



An automatic framework for endoscopic image restoration and enhancement

Muhammad Asif¹ · Lei Chen¹ · Hong Song¹ · Jian Yang² · Alejandro F. Frangi³

Accepted: 1 September 2020 / Published online: 22 October 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Despite its success in the field of minimally invasive surgery, endoscopy image analysis remains challenging due to limited image settings and control conditions. The low resolution and existence of large number of reflections in endoscopy images are the major problems in the automatic detection of objects. To address these issues, we presented a novel framework based on the convolutional neural networks. The proposed approach consists of three major parts. First, a deep learning (DL)-based image evaluation method is used to classify the input images into two groups, namely, specular highlights and weakly illuminated groups. Second, the specular highlight is detected using the DL-based method, and the reflected areas are recovered through a patch-based restoration operation. Lastly, gamma correction with optimized reflectance and illumination estimation is adopted to enhance the weakly illuminated images. The proposed method is compared against the existing ones, and the experimental results demonstrate that the former outperforms the latter in terms of subjective and objective assessments. This finding indicates that the proposed approach can serve as a potential tool for improving the quality of the endoscopy images used to examine internal body organs.

Keywords Endoscopy · Image restoration · Image enhancement · Specular highlights · Weak illuminance

1 Introduction

The development of endoscopy image processing technology has received increasing attention due to the widespread use of minimally invasive treatments [1]. These computer vision-based technologies provide an observable endoscopy view of the internal organs to help physicians make highly accurate diagnosis [2]. Despite the great progress of natural image processing techniques, such as image restoration and enhancement, only few methods can be applied to

endoscopy scenes due to the unique acquisition processes and imaging environment.

Two common situations affect the quality of endoscopy images. One is the bright spots produced by the light reflections in the smooth organs' surface (Fig. 1). These spots, which are caused by specular reflections, can result in the loss of image texture and color information and leads to significant discontinuities in endoscopy imaging and affects the physician's vision, which are not conducive to diagnosis tasks [3, 4]. The automatic detection and restoration of specular reflections are popular processes, and many researchers promote this stream by proposing different kinds of methods. Alsaleh et al. [5] proposed an adaptive threshold-based method to capture the specular reflection regions. After detecting the required regions, they used a mask-based approach to correct the bright spots. Similarly, Guo et al. [6] used a threshold-based algorithm to detect the reflection regions and developed an improved energy function to recover such regions with improved visual quality. Hsia et al. [7] designed a mask-based method to extract the textural information in endoscopy images and subsequently restored the specular reflections using the extracted features. Saint-Pierre et al. [8] argued that specular highlights appear as a convex

✉ Hong Song
anniesun@bit.edu.cn

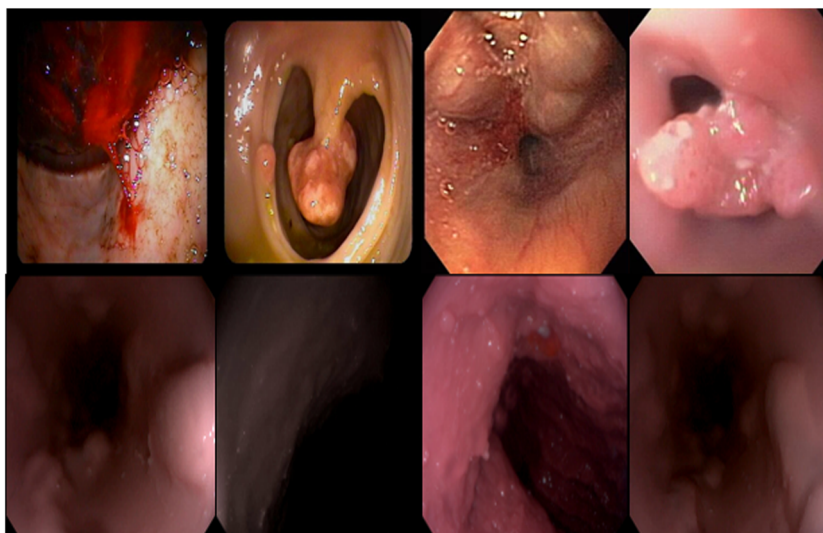
✉ Jian Yang
jyang@bit.edu.cn

¹ School of Computer Science and Technology, Beijing Institute of Technology, Beijing, China

² School of Optics and Electronics, Beijing Institute of Technology, Beijing, China

³ School of Computing and School of Medicine, University of Leeds, Leeds, UK

Fig. 1 1st row reflected endoscopy images and 2nd row weak illumination endoscopy images



shape in the pixel histogram. They detected the reflection regions by isolating the peak component in the histogram and extracted the region of interest using the relevant neighbor components, thereby resulting in the acquisition of a mask of the reflection position. They further utilized an inpainting model to correct the reflection areas on the basis of the detected mask information. Meslouhi et al. [9] used a dichromatic reflection model to detect specular highlights and utilized local information, along with a multiresolution inpainting approach, to recover lost color information in reflection areas. Zimmerman-Moreno et al. [10] combined the gradient, saturation, and intensity information to detect reflection areas and designed a cascade detection method, which includes a coarse region detection approach and a probabilistic model for result optimization. The neighborhood color information was used in the restoration process. The abovementioned studies indicate that although the specular reflection can be detected and compensated to some extent (i.e., limited to complex imaging environment), the results still have obvious artifacts.

Another factor that affects the quality of endoscopy images is weak illumination (Fig. 1), which is caused by the absence of extra light illumination inside the body except for the unidirectional light source emitted from the moving capsule. Such a dynamic lighting process easily creates dark areas that affect the surgical environment. Therefore, developing an image enhancement algorithm embedded in the lighting system is advantageous for enhancing the visual effect and surgical accuracy of surgeons. Imtiaz and Khan [11] used a conversion matrix to transform the color image into luminance and chrominance components. They also applied a sigmoid function to the luminance pixels and utilized old texture-based chrominance information to acquire new chrominance components. The new luminance

and chrominance components were then converted into a color image again to highlight the tissue characteristics. Imtiaz and Wahid [12] converted endoscopy images into three spectral images, in which the one with the largest entropy was selected as the benchmark image. By using the neighborhood method, they matched the luminance and textural information to acquire the chromaticity diagram of the original color image. Subsequently, they added the diagram to the benchmark image for color recovery to enhance the detail of some tissues. On the basis of local image information, Li and Meng [13] proposed a contrast diffusion algorithm that can automatically select parameters to enhance endoscopy images, and the experimental results demonstrated the effectiveness of this method. The above studies and the corresponding solutions only focus on one specific problem, which is either correcting the bright spots or highlighting the dark areas in the images.

At present, deep learning (DL) methods are widely used in many computer vision applications, such as image classification [14]; image segmentation [15]; object detection [16]; image reconstruction; speech, face, and text recognition [17–19]; drug discovery [20]; and lip reading [21]. DL promoted the identification of various diseases using X-ray, computed tomography scan, magnetic resonance imaging scan, and endoscopy images in the field of medical image processing and analysis. The detection accuracy mainly depends on the image acquisition devices, which were improved for the field of image interpretation. In addition, DL resolved the image interpretation issue caused by the large amount of learning features that vary from patient to patient. For instance, convolutional neural networks (CNNs) display state-of-the-art performance due to its rapidness and ability to obtain large amounts of learning features from images [22]. DL methods also learn

the abnormal feature arrangements under the presence of unwanted factors in medical endoscopy images. The endoscopy images are prone to the misdetection of polyps due to the presence of overlay information, specular reflection, flat polyps, light over polyps, overexposed area, and low image resolution [23]. The accurate and automatic detection of polyps from endoscopy images requires normal- and high-resolution images. Therefore, removing reflections and enhancing the image resolution can greatly improve the accuracy of polyp detection.

