π-Diff: Physically-Inspired Low-Light Image Enhancement with Structure Preserving Diffusion Priors

Supplementary Material

1. Additional Experimental Results

The additional qualitative comparisons in Fig. 1 further validate the robustness and superior generalization of π-Diff across diverse challenging scenarios. These additional results reinforce our method's performance advantages:

Structural Preservation and High-Frequency Details. The critical advantage of our decoupled structural proxy ($\boldsymbol{P}_t$) is most explicitly visible in regions containing dense high-frequency textures and rigid geometric structures. Direct-generation diffusion models inherently entangle structural denoising with illumination enhancement in a single black-box mapping, leading to severe representation conflicts. Consequently, in the first and second rows featuring complex bicycle spokes and sharp building facades, GDP[1] completely over-smooths the scenes, heavily degrading fine structural details to compensate for the extreme darkness. Furthermore, recent generative baselines like QuadPrior[8] and AGLLDiff[5] introduce severe hallucination artifacts, distorting straight edges and warping geometric boundaries due to the lack of explicit structural constraints. In stark contrast, by isolating the denoising process from photometric calibration, our proxy acts as a noiseless physical anchor. This ensures our method perfectly preserves the sharp structural outlines, crisp high-frequency edges, and the holistic integrity of the scene layout, closely matching the Ground Truth.

Illumination Recovery and Color Restoration. Traditional methods like MBLLEN[7] and Zero-DCE++[4] consistently struggle to provide sufficient illumination in extremely dark regions due to limited network capacity. Conversely, while EnlightenGAN[3] and RUAS[6] manage to brighten the scenes, they suffer from severe color distortion. This degradation occurs because these models indiscriminately amplify pixel values without separating intrinsic object properties from environmental lighting. For instance, in the bottom row, the natural cool tone of the indoor floor is erroneously shifted to a highly unnatural orange and yellow hue. Our framework elegantly resolves this by extracting and preserving the intrinsic reflectance map ($\boldsymbol{R}_t$). By locking the scene's true colors, our ISP-inspired mechanism can safely and accurately calibrate the exposure on the decoupled illumination map ($\boldsymbol{L}_t$). This guarantees optimal brightness while strictly maintaining the intrinsic scene colors without arbitrary chromatic shifts.

Artifact Suppression in Complex Patterns. Regions with complex, repetitive structures, such as the stadium seating in the third row, pose a severe challenge for modern enhancement models. In these areas, the extremely low signal-to-noise ratio often triggers uncontrolled semantic hallucinations in standard diffusion priors, leading to the noise-like artifacts, structural warping, and patchy color distribution that severely plague methods like PairLIE[2] and AGLLDiff[5]. Our method cleanly enhances the image and entirely avoids these generative flaws. By subjecting the reverse sampling process to our physically-guided joint optimization, the gradients derived from the spatial and color consistency losses strictly regulate the diffusion trajectory, actively suppressing any non-physical structural deviations and ensuring artifact-free restoration.

Overall, these supplementary comparisons definitively prove that our physically-guided approach robustly avoids the hallucination and color-shift pitfalls of existing baselines, producing structurally faithful, artifact-free, and visually pleasing restorations across diverse real-world environments.

