GM-MoE: Low-Light Enhancement with Gated-Mechanism Mixture-of-Experts

Supplementary Material

A. Appendix Section

A.1. Visualization of the results of a comparison experiment

The main text only shows a partially enlarged view of the water bottle area from the Huawei dataset. As shown in Fig. 1, we further give a full view of this photo so readers can more fully observe the results. It can be seen that for areas with patterns and text, the enhancement results of other networks are often blurry, while the text boundaries in the image processed by the GM-MOE method are clearer, the colors are more distinct, and there is no obvious blurring.

As can be seen in Fig. 2, the network enhanced by GM-MOE does not experience color distortion or overexposure compared to other networks.

In addition, we also show a photograph of a building from the Nikon dataset, as shown in Fig. 3, By comparing the enhancement effects of different methods (such as DeepUPE), it can be found that the output of some methods has serious color distortion or loss of detail. In contrast, our method performs better in terms of color reproduction and detail retention, and the enhanced map is closest to the real map (Ground Truth). From a quantitative perspective, our method also achieves the highest scores in metrics such as PSNR and SSIM, which further proves its superiority.

A.2. About the dataset

LOL-v1 [10] is a classic low-light image enhancement dataset that covers a variety of scenes and is used to test the low-light processing effects of models in different real-world scenarios. Compared to LOL-v1 [10], LOLv2-Real [14] provides more diverse lighting scenarios, while LOL-v2-Synthetifutc generates a wider range of scenarios through artificial low-light simulation algorithms, which are mainly used to evaluate the generalization ability of the model. The LSRW-Huawei [4] and LSRW-Nikon [4] datasets, which were captured by Huawei and Nikon devices respectively, contain images of real-world low-light scenes, which require a high level of detail in processing low-light photos.

Taking the LOLv2-Real dataset with enhanced data as an example, our method performs best in recovering the glass surface. As can be seen from the above figure, GM-MOE has better generalization ability on multiple datasets, especially in terms of detail processing, which is superior to other methods.

A.3. Generalisation to SID Dataset

On the SID benchmark [2], GM–MoE attains 24.80 PSNR and 0.69 SSIM, demonstrating strong performance.

A.4. Comparison with Competitive Baselines on LOL-v1

As shown in Tab. 1, our method achieves a PSNR of **26.66 dB**, an SSIM of **0.86**, and a LPIPS of **0.098**.

Methods	PSNR (↑)	SSIM (↑)	LPIPS (↓)
SCI [7]	14.78	0.53	0.392
NeRCo [12]	22.95	0.79	0.311
DiffLLE [13]	22.24	0.79	–
LightenDiffusion [5]	20.45	0.80	0.192
Ours	**26.66**	**0.86**	**0.098**

Table 1. Quantitative comparison on the LOL-v1 dataset among different methods: SCI [7], NeRCo [12], DiffLLE [13] , LightenDiffusion [5], and Ours.

A.5. Perceptual Quality (LPIPS)

As shown in Tab. 2 , we report the LPIPS scores of different models on different benchmarks.

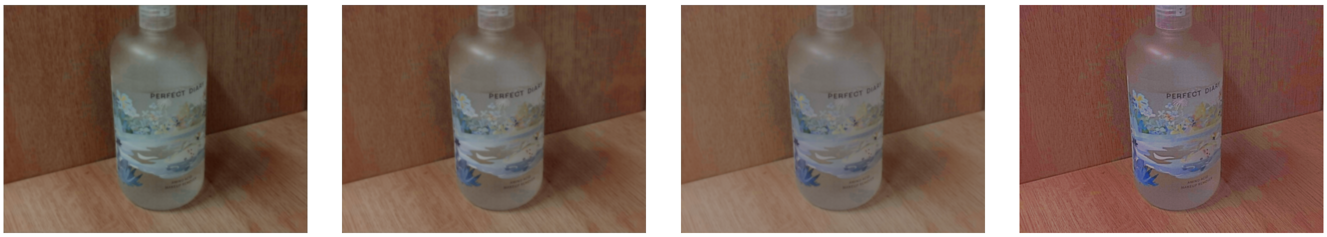
Method	LOL-v1	LOL-v2-Real	LOL-v2-Synthetic
Retinexformer [1]	0.129	0.171	0.059
LLFormer [16]	0.167	0.211	0.066
GM–MoE (Ours)	**0.098**	**0.100**	**0.041**

Table 2. LPIPS (↓) comparison across three LOL datasets. Lower is better.

