

Supplementary materials for "Deep Learning for image deconvolution in the presence of kernel/model uncertainty"

Yuesong Nan and Hui Ji

Department of Mathematics, National University of Singapore, 119076, Singapore

nanyuesong@u.nus.edu and matjh@nus.edu.sg

1. Overview

The supplementary material is organized as follows. In Section 2, we showed some ablation studies on the design of dual-path UNet for predicting correction term, motivated from TLS-based formulation of the problem. Then the visualization of correction term \mathbf{u} for some example is given. In Section 3, more experiments are conducted for better understanding of our proposed methods. Next, Section 4 showed the visual comparison of more real images without ground truth. Section 5 is devoted to visual comparison of some synthetic images with ground truths.

2. Additional study and visualization on DP-Unet for predicting correction term

Discussion on dual inputs for DP-Unet Based on the EIV model, The prediction on correction term is determined by two terms: the residual $\mathbf{r}^{(t)} = \mathbf{y} - \hat{\mathbf{k}} \otimes \mathbf{x}^{(t)}$ and the estimate $\mathbf{x}^{(t)}$. These two inputs are combined after down-sampling by the proposed DP-Unet module. See Section 3.3 and Fig 2 in main paper for the architecture.

In the work proposed in [2], their model also contains a ℓ_1 -norm-regularized term for representing correction term. However, the estimate of such a term is only dependent on the residual. In contrast, ours depends on both the residual and current estimate $\mathbf{x}^{(t)}$. An ablation study is conducted to check whether $\mathbf{x}^{(t)}$ is important to the prediction as our mathematical formulation indicated. See Table 1 for such a comparison on the DP-Unet and the same NN but with the path related to $\mathbf{x}^{(t)}$ being removed. It showed that the DP-Unet with dual inputs does not show large improvement on Levin *et al.* [7] whose kernel error is small, and show large improvement on Lai *et al.*'s dataset whose kernel error is much larger. This study justifies the need for dual-path of the UNet which takes as the input both the residual $\mathbf{r}^{(t)} = \mathbf{y} - \hat{\mathbf{k}} \otimes \mathbf{x}^{(t)}$ and the estimate $\mathbf{x}^{(t)}$.

Visualization of learned correction term See Fig. 1 2 for the visualization of learned correction terms \mathbf{u} of one example, in comparison with ground truth correction term $\Delta\mathbf{k} \otimes \mathbf{x}$. It can be seen that learned correction term really indicates the property of ground truth correction term, which verifies that our proposed DP-UNet is able to reduce the negative effect of erroneous kernels in non-blind deconvolution scheme.

Table 1: Average PSNR(dB)/SSIM of deblurring results for DP-Unet with single input \mathbf{r} vs dual inputs $\mathbf{r} + \mathbf{x}$.

	[1]	[8]	[11]	[13]
Levin <i>et al.</i>	30.98/0.90	31.06/0.92	34.61/0.96	33.30/0.94
Input \mathbf{r}	30.92/0.90	31.14/0.92	34.66/0.96	33.36/0.94
Sun <i>et al.</i>	[1]	[16]	[9]	
Input \mathbf{r}	31.85/0.92	32.57/0.93	31.40/0.90	
Input $\mathbf{r} + \mathbf{x}$	32.12/0.92	32.76/0.93	31.60/0.90	
Lai <i>et al.</i>	[16]	[17]	[13]	[12]
Input \mathbf{r}	24.27/0.80	23.88/0.79	24.30/0.79	22.88/0.75
Input $\mathbf{r} + \mathbf{x}$	24.81/0.81	24.46/0.80	24.78/0.80	23.22/0.76

3. Additional Studies on the proposed algorithm.

Evaluation on the ground truth kernels Our proposed deblurring method aims to tackle on deblurring with *kernel/model errors*. But it is interesting to check its performance when kernels are *perfect*. Table 2 shows such experiments in comparison



Figure 1: Visualization of the output of DP-Unet \mathbf{u} in Stage S_4 of one example, in comparison with the ground truth correction term $\Delta\mathbf{k} \otimes \mathbf{x}$. (a) Ground truth kernel. (b) Noisy kernel. (c) Sharp image \mathbf{x} . (d) Blurry image \mathbf{y} . (e) Recovered image $\hat{\mathbf{x}}$. (f) Ground truth correction term. $\Delta\mathbf{k} \otimes \mathbf{x}$. (g) Predicted correction term $\mathbf{u}^{(4)}$.

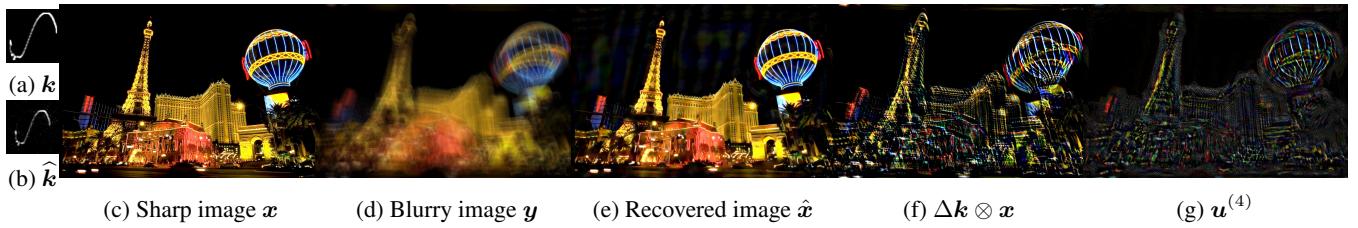


Figure 2: Visualization of the output of DP-Unet \mathbf{u} in Stage S_4 of one example, in comparison with the ground truth correction term $\Delta\mathbf{k} \otimes \mathbf{x}$. (a) Ground truth kernel. (b) Noisy kernel. (c) Sharp image \mathbf{x} . (d) Blurry image \mathbf{y} . (e) Recovered image $\hat{\mathbf{x}}$. (f) Ground truth correction term. $\Delta\mathbf{k} \otimes \mathbf{x}$. (g) Predicted correction term $\mathbf{u}^{(4)}$.

with the deep-learning benchmark Zhang *et al.* [18]. In this study, we use the same datasets as main paper but use ground-truth kernels. The noise level is set to be 1%. We can observe that our method perform roughly the same as Zhang *et al.*'s.

Table 2: Average PSNR(dB)/SSIM of the results, in comparison to the deep learning benchmark Zhang-17's [18] when using ground-truth kernels.

Dataset	Levin <i>et al.</i>	Sun <i>et al.</i>	Lai <i>et al.</i>
Zhang-17 [18]	31.74/0.91	32.12/0.88	22.97/ 0.82
Ours	32.88/0.93	31.81/ 0.88	23.09/0.75

Computational Efficiency The experiments were conducted on a workstation with a 3.2GHz Intel Xeon E5-2620 v4 CPU, 64G RAM and a GeForce GTX 2080 Ti GPU. The training time took around 72 hours. The comparison of average testing time for images with size 256×256 is shown in the Table 3. It is shown that our method is the second fastest method among all, which indicates the practical usage.

