



Low-light image enhancement based on normal-light image degradation

Bai Zhao¹ · Xiaolin Gong¹ · Jian Wang¹ · Lingchao Zhao¹

Received: 27 May 2021 / Revised: 11 November 2021 / Accepted: 14 November 2021 / Published online: 28 January 2022
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2021

Abstract

In recent years, deep learning has demonstrated its impressive performance in image enhancement. A novel image enhancement method based on normal-light image degradation is proposed in this paper. The degraded images generated from normal-light images by gamma transformation are adopted as the reference images in network training process. Besides, we designed a multi-scale fusion network, which connects two encoding–decoding subnetworks in parallel. The network completes repeated multi-scale fusions by exchanging the information across the parallel subnetworks over and over through the training process. The final enhanced images are obtained by performing inverse gamma transformation on the output of the network. Benefiting from good detail preservation of reference images, smaller gap in brightness and contrast of training image pairs, and the multi-scale fusion network, the method is expected to enhance low-light images while preserving naturalness. Experiments demonstrate the superiority of the proposed method over state-of-the-art image enhancement methods.

Keywords Low-light image enhancement · Normal-light image degradation · Multi-scale fusion network

1 Introduction

The development of digital photography greatly improves the images quality. However, the captured images in the low-light environment often suffer from low contrast and low quality due to non-uniform illuminated conditions [1]. These disadvantages may deteriorate the performance and efficiency of relative vision processing system, such as medical examination, monitoring and reconnaissance [2–4]. Upgrading camera sensor could alleviate the problem in some way; however, the high cost limits its application. Increasing exposure time may introduce additional noise or blur [5]. As an effective solution, low-light image enhancement method at the

software end has been proposed for many years [6]. It aims to restore the low-light image into its natural scene of high contrast, vivid color and rich details, where the utilization of information is expected to be improved [1]. Nowadays, there are plenty of image enhancement methods with their own advantages and priorities. Histogram equalization (HE) methods [7,8] perform light enhancement through expanding the dynamic range. Retinex theory [9] assumes that an image can be described as the product of illumination and reflectance. Retinex-based methods [10–14] adaptively adjust the two components to achieve image enhancement. Ying et al. [15] fused the input with the synthesized image according to the estimated weight matrix. Ren et al. [16] selected a camera response model to adjust the pixels exposure values. Fu et al. [17] proposed a fusion method that combines the advantages of sigmoid function and histogram equalization. These methods can effectively enhance the brightness of low-light images. However, they may ignore the correlation between regions, which tends to incur over-exposure or color distortion. Recently, with the development of its own technology, deep learning also inspires its application in low-light image enhancement. It can be broadly categorized into two groups: convolutional neural network (CNN)-based methods and generative adversarial network (GAN)-based methods [18]. CNN-based methods restore the

✉ Xiaolin Gong
gxl@tju.edu.cn

Bai Zhao
zhaobai@tju.edu.cn

Jian Wang
jianwang@tju.edu.cn

Lingchao Zhao
zhaolingchao@tju.edu.cn

¹ Tianjin Key Laboratory of Imaging and Sensing
Microelectronic Technology, School of Microelectronics,
Tianjin University, Tianjin 300072, China

buried information under the guidance of reference images. Wei et al. [19] proposed a Retinex-Net, which includes a Decom-Net for decomposition and an Enhance-Net for illumination adjustment. Ma et al. [20] adjusted saturation (S) and used CNN to enhance intensity component in HSI color space. Huang et al. [21] used the illumination mask to predict the illumination distribution and used the Retinex model to estimate the initial enhanced image, while the final enhanced result could be obtained after color distortion modification and noise suppression. Wang et al. [22] calculated a global illumination estimation and then utilized the estimation and the original input to reconstruct details. Atoum et al. [23] proposed a color-wise attention network to learn an end-to-end mapping between low-light and enhanced images while searching for any useful color cues in the low-light image to aid in the color enhancement process. CNN-based methods can flexibly design modules to denoise or adjust illumination, but existing networks may not perform well in details. In contrast, GAN-based methods are proposed with unpaired supervision. Jiang et al. [24] proposed an unsupervised GAN with a global-local discriminator structure, a self-regularized perceptual loss fusion, and attention mechanism. Hua et al. [25] proposed a joint GAN for image enhancement and an image quality assessment techniques for quality improvement. The enhanced images of GAN-based methods are usually visually consistent with human perception, while problems such as color distortion and inconsistency may be inevitable.

Comprehensively considering the advantages, weakness, and potentials of existing methods, we propose a novel low-light image enhancement approach based on normal-light image degradation in this paper, through which it is expected to obtain the enhancement effect of both well color distributed and detail restructured.

