From Face to Natural Image: Learning Real Degradation for Blind Image Super-Resolution

(Supplementary Materials)

In this supplemental material, we firstly analyze the different losses given in Section 3.4. We take the pixel MSE loss as *Baseline*, and gradually add feature MSE loss $\mathcal{L}^F_{mse}$, style loss $\mathcal{L}_{style}$, adversarial loss $\mathcal{L}_D$ and degradation consistent loss $\mathcal{L}_{cons}$. Their (PSNR, LPIPS, NIQE) on RealSR-Canon are shown in Table A. Note that the MSE loss on the first four variants performs on both face and natural images, including $\ell_{mse}(\hat{I}^L_f, I^{RealL}_f)$, $\ell_{mse}(\hat{I}^{SynL}_n, I^{SynL}_n)$ and $\ell_{mse}(\hat{I}^{SynL}_f, I^{SynL}_f)$, just without switching the degradation representation in Eqn. 13. One can see that pixel and feature MSE loss is crucial for PSNR, while the style and adversarial losses are important for visual quality (LPIPS, NIQE) at the cost of sightly decreasing PSNR. With $\mathcal{L}_{cons}$, our method achieves great improvement.

Table A: Quantitative comparison of different variants on RealSR-Canon.

Variants	PSNR↑	LPIPS↓	NIQE↓
Baseline	**25.74**	0.49	5.67
$Baseline + \mathcal{L}^F_{mse}$	25.73	0.44	5.43
$Baseline + \mathcal{L}^F_{mse} + \mathcal{L}_{style}$	25.60	0.40	5.39
$Baseline + \mathcal{L}^F_{mse} + \mathcal{L}_{style} + \mathcal{L}_D$	25.51	0.37	5.20
$Baseline + \mathcal{L}^F_{mse} + \mathcal{L}_{style} + \mathcal{L}_D + \mathcal{L}_{cons}$	25.57	**0.36**	**5.18**

Then, we analyze the effectiveness of different face restoration methods on the blind natural image super-resolution. Here we consider DFDNet [1], GFPGAN [2] and GPEN [4] used in this paper. Their performance on restoring our real-world LQ face test set is shown in the left column of Table B. We may find that GFPGAN performs on par with GPEN, but both of them are superior to DFDNet. By adopting DFDNet and GFPGAN in ReDegNet, their performance on natural image restoration show a positive correlation with these on face restoration (see right column of Table B). This may indicate that a better face restoration model can contribute better to natural image restoration. We analyze that both GFPGAN and GPEN are GAN-prior based methods, whose results are more photo-realistic and closer to the real-world HQ images, and thus can help our ReDegNet to learn more accurate real-world degradation.

Table B: Quantitative results of three face restoration methods.

FaceSR Methods	Face Restoration		RealSRSet	RealLQSet
	FID↓	NIQE↓	NIQE↓	NIQE↓
DFDNet [1]	75.63	5.52	5.14	5.20
GFPGAN [2]	63.70	**4.24**	**4.85**	4.95
GPEN [4]	**62.54**	4.27	**4.85**	**4.93**

Furthermore, we also analyze the benefit of real-world LQ face images and the synthetic LQ face images from FFHQ that bring to the final results. Here we consider three variants that are trained with 1) only synthetic LQ images from FFHQ, 2) only real-world LQ face images, and 3) the combination of them, respectively. Their NIQE results on real-world LQ natural images (RealSRSet [5] and our RealLQSet) are shown in Table C. We can see that the real-world face images contribute most to final results. By adding FFHQ to the real-world LQ images, the NIQE metric obtains 5.6% and 3.1% improvements, respectively. So combining the FFHQ and real-world LQ face images can not only extend the degradation space, but also improve the generalization to real-world scenarios.

Table C: Quantitative comparison on three types of training data.

Test Sets	Only FFHQ	Only Real-world	FFHQ and Real-world
RealSRSet	6.65	5.14	**4.85**
RealLQSet	6.82	5.09	**4.93**

Finally, we show more visual comparisons on different real-world LQ images in Figures A, B, C, D and E, which cover the most real-world LQ scenarios. We can observe that our method can generate more photo-realistic textures and performs superior against BSRGAN [5] and Real-ESRGAN [3] in most cases.

In addition, more samples of the synthetic LQ natural images with degradation transferring from the face pairs are shown in Figures F and G. These degradation cover different types, *i.e.*, halftone, noise, compression, blur, and even the severe degradation. One can see that our synthetic LQ natural images have nearly the same degradation with the given face pairs, indicating the effectiveness of our method in transferring the degradation from face to natural images.

Figure A: Close-up in the top-left is the face restoration result. Ours* represents our model fine-tuned with the degradation representation extracted from face region.

