HDR Reconstruction Boosting with Training-Free and Exposure-Consistent Diffusion Supplementary Material

Yo-Tin Lin[1] Su-Kai Chen[2] Hou-Ning Hu[2] Yen-Yu Lin[1] Yu-Lun Liu[1]
[1]National Yang Ming Chiao Tung University [2]MediaTek Inc.

1. Additional Implementation Details

Algorithm 1 HDR to LDR Tone Mapping using Inverse CRF

Input: HDR image hdr, inverse CRF $iCRF$
Output: LDR images $X : \{X_0, X_{-1}, X_{-2}, X_{-3}\}$
Define `ldr_indices` as an array from 0 to 255
for $ev = 0$ to -3 **do**
  Scale HDR: `hdr_scaled` $\leftarrow$ `hdr` $\times 2^{ev}$
  Initialize `tm_img` as zeros with the same shape as `hdr`
  for each channel c in {R, G, B} **do**
    Interpolate `hdr_scaled[..., c]` to `tm_img` using `iCRF[..., c]`
  end for
  $X_{ev} \leftarrow$ `tm_img`
end for

Algorithm 2 HDR Merging Algorithm

Input: LDR images $X : \{X_0, X_{-1}, X_{-2}, X_{-3}\}$, exposure times `times`, inverse CRF `iCRF`
Output: HDR image `hdr`
Initialize `hdr` and `total_weight` to zero
for each LDR image X_i **do**
  Normalize X_i to range [0, 1]
  Compute weights based on brightness and noise
  Linearize X_i using `iCRF`
  Accumulate weighted contributions to `hdr`
  Update `total_weight`
end for
Normalize `hdr` by `total_weight` and apply exponential
Handle saturated pixels using middle exposure if necessary
Return: HDR image `hdr`

Algorithm 3 Debevec HDR Merging with Simple CRF Repairing

Input: LDR images $\{X_0, X_{-1}, X_{-2}, X_{-3}\}$, exposure times `times`, inverse CRF `iCRF`
Output: HDR image `hdr`
Initialize `hdr` and `total_weight` to zero arrays
Set saturation and noise thresholds
for each LDR image X_i **do**
  Normalize X_i to range [0, 1]
  Compute pixel-wise weights using `weight_fun`
  Set weights to zero for very dark pixels
  Linearize X_i using inverse CRF (`removeCRF`)
  Accumulate: $hdr = hdr + weight \times (\log(\text{linearized } X_i + \delta) - \log(\text{exposure time}))$
  Update $total\_weight = total\_weight + weight$
end for
Normalize `hdr` by `total_weight`
Apply exponential to obtain final HDR values
if any pixel in `total_weight` is below saturation threshold **then**
  Use middle exposure and saturated image to repair saturated regions
  Replace saturated pixels in `hdr` with values from saturated image
end if
Return: HDR image `hdr`

2. SDEdit Scheduling.

Tab.1 shows the quantitative comparison of the different SDEdit strengths. With the iterative refinement process, our approach can improves the generation diversity, cross-exposure consistency and has the better visual quality scores.

3. User Study

To demonstrate that our method consistently improves perceptual quality, we conducted a user study with 29 partic-

Table 1. **Different SDEdit Strength.** This comparison shows the visual quality score on different SDEdit strengths on VDS dataset.

SDEdit Strength (RH's TMO)	BRISQUE ↓	NIQE ↓	NIMA ↑	MUSIQ ↑	PU21-PIQE ↓
Baseline (CEVR)	72.42	14.33	4.044	2.043	90.44
1.0	**70.41**	**13.31**	4.055	2.039	**80.37**
0.8	71.56	13.50	4.006	2.024	86.21
0.6	72.11	14.42	4.010	2.024	87.39
Ours Scheduling	71.17	13.71	**4.100**	**2.072**	86.12

ipants aged 20–60 (69% male, 31% female). For each of the 12 test cases, participants were shown a pair of images: the baseline output and our refined result based on the same baseline. They were asked to select the image that appeared more realistic and natural. The results are summarized in the table below, where a higher preference score indicates that a larger proportion of participants selected that option.

