

Learning Dual Convolutional Neural Networks for Low-Level Vision

Supplemental Material

Jinshan Pan¹ Sifei Liu² Deqing Sun² Jiawei Zhang³ Yang Liu⁴ Jimmy Ren⁵
Zechao Li¹ Jinhui Tang¹ Huchuan Lu⁴ Yu-Wing Tai⁶ Ming-Hsuan Yang⁷

¹Nanjing University of Science and Technology ²NVIDIA ³City University of Hong Kong
⁴Dalian University of Technology ⁵SenseTime Research ⁶Tencent Youtu Lab ⁷UC Merced

<https://sites.google.com/site/jspanhomepage/dualcnn>

Overview

In this document, we first analyze the effect of the proposed loss function in Section 1, and then add more analysis of the proposed dual network in Section 2. Finally, we show more visual comparisons in Section 3.

1. Effect of the Proposed Loss Function

In Table 7 of the manuscript, we quantitatively evaluate the effects of different loss functions on image dehazing. To better understand the effects, we show more visual comparison results generated by different loss functions. Figure 1 shows the evaluation results on image dehazing. The dehazed results by the proposed method using only the loss function \mathcal{L}_x contain distorted color and the results by the proposed method using \mathcal{L}_x and \mathcal{L}_d or \mathcal{L}_x and \mathcal{L}_s contain halo effects. In contrast, the results generated by the proposed method using the proposed loss functions are much better.

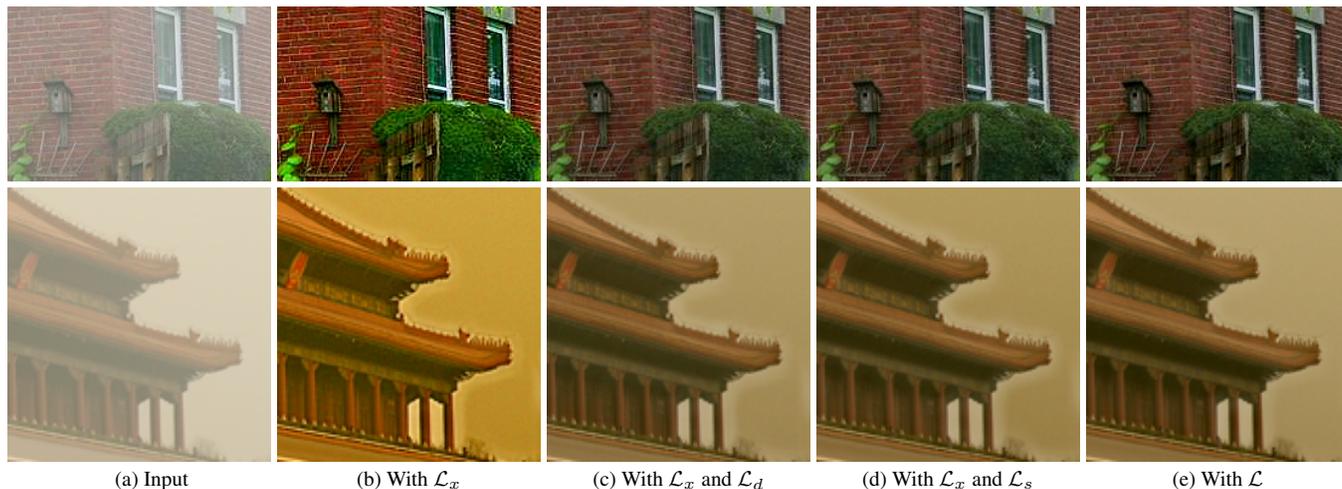


Figure 1. The effect of the proposed loss function on image dehazing. The results generated by the proposed method with only \mathcal{L}_x contain color distortion, while the dehazed results by the method with \mathcal{L}_x and \mathcal{L}_d or \mathcal{L}_x and \mathcal{L}_s contain halo effects (best viewed on high-resolution displays).

2. More Analysis of the Proposed DualCNN

As mentioned in the manuscript, our algorithm is motivated by the decomposition of a signal into structures and details and thus needs to learn both structures and details. To validate whether the proposed algorithm is able to estimate the details and structures, we show some results from the Net-S and Net-D of the proposed algorithm on the image super-resolution, image denoising, image filtering, and image deraining problems. Figures 2, 3, 4 and 5 show that the DualCNN is able to estimate the details and structures, thus facilitating the estimations of the image super-resolution, image denoising, image filtering, and image deraining.