Table 3: The comparison of average testing time (s) for non-blind deconvolution methods when deconvoluting 256×256 images

Method	Krishnan-09[3]	Ji-12[2]	Zoran-11[21]	Whyte-14[15]	Kruse-17[5]	Zhang-17[18]	Zhang-17[19]	Ours
Time(s)	0.26	18.8	105.18	0.75	0.15	0.02	0.31	0.09

4. Visual comparisons on real images

As there is no ground truth for real images. we showed visual comparison of many examples on real images. See Fig 3 - 14 for the comparison of the results on Lai *et al.*'s dataset [6]. It can be seen that our results are noticeably better than those from other methods in terms of visual quality. It justified the value of the proposed method to image deblurring in practice.

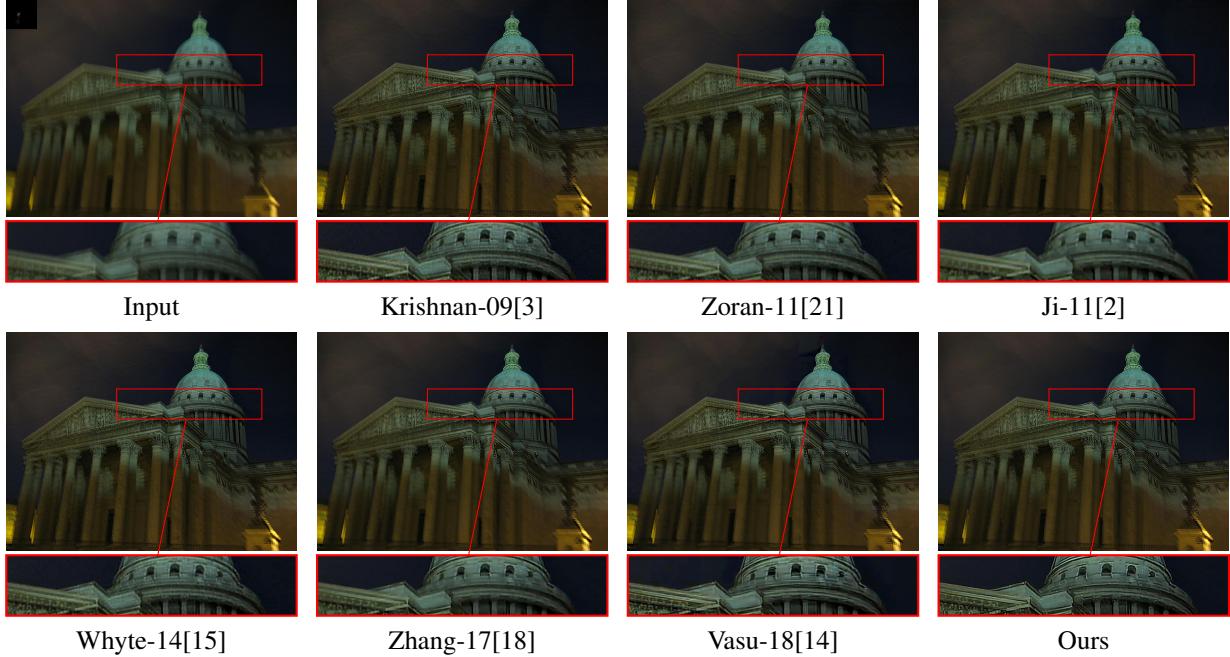


Figure 3: Deblurred results of image "Pantheon" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Zhong *et al.* [20]. Zoom in for better inspection.

5. Visual comparisons on synthetic images

See Fig 16 - 23 for visual comparison of the results of some synthetic images from Levin *et al.*'s dataset and Sun *et al.*'s dataset with noise level 1%. It can be seen that our recovery results are noticeably better than other compared ones in terms of visual quality: recovering more image details while introducing less artifacts. Such an advantage of the proposed method over others in terms of visual quality is consistent with that in terms of quantitative metric, shown in Table 2 in main paper.

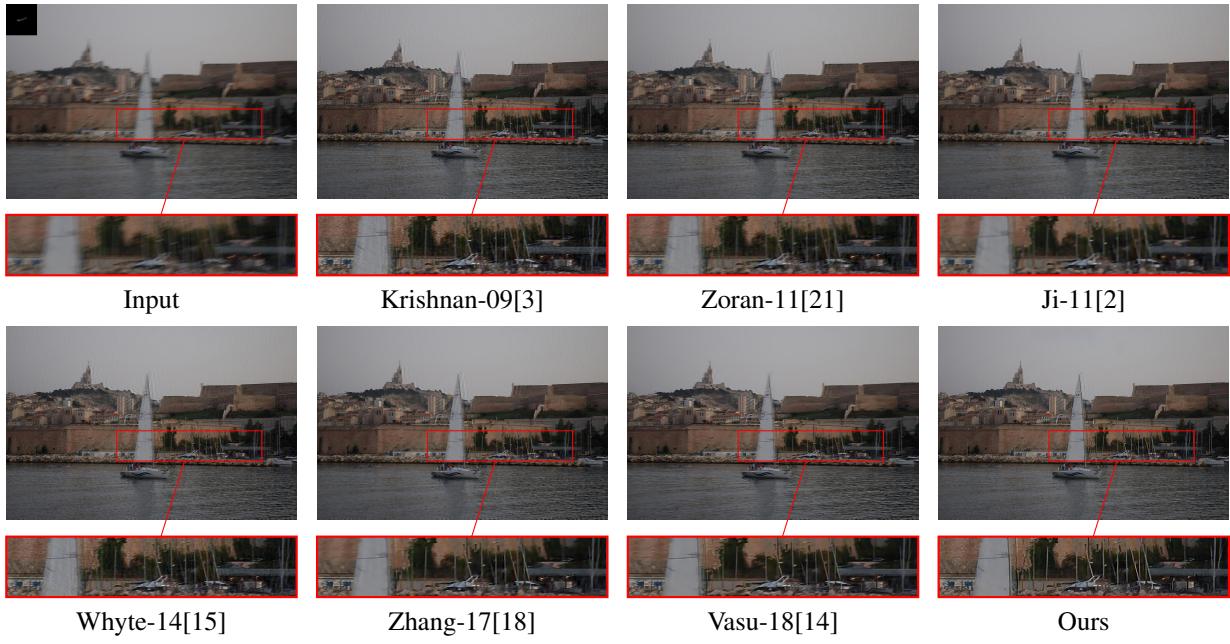


Figure 4: Deblurred results of image "boat1" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Cho and Lee [1]. Zoom in for better inspection.



Figure 5: Deblurred results of image "istanbul" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Pan *et al.* [10]. Zoom in for better inspection.

