Supplementary Material – Burst Image Restoration and Enhancement

Akshay Dudhane[1] Syed Waqas Zamir[2] Salman Khan[1,3]
Fahad Shahbaz Khan[1,4] Ming-Hsuan Yang[5,6,7]
[1]Mohamed bin Zayed University of AI [2]Inception Institute of AI [3]Australian National University
[4]Linköping University [5]University of California, Merced [6]Yonsei University [7]Google Research

Here we describe the architectural details of the proposed BIPNet (Sec. 1), and present additional visual comparisons with existing state-of-the-art approaches for burst SR and burst de-noising (Sec. 2 and Sec. 3).

1. Network Architecture Details

1.1. Edge Boosting Feature Alignment (EBFA)

The proposed feature processing module (FPM) consists of three residual-in-residual (RiR) [?] groups. Each RiR is made up of three RGCAB and each RGCAB contains a basic residual block followed by a global context attention as shown in Fig. 2 (a) of the main paper. Although, the deformable convolution layer is shown only once in the Fig. 2 (b) for simplicity, we apply three such layers to improve the feature alignment ability of the proposed EBFA module.

1.2. Pseudo Burst Feature Fusion (PBFF)

The proposed PBFF is as shown in Fig. 3 (a) in main paper. It consists of multi-scale feature (MSF) extraction module which is made up of a light-weight 3-level U-Net [5]. We employed one FPM (with 2 RiR and 2 RGCAB in each RiR) after each downsample and upsample convolution layer. Number of convolution filters are increased by a factor of 1.5 at each downsampling step and decreased by the rate of 1.5 after each upsampling operation. We simply add features extracted at each level to the upsampled features via skip connections.

1.3. Adaptive Group Up-sampling (AGU)

Our AGU module is shown in Fig. 3 (c) in the main paper. It aggregates the input group of pseudo bursts and pass them through a bottleneck convolution layer of kernel size 1×1 followed by a set of four parallel convolution layers, each with kernel size of 1×1 and 64 filters. Further, the outputs from previous step are passed through the softmax activation to obtain the dense attention maps.

2. Additional Visual Results for Burst SR

The results provided in Fig. S1 and Fig. S2 show that our method performs favorably on both real and synthetic images for the scale factor $\times4$. The true potential of the proposed approach is demonstrated in Fig. S3, where it successfully recovers the fine-grained details from extremely challenging LR burst images (that are down-scaled by a factor of $\times8$).

3. Additional Results for Burst Denoising

The results provided in Fig. S4 and Fig. S5 show that our method performs favorably on both grayscale [4] and color [6] noisy images. Specifically, it can recover fine details in the outputs and is more closer to the ground-truth compared to existing state-of-the-art approaches.

References

[1] Goutam Bhat, Martin Danelljan, Luc Van Gool, and Radu Timofte. Deep burst super-resolution. In *CVPR*, 2021. 2, 3

[2] Goutam Bhat, Martin Danelljan, Fisher Yu, Luc Van Gool, and Radu Timofte. Deep reparametrization of multi-frame super-resolution and denoising. In *ICCV*, 2021. 2, 4

[3] Bruno Lecouat, Jean Ponce, and Julien Mairal. Lucas-kanade reloaded: End-to-end super-resolution from raw image bursts. In *ICCV*, 2021. 2

[4] Ben Mildenhall, J. Barron, Jiawen Chen, Dillon Sharlet, R. Ng, and Robert Carroll. Burst denoising with kernel prediction networks. In *CVPR*, 2018. 1, 4

[5] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: convolutional networks for biomedical image segmentation. In *MICCAI*, 2015. 1

[6] Zhihao Xia, Federico Perazzi, M. Gharbi, Kalyan Sunkavalli, and A. Chakrabarti. Basis prediction networks for effective burst denoising with large kernels. In *CVPR*, 2020. 1, 4

