Supplementary materials for "Uncertainty-Aware Unsupervised Image Deblurring with Deep Residual Prior"

Xiaole Tang[1]   Xile Zhao[1*]   Jun Liu[2*]   Jianli Wang[1]   Yuchun Miao[1]   Tieyong Zeng[3]

[1]University of Electronic Science and Technology of China, Chengdu, China
[2]Northeast Normal University, Changchun, China
[3]The Chinese University of Hong Kong, Shatin, NT, Hong Kong

{sherlock315,xlzhao122003,junliucd,wangjianli123,szmyc1}@163.com,zeng@math.cuhk.edu.hk
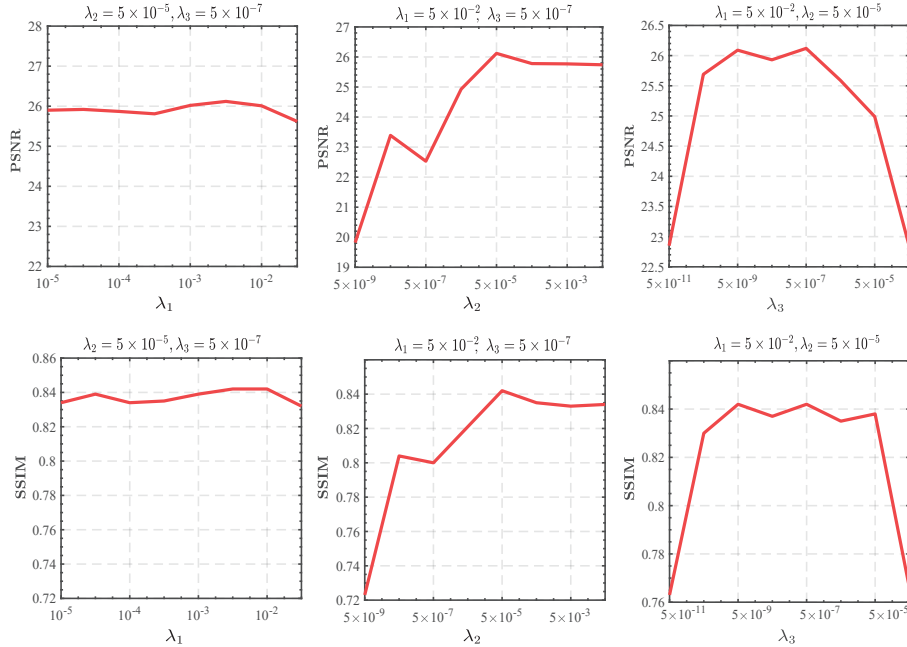
A. Additional implementation details

Figure 1. The PSNR/SSIM values with respect to the values of different hyperparameters on simulated data *Cameraman* blurred by Motion (20, 10) and the inaccurate kernel Motion (20, 20) as the input.

- **Network structures.** In Section 4.2, we compared the customized U-Net with FCN and the standard U-Net in estimating the residual induced by the kernel error. U-Net and FCN are widely used in deep learning methods [2, 13].The customized U-Net replaces the sigmoid activation with a soft-shrinkage activation. The sigmoid which easily causes gradient vanishing problem and is more computationally extensive, which is verified in WGAN [1]. The FCN has a hidden layer of 200 nodes and an input/output layer of nodes with the same size of the input image.

- **Model hyperparameters setting.** The model hyperparameters of the proposed method are $\{\lambda_i\}_{i=1}^3$ and L. In all experiments, we empirically fix $L = 3$. Then $\{\lambda_i\}_{i=1}^3$ are tuned to obtain the best PSNR value. To comprehensively investigate the influences of different model hyperparameters, we conduct experiments (see Figure 1) on the simulated data *Cameraman* blurred by Motion (20, 10) and the inaccurate kernel Motion (20, 20) as the input. Based on the

best PSNR/SSIM values shown in Figure 1, λ_1, λ_2, λ_3 are finally determined to be $5 \times 10^{-2}, 5 \times 10^{-5}, 5 \times 10^{-7}$, respectively.

To circumvent the adversity of convolution in the spatial domain and save computational cost, we transform the kernel $\boldsymbol{k}$ and image $\boldsymbol{x}$ into frequency with fast Fourier transform (FFT) in practical implementation. In this case, the convolution process can be reformulated as

$$\boldsymbol{k} \otimes \boldsymbol{x} = \mathcal{F}^{-1}\left(\mathcal{F}\left(\boldsymbol{k}\right) \odot \mathcal{F}\left(\boldsymbol{x}\right)\right), \tag{1}$$

in which $\mathcal{F}$ and $\odot$ respectively represent operators of FFT and Hadamard product.

