Wavelet High-Frequency Enhancement and Frame Dropout Compensation for Accurate Tracking Guidewire/Catheter in Carotid DSA

Xiaotian Yan, Yuanbo He*, Yuting Guo, Xinyu Zhang, Jinghao Li, Zitao Zhou, Ziyi Pei, Fan Gao, Peng Lun, Shuai Li, Peng Yu

***Abstract*— Digital subtraction angiography (DSA) based interventional surgery is the primary treatment modality for severe carotid artery stenosis (CAS). During the procedure, the 2D position information of guidewire/catheter provided by intraoperative DSA can be registered with the preoperative 3D reconstructed model of the patient's carotid artery, thereby providing real-time 3D visualization of the guidewire/catheter. Therefore, accurate tracking guidewire/catheter in carotid DSA is crucial for ensuring surgical success. However, due to the low contrast of DSA imaging and the potential for occlusion by lesions or other anatomical structures, accurate tracking of the guidewire/catheter remains a challenge. To address these challenges, this paper proposes a method for accurate tracking guidewire/catheter by combining wavelet high-frequency enhancement and frame dropout compensation. Specifically, wavelet enhancement is introduced to improve the detection of sharp high-frequency features, such as the shape contours of guidewire/catheter, and this is combined with interpolation-based compensation for lost frames to achieve continuous, stable, and highly accurate tracking guidewire/catheter in DSA. The proposed method improved the tracking performance of the two superior baseline models by 5.64% and 8.14%, respectively.**

***Clinical relevance*—As illustrated in Fig 4, accurate tracking guidewire/catheter during intraoperative DSA, when registered with the patient's preoperative 3D reconstruction model of the carotid artery, provides real-time 3D visualization guidance for the position of the guidewire/catheter to the surgeon.**

Research supported by the National Key Research and Development Program of China NO.2023YFF1203803, the National Natural Science Foundation of China No. 62272021, the Demonstration Project in Shandong Province No. 2022SFGC0601, the Postdoctoral Fellowship Program of CPSF under grant number GZC20233375.

Xiaotian Yan, Yuting Guo, Jinghao Li, Ziyi Pei, Peng Yu, Fan Gao, Shuai Li are with the State Key Laboratory of Virtual Reality Technology and Systems, Beihang University, Beijing, China.

Peng Lun is with the Neurosurgery Department, Affiliated Hospital of Qingdao University, Qingdao, China.

Yuanbo He is with the Institute of Electrophysiology, Henan Academy of Innovations in Medical Science, Zhengzhou, China (corresponding author email: heluxixue@163.com).

Xinyu Zhang, Zitao Zhou are with the Qingdao Virtual Reality Institute Co., Ltd, Qingdao, China.

I. INTRODUCTION

The carotid artery is the main blood vessel through which blood flows from the heart to the brain and other parts of the head. Atherosclerotic plaques in the carotid artery lumen can lead to CAS, which, when severe, may result in adverse clinical outcomes such as stroke or even death [1]. Interventional surgery based on DSA is one of the primary treatments for severe CAS. This procedure involves injecting contrast agents into the patient's arteries, allowing them to circulate to the lesion site and become visible under X-ray imaging. This process guides physicians to accurately deliver guidewire and catheter to the lesion site for diagnosis and local treatment. However, DSA only provides 2D positional information of the guidewire/catheter. Even experienced physicians find it challenging to discern the 3D spatial positions of these instruments in real time and must rely on their experience to perform the surgery. Registering the 2D positions of the guidewire/catheter provided by intraoperative DSA with the preoperative 3D reconstructed model of the patient's carotid artery can offer real-time 3D visualization of the guidewire/catheter. Therefore, accurate tracking guidewire/catheter in carotid artery DSA is crucial for ensuring surgical success.

