# Lightweight Single-Image Super-Resolution Network with Attentive Auxiliary Feature Learning - Supplementary Material

Xuehui Wang[1], Qing Wang[1], Yuzhi Zhao[2], Junchi Yan[3], Lei Fan[4], and Long Chen[1(✉)]

[1] School of Data and Computer Science, Sun Yat-sen University, Guangzhou, China
wangxh228@mail2.sysu.edu.cn, chenl46@mail.sysu.edu.cn
[2] City University of Hong Kong, Hong Kong, China
[3] Shanghai Jiao Tong University, Shanghai, China
[4] Northwestern University, Evanston, USA

## 1 Ablation Study on a Challengeable Case

To thoroughly validate the effectiveness of the PU and CA block, we conduct experiments on challenge case (e.g. x4) with small network ($A^2F$-S). The results is shown in Table 1. Similar to the ablation study on $A^2F$-L, performance can be slightly improved by using auxiliary features (with projection unit), while the performance of the model can be greatly improved by using auxiliary features and channel attention. We think it is because shallow network with fewer auxiliary features is insufficient to get high frequency information, while the attentive auxiliary features can help (e.g. $A^2F$-S).

**Table 1.** Ablation study on $A^2F$-S.

| Methods | Params | MultiAdd | Set5 | Set14 | B100 | Urban100 | Manga109 |
|---|---|---|---|---|---|---|---|
| Baseline | 312k | 17.98G | 31.75 | 28.35 | 27.39 | 25.50 | 29.66 |
| $A^2$F-S-NOCA | 323k | 18.57G | 31.75 | 28.33 | 27.40 | 25.51 | 29.68 |
| $A^2$F-S | 331k | 18.60G | 31.87 | 28.36 | 27.41 | 25.58 | 29.77 |

## 2 Necessity about designing $A^2$F

To demonstrate the necessity of designing our lightweight SR method, we do a comparision with [1] by reducing its model size through using less residual groups or residual blocks. Thus, RCAN-L and RCAN-S have almost the same number of convolutional layers with $A^2$F-L and $A^2$F-SD, which makes this comparision fair. We use the official code[5] of RCAN and train the model with official configuration. We also calculate the num of parameters and GFLOPS when the size of inputs is $32 \times 32 \times 3$. From Table2, our method has great superiority, which demonstrates the necessity of designing a customized network for lightweight SR task.

[5] https://github.com/yulunzhang/RCAN

**Table 2.** Necessity about designing $A^2F$. These results are obtained with scale x4.

| Model | Param/GFlops | Set5 | Set14 | B100 | Urban100 | Manga109 |
|---|---|---|---|---|---|---|
| RCAN-S | 423k/0.413G | 31.58 | 28.19 | 27.29 | 25.23 | 29.18 |
| **$A^2$F-SD(ours)** | **320k/0.321G** | **32.06** | **28.47** | **27.48** | **25.80** | **30.16** |
| RCAN-L | 1450K/1.29G | 31.90 | 28.40 | 27.42 | 25.65 | 29.76 |
| **$A^2$F-L(ours)** | **1374k/1.37G** | **32.32** | **28.67** | **27.62** | **26.32** | **30.72** |

## 3 Effectiveness of auxiliary features

We remove auxiliary features links from the network ($A^2$F-SD) to prove the effect of auxiliary features. So the input of the attention branch is only the output of last layer. The result shown in Table 3 demonstrates that auxiliary features are useful.

**Table 3.** Effectiveness of auxiliary features. "$(w/o)$" means that no auxiliary features are adopted in $A^2$F-SD. These results are obtained with scale x4.

| Model | Param/GFlops | Set5 | Set14 | B100 | Urban100 | Manga109 |
|---|---|---|---|---|---|---|
| $A^2$F-SD$(w/o)$ | 313k/0.315G | 31.86 | 28.37 | 27.38 | 25.55 | 29.75 |
| **$A^2$F-SD** | **320k/0.321G** | **32.06** | **28.47** | **27.48** | **25.80** | **30.16** |

## 4 Average PSNR/SSIM

According to Table 6 in our paper, we calculate the average PSNR/SSIM among five datasets in Table 4 to avoid the situation that some of these five datasets may be saturated but often the improvement is in the range of 0.05db. Note that we do this on a challengeable task (e.g. scale x4) to illustrate our powerful capacity.

