Supplementary Material: MASA-SR: Matching Acceleration and Spatial Adaptation for Reference-Based Image Super-Resolution

Liying Lu[1*] Wenbo Li[1*] Xin Tao[2] Jiangbo Lu[3] Jiaya Jia[1,3]

[1] The Chinese University of Hong Kong [2] Kuaishou [3] SmartMore

{lylu,wenboli,leojia}@cse.cuhk.edu.hk, jiangsutx@gmail.com, jiangbo@smartmore.com

1. Details of Ablation Study

In the ablation study of block sizes and dilation rates, it should be noted that if the size of the LR block is $m \times n$, then the basic size of the Ref↓ block is set to $\frac{mH_{Ref\downarrow}}{H_{LR}} \times \frac{nW_{Ref\downarrow}}{W_{LR}}$, where $H_{Ref\downarrow}$, $W_{Ref\downarrow}$ and H_{LR}, W_{LR} are the height and width of the Ref↓ feature and the LR feature, respectively. We multiply the basic size of different scale factors in the ablation study, and *use the scale factors to denote the Ref↓ block size* in Fig. 5(b) of body text and in the following descriptions.

Influence of LR block sizes. The FLOPS in Fig. 5(a) is computed by taking as input a 192×192 LR image and a 768×768 Ref image. The Ref↓ block size is 1.5 and the dilation is 1.

Influence of Ref↓ block sizes. The FLOPS in Fig. 5(b) is computed on a 128×128 LR image and a 512×512 Ref image. The LR block size is 8 and the dilation is 1.

Influence of dilation rates. The FLOPS in Fig. 5(c) is computed on a 120×120 LR image and a 480×480 Ref image. The LR block size is 12 and the Ref↓ block size is 1.5.

2. More Visual Results

We show more visual results of the proposed MASA and other state-of-the-art methods, including RCAN [6], HAN [2], ESRGAN [4], SRNTT [7] and TTSR [5]. RCAN and HAN are SISR methods that have achieved the best performance on PSNR, and ESRGAN is a GAN-based SISR method that is considered state-of-the-art in visual quality. SRNTT [7] and TTSR [5] are state-of-the-art RefSR methods. The visual comparison on CUFED5 [7] testing set are shown in Fig. 1 and Fig. 2, and the comparison on Sun80 [3] and Urban100 [1] are shown in Fig. 3 and Fig. 4, respecitively.

It can be observed that our MASA can restore more regular structures and generate photo-realistic details.

*Equal contribution

References

[1] Jia-Bin Huang, Abhishek Singh, and Narendra Ahuja. Single image super-resolution from transformed self-exemplars. In *CVPR*, pages 5197–5206, 2015.

[2] Ben Niu, Weilei Wen, Wenqi Ren, Xiangde Zhang, Lianping Yang, Shuzhen Wang, Kaihao Zhang, Xiaochun Cao, and Haifeng Shen. Single image super-resolution via a holistic attention network. In *ECCV*, pages 191–207. Springer, 2020.

[3] Libin Sun and James Hays. Super-resolution from internet-scale scene matching. In *2012 IEEE International Conference on Computational Photography (ICCP)*, pages 1–12. IEEE, 2012.

[4] 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 *ECCV*, pages 0–0, 2018.

[5] Fuzhi Yang, Huan Yang, Jianlong Fu, Hongtao Lu, and Baining Guo. Learning texture transformer network for image super-resolution. In *CVPR*, pages 5791–5800, 2020.

[6] 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 286–301, 2018.

[7] Zhifei Zhang, Zhaowen Wang, Zhe Lin, and Hairong Qi. Image super-resolution by neural texture transfer. In *CVPR*, pages 7982–7991, 2019.

