Supplementary File for
A Closer Look at Blind Super-Resolution: Degradation Models, Baselines, and Performance Upper Bounds

Wenlong Zhang[1], Guangyuan Shi[1], Yihao Liu[2 3], Chao Dong[2 4], Xiao-Ming Wu[1*]
[1]The HongKong Polytechnic University, [2]Shenzhen Institute of Advanced Technology, CAS
[3]University of Chinese Academy of Sciences, [4]Shanghai AI Laboratory
{wenlong.zhang, guang-yuan.shi}@connect.polyu.hk, {yh.liu4, chao.dong}@siat.ac.cn
xiao-ming.wu@polyu.edu.hk

Abstract

In this supplementary file, we first present more experimental details and more analysis on classical and practical blind SR. Then, we provide more quantitative and visual experimental results by comparing the proposed baselines with state-of-the-art methods.

1. A Comprehensive Analysis of Blind SR with Performance Upper Bounds

1.1. Datasets and Implementation Details

Datasets. Following existing blind SR methods [2, 6, 9, 10, 12, 13], we use DIV2K (800 images) [1] and Flickr2K (2650 images) [8] dataset for training. The training images are randomly cropped to 128×128 patches. Following the setting of IKC [2], DAN [6] and DASR [9], the HR images are blurred by isotropic Gaussian blur [0, 3.0] for the analysis of classical blind SR. The down-sampling adopts ×4 bicubic in the RealESRGAN version. For testing, we use benchmark datasets BSD100 [7] and Urban100 [3]. As the training dataset includes isotropic Gaussian blur [0.1, 3.0], we adopt a *Practical5* test dataset to quantitatively evaluate the blind SR network. *Practical5* consists of {bic, 0.6, 1.2, 1.8, 2.4}. The evaluation metric employs PSNR to evaluate the blind SR network. We adopt the same training settings for the analysis of the practical blind SR.

Training. Similar to FAIG [11], we adopt the representative RRDBNet with 5 blocks as the BSRNet named BSRNet-FAIG. The Adam [4] optimization method with $\beta_1 = 0.9$ and $\beta_1 = 0.99$ is used for training. The initial learning rate is set to 2×10^{-4}, which is reduced by a half for multi-step $[25 \times 10^4, 50 \times 10^4, 75 \times 10^4, 100 \times 10^4]$. A total of 100×10^4 iterations are executed by PyTorch. The loss function adopts L1 loss between SR results and HR images. We use the same setting for the analysis of the classical and practical blind SR.

Table 1. Average PSNR of BSRNet-FAIG with classical degradation models for ×4 super-resolution on BSD100 [7] and Urban100 [3].

Method	BSD100					Urban100				
	bic	0.6	1.2	1.8	2.4	bic	0.6	1.2	1.8	2.4
Bicubic	24.63	25.51	26.01	25.60	24.99	21.89	22.72	23.2	22.75	22.1
BSRNet-FAIG [11])	26.26	26.69	27.20	27.48	27.52	24.53	25.15	25.67	25.84	25.73
Upper bound	26.36	26.81	27.43	27.65	27.69	25.13	25.71	26.41	26.41	26.22

*Corresponding author

1.2. Analysis of Classical Blind SR

Table 1 shows the performance of BSRNet-FAIG on two test datasets - BSD100 [7] and Urban100 [3]. Firstly, we find that BSRNet-FAIG is only 0.2 dB lower than the upper bound on the BSD100 test dataset and 0.4 - 0.6 dB lower than that on Urban100, which is acceptable. Figure 1 shows that BSRNet-FAIG can achieve promising visual results although the quantitative results on the Urban100 test dataset is slightly lower. Both quantitative and qualitative results show that a blind SR network can well handle the classical blind SR problem.