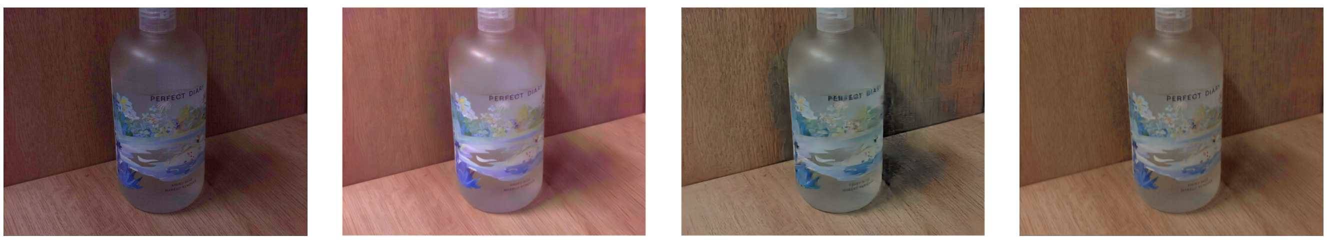
A.6. How are expert modules coordinated?

As shown in Fig. 4 , we enforce the use of only one expert at a time. The heatmap of Expert1 exhibits yellow-to-red contributions in fringe details, indicating its primary role in local chroma restoration. Expert2's heatmap shows predominantly dark blue regions with only faint highlights in limited detailed areas, suggesting its specialization in fine texture recovery. Expert3 demonstrates extensive orange-yellow coverage across both the main fringe and background regions. These three heatmaps reveal spatial complementarity among the experts, enabling the final enhanced results to achieve both rich detail preservation and more realistic color reproduction.

(a) Ground Truth (b) RetinexNet [10] (c) DeepUPE [8] (d) Restormer [15]



(e) SNR-Net [11] (f) Retinexformer [1] (g) LightenDiffusion [6] (h) PairLIE [3]



(i) SCI [7] (j) QuadPrior [9] (k) NeRCo [12] (l) Ours (GM-MOE)

Figure 1. **Visual comparison on the LSRW-Huawei dataset [4].** The models compared include RetinexNet, DeepUPE, Restormer, SNR-Net, Retinexformer, LightenDiffusion, PairLIE, SCI, QuadPrior, NeRCo, Ours (our model), and Ground Truth. Among them, GM-MOE achieves better enhancement compared to other models. Zoom in to see more details of the differences between models.

A.7. Computational Efficiency

Our method delivers superior low-light enhancement quality while incurring a computational cost of 27.2 GFLOPs, which is still practical for real-time deployment. Future work will focus on further reducing latency and memory footprint without sacrificing restoration accuracy.

A.8. Ablation study on network structure and expert interactions

To explore in depth the relationship between the number of parameters and performance, we conducted three sets of experiments: fixed weights, cascaded networks, and constrained parameter growth.

No.	Variant	Params (M)	PSNR (↑)	SSIM (↑)
1	Original GM-MoE	19.99	23.65	0.80
2	Without dynamic gating (fixed weights)	19.86	21.70	0.71
3	Serializing three expert networks	19.86	21.34	0.83
4	Experts' channels concat + 1×1 fusion	20.60	17.84	0.70

