

Seeing Through Smoke: Surgical Desmoking for Improved Visual Perception

Jingpei Lu^[0000–0002–9136–6096], Fengyi Jiang^[0009–0007–9420–5259], Xiaorui Zhang^[0009–0005–3639–3875], Lingbo Jin, and Omid Mohareri

Intuitive Surgical, Sunnyvale, CA 94086, USA
jingpei.lu@intusurg.com

Abstract. Minimally invasive and robot-assisted surgery relies heavily on endoscopic imaging, yet surgical smoke produced by electrocautery and vessel-sealing instruments can severely degrade visual perception and hinder vision-based functionalities. We present a transformer-based surgical desmoking model with a physics-inspired desmoking head that jointly predicts smoke-free image and corresponding smoke map. To address the scarcity of paired smoky-to-smoke-free training data, we develop a synthetic data generation pipeline that blends artificial smoke patterns with real endoscopic images, yielding over 80,000 paired samples for supervised training. We further curate, to our knowledge, the largest paired surgical smoke dataset to date, comprising 5,817 image pairs captured with the da Vinci robotic surgical system, enabling benchmarking on high-resolution endoscopic images. Extensive experiments on both a public benchmark and our dataset demonstrate state-of-the-art performance in image reconstruction compared to existing dehazing and desmoking approaches. We also assess the impact of desmoking on downstream stereo depth estimation and instrument segmentation, highlighting both the potential benefits and current limitations of digital smoke removal methods.

Keywords: Robot-assisted Surgery · Endoscopic Imaging · Smoke Removal.

1 Introduction

Minimally invasive procedures, especially robot-assisted surgery, can improve patient outcomes by reducing incision size, infection risk, and recovery time versus open surgery. However, they introduce visibility challenges when high-energy instruments such as electrocautery and vessel sealers are used. Rapid tissue heating vaporizes intracellular water and degrades organic components, producing surgical smoke made of water vapor and protein/lipid byproducts. Because minimally invasive surgery relies heavily on endoscopic imaging, smoke and lens fog can obscure fine tissue detail, prolong operative time, and increase error risk by limiting informed intraoperative decision-making. Smoke artifacts also degrade downstream computer-assisted functions, including instrument segmentation, tool tracking, and 3D reconstruction [5].

Common smoke evacuation methods include instrument-tip suction, port-side evacuation, and valveless trocar systems, each with limitations [29]. For example, valveless insufflation systems such as AirSeal can draw room air into the abdomen during suction, highlighting trade-offs among smoke clearance, pressure control, and gas composition [25]. Alternatively, digital smoke removal is appealing because it can enhance visibility in real time without interrupting the procedure, reducing the need for manual lens cleaning or frequent smoke-evacuation maneuvers. It may also reduce reliance on specialized hardware, lowering device cost and increasing surgeon autonomy. Nevertheless, desmoking is challenging because smoke varies widely in density and motion and induces non-uniform scattering that can resemble anatomical texture.

Early analytical methods use Bayesian probabilistic graphical models [1, 12] or dark channel priors [8, 22] to suppress smoke and enhance contrast. However, these approaches often rely on strong priors developed for natural hazy images, overlooking key differences between surgical smoke and atmospheric haze. More recently, domain-specific, data-driven methods have been proposed for desmoking in minimally invasive surgery. Several approaches apply convolutional neural networks (CNNs) [2, 11, 23, 19] to produce smoke-free images in real time without explicit physical modeling. Further gains in color fidelity and contrast have been achieved using attention-based networks [9, 20, 17], hybrid CNN–transformer architectures [24], and diffusion models with multi-frequency learning [13]. Progress on surgical desmoking remains limited by the scarcity of paired *in vivo* data with smoke-free ground truth. Although unpaired learning can reduce the need for paired ground truth [27, 16], it often introduces structural distortions or color shifts that reduce clinical reliability. A practical alternative is to train on synthetic data while explicitly addressing the sim-to-real domain gap.