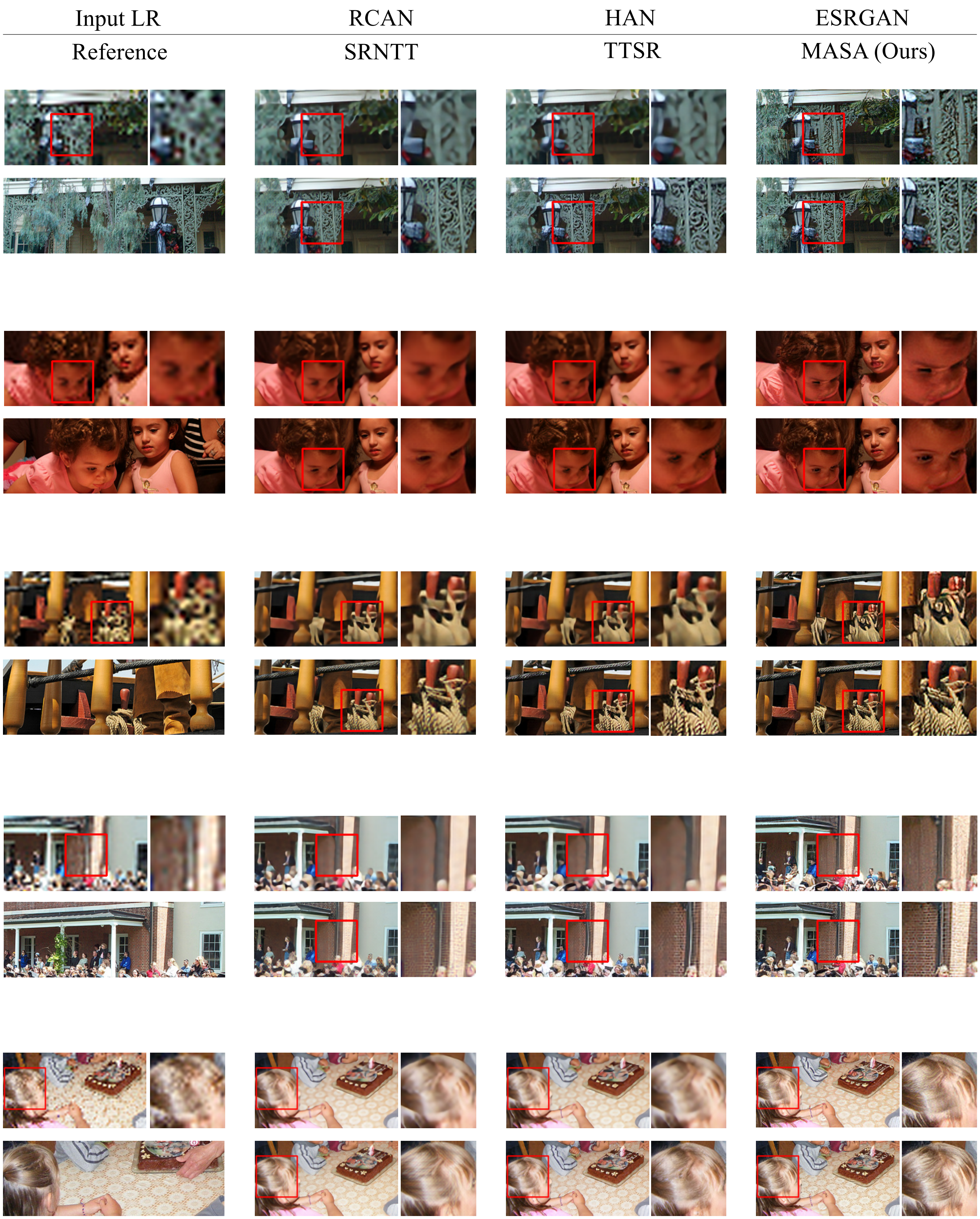

Figure 1: Visual comparison among different SR methods on the CUFED5 [7] testing set.

