Automated Multi-View Planning for Endovascular Aneurysm Repair Procedures

Baochang Zhang[1,2,4(✉)], Yiwen Liu[1], Shuting Liu[1], Heribert Schunkert[2], Reza Ghotbi[3], and Nassir Navab[1]

[1] Computer Aided Medical Procedures, Technical University of Munich, Munich, Germany
[2] German Heart Center Munich, Munich, Germany
[3] German Centre for Cardiovascular Research, Munich Heart Alliance, Munich, Germany
[4] HELIOS Hospital west of Munich, Munich, Germany
baochang.zhang@tum.de

Abstract. During Endovascular aneurysm repair (EVAR) procedures, surgeons always require several views of vessel structures to accurately assess the size, shape, and location of the aneurysm, along with the surrounding vasculature. However, even expert surgeons often require multiple attempts to find a desired view, which leads to increased radiation exposure, high doses of contrast agents for patients, and time-consuming re-positioning of the C-arm. This paper introduces an automatic framework to provide optimal multi-view for the whole EVAR procedure. First, a 3D nnUNet is employed to extract geometric information and semantic information, providing accurate vascular and aneurysm segmentation as well as semantic bifurcation detection. Then, a semantic vessel tree model is built by integrating semantic information and geometric information. A local 3D plane at each critical bifurcation is fitted based on the centerlines surrounding this bifurcation, where we regard the estimated 3D local plane as a good view plane in patient physical space. Next, some 3D points are collected from these centerlines, projected onto the estimated local 3D plane, and transformed to the image domain to get the paired 2D points. Finally, based on the geometric information of the C-arm X-ray imaging device, the most informative view pose for C-arm positioning is solved via RANSAC Perspective-n-Point algorithm with the Levenberg-Marquardt optimization. Our work not only streamlines the surgical planning process, but also helps in customizing the patient-specific strategies to reduce risks and improve surgical outcomes. Our framework has been validated using an in-house dataset collected from 27 patients, which contains preoperative CTA data and intraoperative X-ray angiography images. The qualitative and quantitative results demonstrate the reliability and effectiveness of our approach. Meanwhile, our system achieved an average runtime of 6 min per patient.

Keywords: Multiple View Planning · Abdominal Aortic Aneurysm · EVAR procedures

B. Zhang and Y. Liu — These authors contributed equally to this work.

K. Drechsler et al. (Eds.): CLIP 2024, LNCS 15196, pp. 22–31, 2024.
https://doi.org/10.1007/978-3-031-73083-2_3

1 Introduction

In recent years, the rate of patients treated with Endovascular Aneurysm Repair (EVAR) procedures has increased notably. Typically, successful EVAR procedures demand the acquisition of multiple views to guide the intervention comprehensively. Each view serves a vital role in navigating the intricate vascular anatomy, ensuring precise placement of endovascular devices, and monitoring post-procedural outcomes. However, the safety of patients is still a major concern as imaging surveillance is required. This comes with risks associated with radiation exposure, contrast agent use, as well as increased costs.

With the exponential growth of medical imaging data and advancements in computational power, deep learning algorithms have demonstrated remarkable performance in various healthcare applications [19]. Utilizing preoperative CT angiography images, significant progress has been made in 3D vascular segmentation [3], aneurysm detection [17], aneurysm growth prediction [10], vascular centerline extraction [6], and vessel labeling [20] to assist EVAR procedures. Intra-operative X-ray images have also been the focus of numerous learning-based approaches to aid EVAR procedures, such as X-ray/CT registration [13] and 2D vessel segmentation [8]. However, few studies directly address the acquisition of the optimal surgeon's view. The definition of good views involves ensuring that the imaging provides clear, accurate, and comprehensive visualization of the relevant anatomical structures during the procedure. Fallavollita et al. [2] proposed a user interface concept enabling the surgeon to manually select the desired view before surgery, aiming to alleviate the challenges associated with constantly repositioning the angiographic C-arm during intervention. Tehlan et al. [16] suggested using an augmented reality head-mounted display for the surgeon to choose a desired X-ray view, providing corresponding C-arm configuration as visual feedback. Nevertheless, manually selecting these optimal views can be time-consuming and subjective, potentially leading to suboptimal outcomes and increased patient risks. Recently, Kausch et al. [9] introduced a convolutional neural network regression model to predict five degrees of freedom pose updates directly from the initial X-ray image in orthopedic surgery, facilitating automated C-arm positioning to achieve the desired view. However, this approach necessitates manual annotations of desired views for training, which is a labor-intensive and time-consuming task that significantly restricts its applicability.

