# Unpaired Learning of Deep Image Denoising

Xiaohe Wu[1], Ming Liu[1], Yue Cao[1], Dongwei Ren[2], and Wangmeng Zuo[1,3] (✉)

[1] Harbin Institute of Technology, China
[2] University of Tianjin, China
[3] Peng Cheng Lab, China
csxhwu@gmail.com, csmliu@outlook.com,
{cscaoyue,rendongweihit}@gmail.com, wmzuo@hit.edu.com

## 1 Description for the blind-spot mechanisms.

Here, we provide an example to explain the blind-spot mechanism illustrated in Fig. (3)a in detail. Taking the $7 \times 7$ input as an example, the output feature map of each layer has the same size, due to the fully-convolutional net with paddings and stride one. For simplicity, we adopt the $3 \times 3$ kernel and ignore the number of channels. Denote the input as $\mathbf{y}$, each pixel can be represented as $\mathbf{y}_{i,j}, i \in \{0, 1, 2, 3\}, j \in \{0, 1, 2, 3\}$, where $\mathbf{y}_{0,0}$ is the center position value, called the blind-spot.

1. First, we operate the centrally masked convolution on y according to Eq. (5), leading to the feature map $\mathbf{f}^1$. Obviously, values including $\mathbf{f}^1_{0,0}$ and $\mathbf{f}^1_{i,j}, i \in \{2, 3\}, j \in \{2, 3\}$ are irrelevant to $\mathbf{y}_{0,0}$. On the contrary, $\mathbf{f}^1_{i,j}, i \in \{1\}, j \in \{1\}$ are computed using $\mathbf{y}_{0,0}$. The blind-spot requirement requires to avoid using $\mathbf{f}^1_{i,j}, i \in \{1\}, j \in \{1\}$ when computing the value of center position in the next feature map, which motivates us to adopt the dilated convolution.
2. With a dilation rate of 2, we get feature map $\mathbf{f}^2$. In particular, its center position value $\mathbf{f}^2_{0,0}$ is obtained from $\mathbf{f}^1_{i,j}, i \in \{0, 2\}, j \in \{0, 2\}$. From (1), these values are not affected by $\mathbf{y}_{0,0}$. Moreover, we also find that $\mathbf{f}^2_{i,j}, i \in \{0, 2\}, j \in \{0, 2\}$ are not relevant to $\mathbf{y}_{0,0}$, which further inspires us to adopt the dilated convolution in the following layers.
3. With such principle, we investigate the blind-spot mechanism as shown in Fig. (3)a.

## 2 Additional Visualization Results

More denoising results for AWGN, heteroscedastic Gaussian (HG) and multivariate Gaussian (MG) noise are provided for comprehensive comparison. Specifically, we compare the proposed D-BSN and MWCNN(unpaired) with the benchmark method CBM3D [1], the supervised MWCNN in a Noise2Clean training manner. In particular, we also consider the Laine19 [3] on AWGN with $\sigma = 25$ by using their released model. As for the real noisy images, we provide more visualization results from CC15 [6], DND [7], RNI6 [4] and RNI15 [4] datasets to compare with the state-of-the-arts. For better view, we recommend to zoom in the images on a computer screen.

Denoising results for AWGN are presented in Fig. 1∼ Fig. 4. It can be noted that, our D-BSN achieves comparable visualization performance with the self-supervised approach Laine19 [3]. Particularly, by leveraging the synthetic paired training data from D-BSN, our MWCNN(unpaired) can preserve more texture information in comparison with all the other methods. For one hand, the nearly noisy-free images in the training set $\mathcal{Y}$ are predicted by D-BSN, which guarantees the promising denoising results of MWCNN(unpaired). On the other hand, the clean image set $\mathcal{X}$ contains the truly clean signal, which is beneficial to learn denoising network with fine details. In addition, we provide the denoising results of PG and MG in Fig. 5 and Fig. 6, respectively. Significantly, both our D-BSN and MWCNN(unpaired) outperform the benchmark method CBM3D [1]. With heavy degradation as shown in Fig. 5 and Fig. 6, our MWCNN(unpaired) can restore the images well and achieves competitive performance with the supervised MWCNN(N2C) variant. For the real noisy images in Fig. 7∼Fig. 14, MWCNN(unpaired) performs favorably against the benchmark method CBM3D [1] and the discriminative learning method DnCNN [9]. Even compared with the CBDNet [2] and VDN [8], our method shows comparable visualization results without the consideration of the details of ISP and paired noisy-clean images.