The true residual $\boldsymbol{r}_{gt} = \Delta \boldsymbol{k} \otimes \boldsymbol{x}$ is obtained by the following formulation:

$$\boldsymbol{r}_{gt} = \boldsymbol{y} - \widehat{\boldsymbol{k}} \otimes \boldsymbol{x} \tag{2}$$

where $\boldsymbol{y}$ is the blurry image, $\widehat{\boldsymbol{k}}$ is the inaccurate kernel and $\boldsymbol{x}$ is the clean image.

B. Test images as wild dataset

In the simulated experiments exploring the influence of different combination of priors in our model and the robustness to kernel error, we use 8 sharp images (see Figure 2) and 3 typical blur kernels (see Figure 3) to generate blurry images as toy dataset. By varying the parameter of the kernel, we can obtain a series of inaccurate kernels. For instance, let the Motion (20,10) (a motion blur with 20 pixels of length and $10°$ orientations) be the true kernel, then a series of inaccurate kernels can be set as Motion(5,10), Motion(10,10), Motion(15,10), Motion(25,10), Motion(30,10) (i.e. bias of length varies from -15 to 15 with a stride of 5.)

| *Cameraman* | *Goldhill* | *Lena* | *Peppers* |
| *Cateyes* | *Barbara* | *Baboon* | *Airplane* |

Figure 2. Sharp images used for specialized blur dataset.

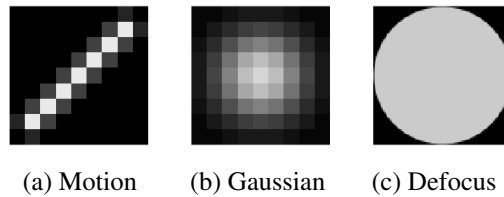(a) Motion (b) Gaussian (c) Defocus

Figure 3. Examples of the 3 typical kernels.

C. Failure case

Due to its better robustness to kernel error, the proposed method generally delivers better results as compared with Fang *et al.* [5]. More specifically, for Levin *et al.*'s dataset the proposed method achieves a performance gain of $+0.74$ dB , with the highest gain of $+2.65$ dB and the lowest gain of $+0.16$ dB. For Lai *et al.*'s dataset, there is an average gain of $+0.53$ dB, with the highest gain of $+1.28$ dB and the lowest gain of -0.22 dB. This indicates there are still some failure cases, e.g., text image in Figure 4. The underlying reason may be that the deep prior and sparse prior are not suitable for the text image.

Blurry	Fang *et al.* [5]	Ours	Ground truth
PSNR 14.15	PSNR **23.01**	PSNR 22.79	PSNR Inf

Figure 4. Visual comparison for a text image from Lai *et al.* [8].

D. Additional visualization on different network architectures for DRP

To further validate the effectiveness of the customized U-Net for DRP as compared to U-Net, we present more results over different blurs and images by using these two architectures.