In this paper, we introduce a complete framework for automatically providing multiple optimal views to guide EVAR procedure, thereby reducing the need for manual selection and minimizing procedural inefficiencies. The main contributions are as follows. (1) We leverage 3D nnUNet to extract geometric information and semantic information, where 3D center distance loss is proposed for accurate semantic bifurcation detection. (2) Our framework effectively integrates semantic geometric information extracted from patient-specific pre-operative CTA data and geometric information of C-arm X-ray imaging device for multi-view C-arm positioning. (3) We validate the feasibility and effectiveness of the proposed framework using CTA data of 27 patients and their corresponding intraoperative

X-ray angiography images, and the entire pipeline achieves an average runtime of 6 min per patient.

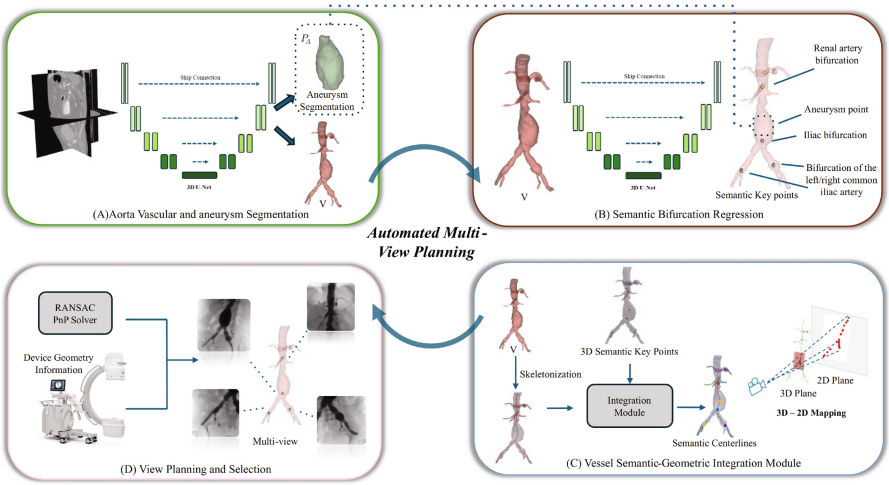

Fig. 1. An overview of the proposed pipeline for automated multi-view planning. The notation V represents the vascular segmentation, P_A is the aneurysm coordinates.

2 Method

2.1 Aorta and Aneurysm Segmentation

We employ a 3D nnU-Net [7] to segment both the aortic vascular and the aneurysms from pre-operative CTA data, trained with the Dice loss function, as shown in Fig. 1(A). The coordinates of the aneurysm P_A are further determined by calculating the centroid of the segmented aneurysm area.

2.2 Semantic Bifurcation Detection

This section focuses on extracting semantic information from pre-operative data. Relying solely on extracting and labeling the centerline from vascular segmentation V can lead to inaccuracies, particularly when the segmentation is not continuous [18]. To improve accuracy, we focus our semantic annotations exclusively on four key bifurcations. We utilize a 3D nnU-Net [7] to detect these bifurcations directly by regressing Gaussian heatmap kernels in a supervised learning manner, as illustrated in Fig 1(B).

