Supplementary Material of NTGAN: Learning Blind Image Denoising without Clean Reference

Rui Zhao
rick10.zhao@connect.polyu.hk
Daniel P.K. Lun
enpklun@polyu.edu.hk
Kin-Man Lam
enkmlam@polyu.edu.hk

Department of Electronic and Information Engineering,
The Hong Kong Polytechnic University,
Hong Kong, China

1 Statistics consistency analysis on the noise

In the paper, we adopt the noise generation model in [8] to synthesize the target observation $\mathbf{z}$ from a real-world image $\mathbf{y}$ as follows:

$$\begin{aligned} \mathbf{z} &= \mathbf{y} + \mathbf{n} \\ \mathbf{n}(\mathbf{y}, \mathbf{M_z}) &= f_{\mathrm{BPD}}(f_{\mathrm{crf}}(\mathbf{L} + \mathbf{n}_s(\mathbf{L}) + \mathbf{n}_c)) - f_{\mathrm{BPD}}(f_{\mathrm{crf}}(\mathbf{L})), \\ \text{with } \mathbf{L} &= f_{\mathrm{icrf}}(\mathbf{y}), \end{aligned} \tag{1}$$

where f_{crf} and f_{icrf} represent the camera response function and the inverse camera response function, respectively. $\mathbf{n}_s$ and $\mathbf{n}_c$ account for the noise components that are dependent and independent of the signal $\mathbf{y}$, respectively. We assume that the noise residing in the real image $\mathbf{y}$ is heterogeneous Gaussian distributed, with the dependency on the clean signal $\mathbf{x}$. Therefore, in the irradiance plane, $\mathbf{L_y}$ can be represented as follows:

$$\begin{aligned} \mathbf{L_y} &= \mathbf{L_x} + \mathbf{n}_s(\mathbf{L_x}) + \mathbf{n}_c, \\ \text{with } \mathbf{L_y} &= f_{\mathrm{icrf}}(\mathbf{y}) \text{ and } \mathbf{L_x} = f_{\mathrm{icrf}}(\mathbf{x}). \end{aligned} \tag{2}$$

In addition, the synthetic observation $\mathbf{z}$ in the irradiance plane can be approximated as:

$$\begin{aligned} \mathbf{L_z} &= \mathbf{L_y} + \mathbf{n}_s(\mathbf{L_y}) + \mathbf{n}_c \approx \mathbf{L_x} + \mathbf{n}_{s'}(\mathbf{L_x}) + \mathbf{n}_{c'}, \\ \text{with } \mathbf{L_z} &= f_{\mathrm{icrf}}(\mathbf{z}), \end{aligned} \tag{3}$$

where $\mathbf{n}_{s'}$ denotes the signal-dependent component in $\mathbf{z}$ with respect to the clean signal, and $\mathbf{n}_{c'}$ denotes the signal-independent component, combining the signal-independent noise from Eqns.(1) and (2). Eqn. (3) indicates that the resultant noise in $\mathbf{z}$ also follows a heterogeneous Gaussian distribution. Therefore, the adopted noise generation model guarantees the consistency of the noise distribution in the source and target observations.

2 Evaluation on synthetic AWGN

To further evaluate the denoising performance of NTGAN, we apply it to the synthetic additive white Gaussian noise (AWGN), and compare its performance with the other state-of-the-art denoisers, including BM3D [3], WNNM [4], DnCNN [15], FFDNet [16], RIDNet [1], N2N [7], DIP [13], N2S [2], and N2V [6]. To make a fair comparison, we change the original noise generation model in the paper, and make it to generate AWGN for producing paired samples in training. Specifically, the training pairs are generated as:

$$\begin{aligned} \mathbf{y} = \mathbf{x} + \mathbf{n_y}, \ \ \mathbf{z} = \mathbf{x} + \mathbf{n_z}, \\ \text{with } \mathbf{n_y} \sim \mathcal{N}(0, \sigma_\mathbf{y}), \ \ \mathbf{n_z} \sim \mathcal{N}(0, \sigma_\mathbf{z}), \end{aligned} \tag{4}$$

