LightsOut: Diffusion-based Outpainting for Enhanced Lens Flare Removal

Supplementary Material

Figure 1. **Comparison of the downstream tasks.** The visual results indicate that LightsOut enhances performance on object detection tasks as well. Our approach not only boosts detection confidence scores but also enables the identification of objects previously undetectable due to flare artifacts.

Algorithm 1 Cropping Algorithm

```
function IMAGECROP(image)
    function LARGESTRECTANGLE(heights)
        heights.append(0)
        stack ← [−1]
        max_area ← 0
        max_bbox ← (0, 0, 0, 0)        ▷ (area, left, right, height)
        for i ← 0 to len(heights) − 1 do
            while heights[i] ¡ heights[stack[-1]] do
                h ← heights[stack.pop()]
                w ← i − stack[−1] − 1
                area ← h × w
                if area > max_area then
                    max_area ← area
                    max_bbox ← (area, stack[−1] + 1, i − 1, h)
                end if
            end while
            stack.append(i)
        end for
        return max_bbox
    end function
    max_area ← 0
    max_bbox ← [0, 0, 0, 0]
    heights ← zeros_like(image.shape[1])
    for row ← 0 to image.shape[0] − 1 do
        temp ← 1 − image[row]
        heights ← (heights + temp) × temp
        (area, left, right, height) ← LargestRectangle(heights)
        if area > max_area then
            max_area ← area
            max_bbox ← [left, right, (row−height+1), row]
        end if
    end for
    return max_bbox
end function
```

A. Appendix Section

A.1. Implementation Details

Dataset and Preprocessing. We use the benchmark dataset Flare7k [1] for both training and testing. Since the dataset was not originally designed for our tasks, we preprocess it to better suit our requirements. Specifically, to handle off-frame or incomplete light source images and define outpainted regions, we first generate YCbCr luminance masks and then apply an algorithm, formalized in Algorithm 1, to identify the largest rectangular area in each image that excludes the light source. Once the bounding box is obtained, we crop the image on-the-fly during training and inference. The cropped region is then masked with a pixel value of 127, defining the area to be outpainted.

Training Details. Our framework comprises three independently trained modules, all implemented on an NVIDIA RTX4090 GPU. The components are optimized independently, allowing each module to specialize in a distinct subtask and enabling them to collectively improve the system's overall performance when integrated. The multitask regression module was trained with a learning rate of 1×10^{-4}, batch size of 32, for 100 epochs, and we set the number of predicted light sources N to 4. The light source condition module was optimized using a learning rate of 1×10^{-5} and a batch size of 8 for 20,000 steps. Finally, the Stable Diffusion inpainting network [7] was fine-tuned using LoRA [3] with a learning rate of 1×10^{-4} and a batch size of 8 for 25,000 steps to achieve optimal performance while maintaining computational efficiency.

Inference Settings. During outpainting process, we set the number of sampling steps to 50, the guidance scale to 7.0, and perform noise reinjection 4 times. Additionally, we uti-

Figure 2. **Failure Cases.**

lize BLIP-2 [6] to automatically generate captions, thereby minimizing human bias.

Evaluation metrics. We evaluate flare removal quality using PSNR, SSIM [9], and LPIPS [10], and assess the accuracy of our light source prediction using mean Intersection over Union (mIoU).

A.2. Downstream Tasks

Lens flare artifacts can negatively impact images in various computer vision tasks. To examine how flare removal affects object detection performance, we utilize the pre-trained YOLOv11 [5] detector to compare two scenarios: images directly processed by SIFR models, and images first enhanced by our proposed outpainting approach before being input to SIFR models. Fig. 1 demonstrates that our proposed approach yields improvements in detection accuracy, particularly for objects located in regions previously compromised by flare artifacts.

A.3. In-the-Wild Images.

We present additional outpainting results on self-collected in-the-wild scenes in Fig. 4, along with flare removal comparisons against baseline methods (Zhou et al.[11], Flare7K++[2], and MFDNet [4]) in Fig. 3. These results highlight our method's effectiveness in outpainting off-frame regions and improving the performance of existing SIFR models, even on challenging in-the-wild images.

