Supplementary Material for Fully Convolutional Pixel Adaptive Image Denoiser

Sungmin Cha[1] and Taesup Moon[1,2]
[1]Department of Electrical and Computer Engineering, [2] Department of Artificial Intelligence
Sungkyunkwan University, Suwon, Korea 16419
{csm9493,tsmoon}@skku.edu

1. Proof of Lemma 1

Lemma 1 *For any $\mathbf{N}$ with the assumptions in [Manuscript, Section 3.1] and $\hat{\mathbf{X}}(\mathbf{Z})$ that has the form [Manuscript, Eq.(2)] with $d \in \{1,2\}$,*

$$\mathbb{E}\mathbf{L}_n(\mathbf{Z}, \hat{\mathbf{X}}(\mathbf{Z}); \sigma^2) = \mathbb{E}\mathbf{\Lambda}_n(\mathbf{x}, \hat{\mathbf{X}}(\mathbf{Z})). \tag{1}$$

Moreover, when $\mathbf{N}$ is white Gaussian, then, $\mathbf{L}_n(\mathbf{Z}, \hat{\mathbf{X}}(\mathbf{Z}); \sigma^2)$ coincides with the SURE [2].

Proof: We note the expectation of the i-th summand in [Manuscript, Eq.(3)] is

$$\begin{aligned}
&\frac{1}{n}\mathbb{E}\Big[(Z_i - \hat{X}_i(\mathbf{Z}))^2 + \sigma^2(\sum_{m=1}^{d} 2^m a_{m,i} Z_i^{m-1} - 1)\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(Z_i - \hat{X}_i(\mathbf{Z}))^2 + \sigma^2(\sum_{m=1}^{d} 2^m a_{m,i} Z_i^{m-1} - 1)\big|\mathbf{Z}^{-i}\big]\Big]
\end{aligned} \tag{2}$$

Now, we divide into two cases, $d=1$ and $d=2$.

1) For $d=1$ (affine mapping), (2) becomes

$$\begin{aligned}
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(Z_i - (a_{1,i}Z_i + a_{0,i}))^2 \\
&\qquad + \sigma^2(\sum_{m=1}^{1} 2^m a_{m,i} Z_i^{m-1} - 1)\big|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(Z_i^2 - \sigma^2) + (a_{1,i}Z_i + a_{0,i})^2 \\
&\qquad - 2a_{1,i}(Z_i^2 - \sigma^2) - 2a_{0,i}Z_i|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(x_i^2 + (a_{1,i}Z_i + a_{0,i})^2 - 2a_{1,i}Z_i x_i - 2a_{0,i}x_i|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(x_i - (a_{1,i}Z_i + a_{0,i}))^2|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(x_i - \hat{X}_i(\mathbf{Z}))^2|\mathbf{Z}^{-i}\big]\Big] \qquad (3) \\
=&\frac{1}{n}\mathbb{E}\Big[x_i - \hat{X}_i(\mathbf{Z})\Big]^2.
\end{aligned}$$

2) For $d=2$ (polynomial mapping), (2) becomes

$$\begin{aligned}
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(Z_i - (a_{2,i}Z_i^2 + a_{1,i}Z_i + a_{0,i}))^2 \\
&\qquad + \sigma^2(\sum_{m=1}^{2} 2^m a_{m,i} Z_i^{m-1} - 1)\big|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(Z_i^2 - \sigma^2) + (a_{2,i}Z_i^2 + a_{1,i}Z_i + a_{0,i})^2 \\
&\qquad - 2a_{2,i}(Z_i^3 - 2Z_i\sigma^2) - 2a_{1,i}(Z_i^2 - \sigma^2) - 2a_{0,i}Z_i|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(x_i^2 + (a_{2,i}Z_i^2 + a_{1,i}Z_i + a_{0,i})^2 \\
&\qquad - 2a_{2,i}Z_i^2 x_i - 2a_{1,i}Z_i x_i - 2a_{0,i}x_i|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(x_i - (a_{2,i}Z_i^2 + a_{1,i}Z_i + a_{0,i}))^2|\mathbf{Z}^{-i}\big]\Big] \\
=&\frac{1}{n}\mathbb{E}\Big[\mathbb{E}\big[(x_i - \hat{X}_i(\mathbf{Z}))^2|\mathbf{Z}^{-i}\big]\Big] \qquad (4) \\
=&\frac{1}{n}\mathbb{E}\Big[x_i - \hat{X}_i(\mathbf{Z})\Big]^2.
\end{aligned}$$