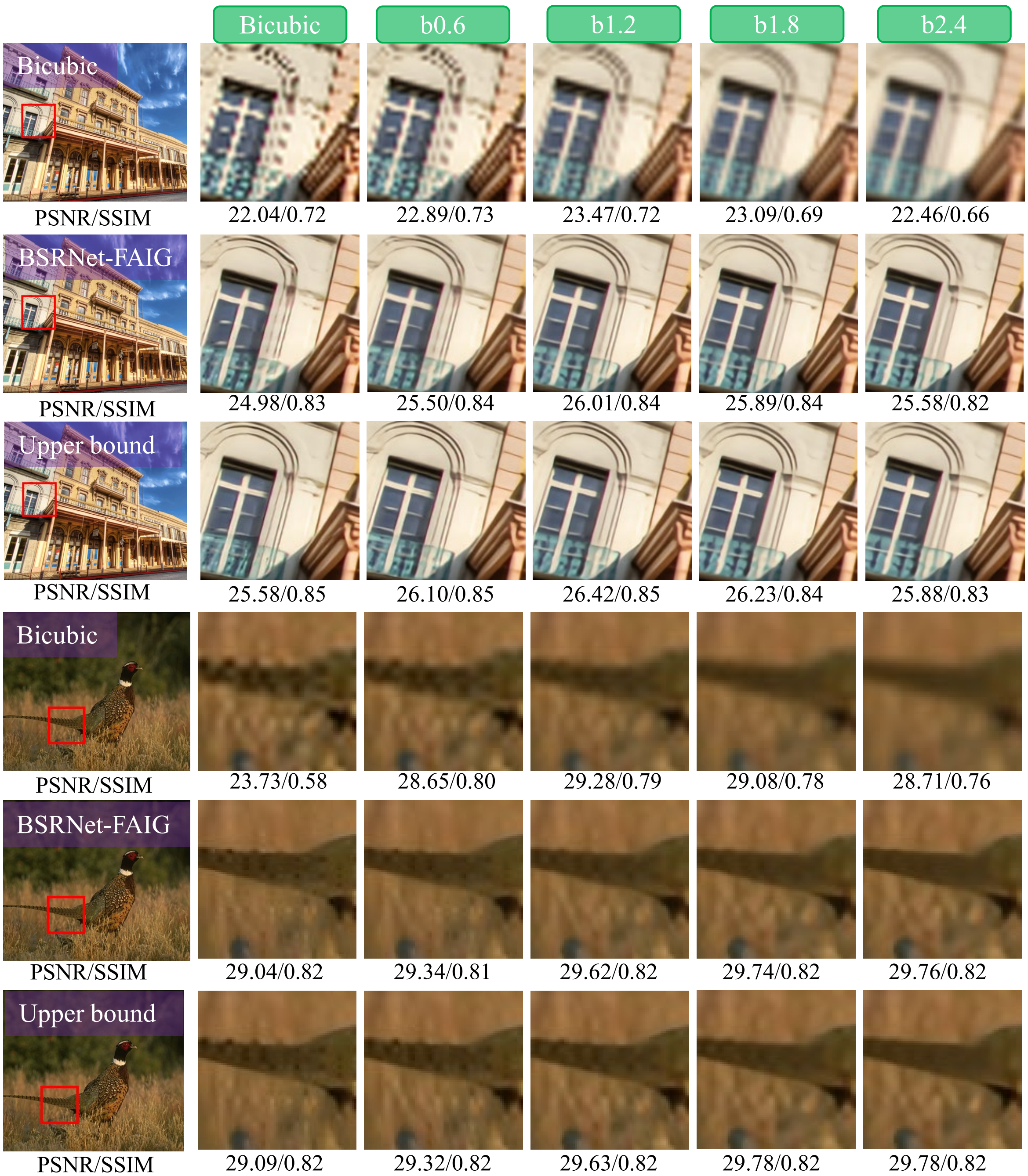

Figure 1. Visual comparisons of the BSRNet-FAIG with upper bound.

1.3. Analysis of Practical Blind SR

We provide more visual results to further compare BSRNet-PD/GD with the upper bounds. Figure 2 shows that BSRNet with our proposed GD model can achieve comparable visual results with the PD model on the most complicated degradation case b2.0n20j60 while performing much better on the other corner degradation cases. This shows that our proposed GD model can handle a broad set of degradation cases.

Figure 2. Visual comparisons of BSRNet-PD/GD with the upper bounds.