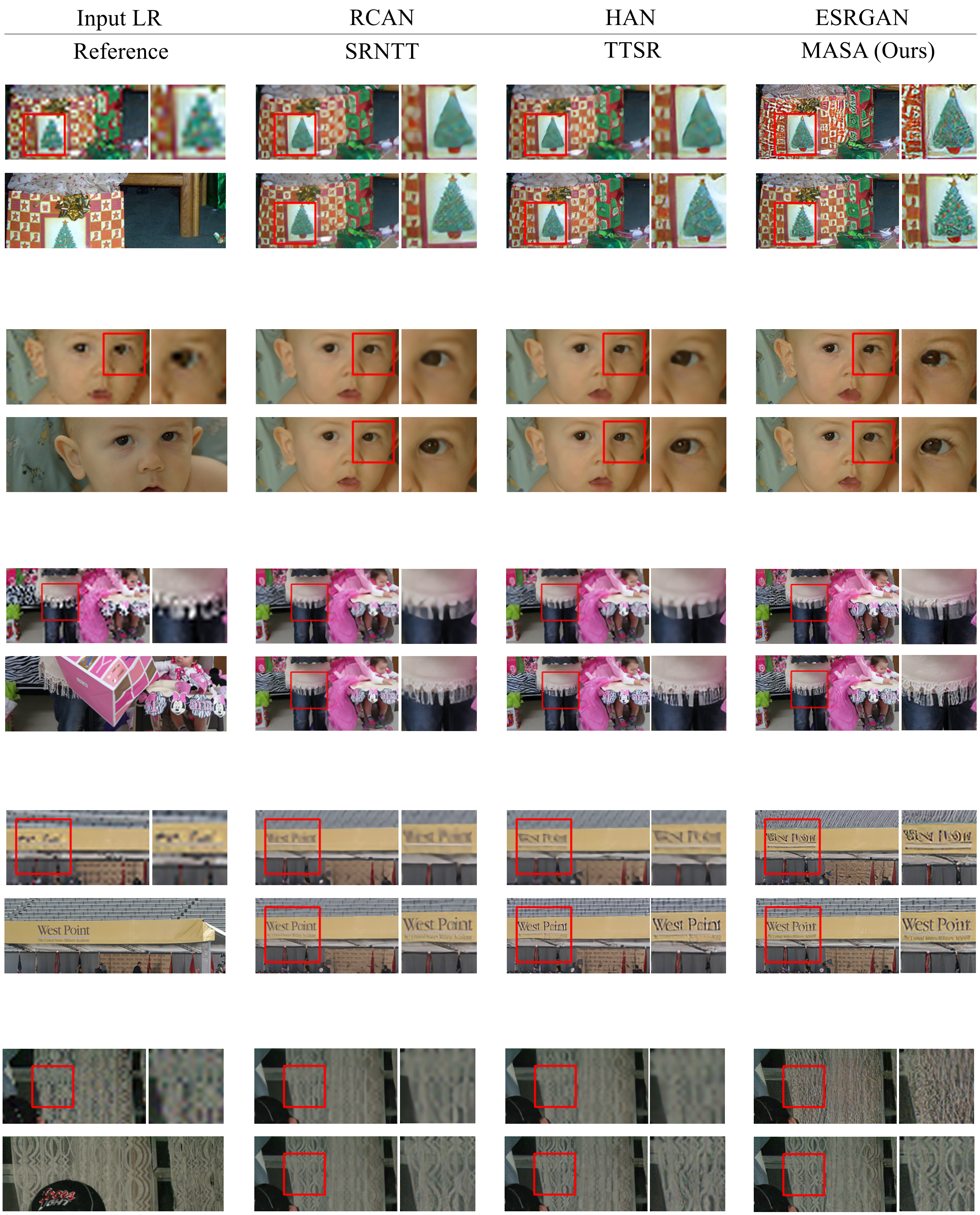

Figure 2: Visual comparison among different SR methods on the CUFED5 [7] testing set.