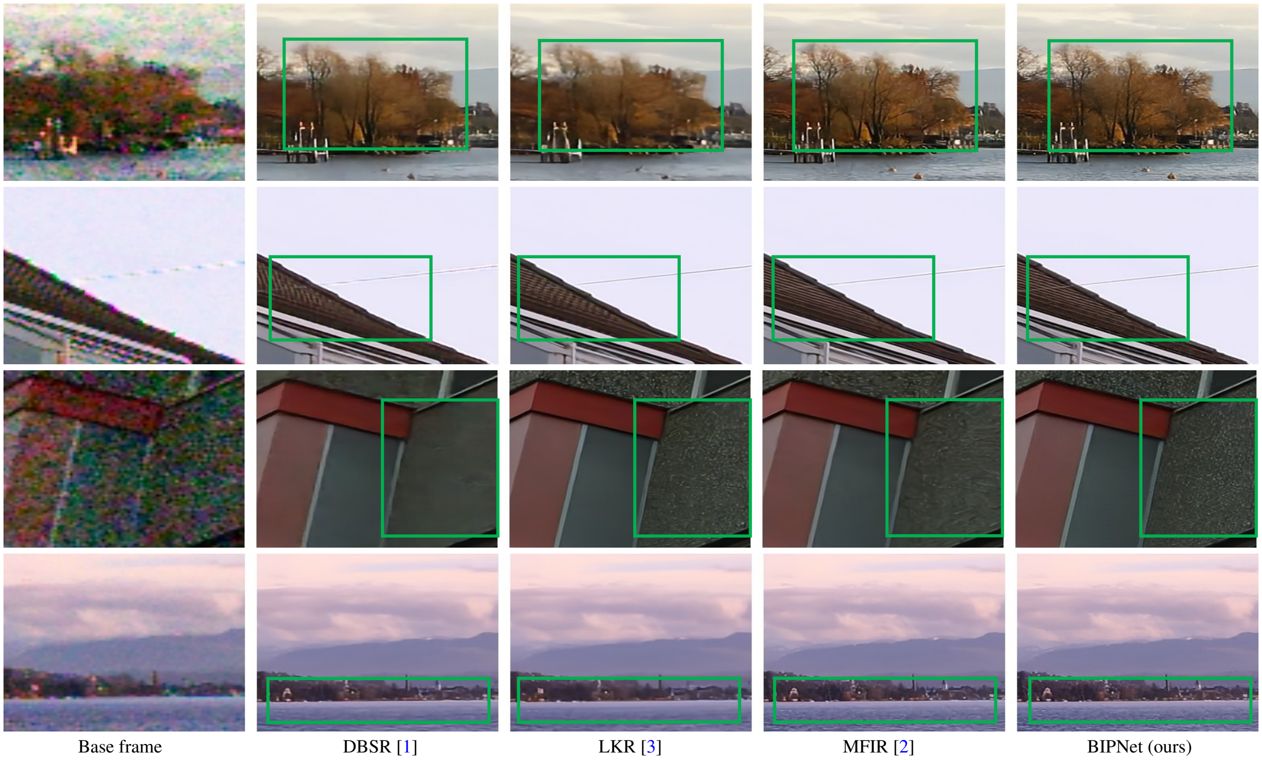


Figure S1. Comparisons for $\times 4$ burst super-resolution on SyntheticBurst dataset [1]. Our BIPNet produces more sharper and clean results than other competing approaches (specifically the marked green box regions).