Table 2. **User Study.**

	Baseline	Ours
Preference	17.5%	82.5%

4. Runtime Breakdown

We analyzed the computational cost of each component in our pipeline. In addition, we evaluated how reducing the number of diffusion timesteps affects the overall runtime. With 50 timesteps and 4 iterations, the average runtime is 210 seconds, whereas using 20 timesteps reduces the runtime to 96 seconds.

Table 3. **Runtime Breakdown.**

Timesteps=50	Inpainting	Merging	Compensation
Runtime (ratio)	96.9%	2.5 %	0.6 %
Runtime (sec)	203.5	5.3	1.2
Timesteps=20	Inpainting	Merging	Compensation
Runtime (ratio)	93.4%	5.2%	1.4 %
Runtime (sec)	89.9	5.0	1.2

5. Complete Version of Ablation Study

We include the censored version of the ablation study in the main paper for clarity. In this section, we present the complete versions of Tab. 2 and 3 in the main paper to Tab. 4 and 5 here.

6. Quantitative Ablation Study for Compensation Pipeline.

In this section, we analyze both the average pixel value and the proportion of pixels falling below the defined lower bound. As shown in Tab. 6, our iterative inpainting approach effectively reduces the regions with insufficient luminance and increases pixel values in saturated areas, leading to more consistent exposure across the image.

Table 4. **Complete Version of Ablation study on diffusion backbones.**

Method	BRISQUE ↓		NIQE ↓		NIMA ↑		PU21-PIQE ↓
TMO	RH's	KK's	RH's	KK's	RH's	KK's	–
CEVR [**?**]	72.42	70.21	14.33	13.48	4.044	4.421	90.44
CEVR + **Ours (SDXL)**	71.17	68.90	13.71	12.97	4.100	4.456	86.12
CEVR + **Ours (SD turbo)**	**69.40**	**66.74**	**12.36**	**11.60**	**4.150**	**4.482**	**78.46**

Table 5. **Full Version of Quantitative ablation studies.**

Components (RH's TMO)	BRISQUE ↓	NIQE ↓	NIMA ↑	MUSIQ ↑	PU21-PIQE ↓
Baseline (CEVR)	72.42	14.33	4.044	2.043	90.44
w/o Merge & Comp.	**70.75**	13.57	4.045	2.050	85.15
w/o Merge	71.02	13.60	4.021	2.027	**84.68**
w/o Comp.	70.97	**13.39**	4.028	2.036	85.07
Ours	71.17	13.71	**4.100**	**2.072**	86.12

Table 6. **Quantitative Ablation Study for Compensation Pipeline.**

Iteration	1	2	3	4
Dark pixels (ratio)	44.3%	10.1 %	4.9 %	1.8 %
Average residual value	16.885	3.419	2.203	0.628

7. More Visual Comparisons

Figures 1 and 2 present additional cases of indirect HDR reconstruction; here, the "gamma" method refers to deriving the LDR stacks directly from the input using a gamma 2.2 function, with Fig. 1 showcasing images captured with a Fujifilm X-T30 and Fig. 2 featuring cases from the VDS and HDReye datasets.

8. Zoom-in Visualization

To further demonstrate the effectiveness of our proposed method, we provide high-resolution visual comparisons in Fig. 3, Fig. 4, and Fig. 6. These examples highlight the ability of our pipeline to generate realistic and consistent textures in over-exposed regions while preserving structural details and maintaining luminance consistency.

