Image Super-Resolution Using Deep Convolutional Networks

Chao Dong, Chen Change Loy

Kaiming He, Xiaoou Tang

IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, pp. 295–307, Feb. 2016.

Outline

1. Introduction
2. Method
3. Experiment
4. Conclusion

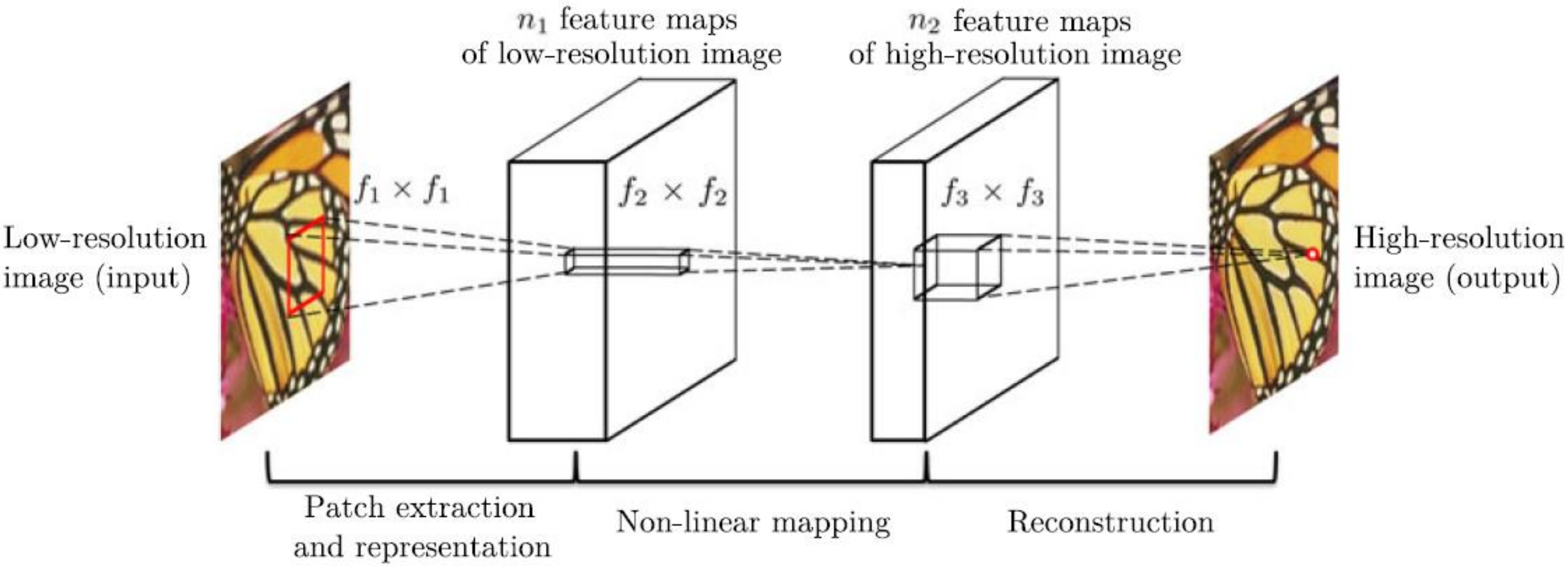
1. Introduction

- We present a fully convolutional neural network for image super-resolution.

- We establish a relationship between our deep-learning-based SR method and the traditional sparse-coding-based SR methods.

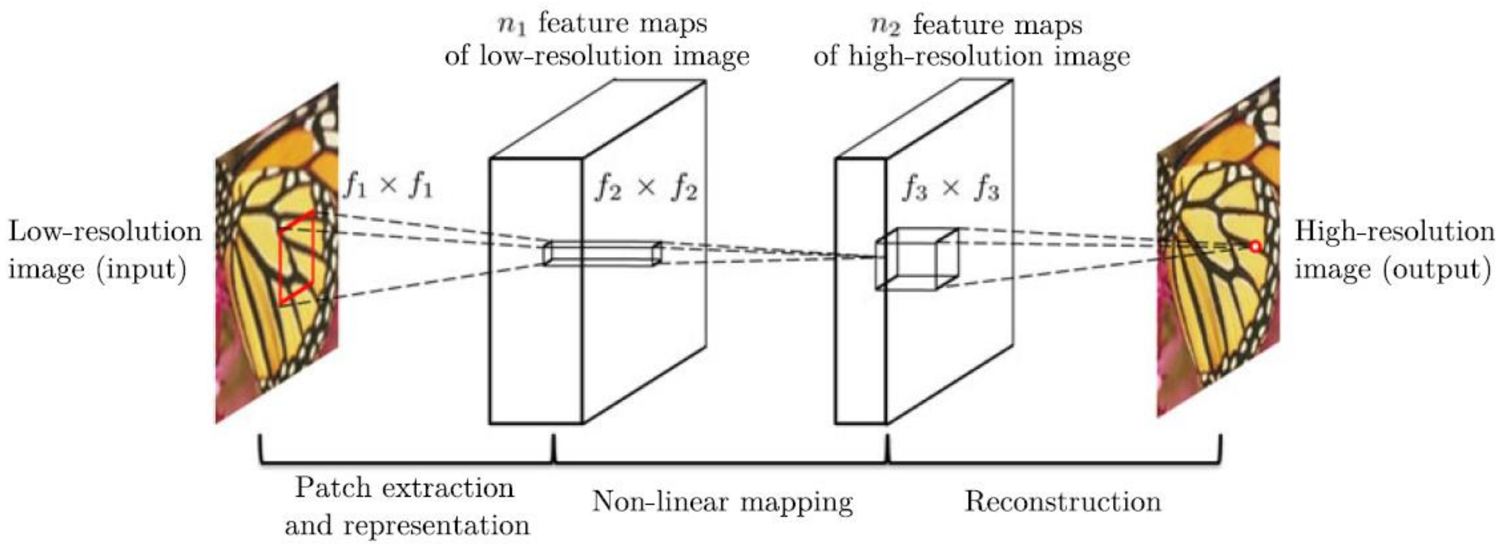
- We demonstrate that deep learning is useful in the classical computer vision problem of super-resolution, and can achieve good quality and speed.