In this work, we propose a transformer-based model for surgical desmoking. Our architecture comprises a Vision Transformer (ViT) [6] backbone, a DPT decoder [18], and a physics-inspired desmoking head that predicts a smoke-free image along with a smoke map. To address data scarcity, we generate a large synthetic training set by blending artificial smoke patterns into real endoscopic images. For evaluation, we also collect 5,817 paired smoky-to-smoke-free images acquired with Intuitive Surgical’s da Vinci robotic surgical system. We evaluate our method on our collected dataset and on the *in vivo* benchmark dataset [28], and demonstrate state-of-the-art performance. Beyond image reconstruction, we further use downstream stereo depth estimation and instrument segmentation as functional evaluations of whether restored images preserve task-relevant structures. These experiments show a nuanced result: desmoking improves instrument segmentation in our benchmark but does not consistently improve stereo depth, highlighting both the promise and limitations of digital smoke removal for surgical perception.

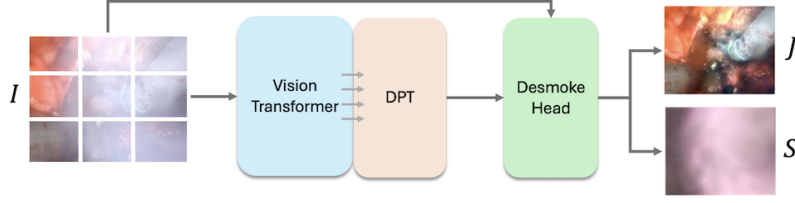


Fig. 1. Overview of the proposed surgical desmoking model.

2 Methodology

In this paper, we introduce an end-to-end model for surgical desmoking (Section 2.1). To train the network, we develop a synthetic smoke data generation pipeline that produces large-scale paired smoky/smoke-free images (Section 2.2). We also collect, to our knowledge, the largest paired surgical smoke dataset to benchmark desmoking performance (Section 2.3). The dataset comprises 5,817 image pairs acquired in an *ex vivo* setup and includes depth maps and segmentation labels to enable extensive downstream evaluation.

2.1 Surgical Desmoking Model

Figure 1 provides an overview of the proposed desmoking model. Given an input smoky image $I \in \mathbb{R}^{H \times W \times 3}$, we use a ViT-Base backbone to extract multi-scale features, initialized with Masked Autoencoder pretrained weights [7] on a large corpus of surgical images. We fuse latent representations from intermediate ViT blocks $\{3, 6, 9, 12\}$ and decode them using DPT blocks. Finally, we attach a lightweight desmoking head to jointly predict the smoke-free image J and the corresponding smoke map S .

To design the desmoking head, we draw inspiration from single-image dehazing methods [8, 21]. We assume a standard atmospheric scattering model [15]:

$$I = Jt + A(1 - t), \quad (1)$$

where I denotes the input smoky image, J the smoke-free image, A the global atmospheric light, and t the medium transmission map. Our objective is to recover J via:

$$J = \frac{I - A + At}{t}. \quad (2)$$

Following [21], we can rewrite Eq. (2) as:

$$J = KI - B + I, \quad (3)$$

where $K = 1/t - 1$ and $B = A(1/t - 1)$. We design the desmoke head to directly predict K and B by regressing a 4D image $O = [K, B] \in \mathbb{R}^{H \times W \times 4}$, such that the network architecture softly constrains the relationship between K and B .

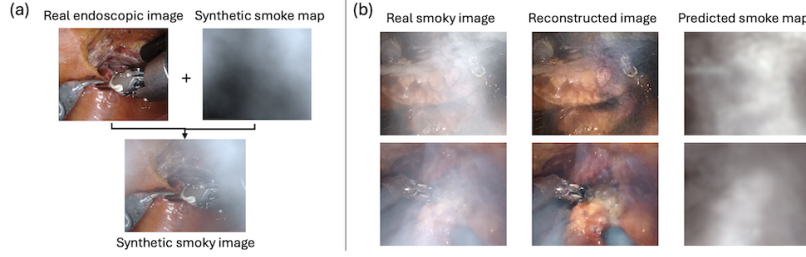


Fig. 2. (a) Synthetic smoke data generation pipeline. (b) Qualitative results for desmoking and smoke-map prediction on real smoky images.

Since I can be viewed as a combination of the attenuated scene radiance Jt and the airlight term $A(1 - t)$ (light scattered by particles), the smoke map is naturally correlated with the airlight component. We therefore predict the smoke map S as a function of $A(1 - t)$:

$$S = f(A(1 - t)) = f\left(\frac{B}{K + 1}\right), \quad (4)$$

where $f(\cdot)$ is approximated as a convolutional layer.