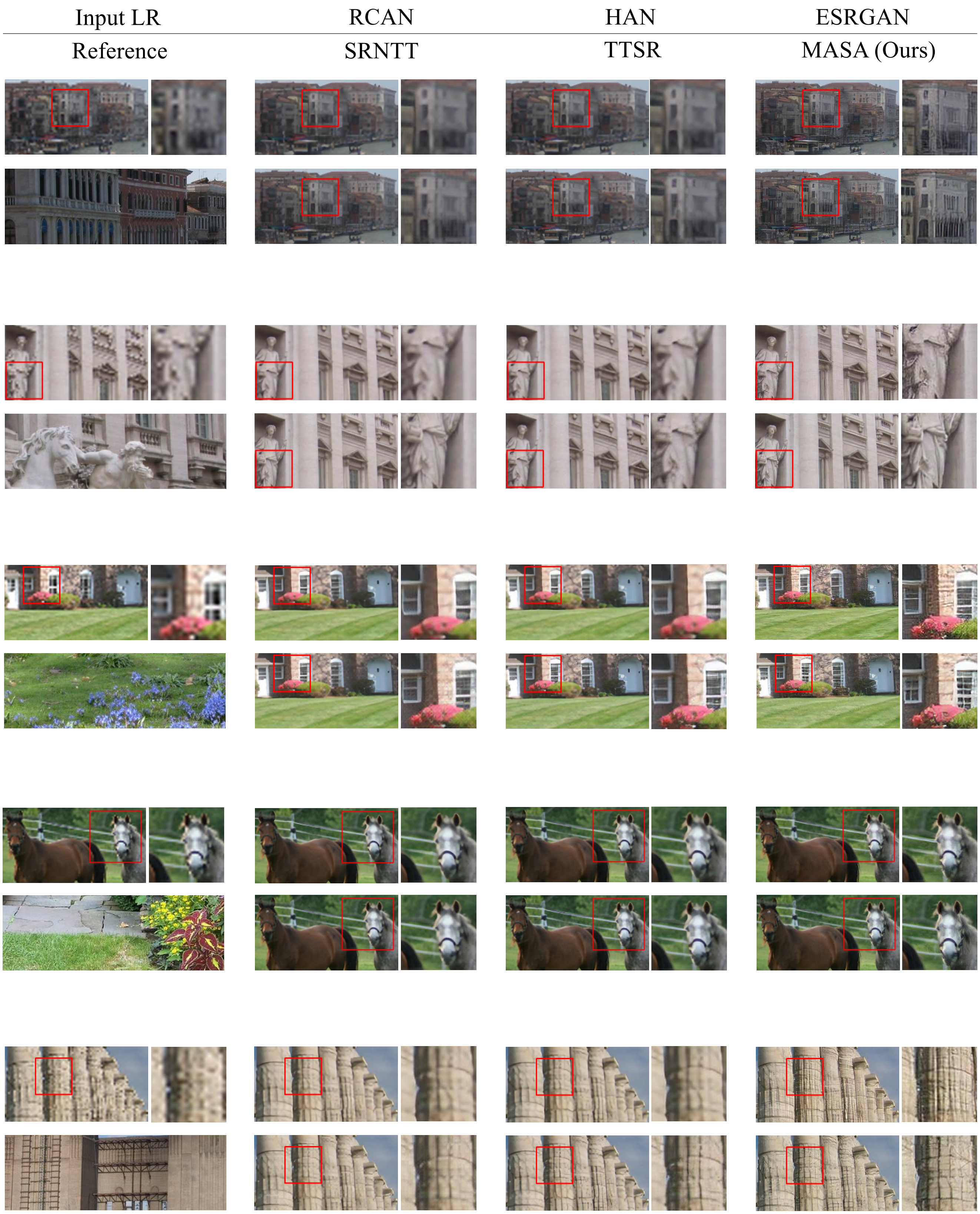

Figure 3: Visual comparison among different SR methods on the Sun80 [3] dataset.