With the widespread application and development of deep learning in medical image analysis, deep learning-based tracking methods have been employed for the tracking guidewire/catheter. Ullah et al. [2] first used the RCNN to detect the region of interest, followed by using the UNet for segmentation to achieve semantic-level tracking of the guidewire. Chen et al. [3] used UNet sequentially to localize the region of the guidewire in fluoroscopic images and extract the shape features within that region, thereby tracking the guidewire. Mei et al. [4] proposed a method combining YOLOv5 and Kalman filtering to track guidewire and catheter. However, current tracking methods often lose the targets of guidewire and catheter in challenging DSA frames, necessitating re-detection and tracking. If the detection algorithm has a low accuracy rate, continuous tracking cannot be maintained. Therefore, high-precision detection of guidewire and catheter is of great importance. Subramanian et al. [5] used UNet to detect catheter in chest X-ray images. Li et al. [6] used YOLOv3 to detect and localize the endpoints of guidewire in fluoroscopic images. Aghasizade et al. [7] implemented a two-stage deep convolutional model based on coordinate regression to detect key points and catheter in fluoroscopic images. This method improves the detection accuracy of catheters by first localizing the region of interest and then refining the search. Although these methods for detecting guidewire and catheter have achieved some success in specific scenarios, detecting these instruments in carotid artery DSA remains challenging due to the very low contrast of the images and the potential for occlusion by lesions or other anatomical structures. Therefore, a tracking method that can not only detect guidewire and catheter stably and accurately but also compensate for lost frames coherently holds significant clinical value.

To address these challenges, this paper proposes a method for accurate tracking guidewire/catheter by combining wavelet high-frequency enhancement with frame dropout compensation.

The specific innovative contributions are as follows:

1. A wavelet enhancement strategy is introduced to improve the detection of sharp high-frequency features, such as the shape and contour of guidewire and catheter, in low-contrast DSA images.

2. A frame interpolation strategy for lost frames in carotid artery DSA is introduced to prevent interruptions in tracking.

3. By combining wavelet high-frequency enhancement with frame interpolation, Our method achieves continuous, stable, and highly accurate tracking guidewire/catheter in carotid DSA.

II. Methodology

A. Overview

The framework of the proposed method is depicted in Fig 1. Specifically, the wavelet enhancement strategy described in Fig 1(A) and Section *B* achieves precise detection of catheter and guidewire in low-contrast DSA by enhancing the extraction of high-frequency information. Subsequently, the frame dropout compensation strategy shown in Fig 1(B) and Section *C* enables stable and continuous tracking guidewire/catheter in DSA.

B. Wavelet High-Frequency Enhancement based Detectinon (***WHFE***)

The wavelet transform possesses [8] excellent time-frequency localization capabilities, enabling it to provide localized information in both time and frequency domains for DSA images, thereby effectively capturing instantaneous features within DSA. This time-frequency localization characteristic allows the wavelet transform to precisely locate high-frequency components in the signal, such as moving and changing guidewire and catheter. Moreover, the multiscale analysis capability of the wavelet transform enables it