2.2 Training and Synthetic Data Generation

Collecting large-scale surgical smoke data is labor-intensive and costly, and obtaining paired smoky-to-smoke-free images is even more challenging. Mining such paired samples from existing laparoscopic datasets is only feasible at limited scale [4, 27]. Consequently, synthetic smoke generation remains the most efficient approach for producing large-scale paired training data.

We curated a dataset of 10k smoke-free endoscopic images spanning both *in vivo* and *ex vivo* scenes. We then generated a diverse set of smoke maps with varying densities and spatial patterns. Finally, we synthesized smoky images by blending smoke maps with real endoscopic images, as illustrated in Fig. 2(a). We use alpha blending with RGB channel correction and illumination enhancement to improve visual realism. To mitigate the sim-to-real gap, we apply domain randomization over smoke color, transparency, and blending coefficients. In total, we generate approximately 80k paired smoky/clean images at 512×640 resolution to train the proposed surgical desmoking network.

During training, we use L_1 (mean absolute error) losses for both smoke-free image reconstruction and smoke-map prediction:

$$\mathcal{L} = \frac{1}{N} \sum_{i=1}^N (|J_i^{gt} - J_i| + \lambda |S_i^{gt} - S_i|), \quad (5)$$

where J^{gt} and S^{gt} denote the ground-truth smoke-free image and smoke map, respectively, and we set $\lambda = 0.1$. We train the network for 300 epochs using

Table 1. Existing surgical smoke datasets.

Dataset	Smoke type	Size	Image type (resolution)	Availability
LSD3K [4]	Synthetic	3,000 pairs	Mono (480×480)	✗
LSVD [27]	Real	786 clips	Mono (1080×1920)	✗
De-Smoking [28]	Real	3,464 pairs	Mono (inconsistent)	✓
Ours	Real	5,817 pairs	Stereo (1024×1280)	✓

AdamW with a cosine-annealed learning rate from 10^{-5} to 10^{-7} . Because ground-truth smoke maps are difficult to obtain for real smoky scenes, we report only qualitative results for these cases in Fig. 2(b). Although we do not evaluate smoke-map estimation quantitatively, we view it as an auxiliary output with potential future uses, such as confidence-aware smoke evacuation control, or guiding downstream tasks by masking high-smoke regions.

2.3 Benchmark Smoke Data Acquisition

To validate desmoking performance on experimental scenes, we collected data using an Intuitive Surgical da Vinci robotic surgical system. We placed porcine tissue specimens (e.g., liver, intestine, stomach, and kidney) in a sealed chamber and used a smoke machine (LENSGO Portable Fog Machine) to generate smoke. At the beginning of each capture, the endoscope was positioned at a fixed viewpoint toward the specimens, and instrument manipulation was introduced. The smoke machine was then activated to introduce smoke into the chamber. After a short interval, we deactivated the smoke machine and allowed the smoke to distribute throughout the scene. Finally, a suction apparatus was activated to gradually evacuate the smoke.

We employed a bidirectional capture strategy—progressive occlusion followed by progressive clearance—which yields paired observations of identical static scenes across a continuous spectrum of smoke densities. To the best of our knowledge, the resulting dataset is the largest publicly available endoscopic smoke dataset to date (Table 1), addressing a critical gap in the availability of large-scale, high-quality paired data for surgical smoke removal. We further derived depth maps and tool segmentation labels from the smoke-free images to support downstream task evaluation and future investigations in surgical desmoking.

3 Experiments and Results

We selected eight representative methods from the dehazing and desmoking literature for comparison, based on their reported performance and the availability of open-source implementations. We reimplemented the algorithms in [8, 21, 17, 19, 14, 20, 10, 27] following the authors’ released code, and evaluated them on the De-smoking dataset [28] as well as our in-house benchmark dataset. We found that most methods failed to produce reasonable results at 1024×1280 resolution. Therefore, we resized all collected images to 512×640 for the evaluations.

