

Low-light image enhancement with contrast regularization

Jinfeng Li*

College of Computer Science & Technology, Qingdao University, Qingdao 266071, Shandong, China.

* Corresponding author: Email: yasuierra@163.com.

Abstract: Because the processing of existing low-light images undergoes multiple sampling processing, there is serious information degradation, and only clear images are used as positive samples to guide network training, low-light image enhancement processing is still a challenging and unsettled problem. Therefore, a multi-scale contrast learning low-light image enhancement network is proposed. First, the image generates rich features through the input module, and then the features are imported into a multi-scale enhancement network with dense residual blocks, using positive and negative samples to guide the network training, and finally using the refinement module to enrich the image details. Experimental results on the dataset show that this method can reduce noise and artifacts in low-light images, and can improve contrast and brightness, demonstrating its advantages.

Keywords: Deep Learning; Low-light enhancement; Residual network; Contrast regularization.

1. Introduction

As an important medium for perceiving the world, the quality of images often determines the performance of algorithms, and the images taken under low-light conditions contain a lot of noise, which affects the information acquisition in the application process of high-level visual tasks, such as semantic segmentation and target recognition. There are many solutions for low-light image enhancement, which are divided into traditional methods and deep learning-based methods according to different algorithmic concepts.

1.1. Low-light enhancement based on traditional methods

Traditional approaches to low-light image enhancement are basically divided into two categories: methods based on distribution mapping [1] and methods based on model optimization. The distribution mapping method uses curve transformation, histogram equalization and other means to improve the pixel distribution of the image to improve the brightness and clarity of the image, but does not consider adjacent pixels, resulting in color distortion and abnormal details. The core of the model-based optimization method is to take the data derived from the physical imaging law as the basic component and further use the existing optimization technology to solve. For example, Retinex theory [2,3] based on color sensitivity consistency, set the original image S is the product of light image L and reflected image R , and then decompose R to eliminate the effect of light unevenness and improve image quality. However, this method does not consider the saturation component and the luminance component, and there is still a problem of blurred detail in low-light images.

1.2. Low-light enhancement based on deep learning

In order to overcome the drawbacks of the above traditional methods, and with the development of deep learning [4-6], the relationship between low-light input and enhanced output has

become a mainstream low-light image enhancement mode by heuristically designing network structures. Two-stage object detection algorithms represented by RCNN [7], Fast RCNN [8], Faster RCNN [9] and single-stage object detection algorithms represented by YOLO [10] and SSD [11] have emerged, but these methods focus on individual performance enhancements. Detail recovery and color correction for low-light images are still a huge challenge.

2. Proposed Method

2.1. The main network structures

The approach we propose is the network architecture shown in Figure 1, which is an end-to-end model. In order for the network to learn more features of multiple dimensions, a 3×3 convolutional layer is used to increase the number of feature dimensions. In order to enable information between the different dimensions to communicate, we borrowed the idea of up sampling and down sampling in U-Net [12] and then used the FA block in FFA-Net [13] with efficient FA blocks as our base blocks. Unlike FFA-Net, in order to improve computational efficiency and reduce memory storage, we reduced the 57 FA blocks in FFA-Net to 6 FA blocks. Specifically, a 4x down sampling operation, such as a regular convolution in step 1 and two convolutional layers with two steps in 2, causes dense FA blocks to learn feature representations in low-resolution spaces, and then uses the corresponding 4x up sampling and a regular convolution to generate the recovered image.

In Figure 1, we represent the low-light image, the corresponding recovered image generated by the low-light enhancement network, and its sharp image as negatives, anchor points, and positives, respectively. Contrast regularization brings the resulting corresponding recovery diagram closer to the sharp image and away from the blurry image. And the anchor image can be constrained to a closed upper and lower bounds by contrast learning, which will help the low-light enhancement network to approach the positive image and away from the negative image faster. Finally, a refinement processing module, the GRDB module, is used to

output the final low-light enhancement image through a convolutional 3×3 .

