# Binarized Low-light Raw Video Enhancement

## Supplementary Material

In this supplementary material, we provide detailed information in various aspects of our work. Section A delves into the specifics of training our BRVD model, highlighting the unique challenges and solutions associated with backward propagation in BNNs. Section B presents our model's performance on the CRVD dataset, showcasing its efficacy in raw video denoising tasks and its generalization capabilities. Section C offers additional visual results of all comparison methods, including BNNs and full precision methods. Finally, in Section D, we discuss the current limitations of our BRVE model and outline potential future research directions to improve its adaptability and practical performance.

## A. Training details for BNN

**Updating weights.** In BNNs, during the forward propagation, both weights and activations are binarized. This is achieved by using the sign function shown in Equations 1 and 2. During backward propagation in training stage, we update the full precision parameter $\boldsymbol{W}^f$ based on gradient descent algorithm. The backward propagation process can be represented as

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^f} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^b}\frac{\partial \boldsymbol{W}^b}{\partial \boldsymbol{W}^f} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^b}\frac{\partial \mathrm{Sign}(\boldsymbol{W}^f)}{\partial \boldsymbol{W}^f}, \tag{12}$$

where $\frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^f}$ is the gradient of loss function $\mathcal{L}$ with respect to the real-valued parameter $\boldsymbol{W}^f$.

However, the derivative of the sign function is an impulse function, which is zero almost everywhere and causes graddient vanising, as shown in Figure 8 (a). To address this issue, BNNs often adopt the straight through estimator (STE) [7, 13] to approximate the gradient of sign function. The STE use the derivative of a clip function to replace sign function, *i.e.*,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^f} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^b}\frac{\partial \mathrm{Sign}(\boldsymbol{W}^f)}{\partial \boldsymbol{W}^f} \approx \frac{\partial \mathcal{L}}{\partial \boldsymbol{W}^b}\frac{\partial \mathrm{Clip}(\boldsymbol{W}^f)}{\partial \boldsymbol{W}^f}, \tag{13}$$

where the clip function shown in Figure 8 (b) can be defined as,

$$\mathrm{Clip}(\boldsymbol{W}^f) = \begin{cases} +1, & \boldsymbol{W}^f \geq 1 \\ \boldsymbol{W}^f, & -1 < \boldsymbol{W}^f < 1 \\ -1, & \boldsymbol{W}^f \leq -1 \end{cases}. \tag{14}$$

As shown in Figure 8 (b), the gradient of clip function w.r.t. the real-valued kernel weight is

$$\frac{\partial \mathrm{Clip}(\boldsymbol{W}^f)}{\partial \boldsymbol{W}^f} = \begin{cases} 0, & \boldsymbol{W}^f \geq 1 \\ 1, & -1 < \boldsymbol{W}^f < 1 \\ 0, & \boldsymbol{W}^f \leq -1 \end{cases}. \tag{15}$$

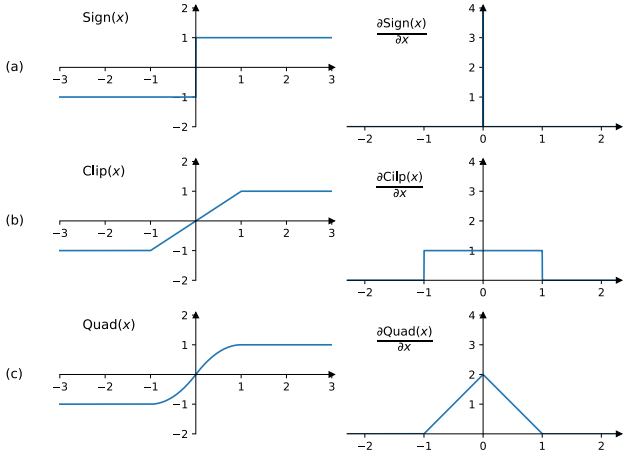


Figure 8. Illustration of the sign, clip, and piecewise quadratic function and their corresponding derivative.