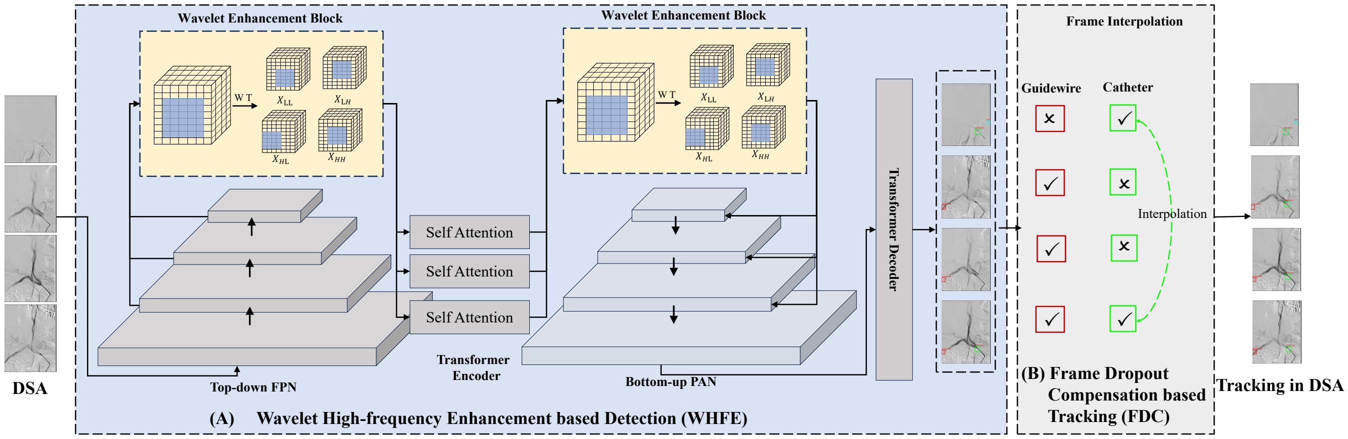

Fig. 1. The pipeline of the Wavelet High-Frequency Enhancement and Frame Dropout Compensation for Accurate Tracking Guidewire/Catheter in Carotid DSA.

to adapt to the features of consecutive DSA frames at different time scales. This multiscale analysis not only handles the dynamic changes between different frames but also better extracts high-frequency information in low-contrast DSA images. Through this approach, the wavelet transform can accurately locate the positions of guidewire and catheter, achieving precise detection of these high-frequency components even in cases of low contrast. Specifically, as shown in the following equation, this paper employs the Haar wavelet [9] to implement the wavelet high-frequency feature enhancement strategy.

$$X_{LL}, X_{LH}, X_{HL}, X_{HH} = \mathrm{WT}(X), \tag{1}$$

$$Y_{HH} = \mathrm{Conv}(W, (X_{HH})), \tag{2}$$

$$Output = \mathrm{Conv}(W, (X)) + \mathrm{Up}\uparrow(Y_{HH}), \tag{3}$$

where X represents a frame of the DSA image, X_{LL}, X_{LH}, X_{HL}, and X_{HH} denote the low-frequency component, horizontal high-frequency component, vertical high-frequency component, and diagonal high-frequency component, respectively. Conv represents the convolution operation, W is the convolutional weight, and WT denotes the Haar wavelet decomposition. Given that guidewire and catheter in the dataset predominantly exist in a diagonal orientation, this paper selects X_{HH} to enhance the high-frequency information extraction of guidewire and catheter. Up $\uparrow$ is the upsampling operation.

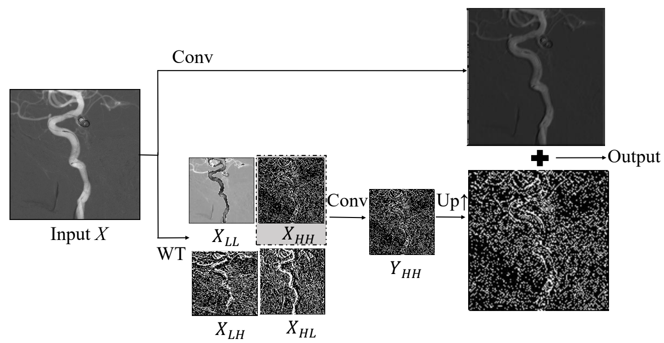

Fig. 2. Wavelet High-Frequency Enhancement.

Given the high detection accuracy of the DETR detection method [10], [11], as shown in Fig 1(A) and Fig 2, this paper combines the wavelet high-frequency enhancement strategy with the DETR detection method to achieve high-precision detection of guidewire and catheter.

*C. Frame Dropout Compensation based Tracking (**FDC**)*

If the detection results of DSA experience unforeseen failures, it will lead to the discontinuity of guidewire and catheter tracking. As a result, we propose a frame interpolation strategy to complete the lost tracking frames in carotid artery DSA based on traditional tracking methods [12]. The state of the trajectory is set to three values: active, lost, and removed. When an active tracking sequence fails to match a detection target in the current frame, the sequence is marked as lost. If it remains unmatched for a certain period, it is then marked as removed. However, if the sequence matches a detection result again before being removed, it will be reactivated, and the lost frames are completed by interpolation. The strategy is shown in algorithm 1.