A.4. Failure Cases

The main failure cases exhibit two characteristic features. First, when the overall image brightness is high, the brightness differential between the flare and other parts of the image becomes less pronounced. Second, when the flare occupies a relatively large proportion of the entire image. Both scenarios make it difficult to delineate the flare region precisely, even with the integration of our proposed method.

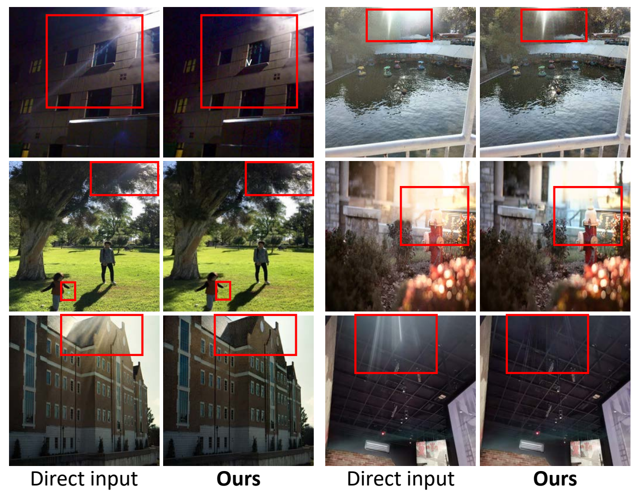

Figure 3. **Flare removal results for in-the-wild scens.** The red boxes indicate flare regions in the images. Our method effectively addresses off-frame light source scenes, which existing SIFR models fail to handle.

A.5. Qualitative Comparisons of Light Source Mask Prediction.

Fig. 5 compares the light source predictions from our multi-task regression module with those generated by U-Net [8]. The results demonstrate that our proposed module predicts the positions and radii of light sources more accurately, both in single and multiple light source scenarios.

A.6. Additional Qualitative Comparisons

We present extensive supplementary visual evidence to demonstrate the efficacy of our approach. Figures Fig. 6, and Fig. 7 showcase additional flare removal results across diverse imaging conditions. Furthermore, we provide comparative analyses between our outpainting results and those produced by both baseline methods and state-of-the-art diffusion-based inpainting and outpainting techniques in Fig. 8, Fig. 9. These comprehensive visual comparisons sub-