- To the best of our knowledge, this is the first attempt to use the degraded images as reference images. The degraded images (as shown in Fig. 1) gamma-transformed from the normal-light images are much closer on brightness and contrast with low-light images. Meanwhile, gamma transform only changes the dynamic range of the image and does not introduce noise. Therefore, the degraded images are more effective for low-light image enhancement network training than the normal-light images.
- We designed a network which repeatedly exchanges information across a high-resolution subnetwork and a symmetric high-to-low and low-to-high subnetwork to boost the feature extraction. Since the two subnetworks are connected in parallel, the exchanging of information results in a rich representation of feature maps.
- With the help of exposure control loss, the output is potentially more natural. Experimental results demonstrate that

our method outperforms several state-of-the-art enhancement methods.

2 Proposed method

The proposed method is illustrated from three parts: data processing, network architecture, and loss function.

2.1 Data processing

The traditional network is trained based on paired low-light/normal-light images. However, due to the great gap in brightness and contrast between them, the network parameters are much more sensitive to the difference of brightness and contrast rather than to the relative small difference of the detail information and thus being unfavorable for the recovery of detail information. By comparison, the brightness and contrast of the degraded images transformed from the normal-light images are closer to that of the low-light images. Therefore, using the degraded images as the reference images, that is, using paired low-light/degraded images to train the network, is helpful to improve the sensitivity of the network to details.

Meanwhile, the images after gamma transformation have been indicated being capable to preserve adequate details information of the original images and performing well in previous work [26], and it can be easily restored to normal-light condition by inverse transformation.

Equation 1 is used to represent that the degraded image I_{deg} is obtained by performing gamma transformation on the normal-light image I_{nor} .

$$I_{\text{deg}} = I_{\text{nor}}^{\gamma} \quad (1)$$

where γ is the degradation coefficient. The degraded image I_{deg} is used as reference image I_{ref} in this paper. An illustration of I_{low} , I_{deg} , I_{nor} is shown as Fig. 1. We do experiments by setting γ within [0.1, 1.1]. The experiment parameters are illustrated in Sects. 2.3 and 3. An example of grayscale histograms and the experiments results is shown in Figs. 2 and 3, respectively. According to the grayscale histograms, it can be found that the degraded images with coefficients of 0.7, 0.8, 0.9 have the similar pixel values range as the low-light image. When the γ is set within [0.1, 0.6], the pixel values of the degraded images gradually concentrate into a very small range, even smaller than the range of pixel values of the low-light image. Since the pixel values are discrete, the information in the degraded images is decreased and the output gradually shows obvious noise. Traditional methods usually use images with γ set to 1 for training, and the large difference in brightness and contrast brings trouble for enhancement. The output in the experiment shows color loss. When the

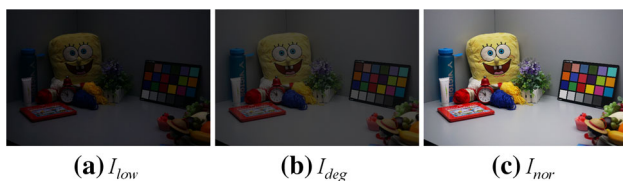


Fig. 1 The example of the low-light image I_{low} , degraded image I_{deg} , and normal-light image I_{nor}

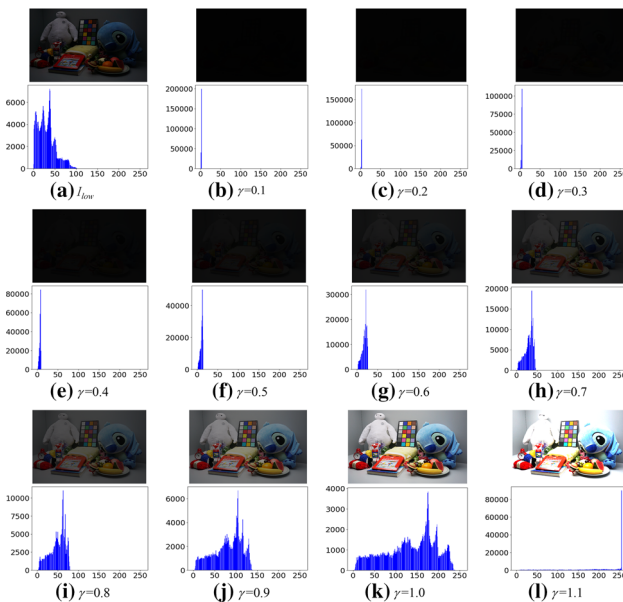


Fig. 2 Low-light image I_{low} , degraded images I_{deg} with coefficients γ taking the values within [0.1, 1.1] and corresponding grayscale histograms

γ is set to 1.1, the gap between the low-light image and the reference image is further expanded. In the experiment, the output shows color loss and the brightness enhancement is also insufficient. The comparison on peak signal-to-noise ratio (PSNR), structural similarity (SSIM) [27] and natural image quality evaluator (NIQE) [28] in Table 1 is also consistent with the analysis. Comprehensively considering the grayscale histograms and the experimental results, gamma is set to 0.8 in this paper. More experimental results on $\gamma = 0.8$ and $\gamma = 1.0$ can be seen in the ablation study in Sect. 3.2.