For the ground truth, it has four channels, and the channel order represents distinct semantic information. Each channel $G \in R^{D\times H\times W}$ contains an unnormalized 3D Gaussian distribution centered on each key bifurcation:

$$G(i,j,k) = \exp\left(-\frac{(i-c_i)^2+(j-c_j)^2+(k-c_k)^2}{2\sigma^2}\right) \tag{1}$$

where (c_i, c_j, c_k) is the IJK coordinates of each bifurcation. Although heatmap-based regression method is commonly used in 2D key point detection [11], it faces challenges in 3D space due to spatial sparsity. To mitigate this, we set the σ value to 28mm to maximize non-overlapping area and minimize sparsity.

We downsampled the vascular segmentation results to help the network recognize global structures more easily and simplify learning, especially for distinguishing symmetric key points, like left common iliac artery bifurcation and right common iliac artery bifurcation. Additionally, we employ specialized loss functions to adapt the sparsity of 3D space.

Weighted Mean Squared Loss. To encourage the network to learn non-zero values, errors associated with non-zero values are given higher weight [4].

$$L_{\text{WMSE}} = \frac{1}{N}\sum_{i=1}^{N} w_i(y_i - \hat{y}_i)^2 \quad \text{where} \quad w_i = \begin{cases} 1.8 & \text{if } y_i > 0 \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

Here, $\hat{y}_i$ and y_i represent the predicted and ground truth values. The weight w_i emphasize foreground values.

3D Center Distance Loss. While the weighted mean squared error loss L_{WMSE} can align predicted voxel values with the ground truth Gaussian distribution, it does not guarantee the accuracy of the predicted center point coordinates. Typically, the Gaussian distribution peaks at the coordinates of annotated bifurcations. While the Argmax function $torch.argmax()$ returns the peak coordinates of predicted Gaussian distributions, it is not differentiable. Inspired by the differentiable 2D SoftArgmax function [12], we implement a 3D variant that reduces the distance between the predicted and real center coordinates. It has been proved that the maximum value location can be approximated by a weighted sum of the predicted heatmap $G \in \mathbb{R}^{D\times H\times W}$ [5], namely taking the expectation of the probability map G. The predicted maximum value coordinate $\hat{c}$ is calculated as:

$$\hat{c} = SoftArgmax(G) = \sum_{n=1}^{3}\sum_{i=1}^{D}\sum_{j=1}^{H}\sum_{k=1}^{W} W_{n,i,j,k} G'_{i,j,k} \tag{3}$$

where the $G'_{i,j,k}$ is the softmax normalized value of G at location (i, j, k). With the location coordinate (i, j, k), we calculate weighted matrix $W \in \mathbb{R}^{3\times D\times H\times W}$, which can be treated as 3D discrete normalized ramps along axis I, J, K. The notation n corresponds to these three channels:

$$W_{1,i,j,k} = \frac{i}{D}, W_{2,i,j,k} = \frac{j}{H}, W_{3,i,j,k} = \frac{k}{W} \tag{4}$$

Our center distance loss calculates the $L1$ Loss between the predicted and ground-truth key bifurcation:

$$L_{center} = |\hat{c} - c_{gt}| \tag{5}$$

Therefore, the composite loss uses coefficiency $\lambda = L_{WMSE}/L_{Center}$ to dynamically combine these two losses:

$$L_{total} = L_{WMSE} + \lambda L_{center} \tag{6}$$

2.3 Automated Optimal View Selection

Until now, the vascular segmentation V and the predicted bifurcation coordinates K are obtained. In order to integrate these geometric and semantic information together, a semantic vascular tree is built. For the centerline extraction from vascular segmentation V, the classic iterative thinning algorithm [14] is adopted. Then, based on topological analysis in 26 neighborhood system, all bifurcations, edges points and end points are identified from extracted centerline model. Predicted bifurcation coordinates K are adjusted to align with nearby bifurcations on the centerline model if they fall within a specified threshold. When we identify a nearest bifurcation point, we accurately determine the location of key points and provide relevant semantic information for those points. Next, we can fit a 3D plane based on the interested bifurcations and the surrounding branches. Of note, due to the multiple bifurcations exist above the renal artery, we apply an outlier exclusion algorithm to discard atypical vessel orientations to ensure that the planes accurately represented the majority of vessel orientations. For further details, please check Algorithm 1.