Blurry

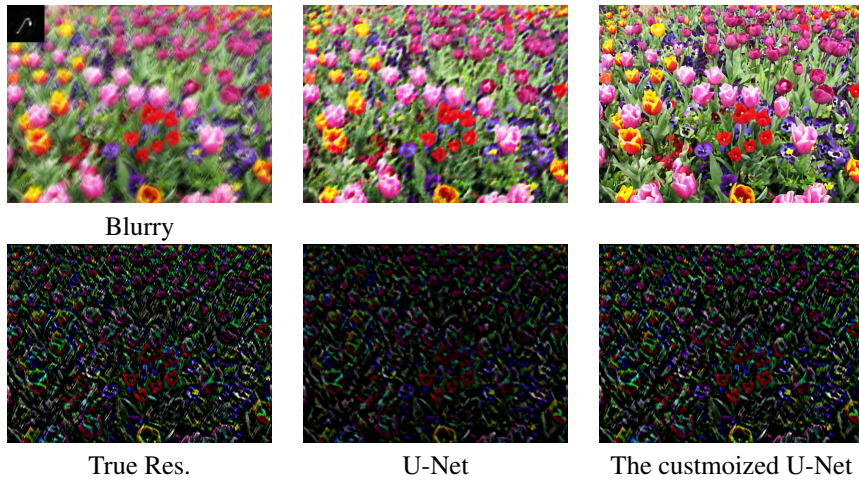True Res.   U-Net   The custmoized U-Net

Figure 5. The comparison of U-Net and the customized U-Net for DRP. Row 1: the deblurred results of image 'natural05' from the dataset of Lai *et al.* [9]. Row 2: the estimated residuals. The input kernel for deblurring is estimated by the method in [4].
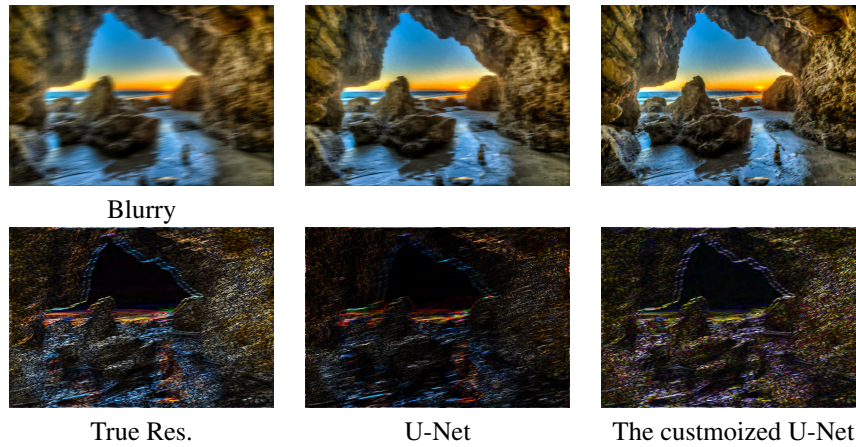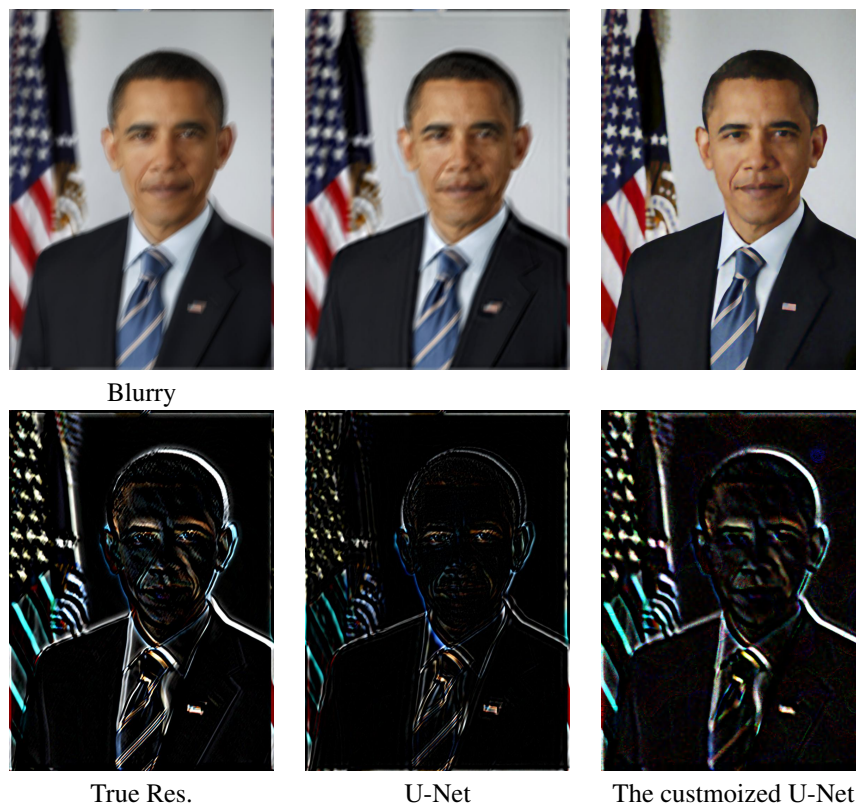
Blurry

True Res.   U-Net   The custmoized U-Net

Figure 6. The comparison of U-Net and the customized U-Net for DRP. Row 1: the deblurred results of image 'natural01' from the dataset of Lai *et al*. [9]. Row 2: the estimated residuals. The input kernel for deblurring is estimated by the method in [4].