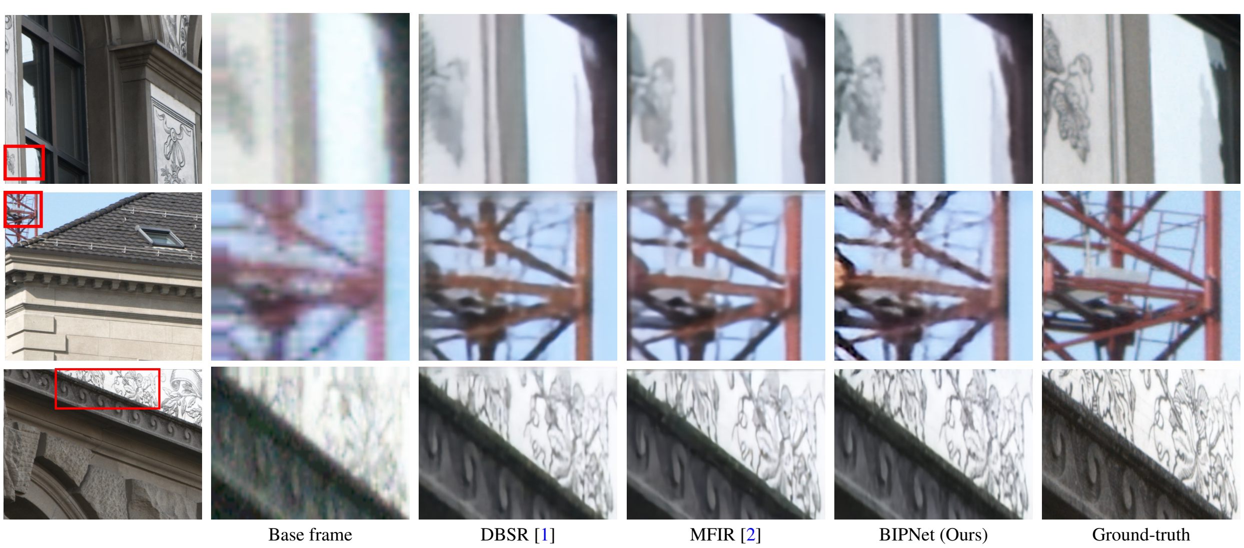