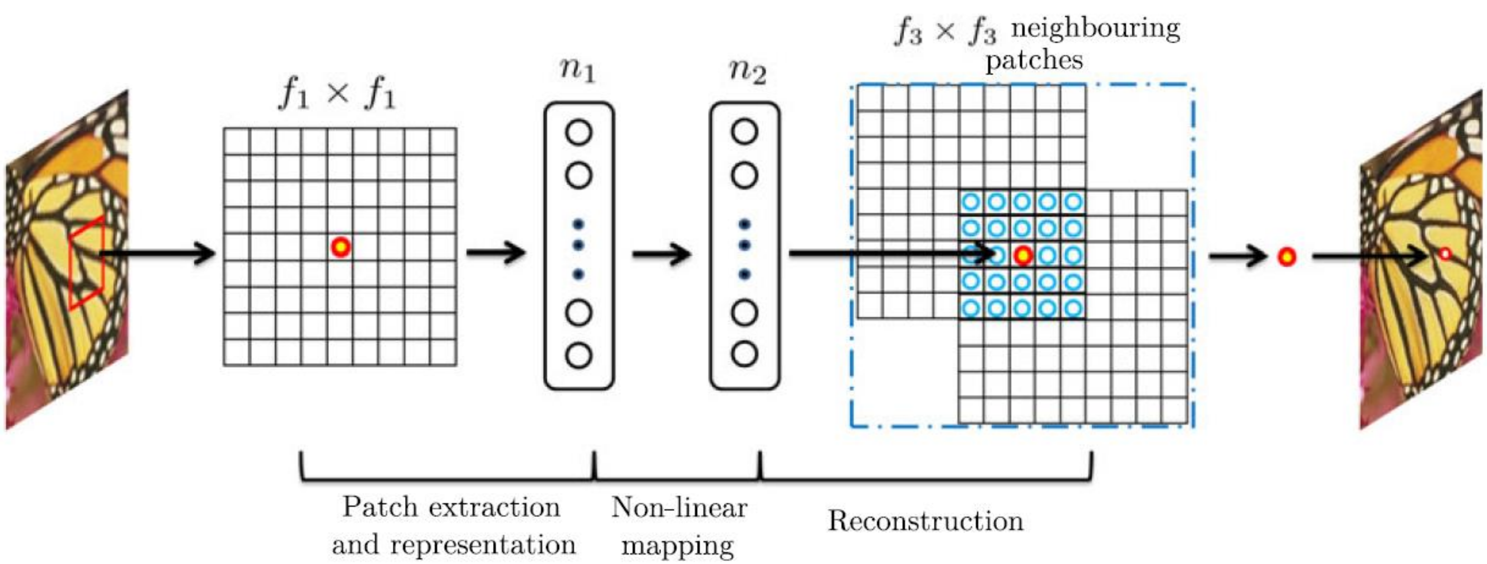
2. Method



- First convolutional layer: $F_1(Y) = \max(0, W_1 * Y + B_1)$
- Second convolutional layer: $F_2(Y) = \max(0, W_2 * F_1(Y) + B_2)$
- Third convolutional layer: $F\ (Y) = \max(0, W_3 * F_2(Y) + B_3)$

- An illustration of sparse-coding-based methods in the view of a convolutional neural network.