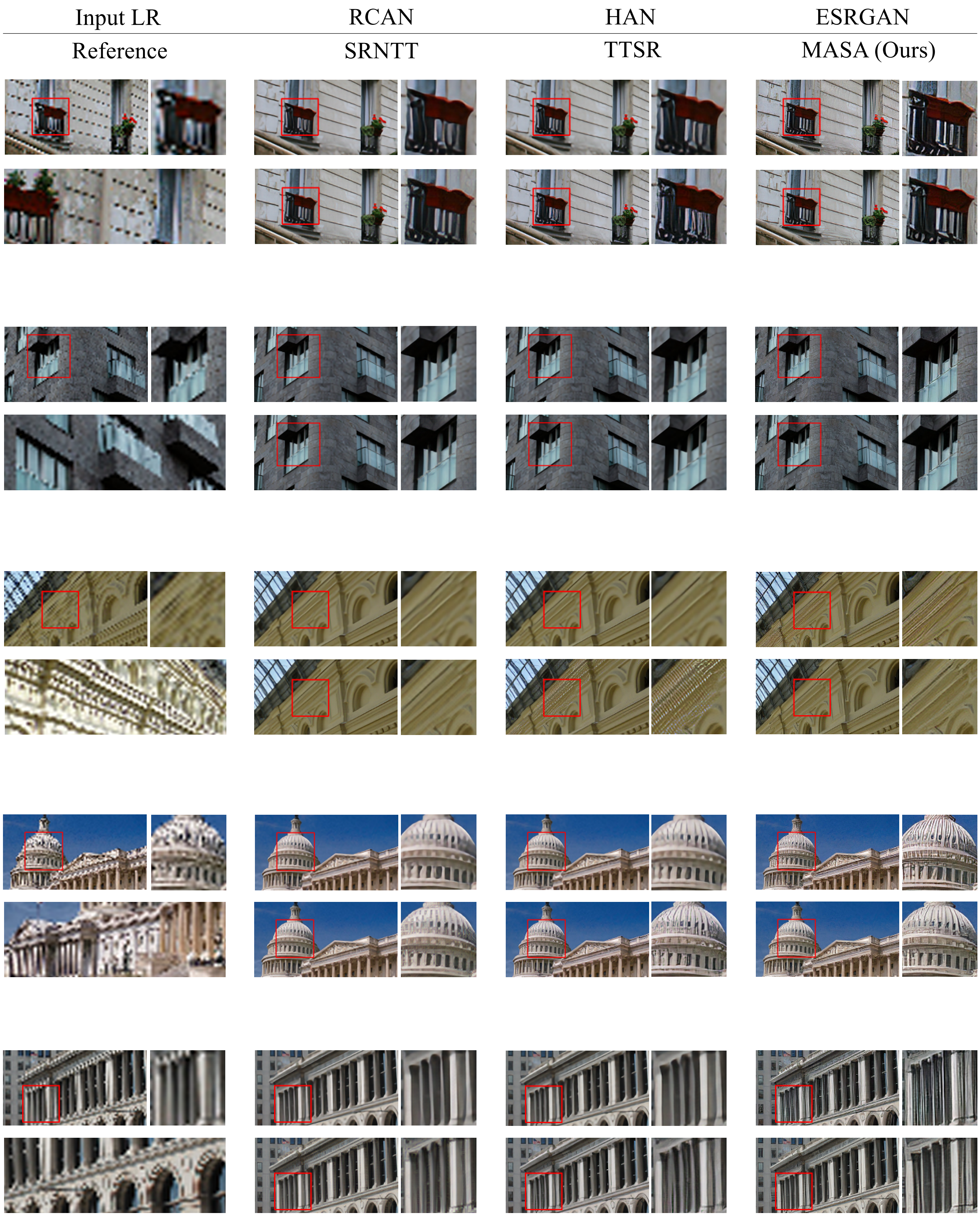

Figure 4: Visual comparison among different SR methods on the Urban100 [1] dataset.