

Supplemental Material: Unsupervised Degradation Representation Learning for Blind Super-Resolution

Longguang Wang¹, Yingqian Wang¹, Xiaoyu Dong^{2,3}, Qingyu Xu¹, Jungang Yang¹, Wei An¹, Yulan Guo^{1*}
¹National University of Defense Technology ²The University of Tokyo ³RIKEN AIP
 {wanglongguang15, yulan.guo}@nudt.edu.cn

Section I presents more analyses on our network. Section II includes additional quantitative and qualitative results achieved on different degradations.

I. Additional Analyses

I.I. Degradation-Aware Convolutions

Our degradation-aware convolution (DA convolution) exploits degradation information by dynamically generating convolutional kernels and channel-wise modulation coefficients from the input degradation representation. We conduct experiments to investigate these two key components.

Convolutional Kernels. We first visualize the convolutional kernels learned for degradations with different kernel widths in Fig. I. It can be observed that the convolutional kernels for $\sigma = 0.2/3.4$ have different patterns. This means that our DA convolution can adapt its convolutional kernels based on the degradation representation.

Channel-Wise Modulation Coefficients. Figure II further visualizes the modulation coefficients learned for degradations with different kernel widths. As we can see, our DA convolution learns to assign different channel-wise significance to the features according to the degradation representation.

By dynamically predicting convolutional kernels based on the degradation representations, our DA convolution

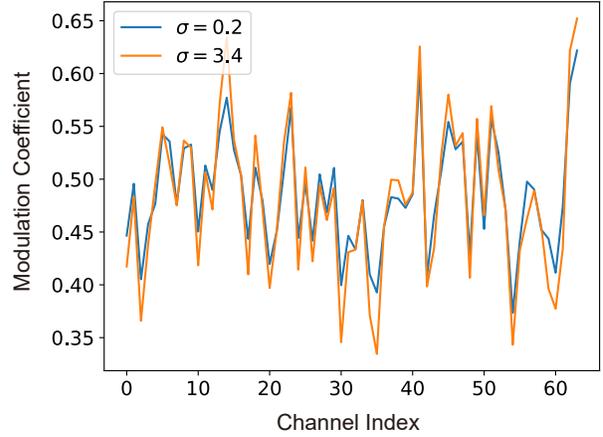


Figure II. Visualization of channel-wise modulation coefficients learned for different kernel widths.

has larger network capacity to handle various degradations. Moreover, learning channel-wise modulation coefficients enables our DA convolution to adapt the features according to the degradation information. Overall, DA convolution facilitates our DASR to better exploit degradation information for flexible adaption to different degradations, as demonstrated in Table 1.

I.II. Image Contents

Intuitively, the accuracy of degradation information provided by degradation representations is related to the image contents. Specifically, degradation representations learned from patches with rich edges and textures are expected to be better than those learned from patches with flat background. We conduct experiments to investigate the degradation representations learned from patches with different image contents. Given an LR image I , a patch was randomly cropped to extract degradation representation to super-resolve I . The relationship between the average gradient of the patch and the PSNR value is presented in Table I. With richer textures in the patch (*i.e.*, higher gradient), learned degradation representations provide more accurate degradation in-

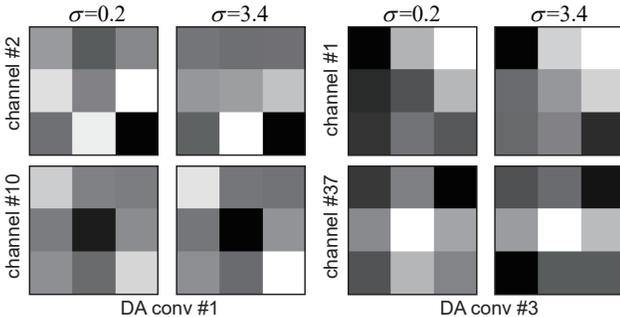


Figure I. Visualization of convolutional kernels learned for different kernel widths.

Table I. PSNR results achieved by our DASR using degradation representations extracted from different patches.

Gradient	8.12	8.75	9.23	9.55	10.14	10.28	10.88	11.58	12.21	12.37	Full Image
PSNR	25.30	25.33	25.39	25.43	25.50	25.52	25.58	25.58	25.59	25.59	25.52

Table III. PSNR results achieved on noise-free degradations with isotropic Gaussian kernels for $\times 4$ SR. Running time is averaged on Set14.

