VPRP: VISUAL PROMPT RESTORATION PIPELINE FOR HIGH-QUALITY NIGHTTIME FLARE REMOVAL

Anonymous authors
Paper under double-blind review

ABSTRACT

Lens flare removal is essential for restoring images degraded by unwanted light scattering or reflection, yet existing methods frequently introduce artifacts in restored outputs. End-to-end approaches struggle with severe degradations where flare completely occludes underlying content, while conventional coarse-to-fine strategies yield marginal improvements when applied directly to flare removal. We identify that the core limitation lies in the lack of explicit guidance for identifying and restoring flare-corrupted regions. To address this, we propose the Visual Prompt Restoration Pipeline (VPRP), a model-agnostic framework that redefines coarse-to-fine processing through a plug-and-play Prompt Restoration Network (PRN). VPRP introduces two synergistic mechanisms: (1) a mask block that leverages predicted flare from the coarse stage to explicitly exclude corrupted pixels, enabling context-aware restoration, and (2) a prompt calibration block that repurposes high-quality decoder features as visual prompts to guide the refinement network, ensuring structural consistency. VPRP seamlessly integrates with arbitrary restoration architectures while introducing negligible computational overhead (minimum 1% parameter increase). Experiments demonstrate that VPRP significantly reduces artifacts and achieves state-of-the-art performance across multiple benchmarks, with PSNR improvements of up to 0.92 dB over strong baselines.

1 INTRODUCTION

Lens flare is an optical artifact caused by light scattering and reflection within camera imaging systems. When strong light enters the lens, it creates bright streaks, halos, or shimmering regions that degrade image quality and impair downstream visual tasks such as semantic segmentation, depth estimation, and object detection Dai et al. [2023a]. Removing flare from images has thus become an important problem in computational photography, with applications spanning smartphone imaging, autonomous driving, and surveillance systems Asha et al. [2019]; Vitoria & Ballester [2019]; Feng et al. [2023]; Sun et al. [2020]; Wu et al. [2021]; Dai et al. [2022].

Recent deep learning approaches have advanced flare removal by learning the mapping from flare-corrupted to flare-free images. Wu et al. Wu et al. [2021], Dai et al. Dai et al. [2022], and Dai et al. Dai et al. [2023b] have proposed datasets and employed end-to-end restoration networks Zamir et al. [2021]; Chen et al. [2021]; Wang et al. [2022]; Zamir et al. [2022] to remove flare. However, these methods share a persistent limitation: they frequently produce visible artifacts in restored images. As shown in Figure 1(a), the Flare7k++ pipeline Dai et al. [2023a] leaves noticeable artifacts when processing real-world flare images. This occurs because flare regions often completely occlude underlying content, leaving the network with insufficient information for faithful reconstruction.

The image restoration community has long employed coarse-to-fine strategies to address similar challenges Zamir et al. [2021]; Chen et al. [2021]; Li et al. [2023]; Nehete et al. [2024]. These methods first perform coarse restoration to remove major degradations, then refine details and suppress artifacts in a second stage. However, directly applying this paradigm to flare removal yields limited improvements. Figure 1(b) compares single-stage methods (UNet Ronneberger et al. [2015], Uformer Wang et al. [2022], Restormer Zamir et al. [2022]) with multi-stage methods (MPRNet Zamir et al. [2021], HINet Chen et al. [2021]) on the Flare7K++ benchmark Dai et al. [2023a]. Surprisingly, multi-stage methods do not outperform single-stage ones, MPRNet even underperforms

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

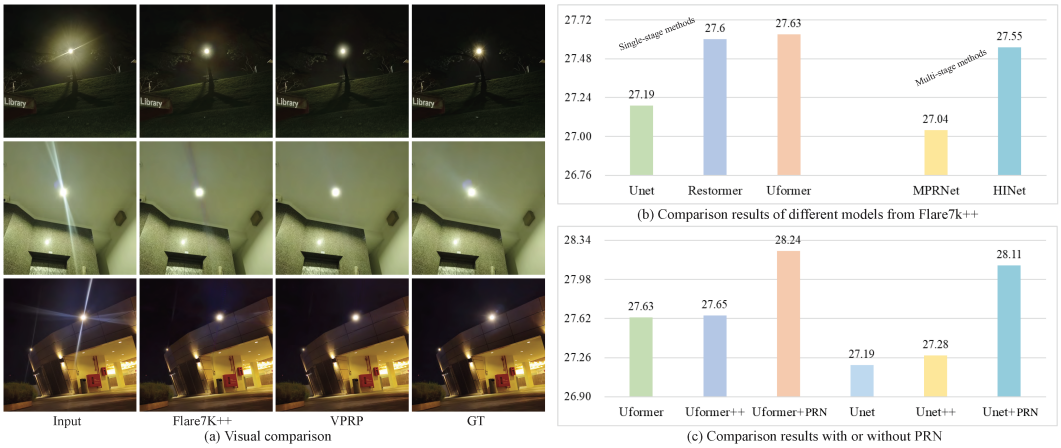


Figure 1: (a) Visual comparison on real-world nighttime flare images using Uformer as the base network. Our VPRP pipeline provides significant improvement in artifact suppression. (b) Comparison of single-stage and multi-stage methods on Flare7k++. Multi-stage methods do not outperform single-stage counterparts. (c) Simply appending a refinement network yields negligible gains.

the simpler UNet. Figure 1(c) further shows that appending a simple U-shape refinement network to pretrained models (denoted as Uformer++ and Unet++) provides negligible gains. These observations raise a critical question: Is the coarse-to-fine paradigm fundamentally unsuitable for flare removal, or does it require a principled redesign?

We identify that the core limitation of conventional coarse-to-fine approaches lies in their lack of explicit guidance for identifying and restoring flare-corrupted regions. Without knowing which pixels are affected, the refinement network cannot effectively leverage contextual information from non-corrupted areas. Inspired by inpainting techniques Dong et al. [2022]; Podell et al. [2023]; Suvorov et al. [2022] and prompt-based learning Brown et al. [2020]; Rombach et al. [2022]; Zhang et al. [2023b], we propose the Visual Prompt Restoration Pipeline (VPRP), which redefines coarse-to-fine processing through a plug-and-play Prompt Restoration Network (PRN). The key insight is to treat flare removal as a guided inpainting problem: we explicitly mask flare-corrupted regions and restore them using structural and semantic cues from the coarse stage.

Figure 2 illustrates the VPRP pipeline. Similar to prior work, we decompose flare removal into a coarse stage and a refinement stage. The coarse stage employs any U-shape network to estimate the flare-free image and the flare component. The refinement stage then introduces two key mechanisms. First, a mask block generates a binary mask from the predicted flare to explicitly exclude corrupted pixels, enabling the network to focus on restoring affected regions. Second, a prompt calibration block repurposes high-quality decoder features from the coarse stage as visual prompts, guiding the refinement network to maintain structural consistency. This design allows VPRP to seamlessly integrate with arbitrary restoration architectures while introducing minimal computational overhead.

We conduct extensive experiments across diverse architectures, from lightweight CNN-based models (UNet Ronneberger et al. [2015]) to heavyweight Transformer-based designs (Uformer Wang et al. [2022], FF-Former Zhang et al. [2023a]). In all cases, VPRP achieves significant performance improvements, with PSNR gains of up to 0.92 dB over strong baselines.

Our contributions are summarized as follows:

- We identify that conventional coarse-to-fine methods fail in flare removal due to the lack of explicit guidance for corrupted regions, and propose the Visual Prompt Restoration Pipeline (VPRP) that addresses this through mask-guided inpainting and visual prompting.
- We design the Prompt Restoration Network (PRN), a model-agnostic module that integrates seamlessly with any U-shape restoration network, whether CNN or Transformer-based, with minimal parameter overhead.
- Extensive experiments demonstrate that VPRP achieves state-of-the-art performance across multiple benchmarks, with consistent improvements on diverse backbone architectures.

2 RELATED WORK

2.1 DATASETS FOR FLARE REMOVAL

The quality of flare removal methods depends heavily on the availability of high-quality paired datasets. Wu et al. Wu et al. [2021] proposed a semi-synthetic dataset containing 2001 captured flare images and 3000 simulated flare images, focusing primarily on daylight scenarios. However, this dataset exhibits limited generalization to real-world nighttime flare, which involves diverse lenses and light sources. To address this gap, Dai et al. Dai et al. [2022] introduced Flare7K, a nighttime flare removal dataset with 5000 scattering and 2000 reflective flare images. Dai et al. Dai et al. [2023a] later extended this to Flare7K++, adding light source annotations and a real-captured test set (Flare-R) with 962 images. Recognizing that existing datasets emphasize scattering flare while giving insufficient attention to reflective flare, Dai et al. Dai et al. [2023b] proposed the BracketFlare dataset, the first dedicated to reflective flare removal, based on the observation that reflective flare and light sources are symmetric around the optical center. Despite these advances, existing methods trained on these datasets still produce artifacts, motivating our approach to improve restoration quality through a principled pipeline design.

2.2 FLARE REMOVAL METHODS

Early deep learning approaches relied on unsupervised learning due to scarce paired training data. He et al. He et al. [2010] employed unsupervised methods, while Qiao et al. Qiao et al. [2021] proposed a GAN-based framework using CycleGAN Zhu et al. [2017] principles to detect light source and flare regions separately. These methods achieve promising results for small flares but struggle with strong light sources and large patterns.

