Supplementary Materials to "Efficient Long-Range Attention Network for Image Super-resolution"

Xindong Zhang[1,2,⋆], Hui Zeng[2,⋆], Shi Guo[1], and Lei Zhang[1,2,⋆⋆]

[1] Dept. of Computing, The Hong Kong Polytechnic University
[2] OPPO Research, China
{csxdzhang, csshiguo, cslzhang}@comp.polyu.edu.hk, cshzeng@gmail.com

In this supplementary, we provide the following materials:

– More visual comparisons of ELAN-light with other state-of-the-art lightweight methods on ×4 upscaling tasks.
– More visual comparisons of ELAN with other state-of-the-art methods on ×4 upscaling tasks.

1 Qualitative results of ELAN-light

More visual comparison examples between our ELAN-light and CARN [1], IDN [4], IMDN[3], EDSR-baseline [8], LAPAR-A [5] and SwinIR-light [7] are shown in Figure 1. We can see that most of the compared methods fail to restore the shape of lattice in the first image, producing undesired artifacts. For the second and third images, the compared methods are prone to generating blurry results or artifacts around the long sharp edges. As for the last image, the competitors fail to preserve the structure of character "M", while our ELAN-light can restore the character very well. These results demonstrate the effectiveness of our modeling of long-range self-attention on the SR task.

2 Qualitative results of ELAN

More visual comparison examples between our ELAN and EDSR [8], RDN [10], SAN [2], RCAN [9], IGNN [6] and SwinIR [7] are shown in Figure 2. Similar to the main paper, several interesting observations can be made from them. CNN-based methods, even with very deep architectures like RCAN [9], often fail to recovering the repeated textures (e.g., first image) and long edges (e.g., third image). The transformer based methods such as SwinIR [7] cannot reproduce the long and sharp edges well (e.g., the last two images) since they can only calculate the self-attention in a small window. Our proposed group-wise multi-scale self-attention (GMSA) enables ELAN to exploit the self-similarity information on different scales with larger windows, thus making ELAN more robust and stable on recovering image structures and details from the low-resolution inputs.

⋆ Equal contribution
⋆⋆ Corresponding author. This work is supported by the Hong Kong RGC RIF grant (R5001-18) and the PolyU-OPPO Joint Innovation Lab.