Method	Time	Set5			Set14			B100			Urban100		
Kernel Width		1.2	2.4	3.6	1.2	2.4	3.6	1.2	2.4	3.6	1.2	2.4	3.6
RCAN [7] + Correction Filter [2] + Predictor [1]	320ms	30.41	28.26	25.75	27.25	26.09	24.17	26.74	25.90	24.09	24.50	23.06	21.47
MZSR [5] + Predictor [1]	75ms	30.06	30.45	27.26	27.34	27.47	25.27	26.47	26.72	25.07	24.24	24.44	22.50
USRnet [6] + Predictor [1]	105ms	30.59	29.30	28.14	27.70	26.81	25.96	26.82	26.30	25.72	24.84	23.89	23.14
DAN [4]	190ms	32.15	31.86	30.44	27.48	27.26	26.51	26.69	26.35	25.98	26.18	25.29	24.82
DASR (Ours)	70ms	31.35	31.60	30.45	27.94	28.17	27.50	27.12	27.37	26.89	25.33	25.40	25.69

formation such that higher PSNR values are achieved by our DASR.

I.III. Unseen Degradations

We conduct experiments to test the generalization of our DASR on unseen degradations. As shown in Fig. III, our DASR achieves promising generalization and outperforms IKC on both seen and unseen degradations.

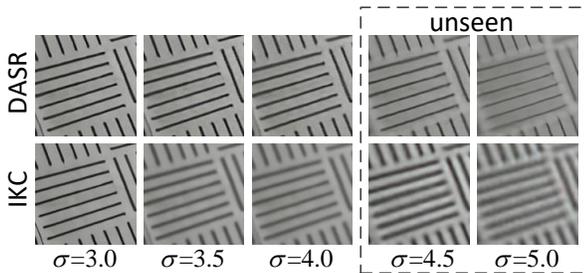


Figure III. Visual comparison on unseen degradations.

I.IV. Lightweight DASR

To compare our DASR to SRMDNF at a similar level of model size, we developed a lightweight version of DASR with fewer DA blocks and channels. As shown in Table II, our DASR-lite outperforms SRMDNF on different kernel widths with comparable model complexity.

Table II. PSNR results achieved on Set14 for $\times 4$ SR.

	Params.	Time	Kernel Width σ				
			0.2	1.0	1.8	2.6	3.4
SRMDNF+Predictor	1.9M	5ms	26.13	26.15	26.19	26.20	26.18
DASR-lite	2.1M	8ms	27.71	27.75	27.51	26.81	26.38

II. Additional Results

II.I. Degradations with s -Fold Downsampler

We re-trained our DASR on degradations with s -fold downsampler and compare its performance with correction filter [2], MZSR [5], USRnet [6] and DAN [4]. Since correction filter, MZSR and USRnet are non-blind SR methods, we used predictor sub-network in IKC [1] to estimate

degradations. Quantitative results are presented in Table III with visual results being provided in Fig. IV.

Quantitative Results. It can be observed from Table III that our DASR outperforms other methods for most scenarios. For example, our network produces notable performance improvements in terms of PSNR on B100 as compared to DAN (27.13/27.37/26.89 vs. 26.69/26.35/25.98). Since correction filter, MZSR and USRnet are sensitive to degradation estimation error, these methods suffer limited performance under blind settings. By iteratively estimating the degradation with the help of previous SR result, DAN produces improved performance. Compared to other methods, DASR benefits from our degradation representation learning scheme and achieves superior performance with better efficiency.

Qualitative Results. From Fig. IV we can see that our DASR achieves better visual quality as compared to other methods. Specifically, our network produces finer and clearer details while other methods suffer notable blurring artifacts, such as the text “ANTIQUES” in the third image.

II.II. Degradations with Bicubic Downsampler

We provide additional visual results achieved on noise-free and general degradations with bicubic downsampler in Figs. V and VI. We can see that our DASR consistently produces results with better perceptual quality and fewer artifacts.

II.III. Real Degradations

We further test our DASR on real images [3] with unknown degradations. Visual results are shown in Fig. VII. It can be observed that DASR achieves much better perceptual quality and recovers clearer details while other methods suffer notable blurring artifacts.

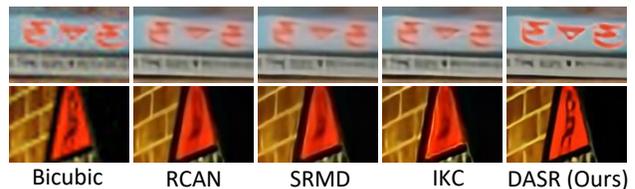


Figure VII. Visual comparison on real images.