We then randomly sample some points on the centerline model around the interested key points. The 2D-3D point pairs are obtained by projecting these 3D points onto the fitted plane and transforming them into image domain, resulting in 2D IJ coordinates. RANSAC Perspective-n-Point (PnP) algorithm [1] is utilized to calculate the pose for virtual C-arm positioning. The goal of using the RANSAC algorithm is to identify and mitigate outlier effects to accurately estimate the object's pose. The intrinsic camera matrix provided to the algorithm is calculated based on the following geometric parameters of the C-arm device: the distance from the X-ray source to the isocenter of device is 742.5 mm; the distance from the detector to the isocenter of device is 517.15 mm; both the width and height of the detector are 432 mm, and the pixel size in detector is 0.3 mm. Meanwhile, to enhance stability across each sampling iterations, we apply the Levenberg-Marquardt algorithm [15] iteratively, calculating the reprojection error from the PnP solution. We then remove outlier points according to the error while ensuring a minimum number of points are maintained. Finally, we visualize our pose quality by rendering X-ray images using Digital Reconstructed Radiography (DRR) method [21].

Algorithm 1

Input: Vascular segmentation V, aneurysm position $P_A \in R^3$, predicted key bifurcations coordinates list $K = \{k_i \in R^3, i = 4\}$

Output: Final key bifurcations coordinate list $K_{final} = \{k_{final_i} \in R^3, i = 4\}$, adjacent points dict $K_{adj} = \{k_{final_i} : k_{adj_j} \in R^3, j = min(len(points), 90)\}$

Get vessel tree T and bifurcation points P_b from V via thinning algorithm
for k_i in K **do**
    Relocate k_i with nearest bifurcation point P_b within threshold $dist = 20mm$ as k_{final_i}, if larger than threshold, choose k_i as k_{final_i}
    Identify vessel sub-tree $T_{sub_k} = T_{k_{final_i}}$ from T associated with k_{final_i}. Sample adjacent points k_{adj_j} on T_{sub_k} for k_{final_i}.
    if k_{final_i} is not kidney or aneurysm key point or degree(k_{final_i} — T_{sub_k}) = 3 **then**
        continue
    else
        for each branch t^* in T_{sub_k} **do**
            Let $S = T_{sub_k} \setminus \{t^*\}$ {Set S contains all branches except t^*}
            Compute the angle θ between the directional vector of t^* and the normal vector of the plane fitted to S
            if all angle is greater than a threshold=25 **then**
                Exclude branch $t*$ from T_{sub_k}
Fit arbitrary local plane L using final key points K_{final} and adjacent points K_{adj}.

3 Experiments and Results

3.1 Dataset

An abdominal dataset is collected from 27 patients diagnosed with aneurysms. For each patient, it includes a preoperative CTA data obtained using a GE Revolution EVO CT scanner and some intraoperative X-ray angiography images. For CTA images, the reconstructed slice thickness ranges from $1mm$ to $3mm$ and in?plane spacing from $0.79mm$ to $1.34mm$. In addition, abdominal vascular mask, aneurysm mask and four key bifurcations from each CTA data are manually annotated using open source 3D-Slicer software. The four key bifurcations are the renal artery bifurcation, iliac bifurcation, and the bifurcations of the left and right common iliac arteries, as shown in Fig. 1(B).

3.2 Results on Vascular and Aneurysm Segmentation

The experiment results of vascular and aneurysm segmentation by five-fold cross validation are shown in Table 1. Obviously, the vascular and aneurysm segmentation show strong performance, laying a solid foundation for subsequent geometric analysis.

