



Coronary Artery Calcification segmentation by using cross-frequency conditioner and geometric priors Learning

Weili Jiang^{1,2}, Yiming Li³, Gadeng Luosang⁴, Gang Peng⁵, Yong Peng³, Yijun Yao³, Zhang Yi², Jianyong Wang² , and Mao Chen³ 


¹ School of Computing and Artificial Intelligence, Southwest Jiaotong University, Chengdu, Sichuan, P. R. China

² Machine Intelligence Laboratory, College of Computer Science, Sichuan University, Chengdu, Sichuan, P. R. China

³ Department of Cardiology, West China Hospital, Sichuan University, Chengdu, Sichuan, P. R. China

⁴ Department of Information Science and Technology, Tibet University, Lhasa, P. R. China

⁵ Department of Cardiology, Chengdu First People's Hospital, Chengdu 610041, Sichuan Province, P. R. China

 *corresponding authors* wjy@scu.edu.cn (Jianyong Wang),
hmaochen@vip.sina.com (Mao Chen)

Abstract. Coronary artery calcification (CAC) is a powerful indicator of cardiovascular disease. Cardiac CT angiography (CCTA) has significant advantages in detecting CAC. However, since the image quality of CCTA can be compromised by cardiac motion or imaging equipment, and the contrast between CAC and surrounding tissue is low, accurate assessment of CAC remains a significant challenge. To address this issue, we propose a model (CAC-Net) for the comprehensive evaluation of CAC to fully exploit the characteristics of clinician annotations. First, inspired by the clinical annotation process, where doctors determine the subject based on boundaries, we propose a cross-frequency regulator module. This module models the interaction between high and low frequencies to distinguish the CAC body and its edges, thereby enhancing edge perception. Then, building on clinicians' anatomical prior knowledge that CAC is confined within coronary arteries, we introduce a geometric prior module to encode their topological relationship, effectively reducing false positives. In experiments, our proposed method is compared with existing state-of-the-art methods on two CAC datasets. The results demonstrate that: (1) our method significantly improves CAC segmentation performance, as evidenced by a higher Dice score compared to U-Net (0.731 vs. 0.659); and (2) it ensures consistency in clinically relevant indicators, including calcium scores.

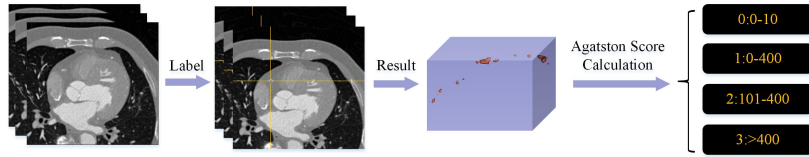
Keywords: Coronary Artery Calcification Segmentation · Frequency Learning · Geometric Priors Constraints · CCTA.

*** the main work of weili jiang was conducted during her phd at sichuan university.

1 Introduction

Cardiovascular disease (CVD) is the leading cause of mortality worldwide, with coronary artery calcification (CAC) as a key predictor [22]. Coronary CT angiography (CCTA) best visualizes CAC, making accurate assessment crucial for CVD prevention [8]. However, manual CAC evaluation is time-consuming, expertise-dependent, and individual heterogeneity [8]. Thus, efficient automated segmentation is essential to streamline diagnosis and support AI-assisted analysis [25].

(A) Clinical Task: Coronary Artery Calcification Segmentation and Scoring



(B) Challenges of Coronary Artery Calcification Segmentation

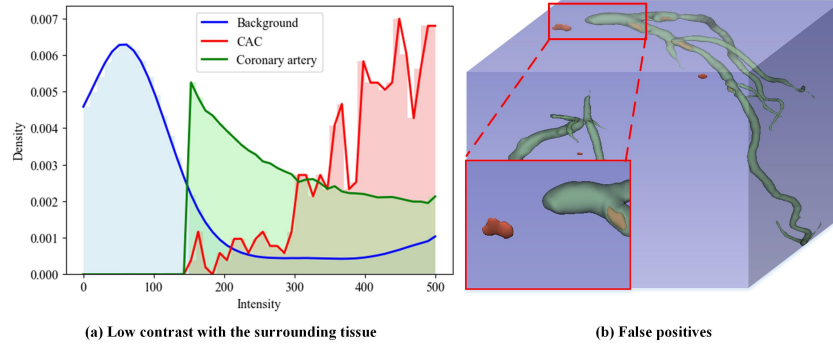


Fig. 1: The process of comprehensively evaluating CAC and the challenges of CAC segmentation