2. More Experimental Results

2.1. The Upper Bounds of SwinIR

In Table 2 of the main paper, we quantitatively evaluate the performance of RRDBNet [10, 12] and the performance upper bounds. For comparison, here we evaluate the performance of a stronger network SwinIR [5] and the corresponding performance upper bounds. The results on BSD100 [7] and Urban100 [3] are provided in Table 2. It can be observed that the average PSNR of the proposed baselines (RRDBNet-GD and SwinIR-GD) is much closer to the corresponding upper bound than RRDBNet and SwinIR respectively.

Table 2. Average PSNR of different methods for $\times 4$ super-resolution on BSD100 [7] and Urban100 [3]. Note that the parameters of all networks are consistent, thus the comparison is fair.

Dataset	Method	Degradation Types								
		bic	b2.0	n20	j60	b2.0n20	b2.0j60	n20j60	b2.0n20j60	Average
BSD100	RRDBNet	25.62	26.76	24.58	25.13	24.33	25.32	24.34	24.11	25.02
	RRDBNet-GD (ours)	26.25	27.31	25.31	25.23	24.95	25.32	24.38	24.07	25.35
	Upper bound (RRDBNet)	26.36	27.68	25.46	25.30	25.34	25.49	24.45	24.15	25.53
	SwinIR	25.84	27.05	24.77	25.27	24.48	25.44	24.44	24.18	25.18
	SwinIR-GD (ours)	26.61	27.58	25.64	25.30	25.30	25.39	24.44	24.14	25.55
	Upper bound (SwinIR)	27.10	27.83	25.76	25.37	25.56	25.55	24.50	24.20	25.73
Urban100	RRDBNet	23.53	24.46	22.89	23.28	22.48	23.17	22.75	22.24	23.10
	RRDBNet-GD (ours)	24.51	25.39	23.57	23.67	23.05	23.18	22.92	22.13	23.55
	Upper bound (RRDBNet)	25.13	26.38	23.91	23.97	23.56	23.62	23.18	22.44	24.02
	SwinIR	24.16	25.10	23.34	23.73	22.86	23.62	23.09	22.53	23.55
	SwinIR-GD (ours)	25.55	26.12	24.40	24.11	23.83	23.56	23.26	22.42	24.16
	Upper bound (SwinIR)	26.16	27.03	24.65	24.55	24.13	23.94	23.58	22.68	24.59

2.2. Light vs. Hard Degradation Models

We apply the proposed GD model to light and hard degradation scenarios to further explore its effectiveness of the proposed GD model. The light degradation model used the section 5.1 includes isotropic Gaussian blur [0.1, 3.0], additive Gaussian noise [1, 30] and JPEG [40, 95]. The order of light degradation model is set to $\{$blur, down-sampling, noise, JPEG$\}$. Based on the light degradation model, the hard degradation model further adds multiple degradation types (RealESRGAN version [10]) consists of multiple blurs: anisotropic Gaussian blur [0.1, 3.0], generalized isotropic/anisotropic Gaussian blur [0.1, 3.0] and plateau isotropic/an-isotropic Gaussian blur [0.1, 3.0]; multiple noises: additive grey/color Gaussian noise [1, 30] and Poisson grey/color noise [0.1, 3.0]. Table 3 shows that the performance of RRDBNet-GD-hard has a slight drop on the light cases, while it has a more significant improvement on the new cases. Figure 3 shows that the RRDBNet-GD-hard can generate the right textures on the an-isotropic degradation type, while RRDBNet-PD-light fails to generate promising results. These results demonstrate that the proposed GD model has the ability to handle complex degradation scenarios.

Table 3. Average PSNR of RRDBNet-GD with light and hard degradation models for $\times 4$ super-resolution on BSD100 [7] and Urban100 [3]. The PSNR distance is the distance between RRDBNet-GD-hard and RRDBNet-GD-light.