With synthetic and real datasets available, supervised end-to-end methods now dominate. Wu et al. Wu et al. [2021] and Dai et al. Dai et al. [2022] used standard restoration networks Ronneberger et al. [2015]; Chen et al. [2021]; Zamir et al. [2022]; Wang et al. [2022] to predict flare-free images directly, treating the task as image-to-image translation. Dai et al. Dai et al. [2023a] proposed an improved pipeline that separately estimates the flare component and restores the image, allowing the network to preserve light sources while better localizing flare. However, this approach burdens the network with simultaneous flare analysis and image generation, still producing artifacts especially when strong streaks are present.

Coarse-to-fine strategies have proven effective in general image restoration Zamir et al. [2021]; Chen et al. [2021]; Li et al. [2023]; Nehete et al. [2024], typically decomposing restoration into coarse correction followed by detail refinement. However, as we demonstrate in this work, directly applying this paradigm to flare removal yields marginal improvements because it lacks explicit guidance for identifying and restoring corrupted regions. Our VPRP addresses this limitation by introducing mask-guided inpainting and visual prompting, enabling effective coarse-to-fine processing specifically designed for flare removal characteristics.

3 METHODOLOGY

We first introduce the Flare7k++ pipeline to establish the baseline, then present our Visual Prompt Restoration Pipeline (VPRP) and its core component, the Prompt Restoration Network (PRN). Figure 2 (a) and (b) illustrates the comparison between the two pipelines.

3.1 FLARE7K++ PIPELINE

The Flare7k++ pipeline Dai et al. [2023a], shown in Figure 2(a), employs a U-net backbone to simultaneously predict the flare-free image and the flare component. This design allows the network to preserve light sources while localizing flare regions. However, the approach faces a fundamental challenge: when flare streaks completely occlude underlying content due to sensor saturation, the network lacks sufficient information to reconstruct the blocked regions. The requirement to simultaneously extract flare and generate flare-free images places a heavy burden on the network, leading to artifacts in the restored output. This observation motivates our two-stage approach that explicitly addresses the information gap in flare-corrupted regions.

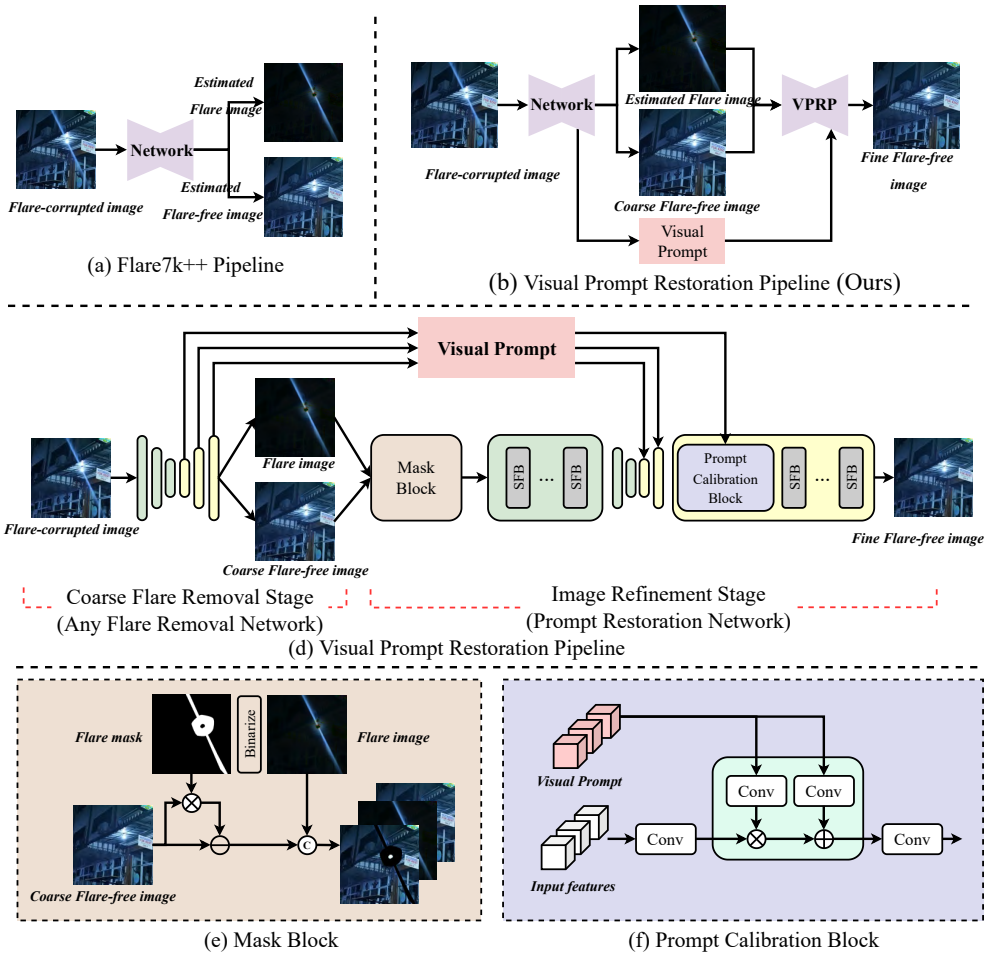


Figure 2: Overview of the VPRP pipeline. The coarse flare removal stage estimates the flare-free image and flare component. The image refinement stage employs PRN to restore flare-corrupted regions under the guidance of visual prompts from the coarse stage.

3.2 Visual Prompt Restoration Pipeline

Our VPRP pipeline reconceptualizes flare removal as a guided inpainting problem. Rather than directly translating flare-corrupted images to flare-free outputs, we identify which regions are corrupted and restore them using contextual information from non-corrupted areas. This approach is motivated by the observation that flare regions often completely block underlying content, making direct restoration unreliable without explicit guidance. Figure 2 presents the VPRP pipeline overview. The pipeline consists of two stages: a coarse flare removal stage that estimates the flare-free image and flare component, followed by an image refinement stage that uses the Prompt Restoration Network (PRN) to restore corrupted regions. VPRP is model-agnostic that any U-shape restoration network can serve as the coarse stage backbone.

3.2.1 Coarse Flare Removal Stage

The coarse stage employs a U-shape network to estimate the flare-free image I_{FF}^{Coarse} and the flare component F_{DF} from the input flare-corrupted image. We adopt the same training objective as Flare7k++ Dai et al. [2023a]:

$$\mathscr{L} = \frac{1}{N} \begin{cases} \sum_{i=1}^{N} |I_{FF}^{i,Coarse} - I_{gt}^{i}| \\ \alpha \sum_{i=1}^{N} |F_{DF}^{i} - F_{gt}^{i}| \end{cases} \tag{1}$$

where $I_{FF}^{i,Coarse}$, F_{DF}^{i}, and F_{gt}^{i} represent the i^{th} pixel in the coarse flare-free image, predicted flare, and ground truth flare respectively. The coarse stage provides two essential outputs for the refinement stage: the flare estimate for mask generation, and decoder features for visual prompting.

3.2.2 IMAGE REFINEMENT STAGE

The refinement stage addresses the artifacts produced by the coarse stage through a guided inpainting approach. Rather than applying another image-to-image translation, we explicitly identify corrupted regions and restore them using contextual information. This design is motivated by the insight that standard restoration networks struggle when degradation completely occludes content, a common scenario in severe flare cases.

Mask Block: The mask block generates the input for the refinement network by explicitly excluding flare-corrupted pixels. Given the predicted flare F_{DF}, we first binarize it with a threshold to obtain a binary mask:

$$\begin{aligned} Mask = \text{Binarize}(F_{DF} > threshold), I_{FF}^{mask} = I_{FF}^{Coarse} \times Mask, \\ I_{FF}^{mask} = I_{FF}^{Coarse} - I_{FF}^{mask}, I_{FF}^{Input} = \text{Concat}(I_{FF}^{mask}, F_{DF}, I_{FF}^{Coarse}). \end{aligned} \tag{2}$$

We use the coarse flare-free image rather than the original corrupted image because the estimated mask may omit low-intensity flare, and using the corrupted image would leave residual flare in the output. The mask block enables the refinement network to focus specifically on corrupted regions while preserving non-affected areas.

U-shape Structure: The refinement network employs a U-net architecture with SFB blocks Zhang et al. [2023a; 2022] at each encoder level to extract multi-scale features from I_{FF}^{Input}. The decoder uses these features along with visual prompts from the coarse stage to restore the masked regions. This design leverages the strong receptive field of SFB blocks for effective context aggregation.

Prompt Calibration Block: The prompt calibration block conditions the refinement network using decoder features from the coarse stage as visual prompts. As shown in Figure 2(e), the block operates at each decoder level:

$$X^i = \text{Conv}(X_{FF}^{i,C}) \times \text{Conv}(X^i) + \text{Conv}(X_{FF}^{i,C}) \tag{3}$$

where X^i is the feature at the i^{th} decoder level and $X_{FF}^{i,C}$ is the corresponding coarse-stage feature. This mechanism enables the refinement network to leverage structural and semantic information extracted during coarse restoration, ensuring consistency between stages.