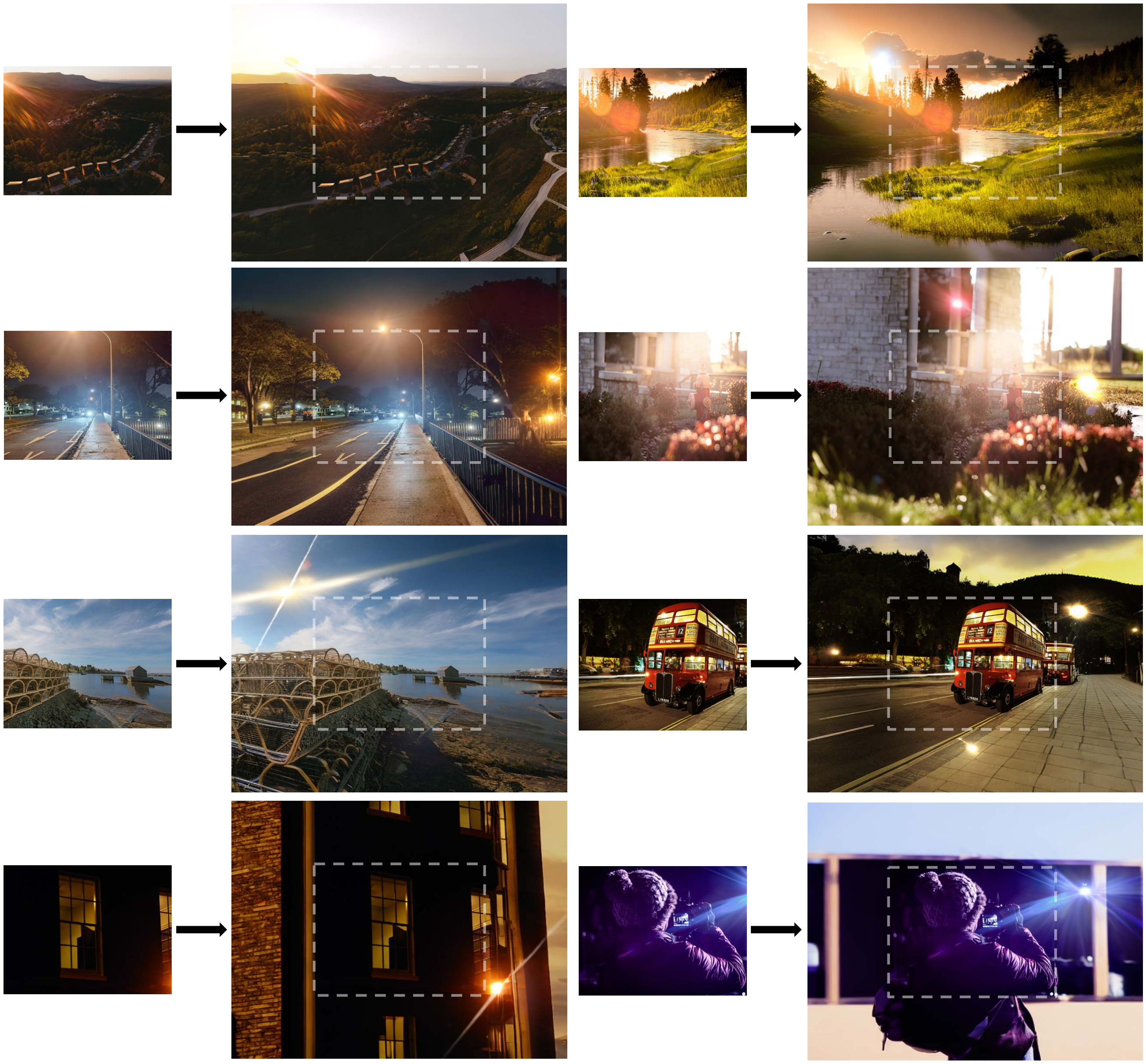

Figure 4. **Outpainting results for in-the-wild scens.**

stantiate the superior robustness and effectiveness of our proposed methodology across a wide spectrum of challenging scenarios.

References

[1] Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy. Flare7k: A phenomenological nighttime flare removal dataset. *NeurIPS*, 2022. 1

[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. *IEEE TPAMI*, 2024. 2

[3] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, Weizhu Chen, et al. Lora: Low-rank adaptation of large language models. *ICLR*, 2022. 1

[4] Yiguo Jiang, Xuhang Chen, Chi-Man Pun, Shuqiang Wang, and Wei Feng. Mfdnet: Multi-frequency deflare network for efficient nighttime flare removal. *The Visual Computer*, 2024. 2

[5] Glenn Jocher and Jing Qiu. Ultralytics yolo11, 2024. 2

[6] Junnan Li, Dongxu Li, Silvio Savarese, and Steven Hoi. Blip-2: Bootstrapping language-image pre-training with frozen