Fig. 1: Denoising results of different methods on AWGN with $\sigma = 25$.

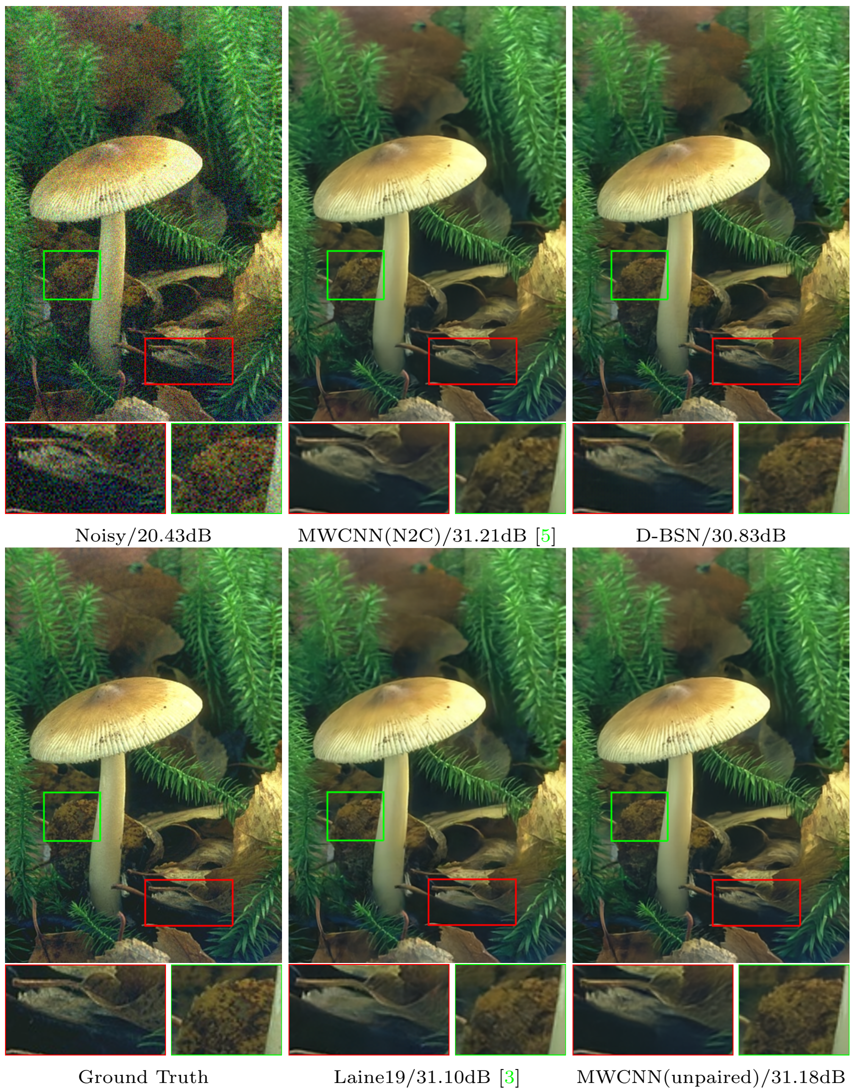


Fig. 2: Denoising results of different methods on AWGN with $\sigma = 25$.

