

# Residual Feature Aggregation Network for Image Super-Resolution (Supplementary Material)

## A. Study of the strided convolution in ESA block

As shown in Table 1, the strided convolution has a higher PSNR than the plain convolution, which indicates that a large receptive field is essential for image SR.

	PSNR
strided conv	32.65
plain conv	32.61

Table 1. Investigation of the effect of strided convolution in the ESA block with scale factor of  $\times 4$  on Set5.

## B. Study of the window size for max-pooling in ESA block

As shown in Table 2, we can achieve a higher PSNR by using a pooling window size of  $7 \times 7$ , further proving the critical importance of a large receptive field for image SR.

	PSNR
$7 \times 7$	32.65
$3 \times 3$	32.59

Table 2. Investigation of the effects of different window sizes for max-pooling in the ESA block with scale factor of  $\times 4$  on Set5.

## C. Running time comparison

In Table 3, we compare average forward time of our proposed RFANet with RCAN [2] and SAN [1] on Urban100 with scale factor  $\times 4$ . The forward time of all the networks is evaluated on the same machine with 4.3GHz Intel i7 CPU (32G RAM) and an NVIDIA 1080Ti GPU using their official codes. Our RFANet runs the fastest while achieving the best PSNR which demonstrates the effectiveness of our method.

Model	Forward Time (s)	PSNR
RCAN [2]	0.3197	26.82
SAN [1]	2.5689	26.79
RFANet (Ours)	0.2109	26.92

Table 3. Average forward time comparison on Urban100 with scale factor 4 on an NVIDIA 1080Ti GPU

## D. More Qualitative Results

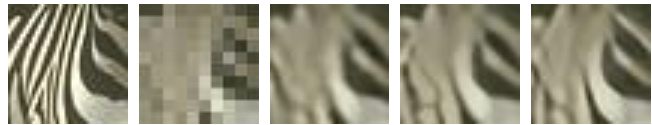
In Fig. 1 and Fig. 2, we provide more visual results of BI and BD degradation models to prove the superiority of the proposed RFANet.

## References

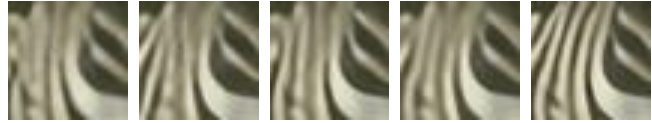
- [1] Tao Dai, Jianrui Cai, Yongbing Zhang, Shu-Tao Xia, and Lei Zhang. Second-order attention network for single image super-resolution. In *CVPR*, pages 11065–11074. Computer Vision Foundation / IEEE, 2019.
- [2] Yulun Zhang, Kunpeng Li, Kai Li, Lichen Wang, Bineng Zhong, and Yun Fu. Image super-resolution using very deep residual channel attention networks. In *ECCV (7)*, volume 11211 of *Lecture Notes in Computer Science*, pages 294–310. Springer, 2018.



253027 from BSD100



HR PSNR/SSIM 20.86/0.5927 Bicubic 21.79/0.6768 FSRCNN 22.39/0.6965 LapSRN 22.70/0.7099 EDSR



DBPN 22.71/0.7128 SRFBN 22.66/0.7113 RCAN 23.12/0.7211 SAN 23.06/0.7198 RFANet (Ours) 23.07/0.7215



img046 from Urban100



HR PSNR/SSIM 22.38/0.6611 Bicubic 23.04/0.7118 FSRCNN 22.39/0.6965 LapSRN 23.32/0.7378 EDSR



DBPN 23.68/0.7644 SRFBN 23.76/0.7688 RCAN 23.96/0.7807 SAN 23.83/0.7729 RFANet (Ours) 24.25/0.7924



img061 from Urban100



HR PSNR/SSIM 22.15/0.5846 Bicubic 23.65/0.6912 FSRCNN 23.95/0.7166 LapSRN 24.65/0.7559 EDSR



