

# Simultaneous Fidelity and Regularization Learning for Image Restoration

## *Supplementary Material*



In this supplementary file, we present the results of Gaussian denoising, visualization of learned filters, interpretability and flexibility of fidelity term, discussion of hyper-parameter settings as well as more results of SFARL.

### 1 GAUSSIAN DENOISING

For Gaussian denoising, when the upper-level objective in Eqn. (11) is visual perception metric, the intermediate denoising result  $x^t$  still has noises but is not Gaussian distributed. Therefore, our SFARL can be used to address this issue by learning proper fidelity term, and benefits visual perception metric and visual quality of the denoising result. To this end, we train three SFARL models for Gaussian denoising with different noise levels, i.e.,  $\sigma = 15, 25, 50$ , respectively. The training dataset consists of 500 images, where 400 are from the BSD dataset [6] and 100 are randomly selected from the UCID dataset [7].

We compare the SFARL model with the state-of-the-art denoising algorithms based on patches, i.e., BM3D [1], EPLL [5] and WNNM [2] approaches, fields of experts, i.e., TNRD [3], and deep CNN, i.e., DnCNN [4]. Table s1

shows the average SSIM values by different methods on the BSD68 test images. The SFARL algorithm achieves the best average SSIM values at all the noise levels. Note that the SSIM metric is known to be more consistent with human visual perception on image, and the SFARL algorithm performs favorably against the other methods. Fig. s1 shows the denoising results with  $\sigma = 50$ . As shown in the red close-ups, the results by the BM3D and WNNM methods contain visible artifacts, while the results by the SFARL algorithm is visually more pleasant. Compared with the TNRD and DnCNN, the SFARL model recovers more texture details, indicating that the fidelity term should not be the  $\ell_2$ -norm when the objective is to minimize the negative SSIM loss.

### 2 VISUALIZATION OF FILTERS

We in this supplementary file visualize the learned filters in both fidelity and regularization terms on these three tasks. On one hand, as shown in Fig. s2, filters in regularization term are not visually dependent with restoration task, mainly attributing to that the regularization term is used to model natural image

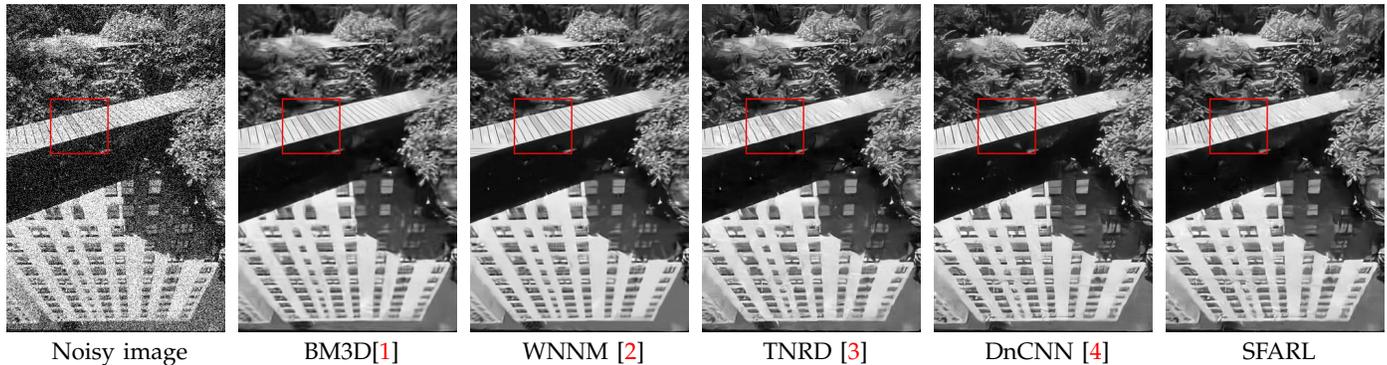


Figure s1: Gaussian denoising ( $\sigma = 50$ ) results by all evaluated methods on a test image.

Table s1: Gaussian denoising on the BSD68 test images in terms of SSIM.