The overall model of the end-to-end single-image low-light enhancement approach can be represented as follows:

$$\arg \min_w \|J - \phi(I, w)\| + \beta \rho(\phi(I, w)) \quad (1)$$

where I is the low-light image, J is the corresponding sharp image, and $\phi(\cdot, w)$ is the low-light enhancement network

with the parameter w . $\|J - \phi(I, w)\|$ is a data fidelity term that typically uses losses based on the L1/L2 norm. $\rho(\cdot)$ is a regularization term that produces naturally smooth defogging images, where TV norm [14, 15] and DCP prior [14, 15] are widely used for regularization terms. β is a penalty parameter used to balance the data fidelity and regularization terms.

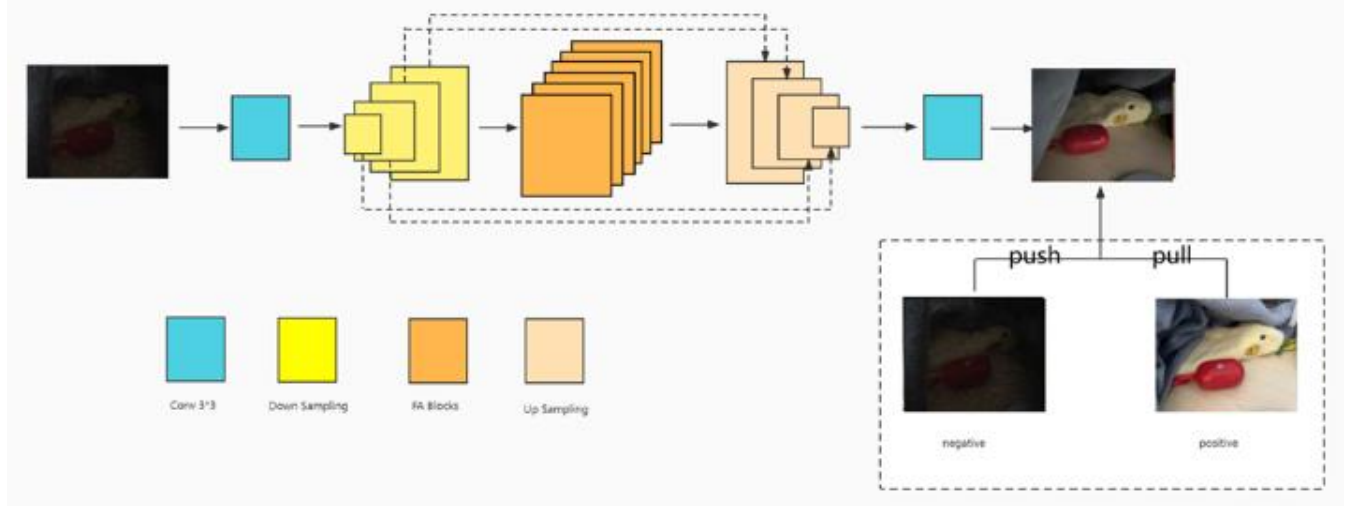


Figure 1. The network structure diagram of the algorithm in this paper

2.2. Loss function

In this paper, SSIM loss and Smooth L1 loss are used to effectively restore low-light images.

As SSIM refers to the overall visual effect, it measures the difference between the original and enhanced images for the purpose of improving their quality.

$$L_{SSIM} = 1 - \sum_{m=1}^M \frac{2\mu_x\mu_y + C_1}{\mu_{x^2} + \mu_{y^2} + C_1} \cdot \frac{2\sigma_{xy} + C_2}{\sigma_{x^2} + \sigma_{y^2} + C_2} \quad (2)$$

Smooth L1 loss can limit the gradient in two ways: the gradient value is not too large when the difference between the prediction frame and the ground truth is too large. The gradient value is small enough when the difference between the prediction frame and ground truth is small.

$$L_{smoothL1} = \begin{cases} \frac{1}{N} \sum_{i=1}^N \|x, y\|_1 \\ \frac{1}{N} \sum_{i=1}^N \|x, y\|_2 \end{cases} \quad (3)$$

The total loss function of the network is:

$$L_{total} = L_{SSIM} + L_{smoothL1} \quad (4)$$

3. Experiments and Results Analysis

3.1. Experimental details

The experiment was conducted on Ubuntu 20.04.3 with a NVIDIA RTX 3090 24GB \times 2. Pytorch-gpu is used for the deep learning architecture, epochs are 150, and batch size is 8. The models are all optimized using the ADAM optimizer with an initial learning rate of 0.0005.