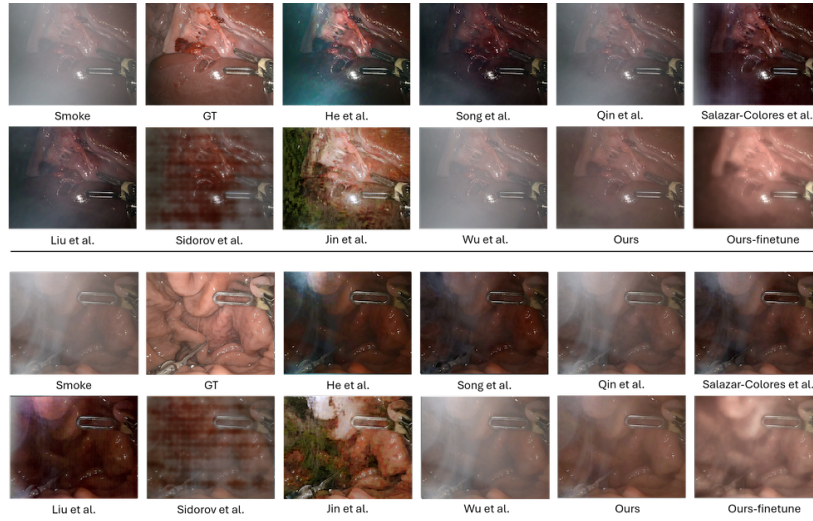


Fig. 3. Qualitative comparisons on our test set across representative scenes (top: liver; bottom: stomach)

in this section. Our model has a size of 111.36 MB and runs at 52 FPS on an NVIDIA RTX 6000 GPU at 512×640 resolution. We split our benchmark data into a fine-tuning set (3,097 image pairs) and a test set (2,720 image pairs) with disjoint anatomical structures. All numbers reported in Tables 3 are computed on the test set. We use the fine-tuning set to further fine-tune our model for an additional 50 epochs; these results are reported as *Ours-finetune*. The model without fine-tuning on real smoke data is reported as *Ours*.

3.1 Surgical Desmoking Benchmark

Table 2 reports the performance of our desmoking model against eight selected dehazing/desmoking baselines on the De-Smoking dataset [28]. The dataset comprises 1,063 paired images from cholecystectomy and 2,401 paired images from prostatectomy procedures, curated from *in vivo* video clips using motion-tracking techniques. We evaluate smoke-free image reconstruction using structural similarity (SSIM) and peak signal-to-noise ratio (PSNR). The values in Table 2 may differ slightly from those reported in [28] due to implementation details. As shown, our method consistently outperforms existing approaches on both the cholecystectomy and prostatectomy subsets.

Table 3 reports performance on our benchmark dataset, and Fig. 3 provides qualitative comparisons, showing that our method can recover scene content behind dense smoke with minimal color distortion and fewer artifacts. Because our dataset is evaluated at higher resolution, several competing methods degrade noticeably relative to their performance on lower-resolution benchmarks. We report results for *Ours-finetune* for completeness, but exclude it from direct

Table 2. Quantitative smoke-free image reconstruction results on the public real in vivo De-Smoking dataset.

Method	Cholecystectomy		Prostatectomy	
	SSIM	PSNR	SSIM	PSNR
He et al. [8]	0.73 ± 0.09	18.44 ± 3.03	0.72 ± 0.12	19.39 ± 3.38
Song et al. [21]	0.83 ± 0.10	23.25 ± 4.51	0.73 ± 0.15	19.22 ± 5.93
Qin et al. [17]	0.84 ± 0.10	24.49 ± 5.88	0.72 ± 0.15	19.13 ± 6.68
Salazar et al. [19]	0.75 ± 0.12	20.06 ± 4.00	0.72 ± 0.13	19.94 ± 4.10
Liu et al. [14]	0.81 ± 0.09	22.21 ± 3.79	0.73 ± 0.13	20.08 ± 5.70
Sidorov et al. [20]	0.76 ± 0.09	20.69 ± 3.89	0.66 ± 0.11	20.09 ± 2.82
Jin et al. [10]	0.74 ± 0.13	21.53 ± 3.89	0.62 ± 0.14	18.92 ± 3.79
Wu et al. [27]	0.83 ± 0.11	23.88 ± 6.89	0.67 ± 0.15	17.99 ± 6.13
Ours	0.86 ± 0.09	25.51 ± 4.69	0.75 ± 0.13	21.71 ± 4.41

Table 3. Quantitative results on our test set. We report 2D image reconstruction quality (SSIM/PSNR) and downstream-task consistency after desmoking: stereo depth estimation (MAE) and instrument segmentation (IoU).

