ISSN 2706-655X Vol.3, Issue 10: 74-77, DOI: 10.25236/IJFET.2021.031009

Recent Advances in Deep Learning-based Smoke Removal Techniques for Laparoscopic Images

Zhengxuan Cao[1,2], Zhengyi Zhong[1,3], Hanze Wei[1], Rongrong Liu[1], Qun Wang[1*]

[1]Suzhou Institute of Biomedical Engineering and Technology Chinese Academy of Science, Suzhou, China
[2]School of Electrical Engineering, Shandong University, Jinan, China
[3]School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai, China
**Corresponding author*

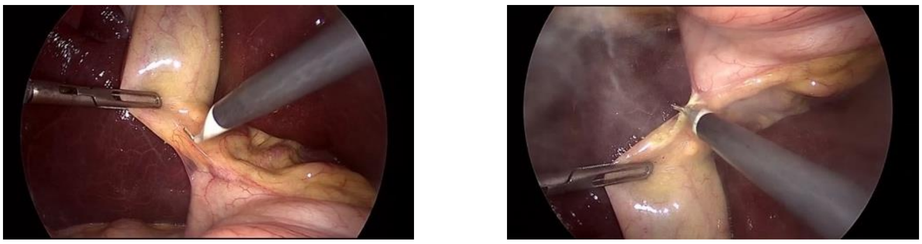
Abstract: *The visibility of the operating fields can be severely deteriorated by endoscopic smoke generated during laparoscopic surgery due to laser ablation and cauterization. Clinical studies have shown that removing smoke effects to laparoscopic images from the operating room reduces operating time and makes surgeons more comfortable during the procedure. Desmoking approaches based on deep learning have been found to be effective in the removal of laparoscopic smoke. This research will review several cutting-edge strategies for underlying theory and performance evaluations that have been developed in recent years.*

Keywords: *Laparoscopic surgery, Smoke removal, Deep learning, Defogging, Supervised learning, Unsupervised learning*

1. Introduction to Smoke Removal in Laparoscopic Surgery

Endoscopic smoke generated during laparoscopic surgery due to laser ablation and cauterization can severely deteriorate the visibility of the operating fields. As the visibility decreases, operations take longer, and the likelihood of surgical errors increases, all of which have severe consequences for patients. Not only that, but it also impairs image content-based analysis, such as segmentation, 3D reconstruction, target tracking, etc.

Clinical studies have discovered that eliminating smoke effects to laparoscopic images shortens operations and makes surgeons more comfortable during the surgery [2-4]. Therefore, it is necessary to employ a desmoking algorithm to remove smoke from endoscopic images.



(a) *(b)*

Figure 1: Laparoscopic smoke-free image (a) and smoke image (b) selected from dataset Cholec80 [1].

2. Introduction to Deep Learning

Deep learning has been broadly used in various applications in different research fields, such as image segmentation, image restoration, surgical smoke removal, etc. Theoretically, Deep learning is a sub-category of machine learning that involves multiple levels of representation, structured in layers of a combination of artificial neurons. The higher layer represents slightly more abstract features than the one below, incorporating non-linear but straightforward transformations. As a result, intricate representations can be learned by composing enough such transformations. Deep learning techniques usually fall into supervised, unsupervised, and semi-supervised categories.

ISSN 2706-655X Vol.3, Issue 10: 74-77, DOI: 10.25236/IJFET.2021.031009

3. Taxonomy of Techniques

Researchers have made numerous attempts to minimize the smoke effects in the surgical field of view. Generally, there are two types of endoscopic desmoking algorithms: conventional approaches and deep learning-based approaches. The traditional method can be considered a smoke-removal procedure known as image de-hazing or defogging [14-15]. On the other hand, deep learning-based solutions have shown promising effects in practice, therefore researchers have integrated various deep learning approaches with image processing techniques to build new artificial intelligence-based desmoking algorithms.

The smoke removal algorithms can be divided into supervised-based learning and unsupervised-based learning based on the algorithms. In the following sections, we will go over a variety of laparoscopic desmoking algorithms that have emerged in recent years, depending on the deep learning approaches chosen.

3.1. Supervised Learning

Supervised learning is the most straightforward type of learning method. It builds a predictive model out of numerous training examples. The model adjusts its internal parameters, called weights, to reduce the error between the output and target scores during the training process. This error correction-learning process trains the network based on hundreds of thousands of input-output labeled pairs.

One of the most successful techniques in supervised learning is Convolution Neural Network, called CNN, which has been broadly used to cope with high-dimensional data, such as medical images and videos. It constructs a model on labeled data, whose structure is inspired by the neurobiology of the visual cortex, bearing a similarity to a conventional neural network. A typical CNN architecture contains convolutional layer(s) followed by a fully connected (FC) layer(s). Pooling layers exist in between convolutional layers. CNN iterates between forward-propagation, where input data are transformed into output, and backpropagation, where a gradient descent algorithm is involved according to the loss value.

