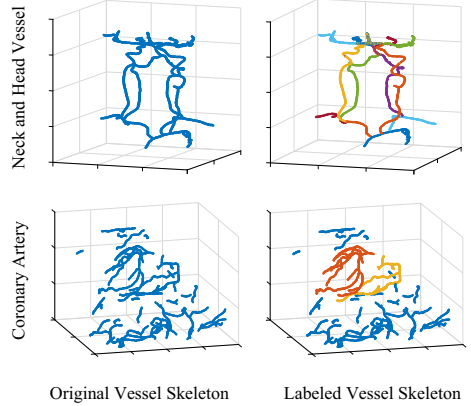
Fig. 3. Generated vessel skeletons. Head and neck arteries (the top row) are divided into 17 categories distinguished by color. Coronary artery (the bottom row) are divided into 3 categories including a category for false positives.

UNet [15] backbone architecture with an encoding-decoding module is selected to transform the input image to the segmentation mask. Moreover, to explore long-range contextual information inside 3D patches, we embed a dual attention module [3] on top of the UNet backbone. Finally, a combination of a binary cross-entropy loss and a Dice loss is used as the total segmentation loss. Please refer to the supplementary for more details.

2.2   Vessel Labeling

Due to the lack of the global information in the patch-wise 3D CNNs, the vessel segments obtained from the vessel segmentation procedure tend to contain some false-positive results like veins and also miss some parts of tiny or tortuous vessels. We propose to perform a vessel labeling procedure that classifies the segmented vessels into different branches. Then such semantic labels can be used to remove non-vascular segments and group discontinuous vessel segments, as shown in Fig. 2. The vessel labeling procedure is implemented by first generating a set of points which represent vascular skeleton from the vessel segmentation results and then using a point-cloud network to predict labels of these skeleton points.

Point-Cloud Generation. As it is inefficient to directly label vessel segments in 3D volumes using CNNs, we propose to perform the labeling on vascular skeletons represented by a set of points, which can reduce complexity and preserve original geometric information of vessels. To generate vascular skeletons from
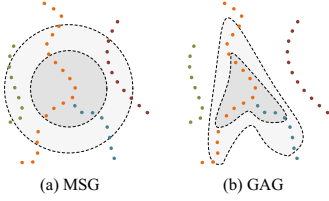
(a) MSG   (b) GAG

(a) centerline heatmap   (b) cost map   (c) minimal cost path

Fig. 4. Comparison of points grouping methods. (a) **Multi-scale grouping (MSG) in the PointNet++**: the grouping areas are multi-scale ϵ-sphere and grouped points may come from different skeleton components. (b) **Geometry-aware grouping (GAG)**: the grouping areas trend to stretch along the lines and group points from the same islands. (Color figure online)
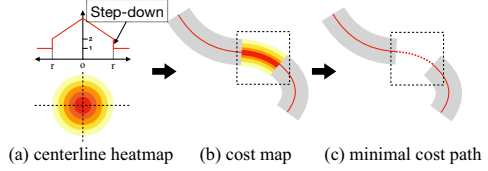
Fig. 5. Centerline Extraction. (a) **Centerline heatmap**: the heatmap has larger value in point closer to the centerline and the points laying outside the vascular radius are set to a step-down low value. (b) **Cost map**: the cost map is built by combining the heatmap and the skeleton lines. (c) **Minimal cost path**: based on the cost map, a minimal cost path can be extracted to connect the broken lines as the complete vessel centerline. (Color figure online)

the vessel segments, a 3D thinning algorithm [11] is applied to erode the vessel segments and finally obtain single-voxel-width skeleton points, as shown in Fig. 3. The generated skeletons are composed of discrete and unordered points lying on the center of the vascular lumen, which represent vascular geometry.

Vessel Labeling Network. Given a set of vascular skeleton points generated from vessel segments $\{Q_i \in \mathbb{R}^3, i \in 1, 2, ..., M\}$, the target of the vessel labeling network is to predict the label of each point, $f(Q_i) \rightarrow Y_j$ $(Y_j \in \mathbb{Z}, j \in 1, 2, ..., K)$, as shown in Fig. 3. It is similar to point-cloud semantic segmentation tasks in 3D computer vision. Any point-cloud network can be adopted, such as the state-of-the-art PointNet++ [14] and dynamic graph CNNs (DGCNN) [21]. We will provide a comparison between them in experiments.