3. Experiment

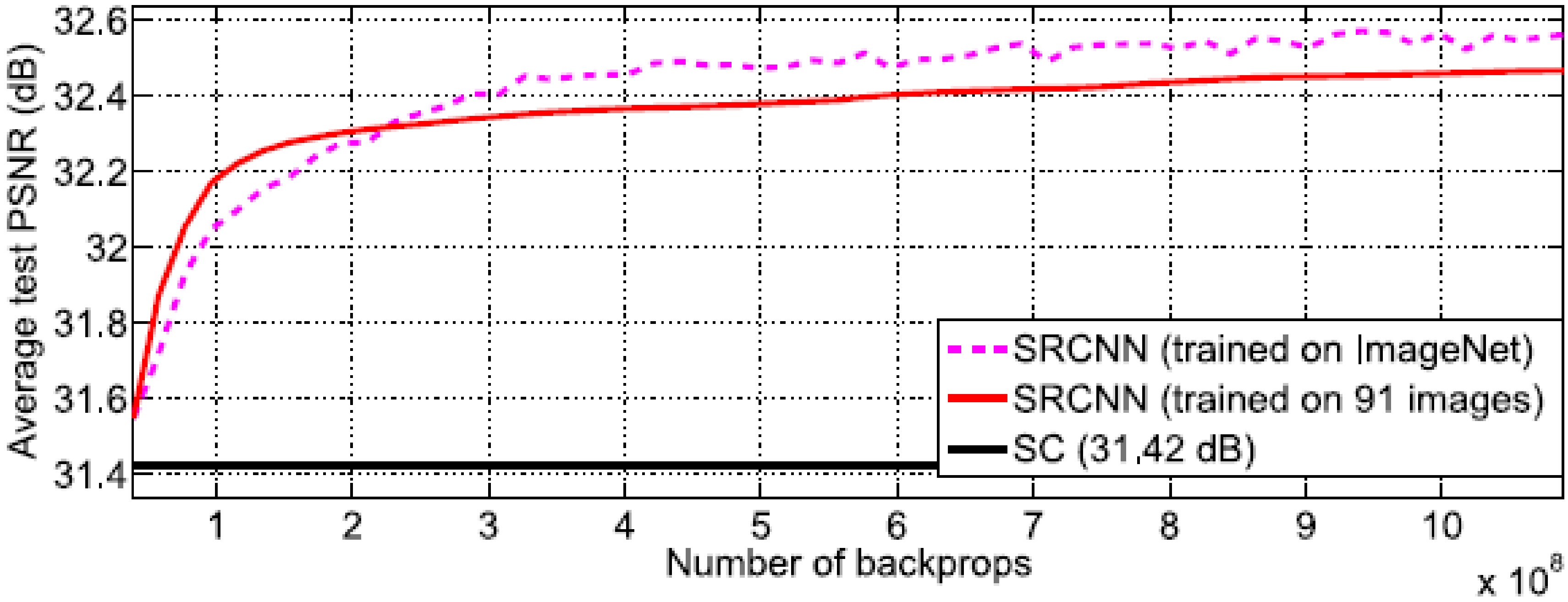
- Investigate the impact of using different datasets on the model performance

- Examine the filters learned by our approach.

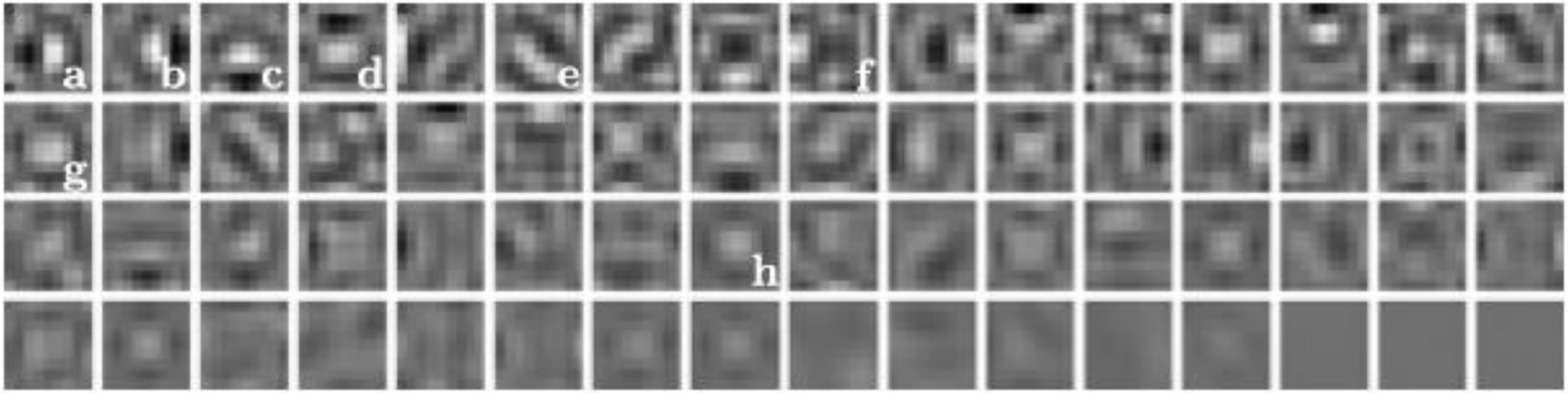
- Explore different architecture designs of the network, like depth, number of filters, and filter sizes.

- Compare our method with recent state-of-the-arts

- Training with the much larger ImageNet dataset improves the performance over the use of 91 images.



