# Can Image Enhancement be Beneficial to Find Smoke Images in Laparoscopic Surgery?

*Congcong Wang* [1,*], *Vivek Sharma* [2,*], *Yu Fan*[1], *Faouzi Alaya Cheikh* [1], *Azeddine Beghdadi* [3], *Ole Jacob Elle* [4,5], *and Rainer Stiefelhagen* [2]

[1] *Norwegian University of Science & Technology, Norway.* [2] *Karlsruhe Institute of Technology, Germany.* [3] *University Paris 13, France.* [4] *Oslo University Hospital, Norway.* [5] *University of Oslo, Norway.*

{congcong.wang,yu.fan,faouzi.cheikh}@ntnu.no, {vivek.sharma,rainer.stiefelhagen}@kit.edu, oelle@ous-hf.no, azeddine.beghdadi@univ-paris13.fr

## Abstract

*Laparoscopic surgery has a limited field of view. Laser ablation in a laproscopic surgery causes smoke, which inevitably influences the surgeon's visibility. Therefore, it is of vital importance to remove the smoke, such that a clear visualization is possible. In order to employ a desmoking technique, one needs to know beforehand if the image contains smoke or not, to this date, there exists no accurate method that could classify the smoke/non-smoke images completely. In this work, we propose a new enhancement method which enhances the informative details in the RGB images for discrimination of smoke/non-smoke images. Our proposed method utilizes weighted least squares optimization framework (WLS). For feature extraction, we use statistical features based on bivariate histogram distribution of gradient magnitude (GM) and Laplacian of Gaussian (LoG). We then train a SVM classifier with binary smoke/non-smoke classification task. We demonstrate the effectiveness of our method on Cholec80 dataset. Experiments using our proposed enhancement method show promising results with improvements of 4% in accuracy and 4% in F1-Score over the baseline performance of RGB images. In addition, our approach improves over the saturation histogram based classification methodologies Saturation Analysis (SAN) and Saturation Peak Analysis (SPA) by 1/5% and 1/6% in accuracy/F1-Score metrics.*

*We can employ our enhancement method in replacement of RGB images for classifier training e.g., CNN architectures, which in turn can lead to more accurate classification. Code will be released for public use.*

## KEYWORDS

Image Enhancement, Weighted Least Squares Framework (WLS), Smoke/Non-Smoke Image Classification.

## INTRODUCTION

Over the last decade, we have seen an increase in the number of laparoscopic surgeries [1]. During the surgery, such as in cavitary treatment, laser ablation causes smoke [2] which significantly degrades the perceptual quality of the images which inevitably influences the surgeon's visibility, further it also influences the performance of computer vision based navigation systems [3]. Moreover, surgical smoke is composed of chemical, physical or biological particles, which may be harmful for surgeons and patients [4–6]. Therefore, it is of vital importance to remove the smoke by computer vision algorithms [7] and by smoke evacuation techniques [8,9]. In order to employ a desmoking technique, as a prior knowledge it is essential to know if the image contains smoke or not. In this work, we propose a method to enhance the images for better classification of smoke and non-smoke images. Our goal is to enhance the images, such that the extracted features from the enhanced images are informative for discrimination that can lead to improved smoke/non-smoke image classification. Note that our goal is not to enhance the images for visual pleasantness of observers' perception, but rather enhance the images features for improved classification.

Our work is inspired from [10,11]. Sharma *et al.* in [10] enhance the visible (RGB) images using near-infrared (NIR) counterparts and show improvement in the image feature quality for biometric verification tasks, further in [11], Sharma *et al.* emulates several image enhancement methods in convolutional neural networks for an accurate image classification. We have a similar goal, though our work differs substantially in technical approach and the application scope. Specifically, we utilize weighted least squares optimization framework (WLS) [12] to decompose an image to fine and coarse enhanced images, and then combine them in a more meaningful way such that the combined image have better image features for our classification task.

Our proposed approach is evaluated on Cholec80 dataset for smoke/non-smoke image classification. We experimentally show that our proposed method consistently improve the classification performance over the baseline RGB images, popular state-of-the-art enhancement methods, and the saturation histogram based classification methodologies Saturation Analysis (SAN) and Saturation Peak Analysis (SPA).