Detected targets in DSA typically do not exhibit significant displacement over short periods. Therefore, we use IoU as the primary criterion for matching, and incorporate time distance as a compensation factor. For each detection result in the most recent frame, we calculate the optimal matching score with historical trajectories, as shown in eq 4, where box_t and box_d is the bounding box of the last frame of trajectory and the detection result, and T_t and T_d is the time stamps of the trajectory and the detection. Whether a match is established is determined by a predefined threshold, and the final matching relationship is based on the score magnitude, ensuring a one-to-one correspondence between trajectories and detection results.

$$matching\ score = IoU(box_t, box_d) + \alpha(T_t - T_d). \quad (4)$$

III. Experiments

A. Dataset

In this paper, a self-made dataset was adopted. The intraoperative DSA images of 200 carotid artery interventional surgeries in the Affiliated Hospital of Qingdao University in the past two years were used as the original data [1]. The labeling team was led by neurosurgeons and medical postgraduates. After converting the DICOM format images frame by frame into JPG format, LabelMe was used for labeling. There were three types of labels, namely catheter, guidewire, and spring. Since the changes between the front and rear frames of DSA images were small, key frame annotation was adopted to save manpower. The team annotated 6,900 key frames by skipping frames. After the labels of the intermediate frames were completed by code, a total of 46,475 DSA image datasets were

[1]The Ethical Approval Number is QYFYEC2023-53.

Algorithm 1 Tracking Strategy

Input: Current trajectories $curTracks$, current detection results $dets$

Output: Next trajectories $nextTracks$

$nextTracks \leftarrow []$
for $det \in dets$ **do**
  $track \leftarrow bestMatch(det, curTracks)$
  if $track\ is\ None$ **then**
    $nextTracks.add(newTrack(det))$
  else if $track.state == LOSS$ **then**
    $tracks \leftarrow interpolation(track.lastframe, det)$
    $nextTracks.addAll(tracks)$
    $track.state \leftarrow ACTIVATE$
  else if $track.state == ACTIVATE$ **then**
    $track.add(det)$
    $nextTracks.add(track)$
  end if
end for
for $track \in curTracks$ **do**
  if $track \notin nextTracks$ **then**
    $track.state \leftarrow LOSS$
    $nextTracks.add(track)$
  end if
end for
return $nextTracks$

obtained. There was little difference in the training results between the 6,900 and 46,475 datasets. Considering the time cost of training, after deleting the interfering images, the finally produced DSA dataset contained a total of 6,731 images. When designing the comparative experiment, the 6,731 datasets were divided into training sets and test sets at a ratio of 8:2 in units of DICOM. 5,407 images were used as the training set, and 1,324 images were used as the validation and test set for training. This dataset was used for all training methods and ablation experiments to conduct fair comparison and verification.

B. Implementation

Our algorithm was implemented under the PyTorch framework and the training was completed on

TABLE I

TRACKING RESULTS ON DIFFERENT METHODS. THE BOLD MEANS THE BEST.

Methods	HOTA ↑		MOTA ↑		IDF1 ↑	
	catheter	guidewire	catheter	guidewire	catheter	guidewire
RT-DETRV2[11]	61.54	60.07	57.96	53.01	74.80	71.42
RT-DETRV2 + WHFE	60.71	61.11	60.00	54.21	77.90	72.18
RT-DETRV2 + FDC	62.21	61.72	56.79	54.86	74.65	74.00
RT-DETRV2 + WHFE + FDC	**63.36**	**63.22**	**63.07**	**55.96**	**80.43**	**74.07**
Salience-DETR[13]	61.86	59.66	61.75	56.50	77.86	73.07
Salience-DETR + WHFE	62.68	60.39	66.72	58.08	79.81	74.16
Salience-DETR + FDC	63.15	60.91	61.61	56.27	78.30	73.88
Salience-DETR + WHFE + FDC	**64.01**	**64.21**	**68.47**	**64.86**	**81.19**	**78.75**

an NVIDIA GeForce RTX 4080 GPU. The original size of the dataset images was 256×256 pixels, and the input image size for the algorithm was 640×640 pixels. The batch size was set to 8, and the batch size in the non-freezing stage was 4. The learning rate was 1×10^{-5}. The evaluation metrics of the tracking included HOTA (Higher Order Tracking Accuracy), MOTA (Multiple Object Tracking Accuracy), and IDF1 (Identification F1 Score). A higher evaluation value indicates a better tracking result.