The refinement stage is trained with L1 loss and perceptual loss:

$$\mathscr{L} = \begin{cases} \frac{1}{N} \sum_{i=1}^{N} |I_{FF}^{i,fine} - I_{gt}^{i}| \\ \mathscr{L}_{per}(I_{FF}^{i,fine}, I_{gt}^{i}) \end{cases} \tag{4}$$

where $I_{FF}^{i,fine}$ is the refined output and $\mathscr{L}_{per}$ denotes perceptual loss.

4 EXPERIMENTS

4.1 EXPERIMENTAL SETUP

Datasets. For the training set, we adopt Flare7K++ Dai et al. [2023a] datasets. Notably, Flare7K and Flare7K++ provide a quantity of flare images, while giving no flare-free images in the training set. We follow the experimental setting in Flare7K which adopts the Flickr24k dataset Zhang et al. [2018] as flare-free images. To form the data pair for supervised training, we add compound flare images on top of flare-free images to generate nighttime flare-corrupted images. As for flare-free images, we add light source annotations on flare-free images. We use horizontal and vertical flips and channel shuffle to enlarge the training set. And other configurations are consistent with Flare7K++ Dai et al. [2023a]. For the testing set, we adopt the real testing set provided by Flare7K++, as the Flare7K++ real set mainly contains scattering and reflective flare. We conduct VPRP pipeline on them for ablation experiments and comparison experiments.

Table 1: Comparison of flare removal methods on Flare7k++ real dataset. Best results are in **bold**. Our VPRP pipeline consistently improves performance across all backbone architectures.

Model	PSNR ↑	SSIM ↑	LPIPS ↓	G-PSNR ↑	S-PSNR ↑
Zhang et al. Zhang et al. [2020]	21.022	0.784	0.1738	19.868	13.062
Sharma et al. Sharma & Tan [2021]	20.492	0.826	0.1115	17.790	12.648
Wu et al. Wu et al. [2021]	24.613	0.871	0.0598	21.772	16.728
UNet Ronneberger et al. [2015]	27.189	0.894	0.0452	23.527	22.647
HINet Chen et al. [2021]	27.548	0.892	0.0464	24.081	22.907
MPRNet* Zamir et al. [2021]	27.036	0.893	0.0481	23.490	22.267
Restormer* Zamir et al. [2022]	27.597	0.897	0.0447	23.828	22.452
Uformer Wang et al. [2022]	27.633	0.894	0.0428	23.949	22.603
FF-Former Zhang et al. [2023a]	27.900	0.900	0.0410	24.136	23.054
MQANet Li et al. [2025]	26.600	0.890	0.0510	-	-
SD-Inpainting Rombach et al. [2022]	28.020	0.894	0.0431	23.810	21.710
SDXL-Inpainting Podell et al. [2023]	26.990	0.891	0.0453	22.410	20.120
LightsOut Tsai et al. [2025]	28.150	0.896	0.0409	24.200	22.240
Difflare Zhou et al. [2024]	26.063	0.898	-	-	-
MFDNet Jiang et al. [2024]	26.980	0.895	0.0510	-	-
SFNet Vasluianu et al. [2024]	27.082	0.896	0.0460	23.640	22.013
PromptIR Potlapalli et al. [2023]	26.843	0.893	0.0470	23.280	21.545
Qi et al. Qi et al. [2025]	28.033	0.903	0.0420	24.537	23.614
Sparse-UFormer Wu et al. [2024]	27.976	0.906	0.0413	24.243	23.529
Kotp et al. Kotp & Torki [2024]	27.662	0.897	0.0422	23.987	22.847
SGSFT Ma et al. [2025]	28.078	0.904	0.0416	24.477	23.306
VPRP Pipeline (Ours)					
UNet + VPRP	28.109	0.901	0.0412	24.592	23.539
Uformer + VPRP	28.241	0.903	0.0404	24.633	23.777
HINet + VPRP	28.298	0.903	0.0406	**24.964**	**24.203**
Restormer* + VPRP	28.089	0.902	0.0427	24.768	23.432
FF-Former + VPRP	**28.443**	**0.906**	**0.0392**	24.837	24.111

Implementation Details. We set 4 levels in the PRN network. The channel of the immediate layer is 16. Since SFB Zhang et al. [2023a] has a wider receptive field, we use it as main feature extraction module for PRN. The optimizer for the VPRP pipeline adopts the same setting as the Flare7K++ Dai et al. [2023a] and FF-Former Zhang et al. [2023a].

4.2 Comparison with State-of-the-Art Methods

In this section, we conduct comparison results on the real testing set of the Flare7K++ dataset to demonstrate the functions of our VPRP pipeline. We employ flare removal methods Wu et al. Wu et al. [2021] and FF-Former Zhang et al. [2023a] and image restoration methods Unet Ronneberger et al. [2015], HINet Chen et al. [2021], MPRNet Zamir et al. [2021], Restormer Zamir et al. [2022], and Uformer Wang et al. [2022] for baselines.

4.2.1 Results on Flare7k++ datasets

To validate the performance of our method on complex scenarios which include both the scattering flare and reflective flare, we show the quantitative results (*e.g.*, PSNR, SSIM, LPIPS) on Table 1 for experiments on Flare7k++ real testing set, correspondingly. Furthermore, we also represent the visual comparison results on Figure 3 for experiments on the Flare7k++ real testing set (we employ Uformer Wang et al. [2022] in coarse flare removal network, same as Flare7k++).

From the observation of Table 1, our VPRP equipped with FF-Former Zhang et al. [2023a] has achieved state-of-the-art PSNR and SSIM. Concretely, from the perspective of pixel level, we significantly outperform the state-of-the-art method FF-Former with 0.543dB on PSNR score, 0.006 on SSIM, 0.0018 on LPIPS, 0.701dB on G-PSNR score, and 1.057dB on S-PSNR score. As for visual performance comparison, as shown in Figure 3 and Figure 4, we figure that methods trained in a supervised manner achieve better results and our VPRP equipped with Uformer has essentially eliminated the flare and introduced the least artifacts in most situation. Despite the shape of the flare, round flare with a large radius, flare with a long streak, or the type of the flare, reflective flare,

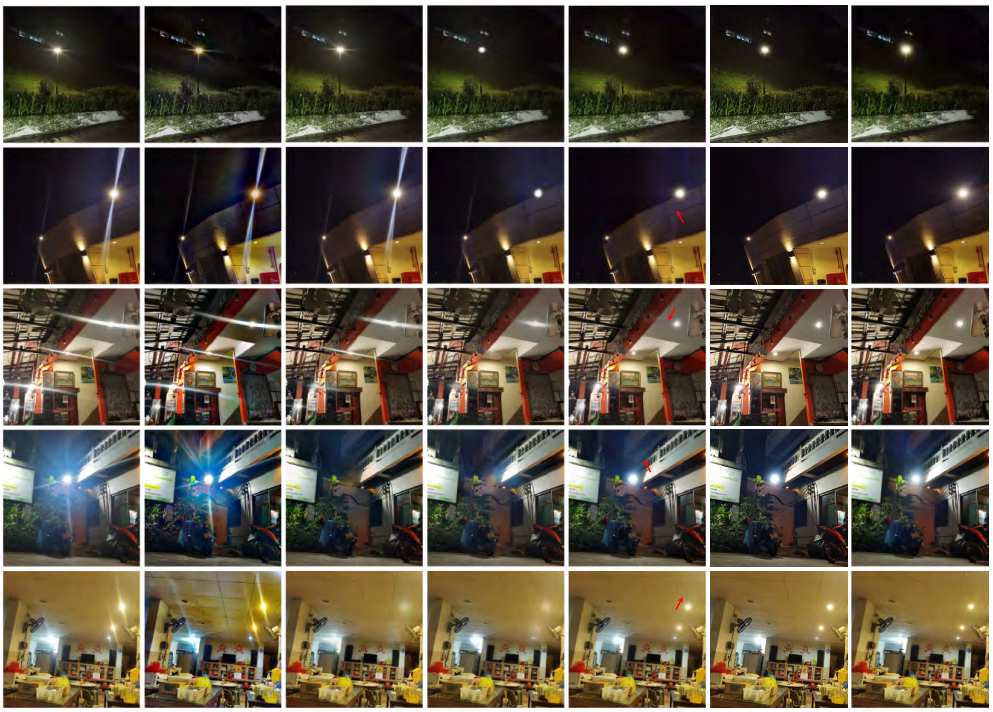


Figure 3: Visual comparison on real-world nighttime flare images. Flare7K indicates training with the pipeline from Wu et al. Wu et al. [2021]; Flare7K++ indicates training with the pipeline from Dai et al. Dai et al. [2023a]. Our method eliminates flare more completely while producing fewer artifacts.

or scattering flare, our method can achieve more realistic and natural results than other baselines in most scenarios.

4.3 Ablation Studies

4.3.1 Impact of PRN Components

Table 2: Ablation: impact of mask block and prompt calibration block (PCB).

Method	PRN	C	Mask	PCB	PSNR
Uformer	✗	-	✗	✗	27.63
	✓	16	✗	✗	27.65
	✓	16	✗	✓	28.06
	✓	16	✓	✓	28.24

A systematic evaluation of the impact of PRN's key components on flare removal is presented in Table 2. For fair comparison, we employ Uformer (first row) as the base network in the coarse flare removal stage and test on the flare7k++ real set. In the second row, a simple U-shape refinement network is added to the Uformer, while achieving marginal improvement. It indicates that using a simple U-shape network as a refinement network does not yield significant performance gains. Comparing the second row and the third row, we can see that adding the Prompt Calibration Block (PCB) significantly boosts the PSNR from 27.65 to 28.06. This improvement demonstrates that the visual prompt is non-substitutable. The fourth row, with all components of PRN, indicates that using high-quality features to rewrite regions contaminated with flare can achieve better performance.