(a) Real-world LQ Image
(b) RealSR
(c) BSRGAN
(d) Real-ESRGAN
(e) Ours

(a) Real-world LQ Image
(b) RealSR
(c) BSRGAN
(d) Real-ESRGAN
(e) Ours

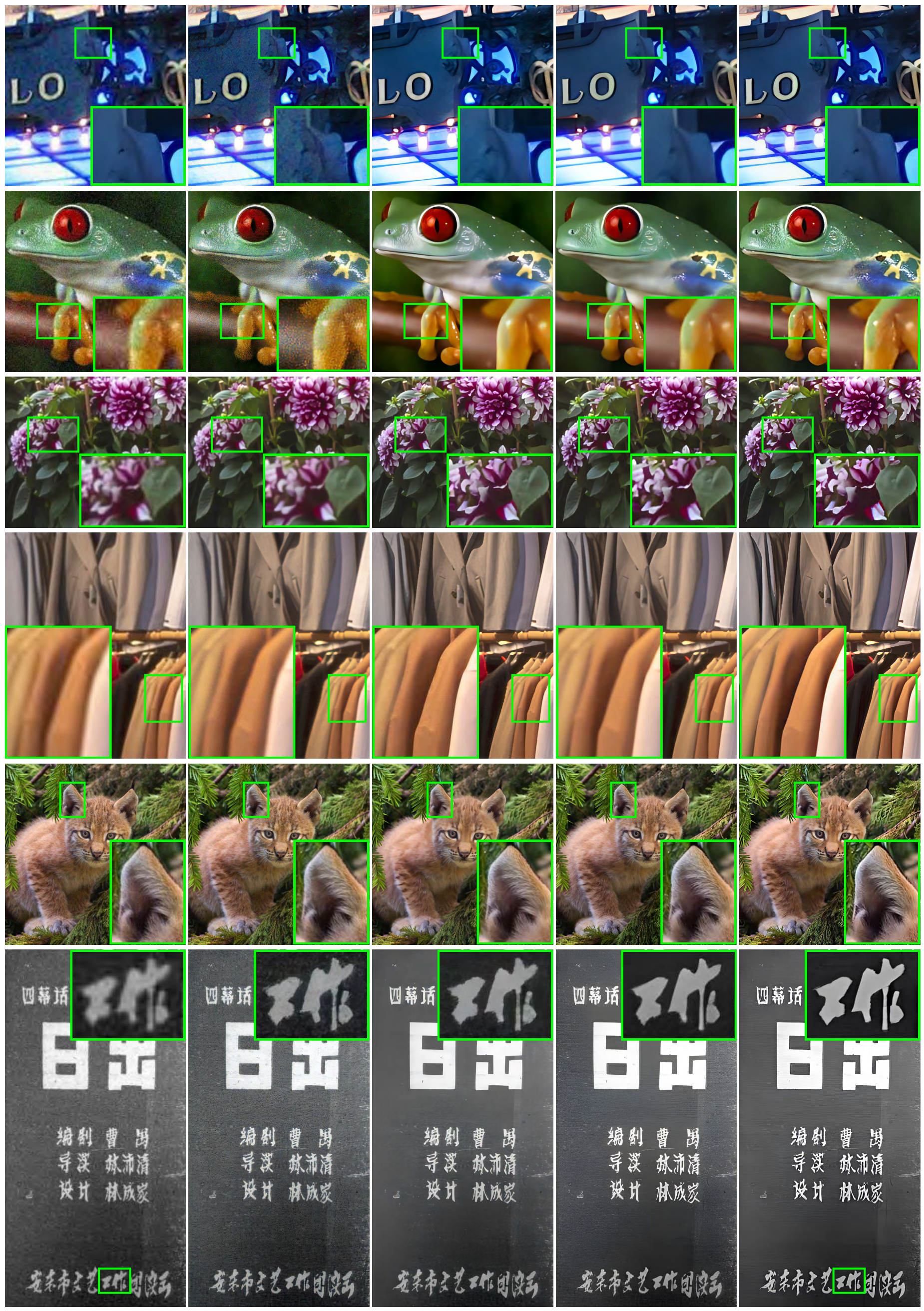
Figure B: More visual comparison of competing methods on rel-world LQ images.

(a) Real-world LQ Image
(b) RealSR
(c) BSRGAN
(d) Real-ESRGAN
(e) Ours

(a) Real-world LQ Image
(b) RealSR
(c) BSRGAN
(d) Real-ESRGAN
(e) Ours

Figure C: More visual comparison of competing methods on rel-world LQ images.

