A. Comparison with Previous Pipelines

To prove the performance of our dataset, we retrain Wu et al. [7] and Dai et al. [2] on our BracketFlare dataset and conduct additional experiments as shown in Table 1. Wu et al. [7] use U-Net as a baseline model, but we find it struggled to locate reflective flares without our prior, sometimes reducing PSNR compared to the input. Dai et al. [2] set Uformer as a baseline. Since it does not encode the prior into the network, it has an obvious gap between our baseline method. To show these methods’ effects on real data, we also conduct a user study with 20 real-captured flare-corrupted images and 20 participants. The participants are asked to identify which of these models did best at removing the reflective flares. The experiments presented in Table 2 illustrates that users strongly prefer our methods over previous methods.

Table 1. Comparison of previous models retrained on our dataset.

<table>
<thead>
<tr>
<th>Model (Retrained)</th>
<th>PSNR</th>
<th>SSIM</th>
<th>LPIPS</th>
<th>Masked PSNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Input</td>
<td>37.30</td>
<td>0.990</td>
<td>0.025</td>
<td>21.68</td>
</tr>
<tr>
<td>Wu et al. [7]</td>
<td>31.84</td>
<td>0.912</td>
<td>0.032</td>
<td>24.87</td>
</tr>
<tr>
<td>Dai et al. [2]</td>
<td>42.70</td>
<td>0.987</td>
<td>0.010</td>
<td>27.46</td>
</tr>
<tr>
<td>Ours</td>
<td>48.41</td>
<td>0.994</td>
<td>0.004</td>
<td>32.09</td>
</tr>
</tbody>
</table>

B. Flare Removal for Downstream Tasks

Fig. 1 demonstrates that it can benefit downstream tasks such as image segmentation. In the figure, it can be observed that the reflective flares are misclassified as a pole, which may pose potential risks for nighttime driving. Removing such reflective flares can help achieve more robust and reliable results, and ultimately benefit the users.

C. Potential in Overexposure Restoration

Due to the limited dynamic range of some cameras, image details especially in light source regions are always saturated and difficult to recover. To achieve the details of the overexposed regions, the mainstream solution is to use HDR (high-dynamic range) imaging with multi-exposure capture [3, 4]. Our main idea may provide a potential solution to this issue.

Specifically, since smartphone reflective flare can be considered as a short-exposure image, our method also provides the potential in implementing overexposed regions restoration. As shown in Fig. 2, we can separate a reflective flare and rotate the estimated flare 180 degrees around the optical center to match the flare with the light source. We suppose the exposure step between the flare and the original input is 12 EV. Then, these two images with low-dynamic ranges will be merged to generate an HDR image with clear details in the light source. The visualization results in Fig. 2 show that our method can recover the details of the saturated regions well.

D. More Visual Results

To further demonstrate the effectiveness of our dataset, we retrain different neural networks including Uformer [5],
Restormer [8], HINet [1], and MPRNet [9] using our proposed dataset. The visual results of different networks for restoring real-world images are presented in Fig. 3. From the visual comparison, we can observe that all the networks retrained on our dataset can remove the flares. Among them, MPRNet [9] obtains the best performance on flare-corrupted regions, as highlighted in the red boxes of Fig. 3. The results manifest the effectiveness of our proposed dataset.

To show our dataset and prior’s generalization ability in different real-world scenes, we show more results of the flare-corrupted images captured by iPhone 13 Pro that is known for severe reflective flares in Fig. 4 and Fig. 5. The results show that our proposed method can tackle a huge diversity of light sources and achieve good visual performance in different scenes. It is owing to the versatility of our proposed prior, as well as the diversity and balanced distribution of our proposed dataset.

We also capture many flare-corrupted video clips by Huawei P40, iPhone 13 Pro, and ZTE Axon 20 5G. To prevent triggering the anti-shake module of the smartphone, we try to keep the smartphone from large movement. Then, we process these videos with our method frame by frame. The video results can be found in a separate video demo. The video demo shows that our method can achieve robust and consistent flare-removal results even for real-world videos.

E. Limitation

As shown in Fig. 10 of the main paper, our method can generalize well to different types of smartphones. However, it is based on the fact that most smartphone cameras satisfy the optical center symmetry prior. For some professional cameras, this prior does not always hold. Besides, due to the dispersion of thick lenses in professional cameras, the main flare spot of the professional cameras are not always the same as the light source’s pattern. The solutions for these cases will be left for future work.

References

Figure 3. Visual comparison of different networks retrained on our dataset for restoring the real-world flare-corrupted images. These image restoration networks include Uformer [5], Restormer [8], HINet [1], and MPRNet [9].
Figure 4. Our results on real-world nighttime flare-corrupted images. In this figure, we use MPRNet [9] as the baseline method to estimate the flare and output image.
Figure 5. More results on real-world nighttime flare-corrupted images. In this figure, we use MPRNet [9] as the baseline method to estimate the flare and output image.