LOW-LIGHT IMAGE ENHANCEMENT VIA FEATURE RESTORATION

Yang Yang⋆ *Yonghua Zhang*† *Xiaojie Guo*⋆,§

⋆ College of Intelligence and Computing, Tianjin University
†College of Artificial Intelligence, Henan University

ABSTRACT

Besides poor visibility, under-exposed images often suffer from severe noise and color distortion. Most existing Retinex-based methods deal with the noise and color distortion via some careful designs to denoising and/or color correction. In this paper, we propose a simple yet effective network from the perspective of feature map restoration to mitigate such issues without constructing any explicit modules. More concretely, we build an encoder-decoder network to reconstruct images, while a feature restoration subnet is introduced to transform the features of low-light images to those of corresponding clear ones. The enhanced images are consequently acquired through assembling the restored features by the decoder, in which, the noise and possible color distortion can be greatly remedied. Extensive experiments on widely-used datasets are conducted to validate the superiority of our design over other state-of-the-art alternatives both quantitatively and qualitatively. Our code is available at `https://github.com/YaN9-Y/FRLIE`.

***Index Terms*—** Low-light image enhancement, feature restoration

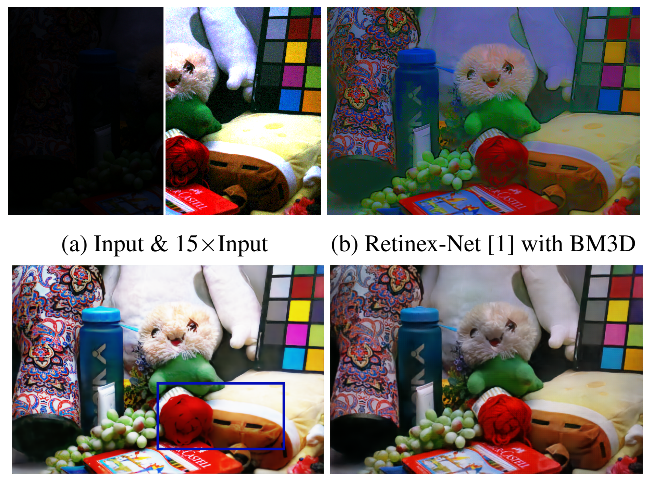
1. INTRODUCTION

Images captured in low-light conditions are commonly of poor visual quality. Enhancing the low-light images in the wild is a challenging task due to the variety of degradation. As shown in Fig.1, when simply increasing the intensity of a low-light image, the low-contrast, color-distortion and heavy noise will be considerably amplified. Comparing to the low-brightness, these kinds of degradation as side-effect of under-exposure are also difficult to deal with.

1.1. Related Work

A variety of algorithms have been proposed to enhance the visual perception of low-light images. In what follows, we provide a brief review of existing methods related to ours.

Traditional methods. Inspired by Retinex theory [3], which assumes that sensations of color have a strong correlation with reflectance and a visible image can be modeled as the product of reflectance and illumination. Wang *et al.* proposed a method to preserve the naturalness of enhanced image with lightness-order-error measure [4]. Guo *et al.* focused on only estimating the illumination map from an initial one with structure prior [5]. These methods do not explicitly consider the degradations like noise and color distortion previously hidden in the darkness.



(a) Input & 15×Input (b) Retinex-Net [1] with BM3D

(c) KinD [2] (d) Ours

Fig. 1: Visual comparison with Retinex-based low-light image enhancement methods. The boxed area of (c) suffers from over-smoothness and overexposure while our method recovers the details faithfully in (d).

Deep learning based methods. With the promotion of deep learning, lots of approaches based on deep models for low-light image enhancement have been proposed. Lore *et al.* developed a deep LLNet [6] to simultaneously enhance and denoise low-light images basing on a synthetic dataset. Wei *et al.* proposed a deep network Retinex-Net [1], to decompose the low-light image into reflectance map and illumination map. However, they use BM3D to relieve the noise effect and the enhanced results are visually unnatural. Zhang *et al.* introduced an extra restoration net to deal with the noise and color distortion and an illumination adjustment net to change the light level [2]. Chen *et al.* developed a pipeline for processing low-light images in RAW format based on end-to-end training of an encoder-decoder network [7], which can jointly

§ Corresponding author.