4.3.2 Model Agnostic Property

To demonstrate that our VPRP is model-agnostic and can be applied with any other flare removal network, we conduct quantitative and qualitative ablation experiments on Unet Ronneberger et al. [2015], Uformer Wang et al. [2022], and FF-Former Zhang et al. [2023a]. The network in the coarse flare removal stage is pretrained for predicting the coarse flare-free image and the flare image on Flare7K++ Dai et al. [2023a] dataset.

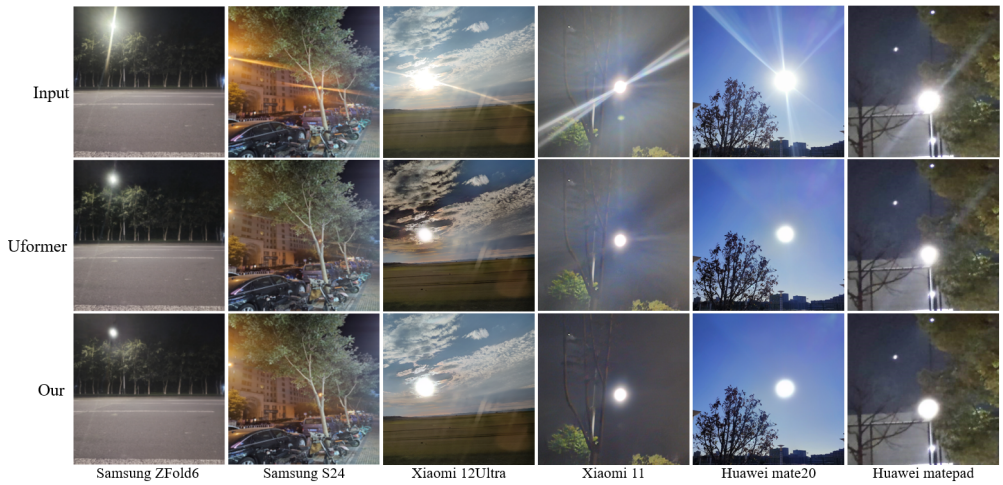


Figure 4: Visual comparison on real nighttime and daytime flare-corrupted images (without ground truth).

For quantitative comparison, based on Table 1 and Table 3, we can observe that with VPRP pipeline, Unet (the channel number of VPRP is 16) Ronneberger et al. [2015] gains 0.92dB, 0.007, 0.004 on PSNR, SSIM, and LPIPS scores which surpass the state-of-the-art method FF-Former with 0.209dB on PSNR score. Moreover, our VPRP pipeline helps Uformer with 0.608dB, 0.009, 0.0024, 0.684dB, and 1.174dB gains on PSNR, SSIM, LPIPS, G-PSNR and S-PSNR scores. As for the state-of-the-art method FF-Former Zhang et al. [2023a], equipped with our pipeline, it manages to obtain new state-of-the-art results, which achieve 28.443dB, 0.906, 0.0392, 24.837dB and 24.111dB on PSNR, SSIM, LPIPS, G-PSNR and S-PSNR scores and our pipeline provides 0.543dB, 0.006, 0.0018, 0.701dB and 1.057dB promotions with a very small increase in the number of parameters, as shown in Table 3.

Table 3: Efficiency analysis: parameters, FLOPs, and inference time (512×512, NVIDIA 3090).

Method	PRN	C	Param	FLOPs	Time	PSNR
UNet	✗	-	9.0M	62.4G	0.0115s	27.19
	✓	4	9.1M	65.7G	0.0126s	27.53
	✓	8	10.2M	70.5G	0.0126s	27.73
	✓	16	11.8M	84.9G	0.0127s	28.11
Uformer	✗	-	20.4M	162.1G	0.0723s	27.63
	✓	4	20.9M	167.8G	0.0817s	27.91
	✓	8	21.5M	175.1G	0.0818s	28.05
	✓	16	23.6M	194.3G	0.0820s	28.24
FF-Former	✗	-	46.5M	407.8G	0.2109s	27.90
	✓	4	47.0M	413.6G	0.2146s	28.26
	✓	8	47.6M	420.8G	0.2175s	28.35
	✓	16	49.6M	440.0G	0.2185s	28.44

As for qualitative comparison, we show the visual comparisons on Figure 5. Compared with the original Unet, Unet + VPRP pipeline removes more shimmer within the flare-corrupted images. As for Uformer, more streaks have been eliminated. Previous pipelines hardly identify and eliminate all the flares added to the image and tend to generate severe artifacts when handling strong flares with large overexposed streaks and shimmers. Our VPRP pipeline suppresses this phenomenon by rewriting the image details polluted by flare with the extracted semantic information within the non-polluted area and the multi-scale features obtained in the coarse flare removal stage. The experiments demonstrate that our pipeline is model-agnostic.

4.3.3 Channel Number Analysis

Integrating the Visual Prompt Restoration Pipeline (VPRP) significantly enhances nighttime flare removal across various neural network architectures (Table 3 and Figures 5 and 8). VPRP achieves this by utilizing multi-scale features from the initial coarse flare removal stage to refine flare-corrupted details. Increasing channel numbers stably enhances both objective PSNR and subjective visual quality. This underscores the effectiveness of VPRP in leveraging deeper networks for image restoration in challenging, flare-prone low-light conditions.

Furthermore, Figure 8 highlights the computational efficiency of PRN, showing inference times remain nearly unchanged despite consistent PSNR gains. As reported in Table 3, Uformer inference

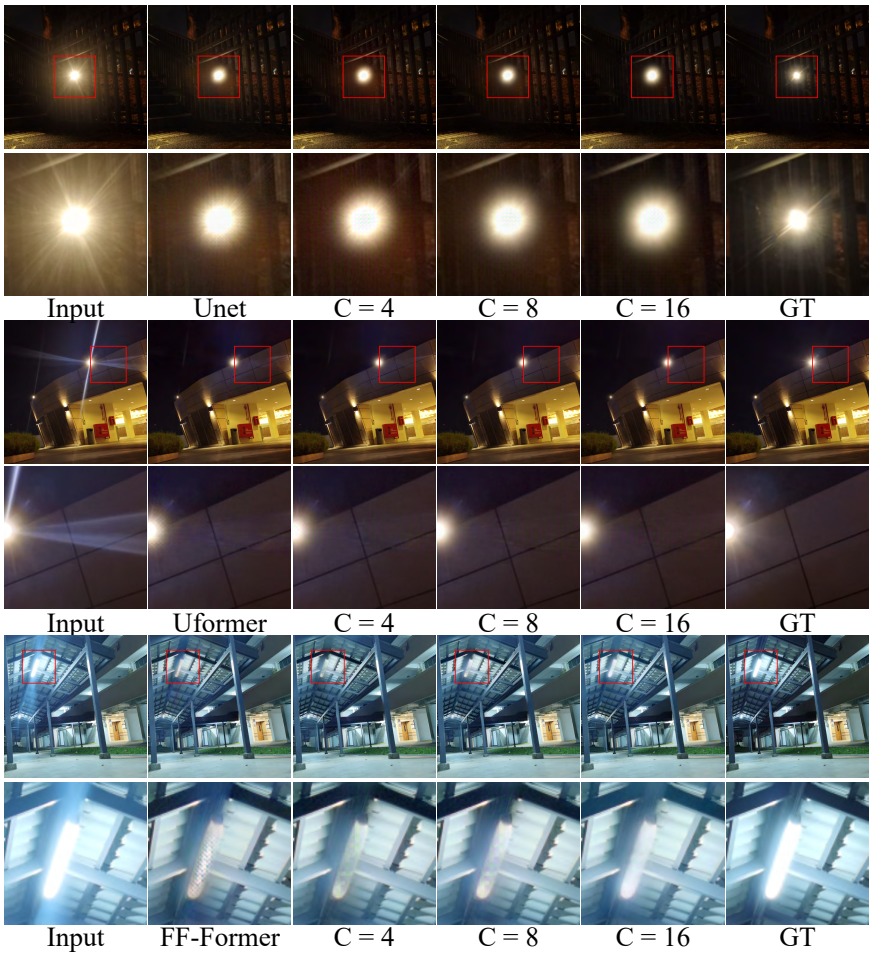


Figure 5: Ablation: visual comparison with and without VPRP. VPRP suppresses artifacts more effectively, especially with weaker networks.

time increases minimally from 0.0723 s to 0.0820 s even at $C = 16$. UNet and FF-Former exhibit similar patterns, where modestly growing parameters and FLOPs minimally impact runtime. Thus, PRN is computationally lightweight, yielding measurable improvements with negligible overhead. This property is crucial for practical deployments in real-time or resource-constrained environments demanding both efficiency and accuracy.

