Learning Frequency-aware Dynamic Network for Efficient Super-Resolution Supplementary Material

Wenbin Xie[1,2*], Dehua Song[1*], Chang Xu[3], Chunjing Xu[1], Hui Zhang[2], Yunhe Wang[1]

[1] Noah's Ark Lab, Huawei Technologies. [2] School of Software, Tsinghua University. [3] The University of Sydney.

{dehua.song, yunhe.wang, xuchunjing}@huawei.com; xiewb18@mail.tsinghua.edu.cn;

1. More State-of-the-art SISR models

In addition to EDSR [4] and RCAN [6], we apply our dynamic mechanism to more state-of-the-art SISR models, *i.e.*, SRCNN [2] and VDSR [3]. SRCNN and VDSR consist of three and twenty convolution layers respectively. Neither SRCNN nor VDSR includes residual blocks. Even so, the dynamic block is still applicable to the two models. For SRCNN, all the three convolution layers are regarded as a block and a 3×3 convolution layer is used as the mask predictor. For VDSR, each two convolution layers are regarded as a block. Basically we follow the training and evaluation settings in Section 4 which are slightly different from the origin papers [2, 3], *e.g.*, using rgb images as model inputs instead of y-channel images in YCbCr color space. Besides, the training iteration number is reduced to 300,0000 and learning rates decay to half every 100,000 iterations. The proposed models based on SRCNN and VDSR are called FAD-SRCNN and FAD-VDSR respectively. Quantitative results are shown in Table 1. Though the model capacities and computational redundancies of SRCNN and EDSR are much smaller than that of EDSR, FAD-SRCNN and FAD-VDSR reduce FLOPs significantly while keep similar super-resolution performance compared to their counterparts.

Table 1: Quantitative results in comparison with the state-of-the-art methods on four benchmark databases.

Scale	Method	Set5			Set14			B100			Urban100		
		PSNR ↑ (dB)	SSIM ↑	FLOPs ↓ (G)	PSNR ↑ (dB)	SSIM ↑	FLOPs ↓ (G)	PSNR ↑ (dB)	SSIM ↑	FLOPs ↓ (G)	PSNR ↑ (dB)	SSIM ↑	FLOPs ↓ (G)
×2	SRCNN [1]	36.70	0.9544	7.9	32.45	0.9063	15.9	31.27	0.8874	10.6	29.38	0.8945	53.6
	FAD-SRCNN	36.67	0.9542	**5.2**	32.42	0.9062	**10.4**	31.27	0.8874	**6.0**	29.36	0.8944	**34.6**
	VDSR [3]	37.67	0.9590	75.8	33.24	0.9138	153.0	31.95	0.8964	102.6	31.29	0.9191	518.0
	FAD-VDSR	37.65	0.9589	**43.0**	33.24	0.9138	**91.7**	31.96	0.8966	**62.8**	31.27	0.9189	**320.6**
×3	SRCNN [1]	32.67	0.9079	7.9	29.28	0.8204	15.9	28.35	0.7845	10.6	26.10	0.7979	53.6
	FAD-SRCNN	32.64	0.9075	**5.3**	29.27	0.8203	**10.5**	28.34	0.7843	**6.4**	26.09	0.7978	**35.6**
	VDSR [3]	33.85	0.9217	75.8	29.93	0.8345	153.0	28.83	0.7994	102.6	27.41	0.8345	518.0
	FAD-VDSR	33.84	0.9217	**47.0**	29.94	0.8346	**96.6**	28.82	0.7984	**66.6**	27.39	0.8343	**327.8**
×4	SRCNN [1]	30.40	0.8626	7.9	27.49	0.7535	15.9	26.85	0.7116	10.6	24.36	0.7205	53.6
	FAD-SRCNN	30.37	0.8622	**5.3**	27.47	0.7532	**10.6**	26.83	0.7115	**6.4**	24.33	0.7203	**35.8**
	VDSR [3]	31.48	0.8827	75.8	28.13	0.7698	153.0	27.26	0.7254	102.6	25.32	0.7275	518.0
	FAD-VDSR	31.46	0.8824	**48.6**	28.15	0.7701	**100.0**	27.26	0.7253	**67.3**	25.29	0.7271	**333.7**

2. More Qualitative Results

We show more qualitative super-resolution resulting images based on EDSR [4] backbone in Fig. 1. High frequency signals are seriously damaged during the downsample procedure. Therefore, our dynamic network uses a specific heavy branch to recover them and thus generates high quality images. One can see that our FAD-EDSR generates resulting images of similar and even better visual quality to EDSR [4]. Besides, FAD-EDSR outperforms its counterpart, *i.e.*, AdaEDSR [5] significantly.