3.3 Results on Semantic Key Bifurcation Detection

The performance of semantic key bifurcation detection is evaluated by calculating the mean distance and variance between predicted coordinates and ground

Table 1. Experiment results on vascular and aneurysm segmentation by five-fold cross validation

	Dice	Precision	Recall	F1-Score	IoU
Vessel Segmentation	0.975	0.964	0.986	0.975	0.952
Aneurysm Segmentation	0.866	0.933	0.831	0.879	0.770

Table 2. Ablation study on semantic bifurcation detection evaluated by L1 distance error

L_{WMSE}	L_{center}	Input	Mean/mm ↓	Variance/mm^2 ↓
✓	✓	Segmentation	**10.968**	**8.954**
✓	✓	Raw CTA	22.313	165.515
✓		Segmentation	18.811	89.561
	✓	Segmentation	99.506	12057.061

truth coordinates of each key bifurcation. The performance of our method is shown in the first row of Table 2. Meanwhile, Table 2 also reflects outcomes from ablation studies testing various combinations of inputs and loss functions. Compared with taking raw CTA data as input, taking vascular segmentation as input has more advantages, resulting in less errors in the mean distance and distance criterions. While employing L_{WMSE} alone yields commendable regression outcomes, it is crucial to note that this loss function does not directly target bifurcation coordinates. Therefore, the integration of L_{center} loss can further enhances our results. Interestingly, when only L_{center} is used, it leads to the worst performance.

3.4 Results on View Planning

An ablation study is conducted on view pose solution, as shown in Table 3. The PnP re-projection error is employed as evaluation metric. From Table 3, traditional PnP algorithm fails in solving view pose. And our experiment results demonstrate that the RANSAC PnP algorithm [1] combined with the Levenberg-Marquardt (LM) algorithm [15] achieves very stable and accurate view pose solution.

Table 3. Ablation study on view pose solution evaluated by re-projection error

Algorithms	Mean /mm ↓	Variance/mm^2 ↓
Traditional PnP	3333.945	2.803×10^7
RANSAC PnP	1.060	0.060
RANSAC PnP + LM	**0.565**	**0.043**

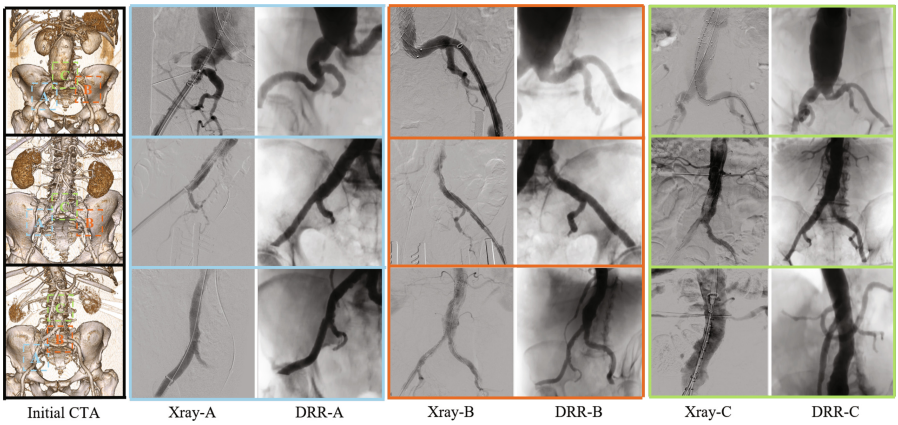

Fig. 2. Visual comparison between real X-rays and DRRs generated from planned view poses in three patients. The locations of the comparison correspond to the 3D visualization of the CTA and are marked with different colors.