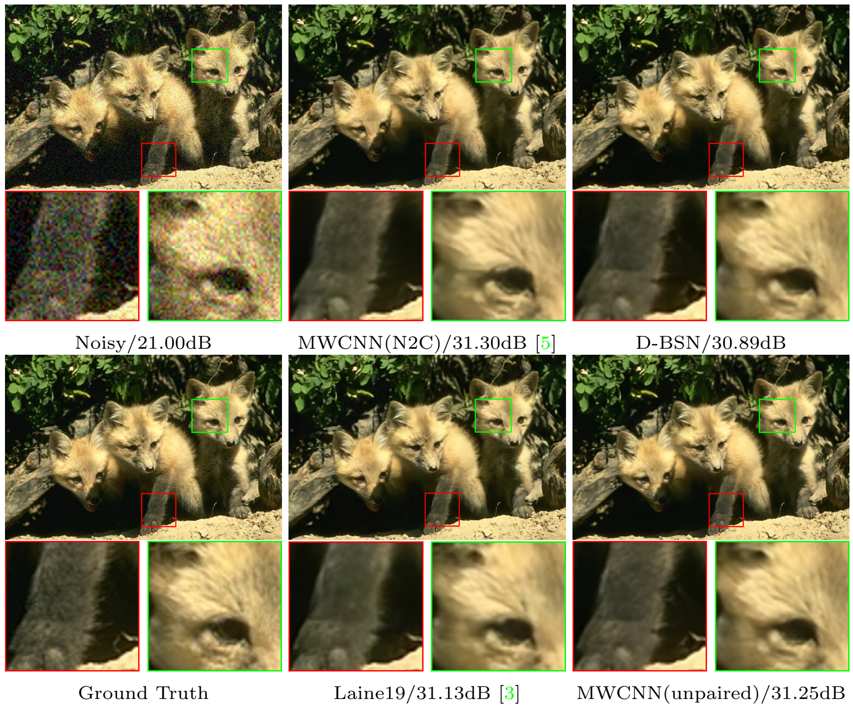


Fig. 3: Denoising results of different methods on AWGN with $\sigma = 25$.

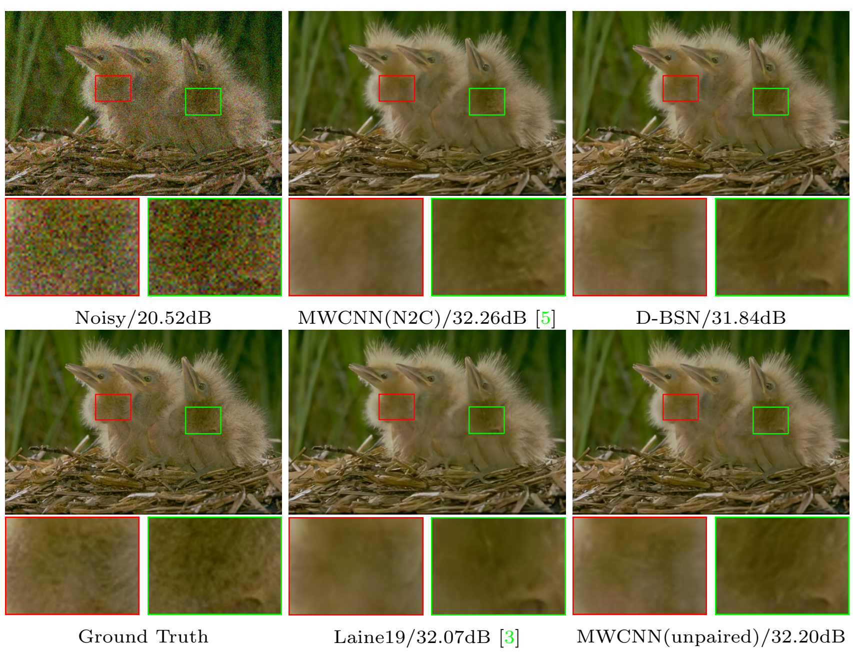


Fig. 4: Denoising results of different methods on AWGN with $\sigma = 25$.