Note (3) and (4) are from the specific form of $\hat{X}_i(\mathbf{Z})$, the fact $\{a_{m,i}\}$'s are independent of Z_i given $\mathbf{Z}^{-i}$, and

$$\begin{aligned}
\mathbb{E}(Z_i^3 - 2Z_i\sigma^2|\mathbf{Z}^{-i}) &= \mathbb{E}(x_i Z_i^2|\mathbf{Z}^{-i}) \\
\mathbb{E}(Z_i^2 - \sigma^2|\mathbf{Z}^{-i}) &= \mathbb{E}(x_i Z_i|\mathbf{Z}^{-i}) \\
&= \mathbb{E}(x_i^2|\mathbf{Z}^{-i}) \\
\mathbb{E}(Z_i|\mathbf{Z}^{-i}) &= \mathbb{E}(x_i|\mathbf{Z}^{-i}),
\end{aligned}$$

which hold due to the assumptions on $\mathbf{N}$ in Section 3.1 and $\mathbf{x}$ being an individual image. Thus, we obtain the Lemma by obtaining the unbiasedness for all $i = 1, \ldots, n$.

Furthermore, when $\mathbf{N}$ is i.i.d. Gaussian, then the SURE of $\mathbf{\Lambda}_n(\mathbf{x}, \hat{\mathbf{X}}(\mathbf{Z}))$ becomes

$$-\sigma^2 + \frac{1}{n}\|\mathbf{Z} - \hat{\mathbf{X}}(\mathbf{Z})\|_2^2 + \frac{2\sigma^2}{n}\sum_{i=1}^{n}\frac{\partial \hat{X}_i(\mathbf{Z})}{\partial Z_i}, \tag{5}$$

which is equivalent to [Manuscript, Eq.(3)] when $\hat{X}_i(\mathbf{Z}) = \sum_{m=0}^{d} a_m(\mathbf{Z}^{-i})Z_i^m$ with $d \in \{1,2\}$. ■

Table 1. PSNR(dB) on *Image13* and *BSD68*.

Data\Alg.	FC-AIDE$_{\text{S+FT}}$	FC-AIDE$_{\text{S+FT}}$(d=2, a_0-only)	FC-AIDE$_{\text{S}}$(d=0)	FC-AIDE$_{\text{S+FT}}$(d=0)
Image13	**29.33**	20.46	26.99	27.69
BSD68 (avg.)	**29.31**	19.20	27.66	27.68

2. The unbiasedness of $\mathbf{L}_n(\cdot)$

Here, we also experimentally verify the unbiasedness of $\mathbf{L}_n(\cdot)$ that is analytically shown in Lemma 1. Figure 1 shows the histograms of differences between MSE and [Manuscript, Eq.(3)] of the FC-AIDE$_{\text{B}}$ model, for 100 independent noise realizations ($\sigma = 25$) on two randomly selected images in BSD68. The mean of the difference clearly concentrates on 0 (*i.e.*, unbiased), and the standard deviation is also extremely small, for both images.

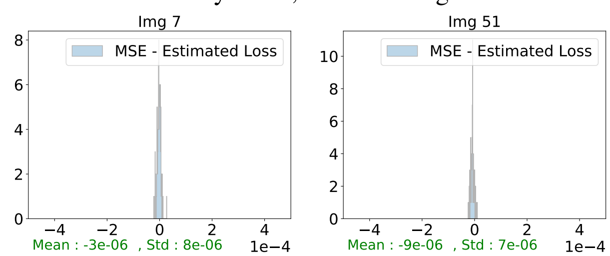

Figure 1. The difference between MSE and Eq.(3) for FC-AIDE$_{\text{B}}$.

