Perception-Oriented Single Image Super-Resolution using Optimal Objective Estimation - Supplementary Material

Seung Ho Park[1,2], Young Su Moon[2], Nam Ik Cho[1,3]
[1]Department of Electrical and Computer Engineering, INMC, Seoul National University, Korea
[2]Visual Display Division, Samsung Electronics, Korea
[3]IPAI, Seoul National University, Korea

1. More Visual Comparisons of SISR 4×

We provide more visual comparisons of 4× SISR between our proposed SROOE and other state-of-the-art methods including a distortion-oriented method, RRDB [7], and perception-oriented methods, such as SRGAN [2], ESRGAN [7], SFTGAN [6], RankSRGAN [8], SRFlow [3], SPSR [4], in Fig. from 1 to 6. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

2. Performance Comparison of SISR 8×

Quantitative Comparison. Table 1 shows the quantitative performance comparison for 8× SISR. We compared it with a distortion-oriented method, RRDB [7], and perception-oriented methods, such as ESRGAN [7], SRFlow [3], and FxSR [5]. The table shows that our method yields the best results among the perception-oriented methods on all datasets, in terms of PSNR, SSIM, LPIPS and DISTS [1]. However, in LR-PSNR, unlike SR 4×, less than 45 dB indicates that LR consistency needs improvement.

Qualitative Comparison. Fig. 7 and Fig. 8 show visual comparisons, where we can observe that SROOE generates more accurate and cleaner structures with less unnatural details. Among the four perception-oriented SR methods, the best performances are highlighted in red.

3. Performance comparison of SR results of ESRGAN models with different weight vectors for perceptual loss

Table 2 shows the data befor normalization. Min-max normalization used for Table 1 in the main paper used is as follows:

$$x' = (x - \min(x))/(\max(x) - \min(x)). \quad (1)$$

References

[1] Keyan Ding, Kede Ma, Shiqi Wang, and Eero P. Simoncelli. Image quality assessment: Unifying structure and texture similarity. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, pages 1–1, 2020. 1

[2] Christian Ledig, Lucas Theis, Ferenc Huszár, Jose Caballero, Andrew Cunningham, Alejandro Acosta, Andrew Aitken, Alykhan Tejani, Johannes Totz, Zehan Wang, et al. Photo-realistic single image super-resolution using a generative adversarial network. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 4681–4690, 2017. 1

[3] Andreas Lugmayr, Martin Danelljan, Luc Van Gool, and Radu Timofte. Srflow: Learning the super-resolution space with normalizing flow. In *European Conference on Computer Vision*, pages 715–732. Springer, 2020. 1, 8

[4] Cheng Ma, Yongming Rao, Yean Cheng, Ce Chen, Jiwen Lu, and Jie Zhou. Structure-preserving super resolution with gradient guidance. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 7769–7778, 2020. 1

[5] Seung Ho Park, Young Su Moon, and Nam Ik Cho. Flexible style image super-resolution using conditional objective. *IEEE Access*, 10:9774–9792, 2022. 1, 8

[6] Xintao Wang, Ke Yu, Chao Dong, and Chen Change Loy. Recovering realistic texture in image super-resolution by deep spatial feature transform. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 606–615, 2018. 1

[7] Xintao Wang, Ke Yu, Shixiang Wu, Jinjin Gu, Yihao Liu, Chao Dong, Yu Qiao, and Chen Change Loy. Esrgan: Enhanced super-resolution generative adversarial networks. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 0–0, 2018. 1, 8

[8] Wenlong Zhang, Yihao Liu, Chao Dong, and Yu Qiao. Ranksrgan: Generative adversarial networks with ranker for image super-resolution. In *2019 IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 3096–3105, 2019. 1