Method	Image Reconstruction		Depth	Segmentation
	SSIM $\uparrow$	PSNR $\uparrow$	MAE (mm) $\downarrow$	% IoU $\uparrow$
None	—	—	47.22 ± 61.48	81.15 ± 17.01
He et al. [8]	0.53 ± 0.07	12.83 ± 1.53	51.53 ± 66.71	80.81 ± 17.03
Song et al. [21]	0.61 ± 0.09	14.45 ± 1.83	49.00 ± 66.45	80.90 ± 16.27
Qin et al. [17]	0.72 ± 0.09	19.57 ± 3.20	49.03 ± 63.84	81.15 ± 17.15
Salazar et al. [19]	0.54 ± 0.12	14.21 ± 2.05	53.73 ± 66.50	79.46 ± 16.85
Liu et al. [14]	0.65 ± 0.11	16.43 ± 2.63	48.02 ± 63.62	81.31 ± 16.18
Sidorov et al. [20]	0.59 ± 0.08	16.17 ± 1.86	59.44 ± 74.02	77.78 ± 16.70
Jin et al. [10]	0.50 ± 0.15	17.41 ± 2.75	56.48 ± 69.13	74.98 ± 17.36
Wu et al. [27]	0.70 ± 0.09	19.73 ± 4.26	53.03 ± 65.20	79.30 ± 17.13
Ours	0.73 ± 0.09	20.61 ± 3.12	49.42 ± 64.48	81.48 ± 16.85
Ours-finetune	0.75 ± 0.09	23.57 ± 3.35	47.64 ± 62.03	81.81 ± 16.08

comparison because the competing models were not fine-tuned on real smoke images. Nevertheless, our approach achieves the best overall performance, with an average SSIM of 0.73 and PSNR of 20.61.

3.2 Impacts on Downstream Computer Vision Tasks

Beyond image reconstruction, we investigate how surgical desmoking affects downstream computer vision tasks. We generate pseudo-ground-truth depth maps by running FoundationStereo [26] on the smoke-free images, and obtain instrument segmentation labels using Segment Anything 3 (SAM3) [3]. To assess the effect of desmoking, we first apply each method to the smoky images and then run FoundationStereo and SAM3 on the corresponding desmoked outputs. We evaluate depth estimation using mean absolute error (MAE) with respect to the pseudo-ground-truth depth, and evaluate segmentation using intersection-over-

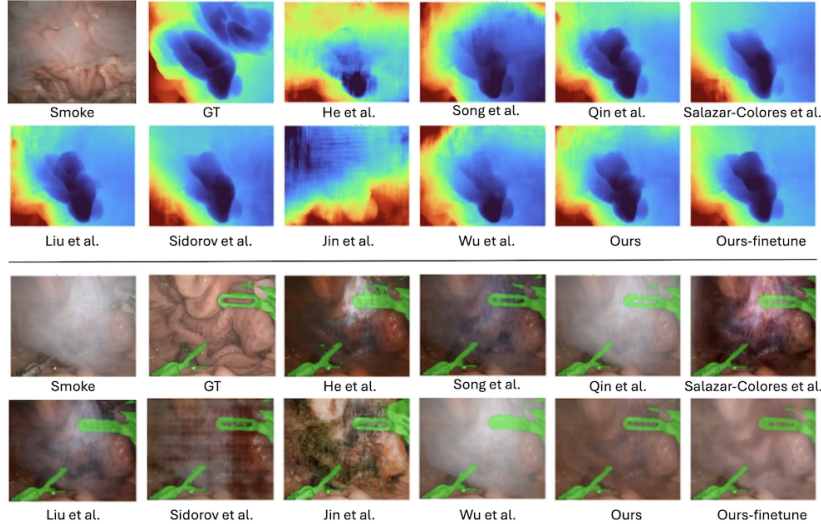


Fig. 4. Qualitative results on downstream stereo depth estimation (top) and instrument segmentation (bottom) after different desmoking methods.

union (IoU) against the generated labels. Table 3 summarizes the quantitative results, while Fig. 4 provide qualitative comparisons.