C. Results

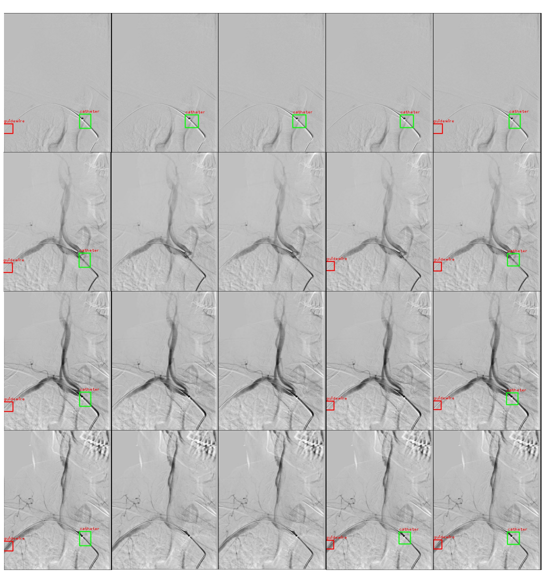

Fig. 3. Visualization of tracking results in ablation experiments. Red rectangle: guidewire, Green rectangle: catheter.

To verify the generalization ability of the proposed method, we employed two versions of the DETR detection method: RT-DETRV2[11] and Salience-DETR[13] , as detection baseline models. As shown in Table I, the introduction of **WHFE** generally improved the tracking performance of the methods. Specifically, while the tracking performance of **RT-DETRV2+WHFE** in terms of HOTA for the catheter slightly decreased (60.71 < 61.54), **RT-DETRV2+WHFE** achieved enhanced performance in both MOTA and IDF1 metrics. Additionally, **Salience-DETR + WHFE** demonstrated improved tracking performance in HOTA, MOTA, and IDF1 for both guidewire and catheter. These results collectively validate the effectiveness of the wavelet high-frequency enhancement proposed in this paper. Similarly, the introduction of **FDC** led to an overall improvement in tracking performance for both **RT-DETRV2+FDC** compared to **RT-DETRV2** and **Salience-DETR+FDC** compared to **Salience-DETR**. This indicates that the incorporation of the frame dropout compensation mechanism can mitigate the discontinuity in tracking. Furthermore, as indicated by the bold font in Table I, the proposed method in this paper, which integrates both **WHFE** and **FDC**, achieved significant improvements in tracking performance for both catheter and guidewire across HOTA, MOTA, and IDF1 metrics. This demonstrates that the proposed method can achieve highly accurate and continuous tracking of guidewire and catheter.

Figure 3 illustrates the tracking results of selected frames from a continuous DSA sequence. As shown in Figure 3(b), the baseline method only successfully

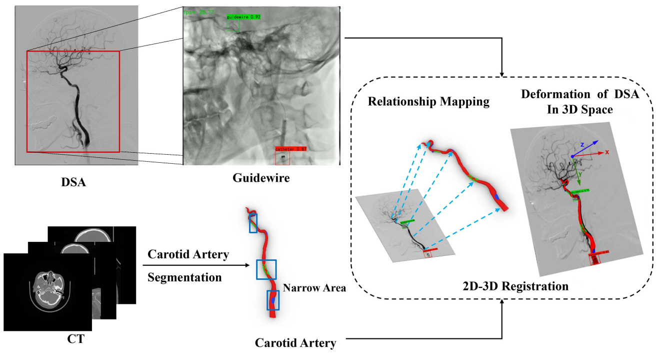

Fig. 4. Intraoperative DSA registration with preoperative 3D carotid artery reconstruction model provides intraoperative visual 3D guidance.

tracked the catheter in the initial frame (indicated by the green rectangular box), but failed to track it in subsequent frames. Due to the loss of catheter information in the preceding frames, the introduction of **FDC** (Frame Dropout Compensation) alone was insufficient to address this issue, resulting in the continued loss of catheter tracking in the subsequent frames, as depicted in Figure 3(c). Figure 3(d) demonstrates that the introduction of **WHFE** (Wavelet High-Frequency Enhancement) alone significantly improved tracking performance in some frames, although tracking failures still occurred, as shown in the T1 frame in Figure 3(d). However, as illustrated in Figure 3(e), the simultaneous incorporation of **WHFE** and **FDC** enabled precise, continuous, and stable tracking of both the guidewire and catheter.