3. Supplementary for Section 4.2

3.1. Ablation study for data augmentation

Figure 2 shows two additional results that use the *self-ensemble* data augmentation only for the training (Tr Aug) or testing (Te Aug) of the fine-tuning. Note "Te Aug" leads to more improvement, while the full augmentation, which is employed by our FC-AIDE$_{\text{S+FT}}$, leads to the highest PSNR and stable convergence.

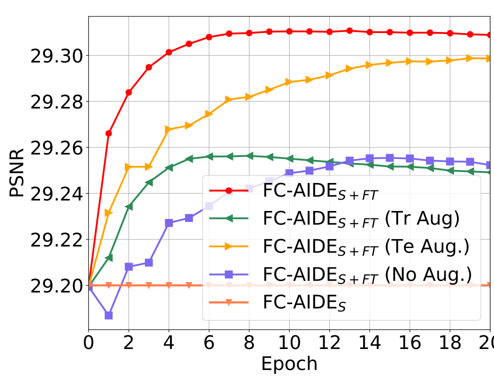

Figure 2. Ablation study on augmentations on BSD68 ($\sigma = 25$).

From Figure 2, we observe that the maximum PSNR of FC-AIDE$_{\text{S+FT}}$ is achieved around epoch 5, but even with a *single* epoch, the PSNR significantly improves over FC-AIDE$_{\text{S}}$. Each epoch takes about 3 seconds, and early stopping can lead to the accuracy-complexity trade-off. Moreover, the running time of each epoch for "No Aug" was 1.8 second, hence, the running time of the partial augmentation schemes lie in between 1.8s and 3s.

3.2. Hyperparameter selection for ℓ_2-SP

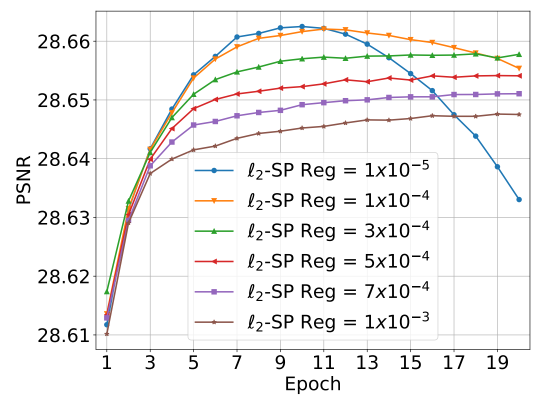

Figure 3. Fine-tuning result on the validation set ($\sigma = 25$)

Table 2. Selected regularization strength for ℓ_2-SP and the stopping epoch for adaptive fine-tuning for each σ.

	λ for ℓ_2-SP	Stopping epoch
$\sigma = 15$	1×10^{-4}	5
$\sigma = 25$	3×10^{-4}	4
$\sigma = 30$	5×10^{-4}	3
$\sigma = 50$	2×10^{-3}	2
$\sigma = 75$	5×10^{-3}	1

As mentioned in [Manuscript, Section 5.1], we used a separate validation set that consists of 32 natural images from BSD300 [1] for selecting the hyper-parameters in our fine-tuning step (*i.e.*, the stopping epoch and the regularization parameter for ℓ_2-SP). Note the validation images do not overlap with our training and test images. We carried out the validation for each noise level $\sigma = \{15, 25, 30, 50, 75\}$ separately and selected the best hyper-parameters that gave the best trade-off between the PSNR and the robustness of the curve. Figure 3 shows the results for $\sigma = 25$, for example. Note the PSNR results are not very different among the hyper-parameter choices, and the selection results for all noise levels are given in Table 2. These hyper-parameters were used for all our experiments in the paper.

4. Supplementary for Section 5.3

Figure 4 shows the PSNR differences between `FC-AIDE`$_{\texttt{S+FT}}$ and `FC-AIDE`$_{\texttt{S}}$ for each test image in BSD68 with $\sigma = 25$. Note the adaptive fine-tuning gives positive PSNR gains for *all* the images, and the four red bars indicate the images with the most PSNR improvements that are visualized in [Manuscript, Figure 4].

