# **Supplementary Material – Multi-Stage Progressive Image Restoration**

Syed Waqas Zamir<sup>\* 1</sup> Aditya Arora<sup>\* 1</sup> Salman Khan<sup>2</sup> Munawar Hayat<sup>3</sup> Fahad Shahbaz Khan<sup>2</sup> Ming-Hsuan Yang<sup>4,5,6</sup> Ling Shao<sup>1,2</sup> <sup>1</sup>Inception Institute of AI <sup>2</sup>Mohamed bin Zayed University of AI <sup>3</sup>Monash University <sup>4</sup>University of California, Merced <sup>5</sup>Yonsei University <sup>6</sup>Google Research

## 1. Additional Ablation on ORSNet design

In the last stage of the proposed architecture, we employ the original-resolution subnetwork (ORSNet). Table 1 shows the impact of changing the original-resolution blocks (ORBs) and channel-attention blocks (CABs). While increasing the number of ORBs and CABs in ORSNet consistently improves accuracy, we use 3 ORBs and 8 CABs so to have a good trade-off between speed and accuracy.

#### 2. Detailed Diagram of CSFF

Cross-stage feature fusion (pink arrows in Fig. 2 of the main paper) is shown at one level for brevity. However, this process is repeated at each level of the encoder-decoder, as shown in Fig 1.

### 3. Training Times

On two NVIDIA Tesla V100 GPUs, our MPRNet takes training time of approximately 37 hours for deraining, 62 hours for denoising, and 86 hours for deblurring.

## 4. Image Deraining Results

Figures 2, 3 and 4 show deraining results of our MPR-Net and those of the state-of-the-art on several challenging images from different datasets. Our method effectively removes rain streaks and yields good quality images both visually and in terms of PSNR.

## 5. Image Deblurring Results

Here we test the performance of different image deblurring methods on synthetic as well as real datasets. For the case of synthetic datasets, the visual results are shown in Figures 5 and 6 on the GoPro dataset [7], and in Figures 7 and 8 on the HIDE dataset [12]. We further evaluate the competing methods on real-world images of the RealBlur dataset [11] and results are presented in Figures 9 and 10 for the RealBlur-J subset and in Figures 11 and 12 for the RealBlur-R subset.

Table 1: Ablation: the impact of increasing ORBs and CABs in the ORSNet subnetwork.

#ORBs	#CABs	PSNR
3	4	28.39
2	8	28.70
3	8	28.96
4	8	29.07
4	10	29.10



Figure 1: Cross-stage feature fusion between the encoderdecoders of stage 1 and stage 2.

#### 6. Image Denoising Results

We provide additional denoising comparisons of our method with the state-of-the-art. Results on the SIDD dataset [1] are shown in Figures 13 and 14. And visual examples on the DND dataset [9] are illustrated in Figure 15.

#### References

- Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A high-quality denoising dataset for smartphone cameras. In *CVPR*, 2018. 1, 14, 15
- [2] Meng Chang, Qi Li, Huajun Feng, and Zhihai Xu. Spatialadaptive network for single image denoising. In *ECCV*, 2020. 14, 15, 16
- [3] Kui Jiang, Zhongyuan Wang, Peng Yi, Baojin Huang, Yimin Luo, Jiayi Ma, and Junjun Jiang. Multi-scale progressive