This work was supported by the National Natural Science Foundation of China under Grant no. 62072327, and TSTC under Grant no. 20JC-QNJC01510.

deal with noise and color distortion. Wang *et al.* presented DUPE to enhance [8] underexposed photos by only estimating illumination maps. Jiang *et al.* proposed a deep light enhancement net without paired supervision by using GANs [9].

1.2. Consideration and Contribution

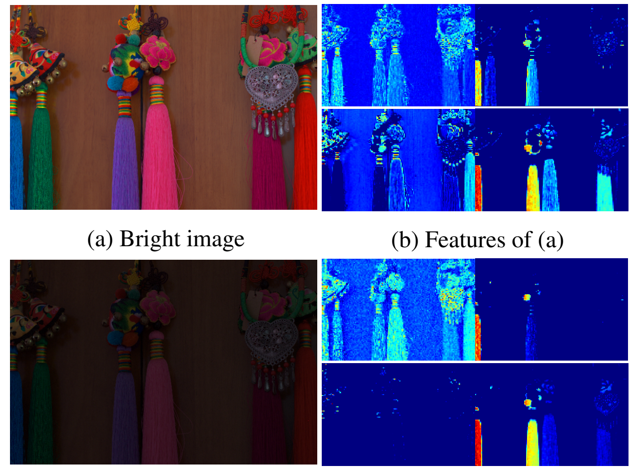
In the literature, Retinex-based methods have dominated the mainstream of both traditional and deep-learning based low-light image enhancement techniques. Nevertheless, such processing steps do not take other degradation factors into consideration. As a result, these methods tend to produce noisy and color-distorted images or have to apply additional denoising and color correction steps [2, 10], which would increase the complexity of the model.

Since Retinex models suffer from the limitations above, we propose to handle the low-light enhancement from the perspective of feature-level restoration. As shown in Fig.2 (a) and (c), the contents in the image space are often overwhelming. And the gap between low-light and normal-light images is large so it is hard to learn a direct end-to-end mapping. However, things become totally different in the deep feature space. As can be seen in Fig.2 (b) and (d), each feature channel mostly concentrates on a certain part of image rather than the whole image content. Besides, the degradation also separately exists on the feature maps. For instance, the left-top feature map in (d) suffers from severe noise and features on right-top and right-bottom are less significant than in (b), which are similar to the low-light effect. Generally speaking, the complex contents and hybrid degradation are decoupled in the feature space. Thus it could be easier for the network to learn the restoration in feature-level. Moreover, by learning the mapping between low-light and normal-light deep features, the side degradation like noise and color-distortion will be solved altogether. To make full use of such characteristics of CNNs, we propose a network trained in two-stage. In the first stage, an encoder-decoder network learns to reconstruct the normal-light images. In the second stage, the low-light and normal-light features are extracted using same encoder in the stage one from the corresponding images. Then a feature restoration subnet is applied to learn the mapping from low-light features to the normal-light ones. Finally the 'repaired' features are upsampled by the same decoder in the stage one to achieve the enhanced results.

The key contributions of this work can be summarized as, we propose a new method for low-light enhancement via feature restoration. Unlike Retinex-based methods, our approach can produce degradation-free enhanced results without explicit degradation removal design. Extensive experiments on public datasets show that our method achieves state-of-the-arts performance both quantitatively and qualitatively.

2. METHODOLOGY

As schematically illustrated in Fig.3, the network enhances the observed low-light images by mapping the dark features to the normal-light ones. We will detail the network in the following subsections.



(a) Bright image (b) Features of (a)

(c) Low-light image (d) Features of (c)

Fig. 2: Feature visualization. The features are extracted by the trained encoder as described in Sec.2.1. The features are normalized by min-max method on each channel.