During the network training process, paired low-light/degraded images are taken as input images. Once the network training ends, the mapping model is determined, and the low-light image I_{low} could be transposed into corrected image I_{cor} through the mapping model. Specifically, the correction result I_{cor} is represented as:

$$I_{cor} = G(I_{low}) \tag{2}$$

where I_{cor} and $G(\cdot)$ are the output and corresponding mapping process for the network, respectively. After performing the inverse gamma transformation on I_{cor} , the final enhanced

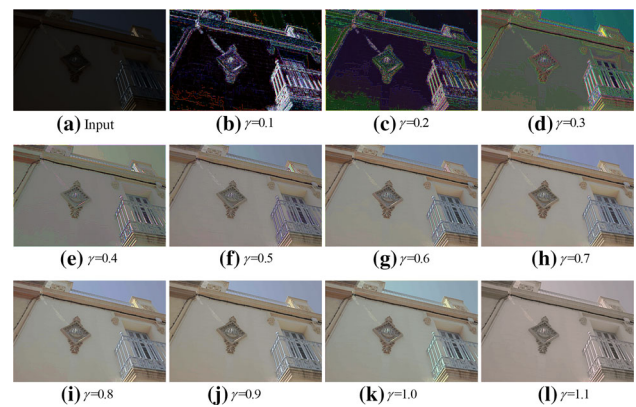


Fig. 3 The results of experiments with different gamma values

Table 1 Comparison of performance metrics for experiments using different gamma values

γ	PSNR	SSIM	NIQE
0.1	5.983	0.027	11.9096
0.2	7.173	0.061	11.0334
0.3	10.781	0.123	7.2164
0.4	14.671	0.253	6.1037
0.5	17.479	0.490	5.3485
0.6	18.707	0.670	4.6330
0.7	19.577	0.726	3.5603
0.8	20.298	0.769	3.1959
0.9	20.066	0.767	3.3358
1.0	19.630	0.767	3.8611
1.1	19.019	0.757	4.5077

The best results are in bold

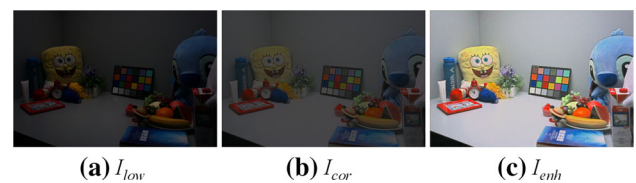


Fig. 4 The example of the corrected image I_{cor} and the enhanced image I_{enh}

image I_{enh} could be gained:

$$I_{enh} = \frac{1}{I_{cor}^\gamma} = (G(I_{low}))^{\frac{1}{\gamma}} \tag{3}$$

An example of I_{cor} and I_{enh} is provided in Fig. 4.

2.2 Network architecture

Figure 5 shows the whole architecture of our design. As illustrated in Fig. 5, the network consists of two subnetworks, a typical symmetric high-to-low and low-to-high network (HL-net) inspired by the U-Net [29], and a high-resolution network (H-net) with feature maps of the same resolution as their input image.

