

Real-Time 3D Instrument Tip Tracking Using 2D X-Ray Fluoroscopy With Vessel Deformation Correction Under Free Breathing

Shuo Yang , Deqiang Xiao , Haixiao Geng , Danni Ai , Jingfan Fan , Tianyu Fu , Hong Song , Feng Duan , and Jian Yang 

Abstract— **Objective:** Accurate localization of the instrument tip within the hepatic vein is crucial for the success of transjugular intrahepatic portosystemic shunt (TIPS) procedures. Real-time tracking of the instrument tip in X-ray images is greatly influenced by vessel deformation due to patient's pose variation, respiratory motion, and puncture manipulation, frequently resulting in failed punctures. **Method:** We propose a novel framework called deformable instrument tip tracking (DITT) to obtain the real-time tip positioning within the 3D deformable vasculature. First, we introduce a pose alignment module to improve the rigid matching between the preoperative vessel centerline and the intraoperative instrument centerline, in which the accurate matching of 3D/2D centerline features is implemented with an adaptive point sampling strategy. Second, a respiration compensation module using monoplane X-ray image sequences is constructed and provides the motion prior to predict intraoperative liver movement. Third, a deformation correction module is proposed to rectify the vessel deformation during procedures, in which a manifold regularization and the maximum likelihood-based acceleration are introduced to obtain the accurate and fast deformation learning. **Results:** Experimental results on simulated and clinical datasets show an average tracking error of 1.59 ± 0.57 mm and 1.67 ± 0.54 mm, respectively. **Conclusion:**

Our framework can track the tip in 3D vessel and dynamically overlap the branch roadmapping onto X-ray images to provide real-time guidance. **Significance:** Accurate and fast (43ms per frame) tip tracking with the proposed framework possesses a good potential for improving the outcomes of TIPS treatment and minimizes the usage of contrast agent.

Index Terms—X-ray image-guided intervention, adaptive sampling, tip tracking, respiration compensation, deformation correction.

I. INTRODUCTION

IN TRANSJUGULAR intrahepatic portosystemic shunt (TIPS) procedures, accurate positioning of the instrument (such as catheter, guidewire, and needle) tip significantly enhances puncture accuracy and efficiency. TIPS represents a comprehensive interventional therapy for addressing portal hypertension and gastric variceal bleeding. This intervention entails puncturing from the hepatic vein (HV) to the portal vein (PV) and establishing an artificial connection channel between them. Based on the segmentation results of HV and PV from preoperative computed tomography angiography (CTA) volume (Fig. 1), the preoperative planning is conducted to determine the optimal HV branch for needle entry and the puncture point. During the puncture stage, the contrast agents are used to enhance the current HV branch around the instrument's location. However, the rapid loss of the contrast agent leads to the branch becoming invisible. Radiologists must rely on their imagination associated with preoperative planning to track the tip and advance it to the puncture point, ensuring the safe connection between HV and PV is established. In this phase, reckless operations can lead to vascular rupture and liver tissue injury, thereby jeopardizing patient well-being. The success of the puncture procedure heavily depends on accurate tip tracking, which significantly enhances the clinical treatment outcome.

X-ray fluoroscopy serves as a commonly used guidance modality in TIPS procedures, offering valuable imaging of the needle tip's advancement towards the target point. Traditionally, 2D X-ray images are enhanced with contrast agents to visualize the vasculature. However, angiography and puncture cannot be performed simultaneously, and the contrast agents rapidly flow into the superior vena cava (SVC) via the HV. The puncture manipulation is performed without visible vasculature, presenting a challenge to the radiologist. Furthermore, the patient's

Received 16 May 2024; revised 7 October 2024 and 30 October 2024; accepted 27 November 2024. Date of publication 28 November 2024; date of current version 21 March 2025. This work was supported in part by the National Science Foundation Program of China under Grant 62025104, Grant 62071048, Grant 62202045, Grant 82330061, and Grant 62271057, in part by the Beijing Natural Science Foundation under Grant JQ22030, and in part by the Beijing Municipal Science and Technology Project under Grant Z231100004823009. (Corresponding authors: Deqiang Xiao; Feng Duan; Jian Yang.)

Shuo Yang, Haixiao Geng, Danni Ai, and Jingfan Fan are with the School of Optics and Photonics, Beijing Institute of Technology, China.

Deqiang Xiao is with the School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China, and also with the Zhengzhou Research Institute, Beijing Institute of Technology, Zhengzhou 450003, China (e-mail: deqiang_xiao@bit.edu.cn).

Tianyu Fu is with the School of Medical Technology, Beijing Institute of Technology, China.

Hong Song is with the School of Computer Science and Technology, Beijing Institute of Technology, China.

Feng Duan is with the Department of Interventional Radiology, Chinese PLA General Hospital First Medical Center, Beijing 100853, China (e-mail: duanfeng@vip.sina.com).

Jian Yang is with the School of Optics and Photonics, Beijing Institute of Technology, Beijing 100081, China (e-mail: jyang@bit.edu.cn).

Digital Object Identifier 10.1109/TBME.2024.3508840

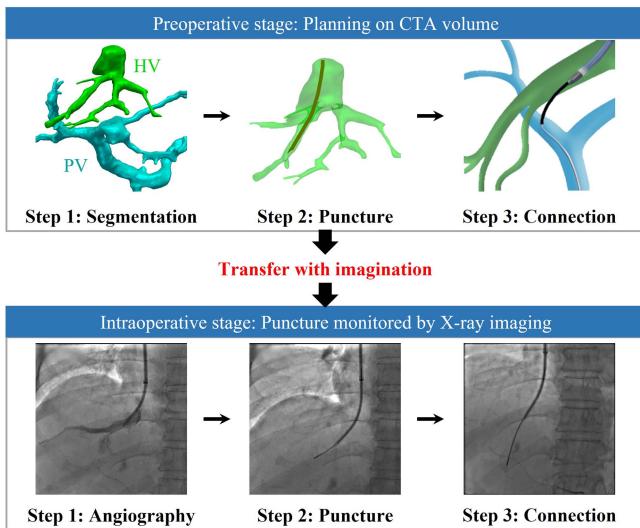


Fig. 1. The TIPS procedure involves two stages. In the preoperative stage, planning is conducted to determine the optimal puncture path and establish a connection based on vessel segmentation models extracted from the CTA volume. In the intraoperative stage, once the HV branch is visualized via angiography, the planning is transferred to practical punctures based on the operator's imagination, ensuring a safe puncture between HV and PV without harming adjacent vessels and tissues.

vasculature undergoes unpredictable deformation during procedures, influencing the intraoperative judgement based on preoperative planning and interfering the tracking of the instrument tip during intervention.

Several methods have been proposed to track the instrument tip in 2D X-ray images using vessel roadmapping. Vessel roadmapping is typically achieved by aligning 3D vasculature with 2D X-ray images through CTA/X-ray registration [1]. However, the accuracy of roadmapping is seriously impacted by disparities between preoperative and intraoperative poses [2]. Baert et al. [3] address this challenge by aligning the intraoperative pose reconstructed via biplane X-ray imaging with the preoperative pose. Similarly, Wagner et al. [4] standardize patient poses and project the 3D vasculature onto the biplane X-ray imaging plane, providing various views for guiding the puncture. Researchers in [5] increase accuracy by employing the self-calibration of C-arm imaging system. Nonetheless, this imaging technique entails a doubling of the X-ray radiation dose, and its high equipment cost remains a challenge for clinical implementation [6], [7], [8]. Consequently, monoplane X-ray image guidance remains the predominant approach in vessel interventional procedures. Nevertheless, accurate tip tracking remains a primary challenge caused by pose variation, respiratory motion, and puncture manipulation.

The purpose of this work is to accurately and efficiently track the 3D instrument tip from the monoplane X-ray imaging for TIPS procedure guidance. Specifically, we propose a deformable instrument tip tracking framework, namely DITT, to map the 2D instrument tip position extracted from the X-ray image to the corresponding 3D point of the 3D vessel centerline extracted from the CTA volume in real time. First, for the problem of pose variation, we introduce a 3D/2D pointset registration module

based on the hidden Markov model (HMM) [9] to rigidly align the preoperative 3D vessel centerline with the intraoperative 2D instrument centerline. As the pointset registration is prone to error on the centerline segment with high curvatures, a curvature-based adaptive point sampling strategy is proposed to increase the number of semantic 3D/2D features to alleviate the registration difficulties. Second, to compensate for respiratory motion, a patient-specific respiration compensation module using the liver masks extracted from X-ray image sequences is designed. Considering that the strategies of respiration compensation in related methods rely on the nearest previous X-ray frame only, which often leads to unstable respiration prediction and failed tip tracking, our strategy is to learn a respiratory motion model from pre-obtained 2D + t X-ray images and take the model as a motion prior to predict the liver movement intraoperatively. Third, to correct the vessel deformation caused by the puncture operations, we present a non-rigid pointset registration module, which calculates a deformation from the projected preoperative vessel to the intraoperative instrument after pose alignment and respiration compensation. In addition, a manifold regularization and an acceleration strategy are introduced to obtain a reliable and efficient deformation learning. Our main contributions are summarized as follows:

- We propose a novel multi-module framework for real-time deformable instrument tip tracking to guide TIPS procedures. This framework effectively addresses the challenges of tracking posed by variations in instrument position, respiratory motion, and puncture manipulation. Our framework provides both the 2D position of the instrument tip within the deformed vessel branch and the 3D lumen structure surrounding the tracking tip.
- We propose a pose alignment (PA) module to address the problem of low registration accuracy caused by the limited instrument centerline features. The PA module improves the rigid alignment of vessel/instrument features through a curvature-aware adaptive sampling strategy.
- We propose a respiration compensation (RC) module to estimate the motion error curve from the X-ray sequence to predict the liver movement in dynamic roadmapping. The RC module addresses the challenge of tracking dependence on adjacent X-ray frames through learning a prior respiratory motion model from 2D + t datasets, which improves the stability of vessel/instrument alignment.
- We introduce a deformation correction (DC) module to rectify vessel deformation during the puncture process. This module contains a non-rigid pointset registration algorithm, in which the manifold regularization forcing the optimal smooth and the geometry preservation is incorporated to ensure high registration accuracy, the maximum likelihood estimation is employed to optimize correspondence matching, enabling rapid correction.

II. RELATED WORKS

In order to overlay vessel roadmapping onto the X-ray image sequence, it is necessary to compute a transformation matrix via

3D/2D registration, and project the segmented 3D vessel branch onto the corresponding 2D position. From the review [10] in 3D/2D vessel registration, various techniques have been proposed in X-ray image-guided interventional procedures among different regions, including the cranial [11], [12], [13], [14], cardiac [15], [16], [17], [18], and abdomen areas [19], [20], [21], [22], [23]. These methods utilize different anatomical features (such as bone, diaphragm, and vessel), employing rigid 3D/2D registration to form 2D vessel roadmapping. Several researches [24], [25], [26], [27] improve registration accuracy by using 3D/2D non-rigid pointset registration between 3D and 2D features from each X-ray frame, but it results in a heavy time consumption and cannot provide real-time guidance in intervention. To address this issue, King et al. [28] consider the diaphragm motion correlation between frames in 3D/2D registration, aiming to reduce pose differences before registration and facilitate the rapid update of registration transformation. Metz et al. [29] simultaneously register 3D + t vessels to multiple frames of 2D + t X-ray sequence, thus obtaining the sequential 2D + t roadmapping on X-ray images. While these enhance the robustness of 3D/2D registration in interventional navigation, they also increase the burden of image acquisition and preprocessing, including diaphragm segmentation and 3D + t scanning. Ambrosini et al. [30] propose a catheter tip tracking (CTT) method that focuses on registration between a segmented 3D vessel centerlines at single time-phase and 2D instrument centerlines from sequential frames, it provides intuitive and continuous roadmapping for interventional procedure. However, this method will fail when the length of instrument within the branch is short (less than 80 mm), and the lack of respiration compensation in the registration transformation can affects the overlay results of roadmapping.

To ensure accurate respiration compensation during interventions, Ma et al. [31] propose a novel dynamic coronary roadmapping overlay method based on Bayesian filtering. Specifically, they utilize electrocardiogram (ECG) data acquired from percutaneous coronary intervention (PCI) procedures to establish pixel correspondences between frames at different motion states and dynamically predict coronary roadmapping in subsequent frames. Vernikouskaya et al. [32] correlate the ECG signal with the catheter's motion using a convolutional neural network (CNN). Compared with the ECG-based methods, image-based method is more direct and convenient to estimate respiration from image sequence. Wagner et al. [33] propose a respiratory estimation model based on the centroid of the image intensity distribution to estimate the breathing signal in liver interventions, achieving dynamic roadmapping of liver vessel. However, this method will fail when the diaphragm is not visible or incomplete in the X-ray image. None of the aforementioned methods account for vessel deformation during the overlay process, particularly the changes in branch pose caused by puncture manipulation. This presents a severe challenge for radiologists, who must locate the instrument tip within the deformed vessel branch and complete the puncture while the patient breathes freely. With the proposed DITT, we provide the true pose of the branch during the intervention and accurately locate the instrument tip within the lumen.

