Endovascular thrombectomy remains the gold standard for treating acute ischemic stroke, relying heavily on the navigation of guidewires and microcatheters through the complex vascular tortuosity of the brain. Currently, interventionalist surgeons rely on pre-operative Computed Tomography Angiography (CTA) scans for roadmap planning. However, a critical disconnect exists between these static pre-operative maps and the intra-operative reality. The insertion of stiff endovascular tools induces significant mechanical deformation in patient-specific vasculature, rendering rigid pre-operative CTA models inaccurate for live navigation and safety monitoring. To compensate, surgeons must mentally reconstruct the 3D anatomy from 2D planar X-ray fluoroscopic imaging, a process that is cognitively demanding during the high stakes of surgery, leading to situations where the surgeon may be prone to error.

While recent advancements in zero-shot deep reinforcement learning have shown promise in automating guidewire navigation across unseen vascular anatomies (Scarponi et al., 2024), these control systems still rely on accurate real-time state estimation. Consequently, the accurate yet real-time estimation of 3D vessel deformation from 2D fluoroscopy remains an open challenge.

Traditionally, aligning 3D pre-operative volumes with 2D intra-operative images has relied on rigid intensity-based or feature-based registration. While effective for static anatomy, early iterative reconstruction algorithms, such as those based on compressed sensing (Li & Luo, 2011), struggled to account for the non-rigid, elastic nature of the vascular tree during surgical manipulation without incurring high computational costs.

To address the dynamic nature of endovascular procedures, researchers have increasingly turned to deep learning methodologies. Recent efforts have utilized fully convolutional networks to estimate deformation. Lecomte et al. (2023) utilized synthetic digitally reconstructed radiographs to train networks capable of predicting vessel deformation from contrast-free 2D fluoroscopy. While this approach achieved a reduction in reprojection error and maintained real-time performance, it suffered from limited out-of-plane accuracy. This limitation is critical since a model that minimizes 2D reprojection error may still result in a 3D shape that is physically impossible.

Parallel to deformation estimation, significant progress has been made in sparse-view 3D reconstruction. The AutoCAR framework (Zhu et al., 2025) demonstrates that combining pre-operative priors with learned acquisition models can recover dynamic 3D vasculature from limited angiographic views. Similarly, NeRF-CA (Maas et al., 2025) applied Neural Radiance Fields to reconstruct coronary anatomies from extremely sparse views by decoupling static backgrounds from dynamic vessels.

More recently, diffusion models have set new benchmarks in this domain. DiffusionBlend (Song et al., 2024) introduced a position-aware score blending technique that enforces 3D consistency in CT reconstruction without hand-crafted regularization, significantly outperforming traditional methods. Simultaneously, the field is witnessing the emergence of Large Multimodal Models (LMMs) Menon (2025) demonstrated a generative AI pipeline using Gemini 2.5 Flash to perform interactive, prompt-driven 2D-to-3D vascular reconstruction. However, while these generative and diffusion-based methods produce plausible geometries, they remain computationally expensive and lack the explicit biomechanical constraints necessary to guarantee safety during live instrument-anatomy interaction.

A recurring failure mode in standard computer vision approaches is the misinterpretation, causing anatomically invalid shapes. Standard CNNs often predict vessel topologies that violate the elastic limits of biological tissues. To mitigate this, recent works have shifted toward incorporating biomechanical constraints directly into the learning process. Yang et al. (2025) proposed the Deformable Instrument Tip Tracking framework, using manifold regularization to correct deformation, though it relies on explicit segmentation.

Two distinct physics-based approaches have recently emerged to address these limitations. Harper et al. (2025) utilized Physics-Informed Neural Networks for augmented reality based surgical tracking, embedding finite element method elasticity constraints directly into the loss function. On the other hand, Amiri-Hezaveh et al. (2025) proposed a deformable registration method based on neural ordinary differential equations. By formulating registration as a minimum potential energy problem solved via a predictor-corrector scheme, they achieved diffeomorphic, physically grounded mappings. However, both FEM-based PINNs and Neural ODEs often struggle to balance the computational density required for physics solving with the 15–30 fps inference speed required for live fluoroscopy.