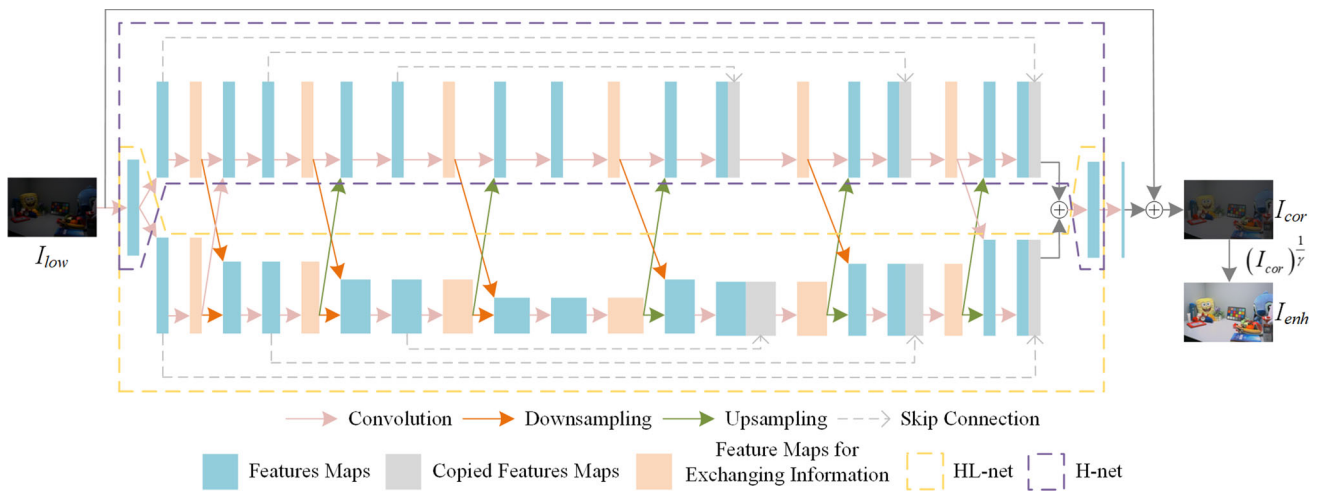


Fig. 5 Architecture of the proposed method

HL-net has 21 convolutional layers, 3 downsampling steps, and 3 upsampling steps. Each downsampling step is a convolution operation with stride 2. And each upsampling step contains a bilinear interpolation to expand the height and width of the feature map to twice the original, which also enables the final model to process images of any size [22]. Besides, three cascaded convolutional layers are included between two spatial resolution regulation operations. Each convolutional layer consists of a 3×3 convolution operation with padding, followed by a rectified linear unit (ReLU) activation function. In addition, skip connections directly concatenate the feature map in the downsampling layer to its corresponding upsampling layer according to space resolution to increase the amount of information in the upsampling steps.

To get more precise feature maps [30], we designed H-net to connect with HL-net in parallel. H-net has the same numbers of feature maps as HL-net and uses skip connections at the same depth as HL-net. We introduce exchange units across parallel H-net and HL-net such that each subnetwork repeatedly receives the information from the other parallel subnetwork.

The exchange unit contains convolution or upsampling or downsampling, so the feature maps that used to exchange information are converted to have the same resolution and channels. The upsampling contains a bilinear interpolation following a 3×3 convolution. Both upsampling and downsampling are used only once in an exchange unit. An example of exchange unit is shown in Fig. 6. Specially, the first convolution layer and the last convolution layer are shared by HL-net and H-net. All feature maps in H-net are 32 channels, while the number of channels of feature maps with different resolutions in HL-net are 32, 64, 128, 256, respectively.

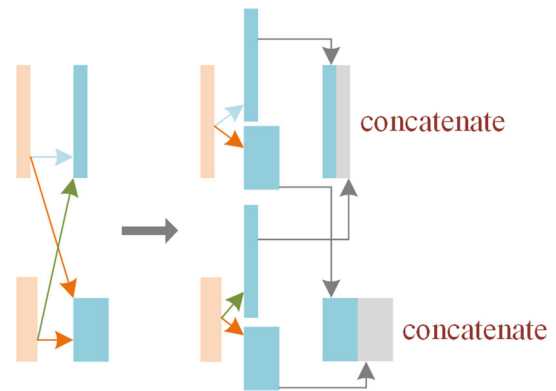


Fig. 6 The illustration of the exchange unit

2.3 Loss function

The loss function of the proposed method consists of three components: global loss L_{mse} for overall adjustment, structural similarity loss L_{ssim} for structural adjustment, and exposure control loss $L_{delight}$ for overexposure suppression. The total loss function for the proposed network is shown as follows:

$$L_{total} = L_{mse} + L_{ssim} + \lambda L_{delight} \tag{4}$$

where λ is used to control the degree of exposure suppression.

2.3.1 Global loss

MSE (mean square error) is to average the squared sum of the corresponding pixel errors between the correction image and the degraded image (which is used as the reference image in the proposed method). We use it to evaluate the change degree of the image. The smaller the value of MSE, the closer the

correction image is to the degraded image. Therefore, the global loss L_{mse} is expressed as:

$$L_{mse} = \frac{1}{H \times W} \|I_{cor} - I_{deg}\|_2^2 \quad (5)$$

where I_{cor} is the correction image, I_{deg} is the degraded image, $\|\cdot\|_2$ means L_2 norm, H and W are the height and width of the image.

2.3.2 Structural similarity loss

Images captured in the low-light condition often suffer from structure distortion problems [31]. In order to improve the quality of the enhanced image, we introduce structural similarity loss. The structure similarity (SSIM) [27] evaluates the similarity between two images in terms of luminance, contrast, and structure. The definition is shown below:

$$SSIM(cor,deg) = \frac{(2\mu_{cor}\mu_{deg} + C_1)(2\sigma_{cor,deg} + C_2)}{(\mu_{cor}^2 + \mu_{deg}^2 + C_1)(\sigma_{cor}^2 + \sigma_{deg}^2 + C_2)} \quad (6)$$

where the parameters cor and deg are simple representations of the correction image I_{cor} and the degraded image I_{deg} , μ_{cor} is the mean of the I_{cor} , μ_{deg} is the mean of the I_{deg} , σ_{cor} is the variance of the I_{cor} , σ_{deg} is the variance of the I_{deg} , $\sigma_{cor,deg}$ is the covariance of the I_{cor} and the I_{deg} , C_1 and C_2 are constants and take the default values ($C_1 = 0.0001$, $C_2 = 0.0009$).

