PBFG: A New Physically-Based Dataset and Removal of Lens Flares and Glares Supplementary Material

Jie Zhu[1] Sungkil Lee[1,2,*]
[1]Department of Electrical and Computer Engineering, Sungkyunkwan University, South Korea
[2]Department of Computer Science and Engineering, Sungkyunkwan University, South Korea
{jzhu, sungkil}@skku.edu

1. Flare and Glare Templates and Comparison

Fig. 1 presents five real camera systems from the 15 real camera systems in our PBFG dataset, illustrating their lens layouts and corresponding flare patterns. Fig. 2 shows synthetic glare patterns generated by combining glow, shimmer, streak, and starburst with random weights. Our PBFG dataset provides rich annotations for glow, shimmer, streak, starburst, light sources, and flares, facilitating tasks such as night glare component segmentation and light source extraction. Integrating these annotated images into training datasets for other night vision models enhances their robustness and generalization, particularly in handling complex night scenes with diverse glare and flare patterns. Thus, our PBFG dataset serves as a vital resource for advancing nighttime computer vision applications.

2. Implementation of Histogram-Matching Processing Module

Histogram-matching (HM) processing module aligns the cumulative distribution function (CDF) of the target image X with that of a reference Y via a monotonic mapping. Given normalized histograms $p_X(x)$ and $p_Y(y)$, the CDFs are:

$$F_X(i) = \sum_{x \le i} p_X(x), \tag{1}$$

$$F_Y(j) = \sum_{y \le j} p_Y(y). \tag{2}$$

The mapping is:

$$h(i) = \min \{ j \mid F_Y(j) \ge F_X(i) \}, \tag{3}$$

refined via interpolation to produce the transformed image $X' = h(X)$. This is applied per channel for accurate histogram alignment. To clarify, the HM module and ground truth images were not used during training or validation, so no data leakage occurred. HM is applied only post hoc to improve stylistic consistency with the GT.

3. Implementation, Rationality and Validity of Enhanced Ground Truth (GT)

The public GT has known limitations [2], including residual artifacts near light sources. To ensure a more accurate evaluation, we manually refine these low-quality GT images. In Photoshop, we mask residual areas, adjust saturation, and apply dehazing to reduce highlights. Then, we sample colors to identify shifts and fine-tune brightness and saturation to restore natural tones, eliminating artifacts while preserving structure. Table 1 compares two networks evaluated with both the original and enhanced GTs. In the first row, PSNR drops from 30.142 dB (original GT) to 26.890 dB (enhanced GT) because Flare7K++ fails to remove these artifacts, making its output similar to the original GT and inflating PSNR (blue grids). Our method removes these artifacts more effectively, leading to higher PSNR with the enhanced GT (green grids). The enhanced GT does not simply favor our method—Flare7K++ also aligns better in some cases (yellow grids), showing it offers a fairer evaluation. Image indices 004 and 045 correspond to the top and bottom samples in Fig. 7 in the main paper.

4. More Experimental Results

4.1. Similarity Between PBFG and Real Datasets

To verify the similarity between our PBFG dataset and real-world datasets, we conduct a t-SNE analysis. As shown in Fig. 3, the left plot shows strong overlap between PBFG and real datasets, indicating a similar feature distribution. PBFG's scattered points reflect greater diversity due to its use of various aperture blades, unlike Flare-R and Flare600. t-SNE results on PBFG labels (glow, shimmer, streak) show clear clusters, confirming its quality, diversity, and utility for glare-related tasks.

*Corresponding author.