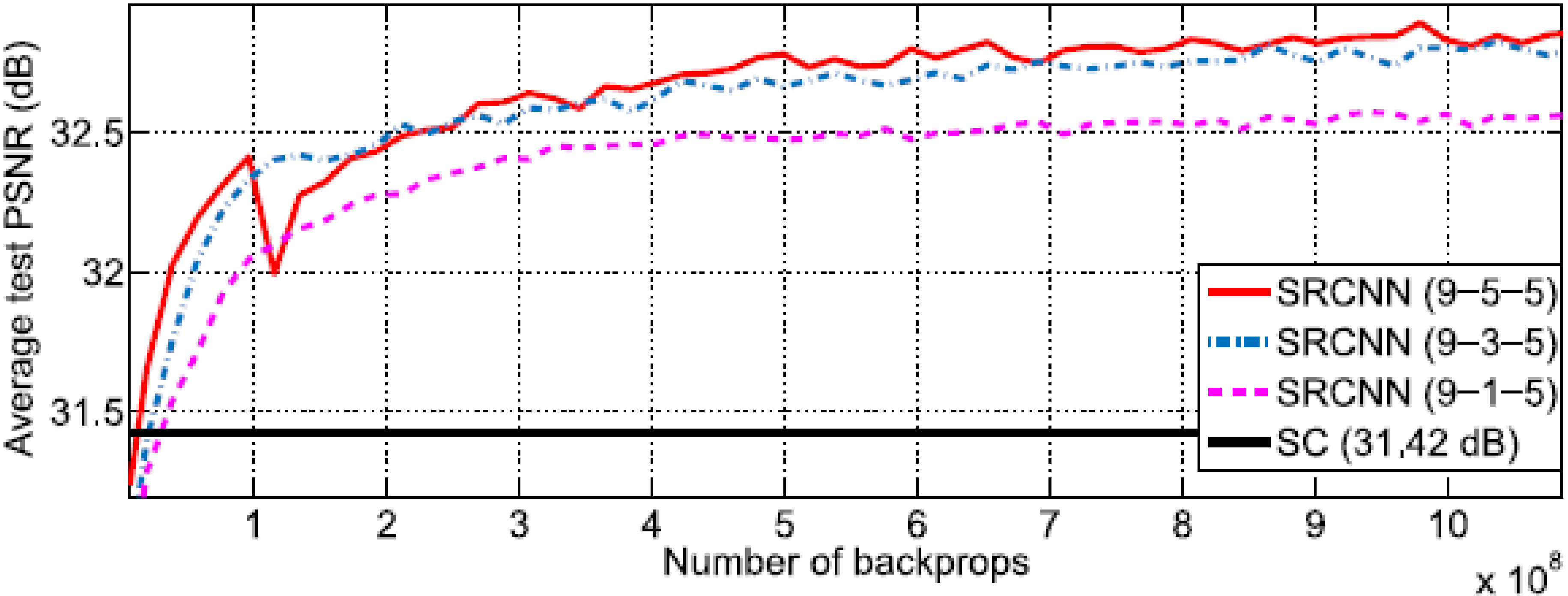
- The filters g and h are like Laplacian/Gaussian filters.
- The filters a - e are like edge detectors at different directions.
- The filter f is like a texture extractor.



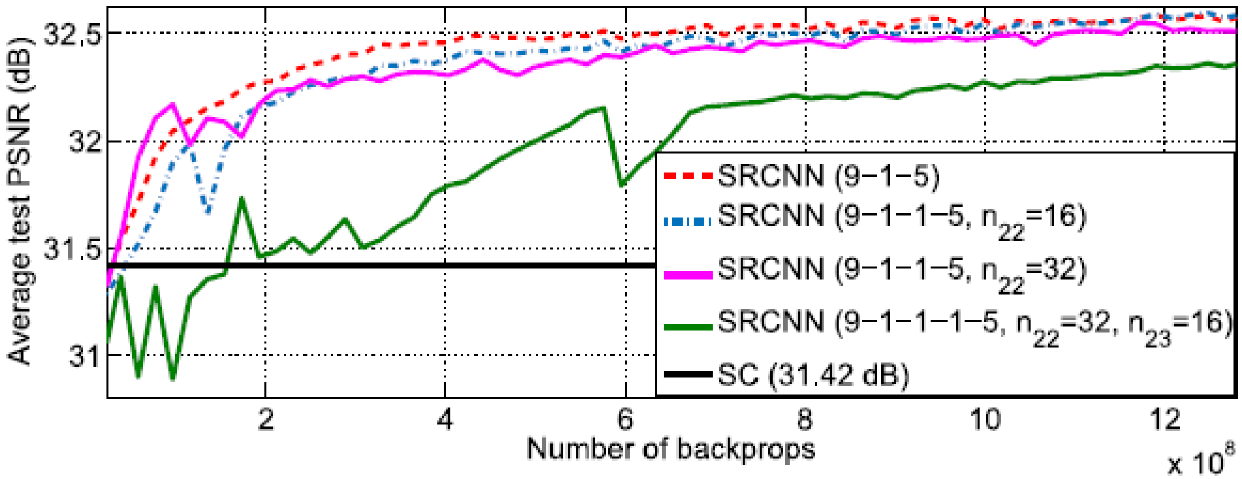
- The Results of Using Different Filter Numbers in SRCNN

