Proceedings of Machine Learning Research 116:81–92, 2020 Machine Learning for Health (ML4H) at NeurIPS 2019

Generative Smoke Removal

Oleksii Sidorov OLEKSIIS@FB.COM
Congcong Wang CONGCONG.WANG@NTNU.NO
Faouzi Alaya Cheikh FAOUZI.CHEIKH@NTNU.NO
The Norwegian Colour and Visual Computing Laboratory
Norwegian University of Science and Technology

Editors: Adrian V. Dalca, Matthew Mcdermott, Emily Alsentzer, Sam Finlayson, Michael Oberst, Fabian Falck, and Brett Beaulieu-Jones

Abstract

In minimally invasive surgery, the use of tissue dissection tools causes smoke, which inevitably degrades the image quality. This could reduce the visibility of the operation field for surgeons and introduces errors for the computer vision algorithms used in surgical navigation systems. In this paper, we propose a novel approach for computational smoke removal using supervised image-to-image translation. We demonstrate that straightforward application of existing generative algorithms allows removing smoke but decreases image quality and introduces synthetic noise (grid-structure). Thus, we propose to solve this issue by modification of GAN's architecture and adding perceptual image quality metric to the loss function. Obtained results demonstrate that proposed method efficiently removes smoke as well as preserves perceptually sufficient image quality.

1. Introduction

In laparoscopic surgery, patient's abdomen is visualized by a camera which is inserted into the body through small incisions. High quality of the captured video is necessary to keep a clear visualization for the operating surgeons as well as for navigation systems (Stoyanov, 2012; Andrea et al., 2018). However, the perceptual quality can be significantly degraded by smoke caused by such tool as laser ablation. The surgeons' visibility is inevitably impacted by this degradation. Furthermore, the computer vision technique based surgical navigation systems are mainly designed for clear videos (Andrea et al., 2018; Wang et al., 2018c), smoke would influence the performance. Therefore, in order to maintain a clear operation field, it becomes necessary to remove the smoke from laparoscopic images by smoke evacuation techniques (Lawrentschuk et al., 2010) and by computer vision algorithms (Wang et al., 2018b). Especially, a real time automatic image processing based method is desired which would not introduce any extra hardware to the surgical procedure. Moreover, the algorithm can be embedded to computer assisted surgical navigation workflow easily.

The majority of existing smoke removal algorithms are based on simplified physical models or assumptions about input image data which limit their practical application. In this paper, our goal is to avoid using any assumptions or simplified models and perform real-time smoke removal end-to-end. It became possible due to the development of deep learning algorithms oriented on conditional image generation. These algorithms utilize the

adversarial process to learn mapping between two image domains in a supervised manner. In case of smoke removal task, these two domains are images with and without smoke correspondingly.

Thus, in our work, we perform analysis of image generation errors and propose a new loss function which allows to optimize image quality during the training and produces data without noise and artifacts. We utilize a set of smoke-free images cast with synthetic smoke to train the network, and further evaluate its performance on real-world data.

The remainder of this paper is structured as follows. In Section 2, we review the related work on smoke removal and image generation using GANs. Next, in Section 3, we describe our proposed method. Section 4 presents the training of the network and discusses the experimental results. Finally, the conclusions are drawn in Section 5.

2. Related works

2.1. Laparoscopic Smoke Removal

In this part, we group the smoke removal methods to traditional approaches and deep learning approaches.

Traditional approaches: As dehazing and desmoking problem share some similarity, traditional desmoking approaches (Wang et al., 2018b; Kotwal et al., 2016; Baid et al., 2017; Tchakaa et al., 2017; Luo et al., 2017) follow similar strategy as dehazing approaches. In those literatures, the atmospheric scattering model (Narasimhan and Nayar, 2002) described in Equation (1) is widely used.