**Table 4.** The average PSNR and SSIM among five datasets and the FLOPs with input size $32 \times 32$ for scale x4.

| $Parameters < 1000K$ | | | | | | |
|---|---|---|---|---|---|---|
| Metric | SRCNN | DRRN | **$A^2$F-SD** | AWSRN-S | VDSR | LapSRN |
| Params/Mul-Adds | 57K/52.7G | 297K/6797G | **320K/18.2G** | 588K/37.7G | 665K/612.6G | 813K/149.4G |
| FLOPs | 0.332G | 3.024G | **0.321G** | 0.601G | 2.728G | 1.988G |
| PSNR/SSIM | 27.41/0.7792 | 28.434/0.8098 | **28.794/0.8179** | 28.566/0.8124 | 28.132/0.8019 | 28.27/0.8051 |
| $Parameters \geq 1000K$ | | | | | | |
| Metric | SRFBN-S | **$A^2$F-M** | AWSRN-M | **$A^2$F-L** | AWSRN | DRCN |
| Params/Mul-Adds | 1000K/- | **1010K/56.7G** | 1254K/72G | **1374K/77.2G** | 1587K/91.1G | 1774K/17974G |
| FLOPs | 0.323G | **1.010G** | 1.280G | **1.370G** | 1.620G | - |
| PSNR/SSIM | 28.698/0.8148 | **29.044/0.8228** | 29.034/0.8226 | **29.13/0.8246** | 29.122/0.8245 | 28.18/0.8017 |

## 5 Visual Comparison on Other Datasets

Besides the qualitative comparison on Set14 [2] and Urban100 [3], we also do this on other datasets that are Set5 [4] and Manga109 [5] on scale $\times 4$. We do not make comparisons on $\times 2$ and $\times 3$ due to that as the scale becomes larger, the difference between methods will be more noticeable. Please see Figure 1 for visual details.


**Fig. 1.** Qualitative comparison over two datasets for scale $\times 4$. The red rectangle on the image of ours indicates visible difference compared with others.

## 6 Comparision with Non-Lightweight SOTAs

In our work, we aim to provide a fast, low-parameters and accurate method, which is appropriate for realistic applications but with a little performance droping off. We make a comparison with non-lightweight SOTAs that are published in ICCV or CVPR recently including OISR-LF-s [6], OISR-LF [6], OISR-RK2-s [6], OISR-RK2 [6], OISR-Rk3 [6], SRFBN [7], MSRN [8], MDSR [9], EBRN [10], SAN [11], RCAN [1], RDN [12], FRSR [13].

From Table 5, we can find that our A$^2$F-L model has the fewest parameters and multi-adds operations on each scale, but it can outperform or comparable with some methods that have more parameters and multi-adds operations (e.g. OISR-LF-s(1370K), OISR-RK2-s(1370K), MSRN(6300K), SRFBN(3631K), OISR-LF(4970K), OISR-RK2(4970K), MDSR(6920K)). Until the parameters is beyond about 7500K, there exists methods outperforming our model obviously.

**Table 5.** The quantitative comparison with non-lightweight methods on four datasets among ×2, ×3, ×4 scale. Red implys that the parameters and the multi-adds operations are both bigger than our model A$^2$F-L. "-" means we do not find an official report or the code to calculate about one method. "**" in SRFBN for ×2 scale indicates that there appears an out-of-memory when we calculates its flops by the same codes, which means its flops may be very huge. We infer 1900G based on the numerical relation from the Multi-Adds of ×3 and ×4.