$n_1 = 128$, $n_2 = 64$		$n_1 = 64$, $n_2 = 32$		$n_1 = 32$, $n_2 = 16$	
PSNR	Time (sec)	PSNR	Time (sec)	PSNR	Time (sec)
32.60	0.60	32.52	0.18	32.26	0.05

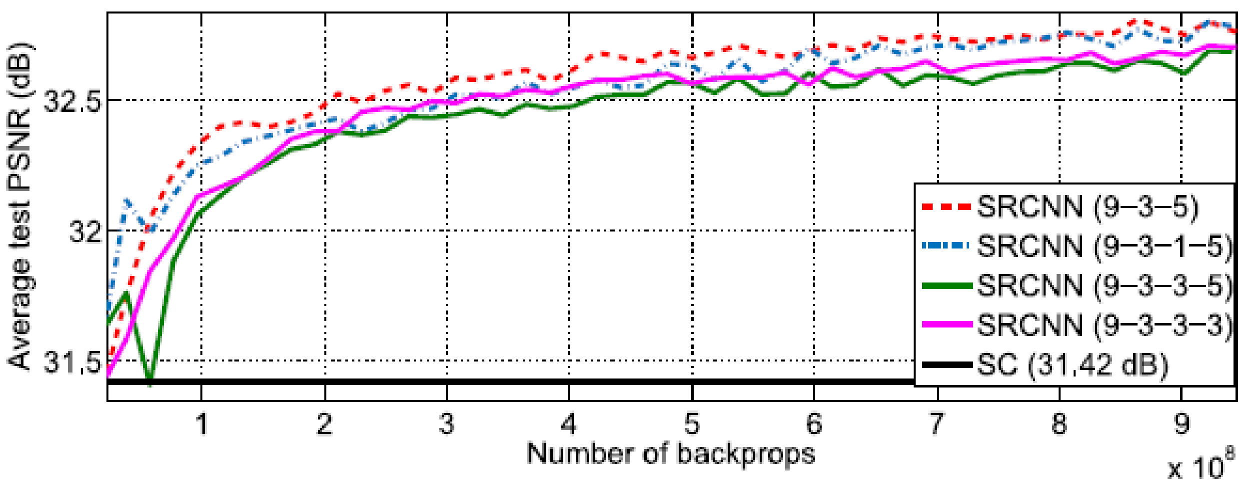
- A larger filter size leads to better results.



- Deeper structure does not always lead to better results.