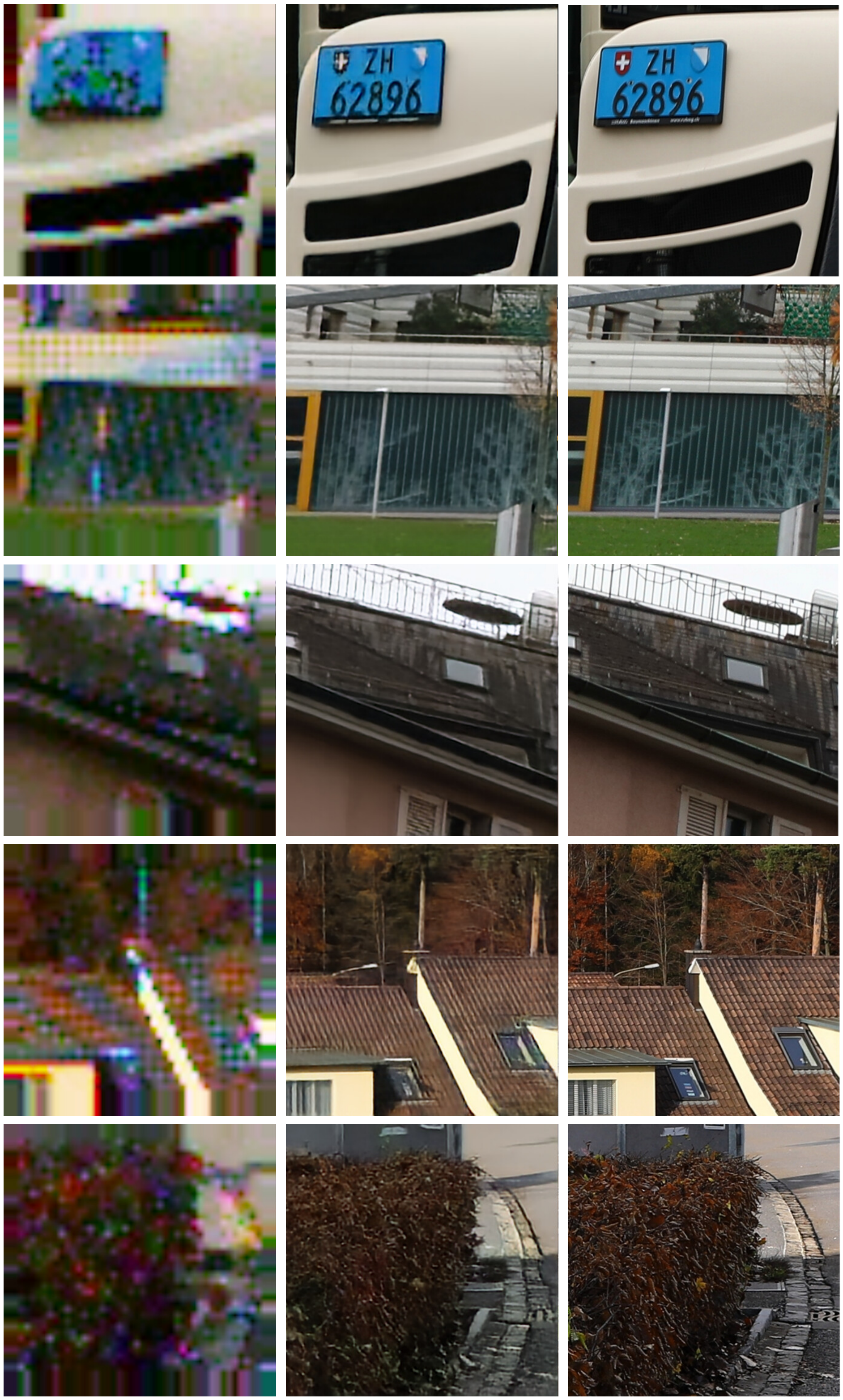