2.1. Image reconstructor

To obtain the normal-light features as supervision, we train an image reconstructor with normal-light images I_h in the first place. The reconstructor has an encoder-decoder architecture. As noted by gray arrows in Fig.3, the reconstructor first extracts deep features of normal-light images, and then maps them back to the RGB space to reconstruct the original image. Denoting the encoder as $\mathcal{E}$ and the decoder as $\mathcal{D}$, the reconstruction is learned through a simple ℓ_2 loss as follows:

$$L_{rec} := \|\mathcal{D}(\mathcal{E}(I_h)) - I_h\|_2^2, \tag{1}$$

where $\|\cdot\|_2$ denotes the ℓ_2 norm. The architecture is also very simple. The encoder downsamples images with a few convolution blocks. The features are then upsampled by a series of pixelshuffle blocks in the decoder.

2.2. Feature restoration network

After the first-stage training, we obtain a feature extractor $\mathcal{E}$ and feature decoder $\mathcal{D}$. For paired normal light and low-light images I_h and I_l, corresponding bright features F_h and dark features F_l can be acquired through the trained encoder $\mathcal{E}$. To further enhance the low-light images, we build a feature restoration network aiming to learn the transformation from the dark and noisy features to its bright and clear form, which can be formulated as:

$$F_r = \text{FRNet}(F_d), \tag{2}$$

where FRNet$(\cdot)$ stands for the feature restoration network. The FRNet comprises of 4 Feature Restoration Blocks (FRBlock). As illustrated in Fig.3, the FRBlock is a stacking of a

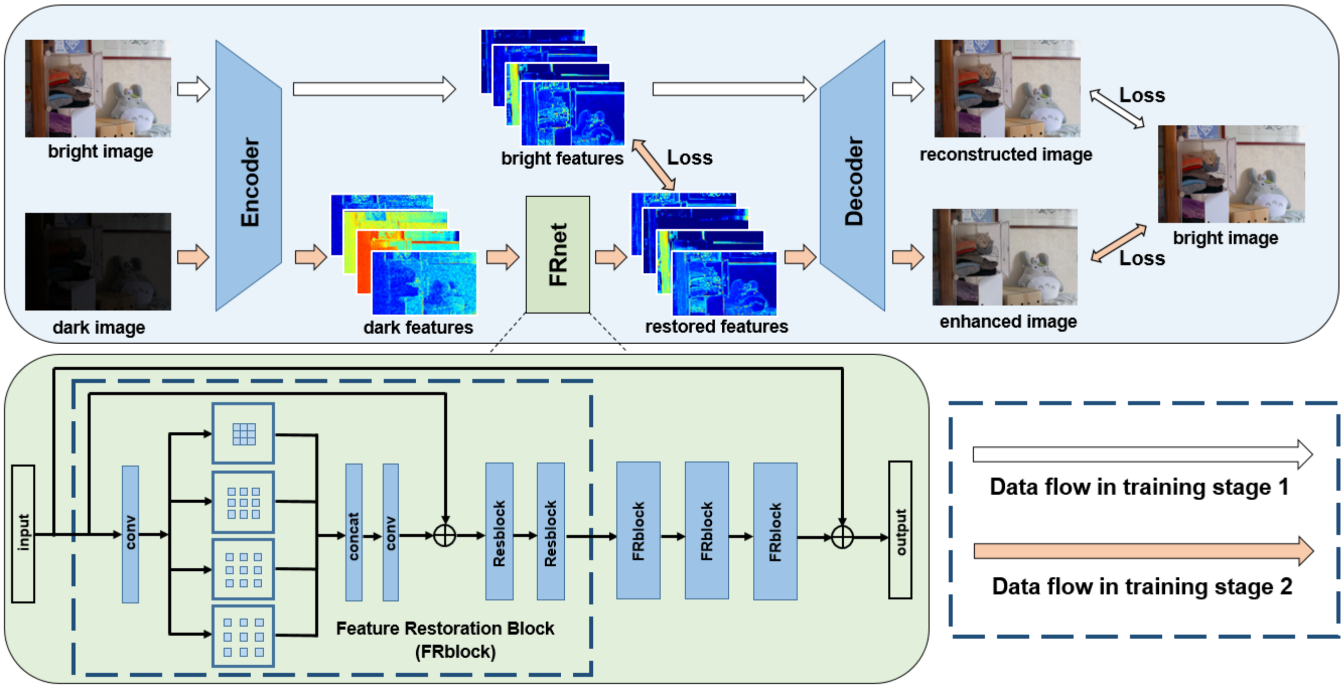


Fig. 3: The architecture of the proposed model.