**Approximated Derivative Function for Activations.** Following previous work [20, 41], we use a piecewise quadratic function to approximate the sign function during backward propagation. The gradient of loss function $\mathcal{L}$ w.r.t. $\boldsymbol{A}^{'}$ can be calculated as,

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{A}^{'}} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}^{b}}\frac{\partial \boldsymbol{A}^{b}}{\partial \boldsymbol{A}^{'}} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}^{b}}\frac{\partial \mathrm{Sign}(\boldsymbol{A}^{'})}{\partial \boldsymbol{A}^{'}} \approx \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}^{b}}\frac{\partial \mathrm{Quad}(\boldsymbol{A}^{'})}{\partial \boldsymbol{A}^{'}}, \tag{16}$$

where $\mathrm{Quad}(\boldsymbol{A}^{'})$ the piecewise quadratic function. As shown in Figure 8 (c), it is defined as

$$\mathrm{Quad}(\boldsymbol{A}^{'}) = \begin{cases} 1, & \boldsymbol{A}^{'} \geq 1 \\ 2\boldsymbol{A}^{'} - (\boldsymbol{A}^{'})^2, & 0 \leq \boldsymbol{A}^{'} < 1 \\ 2\boldsymbol{A}^{'} + (\boldsymbol{A}^{'})^2, & -1 \leq \boldsymbol{A}^{'} < 0 \\ -1, & \boldsymbol{A}^{'} < -1 \end{cases}. \tag{17}$$

Its gradient w.r.t. $\boldsymbol{A}^{'}$ can be represented as

$$\frac{\partial \mathrm{Quad}(\boldsymbol{A}^{'})}{\partial \boldsymbol{A}^{'}} = \begin{cases} 0, & \boldsymbol{A}^{'} \geq 1 \\ 2 - 2\boldsymbol{A}^{'}, & 0 \leq \boldsymbol{A}^{'} < 1 \\ 2 + 2\boldsymbol{A}^{'}, & -1 \leq \boldsymbol{A}^{'} < 0 \\ 0, & \boldsymbol{A}^{'} < -1 \end{cases}. \tag{18}$$

For the RSign function we use the approximated gradient in Equation 16 to update the learnable parameter $\boldsymbol{\alpha}$,

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \boldsymbol{\alpha}_i} &= \sum_{H,W} \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}_i^{'}}\frac{\partial \boldsymbol{A}_i^{'}}{\partial \boldsymbol{\alpha}_i} \\ &= \sum_{H,W} \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}_i^{'}}\frac{\partial (\boldsymbol{A}_i^{f} - \boldsymbol{\alpha}_i)}{\partial \boldsymbol{\alpha}_i} = -\sum_{H,W} \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}_i^{'}}, \end{aligned} \tag{19}$$

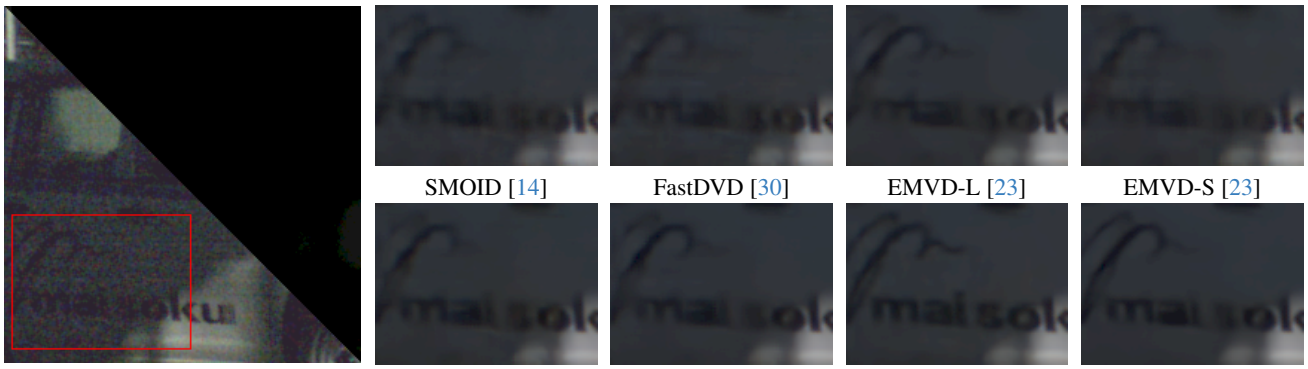


Figure 9. Additional visualization results of full precision methods on SMOID dataset.

