ISALux: Illumination and Semantics-Aware Transformer Employing Mixture of Experts for Low Light Image Enhancement
Supplementary Materials

Raul Balmez[M] Alexandru Brateanu[M] Ciprian Orhei[PT, WT] Codruta O. Ancuti[PT] Cosmin Ancuti[PT, WT]

[M] The University of Manchester [PT] Polytechnic University Timisoara [WT] West University Timisoara

{raul.balmez, alexandru.brateanu}@student.manchester.ac.uk

{ciprian.orhei, codruta-o.ancuti, cosmin.ancuti}@upt.ro

1. Computational analysis

In order to support the argument that ISALux is a computationally efficient we also provide a computational efficiency comparison between us and other methods, as in Table 1.

Method	**FLOPS (G)**	**PSNR**	**SSIM**	**FLOPS/PSNR**
RetinexNet [8]	587.47	18.92	0.427	31.06
MIRNet [11]	785.00	26.52	0.856	29.60
Restormer [12]	144.25	26.68	0.853	5.41
EnGAN [3]	61.01	20.00	0.691	3.05
SNR-Net [9]	26.35	26.72	0.851	0.99
Retinexformer [1]	15.57	27.14	0.850	0.57
ISALux (Ours)	37.82	27.63	0.881	1.36

Table 1. Comparison of different methods in terms of FLOPS(G), PSNR, and SSIM measured on LOL. To highlight the FLOPS required to achieve certain performances, we also introduce the ratio between the FLOPS and the PSNR.

2. Experts distribution

One might argue that the gating mechanism might overfit and that some experts might be employed more than others. To prove that all of the experts are used in the enhancement process, we analyze experts' conditional activations when testing on the LOL-v1 [8] dataset, , as in Table 2.

Expert	**Activation Percentage (%)**
Expert 0	23.81
Expert 1	20.95
Expert 2	23.33
Expert 3	31.90

Table 2. Activation probabilities of each expert based on selection data.

3. The accuracy of semantic segmentation maps under low light conditions:

The quality of semantic segmentation maps might be questioned when these feature maps are generated from a low-light image. We approached this conundrum by assuming two important aspects:

1. For severely light-deficient images, we assumed poor segmentation of objects, which is perceived by the model as a need for uniform light enhancement across the entire scene.
2. The most commonly used datasets for low-light image enhancement (those described in our paper) also contain images with scenes where the lighting level is higher and does not entirely hinder the correct segmentation of objects. This information is incorporated into the attention mechanism to preserve object separation in the feature maps used during the enhancement process.

4. Attention feature maps

The HISA-MSA attention mechanism is central to ISALux. While one could use a single multi-head attention module with point-wise multiplication by the prior maps, we opt for two separate modules. To justify this, we present four feature maps: (1) the first attention map, (2) its product with the illumination prior, (3) the second attention map, and (4) its product with the segmentation prior. Figure 1 highlights the distinct feature weighting and the impact of prior-guided modulation.

