Flare7K++: Mixing Synthetic and Real Datasets for Nighttime Flare Removal and Beyond

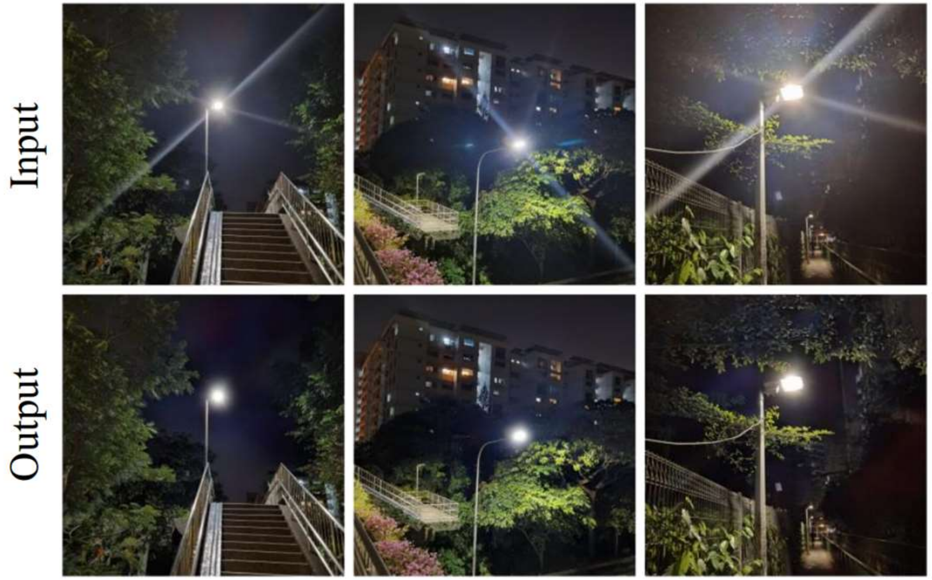
Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng,
Yihang Luo, Chen Change Loy, Senior Member, IEEE

TPAMI, 2024

Nighttime Flare Removal

Lens flare is an optical phenomenon in which intense light is scattered and/or reflected in an optical system. It leaves a radial-shaped bright area and light spots on the captured photo. The effects of flares are more severe in the nighttime environment due to the existence of multiple artificial lights. This phenomenon may lead to low contrast and suppressed details around the light sources, degrading the image's visual quality and the performance of vision algorithms.

Furthermore, the spectrum of artificial lights often differs significantly from that of the sun, leading to different diffraction patterns that further complicate flare modeling. The differences between existing synthetic flares and real-world nighttime flares make it difficult to develop models that can generalize well across a variety of nighttime situations.