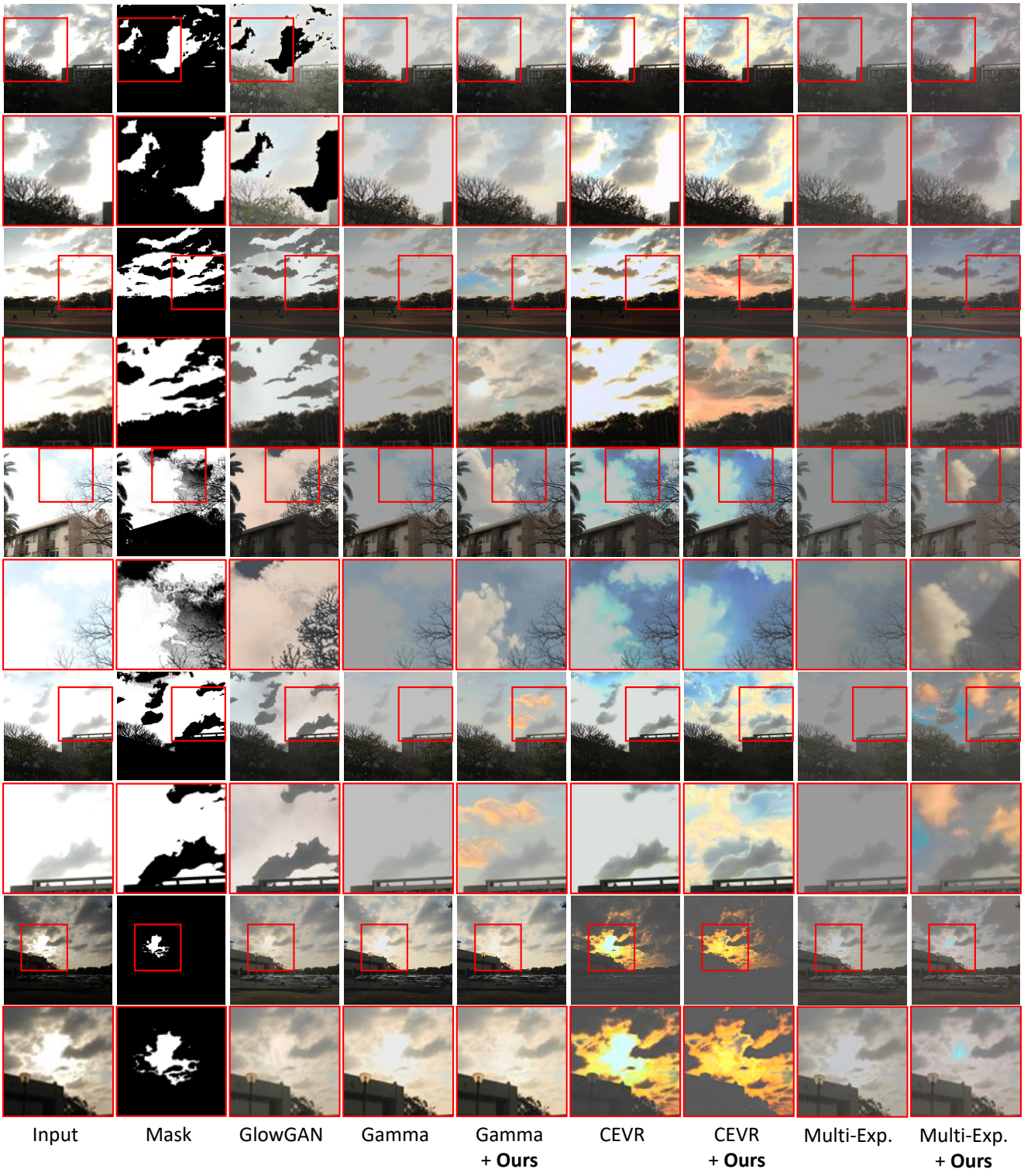

Figure 1. **In-the-wild Qualitative Results.** Our approach enhances various baseline methods (indicated by "+ **Ours**") across diverse scenes. There are *in-the-wild* cases captured by the Fujifilm X-T30.