5 Conclusion

This paper addresses the artifact problem in flare removal by proposing the Visual Prompt Restoration Pipeline (VPRP). The key insight is that conventional coarse-to-fine methods fail because they lack explicit guidance for identifying and restoring flare-corrupted regions. VPRP addresses this through two mechanisms: a mask block that explicitly excludes corrupted pixels, and a prompt calibration block that leverages coarse-stage features as visual prompts for guided restoration. The Prompt Restoration Network (PRN) is model-agnostic, integrating seamlessly with any U-shape restoration network while introducing minimal computational overhead. Experiments demonstrate that VPRP achieves state-of-the-art performance across multiple benchmarks, with consistent improvements on diverse backbone architectures. A current limitation is that PRN adopts generic restoration modules without explicit modeling of flare-specific physical properties. Future work could explore specialized architectures that incorporate the spatial distribution and spectral characteristics of different flare types for further performance gains.

REFERENCES

CS Asha, Sooraj Kumar Bhat, Deepa Nayak, and Chaithra Bhat. Auto removal of bright spot from images captured against flashing light source. In *2019 IEEE International Conference on Distributed Computing, VLSI, Electrical Circuits and Robotics (DISCOVER)*, pp. 1–6. IEEE, 2019.

Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.

Liangyu Chen, Xin Lu, Jie Zhang, Xiaojie Chu, and Chengpeng Chen. Hinet: Half instance normalization network for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 182–192, 2021.

Yuekun Dai, Chongyi Li, Shangchen Zhou, Ruicheng Feng, and Chen Change Loy. Flare7k: A phenomenological nighttime flare removal dataset. *Advances in Neural Information Processing Systems*, 35:3926–3937, 2022.

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*, 2023a.

Yuekun Dai, Yihang Luo, Shangchen Zhou, Chongyi Li, and Chen Change Loy. Nighttime smartphone reflective flare removal using optical center symmetry prior. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 20783–20791, 2023b.

Yuekun Dai, Dafeng Zhang, Xiaoming Li, Zongsheng Yue, Chongyi Li, Shangchen Zhou, Ruicheng Feng, Peiqing Yang, Zhezhu Jin, Guanqun Liu, and Chen Change Loy. Mipi 2024 challenge on nighttime flare removal: Methods and results. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 1144–1152, June 2024.

Qiaole Dong, Chenjie Cao, and Yanwei Fu. Incremental transformer structure enhanced image inpainting with masking positional encoding. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11358–11368, 2022.

Zheng Dong, Ke Xu, Yin Yang, Hujun Bao, Weiwei Xu, and Rynson WH Lau. Location-aware single image reflection removal. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 5017–5026, 2021.

Ruicheng Feng, Chongyi Li, Huaijin Chen, Shuai Li, Jinwei Gu, and Chen Change Loy. Generating aligned pseudo-supervision from non-aligned data for image restoration in under-display camera. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 5013–5022, 2023.

Kaiming He, Jian Sun, and Xiaoou Tang. Single image haze removal using dark channel prior. *IEEE transactions on pattern analysis and machine intelligence*, 33(12):2341–2353, 2010.

Yuwen He, Wei Wang, Wanyu Wu, and Kui Jiang. Disentangle nighttime lens flares: Self-supervised generation-based lens flare removal. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 39, pp. 3464–3472, 2025.

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*, pp. 1–14, 2024.

Soomin Kim, Yuchi Huo, and Sung-Eui Yoon. Single image reflection removal with physically-based training images. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5164–5173, 2020.

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)*, pp. 2565–2569. IEEE, 2024.

Yu Li, Ming Liu, Yaling Yi, Qince Li, Dongwei Ren, and Wangmeng Zuo. Two-stage single image reflection removal with reflection-aware guidance. *Applied Intelligence*, 53(16):19433–19448, 2023.

Zihao Li, Junming Feng, Siyao Hao, Yuze Wang, and Weibang Bai. Mask-q attention network for flare removal. *Neurocomputing*, 637:130100, 2025.

Tianlei Ma, Zhiqiang Kai, Xikui Miao, Jing Liang, Jinzhu Peng, Yaonan Wang, Hao Wang, and Xinhao Liu. Self-prior guided spatial and fourier transformer for nighttime flare removal. *IEEE Transactions on Automation Science and Engineering*, 2025.

Hemkant Nehete, Amit Monga, Partha Kaushik, and Brajesh Kumar Kaushik. Fourier prior-based two-stage architecture for image restoration. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 6014–6023, 2024.

Dustin Podell, Zion English, Kyle Lacey, Andreas Blattmann, Tim Dockhorn, Jonas Müller, Joe Penna, and Robin Rombach. Sdxl: Improving latent diffusion models for high-resolution image synthesis. *arXiv preprint arXiv:2307.01952*, 2023.

Vaishnav Potlapalli, Syed Waqas Zamir, Salman H Khan, and Fahad Shahbaz Khan. Promptir: Prompting for all-in-one image restoration. *Advances in Neural Information Processing Systems*, 36:71275–71293, 2023.

Kejing Qi, Bo Wang, and Yun Liu. A self-prompt based dual-domain network for nighttime flare removal. *Engineering Applications of Artificial Intelligence*, 144:110103, 2025.

Xiaotian Qiao, Gerhard P Hancke, and Rynson WH Lau. Light source guided single-image flare removal from unpaired data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 4177–4185, 2021.

Lishen Qu, Zhihao Liu, Jinshan Pan, Shihao Zhou, Jinglei Shi, Duosheng Chen, and Jufeng Yang. Flarex: A physics-informed dataset for lens flare removal via 2d synthesis and 3d rendering. *arXiv preprint arXiv:2510.09995*, 2025.

Robin Rombach, Andreas Blattmann, Dominik Lorenz, Patrick Esser, and Björn Ommer. High-resolution image synthesis with latent diffusion models. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 10684–10695, 2022.

Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18*, pp. 234–241. Springer, 2015.

Aashish Sharma and Robby T Tan. Nighttime visibility enhancement by increasing the dynamic range and suppression of light effects. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 11977–11986, 2021.

Qilin Sun, Ethan Tseng, Qiang Fu, Wolfgang Heidrich, and Felix Heide. Learning rank-1 diffractive optics for single-shot high dynamic range imaging. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 1386–1396, 2020.

Roman Suvorov, Elizaveta Logacheva, Anton Mashikhin, Anastasia Remizova, Arsenii Ashukha, Aleksei Silvestrov, Naejin Kong, Harshith Goka, Kiwoong Park, and Victor Lempitsky. Resolution-robust large mask inpainting with fourier convolutions. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 2149–2159, 2022.

Shr-Ruei Tsai, Wei-Cheng Chang, Jie-Ying Lee, Chih-Hai Su, and Yu-Lun Liu. Lightsout: Diffusion-based outpainting for enhanced lens flare removal. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 6353–6363, 2025.

Florin Vasluianu, Zongwei Wu, and Radu Timofte. Sfnet-a spatial-frequency domain neural network for image lens flare removal. In *2024 IEEE International Conference on Image Processing (ICIP)*, pp. 1711–1717. IEEE, 2024.

Patricia Vitoria and Coloma Ballester. Automatic flare spot artifact detection and removal in photographs. *Journal of Mathematical Imaging and Vision*, 61(4):515–533, 2019.

Zhendong Wang, Xiaodong Cun, Jianmin Bao, Wengang Zhou, Jianzhuang Liu, and Houqiang Li. Uformer: A general u-shaped transformer for image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 17683–17693, 2022.

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.

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 *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 2239–2247, 2021.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Multi-stage progressive image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 14821–14831, 2021.

Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, and Ming-Hsuan Yang. Restormer: Efficient transformer for high-resolution image restoration. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 5728–5739, 2022.

Dafeng Zhang, Feiyu Huang, Shizhuo Liu, Xiaobing Wang, and Zhezhu Jin. Swinfir: Revisiting the swinir with fast fourier convolution and improved training for image super-resolution. *arXiv preprint arXiv:2208.11247*, 2022.

Dafeng Zhang, Jia Ouyang, Guanqun Liu, Xiaobing Wang, Xiangyu Kong, and Zhezhu Jin. Ff-former: Swin fourier transformer for nighttime flare removal. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2823–2831, 2023a.

Jing Zhang, Yang Cao, Zheng-Jun Zha, and Dacheng Tao. Nighttime dehazing with a synthetic benchmark. In *Proceedings of the 28th ACM international conference on multimedia*, pp. 2355–2363, 2020.

Lvmin Zhang, Anyi Rao, and Maneesh Agrawala. Adding conditional control to text-to-image diffusion models. In *Proceedings of the IEEE/CVF international conference on computer vision*, pp. 3836–3847, 2023b.

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*, pp. 4786–4794, 2018.

Tianwen Zhou, Qihao Duan, and Zitong Yu. Difflare: Removing image lens flare with latent diffusion model. *arXiv preprint arXiv:2407.14746*, 2024.

Jun-Yan Zhu, Taesung Park, Phillip Isola, and Alexei A Efros. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proceedings of the IEEE international conference on computer vision*, pp. 2223–2232, 2017.

A Additional Ablation Study

A.0.1 Training Process

The process of our work involves two main steps. Step1: Train flare removal network with Flare7k++ pipline (or download pretrained model like Uformer in Flare7K's github repo). Step2: Frozen flare removal network and train PRN (our training configuration is consistent with Flare7k++ and does not require special design). FF-Former has achieved SOTA performance and SFB Zhang et al. [2023a] has a large receptive field, so we use it as our core module of PRN.