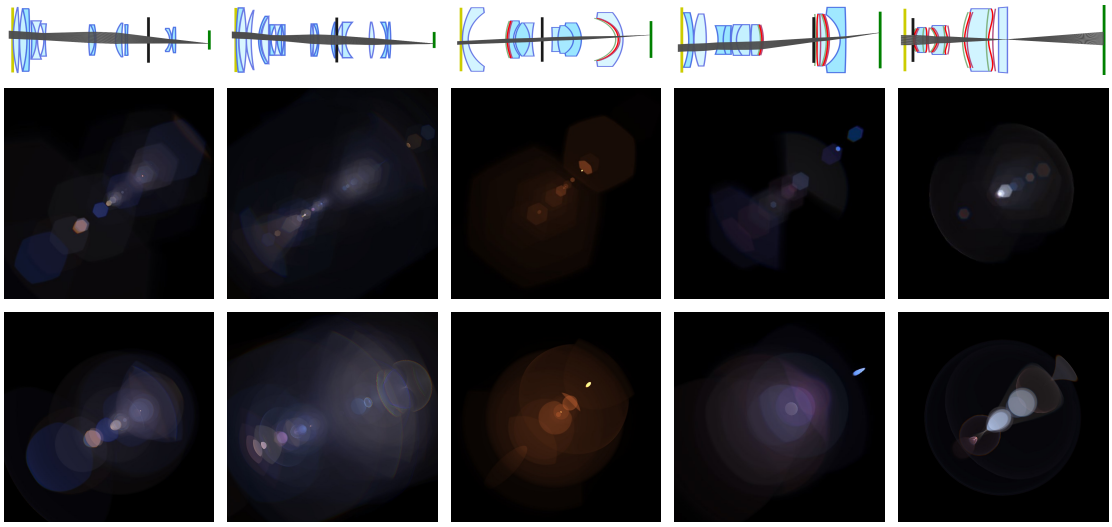

Figure 1. Flare patterns in the PBFG dataset for different lens systems. Top: lens layouts with black lines for apertures, yellow for light entrance, and green for sensors. Middle: flare patterns with six-blade apertures. Bottom: flare patterns with circular apertures.

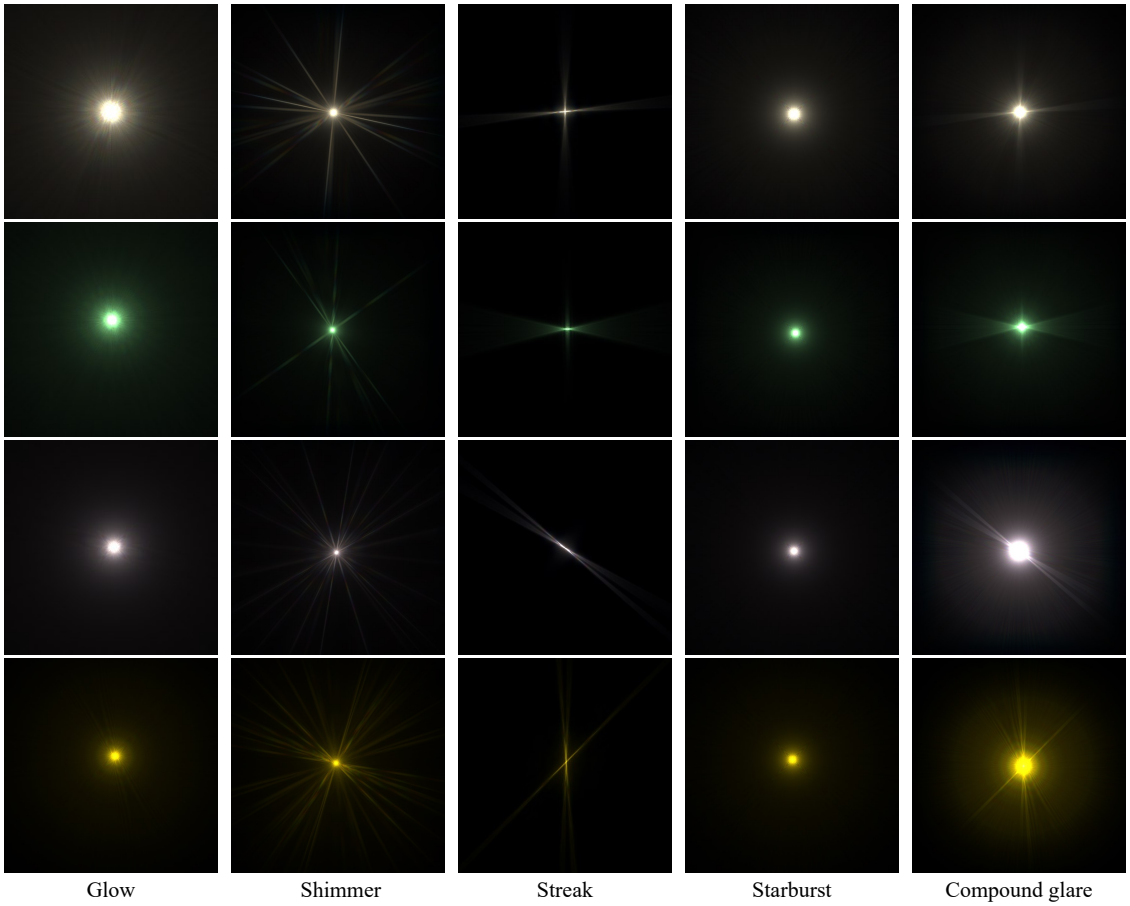

Figure 2. Glow, shimmer, streak, starburst, and compound glare patterns synthesized with our computational rendering scheme.