The remainder of this paper is structured as follows. First, we review the related work on image enhancement and smoke detection methods. Next, we describe our proposed method, and discuss the experimental results. Finally, the conclusions are drawn.

## RELATED WORK

**Image Enhancement.** Image enhancement or filtering techniques enhance the contrast, boost the image details, and produce more vivid colors, and at the same time removes the effects of blur, noise, and compression artifacts. Examples of

---

* Congcong Wang and Vivek Sharma contributed equally to this work and listed in alphabetical order.

such filtering methods include weighted least squares (WLS) [12], bilateral filtering [13], image sharpening, guided filtering [14], BFWLS_AVG [10] and more. Filtering using (1) RGB and (2) RGB-NIR images are used for several applications in computer vision and computational photography applications, such as to improve the contrast of the haze-degraded color images [15]; tone mapping and detail enhancement [12]; denoising in RGB videos [16] and images [17]; multi-modal medical image fusion from MRI-CT [18]; illumination transfer from reference to target images [19]; feature matching [20]; object recognition and image classification tasks [11]; biometrics verification tasks [10]; and the list goes on. To the best of our knowledge, our work is the first to show that the image enhancement can be beneficial for smoke/non-smoke image classification. We compare against various of these filtering methods discussed above in our experimental section.

**Smoke Detection.** Considerable progress has been seen in the development of "*in the wild*" video smoke detection techniques over the last decade [21]. The traditional detection methods exploit smoke features, which are intensity, color, motion, and texture attributes. These features are used to train classifiers for smoke region detection [22] or frame detection [23, 24]. It has been found that smoke can reduce the sharpness of edges in order to mitigate over it, several descriptors have been utilized for feature extraction, such as wavelet coefficients [25], local binary pattern (LBP) [26,27], textural features are estimated from the smoke region only [28]. Further, in [29], the authors propose to separate smoke and background from a single image by the dual-dictionary approach, and the estimated sparse coefficients are used as features for smoke frames detection.

In the medical imaging community, researchers have shown that the formation of smoke and the lighting condition for laparoscopic surgery images are very different, which limits the "*in the wild*" smoke detection methodologies applicability in medical domain [9, 30]. Recently, Chou *et al.* [31] explore temporal differences information that is motion blur and block analysis between current and previous frames as features for smoke detection. Loukas *et al.* [32], propose a method to detect electrocautery smoke for surgical events retrieval. They extract different features from optical flow estimated by Kanade-Lucas-Tomasi (KLT) algorithm [33], and then a support vector machine (SVM) is trained to classify each shot. Leibetseder *et al.* in [9, 30] propose saturation histogram based classification methodologies Saturation Peak Analysis (SPA) and Saturation Analysis (SAN) and deep learning (DL) techniques. In [30], Leibetseder *et al.* train three variants of convolutional neural network (CNN) models: GoogleLeNet [34] trained with RGB images (GLN RGB) and saturation channels only (GLN SAT), and a modified AlexNet [35] trained on RGB images (ALEX RGB). The performance of their CNN models on Cholec80 [36] dataset are very similar to that of SAN or SPA. While DL shows its promising performance on smoke classification [9,30], to the best of our knowledge there is no work that does image enhancement for smoke classification. Therefore, we limit to comparisons with the non deep learning based methods only in our experimental section. Moreover, one can consistently improve the classification performance by employing our enhancement method and then train the CNN architectures, following the ideas proposed in Sharma *et al.* [11].
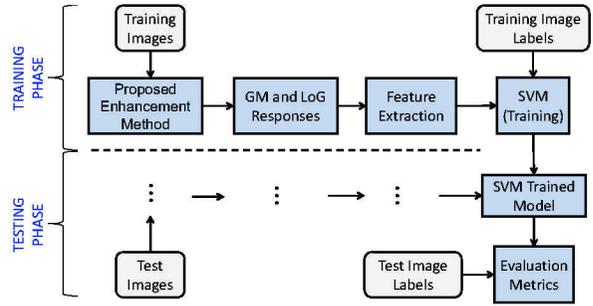


Figure 1: Pipeline of the proposed method.