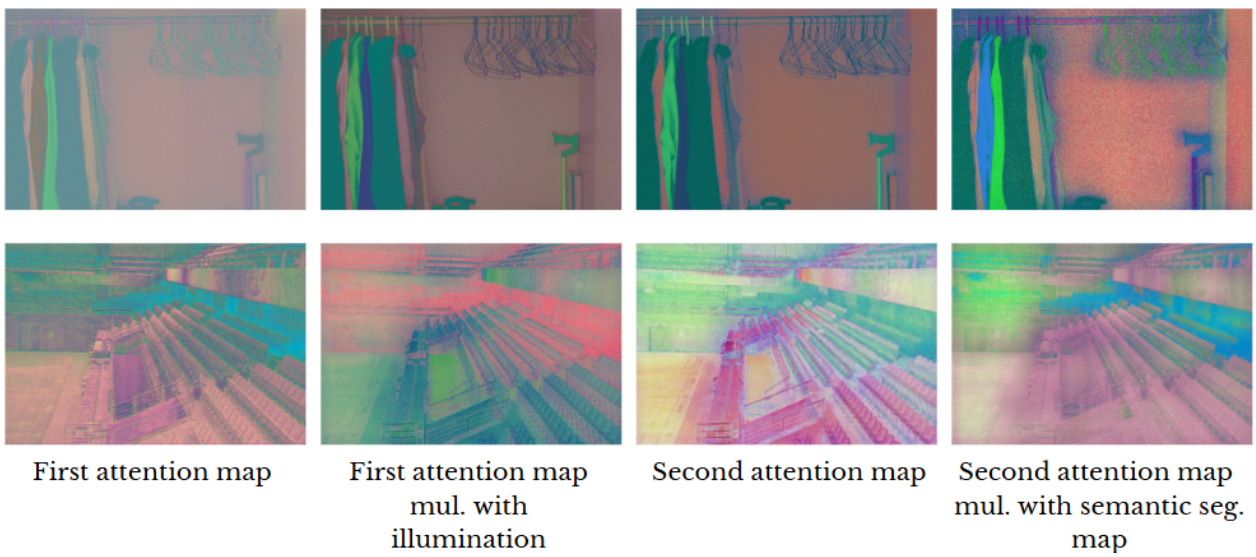

Figure 1. The 4 feature maps extracted from the HISA-MSA module. To produce them we have used Principal Component Analysis (PCA) due to the depth of the maps which was much higher than 3, the RGB standard.

5. Qualitative results

We provide further qualitative results on some of the datasets we have used as benchmarks in the paper: the LOL datasets, the non-reference datasets (MEF [5], LIME[2], DICM [4], NPE [6]), and LOL-Blur [14]. The comparison is made against the following methods: LLFlow [7], SNR-Net [9], Retinexformer [1], and GLARE [13], LL-SKF [10].

