Rakuten Institute of Technology
http://rit.rakuten.co.jp
rakutenRIT

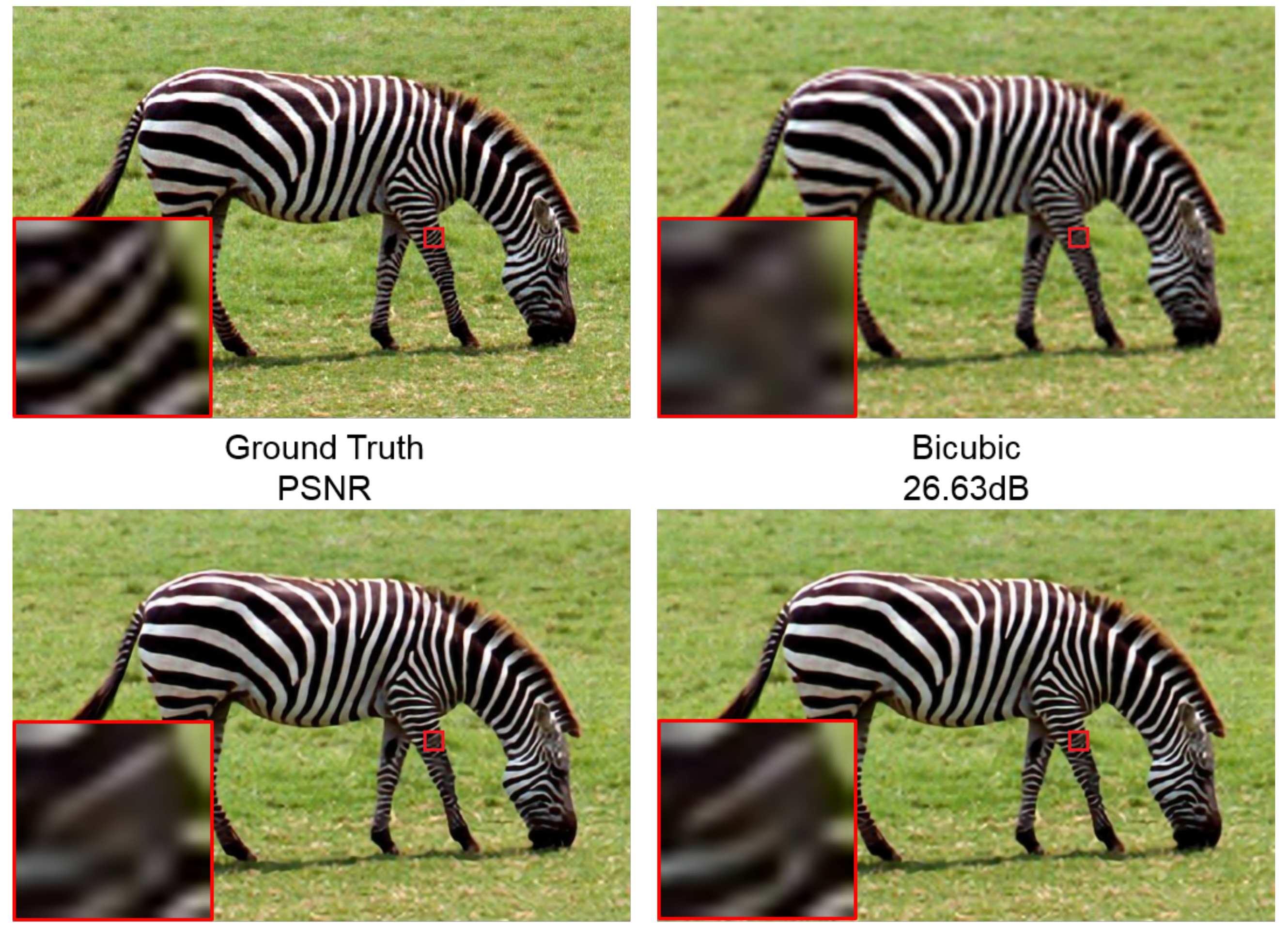
ByNet-SR: Image Super Resolution with a Bypass Connection Network

Jiu Xu *Yeongnam Chae* *Björn Stenger*

Rakuten Institute of Technology, Boston *Rakuten Institute of Technology, Tokyo*

Abstract

This paper proposes a deep residual network, ByNet, for the single image super resolution task. Two original network components are introduced, which increase performance and speed compared to VDSR [10] and are easy to implement. Experiments on standard benchmarks show that the proposed method achieves state of the art results over multiple scales in terms of PSNR and structural similarity (SSIM).