*Equal contribution

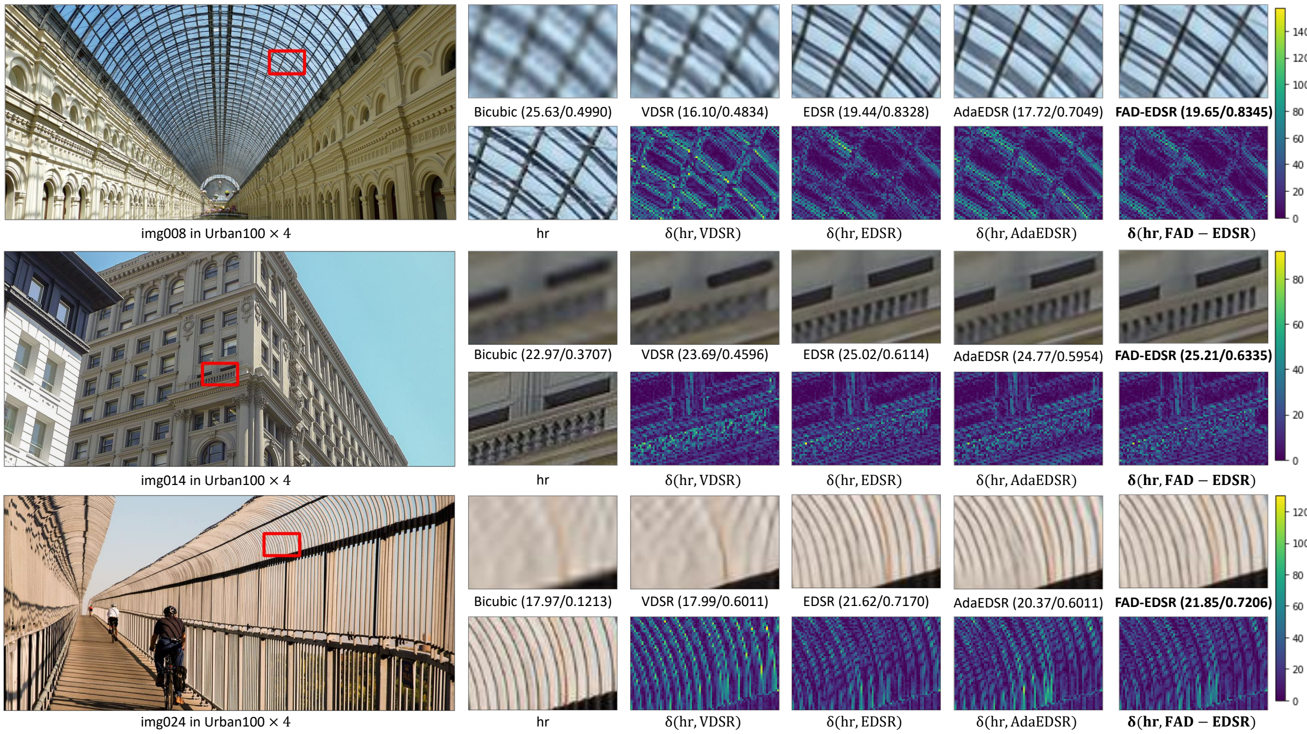

Figure 1: More qualitative results based on EDSR backbone.

References

[1] Chao Dong, Chen Change Loy, Kaiming He, and Xiaoou Tang. Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2):295–307, 2015. 1

[2] Chao Dong, Chen Change Loy, and Xiaoou Tang. Accelerating the super-resolution convolutional neural network. In *ECCV*, pages 391–407. Springer, 2016. 1

[3] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee. Accurate image super-resolution using very deep convolutional networks. In *CVPR*, pages 1646–1654, 2016. 1

[4] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah, and Kyoung Mu Lee. Enhanced deep residual networks for single image super-resolution. In *CVPR Workshops*, pages 136–144, 2017. 1

[5] Ming Liu, Zhilu Zhang, Liya Hou, Wangmeng Zuo, and Lei Zhang. Deep adaptive inference networks for single image super-resolution. In *European Conference on Computer Vision Workshop*, 2020. 1

[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 *Proceedings of the European Conference on Computer Vision*, pages 286–301, 2018. 1