Acknowledge

The authors would like to thank anonymous reviewers for their insightful suggestions. Xiaoyu Dong was supported by RIKEN Junior Research Associate Program.

References

- [1] Jinjin Gu, Hannan Lu, Wangmeng Zuo, and Chao Dong. Blind super-resolution with iterative kernel correction. In *CVPR*, 2019. [2](#)
- [2] Shady Abu Hussein, Tom Tirer, and Raja Giryes. Correction filter for single image super-resolution: Robustifying off-the-shelf deep super-resolvers. In *CVPR*, 2020. [2](#)
- [3] Andreas Lugmayr, Martin Danelljan, Radu Timofte, Manuel Fritsche, Shuhang Gu, Kuldeep Purohit, Praveen Kandula, Maitreya Suin, AN Rajagoapalan, Nam Hyung Joon, et al. Aim 2019 challenge on real-world image super-resolution: Methods and results. In *ICCVW*, pages 3575–3583. IEEE, 2019. [2](#)
- [4] Zhengxiong Luo, Yang Huang, Shang Li, Liang Wang, and Tieniu Tan. Unfolding the alternating optimization for blind super resolution. In *NeurIPS*, 2020. [2](#)
- [5] Jae Woong Soh, Sunwoo Cho, and Nam Ik Cho. Meta-transfer learning for zero-shot super-resolution. In *CVPR*, 2020. [2](#)
- [6] Kai Zhang, Luc Van Gool, and Radu Timofte. Deep unfolding network for image super-resolution. In *CVPR*, 2020. [2](#)
- [7] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *ECCV*, pages 1646–1654, 2018. [2](#)