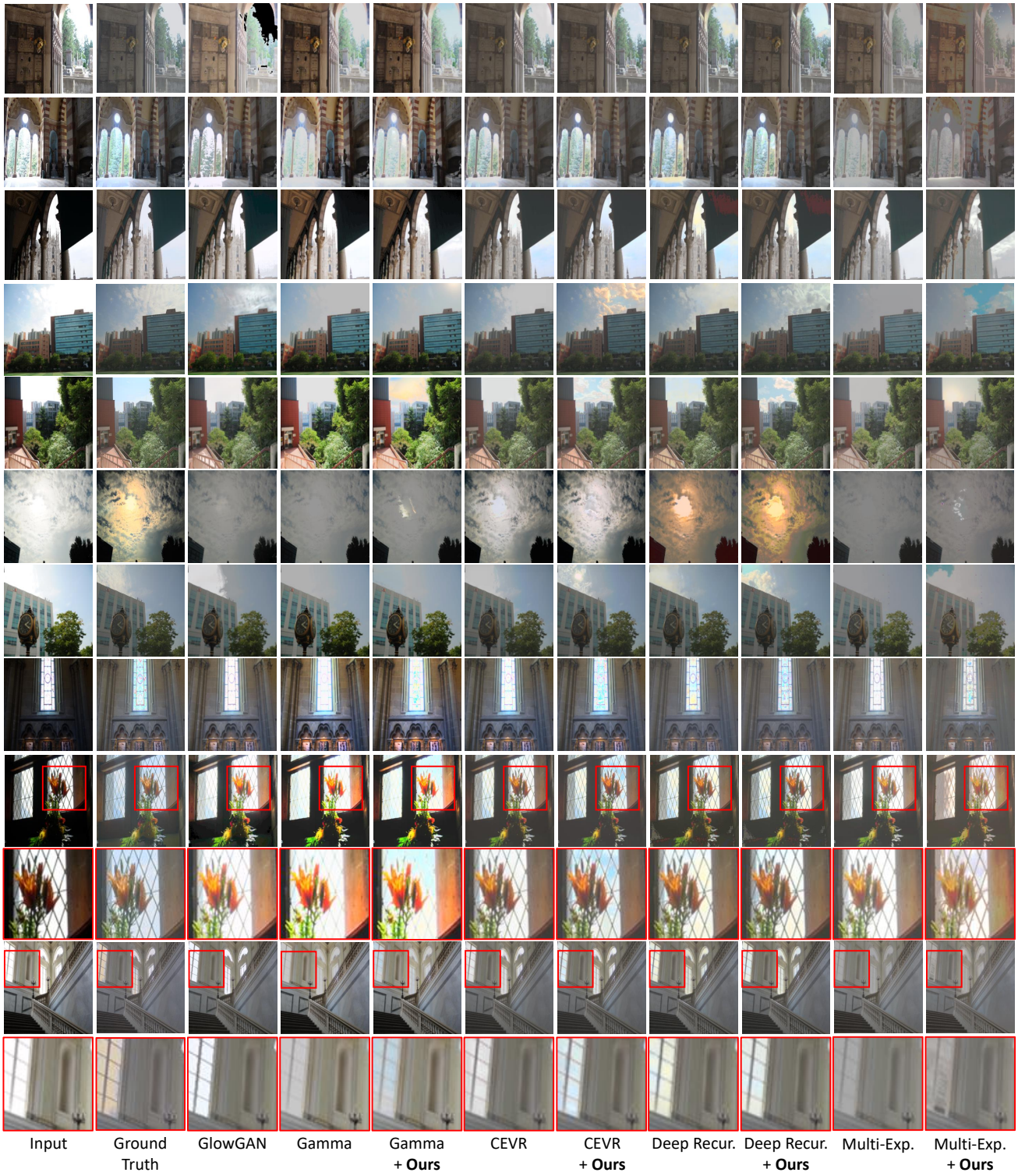

Figure 2. **Extreme Qualitative Results.** Our approach enhances various baseline methods (indicated by "**+ Ours**") across diverse scenes. There are *extreme* cases from public HDR dataset.

