1. Insufficient quantitative experiments raised by the reviewers 3E15, 22B2, 2A17 and 067D.

We tested and compared several advanced methods on the additional DICM[1], LIME[2], NPE[3] and MEF[4] datasets. The details are shown in the Table I:

TABLE I

OUR METHOD IS IN AN ADVANTAGEOUS POSITION IN COMPARING MUSIQ, DBCNN, AND TRES INDICATORS ON FOUR UNPAIRED DATASETS: DICM, LIME, NPE, AND MEF. RED IS THE BEST, GREEN COMES SECOND, AND BLUE COMES THIRD.

Method	DICM			LIME		
	MUSIQ↑	DBCNN↑	TRES↑	MUSIQ↑	DBCNN↑	TRES↑
RUAS	47.4793	0.3349	47.5408	55.4249	0.4815	61.5105
LLformer	56.4798	0.4198	59.6249	59.1557	0.4381	61.7582
SGRDR	53.4054	0.4243	61.6343	54.8332	0.5182	68.5986
SCI	53.8715	0.4828	59.2528	59.1134	0.5387	67.4468
EnlightenGAN	57.4258	0.4658	64.3612	59.2464	0.5333	71.0170
ours	57.6152	0.4331	63.4235	63.4894	0.5737	78.3176
Method	NPE			MEF		
	MUSIQ↑	DBCNN↑	TRES↑	MUSIQ↑	DBCNN↑	TRES↑
RUAS	49.2164	0.3960	52.8164	55.5759	0.4874	67.3862
LLformer	60.2655	0.4597	67.8456	59.2560	0.4469	69.5133
SGRDR	57.4498	0.5068	69.9917	60.4999	0.5648	77.8791
SCI	56.9149	0.5260	67.2868	62.5313	0.5762	74.2538
EnlightenGAN	59.9291	0.5316	70.4285	62.3360	0.5459	75.9656
ours	60.3305	0.4999	69.5850	63.2980	0.5474	80.1557

We utilize the MUSIQ metric[5], the TRES metric[6] and the DBCNN metric[7] for the evaluation of the compared methods. MUSIQ is an indicator used to evaluate image quality, typically used to calculate the similarity or quality between the generated image and the original image. TRES is an indicator for evaluating the quality of unreferenced images through transformers, relative ranking, and self-consistency. DBCNN is a blind image quality evaluation metric based on deep bilinear convolutional neural networks. Large MUSIQ, TRES and DBCNN means a good performance in image quality.

As shown in Table I, the MUSIQ metric of our method outperforms other methods in each dataset. The TRES metric is the best in two datasets, and the DBCNN metric is the best in LIME dataset. Moreover, the TRES and DBCNN metrics rank top 3 performance among most datasets. Overall, our method takes a leading position in these quantitative metrics, indicating the superior performance of our approach.

2. Reviewer 067D proposed the use of DeltaE metrics to evaluate the color recovery performance of our method.

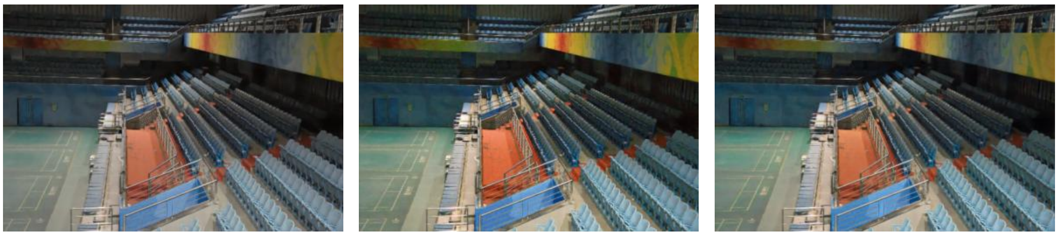
We selected ten methods to compare with our method, including one newly added method (LLformer AAAI,2023). In Table II, it can be seen that our method has achieved the best performance in four indicators: PSNR, SSIM, LPIPS, and DeltaE. Larger PSNR and SSIM values mean better image quality, smaller LPIPS metrics mean better perceptual quality of images, and smaller DeltaE indicate smaller color difference between the enhanced image and the actual values. Among them, LLformer (AAAI 2023), URetinex (CVPR 2022), and SCI (CVPR 2022) are the most advanced methods in the past two years, and our method has achieved higher performance compared to them, which also demonstrates the superiority of our method.