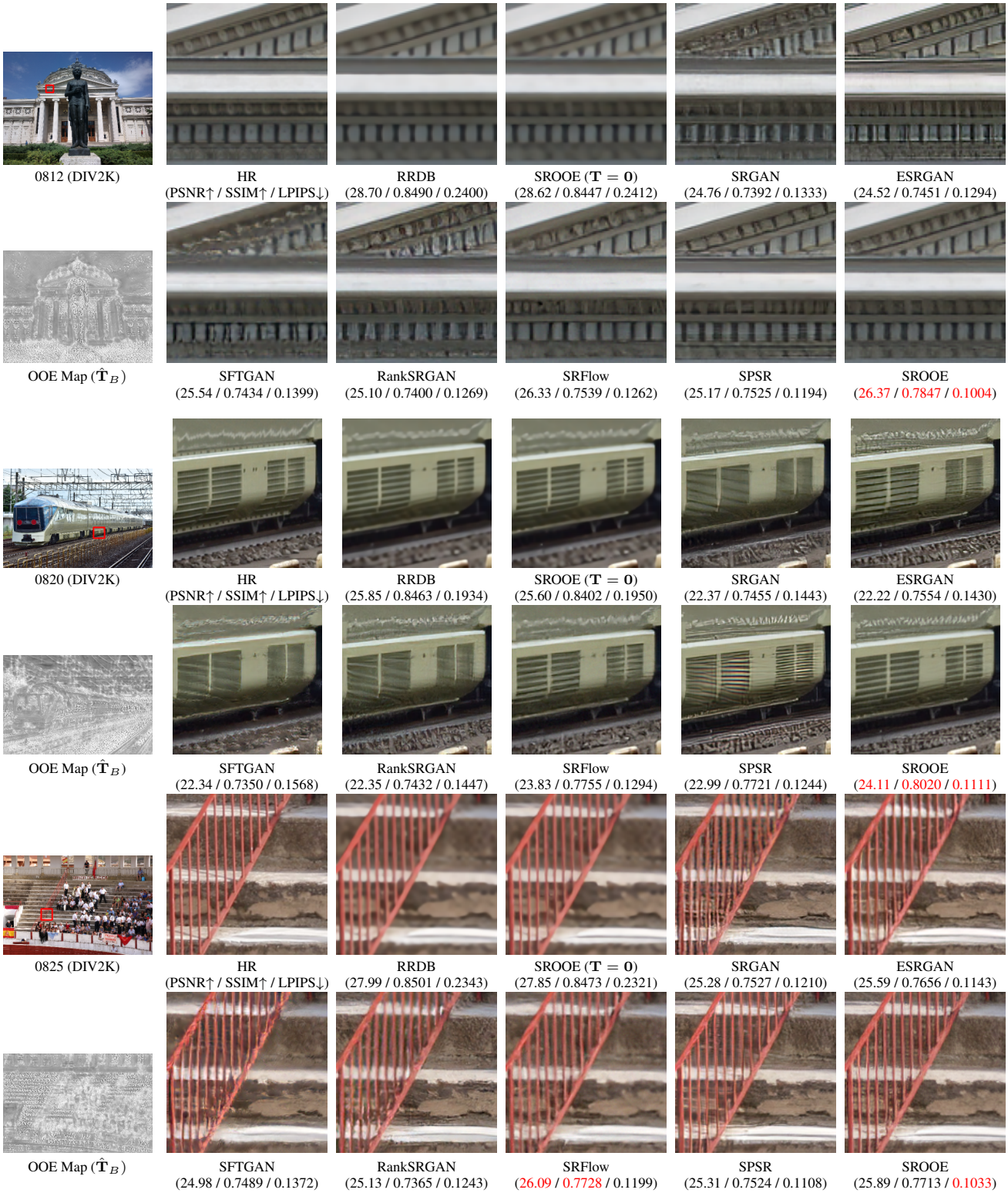

Figure 1. Visual comparison with state-of-the-art methods for $4\times$ SR results. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