III. METHOD

The architecture of DITT, depicted in Fig. 2, is designed to track the instrument tip using both the CTA volume and the X-ray fluoroscopic sequence. It consists of three modules, i.e., PA, RC, and DC modules. In the PA module, feature points are extracted from the segmented 3D HV and 2D instrument centerlines using an adaptive sampling strategy, facilitating the pose alignment through centerline pointset matching. This module aligns the 3D branch with the 2D instrument to initially overlay branch roadmapping. Subsequently, the RC module extracts the respiratory motion prior from the 2D X-ray sequence to synchronize the roadmapping with breathing. In the DC module, the sampled 3D/2D feature points act as geometric structures for non-rigid pointset registration to correct the vessel deformation. Finally, the 2D deformed branch roadmapping and 3D tip position are obtained.

A. Pose Alignment With Adaptive Sampling

To ensure accurate pose alignment, we propose an adaptive point sampling strategy in the PA module to improve 3D/2D centerline feature matching (Fig. 3). This strategy calculates the curvature of each point along the input 3D/2D centerlines, with more features being sampled in high-curvature regions and fewer in smoother areas. Subsequently, the sampled 3D/2D feature points are registered using HMM, and the Viterbi algorithm [9] is employed to optimize 3D/2D transformation.

1) Curvature-Aware Adaptive Sampling: The aim of pre-operative and intraoperative pose alignment is to establish correct correspondences between the 3D vessel centerline pointset $V_{3D} = \{p_1, p_2, \dots, p_N\}$ and the 2D instrument centerline pointset $C_{2D} = \{c_1, c_2, \dots, c_M\}$ in each frame of the X-ray image sequence. Here, N and M represent the number of points in the two sets, respectively.

To register C_{2D} and V_{3D} , we propose an adaptive sampling algorithm for the vessel/instrument centerline to extract more features around high-curvature regions, guiding the optimizer to achieve better alignment of the 3D/2D pointsets. The high-curvature regions encompass end-branch points, bifurcation points and yield points. When extracting 2D instrument features in X-ray images, we employ Bezier curve estimation [34] to calculate the curvature of each point from the 2D centerline, as depicted in Fig. 3, we select P_0 and P_2 as the initial points and extend them along the instrument to obtain the intersection point P_1 . As point Q_0 moves along the line segment P_0P_1 , the segment is divided into two parts, i.e., l and $(1 - l)$. Upon connecting Q_0 and Q_1 , the trajectory of point B is expressed by the Bezier curve equation:

$$B(l) = (1 - l)^2 p_0 + 2l(1 - l)p_1 + l^2 p_2, \quad (1)$$

where the $p_0(x_0, y_0)$, $p_1(x_1, y_1)$, and $p_2(x_2, y_2)$ are the position of P_0 , P_1 , and P_2 , respectively. On this basis, the curvature of any point on the Bezier curve is expressed as:

$$k_{2D} = \frac{dy/dl}{dx/dl} = \frac{-(1 - l)y_0 + (1 - 2l)y_1 + ly_2}{-(1 - l)x_0 + (1 - 2l)x_1 + lx_2}. \quad (2)$$

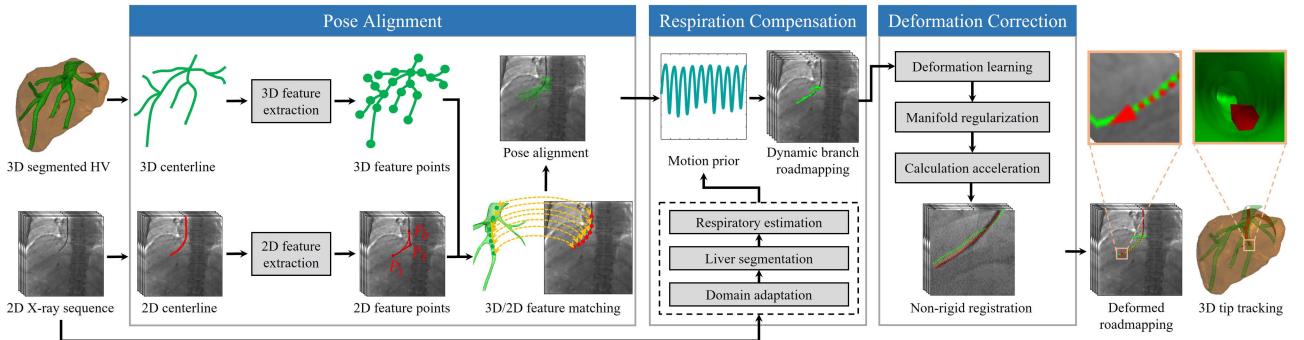


Fig. 2. Three modules and the inference process of DITT. In the PA module, 3D and 2D centerline feature points are extracted via point adaptive sampling, these points engage in 3D/2D rigid registration to facilitate preoperative and intraoperative pose alignment. The RC module estimates the breathing curve from the 2D X-ray image sequence and compensates for respiratory motion during dynamic roadmapping. Subsequently, the DC module utilizes deformation learning and acceleration algorithm to rectify the deformation of HV branch. During the inference phase, real-time 2D deformed roadmapping is produced, and the 3D instrument tracking is achieved through rendering.

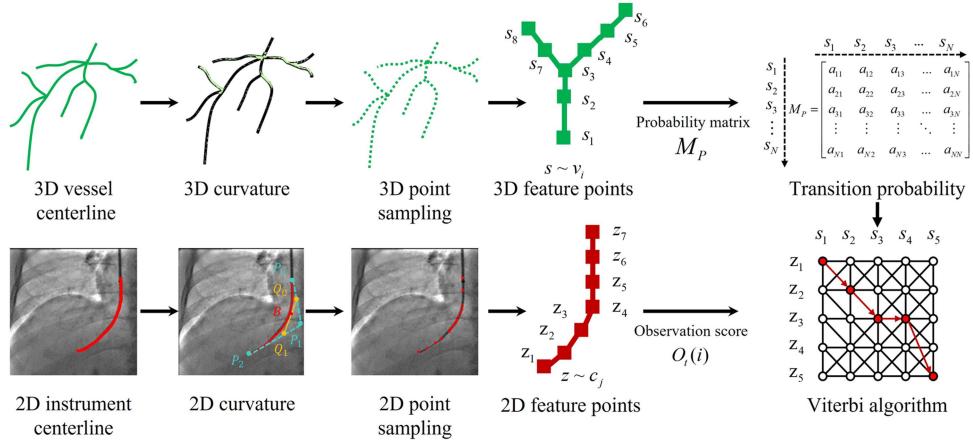


Fig. 3. In the PA module, adaptive point sampling is conducted separately for 3D and 2D centerlines based on 3D/2D curvature calculation. These sampled feature points aid in HMM-based pose alignment. The Viterbi algorithm is employed to determine the optimal transformation using a transition probability matrix and observation scores.

Next, we apply a threshold to filter out high-curvature points. Unlike the 2D instrument, the 3D vessel tree exhibits more complex branching and curvature variations. Therefore, the extraction of 3D vessel centerlines and calculation of curvature k_{3D} at each 3D centerline point are conducted during preprocessing. The processes for 3D/2D sampling are expressed as follows:

$$\begin{aligned} V^{AS} &= \{V_{3D} \mid k_{3D}(v_i) \mid_{i=1}^n > \mu_1\}, \\ C^{AS} &= \{C_{2D} \mid k_{2D}(c_j) \mid_{j=1}^m > \mu_2\}, \end{aligned} \quad (3)$$

where V^{AS} and C^{AS} represent the 3D and 2D feature pointsets after adaptive sampling. v_i and c_j represent each 3D and 2D point, respectively, while μ_1 and μ_2 denote the curvature thresholds used to constrain feature point sampling. Subsequently, V^{AS} and C^{AS} serve as the 3D and 2D centerline pointsets, facilitating pose alignment through HMM-based registration.

2) HMM-Based Registration: For the HMM-based vessel/instrument centerline registration, a probability distribution of the 3D tip position is computed using an HMM and a 3D/2D registration is performed for the most probable tip locations.

A hidden Markov model is a system with a set of N states $\{s_1, \dots, s_N\}$ (Fig. 3), the HMM state changes at each time point t according to the probabilities associated with state transitions and the current observation. The transition probability between states are defined in a matrix M_P ($N \times N$), reflecting the likelihood of the tip transition between 3D centerline points. An observation likelihood (observation score) $O_t(i) = P(Z_t | s_i)$ is defined, representing the likelihood that the state s_i at time point t produced the observation Z_t . In HMM-based registration, the observation of HMM is the 2D X-ray image, and the observation score is evaluated via 3D/2D registration. The Viterbi algorithm is employed to determine the highest score from the initial state to the end state. The score is computed from the registration between the 2D instrument centerline pointset C_{2D} in each frame and the 3D vessel centerline pointset V_{3D} . The observation score $O_t(i)$ is a likelihood between 0 and 1 that indicates how well the 3D/2D registration performs at time t given that the tip is at the position p_i in the 3D vessel centerline. This relationship is defined as a Gaussian function $O_t(i) = e^{-(M^2/2\sigma_o^2)}$, where M is the metric of 3D/2D registration and σ_o controls the scaling.

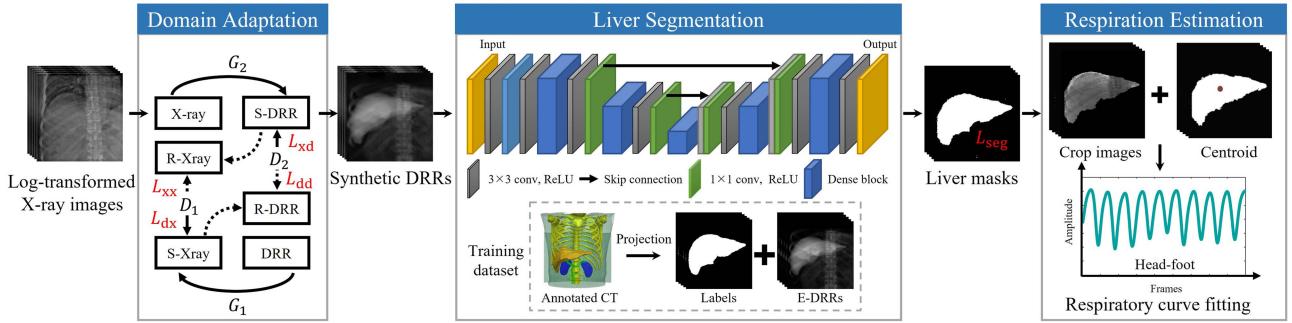


Fig. 4. The RC module involves translating X-ray images into synthetic DRRs, which serve as the input fed into the segmentation network, generating liver masks. Utilizing the respiratory motion model, we estimate the continuous centroid positions of these liver masks to derive the breathing curve.

of the metric. Thus, the 3D/2D registration between the V_{3D} and C_{2D} yields the registration transformation matrix τ that minimizes the metric, and the metric M quantifies the alignment of the 3D vessel with each point of the 2D instrument:

$$M(V_{3D}, C_{2D}) = \min_{\tau} \left(\sum F_{\text{cost}}(v_i, c_j, \tau) \right), \quad (4)$$

where $v_i \in V_{3D}$, $c_j \in C_{2D}$, and $\tau = T_{\text{rot}_z} T_{\text{rot}_y} T_{\text{rot}_x} T_{\text{trans}_{xyz}}$. We use the cost function F_{cost} proposed in [30] to evaluate the similarity between 2D instrument features and 3D vessel features after transformation τ .

In the process of computing metric M in (4), we utilize the cost function F_{cost} that takes into account the vessel diameter at vessel V_{3D} . Using the lumen radius at each 3D point, associated with point position v_i . The cost function based on Sigmoid function is:

$$F_{\text{cost}}(v_i, c_j, \tau) = \frac{F_{\text{min}}}{1 + e^{\gamma_1(-F_{\text{min}} + \gamma_2 \cdot r_i)}}, \quad (5)$$

where $F_{\text{min}} = \min_{v_i \in V_{3D}} \|c_j - F_{\text{proj}}(v_i, \tau)\|$, F_{proj} denotes the projection between a 3D point v_i and the corresponding 2D point c_j . γ_1 and γ_2 control the cost inside the vessel, it means that $F_{\text{cost}}(\cdot)$ start to penalize when the minimum distance reaches the radius r_i , where r_i is the radius at the point v_i .