<sup>\*</sup>Equal contribution

fusion network for single image deraining. In *CVPR*, 2020. 3, 4, 5

- [4] Yoonsik Kim, Jae Woong Soh, Gu Yong Park, and Nam Ik Cho. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization. In *CVPR*, 2020. 14, 15, 16
- [5] Orest Kupyn, Tetiana Martyniuk, Junru Wu, and Zhangyang Wang. DeblurGAN-v2: Deblurring (orders-of-magnitude) faster and better. In *ICCV*, 2019. 6, 7, 8, 9, 10, 11, 12, 13
- [6] Xia Li, Jianlong Wu, Zhouchen Lin, Hong Liu, and Hongbin Zha. Recurrent squeeze-and-excitation context aggregation net for single image deraining. In ECCV, 2018. 3, 4, 5
- [7] Seungjun Nah, Tae Hyun Kim, and Kyoung Mu Lee. Deep multi-scale convolutional neural network for dynamic scene deblurring. In *CVPR*, 2017. 1, 6, 7
- [8] Dongwon Park, Dong Un Kang, Jisoo Kim, and Se Young Chun. Multi-temporal recurrent neural networks for progressive non-uniform single image deblurring with incremental temporal training. In *ECCV*, 2020. 6, 7, 8, 9
- [9] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *CVPR*, 2017. 1, 16
- [10] Dongwei Ren, Wangmeng Zuo, Qinghua Hu, Pengfei Zhu, and Deyu Meng. Progressive image deraining networks: A better and simpler baseline. In *CVPR*, 2019. 3, 4, 5
- [11] Jaesung Rim, Haeyun Lee, Jucheol Won, and Sunghyun Cho. Real-world blur dataset for learning and benchmarking deblurring algorithms. In ECCV, 2020. 1, 10, 11, 12, 13
- [12] Ziyi Shen, Wenguan Wang, Xiankai Lu, Jianbing Shen, Haibin Ling, Tingfa Xu, and Ling Shao. Human-aware motion deblurring. In *ICCV*, 2019. 1, 8, 9
- [13] Maitreya Suin, Kuldeep Purohit, and A. N. Rajagopalan. Spatially-attentive patch-hierarchical network for adaptive motion deblurring. In *CVPR*, 2020. 6, 7, 8, 9
- [14] Xin Tao, Hongyun Gao, Xiaoyong Shen, Jue Wang, and Jiaya Jia. Scale-recurrent network for deep image deblurring. In *CVPR*, 2018. 6, 7, 8, 9, 10, 11, 12, 13
- [15] Wei Wei, Deyu Meng, Qian Zhao, Zongben Xu, and Ying Wu. Semi-supervised transfer learning for image rain removal. In *CVPR*, 2019. 3, 4, 5
- [16] Wenhan Yang, Robby T Tan, Jiashi Feng, Jiaying Liu, Zongming Guo, and Shuicheng Yan. Deep joint rain detection and removal from a single image. In *CVPR*, 2017. 4, 5
- [17] Rajeev Yasarla and Vishal M Patel. Uncertainty guided multi-scale residual learning-using a cycle spinning cnn for single image de-raining. In CVPR, 2019. 3, 4, 5
- [18] Zongsheng Yue, Qian Zhao, Lei Zhang, and Deyu Meng. Dual adversarial network: Toward real-world noise removal and noise generation. In *ECCV*, 2020. 14, 15, 16
- [19] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. CycleISP: Real image restoration via improved data synthesis. In *CVPR*, 2020. 14, 15, 16
- [20] Hongguang Zhang, Yuchao Dai, Hongdong Li, and Piotr Koniusz. Deep stacked hierarchical multi-patch network for image deblurring. In *CVPR*, 2019. 6, 7, 8, 9, 10, 11, 12, 13
- [21] He Zhang and Vishal M Patel. Density-aware single image de-raining using a multi-stream dense network. In CVPR, 2018. 3, 4

- [22] He Zhang, Vishwanath Sindagi, and Vishal M Patel. Image de-raining using a conditional generative adversarial network. *TCSVT*, 2019. 3, 5
- [23] Kaihao Zhang, Wenhan Luo, Yiran Zhong, Lin Ma, Bjorn Stenger, Wei Liu, and Hongdong Li. Deblurring by realistic blurring. In *CVPR*, 2020. 6, 7, 8, 9



PSNR Reference 21.07 dB Rainy

17.81 dB SEMI [15]

18.90 dB UMRL [17]



21.06 dB RESCAN [6]



17.96 dB PreNet [10]





31.05 dB MPRNet (Ours)



PSNR





23.90 dB

MSPFN [3]





Figure 2: Visual examples for image deraining. Top image is from Test100 [22], and the bottom is from Test1200 [21].







16.46 dB Rainy



21.39 dB SEMI [15]



28.67 dB UMRL [17]



25.76 dB RESCAN [6]





25.09 dB MSPFN [3]



32.06 dB MPRNet (Ours)



PSNR Reference





14.86 dB SEMI [15]



25.71 dB UMRL [17]



Figure 3: Visual examples for image deraining. Top image is from Test1200 [21], and the bottom is from Rain100H [16].







19.41 dB RESCAN [6]



8.39 dB

21.33 dB PreNet [10]



9.52 dB

20.72 dB MSPFN [3]



20.79 dB

23.51 dB



PSNR Reference



21.52 dB Rainy



18.12 dB SEMI [15]



19.25 dB UMRL [17]



20.15 dB RESCAN [6]



18.82 dB PreNet [10]



20.77 dB MSPFN [3]



29.25 dB MPRNet (Ours)

Figure 4: Visual examples for image deraining. Top image is from Rain100H [16], and the bottom is from Test100 [22].







28.68 dB DeblurGAN-V2 [5]

28.60 dB DBGAN [23]

28.46 dB MTRNN [8]



28.80 dB



Suin *et al*. [13]



24.91 dB

PSNR Reference 19.59 dB Blurry



26.30 dB

26.15 dB



25.02 dB



Figure 5: Image deblurring comparisons on the GoPro dataset [7]. The full-resolution versions of the images provided in Figure 6 of the main paper.