$$\mathbf{I}(x,y) = \mathbf{J}(x,y)t(x,y) + \mathbf{A}(1 - t(x,y)), \tag{1}$$

where $\mathbf{I}$ is the observed haze image, $\mathbf{J}$ represents the haze-free image, $\mathbf{A}$ is the global atmospheric light and t is the medium transmission map. In (Kotwal et al., 2016), desmoking and denoising problem is formulated to probabilistic graphical model and then it is extended in (Baid et al., 2017) for desmoking, denoising and specular removal. In (Tchakaa et al., 2017), a dark channel prior dehazing method originally proposed in (He et al., 2011) is modified for smoke removal purpose. In (Luo et al., 2017), Luo *et al.* propose to estimate atmospheric veil ($\mathbf{A}(1 - t(x,y))$) directly instead of calculating t. In (Wang et al., 2018b), Wang *et al.* present a variational method to estimate the atmospheric veil. The approaches proposed in (Luo et al., 2017; Wang et al., 2018b) show promising results, but the performance degraded for dense and heterogeneous smoke. Moreover, all the methods rely on some assumptions, therefore, the methods' performance degenerates when the assumptions are wrong.

Deep Learning approaches: In (Bolkar et al., 2018), Bolkar *et al.* propose the first deep learning desmoking approach. Synthetic dataset created by Perlin noise is generated and used for fine tuning AODNet (Li et al., 2017). Later, in (Chen et al., 2018), a conditional Generative Adversarial Network is trained by Blender[1] generated synthetic dataset for desmoking. These deep learning based methods show promising direction for developing real-time smoke removal algorithms.

1. https://www.blender.org/

2.2. Image generation using GANs

The deconvolutional ("transposed convolutional") layers of a CNN (Convolutional Neural Network) have made possible the generation of an output of the same size as an input image. However, L1- or L2-loss used as similarity metric leads to a prediction of blurred images. Generative Adversarial Networks (GANs) (Goodfellow et al., 2014) solve this issue by adding an adversarial loss, implemented as separate CNN with a binary output (discriminator), that allows achieving a photo-realistic quality of a synthesis. Recently, GANs demonstrated state-of-the-art performance in numerous computer vision tasks, such as image mapping (Isola et al., 2017), video generation (Wang et al., 2018d), segmentation (Luc et al., 2016), inpainting (Iizuka et al., 2017), *etc.* Furthermore, a few GAN-based models were proposed recently for image dehazing (Bharath Raj and Venkateswaran, 2018; Li et al., 2018).

Isola *et al.* (Isola et al., 2017) first demonstrated the great potential of supervised image-to-image translation with the algorithm called pix2pix. It was successfully applied to image colorization, segmentation, generation of images from edges and segmented labels. Pix2pix utilizes L1-loss of U-Net (Ronneberger et al., 2015) shaped generator together with an adversarial loss of discriminator. On the other hand, Perceptual Adversarial Network (PAN) (Wang et al., 2018a) and pix2pixHD (Wang et al., 2017), apply similar technique complemented by "perceptual loss". The latter was first presented by Johnson *et al.* (Johnson et al., 2016) with an objective to cover not only a pixel-wise similarity but also a similarity of high-level features. A perceptual loss is created as a concatenation of activations extracted from different layers of pretrained VGG-16 (pix2pixHD) or directly from discriminator (PAN). Nevertheless, despite being called "perceptual", this metric has no relation to the human visual system that we discuss further in the text.

While pix2pix and PAN learn mapping in a supervised manner, algorithms like CycleGAN (Zhu et al., 2017), DiscoGAN (Kim et al., 2017), and UNIT (Liu et al., 2017) utilize cycle loss which allows using unpaired data. This approach is not covered in the scope of this work, but it has high potential to be applied for smoke removal when trained using unrelated batches of images with and without smoke.

3. Methodology

3.1. Model architecture and loss function

As a baseline for the implementation of our approach, we used PAN proposed by Wang *et al.* (Wang et al., 2018a). It has similar architecture to pix2pix except it employs a more efficient loss function. Fig. 1 (b) illustrates the main drawback of the results obtained when PAN is applied to desmoking straightforward: strongly manifested periodic noise (grid structure). It also manifests in Fourier spectrum of the images (Fig. 1 (c)). Considering that standard approach to removing spatial periodic noise is to apply image filtering in the frequency space, we designed a mask which maximally covers undesired peaks. Nevertheless, this approach did not improve the image quality significantly, even with various shapes of the mask tested.