B. Respiration Compensation

In order to compensate for respiratory motion and dynamically overlap rodmmaps on X-ray sequences, we design the RC module (Fig. 4). The input X-ray images are translated to synthetic digitally reconstructed radiographs (DRRs) through domain adaptation [35]. These synthetic DRRs are then fed into the segmentation network pretrained using pairwise labeled DRR datasets to obtain corresponding liver masks [36]. Subsequently, we derive the periodic respiration prior by computing the relative motion across multiple frames and fitting a respiratory curve.

1) Fluoroscopy to DRR Domain Adaptation: To mitigate the domain gap between X-ray image $x \sim p(x)$ and DRR $d \sim p(d)$, a generative synthesis network based on Cycle-GAN [37] is trained. A dataset comprising 2600 real X-ray images and 2600 DRRs from 26 patients is utilized. Each X-ray image undergoes a log-transform to ensure the same grayscale distribution as

the DRR. The DRR is manually aligned with the corresponding patient's X-ray image according to the diaphragm to ensure the stability of domain adaptation. The network comprises various paths for translating and reconstructing images between the two protocols, with different losses in each path. For a given real DRR, the generator G_1 generates the synthetic X-ray (S-Xray) image, while the discriminator D_1 distinguishes whether the generated X-ray image is real. The loss function is expressed as follows:

$$\mathcal{L}_{\text{dx}} = \mathbb{E}_{x \sim p(x)} [\log D_1(x)] + \mathbb{E}_{d \sim p(d)} [1 - \log D_1(G_1(d))]. \quad (6)$$

Generator G_2 tries to generate synthetic DRR (S-DRR), and discriminator D_2 help G_2 to finish the domain adaptation from X-ray image to DRR. The S-DRR is then fed into the segmentation network described in the following section to obtain the liver mask. The loss function for this path is expressed as follows:

$$\mathcal{L}_{\text{xd}} = \mathbb{E}_{d \sim p(d)} [\log D_2(d)] + \mathbb{E}_{x \sim p(x)} [1 - \log D_2(G_2(x))]. \quad (7)$$

For the cyclic-consistency of domain adaptation, the S-DRR generated by G_2 is translated back through G_1 to obtain reconstructed X-ray (R-Xray), the corresponding loss function is expressed as follows:

$$\mathcal{L}_{\text{xx}} = \mathbb{E}_{x \sim p(x)} [\|G_1(G_2(x)) - x\|_1]. \quad (8)$$

Similarly, the S-Xray is translated back to the reconstructed DRR (R-DRR), with corresponding backward cycle-consistency loss

$$\mathcal{L}_{\text{dd}} = \mathbb{E}_{d \sim p(d)} [\|G_2(G_1(d)) - d\|_1]. \quad (9)$$

The total loss function is a summation of each weighted loss

$$\mathcal{L}_{\text{all}} = \alpha_1 \mathcal{L}_{\text{dx}} + \alpha_2 \mathcal{L}_{\text{xd}} + \alpha_3 \mathcal{L}_{\text{xx}} + \alpha_4 \mathcal{L}_{\text{dd}}, \quad (10)$$

where $\{\alpha_1, \alpha_2, \alpha_3, \alpha_4\}$ are constant weights to control the balance among the losses. In the test, the predicted liver masks are obtained from X-ray images directly. The same parameter settings from [38] are employed in the implementation, and the respiratory motion is calculated based on the segmented liver masks.

2) DRR Segmentation Network: To achieve accurate liver segmentation results to estimate the breathing curve, we utilize a deep image-to-image network based on U-Net [39]. This network is trained on a dataset comprising 1967 cases of labeled DRR data generated from annotated 3D CT scans. Specifically, we employ the Dense-UNet [40] with a reduced dense block [41]. During training, we adjust the Hounsfield units (HU) to enhance the DRR, thereby improving the contrast between the liver region and the background. Our loss function is defined based on the Dice metric [42]:

$$\mathcal{L}_{\text{seg}} = 1 - \frac{2 * (B^P \cap B^{\text{GT}})}{B^P \cup B^{\text{GT}}}, \quad (11)$$

where B^{GT} and B^P represent the ground truth (GT) and the predicted liver label, respectively. They have the same size as the input image.

3) Respiratory Motion Calculation: After obtaining the sequential liver masks $I_s(x, y)$, where s denotes the respiratory state in a breathing cycle. Each mask is represented as a binary image defined on a regular 2D grid with coordinates x and y . The respiratory motion r_s of each frame is estimated based on the position of the liver centroid, and the trajectory of the centroid position is utilized to estimate the respiratory signal curve.

$$\begin{aligned} r_{s_i}^p &= \left[\sum_{\forall x, y} y \cdot I_{s_i}(x, y) \right] / \left[\sum_{\forall x, y} I_{s_i}(x, y) \right], \\ r_{s_i}^v &= \left[\sum_{\forall x, y} x \cdot I_{s_i}(x, y) \right] / \left[\sum_{\forall x, y} I_{s_i}(x, y) \right], \end{aligned} \quad (12)$$

where $r_{s_i}^p$ and $r_{s_i}^v$ represent the respiration along head-foot and left-right directions during abdominal intervention [43]. Aims to apply respiration compensation in dynamic overlapping, it's necessary to calculate the relative displacement between multiple frames within a breathing cycle:

$$D_{s_{i+1}}^p = r_{s_{i+1}}^p - r_{s_i}^p, \quad D_{s_{i+1}}^v = r_{s_{i+1}}^v - r_{s_i}^v, \quad (13)$$

where $r_{s_{i+1}}$ and r_{s_i} represent the respiration at two adjacent breathing states. During the registration, the respiration compensation is equivalent to an affine transform to describe the respiratory motion of the HV. Each element of registration transformation matrix T_r can be described by $D_{s_i}^p$ and $D_{s_i}^v$ in different respiratory state s_i . Therefore, the registration transformation matrix τ mentioned above is updated as $\tau^* = T_{\text{rot}_z} T_{\text{rot}_y} T_{\text{rot}_x} T_{\text{trans}_{xyz}} T_r$.

C. Vessel Deformation Correction

During the puncture process, the HV branch dynamically deforms due to the instrument shape change when advancing the instrument in the vessel with breathing. The deformation learning from the HV branch to the instrument can improve the accuracy of tip tracking within the branch lumen. The geometric points sampled from the PA module are used in this procedure, and the motion prior from the RC module dynamically aligns the branch with the instrument, thus ensuring the stable non-rigid registration during deformation correction.

1) Deformation Learning: In this section, we establish two pointsets $\{v_i\}_{i=1}^n$ and $\{c_j\}_{j=1}^m$ at time t , which represent the vessel/instrument centerline feature points after adaptive point sampling. We employ corresponding shape contexts (SCs) [44] and the Hungarian method [45] to construct feature descriptors and determine correspondence between $\{v_i\}_{i=1}^n$ and $\{c_j\}_{j=1}^m$. Subsequently, a putative correspondence set $S = \{(v_i, c_i)\}_{i=1}^L$ is obtained, where $L \leq \min\{m, n\}$ represents the number of correspondences. For learning the non-rigid transformation \mathcal{T} :

$$\mathcal{T}^* = \arg \min_{\mathcal{T}} \sum_{i=1}^L \|c_i - \mathcal{T}(v_i)\|^2, \quad c_i, v_i \in S, \quad (14)$$

its solution is not unique without an additional constraint. The Tikhonov regularization [46] defined in a reproducing kernel Hilbert space (RKHS) \mathcal{H} serves as a constraint, which is used to minimize the following cost function to obtain the transformation \mathcal{T} defined in \mathcal{H} .

$$\mathcal{T}^* = \arg \min_{\mathcal{T} \in \mathcal{H}} \sum_{i=1}^L \|c_i - \mathcal{T}(v_i)\|^2 + \lambda_1 \|\mathcal{T}\|_{\mathcal{H}}^2, \quad (15)$$

where $\|\cdot\|_{\mathcal{H}}$ represents the RKHS norm, and λ_1 balances the two terms. The transformation \mathcal{T} defined in RKHS can be expressed as a reproducing kernel function as follows.

$$\mathcal{H} = \left\{ \mathcal{T} | \mathcal{T} = \sum_{i=1}^L \mathbf{m}_i \mathbf{k}(c_i, \cdot), \mathbf{m}_i \in \mathbb{R}^{3 \times 1} \right\}, \quad (16)$$

where $\mathbf{k}(\cdot)$ is a positive definite kernel function, and the RKHS norm is calculated with an inner product $\|\cdot\|_{\mathcal{H}} = \sqrt{\langle \cdot, \cdot \rangle_{\mathbf{k}}}$. In this work, we take the Gaussian function ϕ as $\mathbf{k}(\cdot)$, which is formulated as $\phi(c_i, c_j) = \exp(-\|(c_i - c_j)/\delta\|^2/2)$, δ represents the kernel size, $\mathbf{m}_i \in \mathbb{R}^{3 \times 1}$ is the coefficient of transformation \mathcal{T} . According to the RKHS definition, the transformation norm $\|\mathcal{T}\|_{\mathcal{H}}^2$ is calculated as follows.

$$\begin{aligned} \|\mathcal{T}\|_{\mathcal{H}}^2 &= \langle \mathcal{T}(\cdot), \mathcal{T}(\cdot) \rangle_{\mathbf{k}} \\ &= \left\langle \sum_{i=1}^L \mathbf{m}_i \phi(c_i, \cdot), \sum_{j=1}^L \mathbf{m}_j \phi(c_j, \cdot) \right\rangle \\ &= \sum_{i=1}^L \sum_{j=1}^L \mathbf{m}_i^T \mathbf{m}_j \langle \phi(c_i, \cdot), \phi(c_j, \cdot) \rangle \\ &= \sum_{i=1}^L \sum_{j=1}^L \mathbf{m}_i^T \mathbf{m}_j \phi(c_i, c_j) \\ &= \text{tr}(\mathbf{M}^T \mathbf{U} \mathbf{M}), \end{aligned} \quad (17)$$

where the matrix $\mathbf{U} \in \mathbb{R}^{L \times L}$, each element $U_{i,j} = \phi(c_i, c_j)$, the coefficient matrix $\mathbf{M} = [\mathbf{m}_1 \mathbf{m}_2 \dots \mathbf{m}_L]^T \in \mathbb{R}^{L \times 3}$, and $\text{tr}(\cdot)$ is the matrix trace. $\|\mathcal{T}\|_{\mathcal{H}}^2$ restrains the smoothness of the deformation field described by the control points from the 2D instrument and the coefficient matrix.

2) Manifold Regularization: Manifold regularization has been used in 3D/2D vessel non-rigid pointset registration [10], [47]. In the matching problem, the number of points of source-set and target-set is often different. After constructing

correspondence between $\{v_i\}_{i=1}^n$ and $\{c_j\}_{j=1}^m$, only L pairs $\{c_1, c_2, \dots, c_L\}$ and $\{v_1, v_2, \dots, v_L\}$ participate in the estimation of transformation \mathcal{T} . It means that only L input points make contribution to geometry registration, and the remaining points without matching may contain extra geometric structure information about the input shape data. Using this information is the advantage of manifold regularization, it defines an extra regularization term $\|\mathcal{T}\|_{\mathcal{I}}^2$ on the input pointset to constrain \mathcal{T} and the predicted transformation \mathcal{T}^* is constrained by cost function:

$$\mathcal{T}^* = \arg \min_{\mathcal{T} \in \mathcal{H}} \sum_{i=1}^L \|c_i - \mathcal{T}(v_i)\|^2 + \lambda_1 \|\mathcal{T}\|_{\mathcal{H}}^2 + \lambda_2 \|\mathcal{T}\|_{\mathcal{I}}^2, \quad (18)$$

where the extra regularization term maintains the geometry completeness of input data. It's defined based on graph Laplacian, which is a discrete analog of the manifold Laplacian [48]. The graph Laplacian models the manifold according to a weighted neighborhood graph G , which is obtained by constructing the vertex set $\text{Ver} = \{g_1, g_2, \dots, g_L\}$ with edges (g_i, g_j) . The following weight W_{ij} is assigned to edge (g_i, g_j) :

$$W_{ij} = e^{-\frac{1}{\epsilon} \|g_i - g_j\|^2}. \quad (19)$$

According to W , a matrix A is constructed with each element expressed as follows:

$$A_{ij} = D_{ij} - W_{ij}, \quad (20)$$

where $D = \text{diag}(\sum_{j=1}^L W_{ij})_{i=1}^L$ (i.e., the diagonal matrix whose i -th entry is the sum of the weights of edges leaving g_i). Denote $t = (\mathcal{T}(g_1), \dots, \mathcal{T}(g_L))^T$. The manifold regularization term is defined as follows:

$$\|\mathcal{T}\|_{\mathcal{I}}^2 = \sum_{i=1}^L \sum_{j=1}^L W_{ij} (\mathcal{T}(v_i) - \mathcal{T}(v_j))^2 = \text{tr}(t^T At), \quad (21)$$

where $\text{tr}(\cdot)$ indicates the trace of matrix and the solution of cost function is used to update the transformation matrix as $\tau^* = T_{\text{rot}_z} T_{\text{rot}_y} T_{\text{rot}_x} T_{\text{trans}_{xyz}} T_r \cdot \mathcal{T}^*$.