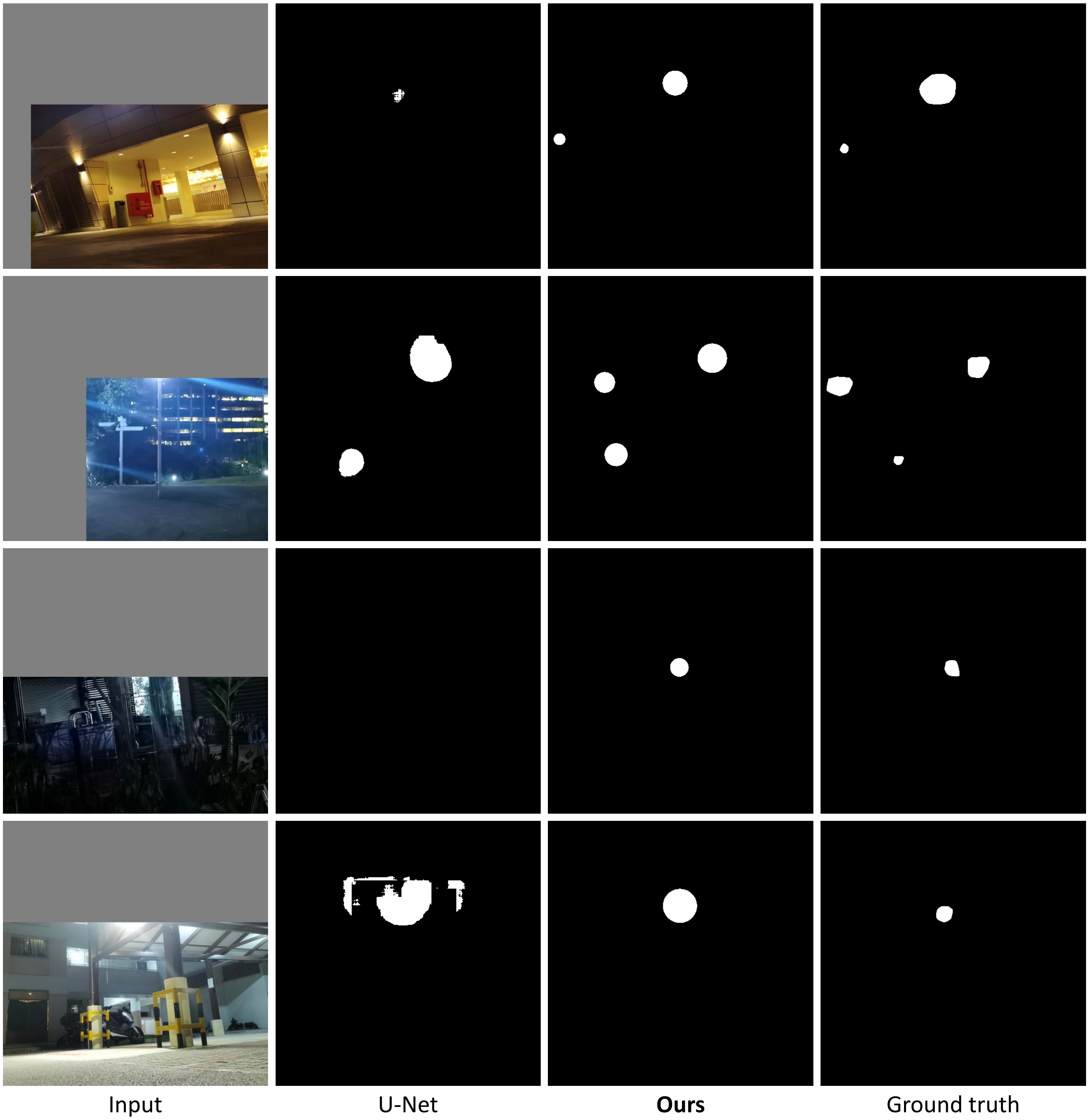

Figure 5. **Qualitative comparisons of light source mask prediction.** .

image encoders and large language models. In *ICML*, 2023. 2

[7] Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *CVPR*, 2022. 1

[8] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *MICCAI*, 2015. 2

[9] Zhou Wang, Alan C Bovik, Hamid R Sheikh, and Eero P Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE TIP*, 13(4):600–612, 2004. 2

[10] Richard Zhang, Phillip Isola, Alexei A Efros, Eli Shechtman, and Oliver Wang. The unreasonable effectiveness of deep features as a perceptual metric. In *CVPR*, 2018. 2

[11] Yuyan Zhou, Dong Liang, Songcan Chen, Sheng-Jun Huang, Shuo Yang, and Chongyi Li. Improving lens flare removal