The failure of classical image processing in Fourier domain and the desire to keep the model end-to-end motivated us to improve the original PAN algorithm. Since the main

issue is the quality of the output images (structural artifacts), the logical step was to add perceptual image quality metric to the loss function and minimize it during training. Peak Signal-to-Noise Ratio (PSNR) and Structural Similarity Index (SSIM) (Wang et al., 2004) are two most commonly used full-reference image quality metrics. The significant advantage of them compared to less popular image quality metrics is their differentiability which allows to use them for gradient computation. However, PSNR (Eq. (2)) was proven (Zhang et al., 2012) to not correlate well with human perception. Moreover, it is highly related to L2 metric (RMSE) which is not suitable for image translation since it produces blurry results (Isola et al., 2017).

$$PSNR = 20log_{10}(max(I)) - 10log_{10}MSE \tag{2}$$

SSIM is a perceptual image similarity metric which was proposed as an alternative to MSE (Mean Square Error) and PSNR in order to increase correlation with subjective evaluation. For original and reconstructed images I and J, SSIM is defined as:

$$SSIM(I,J) = \frac{(2\mu_I\mu_J + c_1)(2\sigma_{IJ} + c_2)}{(\mu_I^2 + \mu_J^2 + c_1)(\sigma_I^2 + \sigma_J^2 + c_1)}, \tag{3}$$

where μ_I, μ_J and σ_I, σ_J are mean and variance of images I and J correspondingly, while σ_{IJ} is covariance of the images. Therefore, corresponding loss function was defined as:

$$L_{SSIM} = -mean(SSIM(I,J)). \tag{4}$$

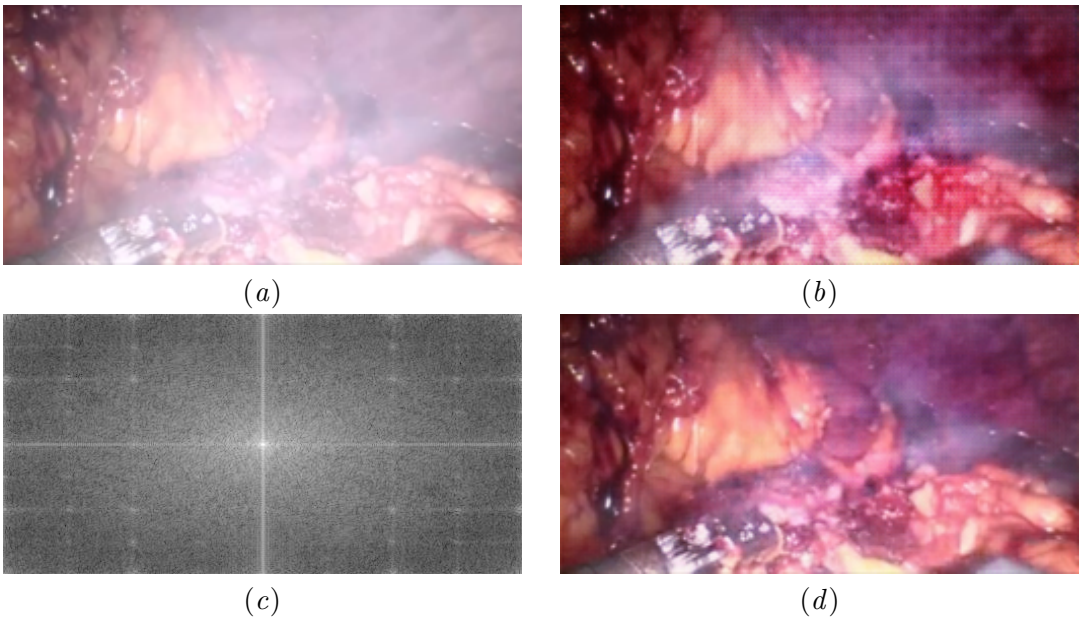

Figure 1: (a) input image; (b) raw PAN output; (c) a magnitude spectrum of the Fourier transform of (b); (d) output generated by PAN with added MS-SSIM loss.

Since SSIM changes in range $0 \sim 1$, where higher value corresponds to higher similarity, it has to be inverted in order to join it with other losses which are minimized during an optimization.

While SSIM processes windows of specified size $n \times n$, its extension Multi-Scale SSIM (MS-SSIM) (Zhang et al., 2012) takes into account windows of different sizes and allows covering a bigger range of spatial frequencies that lead to better results. So, in addition to SSIM we also used MS-SSIM loss which is computed in the analogous way.