## PROPOSED METHOD

In this section, we illustrate our proposed approach, starting with the proposed enhancement method, then the feature extraction method, and finally classifier training. Figure 1 shows the schematic layout of our framework.

**Proposed Enhancement Method.** We denote our WLS-based filtering method as FC. To enhance the visible RGB image, we first transform color space from RGB to YCbCr space that is a luminance-chrominance color space. Our method operates on the luminance component [37] such that to modify the overall contrast and sharpness of the image without sensibly affecting the color. Where the chrominance is simply re-combined in the final enhanced image.

In our work, we employ an edge-preserving filter, weighted least squares optimization framework (WLS) [12]. The WLS is a non-linear method that captures image details at a variety of scales via multi-scale decompositions. The WLS helps to find an approximate enhanced image $g^{Filtered}$ that is close to the input image $g$, and also at the same time, is smooth along significant gradients, thus resulting to sharper preserved edges. Formally, it is defined as:

$$g^{Filtered} = F_\lambda(g) = (I + \lambda L_g)^{-1} g \qquad (1)$$

where $L_g = D_x^T A_x D_x + D_y^T A_y D_y$ with $D_x$ and $D_y$ are discrete differentiation operators. $A_x$ and $A_y$ contain the smoothness weights, the smoothness requirement is enforced in a spatially varying manner which depend on $g$. $\lambda$ is the balance factor that maintains a balance between the data term and the smoothness term. Increasing $\lambda$ value produces progressively smoother images.

Given an input image, the WLS filter decomposes an image into base and detail layers. The detail layer is simply obtained by subtracting the base layer from the the original image. The base layer comprises of low frequency contents with general appearance of the image over smooth areas, while the detail layer comprises of high frequency contents with sharp edges.

In our method, we apply WLS-based two-level decomposition of the luminance component of an RGB image for extraction of fine and coarse enhanced sharp images. We retain, for each pixel an average value between the fine and coarse detail layers (Step 1). We chose $\lambda_1 = 0.125$ (WLS$_1$), $\lambda_2 = 0.5$ (WLS$_2$) for WLS in our experiments. The fusion criteria is based on the following observations: the WLS filter is very good at preserving fine and coarse details at arbitrary scales. Taking an average of
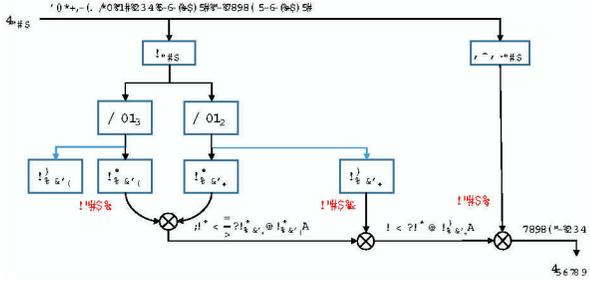
Figure 2: Proposed enhancement method.

two, allows to retain the informative content from both, thus allowing to preserve the image structures and also moderately boost the image details. This fusion criteria is denoted as FC_AVG. In addition, we also tried to retain the maximum values between the two as a fusion criterion. We denote this fusion criteria as FC_MAX. Finally, we combine the new fused detail layer with the base layer of RGB image obtained using $WLS_1$ (Step 2), we consider $WLS_1$, as smaller $\lambda$ values are better for the base layer because of less smoothing. And which is then re-combined with the chrominance of RGB image and to reconstruct the final enhanced image (Step 3). Figure 2 illustrates our proposed method.

**Feature Extraction.** In our work, we exploit Xue *et al.*'s method [38] for feature extraction. Xue *et al.* utilize gradient magnitude (GM) and Laplacian of Gaussian (LoG) maps to describe the structural information for image perceptual quality assessment. Their method is driven by the image statistics, and exploits the histogram information, which is perfect for low-level vision tasks, such as ours. Motivated by this observation, we employ GM and LoG features to represent the local spatial contrast information in images.