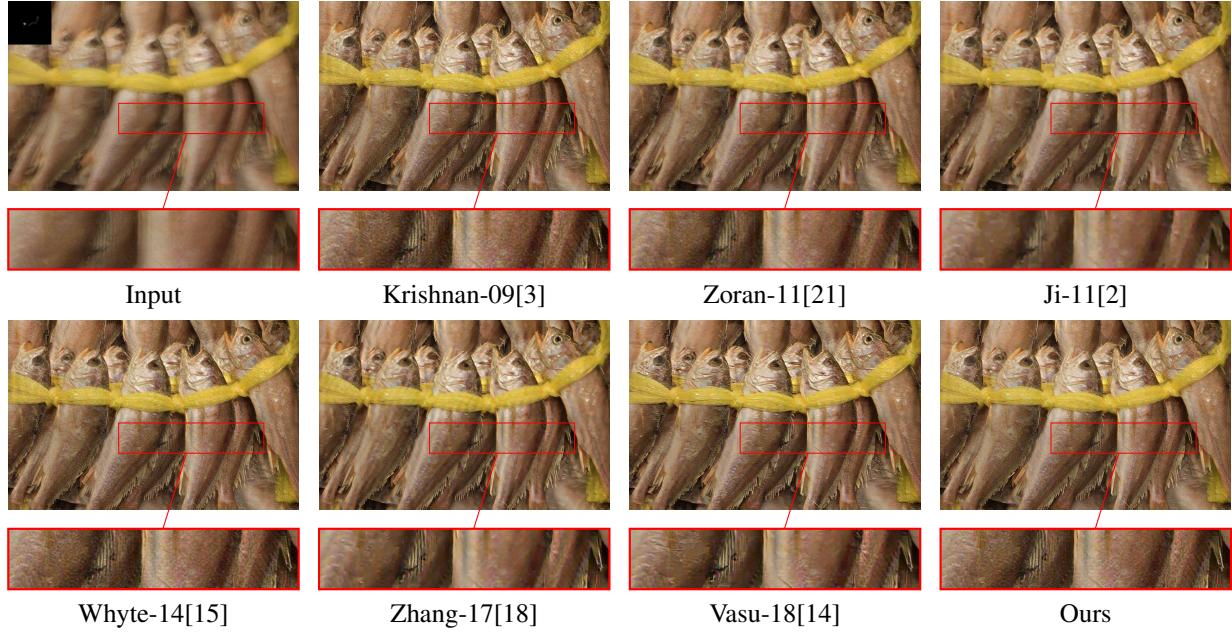


Figure 6: Deblurred results of image "fishes" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Cho and Lee [1]. Zoom in for better inspection.

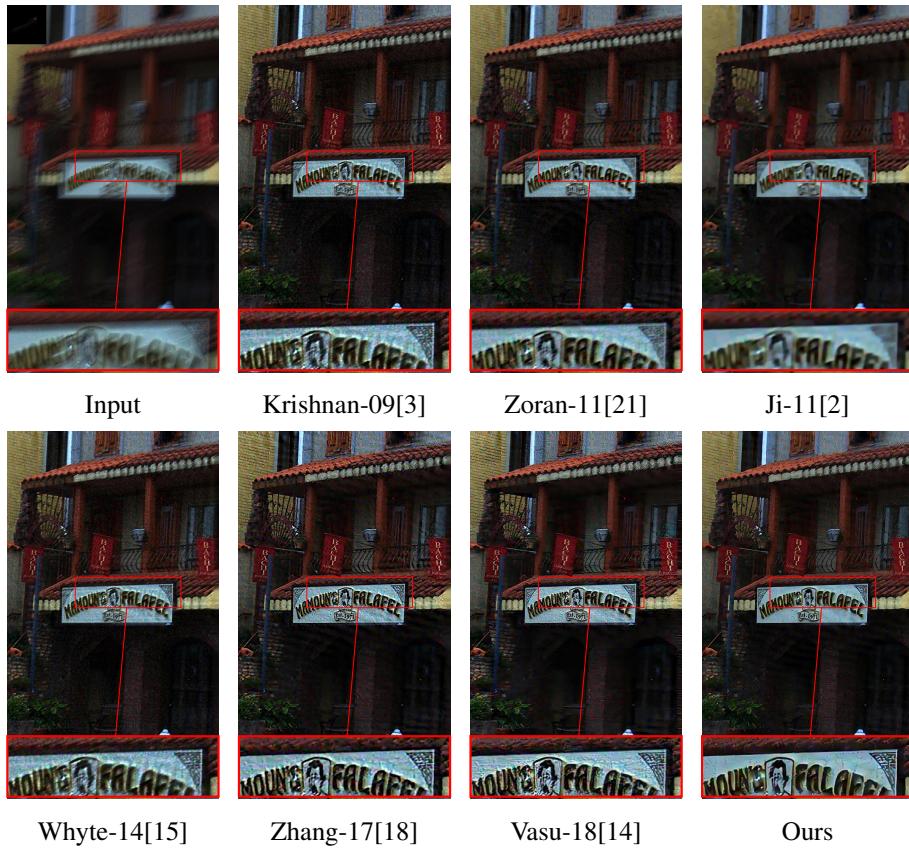


Figure 7: Deblurred results of image "house3" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Pan *et al.* [10]. Zoom in for better inspection.



Figure 8: Deblurred results of image "fountain1" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Sun *et al.* [13]. Zoom in for better inspection.



Figure 9: Deblurred results of image "lyndsey" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Xu *et al.* [17]. Zoom in for better inspection.

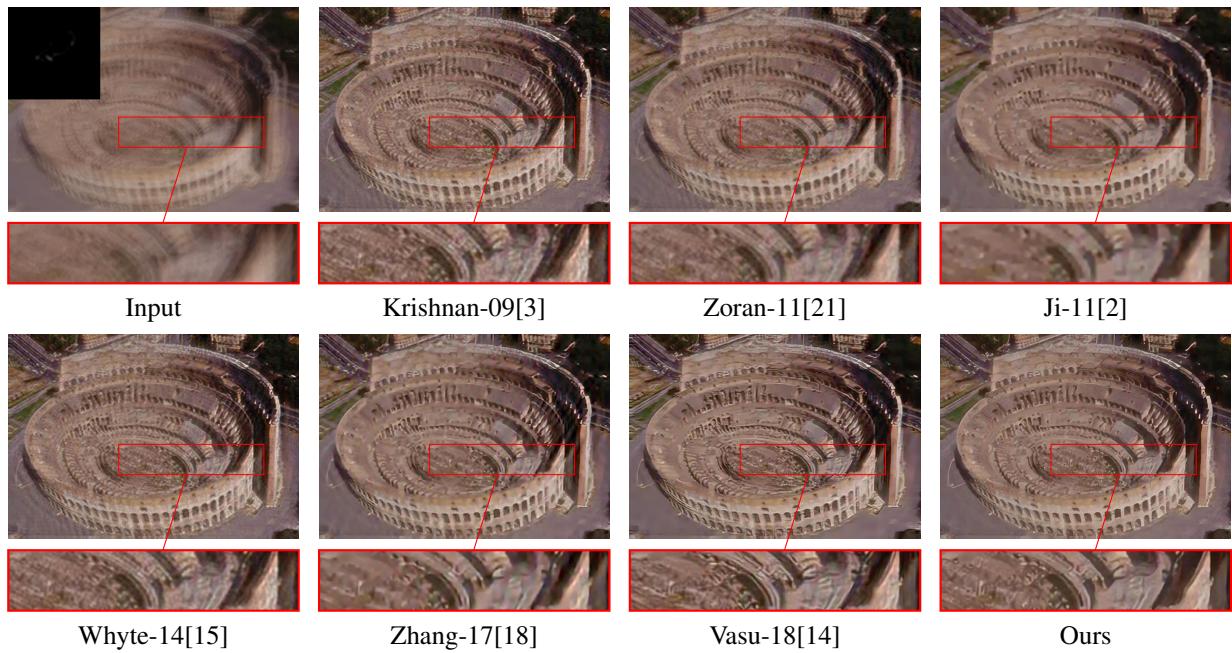


Figure 10: Deblurred results of image "roma" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Xu *et al.* [17]. Zoom in for better inspection.