The complete framework of the algorithm is illustrated in Figure 2. The generator G is a U-Net-like convolutional network with skip connections. The discriminator D is a conventional CNN-based binary classifier. The encoding layers of both G and D consist of $Convolution$ layer followed by $BatchNorm$, and $LeakyReLU$. The size of filters in $Convolution$ layer is 3×3, whereas their number and stride are marked in Figure 2. The decoding layers follows the same template but with $DeConvolution$ of size 4×4 instead of $Convolution$. The complex loss function is constructed as a linear combination of the SSIM (MS-SSIM) loss, "perceptual" loss extracted from the discriminator's layers, and an adversarial loss.

3.2. Implementation Details

The model is implemented in Pytorch 0.4 and its code is publicly accessible[2]. The training was performed using single 4GB GTX980 GPU. Original images were re-sized and zero-padded in order to match 256x256 input size. In our experiments we used ADAM optimizer with a learning rate of 0.0002 and momentum 0.5, batch size 4, and 50 training epochs. All the other parameters' values can be found in the code provided.

4. Results and discussion

4.1. Dataset

There exist no labeled datasets for desmoking. Therefore, the synthetic dataset from (Wang et al., 2019) [3] is used to train our network. Manually selected smoke-free images from (Chen et al., 2018) are used as the groundtruth images, then Adobe Photoshop[4] is used to render clouds to simulate the smoke images appearances. As a result, 7,500 smoke-free images with smoke of three different densities produced 22,500 synthetic image pairs which were used for training. The evaluation, however, was performed using both types of images: with synthetic smoke (ground truth available) and with real smoke (no ground truth)[5].

4.2. Experimental results

According to the availability of the source code and the suitability of it for smoke removal, we compare the proposed method with four following methods: physical model based dark channel prior (DCP) (He et al., 2011) and refined dark channel prior (R-DCP) (Tchakaa

2. https://github.com/acecreamu/ssim-pan
3. http://hamlyn.doc.ic.ac.uk/vision/
4. https://www.adobe.com/products/photoshop.html
5. The data with real smoke has been captured dynamically during surgery.