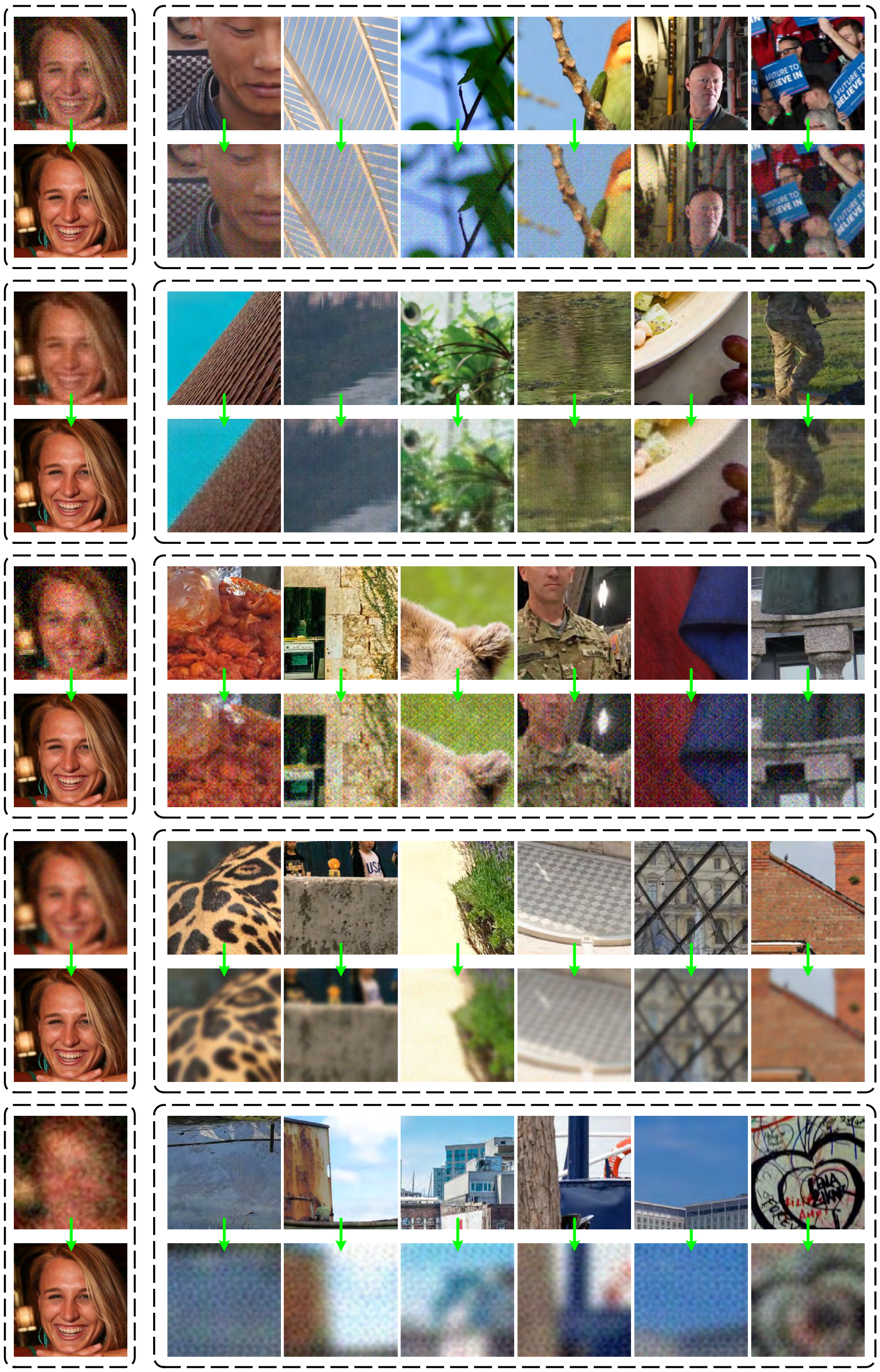
Figure D: Visual comparison of competing methods on restoring old video frames.



Real-world LQ Input RealSR BSRGAN Real-ESRGAN Ours

Figure E: Visual comparison of these competing methods on real-world LQ images.

Figure F: More examples of the synthetic LQ natural images in Figures 1 and 4 in the main text. Best view it by zooming in on the screen.



(a) Face Pairs (b) The Synthetic Natural Pairs

Figure G: More examples of the degradation transferring from face to natural pairs.

References

1. Li, X., Chen, C., Zhou, S., Lin, X., Zuo, W., Zhang, L.: Blind face restoration via deep multi-scale component dictionaries. In: ECCV (2020)
2. Wang, X., Li, Y., Zhang, H., Shan, Y.: Towards real-world blind face restoration with generative facial prior. In: CVPR (2021)
3. Wang, X., Xie, L., Dong, C., Shan, Y.: Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In: ICCVW (2021)
4. Yang, T., Ren, P., Xie, X., Zhang, L.: Gan prior embedded network for blind face restoration in the wild. In: CVPR (2021)
5. Zhang, K., Liang, J., Van Gool, L., Timofte, R.: Designing a practical degradation model for deep blind image super-resolution. In: ICCV (2021)