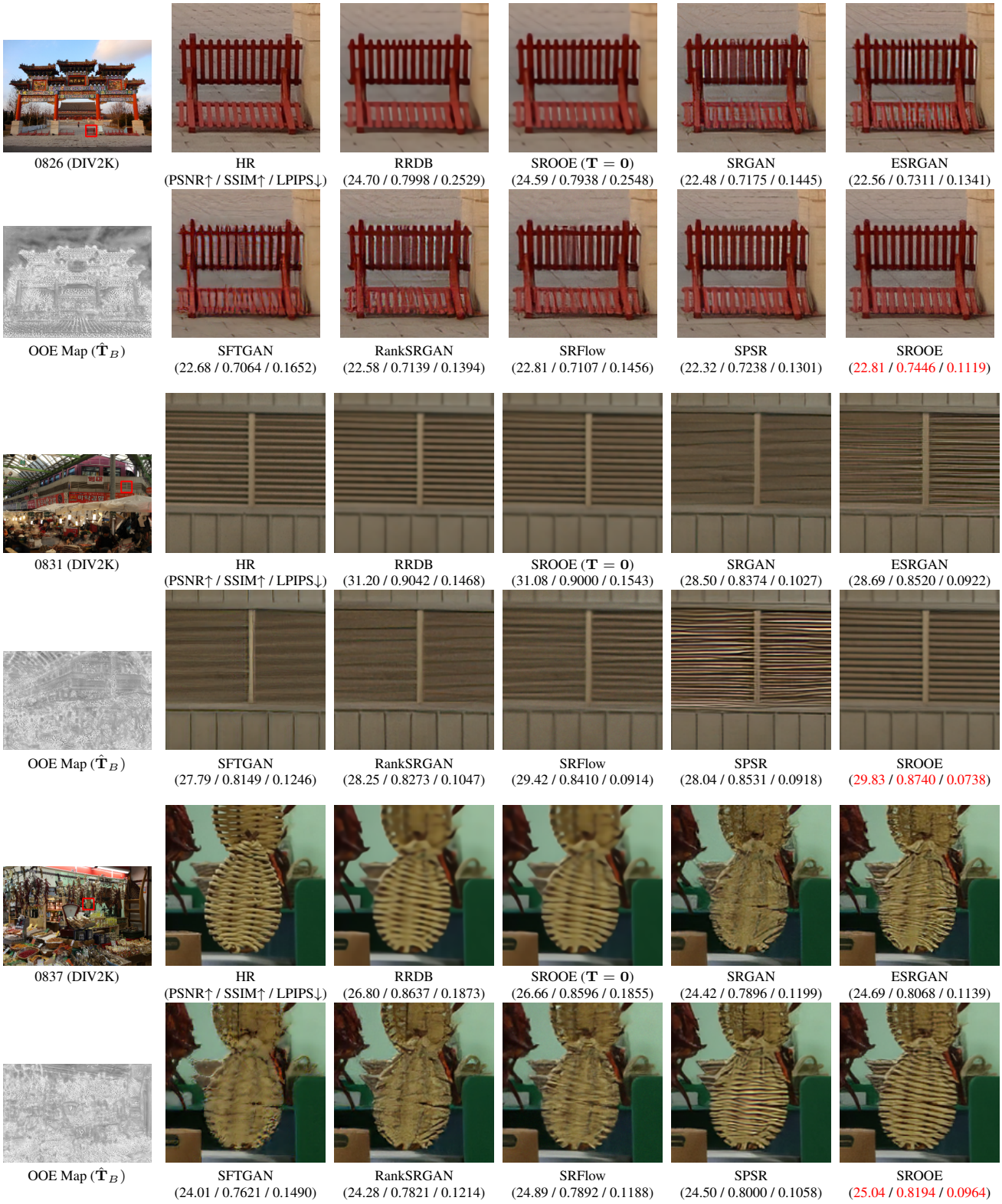

Figure 2. Visual comparison with state-of-the-art methods for $4\times$ SR results. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