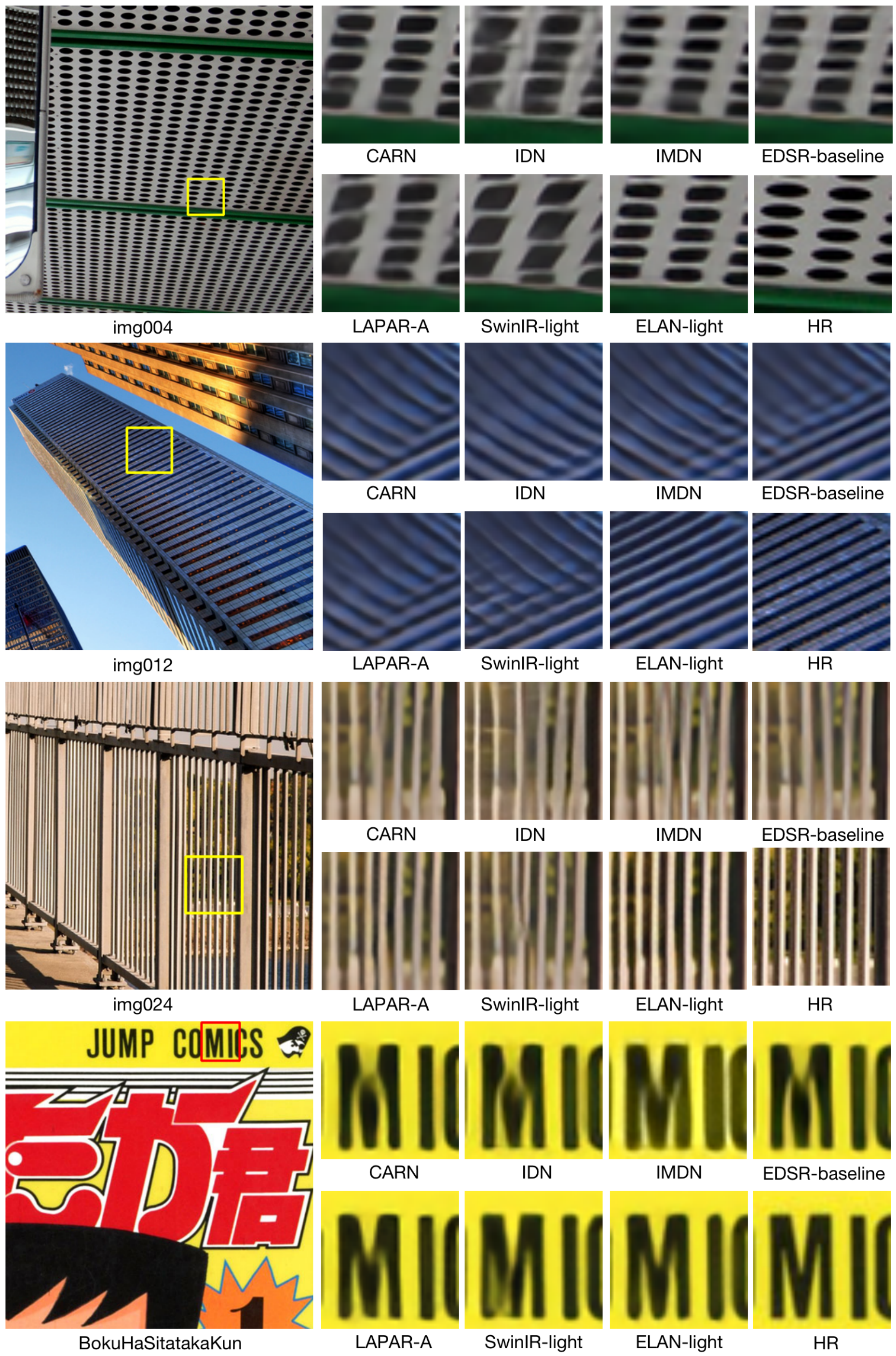

Fig. 1. Qualitative comparison of state-of-the-art light-weight SR models for the $\times 4$ upscaling task.