Noise $\sigma$	BM3D [1]	EPLL [5]	WNNM [2]	TNRD [3]	DnCNN [4]	SFARL
15	0.8604	0.8713	0.8756	0.8773	0.8860	0.8869
25	0.8028	0.8123	0.8135	0.8160	0.8183	0.8197
50	0.6864	0.6948	0.7010	0.7024	0.7104	0.7120

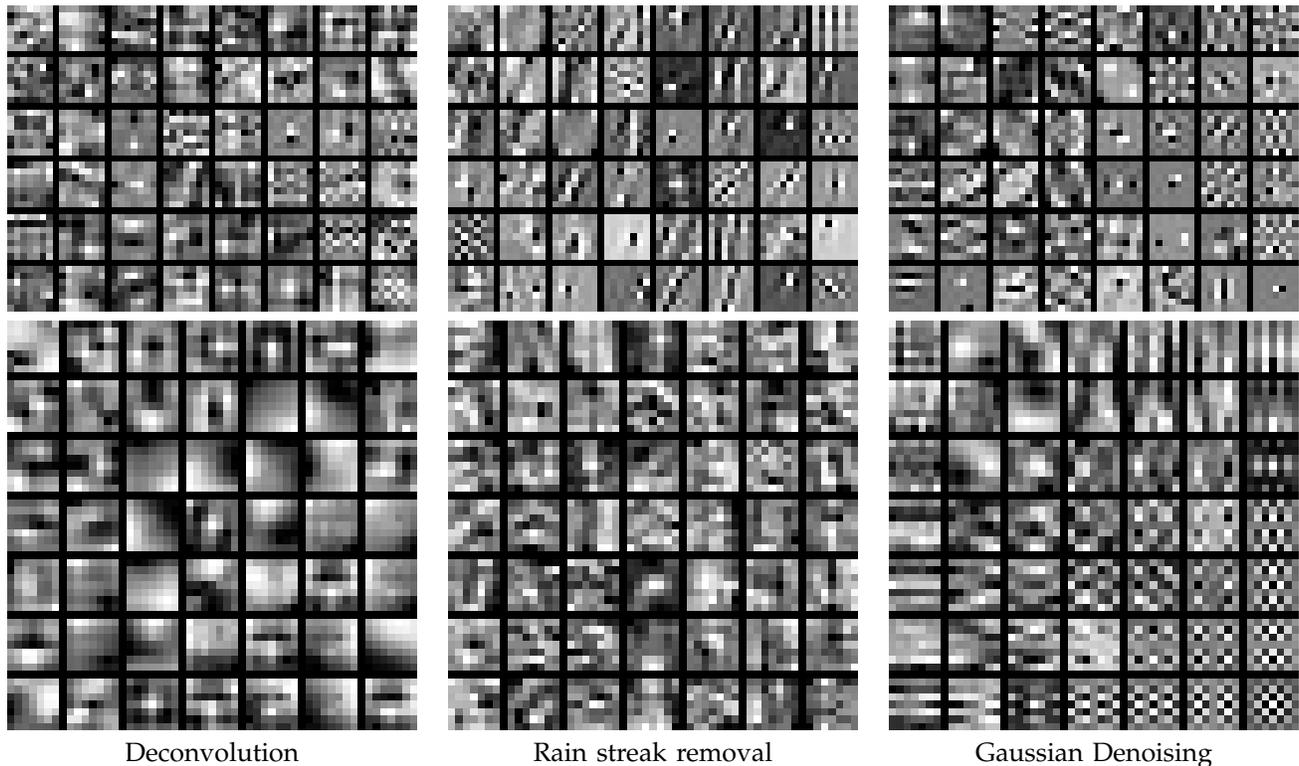


Figure s2: Visualization of learned filters for 3 restoration tasks. The first row is 48 filters in regularization term, while the second row presents 49 filters in fidelity term.

priors. On the other hand, filters in fidelity term are quite different for these three tasks, and have task-specific patterns correlated with degradation types. As for deconvolution with inaccurate blur kernels, mildly blurry filters are learned to model complex patterns of camera

motion blur kernels. As for rain streak removal, learned filters have complex and sharp patterns to model rain streaks with different scales and orientations. Since Gaussian denoising is a much simple task, learned filters are not as complex as those in deconvolution and rain