IV. CONCLUSIONS

This paper proposes a high-precision and stable tracking method for guidewire/catheter in DSA based on wavelet high-frequency enhancement and frame dropout compensation. The proposed method improved the tracking performance of the two superior baseline models by 5.64% and 8.14%, respectively. Therefore the proposed method can support real-time visualization guidance for CAS surgery based on 2D-3D registration.

REFERENCES

[1] R. Arasu, A. Arasu, and J. Muller, "Carotid artery stenosis: An approach to its diagnosis and management," *Australian Journal of General Practice*, vol. 50, no. 11, pp. 821–825, 2021.

[2] I. Ullah, P. Chikontwe, and S. H. Park, "Real-time tracking of guidewire robot tips using deep convolutional neural networks on successive localized frames," *IEEE Access*, vol. 7, pp. 159 743–159 753, 2019.

[3] K. Chen, W. Qin, Y. Xie, and S. Zhou, "Towards real time guide wire shape extraction in fluoroscopic sequences: a two phase deep learning scheme to extract sparse curvilinear structures," *Computerized Medical Imaging and Graphics*, vol. 94, p. 101989, 2021.

[4] Z. Mei, H. Wang, S. Pan, H. Chen, J. Wei, Q. Zhang, J. Mao, G. Liu, and Y. Zhao, "Real time detection and tracking of guide wire/catheter for interventional embolization robot based on deep learning," in *2023 IEEE International Conference on Mechatronics and Automation (ICMA)*. IEEE, 2023, pp. 778–783.

[5] V. Subramanian, H. Wang, J. T. Wu, K. C. Wong, A. Sharma, and T. Syeda-Mahmood, "Automated detection and type classification of central venous catheters in chest x-rays," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part VI 22*. Springer, 2019, pp. 522–530.

[6] R.-Q. Li, G. Bian, X. Zhou, X. Xie, Z. Ni, and Z. Hou, "A two-stage framework for real-time guidewire endpoint localization," in *Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, October 13–17, 2019, Proceedings, Part V 22*. Springer, 2019, pp. 357–365.

[7] M. Aghasizade, A. Kiyoumarsioskouei, S. Hashemi, M. Torabinia, A. Caprio, M. Rashid, Y. Xiang, H. Rangwala, T. Ma, B. Lee *et al.*, "A coordinate-regression-based deep learning model for catheter detection during structural heart interventions," *Applied Sciences*, vol. 13, no. 13, p. 7778, 2023.

[8] S. E. Finder, R. Amoyal, E. Treister, and O. Freifeld, "Wavelet convolutions for large receptive fields," in *European Conference on Computer Vision*, 2024.

[9] S. E. Finder, Y. Zohav, M. Ashkenazi, and E. Treister, "Wavelet feature maps compression for image-to-image cnns," *Advances in Neural Information Processing Systems*, vol. 35, pp. 20 592–20 606, 2022.

[10] N. Carion, F. Massa, G. Synnaeve, N. Usunier, A. Kirillov, and S. Zagoruyko, "End-to-end object detection with transformers," in *European conference on computer vision*. Springer, 2020, pp. 213–229.

[11] Y. Zhao, W. Lv, S. Xu, J. Wei, G. Wang, Q. Dang, Y. Liu, and J. Chen, "Detrs beat yolos on real-time object detection," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 16 965–16 974.

[12] M. Yang, G. Han, B. Yan, W. Zhang, J. Qi, H. Lu, and D. Wang, "Hybrid-sort: Weak cues matter for online multi-object tracking," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 38, no. 7, 2024, pp. 6504–6512.

[13] X. Hou, M. Liu, S. Zhang, P. Wei, and B. Chen, "Salience detr: Enhancing detection transformer with hierarchical salience filtering refinement," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2024, pp. 17 574–17 583.