Reference

30.03 dB



30.53 dB

SRN [14]



30.84 dB DMPHN [20]



29.41 dB

31.27 dB Suin *et al*. [13]



32.47 dB MPRNet (Ours)



Reference



26.35 dB





29.49 dB

SRN [14]

32.35 dB DeblurGAN-V2 [5] 32.39 dB DBGAN [23]

31.49 dB MTRNN [8]



DMPHN [20]

Suin *et al*. [13]

MPRNet (Ours)

Figure 6: Image deblurring comparisons on the GoPro dataset [7].



PSNR Reference

21.63 dB





25.78 dB SRN [14]

26.95 dB

25.04 dB DeblurGAN-V2 [5] 25.54 dB



25.83 dB DMPHN [20]





31.04 dB

PSNR Reference 25.25 dB Blurry



31.10 dB MTRNN [8]



31.24 dB

DMPHN [20]

29.55 dB

31.85 dB Suin *et al*. [13]

MPRNet (Ours)

Figure 7: Image Deblurring results on the HIDE dataset [12].



Reference

20.89 dB



SRN [14]

26.65 dB

25.13 dB DeblurGAN-V2 [5]



25.79 dB DBGAN [23]

MTRNN [8] 29.37 dB

26.75 dB DMPHN [20]





PSNR





19.70 dB

SRN [14]

23.73 dB

21.95 dB DeblurGAN-V2 [5]

21.96 dB DBGAN [23]





23.99 dB DMPHN [20]

25.03 dB Suin *et al*. [13]

26.60 dB MPRNet (Ours)

Figure 8: Image Deblurring results on the HIDE dataset [12].



SRN [14]



24.90 dB DeblurGAN-V2 [5]

PSNR



22.22 dB DMPHN [20]





31.23 dB

Figure 9: Image Deblurring results on RealBlur-J subset[11].



Blurry

27.72 dB DMPHN [20]

114

28.94 dB SRN [14]



28.87 dB DeblurGAN-V2 [5]



29.31 dB MPRNet (Ours)



PSNR Reference 24.64 dB Blurry









Blurry

37.29 dB SRN [14]



37.27 dB DeblurGAN-V2 [5]

PSNR



DMPHN [20]



36.54 dB

40.59 dB SRN [14]



Figure 11: Image Deblurring results on RealBlur-R subset[11].



Blurry

43.93 dB SRN [14]



40.90 dB DeblurGAN-V2 [5]

44.53 dB DMPHN [20]



MPRNet (Ours)

40.42 dB



PSNR

41.23 dB



Figure 12: Image Deblurring results on RealBlur-R subset[11].



Figure 13: Denoising comparisons on the SIDD dataset [1].

ogrammers, 3D Grap	grammers, 3D Grap	ogrammers, 3D Grap	ogrammers, 3D Grap ud Bevend is a come	ogrammers, 3D Gray	ogrammers, 3D Grap	ogrammers. 3D Gray
graphics that not of is used in computer the key mathematica uter graphics librari	graphics that not or s used in computer be key mathematica iter graphics librari	graphics that not or s used in computer he key mathematics mer graphics librari	graphics that not on a used in computer he key mathematics atter graphics library	r graphics that not or es used in computer the key mathematica after graphics library	graphics that not or is used in computer the key mathematics uter graphics library	graphics that not or is used in computer the key mathematics uter graphics library
PSNR	20.37 dB	35.13 dB	35.70 dB	35.46 dB	35.40 dB	35.95 dB
		1	1	1	1	1
PSNR	18.01 dB	34.04 dB	34.32 dB	34.36 dB	34.39 dB	34.64 dB
PSNR	17.69 dB	37.84 dB	38.43 dB	37.83 dB	38.42 dB	38.70 dB
PSNR	18.20 dB	31.17 dB	31.82 dB	31.87 dB	31.86 dB	32.26 dB
PSNR	27.67 dB	41.88 dB	42.62 dB	42.65 dB	42.61 dB	43.21 dB
PSNR	25.95 dB	40.94 dB	41.2 dB	41.24 dB	41.30 dB	41.48 dB
PSNR Reference	17.92 dB Noisy	30.98 dB AINDNet [4]	31.61 dB SADNet [2]	31.79 dB DANet [18]	31.66 dB CycleISP [19]	32.15 dB MPRNet (Ours)

Figure 14: Denoising comparisons on the SIDD dataset [1].



Figure 15: Denoising examples from the DND benchmark dataset [9].