Figure 1(A) illustrates a fully automated CAC analysis pipeline for CCTA. The primary challenge in automation lies in CAC segmentation. Clinical experts rely on both local intensity features and contextual dependencies when interpreting CCTA images (Figure 1(B)). Locally, CAC appears as high-density bright regions, strictly confined to coronary arteries. Traditional manual review leverages these geometric and semantic priors for accurate identification. Earlier methods, including support vector machines [24] and decision trees [20], were applied to CAC detection in calcium scoring CT but are unsuitable for CCTA due to their reliance on a fixed 130 HU threshold. In CCTA, distinguishing CAC from attenuating cavities requires a more adaptive approach [25]. Deep learning methods have advanced CAC segmentation, with prior work leveraging Unet [1],

SegNet [2], and 3D CNN-based approaches [28,16]. Recently, transformer-based architectures have further improved segmentation performance [17,15].

Despite these advancements, challenges remain: 1) CAC intensity in CCTA varies significantly, complicating differentiation from background structures. Prior methods [9,28,11,4] enhance grayscale contrast but often ignore contextual cues, reducing robustness against noise and contrast agent interference. 2) Many approaches [6,21,13,17] lack geometric and semantic priors, leading to increased false positives in CAC segmentation (Figure 1(B)). 3) Existing methods often provide image-based classifications [9,12] but lack precise volume and location information, limiting clinical applicability.

To address these limitations, we propose CAC-Net, which mimics clinicians' annotation process by first locating CAC within coronary arteries and then refining its boundaries for improved segmentation accuracy. The main contributions of this work are: • A cross-frequency conditioner module (CFC) models the interaction between high- and low- frequencies to represent the CAC body and its edges, effectively simulating how clinicians determine the structure based on boundaries. • A geometric prior module encodes the topological relationship between the coronary arteries and CAC, simulating clinicians' anatomical prior knowledge that CAC is confined to the coronary arteries, thus reducing false positives. • A comprehensive evaluation across pixel-, image-, and lesion-level metrics, assessing segmentation accuracy and clinical relevance, including calcium scoring.

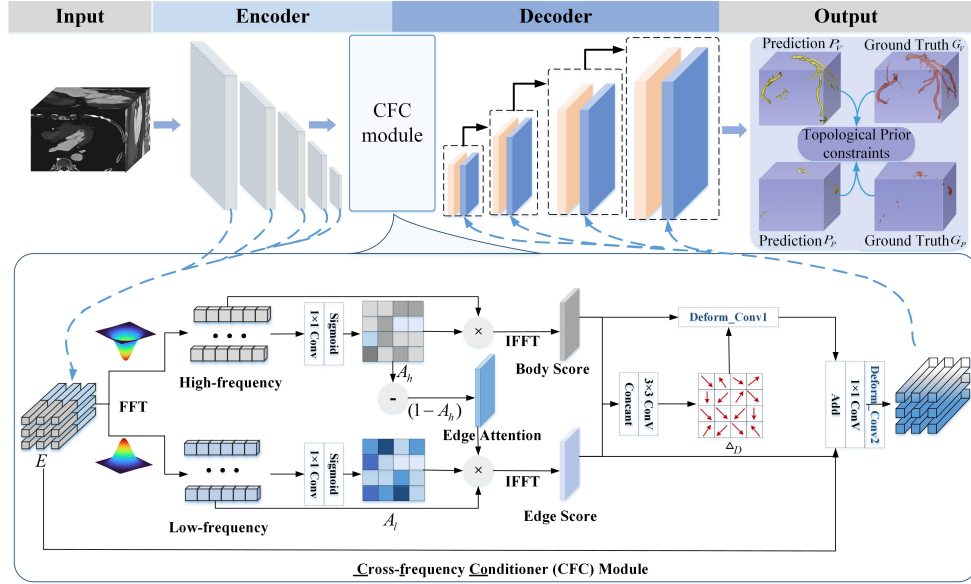


Fig. 2: Overview of the proposed method.