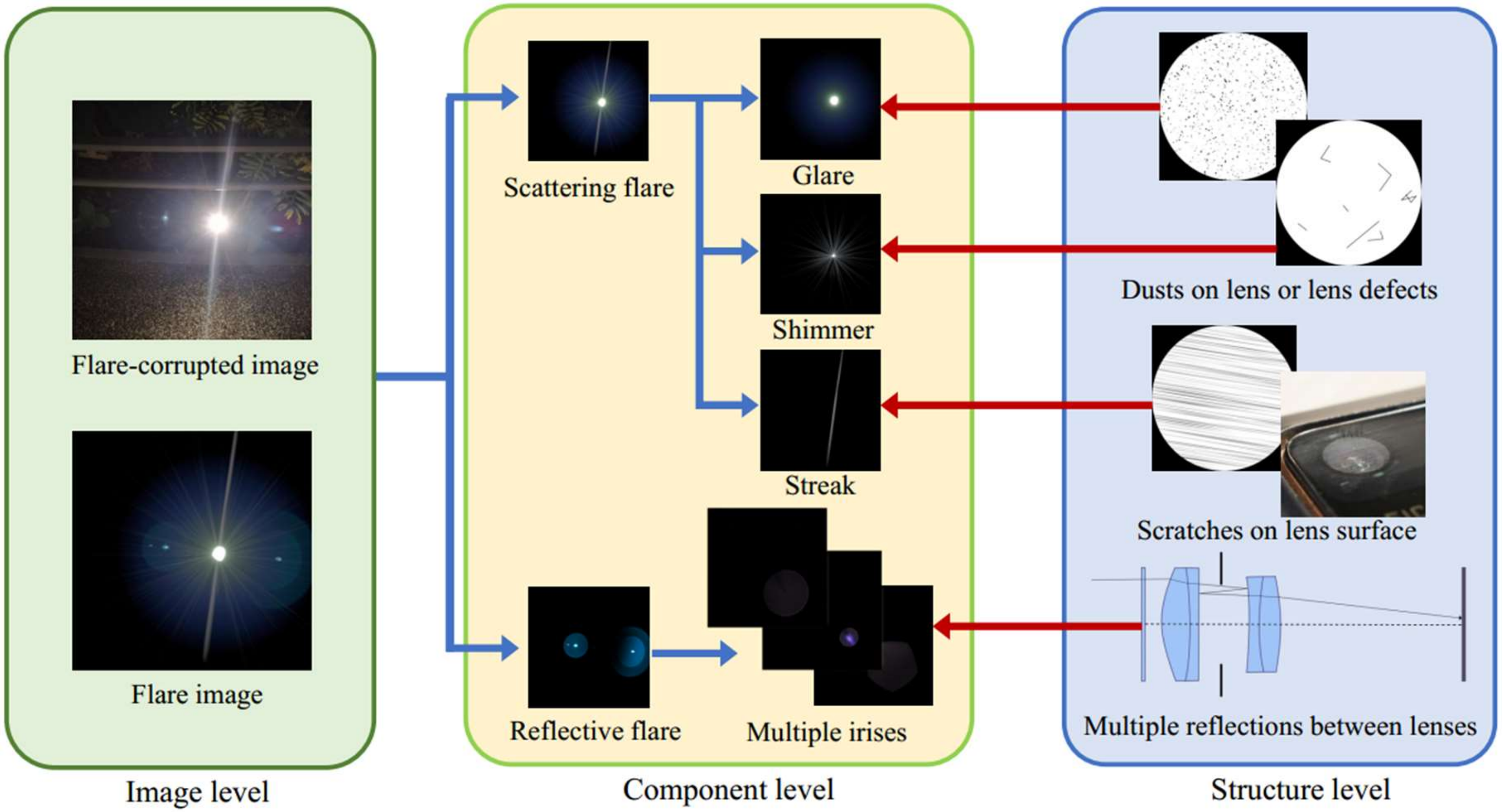


Contributions

- We create a mixing dataset called Flare7K++ that augments the synthetic **Flare7K** dataset (conference version) with a new **real-captured Flare-R** dataset. Utilizing this mixing dataset to train deep models can significantly improve the effectiveness of removing strong degradation.
- We design a new **end-to-end framework** for flare removal, which preserves the light source while removing lens flares.
- We manually label masks for different flare components of test flare-corrupted images and introduce two new metrics, **GPSNR and S-PSNR**, to reflect restoration results on glare and streak corrupted regions.