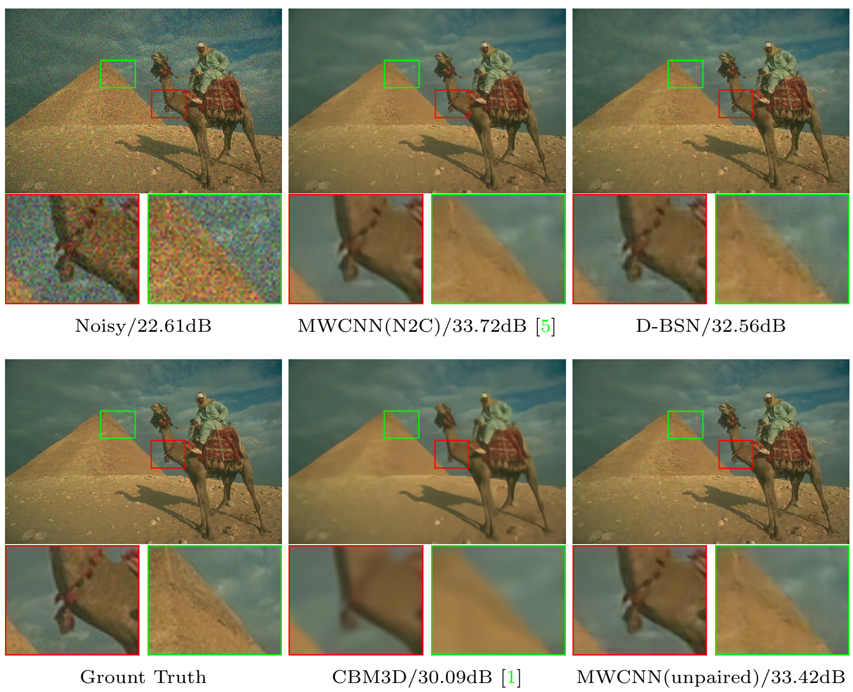


Fig. 5: Denoising results of different methods on heteroscedastic Gaussian (HG) noise.

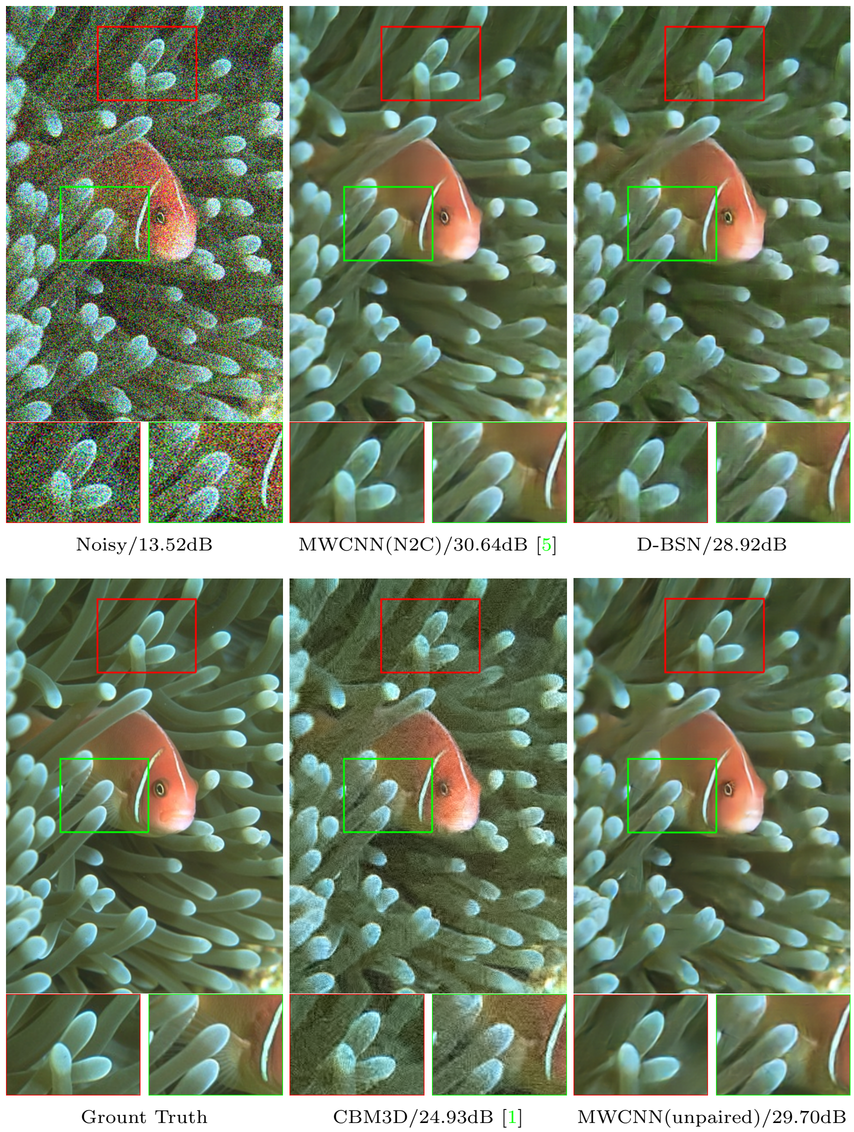


Fig. 6: Denoising results of different methods on multivariate Gaussian (MG) noise.