2 Methods

As illustrated in Figure 2, the proposed network adopts an encoder-decoder structure similar to 3D U-Net. To enhance the CAC boundary detail perception, we introduce the cross-frequency conditioner (CFC) module. Additionally, a geometric prior module is designed to fully leverage the spatial relationship between CAC and coronary arteries.

2.1 Cross-frequency Conditioner Module

Frequency disentanglement. Segmenting CAC in CCTA is challenging due to contrast agent interference. To improve accuracy, we simulate clinicians’ annotation habits by emphasizing CAC edge regions. To tackle this, we introduce a frequency unwrapping module (Figure 2), which applies high-pass and low-pass filters to decompose input features in the frequency domain. In practice, low-level features capture fine-grained details, beneficial for high-quality segmentation. Following [26], we transform the input image $X \in \mathbb{R}^{w \times h \times d}$ into multi-scale encoder features $E \in \mathbb{R}^{C \times \frac{w}{16} \times \frac{h}{16} \times \frac{d}{16}}$. The frequency disentanglement module then performs a 3D FFT to convert E to the frequency domain, splitting it into high-frequency (E_h) and low-frequency (E_l) components:

$$\begin{cases} E_l = \begin{cases} (\mathcal{F}(E))_{uvz}, & \frac{U,V,Z}{4} < u, v, z \leq \frac{3 \times (U,V,Z)}{4} \\ 0, & \text{others} \end{cases} \\ E_h = \begin{cases} 0, & \frac{U,V,Z}{4} < u, v, z \leq \frac{3 \times (U,V,Z)}{4} \\ (\mathcal{F}(E))_{uvz}, & \text{others} \end{cases} \end{cases} \quad (1)$$

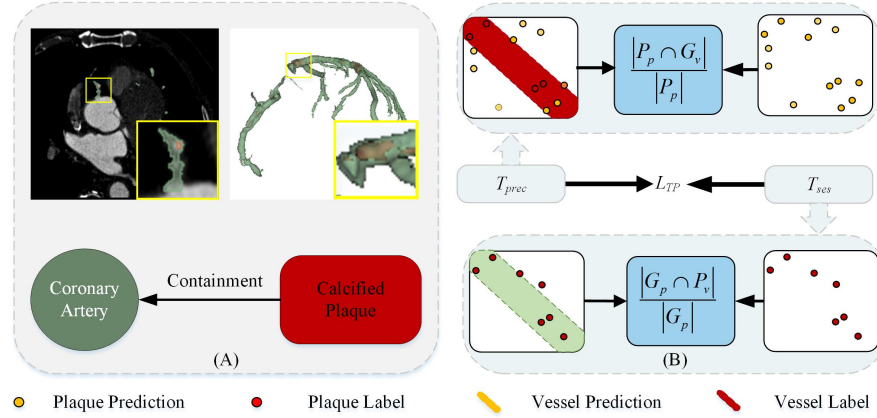


Fig. 3: Overview of the Geometric Prior Module.

Cross-frequency interaction. To refine segmentation, we propose a cross-frequency conditioner module leveraging frequency-aware attention to suppress

irrelevant features. The attention weights A_h and A_l are computed using linear transformers and normalized via a sigmoid function:

$$\begin{cases} A_l = \sigma(C_{1 \times 1}(E_l \| \theta_l)) \\ A_h = \sigma(C_{1 \times 1}(E_h \| \theta_h)) \end{cases}, \quad (2)$$

where $C_{1 \times 1}$ represents a 1×1 convolution, and θ_l, θ_h are learnable parameters for E_l and E_h . CCTA images can be decomposed into low-frequency smooth structures and high-frequency fine details. Since CAC segmentation relies on body and edge information, we define:

$$\begin{cases} E_h^{bod} = \mathcal{F}^{-1}(A_h \times E_h) \\ E_l^{edg} = \mathcal{F}^{-1}((1 - A_h) \times A_l \times E_l) \end{cases}, \quad (3)$$

where \mathcal{F}^{-1} denotes the inverse 3D FFT, converting spectral features back to image space.