5.1. LOL datasets

Figure 2. Qualitative results on the LOL dataset

5.2. Non-reference datasets

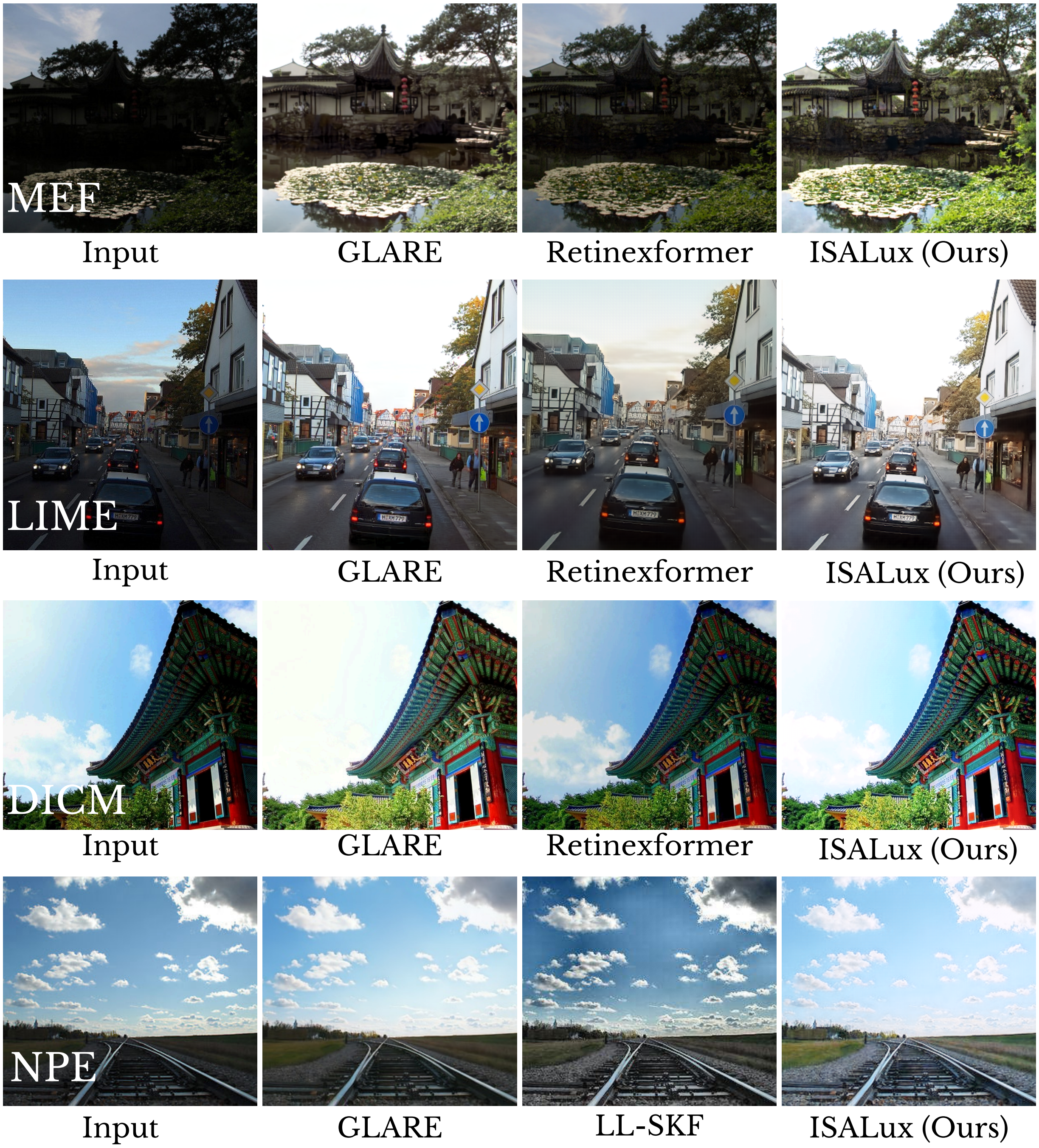

Figure 3. Qualitative results on the non-reference datasets

5.3. LOL-Blur

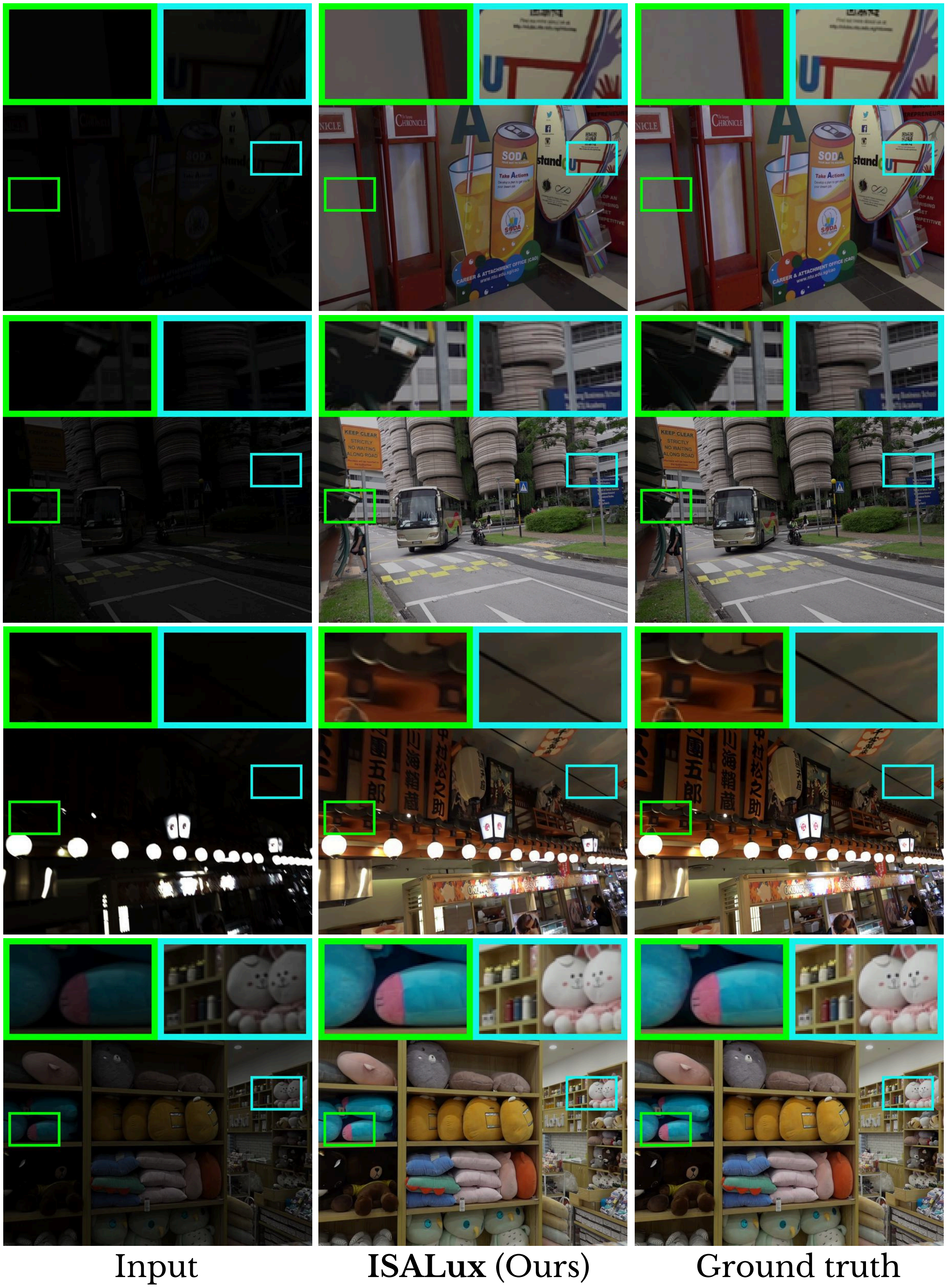

Figure 4. Qualitative results on the LOL-Blur dataset

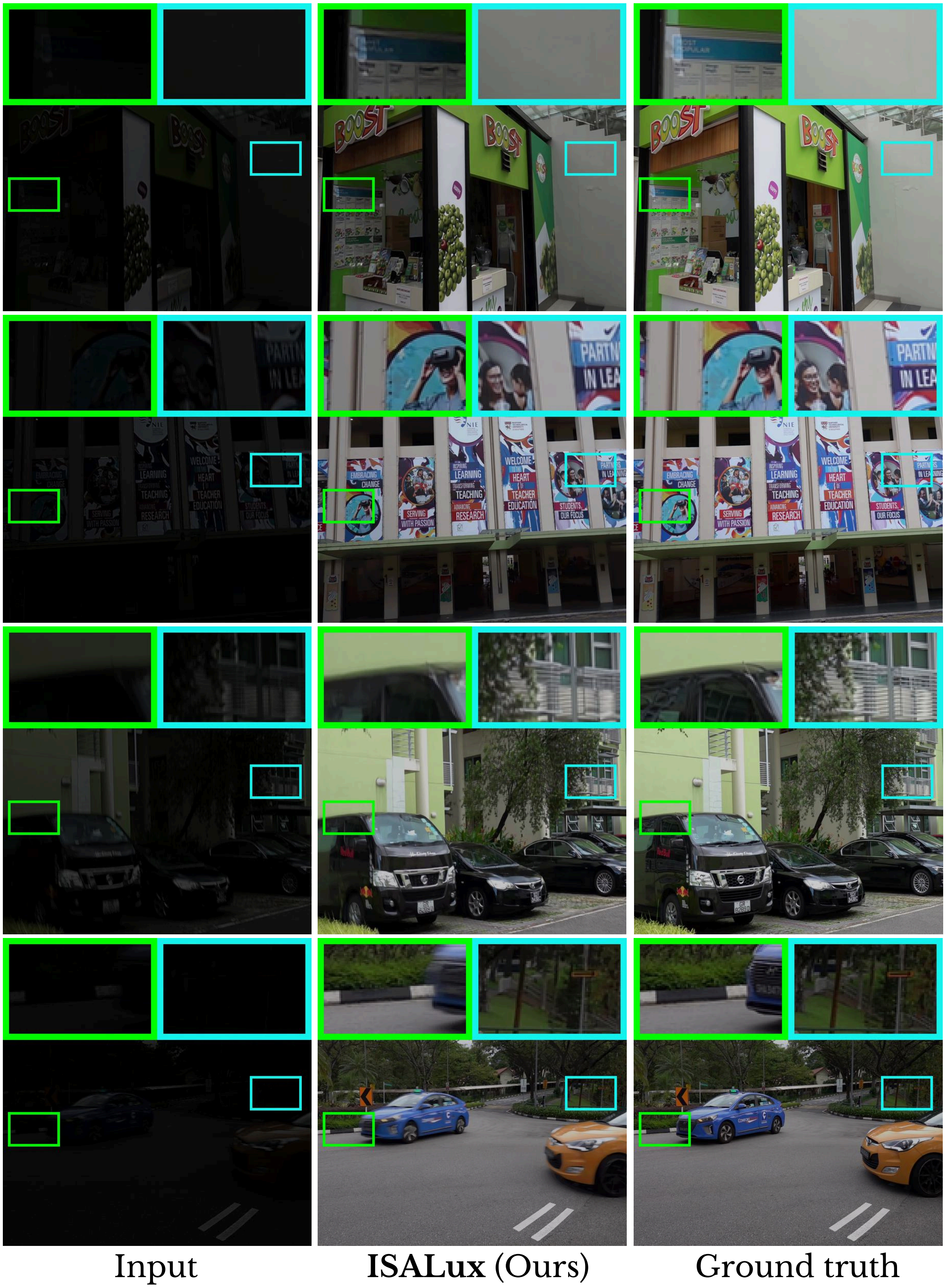

Figure 5. Qualitative results on the LOL-Blur dataset

References

[1] Yuanhao Cai, Hao Bian, Jing Lin, Haoqian Wang, Radu Timofte, and Yulun Zhang. Retinexformer: One-stage retinex-based transformer for low-light image enhancement. In *IEEE/CVF International Conference on Computer Vision (ICCV)*, 2023. 1, 3

[2] Xiaojie Guo, Yu Li, and Haibin Ling. Lime: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, 2016. 3

[3] Yifan Jiang, Xinyu Gong, Ding Liu, Yu Cheng, Chen Fang, Xiaohui Shen, Jianchao Yang, Pan Zhou, and Zhangyang Wang. Enlightengan: Deep light enhancement without paired supervision. *IEEE Transactions on Image Processing*, 30:2340–2349, 2021. 1

