Supplementary: Wiener Guided DIP for Unsupervised Blind Image Deconvolution

Gustav Bredell[1]   Ertunc Erdil[1]   Bruno Weber[2]   Ender Konukoglu[1]

[1]Computer Vision Laboratory, ETH Zürich   [2]University of Zürich

{gustav.bredell, ertunc.erdil, ender.konukoglu}@vision.ee.ethz.ch

1. Overview of Supplementary

In this supplementary we will show a visual example of the SSIM favoring over-smoothed images and we will provide a more in depth explanation regarding the kernel shift during optimization as well as for our hyper-parameter tuning. Furthermore, the microscopy dataset will be introduced with visual examples. In addition, we will provide qualitative examples for non-uniform blur [3] and for all five datasets that were introduced in the main article namely, Levin *et al*. [5], Lai *et al*. [4], the generated microscopy dataset, Sun *et al*. [10] and on real-world images [4]. Lastly, we also include two more visual examples of the high frequency artifact suppression by DIP.

2. SSIM favoring over-smoothed images

As pointed out in the main text, the SSIM metric tends to provide higher scores to images that are visually over-smoothed and could be an explanation why W-DIP had a lower score for the SSIM metric, compared to other baseline methods such as Pan *et al*. [6]. In Fig. 1 an image that was deblurred with W-DIP and Pan *et al*. [6], respectively, is compared. It can be seen that especially objects with many fine details, such as the pullover of the child, are over-smoothed by Pan *et al*. [6].

Pan *et al*. [6], SSIM:0.9350   W-DIP, SSIM:0.9180

Figure 1: Visual example of SSIM favoring over-smoothing of image on Levin *et al*. [5] dataset

3. Kernel Shift

As we mentioned in the main article, k and $g_\phi(z_k)$ could represent the same kernel, but a shifted version to each other. In such cases, the kernel matching term weighted by β in Eq. (6) in the main article produces large loss due to translation between the kernels. During optimization, we shift the auxiliary kernel k with regard to the generated kernel g_ϕ to prevent inducing a loss for shifted kernels as can be seen in Fig. 2. The shift is performed by first padding k with zeros on each side equal to half of the estimated kernel size. As a next step, patches with stride equal to one and with a patch size equal to estimated kernel size is extracted from the padded k. The patch is then selected, which has the lowest summed squared error with the generated kernel, g_ϕ. Based on the shift of the selected patch to the generated kernel the deconvolved image $W(I_B, k)$ is then also shifted with regard to the generated image $f_\theta(z_I)$ by an equal amount.

4. Microscopy Dataset

The microscopy dataset is composed of four sharp images and four kernels as can be seen in Fig. 3. The images were generated by Schneider *et al*. [9] and four 2D slices were selected from the 3D volume and were cropped to 255×255. The selected images are displayed in the top row of Fig. 3. These sharp images were then blurred with one kernel form Levin *et al*. [5] (left-side) and 3 generated kernels shown in the bottom row of Fig. 3. The size of the kernels range between 13×13 to 41×41 and were chosen to be in the range of point spread function sizes expected in two-photon microscopy [1]. The shown kernels are the "bad" kernel size estimates, as can be seen by the large zero space around each kernel, especially kernel 3. The full dataset including the blurred images can be found in the provided data.

| Auxiliary Kernel, k | Generated Kernel, g_ϕ | Shifted Auxiliary Kernel |

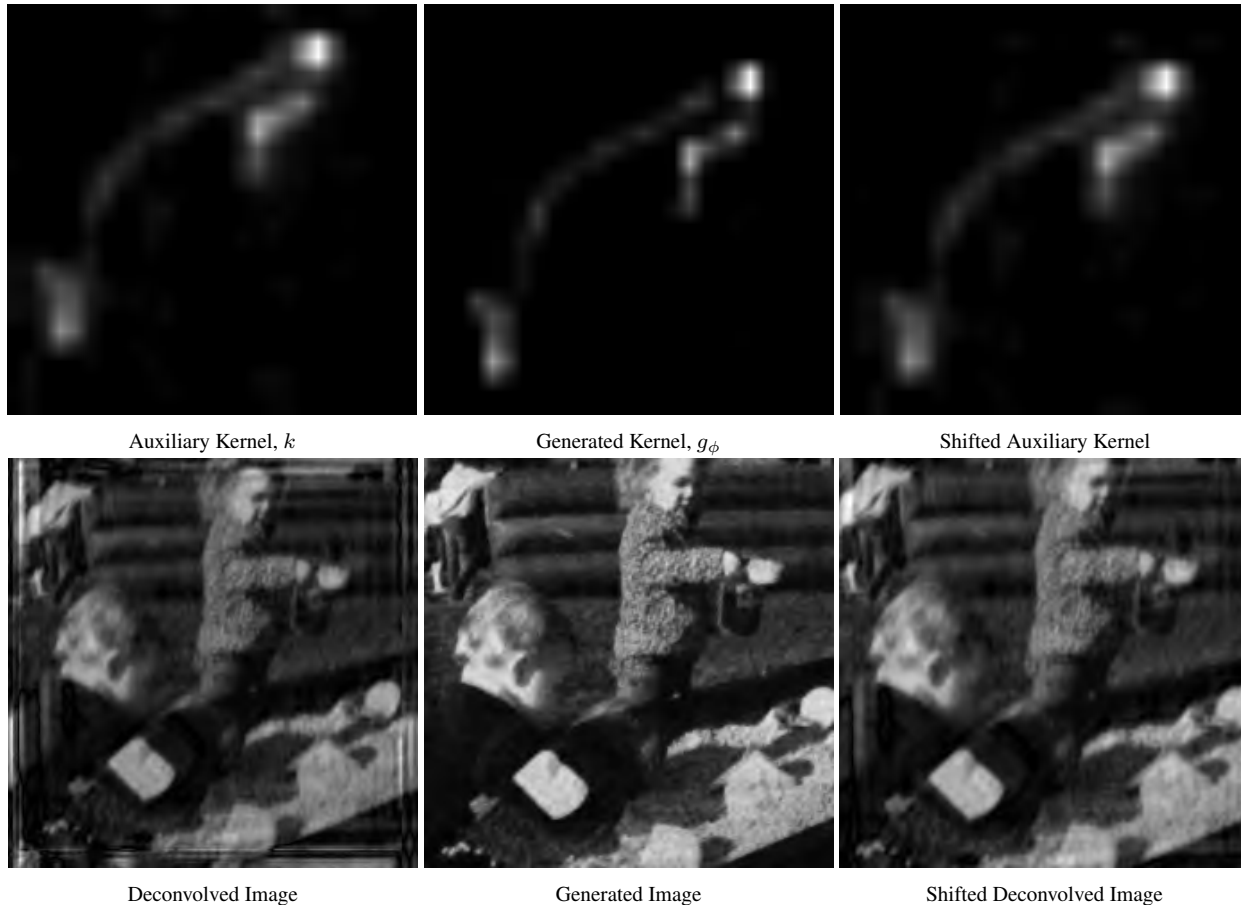| Deconvolved Image | Generated Image | Shifted Deconvolved Image |

Figure 2: Visual example showing how the auxiliary kernel and generated kernel can be shifted to each other and how the shift is corrected for the kernel as well as the generated image. Note that in the process of shifting the image, the images needs to be cropped to allow space for shift.

