

# Supplementary Material of PANS

## 1 Literary Comparison

Vision-based bronchoscope navigation techniques are limited by real-time performance, adaptability, and biased dataset validation, as detailed in Table S1.

**Table S1:** Comparison of recent bronchoscope localization methods

Method	Year	Real-time Speed	Case-Specific Training	Data Source (Volume) for Testing	Max Gen.	Trajectory Type	Supplementary Equipment
Sganga et al. [1]	2019	Yes	Yes	Phantom (13 seq.); Cadaver (8 cases)	8	Insertion to targets	No
Shen et al. [2]	2019	No	No	In-vivo (2 cases)	-	-	No
Zhao et al. [3]	2020	Yes	Yes	Phantom (5 seq.)	-	-	No
Banach et al. [4]	2021	No	Yes	Phantoms (5 seq.); Porcine (2 cases)	3	Insertion to targets	No
Gu et al. [5]	2022	No	No	Phantom (3 seq.); Porcine (2 seq.)	5	Standard clinical routine	Yes (Kinematics)
Luo et al. [6]	2023	No	No	In-vivo (9 cases)	5	Standard clinical routine	Yes (EM)
DD-VNB [7]	2024	Yes	No	Phantom (5 seq.); In-vivo (3 cases)	3	Insertion to targets	No
Ours	-	Yes	No	In-vivo (10 cases)	5	Standard clinical routine	No

\* seq. represents video sequence, and Gen. is generation of airway branches.

References: [1] Sganga, J., et al. : Autonomous driving in the lung using deep learning for localization. arXiv, 2019. [2] Shen, M., et al. : Context-aware depth and pose estimation for bronchoscopic navigation. IEEE RA-L, 2019. [3] Zhao, C., et al. : Generative localization with uncertainty estimation through video-CT data for bronchoscopic biopsy. IEEE RA-L, 2020. [4] Banach, A., et al. : Visually navigated bronchoscopy using three cycle-consistent generative adversarial network for depth estimation. MedIA, 2021. [5] Gu, Y., et al. : Vision-kinematics interaction for robotic-assisted bronchoscopy navigation. IEEE TMI, 2022. [6] Luo, X., et al. : Monocular endoscope 6-dof tracking with constrained evolutionary stochastic filtering. MedIA, 2023. [7] Tian, Q., et al. : DD-VNB: A depth-based dual-loop framework for real-time visually navigated bronchoscopy. arXiv, 2024.

## 2 Details of Bronchoscope Localization Dataset



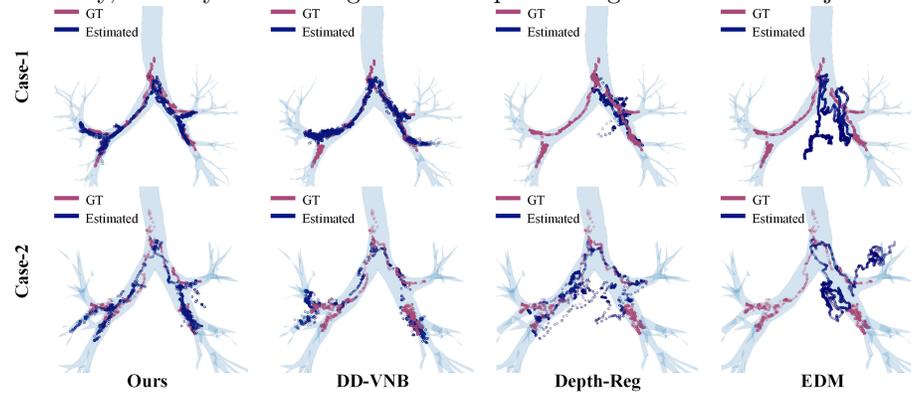
**Fig. S1:** Bronchoscopic frames featuring lack of informative content, motion blur or occlusion.

Bronchoscopy videos are filmed during regular inspection procedure, containing challenging images as presented in Fig. S1, capture various conditions such as

pulmonary nodules, pneumonia, and lung cancer. We calibrate the camera using checkerboard images to correct frame distortions. Localization annotations are made by our OpenGL-based toolkit, aligning virtual camera intrinsics with the actual bronchoscope. Three experts manually labeled the dataset by registering virtual views to real data. To assess annotation accuracy, two cases have been independently labeled by the experts, resulting in a group variance of 0.58 mm.

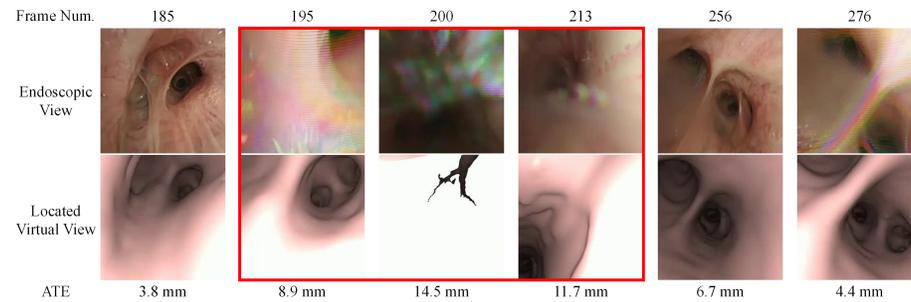
### 3 Qualitative Localization Comparison

In Fig. S2, we report localization trajectories of two cases. Scale of EDM is aligned with ground truth before evaluation. PANS exhibits superior localization accuracy, robustly maintaining bronchoscope tracking over extended trajectories.



**Fig. S2:** Localization results against the ground truth trajectory.

### 4 Localization in Challenging Frames



**Fig. S3:** PANS localization in challenging sequence.

Fig. S3 demonstrates PANS's localization results in challenging frames, with occlusion, motion blur, and limited information (highlighted in red box), where PANS successfully resumes tracking upon the return of informative frames.