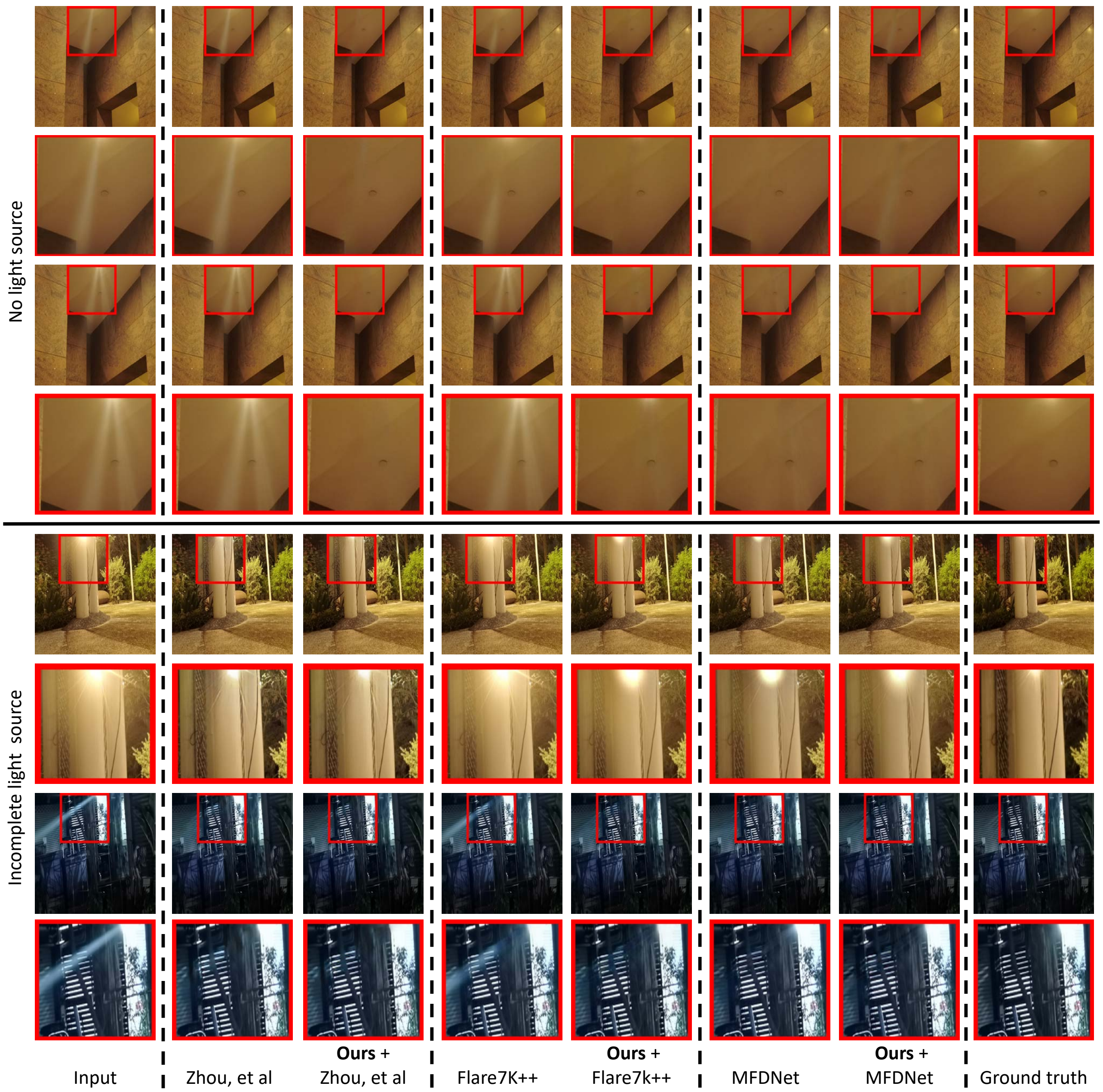

Figure 6. **Additional Qualitative Comparisons.** .