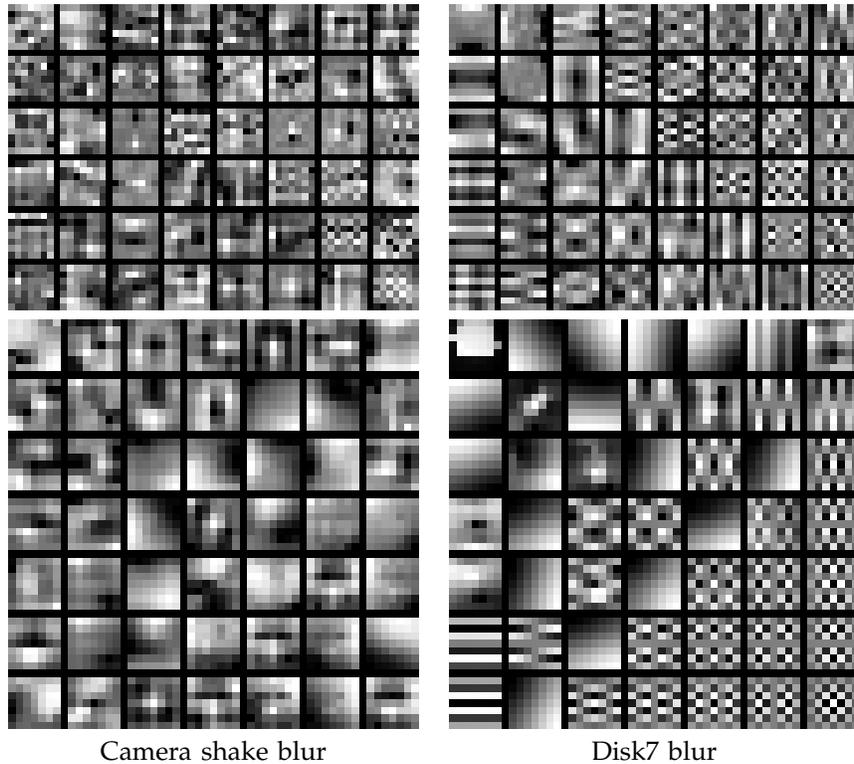


Figure s3: Visualization of learned filters of SFARL trained on camera shake blur dataset and Disk7 blur dataset. The first row is 48 filters in regularization term, while the second row presents 49 filters in fidelity term.

streak removal. Also it is interesting to see that filters in the right bottom corner are nearly the same with DCT filters, indicating that they are not changed much during training.

Furthermore, the learned filters in fidelity term are also associated with diversity of synthetic training dataset. We in Fig. s3 show the learned filters of two SFARL models for deconvolution, which are trained on camera motion blur datasets and Disk7 blur datasets [15], respectively. In Disk7 blur dataset, only one disk kernel with radius 7 is used to generate blurry images (Please see Section 5.2 for details). Compared with the filters in fidelity term for camera motion blur, the ones for Disk7 blur have less patterns. For example, the first filter is very similar to a disk kernel.

### 3 INTERPRETABILITY AND FLEXIBILITY OF FIDELITY TERM

Using rain streak removal as an example, we discuss the interpretability and flexibility of non-linear functions in fidelity term. In the fidelity term (9), the filters  $\mathbf{p}_i$  are used to extract features of rain streaks from residual images, while the distribution of filter responses can be characterized with the penalty function  $\mathcal{D}_i$ . Using the 8-th filter  $\mathbf{p}_8$  as an example, Fig. s4 shows an response image, penalty function  $\mathcal{D}_8$  and its derivative (i.e., influence function)  $\varphi_8$ , and the distribution of  $\mathbf{p}_8 \otimes \mathbf{r}_s$  on all samples. It can be seen that the response image is closely dependent with the rain streaks. Moreover, the influence function is much more complex and flexible than the nonlinear activations in CNNs, e.g., ReLU and its variants. Furthermore, the shape of  $\mathcal{D}_8$  is similar to that of the distribution of responses  $\mathbf{p}_8 \otimes \mathbf{r}_s$ , indicating that that the

learned  $\mathcal{D}_8$  and  $\varphi_8$  can be used to characterize the distribution of responses intuitively. Fig. s5 presents several learned influence functions  $\varphi_i$  (15) and their corresponding  $\mathcal{D}_i$  (9), further demonstrating their complexity and flexibility.

To sum up, in comparison to conventional CNNs, the fidelity term (9) can present better interpretability and flexibility in characterizing the residual. In terms of interpretability, the response image of  $\mathbf{p}_i$  is visually dependent with the rain streaks, and the learned  $\mathcal{D}_i$  and  $\varphi_i$  can intuitively characterize the distribution of  $\mathbf{p}_i \otimes \mathbf{r}_s$ . In terms of flexibility, the distribution of  $\mathbf{p}_i \otimes \mathbf{r}_s$  is much more complex, and cannot be simply characterized by ReLU and its variants.

## 4 HYPER-PARAMETERS SETTING

The hyper-parameters in our SFARL model only include filter size  $s \times s$  and stage number  $T$ , based on which the numbers of filters and penalty functions in regularization term  $N_r$  and in fidelity term  $N_f$  can be accordingly set.