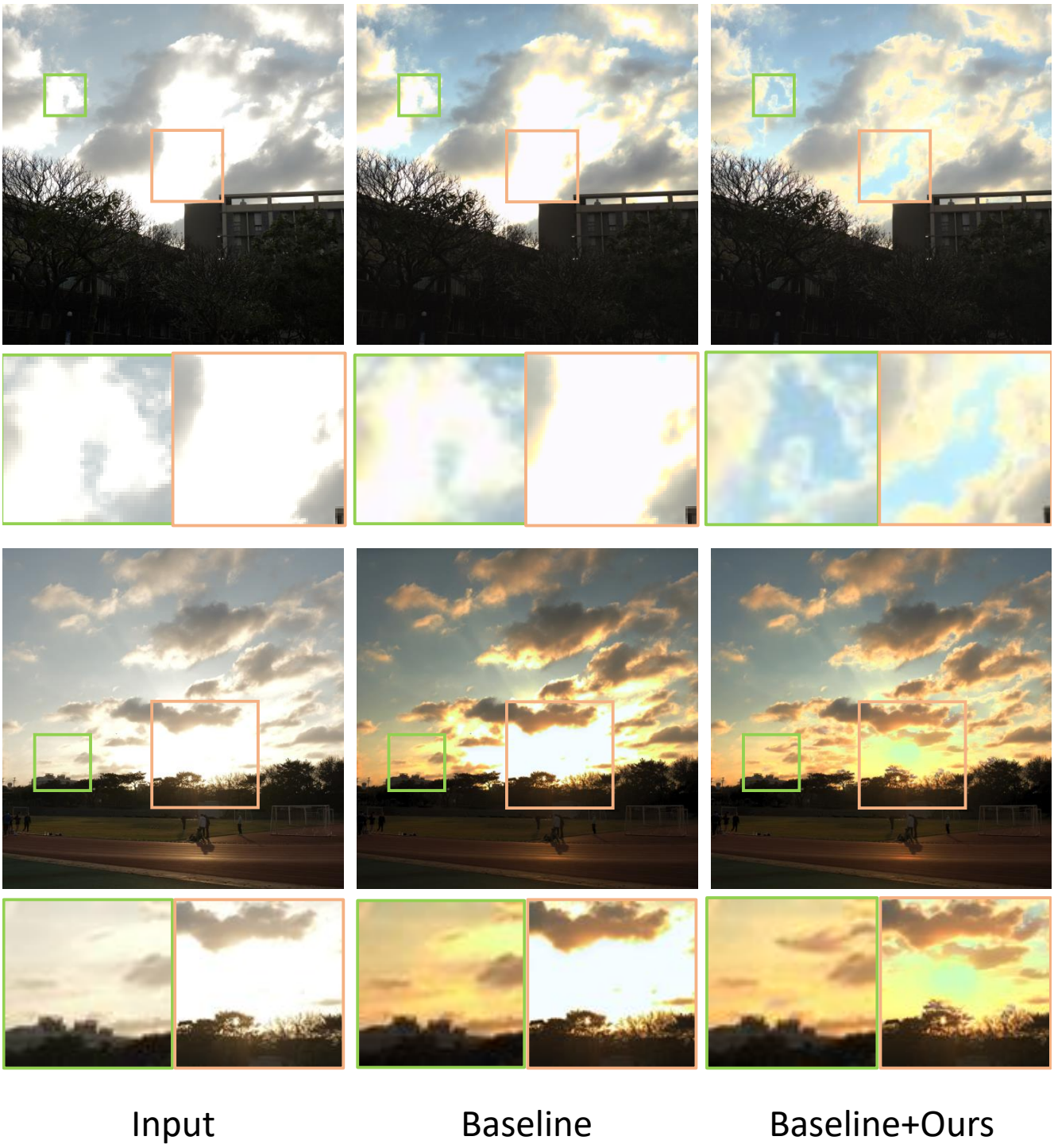

Figure 3. **More Visual Comparisons.** We also provide high-resolution comparisons here.