Feature Alignment. Frequency disentanglement can introduce spatial misalignment between the extracted frequency-sensitive features and original encoder features, degrading segmentation performance. To address this, we introduce a feature alignment module for adaptive feature fusion (Figure 2). It consists of two deformable convolutions [7] (Deform_Conv1 aligns D_h^{bod} with D_l^{edg} by computing an offset map Δ_D and Deform_Conv2 refines the fused features):

$$\begin{cases} \Delta_D = \text{Conv}_{3 \times 3}(D_h^{Bod} \otimes D_l^{edg}) \\ E'_f = \text{De_Conv}_{3 \times 3}(\text{De}(\Delta_D * D_h^{Bod}) \oplus D_l^{edg} \oplus E) \end{cases}, \quad (4)$$

where \otimes denotes concatenation, \oplus represents addition, and De denotes deformable convolutions. The final fused feature representation is $E' = E'_e + E'_f$. This approach integrates frequency-domain insights while preserving CNN-based spatial cues, ensuring accurate CAC segmentation.

2.2 Geometric Prior Module

Geometric Interactions Constraint. Although frequency-domain analysis helps distinguish CAC and distractors from the background, their structural similarity leads to high false positives in segmentation. To address this, we simulate the prior knowledge of clinicians that plaques are located in coronary arteries (Figure 3) and propose topological interaction constraints to model this relationship, effectively reducing false positives. Specifically, we define four binary masks: plaque segmentation mask Gp , coronary artery mask Gv , plaque prediction mask Pp , and coronary artery prediction mask Pv . To quantify the alignment between predicted plaques and real vessels, we define geometric prior accuracy (T_{prec}) and geometric prior sensitivity (T_{sens}) as:

$$\begin{cases} T_{\text{prec}}(P_p, G_v) = \frac{|P_p \cap G_v|}{|P_p|} \\ T_{\text{sens}}(G_p, P_v) = \frac{|G_p \cap P_v|}{|G_p|} \end{cases}. \quad (5)$$

Geometric-prior Loss. A high T_{prec} indicates a tendency for false positives, while a high T_{sens} suggests potential false negatives. To mitigate these errors, we maximize both metrics by defining the geometric-prior loss L_{gp} as their harmonic mean:

$$L_{gp} = 2 \times \frac{T_{\text{prec}} (P_p, G_v) \times T_{\text{sens}} (G_p, P_v)}{T_{\text{prec}} (P_p, G_v) + T_{\text{sens}} (G_p, P_v)}. \quad (6)$$

This loss function enforces topological consistency, improving segmentation precision. To further refine predictions, we combine it with standard soft Dice and cross-entropy losses, resulting in the final objective function:

$$L_{\text{total}} = L_{ce} + \lambda_{\text{dice}} L_{\text{dice}} + \lambda_{gp} L_{gp}, \quad (7)$$

where L_{ce} and L_{dice} are Cross-entropy and Dice losses, respectively, with weights λ_{dice} and λ_{gp} controlling their contributions.

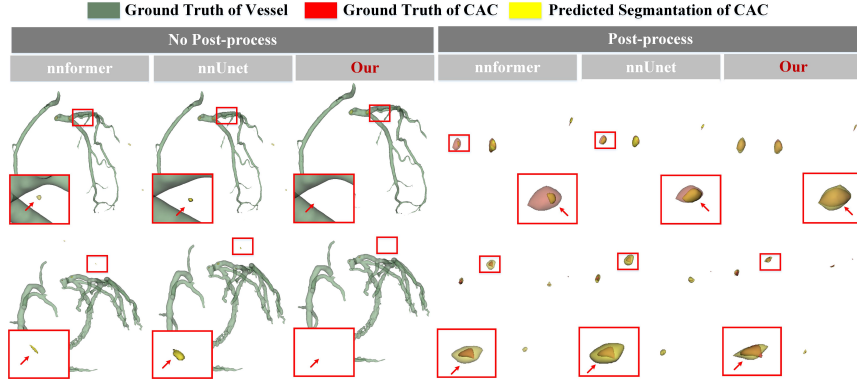


Fig. 4: Visualization of CAC segmentation results before and after utilizing coronary artery containing CAC post-processing.