[4] Chulwoo Lee, Chul Lee, and Chang-Su Kim. Contrast enhancement based on layered difference representation of 2d histograms. *IEEE Transactions on Image Processing*, 22(12):5372–5384, 2013. 3

[5] Kede Ma, Kai Zeng, and Zhou Wang. Perceptual quality assessment for multi-exposure image fusion. *IEEE Transactions on Image Processing*, 24(11):3345–3356, 2015. 3

[6] Shuhang Wang, Jin Zheng, Hai-Miao Hu, and Bo Li. Naturalness preserved enhancement algorithm for non-uniform illumination images. *IEEE Transactions on Image Processing*, 22(9):3538–3548, 2013. 3

[7] Y. Wang, R. Wan, W. Yang, H. Li, Lap-Pui Chau, and A. Kot. Low-light image enhancement with normalizing flow. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 2604–2612, 2022. 3

[8] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *Proceedings of the British Machine Vision Conference (BMVC)*, 2018. 1

[9] X. Xu, R. Wang, C.-W. Fu, and J. Jia. SNR-aware low-light image enhancement. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. 1, 3

[10] Wu Yuhui, Pan Chen, Wang Guoqing, Yang Yang, Wei Jiwei, Li Chongyi, and Heng Tao Shen. Learning semantic-aware knowledge guidance for low-light image enhancement. In *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. 3

[11] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, M.-H. Yang, and L. Shao. Learning enriched features for real image restoration and enhancement. In *European Conference on Computer Vision (ECCV)*, 2020. 1

[12] S. W. Zamir, A. Arora, S. Khan, M. Hayat, F. S. Khan, and M.-H. Yang. Restormer: Efficient transformer for high-resolution image restoration. In *IEEE/CVF Conf. on Computer Vision and Pattern Recog. (CVPR)*, 2022. 1

[13] H. Zhou, W. Dong, X. Liu, S. Liu, X. Min, G. Zhai, and J. Chen. Glare: Low light image enhancement via generative latent feature based codebook retrieval. In *European Conference on Computer Vision (ECCV)*, 2024. 3

[14] S. Zhou, C. Li, and C. C. Loy. Lednet: joint low-light enhancement and deblurring in the dark. *Lecture Notes in Computer Science*, pages 573–589, 2022. 3