Dataset	Method	Degradation Types											
		bic	b2.0	n20	j60	b2.0n20	b2.0j60	n20j60	b2.0n20j60	color-n20	poisson-n20		
BSD100	Bicubic	24.63	25.40	21.56	24.06	21.90	24.65	21.22	21.72	22.90	22.82	25.51	25.50
	RRDBNet-GD-light	26.25	27.31	25.31	25.23	24.95	25.32	24.38	24.07	24.62	24.46	26.45	26.49
	RRDBNet-GD-hard	26.20	27.25	25.21	25.24	24.98	25.32	24.33	24.02	25.17	25.11	26.61	26.60
	PSNR distance	-0.05	-0.06	-0.10	0.01	0.03	0.00	-0.05	-0.05	0.55	0.65	0.16	0.11
Urban100	Bicubic	21.89	22.54	19.99	21.50	20.36	22.02	19.74	20.19	20.89	20.80	22.64	22.62
	RRDBNet-GD-light	24.51	25.39	23.57	23.66	23.05	23.18	22.92	22.13	23.14	22.98	24.06	24.20
	RRDBNet-GD-hard	24.45	25.20	23.44	23.66	22.91	23.16	22.89	22.06	23.70	23.62	24.29	24.33
	PSNR distance	-0.06	-0.19	-0.13	0.00	-0.14	-0.02	-0.03	-0.07	0.56	0.64	0.23	0.13

2.3. More Visual Results

In this section, we provide additional visual results ($\times 4$ scale) to demonstrate the high effectiveness of our proposed gated degradation model. We compare the proposed baselines RRDDBNet-GD and SwinIR-GD with state-of-the-art blind SR networks. Figure 4 shows our model can achieve much better results than existing methods on various corner degradation cases, such as b2.0 and b2.0n20 while achieving comparable results on the complicated degradation b2.0n20j60. The same observation can be made for GAN-based methods, as shown in Figure 5. These visual results demonstrate the clear advantage of the proposed baselines in dealing with different degradation cases.

Figure 3. Visual results comparisons of RRDBNet-GD on light and hard degradation model. Please zoom in for a better view.