with general-purpose pipeline and multiple light sources recovery. In *ICCV*, 2023. 2

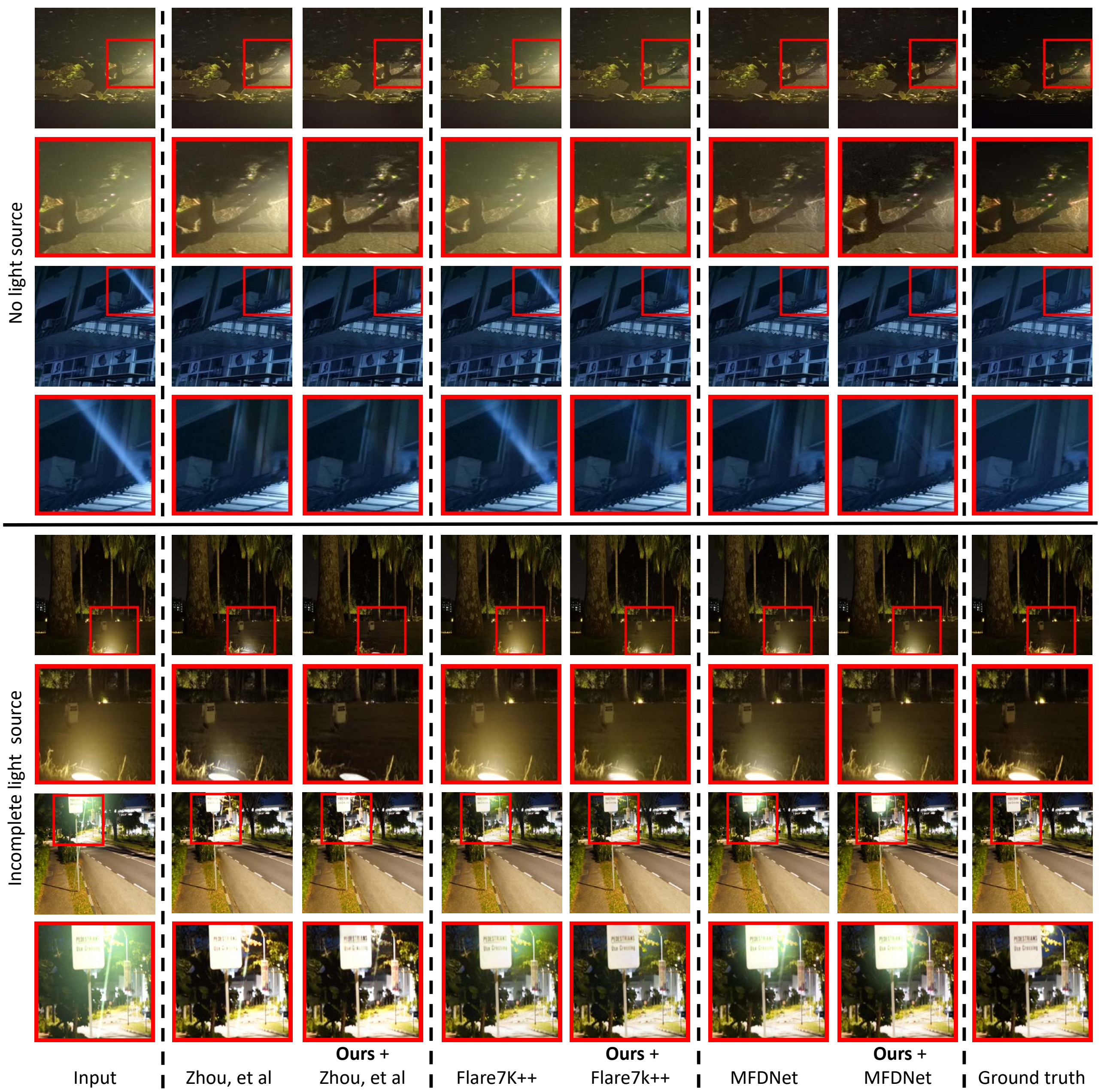

Figure 7. **Additional Qualitative Comparisons.** .