In this study, we proposed an automatic framework that can simultaneously address the issues regarding the two previously mentioned artifacts, namely, light reflection and weak illuminance. The main contributions of this work are as follows. First, an image evaluation algorithm is designed using the DL classification approach. After the evaluation, two groups of images are acquired, namely, specular reflection and weakly illumination. Second, the DL-based bright spot detection method and patch-based restoration model are combined to correct the reflection areas. Third, image enhancement is performed by estimating the reflectance and illumination of the image, followed by gamma correction to improve the illumination component. To the best of our knowledge, the proposed scheme is the first automatic framework that addresses the two problems in endoscopy images. The findings revealed that the proposed method achieved better subjective and objective performances compared with other existing techniques.

2 Materials and methods

Figure 2 shows the flowchart of the proposed framework. First, we classified the input images into two categories on the basis of DL network. The details of this process are discussed in Section 2.1. The image that contains specular reflections is then compensated using a patch-based image restoration model (Section 2.2). The weakly illuminated image is corrected by performing reflectance and illumination estimations (Section 2.3). If the image does not belong to the above two situations, then it is considered as a normal image and no modification is applied.

2.1 Image classification

2.1.1 Image classification material

We used 1000 endoscopy images from the KVASIR database (polyp category) [24], 379 images from CVC colon DB [25] and 116 images from ETIS-Larib Polyp DB [26], and 95 images from the North American Society for Pediatric Gastroenterology, Hepatology, and Nutrition (NASPGHAN) [27] to train, validate, and test the reflected images, respectively. We also collected 90 endoscopy images from various studies and hospitals [1, 5, 9–12], [26]. All datasets used for training, validation, and testing would be confident for conclusive and object results. The collected original images are presented in Fig. 3.

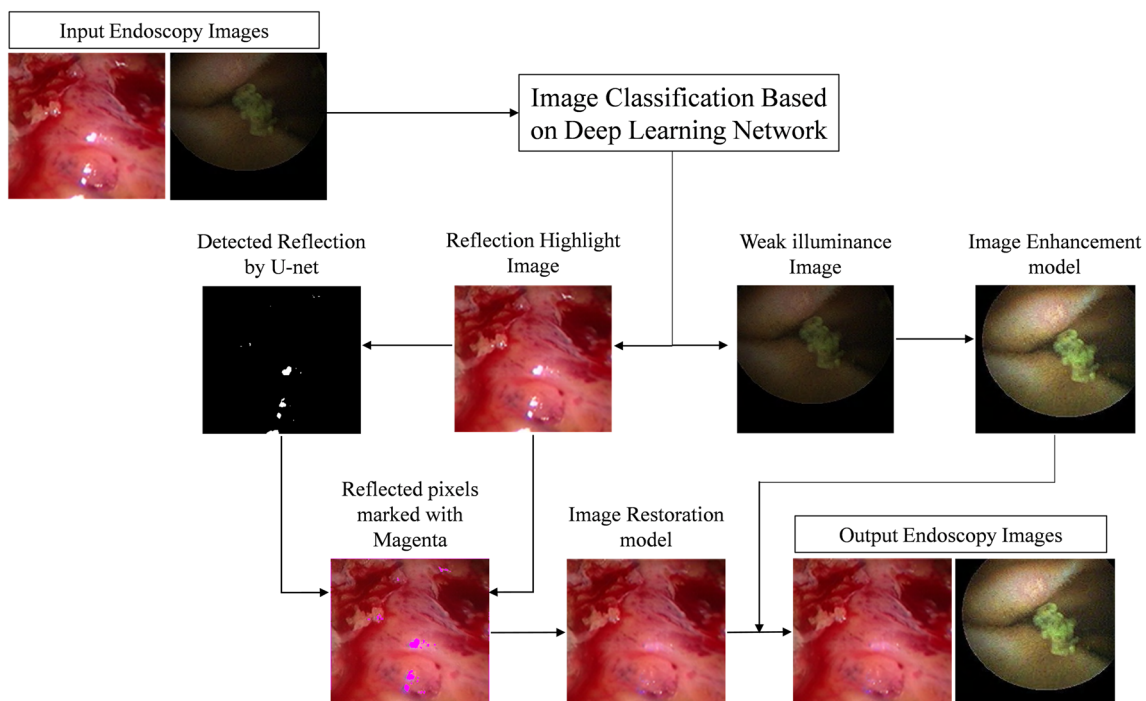


Fig. 2 Flowchart of the proposed framework. It consists of three parts (i) classification network classify reflected and weak illumination endoscopy images and in (ii) reconstruct to detected reflected area from reflected endoscopy images and (iii) enhance to weak illumination endoscopy images

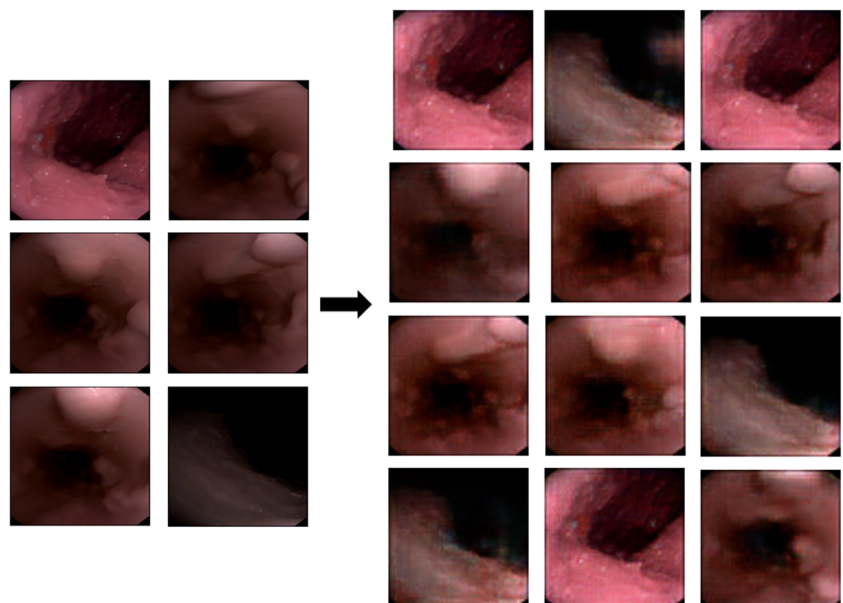


Fig. 3 Selected endoscopy images from Kvasir DB [24], CVC colon DB [25] and ETIL Larib PolypDB [26]

The proposed framework consists of three parts: (i) classification network, which classifies the reflected and weakly illuminated endoscopy images, (ii) reconstruction to detect the reflected area from the reflected endoscopy images, and (iii) enhancement of the weak illuminance of the endoscopy images. To the best of our knowledge, no weak illuminance dataset is available. The 34 weakly

illuminated endoscopy images are collected from Beijing Tongren Hospital, and a generative adversarial network (GAN) is adopted to augment the dataset [28]. The images are used to generate the 250 weakly illuminated endoscopy images shown in Fig. 4. The GAN network takes endoscopy images with sizes of 128×128 as the training input for over 4000 iterations and then generates

Fig. 4 Left: The collected original endoscopy weak illuminance images from Beijing Tongren Hospital. Right: The Generated weak illuminance endoscopy images by using Generative Adversarial Network (GAN)



weakly illuminated endoscopy images with the same size. Figure 4 shows some of the endoscopy images used for GAN training and their corresponding weakly illuminated images.