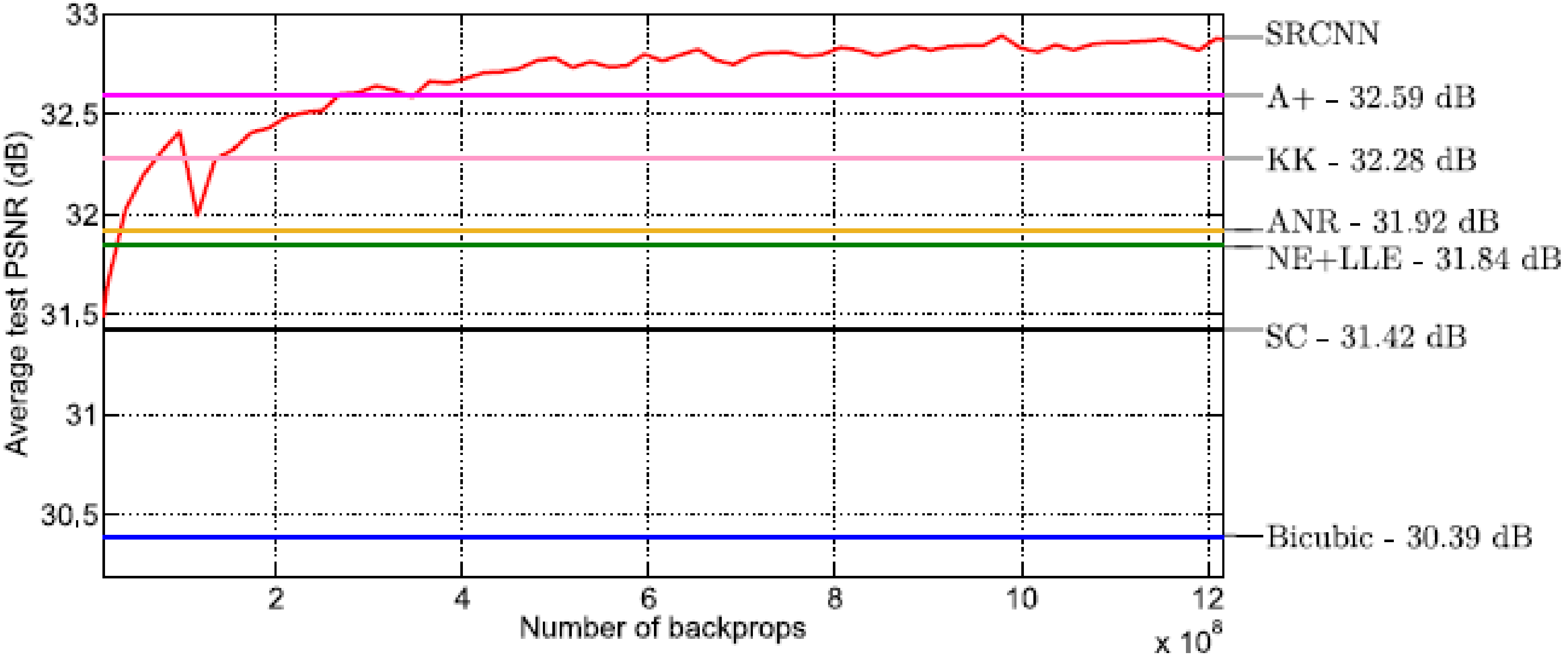


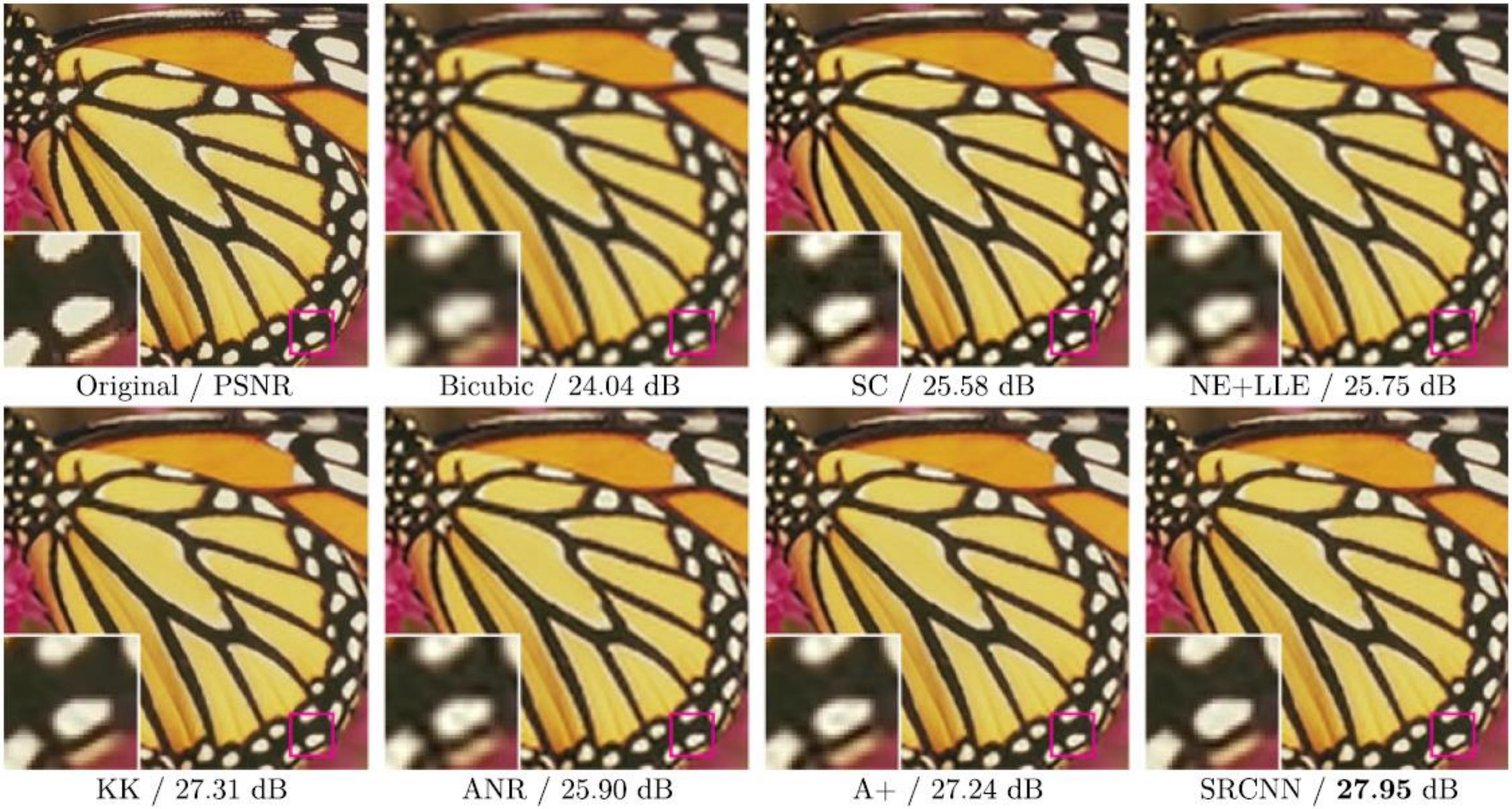
(a) 9-1-1-5 ($n_{22} = 32$) and 9-1-1-1-5 ($n_{22} = 32, n_{23} = 16$)