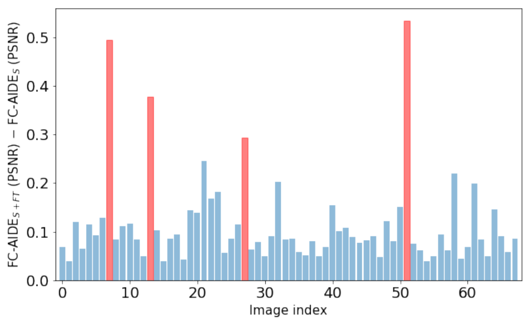

Figure 4. Improvement on BSD68

5. Supplementary for Section 5.4

Here, we emphasize the importance of the polynomial coefficients of `FC-AIDE`$_{\texttt{S+FT}}$ for denoising. In Table 1, we report the PSNRs on *Image13* (of BSD68) as well as the average PSNRs on the entire BSD68. Note the visualizations of the pixelwise coefficients are given in [Manuscript, Figure 7]. In the table, we compare `FC-AIDE`$_{\texttt{S+FT}}$ with several other baseline models; `FC-AIDE`$_{\texttt{S+FT}}$`(d=2,` a_0`-only)` is a scheme that denoises only with the a_0 terms after learning `FC-AIDE`$_{\texttt{S+FT}}$, `FC-AIDE`$_{\texttt{S}}$`(d=0)` is a supervised-trained model with setting $a_1 = a_2 = 0$, and `FC-AIDE`$_{\texttt{S+FT}}$`(d=0)` is the model obtained by fine-tuning `FC-AIDE`$_{\texttt{S}}$`(d=0)` using $\mathbf{L}_n(\cdot)$ (Manuscript, Eq.(3)).

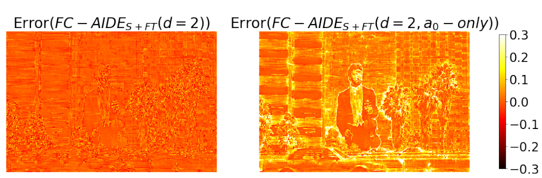

Figure 5. Pixelwise errors of `FC-AIDE`$_{\texttt{S+FT}}$ and `FC-AIDE`$_{\texttt{S+FT}}$`(d=2,` a_0`-only)`.

Note `FC-AIDE`$_{\texttt{S+FT}}$`(d=2,` a_0`-only)` and `FC-AIDE`$_{\texttt{S+FT}}$`(d=0)` are different schemes, and they are *not* equivalent to the regular end-to-end scheme, since they both do not use Z_i and are adaptively fine-tuned. From the table, we note that `FC-AIDE`$_{\texttt{S+FT}}$`(d=2,` a_0`-only)` hardly does any denoising (as the PSNR of the noisy *Image13* is 20.16dB), and `FC-AIDE`$_{\texttt{S+FT}}$`(d=0)` is also much worse than `FC-AIDE`$_{\texttt{S+FT}}$. Figure 5 shows the pixelwise errors on *Image13*, further demonstrating the importance of a_1 and a_2 in our polynomial model.

6. Visualization

Figure 10 and 11 show the clean images used for Set5 and Set12. Moreover, in Figures 6∼9, we visualized the denoising results on sample images from our evaluation datasets, *i.e.*, Set12, BSD68, Urban100, Manga109, *BSD68/Laplacian* and *Medical/Gaussian*. We compare our `FC-AIDE`$_{\texttt{S+FT}}$ with the most competitive state-of-the-art baselines and show the superiority of `FC-AIDE`$_{\texttt{S+FT}}$ both quantitatively and qualitatively.