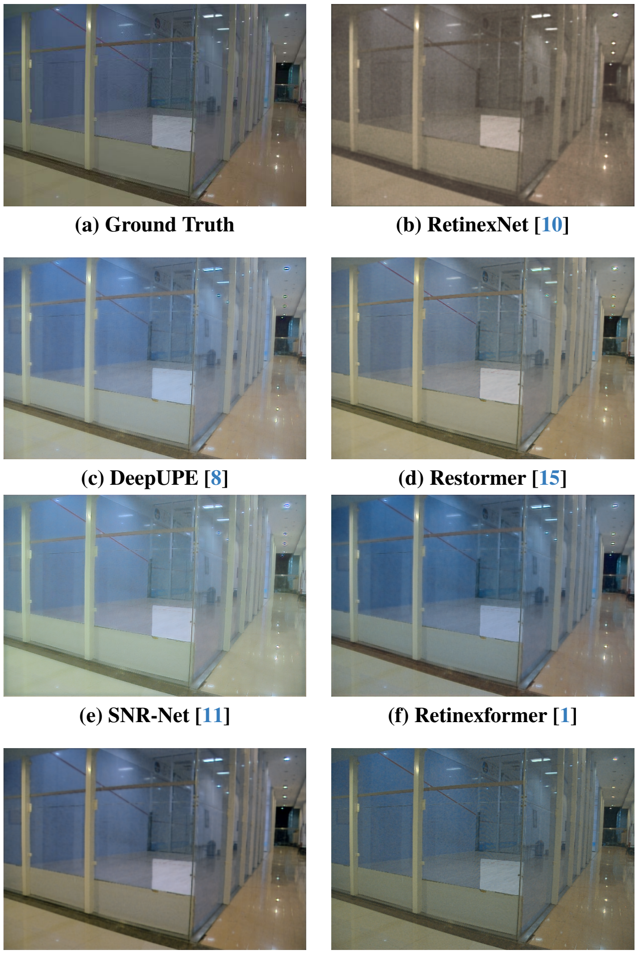
Table 3. Ablation study on network structure and expert interactions evaluated on the LOLv2-Real dataset.

A.9. Limitation and future works

Increase in Sub-Expert Networks and Its Impact on Performance. Increasing the number of sub-expert networks may improve the model's performance, but it also introduces additional computational complexity. In GM-MoE, the role of the sub-expert networks is to tackle different low-light enhancement tasks, allowing the model to process various image features more specifically. Each sub-expert network focuses on different aspects of low-light image enhancement, which can lead to better performance, particularly when the tasks are well-defined and complementary.

Scalability to Downstream Tasks. Currently, we have applied GM-MoE to enhance low-light images and used these enhanced images for object detection. However, future work should explore extending the GM-MoE framework to other downstream tasks. For example, video enhancement processing is a promising avenue for application. The framework may also be applicable to other tasks such as image segmentation or visual recognition, which could further demonstrate the versatility of GM-MoE.

Adjustability of Loss Functions and Their Impact on the Model. Currently, we use PSNR as the primary met-



(a) Ground Truth (b) RetinexNet [10]
(c) DeepUPE [8] (d) Restormer [15]
(e) SNR-Net [11] (f) Retinexformer [1]
(g) LightenDiffusion [6] (h) Ours(GM-MOE)

Figure 2. **Visual comparison on the LOL-v2-Real dataset.**The enhanced effect of GM-MOE is better than other models in terms of detail processing.

ric for image quality. However, in future work, we should investigate how adjusting the loss function impacts the overall performance of the model. Experimenting with alternative loss functions, such as perceptual loss or adversarial loss, may provide better results in preserving image details and enhancing visual quality, especially for more complex tasks. The choice of loss function can significantly affect the model's ability to generalize across different datasets and tasks.

In future work, we will further explore the application of the GM-MoE model. First, we will study how to improve the computational efficiency of the model. Second, we will explore the use of GM-MoE-enhanced images in downstream tasks such as image segmentation and video enhancement to verify its versatility and adaptability.

(a) Ground Truth (b) RetinexNet [10]
(c) DeepUPE [8] (d) Restormer [15]
(e) SNR-Net [11] (f) Retinexformer [1]
(g) LightenDiffusion [6] (h) Ours(GM-MOE)

Figure 3. **Visual comparison on the Nikon dataset.** The enhancement effect of GM-MOE is closer to the original picture than other networks.