3) Fast Deformation Learning: In order to apply our non-rigid pointset registration to the real-time interventional navigation, false correspondence and long time-consuming should be solved. Therefore, we introduce an acceleration strategy to get the fast deformation learning.

False correspondence during matching often occurs due to the noisy points around candidate points. We apply isotropic Gaussian model $\mathcal{N}(0, \sigma^2)$ to fit the distribution of noisy points [49], and set $\mathbf{v} = (v_1, v_2, \dots, v_L)^T$ and $\mathbf{c} = (c_1, c_2, \dots, c_L)^T$ as two pointsets, respectively. The likelihood of matching rate is a mixture model as $\text{pd}(\mathbf{c}|\mathbf{v}, \theta)$ where pd represents a prior probability distribution, the $\theta = \{\mathcal{T}, \sigma^2\}$ contains the variables to be solved, the $\text{pd}_i = (c_i|v_i, \theta)$ represents the probability that v_i and c_i have a correspondence matching under the constraint of variables θ , which is estimated via the Bayes rule [50]. We use the maximum likelihood estimation (MLE) algorithm to optimize the cost function, which approximates the log posterior $\mathcal{Q}(\theta)$ maxima by maximizing the surrogate function. Considering the relationship between $\mathcal{Q}(\theta)$ and \mathcal{T} , the manifold regularized risk function is

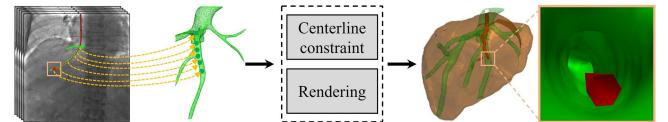


Fig. 5. After completing the deformation learning, we project the tip from the deformed roadmap back onto the 3D branch centerline point. This process enables us to render the instrument tip and visualize its surrounding internal lumen area for intuitive viewing.

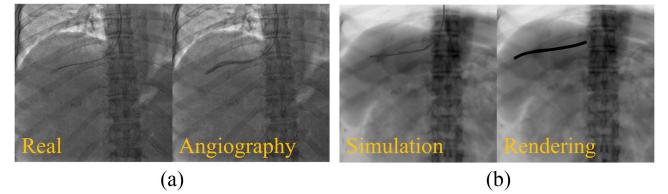


Fig. 6. (a) We collect intraoperative X-ray fluoroscopy and angiography, including the guidewire and enhanced branch. (b) The guidewire is simulated using DRRs, and the branch is rendered in black.

expressed as follows:

$$\begin{aligned} \varepsilon(\mathcal{T}, \sigma^2) = & \frac{1}{2\sigma^2} \sum_{i=1}^L \text{pd}_i \|c_i - \mathcal{T}(v_i)\|^2 \\ & + \frac{\lambda_1}{2} \|\mathcal{T}\|_{\mathcal{H}}^2 + \frac{\lambda_2}{2} \|\mathcal{T}\|_{\mathcal{I}}^2, \end{aligned} \quad (22)$$

the utilization of risk function reduces the mismatch rate between points. For fast implementation, we compute the gradient at each point from \mathbf{v} and \mathbf{c} , determining the step size of down-sampling. The downsampled pointset are $\{v_i\}_{i=1}^{n_s}$ and $\{c_i\}_{i=1}^{m_s}$, respectively, where $n_s, m_s \ll L \leq \min(m, n)$. These sets participate in the registration process to ensure rapid deformation learning for real-time navigation. When we complete the deformation correction, the tip in the X-ray image is back-projected through the previously established correspondence matching between 3D centerline points and 2D projection points (Fig. 5). Then, we connect the tip with the branch root point and render the surrounding area of tip, obtaining an internal view of the 3D lumen around the tip.

IV. EXPERIMENTS AND RESULTS

A. Dataset and Parameters Setting

1) Simulated Dataset: We track the advancing instrument tip in a sequential DRR frames generated from CTA. To ensure our DRR sequence closely approximates the real X-ray image sequence (Fig. 6), we manually determine the length of the instrument within the lumen based on the 3D vessel centerline. Subsequently, we simulate the guidewire and angiographic HV branch by connecting the tip to the HV branch root point and rendering the line segment with lumen diameter at centerline points. Considering the interventional operations, we set the tip to advance at a random speed from 1 mm to 10 mm per frame. We model the normal liver as incompressible tissue [51], and the movement of intrahepatic vessels under free breathing is

TABLE I
BREATHING PARAMETERS FOR DIFFERENT SIMULATED SEQUENCES

Case	Frame	β (s)	A_1 (mm)	A_2 (mm)	σ_n (mm)	σ_{Iso} (mm)
1	90	6.00	20.62	2.50	[1, 2]	[2.02, 2.02]
2	90	6.00	16.75	2.50	[1, 2]	[2.02, 2.02]
3	90	8.00	16.75	2.50	[1, 2]	[2.02, 2.02]
4	90	8.00	20.62	2.50	[1, 2]	[2.02, 2.02]
5	90	10.00	20.62	2.50	[1, 2]	[2.02, 2.02]
6	100	6.00	16.41	2.50	[3, 5]	[2.96, 2.96]
7	100	6.00	16.41	3.50	[4, 6]	[2.96, 2.96]
8	100	8.00	19.39	3.50	[4, 6]	[2.96, 2.96]
9	100	8.00	19.39	2.50	[3, 5]	[2.02, 2.02]
10	100	10.00	19.39	3.50	[3, 5]	[2.96, 2.96]
11	120	6.00	19.53	3.50	[4, 6]	[2.96, 2.96]
12	120	6.00	19.53	3.50	[3, 5]	[2.96, 2.96]
13	120	8.00	19.53	2.50	[4, 6]	[2.96, 2.96]
14	120	8.00	16.75	3.50	[4, 6]	[2.96, 2.96]
15	120	10.00	16.75	2.50	[4, 6]	[2.96, 2.96]
16	110	4.00	23.94	4.50	[1, 2]	[1.23, 1.23]
17	110	4.00	23.94	3.50	[1, 2]	[1.23, 1.23]
18	110	6.00	23.94	3.50	[3, 5]	[1.85, 1.85]
19	110	6.00	19.11	3.50	[3, 5]	[1.85, 1.85]
20	110	6.00	14.28	4.50	[4, 6]	[2.46, 2.46]

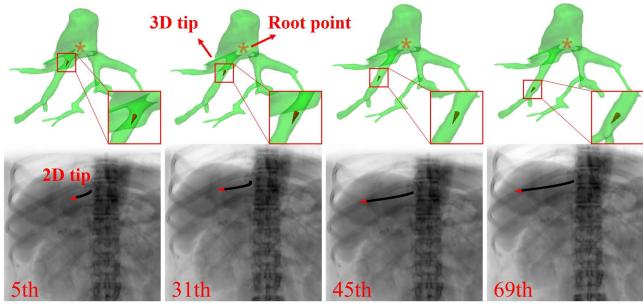


Fig. 7. We simulate the continuous advancement of the 3D tip (red) alongside the enhanced HV branch (black). The upper row displays various positions of the 3D tip within the HV branch, the bottom row illustrates the enhanced HV branch achieved by thickening the connection between the tip and the branch root point.

primarily influenced by the liver parenchyma [52]. Therefore, the respiratory motion is simulated in the head-foot and left-right directions [53], defining the amplitudes of respiration as follows:

$$\begin{aligned} D_p(i) &= A_1 \cdot \sin(2\pi i \Delta_t / \beta), \\ D_v(i) &= A_2 \cdot \sin(2\pi i \Delta_t / \beta), \end{aligned} \quad (23)$$

where D_p and D_v are the displacements in two directions (head-foot and left-right). The parameters encompass the peak amplitude A_1 and A_2 , a cycle time β , frames i , and interval time Δ_t . To simulate the deformation of the HV branch, we apply a Gaussian smoothing model with a random $\sigma_n \in [1, 6]$ mm on the instrument. We set σ_{iso} to simulate the center offset in projection, while the other parameters are obtained from the DICOM file, including magnification ratio, field size, and positioner angle. The details of the simulated breathing parameters, derived from the respiratory motion of real patients, are presented in Table I. Fig. 7 illustrates the simulated X-ray image sequence, including a continuous angiography procedure as the instrument approaches through the HV branch and releases contrast agents.

2) Clinical Dataset: We retrospectively collect a clinical dataset from the Department of Interventional Radiology at the First and Fifth Medical Center of PLA General Hospital. Informed consent to participate in the study was obtained from all participants, this study is approved by the Ethics Committee of PLA General Hospital (No. S2023-313-01) and is conducted in accordance with the Declaration of Helsinki. Each patient underwent a 3D CT scan with a resolution of 512×512 pixels and slice thickness ranging from 1.2 mm to 5 mm. The 3D vessel tree centerlines are extracted from CTA volume using morphological operations during preprocessing, providing the coordinates and lumen radius of each point of the 3D centerline. X-ray image sequences are obtained using an angiographic C-arm system. In total, we collect 106 short X-ray sequences (27–53) frames and 40 long X-ray sequences (65–120) frames from various TIPS cases. A total of 24 cases are utilized for framework evaluation and result presentation. For each X-ray fluoroscopic image Z_t in every case, the 2D instrument centerline was segmented automatically in advance [41]. Additionally, the 3D tip location is manually annotated on the 3D vessel tree for the initial frame of each X-ray image sequence to initialize the 3D instrument path V_{3D} .

3) Parameters Setting: In our experiment, we set the parameters as follows: $\mu_1 = 0.2$, $\mu_2 = 0.26$ in (3). Metric scaling $\sigma_o = 0.25$, $\gamma_1 = 2$ and $\gamma_2 = 0.85$ in (5). $\alpha_1 = \alpha_2 = 1$, $\alpha_3 = 4$, and $\alpha_4 = 6$ in (10). Kernel size $\delta = 5$. Additionally, in (18), λ_1 and λ_2 are employed to constrain the transformation to exploit the geometric completeness. In (19), parameter ϵ establishes the graph Laplacian and computes the weight of its edges, with settings as follows: $\lambda_1 = 2$, $\lambda_2 = 0.08$, $\epsilon = 0.05$. The σ in (22) is set to 0.86. In this study, our generation and segmentation models are implemented with PyTorch. The two models use the Adam optimizer with the initial learning rates of 2e-4 and 5e-4, respectively, and the same decay rate of 0.95. The two models are trained for 300 and 100 epochs, respectively. The plot of training loss curves for the two models is shown in Fig. 8. The experiments are conducted on a desktop equipped with an Intel Core i7-9700K CPU (3.6 GHz), a 32GB RAM, and an NVIDIA GeForce RTX-2060 GPU (6GB).

B. Evaluation Metrics

In simulated experiments, where the 3D position of the instrument tip within the HV lumen is known, we utilize the tip positioning error E_p to quantitatively evaluate the sequential simulated images. However, in clinical cases, the ground truth of 3D instrument tip position is unavailable. Hence, we resort to employing the 2D HV branch segmentation obtained from angiographic images as a reference to evaluate the dynamic deformed branch roadmapping in sequential clinical images. In this process, we use the average closest distance error E_c [54] as a quantitative assessment metric.

C. Comparison Methods

In the comparison of different methods, CTT [30] is used as a baseline method for improving tip tracking. Additionally, we incorporate Res-Dia (Respiratory-Diaphragm) and Res-Cen

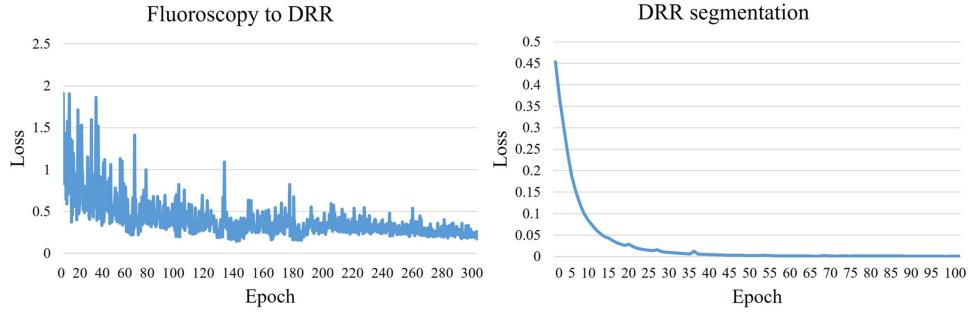


Fig. 8. Two training loss curves for our generation and segmentation models, respectively.