Since our work is the first to propose multi-view planning for EVAR procedures, there are no existing works available for direct comparison. To validate the effectiveness of our proposed framework, we compared our results with intra-operative X-Ray angiography images, as illustrated in Fig. 2. Compared to the X-ray images used by surgeons during intervention, the DRR images generated using our planned views are very similar to them and show clearer vascular projection anatomy. The surgeons further evaluated the planning views generated by our proposed framework in these 27 patients and they were confident that these planning views were sufficient to guide the surgery. Additionally, more planning views are shown in Fig. 3. Our proposed method can not only provide a view based on a single key point of interest, but also coordinate multiple key points

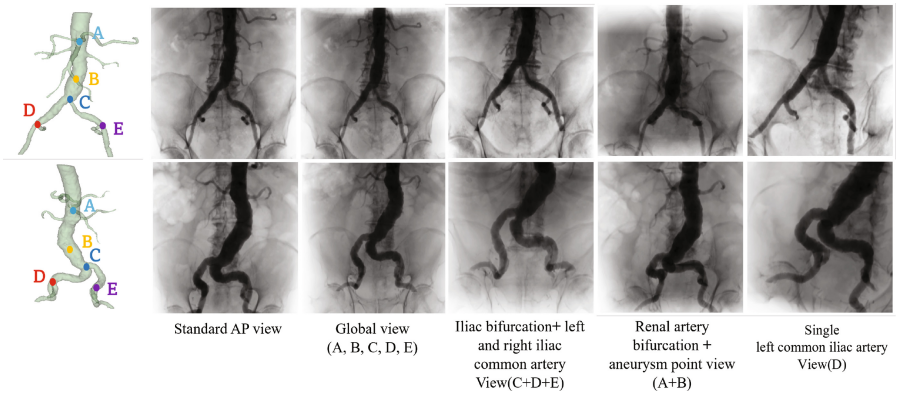

Fig. 3. Different views based on various key points of interest.

of interest to plan the view. Compared to standard AP view, the planned views help surgeons understand complex vascular anatomy more easily and quickly, showing less overlap and clearer branching structures.

4 Conclusion

In this paper, we present an efficient framework that automatically provides multiple optimal views for guiding EVAR procedures. By utilizing 3D nnUNet for precise vascular segmentation and semantic bifurcation detection, we construct a semantic vessel tree model integrating geometric and semantic information. This model assists in identifying optimal viewing planes at critical bifurcations. We then accurately determine the C-arm pose using the RANSAC Perspective-n-Point algorithm. Our framework helps address challenges such as excessive radiation exposure, high contrast agent doses, and time-consumption repositioning, streamlining the surgical planning process and enabling patient-specific strategies. Validated with an in-house dataset from 27 patients, our system demonstrates reliability, effectiveness, and a practical runtime of 6 min per patient. Future work will focus on further refining the framework to enhance its adaptability to a wider range of vascular surgeries and integrating preoperative and intra-operative registration modules to smoothly apply preoperative view planning to intra-operative settings.

Acknowledgments. The project was supported by the Bavarian State Ministry of Science and Arts within the framework of the"Digitaler Herz-OP" project under the grant number 1530/891 02 and the China Scholarship Council (File No.202004910390). We also thank BrainLab AG for their partial support.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Chum, O., Matas, J., Kittler, J.: Locally optimized RANSAC. In: Michaelis, B., Krell, G. (eds.) DAGM 2003. LNCS, vol. 2781, pp. 236–243. Springer, Heidelberg (2003). https://doi.org/10.1007/978-3-540-45243-0_31
2. Fallavollita, P., et al.: Desired-view controlled positioning of angiographic C-arms. In: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (eds.) MICCAI 2014. LNCS, vol. 8674, pp. 659–666. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-10470-6_82
3. Fantazzini, A., et al.: 3D automatic segmentation of aortic computed tomography angiography combining multi-view 2D convolutional neural networks. Cardiovasc. Eng. Technol. **11**, 576–586 (2020)
4. Geng, Z., Sun, K., Xiao, B., Zhang, Z., Wang, J.: Bottom-up human pose estimation via disentangled keypoint regression. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 14676–14686 (2021)
5. Goroshin, R., Mathieu, M.F., LeCun, Y.: Learning to linearize under uncertainty. Adv. Neural Inf. Process. Syst. **28** (2015)