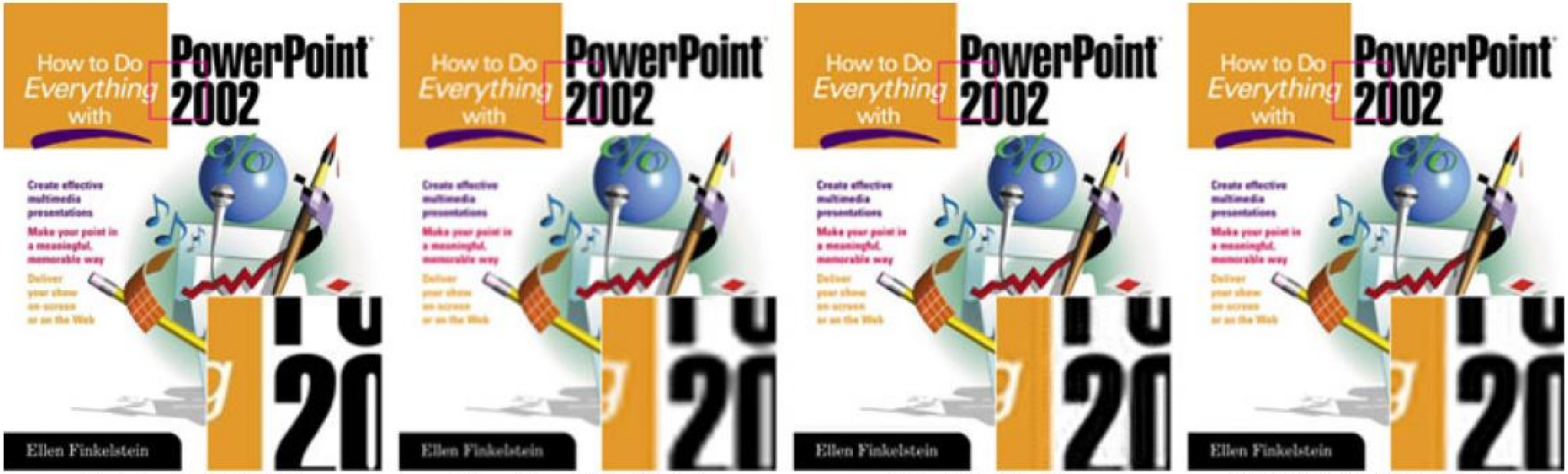


(b) 9-3-3-5 and 9-3-3-3

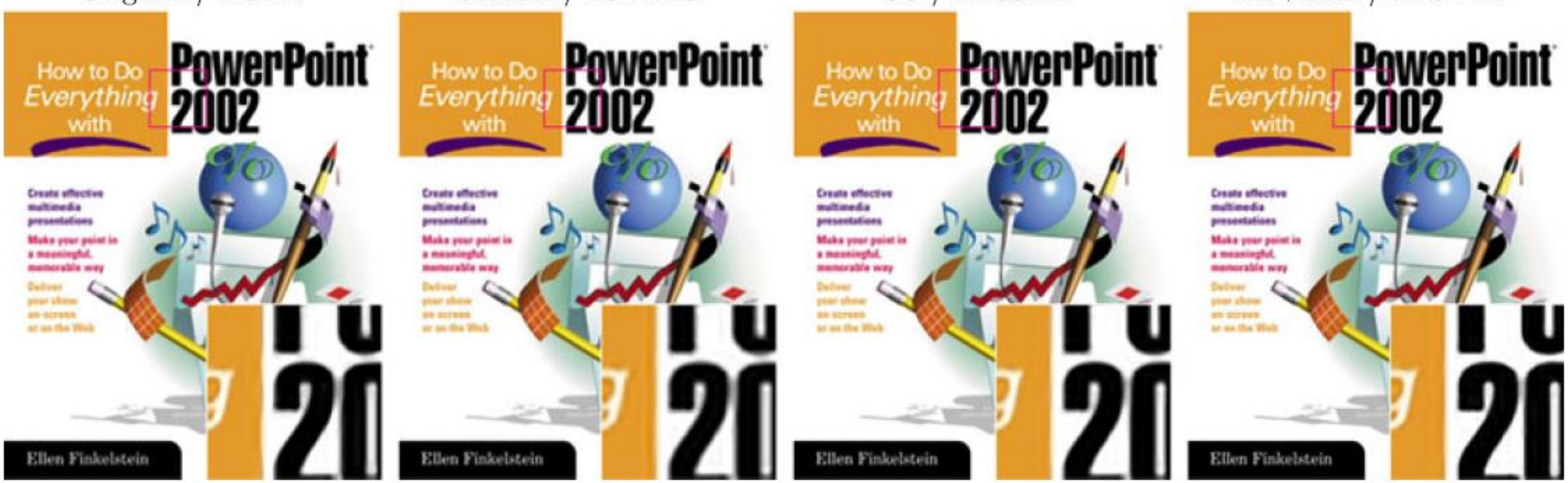
- Compare our method with recent state-of-the-arts