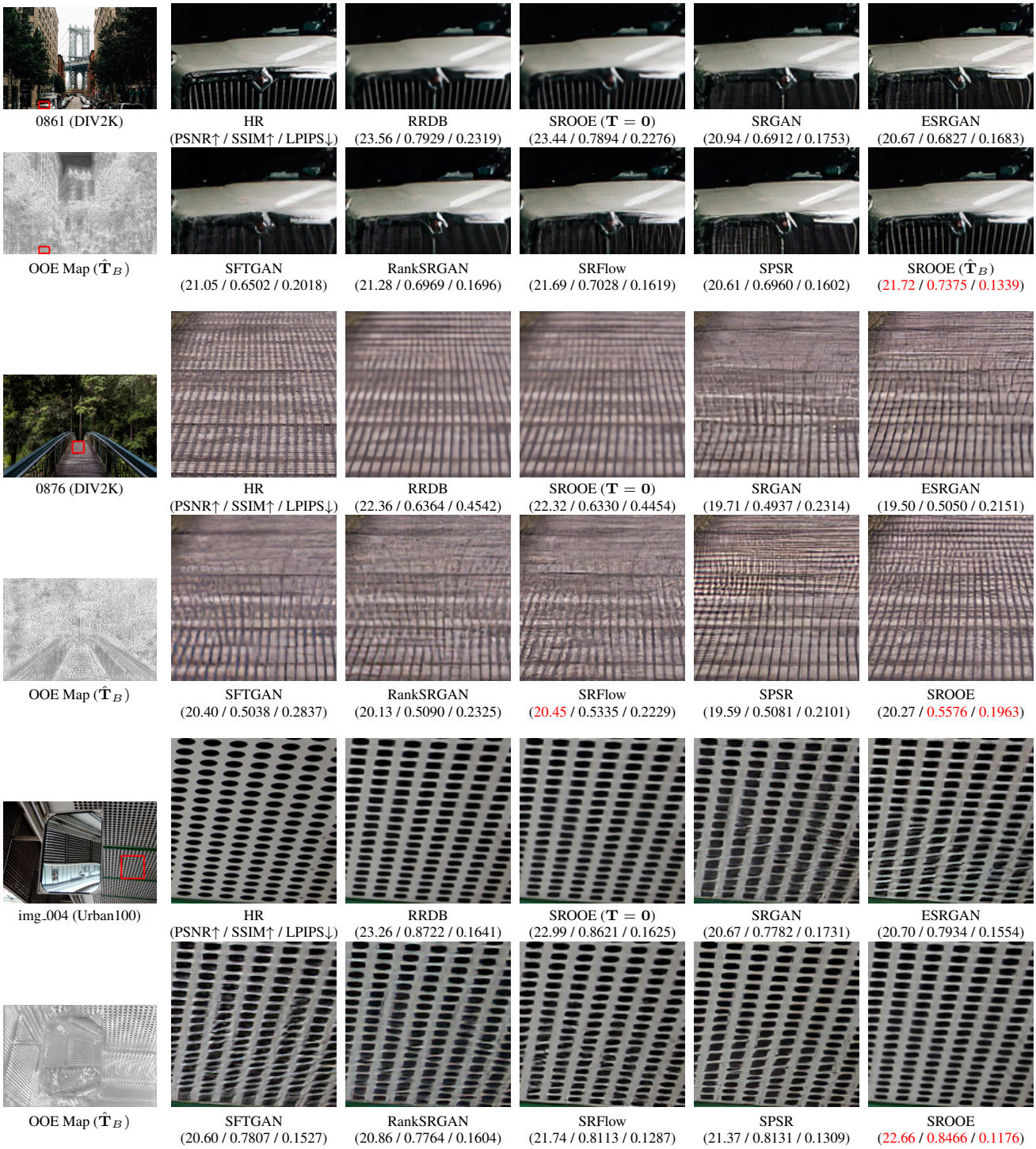

Figure 3. Visual comparison with state-of-the-art methods for $4\times$ SR results. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

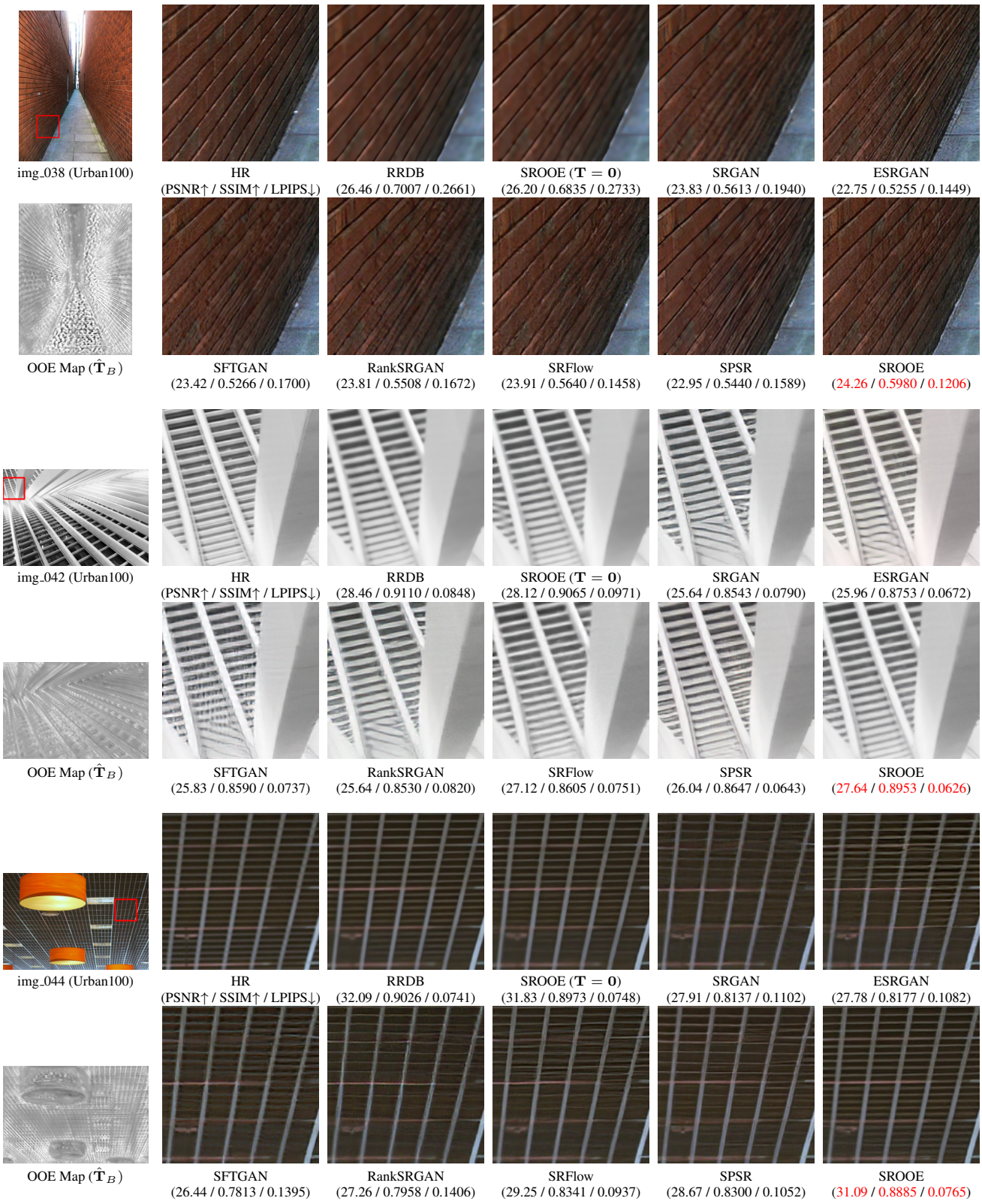

Figure 4. Visual comparison with state-of-the-art methods for $4\times$ SR results. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