TABLE II
AVERAGE POSITIONING ERROR E_p (MM) FOR 20 SIMULATED CASES USING DIFFERENT COMPARISON METHODS

Case	Rigid Methods						Non-rigid Methods		
	DT [24]	ICC [16]	CTT [30]	CRD [55]	ICGM [18]	CRC [33]	HTSR [17]	MCTS [27]	DITT (Ours)
1	13.25 ± 2.62	11.49 ± 1.32	1.89 ± 1.16	1.73 ± 0.94	1.69 ± 0.76	1.62 ± 0.73	1.56 ± 0.52	1.49 ± 0.72	1.43 ± 0.89
2	11.07 ± 3.14	10.15 ± 1.74	1.82 ± 1.08	1.66 ± 1.02	1.62 ± 0.69	1.58 ± 0.99	1.52 ± 0.59	1.40 ± 0.54	1.35 ± 0.70
3	13.13 ± 2.78	11.28 ± 2.46	1.64 ± 0.46	1.57 ± 0.37	1.54 ± 0.38	1.53 ± 0.54	1.45 ± 0.67	1.33 ± 0.19	1.29 ± 0.29
4	12.99 ± 2.48	11.08 ± 1.26	2.03 ± 1.25	1.95 ± 0.64	1.94 ± 0.52	1.89 ± 0.35	1.83 ± 0.59	1.58 ± 0.46	1.49 ± 0.68
5	13.04 ± 2.87	11.29 ± 1.13	2.28 ± 1.18	2.21 ± 0.48	2.16 ± 0.54	2.13 ± 0.46	2.09 ± 0.31	1.46 ± 0.53	1.39 ± 0.36
6	12.51 ± 2.31	12.03 ± 1.96	3.18 ± 1.34	3.14 ± 0.99	3.10 ± 0.84	3.08 ± 0.35	3.05 ± 0.62	1.56 ± 0.54	1.34 ± 0.40
7	11.68 ± 2.88	10.94 ± 1.25	4.26 ± 1.29	4.18 ± 1.04	4.12 ± 0.95	4.09 ± 0.64	4.02 ± 0.40	1.59 ± 0.35	1.35 ± 0.48
8	15.91 ± 2.60	13.48 ± 1.93	7.26 ± 0.59	7.21 ± 0.51	7.19 ± 0.51	7.19 ± 0.77	7.02 ± 0.40	2.46 ± 0.95	2.25 ± 1.46
9	13.45 ± 2.16	12.79 ± 2.01	6.04 ± 1.56	5.86 ± 0.48	5.61 ± 0.49	5.24 ± 0.32	5.01 ± 0.46	2.11 ± 0.38	1.86 ± 0.37
10	14.24 ± 2.68	13.47 ± 1.64	6.53 ± 1.42	6.37 ± 0.57	6.28 ± 0.48	6.20 ± 0.57	5.24 ± 0.68	2.12 ± 0.62	1.91 ± 0.47
11	13.82 ± 4.08	11.31 ± 2.14	5.07 ± 1.20	5.04 ± 0.19	5.03 ± 0.38	5.02 ± 0.21	4.74 ± 0.78	1.98 ± 0.52	1.85 ± 0.36
12	12.40 ± 3.37	12.15 ± 2.16	6.43 ± 0.27	6.42 ± 0.27	6.42 ± 0.35	6.41 ± 0.29	5.64 ± 0.34	2.03 ± 0.40	1.45 ± 0.56
13	13.91 ± 3.68	13.35 ± 3.10	4.28 ± 1.34	4.21 ± 0.74	3.73 ± 0.85	3.21 ± 0.34	2.28 ± 0.54	1.69 ± 0.55	1.52 ± 0.66
14	11.65 ± 2.96	10.68 ± 2.04	4.13 ± 0.93	4.01 ± 0.73	3.57 ± 0.71	2.84 ± 0.85	2.15 ± 0.67	1.78 ± 0.39	1.66 ± 0.48
15	11.05 ± 2.18	10.02 ± 2.14	3.94 ± 0.85	3.44 ± 1.05	3.09 ± 0.55	2.76 ± 0.67	2.34 ± 0.41	1.68 ± 0.62	1.49 ± 0.53
16	13.94 ± 2.68	13.28 ± 2.23	5.75 ± 1.35	5.37 ± 1.14	4.61 ± 0.58	4.19 ± 0.66	4.10 ± 0.74	2.07 ± 0.76	1.63 ± 0.69
17	13.88 ± 2.55	13.12 ± 2.05	5.26 ± 1.53	5.19 ± 0.85	4.43 ± 0.96	4.12 ± 0.85	4.08 ± 0.88	1.70 ± 0.48	1.57 ± 0.56
18	14.31 ± 2.74	13.52 ± 2.57	6.98 ± 1.49	6.45 ± 0.94	6.03 ± 1.15	5.46 ± 0.72	4.75 ± 0.64	2.24 ± 0.63	1.73 ± 0.64
19	12.25 ± 3.10	11.96 ± 2.68	6.28 ± 1.52	6.03 ± 1.41	5.84 ± 1.13	5.13 ± 1.05	4.52 ± 1.21	2.08 ± 0.94	1.66 ± 0.51
20	11.49 ± 2.53	10.37 ± 2.47	4.02 ± 1.05	3.76 ± 0.89	3.43 ± 0.56	2.80 ± 0.35	2.26 ± 0.34	1.82 ± 0.51	1.71 ± 0.41
Mean	12.99 ± 2.89	11.89 ± 2.03	4.45 ± 1.26	4.28 ± 0.77	4.07 ± 0.68	3.82 ± 0.60	3.48 ± 0.63	1.81 ± 0.57	1.59 ± 0.57

(Respiratory-Centroid), which respectively utilize diaphragm-based [55] and centroid-based [33] respiration compensation algorithms as enhancement modules. These are combined with CTT to form CRD and CRC, respectively. Alongside, we include 3D/2D vessel pointset registration algorithms as comparison methods, such as DT [24], ICC [16], HTSR [17], ICGM [18], and MCTS [27]. The maximum number of iterations for DT, ICC, and ICGM is set to 300, 10, and 10, respectively. In addition, three radiologists, each with over 13 years of experience in TIPS procedures, are invited to evaluate our framework qualitatively.

D. Evaluation on Simulated Dataset

We evaluate tip tracking accuracy using the average positioning error across 20 simulated cases. Table II shows average tip positioning errors for different methods, with each column containing the mean value and the standard deviation (STD). The best-performing method is highlighted in bold. DITT achieves a mean E_p of 1.59 ± 0.57 mm, representing a notable 12.14% reduction in tracking error compared with the state-of-the-art method. Our framework yields statistically significant

improvements (p -value < 0.05 , Wilcoxon signed-rank test) compared to related methods evaluated. Both DITT and MCTS employ non-rigid pointset registration, leading to significantly enhanced accuracy compared to rigid registration methods. Moreover, experimental results reveal that breathing amplitude and deformation are primary factors influencing accuracy, particularly head-foot motion.

In order to provide radiologists with reliable 2D vessel branch postures on X-ray images without using contrast agents, we utilize common segmentation evaluation metrics to verify the performance of different methods. These metrics contain Dice similarity coefficient (DSC), Accuracy (ACC), Sensitivity (SEN), and Jaccard similarity coefficient (JSC). A total of 20 simulated cases are applied to test the performance of four competing methods, i.e., CTT, CRD, CRC, and DITT. The experimental results are summarized in Table III, DITT achieves the best branch postures overlapping in all simulated cases, statistically significant differences (p -value < 0.01 , Wilcoxon signed-rank test) are obtained. It shows that the deformed roadmaps provided by our proposed framework is well consistent with the morphology of intraoperative vessel branches.

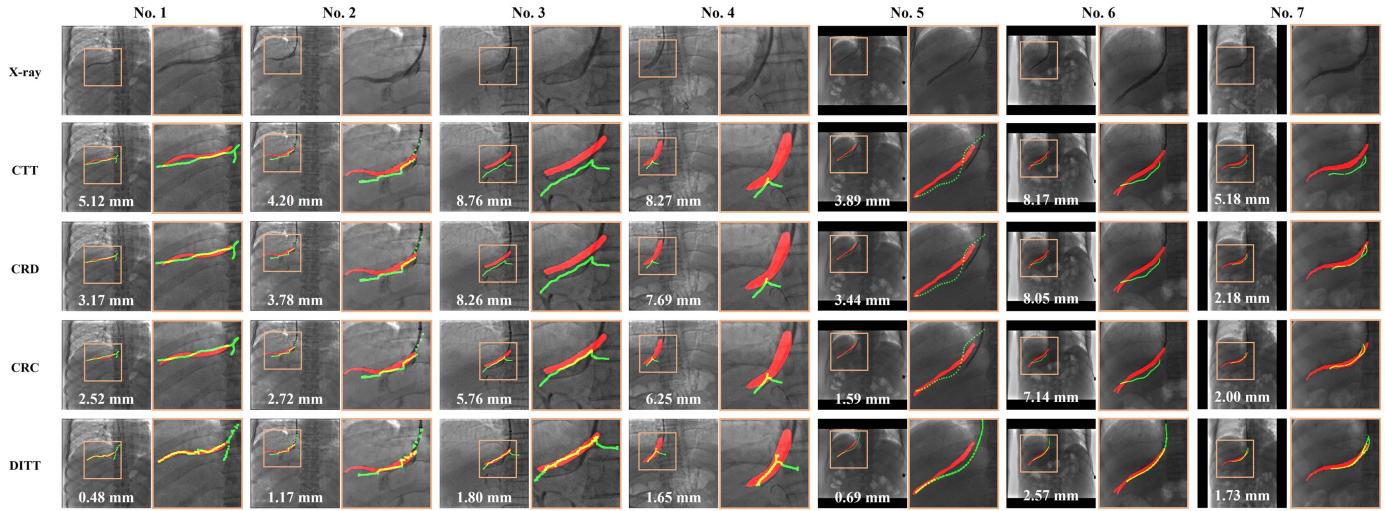


Fig. 9. The projected 3D branch (green) are aligned with the 2D instrument. During the intraoperative guidance, the branch roadmapping is dynamically overlapped on the 2D image plane. It is evaluated with 2D angiographic branch (red) with the average closest distance error at the bottom of each frame. The single-frame results of 7 random clinical cases obtained by different methods are shown.

TABLE III
THE PERFORMANCE ON BRANCH POSTURE WITH MEAN AND STD FOR COMPARISON METHODS ACROSS 20 SIMULATED CASES

Methods	DSC (%)	ACC (%)	SEN (%)	JSC (%)
CTT [30]	36.49 ± 8.24	45.25 ± 6.42	50.81 ± 9.54	45.84 ± 4.92
CRD [55]	51.11 ± 5.64	50.48 ± 6.37	49.35 ± 3.84	48.89 ± 5.57
CRC [33]	72.42 ± 4.67	74.63 ± 5.01	75.43 ± 6.01	67.96 ± 5.58
DITT (Ours)	90.87 ± 3.96	92.97 ± 3.85	91.04 ± 3.44	85.78 ± 5.67

E. Evaluation on Clinical Dataset

The clinical dataset comprises X-ray fluoroscopic images and X-ray angiographic images. The enhanced HV branches extracted from the angiographic images serve as the ground truth, which is utilized to evaluate the deformation correction and tip tracking. Accordingly, we designed two distinct experiments as follows:

1) Quantitative Evaluation: The quantitative experiments are conducted using sequential fluoroscopic images and evaluated using sequential X-ray angiographic images. The E_c between 3D projected HV branch and 2D enhanced HV branch around the tip is calculated.

2) Qualitative Evaluation: The qualitative experiments are conducted using sequential fluoroscopic images with 2D deformed dynamic roadmapping. Several experienced radiologists evaluate the clinical effectiveness of our framework using roadmapping as the interventional guidance.