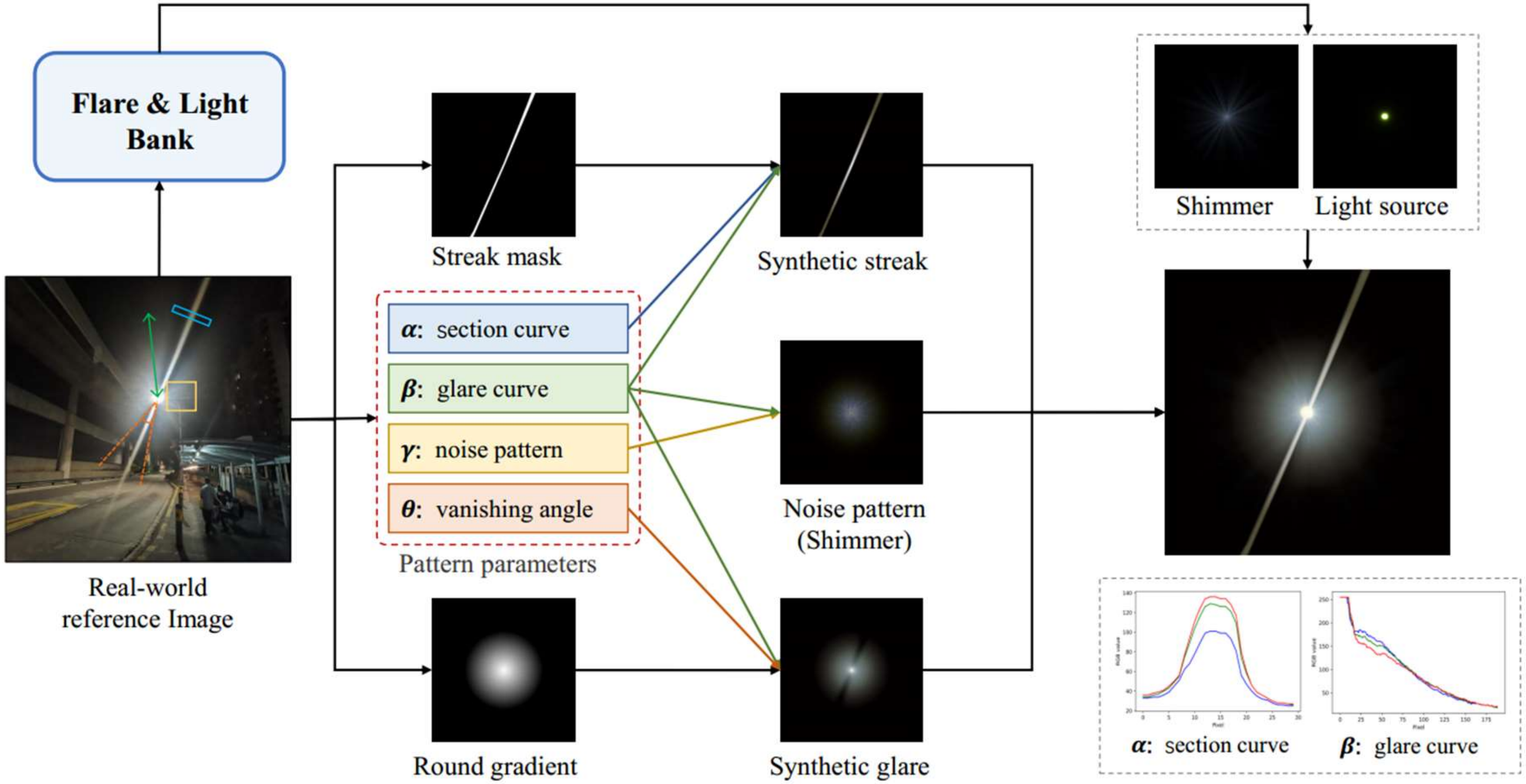
Flare7K++ Dataset

Formulation of nighttime lens flare.



Flare7K++ Dataset

Flare7K Dataset



Flare7K++ Dataset

Flare-R Dataset

Since the synthetic dataset does not contain complicated degradation caused by diffraction and dispersion in the lens system, we also capture a real-world flare dataset, Flare-R, with 962 flare patterns.