| Method | Scale | Param | MultiAdds | Set5 | Set14 | B100 | Urban100 |
|---|---|---|---|---|---|---|---|
| OISR-LF-s (CVPR2019) | 2 | 1370K | 316.2G | 38.02/0.9605 | 33.62/0.9178 | 32.20/0.9000 | 32.21/0.9290 |
| OISR-RK2-s (CVPR2019) | | 1370K | 316.2G | 37.98/0.9604 | 33.58/0.9172 | 32.18/0.8996 | 32.09/0.9281 |
| SRFBN (CVPR2019) | | 2140K | 1900G** | 38.11/0.9609 | 33.82/0.9196 | 32.29/0.9010 | 32.62/0.9328 |
| OISR-LF (CVPR2019) | | 4970K | 1145.7G | 38.12/0.9609 | 33.78/0.9196 | 32.26/0.9007 | 32.52/0.9320 |
| OISR-RK2 (CVPR2019) | | 4970K | 1145.7G | 38.12/0.9609 | 33.80/0.9193 | 32.26/0.9006 | 32.48/0.9317 |
| EBRN (ICCV2019) | | 5778K | - | 38.35/0.9620 | 34.24/0.9226 | 32.47/0.9033 | 33.52/0.9402 |
| MSRN (ECCV2018) | | 6300K | - | 38.08/0.9605 | 33.74/0.9170 | 32.23/0.9013 | 32.22/0.9326 |
| MDSR (CVPR2017) | | 6920K | 1592.2G | 38.11/0.9602 | 33.85/0.9198 | 32.29/0.9007 | 32.84/0.9347 |
| SAN (CVPR2019) | | 15700K | - | 38.31/0.9620 | 34.07/0.9213 | 32.42/0.9028 | 33.10/0.9370 |
| RCAN (ECCV2018) | | 16000K | 1570.4G | 38.27/0.9614 | 34.12/0.9216 | 32.41/0.9027 | 33.34/0.9384 |
| RDN (CVPR2018) | | 22120K | 5096.2G | 38.24/0.9614 | 34.01/0.9212 | 32.34/0.9017 | 32.89/0.9353 |
| OISR-RK3 (CVPR2019) | | 41910K | 9656.5G | 38.21/0.9612 | 33.94/0.9206 | 32.36/0.9019 | 33.03/0.9365 |
| **A2F-L (ours)** | **2** | **1363K** | **306.1G** | **38.09/0.9607** | **33.78/0.9192** | **32.23/0.9002** | **32.46/0.9313** |
| OISR-LF-s (CVPR2019) | 3 | 1550K | 160.1G | 34.39/0.9272 | 30.35/0.8426 | 29.11/0.8058 | 28.24/0.8544 |
| OISR-RK2-s (CVPR2019) | | 1550K | 160.1G | 34.43/0.9273 | 30.33/0.8420 | 29.10/0.8053 | 28.20/0.8534 |
| SRFBN (CVPR2019) | | 2833K | 1431.9G | 34.70/0.9292 | 30.51/0.8461 | 29.24/0.8084 | 28.73/0.8641 |
| OISR-LF (CVPR2019) | | 5640K | 578.6G | 34.56/0.9284 | 30.46/0.8450 | 29.20/0.8077 | 28.56/0.8606 |
| OISR-RK2 (CVPR2019) | | 5640K | 578.6G | 34.55/0.9282 | 30.46/0.8443 | 29.18/0.8075 | 28.50/0.8597 |
| MSRN (ECCV2018) | | 6300K | - | 34.38/0.9262 | 30.34/0.8395 | 29.08/0.8041 | 28.08/0.8554 |
| MDSR (CVPR2017) | | 7510K | 768.1G | 34.66/0.9280 | 30.44/0.8452 | 29.25/0.8091 | 28.79/0.8655 |
| SAN (CVPR2019) | | 15700K | - | 34.75/0.9300 | 30.59/0.8476 | 29.33/0.8112 | 28.93/0.8671 |
| RCAN (ECCV2018) | | 16000K | 1590G | 34.74/0.9299 | 30.65/0.8482 | 29.32/0.8111 | 29.09/0.8702 |
| RDN (CVPR2018) | | 22310K | 2281.2G | 34.71/0.9296 | 30.57/0.8468 | 29.26/0.8093 | 28.80/0.8653 |
| OISR-RK3 (CVPR2019) | | 44860K | 4590.1G | 34.72/0.9297 | 30.57/0.8470 | 29.29/0.8103 | 28.95/0.8680 |
| **A2F-L (ours)** | **3** | **1367K** | **136.3G** | **34.54/0.9283** | **30.41/0.8436** | **29.14/0.8062** | **28.40/0.8574** |
| OISR-LF-s (CVPR2019) | 4 | 1520K | 114.2G | 32.14/0.8947 | 28.63/0.7819 | 27.60/0.7369 | 26.17/0.7888 |
| OISR-RK2-s (CVPR2019) | | 1520K | 114.2G | 32.21/0.8950 | 28.63/0.7822 | 27.58/0.7364 | 26.14/0.7874 |
| SRFBN (CVPR2019) | | 3631K | 1128.7G | 32.47/0.8983 | 28.81/0.7868 | 27.72/0.7409 | 26.60/0.8015 |
| FRSR (CVPR2019) | | 4800K | - | 32.20/0.8939 | 28.54/0.7808 | 27.60/0.7366 | 26.21/0.7904 |
| OISR-LF (CVPR2019) | | 5500K | 412.2G | 32.33/0.8968 | 28.73/0.7845 | 27.66/0.7389 | 26.38/0.7953 |
| OISR-RK2 (CVPR2019) | | 5500K | 412.2G | 32.32/0.8965 | 28.72/0.7843 | 27.66/0.7390 | 26.37/0.7953 |
| MSRN (ECCV2018) | | 6300K | - | 32.07/0.8903 | 28.60/0.7751 | 27.52/0.7273 | 26.04/0.7896 |
| MDSR (CVPR2017) | | 7880K | 480.4G | 32.50/0.8973 | 28.72/0.7857 | 27.72/0.7418 | 26.67/0.8041 |
| EBRN (ICCV2019) | | 8186K | - | 32.79/0.9032 | 29.01/0.7903 | 27.85/0.7464 | 27.03/0.8114 |
| SAN (CVPR2019) | | 15700K | - | 32.64/0.9003 | 28.92/0.7888 | 27.78/0.7436 | 26.79/0.8068 |
| RCAN (ECCV2018) | | 16000K | 919.9G | 32.63/0.9002 | 28.87/0.7889 | 27.77/0.7436 | 26.82/0.8087 |
| RDN (CVPR2018) | | 22270K | 1309.2G | 32.47/0.8990 | 28.81/0.7871 | 27.72/0.7419 | 26.61/0.8028 |
| OISR-RK3 (CVPR2019) | | 44270K | 2962.5G | 32.52/0.8992 | 28.86/0.7878 | 27.75/0.7428 | 26.79/0.8068 |
| **A2F-L (ours)** | **4** | **1374K** | **77.2G** | **32.32/0.8964** | **28.67/0.7839** | **27.62/0.7379** | **26.32/0.7931** |