As for filter size  $s \times s$ , we have  $s^2$  complete DCT basis. And thus, we adopt  $N_f = s^2$  filters in fidelity term. But in regularization term, it is suggested that only high frequency filters are useful to model natural image priors [3,12]. Thus, the first DC component in DCT filters is excluded, and the number of filters in regularization is  $N_r = s^2 - 1$ . In our implementation, SFARL can take advantage of recent advanced optimization algorithms, e.g., ADAM, and GPU parallel acceleration, which allows us to take large size filters, i.e.,  $7 \times 7$  filters, in both fidelity term and regularization term.

As for stage number, we set  $T$  by stopping adding new stage when greedy training only makes a small contribution. And thus we empirically set the stage numbers to  $T = 10$  for

deconvolution,  $T = 5$  for rain streak removal, and  $T = 5$  for Gaussian denoising. Using rain streak removal as an example, we in Fig. s6 plot average PSNR curve of each epoch during greedy training and joint fine-tuning. From Fig. s6, it only makes a very marginal PSNR improvement by adding 5-th stage, and thus it is reasonable to set the stage number as 5 for rain steak removal. To support it, we also train a 4-stage SFARL model on the training dataset [16]. Table s2 reports average PSNR and SSIM on 1,400 rainy images [16], from which one can see that 4-stage SFARL only performs slightly inferior to 5-stage SFARL.

Table s2: Comparison of SFARL models with different stages for rain streak removal.

Model	4-stage SFARL	5-stage SFARL
PSNR	30.56	31.37
SSIM	0.9103	0.9188

## 5 MORE RESULTS

We hereby show more results on image deconvolution with multiple degradations in Fig. s7, rain streak removal in Fig. s8 and Gaussian denoising with  $\sigma = 50$  in Fig. s9.

## REFERENCES

- [1] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *SPIE Electronic Imaging*, 2006. 1, 2
- [2] S. Gu, L. Zhang, W. Zuo, and X. Feng, "Weighted nuclear norm minimization with application to image denoising," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2014, pp. 2862–2869. 1, 2
- [3] Y. Chen, W. Yu, and T. Pock, "On learning optimized reaction diffusion processes for effective image restoration," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 5261–5269. 1, 2
- [4] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising," *IEEE Transactions on Image Processing*, vol. 26, no. 7, pp. 3142–3155, 2017. 1, 2

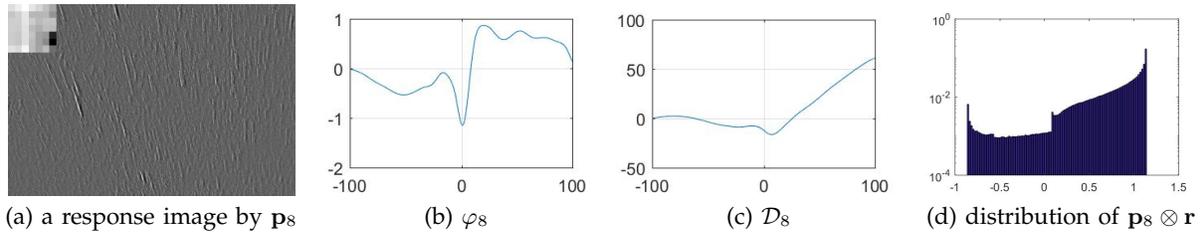


Figure s4: Illustration of fidelity term in rain streak removal.