A.0.2 Threshold Selection

To determine the optimal binarization threshold for the Mask Block, we conducted a sensitivity analysis by evaluating the Intersection over Union (IoU) between the predicted flare masks and the ground truth masks on the test set. As illustrated by the ground truth mask in Figure 7, our evaluation specifically targets the flare regions (highlighted in red and yellow) rather than the entire image, ensuring the metric accurately reflects the model's capability to localize flare artifacts. Figure 6 quantitatively presents the Average IoU across varying threshold values. The curve exhibits a distinct peak around 0.1, indicating the optimal operating point. Consequently, we adopt a threshold value around 0.1 for the VPRP pipeline to strike the best balance between effective artifact suppression and the preservation of non-polluted semantic details.

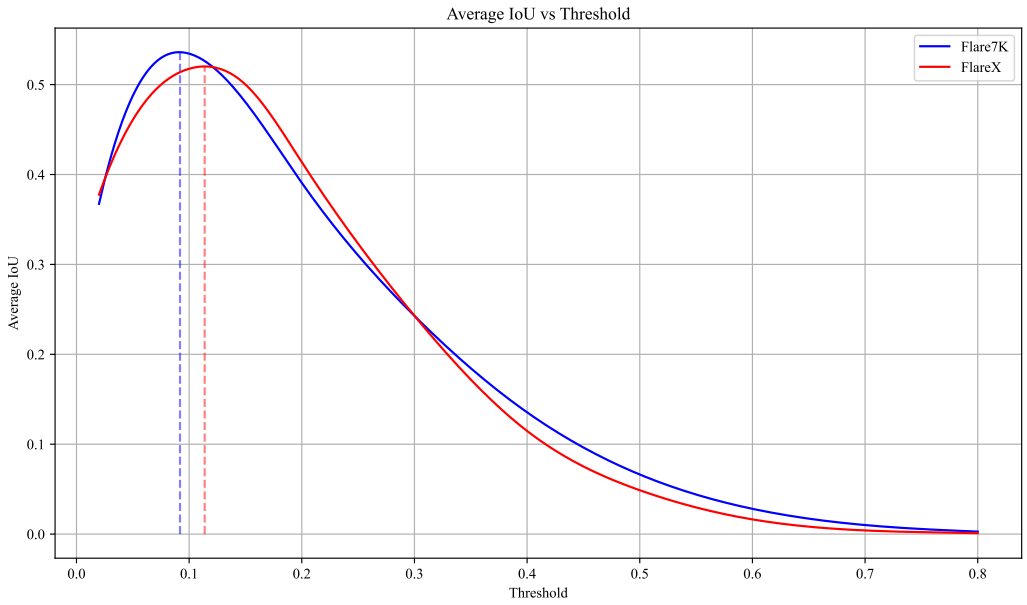


Figure 6: Average IoU vs Threshold (th).

A.0.3 Effectiveness and Efficiency Analysis of VPRP

Based on the results in Figure 8 and the quantitative evidence in Table 3, several insights can be drawn regarding the effectiveness and efficiency of the proposed Prompt Restoration Network (PRN). First, the left plot of Figure 8 clearly illustrates a consistent upward trend in PSNR as the number of channels C in PRN increases. When $C = 0$, i.e., without employing PRN, the baseline models exhibit comparatively lower PSNR. As channels are gradually increased (e.g., $C = 4, 8, 16$), PSNR improves steadily across different backbones (Unet, Uformer, FF-Former). This trend demonstrates that the inclusion of PRN channels enhances the model's capacity to refine flare-corrupted regions by leveraging multi-scale contextual information from the coarse stage. Importantly, the gains are not marginal: for instance, Table 3 shows that Uformer improves from 27.63 dB (without PRN) to 28.24 dB when PRN with 16 channels is adopted, corresponding to a substantial improvement of 0.61 dB. Similar improvements are observed for Unet (+0.92 dB) and FF-Former (+0.54 dB). These results collectively validate that the visual prompts provided by PRN play a decisive role

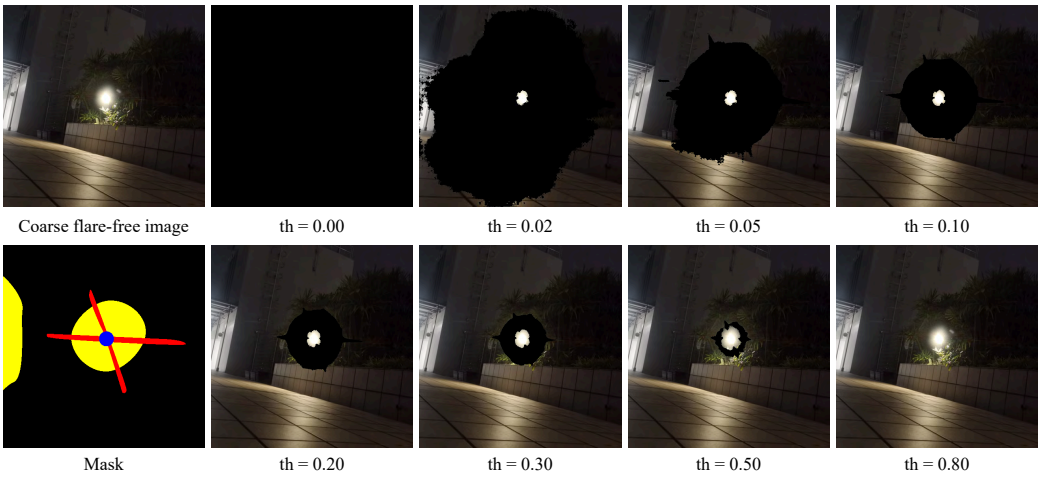


Figure 7: Mask vs Threshold (th).

in guiding the refinement process and reducing artifacts that simple coarse-to-fine pipelines fail to address.

Equally important is the observation from the right plot of Figure 8, which highlights the computational efficiency of PRN. Despite consistent increases in PSNR, the inference time remains nearly unchanged when PRN is integrated. For example, the inference time of Uformer increases only minimally from 0.0723s to 0.0820s even at $C = 16$, as reported in Table 3. A similar pattern is found with Unet and FF-Former, where additional parameters and FLOPs grow modestly but without a significant impact on runtime. This indicates that PRN is computationally lightweight, introducing only a negligible overhead while yielding measurable performance improvements. Such a property is crucial for practical deployment, especially in real-time or resource-constrained environments, where both efficiency and accuracy are indispensable.

Figure 8 and Table 3 jointly demonstrate that PRN effectively balances performance enhancement and computational efficiency. The results confirm that PRN's visual prompt mechanism not only scales with the number of channels to deliver progressive improvements in PSNR but also maintains lightweight inference costs. This dual advantage underscores the robustness and practicality of the proposed approach, affirming its value as a plug-and-play module for flare removal tasks across different backbone architectures.

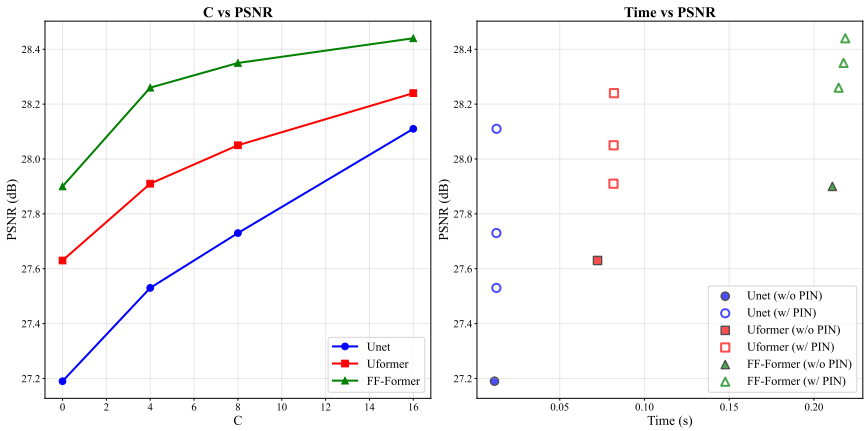


Figure 8: Relationship between the number of PRN channels and PSNR (left), and inference time versus PSNR (right), demonstrating both performance gains and lightweight efficiency of PRN.



Figure 9: Visual comparison of BraceKet flare dataset. Zoom in for a better view. We mark the area around the reflective flare for a better look.

Table 4: Comparison of flare removal methods on BracKetFlare dataset. The highest score has been shown on **bold**.

Method	BracKetflare		
	PSNR ↑	SSIM ↑	LPIPS ↓
Wu et al. Wu et al. [2021]	26.13	0.895	0.055
Unet Ronneberger et al. [2015]	47.00	0.897	0.006
HINet Chen et al. [2021]	48.03	0.994	0.003
MPRNet* Zamir et al. [2021]	48.41	0.994	0.004
Restormer* Zamir et al. [2022]	48.11	0.994	0.004
Uformer Wang et al. [2022]	47.47	0.991	0.003
FF-Former Zhang et al. [2023a]	47.98	0.992	0.003
Zhang et al.Zhang et al. [2018]	-	0.830	0.074
Dong et al.Dong et al. [2021]	-	0.907	0.041
Soomin Kim et al.Kim et al. [2020]	-	0.857	0.069
He et al.He et al. [2025]	-	0.967	0.038
Unet + VPRP (Ours)	48.50	0.992	0.004
Uformer + VPRP (Ours)	48.69	0.994	0.002
FF-former + VPRP (Ours)	**49.18**	**0.994**	**0.002**

A.0.4 Results on BracketFlare datasets

To further validate the robustness of our VPRP framework, we conduct experiments on the Bracket-Flare dataset Dai et al. [2023b], which specifically focuses on reflective flare, a particularly challenging scenario due to its symmetry around the optical center and high intensity. Unlike scattering flare, reflective flare often overlaps with important scene structures, making its removal difficult without introducing structural distortions or residual artifacts. As reported in Table 4, our VPRP-equipped

Table 5: Quantitative comparison of different flare removal methods trained on FlareX Qu et al. [2025] and FlareReal600 Dai et al. [2024] datasets and evaluated on Flare7K, FlareX and FlareReal600 test sets.