2.1.2 Image classification based on deep-learning network

The neural network architecture for distinguishing specular reflection and weak illuminance endoscopy images can be viewed in Fig. 5. The architecture comprises eight convolution layers, four max pooling layers, three dense layers, two dropout layers, and one flat layer. The image classification network is trained in 1495 reflected images, in which 1000 images are selected for training and 495 are utilized for validation. Meanwhile, a total of 250 weak illuminance images examples are used to train the network for classifying weak illuminance endoscopy images, which divided into training (200 images) and validation (50 images) respectively. The input image size 200×200 is used for the neural network. Initially, a group of two CNN followed by a Max-pooling layer for down-sampling to the endoscopy images used four times in a sequence. The factor of 2 makes used for image down-sampling that halves the image resolution but the number of features is doubled in each group (32,64,128 and 256). Moreover, the flattening layer is applied for flattening the data for feeding to the fully connected layer. Subsequently, two groups of (fully connected layer + dropout layer) next to the final fully connected layer are used to classify the images into their respective categories. The rectified linear unit activation layer is used in all layers except for the last one, wherein a sigmoid activation function is used. A total of 10,675,745 trainable parameters are used in the entire network. The RMSprop optimizer with a learning rate of 0.0001 and

the binary cross-entropy loss function are utilized to train the network. An image data generator is used for data augmentation, in which 8000 images were used per epoch. Ten epochs with batch sizes of 32 are used, and the NVIDIA GTX 1070 GPU are utilized to train the network.

2.2 Reflection elimination method

To restore a particular region from the reflected endoscopy images, the reflected part of the image must be identified because the reflection detection results facilitate the subsequent restoration work. The reflection elimination method consists of two steps, namely, reflection detection and image restoration.

2.2.1 Reflection detection based on deep learning

The reflection detection calculates the entire reflection in the endoscopy images. DL methods require labels for each input image to determine the corresponding outputs. In this study, we adopted the threshold-based reflection detection approach proposed in [5] to generate the labels. The CVC colon DB images [25] are used to produce their corresponding labels (Fig. 6). In addition, the U-net [29] CNN for medical image segmentation networks is adopted to train the endoscopy images in their corresponding generated reflection masks. The output of the U-net is illustrated in Fig. 6.

2.2.2 Image restoration

The above detection results are used to apply a magenta color to the original images by setting the pixel values of $R(x, y)$, $G(x, y)$, and $B(x, y)$ as 255, 0, and 255,

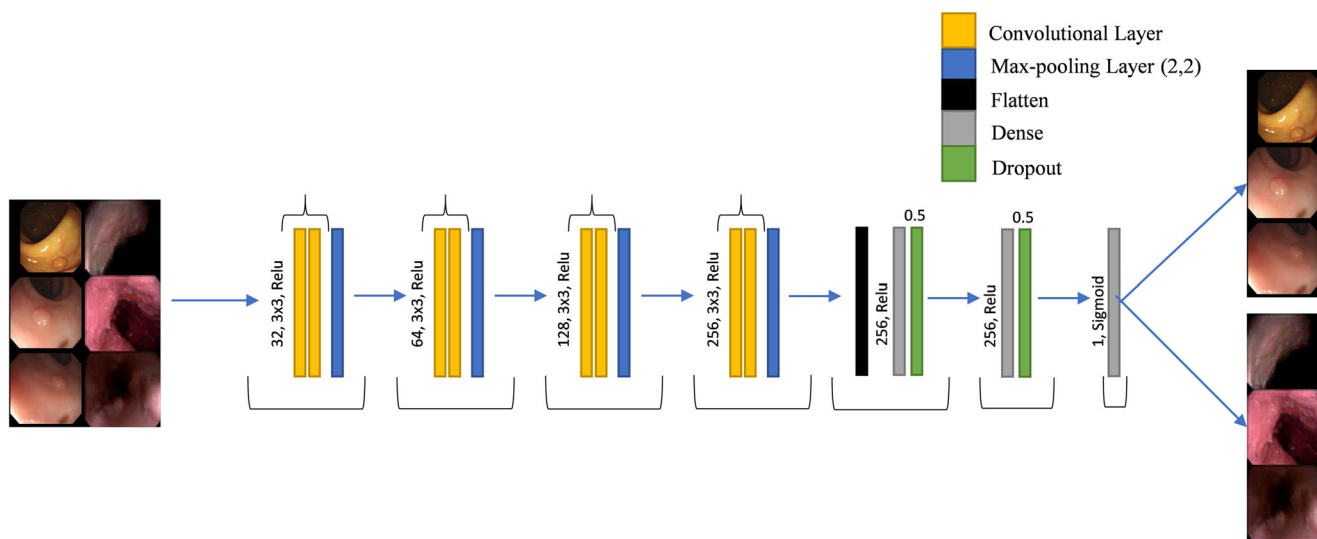
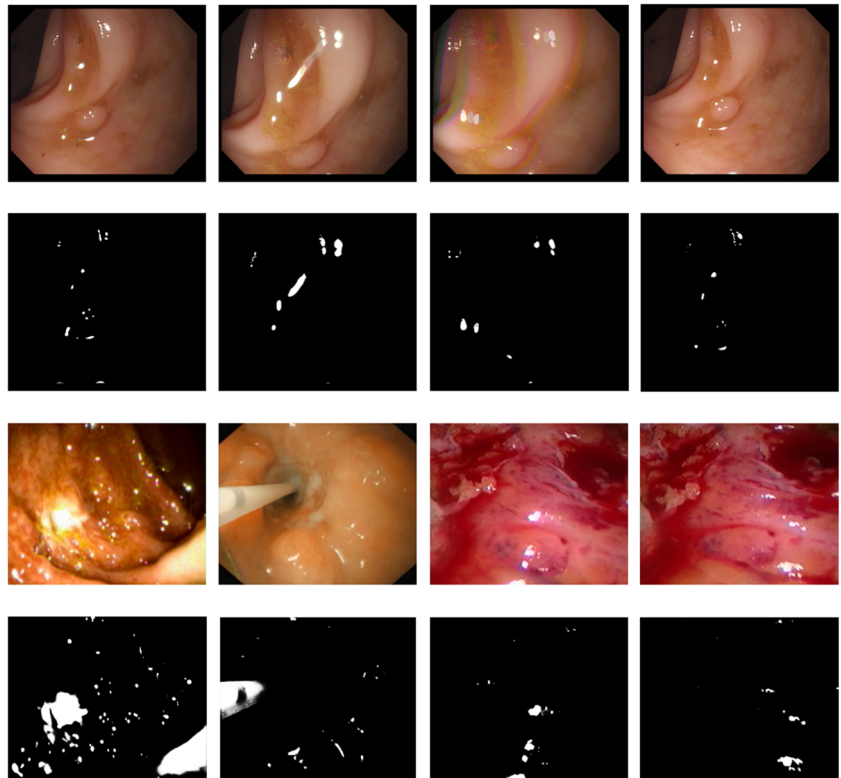