Geometry-Aware Grouping. A particular property of the skeleton points compared to a general point cloud is the given connectivity among adjacent skeleton points. Specifically, skeleton points can be divided into separated components using a connected-component labeling (CCL) algorithm [12]. However, both PointNet++ and DGCNN are realized to group points by k-nearest neighbor (k-NN) algorithm based on L2 distance, which ignores the given geometry of the vessel skeletons. As demonstrated in Fig. 4(a), the L2-based methods are more likely to group points belonging to different components. To address this issue and leverage the skeletons' connection property, we propose a geometry-aware grouping method (GAG) to modify the distance based on the connection relationship. The modified distance between two points is computed by

$$\mathbf{D}_{i,j} = \begin{cases} \lambda_n \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_2 & \text{if } \mathbf{x}_i, \mathbf{x}_j \in C_n. \\ (1 - \lambda_n) \left\| \mathbf{x}_i - \mathbf{x}_j \right\|_2 & \text{otherwise.} \end{cases} \qquad (1)$$

where λ_n ($\lambda_n < 0.5$) is a weight, C_n is the n-th component. As shown in Fig. 4(b), the grouping area in GAG is prone to stretching along the skeleton lines and the points belonging to the same connected component are more likely to be grouped. This design facilitates local feature consistency in the same component and consequently improves the accuracy of vessel labeling as evaluated in the experiments.

2.3   Centerline Extraction

After vessel labeling, each vessel segment is assigned a semantic label. The semantic label can be used to remove non-vascular tissues and provide the guidance to connect disjointed vessel segments. Specifically, we first determine which segments should be connected and then use a minimal cost path method to connect them and output final centerlines based on a cost map. The cost map is constructed by a centerline heatmap regressed from another 3D CNNs and the labeled vessel skeletons.

Centerline Heatmap Regression. The centerline heatmap is defined as the opposite of a distance map, where points closer to the vessel centerline have larger values and points outside the vascular radius have a step-down low value, as shown in Fig. 5(a). Considering that the probability map obtained from the segmentation network has only learned the difference between background and the vessel (e.g. edge features) rather than the centerline feature within the vessel lumen, we adopt another 3D CNN similar to the network used in vessel segmentation to regress the centerline heatmap. The mean square error (MSE) loss is used to train the network. Based on the centerline heatmap, we construct a cost map to guide the minimal cost path search algorithm as shown in Fig. 5(b). In areas where vessel segments exist, the cost map directly assigns a large value to the skeleton points (red lines in Fig. 5(b)) and a small value to the points elsewhere (gray areas in Fig. 5(b)).

Minimal Cost Path. Given vessel skeletons and their labels, we successively merge the disjointed segments with the same label. Paired boundary points with the same label are put in a priority queue according to the distance between the two points in the pair. We take a pair successively from the queue and then the minimal cost path is found by Dijkstra algorithm [20] to connect the two boundary points, as shown in Fig. 5(c). If the two segments represented by the two points are connected in the previous step, we skip the pair and go on until the queue is empty.

3   Experiments

We evaluate the proposed method on two datasets: a public coronary artery dataset and a private head and neck artery dataset. The first dataset is mainly

Table 1. Performance comparison of automatic coronary artery centerline extraction methods. OV represents the completeness of extracted vessel centerline and is similar to Dice coefficient. OF determines the ratio of a coronary artery that has been extracted before making an error. OT gives an indication of how much the extracted centerline overlaps with the clinically relevant centerline reference (radius $\geq 0.75\,\mathrm{mm}$).

Methods	OV (%)	OF (%)	OT (%)
ModelDrivenCenterline [28]	92.4	80.6	93.4
SupervisedExtraction [7]	90.6	70.9	92.5
GFVCoronaryExtractor [23]	93.7	74.2	95.9
DepthFirstModelFit [25]	84.7	65.3	87.0
AutoCoronaryTree [19]	84.7	59.5	86.2
Our approach	**95.8**	72.3	**96.3**