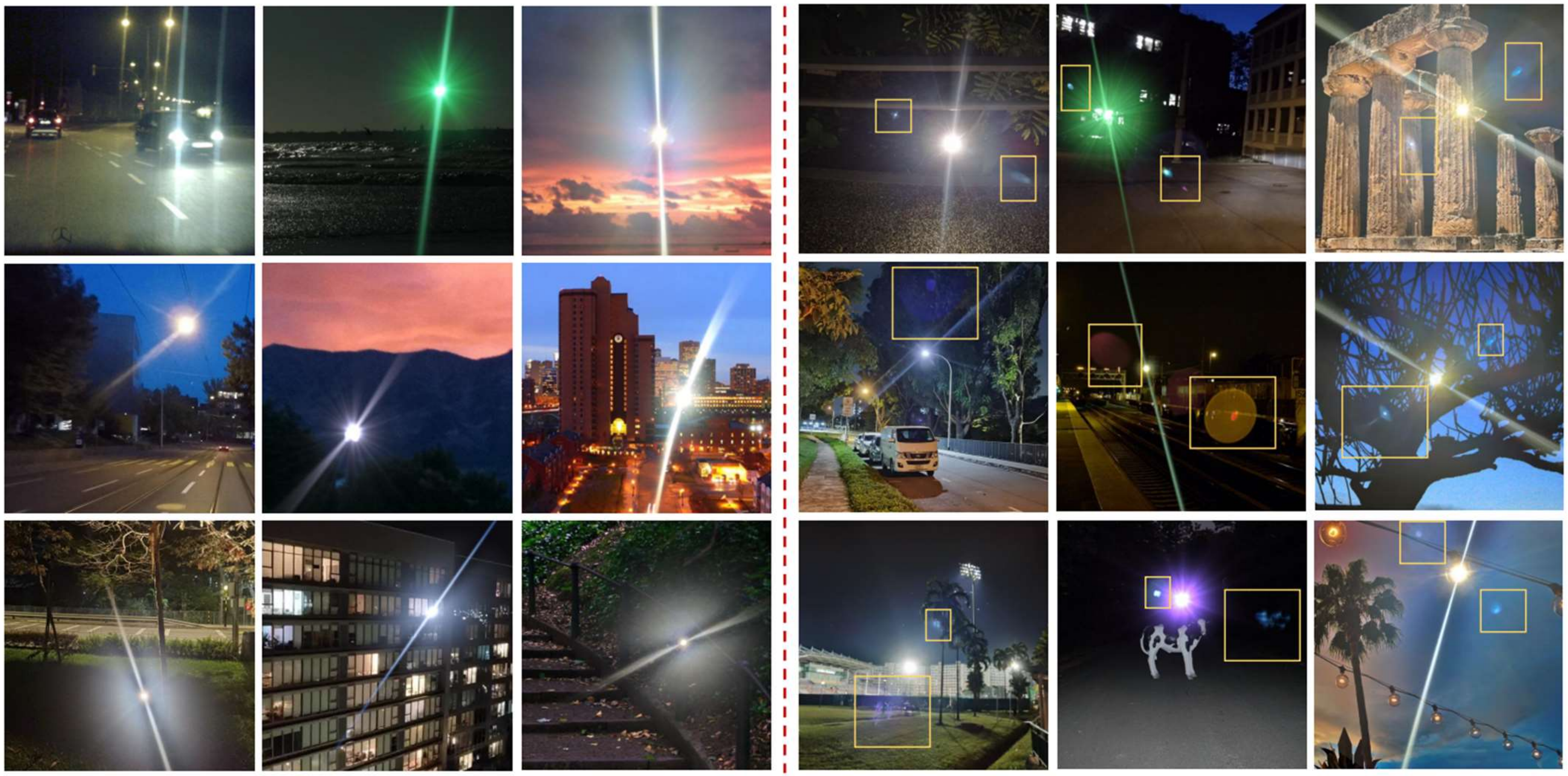
• We dip **different types of liquid** including water, oil, ethyl alcohol, and carbonated drinks on the lens surface and **wipe it with our fingers and different types of clothes**. After each wipe, we capture a new lens flare image.

• We disable the automatic white balance of the phone cameras and **obtain different-colored lens flare** images by changing the color temperature of the light source.

• The Huawei P40, iPhone 13 Pro, and ZTE AXon20 5G all have three rear lenses with **different focal lengths**.

• For each lens, we **vary the distance** from the camera to the light source and capture approximately 100 images.