Blurry

True Res.   U-Net   The custmoized U-Net

Figure 7. The comparison of U-Net and the customized U-Net for DRP. Row 1: the deblurred results of image 'people03' from the dataset of Lai *et al*. [8]. Row 2: the estimated residuals. The input kernel for deblurring is estimated by the method in [4].
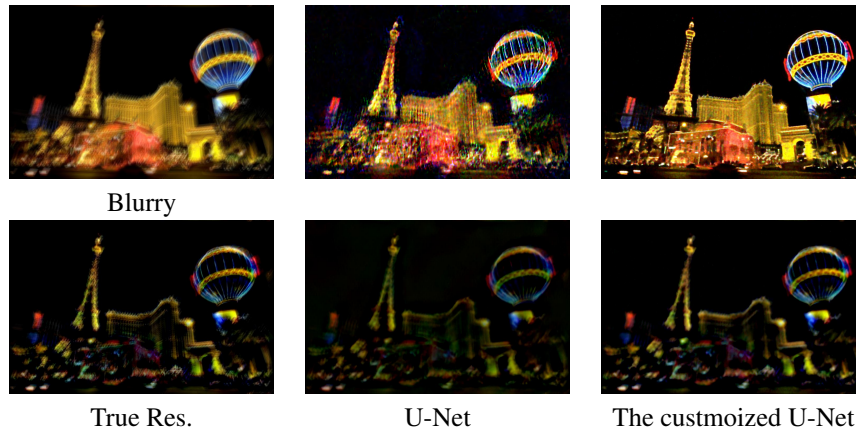
Blurry   **True Res.**   **U-Net**   **The custmoized U-Net**

Figure 8. The comparison of U-Net and the customized U-Net for DRP. Row 1: the deblurred results of image 'saturated01' from the dataset of Lai *et al*. [8]. Row 2: the estimated residuals. The input kernel for deblurring is estimated by the method in [4].
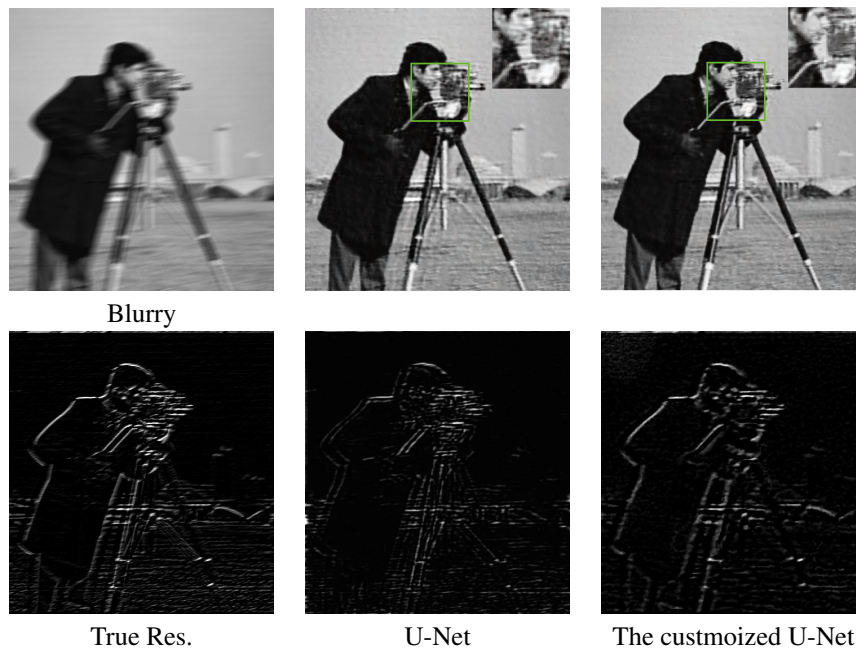

Blurry   **True Res.**   **U-Net**   **The custmoized U-Net**

Figure 9. The comparison of U-Net and the customized U-Net for DRP. Row 1: the deblurred results of image '*cameraman*'. Row 2: the estimated residuals. The input inaccurate kernel is Motion (20,20), and the true kernel is Motion (20,10).
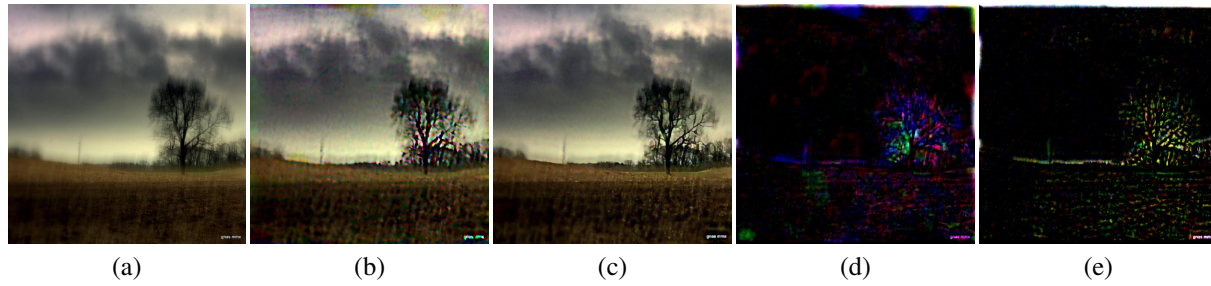
Figure 10. The comparison of U-Net and the customized U-Net for DRP. the deblurred results of real image 'outdoor' from the dataset of Lai *et al*. [8]. (a) Blurry image; (b) deblurred reuslt by using U-Net for DRP; (c) deblurred reuslt by using the customized U-Net for DRP; (d) estimated residual by U-Net; (e) estimated residual by the customized U-Net.