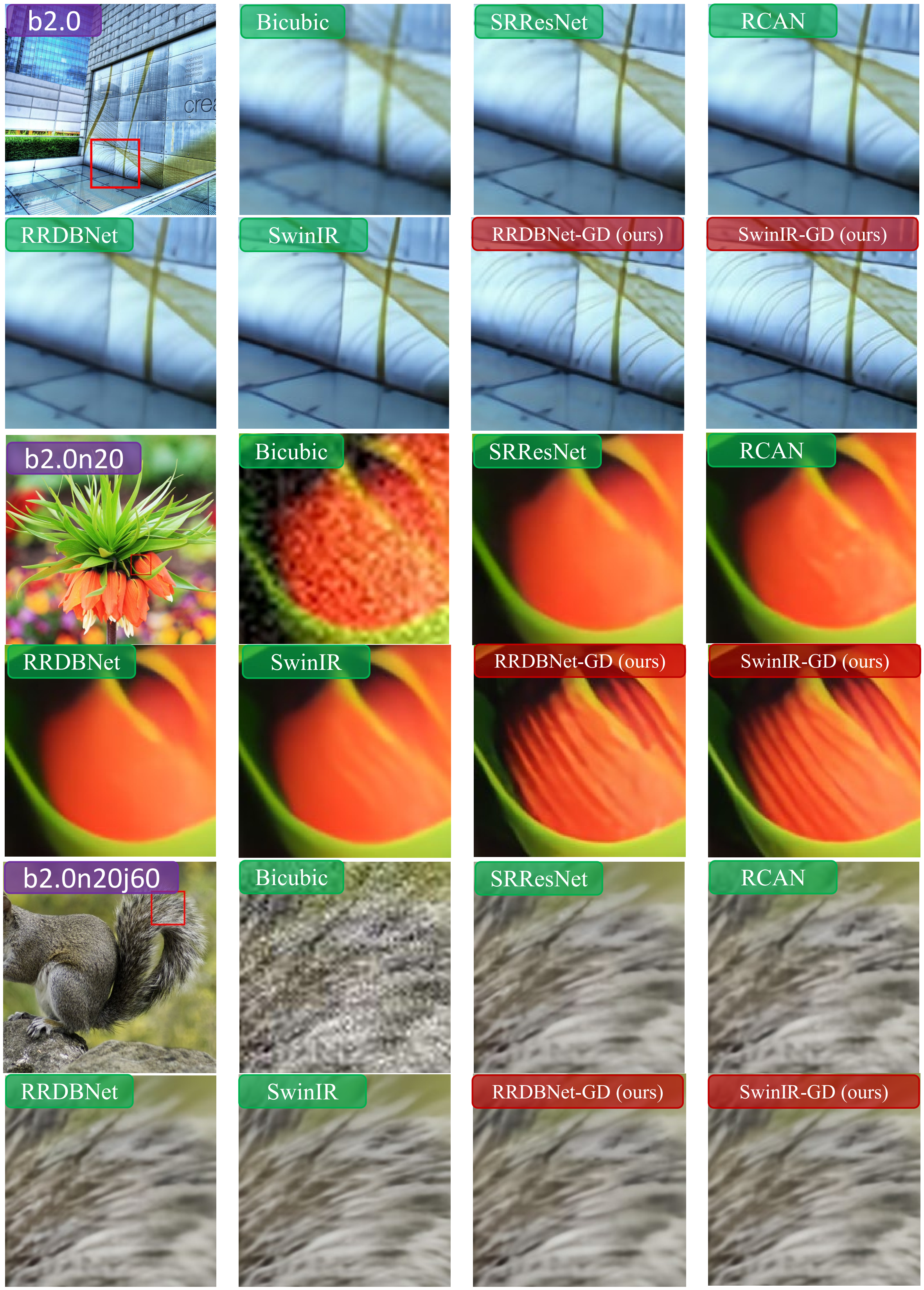

Figure 4. Visual results comparisons of our model with existing works on $\times 4$ super-resolution. Please zoom in for a better view.

Figure 5. Visual results comparisons of our model with existing GAN-based works on $\times 4$ super-resolution. Please zoom in for a better view.

References

[1] Eirikur Agustsson and Radu Timofte. Ntire 2017 challenge on single image super-resolution: Dataset and study. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, volume 3, page 2, 2017. 1

[2] Jinjin Gu, Hannan Lu, Wangmeng Zuo, and Chao Dong. Blind super-resolution with iterative kernel correction. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 1604–1613, 2019. 1

[3] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 5197–5206, 2015. 1, 2, 4

[4] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 1

[5] Jingyun Liang, Jiezhang Cao, Guolei Sun, Kai Zhang, Luc Van Gool, and Radu Timofte. Swinir: Image restoration using swin transformer. In *IEEE International Conference on Computer Vision Workshops*, 2021. 4

[6] Zhengxiong Luo, Yan Huang, Shang Li, Liang Wang, and Tieniu Tan. Unfolding the alternating optimization for blind super resolution. *arXiv preprint arXiv:2010.02631*, 2020. 1

[7] David Martin, Charless Fowlkes, Doron Tal, and Jitendra Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Computer Vision, 2001. ICCV 2001. Proceedings. Eighth IEEE International Conference on*, volume 2, pages 416–423. IEEE, 2001. 1, 2, 4

[8] Radu Timofte, Eirikur Agustsson, Luc Van Gool, Ming-Hsuan Yang, and Lei Zhang. Ntire 2017 challenge on single image super-resolution: Methods and results. In *Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, pages 114–125, 2017. 1

[9] Longguang Wang, Yingqian Wang, Xiaoyu Dong, Qingyu Xu, Jungang Yang, Wei An, and Yulan Guo. Unsupervised degradation representation learning for blind super-resolution. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 10581–10590, 2021. 1

[10] Xintao Wang, Liangbin Xie, Chao Dong, and Ying Shan. Real-esrgan: Training real-world blind super-resolution with pure synthetic data. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 1905–1914, 2021. 1, 4

[11] Liangbin Xie, Xintao Wang, Chao Dong, Zhongang Qi, and Ying Shan. Finding discriminative filters for specific degradations in blind super-resolution. *arXiv preprint arXiv:2108.01070*, 2021. 1

[12] Kai Zhang, Jingyun Liang, Luc Van Gool, and Radu Timofte. Designing a practical degradation model for deep blind image super-resolution. *arXiv preprint arXiv:2103.14006*, 2021. 1, 4

[13] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Learning a single convolutional super-resolution network for multiple degradations. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3262–3271, 2018. 1