5. Qualitative Results Microscopy Dataset

As mentioned in the main paper, we observed a performance fluctuation for neural blind deconvolution without Wiener-guidance [8] if the kernel size estimate was changed for the microscopy dataset. This is shown qualitatively in Fig. 4. In the top row the results for Ren *et al.* [8], W-DIP and W-DIP fine-tuned is shown when the kernel estimate is "bad", meaning the size of the kernels are estimated too large. In the bottom row the results for the same methods can be seen if the kernel size esimate is more accurate, denoted as "good". For this particular kernel the size estimate decreased from 31×31 ("bad") too 21×21 ("good"). Whereas Ren *et al.* [8] seems to profit from the more accurate kernel size estimate the results of our method is more stable to this change.

6. Hyper-parameter Tuning

To select the weighting parameters (α, β, λ) four images from the Levin *et al.* [5] dataset were used, namely: I1K4,

I2K2, I2K7 and I4K8, where "I" and "K" represent the respective sharp image and kernel that was used to blur it. The examples can be found in the provided data. An initial random weight assignment in the range of $[1, 1e^{-1}, 1e^{-2}, 1e^{-3}, 1e^{-4}, 1e^{-5}]$ was performed for the weights and for the learning rate a range of $[1e^{-1}, 1e^{-2}, 1e^{-3}, 1e^{-4}, 1e^{-5}, 1e^{-6}, 1e^{-7}, 1e^{-8}]$ was investigated. Based on the results of the initial random parameter search, the search space was condensed to $[1e^{-2}, 1e^{-3}, 1e^{-4}]$ for the weights and $[1e^{-5}, 1e^{-6}, 1e^{-7}]$ for the learning rate. Within this parameter range the results on the selected images was stable and a final selection was done leading to the parameter weights, $\alpha = 1e^{-3}$, $\beta = 1e^{-4}$, $\lambda = 1e^{-3}$ and setting the learning rate for the auxiliary kernel to $1e^{-6}$.

The fine-tuning of the weights for the microscopy dataset was done within the range of the condensed search space but with the addition of $[1e^{-1}]$ for the parameter weights. The images I0K0 and I0K1 were used to evaluate on, where "I" and "K" represent the respective sharp image and kernel that was used to blur it. The examples can be found in
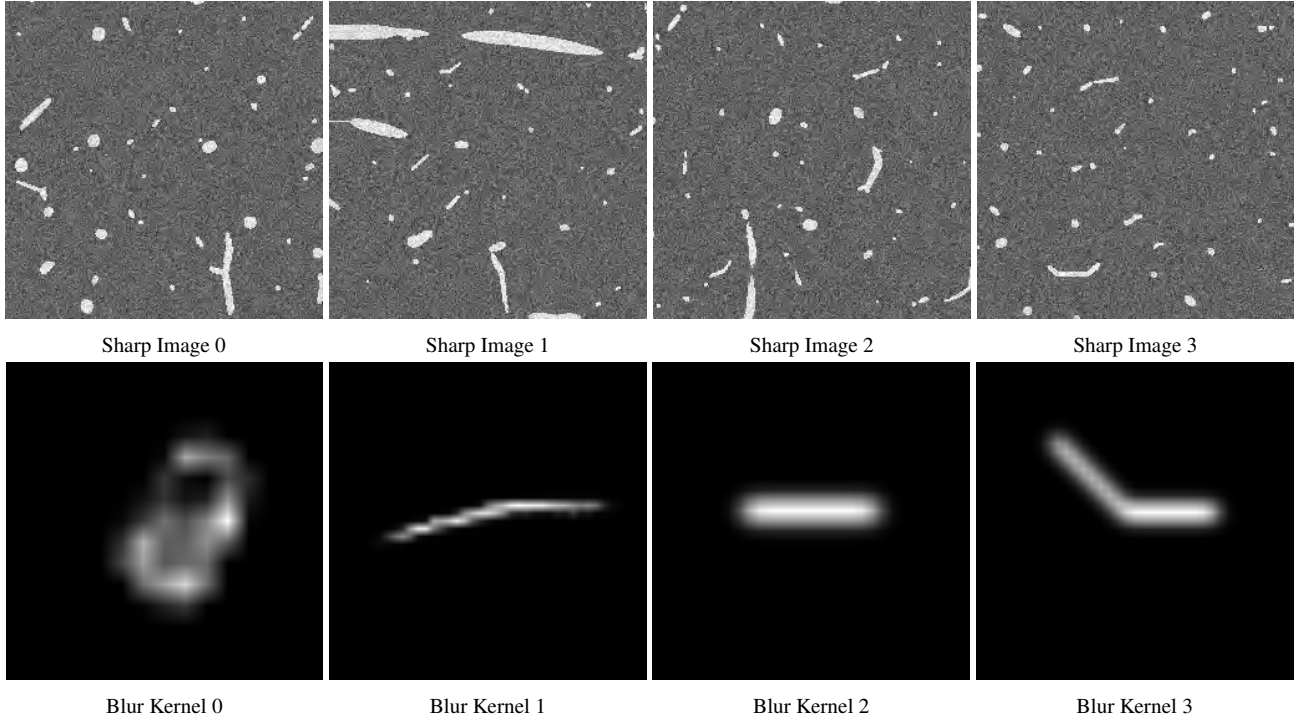
Sharp Image 0   Sharp Image 1   Sharp Image 2   Sharp Image 3

Blur Kernel 0   Blur Kernel 1   Blur Kernel 2   Blur Kernel 3

Figure 3: The sharp images and blur kernels of the microscopy dataset is shown.