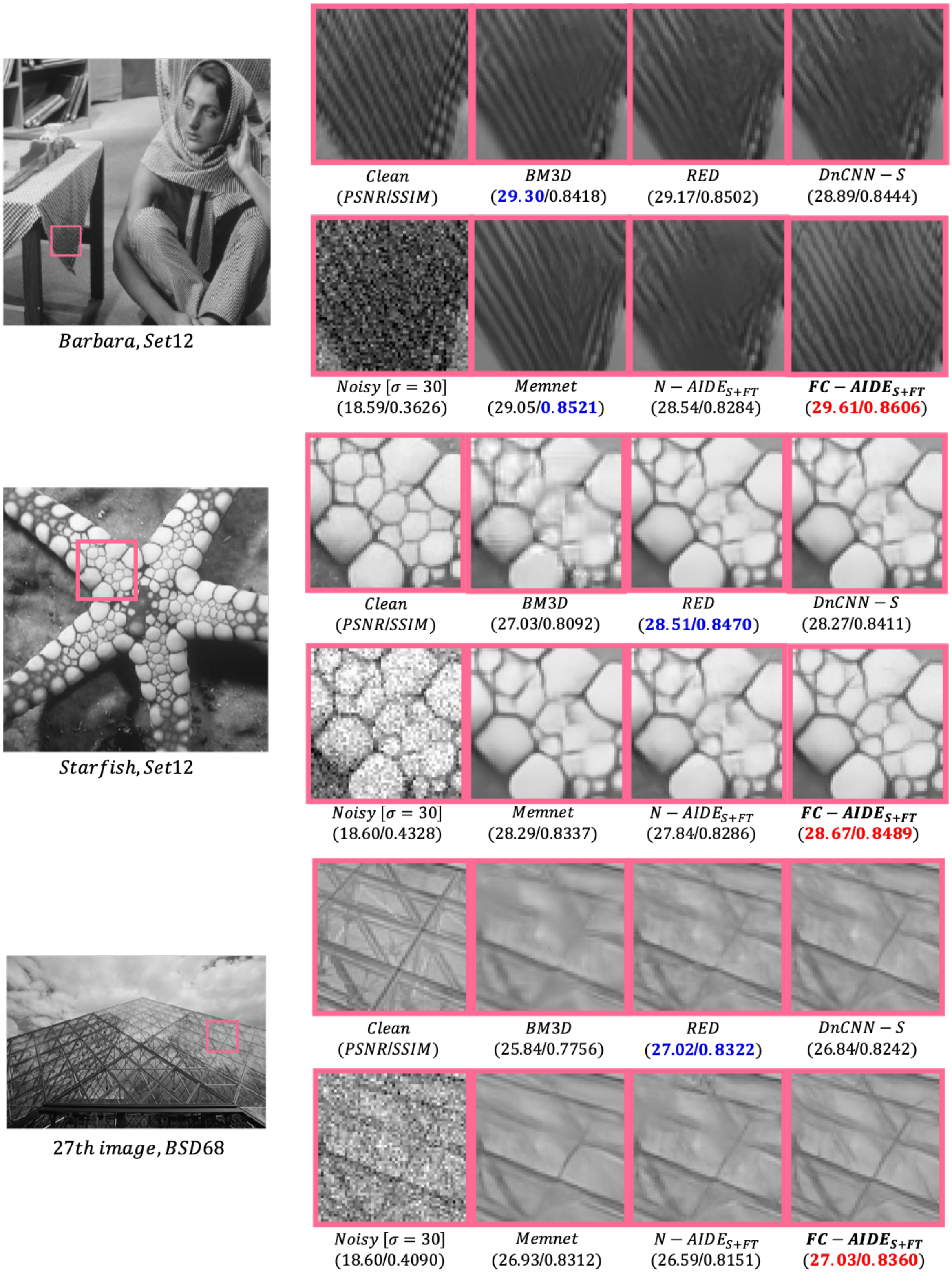

Figure 6. Denoising results on Set12 and BSD68

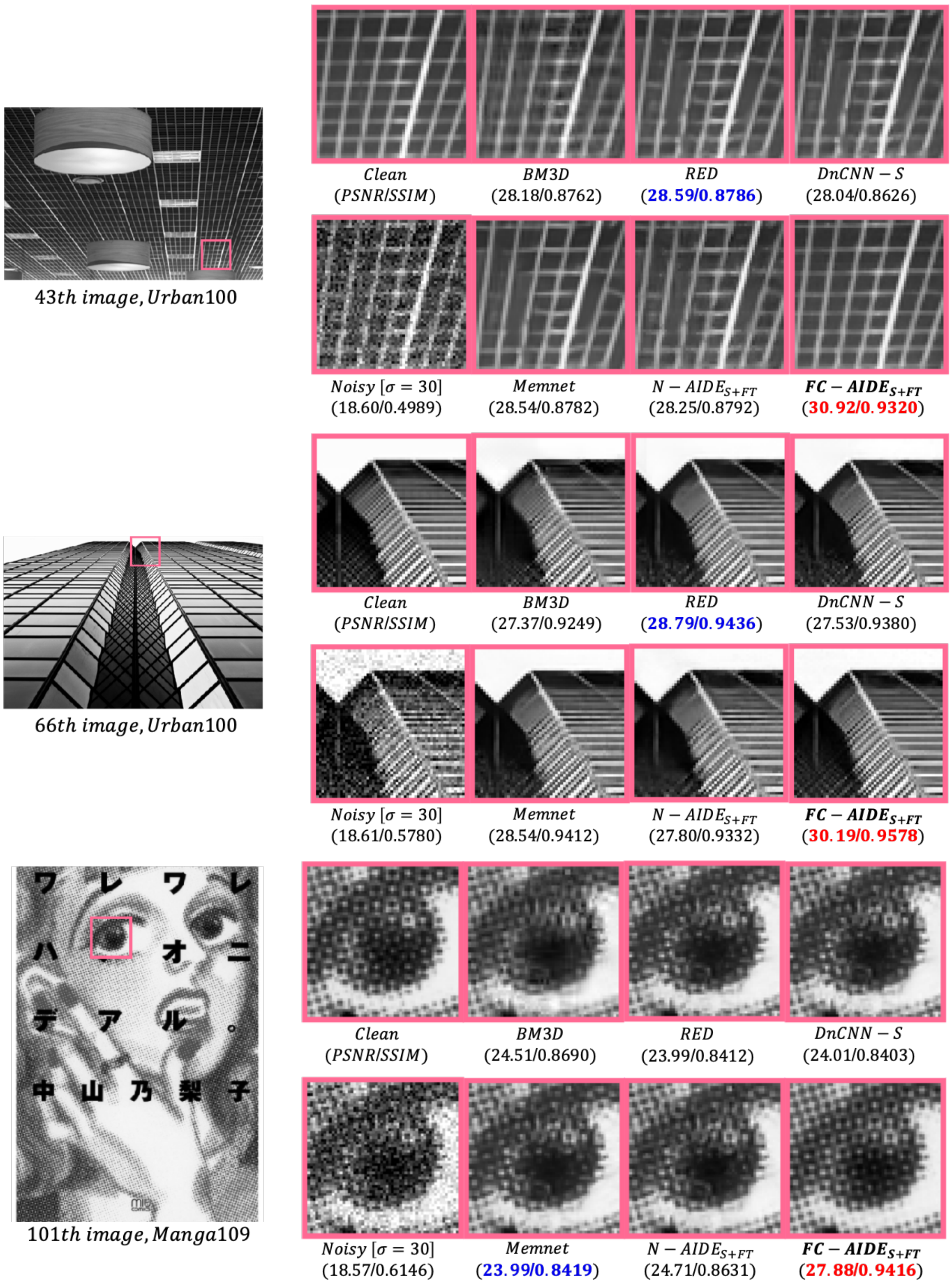

Figure 7. Denoising results on Urban100 and Manga109