Formally, gradient magnitude (GM) is defined as:

$$GM = \sqrt{(I \otimes \frac{\partial G}{\partial x})^2 + (I \otimes \frac{\partial G}{\partial y})^2}, \qquad (2)$$

where $I$ is the gray scale image, $\otimes$ denotes the convolution operation. $\frac{\partial G}{\partial x}$ and $\frac{\partial G}{\partial y}$ denote the Gaussian partial derivative along $x$ (horizontal) and $y$ (vertical) directions, respectively. They are computed as:

$$\frac{\partial G(x,y,\sigma)}{\partial d} = -\frac{1}{2\pi\sigma^2}\frac{d}{\sigma^2}e^{-\frac{x^2+y^2}{2\sigma^2}}, \qquad (3)$$

where $d \in \{x,y\}$, and $\sigma$ is a scalar parameter. And the Laplacian of Gaussian (LoG) is defined as:

$$LoG = I \otimes h_{LoG} \qquad (4)$$

$$\begin{aligned} h_{LoG}(x,y,\sigma) &= \frac{\partial^2 G(x,y,\sigma)}{\partial^2 x} + \frac{\partial^2 G(x,y,\sigma)}{\partial^2 y} \\ &= -\frac{1}{\pi\sigma^4}(1 - \frac{x^2+y^2}{2\sigma^2})e^{-\frac{x^2+y^2}{2\sigma^2}}. \end{aligned} \qquad (5)$$

Subsequently, a joint adaptive normalization step is applied to adjust the image statistics of the GM and LoG maps, in order to reduce its dependency on local image content. Finally, a bivariate histogram based feature vector is computed from the marginal probability functions and independency distributions of the normalized GM and LoG maps, resulting to a fixed-size feature representation vector of 40 dimensions for an image. We used the default parameters for feature extraction, more details can be found in [38].

**Classifier.** In our work, we use linear SVM [39] [1] for a binary classification task in which the objective is to predict the class $y \in \{0,1\}$ that is smoke/non-smoke image classification. We use linear kernel function with C=10$k$ to train/test SVM classifier with the extracted features.

## EXPERIMENTS

In this section, we demonstrate the use of our proposed enhancement method on a challenging smoke/non-smoke classification dataset [30,36]. We first introduce the dataset, followed by a thorough analysis of the proposed method. The analysis includes, comparison of our proposed method with baseline RGB images, other popular enhancement methods, and ending with a comparison to saturation histogram based classification methodologies.
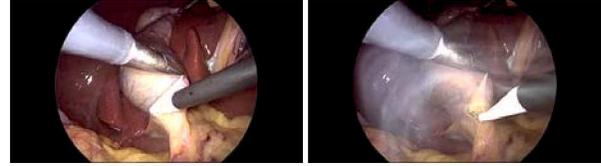


(a)       (b)

Figure 3: Example images from Cholec80 dataset: (a) a smoke free image, (b) image with smoke, covering the surgeon's field of view for visualisation

**Dataset.** We conduct experiments on Cholec80 dataset [36] which contains 80 videos of cholecystectomy surgeries manually labeled with smoke/non-smoke image sequence by [30] [2]. The dataset in overall contains approximately 100K annotated images, in particular between 200-1300 images of smoke/non-smoke in each video. In our experiments, we use a subset of the dataset with three videos for training and nine videos for testing. The videos are randomly chosen. We extract JPEG images, and resize them to a resolution of $427 \times 240$px from the original resolution of $854 \times 480$px, to fasten the image enhancement. In particular, we use 4,381 images obtained from video$\{1, 31, 34\}$ for training, and 10,653 images obtained from video$\{2, 22, 10, 40, 59, 64, 65, 71, 80\}$ for testing. In Figure 3, we show some examples of smoke/non-smoke images obtained from Cholec80 dataset. Also, note that each video represents a unique person in this dataset.

**Evaluation Metrics.** We use accuracy and F1-Score measures computed from a confusion matrix between the predicted labels and the actual ground-truth labels, as the metrics to evaluate the quality of classification.