Figure 11. The comparison of U-Net and the customized U-Net for DRP. the deblurred results of real image 'harubang' from the dataset of Lai *et al*. [8]. (a) Blurry image; (b) deblurred reuslt by using U-Net for DRP; (c) deblurred reuslt by using the customized U-Net for DRP; (d) estimated residual by U-Net; (e) estimated residual by the customized U-Net.

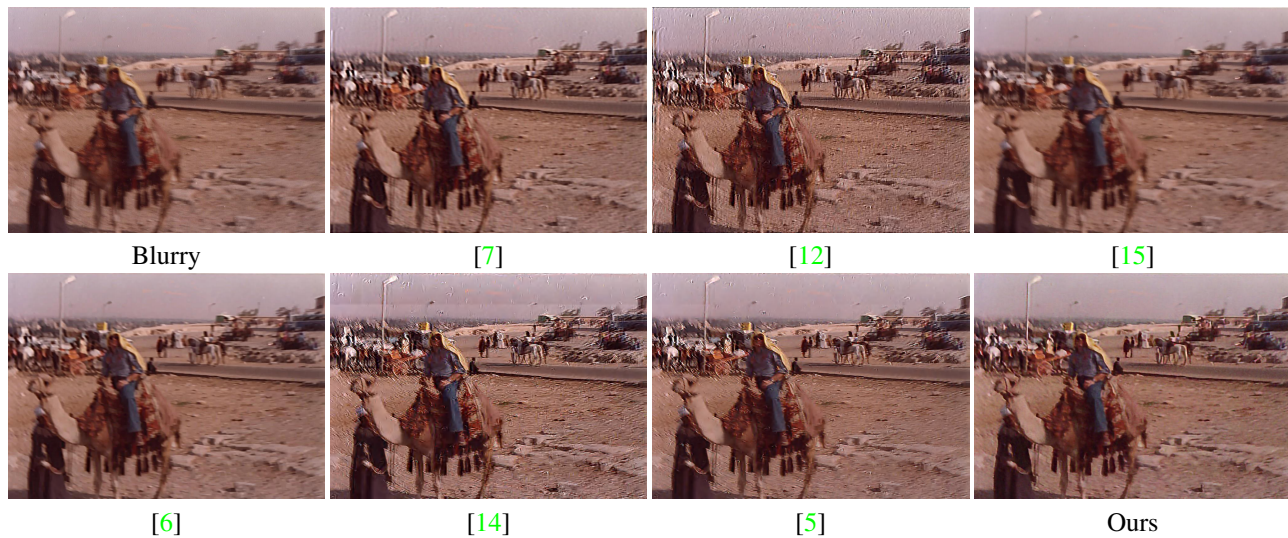# E. Additional visual comparison of deblurred results over real images

Figure 12. The deblurred results by different methods over the real image 'cairo1979' from the dataset of Lai *et al*. [8]. The kernel is estimated by the method in [4].

