Supplementary Material for
Fully Convolutional Pixel Adaptive Image Denoiser

Sungmin Cha¹ and Taesup Moon¹,²
¹Department of Electrical and Computer Engineering, ²Department of Artificial Intelligence
Sungkyunkwan University, Suwon, Korea 16419
{csm9493,tsmoon}@skku.edu

1. Proof of Lemma 1

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

\[ \mathbb{E}_{N}(Z, X(Z); \sigma^2) = \mathbb{E}_{N}(x, \hat{X}(Z)). \] (1)

Moreover, when N is white Gaussian, then, \( \mathbb{E}_{N}(Z, X(Z); \sigma^2) \) coincides with the SURE (2).

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

\[ \frac{1}{n} \mathbb{E} \left[ (Z_i - \hat{X}_i(Z))^2 \right] + \sigma^2 \left( \sum_{m=1}^{d} 2^m a_{m,i} z_i^{m-1} - 1 \right) |Z^{-i}| \]

for \( i = 1, \ldots, n \).

For \( d = 1 \) (affine mapping), (3) becomes

\[ \mathbb{E} \left[ (Z_i - \hat{X}_i(Z))^2 \right] + \sigma^2 (2^m a_{m,i} z_i^{m-1} - 1) |Z^{-i}| \]

for \( i = 1, \ldots, n \) with the assumptions in [Manuscript, Eq.(2)].

Thus, we obtain the Lemma by obtaining the unbiasedness for all \( i = 1, \ldots, n \).

Furthermore, when N is i.i.d. Gaussian, then the SURE of \( \mathbb{E}_{N}(x, X(Z)) \) becomes

\[ -\sigma^2 + \frac{1}{n} \mathbb{E} \left[ \mathbb{E} \right] + \frac{2\sigma^2}{n} \sum_{i=1}^{d} \frac{\partial \hat{X}_i(Z)}{\partial Z_i}, \]

which is equivalent to [Manuscript, Eq.(3)] when \( \hat{X}_i(Z) = \sum_{m=0}^{d} a_m(Z^{-i}) z_i^m \) with \( d \in \{1, 2\} \).
2. The unbiasedness of $L_n(\cdot)$

Here, we also experimentally verify the unbiasedness of $L_n(\cdot)$ that is analytically shown in Lemma 1. Figure 1 shows the histograms of differences between MSE and [Manuscript, Eq.(3)] of the FC-AIDE$_S$ 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$_S$.](image)

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$_{S+FT}$, leads to the highest PSNR and stable convergence.

![Figure 2. Ablation study on augmentations on BSD68 ($\sigma = 25$).](image)

From Figure 2, we observe that the maximum PSNR of FC-AIDE$_{S+FT}$ is achieved around epoch 5, but even with a single epoch, the PSNR significantly improves over FC-AIDE$_S$. 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 $\ell_2$-SP

Table 2 shows the selected regularization strength for $\ell_2$-SP and the stopping epoch for adaptive fine-tuning for each $\sigma$.

<table>
<thead>
<tr>
<th>$\sigma$</th>
<th>$\lambda$ for $\ell_2$-SP</th>
<th>Stopping epoch</th>
</tr>
</thead>
<tbody>
<tr>
<td>15</td>
<td>$1 \times 10^{-4}$</td>
<td>5</td>
</tr>
<tr>
<td>25</td>
<td>$3 \times 10^{-4}$</td>
<td>4</td>
</tr>
<tr>
<td>30</td>
<td>$5 \times 10^{-4}$</td>
<td>3</td>
</tr>
<tr>
<td>50</td>
<td>$2 \times 10^{-3}$</td>
<td>2</td>
</tr>
<tr>
<td>75</td>
<td>$5 \times 10^{-3}$</td>
<td>1</td>
</tr>
</tbody>
</table>

As mentioned in [Manuscript, Section 5.1], we used a separate validation set that consists of 32 natural images from BSD300 for selecting the hyper-parameters in our fine-tuning step (i.e., the stopping epoch and the regularization parameter for $\ell_2$-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.

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

Table 1. PSNR(db) on Image13 and BSD68.

<table>
<thead>
<tr>
<th>Data\Alg.</th>
<th>FC-AIDE$_S$+FT</th>
<th>FC-AIDE$_S$+FT ($d=2$, $a_0$-only)</th>
<th>FC-AIDE$_S$ ($d=0$)</th>
<th>FC-AIDE$_S$+FT ($d=0$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Image13</td>
<td>29.33</td>
<td>20.46</td>
<td>26.99</td>
<td>27.69</td>
</tr>
<tr>
<td>BSD68 (avg.)</td>
<td>29.31</td>
<td>19.20</td>
<td>27.66</td>
<td>27.68</td>
</tr>
</tbody>
</table>
4. Supplementary for Section 5.3

Figure 4 shows the PSNR differences between FC-AIDE<sub>S+FT</sub> and FC-AIDE<sub>S</sub> for each test image in BSD68 with σ = 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<sub>S+FT</sub> for denoising. In Table 5, 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<sub>S+FT</sub> with several other baseline models; FC-AIDE<sub>S+FT</sub> (d=2, a<sub>0</sub>-only) is a scheme that denoises only with the a<sub>0</sub> terms after learning FC-AIDE<sub>S+FT</sub>. FC-AIDE<sub>S</sub> (d=0) is a supervised-trained model with setting a<sub>1</sub> = a<sub>2</sub> = 0, and FC-AIDE<sub>S+FT</sub> (d=0) is the model obtained by fine-tuning FC-AIDE<sub>S</sub> (d=0) using L<sub>α</sub>(·) (Manuscript, Eq.(3)).

**Figure 5. Pixelwise errors of FC-AIDE<sub>S+FT</sub> and FC-AIDE<sub>S+FT</sub> (d=2, a<sub>0</sub>-only)**. Note FC-AIDE<sub>S+FT</sub> (d=2, a<sub>0</sub>-only) and FC-AIDE<sub>S+FT</sub> (d=0) are different schemes, and they are **not** equivalent to the regular end-to-end scheme, since they both do not use Z<sub>i</sub> and are adaptively fine-tuned. From the table, we note that FC-AIDE<sub>S+FT</sub> (d=2, a<sub>0</sub>-only) hardly does any denoising (as the PSNR of the noisy Image13 is 20.16dB), and FC-AIDE<sub>S+FT</sub> (d=0) is also much worse than FC-AIDE<sub>S+FT</sub>. Figure 5 shows the pixelwise errors on Image13, further demonstrating the importance of a<sub>1</sub> and a<sub>2</sub> 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<sub>S+FT</sub> with the most competitive state-of-the-art baselines and show the superiority of FC-AIDE<sub>S+FT</sub> 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
References