Image Index	Network	Original GT / Enhanced GT (Ours)				
		PSNR↑	SSIM↑	LPIPS↓	G-PSNR↑	S-PSNR↑
004	Flare7K++	30.142 / 26.890	0.913 / 0.901	0.0321 / 0.0312	26.522 / 21.483	27.257 / 23.032
	FGRNet (Ours)	29.417 / 32.105	0.914 / 0.943	0.0327 / 0.0274	23.616 / 29.283	24.726 / 28.751
045	Flare7K++	30.023 / 30.125	0.956 / 0.955	0.0211 / 0.0142	29.067 / 28.682	25.443 / 27.520
	FGRNet (Ours)	31.553 / 32.354	0.960 / 0.960	0.0220 / 0.0159	31.049 / 30.969	27.963 / 31.105

Table 1. Comparison of artifact removal results trained using Flare7K++ and our FGRNet, evaluated using both original and enhanced GT images. Image indices 004 and 045 correspond to the top and bottom samples in Fig. 7 of the main paper, drawn from the test set.

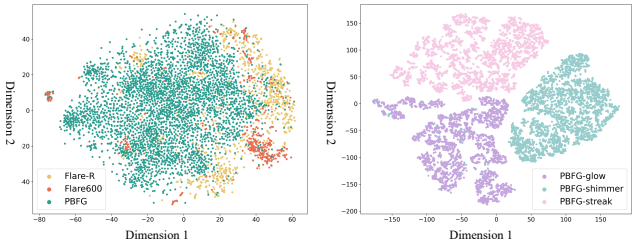

Figure 3. t-SNE analysis of our PBFG dataset and real datasets.

Training Dataset	Model	PSNR↑	SSIM↑	LPIPS↓	G-PSNR↑	S-PSNR↑
Flare7K++ [2]	Uformer	27.316	0.889	0.0453	23.543	22.181
	Uformer + SRU	27.419	0.892	0.0444	24.396	22.451
	Uformer + FEU	27.893	0.896	0.0440	24.576	23.215
	FGRNet (Ours)	**28.318**	**0.898**	**0.0433**	**24.830**	**23.582**
PBFG (Ours)	Uformer	27.920	0.896	0.0455	24.330	23.115
	Uformer + SRU	27.965	0.896	0.0458	24.540	23.197
	Uformer + FEU	28.307	0.896	0.0432	24.641	23.935
	FGRNet (Ours)	**28.659**	**0.898**	**0.0426**	**25.444**	**24.385**

Table 2. Quantitative comparison of FGRNet components without the HM module trained on Flare7K++ and our PBFG datasets. The best results are in **bold**.

4.2. Data Augmentation

Dai et al. [2] present the most relevant and advanced approach to our work. We adopt their data augmentation strategy to generate flare- and glare-corrupted images. Background images are sourced from Flickr24K [7]. To linearize the background, flare, glare, and light source images, we apply inverse gamma correction using $U(1.8, 2.2)$. The background image undergoes random RGB adjustment with $U(0.5, 1.2)$ and Gaussian noise addition, where the noise variance follows a scaled chi-square distribution $\sigma^2 \sim 0.01\chi^2$. We select random flare and glare patterns from our PBFG dataset and combine them to create a compound image. This compound image is subjected to a series of affine transformations, including random rotation $U(0, 2\pi)$, shear $U(-\pi/9, \pi/9)$, scaling $U(0.7, 1.5)$, and translation $U(-100, 100)$. Additionally, we apply random Gaussian blur with a size in $U(0.1, 3)$, color offset $U(-0.02, 0.02)$, and perform random horizontal and vertical flips. Both the adjusted background image and the transformed compound image are then clipped to the range [0, 1]. They are combined to generate the final flare-/glare-corrupted image, which is also clipped to the range [0, 1].