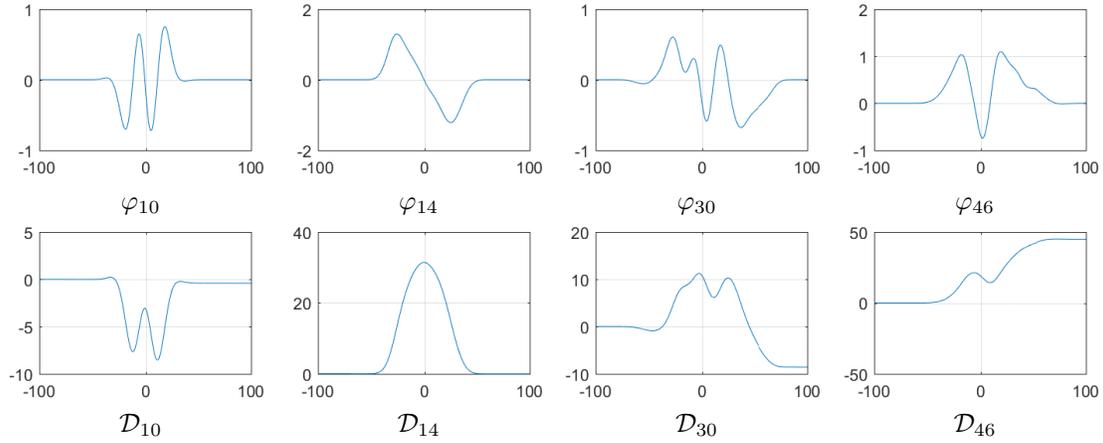
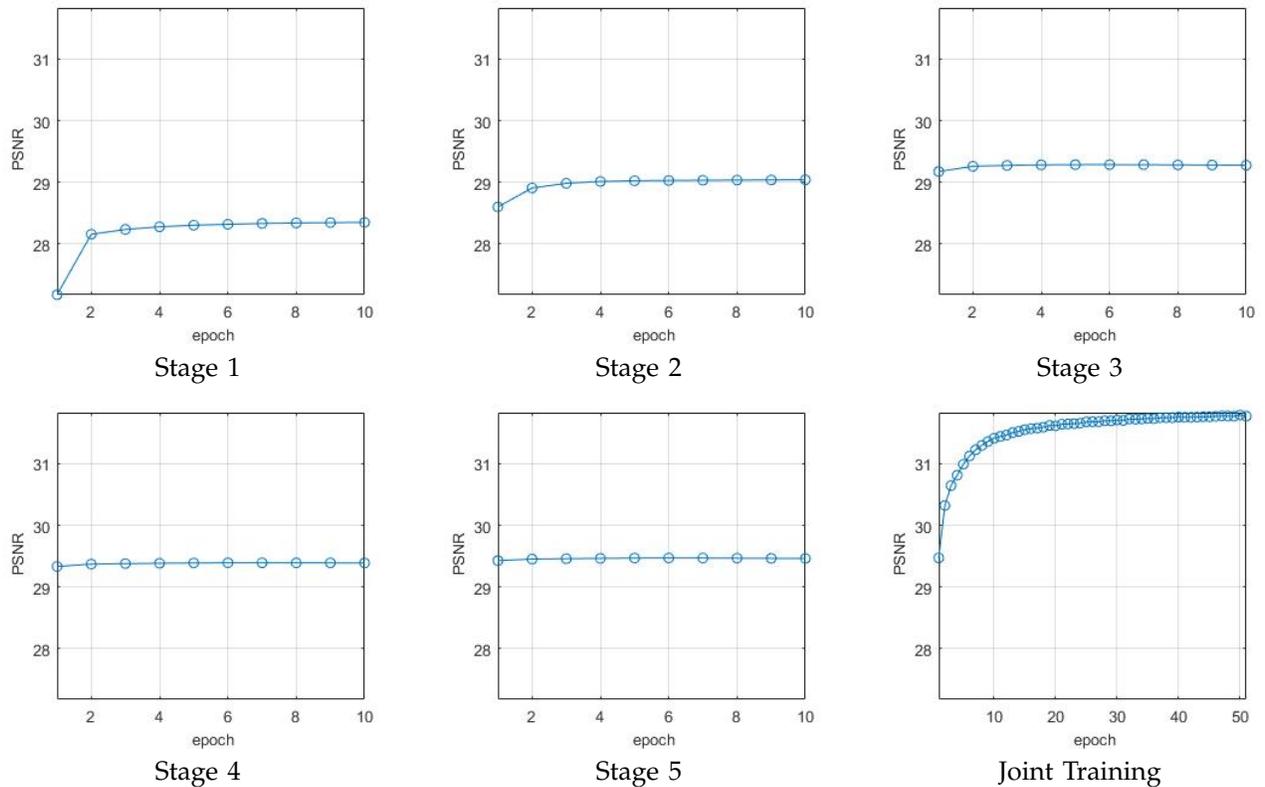
Figure s5: Examples of non-linear functions  $\varphi_i$  and  $\mathcal{D}_i$ .

Figure s6: Convergence curve of SFARL for rain streak removal in greedy training and joint training.



Blurry images

SFARL

Figure s7: More results on image deconvolution with multiple degradations by SFARL.