Training Dataset	PRN	HINet			Uformer			Restormer		
		PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓	PSNR↑	SSIM↑	LPIPS↓
Test Dataset: Flare7K										
FlareX	✗	26.658	0.894	0.052	25.259	0.887	0.057	26.196	0.889	0.055
	✓	26.978	0.901	0.050	25.824	0.889	0.055	26.537	0.900	0.052
Test Dataset: FlareX										
FlareX	✗	39.691	0.652	0.132	39.292	0.658	0.133	39.137	0.658	0.131
	✓	39.900	0.666	0.130	39.648	0.665	0.130	39.323	0.660	0.131
Test Dataset: Flare7K										
FlareReal600	✗	28.782	0.916	0.036	28.608	0.917	0.036	28.744	0.917	0.036
	✓	28.956	0.917	0.033	28.915	0.917	0.033	28.929	0.929	0.032
Test Dataset: FlareReal600										
FlareReal600	✗	23.121	0.706	0.1311	22.644	0.700	0.1316	22.860	0.703	0.1318
	✓	23.266	0.708	0.1307	22.791	0.701	0.1313	22.989	0.705	0.1312

models consistently surpass strong baselines, including MPRNet, Restormer, and FF-Former. Notably, when integrated with FF-Former, our VPRP achieves a PSNR of 49.18 dB, which represents an improvement of 0.77 dB over MPRNet and establishes a new state-of-the-art on this dataset. The SSIM and LPIPS scores further confirm this advantage, showing that our method not only restores images with higher pixel-level fidelity but also achieves better perceptual quality.

Visual results in Figure 9 corroborate these quantitative findings. Competing methods often leave faint residual flare streaks or introduce excessive smoothing in flare regions, leading to the loss of fine structures around light sources. In contrast, our approach effectively suppresses reflective flare while preserving structural details and texture consistency, producing results that are visually closer to the ground truth. The superior performance can be attributed to the prompt-guided inpainting mechanism of PRN, which leverages non-polluted regions and multi-scale semantic features from the coarse stage to reconstruct corrupted areas more faithfully.

These results highlight the generalization ability of our VPRP framework beyond the Flare7K++ dataset. By successfully handling both scattering and reflective flare, VPRP demonstrates its adaptability to diverse real-world flare phenomena. This reinforces its potential as a model-agnostic and practically deployable solution for nighttime imaging applications where various flare types coexist.

A.0.5 Results on FlareX and FlareReal600 datasets

To further substantiate the generalization capability and robustness of our proposed framework beyond the Flare7K++ benchmark, we extend our evaluation to the FlareX Qu et al. [2025] and FlareReal600 Dai et al. [2024] datasets. These experiments are designed to verify whether the Visual Prompt Restoration Pipeline (VPRP) maintains its effectiveness when trained on diverse data distributions (synthetic and real) and evaluated across different domains. For these experiments, the channel number of the Prompt Restoration Network (PRN) is set to 16. Table 5 presents a comprehensive quantitative comparison using three distinct backbone architectures: HINet, Uformer, and Restormer. The results demonstrate that integrating VPRP yields consistent performance improvements across all testing scenarios, regardless of the training dataset used. Specifically, when models are trained on the physics-informed FlareX dataset, our method significantly enhances generalization to the Flare7K test set. For instance, the Uformer backbone integrated with VPRP achieves a PSNR improvement of 0.565 dB (from 25.259 dB to 25.824 dB) and a reduction in LPIPS compared to the baseline.

Furthermore, experiments on the FlareReal600 dataset validate the efficacy of our approach in real-world scenarios. Even with the strong baseline performance of HINet on this dataset, the addition of VPRP further elevates the PSNR from 23.121 dB to 23.266 dB. This consistent enhancement across both synthetic-to-real and real-to-real settings confirms that the coarse-to-fine paradigm of VPRP, augmented by explicit flare masking and visual prompting, is not limited to specific datasets. Instead,

it serves as a robust, universal solution that effectively mitigates artifacts and bridges domain gaps where standard end-to-end baselines often struggle. It is worth noting that unlike synthetic datasets, the FlareReal600 dataset provides only flare-corrupted inputs and flare-free ground truths, lacking isolated flare images. Since our VPRP framework relies on a coarse flare mask for the refinement stage, we require an explicit estimation of the flare component during training. To address this, we estimate the flare component analytically by subtracting the clean image from the corrupted image. However, to ensure physical plausibility and avoid intensity artifacts during subtraction, we first linearize the image signals using gamma correction ($\gamma = 2.2$) before computing the residual flare, as outlined in Algorithm 1.

Algorithm 1 PyTorch-style Pseudocode for Gamma-corrected Flare Estimation

```
def adjust_gamma(image, gamma):
    # image: [B, C, H, W]; gamma: [B]
    # Broadcast gamma to match image dimensions
    gamma_tensor = gamma.view(-1, 1, 1, 1)
    out = torch.pow(image, gamma_tensor)
    out = torch.clamp(out, 0.0, 1.0)
    return out

def adjust_gamma_reverse(image, gamma):
    # image: [B, C, H, W]; gamma: [B]
    # Calculate inverse gamma for correction
    inv_gamma = 1.0 / gamma
    gamma_tensor = inv_gamma.view(-1, 1, 1, 1)
    out = torch.pow(image, gamma_tensor)
    out = torch.clamp(out, 0.0, 1.0)
    return out

def get_flare_mask(output_tensor, input_tensor, gamma):
    # output_tensor: the model-predicted flare-free image (
       coarse prediction)
    # input_tensor: the original flare-corrupted input image

    # Linearize the input and output images
    input_lin = adjust_gamma(input_tensor, gamma)
    output_lin = adjust_gamma(output_tensor, gamma)

    # Compute the flare residual in linear space to avoid
       artifacts
    diff = input_lin - output_lin

    # Clamp to ensure validity before reverse gamma
    diff = torch.clamp(diff, 1e-7, 1.0)

    # Convert back to sRGB space
    flare_pred = adjust_gamma_reverse(diff, gamma)
    return flare_pred
```

B Limitations and Future Work

In this paper, we introduce the Visual Prompt Restoration Pipeline (VPRP), which achieves state-of-the-art performance in nighttime flare removal by leveraging a model-agnostic two-stage framework. While our core contribution lies in redefining the coarse-to-fine paradigm through the integration of visual prompts and inpainting concepts, the current design of the Prompt Restoration Network (PRN) does not include explicit architectural optimizations tailored to the unique characteristics of flare patterns, such as their spatial distribution, spectral properties, or geometric structures. And the

current implementation adopts generic restoration backbones, which may not fully exploit domain-specific priors inherent to flare corruption. For future work, developing specialized neural network components that explicitly model the physical properties and visual signatures of different flare types has the potential to further enhance the pipeline's ability to handle complex flare scenarios and achieve even more refined restoration results.

C Additional Visual Results

We provide additional visual results compared to the Flare7k++ pipeline with Uformer that trained on Flare7k++ dataset.

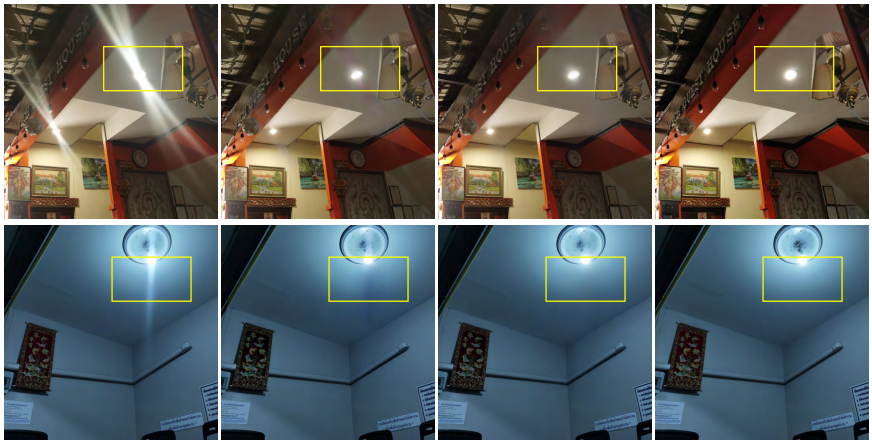


Figure 10: Visual results achieved by different pipelines on the Flare7k real dataset.