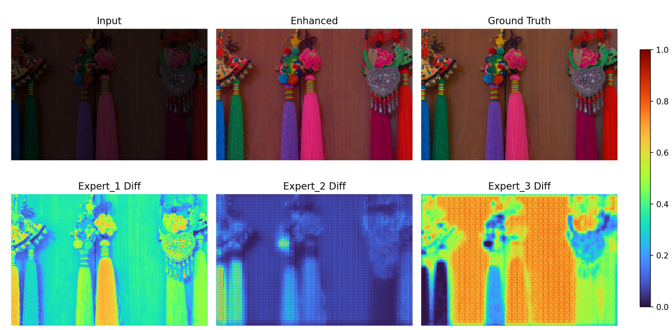


Figure 4. The heatmaps show distinct focus areas across the three experts, forming a coherent, synergistic attention distribution.

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 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 12504–12513, 2023. 1, 2, 3

[2] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 1

[3] Zhenqi Fu, Yan Yang, Xiaotong Tu, Yue Huang, Xinghao Ding, and Kai-Kuang Ma. Learning a simple low-light image enhancer from paired low-light instances. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 22252–22261, 2023. 2

[4] Jiang Hai, Zhu Xuan, Ren Yang, Yutong Hao, Fengzhu Zou, Fang Lin, and Songchen Han. R2rnet: Low-light image enhancement via real-low to real-normal network. *Journal of Visual Communication and Image Representation*, 90: 103712, 2023. 1, 2

[5] Hai Jiang, Ao Luo, Xiaohong Liu, Songchen Han, and Shuaicheng Liu. Lightendiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. In *European Conference on Computer Vision*, 2024. 1

[6] Hai Jiang, Ao Luo, Xiaohong Liu, Songchen Han, and Shuaicheng Liu. Lightendiffusion: Unsupervised low-light image enhancement with latent-retinex diffusion models. In *European Conference on Computer Vision*, pages 161–179. Springer, 2025. 2, 3

[7] Long Ma, Tengyu Ma, Risheng Liu, Xin Fan, and Zhongxuan Luo. Toward fast, flexible, and robust low-light image enhancement. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 5637–5646, 2022. 1, 2

[8] Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, and Jiaya Jia. Underexposed photo enhancement using deep illumination estimation. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1–10, 2019. 2, 3

[9] Wenjing Wang, Huan Yang, Jianlong Fu, and Jiaying Liu. Zero-reference low-light enhancement via physical quadruple priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 26057–26066, 2024. 2

[10] Chen Wei, Wenjing Wang, Wenhan Yang, and Jiaying Liu. Deep retinex decomposition for low-light enhancement. In *British Machine Vision Conference*, pages xx–xx, 2018. 1, 2, 3

[11] Xiaogang Xu, Ruixing Wang, Chi-Wing Fu, and Jiaya Jia. Snr-aware low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 17714–17724, 2022. 2, 3

[12] Shuzhou Yang, Moxuan Ding, Yanmin Wu, Zihan Li, and Jian Zhang. Implicit neural representation for cooperative low-light image enhancement. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 12918–12927, 2023. 1, 2

[13] Shuzhou Yang, Xuanyu Zhang, Yinhuai Wang, Jiwen Yu, Yuhan Wang, and Jian Zhang. Difflle: Diffusion-guided domain calibration for unsupervised low-light image enhancement, 2023. 1

[14] Wenhan Yang, Wenjing Wang, Haofeng Huang, Shiqi Wang, and Jiaying Liu. Sparse gradient regularized deep retinex network for robust low-light image enhancement. *IEEE Transactions on Image Processing*, 30:2072–2086, 2021. 1

[15] Syed Waqas Zamir, A R Rajeev Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 5728–5739, 2022. 2, 3

[16] X. Zhang, Y. Fan, S. Deng, Y. Huang, and Y. Zhang. Llformer: A lightweight transformer for low-light image enhancement with information distillation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 3233–3241, 2022. 1