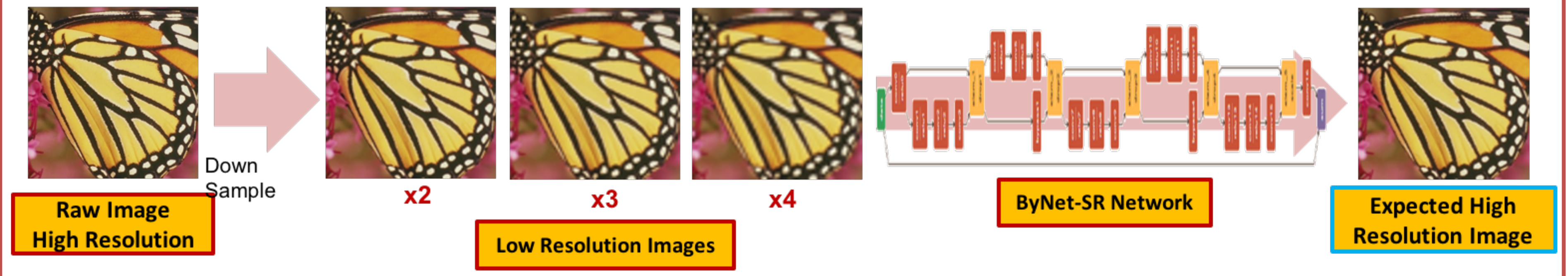


Ground Truth PSNR | Bicubic 26.63dB | VDSR 29.49dB | ByNet5 (Ours) 29.85dB

Example result. *A region in the zebra image shows improved recovery of detail (at 3× upscaling).*