3 Experiments

3.1 Datasets and Data-processing

CAC-CTA dataset. This dataset comprises 150 patients from a top-tier Chinese hospital, encompassing 802 instances of coronary artery calcification (CAC), with an average of 5.36 ± 4.92 lesions per patient. CAC-CTA scans were acquired using a uniform Siemens dual-source CT system under standardized protocols. Patient demographics were recorded, with a mean age of 64.45 ± 11.44 years. The datasets were independently and anonymously annotated by cardiologists and radiologists using 3D Slicer. Annotation consistency was assessed using Dice similarity, with discrepancies rechecked. For training and evaluation, 80% of the images were used for training, and 20% for testing.

Table 1: Quantitative comparison and Cross-validation result

Method	No Post-process			Post-process			No Post-process			Post-process		
	SE	PE	Dice	SE	PE	Dice	SE	PE	Dice	SE	PE	Dice
	Quantitative comparison result						Cross-validation result					
Unet [5]	0.627	0.357	0.385	0.627	0.774	0.659	0.603	0.323	0.231	0.603	0.754	0.629
Vnet [19]	0.646	0.412	0.415	0.646	0.756	0.672	0.625	0.336	0.408	0.625	0.736	0.651
GLIA-Net [3]	0.655	0.467	0.475	0.655	0.77	0.714	0.637	0.409	0.454	0.637	0.753	0.693
nnUnet [14]	0.76	0.685	0.712	0.76	0.782	0.725	0.725	0.696	0.713	0.725	<u>0.761</u>	0.707
U-Transformer [18]	0.656	0.481	0.481	0.656	0.702	0.698	0.641	0.427	0.467	0.641	0.762	0.672
SwinUNETR [10]	0.671	0.531	0.581	0.671	0.531	0.718	0.651	0.527	0.618	0.651	0.511	0.704
nnformer [27]	0.721	<u>0.701</u>	0.612	0.721	0.752	0.719	0.695	0.559	0.709	0.695	0.743	0.698
CACer [15]	<u>0.767</u>	0.699	<u>0.715</u>	<u>0.767</u>	0.758	<u>0.727</u>	<u>0.731</u>	<u>0.713</u>	<u>0.717</u>	<u>0.731</u>	0.759	<u>0.719</u>
Our	0.784	0.759	0.725	0.784	<u>0.763</u>	0.731	0.754	0.737	0.734	0.754	0.756	0.731

Orcascore dataset. We use the MICCAI 2014 Challenge cardiac CT dataset [23], containing scans from four hospitals using different CT scanners. Each patient has both non-contrast CT and contrast-enhanced CCTA, with 32 patients annotated for CAC. CCTA segmentation results are validated against non-contrast CT.

Data-processing of CAC dataset. CTA images are resampled to 0.5 mm thickness. A -224 HU to 600 HU threshold is applied for coarse lung segmentation, followed by seed-filling for refinement. The lung mask is subtracted from the original image to reduce noise and improve CAC segmentation.

3.2 Implementation Details and Evaluation Metric

Network parameters. All experiments were conducted using Python 3.8 and PyTorch 1.8⁶ on NVIDIA Tesla P100 GPUs. The model is optimized using SGD with an initial learning rate of 0.01, weight decay of 2e-4, and ReduceLROnPlateau scheduling (coefficient 0.5, patience and cooldown 3, minimum learning rate 1e-8). During training, a $128 \times 128 \times 128$ block is randomly cropped from CCTA images, normalized, and fed into the network. The batch size is set to 2, and each experiment runs for 120 stages. The final segmentation is obtained by binarizing the model’s probability map using a 0.5 threshold.