Fig. 5 The deep learning based network for classification of reflected endoscopy images and weak illuminance endoscopy images

Fig. 6 Row 1-2: CVC colon DB [25] image and their corresponding reflection labels produced by Aslaleh et al., [5] used for training U-net. Row 3-4: NASPAGHAN [27] images and their corresponding results from U-net



respectively. Then, a hole-filling method [30] is adopted to correct the reflection area in the endoscopy images. During the endoscopy image restoration, we optimized the energy function E between the specular reflection region R and normal region N . E is used to restore the local information of R with enhanced reasonability and similarity to several local neighborhood information of N and defined as (1).

$$E(R, N) = \sum_{q \in R} \min_{p \in N} (D(Q, P) + \lambda D(\nabla Q, \nabla P)) \quad (1)$$

Where $Q = H(q)$ is a rectangle patch located at the upper left corner of reflection pixel q , $P = f(N(p))$ is a transformation f which includes various kinds of operations, such as rotation, scale, and translation to the normal pixel p 's neighborhood region H . Here the notation Q and P , ∇Q and ∇P are also used to represent the color and luminance gradient channels of the corresponding patch, respectively. The sum of squared distances of all channels in the patch is represented by D . In order to decrease the energy function E , we optimized it through the iterative operation of patch searching and color voting [31]. Specifically, during the searching process, all the output overlapping patches are retrieving their nearest neighborhood input patches. And in the voting process, the

final image is acquired by averaging the votes of each color pixel from the above-blended patches in searching results. The colors will become converge with the iterations come to an end. Finally, we repeated this step to realize the image from coarse to fine-scale restoration. More implementation details can be found in [30].

2.3 Image enhancement

The relationship of an image S with its reflectance R and illumination L can be expressed as $S = RL$. Most of the weak illumination situations are caused by a low illumination parameter (L). Therefore, accurately determining R and L and performing illumination correction operations can enhance the image. We then conducted an endoscopy image enhancement on the basis of the estimation method proposed in [32]. The objective of estimating R and L is to solve the objective function below.

$$E(r^k, l^k, d^k, b^k) = \|l^k + r^k - s\|_2^2 + c_2 \|L^{k-1} \cdot \nabla l^k\|_2^2 + c_1 \left\{ \|d^k\|_1 + \lambda \|R^{k-1} \cdot \nabla r^k - d^k + b^k\|_2^2 \right\} \quad (2)$$

s.t. $r^k \leq 0$ and $s \leq l^k$

Table 1 Evaluation results of image classification network

	Reflected Predicted	Predicted Weak illuminance
Actual	95	0
Reflected (95)		
Actual Weak illuminance (34)	0	34

where c_1 and c_2 are parameters larger than zero; r , l , and s are the logarithms of R , L , and S , respectively; and d and b are the auxiliary parameters. To reduce the value of E , the first term $\|l + r - s\|_2^2$ is used to make the estimated image value $r + l$ equal to the original value s . Then, the second term $\|L \cdot \nabla l\|_2^2$ is used to increase the smoothness of the estimated illumination l , and the third term is used to transform the reflectance r into a piecewise constant.

We performed the alternating direction method of multipliers to minimize the objective function and determine the values of r and l ; the details of this method are discussed in [33]. The R and L of an input endoscopy image S can be calculated as $R = e^r$ and $L = e^l$, respectively. To achieve

image enhancement after the illumination component is calculated, we introduced gamma correction [34, 35] to adjust the L as $L' = w(L/w)^{1/r}$, where $w = 250$ and $r = 2.5$. Finally, the enhanced image S' is obtained as $S' = R \cdot L'$, where R is the estimated reflectance and L' is the adjusted illumination.

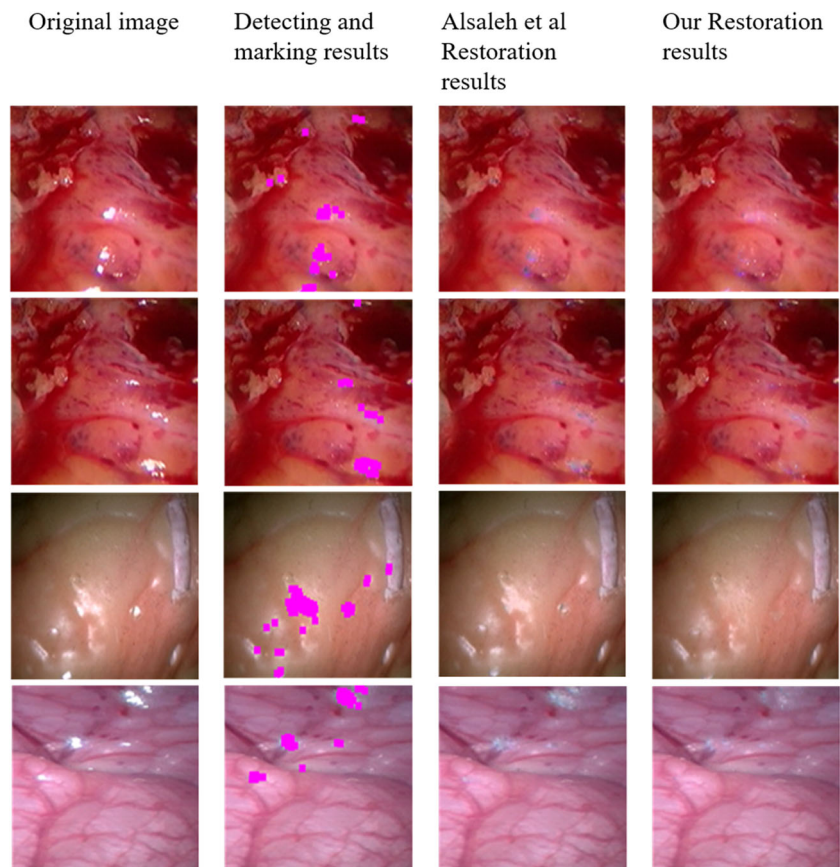
3 Experimental results and analysis

We conducted several experiments to evaluate the performance of the proposed framework.

3.1 Image Classification

As previously mentioned, image classification testing is performed on 129 (34 + 95) images gathered from Beijing Tongren Hospital and NASPGHAN [27]. Some of the images are shown in Fig. 1. The test images are classified using an image classification neural network. The 129 images can be classified into two categories, and the classification ratio for all input endoscopy images in the network is 100% (Table 1). A confusion matrix is used to analyze and compare the incorrect and correct predictions with the actual results.