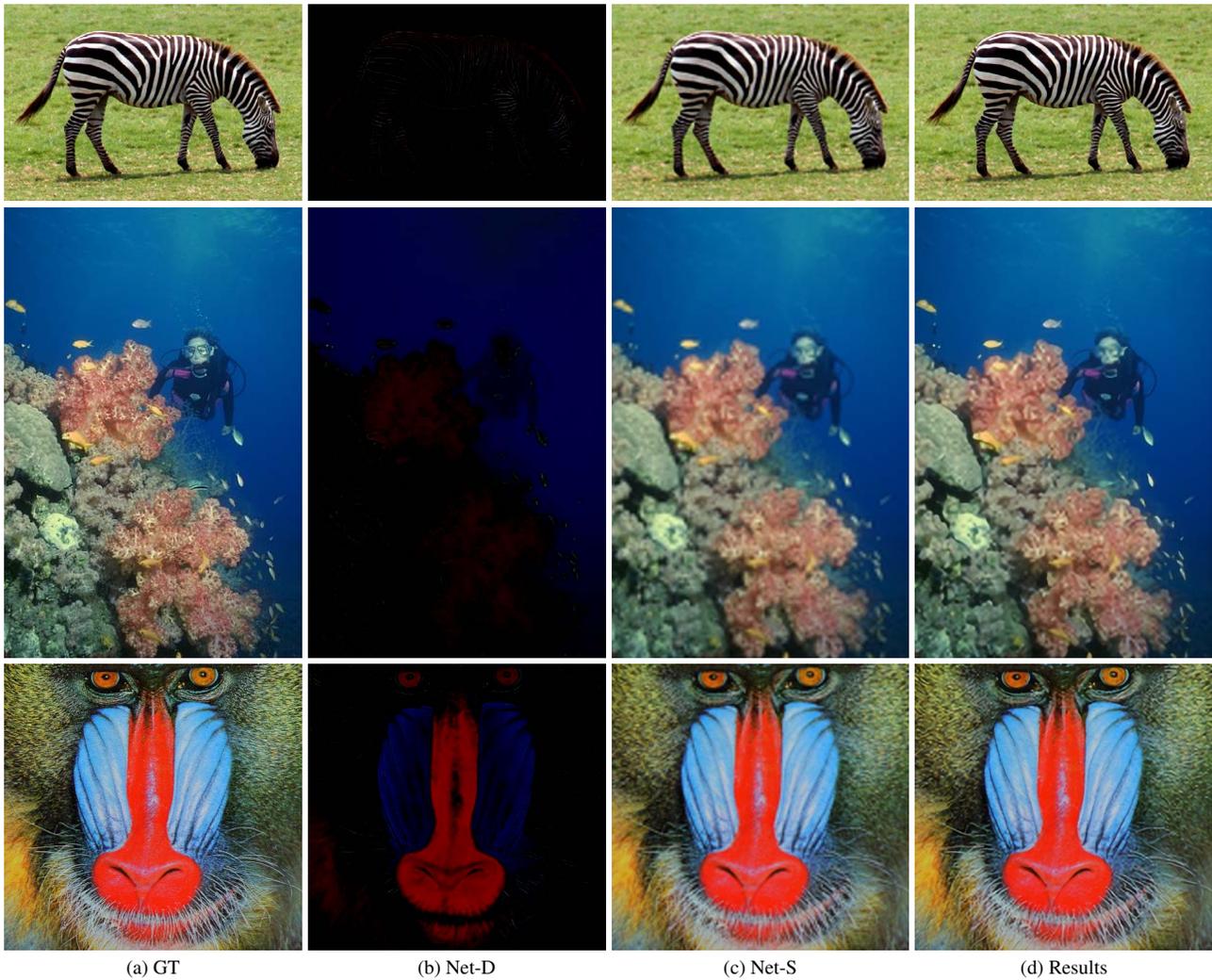


Figure 2. On the image super-resolution problem (×3). The proposed DualCNN is able to estimate the details and structures from the input images, thus leading to the high-quality results with finer details (best viewed on high-resolution displays).