Evaluation metric. We assess performance in two aspects: (1) Segmentation Accuracy (sensitivity (SE), precision (PE), and Dice) and (2) Geometric Prior Influence, evaluated by comparing results before and after post-processing with coronary artery masks.

3.3 Comparison with State-of-the-Art Method

Comparison with SOTA Method. As shown in Table 1, post-processing improves Dice scores by reducing false positives through geometric priors, but precision (PE) does not reach SOTA. This occurs because post-processing, while eliminating false positives outside the coronary arteries, may also filter out small

⁶ <https://pytorch.org/>

true CAC regions, slightly reducing precision. Figure 4 further demonstrates that our method better preserves fine CAC structures compared to others. Overall, our model achieves higher segmentation index, demonstrating the effectiveness of the cross-frequency conditioner module and geometric prior module in improving CAC detection.

Generalization. This paper employs cross-validation to evaluate the model’s generalization. The CAC-CTA dataset was used for training, while the OrcaScore dataset served as the test set. As shown in Table 1, our model outperforms nnU-Net, nnFormer, and CACer, achieving consistently higher scores across all metrics, both before and after pre-processing.

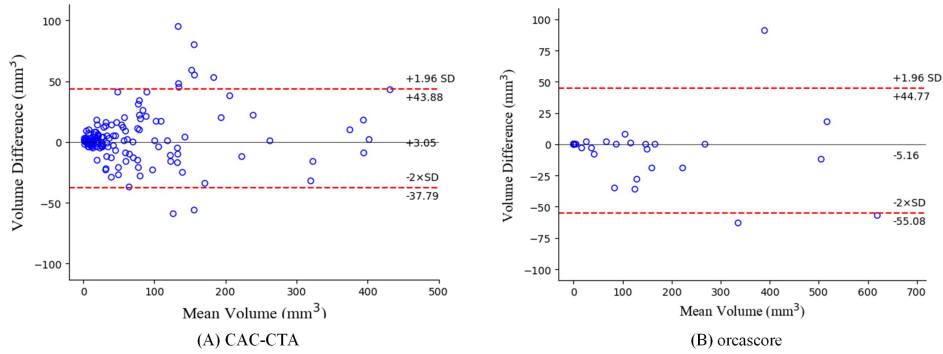


Fig. 5: The Bland-Altman plot compares the automatically acquired CAC volumes from the CCTA segmentation method with reference annotations in CCTA or CT.

Per Patient CAC Quantification. For each patient, cardiologists and radiologists recorded CCTA volume and mass scores based on the Agatston score, which our CAC-Net also predicts. The Bland-Altman plot in Figure 5 illustrates the agreement between predicted and true Agatston scores. Orcascore closely matches the CT gold standard, while on the CAC-CTA dataset, automatic scores tend to be higher than the reference (+3.05 deviation). Conversely, on the Orcascore dataset, automated scores are lower, with a consistency limit of -5.16.

3.4 Ablation Study

To evaluate the effectiveness of each component within the proposed model, we performed an ablation study on the CTA-CAC dataset. As shown in Table 2, each component significantly improves segmentation performance, demonstrating the effectiveness of both the cross-frequency conditioner and geometric prior module.

Table 2: Ablation study of our model on CAC-CTA dataset. ("ONE" denotes one-class task. "TWO" denotes two-class task. "FAS" denotes the cross-frequency conditioner module, and "TP" denotes TP LOSS).

One	Two	FAS	TP	No Post-process			Post-process		
				SE	PE	Dice	SE	PE	Dice
✓				0.627	0.357	0.385	0.627	0.774	0.659
✓		✓		0.676	0.631	0.542	0.676	0.784	0.716
	✓			0.749	0.603	0.625	0.749	0.706	0.712
		✓	✓	0.745	0.664	0.668	0.745	0.749	0.722
	✓		✓	0.757	0.727	0.713	0.757	0.751	0.726
	✓	✓	✓	0.784	0.759	0.725	0.784	0.763	0.731

4 Conclusion

We propose a novel CAC segmentation model, leveraging CAC-Net to assess token connectivity and learn instance-specific attention patterns. Additionally, we introduce instance-aware guided semantic learning in the Fourier domain to enhance long-range feature interactions. Our model achieves state-of-the-art performance on CAC tasks, enabling accurate, automated CAC identification and quantification in CCTA. This advancement could eliminate the need for separate CSCT scans, reducing patient radiation exposure.