Due to the superior performance of CNN, the author in [9] proposed an improved CNN with encoder-decoder architecture for real-time surgical smoke removal. The network takes smoke images and its Laplacian image pyramid decomposition as inputs and outputs a clean image without smoke. They created the synthetic dataset by adding rendered clouds to the clean image with Blender and Adobe Photoshop.

Other researchers presented different approaches. For example, J. Lin et al. [7] proposed a variational U-Net that uses a convolutional block attention module as an embedded guide mask of the decoder component, inspired by the NTIRE 2020 Challenge. U-Net is a supervised model whose architecture is in the shape of "U." The network is symmetric and composed of two major parts: the contracting path on the left and the expansive path on the right. The author used Blender to create graphics from laparoscopic photos with light and dense smoke as the training set. A CBAM module was introduced to the decoder component of the U-Net to improve the models' inter features. The SSIM for test data is up to 0.945, and the PSNR is up to 29.27.

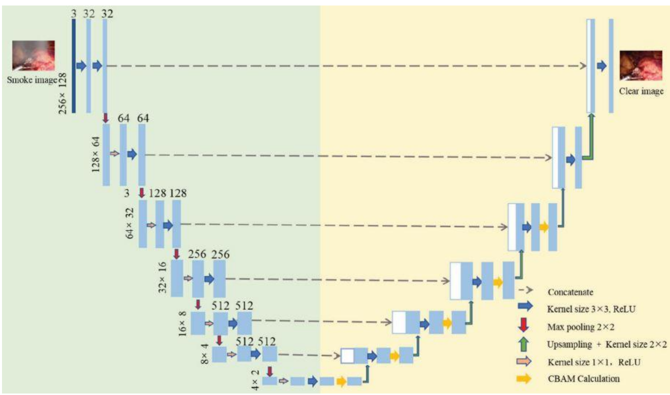


Figure 2: U-Net architecture proposed in [7].

Sabri Bolka et al. [8] construct a synthetic dataset including pairs of images equipped with ground truth clean tissue images and smoke embedded versions regarding surgical smoking. They borrowed the All-in-One Dehazing Network (AOD-Net) model, initially designed for outdoor defogging, and applied it to surgical smoking removal. The supervised model consists of five convolutional layers with ReLU activation units and three concatenating layers. The author manages performance comparisons with previous defogging methods in MSE, PSNR, and MAD.

3.2. Unsupervised Learning

Unlike supervised learning, unsupervised learning does not require specifically labeled data. The learning model can automatically classify the data by discovering inherent attributes, creating labels subsequently used to implement supervised learning tasks.

Generative Adversarial Networks (GAN) is an unsupervised learning method whose objective is to calculate the distribution of original samples. It was first presented as a form of a generative model, although it has been found helpful domains such as supervised learning [13], semi-supervised learning [12], and reinforcement learning. It initially employs a Discriminator and a Generator whose work is to generate new samples from the real data samples. GAN has made significant development in image quality restoration, image processing, etc.

GAN framework has played an essential role in the research field of smoke removal. For instance, Oleksii Sidorov et al. [6] developed a GAN-based technique for smoke removal using an end-to-end model trained on synthetic data. The author argued that the simple use of basic GAN could remove smoke, but it degrades image quality and introduces synthetic noise. Therefore, they proposed a typical pix2pix-like architecture combined with perceptual loss and MS-SSIM loss to create a practical image enhancement approach.

Researchers also integrated classic desmoking algorithms with the GAN framework. S. S. Colores [5] combined a conditional unsupervised network with an embedded dark channel mask. The generative adversarial network includes four channels (RGB channels & dark channel) as input and one output. The smoke-free image is created by the first three channels of the input. The dark channel, aka dark channel prior, is based on "dark pixels," which contains low intensity in at least one color channel. He [11] proposed the defogging technique and constructed the theory on the atmospheric scattering model. Due to its simplicity and effectiveness, DCP immediately gains popularity as soon as it is introduced.

Vishal V. et al. [10] presented a GAN-based network that contains two discriminators and two generators. Two discriminators create synthetic smoke and smoke-free images, respectively. Two discriminators are designed to distinguish synthetic smoke and smoke-free images from real smoke and smoke-free images.