Table IV shows E_c (within a 30 mm radius area around the tip) value for 24 cases after deformation correction. Our framework achieves a 15.23% error reduction compared with the state-of-the-art method. Statistically significant differences (p -value < 0.05 , Wilcoxon signed-rank test) are obtained when comparing DITT with related methods in the experiments. Apart from error E_c , computational efficiency is crucial for real-time navigation. Table V illustrates that the tip tracking frequency of CTT, CRD, and CRC reaches up to 20 Hz, while DITT reaches up

to 23 Hz, meeting the demand for intraoperative real-time X-ray image guidance [56]. Therefore, we present clinical tracking results obtained by CTT, CRD, CRC, and DITT for 7 random cases in Fig. 9. The fluoroscopy displays the overlapped branch roadmap, with the 2D enhanced vessel branch serving as the ground truth for the evaluation of the tracking accuracy. Specifically, for the No. 1, 2, 5, 6, and 7 columns, where the instrument penetrates deeply into the branch, DITT achieve accurate results. Conversely, in No. 3 and 4 columns, where the length of instrument is insufficient, errors are higher. In clinical X-ray images, diaphragm visibility is not guaranteed due to many factors such as a limited field of view or C-arm angulation choice [57]. This leads to the poor results with CRD, which is sensitive to diaphragm dome motion. Compared with DITT, CRD and CRC cannot mitigate the influence of vessel deformation during rigid registration, and thereby the dynamic roadmapping during intervention fails to illustrate vessel branch posture.

We illustrate the performance trends of various methods on both simulated and clinical datasets using error curves, depicted in Fig. 10. The X-axis represents the frame timeline, and the Y-axis indicates the positioning error and the closest distance error. Each point on the curve reflects the mean tracking error at the current frame across different cases. CTT exhibits considerable fluctuations due to its sensitivity to the length of the entering instrument. The result of CRD relies on diaphragm integrity, resulting in increased errors when the diaphragm gradually disappears during breathing in some cases. Additionally, the unstable output of CRC is influenced by changes in image intensity caused by diaphragm alterations. In No. 5, 6, and 7 columns from Fig. 9, umaged areas along the edges further interfere with centroid calculation. In contrast, our DITT yields more stable and accurate results compared to the comparison methods. This indicates the potential of our framework for application across diverse patient scenarios and complex intraoperative situations.

During the qualitative evaluation, three radiologists evaluate the tracking results obtained using different comparison methods

TABLE IV
AVERAGE CLOSEST DISTANCE ERROR E_c (MM) FOR 24 CLINICAL CASES DURING TIP TRACKING USING DIFFERENT COMPARISON METHODS

Case	Rigid Methods							Non-rigid Methods	
	DT [24]	ICC [16]	CTT [30]	CRD [55]	ICGM [18]	CRC [33]	HTSR [17]	MCTS [27]	DITT (Ours)
1	13.62 ± 1.32	12.49 ± 1.64	5.87 ± 0.53	4.22 ± 0.73	3.92 ± 0.64	3.87 ± 0.84	1.68 ± 0.32	0.94 ± 0.35	0.69 ± 0.27
2	13.10 ± 1.95	12.98 ± 1.88	4.45 ± 1.62	4.20 ± 1.64	3.61 ± 0.89	2.80 ± 0.94	2.71 ± 0.84	2.46 ± 0.56	2.05 ± 0.64
3	16.64 ± 2.09	15.84 ± 1.80	9.22 ± 0.71	8.37 ± 0.64	7.15 ± 1.24	5.39 ± 0.88	3.25 ± 0.94	2.16 ± 0.48	1.94 ± 0.62
4	12.95 ± 1.45	12.38 ± 1.49	2.95 ± 0.55	1.51 ± 0.46	1.36 ± 0.52	1.29 ± 0.19	1.27 ± 0.16	1.23 ± 0.25	1.19 ± 0.43
5	12.69 ± 1.84	12.28 ± 2.21	8.17 ± 1.55	7.87 ± 0.61	7.34 ± 1.36	7.12 ± 0.74	5.56 ± 1.98	2.25 ± 0.43	1.92 ± 0.69
6	10.86 ± 2.68	09.89 ± 2.08	3.08 ± 1.20	2.07 ± 0.80	1.95 ± 0.42	1.83 ± 0.58	1.59 ± 0.69	1.37 ± 0.46	1.02 ± 0.47
7	13.28 ± 1.58	12.94 ± 1.62	5.94 ± 1.90	5.75 ± 2.00	5.01 ± 1.69	4.48 ± 1.69	3.80 ± 1.26	2.45 ± 0.85	2.32 ± 0.91
8	14.69 ± 1.75	14.03 ± 1.34	8.16 ± 0.88	8.24 ± 0.85	8.09 ± 1.65	7.84 ± 0.70	5.28 ± 1.59	1.96 ± 0.35	1.88 ± 0.26
9	13.95 ± 1.56	12.86 ± 1.04	6.24 ± 1.82	3.45 ± 0.87	3.08 ± 0.51	2.77 ± 0.71	2.73 ± 0.96	2.61 ± 0.62	2.51 ± 0.47
10	11.36 ± 1.49	11.12 ± 1.43	4.93 ± 0.94	4.61 ± 0.88	3.94 ± 0.82	3.57 ± 0.91	3.29 ± 0.78	2.08 ± 0.44	1.70 ± 0.51
11	12.48 ± 1.66	11.28 ± 1.54	4.51 ± 1.08	4.45 ± 0.96	3.84 ± 1.21	3.19 ± 0.78	2.67 ± 0.85	1.92 ± 0.42	1.51 ± 0.48
12	11.72 ± 1.57	11.05 ± 1.21	4.81 ± 0.87	4.17 ± 1.02	3.41 ± 0.95	2.91 ± 1.06	2.55 ± 1.04	1.84 ± 0.34	1.22 ± 0.44
13	11.85 ± 1.43	10.57 ± 1.09	4.36 ± 1.05	4.08 ± 0.79	3.79 ± 1.08	3.14 ± 1.11	2.94 ± 1.16	1.96 ± 0.37	1.57 ± 0.54
14	12.94 ± 1.68	10.98 ± 1.38	6.98 ± 1.54	6.27 ± 1.25	5.43 ± 1.20	4.25 ± 1.19	3.41 ± 1.23	2.14 ± 0.29	1.85 ± 0.62
15	10.59 ± 1.67	09.88 ± 1.41	4.19 ± 0.58	4.05 ± 0.82	3.49 ± 0.75	3.25 ± 0.78	3.18 ± 0.97	2.05 ± 0.33	1.64 ± 0.49
16	10.95 ± 1.55	10.04 ± 1.32	4.38 ± 1.32	4.02 ± 1.03	3.82 ± 0.98	3.30 ± 0.86	2.96 ± 1.04	2.08 ± 0.41	1.66 ± 0.51
17	11.63 ± 1.46	10.42 ± 1.28	5.01 ± 0.68	4.37 ± 0.94	3.72 ± 1.14	3.24 ± 1.03	2.71 ± 0.92	1.94 ± 0.47	1.72 ± 0.45
18	12.67 ± 1.54	11.09 ± 1.57	5.75 ± 0.99	4.96 ± 0.87	4.12 ± 1.15	3.39 ± 0.79	2.84 ± 0.87	2.03 ± 0.36	1.74 ± 0.46
19	10.77 ± 1.54	09.67 ± 1.21	4.49 ± 0.76	4.11 ± 0.85	3.34 ± 1.04	2.85 ± 0.94	2.36 ± 0.86	1.88 ± 0.47	1.48 ± 0.55
20	11.27 ± 1.33	10.91 ± 1.04	4.63 ± 0.79	4.40 ± 0.86	3.72 ± 0.94	3.06 ± 0.95	2.59 ± 0.96	1.93 ± 0.44	1.60 ± 0.55
21	11.97 ± 1.35	11.06 ± 1.18	5.21 ± 1.94	4.52 ± 0.91	3.44 ± 0.88	3.27 ± 0.97	2.75 ± 1.01	1.95 ± 0.53	1.74 ± 0.49
22	12.75 ± 1.48	11.19 ± 1.25	5.60 ± 0.85	4.83 ± 0.65	4.06 ± 0.85	3.29 ± 0.88	2.97 ± 0.87	2.04 ± 0.49	1.81 ± 0.68
23	11.86 ± 1.28	10.92 ± 1.10	5.19 ± 0.89	4.35 ± 0.95	3.78 ± 0.94	3.13 ± 0.71	2.91 ± 0.77	1.95 ± 0.51	1.70 ± 0.50
24	12.13 ± 1.19	11.04 ± 1.13	5.31 ± 0.92	4.37 ± 0.79	3.85 ± 0.96	3.17 ± 0.93	2.94 ± 0.88	2.03 ± 0.47	1.76 ± 0.48
Mean	12.45 ± 1.50	11.54 ± 1.58	5.39 ± 1.15	4.72 ± 0.96	4.14 ± 1.02	3.60 ± 0.87	2.96 ± 0.98	1.97 ± 0.48	1.67 ± 0.54

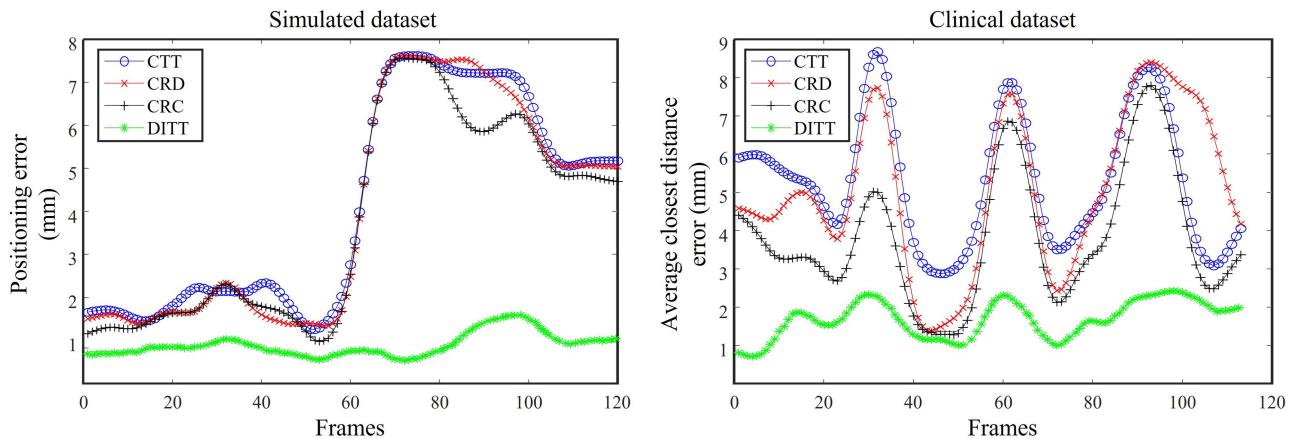


Fig. 10. The error curve illustrates the continuous performance across all simulated and clinical datasets, demonstrating the consistent outperformance of DITT over CTT, CRD, and CRC in terms of accuracy and applicability. Moreover, our framework exhibits efficacy in scenarios involving short instruments and local X-ray imaging.

TABLE V
AVERAGE TIME CONSUMPTION (MS) FOR EACH CLINICAL X-RAY SEQUENCE WITH DIFFERENT IMAGING SIZES

Methods	512 × 512 (Case 1–Case 5)	1024 × 1024 (Case 6–Case 9)
DT [24]	2306	2483
ICC [16]	1915	2183
HTSR [17]	1637	1853
ICGM [18]	1538	1802
MCTS [27]	1161	1527
CTT [30]	45	56
CRD [55]	50	66
CRC [33]	49	62
DITT (Ours)	43	49

on clinical sequences. We establish three different observations and summarized the outcomes in Fig. 11. DITT shows the quickest initiation of effective guidance. In the same clinical sequence, DITT yields more frames with correct tracking and clinically effective observation (Excellent). Although there are still some frames with incorrect tracking but clinically effective observation (Good), because the initial frames lack sufficient data to fit a motion curve at the onset of the procedures. Additionally, frames with incorrect tracking and clinical ineffective observation (Poor) occurred when the instrument initially entered the branch, the best result of pose alignment have not been achieved. Compared with other methods, DITT displays less

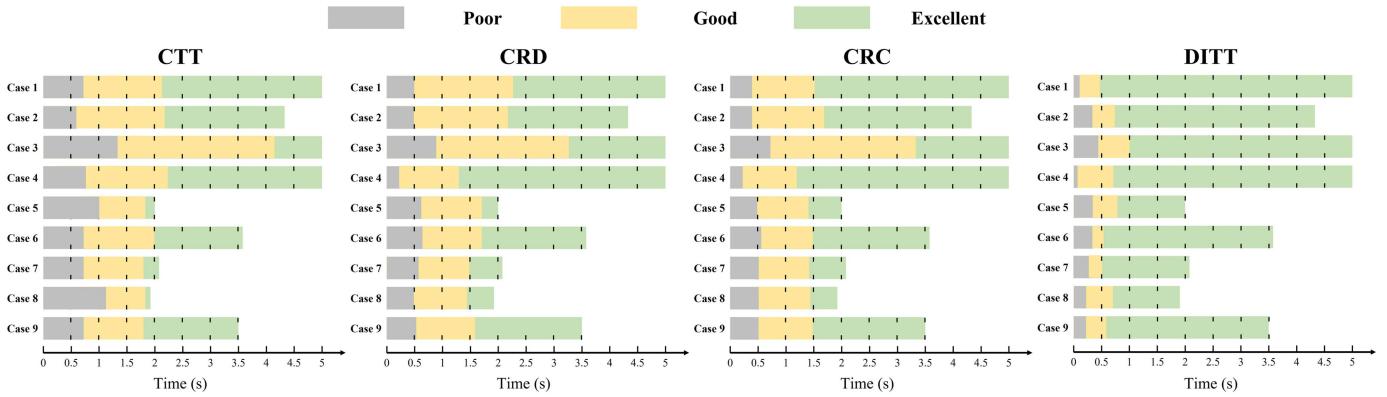


Fig. 11. Three radiologists conduct a qualitative evaluation of tracking performance yielded by four competing methods on 24 clinical X-ray image sequences. Each frame of the tip tracking sequence is scored with one of three observations: Excellent (green), Good (orange), and Poor (gray).