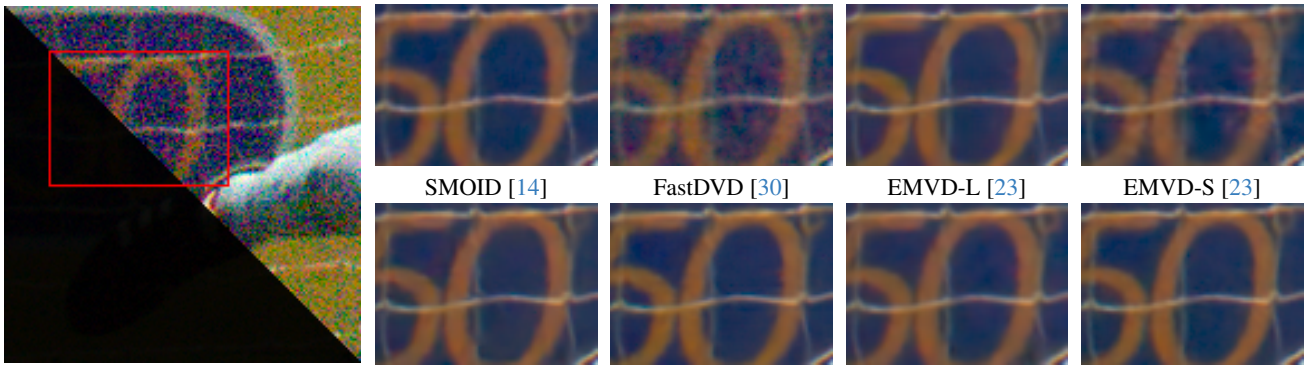


Figure 10. Additional visualization results of full precision methods on LLRVD dataset.

where $i = 1...C$ and $\boldsymbol{A}^f \in \mathbb{R}^{H\times W\times C}$. The gradient of the loss function w.r.t. the full precision activation $\boldsymbol{A}^f$ is

$$\frac{\partial \mathcal{L}}{\partial \boldsymbol{A}^f} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}'}\frac{\partial \boldsymbol{A}'}{\partial \boldsymbol{A}^f} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}'}\frac{\partial(\boldsymbol{A}^f - \boldsymbol{\alpha})}{\partial \boldsymbol{A}^f} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}'}\mathbf{1} = \frac{\partial \mathcal{L}}{\partial \boldsymbol{A}'}. \tag{20}$$

## B. Results on CRVD Dataset

We also train our BRVE model on the CRVD [46] dataset to validate its performance on raw video denoising task. The quantitative results on indoor scenes are shown in Table 1. We also illustrate the visualization of denoising effects on the indoor scenes in Figures 13 and 14. This experiment shows that our BRVE model has a good generalization ability on other raw video restoration tasks.

## C. More Visualization Results

We present the visualization results of the full precision methods in Figures 9 and 10, which correspond to the BNN method visualizations shown in Figures 5 and 6, respectively. In addition, we present more visualization results of all methods on SMOID dataset and LLRVD dataset in Figures 11 and 12 respectively.

| Method | PSNR ↑ | SSIM ↑ |
|---|---|---|
| FastDVD [30] | 39.84 | 0.9703 |
| EMVD-S [23] | 42.63 | 0.9851 |
| RViDeNet [46] | 43.97 | 0.9874 |
| LLRVD [10] | 44.23 | 0.9879 |
| FloRNN [19] | 45.15 | 0.9907 |
| BRVE (ours) | 43.32 | 0.9863 |

Table 1. Quantitative results on CRVD dataset.

## D. Limitations and Future Work

While our BRVE model has achieved promising results in enhancing low-light raw videos, it also has some limitations. On one hand, we use a fixed kernel for all scenes, which is less flexible than the feature alignment modules using optical flow or deformable convolution. Developing an adaptive shift kernel presents a potential solution to address this issue, offering a more flexible approach that could enhance the model's performance across varying scenes. On the other hand, we only analyze the operation number of our BRVE model theoretically. Additionally, deploying it on edge devices to practically validate its acceleration performance is a worthwhile direction for future work.