---

[1] https://www.csie.ntu.edu.tw/~cjlin/libsvm/
[2] http://www.itec.aau.at/ftp/datasets/Smoke_cholec80

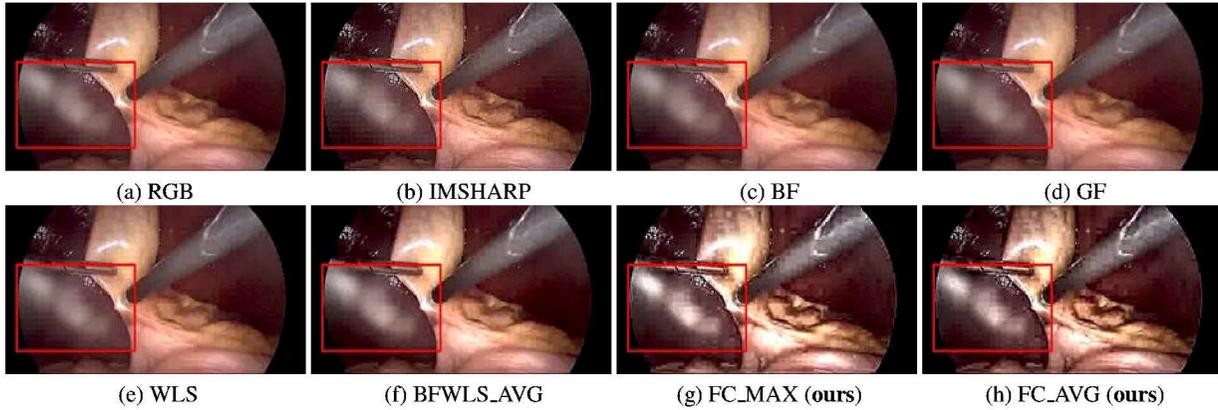|  |  |  |  |
|---|---|---|---|
| (a) RGB | (b) IMSHARP | (c) BF | (d) GF |
| (e) WLS | (f) BFWLS_AVG | (g) FC_MAX (**ours**) | (h) FC_AVG (**ours**) |

Figure 4: Visual comparison of FC_AVG and FC_MAX against other enhancement methods. We can clearly see that the smoke part becomes more perceptually visible after image enhancement. Best viewed in color.

**Comparison with other enhancement methods.** We compare the proposed enhancement method with the baseline RGB images, and popular state-of-the-art enhancement approaches: (1) guided filtering (GF) [14], (2) bilateral filtering (BF) [13,40], (3) image sharpening filter (IMSHARP), (4) WLS [12], (5) fused BF and WLS filter (BFWLS_AVG) [10]. In BFWLS_AVG, Sharma *et al.* [10] apply BF and WLS filters on the NIR channel, and retain an average of two for each pixel, similar to [10], instead we employ their technique in the RGB image. For a fair comparison, we compare all the methods under the same evaluation protocol discussed above. For the evaluation, we use the same parameters for feature extraction and classifier training for all methods. We did not optimize the parameters for the BF, GF, WLS, IMSHARP, and BFWLS_AVG and used the default parameters for each method. The source code for fast BF [3], WLS [4] and BFWLS_AVG [5] are publicly available, and others are available in the Matlab framework. For comprehensive discussion of different methods, we refer the reader to [10,12,14,40].

In Table 1, we quantitatively evaluate the accuracy and F1-Score of our proposed method and other methods. We can clearly observe that FC_AVG performs the best among all methods. Enhancing an image using FC_AVG takes approximately 0.32 seconds. Note that, FC_AVG improves over the baseline performance of RGB images by 4% in accuracy and 4% in F1-Score, giving higher-quality features to learn from, which in turn lead to more accurate classification. We believe our work opens many possibilities for further exploration for its usage in other tasks too. Further the performance gap of FC_AVG is 4/5% better than FC_MAX in accuracy/F1-Score measures. Our method FC_AVG consistently performs better than all other methods: IMSHARP, BF, WLS, and GF enhancement methods. In addition, our methods also performs better than BFWLS_AVG, although we agree that BFWLS_AVG is meant for RGB-NIR fusion and not RGB image enhancement, that may be the reason why it underperforms to our method. Figure 5 (a) shows the ROC curve for all the methods.