Figure S2. Comparison for $\times 4$ burst SR on real BurstSR dataset [1]. The crops shown in red boxes (in the input images shown in the left-most column) are magnified to illustrate the improvements in restoration results. The reproductions of our BIPNet are perceptually more faithful to the ground-truth than those of other methods.



Base frame | **BIPNet (Ours)** | Ground-truth

Figure S3. Results for $\times 8$ SR on images from SyntheticBurst dataset [1]. Our method effectively recovers image details in extremely challenging cases.

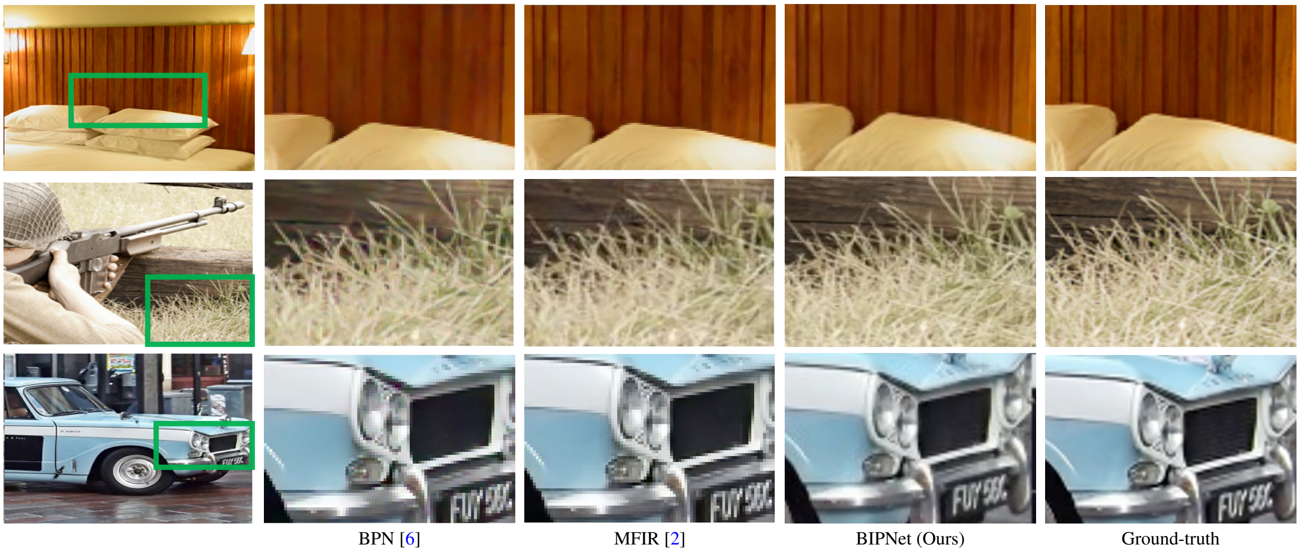


Figure S4. Comparisons for burst denoising on color datasets [6]. The crops shown in green boxes (in the input images shown in the left-most column) are magnified to illustrate the improvements in restoration results. Our proposed BIPNet produces more sharper and clean results than other competing approaches.

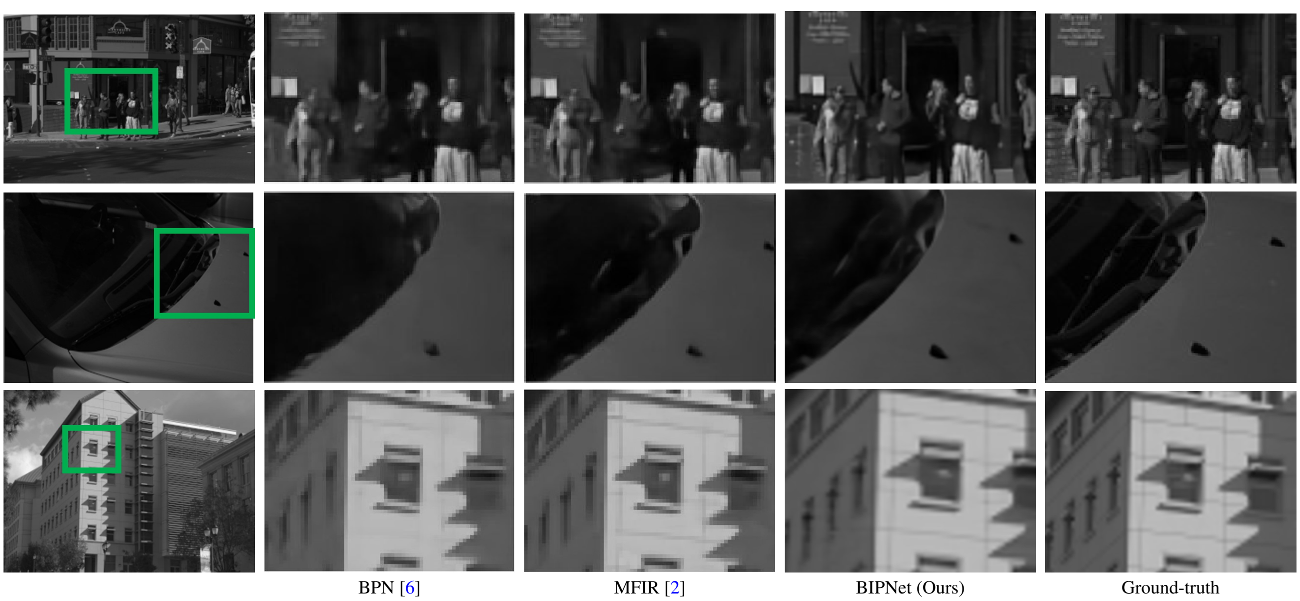


Figure S5. Comparisons for burst denoising on gray-scale [4]. The crops shown in green boxes (in the input images shown in the left-most column) are magnified to illustrate the improvements in restoration results. Our BIPNet produces more sharper and clean results than other competing approaches.