Interestingly, for stereo depth estimation, the no-desmoking baseline (*None*) performs best. One possible explanation is that foundation models such as FoundationStereo are relatively robust to smoke due to large-scale training. In addition, desmoking can alter image statistics and potentially disrupt left-right feature consistency, which is critical for stereo matching, thereby degrading reconstruction accuracy. Overall, our results suggest that current desmoking approaches do not consistently improve stereo depth estimation under smoke. In contrast, although some methods reduce segmentation performance, our approach improves IoU on smoky scenes. We attribute this to enhanced local contrast and clearer instrument boundaries after desmoking. While these findings reveal limitations of using desmoking purely as a preprocessing step, they also highlight its potential to benefit downstream applications (e.g., instrument segmentation) by improving visual perception in the presence of smoke.

4 Conclusion

We presented a transformer-based surgical desmoking framework that combines a ViT backbone with a physics-inspired head to jointly predict smoke-free images and corresponding smoke maps. To mitigate the scarcity of paired training data, we developed a large-scale synthetic data generation pipeline and curated a high-resolution paired dataset for benchmarking, enabling supervised training

and evaluation at unprecedented scale. Our method achieves state-of-the-art reconstruction performance across datasets, improving anatomical visibility under dense smoke while preserving visual fidelity. Analysis of downstream tasks shows that desmoking can improve instrument segmentation, but does not consistently benefit stereo depth estimation, underscoring the nuanced interaction between image reconstruction and geometry-based algorithms. Overall, these results position digital desmoking as a valuable component of surgical vision pipelines; future work will focus on real-time deployment and tighter integration with computer-assisted interventions.

Disclosure of Interests. All authors are employees of Intuitive Surgical. Intuitive Surgical manufactures the da Vinci surgical system used for data acquisition in this study. The authors declare no other competing interests relevant to the content of this article.

References

1. Baid, A., Kotwal, A., Bhalodia, R., Merchant, S., Awate, S.P.: Joint desmoking, specular removal, and denoising of laparoscopy images via graphical models and bayesian inference. In: 2017 IEEE 14th international symposium on Biomedical imaging (ISBI 2017). pp. 732–736. IEEE (2017)
2. Bolkar, S., Wang, C., Cheikh, F.A., Yildirim, S.: Deep smoke removal from minimally invasive surgery videos. In: 2018 25th IEEE International Conference on Image Processing (ICIP). pp. 3403–3407. IEEE (2018)
3. Carion, N., Gustafson, L., Hu, Y.T., Debnath, S., Hu, R., Suris, D., Ryali, C., Alwala, K.V., Khedr, H., Huang, A., et al.: Sam 3: Segment anything with concepts. arXiv preprint arXiv:2511.16719 (2025)
4. Chang, W., Li, Y., Zhu, Z., Yang, Y.: Lsd3k: A benchmark for smoke removal from laparoscopic surgery images. In: 2024 3rd International Conference on Artificial Intelligence, Internet of Things and Cloud Computing Technology (AIoTC). pp. 1–5. IEEE (2024)
5. Ding, H., Zhang, Y., Lu, T., Liang, R., Shu, H., Seenivasan, L., Long, Y., Dou, Q., Gao, C., Leng, Y., et al.: Segstrong-c: Segmenting surgical tools robustly on non-adversarial generated corruptions—an endovis’ 24 challenge. arXiv preprint arXiv:2407.11906 (2024)
6. Dosovitskiy, A.: An image is worth 16x16 words: Transformers for image recognition at scale. arXiv preprint arXiv:2010.11929 (2020)
7. He, K., Chen, X., Xie, S., Li, Y., Dollár, P., Girshick, R.: Masked autoencoders are scalable vision learners. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 16000–16009 (2022)
8. He, K., Sun, J., Tang, X.: Single image haze removal using dark channel prior. IEEE transactions on pattern analysis and machine intelligence **33**(12), 2341–2353 (2010)
9. Hong, T., Huang, P., Zhai, X., Gu, C., Tian, B., Jin, B., Li, D.: Mars-gan: multilevel-feature-learning attention-aware based generative adversarial network for removing surgical smoke. IEEE Transactions on Medical Imaging **42**(8), 2299–2312 (2023)

