PRE-SAM: LoRA-Adapted SAM with Prompt-Routed Experts for Guidewire Segmentation from DSA Images

Anonymized Authors

Anonymized Affiliations
email@anonymized.com

Abstract. Accurate guidewire segmentation is essential for safe and precise navigation in fluoroscopy-guided interventions, yet it remains challenging in Digital Subtraction Angiography (DSA) images due to extreme thinness, severe class imbalance, motion blur, and confusing linear artifacts from bones and contrast-filled vessels. We present PRE-SAM, a domain-adaptive extension of the Segment Anything Model (SAM) tailored for guidewire segmentation. PRE-SAM injects LoRA into the SAM image encoder for parameter-efficient specialization, and augments the mask decoder with class-specific experts for guidewire and vessels. A prompt-routed router conditioned on a learnable implicit prompt token performs conservative residual fusion, injecting expert logits into the global multi-class prediction to better capture thin structures and connectivity while preserving the baseline pathway. To further improve anatomical plausibility, we propose a lightweight anatomy-constrained refinement that leverages the predicted vessel mask to form a tolerance band and repairs small guidewire discontinuities via morphology-aware filtering and endpoint bridging. We curate a clinically collected DSA dataset of over 3,000 images with pixel-level annotations for both guidewires and vessels. Extensive experiments demonstrate that PRE-SAM consistently outperforms strong baselines, achieving higher Dice and substantially lower HD95, while maintaining stable optimization through progressive training.

Keywords: DSA Dataset · SAM · Guidewire Segmentation · Expert Guidance · LoRA · Anatomy Constraints

1 Introduction

Accurate guidewire segmentation plays a critical role in interventional procedures, where guidewires serve as the structural backbone for device navigation and support. Reliable delineation of guidewires is essential for procedural safety, precision control, and robotic assistance [11]. Automated guidewire segmentation can reduce operator dependency and provide real-time visual feedback, facilitating the development of semi-autonomous and fully autonomous interventional systems.

Despite recent advances in deep learning for medical image analysis, guidewire segmentation in fluoroscopic sequences remains highly challenging. Guidewires are extremely thin and occupy only a small fraction of image pixels, leading to severe class imbalance. Motion blur, low signal-to-noise ratios, and interference from visually similar linear structures such as bones or contrast-filled vessels further complicate discrimination. In addition, the limited scale and diversity of publicly available datasets restrict the robustness and generalization of existing models in complex anatomical scenarios.

Early approaches rely on geometric modeling and tracking techniques [20, 2, 17], which often lack robustness under low-contrast fluoroscopy. With the rise of deep learning, U-Net-based architectures [15] have been widely adopted for curvilinear structure segmentation. For example, Ambrosini et al. [1] propose an automatic catheter segmentation framework, and Vlontzos et al. [18] introduce a multi-class twin U-Net for joint vessel and instrument segmentation. FGA-Net [22] incorporates attention mechanisms to alleviate class imbalance and better model irregular guidewire morphology. Zhang et al. [23] propose a Transformer-based guidewire detection framework. However, purely CNN-based models often lack global contextual modeling, limiting their ability to distinguish guidewires from visually similar structures in complex DSA sequences.

Recently, foundation models have gained increasing attention in medical imaging. The Segment Anything Model (SAM) [9] shifts the paradigm toward promptable segmentation, but its pretraining on natural images introduces a substantial domain gap when applied to medical data [4, 14]. MedSAM [13] performs large-scale fine-tuning for medical adaptation, while Medical SAM Adapter [21], SAMed [24] and H-SAM [3] explore parameter-efficient strategies. Nevertheless, the uniform decoding design of SAM limits its capability in thin-structure and motion-blurred tasks such as guidewire segmentation, where class-specific representation and structural constraints are crucial.

Mixture-of-Experts (MoE) architectures dynamically route features to specialized sub-networks [6, 16, 10] and have demonstrated remarkable success in large-scale language models [7, 12]. However, their application in medical image segmentation, particularly for disentangling anatomical (vessel) and instrumental (guidewire) structures, remains underexplored [19]. Achieving structural specialization while maintaining parameter efficiency and training stability remains an open challenge.