Fig. 7 Detection and restoration results of the proposed approach. Left to right 1st column: Original image, 2nd column: Detecting and marking results, 3rd column Alsaleh et al., method and 4th column: Our results



3.2 Image restoration

The detection and restoration results of the proposed approaches are presented in Fig. 7. The figure shows a visual illustration of the results, in which the detected specularity is overlaid to the original areas. The results indicate that the proposed detection and patch-based approaches fill various sizes of holes and complete the image using examples from different orientations, scales, and colors. The comparison of the restoration results obtained through the proposed algorithm and that introduced by Alsaleh et al. [5] is displayed in Fig. 7. In terms of subjective evaluation, the proposed approach can adapt to diverse color variations and eliminate the specular highlight points better than that of Alsaleh et al. The results obtained using the latter exhibit visible artifacts in the restored areas, whereas those obtained through the former are visually plausible and in line with human perception.

Given the absence of ground truth for the datasets, we used the coefficient of variation (COV) to quantitatively evaluate the proposed restoration approach. The COV reflects the intensity homogeneity within a region and

is defined as $COV = \sigma/\mu$, where σ is the standard deviation and μ is the mean of the pixel values. A set of affected regions with specular highlights are shown in Fig. 8, and their corresponding COV values are listed in Table 2.

The images restored using the proposed method are more homogeneous than those restores using the method proposed in [5]. The results in Table 2 suggest that the dispersion probability of the proposed method is relatively lower than that of the method of Alsaleh et al. [5]. This finding indicates that the variation in the mean of the proposed method is acceptable; high values of mean and standard deviation signify high specular reflection, which affects the COV of the image. The standard deviation values of the proposed method in all four examples are significantly lower given in Table 2, which highly contributes to exact restoration on the basis of reflection detection. The evaluated mean values of the first three examples for the proposed method express the good performance, while the fourth example mean value is close to the Alsaleh et al. method but our method does not produce irregular color as [5] (Fig. 8). The final COV results of

Fig. 8 Results of restoration process showing better coefficient of variation (COV) in selected regions, COV of 1st column: Original image, 2nd column: Alsaleh et al.'s result and 3rd column: Our result

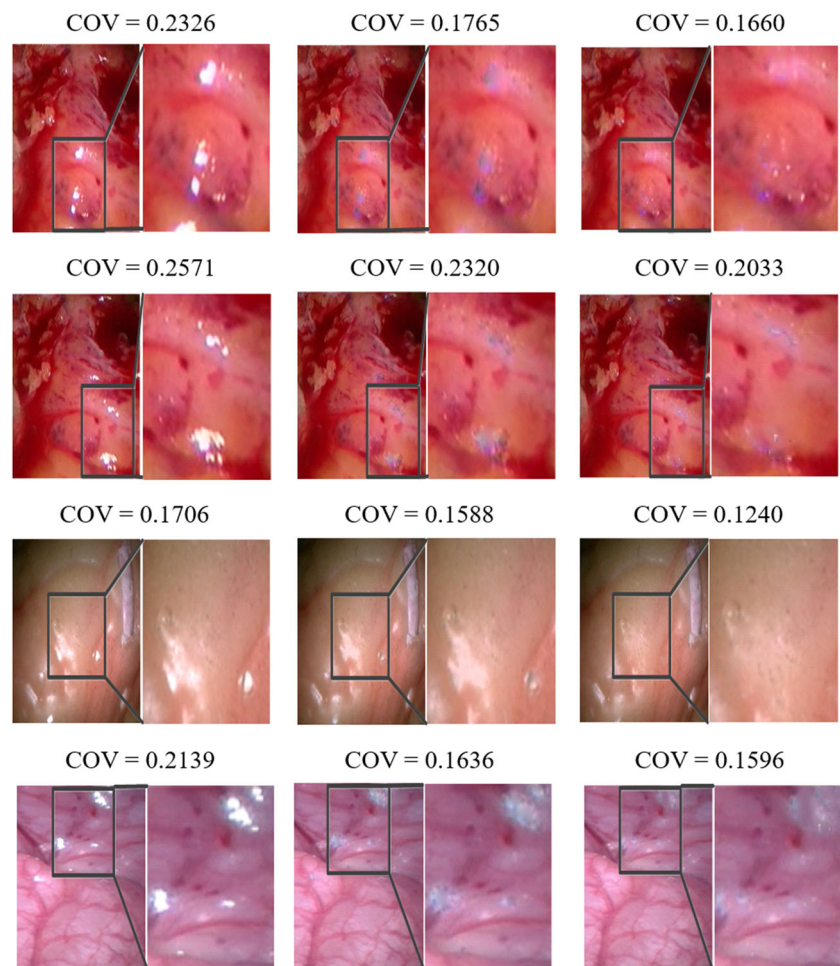


Table 2 Comparison of the COV values for evaluation images given in Fig. 8

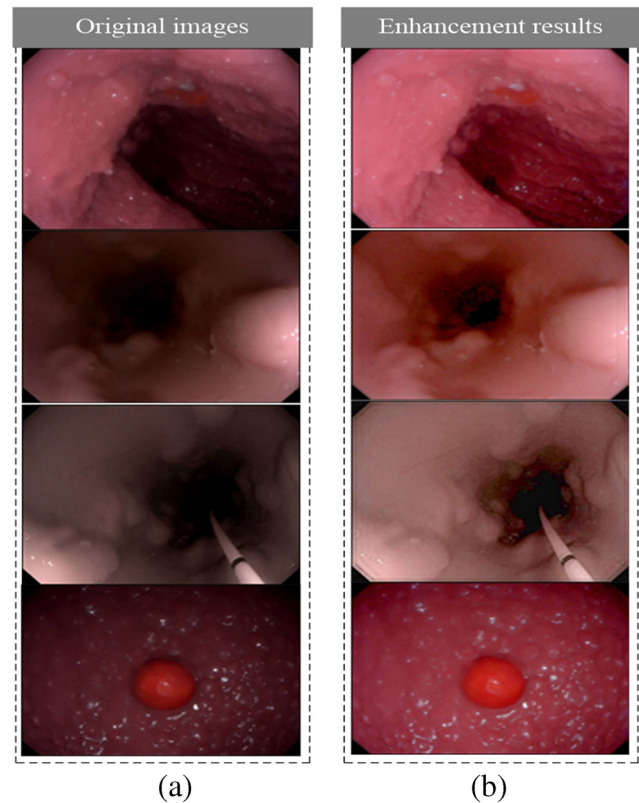
Comparison	The standard deviation of pixel values σ	The mean of pixel values μ	The coefficient of variation (COV) of image $\frac{\sigma}{\mu}$
1st Row			
Original image	28.78	123.71	0.23
Alsaleh et al.	21.29	120.66	0.18
Our result	20.45	123.17	0.17
2nd Row			
Original image	32.06	124.70	0.26
Alsaleh et al.	26.66	114.90	0.23
Our result	24.55	120.76	0.20
3rd Row			
Original image	26.85	161.15	0.17
Alsaleh et al.	24.55	154.58	0.16
Our result	19.59	157.93	0.12
4th Row			
Original image	28.48	133.15	0.21
Alsaleh et al.	21.32	130.40	0.16
Our result	20.47	130.50	0.16

Fig. 8 and Table 2 clearly classify that obtained results by the proposed method are enhanced. Subsequently, we quantitatively investigated the specular highlights by treating them as image noise and determining the signal-to-noise ratio (SNR). The SNR of the restored images in Fig. 8 ranges from 28.7 dB to 30.5 dB, which indicates a plausible measurement. The SNR values of the restored images are 28.9, 29.4, 30.5, and 28.7 dB, which show the minimal influence of noise on the overall signal. In conclusion, the COV and SNR measurements demonstrate the high efficacy of the proposed method.