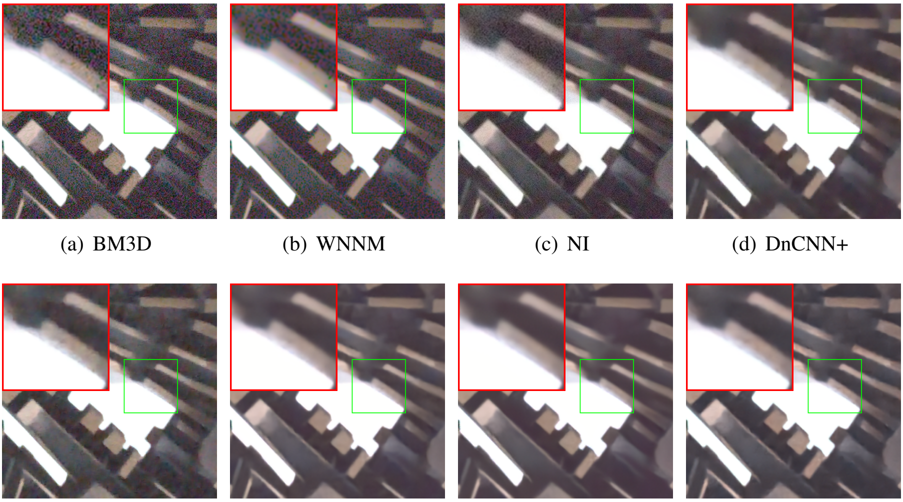
where both $\sigma_\mathbf{y}$ and $\sigma_\mathbf{z}$ are uniformly sampled from the range $(0, 75/255]$, and $\mathbf{x}$ represents the clean signal. In other words, the training pairs are obtained by independently adding two small AWGNs to the clean image. We collect all the 4,744 images from the Waterloo Exploration database [9] to form the training set, and use Eqn. (4) to generate the training pairs. We follow the same settings mentioned in the paper to train up NTGAN and evaluate it on BSD68 [10] with the noise level set to 15, 25, and 50, respectively. The results are summarized in Table 1. It can be observed that NTGAN achieves comparable or even better results, compared to the supervised non-blind denoisers, i.e. DnCNN and FFDNet. However, NTGAN performs slightly worse than RIDNet when facing AWGN.The reason is that the distribution of AWGN is much easier to be learned, and thus the proposed noise transference strategy loses its advantage by serving as an augmentation approach for learning complicated noise distributions. NTGAN achieves about 31.72dB, 29.28dB, and 26.37dB with the noise level set to 15, 25, and 50, respectively, which basically outperforms the other unsupervised deep denoisers by a large margin.

Type	Traditional methods		Supervised CNNs			Unsupervised CNNs				
Method	BM3D	WNNM	DnCNN	FFDNet	RIDNet	N2N	DIP	N2S	N2V	Ours
$\sigma = 15$										
PSNR	31.08	31.32	31.73	31.63	**31.81**	**31.81**	27.07	29.23	29.75	31.72
$\sigma = 25$										
PSNR	28.57	28.83	29.23	29.23	**29.34**	28.67	24.63	27.39	27.76	29.28
$\sigma = 50$										
PSNR	25.62	25.83	26.23	26.29	**26.40**	26.07	22.06	25.17	25.08	26.37

Table 1: The quantitative results on the grayscale BSD68 images corrupted with AWGN. The best results are highlighted in **bold**.