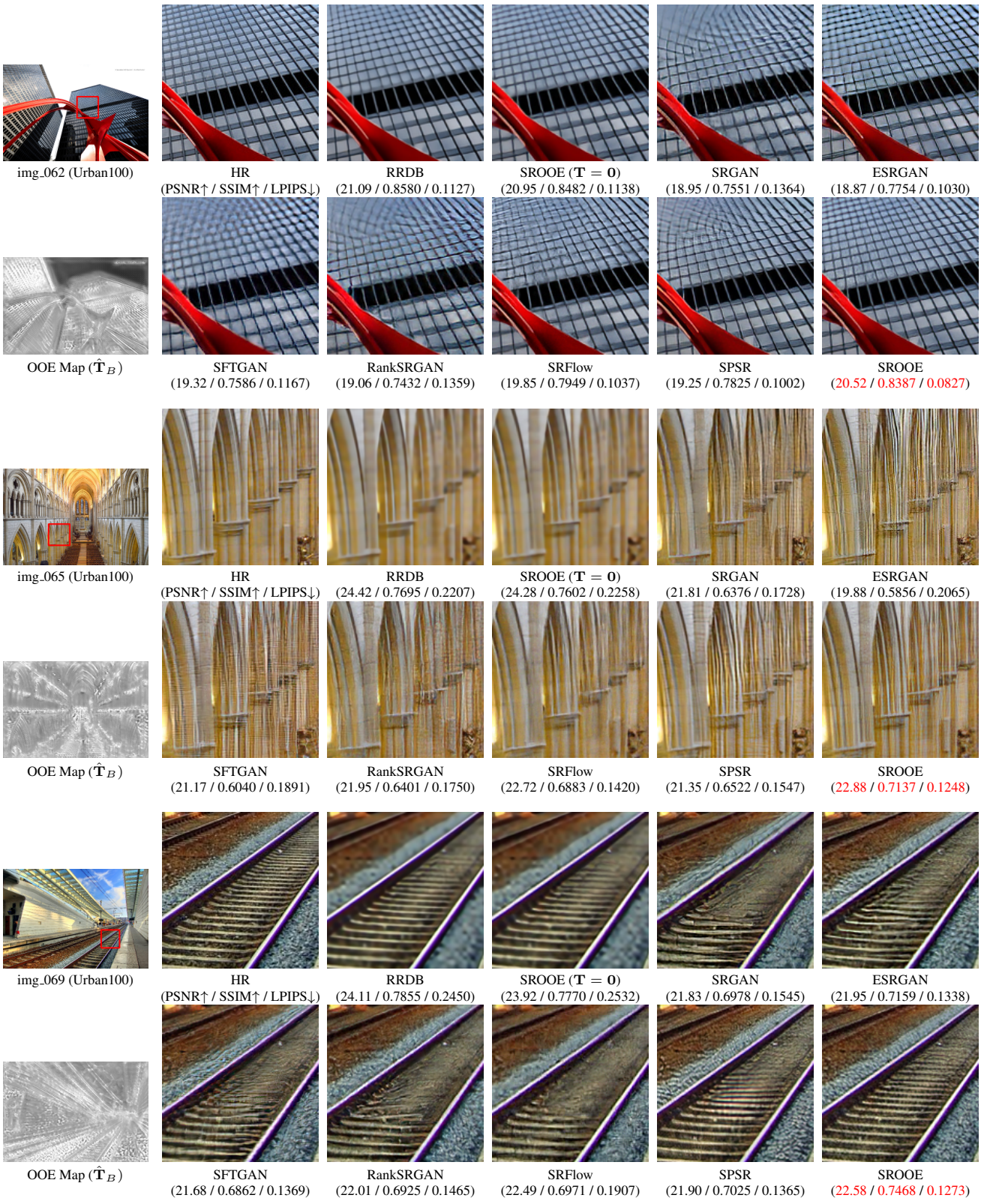

Figure 5. Visual comparison with state-of-the-art methods for $4\times$ SR results. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