4.3. Comparison with Existing Dataset

Fig. 4 shows that models trained on the PBFG dataset outperform those trained on previous datasets in removing real-world flare and glare artifacts, validating the realism and effectiveness of our dataset. The first four rows illustrate that models trained on PBFG effectively remove complex artifacts, including prominent and subtle streaks, and off-frame stray light source-induced veiling glares or streaks. In contrast, models trained on existing datasets struggle with subtle streaks, leaving residual artifacts. This improvement is due to our advanced streak synthesis scheme, which accurately removes streaks of varying sizes and intensities. The last two rows demonstrate that PBFG-trained models effectively remove large glow areas with streaks, thanks to our physically-based synthesis and spectral integration, which realistically simulate glare color transitions. In contrast, existing non-physical datasets replicate only the shape of the glow, failing to capture spectral color variations. These findings highlight our dataset's superior capability to enhance model robustness and accuracy in complex nighttime scenes.

4.4. More Quantitative Comparisons

FGRNet Performance without HM. GT and the HM module are unavailable in real scenarios, but FGRNet alone achieves strong visual results. Table 2 shows the performance of FGRNet components without the HM module. While HM enhances style consistency with GT, its absence does not affect the effectiveness of our method on flare and glare removal.

Results on Synthetic Test Set. Table 3 shows our method outperforms all Flare7K++-trained methods on the synthetic test set, especially in G-PSNR and S-PSNR for glare and streak removal.

Efficiency Analysis. Table 4 shows the efficiency analysis on an NVIDIA RTX 3090 with 512 × 512 images. Featuring layer depths of [1, 2, 2, 2] and a lightweight SFEM module, FGRNet delivers superior removal with marginal overhead, making it suitable for real-world deployment.