*national Conference on Computer Vision.* IEEE, 2011, pp. 479–486. [1](#), [2](#)

Society for Optics and Photonics, 2003, pp. 472–480. [1](#)

- [6] D. Martin, C. Fowlkes, D. Tal, and J. Malik, “A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics,” in *Proceedings of the IEEE International Conference on Computer Vision*, vol. 2. IEEE, 2001, pp. 416–423. [1](#)

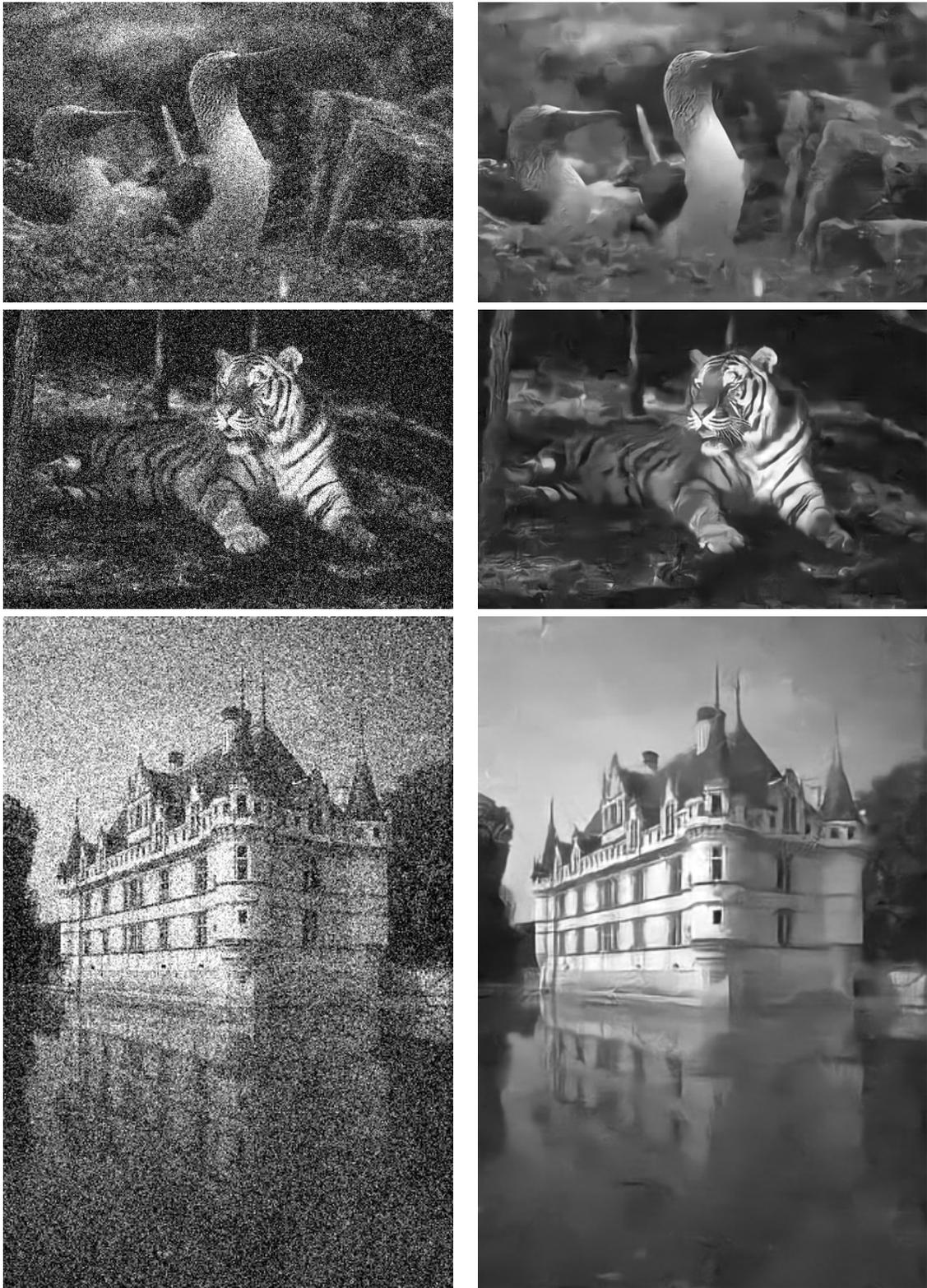
- [7] G. Schaefer and M. Stich, “UCID: An uncompressed color image database,” in *Electronic Imaging*. International



Rainy images

SFARL

Figure s8: More results on rain streak removal by SFARL.



Noisy images

SFARL

Figure s9: More results on Gaussian denoising by SFARL.