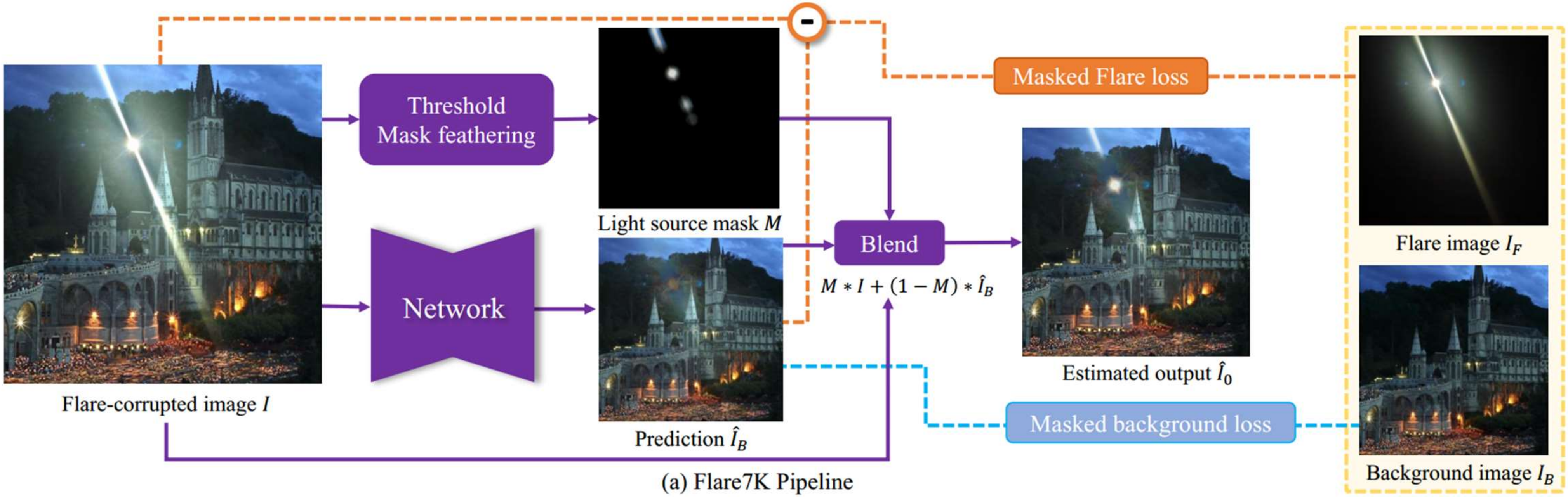
Flare7K++ Dataset



(a) Scattering flare (b) Flare7K (Ours) (c) Flare-R (Ours) (d) Reflective flare (e) Flare7K (Ours) (f) Flare-R (Ours)

Flare7K Pipeline

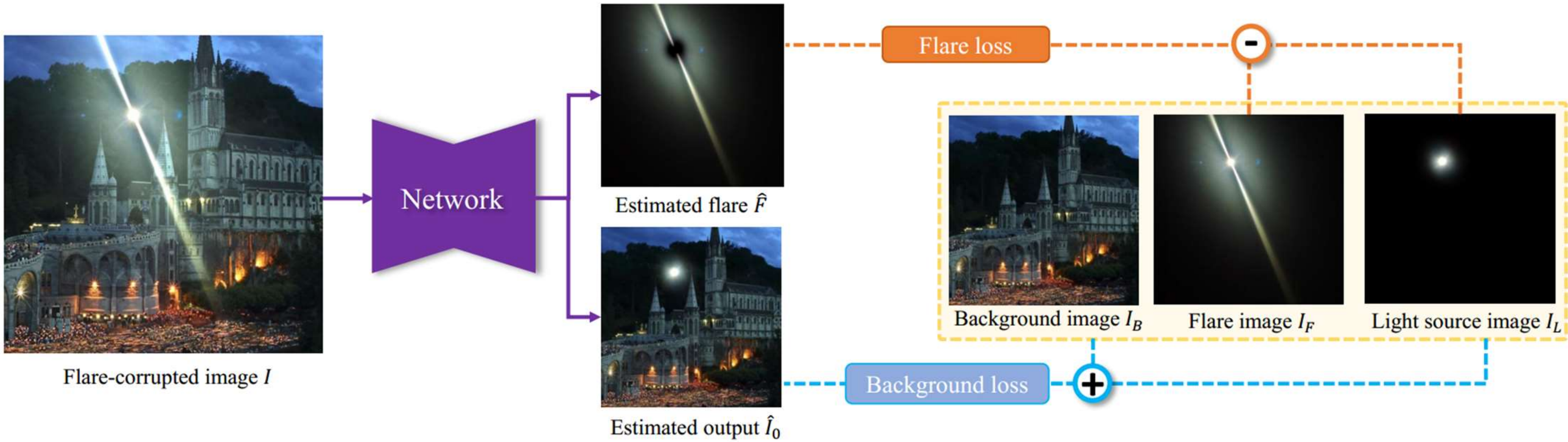
In the inference pipeline, the network outputs a flare-free prediction that lacks a light source. The flare-corrupted image is then processed with a manual threshold value and blur kernel to extract the light source mask. However, the streak region may still get overexposed, leaving parts of the flare on the image. Moreover, thresholding-based algorithms may fail to segment light sources that are not bright enough, and blurring operations may inadvertently erase tiny light sources.