(a) GT

(b) Net-D

(c) Net-S

(d) Results

Figure 3. On the image denoising problem. The proposed DualCNN is able to estimate the details and structures from the input images, thus leading to the clean images with finer details (best viewed on high-resolution displays).

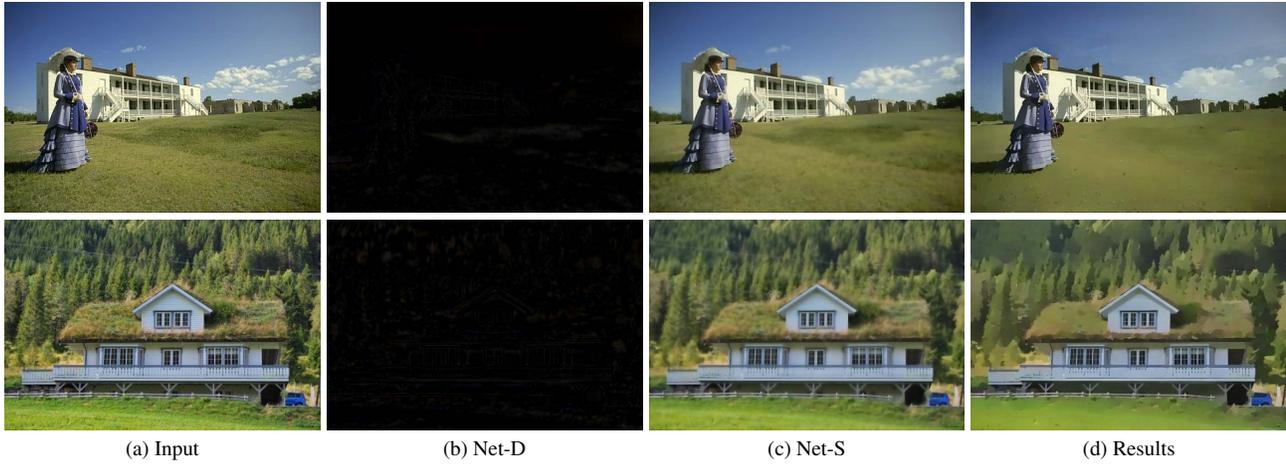


Figure 4. On the image filtering problem [5]. The proposed DualCNN is able to estimate the details and structures from the input images. For this problem, the pixel values of the images from Net-D are almost negative. Thus, based on the formation, the proposed method is able to generate the results which are close to the ground truths.

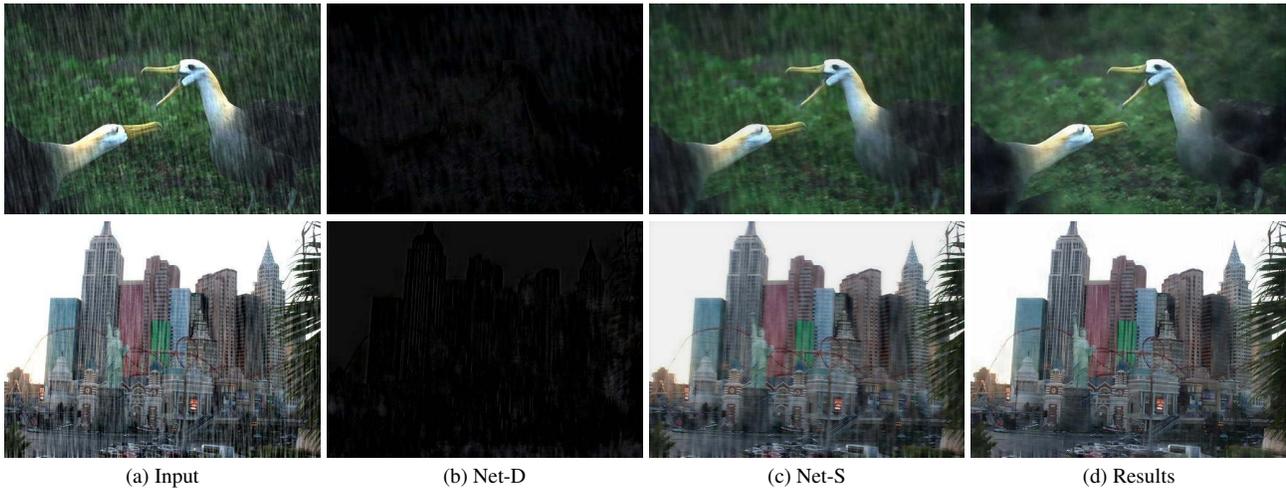


Figure 5. On the image deraining problem. The proposed DualCNN is able to estimate the details (including both rainy streaks and details) and structures from the input images and generates the clean results.