In Figure 4, we compare enhancement methods for an example image obtained from video1. We can observe that FC_AVG has improved features for the smoke-part (highlighted by red rect-

angles), in addition to reduced specular intensity when compared to FC_MAX. We believe this plays an important role in discrimination.

| Method | Accuracy | F1-Score |
|---|---|---|
| RGB | 0.60 | 0.60 |
| IMSHARP | 0.58 | 0.58 |
| BF [40] | 0.60 | 0.59 |
| GF [14] | 0.60 | 0.59 |
| WLS [12] | 0.60 | 0.59 |
| BFWLS_AVG [10] | 0.57 | 0.56 |
| FC_MAX (**Ours**) | 0.60 | 0.59 |
| FC_AVG (**Ours**) | **0.64** | **0.64** |

Table 1: Comparison with the baseline RGB images and other enhancement methods.

**Comparison with the saturation histogram based classification methodologies.** We compare FC_AVG with the saturation histogram based classification methodologies Saturation Analysis (SAN) and Saturation Peak Analysis (SPA) [9,30]. SAN and SPA codes are provided by [9,30] [6]. SAN and SPA take advantage of the saturation channel of the HSV color space. In [9,30], the authors show that the histogram bin curve of saturation channel is strongly correlated with the presence of smoke, as smoke images contain more low-saturation pixels and that is helpful for classifying smoke/non-smoke images. In their method, a threshold $t_c$ is set empirically. For SAN method, smoke image is recognized if majority of the bin values are below $t_c$. For SPA method, smoke image is classified by the number of peaks found below and above $t_c$. We use the default parameters suggested by Leibetseder *et al.* [9,30] for both SAN and SPA where $t_c$ is set to 0.35.
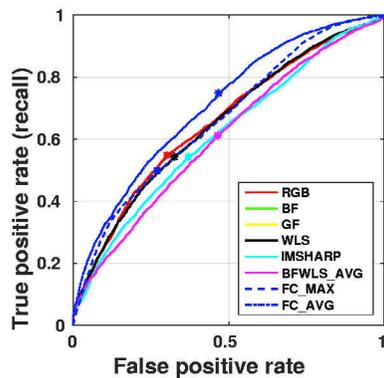
It is evident from Table 2 that FC_AVG shows improved performance over the saturation histogram based classification methodologies: SAN and SPA. Precisely, FC_AVG is 1/5% and 1/6% better than the SAN and SPA methods in accuracy/F1-Score
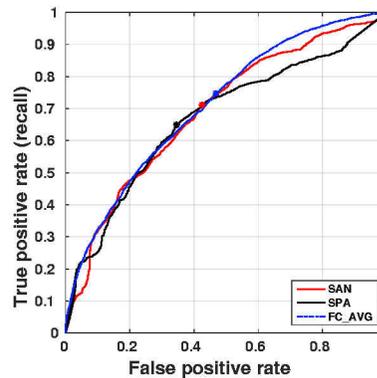
---

[3]http://people.csail.mit.edu/sparis/bf/
[4]http://www.cs.huji.ac.il/~danix/epd/
[5]https://vivoutlaw.github.io/CIC_RGB_NIR_Codes.zip

[6]https://github.com/amplejoe/SaturationPeakAnalysis

(a) Comparison with other enhancement methods.

(b) Comparison with SAN and SPA.

Figure 5: The ROC curves for the Cholec80 dataset using GM-LoG features [38] for smoke/non-smoke classification task. * denotes the EER when the false accept rate is equal to the false reject rate. Best viewed in color.

metrics. Further in Figure 5 (b), we show the ROC curve for FC_AVG, SAN and SPA classification methodologies.

| Method | Accuracy | F1-Score |
|---|---|---|
| SPA [9] | 0.63 | 0.58 |
| SAN [30] | 0.63 | 0.59 |
| FC_AVG (Ours) | **0.64** | **0.64** |

Table 2: Comparison with the saturation histogram based classification methodologies Saturation Analysis (SAN) and Saturation Peak Analysis (SPA)

## CONCLUSION

We present a method to enhance RGB images using weighted least squares filter. Our method successfully enhances the informative features in the images for discrimination of smoke/non-smoke images. We demonstrate our proposed method on Cholec80 dataset. Our proposed method obtains significant gains over the baseline RGB images and other enhancement methods. Our method also outperforms the saturation histogram based classification methodologies. Our enhancement method can be used to perform explicit enhancement of image texture and structure features for other tasks too which in turn can lead to more improved performance.