Acknowledgement. This work was supported by the National Natural Science Foundation of China (Grant No. 62476185, Grant No. 62306192 and Grant No. 62406257). The National Natural Science Foundation of China Major Project (Grant No. 62495064). The Sichuan Province Natural Science Foundation Project (2025ZNSFSC0739).

Disclosure of Interests. The authors have no competing interests.

References

1. Abdolmanafi, A., Cheriet, F., Duong, L., Ibrahim, R., Dahdah, N.: An automatic diagnostic system of coronary artery lesions in kawasaki disease using intravascular optical coherence tomography imaging. *Journal of Biophotonics* **13**(1), e201900112 (2020)
2. Badrinarayanan, V., Kendall, A., Cipolla, R.: Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *IEEE transactions on pattern analysis and machine intelligence* **39**(12), 2481–2495 (2017)
3. Bo, Z.H., Qiao, H., Tian, C., Guo, Y., Li, W., Liang, T., Li, D., Liao, D., Zeng, X., Mei, L., et al.: Toward human intervention-free clinical diagnosis of intracranial aneurysm via deep neural network. *Patterns* **2**(2), 100197 (2021)
4. Chang, C.Y., Hong, Y.C., Tseng, C.h., et al.: A neural network for thyroid segmentation and volume estimation in ct images. *IEEE Computational Intelligence Magazine* **6**(4), 43–55 (2011)

5. Çiçek, Ö., Abdulkadir, A., Lienkamp, S.S., Brox, T., Ronneberger, O.: 3d u-net: learning dense volumetric segmentation from sparse annotation. In: International Conference on Medical Image Computing and Computer-assisted Intervention. pp. 424–432. Springer (2016)
6. Clough, J., Byrne, N., Oksuz, I., Zimmer, V.A., Schnabel, J.A., King, A.: A topological loss function for deep-learning based image segmentation using persistent homology. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2020)
7. Dai, J., Qi, H., Xiong, Y., Li, Y., Zhang, G., Hu, H., Wei, Y.: Deformable convolutional networks. In: Proceedings of the IEEE international conference on computer vision. pp. 764–773 (2017)
8. Föllmer, B., Williams, M.C., Dey, D., Arbab-Zadeh, A., Maurovich-Horvat, P., Volleberg, R.H., Rueckert, D., Schnabel, J.A., Newby, D.E., Dweck, M.R., et al.: Roadmap on the use of artificial intelligence for imaging of vulnerable atherosclerotic plaque in coronary arteries. *Nature Reviews Cardiology* **21**(1), 51–64 (2024)
9. Gessert, N., Lutz, M., Heyder, M., Latus, S., Leistner, D.M., Abdelwahed, Y.S., Schlaefer, A.: Automatic plaque detection in ivoct pullbacks using convolutional neural networks. *IEEE transactions on medical imaging* **38**(2), 426–434 (2018)
10. Hatamizadeh, A., Nath, V., Tang, Y., Yang, D., Roth, H.R., Xu, D.: Swin unetr: Swin transformers for semantic segmentation of brain tumors in mri images. In: International MICCAI Brainlesion Workshop. pp. 272–284. Springer (2021)
11. He, C., Wang, J., Yin, Y., Li, Z.: Automated classification of coronary plaque calcification in oct pullbacks with 3d deep neural networks. *Journal of Biomedical Optics* **25**(9), 095003 (2020)
12. 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. pp. 770–778 (2016)
13. Hu, X., Li, F., Samaras, D., Chen, C.: Topology-preserving deep image segmentation. *Advances in Neural Information Processing systems* **32** (2019)
14. 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. *Nature Methods* **18**(2), 203–211 (2021)
15. Jiang, W., Li, Y., Yi, Z., Wang, J., Chen, M.: Iarcac: Instance-aware representation for coronary artery calcification segmentation in cardiac ct angiography. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 285–295. Springer (2024)
16. Li, C., Jia, H., Tian, J., He, C., Lu, F., Li, K., Gong, Y., Hu, S., Yu, B., Wang, Z.: Comprehensive assessment of coronary calcification in intravascular oct using a spatial-temporal encoder-decoder network. *IEEE Transactions on Medical Imaging* **41**(4), 857–868 (2021)
17. Liu, Y., Nezami, F.R., Edelman, E.R.: A transformer-based pyramid network for coronary calcified plaque segmentation in intravascular optical coherence tomography images. *Computerized Medical Imaging and Graphics* p. 102347 (2024)
18. Petit, O., Thome, N., Rambour, C., Themyr, L., Collins, T., Soler, L.: U-net transformer: Self and cross attention for medical image segmentation. In: International Workshop on Machine Learning in Medical Imaging. pp. 267–276. Springer (2021)
19. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. In: International Conference on Medical Image Computing and Computer-assisted Intervention. pp. 234–241. Springer (2015)
20. Shahzad, R., van Walsum, T., Schaap, M., Rossi, A., Klein, S., Weustink, A.C., de Feyter, P.J., van Vliet, L.J., Niessen, W.J.: Vessel specific coronary artery calcium scoring: an automatic system. *Academic Radiology* **20**(1), 1–9 (2013)