3. More Experimental Results

In this section, we provide more visual comparisons with state-of-the-art methods on each individual application.

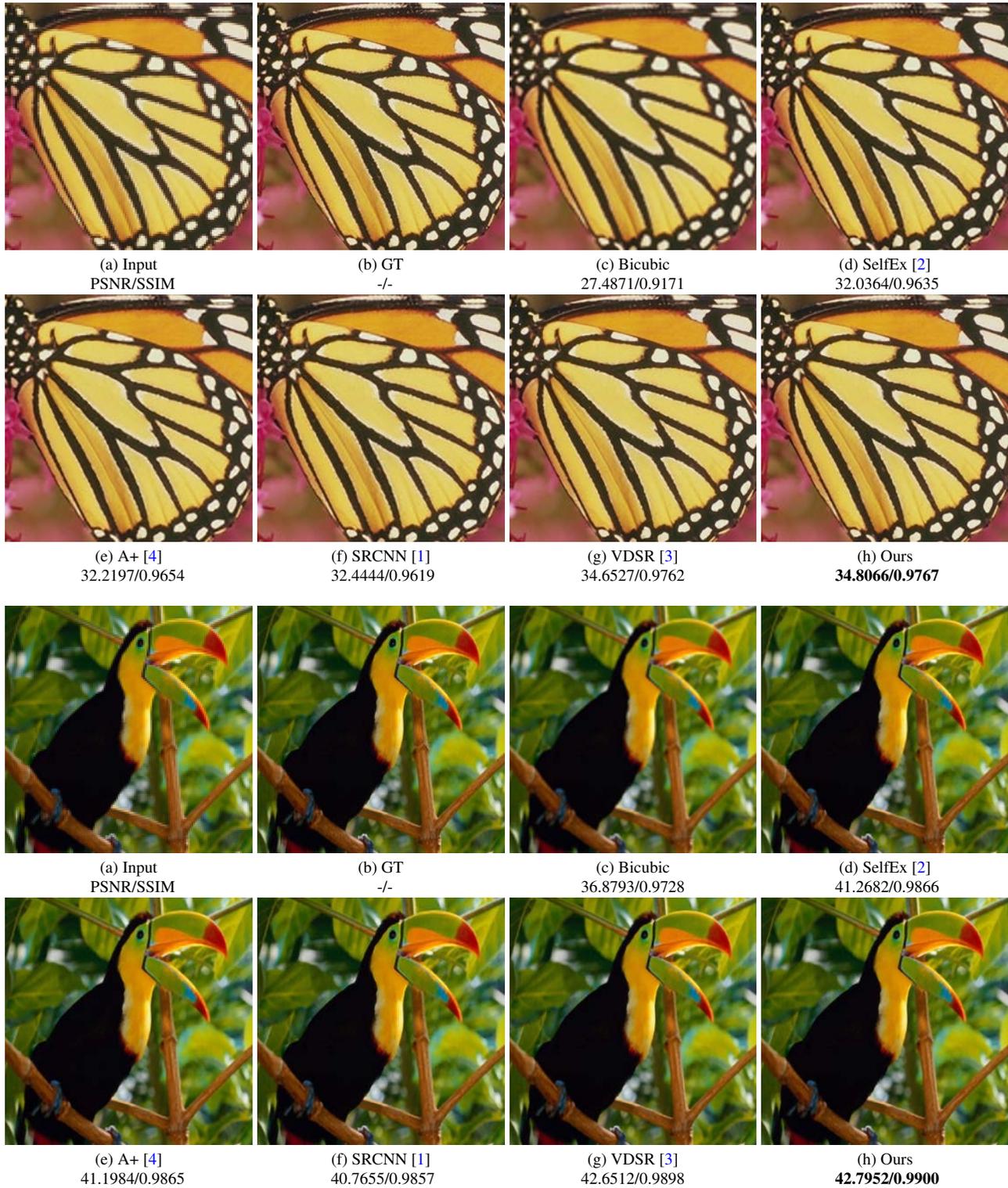


Figure 6. Visual comparisons for super-resolution ($\times 2$). The proposed method generates the results with higher PSNR and SSIM values (best viewed on high-resolution displays).