Input   SelfDeblur$_{bad}$ W-DIP$_{bad}$ W-DIP FT$_{bad}$ GT$_{bad}$

Input   SelfDeblur$_{good}$ W-DIP$_{good}$ W-DIP FT$_{good}$ GT$_{good}$
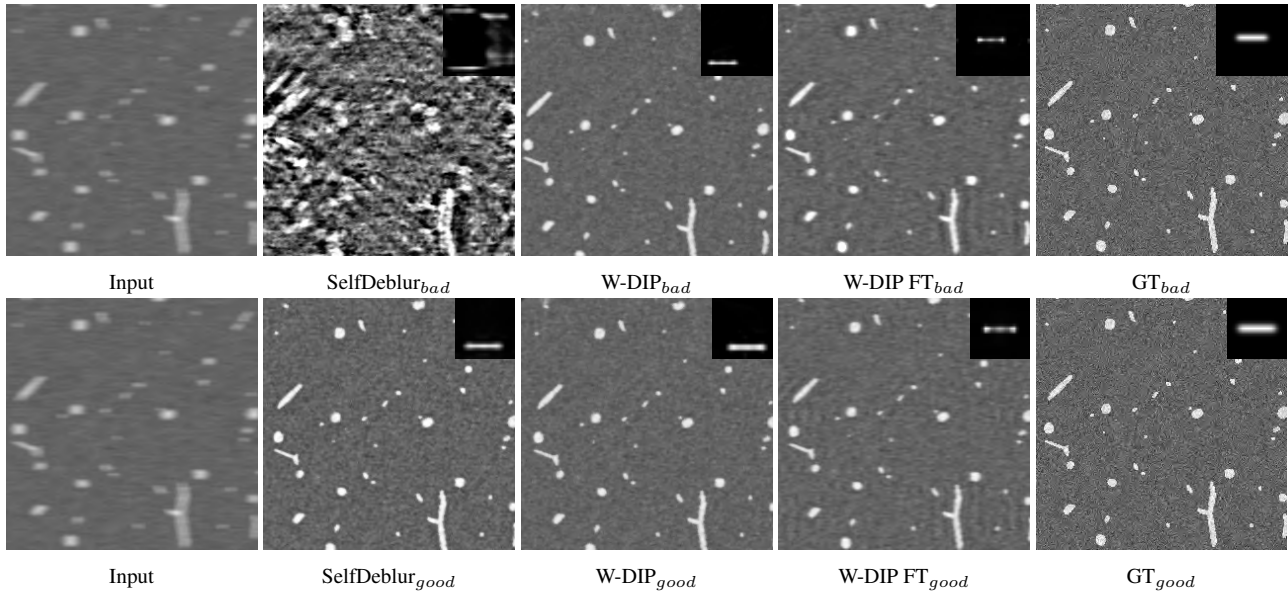
Figure 4: The results of blind deconvolution can be seen for Ren *et al.* [8], W-DIP, W-DIP fine-tuned (FT) for different kernel size estimates. In the bottom row it can be seen that if the kernel size estimate is "good" then all methods perform well with respect to the ground truth (GT). In the top row it is shown that if the kernel size estimate is "bad", that means too large in this case, that only methods with Wiener-deconvolution guidance arrive at reasonable solutions.

the provided data. The final selection for the microscopy dataset was, $\alpha = 1e^{-2}$, $\beta = 1e^{-1}$, $\lambda = 1e^{-2}$ and setting the learning rate for the auxiliary kernel to $1e^{-7}$.

7. Qualitative Examples Non-uniform Blur

Blur in the real-world is not always uniform as many of the benchmark datasets assume. To show that our method

can also handle real-world blur, we included qualitative examples from the Lai *et al*. [4] dataset in the main article. Here we provide qualitative examples of specifically non-uniform blur taken from the dataset Kohler *et al*. [3] and is shown in Fig. 5-6. It can be seen that W-DIP can still provide good results for kernels that vary little spatially, Fig.5. Even for larger kernels W-DIP produce reasonable results, Fig.6, especially compared to neural blind deconvolution without Wiener-guidance [8]. Failure cases have also been observed and more quantitative results will be necessary to determine to what extent W-DIP can address non-uniform blur.

8. Additional Qualitative Examples

In this section we provide additional qualitative examples on the introduced datasets. In Figures 7, 8, 9 qualitative comparisons on the Lai *et al*. [4] dataset is shown from the categories not shown in the main paper namely, "man-made", "saturated" and "text".

Next three examples from the Levin *et al*. [5] is shown in Figures 10, 11, 12. It should be noted that in general the performance of Ren *et al*. [8] can be good and some of the examples taken are failure cases that are run dependent. Nevertheless, all three images from Ren *et al*. [8] and W-DIP are from the same run.

Lastly, we show results on the Sun *et al*. [10] in Figures 13 and 14 and real-world images [4] in Figures 16 and 15.

9. High-frequency Artifact Suppression

To further provide qualitative examples for the delayed reproduction of high-frequency artifacts of deconvolution we show DIP results for the Levin *et al*. [5] and custom microscopy dataset below in Figs. 17 and 18, respectively. The suppression is best appreciated when zoomed in.

References

[1] Emmanuelle Chaigneau, Amanda J Wright, Simon P Poland, John M Girkin, and R Angus Silver. Impact of wavefront distortion and scattering on 2-photon microscopy in mammalian brain tissue. *Optics express*, 19(23):22755–22774, 2011.

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

[3] Rolf Köhler, Michael Hirsch, Betty Mohler, Bernhard Schölkopf, and Stefan Harmeling. Recording and playback of camera shake: Benchmarking blind deconvolution with a real-world database. In *European conference on computer vision*, pages 27–40. Springer, 2012.

[4] 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.

[5] Anat Levin, Yair Weiss, Fredo Durand, and William T Freeman. Understanding and evaluating blind deconvolution algorithms. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1964–1971. IEEE, 2009.

[6] 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.

[7] 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.

[8] Dongwei Ren, Kai Zhang, Qilong Wang, Qinghua Hu, and Wangmeng Zuo. Neural blind deconvolution using deep priors. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 3341–3350, 2020.

[9] Matthias Schneider, Sven Hirsch, Bruno Weber, Gábor Székely, and Bjoern H Menze. Joint 3-d vessel segmentation and centerline extraction using oblique hough forests with steerable filters. *Medical image analysis*, 19(1):220–249, 2015.

[10] 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.

[11] Oliver Whyte, Josef Sivic, and Andrew Zisserman. Deblurring shaken and partially saturated images. *International journal of computer vision*, 110(2):185–201, 2014.

[12] Li Xu and Jiaya Jia. Two-phase kernel estimation for robust motion deblurring. In *European conference on computer vision*, pages 157–170. Springer, 2010.