context block [11] and two Resblocks [12]. The context block combines results of convolution with various dilation rates to enlarge the receptive field. The whole feature restoration net is learned through predicting the residual of each feature.

Normally speaking, there is no 'ground-truth' features as supervision for the network to learn the transformation. But remember we have features F_h extracted from normal-light images, which can be utilized as pseudo ground truth features. Now that the supervision features are acquired, we are able to apply the feature matching loss as follows:

$$L_{FM} := \|F_r - F_h\|_1, \tag{3}$$

where $\|\cdot\|$ designates the ℓ_1 norm. The feature matching loss enforces the restored features be closer to the features directly extracted from normal-light images. Following this way, the enhanced image acquired by decoding the restored features should also be similar to the normal-light images, so we utilize the decoding loss as:

$$L_{\mathcal{D}} := \|\mathcal{D}(F_r) - I_h\|_2^2. \tag{4}$$

Comparing to ℓ_2 distance, the SSIM can better reflect structure similarity. So we also adopt SSIM loss as follows:

$$L_{ssim} := 1 - SSIM(\mathcal{D}(F_r), I_h). \tag{5}$$

Besides, to overcome the color distortion, we propose to apply the hue loss for better color preservation. The hue loss is defined as follows:

$$L_{hue} := \frac{1}{180}\|180 - |180 - \mathcal{H}(\mathcal{D}(F_r)) - \mathcal{H}(I_h))|\,\|_1, \tag{6}$$

where $\mathcal{H}(\cdot)$ denotes the hue channel in the HSV space.

The total loss of the feature restoration network yields:

$$L_{FR} := \lambda_D L_D + \lambda_{FM} L_{FM} + \lambda_{SSIM} L_{SSIM} + \lambda_{Hue} L_{Hue}. \tag{7}$$

We set $\lambda_D = 100$, $\lambda_{FM} = 0.2$, $\lambda_{SSIM} = 0.5$ and $\lambda_{Hue} = 0.05$ in our experiments. Please note that in the second stage, both $\mathcal{D}$ and $\mathcal{E}$ are frozen, and we only optimize parameters of the feature restoration network. On both training stages, the model is optimized by the Adam optimizer [19] with batch-size of 16, $\beta_1 = 0.5$ and $\beta_2 = 0.9$ and the learning rate decays from 2×10^{-4} to 2×10^{-5} after 15000 epochs.

3. EXPERIMENT

3.1. Datasets and metrics

In this section, we evaluate the low-light enhancement performance of our proposed method. We adopt the training set of LOL [1] to train our model and evaluate our methods on the testing set of LOL, DICM [20], LIME [5], MEF [21] and NPE [4]. Several state-of-the-arts methods are selected as competitors. We employ PSNR, SSIM [22] and NIQE [23] as criteria for quantitative evaluation. Higher SSIM and PSNR indicate better quality while for NIQE the lower the better.