(a) Input
PSNR/SSIM

(b) GT
-/-

(c) Bicubic
28.7716/0.8922

(d) SelfEx [2]
31.5731/0.9321



(e) A+ [4]
31.4068/0.9302



(f) SRCNN [1]
31.1108/0.9260



(g) VDSR [3]
32.5045/0.9424



(h) Ours
32.5770/0.9429



(a) Input
PSNR/SSIM

(b) GT
-/-

(c) Bicubic
23.9357/0.6801

(d) SelfEx [2]
35.1135/0.9223



(e) A+ [4]
35.1371/0.9223



(f) SRCNN [1]
25.8626/0.7443



(g) VDSR [3]
26.6169/0.7698



(h) Ours
26.7200/0.7701

Figure 7. Visual comparisons for super-resolution ($\times 3$). The proposed method generates the results with higher PSNR and SSIM values (best viewed on high-resolution displays).



(a) Input
PSNR/SSIM

(b) GT
-/-

(c) Bicubic
25.1355/0.6859

(d) SelfEx [2]
25.6412/0.7273



(e) A+ [4]
25.6964/0.7266



(f) SRCNN [1]
25.6464/0.7208



(g) VDSR [3]
25.7124/0.7377



(h) Ours
25.8678/0.7389



(a) Input
PSNR/SSIM