[13] Wangmeng Zuo, Dongwei Ren, David Zhang, Shuhang Gu, and Lei Zhang. Learning iteration-wise generalized shrinkage–thresholding operators for blind deconvolution. *IEEE Transactions on Image Processing*, 25(4):1751–1764, 2016.
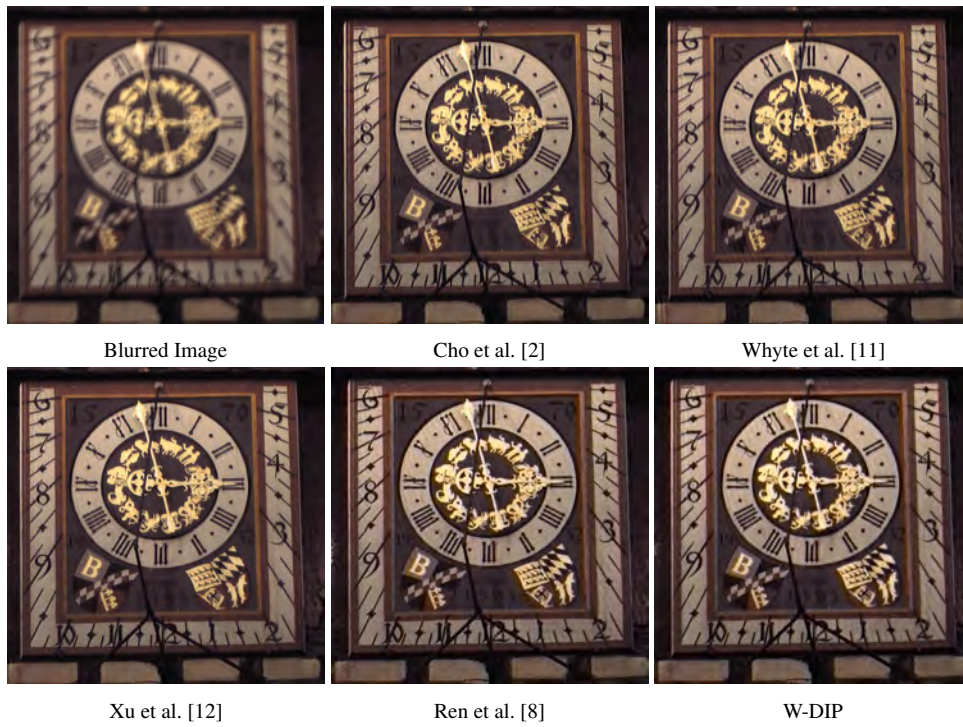
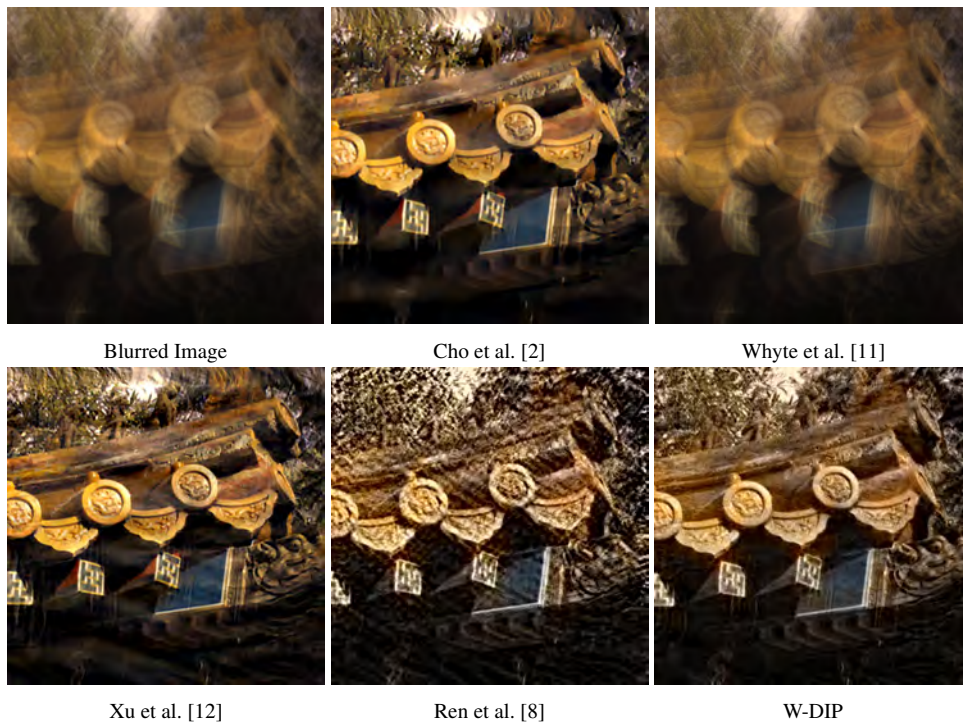| Blurred Image | Cho et al. [2] | Whyte et al. [11] |

| Xu et al. [12] | Ren et al. [8] | W-DIP |

Figure 5: Visual results from the non-uniform blur dataset by Kohler et al. [3]. W-DIP is compared with four baselines.
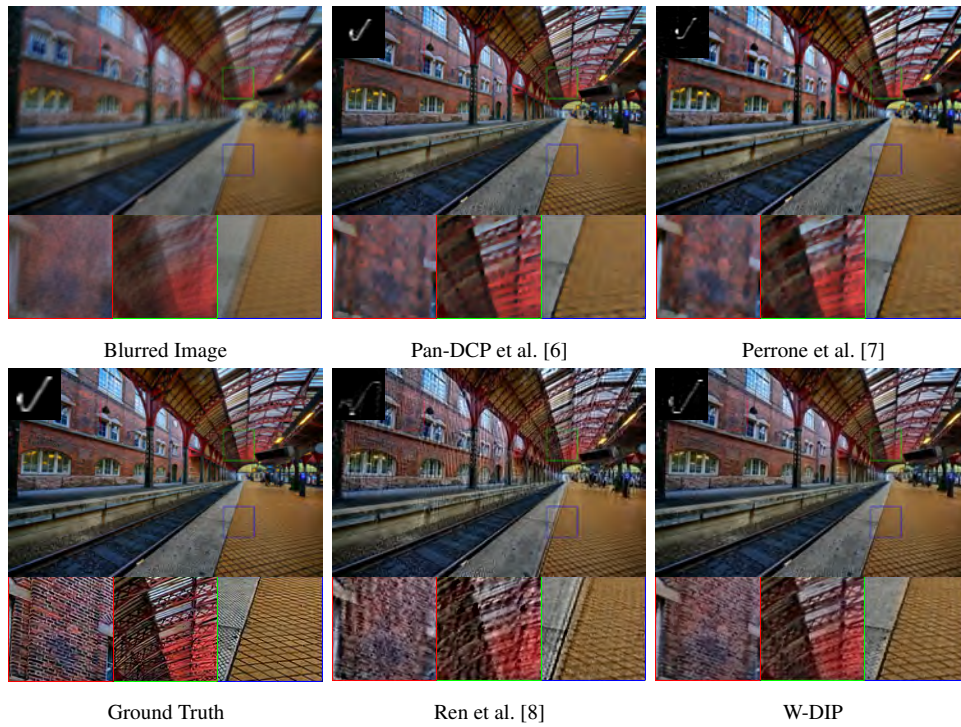

| Blurred Image | Cho et al. [2] | Whyte et al. [11] |

| Xu et al. [12] | Ren et al. [8] | W-DIP |

Figure 6: Visual results from the non-uniform blur dataset by Kohler et al. [3]. W-DIP is compared with four baselines.