3 Qualitative comparison on DND

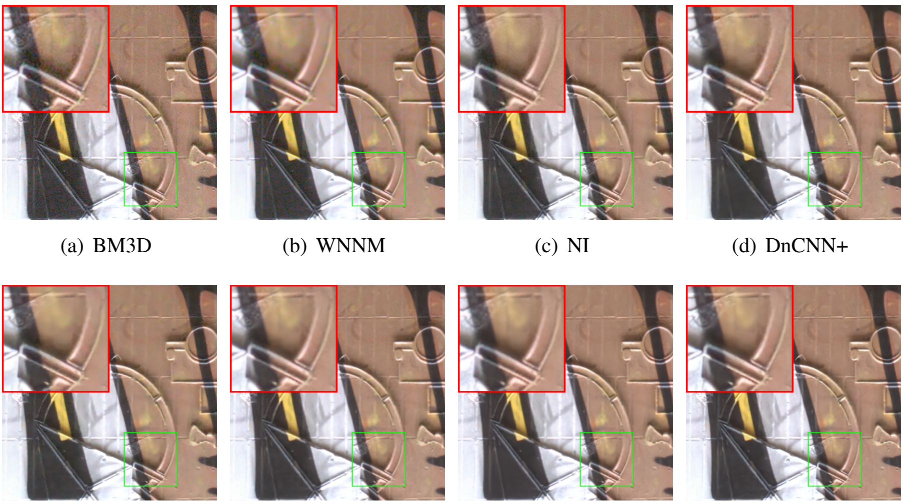
Here, we present the qualitative results on DND [12] in Figs. 1 and 2 to show the perceptual quality of the restored images based on NTGAN. It is clear from the figures that NTGAN acquires a better ability in recovering the texture and edge details. The traditional methods, e.g., BM3D [3], WNNM [4], and NI [11], still retain some noise in the restored images. The supervised denoisers, e.g., DnCNN+ [15], RIDNet [1], and VDNet [14], oversmooth the images. CBDNet [5] produces the images with visible artefacts on the edges. The proposed NTGAN can recover the image with fewer artefacts and more details.



(a) BM3D (b) WNNM (c) NI (d) DnCNN+

(e) CBDNet (f) RIDNet (g) VDNet (h) NTGAN

Figure 1: Evaluation on the perceptual qualities of a real-world noisy image from DND, restored by different methods. BM3D: 23.95dB, WNNM: 25.63dB, NI: 27.28dB, DnCNN+: 32.26dB, CBDNet: 31.40dB, RIDNet: 34.30dB, VDNet: 34.08dB, NTGAN: 35.10dB.



(a) BM3D (b) WNNM (c) NI (d) DnCNN+

(e) CBDNet (f) RIDNet (g) VDNet (h) NTGAN

Figure 2: Evaluation on the perceptual qualities of a real-world noisy image from DND, restored by different methods. BM3D: 30.91dB, WNNM: 30.94dB, NI: 32.23dB, DnCNN+: 33.29dB, CBDNet: 33.62dB, RIDNet: 34.09dB, VDNet: 33.89dB, NTGAN: 34.37dB.

References

[1] Saeed Anwar and Nick Barnes. Real image denoising with feature attention. In *The IEEE International Conference on Computer Vision (ICCV)*, October 2019.

[2] Joshua Batson and Loïc Royer. Noise2self: Blind denoising by self-supervision. *arXiv*, 2019.

[3] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen O. Egiazarian. Image denoising by sparse 3-d transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16:2080–2095, 2007.

[4] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *CVPR*, pages 2862–2869, 2014.

[5] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *CVPR*, pages 1712–1722, June 2019.

[6] Alexander Krull, Tim-Oliver Buchholz, and Florian Jug. Noise2void - learning denoising from single noisy images. In *CVPR*, June 2019.

[7] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In *ICML*, volume 80, pages 2965–2974, July 2018.

[8] Ce Liu, William T. Freeman, Richard Szeliski, and Sing Bing Kang. Noise estimation from a single image. In *CVPR*, pages 901–908, 2006.

[9] Kede Ma, Zhengfang Duanmu, Qingbo Wu, Zhou Wang, Hongwei Yong, Hongliang Li, and Lei Zhang. Waterloo Exploration Database: New challenges for image quality assessment models. *IEEE Transactions on Image Processing*, 26(2):1004–1016, Feb. 2017.

[10] 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 *ICCV*, volume 2, pages 416–423, July 2001.

[11] Neatlab and ABSoft. Neat image. *http://ni.neatvideo.com/home*, 2001.

[12] Tobias Plotz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *CVPR*, pages 2750–2759, 2017.

[13] Dmitry Ulyanov, Andrea Vedaldi, and Victor S. Lempitsky. Deep image prior. In *CVPR*, pages 9446–9454, 2017.

[14] Zongsheng Yue, Hongwei Yong, Qian Zhao, Deyu Meng, and Lei Zhang. Variational denoising network: Toward blind noise modeling and removal. In *NIPS*, pages 1690–1701, 2019.

[15] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. *IEEE Transactions on Image Processing*, 26:3142–3155, 2017.

[16] Kai Zhang, Wangmeng Zuo, and Lei Zhang. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. *IEEE Transactions on Image Processing*, 27:4608–4622, 2017.