3.3 Image enhancement

If an input image is classified as a weakly illuminated case, the image enhancement model is used to enhance the dark areas. The empirical parameters are set as 0.01, 0.1, and 1. To preserve the color information, the gamma correction is processed in the hue, saturation, value (HSV) domain. The overall enhancement results of the proposed approach are shown in Fig. 9.

To evaluate the performance of the image enhancement algorithm, we compared it with the method presented by Selka et al. [36]. The results show that the former has a smoother effect and better natural performance in shadowed areas than the latter (Fig. 10). Given that the proposed method simultaneously estimates the reflectance and illumination, the regularization term in

**Fig. 9** Enhanced results. Left to right **a** Original image, **b** Proposed approach

(2) can effectively suppress the noise in dark areas. Moreover, because the ground truth of an enhanced image is unknown, the Natural Image Quality Evaluator (NIQE) [34] blind image quality assessment based on statistical regularities of natural and undistorted images is adopted to evaluate the enhanced results; a low NIQE signifies high image quality. Table 3 values shows that the proposed approach obtains a lower value than the method of Selka et al. [36], which means that the former achieved high-quality enhanced results. We further compared the proposed algorithm with other methods, including dynamic histogram equalization (DHE) [37], anisotropic diffusion method (ADM) [38], and adaptive anisotropic diffusion (AAD) [39].

The overall enhancement results and NIQE values are shown in Fig. 11 and Table 4, respectively. The experimental results imply that the empirical settings of the parameters generate satisfactory results, and the proposed algorithm effectively enhances the weakly illuminated endoscopy images.

As shown in Fig. 11, the original image is slightly blurred, and some edges of this wireless capsule endoscopy image are too weak to detect. The first row of Table 4 indicates that the ADM algorithm failed to enhance the image and produced a blurrier image than the original. The

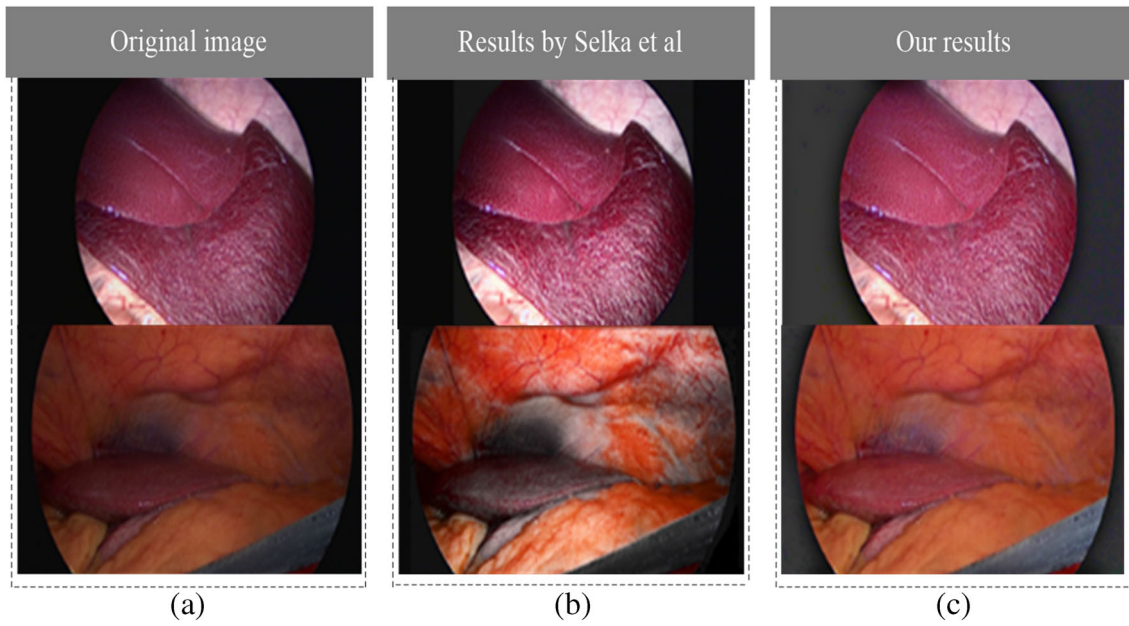


Fig. 10 Comparison of enhanced results: **a** Original image, **b** Results by that Selka et al., **c** Our results

Table 3 Average NIQE values of Fig. 10

Comparison	Original Image	Selka et al.'s	Our Results
1st row	11.9	11.0	10.4
2nd row	5.1	2.9	3.4
Average	8.5	7.0	6.9

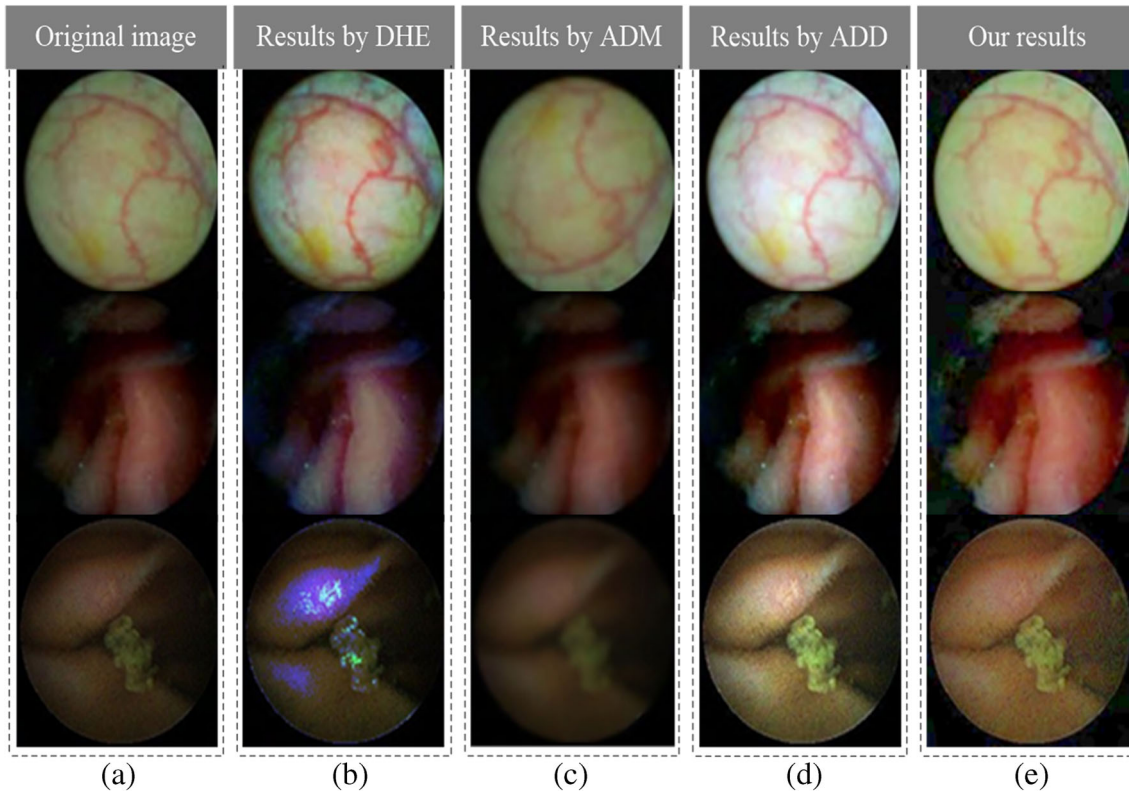


Fig. 11 Comparison of enhancement results, **a** original image, **b** results by DHE, **c** results by ADM, **d** results by AAD, **e** our results