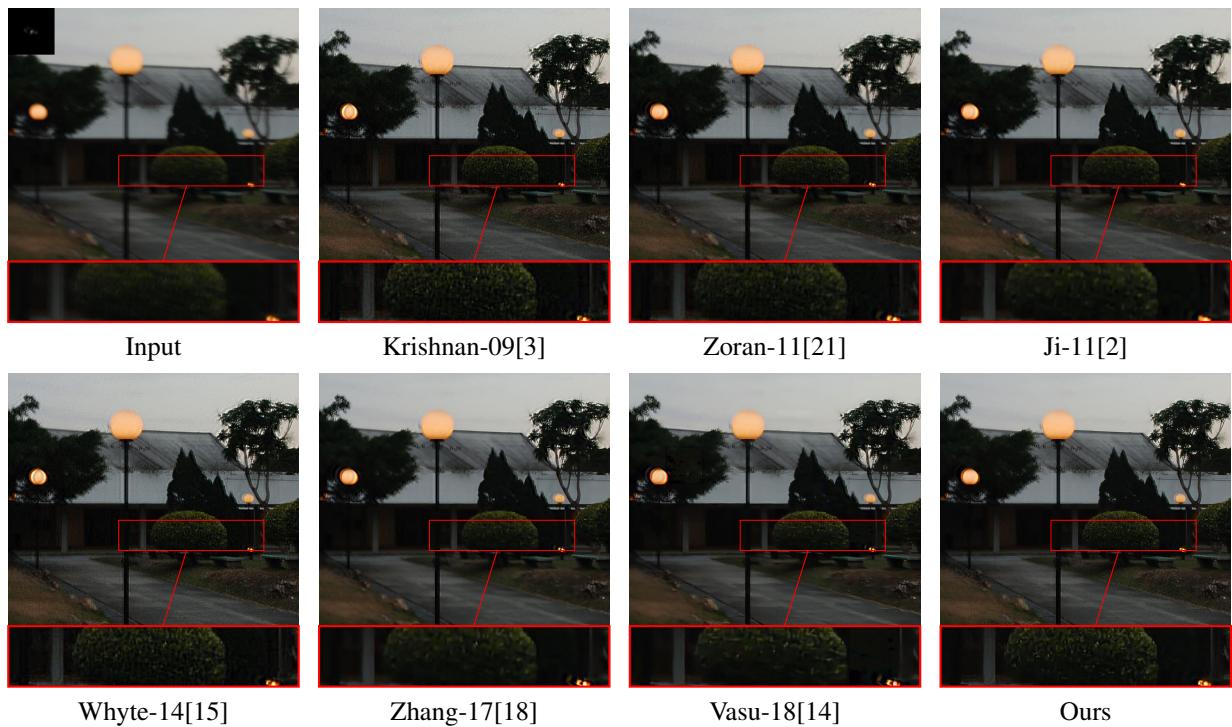


Figure 11: Deblurred results of image "nv" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Xu and Jia [16]. Zoom in for better inspection.

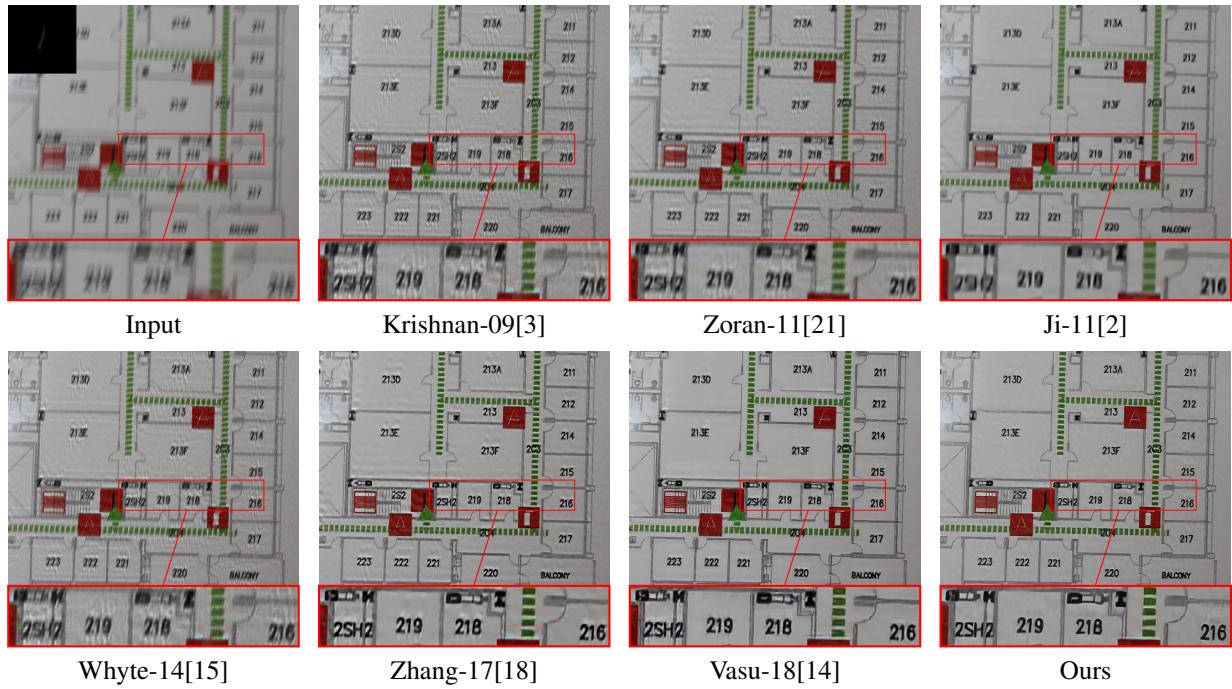


Figure 12: Deblurred results of image "text10" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Pan *et al.* [10]. Zoom in for better inspection.



Figure 13: Deblurred results of image "statue1" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Krishnan *et al.* [4]. Zoom in for better inspection.



Figure 14: Deblurred results of image "toy" from the real dataset in Lai *et al.* [6]. The kernel are estimated by Levin *et al.* [8]. Zoom in for better inspection.