The review reveals a distinct gap. Pure CV methods (Lecomte et al.) are fast but prone to depth inaccuracies. Advanced reconstruction methods (AutoCAR, NeRF-CA, DiffusionBlend) and Generative AI (Menon) offer geometric fidelity but incur prohibitive computational costs or lack biomechanical guarantees. Conversely, rigorous physics-based models (Harper et al., Amiri-Hezaveh et al.) ensure physical plausibility but often face latency bottlenecks. There is currently no method that successfully combines the inference speed of deep learning with the geometric integrity of biomechanics to estimate guidewire-induced deformation from monocular fluoroscopy in real-time.

Our study proposes to fill this gap by introducing a Physics-Informed Neural Network (PINN) specifically designed to estimate guidewire-induced vascular deformation. Unlike standard registration methods that predict dense pixel displacements or solve complex ODEs, this approach constrains the search space by predicting sparse control points that drive a Radial Basis Function deformation field. By anchoring the network around a physically grounded deformation equation we can minimize both 2D projection error and strain energy. This method ensures that inferred deformations are physically plausible and topologically consistent. This contribution extends the work of Lecomte et al. by improving out-of-plane accuracy through physics constraints baked into the loss function that penalizes unrealistic curvature and null space solutions, while projection ambiguity is resolved thorough minimum-energy solutions. This offers a more computationally efficient alternative to the heavy reconstruction pipelines of Song et al. and Amiri-Hezaveh et al.

References

Scarponi, V., Duprez, M., Nageotte, F., & Cotin, S. (2024). A zero-shot reinforcement learning strategy for autonomous guidewire navigation. *International Journal of Computer Assisted Radiology and Surgery*, *19*(6), 1185-1192.

Li, X., & Luo, S. (2011). A compressed sensing-based iterative algorithm for CT reconstruction and its possible application to phase contrast imaging. *Biomedical engineering online*, *10*(1), 73.

Lecomte, F., Scarponi, V., Alvarez, P.A., Verde, J.M., Dillenseger, J.-L., Vibert, E., Cotin, S.: Enhancing Fluoroscopy-Guided Interventions: a Neural Network to Predict Vessel Deformation without Contrast Agents. *Hamlyn Symposium on Medical Robotics (2023)*, pp. 75–76.

Zhu, Y., Wang, Y., Di, C., Liu, H., Liao, F., & Ma, S. (2025). Sparse and transferable three-dimensional dynamic vascular reconstruction for instantaneous diagnosis. *Nature Machine Intelligence*, 1-13.

Maas, K. W., Ruijters, D., Vilanova, A., & Pezzotti, N. (2025). NeRF-CA: Dynamic Reconstruction of X-ray Coronary Angiography with Extremely Sparse-views. *IEEE Transactions on Visualization and Computer Graphics*.

Song, B., Hu, J., Luo, Z., Fessler, J., & Shen, L. (2024). Diffusionblend: Learning 3d image prior through position-aware diffusion score blending for 3d computed tomography reconstruction. *Advances in Neural Information Processing Systems*, *37*, 89584-89611.

Menon, P. G. (2025). Generative AI Pipeline for Interactive Prompt-driven 2D-to-3D Vascular Reconstruction for Fontan Geometries from Contrast-Enhanced X-Ray Fluoroscopy Imaging. *arXiv preprint arXiv:2509.13372*.

Yang, S., Xiao, D., Geng, H., Ai, D., Fan, J., Fu, T., ... & Yang, J. (2024). Real-Time 3D Instrument Tip Tracking Using 2D X-Ray Fluoroscopy With Vessel Deformation Correction Under Free Breathing. *IEEE Transactions on Biomedical Engineering*.

Harper, D. M., Chen, L. J., McKay, R. T., Nguyen, S. L., & Fontaine, M. A. (2025). Physics-Informed Neural Networks for Real-Time Deformation-Aware AR Surgical Tracking. *bioRxiv*, 2025-09.

Amiri-Hezaveh, A., Tan, S., Deng, Q., Umulis, D., Cunniff, L., Weickenmeier, J., & Buganza Tepole, A. (2025). A Physics-Informed Deep Learning Deformable Medical Image Registration Method Based on Neural ODEs: A. Amiri-Hezaveh et al. *International Journal of Computer Vision*, 1-26.