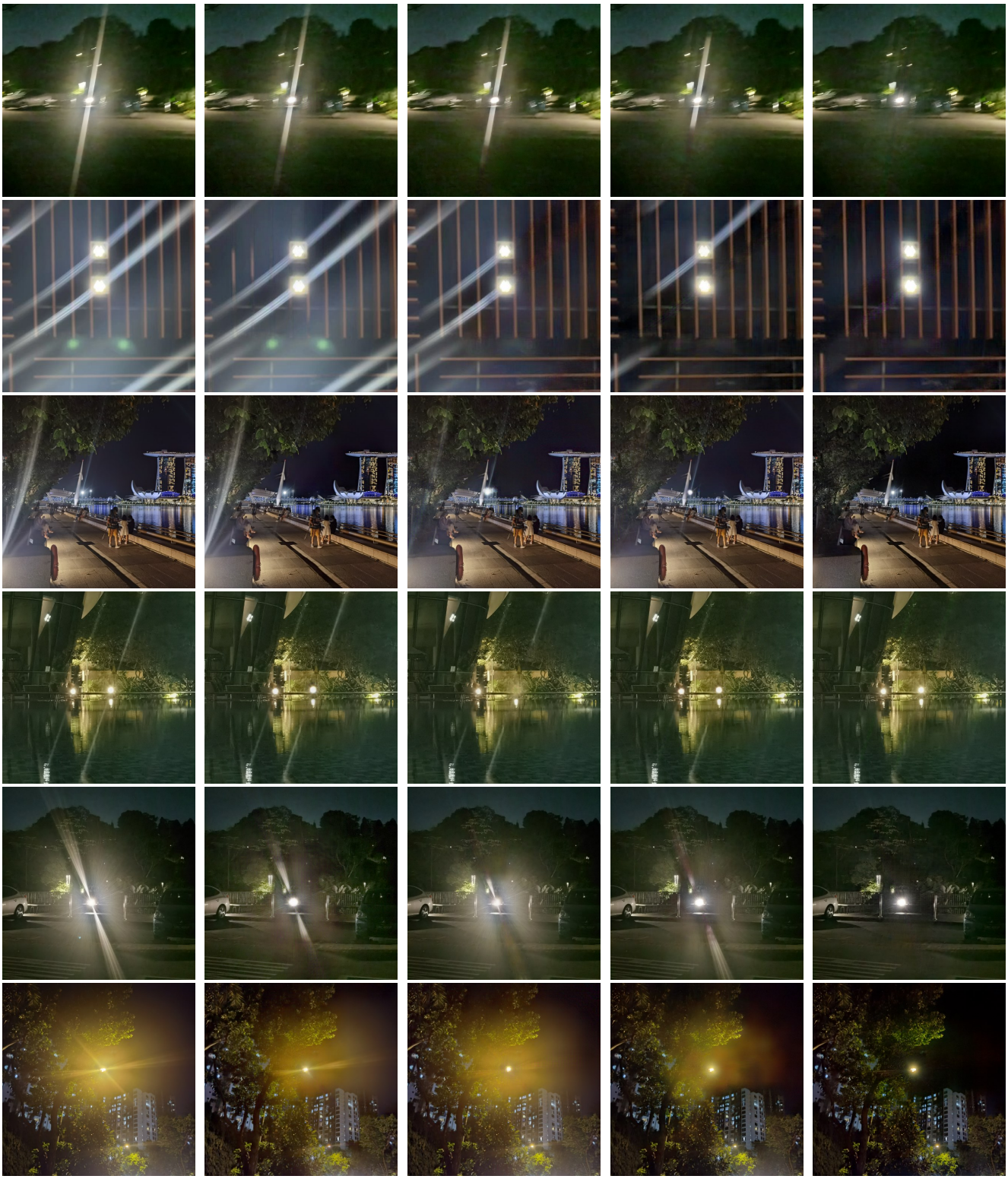

Figure 4. Visual results on real-world images. The same model is trained on different datasets [1, 2, 6] for lens flare and glare removal, with our dataset yielding the best performance.

Method	PSNR↑	SSIM↑	LPIPS↓	G-PSNR↑	S-PSNR↑
Flare7K++ [2]	29.498	0.962	**0.0210**	24.686	24.155
Zou et al. [8]	29.224	0.956	0.0244	24.489	24.401
Flare-Free Vision [4]	29.521	0.960	0.0234	24.513	24.217
Sparse-UFormer [5]	29.584	0.959	0.0229	24.379	24.591
FGRNet (Ours)	29.722	0.960	0.0215	24.945	24.756
FGRNet + HM (Ours)	**30.135**	**0.967**	0.0235	**25.228**	**25.178**

Table 3. Quantitative evaluation on synthetic nighttime images. The best results are in **bold**, and the second-best results are underlined.

Method	Params. (M)	FLOPs (G)	Time (s)	Train Iters. (K)
Wu et al. [5]	34.51	261.95	0.0956	1200
Flare7K [1]	**20.46**	161.08	**0.0556**	1200
Flare7K++ [2]	**20.46**	161.08	**0.0556**	**300**
Flare-Free Vision [4]	154.81	274.32	0.1842	417.5
FGRNet (Ours)	34.39	**160.21**	0.0570	400

Table 4. Model complexity and training efficiency comparison across methods. The best results are in **bold**, and the second-best results are underlined.

4.5. More Qualitative Comparisons

Figs. 5, 6, 7, and 8 show more visual comparisons of our method and state-of-the-art methods [1, 2, 4, 5].

References

[1] Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy. Flare7k: A phenomenological nighttime flare removal dataset. In *Thirty-sixth Conference on Neural Information Processing Systems Datasets and Benchmarks Track*, 2022. 4, 5, 7

[2] Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, Yihang Luo, and Chen Change Loy. Flare7k++: Mixing synthetic and real datasets for nighttime flare removal and beyond. *arXiv preprint arXiv:2306.04236*, 2023. 1, 3, 4, 5, 6, 7, 8

[3] Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, Qingpeng Zhu, Qianhui Sun, Wenxiu Sun, Chen Change Loy, Jinwei Gu, Shuai Liu, et al. Mipi 2023 challenge on nighttime flare removal: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2853–2863, 2023. 6

[4] Yousef Kotp and Marwan Torki. Flare-free vision: Empowering uformer with depth insights. In *ICASSP 2024-2024 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 2565–2569. IEEE, 2024. 5, 7

[5] Siqi Wu, Fei Liu, Yu Bai, Houzeng Han, Jian Wang, and Ning Zhang. Flare removal model based on sparse-uformer networks. *Entropy*, 26(8):627, 2024. 5, 7

[6] Yicheng Wu, Qiurui He, Tianfan Xue, Rahul Garg, Jiawen Chen, Ashok Veeraraghavan, and Jonathan T Barron. How to train neural networks for flare removal. In *IEEE International Conference on Computer Vision*, pages 2239–2247, 2021. 4

[7] Xuaner Zhang, Ren Ng, and Qifeng Chen. Single image reflection separation with perceptual losses. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4786–4794, 2018. 3