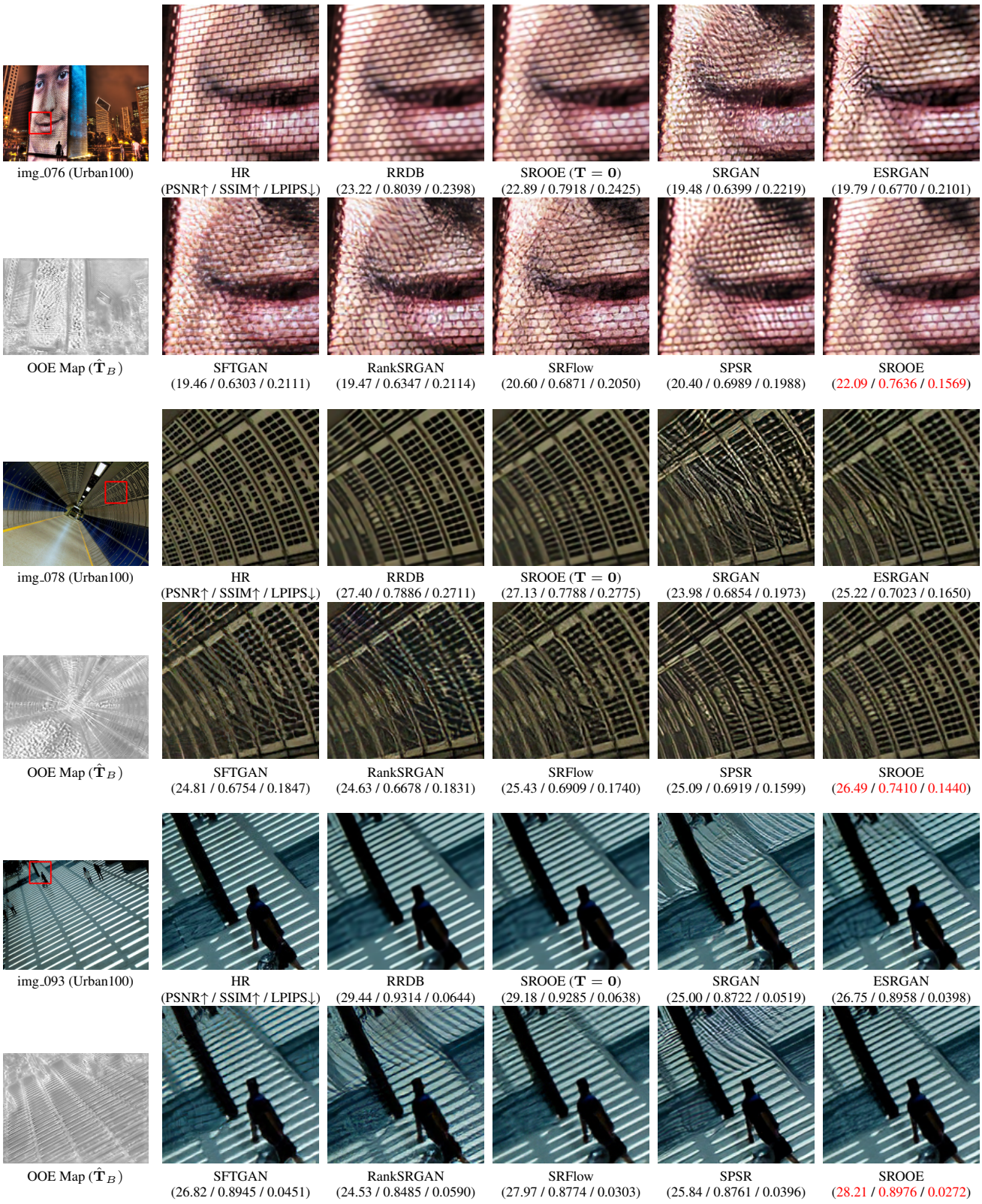

Figure 6. Visual comparison with state-of-the-art methods for $4\times$ SR results. Among the seven perception-oriented SR methods, the best performances are highlighted in red.

Table 1. Comparison with state-of-the-art SR 8× methods on benchmarks. The 1st and the 2nd best performances for each group are highlighted in red and blue, respectively. LR-PSNR values greater than 45dB are underlined.