TABLE II
QUANTITATIVE PERFORMANCE COMPARISONS ON IMAGES FROM LOL DATASET (RED: BEST; GREEN: THE SECOND BEST; BLUE: THE THIRD BEST)

Method	PSNR(↑)	SSIM(↑)	LPIPS(↓)	DeltaE(↓)
LIME	17.1690	0.5592	0.2159	14.8479
RetinexNet	16.7347	0.4013	0.5049	15.9330
EnlightenGAN	17.6079	0.6694	0.3342	14.6331
Kind	20.4340	0.8386	0.1561	9.5590
TBEFN	17.4226	0.7915	0.2219	13.2722
Zero-DCE	14.8789	0.5595	0.3573	18.7785
RUAS	16.4395	0.5093	0.2562	16.8025
SCI	14.7714	0.5208	0.3617	19.4929
URetinexNet	21.4313	0.8440	0.1348	9.4074
LLformer	23.9062	0.8360	0.1622	7.7470
ours	24.4223	0.8571	0.1335	7.7020

3. Reviewer 3E15 proposed whether it is reasonable to use VGG to evaluate the output results of the Multi Scale Recursive Feature Enhancement (MSRFE) module.

We visualized the output results of the three MSRFE modules as shown in Fig.1. Fig.1 (1), (2) and (3) represents the enhanced results of the simulated low, medium and high exposure images by MSRFE module. It can be seen that their brightness is close to the normal light image. Therefore, the VGG module can be used to evaluate the output results of the MSRFE module.



(1) MSRFE_low (2) MSRFE_medium (3) MSRFE_high

Fig.1 Visualization of the enhanced results of simulated multi exposure images by MSRFE module. Enhanced image of (1) Low exposure images, (2) Medium exposure images and (3) High exposure images.

REFERENCES

[1] Chulwoo Lee, Chul Lee, and Chang-Su Kim, “Contrast enhancement based on layered difference representation of 2d histograms,” IEEE Transactions on Image Processing, vol. 22, no. 12, pp. 5372 – 5384, 2013.

[2] Xiaojie Guo, Yu Li, and Haibin Ling, “Lime: Low-light image enhancement via illumination map estimation,” IEEE Transactions on Image Processing, vol. 26, no. 2, pp. 982 – 993, 2017.

[3] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li, “Naturalness preserved enhancement algorithm for non-uniform illumination images,” IEEE Transactions on Image Processing, vol. 22, no. 9, pp. 3538 – 3548, 2013.

[4] Kede Ma, Kai Zeng, and Zhou Wang, “Perceptual quality assessment for multi-exposure image fusion,” IEEE Transactions on Image Processing, vol. 24, no. 11, pp. 3345 – 3356, 2015.

[5] J. Ke, Q. Wang, Y. Wang, P. Milanfar and F. Yang, "MUSIQ: Multi-scale Image Quality Transformer," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada, 2021, pp. 5128-5137, doi: 10.1109/ICCV48922.2021.00510.

[5] S. A. Golestaneh, S. Dadsetan and K. M. Kitani, "No-Reference Image Quality Assessment via Transformers, Relative Ranking, and Self-Consistency," 2022 IEEE/CVF Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 2022, pp. 3989-3999, doi: 10.1109/WACV51458.2022.00404.

[6] W. Zhang, K. Ma, J. Yan, D. Deng and Z. Wang, "Blind Image Quality Assessment Using a Deep Bilinear Convolutional Neural Network," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 1, pp. 36-47, Jan. 2020, doi: 10.1109/TCSVT.2018.2886771.