[8] Guanyu Zou, Hongyang Bai, Yuman Yuan, Tianyu Deng, Zhenpeng Yin, and Jingyi Wei. Research on flare removal network based on channel attention mechanism and depthwise over-parameterized convolution. In *Proceedings of the 4th International Conference on Artificial Intelligence and Computer Engineering*, pages 919–926, 2023. 5

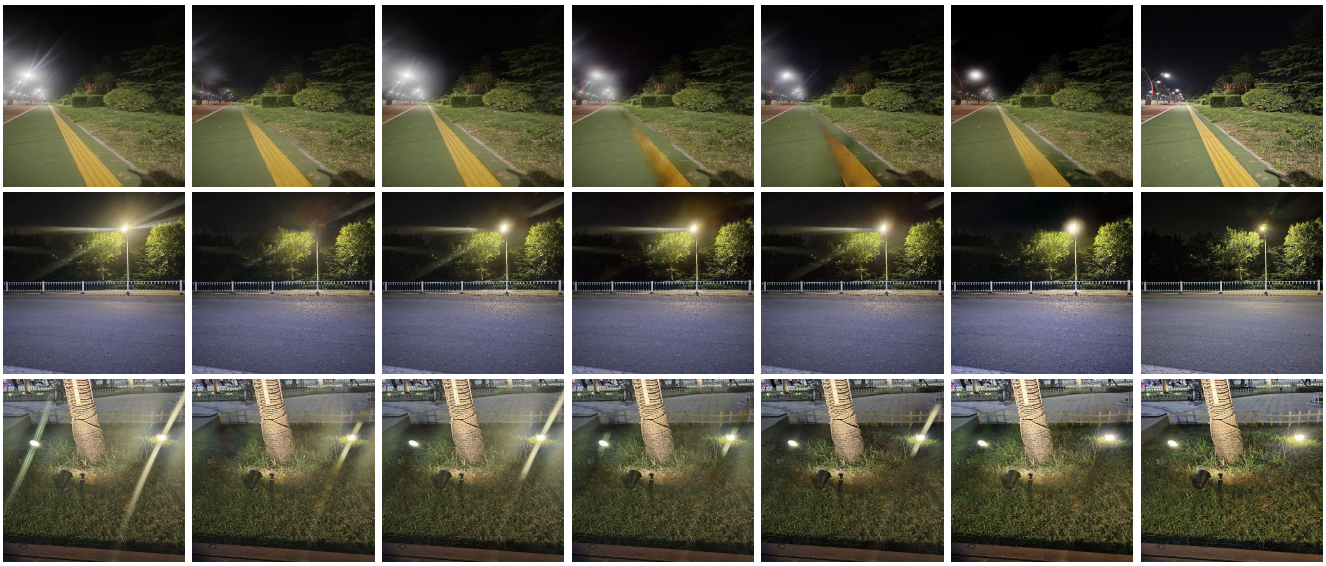

Figure 5. Visual results of flare and glare removal on real-world nighttime dataset [3]. Competing methods leave noticeable residual artifacts, while our approach effectively eliminates all flare and glare.

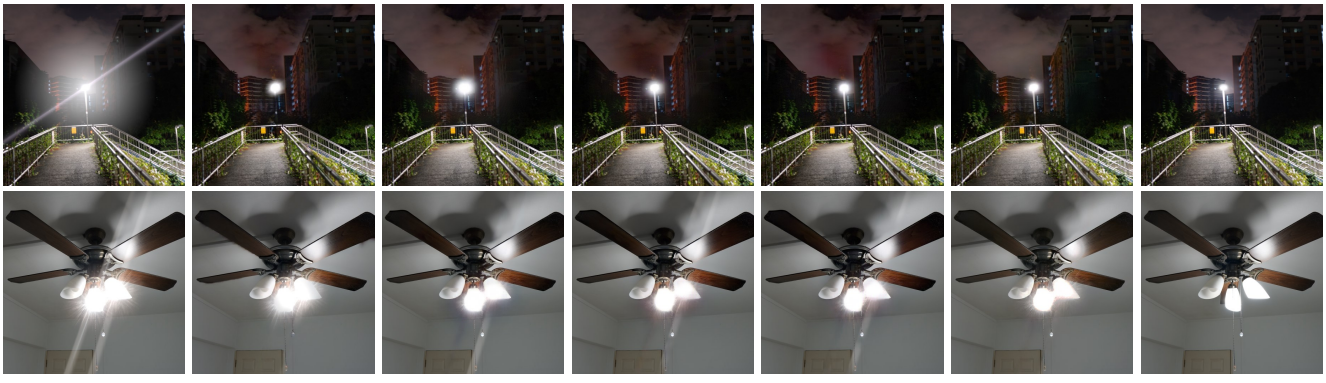

Figure 6. Visual results of flare and glare removal on a synthetic nighttime dataset [2].