DBPN 24.89/0.7714 SRFBN 24.83/0.7715 RCAN 25.10/0.7857 SAN 25.16/0.7836 RFANet (Ours) 25.37/0.7947



PsychoStaff from Manga109



HR PSNR/SSIM 25.45/0.8044 Bicubic 30.77/0.9099 FSRCNN 32.57/0.9402 LapSRN 33.60/0.9499 EDSR



DBPN 34.39/0.9549 SRFBN 34.63/0.9567 RCAN 34.82/0.9582 SAN 34.67/0.9575 RFANet (Ours) 35.01/0.9586



YasassiAkuma from Manga109



HR PSNR/SSIM 22.86/0.7560 Bicubic 26.25/0.8382 FSRCNN 27.63/0.8892 LapSRN 28.71/0.9075 EDSR



DBPN 29.32/0.9135 SRFBN 29.27/0.9149 RCAN 29.32/0.9181 SAN 29.67/0.9186 RFANet (Ours) 29.60/0.9195

Figure 1. Visual comparisons for  $\times 4$  SR with BI degradation model.

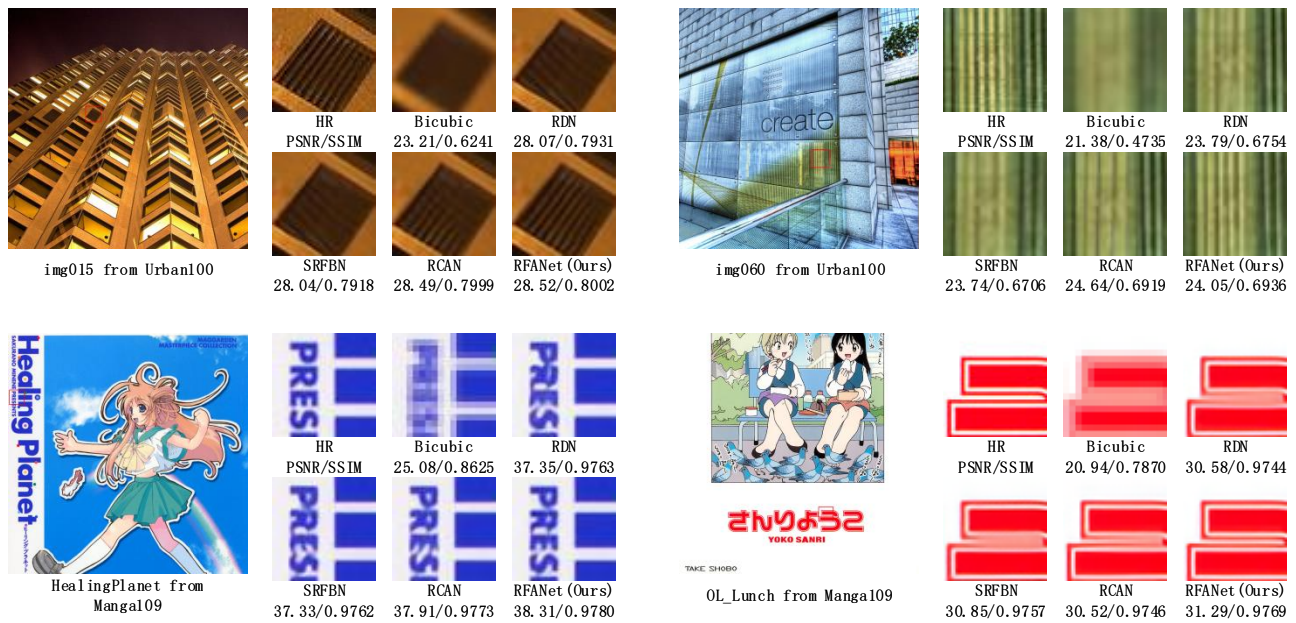


Figure 2. Visual comparisons for  $\times 4$  SR with BD degradation model.