To address these limitations, we propose PRE-SAM, a domain-adaptive extension of SAM tailored for guidewire segmentation. We employ Low-Rank Adaptation (LoRA) [5] for parameter-efficient domain specialization, and introduce a prompt-routed expert decoder that disentangles guidewire and vessel representations through residual gating. Unlike prior approaches that treat vessels and instruments independently, we jointly predict vessel masks to provide structural constraints that assist guidewire refinement. An anatomy-constrained post-processing module further enforces spatial and topological consistency by leveraging the vessel prediction to regularize guidewire continuity.

To support this study, we construct a large-scale DSA dataset comprising over 3,000 clinically collected images with dual pixel-level annotations of guidewires and vessels. Extensive experiments demonstrate that PRE-SAM consistently outperforms existing methods across multiple evaluation metrics.

In summary, our contributions are three-fold:

- We construct the largest clinically annotated DSA guidewire segmentation dataset to date, enabling joint modeling of guidewires and vessels.
- We propose a prompt-routed expert-guided extension of SAM with LoRA-based domain adaptation, specifically designed for thin-structure segmentation.
- We introduce an anatomy-constrained refinement strategy that leverages vessel predictions to enhance guidewire continuity and topological plausibility.

2 Method

2.1 Overview

PRE-SAM model extends the SAM foundation model. The architecture of our model is illustrated in Figure 1. After preprocessing, the input image is fed into the SAM image encoder (ViT), where Low-Rank Adaptation (LoRA) is injected into the encoder for parameter efficient finetuning. The mask decoder, after a shared transformer decoding step, outputs three parallel logits: global multi-class logits, guidewire expert logits, and vessel expert logits. The router takes the implicit prompt embedding as input and generates two gates for guidewire and vessel. These gates are then used in a residual fashion to inject the expert logits into the global logits, resulting in the final logits. Finally we apply an anatomical constraint post-processing module that performs topological correction on guidewire prediction results based on vascular morphological prior knowledge. This method significantly improves the accuracy and spatial rationality of guidewire segmentation through two-stage optimization of frequency-domain feature modulation and anatomical structure.

2.2 LoRA in Image Encoder

In this work, we apply the LoRA (Low-Rank Adaptation) technique to fine-tune the image encoder of SAM for guidewire segmentation.LoRA enables efficient customization by updating a minimal subset of parameters, significantly reducing the computational overhead while maintaining performance. We adopt the following approach for fine-tuning the image encoder: we freeze the majority of the transformer blocks in the encoder and introduce low-rank approximations to the projection layers of the query and value transformations. This strategy not only helps in reducing memory and computational costs but also achieves satisfactory segmentation results on medical images. The projected output of the transformer block is updated using two learnable matrices $A \in \mathbb{R}^{r \times C_{\text{in}}}$ and $B \in \mathbb{R}^{C_{\text{out}} \times r}$, where r is much smaller than C_{in} and C_{out}, the input and output

