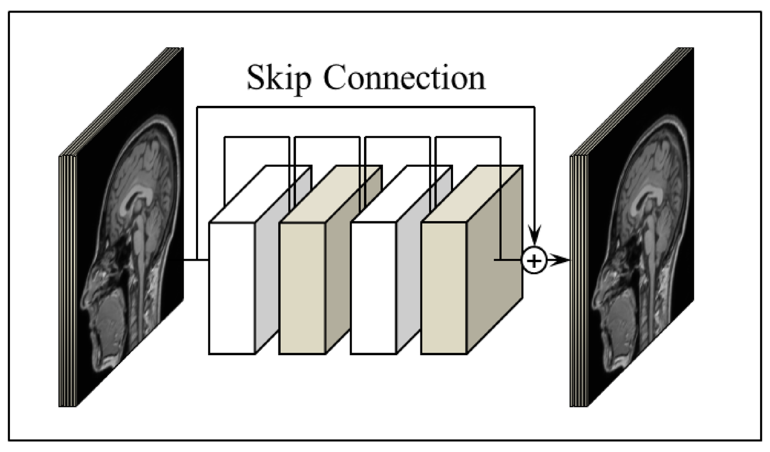


Multilevel Residual Learning for Single Image Super Resolution

Xiaole Zhao, Hangfei Liu, Tao Zhang, Wei Bian, Xueming Zou
zxlation@foxmail.com, taozhangjin@gmail.com, mark.zou@alltechmed.com

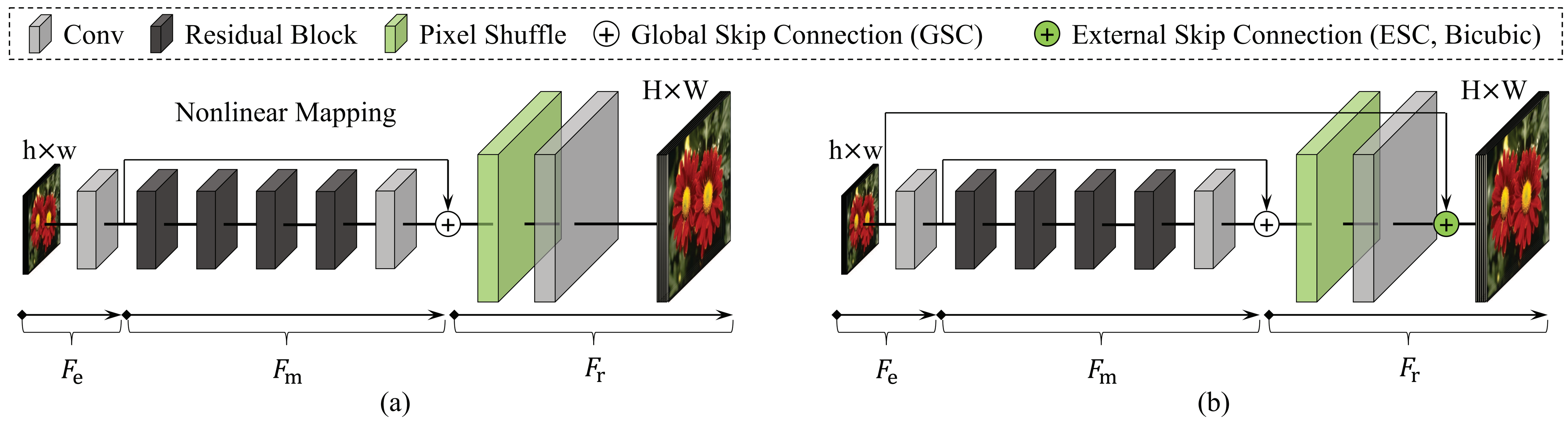
Key Lab of Brainformatics
UESTC

Contributions

1. We introduce residual learning deep into the main path of a traditional residual block, which we term as *fine-grained residual learning*.
2. An external skip connection (ESC) between the interpolated version of the input image and its high resolution (HR) counterpart is applied to form a *coarse-grained residual learning* pattern.
3. Combined with the shortcut of a common residual block, the residual pattern is applied to multiple abstract levels and the entire network displays the characteristic of multilevel residual learning from fine to coarse grain.
4. We empirically show that simply stacking more building blocks to increase the depth of the network does not obtain the expected gain of performance, which may shed some light on the structural design of deep networks.