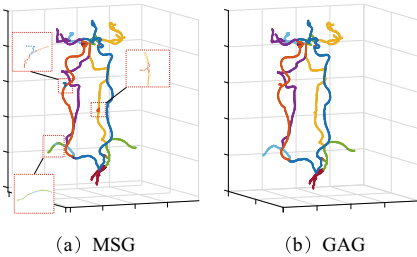(a) MSG   (b) GAG

Low Contrast

Artificial Noise

Fig. 6. Comparison of different grouping methods used in vessel labeling networks. It can be seen that the proposed GAG method achieves better local consistency of the predicted vessel labels.
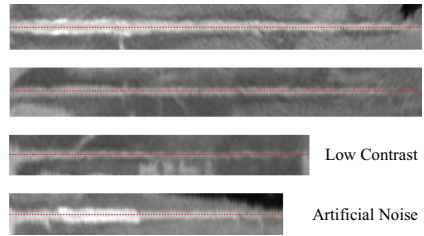
Fig. 7. Vessel curved-planar reconstruction (CPR) of the coronary artery based on the extracted centerline. Our approach can extract complete centerlines under serious imaging artifacts.

used to compare our method with existing baseline methods in literature. The second dataset is mainly used for ablative study to verify our system designs. Following [16], extracted centerlines are evaluated based on three metrics, namely total overlap (OV), overlap until first error (OF), and overlap with the clinically relevant part of the vessel (OT). The stage-wise results for the proposed framework are demonstrated in Fig. 8.

Implementation Details. In vessel segmentation, we randomly crop 3D patches with the size of $256 \times 256 \times 32$ for the head and neck dataset and $32 \times 32 \times 32$ for the coronary artery dataset. The ResNet34 is used as the encoder in the UNet architecture which starts with 32 feature channels that are doubled in each scale, and the max-pooling layer is removed from the original residual network. All convolutions are specified as $3 \times 3 \times 3$ kernels, except the last two ResNet blocks, which are $3 \times 3 \times 1$ to reduce the parameter count. We employ the Adam optimizer with a polynomial learning rate which equals to $0.001 \times (1 - \frac{iter}{total\ iter})^{0.9}$. In vessel labeling, the inputs are skeleton points,

Table 2. Performance comparison of different variants of vessel labeling networks on the private head and neck artery dataset. Four metrics are reported: the accuracy of vessel labeling and OV, OF and OT of four main vessel centerlines including left and right common carotid artery (L/RCCA), left and right vertebral artery (L/RVA).

Vessel Labeling	Accuracy (%)	OV (%)	OF (%)	OT (%)
PointNet [13]	92.2	96.1	84.5	96.6
PointNet++ [14]	95.1	96.5	**85.8**	97.0
PointNet++ (GAG)	97.0	**97.2**	85.7	**97.6**
DGCNN [21]	96.5	96.6	84.4	97.1
DGCNN (GAG)	**97.2**	97.0	84.9	97.4

Vessel Segmentation   Vessel Skeleton Labeling   Labeled Vessel   Vessel Centerline
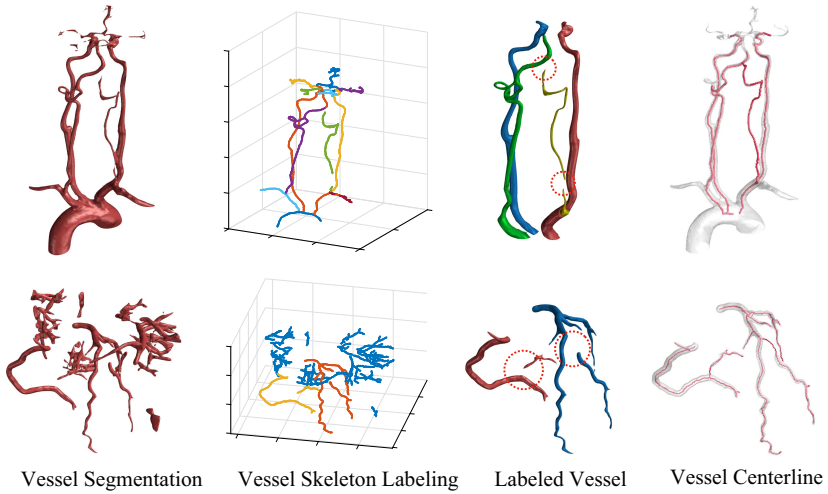
Fig. 8. Vessel segmentation, labeling and centerline extraction. Top: head and neck arteries. Note that the discontinuous RVA is connected in the centerline extraction procedure. **Bottom:** coronary arteries. Note that the non-vascular tissues are removed in the vessel labeling procedure and the discontinuous vessel segments are connected.

which are generated from vessel segments and resampled to 3000 points for each sample. The inital learning rate is set to 0.001, which is reduced by half every 30 epochs. The weight λ_n in GAG is set to be 0.3.