10. Jin, Y., Yan, W., Yang, W., Tan, R.T.: Structure representation network and uncertainty feedback learning for dense non-uniform fog removal. In: Asian Conference on Computer Vision. pp. 155–172. Springer (2022)
11. Kanakatte, A., Seemakurthy, K., Gubbi, J., Saha, J., Ghose, A., Purushothaman, B.: Surgical smoke dehazing and color reconstruction. In: 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). pp. 280–284. IEEE (2021)
12. Kotwal, A., Bhalodia, R., Awate, S.P.: Joint desmoking and denoising of laparoscopy images. In: 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI). pp. 1050–1054. IEEE (2016)
13. Li, H., Zhai, X., Xue, J., Gu, C., Tian, B., Hong, T., Jin, B., Li, D., Huang, P.: Multi-frequency and smoke attention-aware learning based diffusion model for removing surgical smoke. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 47–56. Springer (2024)
14. Liu, X., Suganuma, M., Sun, Z., Okatani, T.: Dual residual networks leveraging the potential of paired operations for image restoration. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 7007–7016 (2019)
15. McCartney, E.J.: Optics of the atmosphere: scattering by molecules and particles. New york (1976)
16. Pan, Y., Bano, S., Vasconcelos, F., Park, H., Jeong, T.T., Stoyanov, D.: Desmoke-lap: improved unpaired image-to-image translation for desmoking in laparoscopic surgery. International Journal of Computer Assisted Radiology and Surgery **17**(5), 885–893 (2022)
17. Qin, X., Wang, Z., Bai, Y., Xie, X., Jia, H.: Ffa-net: Feature fusion attention network for single image dehazing. In: Proceedings of the AAAI conference on artificial intelligence. vol. 34, pp. 11908–11915 (2020)
18. Ranftl, R., Bochkovskiy, A., Koltun, V.: Vision transformers for dense prediction. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 12179–12188 (2021)
19. Salazar-Colores, S., Jiménez, H.M., Ortiz-Echeverri, C.J., Flores, G.: Desmoking laparoscopy surgery images using an image-to-image translation guided by an embedded dark channel. IEEE Access **8**, 208898–208909 (2020)
20. Sidorov, O., Wang, C., Cheikh, F.A.: Generative smoke removal. In: Machine Learning for Health Workshop. pp. 81–92. PMLR (2020)
21. Song, Y., He, Z., Qian, H., Du, X.: Vision transformers for single image dehazing. IEEE Transactions on Image Processing **32**, 1927–1941 (2023)
22. Tchaka, K., Pawar, V.M., Stoyanov, D.: Chromaticity based smoke removal in endoscopic images. In: Medical Imaging 2017: Image Processing. vol. 10133, pp. 463–470. SPIE (2017)
23. Wang, C., Mohammed, A.K., Cheikh, F.A., Beghdadi, A., Elle, O.J.: Multiscale deep desmoking for laparoscopic surgery. In: Medical Imaging 2019: Image Processing. vol. 10949, pp. 505–513. SPIE (2019)
24. Wang, F., Sun, X., Li, J.: Surgical smoke removal via residual swin transformer network. International Journal of Computer Assisted Radiology and Surgery **18**(8), 1417–1427 (2023)
25. Weenink, R., Kloosterman, M., Hompes, R., Zondervan, P., Beerlage, H., Tanis, P., van Hulst, R.: The airseal® insufflation device can entrain room air during routine operation. Techniques in coloproctology **24**(10), 1077–1082 (2020)
26. Wen, B., Trepte, M., Aribido, J., Kautz, J., Gallo, O., Birchfield, S.: Foundation-stereo: Zero-shot stereo matching. In: Proceedings of the Computer Vision and Pattern Recognition Conference. pp. 5249–5260 (2025)

27. Wu, R., Zhang, Z., Zhang, S., Gou, L., Chen, H., Zhang, L., Chen, H., Zuo, W.: Self-supervised video desmoking for laparoscopic surgery. In: European Conference on Computer Vision. pp. 307–324. Springer (2024)
28. Xia, W., Peters, T.M., Fan, V., Sthanunathan, H., Qi, O., Chen, E.C.: In vivo laparoscopic image de-smoking dataset, evaluation, and beyond. *IEEE Transactions on Medical Imaging* (2025)
29. Zhang, J., Zhou, Z., Ho, W.S., Liu, Q., Tian, F., Guo, J.: Methods used for smoke evacuation in laparoendoscopic surgery: a technical review. *Surgery Today* **55**(11), 1513–1523 (2025)