Overall Network Structure

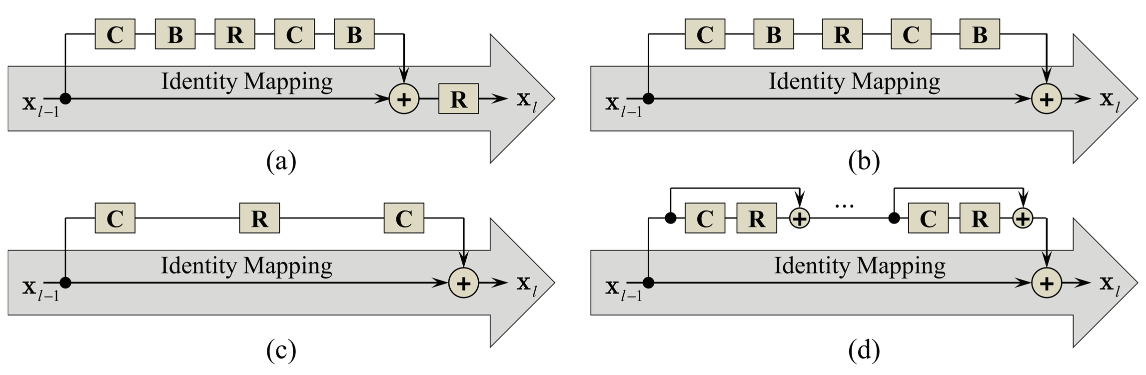
To verify the effectiveness of the strategy of multilevel residual learning, we refer to the structure of EDSR [3]. However, only 4 residual blocks are utilized to build the entire network. The coarse-grained residual learning is implemented by the external skip connection (ESC) between the interpolated input image $\mathbf{x}$ and the predication of the model. Thus, the residual pattern is applied to multiple abstract levels of the network and it shows the feature of multilevel residual learning from fine to coarse grain.



The overall network structure: (a) The reference structure same as EDSR [3] but only 4 residual blocks are used. (b) Extension of EDSR with an external skip connection (ESC).

Building Block

Normally, a residual block consists of a residual path (an identity mapping $\mathbf{x}$) and a main path $\mathcal{F}_\mathrm{B}(\mathbf{x})$. Unlike most of the previous methods, the proposed residual block applies skip connection deep into the main path of a residual block, as shown in sub figure (d).

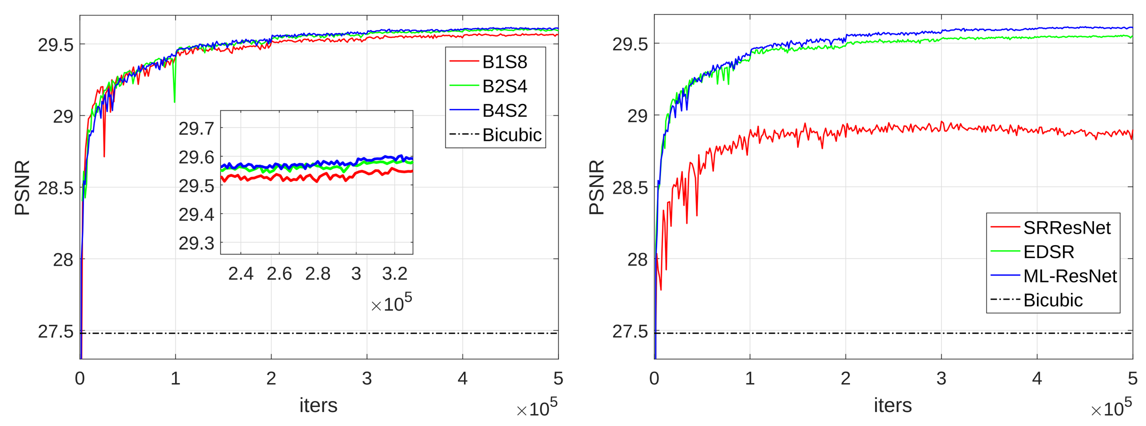


Residual blocks: C, B, R and ⊕ denote conv, batch norm, ReLU and element-wise add, respectively. (a) The original residual block [1]. (b) SRResNet [2]. (c) EDSR/MDSR [3]. (d) Proposed residual block.