Experiments on the Coronary Artery Dataset. This public dataset contains 100 cardiac CT angiography (CCTA) scans collected from the clinic for training and 32 CCTA scans from [16] for evaluation. We train the vessel segmentation network and vessel labeling network on the annotated coronary artery CTA images and the vessel skeletons are labeled as three categories including right arteries, left arteries and false-positive venous vessels. According to the ablation study in the head and neck artery dataset, we use the Pointnet++ with the GAG module as our vessel labeling network. The quantitative com-

parison is listed in Table 1 and the visualization of results is showed in Fig. 7. Our hybrid approach achieves the highest performance in terms of OV and OT, respectively, indicating that the centerlines extracted by the proposed method are more complete than those produced by other methods.

Experiments on the Head and Neck Artery Dataset. This private dataset collected from the clinic contains 450 CTA scans, each of which has a manually annotated vessel mask. The dataset is split into 380 scans for training, 20 for validation and 50 for testing. In the dataset, vessel skeletons are labeled as 17 categories including left and right common carotid artery (L/RCCA), left and right vertebral artery (L/RVA), etc. Table 2 shows the evaluation results of several variants of our system with different point-cloud network designs. It can be seen that, with the geometry-aware grouping (GAG) method, the vessel labeling accuracy of both PointNet++ and DGCNN can be improved. Figure 6 shows the GAG can facilitate local consistency of the skeleton components.

4   Conclusions

We propose an automatic and template-free approach to 3D vessel centerline extraction based on hybrid representations, which ensures the connectivity of extracted centerlines. We show that the hybridization between learning local appearance with patch-based CNNs and learning global geometry with point-cloud networks results in an efficient and robust framework to extract geometric objects from 3D data. We demonstrate superior performance on artery centerline extraction from CTA images and believe that the proposed approach can also be applied in other centerline or skeleton extraction tasks.

References

1. Cetin, S., Unal, G.: A higher-order tensor vessel tractography for segmentation of vascular structures. IEEE Trans. Med. Imaging **34**(10), 2172–2185 (2015)
2. Deschamps, T., Cohen, L.D.: Fast extraction of minimal paths in 3D images and applications to virtual endoscopy. Med. Image Anal. **5**(4), 281–299 (2001)
3. Fu, J., et al.: Dual attention network for scene segmentation. In: CVPR (2019)
4. Gülsün, M.A., Funka-Lea, G., Sharma, P., Rapaka, S., Zheng, Y.: Coronary centerline extraction via optimal flow paths and CNN path pruning. In: Ourselin, S., Joskowicz, L., Sabuncu, M.R., Unal, G., Wells, W. (eds.) MICCAI 2016. LNCS, vol. 9902, pp. 317–325. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-46726-9_37