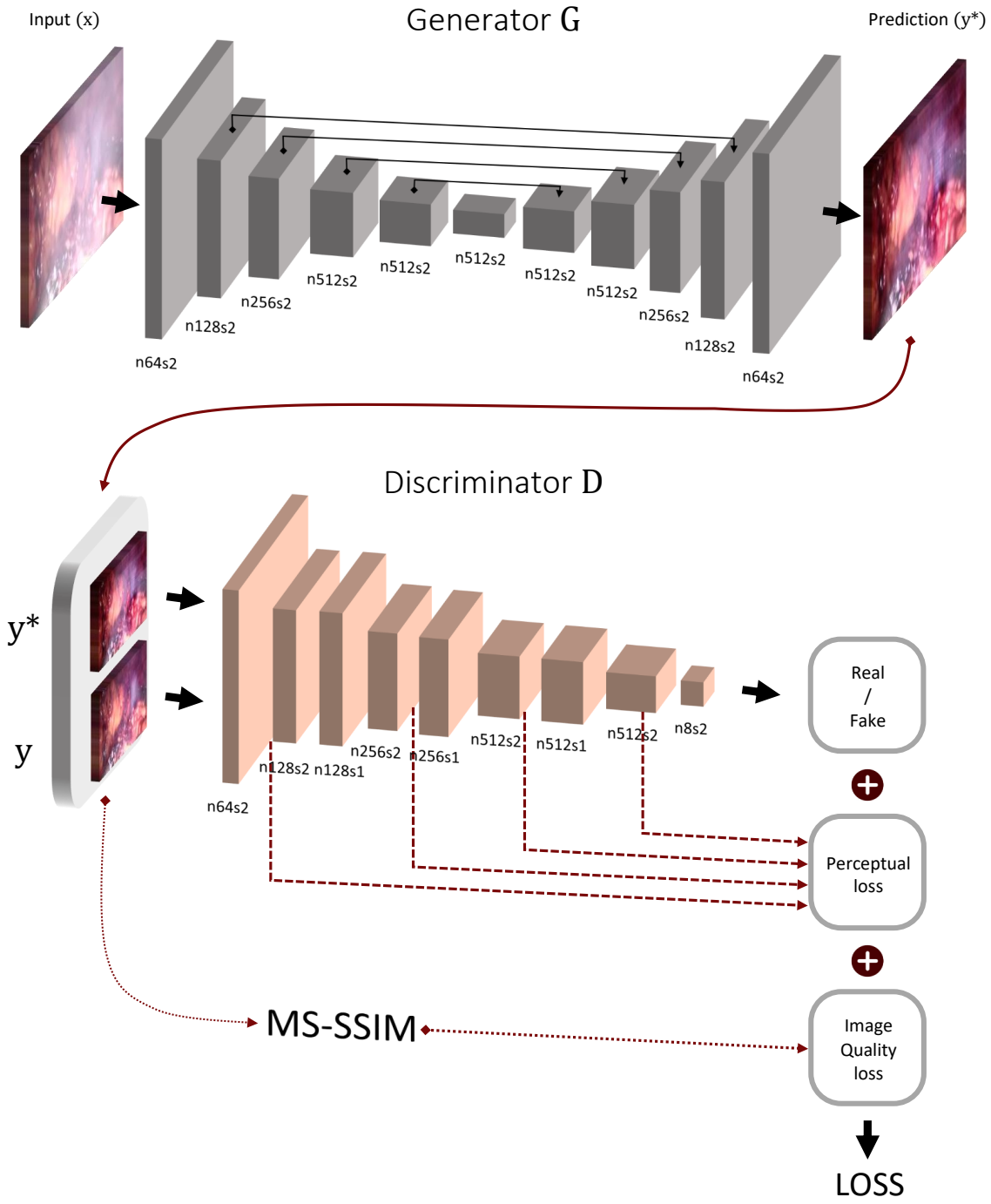

Figure 2: PAN Framework complemented by MS-SSIM image quality loss. "nAsB" denotes A filters of stride B. The detailed description of the layers is in the text.

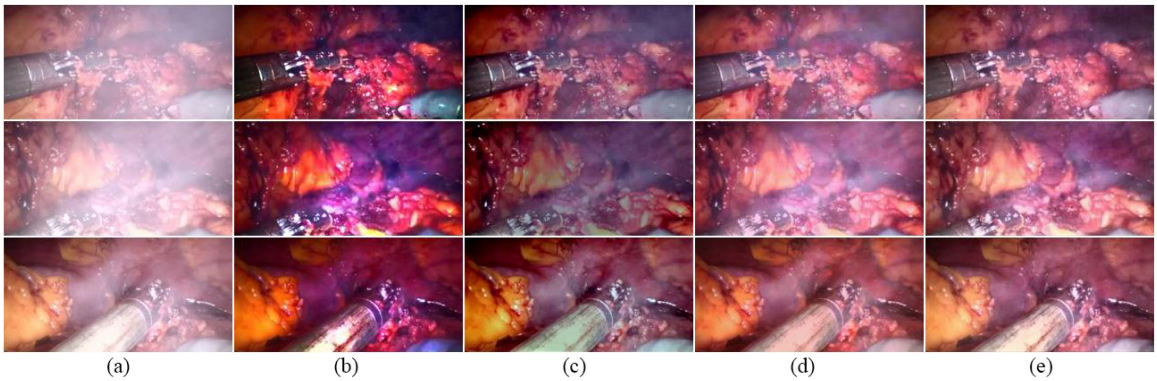

Figure 3: Qualitative comparison. (a) Input smoke images and desmoked ones by: (b) DCP (He et al., 2011), (c) VAR (Wang et al., 2018b), (d) EVID (Galdran et al., 2015), (e) proposed method. Zoom is required.

et al., 2017) methods, mild physical constraint based variational approach EVID (Galdran et al., 2015) and recently proposed desmoking approach VAR (Wang et al., 2018b).

300 real images with added synthetic smoke have been used for quantitative evaluation due to the direct availability of ground truth. As shown in Table 1, we report results in terms of colorimetric difference CIEDE2000 (Luo et al., 2001) (which describes accuracy of color reconstruction for human visual system) as well as RMSE of pixel values (which is important for computational algorithms). Our proposed method outperforms the other approaches in terms of CIEDE2000 and RMSE.

Qualitative result obtained from real smoke images is illustrated in Fig. 3. It can be clearly seen that the proposed approach outperforms all the other techniques and produces significant visibility enhancement even in the cases with very dense smoke. Another beneficial property is preserving of color information similar to original. Additional improvement of the results is expected in case of using larger and more diverse dataset for training.