Block Structure and Density

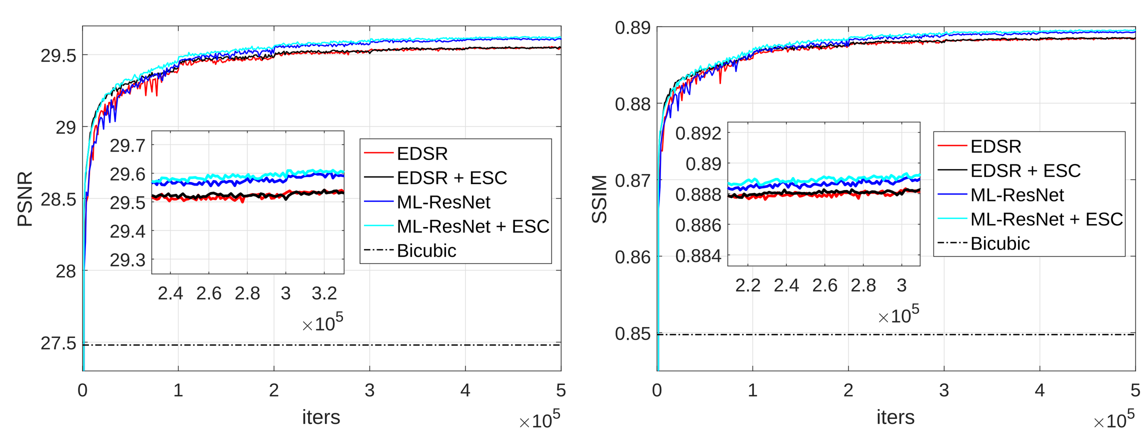
In the proposed residual block, the main path $\mathcal{F}_\mathrm{B}(\mathbf{x})$ consists of several sub residual blocks. Let B and S denote the number of residual blocks and the sub residual blocks on $\mathcal{F}_\mathrm{B}(\mathbf{x})$, the left figure below shows that the equilibrium between B and S can boost the model performance.

The right figure below implies that the proposed ML-ResNet provides the best performance. Note that SRResNet [2] is quantificationally worst but it still can provide good visual perception.



External Skip Connection

We investigate the impact of ESC on the model performance by comparing the models with and without ESC. Interestingly, the ESC does not seem to have obvious effect on the performance of the network with EDSR [3] residual blocks but it slightly improves the performance of the model built with the proposed residual blocks, as shown in the figures below. The number of building blocks for both compared models are set to 4 (left: PSNR, right: SSIM).



References

[1] He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: IEEE conference on Computer Vision and Pattern Recognition, pp.770-778 (2016)

[2] Ledig, C., Theis, L., Huszar, F., Caballero, J., et al.: Photo-realistic single image super-resolution using a generative adversarial network. In: IEEE Conference on Computer Vision and Pattern Recognition, pp.105-114 (2017)

[3] Lim, B., Son, S., Kim, H., Nah, S., Lee, K. M.: Enhanced deep residual networks for single image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp.1132-1140 (2017)

[4] Dong, C., Loy, C. C., He, K., Tang, X.: Image super-resolution using deep convolutional networks. IEEE Transactions on Pattern Analysis & Machine Intelligence, 38(2), 295-307 (2016)

[5] Kim, J., Lee, J. K., Lee, K. M.: Accurate image super-resolution using very deep convolutional networks. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1646-1654 (2016)

[6] Tai, Y., Yang, J., Liu, X.: Image super-resolution via a deep recursive residual network. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 2790-2798 (2017)

[7] Kim, J., Lee, J. K., Lee, K. M. Deeply-recursive convolutional network for image super-resolution. In: IEEE Conference on Computer Vision and Pattern Recognition, pp. 1637-1645 (2016)

[8] Veit, A., Wilber, M., Belongie, S.: Residual networks behave like ensembles of relatively shallow networks. arXiv:1605.06431 [cs.CV] (2016)