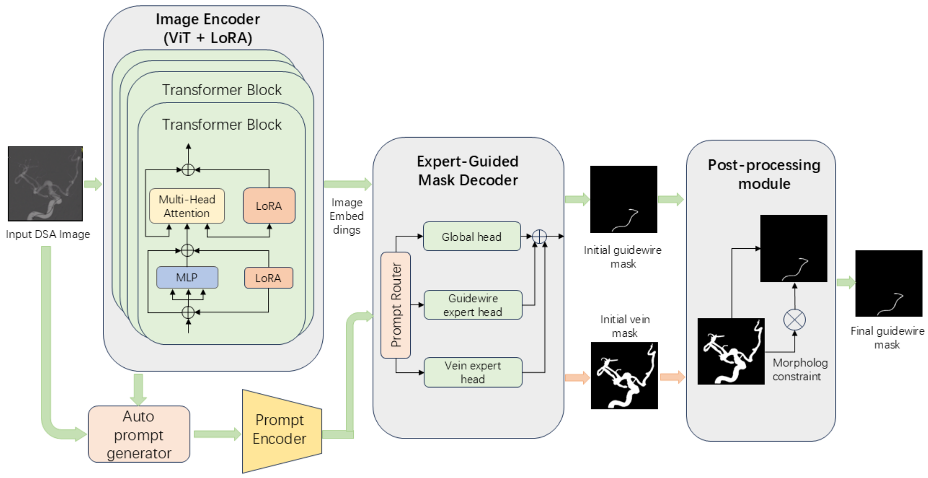

Fig. 1. Overview of the proposed PRE-SAM framework. The model extends SAM with LoRA-based encoder adaptation and introduces class-specific expert heads for guidewire and vessel segmentation. A prompt-conditioned router performs conservative residual fusion, injecting expert logits into the global prediction while preserving the baseline pathway. An anatomy-constrained refinement module further enhances structural continuity and spatial plausibility of the final guidewire mask.

dimensions of the attention mechanism, respectively. LoRA proposes a low-rank approximation to update the projection layer as follows:

$$\hat{F} = WF \tag{1}$$

$$\hat{W} = W + \Delta W = W + BA \tag{2}$$

where $A \in \mathbb{R}^{r \times C_{\text{in}}}$ and $B \in \mathbb{R}^{C_{\text{out}} \times r}$ represent the low-rank matrices, with $r \ll \min(C_{\text{in}}, C_{\text{out}})$. The updated token sequence $\hat{F}$ is computed by the low-rank adaptation of the original projection layer W, allowing efficient fine-tuning with minimal parameter updates.

2.3 Expert Guided Decoder

To handle the segmentation of both thin structures (guidewires) and large connected regions (vessels), we introduce class-specific expert heads into the SAM mask decoder. These expert heads specialize in the segmentation of guidewires and vessel, respectively, enabling the model to focus on the distinct features of each class. The outputs from the expert heads are then fused using a router mechanism that adjusts the contribution of each expert based on the image's content. The router uses the learnable implicit prompt embeddings as input to generate gating values for each expert. These gating values are learned during training and control the weighted combination of expert logits. The overall segmentation output is given by the following fusion mechanism:

$$\mathbf{L}_{\text{final}} = \mathbf{L}_{\text{global}} + \alpha \left(g_{\text{wire}} \cdot \mathbf{L}_{\text{wire}} + g_{\text{vessel}} \cdot \mathbf{L}_{\text{vessel}} \right) \tag{3}$$

where α is a scaling factor initialized to zero and gradually increased during training. For the loss function, we use a combination of cross-entropy loss (CE) for the global output, binary cross-entropy loss (BCE) for the expert outputs, and Dice loss for both the global and expert outputs.

$$\mathcal{L}_{\text{total}} = \mathcal{L}_{\text{global}} + \lambda_{\text{wire}} \cdot \mathcal{L}_{\text{wire}} + \lambda_{\text{vessel}} \cdot \mathcal{L}_{\text{vessel}} \tag{4}$$

2.4 Post-processing Module

To improve anatomical plausibility and structural continuity, we apply a lightweight morphology- and constraint-aware refinement to the predicted label map $\hat{Y} \in \{0,1,2\}^{H\times W}$ (vessels and guidewire). Binary masks are extracted as

$$\hat{M}_c = \mathbf{1}[\hat{Y} = c], \quad c \in \{1,2\}. \tag{5}$$

Isolated noise is removed via connected-component filtering with class-specific thresholds. Vessels undergo mild opening–closing for regional stabilization, while a light closing is applied to the thin guidewire to suppress micro-fragmentation. To enforce anatomical consistency while tolerating boundary ambiguity, a vascular tolerance band is constructed by dilation,

$$\hat{M}_1^{\text{dil}} = \text{Dilate}(\hat{M}_1; r), \tag{6}$$

guidewire predictions are constrained by:

$$\hat{M}_2 \leftarrow \hat{M}_2 \cap \hat{M}_1^{\text{dil}} \tag{7}$$

This suppresses anatomically implausible detections outside the vascular neighborhood. To repair small discontinuities of the elongated guidewire structure, skeleton endpoints are detected and endpoint pairs with $\|p_i - p_j\|_2 < d_{\max}$ are bridged within the tolerance band, followed by mild thickening:

$$\hat{M}_2 \leftarrow \hat{M}_2 \cup \text{Bridge}(\hat{M}_2). \tag{8}$$

The refined masks are finally recomposed into the output label map, yielding topologically coherent and anatomically consistent segmentation.

2.5 Training Strategies

To ensure stable optimization and controlled capacity expansion, we adopt a progressive three-phase training strategy that preserves baseline performance in early stages and gradually activates additional modules.

In the first phase, expert heads and the router are frozen, and only the global branch with the LoRA-adapted backbone is optimized. The router contribution is set to zero, so the model degenerates to the original SAM+LoRA baseline. This guarantees stable initialization and prevents performance degradation caused by randomly initialized expert components.

In the second phase, the expert heads are unfrozen while the router remains fixed. The global branch continues to be optimized jointly. With additional class-specific supervision and exclusivity constraints, the experts learn to specialize in thin guidewire and connected vessel structures without disturbing the global representation.

In the final phase, the router is unfrozen and trained with a smaller learning rate to learn adaptive gating conditioned on the implicit prompt token. Since expert logits are injected in a residual manner, the global pathway is always preserved, ensuring that worst-case performance remains comparable to the baseline. This staged strategy enables a smooth transition from global prediction to dynamic expert-enhanced segmentation.

3 Experiments

3.1 Datasets

The dataset uses in this study comprises more than 3,000 authentic clinical DSA images. These images encompass diverse vascular anatomical structures and catheter insertion angles, sourced from approximately twenty patients undergoing various interventional procedures, including aneurysm repair, internal carotid artery stenosis treatment, and vascular malformation interventions. The dataset images have a resolution of 688×688 pixels. Following a patient-level split, the dataset is divided into a training set of 2,584 images and a test set of 485 images. All image annotations are manually performed at the pixel level using the LabelMe tool. The annotation results undergo cross-verification by two medical professionals to ensure accuracy.

3.2 Results

To comprehensively evaluate the performance of our proposed model, we conduct systematic comparisons against three representative categories of baseline methods. First, we included vascular segmentation models named DARL [8] , which are specifically designed to capture elongated, tree-like anatomical structures. Since guidewires share similar geometric characteristics with vascular branches, these methods serve as strong references for evaluating the capability of modeling thin and topologically complex structures.

Second, we compare against SAM-derived medical adaptation models, including Medical-SAM-adaptor [21], H-SAM [3] and SAMed [24]. These approaches leverage large-scale visual pretraining and contextual representation learning,

Table 1. Quantitative comparison on the test set. The best results are shown in bold.

Method	Dice (%) ↑	HD95 (px) ↓
DARL (ICLR 2023)	76.8	23.25
Medical-SAM-adaptor (MEDIA 2025)	X.X	X.X
H-SAM (CVPR 2024)	83.1	6.89
SAMed (preprint)	77.4	18.25
FGA-Net (QIMS 2025)	78.7	17.09
PRE-SAM (Ours)	**83.9**	**2.98**

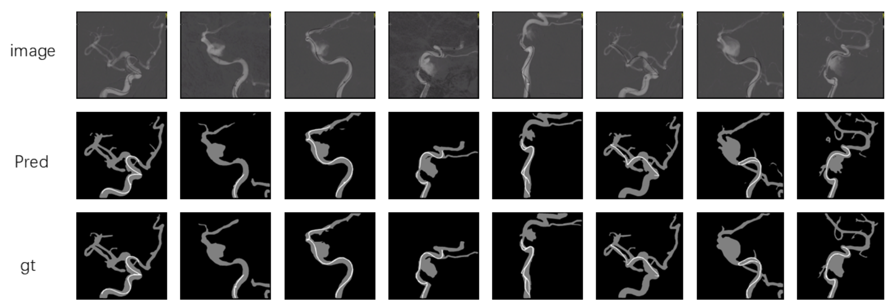

Fig. 2. Training results.

or incorporate domain-specific fine-tuning for medical imaging applications, enabling them to capture spatial relationships among guidewires, catheters, and blood vessels.