DHE algorithm produces an acceptable result but magnifies noises, and the AAD algorithm generates a clear vessel texture but loses the original pixel information. In the second row, the result produced by the ADM algorithm is also inferior to the original image. The image obtained through the DHE algorithm is still weakly illuminated and causes chrominance change. In the third row, the result from the DHE algorithm induces chrominance changes in the abnormal region. The average values presented in Table 4 are low, and relatively clear results with low noise amplification are observed in the dark regions, where many image details can be viewed to achieve an accurate judgment of the images obtained by the proposed method. The NIQE evaluation not fully proximate to the reference image but give us absolute image quality. The quantitative values of the first three rows in Table 4 have minute difference comparatively other methods but the visual representation is supported to the proposed method results, as presented in Fig. 11. The statistical regularities of the resultant images are sharper, clean, without noise and very similar to the raw images, even though other methods generate irregular colors. Which verify that endoscopy image enhancement with the proposed technique is relatively better than other methods.

The proposed automatic framework is the first one, which comprehensively explored DL models for endoscopy images and deals with reflected and low-resolution endoscopy images simultaneously. The DL model for image classification is trained using pretrained ImageNet CNNs because the endoscopy images are affected by several restrictions in medical image analysis. For instance, such analysis assumes that DL requires huge data for learning. Clear medical images require a large amount of data, which is the same as the amount used in ImageNet. Natural images display numerous variations in terms of appearance, geometry, and lighting conditions. Conversely, the variations in medical images are relatively minimal, and these images do not require a large amount of data [40, 41]. The VGGNet was further developed to classify 1000 uncommon medical classes. Medical images do not require large models.

Moreover, the full extracted feature can be obtained by resizing the image in accordance to the training image size of ImageNet. A small image size can increase the computational cost, whereas a large one can reduce the image details [42].

Restoring the reflected areas is difficult without proper detection. Similarly, training the DL model without ground truth is challenging. Therefore, we developed reflection labels by applying the proposed method on CVC colon DB images [25].

On the other hand to medical image quality enhancement, models trained on natural images enhancement are never tested on medical image [43–45]. Because quality requirements and medical images have specific challenges, unlike natural Images [41, 46]. The HSV images are often dark in some regions and natural images are poorly-lit due to underexposed areas. Thus, the natural images are required enhancement in specific regions but medical images are sparse and unfulfilled that require a specific number of labeled training data annotated by experts [46, 47], which is not yet available.

4 Conclusion

In this study, an automatic framework for the simultaneous restoration and enhancement of endoscopy images is proposed. The endoscopy images are classified into two categories using DL techniques. Images with specular highlights are restored by using an automatic highlight detection method and a patch-based optimization restoration model, and the weakly illuminated images are enhanced using the alternating direction method of multipliers and a gamma correction operation. The proposed framework and algorithm are evaluated through a comprehensive experimental analysis. The quantitative and qualitative results show the effectiveness and efficiency of the proposed method in the collected dataset. Future works will include additional clinical studies.

Table 4 Average NIQE values of Fig. 11

	Original image	Abdullah et al.'s (DHE)	Perona et al.'s (ADM)	Li et al.'s (AAD)	Our
1st row	39.8	39.8	39.8	39.85	39.8
2nd row	39.8	39.8	39.8	39.8	39.8
3rd row	5.9	5.8	11.7	6.9	5.5
Average	28.5	28.5	30.4	28.8	28.4