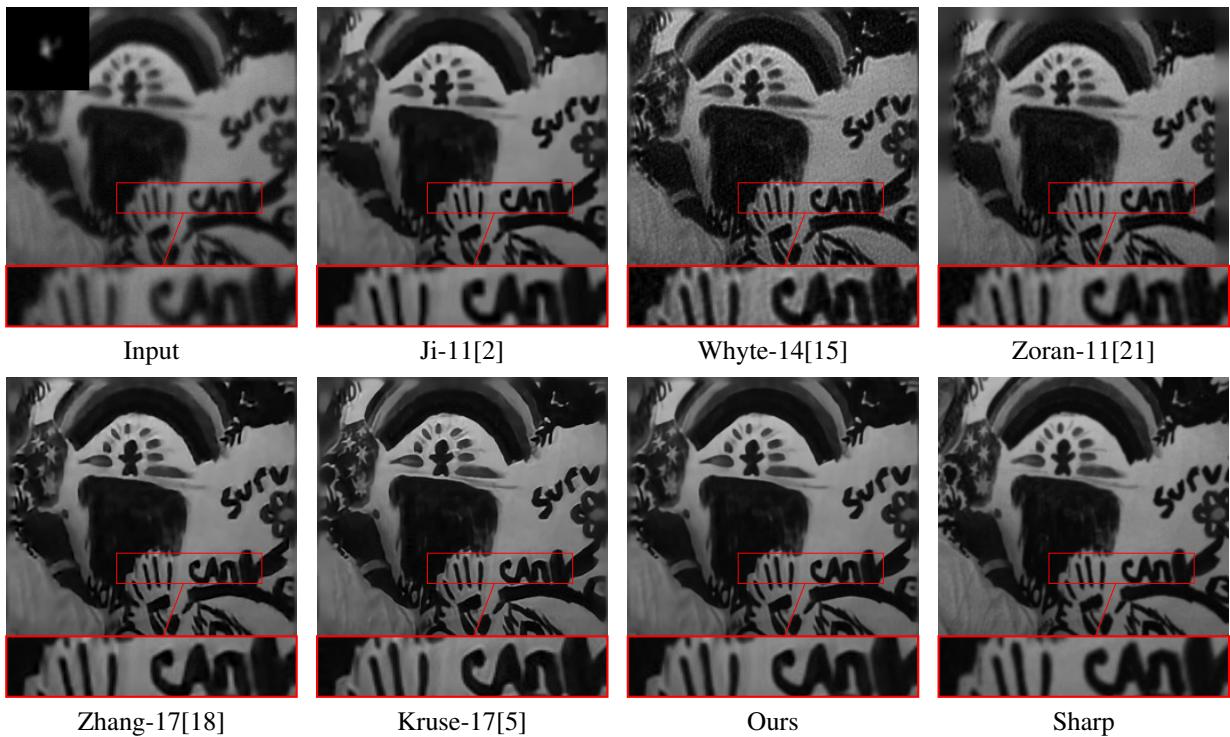


Figure 15: Deblurred results for one example image from Levin *et al.*'s dataset with noise level 1%.

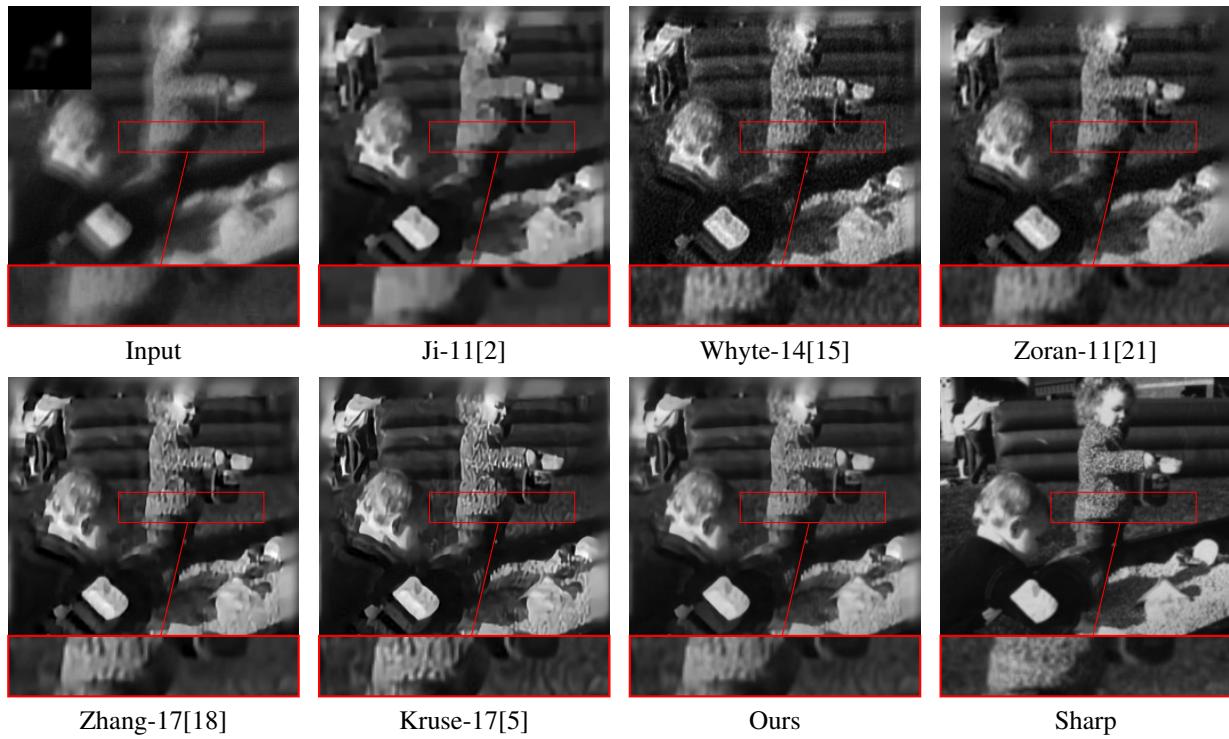


Figure 16: Deblurred results for one example image from Levin *et al.*'s dataset with noise level 1%.