Finally, we include a dedicated guidewire segmentation model, FGA-Net [22], which is specifically designed for guidewire extraction tasks, providing a task-oriented benchmark for performance comparison. As shown in Table 1, PRE-SAM achieves state-of-the-art performance on the test set. In particular, it obtains the highest Dice score and the lowest HD95 value among all compared methods, demonstrating superior capability in preserving structural continuity and reducing boundary errors. Notably, for the highly slender and low-occupancy guidewire structures, PRE-SAM significantly improves segmentation completeness while maintaining precise boundary localization. These results validate the effectiveness of the proposed expert specialization mechanism and conservative router-based fusion strategy in challenging medical imaging scenarios.

Our training results are presented in Figure 2.

3.3 Ablation Study

As shown in Table 2, we conduct a progressive ablation study to quantify the contribution of each component to guidewire and vessel segmentation, while ex-

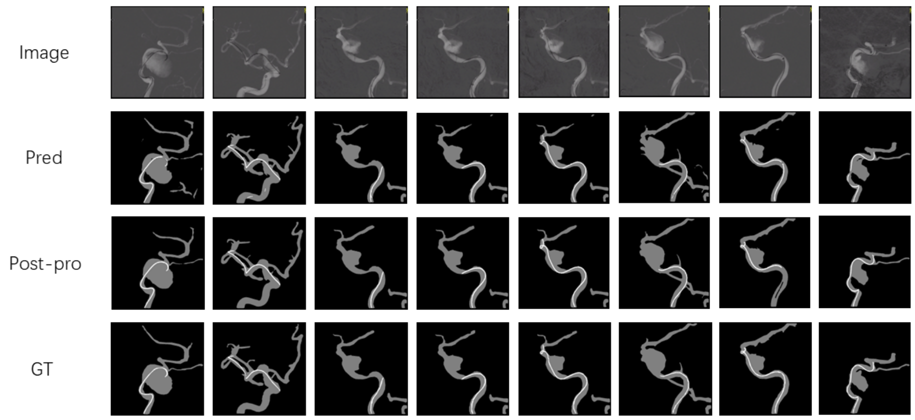

Fig. 3. Visualization of post-processing module.Compared with the raw predictions, the refinement step suppresses anatomically implausible detections outside vessel regions, repairs small discontinuities in elongated guidewire structures, and improves boundary coherence.

plicitly verifying training stability. Our baseline, **SAM+LoRA**, augments the SAM backbone with a learnable implicit prompt embedding to enable prompt-free inference, and applies LoRA-based adaptation to the image encoder for angiography-domain specialization. We then introduce **Experts (no router)** by attaching two lightweight class-specific expert heads (wire/vessel) on top of the shared mask-decoder trunk, trained with additional binary supervision and an exclusivity regularizer to encourage class specialization, while keeping the global branch output unchanged. Next, **Experts+Router** activates a prompt-token router that maps the implicit prompt token to gating coefficients and injects the expert logits into the global logits via residual fusion, thereby improving thin-structure and connectivity modeling while preserving the baseline global pathway. For **Experts+Router+Post-processing**, instead of further quantitative comparison, we present qualitative visualization results (Fig. 3) to highlight structural refinements. The post-processing module primarily enforces anatomical plausibility and topological consistency, which are not fully captured by region-based metrics. Visual comparisons show that it effectively suppresses anatomically implausible detections, repairs small discontinuities in guidewire predictions, and enhances boundary coherence, especially in challenging low-contrast and overlapping regions.