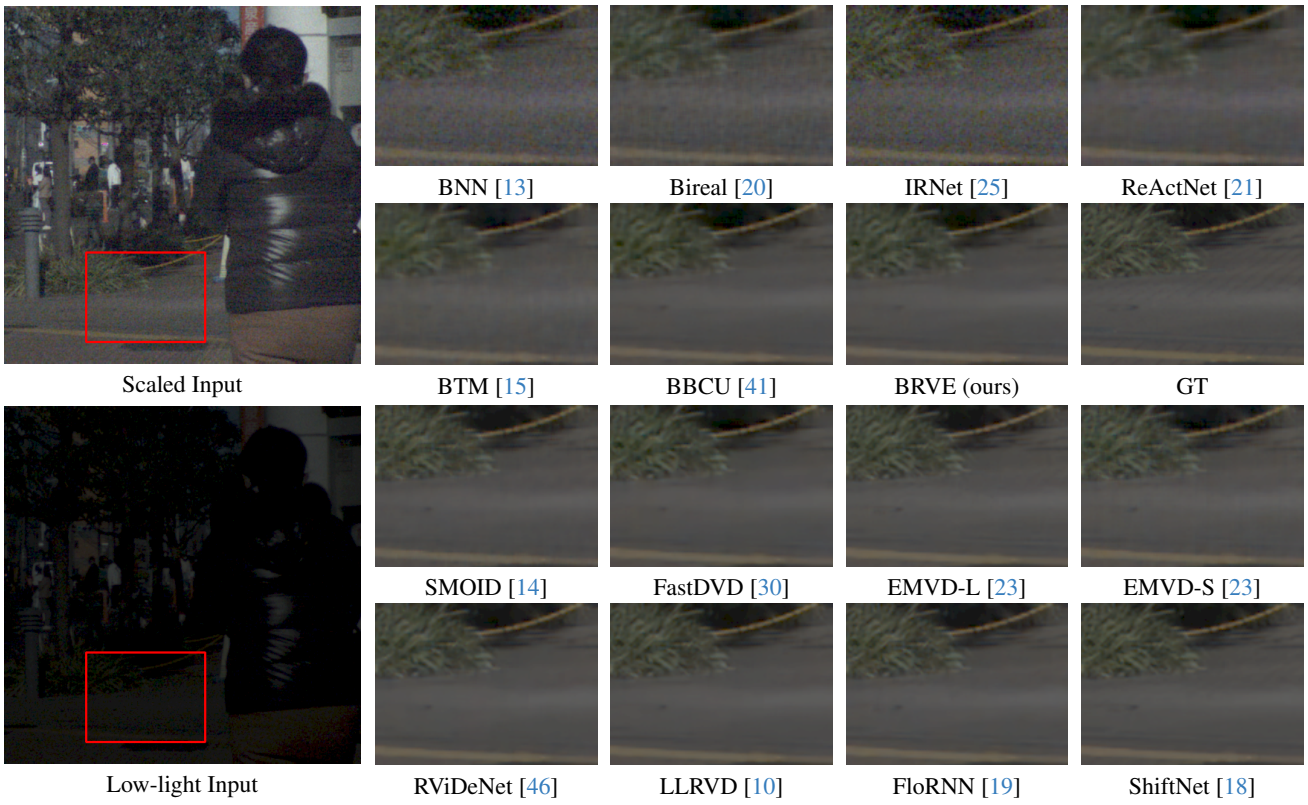


Figure 11. Additional visualization results of full precision methods on SMOID dataset.