Table 1: Comparison between different desmoking methods based on deep Learning

Paper	Supervised/Unsupervised	Model + Algorithm	Evaluation Metrics	Speed	Synthetic or Real?	Training(pairs)	Testing(pairs)
S. Salazar-Colores et al.	Unsupervised	GAN + DCP	PSNR / SSIM	92 fps	Simulated by Python Cloud + Real	20,000	2398
O. Sidorov et al.	Unsupervised	Improved GAN	Mean / STD	~	Adobe PhotoShop 4 + Real	22,500	300
J. Lin et al.	Supervised	U-Net + CBAM	PSNR / SSIM	73.25 fps, 256x128	Blender + Real	7000	556
S. Bolkar et al.	Supervised	AOD-Net	MSE, PSNR, and MAD	20 fps, 512x512	Blender & PhotoShop + Real	19500	100
C. Wang et al.	Supervised	Improved CNN	PSNR / SSIM	26 fps, 512x512	Adobe PS + Real	75530	300
V. Vishal	Unsupervised	Variational GAN	BRISQE PIQUE CEIQ	~	Real	1200	200

4. Conclusion

Surgical smoke removal is critical for surgeons to enhance the visibility of the operating field. This study investigates several deep learning-based smoke removal methods for laparoscopic surgery. As discovered in this review, multiple approaches have been approved effective in laparoscopic image smoke removal. With the continuous development of machine learning technologies, more and more

ISSN 2706-655X Vol.3, Issue 10: 74-77, DOI: 10.25236/IJFET.2021.031009

methods will be developed to take desmoking techniques to new heights.

Acknowledgements

This work is supported by National Key R&D Program of China, 2019YFC0118004

References

[1] A. P. Twinanda, S. Shehata, D. Mutter, J. Marescaux, M. de Mathelin and N. Padoy, "EndoNet: A Deep Architecture for Recognition Tasks on Laparoscopic Videos," in IEEE Transactions on Medical Imaging, vol. 36, no. 1, pp. 86-97, Jan. 2017.

[2] Salimans, Tim; Goodfellow, Ian; Zaremba, Wojciech; Cheung, Vicki; Radford, Alec; Chen, Xi (2016). "Improved Techniques for Training GANs."

[3] Isola, Phillip; Zhu, Jun-Yan; Zhou, Tinghui; Efros, Alexei (2017). "Image-to-Image Translation with Conditional Adversarial Nets." Computer Vision and Pattern Recognition.

[4] Ho, Jonathon; Ermon, Stefano (2016). "Generative Adversarial Imitation Learning." Advances in Neural Information Processing Systems. 29: 4565-C4573.

[5] S. Salazar-Colores, H. M. Jiménez, C. J. Ortiz-Echeverri, and G. Flores, "Desmoking Laparoscopy Surgery Images Using an Image-to-Image Translation Guided by an Embedded Dark Channel," in IEEE Access*, vol. 8, pp. 208898-208909, 2020*

[6] Oleksii Sidorov, Congcong Wang, Faouzi Alaya Cheikh Proceedings of the Machine Learning for Health NeurIPS Workshop, PMLR 116:81-92, 2020.

[7] Lin, J, Jiang, M, Pang, Y, et al. A desmoking algorithm for endoscopic images based on improved U-Net model. Concurrency Computat Pract Exper. 2021.

[8] S. Bolkar, C. Wang, F. A. Cheikh and S. Yildirim, "Deep Smoke Removal from Minimally Invasive Surgery Videos," 2018 25th IEEE International Conference on Image Processing (ICIP)*, 2018, pp. 3403-3407*

[9] Congcong Wang, Ahmed Kedir Mohammed, Faouzi Alaya Cheikh, Azeddine Beghdadi, and Ole Jacob Elle "Multiscale deep desmoking for laparoscopic surgery," Proc. SPIE 10949, Medical Imaging 2019: Image Processing, 109491Y (15 March 2019)

[10] Vishal V., Sharma N., Singh M. "Guided Unsupervised Desmoking of Laparoscopic Images Using Cycle-Desmoke". In: Zhou L. et al. (eds) OR 2.0 Context-Aware Operating Theaters and Machine Learning in Clinical Neuroimaging. OR 2.0 2019, MLCN 2019. Lecture Notes in Computer Science, vol 11796. Springer, Cham.

[11] Kaiming He, Jian Sun and Xiaoou Tang, "Single image haze removal using dark channel prior," 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 1956-1963.

[12] A Odena, "Semi-supervised learning with generative adversarial networks Workshop on Data-Efficient Machine Learning", ICML 2016.

[13] K. S and M. Durgadevi, "Generative Adversarial Network (GAN): a general review on different variants of GAN and applications," 2021 6th International Conference on Communication and Electronics Systems (ICCES), 2021, pp. 1-8.

[14] A. Baid, A. Kotwal, R. Bhalodia, S. N. Merchant, and S. P. Awate, "Joint Desmoking, Specularity Removal, and Denoising of Laparoscopy Images Via Graphical Models and Bayesian Inference" Indian Institute of Technology (IIT) Bombay. University of Utah., in IEEE 14th International Symposium on Biomedical Imaging, 2017, pp. 732–736.

[15] A. Kotwal, "Joint Desmoking and Denoising of Laparoscopy Images" Department of Electrical Engineering Indian Institute of Technology (IIT) Bombay Department of Computer Science and Engineering Indian Institute of Technology (IIT) Bombay," in IEEE 13th International Symposium on Biomedical Imaging (ISBI), 2016, pp. 1050–1054.