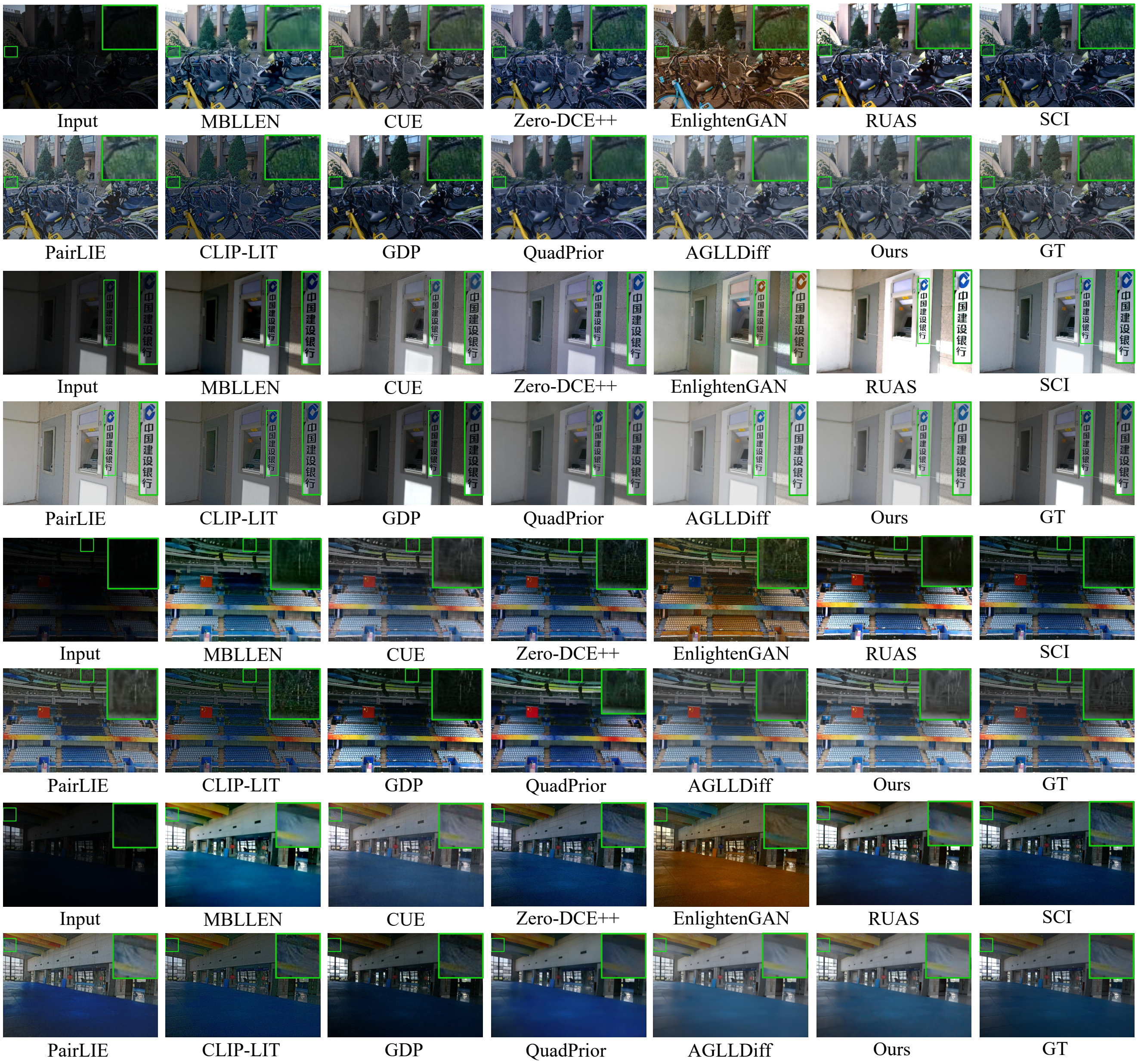

Figure 1. **More comparison with existing methods.** Our method effectively recovers detailed structure and natural appearance while other methods often struggle with detail preservation, denoising, color restoration, and illumination improvement.

References

[1] Ben Fei, Zhaoyang Lyu, Liang Pan, Junzhe Zhang, Weidong Yang, Tianyue Luo, Bo Zhang, and Bo Dai. Generative diffusion prior for unified image restoration and enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 9935–9946, 2023. 1

[2] 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. 1

[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] Chongyi Li, Chunle Guo, and Chen Change Loy. Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE transactions on pattern analysis and machine intelligence*, 44(8):4225–4238, 2021. 1

[5] Yunlong Lin, Tian Ye, Sixiang Chen, Zhenqi Fu, Yingying

Wang, Wenhao Chai, Zhaohu Xing, Wenxue Li, Lei Zhu, and Xinghao Ding. Aglldiff: Guiding diffusion models towards unsupervised training-free real-world low-light image enhancement. In *Proceedings of the AAAI Conference on Artificial Intelligence*, pages 5307–5315, 2025. 1

[6] Risheng Liu, Long Ma, Jiaao Zhang, Xin Fan, and Zhongxuan Luo. Retinex-inspired unrolling with cooperative prior architecture search for low-light image enhancement. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 10561–10570, 2021. 1

[7] Feifan Lv, Feng Lu, Jianhua Wu, and Chongsoon Lim. Mbllen: Low-light image/video enhancement using cnns. In *Bmvc*, page 4. Northumbria University, 2018. 1

[8] 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*, pages 26057–26066, 2024. 1