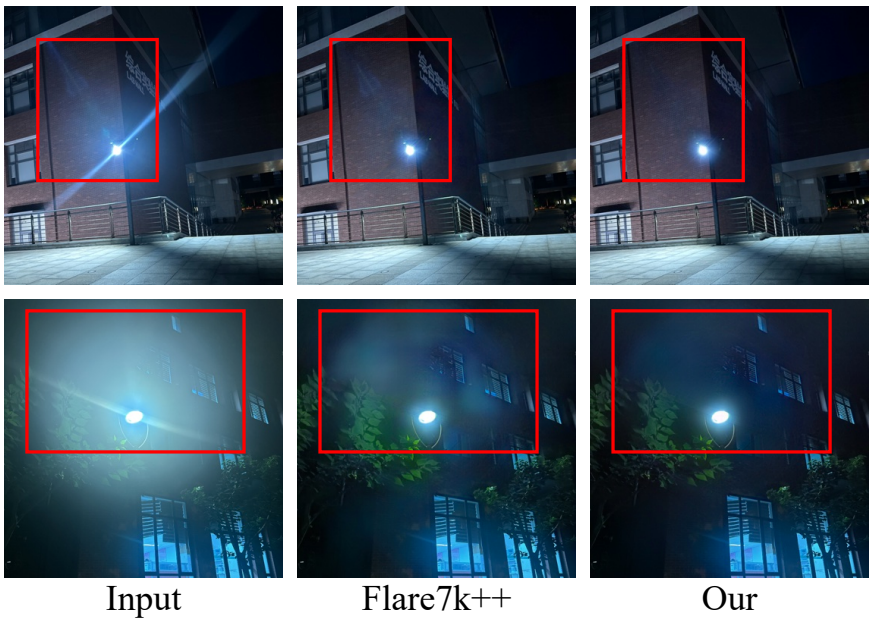
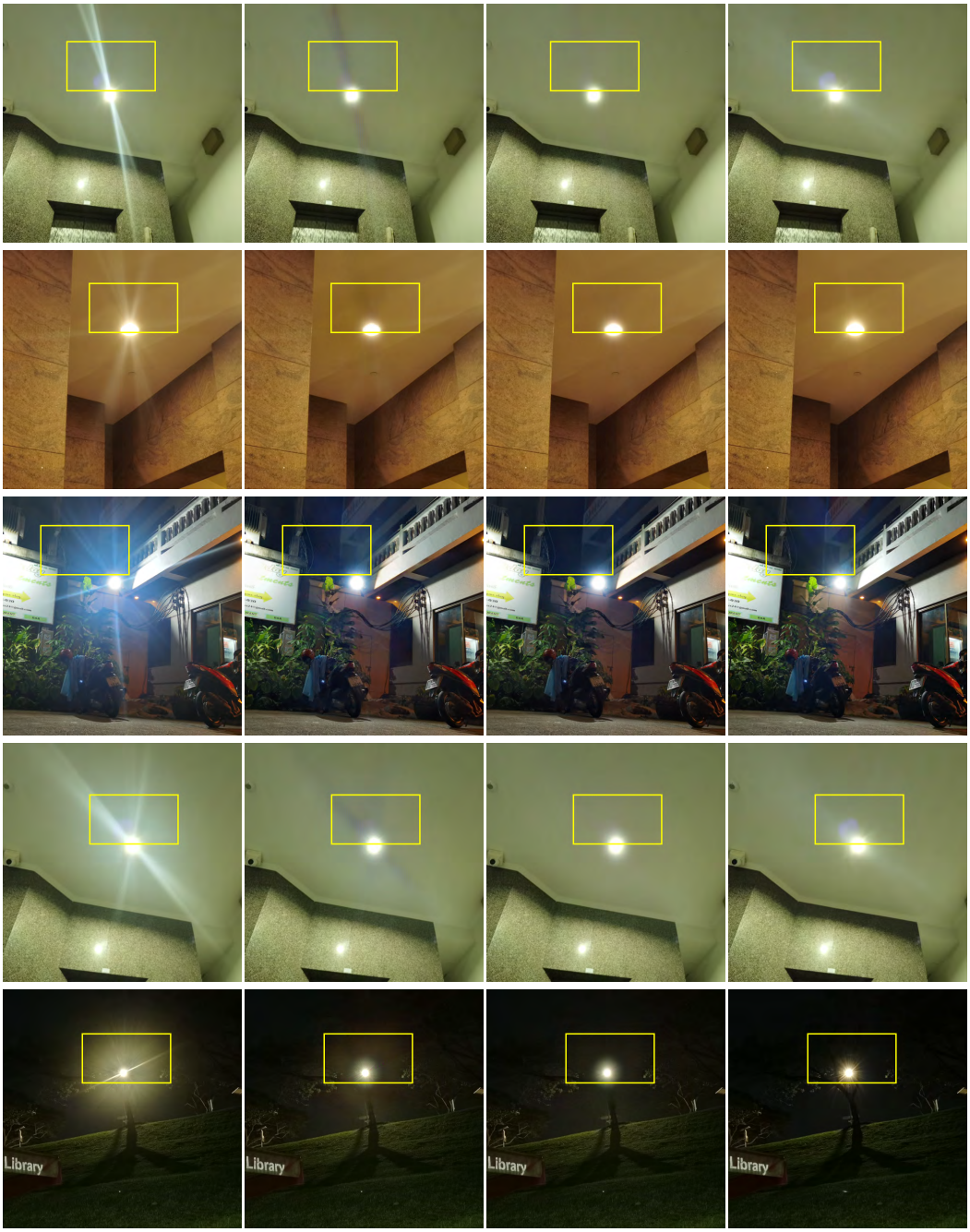


Figure 11: Visual results achieved by different pipelines on the FlareX dataset.



Input Flare7k pipeline Our GT

Figure 12: Visual results achieved by different pipelines on the Flare7k real dataset.

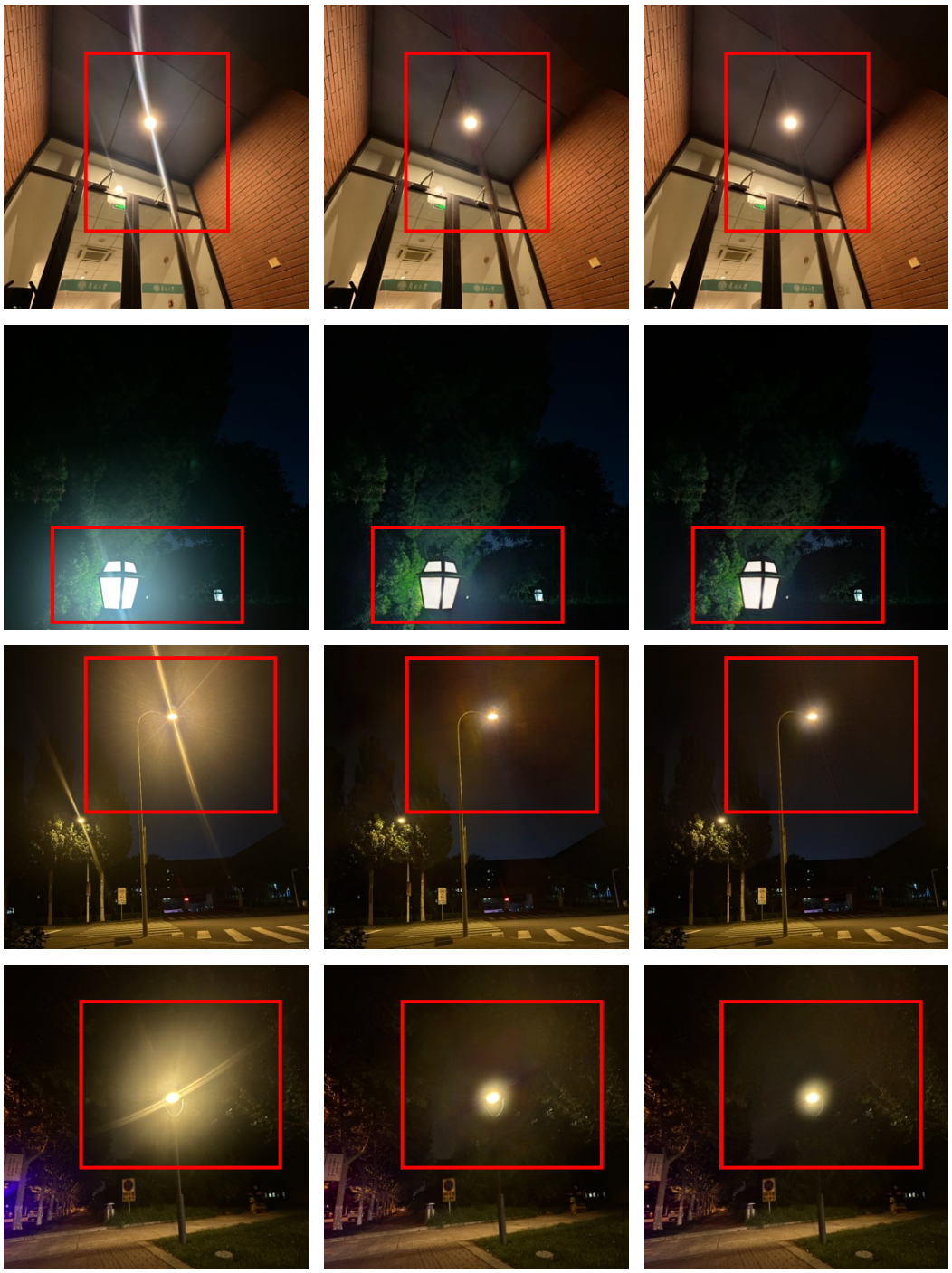


Figure 13: Visual results achieved by different pipelines on the FlareX dataset.



Figure 14: Visual results achieved by different pipelines on the FlareReal600 dataset.