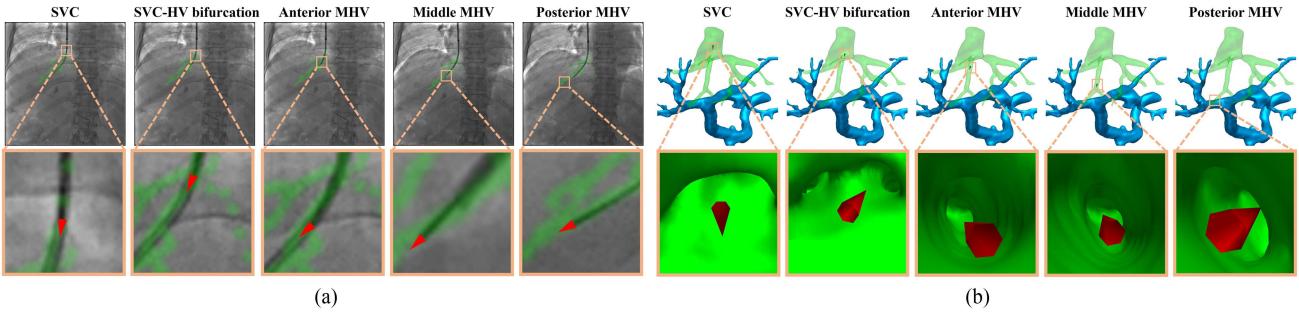


Fig. 12. The guidance outcomes in the TIPS procedure. (a) The advancing tip (red) on 2D vessel roadmapping (green). (b) The corresponding 3D advancing tip (red) at various structures, such as different sections (anterior, middle, and posterior) of the HV middle-branch (HMV).

dependency on the length of entered instrument. Respiration compensation and deformation correction ensure accurate overlapping of dynamic roadmaps, facilitating effective intervention guidance and enhancing operational efficiency.

Our framework renders 3D tip tracking within the HV branch lumen. As the interventional instrument tip enters the HV branch, its 2D in-plane position is determined in X-ray fluoroscopy (Fig. 12(a)). Subsequently, its corresponding 3D position within HV lumen is obtained in real time by DITT. Leveraging the coordinates of the tip and its trajectory, we render the 3D advancing tip at different HV structures (Fig. 12(b)). This approach enhances the control of interventional instrument during puncture, particularly at the SVC-HV bifurcation, which offers multiple branch entrances in different directions at the current view of tip. Such visualization aids in intuitive tracking and mitigates the risk of vessel wall damage.

F. Ablation Study

In order to evaluate the effectiveness of adaptive sampling (AS) in pose alignment module, respiration compensation (RC) module, and deformation correction (DC) module, an ablation study is conducted. Compared with CTT and our full framework (DITT), a total of 6 schemes with different module combinations are examined: three schemes involves integrating each single module (i.e., AS, RC, and DC modules) into CTT, forming CAS,

TABLE VI
THE MEAN, STANDARD DEVIATION (STD), AND MEDIAN OF POSITIONING ERROR E_p (MM) FOR ALL SIMULATED CASES IN THE ABLATION STUDY

Scheme	CTT	AS	RC	DC	Mean	Std	Median
1	✓				4.08	1.92	4.15
2	✓		✓		3.88	1.89	3.84
3	✓			✓	3.73	1.80	3.65
4	✓			✓	1.91	0.44	1.80
5	✓	✓	✓		3.13	1.35	3.25
6	✓	✓		✓	1.77	0.35	1.72
7	✓		✓	✓	1.68	0.31	1.56
8	✓	✓	✓	✓	1.59	0.57	1.49

CRC, and CDC, respectively, the other three schemes integrates every two modules into CTT, forming CASRC, CASDC, and CRCDC, respectively. The results, evaluated by the average tip E_p positioning error of each scheme, are presented in Table VI. The results show that each module is able to improve the accuracy of tip positioning. CAS, CRC, and CASRC contain rigid registration modules only, their performance is lower than CDC, CRCDC, and DITT. Additionally, AS and RC provide a better initial pose for non-rigid registration, improving the feature matching between the instrument and vessel branch. In general, three proposed modules in DITT demonstrate their effectiveness.

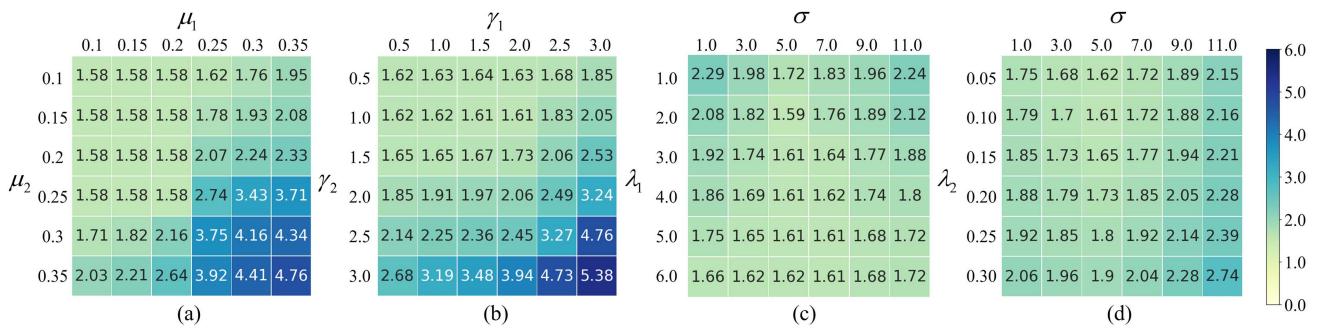


Fig. 13. Mean positioning errors of the 3D tip on the simulated dataset with various parameters. Experimental results for μ_1 and μ_2 (a), γ_1 and γ_2 (b), σ and λ_1 (c), σ and λ_2 (d).

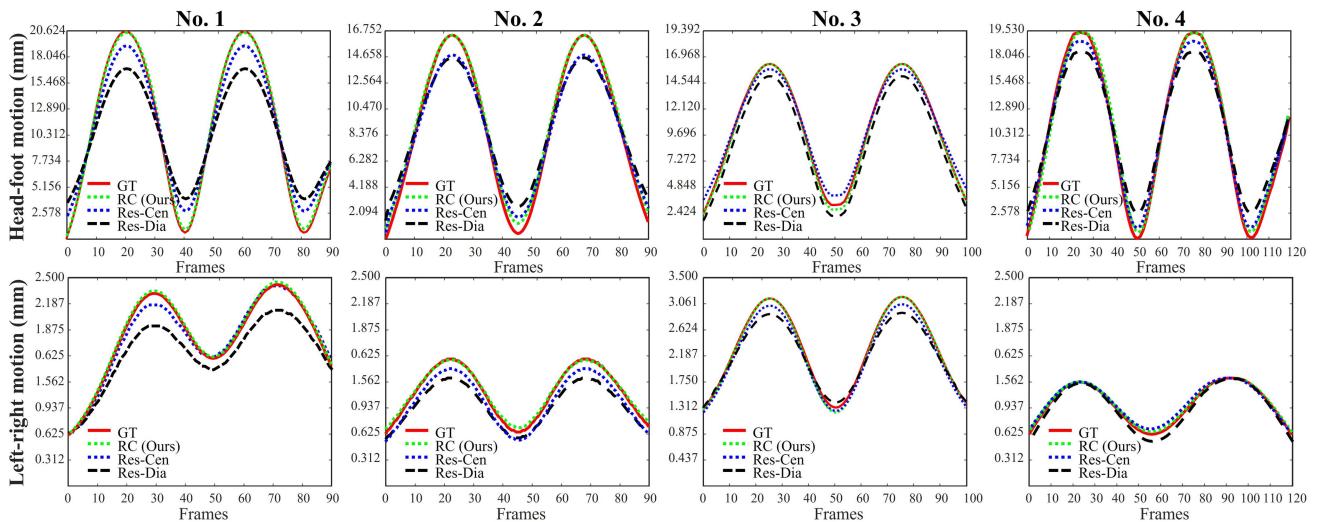


Fig. 14. Comparison of results from different respiratory estimation methods with ground truth along the head-foot and left-right directions in 4 randomly selected simulated cases. The estimation methods include Res-Dia, Res-Cen, and our RC.

V. DISCUSSION

A. Sensitivity to Parameters

Fig. 13 shows the tip tracking results on the simulated dataset for various μ_1 , μ_2 , γ_1 , γ_2 , σ , λ_1 , and λ_2 . The positioning error is sensitive to these parameters differently. For the curvature thresholds μ_1 and μ_2 in the PA module, the μ_2 has a greater impact than μ_1 on the error, as the 2D instrument is more simple and has less feature points to be sampled. In addition, more feature points will lead to the increased time consumption in subsequent computation, thus we utilize the higher curvature threshold to ensure the efficiency while maintaining the accuracy. γ_1 and γ_2 constrain the tip to be tracked around the 3D centerline, allowing the tip to move within the radius of vessel lumen. In (22), the σ controls the contribution between corresponding matching points, λ_1 determines the degree of smoothness, and λ_2 focuses on maintaining the geometric structure. The results in Fig. 13 show that a good balance between these parameters can result in the high tracking accuracy.

B. Respiratory Estimation Accuracy

The accuracy of respiratory estimation directly impacts the quality of vessel/instrument registration. Given the absence of

ground truth of the respiratory curve in clinical X-ray image sequences, we evaluate estimation performance using a simulated dataset. We compare the breathing curve estimations from Res-Dia, Res-Cen, and our method (RC) with ground truth along both the head-foot and left-right directions. As shown in Fig. 14, four randomly selected simulated cases are compared to evaluate differences in breathing curves. The results show that RC achieves superior estimation accuracy, providing the foundation for updating registration transformation between frames.

C. Limitations and Future Work

The tip tracking within our framework may be constrained by the accuracy of respiratory motion estimation. The liver expansion and compression motion introduce significant complexity to the branch deformation learning. This challenge can be addressed by directly estimating the 3D respiratory motion of liver vessels during free breathing. To our knowledge, no study has directly modeled the respiratory dynamics of liver vessels. Previous research has linked the displacement and deformation of 3D intrahepatic vessels with liver parenchyma motion [58], integrating preoperative four-dimensional (4D) CT with intraoperative imaging to estimate 3D + t respiratory motion [59]. Along with the shape and edge features of the liver [60], the bifurcation

and topological features [27] should also be incorporated into respiration modeling. In the future work, we would like to collect 4D CTA datasets of liver vessels and utilize statistical modeling to analyze the motion of each branch under free breathing. Additionally, by dynamically registering statistical templates with intraoperative enhanced branch images, we aim to obtain a patient-specific 3D respiratory motion estimation, thereby enhancing clinical applicability.

VI. CONCLUSION

In this paper, our framework tracks the instrument tip within vessel branch using preoperative CTA volume and intraoperative X-ray sequence for the TIPS procedure guidance. First, we propose a curvature-based adaptive point sampling strategy in the PA module to increase the number of semantic 3D/2D features, thereby alleviating registration difficulties and improving the accuracy of 3D vessel branch and 2D instrument registration. Second, we design an RC module to compensate for respiratory motion using the liver masks extracted from X-ray image sequence. The obtained motion prior drives the initially projected 3D branch to move intraoperatively. Finally, we employ a DC module to correct the HV branch deformation during interventional procedures. This module incorporates manifold regularization and an acceleration strategy to achieve reliable and efficient deformation learning, significantly improving tip tracking accuracy and preventing accidental injury to vessel lumen and tissue. Experimental results on both simulated and clinical datasets demonstrate that our framework achieves better tracking performance. The DITT framework illustrates the inspiring potential in clinical guidance, offering 2D dynamic deformed roadmapping and 3D real-time tip tracking within branch lumen. The framework could potentially be applied to free-breathing X-ray image-guided interventions without the need of contrast agents.