(a) Flare7K Pipeline

Flare7K++ Pipeline

Light source image IL is added to the background images IB to synthesize the ground truth of flare-free images I0. Then, the difference between the flare image IF and light source image IL is calculated to supervise the estimated flare ˆF. This allows us to train end-to-end networks without manually setting thresholds to extract the light source. Thus, our pipeline enables a network to distinguish overexposed light sources and overexposed streaks. Even for non-saturated light sources, our learning-based method can also produce accurate output.



(b) Flare7K++ Pipeline

Loss Function

$$\mathcal{L} = w_1 L_B + w_2 L_F + w_3 L_{rec},$$

w1, w2, w3 are respectively set to 0.5, 0.5, and 1.0

L_B – background loss

L_F – flare loss

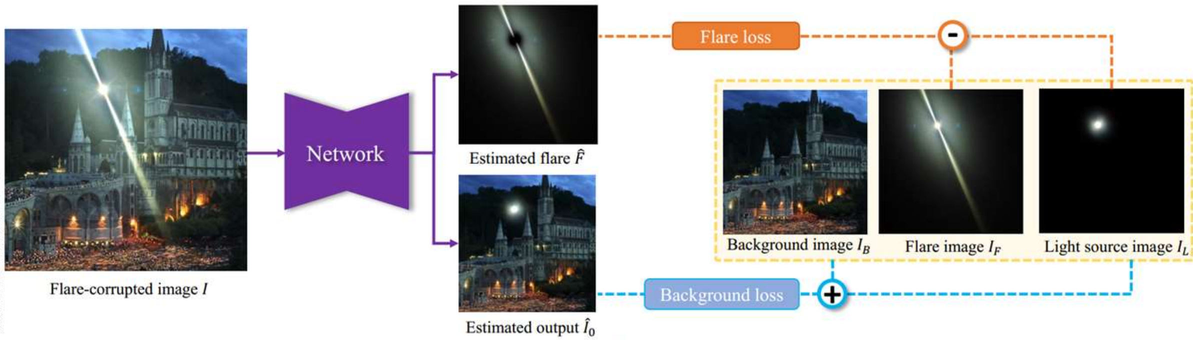
Flare and background images are supervised using L_1 loss and perceptual loss L_{vgg}. The background image loss L_B can be written as:

$$L_B = L_1(\hat{I}_0, I_0) + L_{vgg}(\hat{I}_0, I_0),$$

shares the same expression with the flare loss L_F

L_{rec} – reconstruction loss重建损失

$$L_{rec} = |I - Clip(\hat{I}_0 \oplus \hat{F})|,$$



(b) Flare7K++ Pipeline

Experiments

Comparison with previous works

Metric\Method	Input	Previous work				Network trained on Flare7K++				
		Zhang [42]	Sharma [28]	Wu [37]	Flare7K baseline	U-Net [25]	HINet [4]	MPRNet* [40]	Restormer* [39]	Uformer [33]
PSNR↑	22.561	21.022	20.492	24.613	26.978	27.189	27.548	27.036	27.597	**27.633**
SSIM↑	0.857	0.784	0.826	0.871	0.890	0.894	0.892	0.893	**0.897**	0.894
LPIPS↓	0.0777	0.1738	0.1115	0.0598	0.0466	0.0452	0.0464	0.0481	0.0447	**0.0428**
G-PSNR↑	19.556	19.868	17.790	21.772	23.507	23.527	**24.081**	23.490	23.828	23.949
S-PSNR↑	13.105	13.062	12.648	16.728	21.563	22.647	**22.907**	22.267	22.452	22.603