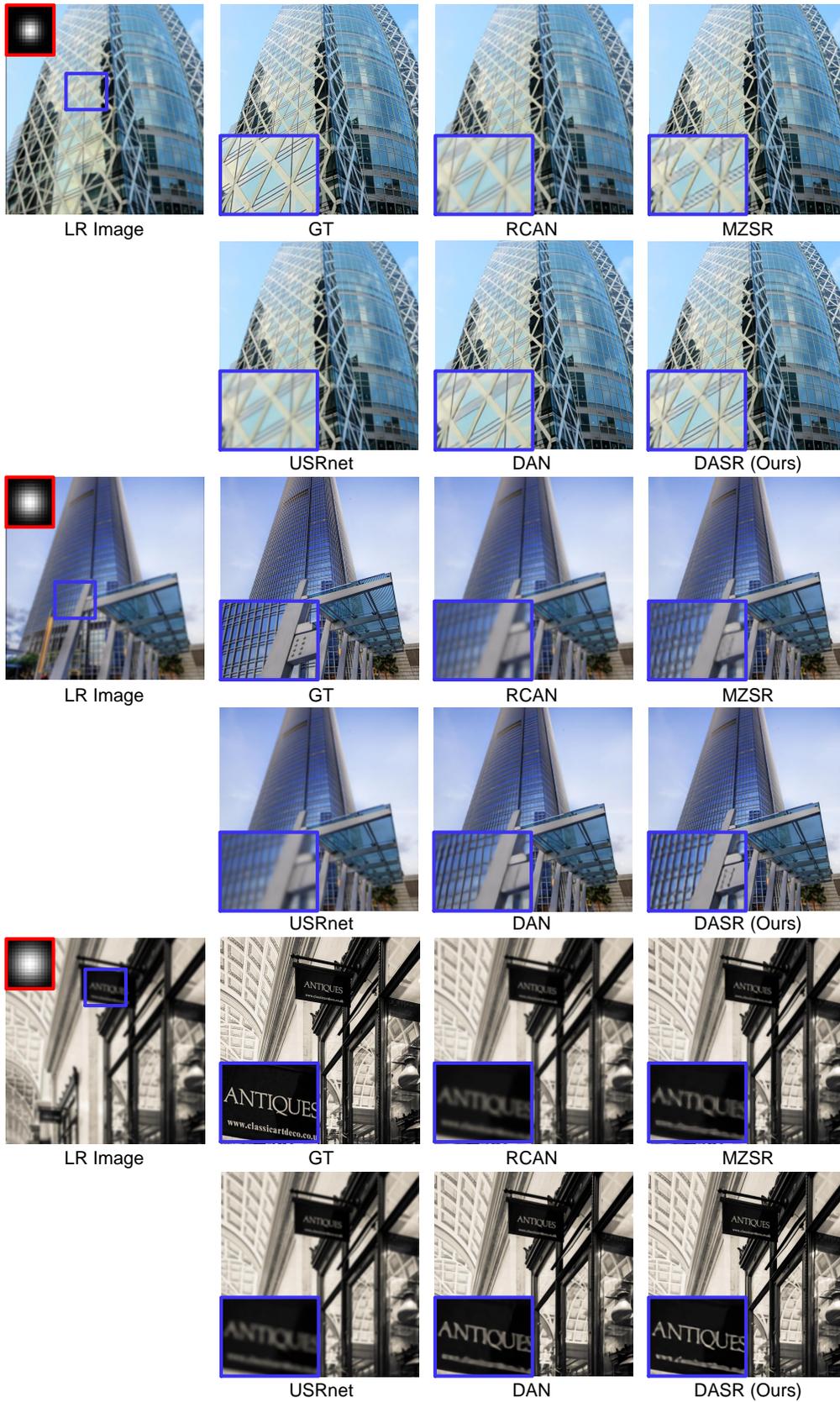


Figure IV. Visual comparison achieved on degradations with s -fold downsampler. The blur kernels are illustrated with red boxes.

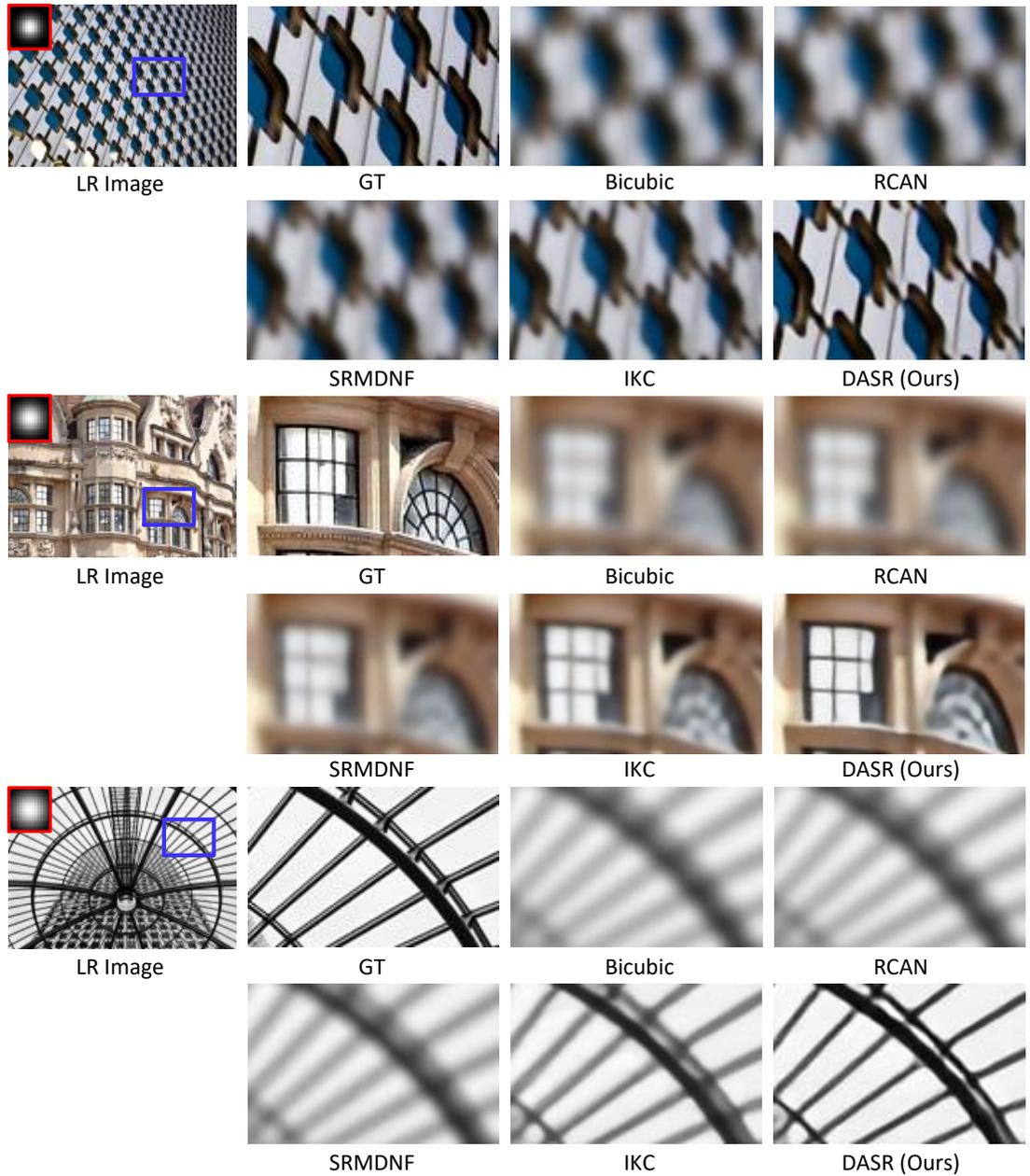


Figure V. Visual comparison achieved on noise-free degradations. The blur kernels are illustrated with red boxes.