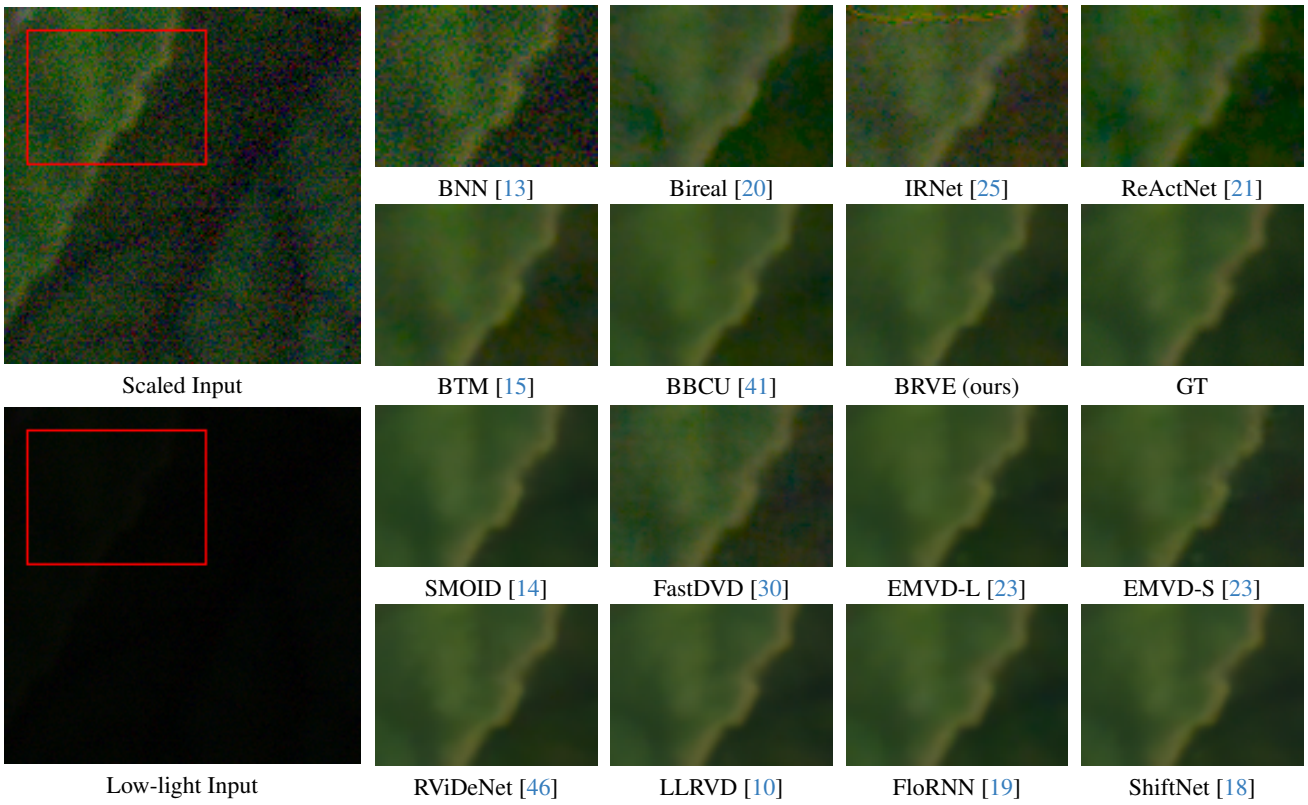


Figure 12. Additional visualization results of full precision methods on LLRVD dataset.