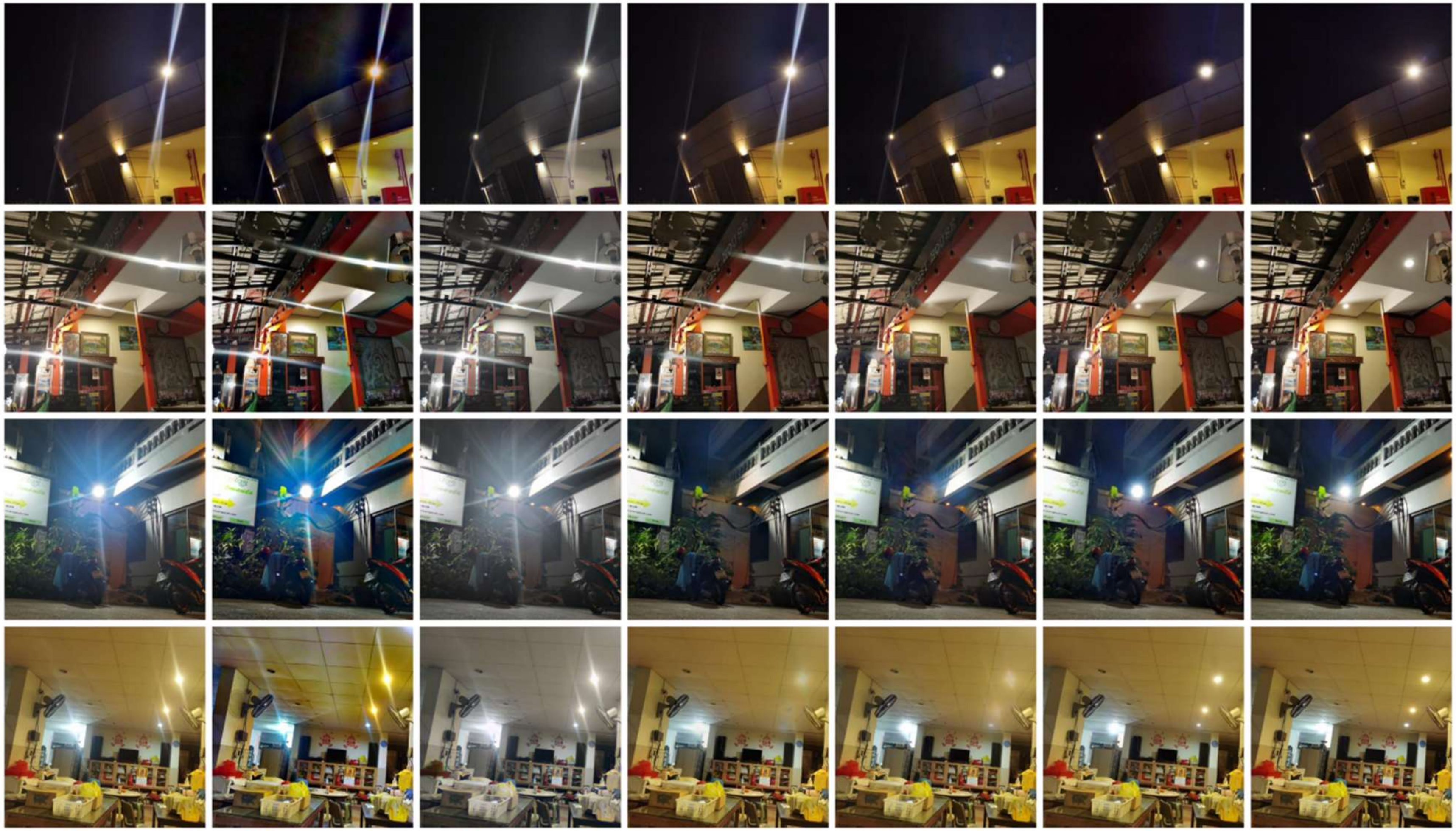
Two new metrics--- GPSNR, S-PSNR

Since the ground truth may still be influenced by the slight flares brought by the lens's defects, global PSNR cannot fully reflect the performance of flare removal methods.
We manually label masks for different flare components of test flare-corrupted images and introduce two new metrics, to reflect restoration results on glare and streak corrupted regions.

Flare-corrupted image Ground truth Annotations

Glare, streak, and light sources are labeled in yellow, red, and blue, respectively. S-PSNR means the PSNR in the red region and G-PNSR represents the PSNR in the sum of the yellow and red regions.