In future work, we plan to further investigate our method on the complete Cholec80 dataset, and also train a CNN architecture with our enhanced RGB images.

## References

[1] C. Tsui, R. Klein, and M. Garabrant, "Minimally invasive surgery: national trends in adoption and future directions for hospital strategy," *Surgical endoscopy*, vol. 27, no. 7, pp. 2253–2257, 2013.

[2] N. Lawrentschuk, N. E. Fleshner, and D. M. Bolton, "Laparoscopic lens fogging: a review of etiology and methods to maintain a clear visual field," *Journal of endourology*, vol. 24, no. 6, pp. 905–913, 2010.

[3] L. Maier-Hein, A. Groch, A. Bartoli, S. Bodenstedt, G. Boissonnat, P.-L. Chang, N. Clancy, D. S. Elson, S. Haase, E. Heim *et al.*, "Comparative validation of single-shot optical techniques for laparoscopic 3-d surface reconstruction," *TMI*, vol. 33, no. 10, pp. 1913–1930, 2014.

[4] O. Al Sahaf, I. Vega-Carrascal, F. Cunningham, J. McGrath, and F. Bloomfield, "Chemical composition of smoke produced by high-frequency electrosurgery," *Irish journal of medical science*, vol. 176, no. 3, pp. 229–232, 2007.

[5] S. H. Choi, T. G. Kwon, S. K. Chung, and T.-H. Kim, "Surgical smoke may be a biohazard to surgeons performing laparoscopic surgery," *Surgical endoscopy*, vol. 28, no. 8, pp. 2374–2380, 2014.

[6] M. Dobrogowski, W. Wesołowski, M. Kucharska, A. Sapota, and L. Pomorski, "Chemical composition of surgical smoke formed in the abdominal cavity during laparoscopic cholecystectomy–assessment of the risk to the patient," *Journal of occupational medicine and environmental health*, vol. 27, no. 2, pp. 314–325, 2014.

[7] C. Wang, F. A. Cheikh, M. Kaaniche, and O. J. Elle, "A smoke removal method for laparoscopic images," *arXiv:1803.08410v1*, 2018.

[8] H. Takahashi, M. Yamasaki, M. Hirota, Y. Miyazaki, J. H. Moon, Y. Souma, M. Mori, Y. Doki, and K. Nakajima, "Automatic smoke evacuation in laparoscopic surgery: a simplified method for objective evaluation," *Surgical endoscopy*, vol. 27, no. 8, pp. 2980–2987, 2013.

[9] A. Leibetseder, M. J. Primus, S. Petscharnig, and K. Schoeffmann, "Image-based smoke detection in laparoscopic videos," in *CARE*. Springer, 2017, pp. 70–87.

[10] V. Sharma, J. Y. Hardeberg, and S. George, "Rgb-nir image enhancement by fusing bilateral and weighted least squares filters," *JIST*, vol. 61, no. 4, pp. 40 409–1, 2017.

[11] V. Sharma, A. Diba, D. Neven, M. S. Brown, L. Van Gool, and R. Stiefelhagen, "Classification driven dynamic image enhancement," in *CVPR*, 2018.

[12] Z. Farbman, R. Fattal, D. Lischinski, and R. Szeliski, "Edge-preserving decompositions for multi-scale tone and detail manipulation," in *TOG*, vol. 27, no. 3. ACM, 2008, p. 67.

[13] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *ICCV*. IEEE, 1998, pp. 839–846.

[14] K. He, J. Sun, and X. Tang, "Guided image filtering," in

ECCV. Springer, 2010, pp. 1–14.

[15] L. Schaul, C. Fredembach, and S. Süsstrunk, "Color image dehazing using the near-infrared," in *ICIP*. IEEE, 2009, pp. 1629–1632.

[16] E. P. Bennett, J. L. Mason, and L. McMillan, "Multispectral video fusion," in *ACM SIGGRAPH*. ACM, 2006, p. 123.

[17] S. Zhuo, X. Zhang, X. Miao, and T. Sim, "Enhancing low light images using near infrared flash images," in *ICIP*. IEEE, 2010, pp. 2537–2540.