4 Conclusion

In this work, we proposed PRE-SAM, a domain-adaptive framework for guidewire segmentation in Digital Subtraction Angiography (DSA) images. Built upon the Segment Anything Model (SAM), our approach integrates LoRA-based parameter-

Table 2. Ablation study of the proposed components. All settings share the same backbone, implicit prompt token, and LoRA configuration.

Method	LoRA+Implicit Prompt	Experts	Router	Dice↑	HD95↓
SAM + LoRA	✓			80.2	9.25
+ Experts	✓	✓		82.8	8.32
+ Experts + Router	✓	✓	✓	83.9	2.98

efficient adaptation, class-specific expert heads, and a conservative prompt-routed fusion mechanism to address the challenges of slender structure segmentation and anatomical inconsistency. The progressive training strategy ensures stable optimization while enabling expert specialization.Our approach was validated on a dataset consisting of over 3,000 DSA images from approximately twenty patients, all collected in clinical settings by experienced medical professionals. Experimental results demonstrate that this method significantly outperforms existing guidewire segmentation models across multiple evaluation metrics.

References

1. Ambrosini, P., Ruijters, D., Niessen, W.J., Moelker, A., van Walsum, T.: Fully automatic and real-time catheter segmentation in x-ray fluoroscopy. In: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (eds.) Medical Image Computing and Computer-Assisted Intervention - MICCAI 2017. pp. 577–585. Springer International Publishing, Cham (2017)
2. Cheng, E., Pang, Y., Zhu, Y., Yu, J., Ling, H.: Curvilinear structure tracking by low rank tensor approximation with model propagation. 2014 IEEE Conference on Computer Vision and Pattern Recognition pp. 3057–3064 (2014), https://api.semanticscholar.org/CorpusID:1051533
3. Cheng, Z., Wei, Q., Zhu, H., Wang, Y., Qu, L., Shao, W., Zhou, Y.: Unleashing the potential of sam for medical adaptation via hierarchical decoding (2024), https://arxiv.org/abs/2403.18271
4. Hu, C., Xia, T., Ju, S., Li, X.: When sam meets medical images: An investigation of segment anything model (sam) on multi-phase liver tumor segmentation (2023), https://arxiv.org/abs/2304.08506
5. Hu, E.J., Shen, Y., Wallis, P., Allen-Zhu, Z., Li, Y., Wang, S., Wang, L., Chen, W.: Lora: Low-rank adaptation of large language models (2021), https://arxiv.org/abs/2106.09685
6. Jacobs, R.A., Jordan, M.I., Nowlan, S.J., Hinton, G.E.: Adaptive mixtures of local experts. Neural Computation **3**(1), 79–87 (1991). https://doi.org/10.1162/neco.1991.3.1.79
7. Jiang, A.Q., Sablayrolles, A., Roux, A., Mensch, A., Savary, B., Bamford, C., Chaplot, D.S., de las Casas, D., Hanna, E.B., Bressand, F., Lengyel, G., Bour, G., Lample, G., Lavaud, L.R., Saulnier, L., Lachaux, M.A., Stock, P., Subramanian, S., Yang, S., Antoniak, S., Scao, T.L., Gervet, T., Lavril, T., Wang, T., Lacroix, T., Sayed, W.E.: Mixtral of experts (2024), https://arxiv.org/abs/2401.04088