Figure 7: Additional visual results from the Lai et al. [4] dataset to compare W-DIP with three baselines.
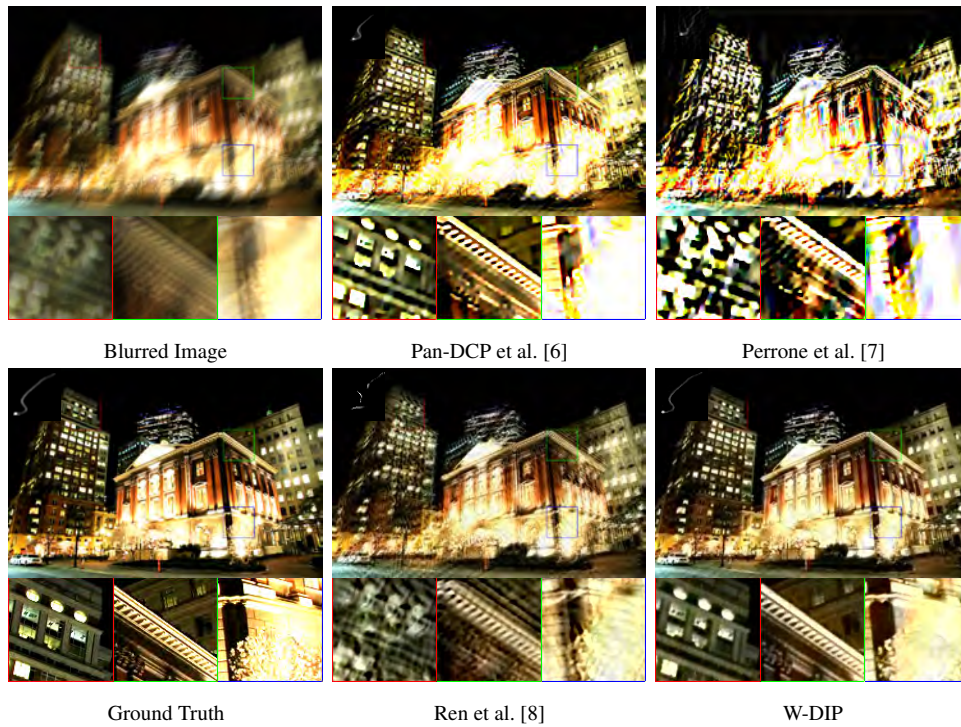
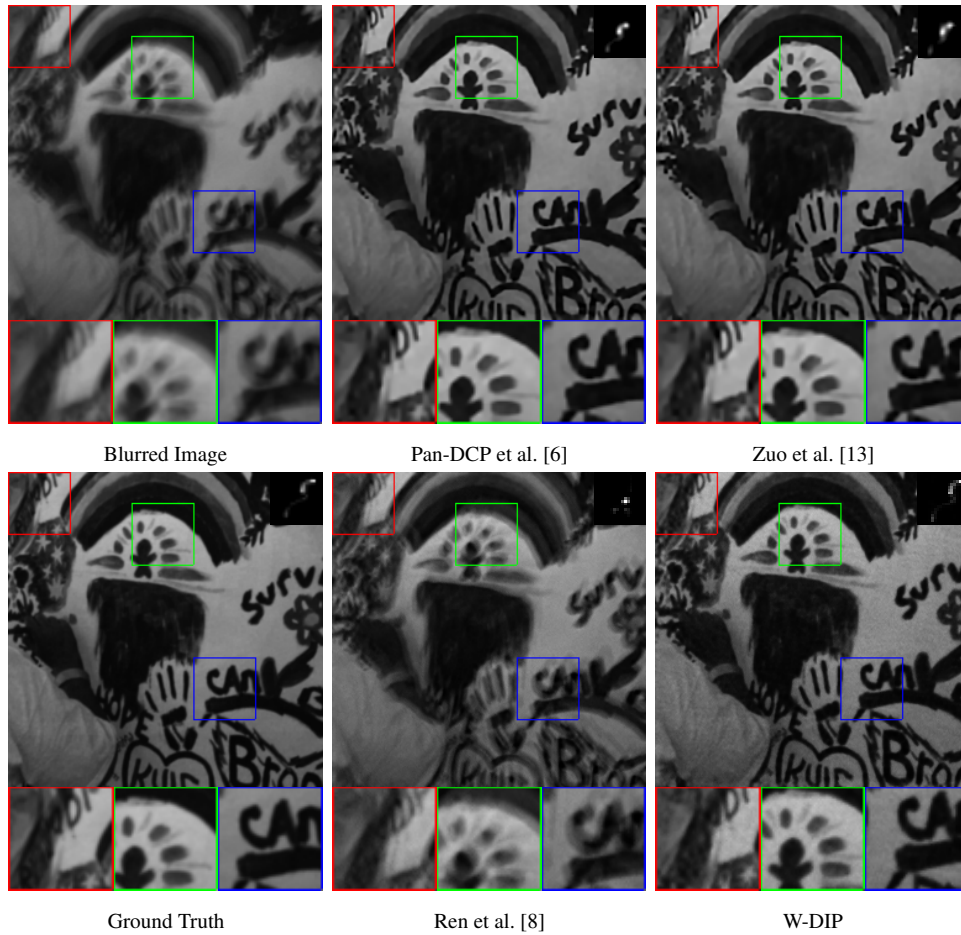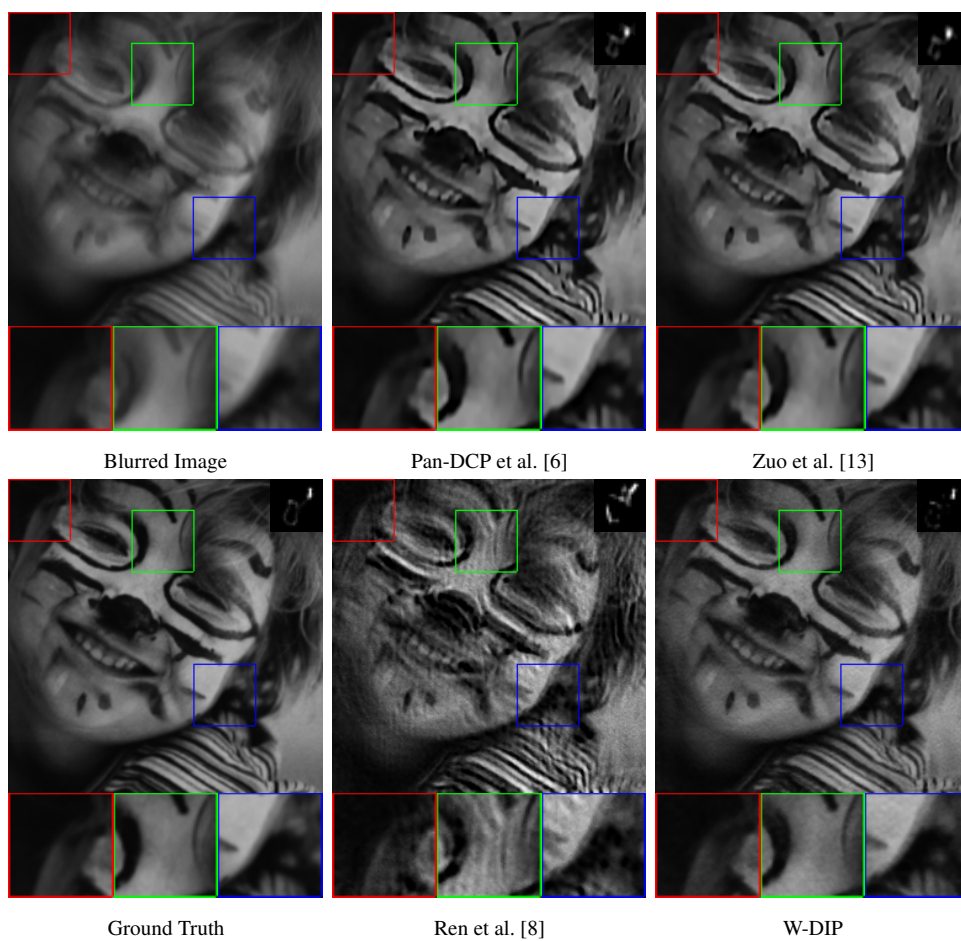Figure 8: Additional visual results from the Lai et al. [4] dataset to compare W-DIP with three baselines.