[18] W. Li, Z. Zhao, J. Du, and Y. Wang, "Edge-preserve filter image enhancement with application to medical image fusion," *Journal of Medical Imaging and Health Informatics*, vol. 7, no. 1, pp. 16–24, 2017.

[19] X. Chen, M. Chen, X. Jin, and Q. Zhao, "Face illumination transfer through edge-preserving filters," in *CVPR*. IEEE, 2011, pp. 281–287.

[20] V. Sharma and L. Van Gool, "Does v-nir based image enhancement come with better features?" *arXiv:1608.06521*, 2016.

[21] Y. Luo, L. Zhao, P. Liu, and D. Huang, "Fire smoke detection algorithm based on motion characteristic and convolutional neural networks," *Multimedia Tools and Applications*, pp. 1–18, 2017.

[22] I. Kolesov, P. Karasev, A. Tannenbaum, and E. Haber, "Fire and smoke detection in video with optimal mass transport based optical flow and neural networks," in *ICIP*. IEEE, 2010, pp. 761–764.

[23] Y. Chunyu, F. Jun, W. Jinjun, and Z. Yongming, "Video fire smoke detection using motion and color features," *Fire technology*, vol. 46, no. 3, pp. 651–663, 2010.

[24] S. Calderara, P. Piccinini, and R. Cucchiara, "Vision based smoke detection system using image energy and color information," *MVA*, vol. 22, no. 4, pp. 705–719, 2011.

[25] B. U. Töreyin, Y. Dedeoğlu, and A. E. Cetin, "Wavelet based real-time smoke detection in video," in *EUSIPCO*. IEEE, 2005, pp. 1–4.

[26] H. Tian, W. Li, P. Ogunbona, D. T. Nguyen, and C. Zhan, "Smoke detection in videos using non-redundant local binary pattern-based features," in *MMSP*. IEEE, 2011, pp. 1–4.

[27] F. Yuan, "Video-based smoke detection with histogram sequence of lbp and lbpv pyramids," *Fire safety journal*, vol. 46, no. 3, pp. 132–139, 2011.

[28] H. Tian, W. Li, L. Wang, and P. Ogunbona, "Smoke detection in video: an image separation approach," *IJCV*, vol. 106, no. 2, pp. 192–209, 2014.

[29] H. Tian, W. Li, P. O. Ogunbona, and L. Wang, "Detection and separation of smoke from single image frames," *TIP*, vol. 27, no. 3, pp. 1164–1177, 2018.

[30] A. Leibetseder, M. J. Primus, S. Petscharnig, and K. Schoefmann, "Real-time image-based smoke detection in endoscopic videos," in *ACM MM, Thematic Workshops*. ACM, 2017, pp. 296–304.

[31] C.-R. Chou and M.-C. Liu, "System and method for smoke detection during anatomical surgery," U.S. Patent No. 9,805,472, issued October 31, 2017.

[32] C. Loukas and E. Georgiou, "Smoke detection in endoscopic surgery videos: a first step towards retrieval of semantic events," *Journal of Medical Robotics and Computer Assisted Surgery*, vol. 11, no. 1, pp. 80–94, 2015.

[33] J. Shi, "Good features to track," in *CVPR*. IEEE, 1994, pp. 593–600.

[34] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich *et al.*, "Going deeper with convolutions," in *CVPR*. IEEE, 2015.

[35] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *NIPS*, 2012, pp. 1097–1105.

[36] 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," *TMI*, vol. 36, no. 1, pp. 86–97, 2017.

[37] C. Fredembach and S. Süsstrunk, "Colouring the near-infrared," in *CIC*, vol. 2008, no. 1, 2008, pp. 176–182.

[38] W. Xue, X. Mou, L. Zhang, A. C. Bovik, and X. Feng, "Blind image quality assessment using joint statistics of gradient magnitude and laplacian features," *TIP*, vol. 23, no. 11, pp. 4850–4862, 2014.

[39] C.-C. Chang and C.-J. Lin, "Libsvm: a library for support vector machines," *TIST*, vol. 2, no. 3, p. 27, 2011.

[40] F. Durand and J. Dorsey, "Fast bilateral filtering for the display of high-dynamic-range images," in *TOG*, vol. 21, no. 3. ACM, 2002, pp. 257–266.