8. Kim, B., Oh, Y., Ye, J.C.: Diffusion adversarial representation learning for self-supervised vessel segmentation (2023), https://arxiv.org/abs/2209.14566
9. Kirillov, A., Mintun, E., Ravi, N., Mao, H., Rolland, C., Gustafson, L., Xiao, T., Whitehead, S., Berg, A.C., Lo, W.Y., Dollár, P., Girshick, R.: Segment anything. In: 2023 IEEE/CVF International Conference on Computer Vision (ICCV). pp. 3992–4003 (2023). https://doi.org/10.1109/ICCV51070.2023.00371
10. Lepikhin, D., Lee, H., Xu, Y., Chen, D., Firat, O., Huang, Y., Krikun, M., Shazeer, N., Chen, Z.: Gshard: Scaling giant models with conditional computation and automatic sharding (2020), https://arxiv.org/abs/2006.16668
11. Li, W., Que, D., Omisore, O.M.: A gaussian-based guidewire segmentation and tracking method for teleoperated robotic intravascular interventions. In: 2021 5th International Conference on Robotics and Automation Sciences (ICRAS). pp. 38–43 (2021). https://doi.org/10.1109/ICRAS52289.2021.9476273
12. Luo, T., Lei, J., Lei, F., Liu, W., He, S., Zhao, J., Liu, K.: Moelora: Contrastive learning guided mixture of experts on parameter-efficient fine-tuning for large language models (2024), https://arxiv.org/abs/2402.12851
13. Ma, J., He, Y., Li, F., Han, L., You, C., Wang, B.: Segment anything in medical images. Nature Communications **15**(1) (Jan 2024). https://doi.org/10.1038/s41467-024-44824-z, http://dx.doi.org/10.1038/s41467-024-44824-z
14. Paranjape, J.N., Nair, N.G., Sikder, S., Vedula, S.S., Patel, V.M.: Adaptivesam: Towards efficient tuning of sam for surgical scene segmentation (2023), https://arxiv.org/abs/2308.03726
15. Ronneberger, O., Fischer, P., Brox, T.: U-net: Convolutional networks for biomedical image segmentation. CoRR **abs/1505.04597** (2015), http://arxiv.org/abs/1505.04597
16. Shazeer, N., Mirhoseini, A., Maziarz, K., Davis, A., Le, Q., Hinton, G., Dean, J.: Outrageously large neural networks: The sparsely-gated mixture-of-experts layer (2017), https://arxiv.org/abs/1701.06538
17. Vandini, A., Glocker, B., Hamady, M., Yang, G.Z.: Robust guidewire tracking under large deformations combining segment-like features (seglets). Medical Image Analysis **38**, 150–164 (2017). https://doi.org/https://doi.org/10.1016/j.media.2017.02.001, https://www.sciencedirect.com/science/article/pii/S1361841517300208
18. Vlontzos, A., Mikolajczyk, K.: Deep segmentation and registration in x-ray angiography video. CoRR **abs/1805.06406** (2018), http://arxiv.org/abs/1805.06406
19. Wang, G., Ye, J., Cheng, J., Li, T., Chen, Z., Cai, J., He, J., Zhuang, B.: Sam-med3d-moe: Towards a non-forgetting segment anything model via mixture of experts for 3d medical image segmentation (2024), https://arxiv.org/abs/2407.04938
20. Wang, P., Chen, T., Zhu, Y., Zhang, W., Zhou, S.K., Comaniciu, D.: Robust guidewire tracking in fluoroscopy. In: 2009 IEEE Conference on Computer Vision and Pattern Recognition. pp. 691–698 (2009). https://doi.org/10.1109/CVPR.2009.5206692
21. Wu, J., Ji, W., Liu, Y., Fu, H., Xu, M., Xu, Y., Jin, Y.: Medical sam adapter: Adapting segment anything model for medical image segmentation (2023), https://arxiv.org/abs/2304.12620
22. Zhang, G., Liu, D., Ji, Z., Xie, B., Chen, Q.: Lightweight wavelet convolutional network for guidewire segmentation. Computerized Medical Imaging and Graphics **124**, 102618 (2025). https://doi.org/https://doi.org/10.1016/j.compmedimag.2025.102618, https://www.sciencedirect.com/science/article/pii/S0895611125001272

23. Zhang, G., Wong, H.C., Wang, C., Zhu, J., Lu, L., Teng, G.: A temporary transformer network for guide- wire segmentation. In: 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI). pp. 1–5 (2021). https://doi.org/10.1109/CISP-BMEI53629.2021.9624350
24. Zhang, K., Liu, D.: Customized segment anything model for medical image segmentation (2023), https://arxiv.org/abs/2304.13785