Blurred Image	Pan-DCP et al. [6]	Perrone et al. [7]
Ground Truth	Ren et al. [8]	W-DIP
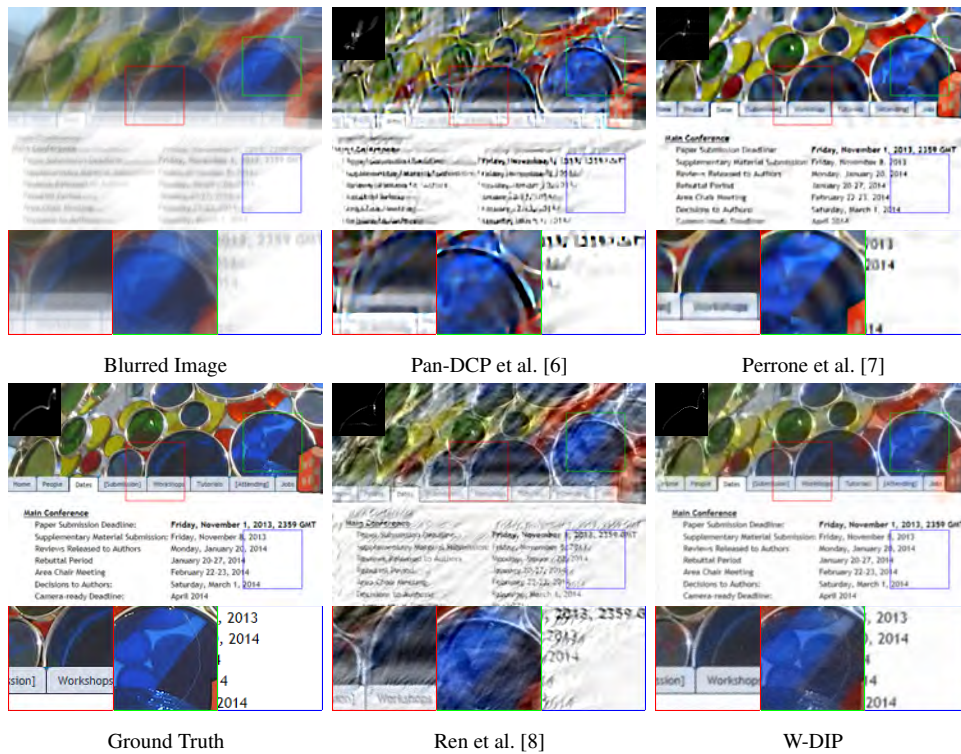
Figure 9: Additional visual results from the Lai et al. [4] dataset to compare W-DIP with three baselines.
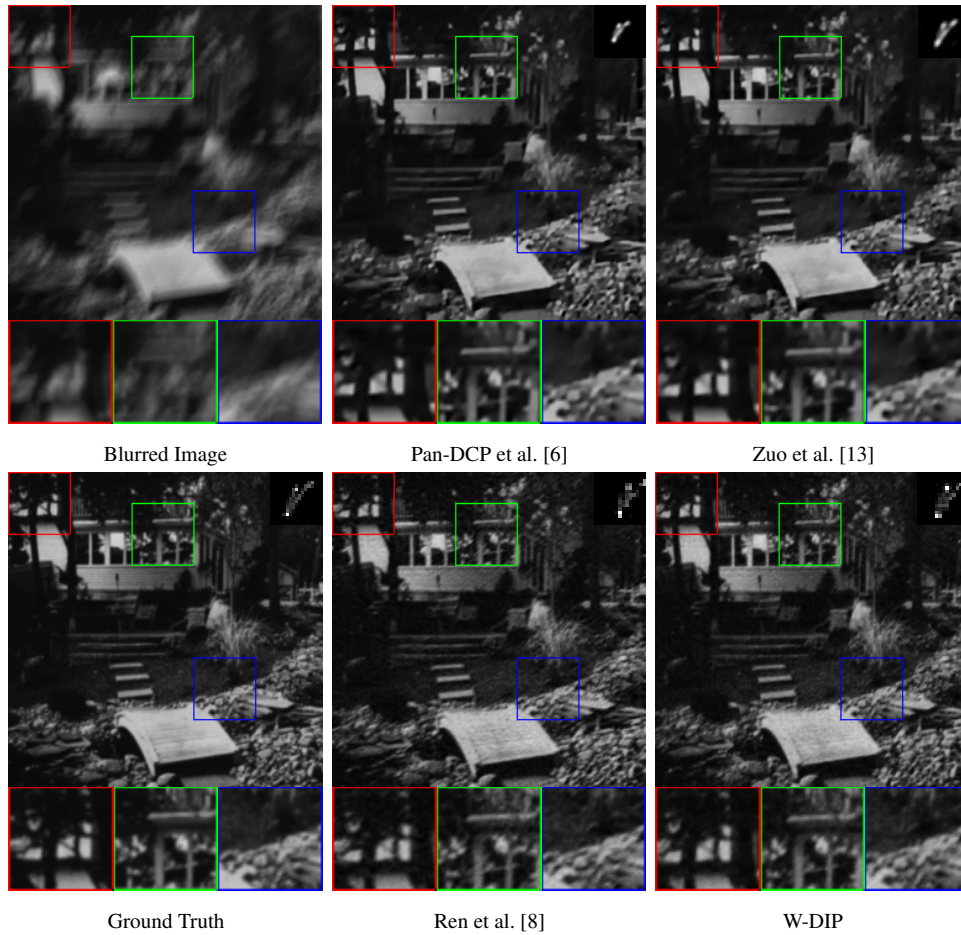
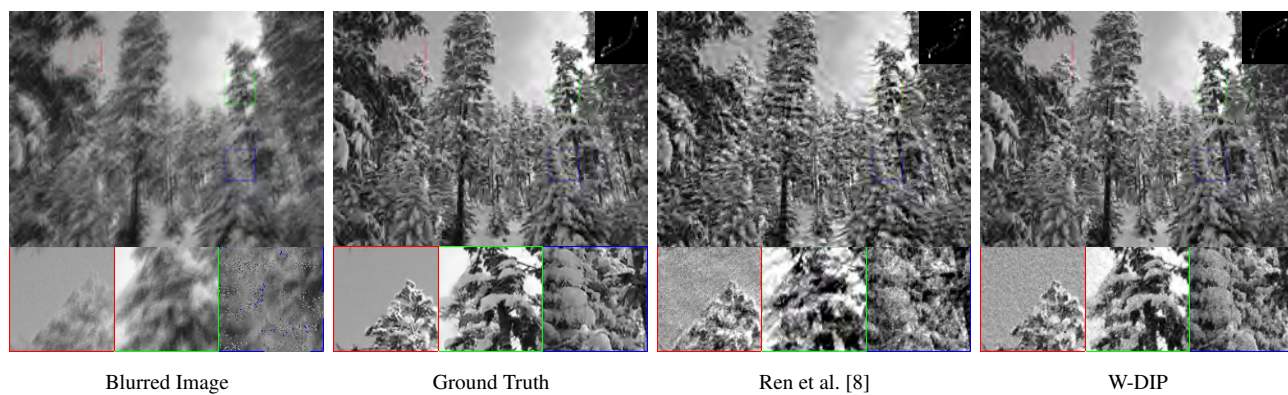Figure 10: Additional visual results from the Levin et al. [5] dataset to compare W-DIP with three baselines.