		Distortion-oriented SR				Perception-oriented SR			
Model		RRDB [7]	SRFlow ($\tau = \mathbf{0}$) [3]	FxSR ($\mathbf{T} = \mathbf{0}$) [5]	SROOE ($\mathbf{T} = \mathbf{0}$)	ESRGAN [7]	SRFlow ($\tau = \mathbf{0.9}$) [3]	FxSR ($\mathbf{T} = \mathbf{0.8}$) [5]	SROOE ($\hat{\mathbf{T}}_B$)
Training dataset		DIV2K	DF2K	DIV2K	DF2K	DF2K+OST	DIV2K	DIV2K	DF2K
BSD100	PSNR↑	23.57	23.37	23.60	23.61	21.11	21.66	21.93	22.23
	SSIM↑	0.5708	0.5428	0.5728	0.5718	0.4658	0.4632	0.5039	0.5164
	LPIPS↓	0.5548	0.5303	0.5079	0.5378	0.3103	0.3238	0.3129	0.2996
	DISTS↓	0.2941	0.3183	0.2753	0.2861	0.1881	0.2068	0.1972	0.1891
	LR-PSNR↑	46.23	52.39	47.12	45.34	33.81	51.09	42.41	43.57
Urban100	PSNR↑	21.10	20.91	21.37	21.37	18.82	19.41	19.88	20.17
	SSIM↑	0.5999	0.5735	0.6122	0.6114	0.5005	0.5013	0.5510	0.5696
	LPIPS↓	0.4082	0.3735	0.3454	0.3720	0.2701	0.2965	0.2461	0.2449
	DISTS↓	0.2599	0.2468	0.2193	0.2344	0.1673	0.2039	0.1695	0.1558
	LR-PSNR↑	42.98	50.90	44.28	42.91	29.93	48.94	40.62	42.45
Manga109	PSNR↑	23.05	22.63	23.23	23.26	20.78	20.65	21.86	22.10
	SSIM↑	0.7516	0.7183	0.7527	0.7545	0.6585	0.6439	0.7001	0.7094
	LPIPS↓	0.2457	0.2419	0.2097	0.2242	0.1823	0.2188	0.1640	0.1578
	DISTS↓	0.1495	0.1627	0.1272	0.1339	0.0999	0.1337	0.0980	0.0862
	LR-PSNR↑	43.03	49.16	42.78	42.40	32.09	45.56	40.89	42.45
DIV2K	PSNR↑	25.50	25.09	25.60	25.61	22.58	23.04	23.56	23.92
	SSIM↑	0.6960	0.6589	0.6989	0.6978	0.5891	0.5728	0.6241	0.6377
	LPIPS↓	0.4227	0.4033	0.3857	0.4071	0.2421	0.2719	0.2402	0.2238
	DISTS↓	0.2185	0.2342	0.1953	0.2073	0.1045	0.1386	0.1190	0.1075
	LR-PSNR↑	46.35	51.28	46.96	45.68	33.15	50.26	42.66	44.08

Table 2. Performance comparison of SR results of ESRGAN models with different weight vectors for perceptual loss. Among the objectives in Sets A and B, except for $\boldsymbol{\lambda_0}$, the 1st and the 2nd best performances for each column are highlighted in red and blue.

Set	objec -tive	$\boldsymbol{\lambda_{per}}$ [λ_{V12},	λ_{V22},	λ_{V34},	λ_{V44},	λ_{V54}]	L_{per_l} L_{V12}	L_{V22}	L_{V34}	L_{V44}	L_{V54}	Metric PSNR	LPIPS
	$\boldsymbol{\lambda_0}$	[0.0,	0.0,	0.0,	0.0,	0.0]	0.6459	1.8793	3.9235	2.2091	1.2767	25.48	0.1960
A	$\boldsymbol{\lambda_1}$	[1.0,	0.0,	0.0,	0.0,	0.0]	0.8320	2.0275	3.8340	1.9804	1.0480	23.95	0.1124
	$\boldsymbol{\lambda_2}$	[0.0,	1.0,	0.0,	0.0,	0.0]	0.8322	1.9523	3.6789	1.9200	1.0171	23.84	0.1125
	$\boldsymbol{\lambda_3}$	[0.0,	0.0,	1.0,	0.0,	0.0]	0.8565	2.0222	3.5480	1.8452	0.9841	23.66	0.1124
	$\boldsymbol{\lambda_4}$	[0.0,	0.0,	0.0,	1.0,	0.0]	0.8844	2.0977	3.7413	1.8777	0.9890	23.28	0.1158
	$\boldsymbol{\lambda_5}$	[0.0,	0.0,	0.0,	0.0,	1.0]	0.9051	2.1593	3.8985	1.9565	1.0085	23.00	0.1232
B	$\boldsymbol{\lambda_1}$	[1.0,	0.0,	0.0,	0.0,	0.0]	0.8320	2.0275	3.8340	1.9804	1.0480	23.95	0.1124
	$\boldsymbol{\lambda_{1\text{-}2}}$	[1/2,	1/2,	0.0,	0.0,	0.0]	0.8112	1.9432	3.6740	1.9114	1.0125	24.08	0.1075
	$\boldsymbol{\lambda_{1\text{-}3}}$	[1/3,	1/3,	1/3,	0.0,	0.0]	0.8468	1.9964	3.5846	1.8532	0.9821	23.81	0.1112
	$\boldsymbol{\lambda_{1\text{-}4}}$	[1/4,	1/4,	1/4,	1/4,	0.0]	0.8474	2.0085	3.6174	1.8489	0.9765	23.68	0.1110
	$\boldsymbol{\lambda_{1\text{-}5}}$	[1/5,	1/5,	1/5,	1/5,	1/5]	0.8439	2.0085	3.6174	1.8515	0.9821	23.74	0.1114