3.2. Results and comparisons

Quantitative comparison: Table 1 reports the performance of our method and other competitors. As can be seen, our methods significantly outperforms other methods in terms of all metrics on LOL testing images. Table 2 shows the

Metrics	BIMEF [13]	CRM [14]	Dong [15]	LIME [5]	MF [16]	RRM [10]	DUPE [8]
PSNR ↑	13.875	17.203	16.716	16.759	16.966	13.876	16.798
SSIM ↑	0.5771	0.6442	0.5824	0.5644	0.6422	0.6577	0.5187
NIQE ↓	7.6992	8.0182	9.1358	9.1272	9.7125	5.9416	8.4736
Metrics	SRIE [17]	Retinex-Net [1]	NPE [4]	GLAD [18]	KinD [2]	Ours w/o FR	Ours
PSNR ↑	11.855	16.774	16.970	19.718	20.726	21.489	**22.319**
SSIM ↑	0.4979	0.5594	0.5894	0.7035	0.8103	0.7944	**0.8256**
NIQE ↓	7.5349	9.7289	9.1352	6.7972	4.1352	3.9101	**3.5280**

Table 1: Quantitative comparison on the LOL dataset in PSNR, SSIM and NIQE. The best results are highlighted in bold.

(a) Input (b) Dong [15] (c) SRIE [17] (d) GLAD [18] (e) CRM [14] (f) LIME [5]

(g) MF [16] (h) Ret-Net [1] (i) NPE [4] (j) DUPE [8] (k) KinD [2] (l) Ours

Fig. 4: Visual comparison on an image from the MEF dataset with state-of-the-art methods.

Metric	NIQE ↓			
Datasets	LIME-data	NPE-data	MEF-data	DICM-data
CRM [14]	3.7431	3.6800	<u>3.1899</u>	3.3624
Dong [15]	4.2465	3.8562	4.5499	4.3412
LIME [5]	4.3473	3.8422	3.8765	3.6642
MF [16]	4.1025	3.6800	3.4256	3.4533
RRM [10]	3.9095	3.9466	3.9385	3.3186
SRIE [17]	**3.4690**	**3.1788**	3.2192	3.0951
Ret-Net [1]	4.9078	4.0676	5.0047	4.7120
NPE [4]	3.8400	3.4455	3.5884	3.4304
GLAD [18]	3.9028	*3.2026*	3.1994	<u>3.0846</u>
DUPE [8]	*3.5318*	3.3327	**3.1025**	3.1628
KinD [2]	<u>3.6292</u>	3.3668	3.3274	*3.0124*
Ours	3.7993	<u>3.2850</u>	*3.1561*	**2.8276**

Table 2: Quantitative comparison in terms of NIQE. The best results are highlighted in bold, the second best are in italic and the third best are underlined.

quantitative results on DICM, LIME, MEF and NPE datasets. Since there are no reference images in these datasets, we simply apply NIQE to compare the performance among different methods. Our method ranks the first place on DICM, the second place on MEF and the third place on NPE. All conducted quantitative comparisons show the superiority of our proposed methods towards the others.

Qualitative comparison: We also conduct qualitative comparison on MEF dataset. As shown in Fig.4, many Retinex-based methods suffer from severe noise, color distortion and under-enhancement, while our results are of proper exposure, and free of noises without explicit denoising procedure.

(a) Results w/o FR (b) Results of our method

Fig. 5: The comparison of results produced by the model with and without feature restoration. Zoom in for better view.

To validate the effectiveness of learning to restore features, we remove the feature matching loss and directly train the whole network in an end-to-end manner. As shown on last two columns in Table 1, the performance of our network without feature restoration drops significantly on all the involved metrics. Figure 5 also corroborates that it suffers from local artifacts and more severe color distortion compared to the original network setting. From this experiment, we can conclude that the feature restoration is crucial in our model.

4. CONCLUSION

In this work, we propose a novel low-light image enhancement method via feature restoration. A feature restore subnet is utilized to recover the noisy and dark features to bright ones. The brightened images are then acquired by upsampling the restored features. Superior to most Retinex-based methods, our approach can produce degradation-free results without explicit degradation removal procedure. The effectiveness of the proposed method has been experimentally confirmed.

5. REFERENCES

[1] C. Wei, W. Wang, W. Yang, and J. Liu, "Deep Retinex decomposition for low-light enhancement," in *BMVC*, 2018.

[2] Y. Zhang, J. Zhang, and X. Guo, "Kindling the darkness: A practical low-light image enhancer," in *ACM MM*, 2019, pp. 1632–1640.