The LOL dataset is the training set for the experiment. A dataset containing 1000 low-light and normal exposure pairs with an image size of 400x400x3, and 985 pairs are used in training and 15 pairs for validation is included. It consists of the evaluation sets of LOL and SICE, as well as the data sets of Exrdark[16], DICM[17], VV[18], NPE[19], MEF[20] and LIME[6].

Peak Signal-to-Noise Ratio (PSNR), Structural Similarity (SSIM), Natural Image Quality Evaluation (NIQE) and Root Mean Square Error (RMSE) are used as objective evaluation

indicators. PSNR and SSIM values are higher when the enhanced image is better. An enhanced image is of higher quality if the NIQE and RMSE values are low.

There are 8 mainstream low-light enhancement algorithms in the current research field: LIME, BIMEF, RetinexNet, SDD, CSDGAN, RUAS, KinD++, EnlightenGAN. Images are tested using models or source code from the original literature.

3.2. Comparison Experiment

The results of the reference LOL and SICE datasets are shown in Table 1. In both PSNR and SSIM metrics, this algorithm achieves the best results, and in LOE metrics, it holds the second place, indicating that this algorithm significantly enhances the quality of low-light images. Table 2 shows the results of NIQE metrics on 6 datasets. Despite not achieving the optimal results on all datasets, the algorithm in this paper has a high advantage on average and does achieve the optimal results on most datasets.

Figure 2 is a comparison chart of the enhancement effect on 4 datasets, and it can be seen that the image color saturation after the enhancement of this method is high, there is no obvious noise and artifacts, and it has a rich detail texture. MBLEEN and SDD methods have the phenomenon of poor brightness, low saturation and blur of the enhanced image. The image enhanced by the Retinex-Net method has obvious color distortion problems, resulting in blurring of some areas in the graph and amplifying the noise in the image; The EnlightenGAN method has a good effect of enhancement.

but there will be some problems of poor exposure, and the halo problem is more serious; Kinder++ has improved on color distortion, but there is still a problem that the enhanced image is not realistic and will amplify the noise in local areas.

In summary, our model has achieved excellent performance in the comprehensive performance of multi-objective enhancement tasks on the test data set, which can adaptively increase the global brightness and avoid overexposure; Can correct the color distortion in low-light images, enhance the

result of rich and full color, and the visual effect is better; Has generalization enhanced for low-light images of real scenes.

Table 1. Comparison of indicators of different methods on reference datasets (Where means higher value is better, means lower value is better, red means the best result, blue means the second-best result)

Dataset	LOL			SICE		
	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
LIME	13.374	0.508	37.2	15.798	0.625	45.6
BIMEF	13.921	0.728	40.3	13.544	0.533	56.1
RetinexNet	17.612	0.653	34.6	15.784	0.616	41.3
SDD	13.352	0.640	60.4	13.388	0.533	59.5
CSDGAN	9.111	0.331	90.1	10.464	0.381	81.9
RUAS	16.532	0.527	51.2	10.578	0.396	43.6
EnlightenGAN	17.546	0.664	38.3	14.768	0.611	50.3
KinD++	17.789	0.766	34.8	14.671	0.513	49.2
Ours	21.365	0.793	20.9	16.453	0.664	42.1

Table 2. Comparison of NIQE indicators of different methods (↑ means higher value is better, means lower value is better, red means the best result, blue means the second-best result)