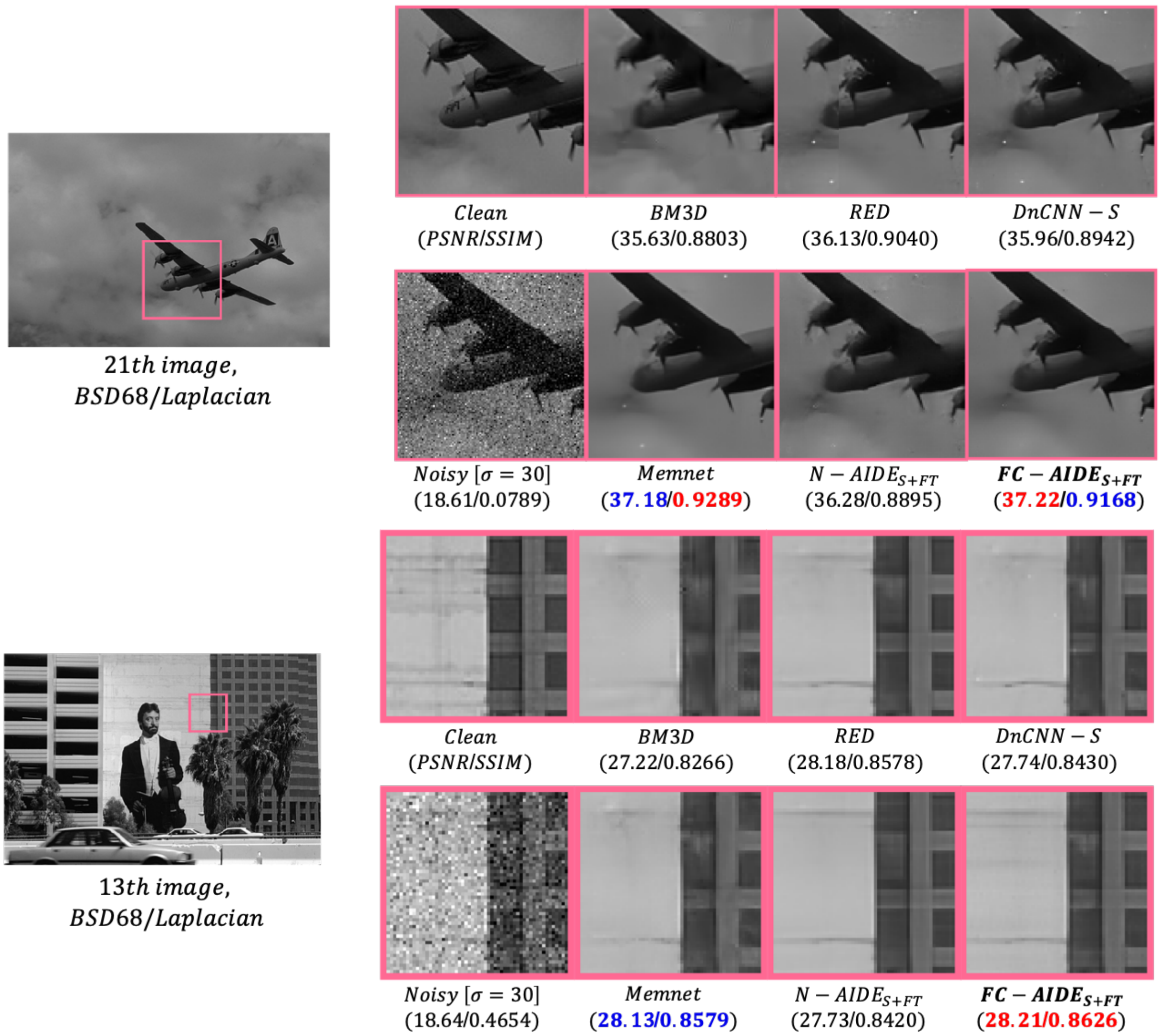

Figure 8. Denoising results on *BSD68/Laplacian*

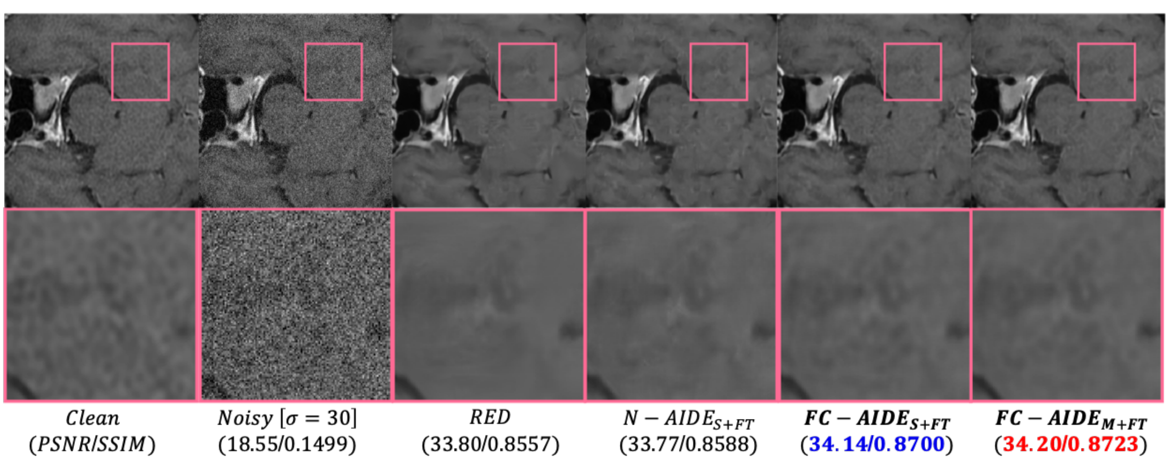

Figure 9. Denoising results on *Medical/Gaussian*

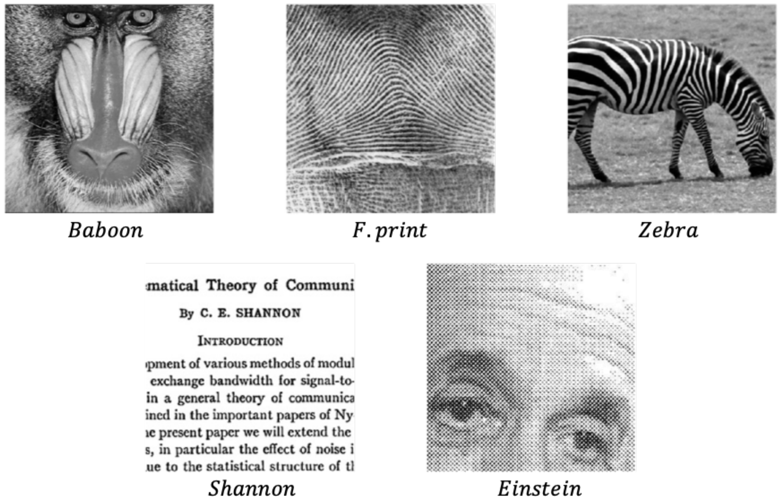