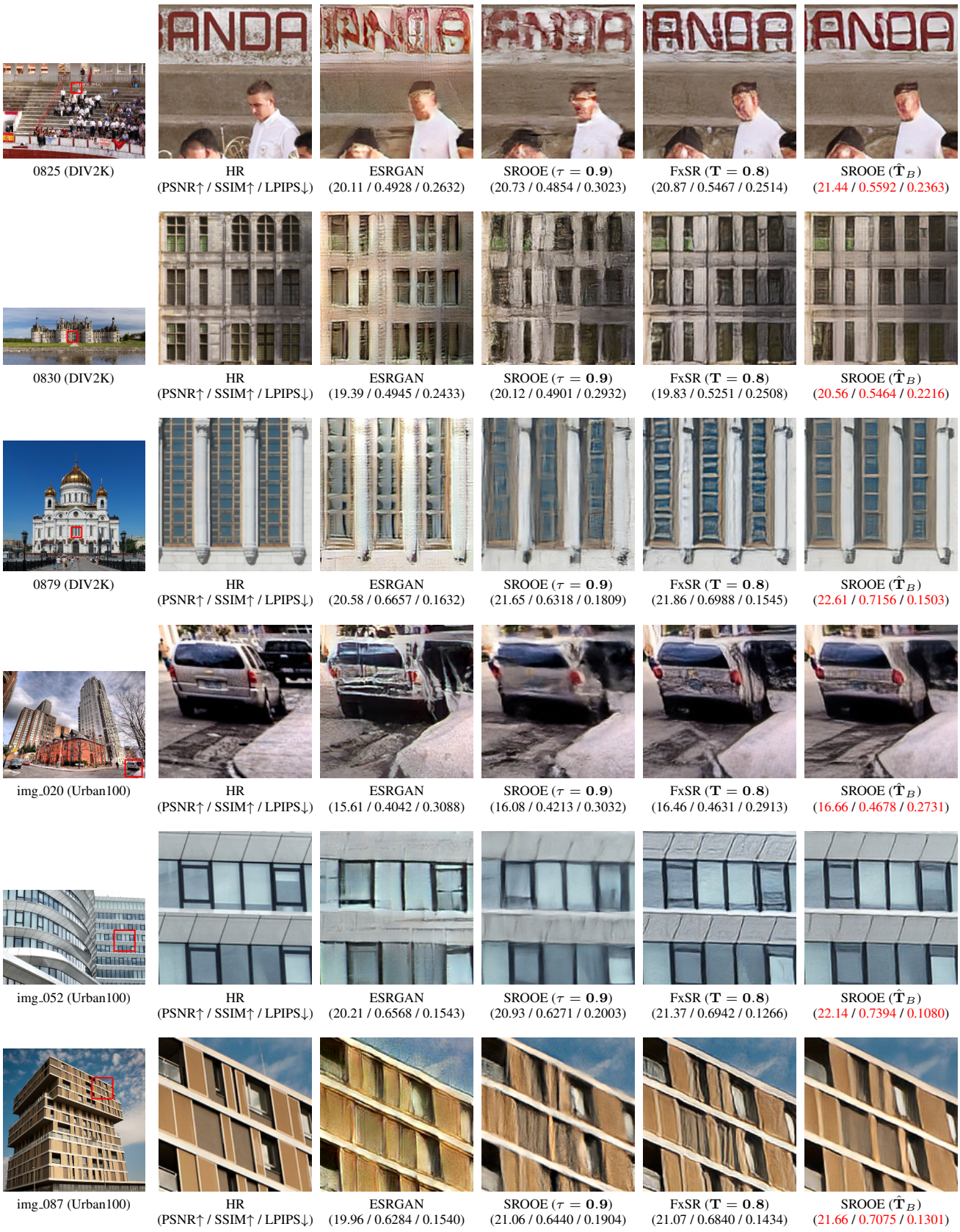

Figure 7. Visual comparison with state-of-the-art methods for $8\times$ SR results. Among the four perception-oriented SR methods, the best performances are highlighted in red.

Figure 8. Visual comparison with state-of-the-art methods for $8\times$ SR results. Among the four perception-oriented SR methods, the best performances are highlighted in red.