The value range of SSIM is 0 to 1, and higher value means better similarity. Therefore, structural similarity loss L_{ssim} is expressed as:

$$L_{ssim} = 1 - \frac{(2\mu_{cor}\mu_{deg} + C_1)(2\sigma_{cor,deg} + C_2)}{(\mu_{cor}^2 + \mu_{deg}^2 + C_1)(\sigma_{cor}^2 + \sigma_{deg}^2 + C_2)} \quad (7)$$

2.3.3 Exposure control loss

In addition to use L_{mse} for global adjustment and L_{ssim} for structural adjustment, we also designed an exposure control loss $L_{delight}$ to restrain the overexposed regions. The I_{corv} and I_{degv} are V channels of HSV images of the corrected image I_{cor} and the degraded image I_{deg} , respectively. And the V channel is always used to represent the illumination of the image. The exposure control loss measures the difference of the average intensity value of the maximum 4% pixels between I_{corv} and I_{degv} , which is also the average difference of the brightest 4% pixels. $L_{delight}$ is defined as follows:

$$L_{delight} = (I_{cor}^{\text{meanmax}} - I_{deg}^{\text{meanmax}})^2 \quad (8)$$

where I_{cor}^{meanmax} is the mean of the maximum 4% pixels of I_{corv} , and I_{deg}^{meanmax} is that of I_{degv} .

In this study, we set $\lambda = 0.5$ experimentally.

3 Experiments

The network is built on TensorFlow framework, and the experiment is completed on a server with Intel(R) Xeon(R) CPU E5-2186 @ 3.80 GHz, Nvidia GeForce GTX 2080TI and 64G RAM. The pixel values of the training images are normalized to [0, 1], and then the images are randomly cropped into 48×48 blocks and fed into the training network. We use ADAM optimizer with default parameters, and the training could be completed within 3 minutes. The initial learning rate is 0.001, which decreases by 90% every 20 epochs. After 100 epochs of training, we get final output.

In this study, the model is trained on a public dataset LOL dataset [19]. The LOL dataset contains 500 low/normal-light image pairs, including 485 image pairs for training and 15 image pairs for evaluation. The LOL dataset is taken in real scenes, and the image resolution is $600 \times 400 \times 3$. This study selects 234 image pairs of different scenes from the 485 image pairs for the training process. In order to show the effectiveness and superiority of the proposed method, images from public datasets LOL dataset, [1,22], DICM [32] are selected to test. These images contain varieties of lighting conditions and include both synthesized and real scenes. In addition to the vision comparison with state-of-the-art methods, three metrics are also adopted to evaluate the performance of the proposed method.

3.1 Comparison with state-of-the-art methods

In this section, we analyze the performance of the proposed method and current state-of-the-art methods: BIMEF [15], LECARM [16], MF [17], LIME [11], Retinex-Net [19], and EnlightenGAN [24]. A list of experiments is conducted, and the enhancement results are shown in Figs. 7 and 8, which include different types of images. The visual comparison of images is elaborated, respectively.

3.1.1 Visual comparison

Figure 7 presents a visual comparison of different methods, and the low-light images are from LOL dataset but outside our training dataset. It can be seen that the brightness of the enhanced results of BIMEF and LECARM is insufficient. In the first row, the bookcase in the output of Retinex-Net has a slight color distortion and that of EnlightenGAN is a bit over-enhanced, while in the second row, the object edges in the results of MF, LIME, Retinex-Net are unnatural. To further demonstrate the superiority of our method, some details of the images in the third row are magnified and shown below. Obviously, the noise can be seen in all output except ours. And the results of the other methods are not smooth enough in details. On the whole, the proposed method could effectively