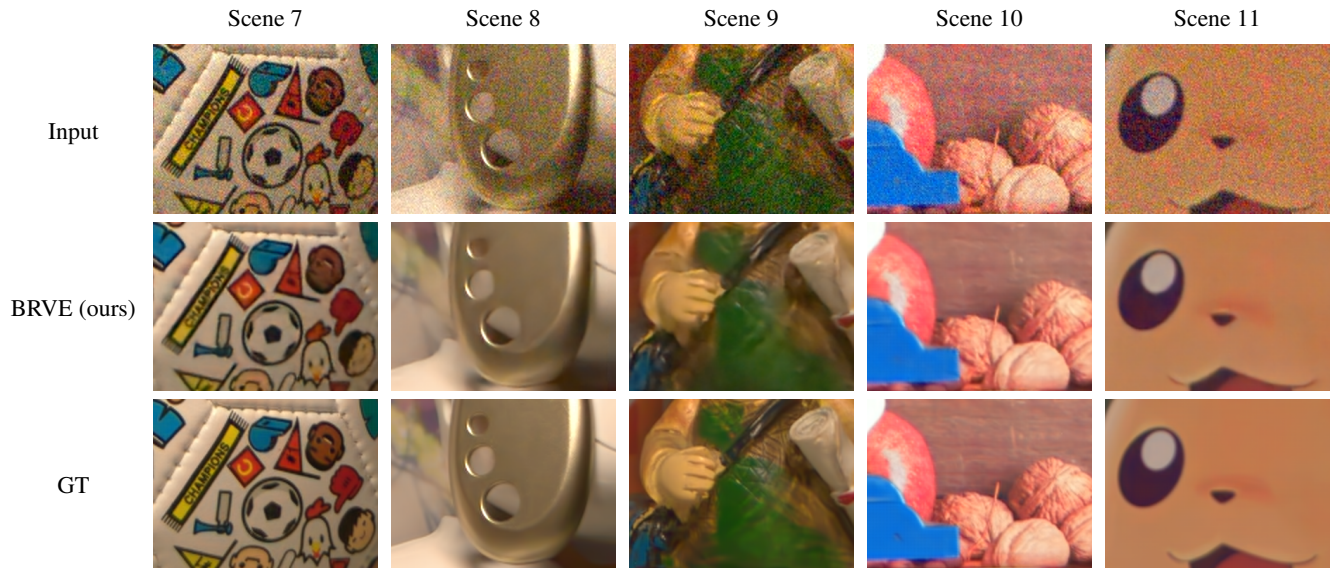


Figure 13. Visualization denoising results of our BRVE model on indoor scenes of CRVD dataset.

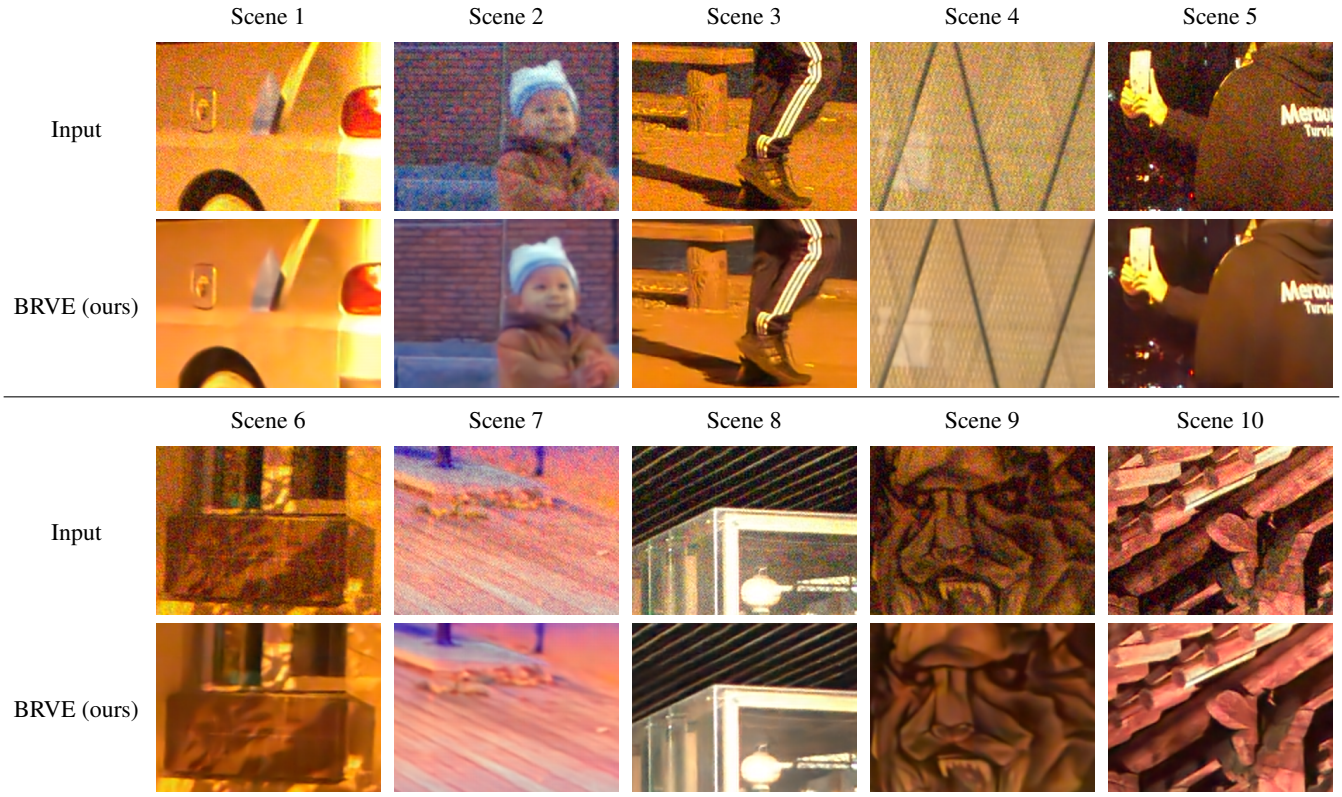


Figure 14. Visualization denoising results of our BRVE model on outdoor scenes of CRVD dataset.