Metric	Exrdark	DICM	LIME	VV	NPE	MEF	AVERAGE
LIME	3.855	3.459	4.077	2.920	3.595	3.025	3.489
BIMEF	3.829	3.541	3.861	3.239	3.607	3.333	3.568
RetinexNet	4.451	4.193	4.825	2.983	4.364	3.982	4.133
SDD	4.120	3.937	4.397	3.281	3.701	3.857	3.882
CSDGAN	5.591	5.782	5.730	4.730	5.179	7.338	5.725
RUAS	4.404	5.204	5.305	5.326	5.380	5.617	5.206
EnlightenGAN	3.691	3.563	3.664	4.237	3.731	3.221	3.684
KinD++	4.110	3.804	4.722	2.806	3.593	3.737	3.795
Ours	3.592	3.512	3.763	3.115	3.359	3.391	3.455

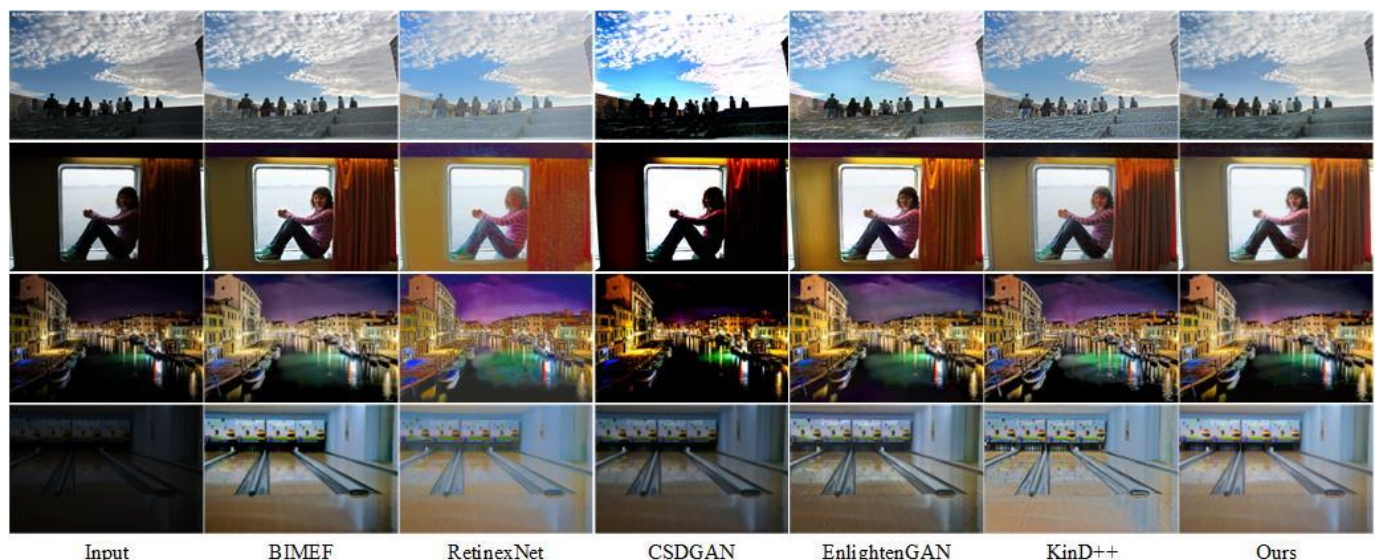


Figure 2. Visual comparison of different methods.

4. Conclusion

Aiming at the problem of low light enhancement, this paper proposes a new multi-scale low-light image enhancement network with regularized contrast. Two positive and negative samples are used to guide the network in learning. Through experimental verification, the proposed algorithm is far superior to most existing low-light image enhancement

algorithms, which enhances the quality of low-light images and provides data for subsequent high-order picture processing tasks. In the next work, we will consider optimizing the computing performance of the network so that it can be combined with more realistic scene information, and try to use the model for low-light video processing.