Figure 11: Additional visual results from the Levin et al. [5] dataset to compare W-DIP with three baselines.

Figure 12: Additional visual results from the Levin et al. [5] dataset to compare W-DIP with three baselines.

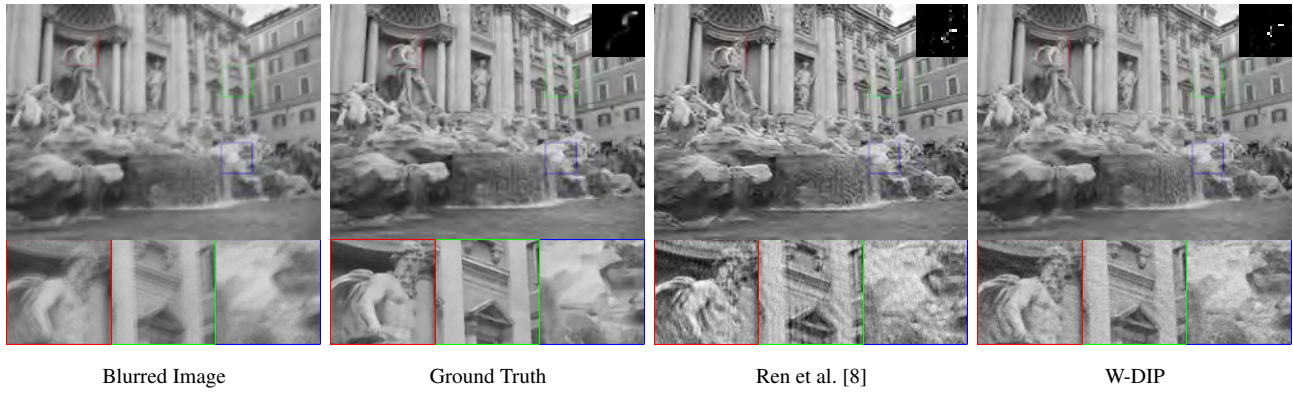Figure 13: Additional visual results from the Sun et al. [10] dataset to compare W-DIP with SelfDeblur [8].

| Blurred Image | Ground Truth | Ren et al. [8] | W-DIP |

Figure 14: Additional visual results from the Sun et al. [10] dataset to compare W-DIP with SelfDeblur [8].

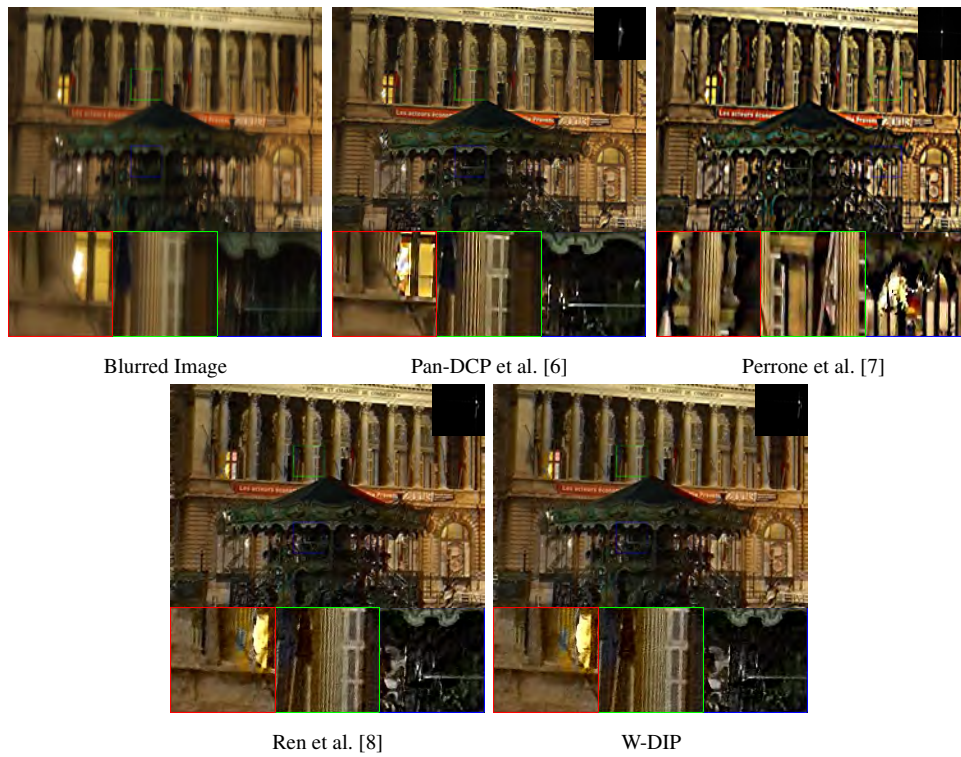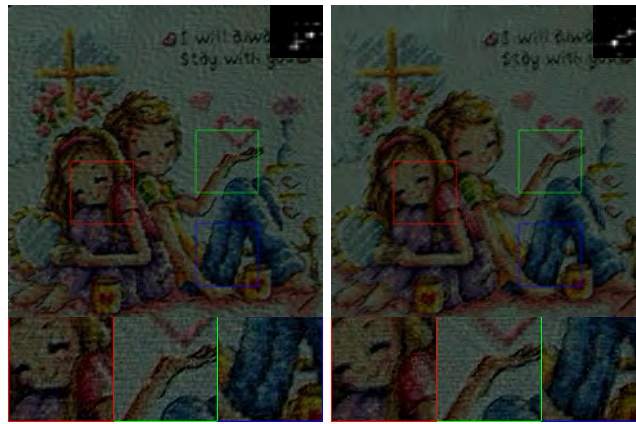| Blurred Image | Pan-DCP et al. [6] | Perrone et al. [7] |

| Ren et al. [8] | W-DIP |

Figure 15: Additional visual results from the Real world images dataset [4] to compare W-DIP with three baselines.