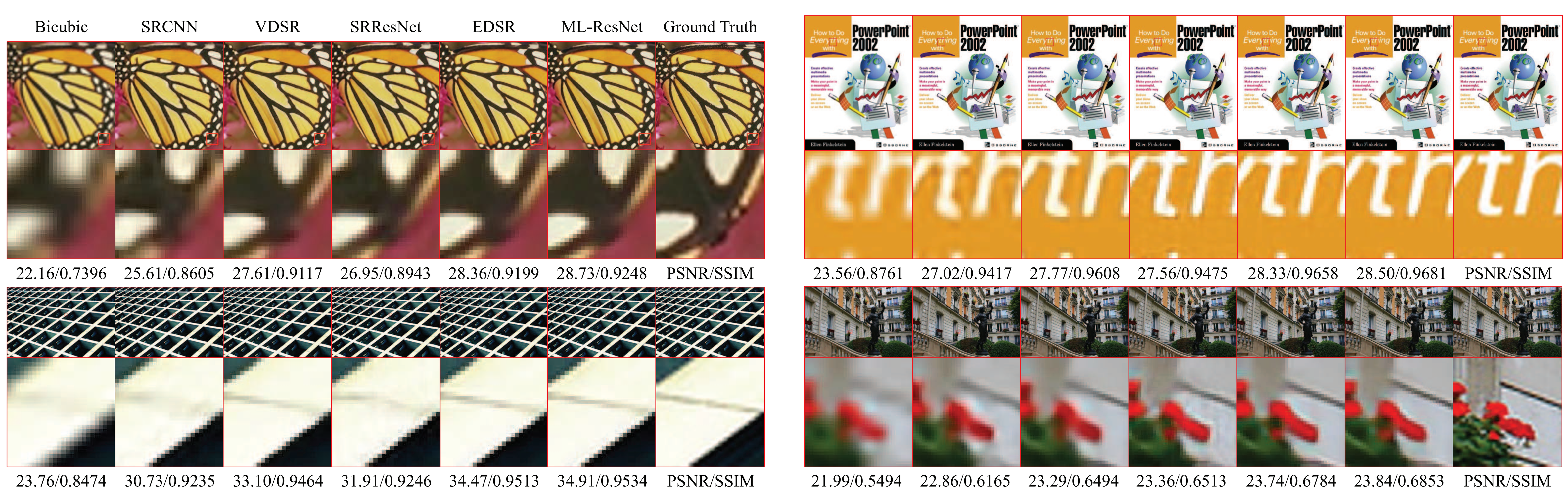
Acknowledgements

This work is supported in part by the National Key Research and Development Program of China (No. 2016YFC0100800 and 2016YFC0100802).

Quantitative and Qualitative Evaluation

Quantitative comparison with other methods: SRResNet (block×4) and EDSR (block×4) are also included here. The maximal values are red, and the second ones are marked as blue (PSNR (dB)/SSIM).

	scale	Bicubic	SRCNN [4]	VDSR [5]	DRRN [6] B1U25	SRResNet [2] block×4	EDSR [3] block×4	ML-ResNet B4S2
Set5	×2	33.66/0.9299	36.66/0.9542	37.53/0.9587	37.74/0.9591	36.94/0.9537	37.64/0.9586	37.78/0.9589
	×3	30.39/0.8682	32.75/0.9090	33.66/0.9213	34.03/0.9244	33.08/0.9080	34.05/0.9234	34.13/0.9248
	×4	28.42/0.8104	30.48/0.8628	31.35/0.8838	31.68/0.8888	31.01/0.8733	31.90/0.8903	32.07/0.8921
Set14	×2	30.24/0.8688	32.42/0.9063	33.03/0.9124	33.23/0.9136	32.74/0.9087	33.29/0.9148	33.32/0.9153
	×3	27.55/0.7742	29.28/0.8209	29.77/0.8314	29.96/0.8349	29.43/0.8232	30.00/0.8367	30.04/0.8371
	×4	26.00/0.7027	27.49/0.7503	28.01/0.7674	28.21/0.7720	27.90/0.7620	28.37/0.7772	28.45/0.7786
B100	×2	29.56/0.8431	31.36/0.8879	31.90/0.8960	32.05/0.8973	31.58/0.8900	32.01/0.8975	32.11/0.8980
	×3	27.21/0.7385	28.41/0.7863	28.82/0.7976	28.95/0.8004	28.42/0.7853	28.91/0.8012	28.95/0.8021
	×4	25.96/0.6675	26.90/0.7101	27.29/0.7251	27.38/0.7284	27.11/0.7185	27.48/0.7329	27.54/0.7346
Urban 100	×2	26.88/0.8403	30.75/0.9133	30.76/0.9140	31.23/0.9188	29.84/0.9106	30.94/0.9262	31.17/0.9278
	×3	24.46/0.7349	26.24/0.7989	27.14/0.8279	27.53/0.8378	26.69/0.8141	27.59/0.8409	27.70/0.8440
	×4	23.14/0.6577	24.52/0.7221	25.18/0.7524	25.44/0.7638	25.01/0.7450	25.77/0.7755	25.94/0.7808



We quantitatively and qualitatively compare the proposed multilevel residual network with several typical methods. We followed the way of DRCN [7] to compute PSNR (dB) and SSIM on several benchmark datasets, i.e., Set5, Set14, B100 and Urban100. It can be observed that the proposed method shows its superiority to other compared methods. However, when we deepen the network and make it have the same depth and model parameters as the original EDSR, the performance of the proposed method is slightly worse than the original EDSR. This implies that directly increasing the number of residual blocks to deepen the network will not get the desired performance improvement, and the multilevel residual structure promotes the propagation and the equilibrium of information flow through the network just when the network is relatively shallow [8].