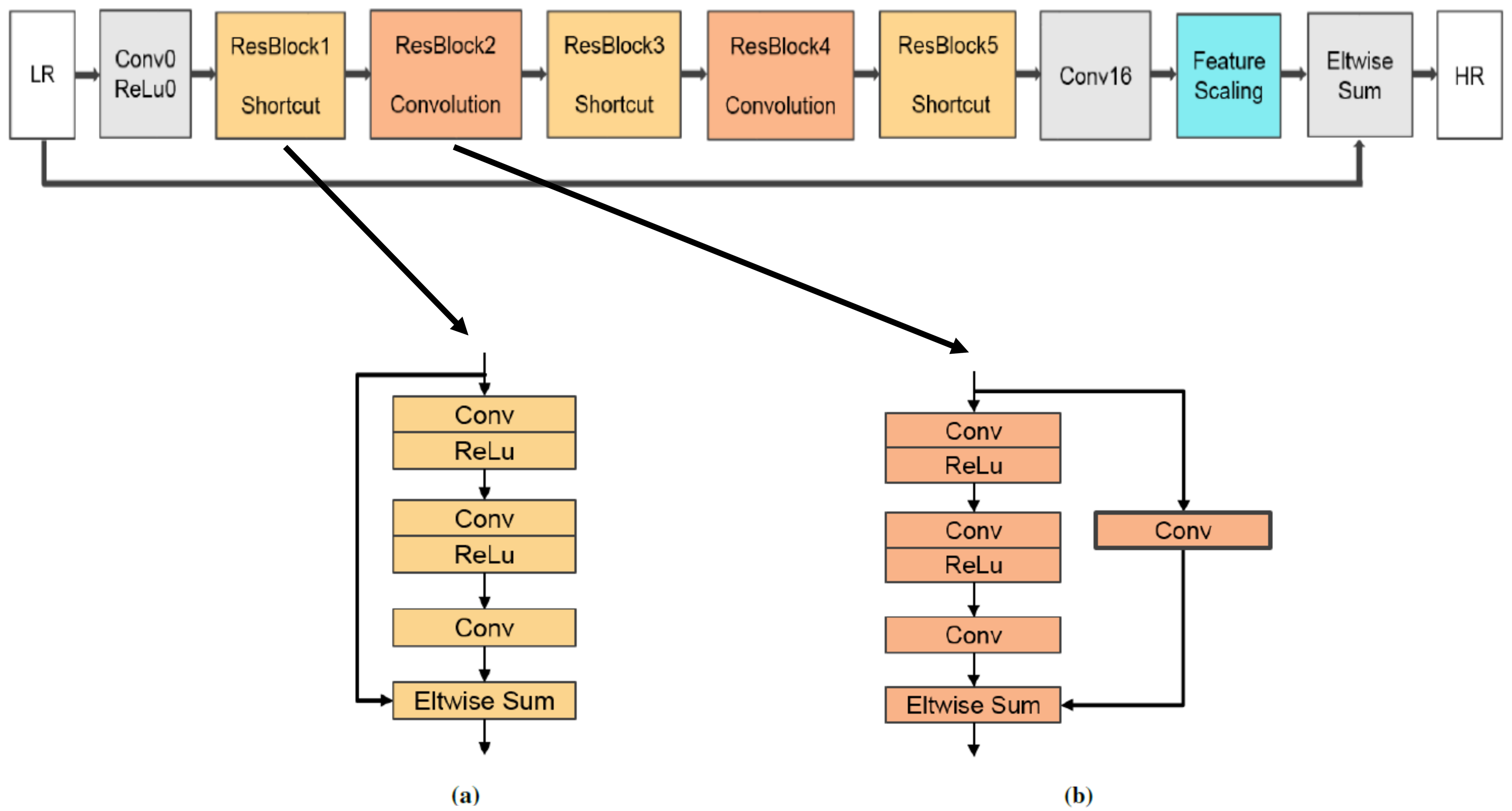
Training

LR images are obtained by down-sampling HR images with the scale factors of 2,3, and 4. These sets of LR images are merged and shuffled for training which allows our model to naturally handle multiple scale factors.



The mean squared error (MSE) between raw high resolution and expected high resolution image is defined as loss

ByNet5 Network Architecture：



$$y = f(h_t(x)) + h_t(x), \qquad y = f(h_t(x)) + w_k * h_t(x).$$

$$f(h_t(x)) = w_{t+3} * (\max(0, w_{t+2} * \max(0, w_{t+1} * h_t(x))))$$

Bypass Connections

(a) Feature bypass with shortcut connection

Merit: Improve convergence properties and achieve higher accuracy within the same training epoch.

(b) Feature bypass with convolutional connection

Merit: Element-wise addition layer sums the features learned from different receptive fields of input help to further performance.

Feature Scaling

One additional parameter that is learned during training for scaling the weights value.

Merit: Feature scaling improves convergence by scaling the network output to efficiently fit the distribution of image residuals.

Analysis of Residual Block Arrangement

The two types of residual blocks can be arranged in various ways.

Dataset	No ByPass Connections	00100 Connection Order	01010 Connection Order	10101 Connection Order	11111 Connection Order
Set5	*37.55dB* *33.80dB* *31.42dB*	*37.60dB* *33.79dB* *31.46dB*	*37.60dB* *33.81dB* *31.47dB*	*37.59dB* *33.77dB* *31.44dB*	*37.59dB* *33.78dB* *31.46dB*

Evaluate of different symmetric arrangements of five bypass blocks in sequence, with "0" represents shortcut connection and "1" represents convolutional connection

Comparison with State of the Art

We compare ByNet with the reported results of the following methods on the same training and test sets: A+ [24], RFL [25], SRCNN [16], and VDSR [10]

Table 2: Performance Comparison on benchmark datasets: PSNR and SSIM are averaged over all images for each scale. The proposed ByNet model consistently achieves the best PSNR and SSIM results. Adding more blocks improves performance. Results for the recent VDSR method [10] are highlighted in blue for easy visual comparison.

Dataset	Scale	Bicubic PSNR	A+[24] PSNR	RFL[25] PSNR	SRCNN[16] PSNR	VDSR[10] PSNR	ByNet5 PSNR	ByNet7 PSNR	ByNet9 PSNR	Bicubic SSIM	A+[24] SSIM	RFL[25] SSIM	SRCNN[16] SSIM	VDSR[10] SSIM	ByNet5 SSIM	ByNet7 SSIM	ByNet9 SSIM
Set5	×2	33.66	36.54	36.54	36.66	37.53	37.60	37.69	**37.74**	0.9299	0.9544	0.9537	0.9542	0.9587	0.9594	0.9598	**0.9599**
	×3	30.39	32.58	32.43	32.75	33.66	33.85	33.89	**33.96**	0.8682	0.9088	0.9057	0.9090	0.9213	0.9237	0.9241	**0.9246**
	×4	28.42	30.28	30.14	30.48	31.35	31.49	31.54	**31.60**	0.8104	0.8603	0.8548	0.8628	0.8838	0.8862	0.8871	**0.8886**
Set14	×2	30.24	32.28	32.26	32.42	33.03	33.13	33.21	**33.24**	0.8688	0.9056	0.9040	0.9063	0.9124	0.9138	0.9144	**0.9147**
	×3	27.55	29.13	29.05	29.28	29.77	29.92	29.94	**29.96**	0.7742	0.8188	0.8164	0.8209	0.8314	0.8342	0.8350	**0.8354**
	×4	26.00	27.32	27.24	27.49	28.01	28.20	28.20	**28.24**	0.7027	0.7491	0.7451	0.7503	0.7674	0.7716	0.7722	**0.7732**
BSD100	×2	29.56	31.21	31.16	31.36	31.90	31.92	31.96	**32.00**	0.8431	0.8863	0.8840	0.8879	0.8960	0.8967	0.8973	**0.8977**
	×3	27.21	28.29	28.22	28.41	28.82	28.86	28.89	**28.91**	0.7385	0.7835	0.7806	0.7863	0.7976	0.7993	0.8001	**0.8808**
	×4	25.96	26.82	26.75	26.90	27.29	27.31	27.34	**27.37**	0.6675	0.7087	0.7054	0.7101	0.7251	0.7267	0.7277	**0.7285**