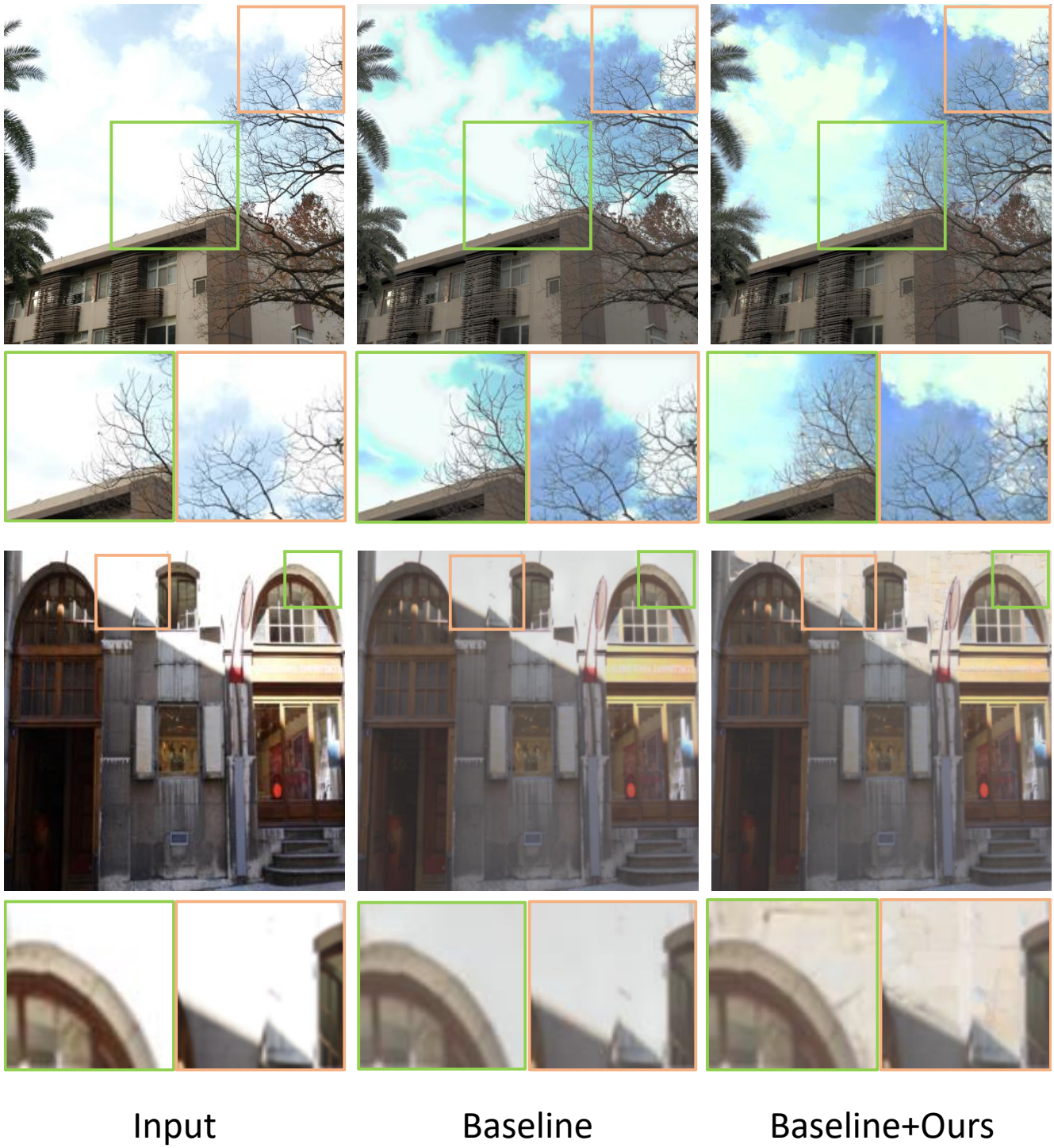

Figure 4. **More Visual Comparisons.** We also provide high-resolution comparisons here.