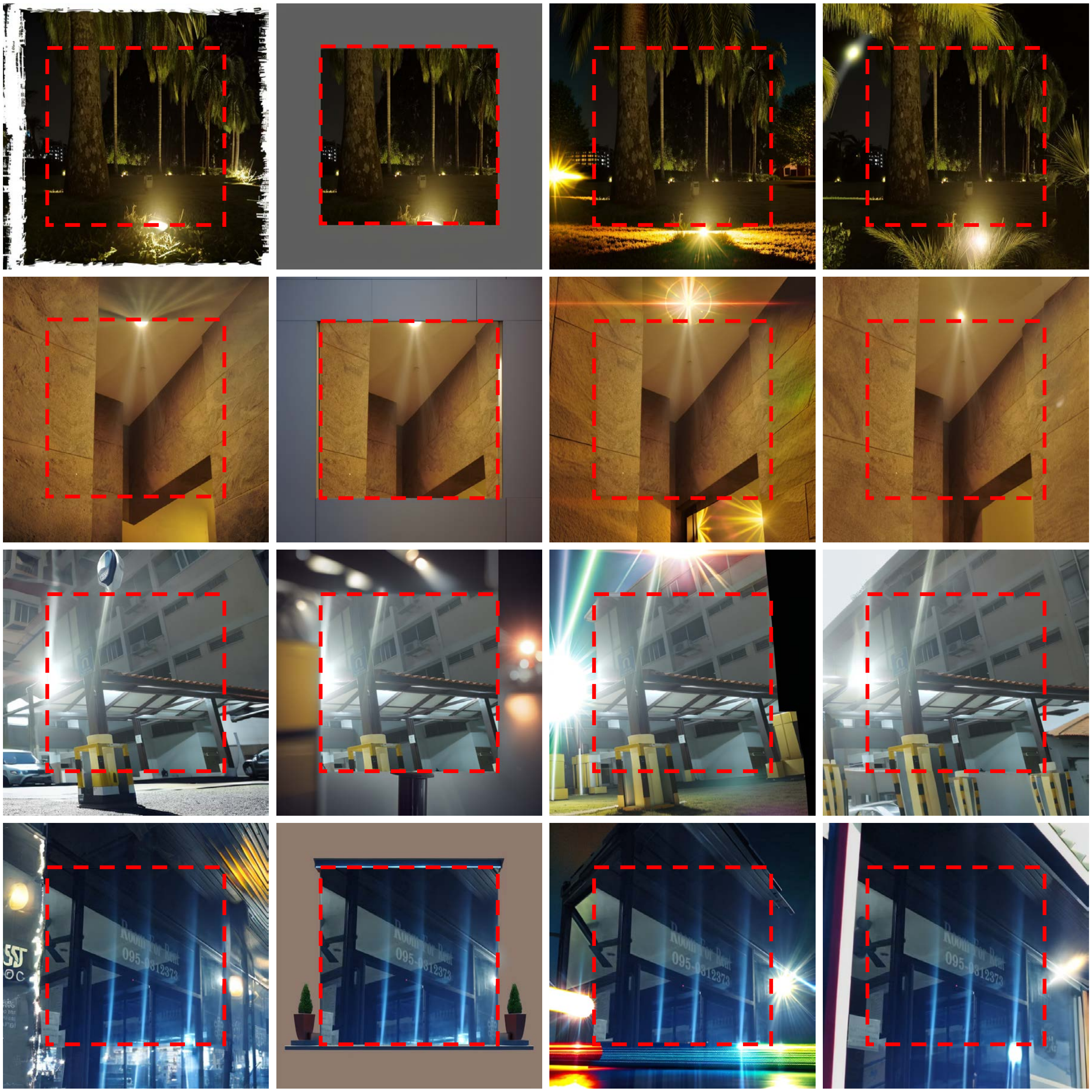

Figure 8. **Additional Qualitative Comparisons.** .