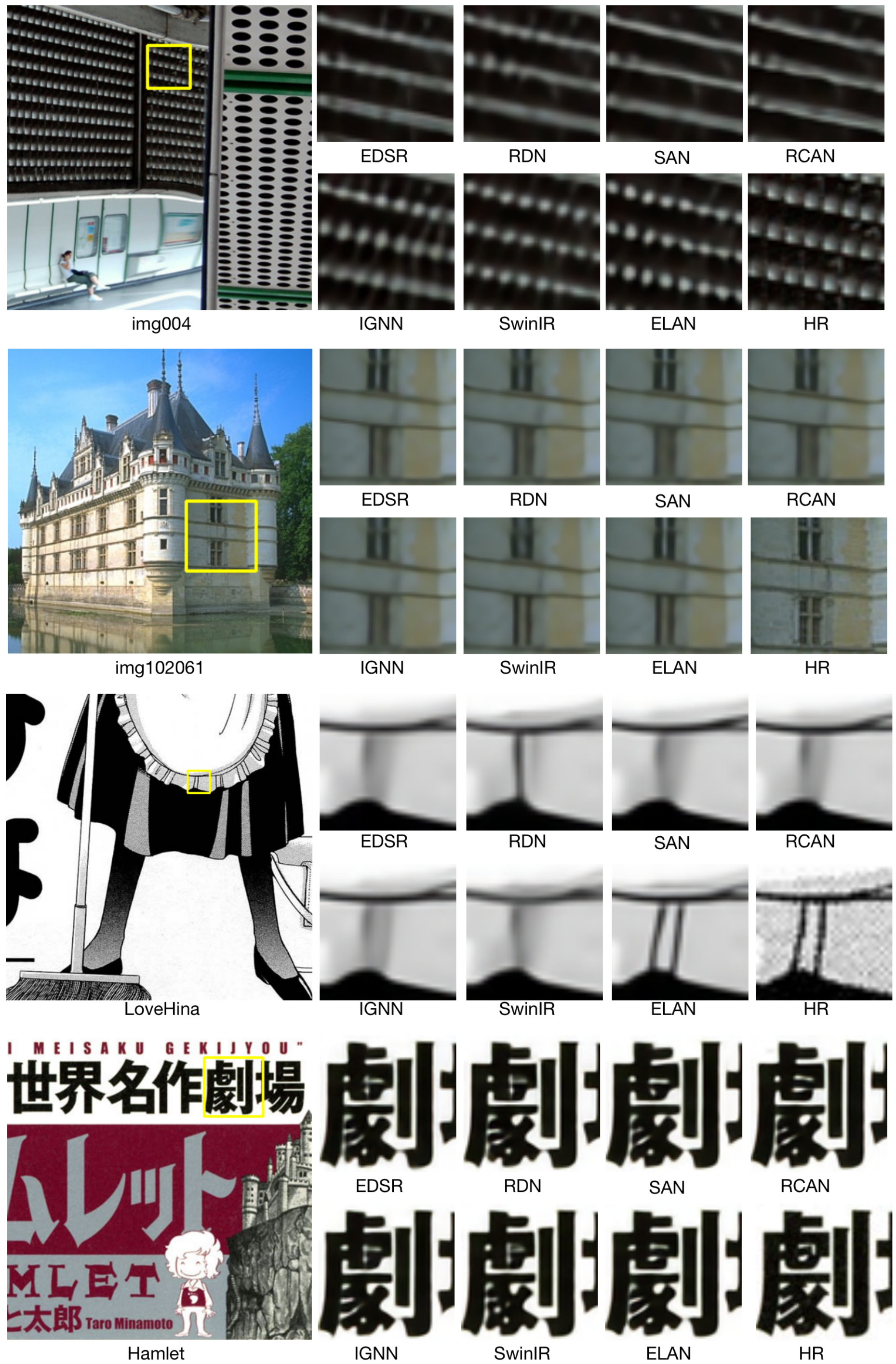

Fig. 2. Qualitative comparison of state-of-the-art classic SR models for the $\times 4$ upscaling task.

References

1. Ahn, N., Kang, B., Sohn, K.A.: Fast, accurate, and lightweight super-resolution with cascading residual network. In: Proceedings of the European Conference on Computer Vision (ECCV). pp. 252–268 (2018)
2. Dai, T., Cai, J., Zhang, Y., Xia, S.T., Zhang, L.: Second-order attention network for single image super-resolution. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 11065–11074 (2019)
3. Hui, Z., Gao, X., Yang, Y., Wang, X.: Lightweight image super-resolution with information multi-distillation network. In: Proceedings of the 27th ACM International Conference on Multimedia. pp. 2024–2032 (2019)
4. Hui, Z., Wang, X., Gao, X.: Fast and accurate single image super-resolution via information distillation network. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 723–731 (2018)
5. Li, W., Zhou, K., Qi, L., Jiang, N., Lu, J., Jia, J.: Lapar: Linearly-assembled pixel-adaptive regression network for single image super-resolution and beyond. Advances in Neural Information Processing Systems **33**, 20343–20355 (2020)
6. Li, Z., Yang, J., Liu, Z., Yang, X., Jeon, G., Wu, W.: Feedback network for image super-resolution. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 3867–3876 (2019)
7. Liang, J., Cao, J., Sun, G., Zhang, K., Van Gool, L., Timofte, R.: Swinir: Image restoration using swin transformer. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 1833–1844 (2021)
8. 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. pp. 136–144 (2017)
9. Zhang, Y., Li, K., Li, K., Zhong, B., Fu, Y.: Residual non-local attention networks for image restoration. arXiv preprint arXiv:1903.10082 (2019)
10. 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. pp. 2472–2481 (2018)