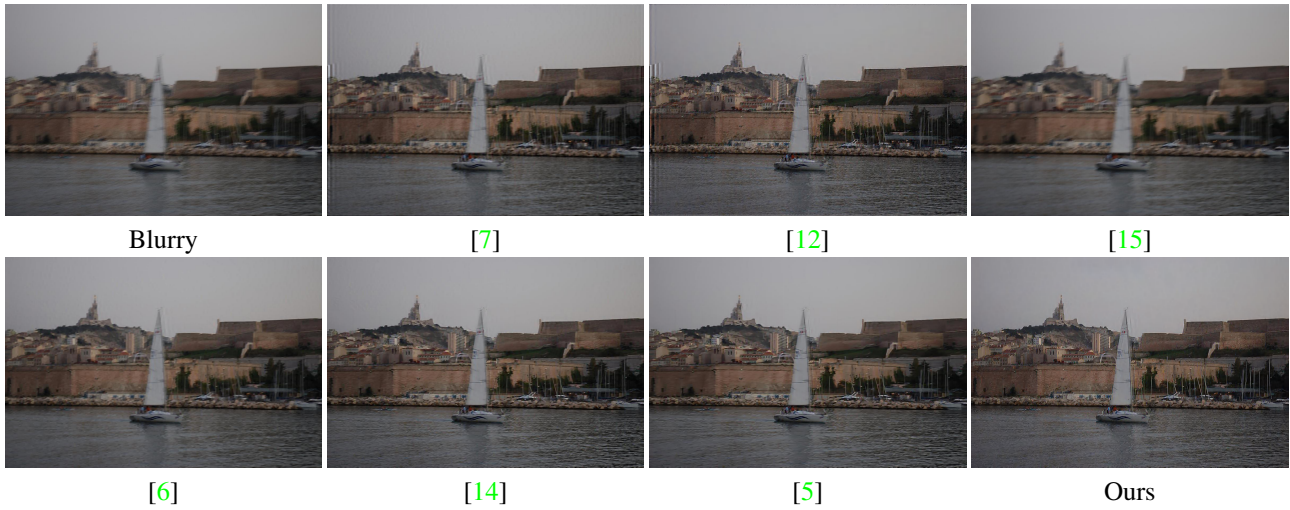|			
Blurry	[7]	[12]	[15]
[6]	[14]	[5]	Ours

Figure 13. The deblurred results by different methods over the real image 'boat1' from the dataset of Lai *et al.* [8]. The kernel is estimated by the method in [3].
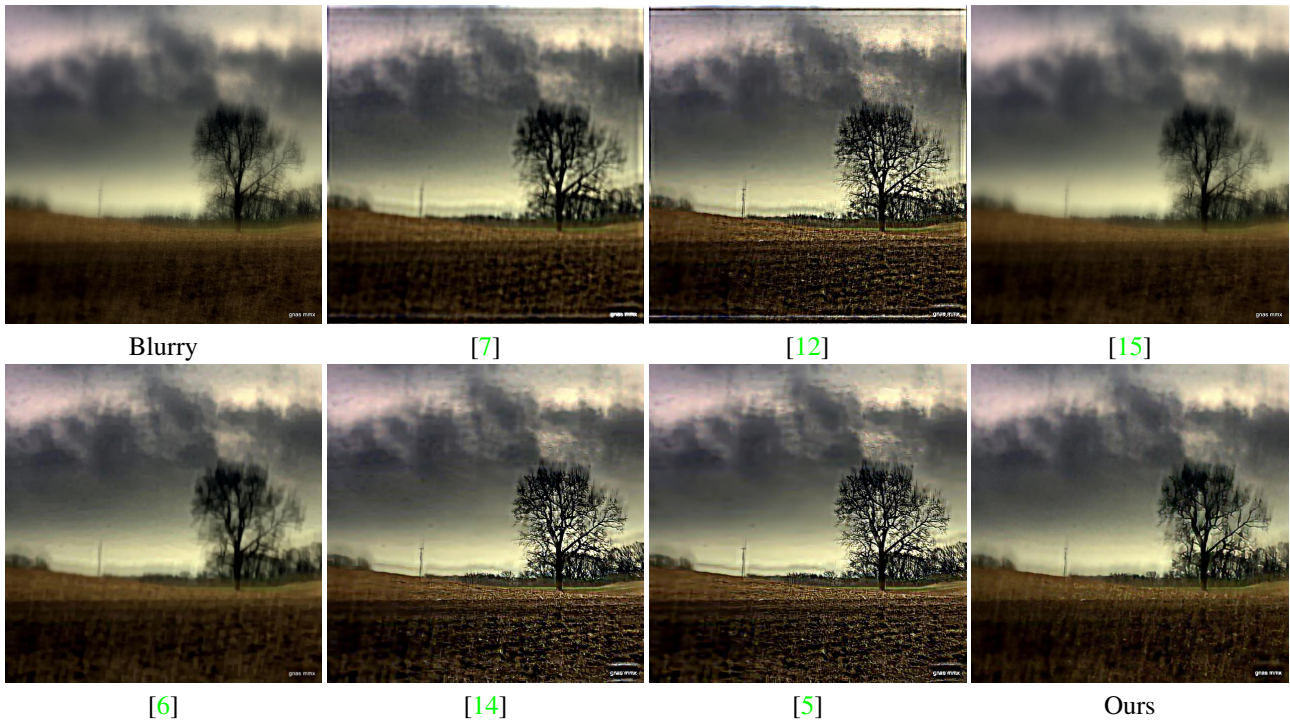