Perceptual evaluation. The real data does not contain ground truth smoke-free images which makes quantitative comparison of the results troublesome. However, since our pri-

Table 1: Quantitative evaluation results.

	CIEDE2000		RMSE	
	mean	std	mean	std
RDCP (Tchakaa et al., 2017)	12.9	1.32	35.0	**4.55**
DCP (He et al., 2011)	11.5	2.13	40.6	6.98
VAR (Wang et al., 2018b)	10.8	2.18	38.1	8.00
EVID (Galdran et al., 2015)	7.41	1.40	24.2	4.81
Proposed	**3.89**	**1.15**	**4.6**	4.71

mary goals is to enhance visual data used by clinicians for simplification of the surgery process, the most relevant evaluation possible is a perceptual experiment with real doctors. This motivated us to gather subjective responses and define which algorithm is more likely to be chosen for real application.

Surgeons' time is expensive, so we designed our experiment as a short online-survey for the sake of easier accessibility and reaching a larger number of participants. The survey is available online[6] and can be used to find more outputs and evaluate them personally. The participant base consisted of 45 surgeons who kindly followed email-invitation. Each trial presented 10 image-choice questions split into two questions: *"Which image do you prefer?"* and *"Which image would be the most useful during a surgery?"*. Figure 4 illustrates that our approach has earned the largest number of votes in both tasks, even though EVID method by (Galdran et al., 2015) is a strong competitor.

4.3. Discussion

In this paper, we show how image quality metrics can improve the results of supervised image-to-image translation in medical domain. This is not a unique case. The usage of image quality metrics in deep learning was first discussed by Dosovitskiy *et al.* (Dosovitskiy and Brox, 2016). Following works (Snell et al., 2017; Zhao et al., 2017) illustrated its application to image restoration and super-resolution. However, it has never been applied to GANs where it is especially useful. From the other side, Odena *et al.* (Odena et al., 2016) relate above-mentioned artifacts (discussed in 3.1) to the uneven overlap of deconvolutional filters and propose to solve it by preliminary resizing of an image. Isola[7] stated that this operation applied to pix2pix can increase training time up to 4 times. On the other hand,

6. https://www.surveymonkey.com/r/2XL9JPH
7. https://github.com/junyanz/pytorch-CycleGAN-and-pix2pix/issues/78#issuecomment-322908732

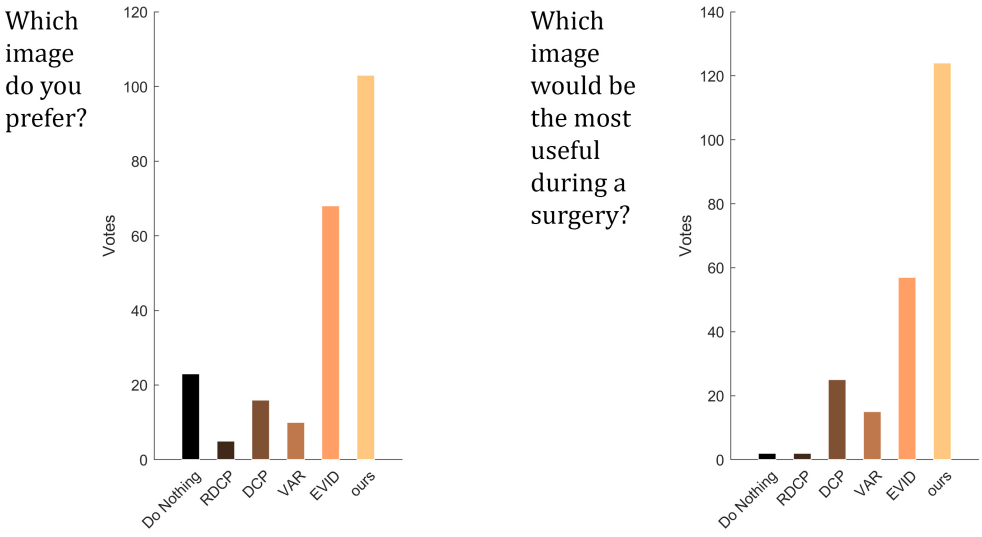

Figure 4: The results of perceptual evaluation by real surgeons.