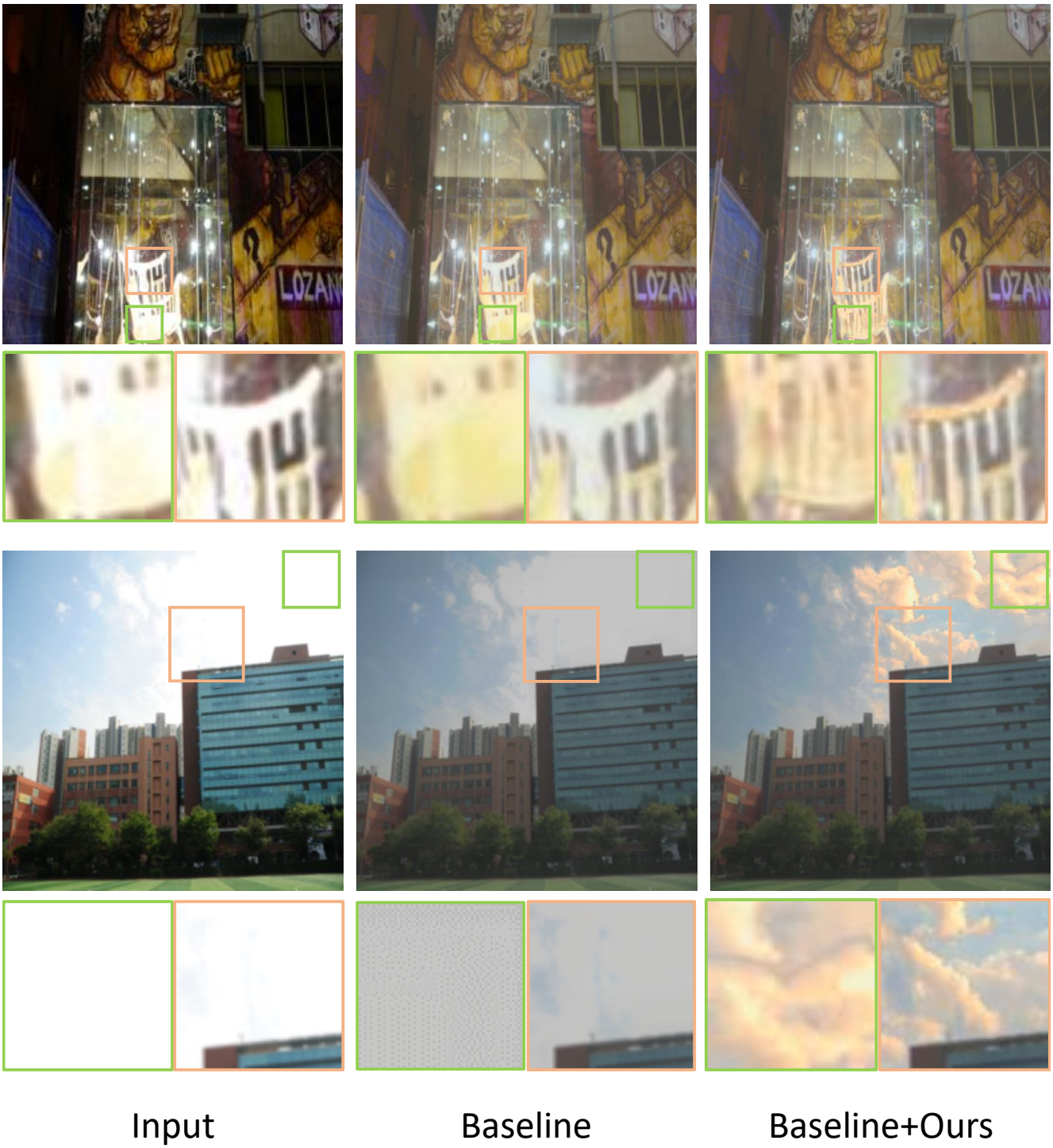

Figure 5. **More Visual Comparisons.** We also provide high-resolution comparisons here.

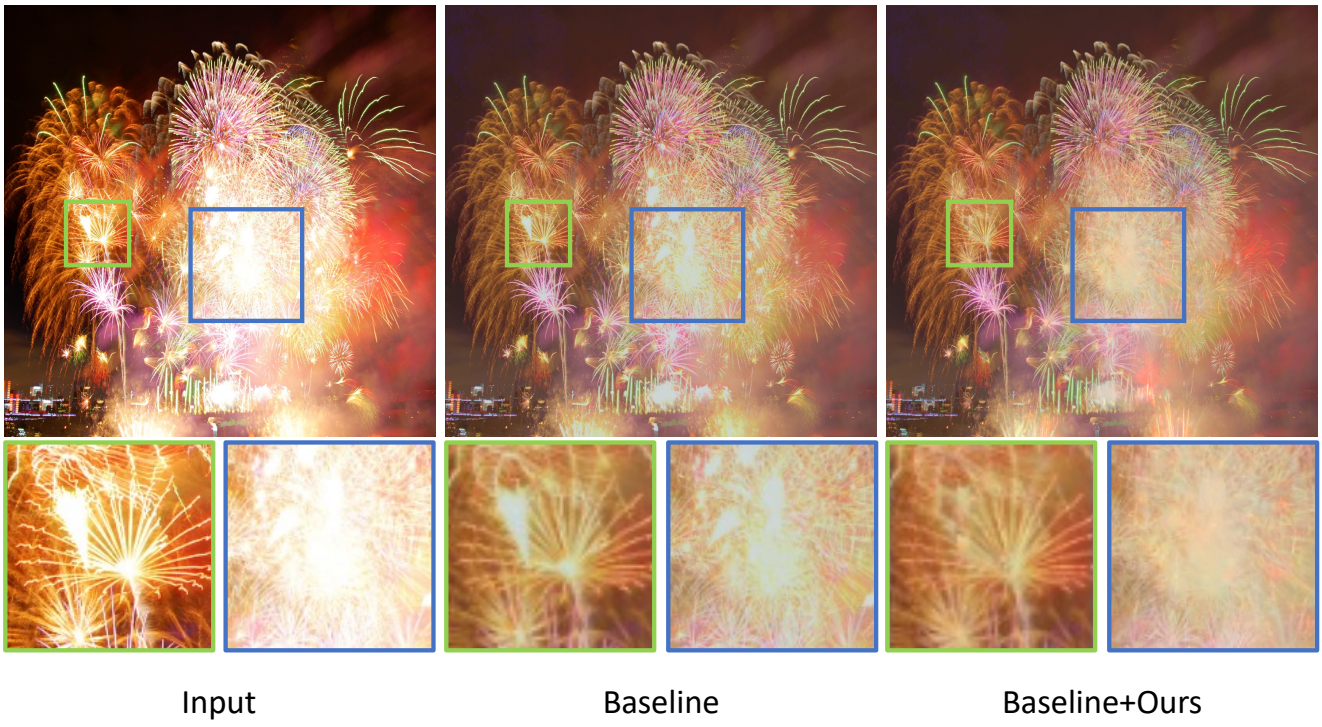

Figure 6. **More Visual Comparisons.** This in-the-wild extreme case also show the adaptability of our method to the large and small regions.