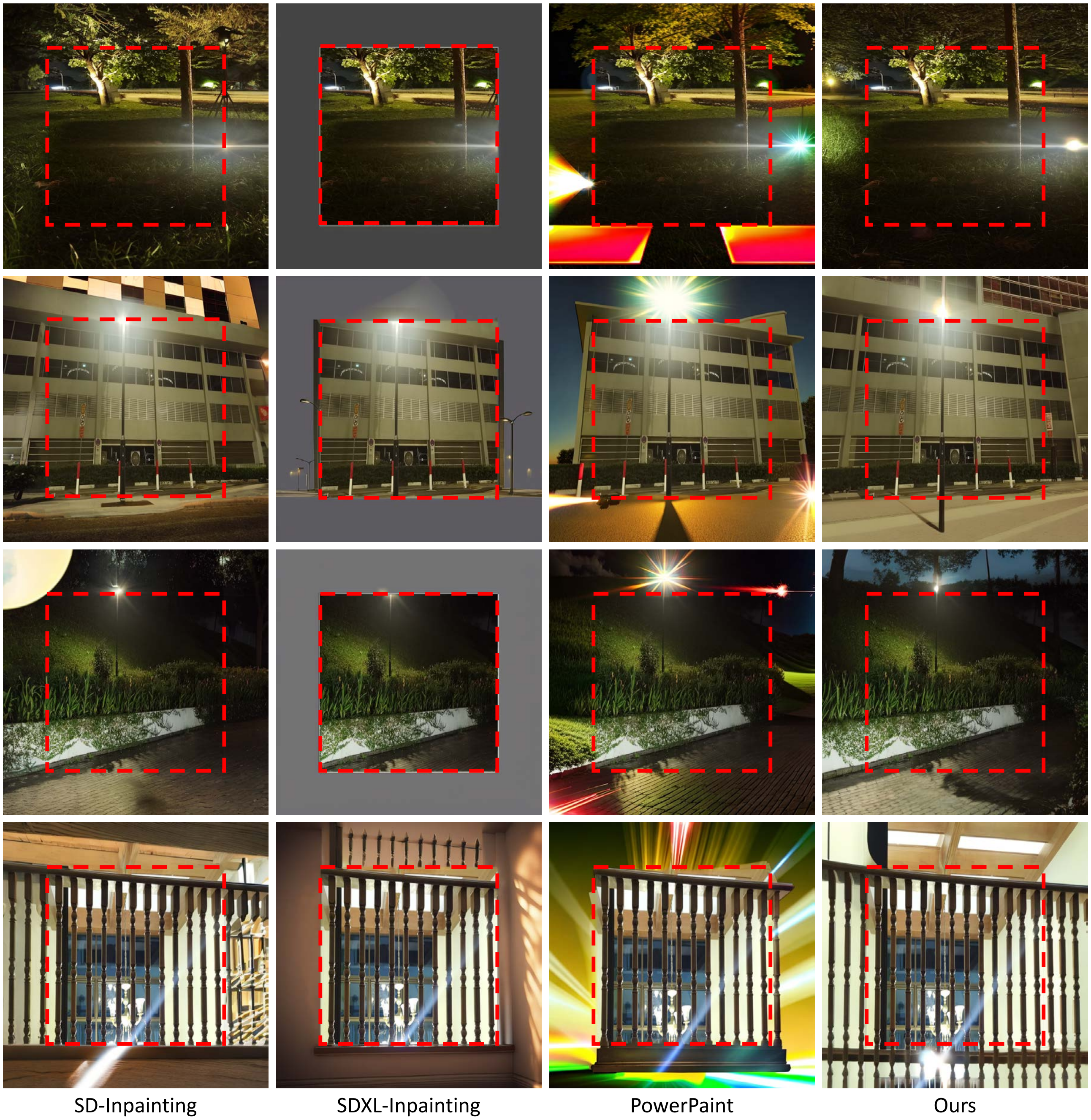

Figure 9. **Additional Qualitative Comparisons.** .