(b) GT
-/-



(c) Bicubic
33.8572/0.9039



(d) SelfEx [2]
26.2477/0.7603



(e) A+ [4]
25.8823/0.7537



(f) SRCNN [1]
34.9628/0.9212

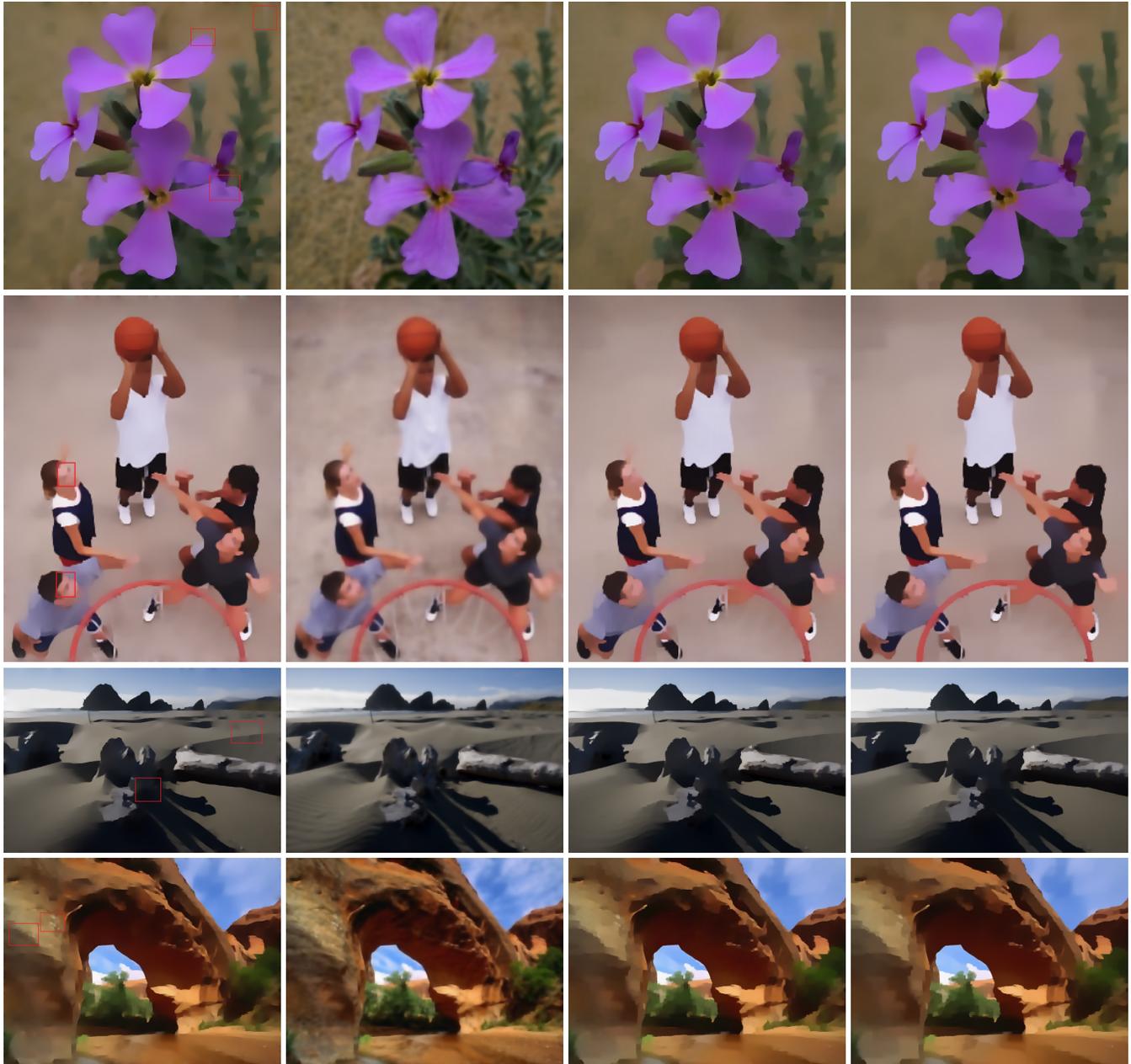


(g) VDSR [3]
35.3354/0.9264



(h) Ours
35.3700/0.9268

Figure 8. Visual comparisons for super-resolution ($\times 4$). The proposed method generates the results with higher PSNR and SSIM values (best viewed on high-resolution displays).



(a) Xu et al. [6]

(b) Net-S

(c) Ours

(d) RTV

Figure 9. Visual comparisons for learning RTV image filters on the test dataset of [6]. The deep learning based method [6] is not able to remove the details and structures that are supposed to be removed (the red boxes in (a)). The proposed method generates the results that are visually similar to the ground truths in (d) (best viewed on high-resolution displays).