21. Van Herten, R.L., Hampe, N., Takx, R.A., Franssen, K.J., Wang, Y., Suchá, D., Henriques, J.P., Leiner, T., Planken, R.N., Išgum, I.: Automatic coronary artery plaque quantification and cad-rads prediction using mesh priors. *IEEE Transactions on Medical Imaging* (2023)
22. Wang, X., Matsumura, M., Mintz, G.S., Lee, T., Zhang, W., Cao, Y., Fujino, A., Lin, Y., Usui, E., Kanaji, Y., et al.: In vivo calcium detection by comparing optical coherence tomography, intravascular ultrasound, and angiography. *JACC: Cardiovascular Imaging* **10**(8), 869–879 (2017)
23. Wolterink, J.M., Leiner, T., De Vos, B.D., Coatrieux, J.L., Kelm, B.M., Kondo, S., Salgado, R.A., Shahzad, R., Shu, H., Snoeren, M., et al.: An evaluation of automatic coronary artery calcium scoring methods with cardiac ct using the orcascore framework. *Medical physics* **43**(5), 2361–2373 (2016)
24. Wolterink, J.M., Leiner, T., Viergever, M.A., Išgum, I.: Automatic coronary calcium scoring in cardiac ct angiography using convolutional neural networks. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*. pp. 589–596. Springer (2015)
25. Wolterink, J.M., Leiner, T., de Vos, B.D., van Hamersvelt, R.W., Viergever, M.A., Išgum, I.: Automatic coronary artery calcium scoring in cardiac ct angiography using paired convolutional neural networks. *Medical Image Analysis* **34**, 123–136 (2016)
26. Zhang, W., Huang, Z., Luo, G., Chen, T., Wang, X., Liu, W., Yu, G., Shen, C.: Topformer: Token pyramid transformer for mobile semantic segmentation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. pp. 12083–12093 (2022)
27. Zhou, H.Y., Guo, J., Zhang, Y., Han, X., Yu, L., Wang, L., Yu, Y.: nnformer: Volumetric medical image segmentation via a 3d transformer. *IEEE Transactions on Image Processing* (2023)
28. Zhou, R., Guo, F., Azarpazhooh, M.R., Spence, J.D., Ukwatta, E., Ding, M., Fenster, A.: A voxel-based fully convolution network and continuous max-flow for carotid vessel-wall-volume segmentation from 3d ultrasound images. *IEEE Transactions on Medical Imaging* **39**(9), 2844–2855 (2020)