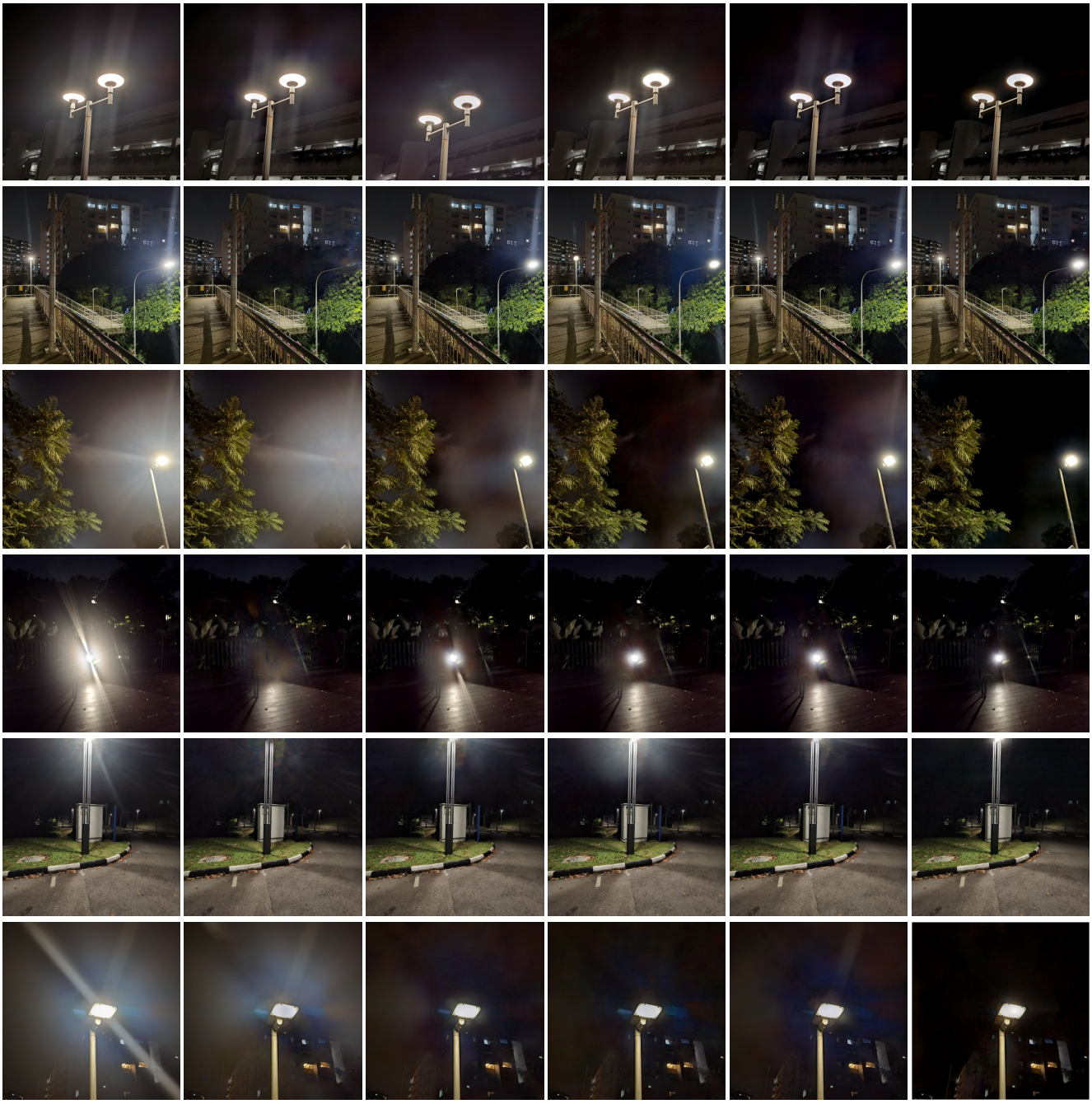

Figure 7. Visual comparison of flare and glare removal on the real-world nighttime dataset [2]. State-of-the-art methods [1, 2, 4, 5] struggle to remove extensive glow and restore affected areas. In contrast, our method achieves more complete artifact removal and clearer background restoration.