REFERENCES

- [1] F. Pernus et al., “3D-2D registration of cerebral angiograms: A method and evaluation on clinical images,” *IEEE Trans. Med. Imag.*, vol. 32, no. 8, pp. 1550–1563, Aug. 2013.
- [2] P. Markelj et al., “A review of 3D/2D registration methods for image-guided interventions,” *Med. Image Anal.*, vol. 16, no. 3, pp. 642–661, 2012.
- [3] S. A. Baert et al., “Three-dimensional guide-wire reconstruction from biplane image sequences for integrated display in 3-D vasculature,” *IEEE Trans. Med. Imag.*, vol. 22, no. 10, pp. 1252–1258, Oct. 2003.
- [4] M. Wagner et al., “4D interventional device reconstruction from biplane fluoroscopy,” *Med. Pyhs.*, vol. 43, no. 3, pp. 1324–1334, 2016.
- [5] M. G. Wagner et al., “Feasibility of 3D motion-compensated needle guidance for TIPS procedures,” *Proc. SPIE*, vol. 11315, pp. 30–36, 2020.
- [6] N. Chabi et al., “Self-calibration of C-arm imaging system using interventional instruments during an intracranial biplane angiography,” *Int. J. Comput. Assist. Radiol. Surg.*, vol. 17, pp. 1355–1366, 2022.
- [7] X. Duan et al., “A novel robotic bronchoscope system for navigation and biopsy of pulmonary lesions,” *Cyborg Bionic Syst.*, vol. 4, 2023, Art. no. 0013.
- [8] H. Mo et al., “Task autonomy of a flexible endoscopic system for laser-assisted surgery,” *Cyborg Bionic Syst.*, vol. 2022, 2022, Art. no. 9759504.
- [9] L. R. Rabiner, “A tutorial on hidden Markov models and selected applications in speech recognition,” *Proc. IEEE*, vol. 77, no. 2, pp. 257–286, Feb. 1989.
- [10] S. Matl et al., “Vascular image registration techniques: A living review,” *Med. Image Anal.*, vol. 35, pp. 1–17, 2017.
- [11] J. H. Hipwell et al., “Intensity-based 2D/3D registration of cerebral angiograms,” *IEEE Trans. Med. Imag.*, vol. 22, no. 11, pp. 1417–1426, Nov. 2003.
- [12] J. Wang et al., “Dynamic 2D/3D rigid registration framework using point-to-plane correspondence model,” *IEEE Trans. Med. Imag.*, vol. 36, no. 9, pp. 1939–1954, Sep. 2017.
- [13] J. Fan et al., “BIRNet: Brain image registration using dual-supervised fully convolutional networks,” *Med. Image Anal.*, vol. 54, pp. 193–206, 2019.
- [14] R. Frysch et al., “A novel approach to 2D/3D registration of X-ray images using Grangeat’s relation,” *Med. Image Anal.*, vol. 67, 2021, Art. no. 101815.
- [15] N. Baka et al., “Statistical coronary motion models for 2D + t/3D registration of X-ray coronary angiography and CTA,” *Med. Image Anal.*, vol. 17, no. 6, pp. 698–709, 2013.
- [16] T. Benseghir et al., “Iterative closest curve: A framework for curvilinear structure registration application to 2D/3D coronary arteries registration,” in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2013, pp. 179–186.
- [17] J. Zhu et al., “Heuristic tree searching for pose-independent 3D/2D rigid registration of vessel structures,” *Phys. Med. Bio.*, vol. 65, no. 5, 2020, Art. no. 055010.
- [18] J. Zhu et al., “Iterative closest graph matching for non-rigid 3D/2D coronary arteries registration,” *Comput. Meth. Programs Biomed.*, vol. 199, 2021, Art. no. 105901.
- [19] A. Varnavas et al., “Fully automated 2D-3D registration and verification,” *Med. Image Anal.*, vol. 26, no. 1, pp. 108–119, 2015.
- [20] S. Miao, Z. J. Wang, and R. Liao, “A CNN regression approach for real-time 2D/3D registration,” *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1352–1363, May 2016.
- [21] R. Schaffert et al., “Learning an attention model for robust 2D/3D registration using point-to-plane correspondences,” *IEEE Trans. Med. Imag.*, vol. 39, no. 10, pp. 3159–3174, Oct. 2020.
- [22] H. Geng et al., “CT2X-IRA: CT to X-ray image registration agent using domain-cross multi-scale-stride deep reinforcement learning,” *Phys. Med. Biol.*, vol. 68, no. 17, 2023, Art. no. 175024.
- [23] S. Yang et al., “Augmented reality navigation system for biliary interventional procedures with dynamic respiratory motion correction,” *IEEE Trans. Biomed. Eng.*, vol. 71, no. 2, pp. 700–711, Feb. 2024.
- [24] D. Rivest-Henault, H. Sundar, and M. Cheriet, “Nonrigid 2D/3D registration of coronary artery models with live fluoroscopy for guidance of cardiac interventions,” *IEEE Trans. Med. Imag.*, vol. 31, no. 8, pp. 1557–1572, Aug. 2012.
- [25] N. Baka et al., “Oriented Gaussian mixture models for nonrigid 2D/3D coronary artery registration,” *IEEE Trans. Med. Imag.*, vol. 33, no. 5, pp. 1023–1034, May 2014.
- [26] S. Yoon et al., “Topological recovery for non-rigid 2D/3D registration of coronary artery models,” *Comput. Methods Programs Biomed.*, vol. 200, 2021, Art. no. 105922.
- [27] J. Zhu et al., “3D/2D vessel registration based on Monte Carlo tree search and Manifold regularization,” *IEEE Trans. Med. Imag.*, vol. 43, no. 5, pp. 1727–1739, May 2024.
- [28] A. P. King et al., “A subject-specific technique for respiratory motion correction in image-guided cardiac catheterisation procedures,” *Med. Image Anal.*, vol. 13, no. 3, pp. 419–431, 2009.
- [29] C. T. Metz et al., “Registration of 3D+t coronary CTA and monoplane 2D+t X-ray angiography,” *IEEE Trans. Med. Imag.*, vol. 32, no. 5, pp. 919–931, May 2013.
- [30] P. Ambrosini et al., “A hidden Markov model for 3D catheter tip tracking with 2D X-ray catheterization sequence and 3D rotational angiography,” *IEEE Trans. Med. Imag.*, vol. 36, no. 3, pp. 757–768, Mar. 2017.
- [31] H. Ma et al., “Dynamic coronary roadmapping via catheter tip tracking in X-ray fluoroscopy with deep learning based bayesian filtering,” *Med. Image Anal.*, vol. 61, 2020, Art. no. 101634.
- [32] I. Vernikouskaya et al., “Deep learning-based framework for motion-compensated image fusion in catheterization procedures,” *Comput. Med. Imag. Graph.*, vol. 98, 2022, Art. no. 102069.
- [33] M. G. Wagner et al., “Real-time respiratory motion compensated roadmaps for hepatic arterial interventions,” *Med. Pyhs.*, vol. 48, no. 10, pp. 5661–5673, 2021.
- [34] R. Rockenfeller and A. Müller, “Augmenting the cobb angle: Three-dimensional analysis of whole spine shapes using Bézier curves,” *Comput. Methods Programs Biomed.*, vol. 225, 2022, Art. no. 107075.

- [35] B. Aubert et al., "X-ray to DRR images translation for efficient multiple objects similarity measures in deformable model 3D/2D registration," *IEEE Trans. Med. Imag.*, vol. 42, no. 4, pp. 897–909, Apr. 2023.
- [36] Y. Zhang et al., "Unsupervised X-ray image segmentation with task driven generative adversarial networks," *Med. Image Anal.*, vol. 62, 2020, Art. no. 101664.
- [37] Y. Zhao et al., "Unpaired image-to-image translation using adversarial consistency loss," in *Proc. Eur. Conf. Comput. Vis.*, 2020, pp. 800–815.
- [38] J. Zhu et al., "Unpaired image-to-image translation using cycle-consistent adversarial networks," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2017, pp. 2223–2232.
- [39] O. Ronneberger et al., "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Intervention*, 2015, pp. 234–241.
- [40] S. Jégou et al., "The one hundred layers tiramisu: Fully convolutional densenets for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2017, pp. 11–19.
- [41] S. Guo et al., "Improved U-net for guidewire tip segmentation in X-ray fluoroscopy images," in *Proc. Int. Conf. Advan. Image Process.*, 2019, pp. 55–59.
- [42] C. H. Sudre et al., "Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations," in *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Berlin, Germany: Springer, 2017, pp. 240–248.
- [43] F. P. Vidal and P. F. Villard, "Development and validation of real-time simulation of X-ray imaging with respiratory motion," *Comput. Med. Imag. Grap.*, vol. 49, pp. 1–15, 2016.
- [44] S. Belongie, J. Malik, and J. Puzicha, "Shape matching and object recognition using shape contexts," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 4, pp. 509–522, Apr. 2002.
- [45] C. H. Papadimitriou and K. Steiglitz, *Combinatorial Optimization: Algorithms and Complexity*. North Chelmsford, MA, USA: Courier Corporation, 1998.
- [46] J. Ma et al., "Nonrigid point set registration with robust transformation learning under manifold regularization," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 30, no. 12, pp. 3584–3597, Dec. 2019.
- [47] J. Wang et al., "Gaussian field estimator with manifold regularization for retinal image registration," *Signal Process.*, vol. 157, pp. 225–235, 2019.
- [48] M. Belkin et al., "Manifold regularization: A geometric framework for learning from labeled and unlabeled examples," *J. Mach. Learn. Res.*, vol. 7, no. 85, pp. 2399–2434, 2006.
- [49] J. Ma et al., "Robust point matching via vector field consensus," *IEEE Trans. Image Process.*, vol. 23, no. 4, pp. 1706–1721, Apr. 2014.
- [50] R. A. Waller and D. B. Duncan, "A Bayes rule for the symmetric multiple comparisons problem," *J. Amer. Statist. Assoc.*, vol. 64, no. 328, pp. 1484–1503, 1969.
- [51] T. Fu et al., "Local incompressible registration for liver ablation surgery assessment," *Med. Phys.*, vol. 44, no. 11, pp. 5873–5888, 2017.
- [52] K. A. Saddi et al., "Large deformation registration of contrast-enhanced images with volume-preserving constraint," *Proc. SPIE*, vol. 6512, pp. 33–42, 2007.
- [53] C. Buerger et al., "Nonrigid motion modeling of the liver from 3D undersampled self-gated golden-radial phase encoded MRI," *IEEE Trans. Med. Imag.*, vol. 31, no. 3, pp. 805–815, Mar. 2012.
- [54] P. Ambrosini et al., "Continuous roadmapping in liver TACE procedures using 2D-3D catheter-based registration," *Int. J. Comput. Assist. Radiol. Surg.*, vol. 10, no. 9, pp. 1357–1370, 2015.
- [55] J. C. Ross et al., "Motion correction for augmented fluoroscopy-application to liver embolization," in *Proc. IEEE Int. Symp. Biomed. Imag.*, 2008, pp. 1553–1556.
- [56] A. Vandini et al., "Robust guidewire tracking under large deformations combining segment-like features (SEGlets)," *Med. Image Anal.*, vol. 38, pp. 150–164, 2017.
- [57] P. Fischer et al., "An MR-based model for cardio-respiratory motion compensation of overlays in X-ray fluoroscopy," *IEEE Trans. Med. Imag.*, vol. 37, no. 1, pp. 47–60, Jan. 2018.
- [58] T. Fu et al., "Divergence-free fitting-based incompressible deformation quantification of liver," *IEEE J. Biomed. Health. Inf.*, vol. 25, no. 3, pp. 720–736, Mar. 2021.
- [59] M. Nakao, M. Nakamura, and T. Matsuda, "Image-to-graph convolutional network for 2D/3D deformable model registration of low-contrast organs," *IEEE Trans. Med. Imag.*, vol. 41, no. 12, pp. 3747–3761, Dec. 2022.
- [60] M. Nakao, M. Nakamura, and T. Matsuda, "Statistical deformation reconstruction using multi-organ shape features for pancreatic cancer localization," *Med. Image Anal.*, vol. 67, 2021, Art. no. 101829.