Blurry	[7]	[12]	[15]
[6]	[14]	[5]	Ours

Figure 14. The deblurred results by different methods over the real image 'outdoor' from the dataset of Lai *et al.* [8]. The kernel is estimated by the method in [11].
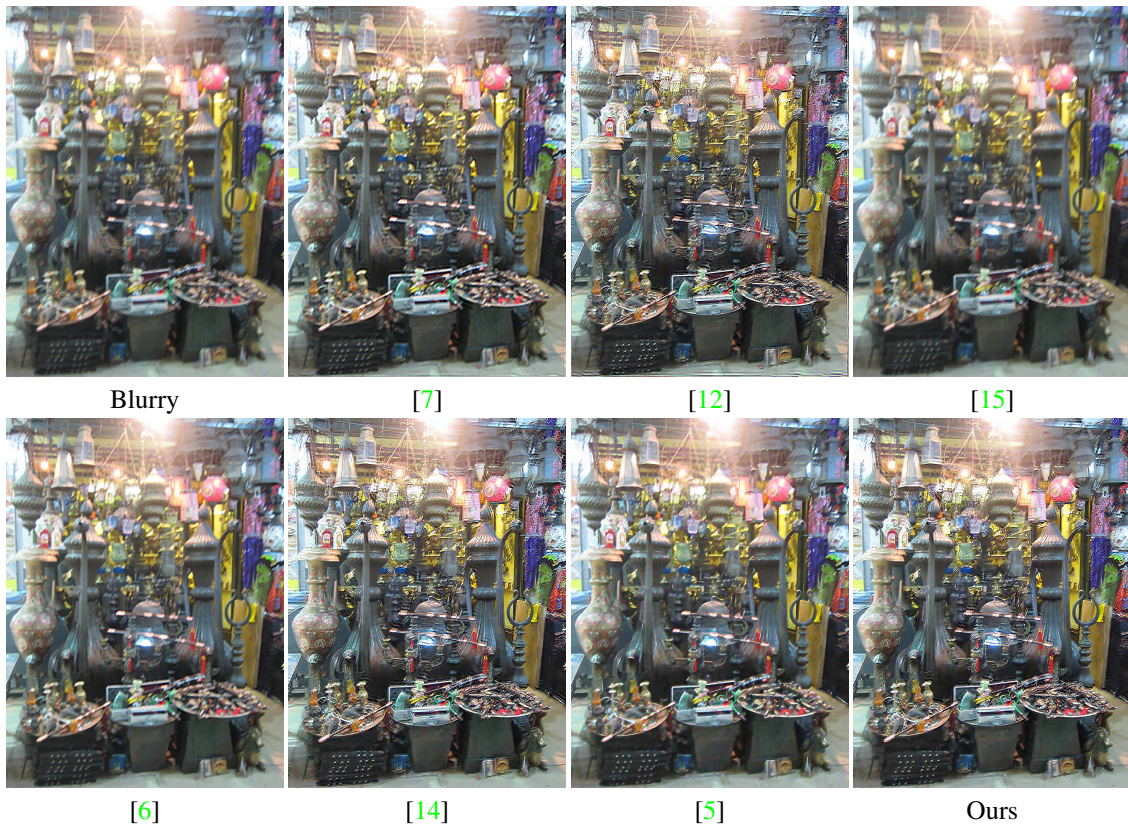
Figure 15. The deblurred results by different methods over the real image 'istanbul' from the dataset of Lai *et al*. [8]. The kernel is estimated by the method in [11].
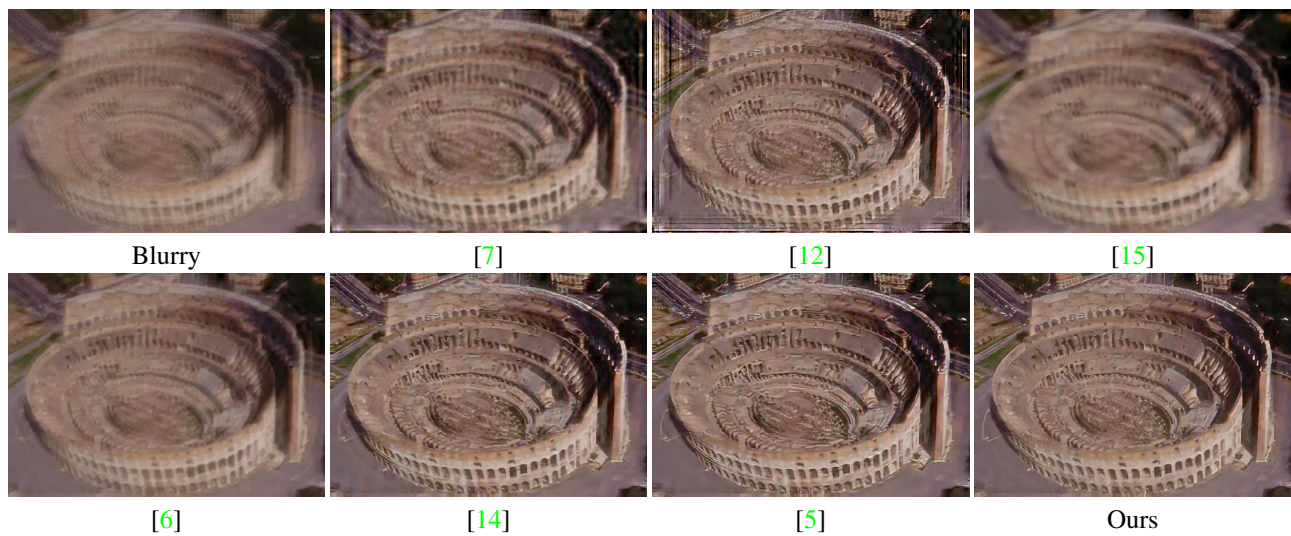