Blurred Image   Pan-DCP et al. [6]   Perrone et al. [7]

Ren et al. [8]   W-DIP

Figure 16: Additional visual results from the Real world images dataset [4] to compare W-DIP with three baselines.
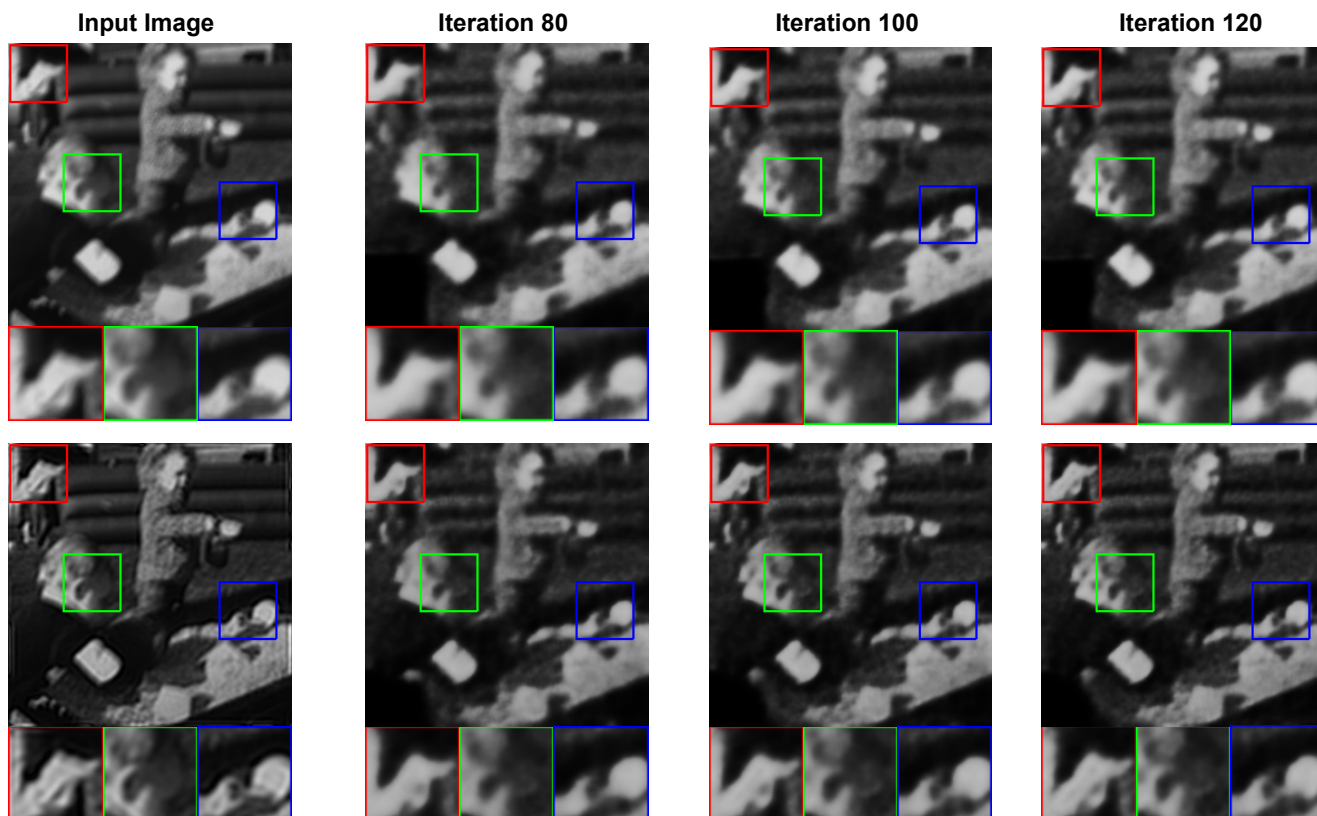
Figure 17: Visual results that show the high-frequency artifact suppression by DIP for the Levin et al. [5] dataset. In the top row the reconstruction result of DIP with the blurred image as target is shown. No transition over the sharp version of the blurred image can be seen. In the bottom row the target image of DIP is the blurry image after it was deconvolved with a Gaussian kernel. It can be seen that the introduced artifacts are suppressed, especially in the blue rectangle.
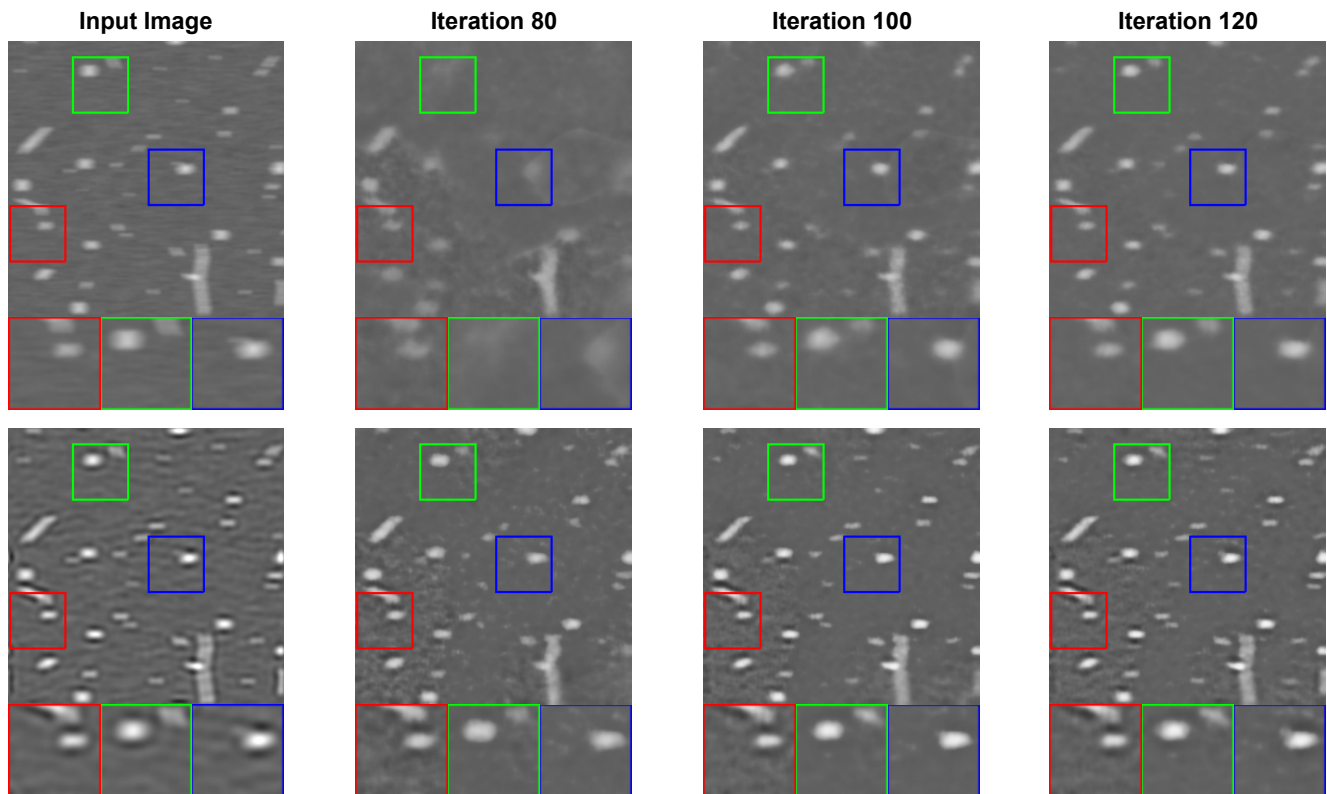
Figure 18: Visual results that show the high-frequency artifact suppression by DIP for the custom microscopy dataset. In the top row the reconstruction result of DIP with the blurred image as target is shown. No transition over the sharp version of the blurred image can be seen. In the bottom row the target image of DIP is the blurry image after it was deconvolved with a Gaussian kernel. It can be seen that the introduced artifacts are suppressed, especially in the green rectangle.