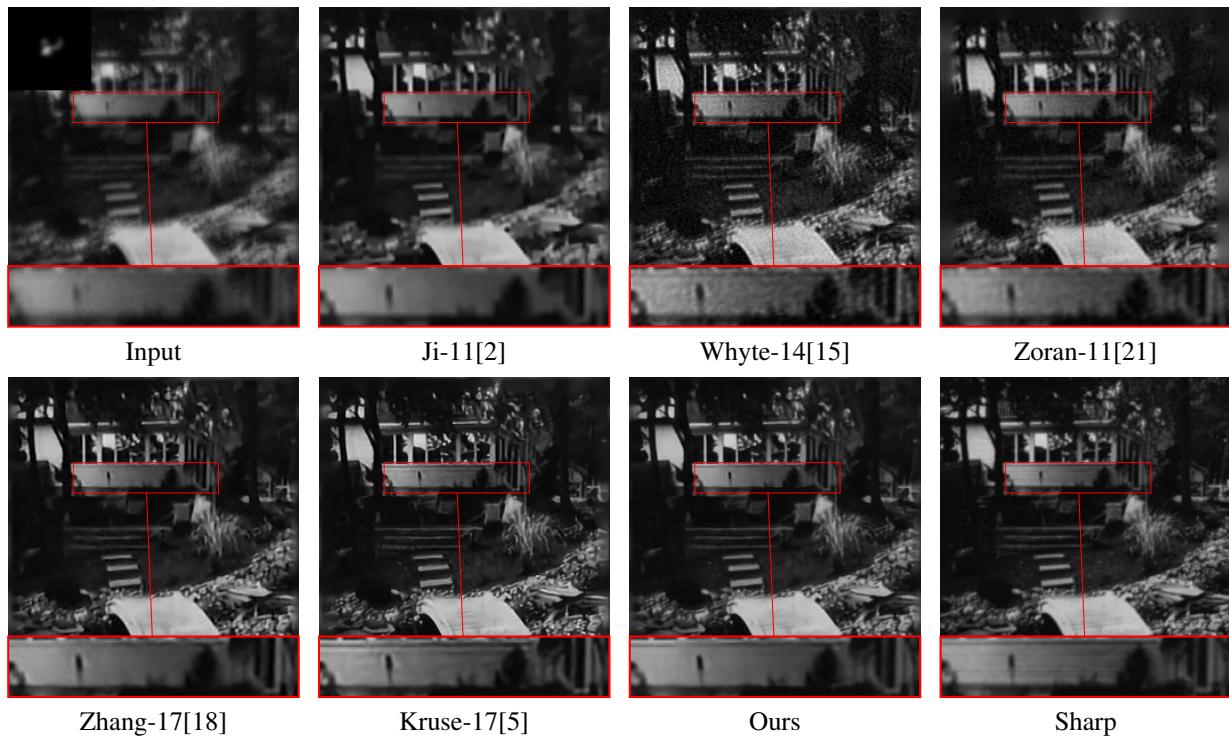


Figure 17: Deblurred results for one example image from Levin *et al.*'s dataset with noise level 1%.

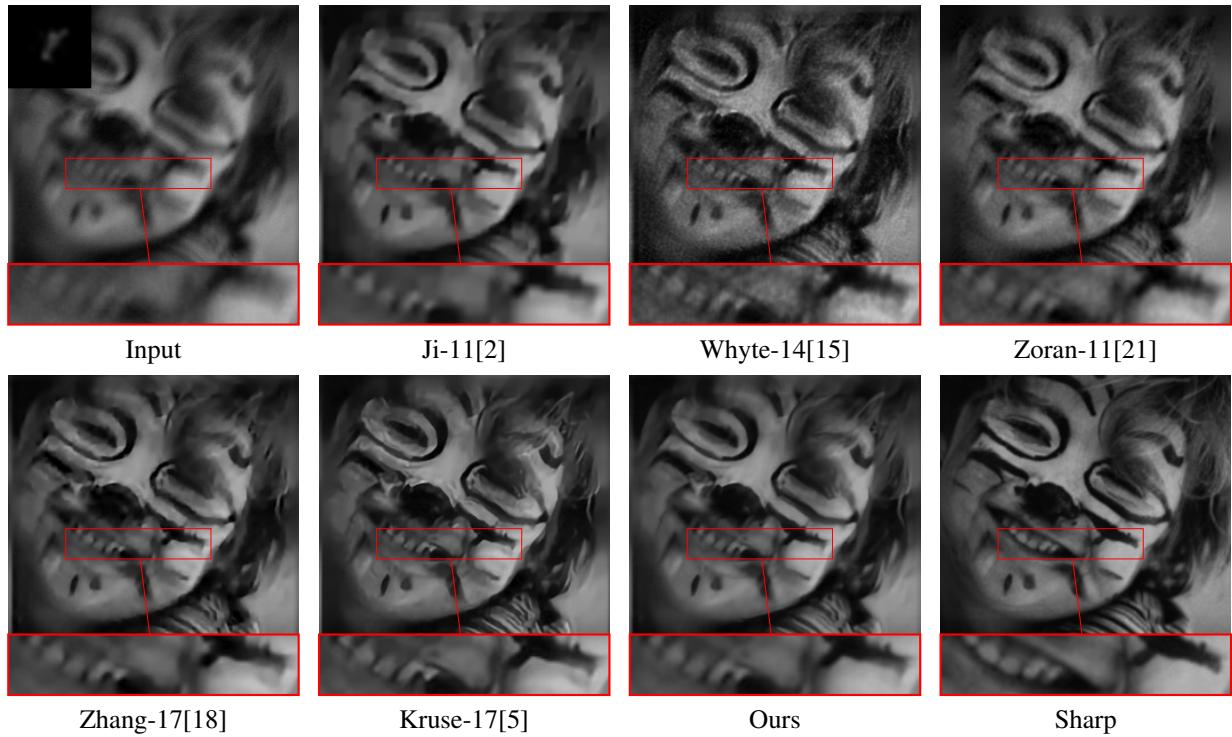


Figure 18: Deblurred results for one example image from Levin *et al.*'s dataset with noise level 1%.

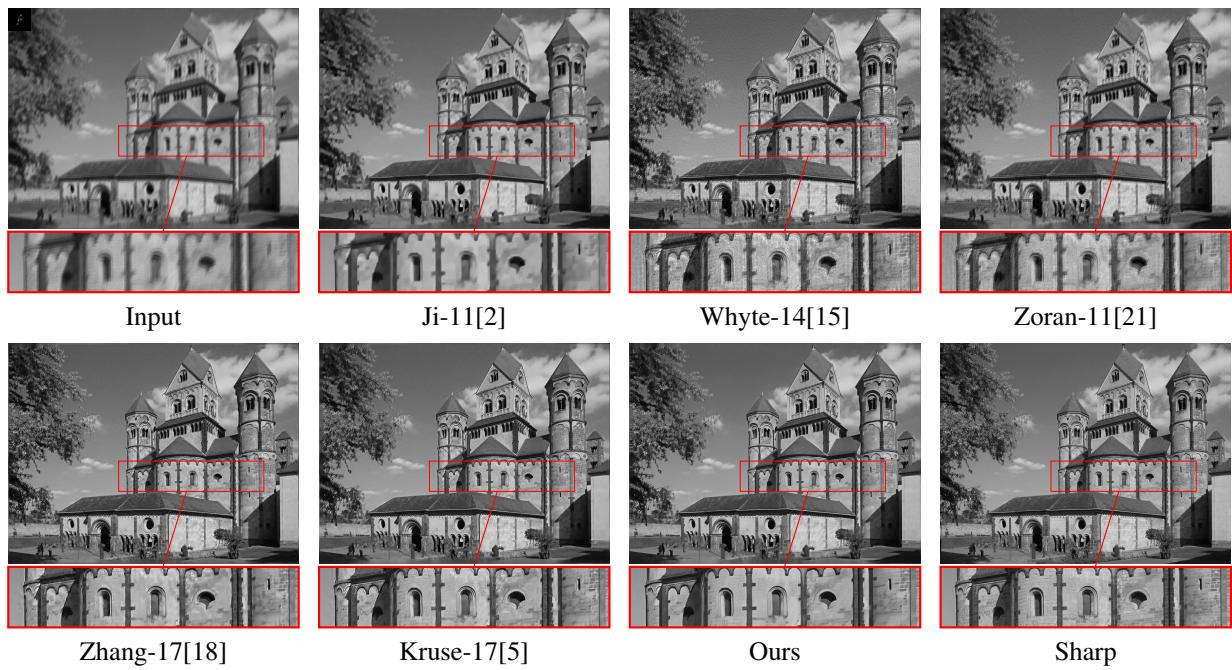


Figure 19: Deblurred results for one example image from Sun *et al.*'s dataset with noise level 1%.