Figure 16. The deblurred results by different methods over the real image 'roma' from the dataset of Lai *et al*. [8]. The kernel is estimated by the method in [10].

References

[1] Martin Arjovsky, Soumith Chintala, and Léon Bottou. Wasserstein generative adversarial networks. In *ICML*, pages 214–223, 2017. 1

[2] Muhammad Asim, Fahad Shamshad, and Ali Ahmed. Blind image deconvolution using deep generative priors. *IEEE TCI*, 6:1493–1506, 2020. 1

[3] Sunghyun Cho and Seungyong Lee. Fast motion deblurring. In *ACM SIGGRAPH*, pages 1–8. 2009. 7

[4] Perrone Daniele and Favaro Paolo. Total variation blind deconvolution: The devil is in the details. In *CVPR*, pages 2909–2916, 2014. 3, 4, 5, 6

[5] Yingying Fang, Hao Zhang, Hok Shing Wong, and Tieyong Zeng. A robust non-blind deblurring method using deep denoiser prior. In *CVPRW*, pages 735–744, June 2022. 3, 6, 7, 8

[6] Hui Ji and Kang Wang. Robust image deblurring with an inaccurate blur kernel. *IEEE TIP*, 21(4):1624–1634, 2011. 6, 7, 8

[7] Dilip Krishnan and Rob Fergus. Fast image deconvolution using hyper-laplacian priors. *NIPS*, 22, 2009. 6, 7, 8

[8] Wei-Sheng Lai, Jia-Bin Huang, Zhe Hu, Narendra Ahuja, and Ming-Hsuan Yang. A comparative study for single image blind deblurring. In *CVPR*, pages 1701–1709, 2016. 3, 4, 5, 6, 7, 8

[9] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Understanding and evaluating blind deconvolution algorithms. In *CVPR*, pages 1964–1971, 2009. 3, 4

[10] Xu Li, Zheng Shicheng, and Jia Jiaya. Unnatural l_0 sparse representation for natural image deblurring. In *CVPR*, pages 1107–1114, 2013. 8

[11] Jinshan Pan, Deqing Sun, Hanspeter Pfister, and Ming-Hsuan Yang. Blind image deblurring using dark channel prior. In *CVPR*, pages 1628–1636, 2016. 7, 8

[12] Dongwei Ren, Hongzhi Zhang, David Zhang, and Wangmeng Zuo. Fast total-variation based image restoration based on derivative alternated direction optimization methods. *Neurocomputing*, 170:201–212, 2015. 6, 7, 8

[13] Dongwei Ren, Kai Zhang, Qilong Wang, Qinghua Hu, and Wangmeng Zuo. Neural blind deconvolution using deep priors. In *CVPR*, pages 3341–3350, 2020. 1

[14] Subeesh Vasu, Venkatesh Reddy Maligireddy, and AN Rajagopalan. Non-blind deblurring: Handling kernel uncertainty with cnns. In *CVPR*, pages 3272–3281, 2018. 6, 7, 8

[15] Xile Zhao, Wei Wang, Tieyong Zeng, Tingzhu Huang, and Michael K. Ng. Total variation structured total least squares method for image restoration. *SIAM J. Imaging Sci.*, 35, 2013. 6, 7, 8