[3] E. H. Land, "The Retinex theory of color vision," *Scientific American*, vol. 237, no. 6, pp. 108–128, 1977.

[4] S. Wang, J. Zheng, H. Hu, and B. Li, "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE TIP*, vol. 22, no. 9, pp. 3538–3548, 2013.

[5] X. Guo, Y. Li, and H. Ling, "LIME: Low-light image enhancement via illumination map estimation," *IEEE TIP*, vol. 26, no. 2, pp. 982–993, 2017.

[6] K. G. Lore, A. Akintayo, and S. Sarkar, "LLNet: A deep autoencoder approach to natural low-light image enhancement," *Pattern Recognition*, vol. 61, pp. 650–662, 2017.

[7] C. Chen, Q. Chen, J. Xu, and V. Koltun, "Learning to see in the dark," in *CVPR*, 2018, pp. 3291–3300.

[8] R. Wang, Q. Zhang, C. Fu, X. Shen, W. Zheng, and J. Jia, "Underexposed photo enhancement using deep illumination estimation," in *CVPR*, 2019, pp. 6849–6857.

[9] Y. Jiang, X. Gong, D. Liu, Y. Cheng, C. Fang, X. Shen, J. Yang, P. Zhou, and Z. Wang, "EnlightenGAN: Deep light enhancement without paired supervision," *arXiv: 1906.06972*, 2019.

[10] M. Li, J. Liu, W. Yang, X. Sun, and Z. Guo, "Structure-revealing low-light image enhancement via robust Retinex model," *IEEE TIP*, vol. 27, no. 6, pp. 2828–2841, 2018.

[11] Meng Chang, Qi Li, Huajun Feng, and Zhihai Xu, "Spatial-adaptive network for single image denoising," in *ECCV*, 2020, pp. 171–187.

[12] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Deep residual learning for image recognition," in *CVPR*, 2016, pp. 770–778.

[13] Z. Ying, L. Ge, and W. Gao, "A bio-inspired multi-exposure fusion framework for low-light image enhancement," *arXiv: 1711.00591*, 2017.

[14] Z. Ying, L. Ge, Y. Ren, R. Wang, and W. Wang, "A new low-light image enhancement algorithm using camera response model," in *ICCVW*, 2018, pp. 3015–3022.

[15] X. Dong, Y. Pang, and J. Wen, "Fast efficient algorithm for enhancement of low lighting video," in *ICME*, 2011, pp. 1–6.

[16] X. Fu, D. Zeng, H. Yue, Y. Liao, X. Ding, and J. Paisley, "A fusion-based enhancing method for weakly illuminated images," *Signal Processing*, vol. 129, pp. 82–96, 2016.

[17] X. Fu, D. Zeng, Y. Huang, X. Zhang, and X. Ding, "A weighted variational model for simultaneous reflectance and illumination estimation," in *CVPR*, 2016, pp. 2782–2790.

[18] W. Wang, W. Chen, W. Yang, and J. Liu, "GLADNet: Low-light enhancement network with global awareness," in *FG*, 2018.

[19] Diederik P Kingma and Jimmy Ba, "Adam: A method for stochastic optimization," *arXiv:1412.6980*, 2014.

[20] C. Lee, C. Lee, and C. S. Kim, "Contrast enhancement based on layered difference representation of 2D histograms," *IEEE TIP*, vol. 22, no. 12, pp. 5372–5384, 2013.

[21] K. Ma, K. Zeng, and Z. Wang, "Perceptual quality assessment for multi-exposure image fusion," *IEEE TIP*, vol. 24, no. 11, pp. 3345–3356, 2015.

[22] Zhou Wang, Alan C Bovik, Hamid R Sheikh, Eero P Simoncelli, et al., "Image quality assessment: from error visibility to structural similarity," *IEEE TIP*, vol. 13, no. 4, pp. 600–612, 2004.

[23] A. Mittal, R. Soundararajan, and A. Bovik, "Making a completely blind image quality analyzer," *IEEE SPL*, vol. 20, no. 3, pp. 209–212, 2013.