Figure 5: The comparison of original and modified versions of PAN illustrated on other datasets.

our solution increases time of one training epoch only by 13%, which is annihilated by faster convergence in a fewer number of epochs.

The additional results of applying the proposed method to the datasets from other domains (edge2shoes (Isola et al., 2017) and SRD (Qu et al., 2017)) are demonstrated in Fig. 5 and compared to the output of the original PAN from the same epoch. As can be seen, the proposed approach achieves significantly better perceptual image quality regardless of the domain of application.

5. Conclusions

In this work, we present the novel GAN-based approach for smoke removal from laparoscopic images. We show how the end-to-end model trained on synthetic data can demonstrate a remarkable performance on real-world images. We used the standard pix2pix-like architecture complemented by "perceptual" loss and MS-SSIM loss to obtain an effective image enhancement method which is useful for clinicians operating surgery as well as for computer vision algorithms.

Further development may include gathering larger real-world dataset, modification of the code for a real-time application, and application of similar unsupervised methods (CycleGAN, UNIT, *etc.*).

References

Teatini Andrea, Wang Congcong, Palomar Rafael, Alaya Cheikh Faouzi, Beghdadi Azeddine, Edwin Bjørn, and Elle Ole Jakob. Validation of stereo vision based liver surface reconstruction for image guided surgery. In *Colour and Visual Computing Symposium (CVCS)*, pages 1–6. IEEE, 2018.

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. In *International Symposium on Biomedical Imaging (ISBI)*, pages 732–736. IEEE, 2017.

N Bharath Raj and N Venkateswaran. Single image haze removal using a generative adversarial network. *arXiv preprint arXiv:1810.09479*, 2018.

Sabri Bolkar, Congcong Wang, Faouzi Alaya Cheikh, and Sule Yildirim. Deep smoke removal from minimally invasive surgery videos. In *International Conference on Image Processing (ICIP)*. IEEE, 2018.

L. Chen, T. Wen, and N.W. John. Unsupervised learning of surgical smoke removal from simulation. In *Hamlyn Symposium on Medical Robotics*. Imperial College London, 2018.

Alexey Dosovitskiy and Thomas Brox. Generating images with perceptual similarity metrics based on deep networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 658–666, 2016.

Adrian Galdran, Javier Vazquez-Corral, David Pardo, and Marcelo Bertalmio. Enhanced variational image dehazing. *SIAM Journal on Imaging Sciences*, 8(3):1519–1546, 2015.

Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in Neural Information Processing Systems (NIPS)*, pages 2672–2680, 2014.

K. He, J. Sun, and X. Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2011.

Satoshi Iizuka, Edgar Simo-Serra, and Hiroshi Ishikawa. Globally and locally consistent image completion. *ACM Transactions on Graphics (TOG)*, 36(4):107, 2017.

Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A Efros. Image-to-image translation with conditional adversarial networks. *arXiv preprint*, 2017.

Justin Johnson, Alexandre Alahi, and Li Fei-Fei. Perceptual losses for real-time style transfer and super-resolution. In *European Conference on Computer Vision (ECCV)*, pages 694–711. Springer, 2016.

Taeksoo Kim, Moonsu Cha, Hyunsoo Kim, Jung Kwon Lee, and Jiwon Kim. Learning to discover cross-domain relations with generative adversarial networks. *arXiv preprint arXiv:1703.05192*, 2017.

A. Kotwal, R. Bhalodia, and S. P. Awate. Joint desmoking and denoising of laparoscopy images. In *International Symposium on Biomedical Imaging (ISBI)*, pages 1050–1054. IEEE, 2016.

Nathan Lawrentschuk, Neil E Fleshner, and Damien M Bolton. Laparoscopic lens fogging: a review of etiology and methods to maintain a clear visual field. *Journal of endourology*, 24(6):905–913, 2010.

Boyi Li, Xiulian Peng, Zhangyang Wang, Jizheng Xu, and Dan Feng. Aod-net: All-in-one dehazing network. In *International Conference on Computer Vision (ICCV)*, pages 4770–4778, 2017.

Runde Li, Jinshan Pan, Zechao Li, and Jinhui Tang. Single image dehazing via conditional generative adversarial network. *methods*, 3:24, 2018.