5. Kamnitsas, K., Ledig, C., Newcombe, V.F., Simpson, J.P., Kane, A.D., Menon, D.K., Rueckert, D., Glocker, B.: Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation. Medical image analysis **36**, 61–78 (2017)
6. Kanitsar, A., Fleischmann, D., Wegenkittl, R., Felkel, P., Gröller, M.E.: CPR: curved planar reformation. In: VIS. IEEE Computer Society (2002)
7. Kitamura, Y., Li, Y., Ito, W.: Automatic coronary extraction by supervised detection and shape matching. In: ISBI. IEEE (2012)
8. Li, H., Yezzi, A.: Vessels as 4-D curves: global minimal 4-D paths to extract 3-D tubular surfaces and centerlines. IEEE Trans. Med. Imaging **26**(9), 1213–1223 (2007)
9. Litjens, G., et al.: A survey on deep learning in medical image analysis. Med. Image Anal. **42**, 60–88 (2017)
10. Moriconi, S., Zuluaga, M.A., Jäger, H.R., Nachev, P., Ourselin, S., Cardoso, M.J.: VTrails: inferring vessels with geodesic connectivity trees. In: Niethammer, M., et al. (eds.) IPMI 2017. LNCS, vol. 10265, pp. 672–684. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-59050-9_53
11. Palágyi, K., et al.: A sequential 3D thinning algorithm and its medical applications. In: Insana, M.F., Leahy, R.M. (eds.) IPMI 2001. LNCS, vol. 2082, pp. 409–415. Springer, Heidelberg (2001). https://doi.org/10.1007/3-540-45729-1_42
12. Park, J.M., Looney, C.G., Chen, H.C.: Fast connected component labeling algorithm using a divide and conquer technique. Comput. Their Appl. **4**, 4–7 (2000)
13. Qi, C.R., Su, H., Mo, K., Guibas, L.J.: Pointnet: deep learning on point sets for 3D classification and segmentation. In: CVPR (2017)
14. Qi, C.R., Yi, L., Su, H., Guibas, L.J.: Pointnet++: deep hierarchical feature learning on point sets in a metric space. In: NIPS (2017)
15. Ronneberger, O., Fischer, P., Brox, T.: U-Net: convolutional networks for biomedical image segmentation. In: Navab, N., Hornegger, J., Wells, W.M., Frangi, A.F. (eds.) MICCAI 2015. LNCS, vol. 9351, pp. 234–241. Springer, Cham (2015). https://doi.org/10.1007/978-3-319-24574-4_28
16. Schaap, M., et al.: Standardized evaluation methodology and reference database for evaluating coronary artery centerline extraction algorithms. Med. Image Anal. **13**(5), 701–714 (2009)
17. Selvan, R., Petersen, J., Pedersen, J.H., de Bruijne, M.: Extracting tree-structures in CT data by tracking multiple statistically ranked hypotheses (2018). arXiv preprint arXiv:1806.08981
18. Stefancik, R.M., Sonka, M.: Highly automated segmentation of arterial and venous trees from three-dimensional magnetic resonance angiography (MRA). Int. J. Cardiovasc. Imaging **17**(1), 37–47 (2001)
19. Tek, H., Gulsun, M.A., Laguitton, S., Grady, L., Lesage, D., Funka-Lea, G.: Automatic coronary tree modeling. Insight J., 1–8 (2008)
20. Verscheure, L., Peyrodie, L., Makni, N., Betrouni, N., Maouche, S., Vermandel, M.: Dijkstra's algorithm applied to 3D skeletonization of the brain vascular tree: evaluation and application to symbolic. In: EMBC. IEEE (2010)
21. Wang, Y., Sun, Y., Liu, Z., Sarma, S.E., Bronstein, M.M., Solomon, J.M.: Dynamic graph cnn for learning on point clouds (2018). arXiv preprint arXiv:1801.07829
22. Wolterink, J.M., van Hamersvelt, R.W., Viergever, M.A., Leiner, T., Išgum, I.: Coronary artery centerline extraction in cardiac CT angiography using a CNN-based orientation classifier. Med. Image Anal. **51**, 46–60 (2019)

23. Yang, G., Kitslaar, P., Frenay, M., Broersen, A., Boogers, M.J., Bax, J.J., Reiber, J.H., Dijkstra, J.: Automatic centerline extraction of coronary arteries in coronary computed tomographic angiography. Int. J. Cardiovasc. Imaging **28**(4), 921–933 (2012)

24. Yu, Q., Xie, L., Wang, Y., Zhou, Y., Fishman, E.K., Yuille, A.L.: Recurrent saliency transformation network: incorporating multi-stage visual cues for small organ segmentation. In: CVPR (2018)

25. Zambal, S., Hladuvka, J., Kanitsar, A., Bühler, K.: Shape and appearance models for automatic coronary artery tracking. The MIDAS Journal. In: MICCAI Workshop-Grand Challenge Coronary Artery Tracking (2008). http://hdl.handle.net/10380/1420

26. Zhai, Z., et al.: Automatic quantitative analysis of pulmonary vascular morphology in CT images. Med. Phys. **46**(9), 3985–3997 (2019)

27. Zhang, D.P., et al.: Coronary motion estimation from CTA using probability atlas and diffeomorphic registration. In: Liao, H., Edwards, P.J.E., Pan, X., Fan, Y., Yang, G.-Z. (eds.) MIAR 2010. LNCS, vol. 6326, pp. 78–87. Springer, Heidelberg (2010). https://doi.org/10.1007/978-3-642-15699-1_9

28. Zheng, Y., Tek, H., Funka-Lea, G.: Robust and accurate coronary artery centerline extraction in CTA by combining model-driven and data-driven approaches. In: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (eds.) MICCAI 2013. LNCS, vol. 8151, pp. 74–81. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40760-4_10