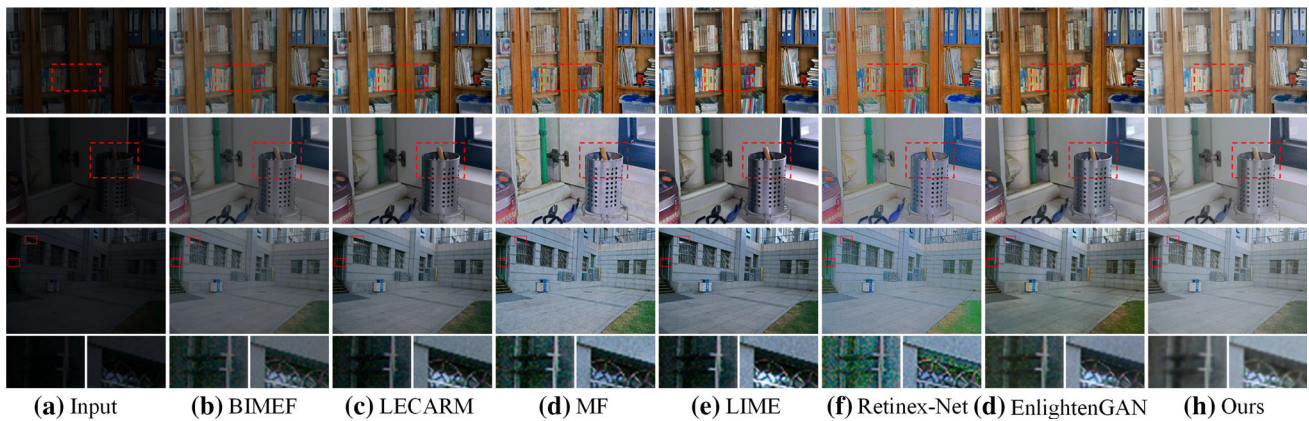


Fig. 7 Visual comparison of the proposed method and the state-of-the-art methods on images from LOL dataset but outside our training dataset

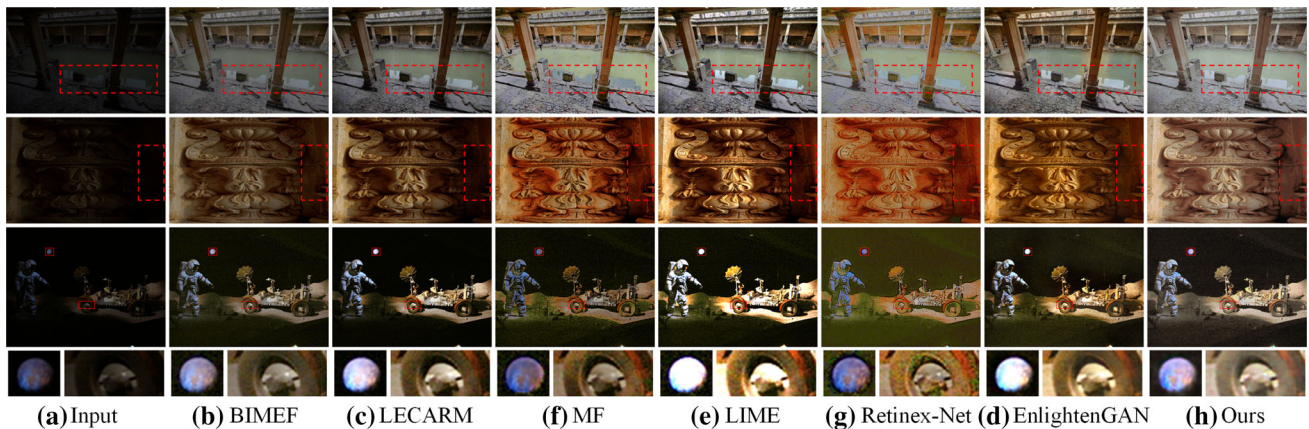


Fig. 8 Visual comparison of the proposed method and the state-of-the-art methods on images from [1,22] and DICM [31]

remove noise and restore color, and the enhanced images are also smoother.

Figure 8 shows the results on low-light images from [1,22] and DICM [32] datasets. In the first row, the MF and Retinex-Net over-enhance the input image. And the pillar in the output of EnlightenGAN has unexpected enhancement. In the second row, we can find insufficient enhancement in the results of BIMEF and LECARM. And the color distortion in the output of MF, Retinex-Net, and EnlightenGAN can be seen clearly. Some details of the images in the third row are also magnified and shown below. It can be seen that the result of Retinex-Net has a little color distortion in dark area. The details of the earth and car in the enhanced images of LECARM, LIME, and EnlightenGAN are lost. For BIMEF and MF, the enhanced results still exist areas of insufficient brightness. In summary, the proposed method has better comprehensive performance in denoising, preserving details, fully enhancing dark areas, and the enhanced results are also natural.

Table 2 Comparison of BIMEF, LECARM, MF, LIME, Retinex-Net, EnlightenGAN, and ours in PSNR, SSIM, NIQE

Method	PSNR	SSIM	NIQE
BIMEF	13.756	0.615	4.2584
LECARM	14.653	0.587	4.3221
MF	17.321	0.543	4.4578
LIME	17.480	0.579	4.4598
Retinex-Net	17.426	0.515	5.0954
EnlightenGAN	17.859	0.689	3.1526
Ours	20.298	0.769	3.1959

The best results are in bold

3.1.2 Evaluation

In order to further test the performance of the proposed method, we adopt peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and natural image quality evaluator (NIQE) to evaluate the image quality. A higher PSNR and SSIM values mean that the enhanced image is closer to the reference image, while a lower NIQE values indicate better

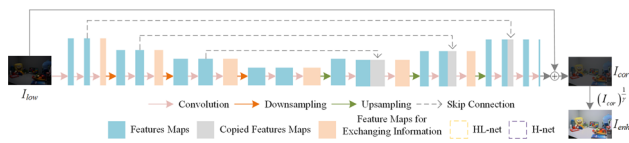


Fig. 9 The architecture of the proposed network without high-resolution subnetwork