References

- Sdiri B, Cheikh FA, Dragusha K, Beghdadi A (2015) Comparative study of endoscopic image enhancement techniques. In: 2015 Colour and Visual Computing Symposium (CVCS), pp 1–5
- Domingues I, Sampaio IL, Duarte H, Santos JAM, Abreu PH (2019) Computer vision in esophageal cancer: a literature review. *IEEE Access* 7:103080–103094
- Fu G, Zhang Q, Song C, Lin Q, Xiao C (2019) Specular Highlight Removal for Real-world Images. *Comput Graph Forum* 38(7):253–263
- Son M, Lee Y, Chang HS (2020) Toward specular removal from natural images based on statistical reflection models. *IEEE Trans Image Process* 29:4204–4218
- Alsaleh SM, Aviles AI, Sobrevilla P, Casals A, Hahn JK (2016) Adaptive Segmentation and Mask-Specific Sobolev Inpainting of Specular Highlights for Endoscopic Images, 38th Annu. Int. Conf. *IEEE Eng. Med. Biol. Soc.*, pp 1196–1199
- Guo J, Shen DF, Lin GS, Huang JC, Liu KC, Lie WN (2016) A specular reflection suppression method for endoscopic images, *Proc. - 2016 IEEE 2nd Int. Conf. Multimed. Big Data, BigMM 2016*, pp 125–128
- Hsia C, Chiang J, Li H, Lin C, Chou K (2016) A 3D endoscopic imaging system with Content-Adaptive filtering and hierarchical similarity analysis. *IEEE Sens J* 16(11):4521–4530
- Saint-Pierre CA, Boisvert J, Grimard G, Cheriet F (2011) Detection and correction of specular reflections for automatic surgical tool segmentation in thoracoscopic images. *Mach Vis Appl* 22(1):171–180
- Meslouhi O, Kardouchi M, Allali H, Gadi T, Benkaddour Y (2011) Automatic detection and inpainting of specular reflections for colposcopic images. *Open Comput Sci* 1(3):341–354
- Zimmerman-Moreno G, Greenspan H (2006) Automatic Detection of Specular Reflections in Uterine Cervix Images. *SPIE Med. imaging*, pp 61446E—61446E—9
- Imtiaz MS, Wahid K (2014) Image enhancement and space-variant color reproduction method for endoscopic images using adaptive sigmoid function. In: 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp 3905–3908
- Imtiaz MS, Wahid K (2014) A color reproduction method with image enhancement for endoscopic images. In: 2nd Middle East Conference on Biomedical Engineering, pp 135–138
- Li B, Meng MQ-H (2012) Wireless capsule endoscopy images enhancement via adaptive contrast diffusion. *J Vis Commun Image Represent* 23(1):222–228
- Huang S, Lee F, Miao R, Si Q, Lu C, Chen Q (2020) A deep convolutional neural network architecture for interstitial lung disease pattern classification. *Med Biol Eng Comput*:1–13
- Hesamian MH, Jia W, He X, Kennedy P (2019) Deep learning techniques for medical image segmentation: Achievements and challenges. *J Digit Imaging* 32(4):582–596
- Zhao Z-Q, Zheng P, Xu S, Wu X (2019) Object detection with deep learning: a review. *IEEE Trans Neural Netw Learn Syst* 30(11):3212–3232
- Nassif AB, Shahin I, Attili I, Azzeh M, Shaalan K (2019) Speech recognition using deep neural networks: a systematic review. *IEEE Access* 7:19143–19165
- Guo G, Zhang N (2019) A survey on deep learning based face recognition. *Comput. Vis. Image Underst.* 189:102805
- Lin H, Yang P, Zhang F (2019) Review of scene text detection and recognition. *Arch Comput Methods Eng*:1–22
- Lavecchia A (2019) Deep learning in drug discovery: opportunities, challenges and future prospects. *Drug Discov Today* 24(10):2017–2032
- Adeel A, Gogate M, Hussain A, Whitmer WM (2019) Lip-reading driven deep learning approach for speech enhancement. *IEEE Trans Emerg Top Comput Intell*
- Shrestha A, Mahmood A (2019) Review of deep learning algorithms and architectures. *IEEE Access* 7:53040–53065
- Bernal J et al (2017) Comparative validation of polyp detection methods in video colonoscopy: Results from the MICCAI 2015 endoscopic vision challenge. *IEEE Trans Med Imaging* 36(6):1231–1249
- Pogorelov PT et al (2017) KVASIR. In: *Proceedings of the 8th ACM on Multimedia Systems Conference*, pp 164–169
- Bernal F, Sanchez J, Vilarino J (2012) Towards automatic polyp detection with a polyp appearance model. *Pattern Recognit* 45(9):3166–3182
- Silva B, Histace J, Romain A, Dray O, Granado X (2014) Towards embedded detection of polyps in WCE images for early diagnosis of colorectal cancer To cite this version. *Int J Comput Assist Radiol Surg* 9(2):283–293
- North American Society for Pediatric Gastroenterology, Hepatology and Nutrition. [Online]. Available: <https://www.naspghan.org/content/97/en/professional-education/resources/endoscopy-photo-gallery>
- Goodfellow Y, Pouget-Abadie I, Mirza J, Xu M, Warde-Farley B, Ozair D, Courville S, Bengio A (2014) Generative adversarial networks. *Adv Neural Inf Process Syst.*:2672–2680
- Ronneberger O, Fischer P, Brox T (2015) U-net: Convolutional networks for biomedical image segmentation, in *International Conference on Medical image computing and computer-assisted intervention*, pp 234–241
- Darabi S, Shechtman E, Barnes C, Goldman DB, Sen P (2012) Image melding: combining inconsistent images using patch-based synthesis. *ACM Trans Graph* 31(4):1–10
- Wexler Y, Shechtman E, Irani M (2004) Space-Time Video Completion, in *Computer Vision and Pattern Recognition (CVPR)*, pp 120–127
- Fu X, Zeng D, Huang Y, Zhang X-P, Ding X (2016) A weighted variational model for simultaneous reflectance and illumination estimation. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp 2782–2790
- Goldstein T, Osher S (2009) The split Bregman method for L1-regularized problems. *SIAM J Imaging Sci* 2(2):323–343
- Mittal A, Soundararajan R, Bovik AC (2012) Making a ‘completely blind’ image quality analyzer. *IEEE Signal Process Lett* 20(3):209–212
- Kimmel R, Elad M, Shaked D, Keshet R, Sobel I (2003) A variational framework for retinex. *Int J Comput Vis* 52(1):7–23
- Selka F, Nicolau SA, Agnus V, Bessaid A, Marescaux J, Soler L (2013) Evaluation of Endoscopic Image Enhancement for Feature Tracking: A New Validation Framework. In: *Augmented Reality Environments for Medical Imaging and Computer-Assisted Interventions*, pp 75–85
- Abdullah-Al-Wadud M, Kabir MH, Dewan MAA, Chae O (2007) A dynamic histogram equalization for image contrast enhancement. *IEEE Trans Consum Electron* 53(2):593–600
- Perona P, Malik J (1990) Scale-space and edge detection using anisotropic diffusion. *IEEE Trans Pattern Anal Mach Intell* 12(7):629–639
- Li L, Zouthe YX, Li Y (2013) Wireless capsule endoscopy images enhancement based on adaptive anisotropic diffusion. In: 2013 IEEE China Summit and International Conference on Signal and Information Processing, pp 273–277
- Erickson BJ, Korfiatis P, Kline TL, Akkus Z, Philbrick K, Weston AD (2018) Deep learning in radiology: does one size fit all? *J Am Coll Radiol* 15(3):521–526

41. Sahiner B et al (2019) Deep learning in medical imaging and radiation therapy. *Med Phys* 46(1):e1–e36
42. Wong KCL, Syeda-Mahmood T, Moradi M (2018) Building medical image classifiers with very limited data using segmentation networks. *Med Image Anal* 49:105–116
43. Li C, Guo J, Porikli F, Pang Y (2018) Lightnet: A convolutional neural network for weakly illuminated image enhancement. *Pattern Recognit Lett* 104:15–22
44. Wang W, Wei C, Yang W, Liu J (2018) GLADNet: Low-light enhancement network with global awareness, in 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018), pp 751–755
45. Lv F, Lu F, Wu J, Lim C (2018) MBLLEN: Low-Light Image/Video Enhancement Using CNNs. In: *BMVC*, pp 220
46. Litjens G et al (2017) A survey on deep learning in medical image analysis. *Med Image Anal* 42:60–88
47. Greenspan H, Van Ginneken B, Summers RM (2016) Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique. *IEEE Trans Med Imaging* 35(5):1153–1159

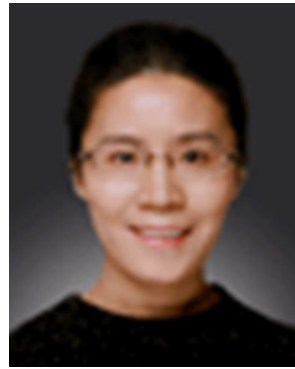
Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Muhammad Asif Ph.D. student at the School of Computer Science and Technology, Beijing Institute of Technology, China. His Major is Software Engineering. His research interest in Medical Image Processing and currently working on Endoscopy images. He is graduated from the National University of Science and Technology (NUST), Pakistan, in 2016.



Lei Chen is a Ph.D. student at the School of Computer Science and Technology, Beijing Institute of Technology, China. His major is software engineering. His research interests lie in medical image processing using machine learning and deep learning methods. He graduated from the Tianjin Institute of Technology, China, in 2016.



Hong Song is a professor of the School of Computer Science and Technology, Beijing Institute of Technology. Her focus is on the application of machine learning and deep learning methods for medical image analysis and improved computer-aided diagnosis. She received her Ph.D. at the School of Computer Science and Technology, Beijing Institute of Technology.



Jian Yang is a professor of the School of Optics and Photonics, Beijing Institute of Technology. His focus is the application of virtual reality, medical image processing and computer vision. He received his Ph.D. at the School of Optics and Photonics, Beijing Institute of Technology, and was a Research Fellow at the Hospital for Sick Children, Canada, from 2007 to 2009.



Alejandro F. Frangi is a Professor of Biomedical Image Computing at the University of Leeds and affiliated to the School of Computing and School of Medicine, UK. He is also Director of the Center for Computational Imaging and Simulation Technologies in Biomedicine. Prof Frangi is a Fellow of IEEE.