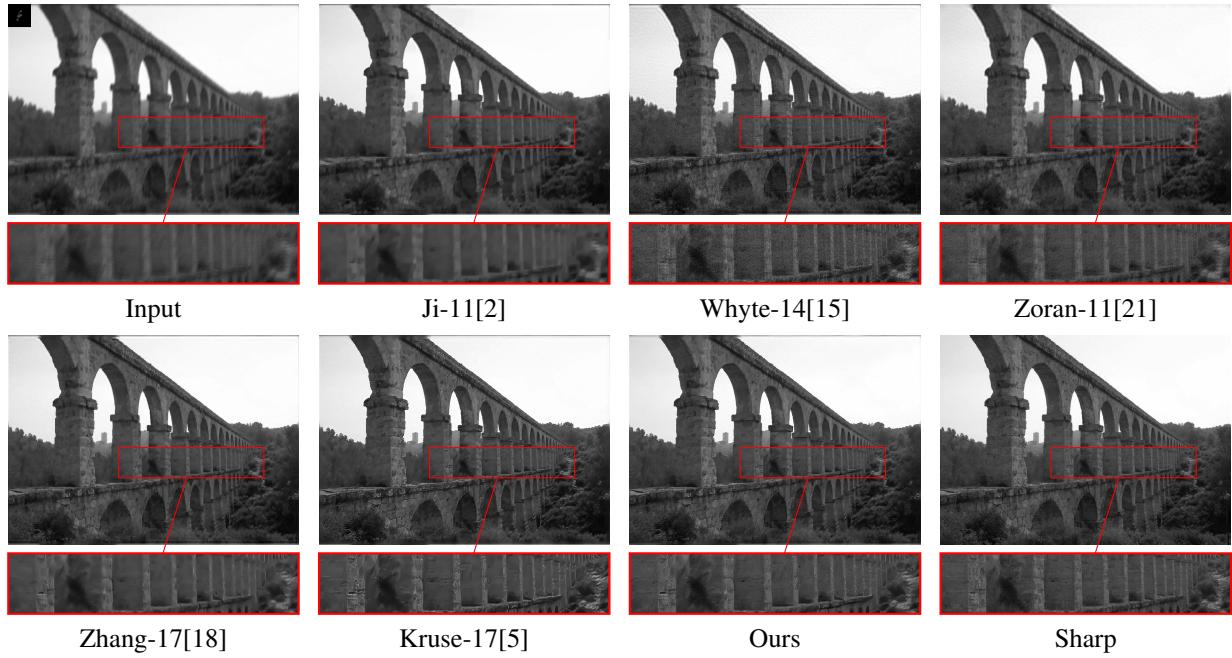


Figure 20: Deblurred results for one example image from Sun *et al.*'s dataset with noise level 1%.

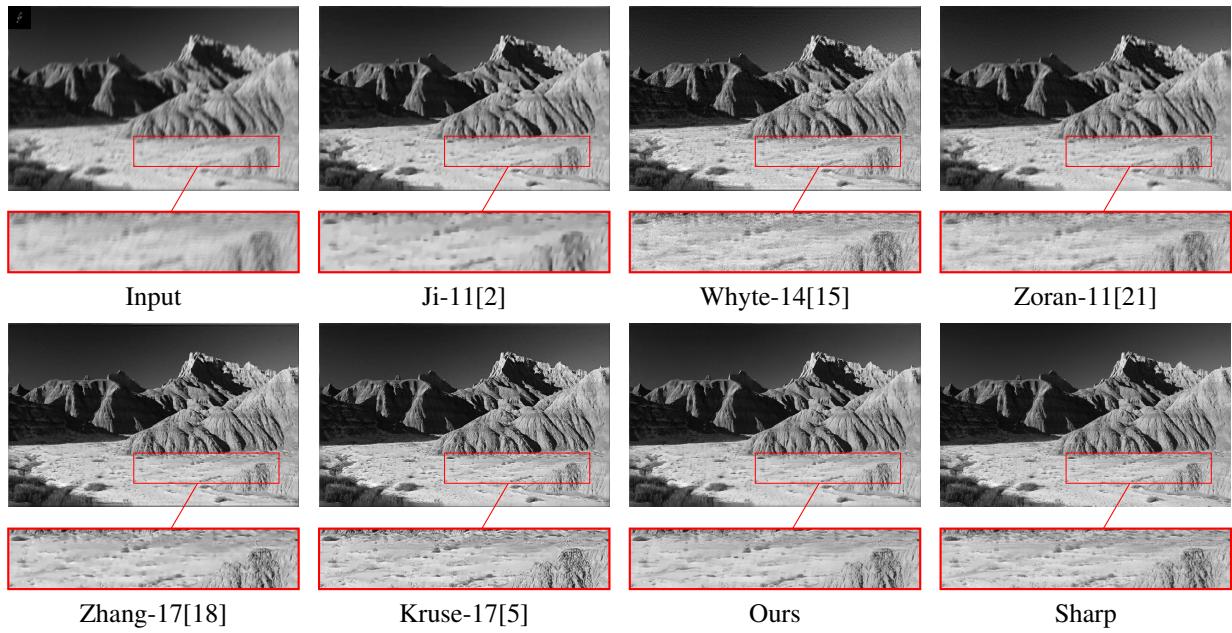


Figure 21: Deblurred results for one example image from Sun *et al.*'s dataset with noise level 1%.

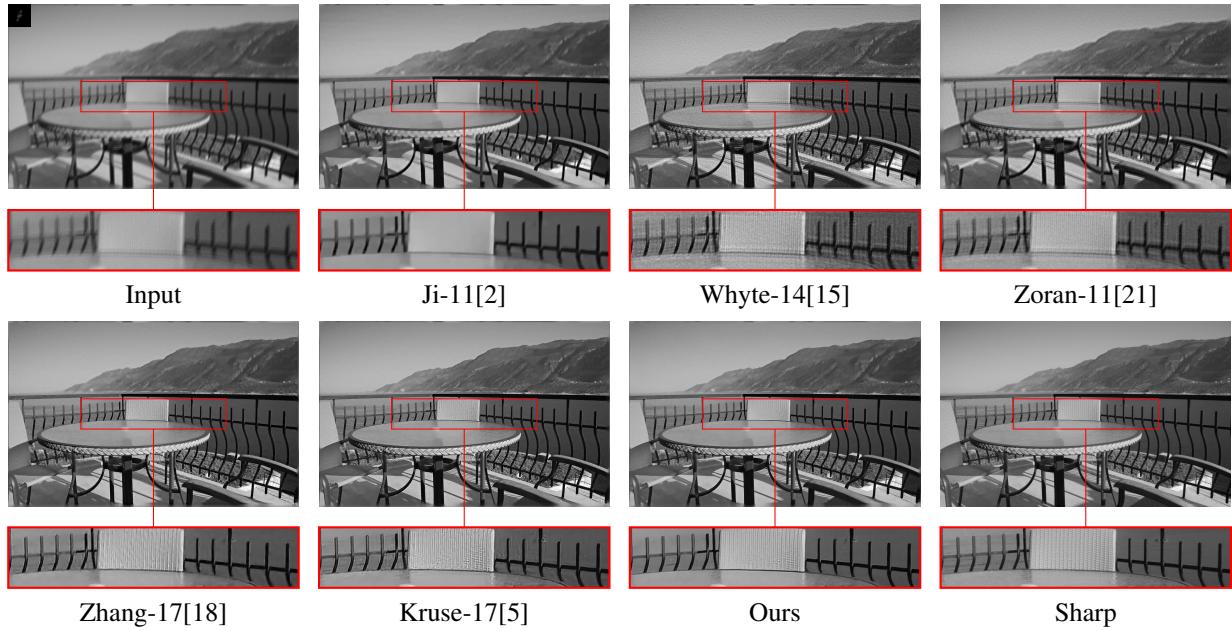


Figure 22: Deblurred results for one example image from Sun *et al.*'s dataset with noise level 1%.