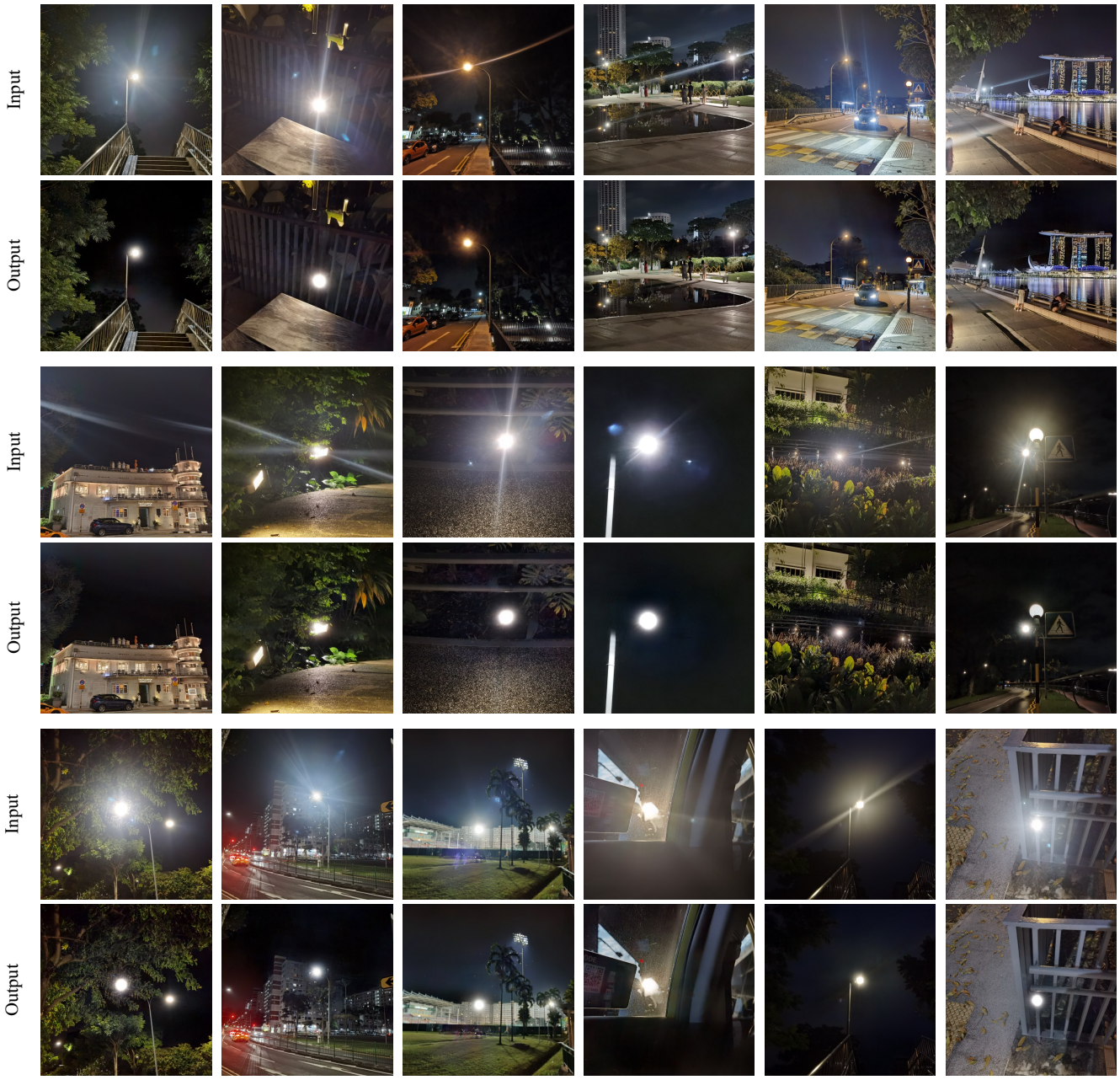

Figure 8. Generalization capability of our dataset and method for flare and glare removal on real-world nighttime images [2]. The diversity and physical realism of our PBFG dataset, combined with the comprehensive local and global feature extraction of FGRNet, ensure robust generalization across diverse scenes.