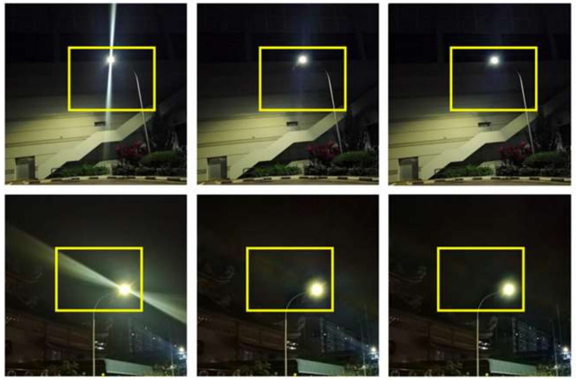
Experiments



Input | Zhang et al. | Sharma et al. | Wu et al. | Flare7K | Flare7K++ | GT

Ablation Study

Dataset

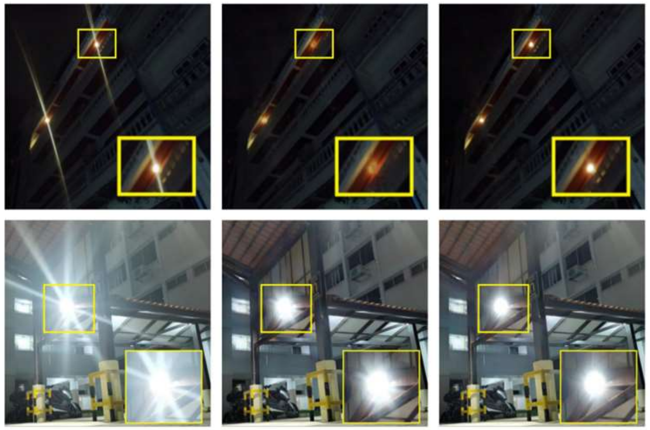


Input Flare7K Flare7K++

Training sets		Real captured test set				
Flare7K	Flare-R	PSNR	SSIM	LPIPS	G-PSNR	S-PSNR
✓		27.257	0.890	0.0471	23.762	21.294
✓	✓	**27.633**	**0.894**	**0.0428**	**23.949**	**22.603**

Light Source Completion

We train a Uformer without using light source annotations.



Input No light source With light source

Light source	PSNR	SSIM	LPIPS	G-PSNR	S-PSNR
w/o light source	26.850	**0.895**	0.0473	23.441	21.909
with light source	**27.633**	0.894	**0.0428**	**23.949**	**22.603**

Ablation Study

Training Pipeline

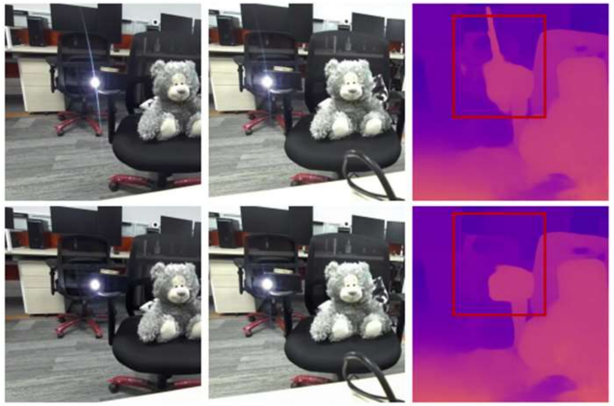
Comparison for training networks with Flare7K's pipeline and our new pipeline. While training the network, we do not change the loss function. All networks are trained with Flare7K++.

Input Flare7K Flare7K++

Pipeline	PSNR	SSIM	LPIPS	G-PSNR	S-PSNR
Wu *et al.* [37]	27.166	0.861	0.0432	23.598	22.118
Ours	**27.633**	**0.894**	**0.0428**	**23.949**	**22.603**

Flare Removal for Downstream Tasks

Stereo Matching.

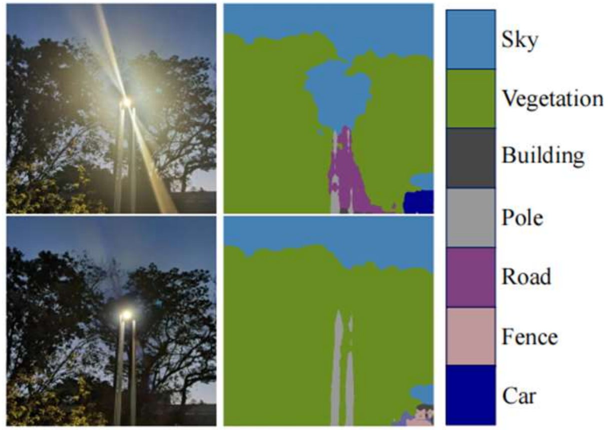


Left image Right image Disparity map

Optical Flow.

Frame 0 Frame 1 Optical flow

Semantic Segmentation.



Input/output Segmentation map