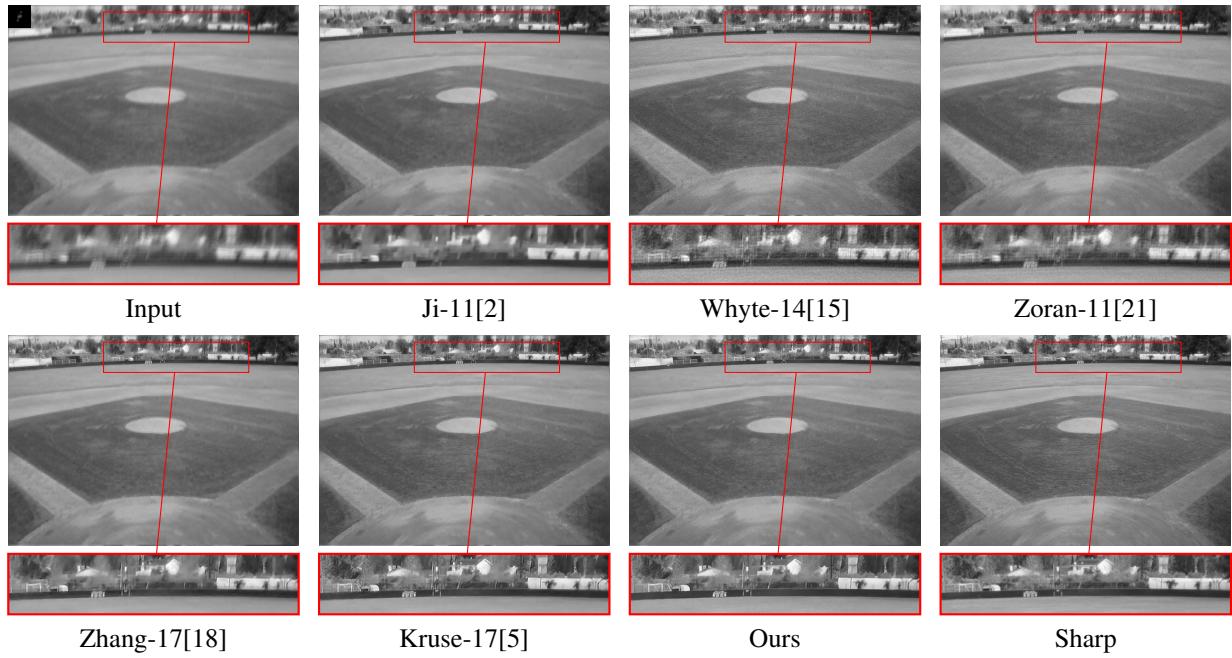


Figure 23: Deblurred results for one example image from Sun *et al.*'s dataset with noise level 1%.

References

- [1] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. *ACM Transactions on graphics (TOG)*, 28(5):145, 2009.
- [2] Hui Ji and Kang Wang. Robust image deblurring with an inaccurate blur kernel. *IEEE Transactions on Image processing*, 21(4):1624–1634, 2011.
- [3] Dilip Krishnan and Rob Fergus. Fast image deconvolution using hyper-laplacian priors. In *NIPS*, pages 1033–1041, 2009.
- [4] Dilip Krishnan, Terence Tay, and Rob Fergus. Blind deconvolution using a normalized sparsity measure. In *CVPR 2011*, pages 233–240. IEEE, 2011.
- [5] Jakob Kruse, Carsten Rother, and Uwe Schmidt. Learning to push the limits of efficient fft-based image deconvolution. In *CVPR*, pages 4586–4594, 2017.
- [6] Wei-Sheng Lai, Jia-Bin Huang, Zhe Hu, Narendra Ahuja, and Ming-Hsuan Yang. A comparative study for single image blind deblurring. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1701–1709, 2016.
- [7] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, pages 1964–1971. IEEE, 2009.
- [8] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Efficient marginal likelihood optimization in blind deconvolution. In *CVPR 2011*, pages 2657–2664. IEEE, 2011.
- [9] Tomer Michaeli and Michal Irani. Blind deblurring using internal patch recurrence. In *European Conference on Computer Vision*, pages 783–798. Springer, 2014.
- [10] Jinshan Pan, Zhe Hu, Zhixun Su, and Ming-Hsuan Yang. Deblurring text images via l0-regularized intensity and gradient prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2901–2908, 2014.
- [11] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring using dark channel prior. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 1628–1636, 2016.
- [12] Daniele Perrone and Paolo Favaro. Total variation blind deconvolution: The devil is in the details. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 2909–2916, 2014.
- [13] Libin Sun, Sunghyun Cho, Jue Wang, and James Hays. Edge-based blur kernel estimation using patch priors. In *IEEE International Conference on Computational Photography (ICCP)*, pages 1–8. IEEE, 2013.
- [14] Subeesh Vasu, Venkatesh Reddy Maligireddy, and AN Rajagopalan. Non-blind deblurring: Handling kernel uncertainty with cnns. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 3272–3281, 2018.
- [15] Oliver Whyte, Josef Sivic, and Andrew Zisserman. Deblurring shaken and partially saturated images. *International journal of computer vision*, 110(2):185–201, 2014.
- [16] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *European conference on computer vision*, pages 157–170. Springer, 2010.
- [17] Li Xu, Shicheng Zheng, and Jiaya Jia. Unnatural l0 sparse representation for natural image deblurring. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1107–1114, 2013.
- [18] Jiawei Zhang, Jinshan Pan, Wei-Sheng Lai, Rynson WH Lau, and Ming-Hsuan Yang. Learning fully convolutional networks for iterative non-blind deconvolution. In *CVPR*, pages 3817–3825, 2017.
- [19] Kai Zhang, Wangmeng Zuo, Shuhang Gu, and Lei Zhang. Learning deep cnn denoiser prior for image restoration. In *CVPR*, volume 2, 2017.
- [20] Lin Zhong, Sunghyun Cho, Dimitris Metaxas, Sylvain Paris, and Jue Wang. Handling noise in single image deblurring using directional filters. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 612–619, 2013.
- [21] Daniel Zoran and Yair Weiss. From learning models of natural image patches to whole image restoration. In *ICCV*, pages 479–486. IEEE, 2011.