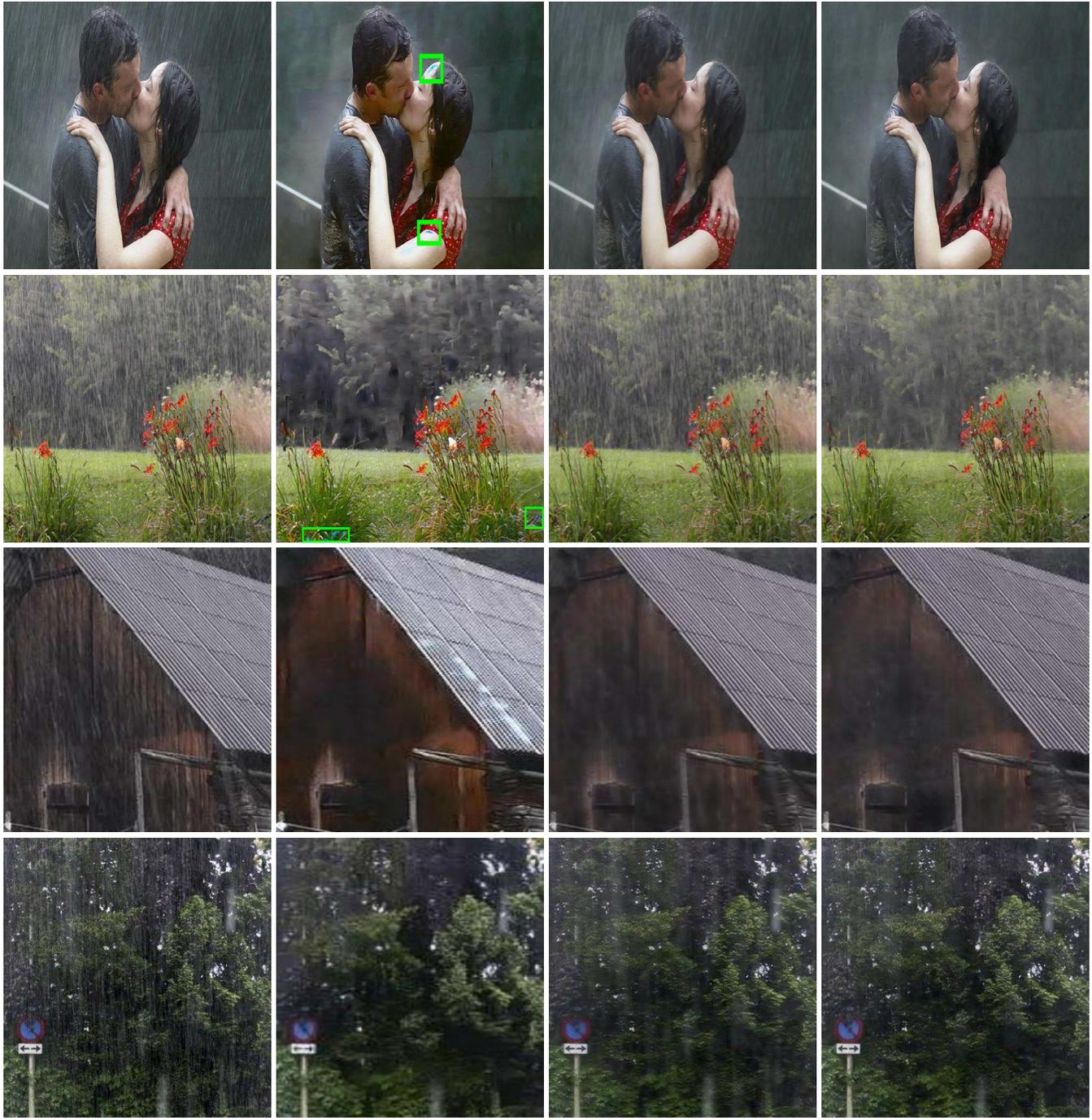
(a) Inputs

(b) ID-GAN [7]

(c) Net-S

(d) Ours

Figure 10. Visual comparisons for image deraining on the test dataset by Zhang et al. [7]. The deep learning based method [7] is not able to remove the rainy streaks. The details of the recovered images are not preserved well. The proposed method generates the much clearer images with finer details (best viewed on high-resolution displays).