Figure 10. Visualization on Set5.

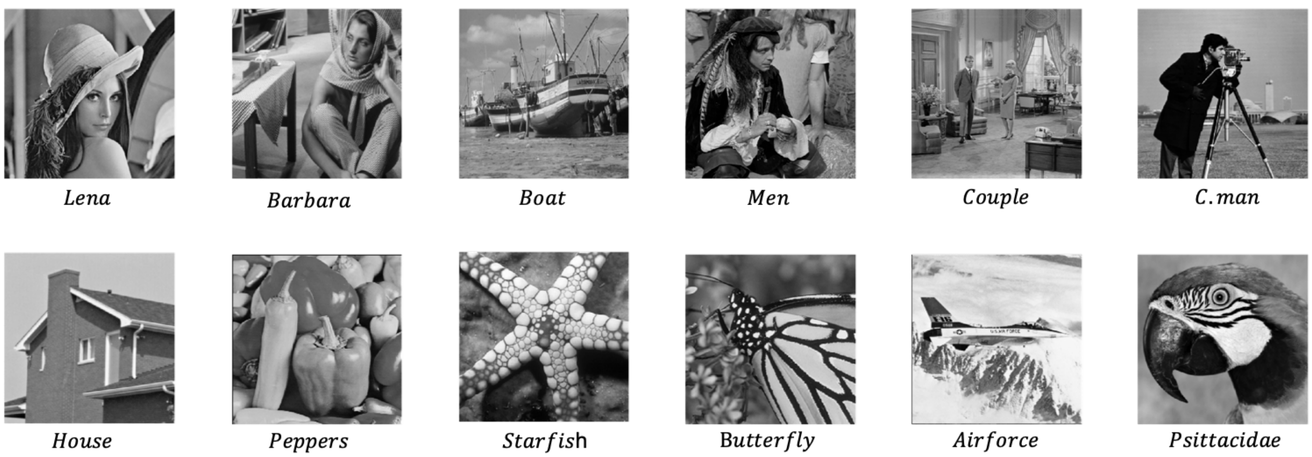

Figure 11. Visualization on Set12.

References

[1] David Martin, Charless Fowlkes, Doron Tal, Jitendra Malik, et al. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. 2001.

[2] Charles M Stein. Estimation of the mean of a multivariate normal distribution. *The annals of Statistics*, pages 1135–1151, 1981.