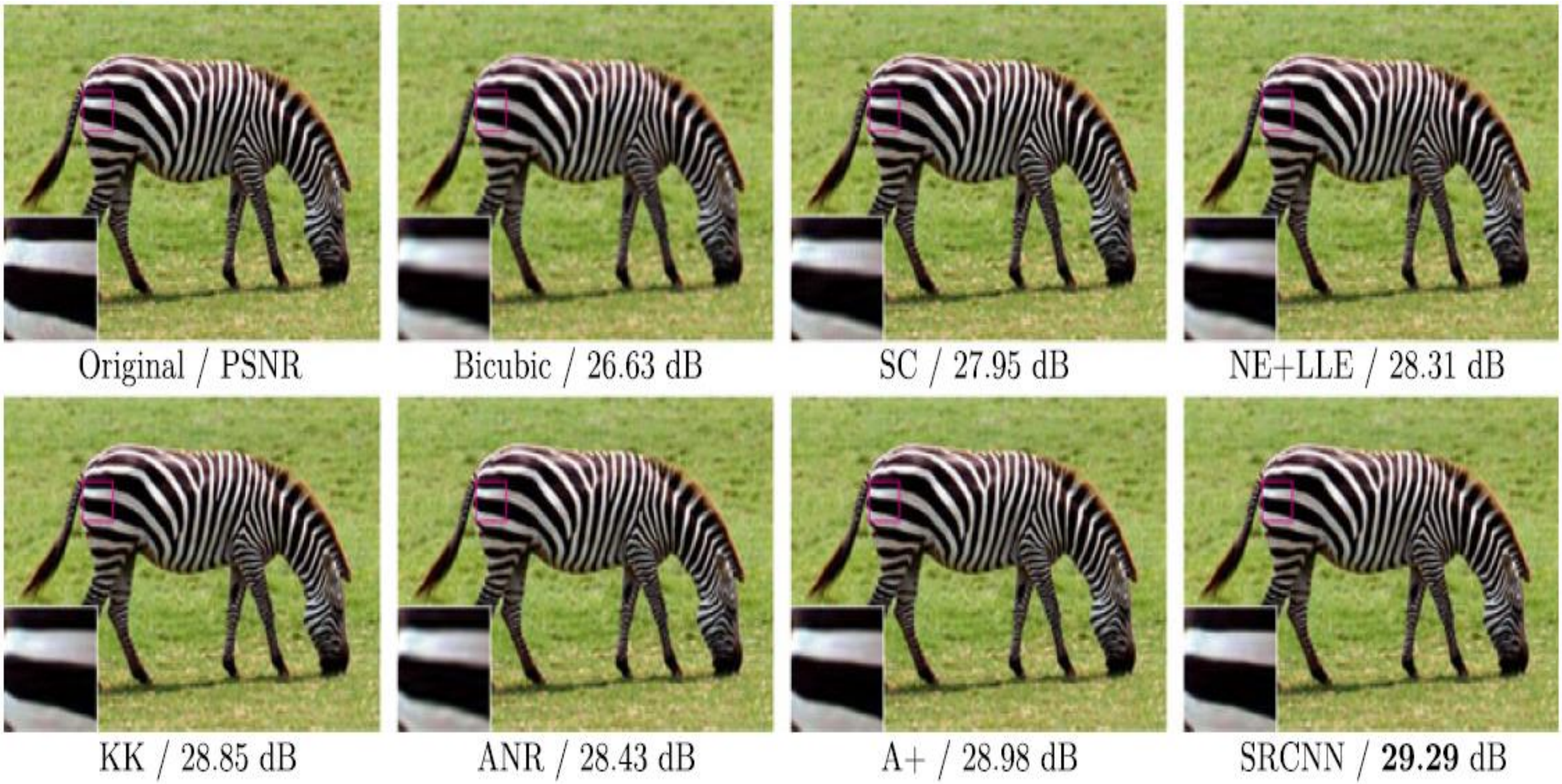

Original / PSNR
Bicubic / 24.04 dB
SC / 25.58 dB
NE+LLE / 25.75 dB
KK / 27.31 dB
ANR / 25.90 dB
A+ / 27.24 dB
SRCNN / **27.95** dB



Original / PSNR | Bicubic / 23.71 dB | SC / 24.98 dB | NE+LLE / 24.94 dB



KK / 25.60 dB | ANR / 25.03 dB | A+ / 26.09 dB | SRCNN / **27.04** dB


Original / PSNR
Bicubic / 26.63 dB
SC / 27.95 dB
NE+LLE / 28.31 dB
KK / 28.85 dB
ANR / 28.43 dB
A+ / 28.98 dB
SRCNN / **29.29** dB

4. Conclusion

- We show that conventional sparse-coding-based SR methods can be reformulated into a deep convolutional neural network.

- Besides, the proposed structure, with its advantages of simplicity and robustness, could be applied to other low-level vision problems, such as image deblurring or simultaneous SR+denoising