References

- [1] Kim, Y.T., "Contrast enhancement using brightness preserving bihistogram equalization," *IEEE transactions on Consumer Electronics*, 1-8(1997).
- [2] Land, E. and McCann, J., "Lightness and retinex theory," *Josa*, vol. 61, no. 1, pp. 1-11, (1971).
- [3] Guo, X., Li, Y. and Ling, H., "Lime: Low-light image enhancement via illumination map estimation," *IEEE Transactions on image processing*, vol. 26, no. 2, pp. 982-993(2016).
- [4] Zhao, A., Dong, J., Li, J., Qi, L. and Zhou, H., "Associated Spatio-Temporal Capsule Network for Gait Recognition," in *IEEE Transactions on Multimedia*, doi: 10.1109/TMM.3060280(2021) .
- [5] Wang, Y., Li, J., Zhao, A., Lv, Z., and Lu, G., "Temporal Attention-Based Graph Convolution Network for Taxi Demand Prediction in Functional Areas," In *International Conference on Wireless Algorithms, Systems, and Applications*, 203-214 Springer (2021).
- [6] Zhao, A., Li, J., Dong, J., Qi, L., Zhang, Q., Li, N., Wang, X. and Zhou, H., "Multimodal Gait Recognition for Neurodegenerative Diseases," in *IEEE Transactions on Cybernetics*, doi: 10.1109/TCYB.3056104(2021).
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and Semantic segmentation[J]. *Proceedings of 2014 IEEE Conference on Computer Vision and Pattern Recognition*. Columbus: IEEE, 2014: 580-587.
- [8] GIRSHICK R. Fast R-CNN [C]//*Proceedings of 2015 IEEE International Conference on Computer Vision*. Santiago: IEEE, 2015: 1440-1448 Hao, S., Han, X., Guo, Y., Xu, X. and Wang, M., "Low-light image enhancement with semi-decoupled decomposition," *IEEE transactions on multimedia*, vol. 22, no. 12, pp. 3025-3038(2020).
- [9] REN S Q, HE K M, GIRSHICK R, et al. Faster R-CNN; towards real-time object detection with region proposal network[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017, 39(6): 1137-1149.
- [10] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]//*Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016: 779-788.
- [11] LIU W, ANGUELOV D, ERHAN D, et al. SSD: Single shot multibox detector [C]//*European conference on computer vision*. Springer, Cham, 2016: 21-37.
- [12] Ma, L., Liu, R., Zhang, J., Fan, X. and Luo, Z., "Learning deep context-sensitive decomposition for low-light image enhancement," *IEEE Transactions on Neural Networks and Learning Systems* (2021).
- [13] Sun, Y., Zang, Y. and Liu, S., "Image Super-Resolution Using Supervised Multi-scale Feature Extraction Network," *Multimedia Tools and Applications*, 80(2): 1995-2008(2020).
- [14] Yuanjie Shao, Lerenhan Li, Wenqi Ren, Changxin Gao, and Nong Sang. Domain adaptation for image dehazing. In *CVPR*, 2020. 1, 2, 4.
- [15] L. Li, Y. Dong, W. Ren, J. Pan, C. Gao, N. Sang, and M. Yang. Semi-supervised image dehazing. *TIP*, 29:2766-2779, 2020. 4.
- [16] Lee, C., Lee, C. and Kim, C., "Contrast enhancement based on layered difference representation of 2d histograms," *IEEE transactions on image processing*, vol. 22, no. 12, pp. 5372-5384(2013).
- [17] Cai, J., Gu, S. and Zhang, L., "Learning a deep single image contrast enhancer from multi-exposure images," *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 2049-2062(2018).
- [18] Wang, S., Zheng, J., Hu, H. and Li, B., "Naturalness preserved enhancement algorithm for non-uniform illumination images," *IEEE transactions on image processing*, vol. 22, no. 9, pp. 3538-3548 (2013).
- [19] Ma, K., Zeng, K. and Wang, Z., "Perceptual quality assessment for multi-exposure image fusion," *IEEE Transactions on Image Processing*, vol. 24, no. 11, pp. 3345-3356(2015).
- [20] Mittal, A., Sundararajan, R. and Bovik, A., "Making a "completely blind" image quality analyzer," *IEEE Signal processing letters*, vol. 20, no. 3, pp. 209-212(2012).