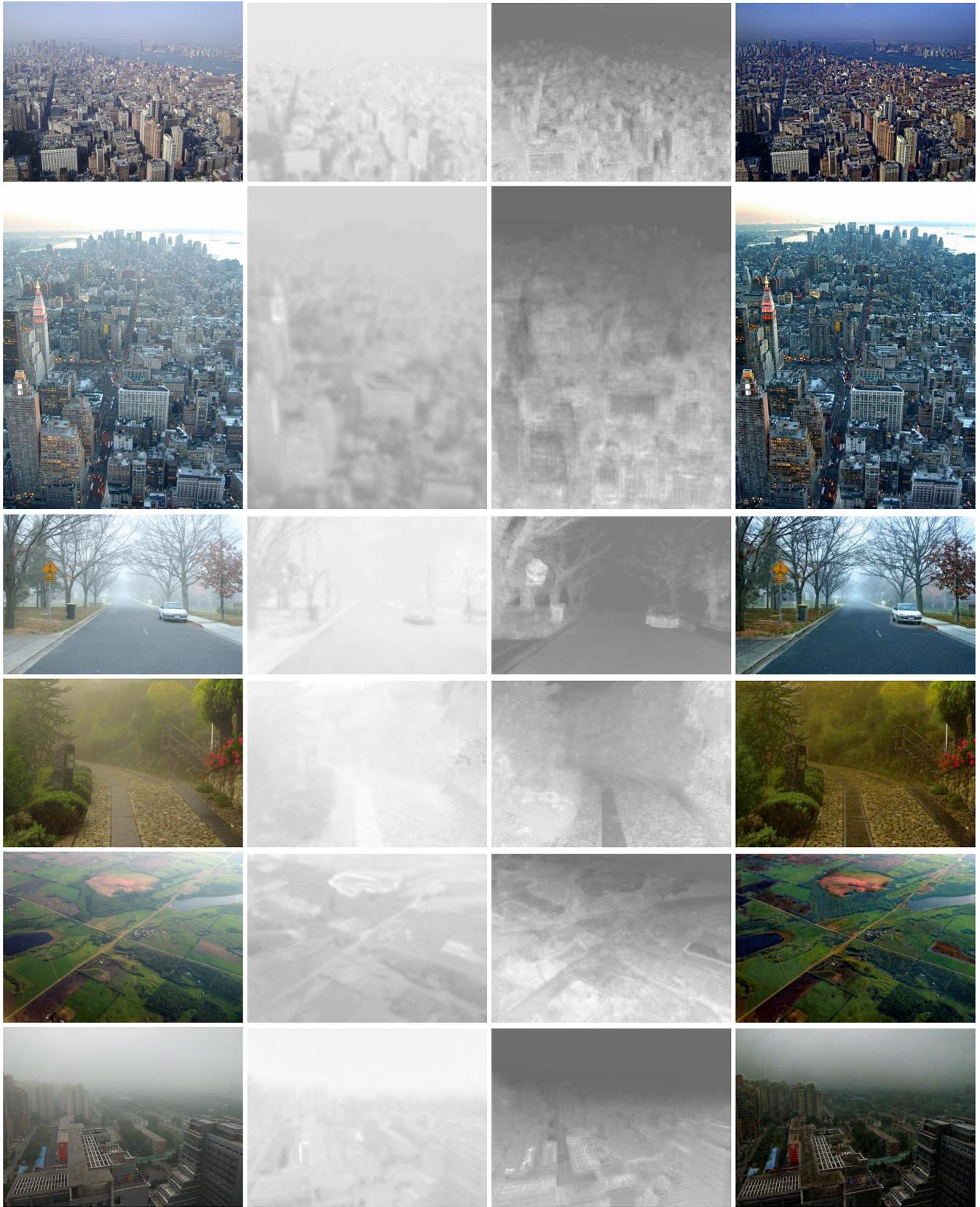
(a) Inputs

(b) ID-GAN [7]

(c) Net-S

(d) Ours

Figure 11. Visual comparisons for image deraining on the real images. The deep learning based method [7] does not preserve the details well and introduces some fake structures (the green boxes in (b)). The proposed method generates much clearer images with finer details (best viewed on high-resolution displays).



(a) Inputs

(b) Atmospheric light

(c) Transmission map

(d) Ours

Figure 12. Image dehazing on the real images. Our method directly estimates the atmospheric light (b) and the transmission map (c) and generates clear images.

References

- [1] C. Dong, C. C. Loy, K. He, and X. Tang. Learning a deep convolutional network for image super-resolution. In *ECCV*, pages 184–199, 2014. [5](#), [6](#), [7](#)
- [2] J.-B. Huang, A. Singh, and N. Ahuja. Single image super-resolution from transformed self-exemplars. In *CVPR*, pages 5197–5206, 2015. [5](#), [6](#), [7](#)
- [3] J. Kim, J. K. Lee, and K. M. Lee. Accurate image super-resolution using very deep convolutional networks. In *CVPR*, pages 1646–1654, 2016. [5](#), [6](#), [7](#)
- [4] R. Timofte, V. D. Smet, and L. J. V. Gool. A+: adjusted anchored neighborhood regression for fast super-resolution. In *ACCV*, pages 111–126, 2014. [5](#), [6](#), [7](#)
- [5] L. Xu, C. Lu, Y. Xu, and J. Jia. Image smoothing via L_0 gradient minimization. *ACM TOG*, 30(6):174:1–174:12, 2011. [4](#)
- [6] L. Xu, J. S. J. Ren, Q. Yan, R. Liao, and J. Jia. Deep edge-aware filters. In *ICML*, pages 1669–1678, 2015. [8](#)
- [7] H. Zhang, V. Sindagi, and V. M. Patel. Image de-raining using a conditional generative adversarial network. *CoRR*, abs/1701.05957, 2017. [9](#), [10](#)