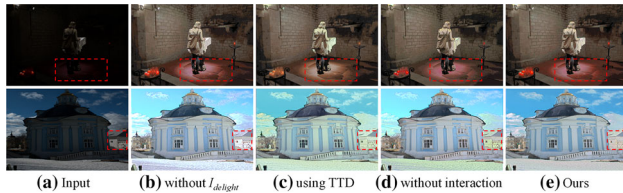


Fig. 10 Ablation study of the effect of $L_{delight}$, training data, and the interaction between the subnetworks

visual quality. In addition to the 15 images used for evaluation in the LOL dataset, another 8 images from LOL dataset (outside our training dataset) and [1] are selected and used as a synthesized testing dataset. The images from [1] are resized to $900 \times 600 \times 3$. The synthesized testing dataset contains more scenes, and the images in it have reference images. Therefore, all three metrics are used for quantitative evaluation. The comparison between our method and state-of-the-art methods is summarized in Table 2. It can be seen that our method has higher PSNR and SSIM values, and only EnlightenGAN has a little better NIQE values than ours. Overall, the proposed method has better enhancement results.

3.2 Ablation study

Ablation study is conducted to analyze the effect of exposure control loss $L_{delight}$, the low-light/degraded image pairs that used for training and the interaction between the subnetworks (the architecture with high-resolution subnetwork removed can be seen in Fig. 9). As shown in Fig. 10, without using $L_{delight}$, overexposure can be observed in brighter areas. For example, the details in the red region in the second row are lost due to overexposure. Furthermore, the results of using low/normal-light image pairs during training (denoted as using TTD, traditional training data, $\gamma = 1.0$) have serious color loss, while the output of the proposed method is more natural. In addition, we can see that without the interaction between subnetworks, the results have unsatisfactory performance in the details. The PSNR, SSIM, and NIQE values in Table 3 also show that using $L_{delight}$, the low-light/degraded image pairs and high-resolution subnetwork could produce better comprehensive performance on the testing dataset.

Table 3 Comparison of performance metrics for ablation study

Method	PSNR	SSIM	NIQE
without $L_{delight}$	20.107	0.766	3.3370
using TTD	19.630	0.767	3.8611
without interaction	19.944	0.735	3.2290
Ours	20.298	0.769	3.1959

The best results are in bold



Fig. 11 Results of Google Cloud Vision API. **a** Recognizing result of low-light image; **b** recognizing result of our enhanced image

3.3 Application

To further demonstrate the performance of our method in improving the accuracy of object recognition, we test our output on Google Cloud Vision API (<https://cloud.google.com/vision/>). As shown in Fig. 11, the API can recognize person and umbrella from our enhanced image, but not in low-light image. The original image is from [22].

4 Conclusion

In this paper, a new low-light image enhancement method based on normal-light image degradation is proposed. By replacing the traditional normal-light images with degraded images as the reference images, as well as building the multi-scale fusion network to exchange information across two parallel subnetworks, the proposed method is deemed to produce better effect in both color and detail recovery. Additionally, thanks to the exposure control loss, the overexposed regions are well restrained. Extensive experiments also demonstrate the superiority of the proposed method against state-of-the-art methods. The future work will focus on improving the robustness and optimizing the network to improve the generalization ability of the enhancement model.

This work was supported by Tianjin Intelligent Security Industry Chain Technology Adaptation and Application Project under Grant 18ZXZNGX00320.