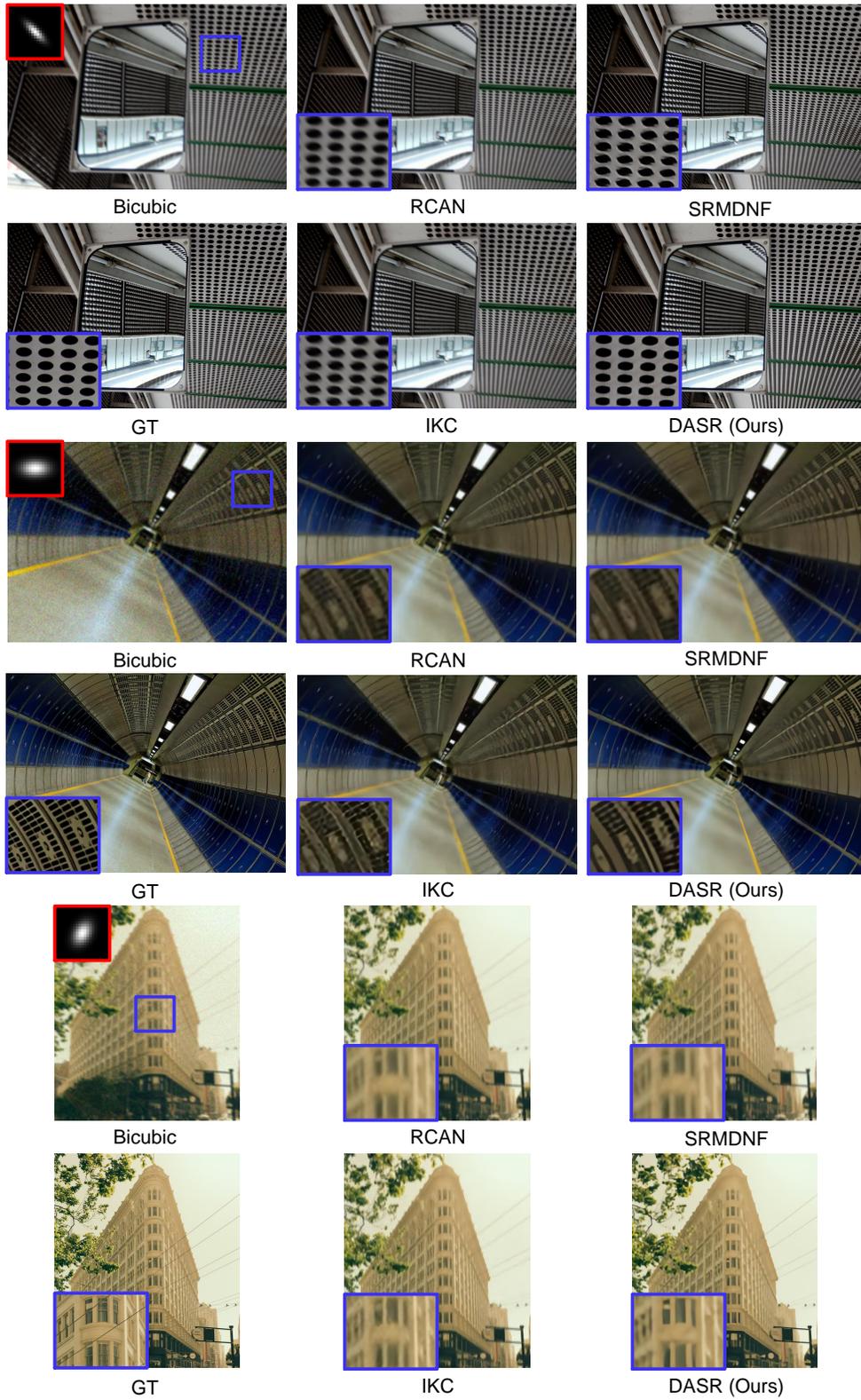


Figure VI. Visual comparison achieved on general degradations. Noise levels are set to 0, 5 and 10 for these three images, respectively.