## References


1. Zhang, Y., Li, K., Li, K., Wang, L., Zhong, B., Fu, Y.: Image super-resolution using very deep residual channel attention networks. In: Proceedings of the European Conference on Computer Vision (ECCV). (2018) 286–301
2. Yang, J., Wright, J., Huang, T.S., Ma, Y.: Image super-resolution via sparse representation. IEEE transactions on image processing **19** (2010) 2861–2873
3. Huang, J.B., Singh, A., Ahuja, N.: Single image super-resolution from transformed self-exemplars. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2015) 5197–5206
4. Bevilacqua, M., Roumy, A., Guillemot, C., Alberi-Morel, M.L.: Low-complexity single-image super-resolution based on nonnegative neighbor embedding. (2012)
5. Matsui, Y., Ito, K., Aramaki, Y., Fujimoto, A., Ogawa, T., Yamasaki, T., Aizawa, K.: Sketch-based manga retrieval using manga109 dataset. Multimedia Tools and Applications **76** (2017) 21811–21838
6. He, X., Mo, Z., Wang, P., Liu, Y., Yang, M., Cheng, J.: Ode-inspired network design for single image super-resolution. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2019)
7. Li, Z., Yang, J., Liu, Z., Yang, X., Jeon, G., Wu, W.: Feedback network for image super-resolution. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2019)
8. Li, J., Fang, F., Mei, K., Zhang, G.: Multi-scale residual network for image super-resolution. In: The European Conference on Computer Vision (ECCV). (2018)
9. Lim, B., Son, S., Kim, H., Nah, S., Mu Lee, K.: Enhanced deep residual networks for single image super-resolution. In: Proceedings of the IEEE conference on computer vision and pattern recognition workshops. (2017) 136–144
10. Qiu, Y., Wang, R., Tao, D., Cheng, J.: Embedded block residual network: A recursive restoration model for single-image super-resolution. In: The IEEE International Conference on Computer Vision (ICCV). (2019)
11. Dai, T., Cai, J., Zhang, Y., Xia, S.T., Zhang, L.: Second-order attention network for single image super-resolution. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2019)
12. Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y.: Residual dense network for image super-resolution. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. (2018) 2472–2481
13. Soh, J.W., Park, G.Y., Jo, J., Cho, N.I.: Natural and realistic single image super-resolution with explicit natural manifold discrimination. In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). (2019)