References

1. Cai, J., Gu, S., Zhang, L.: Learning a deep single image contrast enhancer from multi-exposure images. *IEEE Trans. Image Process.* **27**(4), 2049–2062 (2018)
2. Jiang, Z., Li, H., Liu, L., Men, H., Wang, H.: A switched view of Retinex: deep self-regularized low-light image enhancement. *Neurocomputing* **454**, 361–372 (2021)
3. Liu, J., Xu, D., Yang, W., Fan, M., Huang, H.: Benchmarking low-light image enhancement and beyond. *Int. J. Comput. Vision* **129**(4), 1153–1184 (2021)
4. Asmare, M.H., Asirvadam, V.S., Hani, A.F.M.: Image enhancement based on contourlet transform. *SIViP* **9**(7), 1679–1690 (2015)
5. Chen, C., Chen, Q., Xu, J., Koltun, V.: Learning to see in the dark. In: *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 3291–3300 (2018)
6. Wang, W., Chen, Z., Yuan, X., Wu, X.: Adaptive image enhancement method for correcting low-illumination images. *Inf. Sci.* **496**, 25–41 (2019)
7. Cheng, H.D., Shi, X.J.: A simple and effective histogram equalization approach to image enhancement. *Digit. Signal Process.* **14**(2), 158–170 (2004)
8. Abdullah-Al-Wadud, M., Kabir, M.H., Dewan, M.A.A., Chae, O.: A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **53**(2), 593–600 (2007)
9. Land, E.H.: The Retinex theory of color vision. *Sci. Am.* **237**(6), 108–129 (1977)
10. Jobson, D.J., Rahman, Z., Woodell, G.A.: A multiscale Retinex for bridging the gap between color images and the human observation of scenes. *IEEE Trans. Image Process.* **6**(7), 965–976 (1997)
11. Guo, X., Li, Y., Ling, H.: LIME: low-light image enhancement via illumination map estimation. *IEEE Trans. Image Process.* **26**(2), 982–993 (2016)
12. Yang, J., Xu, Y., Yue, H., Jiang, Z., Li, K.: Low-light image enhancement based on Retinex decomposition and adaptive gamma correction. *IET Image Process.* **15**(5), 1189–1202 (2021)
13. Song, X., Huang, J., Cao, J., Song, D.: Multi-scale joint network based on Retinex theory for low-light enhancement. *Signal Image Video Process.* **15**, 1257–1264 (2021)
14. Jiang, X., Yao, H., Liu, D.: Nighttime image enhancement based on image decomposition. *SIViP* **13**(1), 189–197 (2019)
15. Ying, Z., Li, G., Gao, W.: A bio-inspired multi-exposure fusion framework for low-light image enhancement (2017). [arXiv preprint arXiv:1711.00591](https://arxiv.org/abs/1711.00591)
16. Ren, Y., Ying, Z., Li, T.H., Li, G.: LECARM: low-light image enhancement using the camera response model. *IEEE Trans. Circuits Syst. Video Technol.* **29**(4), 968–981 (2018)
17. Fu, X., Zeng, D., Huang, Y., Liao, Y., Ding, X., Paisley, J.: A fusion-based enhancing method for weakly illuminated images. *Signal Process.* **129**, 82–96 (2016)
18. Guo, C., Li, C., Guo, J., Loy, C. C., Hou, J., Kwong, S., Cong, R.: Zero reference deep curve estimation for low-light image enhancement. In: *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1780–1789 (2020)
19. Wei, C., Wang, W., Yang, W., Liu, J.: Deep retinex decomposition for low-light enhancement. In: *Proceeding of British Machine Vision Conference*, pp. 1–12 (2018)
20. Ma, S., Ma, H., Xu, Y., Li, S., Lv, C., Zhu, M.: A low-light sensor image enhancement algorithm based on HSI color model. *Sensors* **18**(10), 3583 (2018)
21. Huang, W., Zhu, Y., Huang, R.: Low light image enhancement network with attention mechanism and Retinex model. *IEEE Access* **8**, 74306–74314 (2020)
22. Wang, W., Wei, C., Yang, W., Liu, J.: GLADNet: Lowlight enhancement network with global awareness. In: *Proceeding of IEEE International Conference on Automatic Face & Gesture Recognition*, pp. 751–755 (2018)
23. Atoum, Y., Ye, M., Ren, L.: Color-wise attention network for low-light image enhancement. In: *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp. 506–507 (2020)
24. Jiang, Y., Gong, X., Liu, D., Cheng, Y., Fang, C., Shen, X., Yang, J., Zhou, P., Wang, Z.: EnlightenGAN: deep light enhancement without paired supervision. *IEEE Trans. Image Process.* **30**, 2340–2349 (2019)
25. Hua, W., Xia, Y.: Low-Light Image enhancement based on joint generative adversarial network and image quality assessment. In: *Proceeding of International Congress on Image and Signal Processing, BioMedical Engineering and Informatics*, pp. 1–6 (2018)
26. Cheng, Y., Yan, J., Wang, Z.: Enhancement of weakly illuminated images by deep fusion networks. In: *Proceeding of IEEE International Conference on Image Processing*, pp. 924–928 (2019)
27. Wang, Z., Bovik, A.C., Sheikh, H.R., Simoncelli, E.P.: Image quality assessment: from error visibility to structural similarity. *IEEE Trans. Image Process.* **13**(4), 600–612 (2004)
28. Mittal, A., Soundararajan, R., Bovik, A.C.: Making a ‘completely blind’ image quality analyzer. *IEEE Signal Process. Lett.* **20**(3), 209–212 (2013)
29. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: *Proceeding of Medical Image Computing and Computer-Assisted Intervention*, pp. 234–241 (2015)
30. Sun, K., Xiao, B., Liu, D., Wang, J.: Deep high-resolution representation learning for human pose estimation. In: *Proceeding of IEEE Conference on Computer Vision and Pattern Recognition*, pp. 5693–5703 (2019)
31. Lv, F., Lu, F., Wu, J., Lim, C.: MBLLN: Low-light image-video enhancement using CNNs. In: *Proceeding of British Machine Vision Conference*, pp. 1–13 (2018)
32. Lee, C., Lee, C., Kim, C. S.: Contrast enhancement based on layered difference representation. In: *Proceeding of IEEE International Conference on Image Processing*, pp. 965–968 (2012)

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