Ming-Yu Liu, Thomas Breuel, and Jan Kautz. Unsupervised image-to-image translation networks. In *Advances in Neural Information Processing Systems (NIPS)*, pages 700–708, 2017.

Pauline Luc, Camille Couprie, Soumith Chintala, and Jakob Verbeek. Semantic segmentation using adversarial networks. *arXiv preprint arXiv:1611.08408*, 2016.

M. R. Luo, G. Cui, and B. Rigg. The development of the cie 2000 colour-difference formula: Ciede2000. *Color Research & Application*, 26(5):340–350, 2001. doi: 10.1002/col.1049. URL https://onlinelibrary.wiley.com/doi/abs/10.1002/col.1049.

Xiongbiao Luo, A Jonathan McLeod, Stephen E Pautler, Christopher M Schlachta, and Terry M Peters. Vision-based surgical field defogging. *IEEE transactions on medical imaging*, 36(10):2021–2030, 2017.

S. G. Narasimhan and S. K. Nayar. Vision and the atmosphere. *International Journal of Computer Vision*, 48(3):233–254, 2002.

Augustus Odena, Vincent Dumoulin, and Chris Olah. Deconvolution and checkerboard artifacts. *Distill*, 1(10):e3, 2016.

Liangqiong Qu, Jiandong Tian, Shengfeng He, Yandong Tang, and Rynson WH Lau. Deshadownet: A multi-context embedding deep network for shadow removal. In *International Conference on Computer Vision and Pattern Recognition (CVPR)*, volume 1, page 3, 2017.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention (MICCAI)*, pages 234–241. Springer, 2015.

Jake Snell, Karl Ridgeway, Renjie Liao, Brett D Roads, Michael C Mozer, and Richard S Zemel. Learning to generate images with perceptual similarity metrics. In *International Conference on Image Processing (ICIP)*, pages 4277–4281. IEEE, 2017.

Danail Stoyanov. Surgical vision. *Annals of biomedical engineering*, 40(2):332–345, 2012.

K. Tchakaa, V. M. Pawara, and D. Stoyanova. Chromaticity based smoke removal in endoscopic images. In *Proc. of SPIE Vol*, volume 10133, pages 101331M–1, 2017.

Chaoyue Wang, Chang Xu, Chaohui Wang, and Dacheng Tao. Perceptual adversarial networks for image-to-image transformation. *IEEE Transactions on Image Processing*, 27(8):4066–4079, 2018a.

Congcong Wang, Faouzi Alaya Cheikh, Mounir Kaaniche, Azeddine Beghdadi, and Ole Jacob Elle. Variational based smoke removal in laparoscopic images. *Biomedical engineering online*, 17(1):139, 2018b.

Congcong Wang, Faouzi Alaya Cheikh, Mounir Kaaniche, and Ole Jakob Elle. Liver surface reconstruction for image guided surgery. In *Medical Imaging: Image-Guided Procedures, Robotic Interventions, and Modeling*, volume 10576, page 105762H. International Society for Optics and Photonics, 2018c.

Congcong Wang, Ahmed Kedir Mohammed, Faouzi Alaya Cheikh, Azeddine Beghdadi, and Ole Jakob Elle. Multiscale deep desmoking for laparoscopic surgery. In *Medical Imaging: Image processing*. International Society for Optics and Photonics, 2019.

Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. High-resolution image synthesis and semantic manipulation with conditional gans. *arXiv preprint arXiv:1711.11585*, 2017.

Ting-Chun Wang, Ming-Yu Liu, Jun-Yan Zhu, Guilin Liu, Andrew Tao, Jan Kautz, and Bryan Catanzaro. Video-to-video synthesis. *arXiv preprint arXiv:1808.06601*, 2018d.

Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE transactions on image processing*, 13(4):600–612, 2004.

Lin Zhang, Lei Zhang, Xuanqin Mou, and David Zhang. A comprehensive evaluation of full reference image quality assessment algorithms. In *International Conference on Image Processing (ICIP)*, pages 1477–1480. IEEE, 2012.

Hang Zhao, Orazio Gallo, Iuri Frosio, and Jan Kautz. Loss functions for image restoration with neural networks. *IEEE Transactions on Computational Imaging*, 3(1):47–57, 2017.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. *arXiv preprint*, 2017.