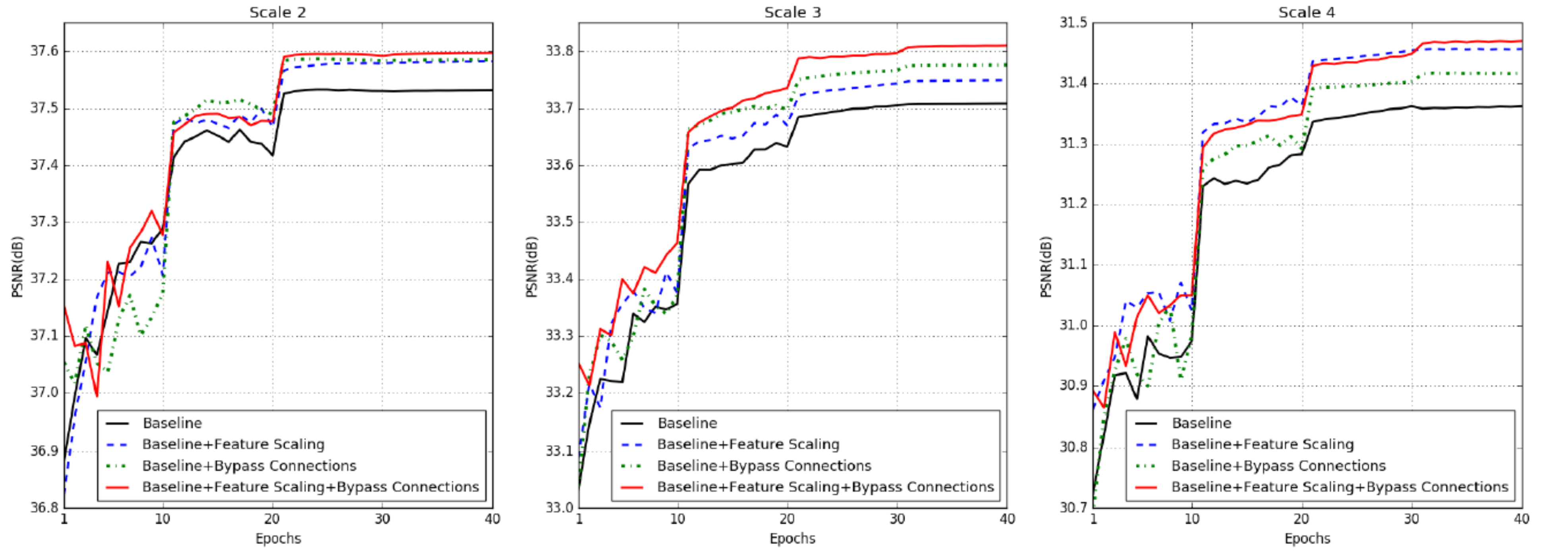


Fig. 4: Ablation study. The convergence curves for the *Set5* training set show that adding both feature scaling and bypass connections consistently lead to increased PSNR across different scales.

[24] R. Timofte, V. De Smet, and L. Van Gool, "A+: Adjustedanchored neighborhood regression for fast superresolution,"in ACCV, 2014.
[25] S. Schulter, C. Leistner, and H. Bischof, "Fast and accurate image upscaling with super-resolution forests," in CVPR, 2015.
[16] C. Dong, C.C. Loy, K. He, and X. Tang, "Image super-resolution using deep convolutional networks," in TPAMI, 2015, vol. 38, pp. 295–307.
[10] J. Kim, J.K. Lee, and K.M. Lee, "Accurate image superresolution using very deep convolutional networks," in CVPR, 2016.