6. He, J., et al.: Learning hybrid representations for automatic 3D vessel centerline extraction. In: Martel, A.L., et al. (eds.) MICCAI 2020. LNCS, vol. 12266, pp. 24–34. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-59725-2_3
7. 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 **18**(2), 203–211 (2021)
8. Kappe, K.O., Smorenburg, S.P., Hoksbergen, A.W., Wolterink, J.M., Yeung, K.K.: Deep learning-based intraoperative stent graft segmentation on completion digital subtraction angiography during endovascular aneurysm repair. J. Endovasc. Ther. **30**(6), 822–827 (2023)
9. Kausch, L., et al.: Toward automatic C-arm positioning for standard projections in orthopedic surgery. Int. J. Comput. Assist. Radiol. Surg. **15**, 1095–1105 (2020)
10. Kim, S., et al.: Deep learning on multiphysical features and hemodynamic modeling for abdominal aortic aneurysm growth prediction. IEEE Trans. Med. Imag. **42**(1), 196–208 (2022)
11. Li, J., Su, W., Wang, Z.: Simple pose: rethinking and improving a bottom-up approach for multi-person pose estimation. In: Proceedings of the AAAI Conference on Artificial Intelligence, vol. 34, pp. 11354–11361 (2020)
12. Luvizon, D.C., Tabia, H., Picard, D.: Human pose regression by combining indirect part detection and contextual information. Comput. Graph. **85**, 15–22 (2019)
13. Meng, C., Wang, Q., Guan, S., Sun, K., Liu, B.: 2d–3d registration with weighted local mutual information in vascular interventions. IEEE Access **7**, 162629–162638 (2019)
14. 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
15. Ranganathan, A.: The Levenberg-Marquardt algorithm. Tutoral LM Algorithm **11**(1), 101–110 (2004)
16. Tehlan, K., Winkler, A., Roth, D., Navab, N.: X-ray device positioning with augmented reality visual feedback. In: 2022 IEEE Conference on Virtual Reality and 3D User Interfaces Abstracts and Workshops (VRW), pp. 870–871. IEEE (2022)
17. Timmins, K.M., Van der Schaaf, I.C., Vos, I.N., Ruigrok, Y.M., Velthuis, B.K., Kuijf, H.J.: Geometric deep learning using vascular surface meshes for modality-independent unruptured intracranial aneurysm detection. IEEE Trans. Med. Imag. **42**(11), 3451–3460 (2023)
18. Virga, S., Dogeanu, V., Fallavollita, P., Ghotbi, R., Navab, N., Demirci, S.: Optimal C-arm positioning for aortic interventions. In: Handels, H., Deserno, T.M., Meinzer, H.P., Tolxdorff, T. (eds.) Bildverarbeitung für die Medizin 2015. I, pp. 53–58. Springer, Heidelberg (2015). https://doi.org/10.1007/978-3-662-46224-9_11
19. Wang, Y., et al.: Deep learning model for predicting the outcome of endovascular abdominal aortic aneurysm repair. Indian J. Surg. **85**(Suppl 1), 288–296 (2023)
20. Yao, L., et al.: TaG-Net: topology-aware graph network for centerline-based vessel labeling. IEEE Trans. Med. Imag. 42(11), 3155–3166 (2023)
21. Zhang, B., et al.: A patient-specific self-supervised model for automatic X-ray/CT registration. In: Greenspan, H., et al. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2023. MICCAI 2023. LNCS, vol. 14228. Springer, Cham (2023). https://doi.org/10.1007/978-3-031-43996-4_49