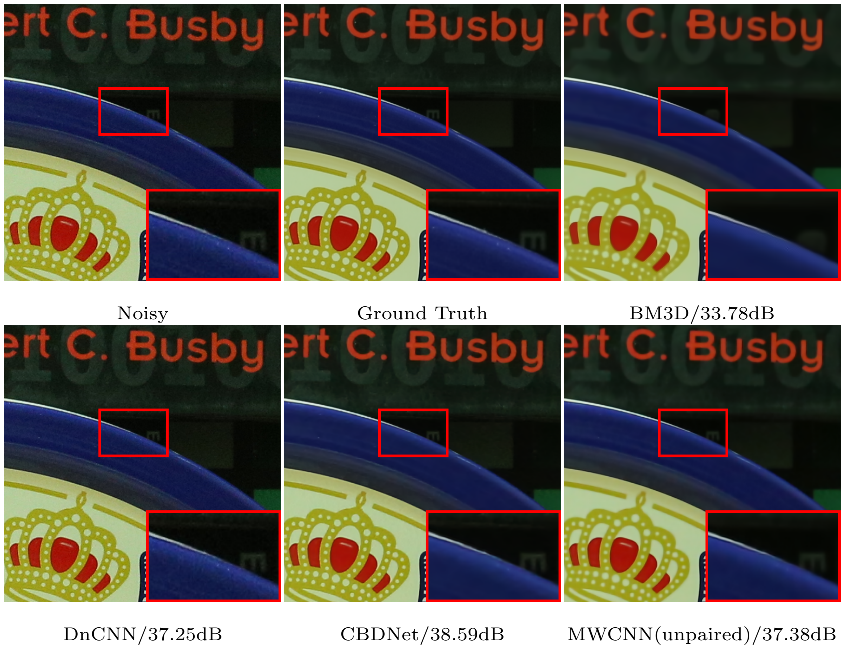


Fig. 7: Denoising results of different methods on real-world images from CC15.

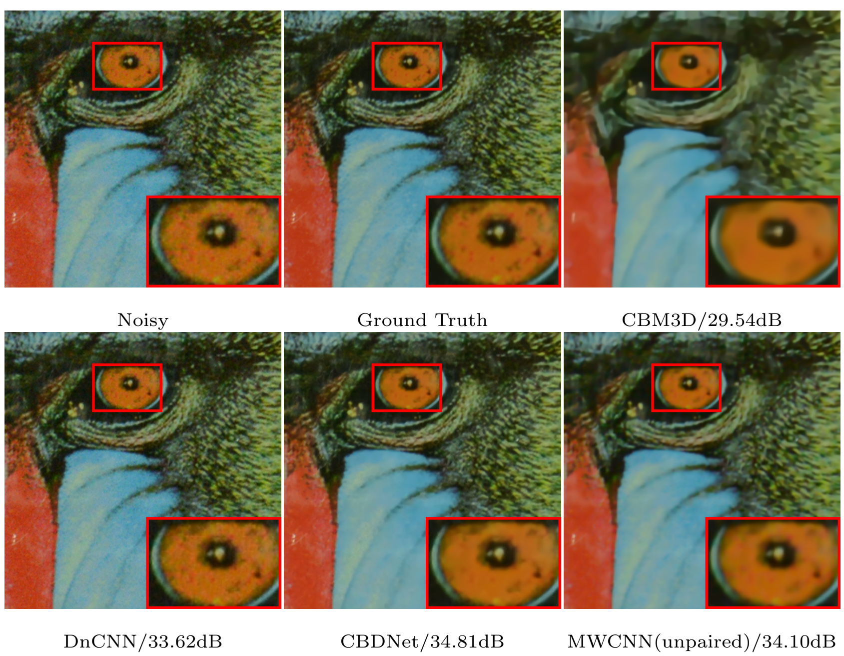


Fig. 8: Denoising results of different methods on real-world images from CC15.

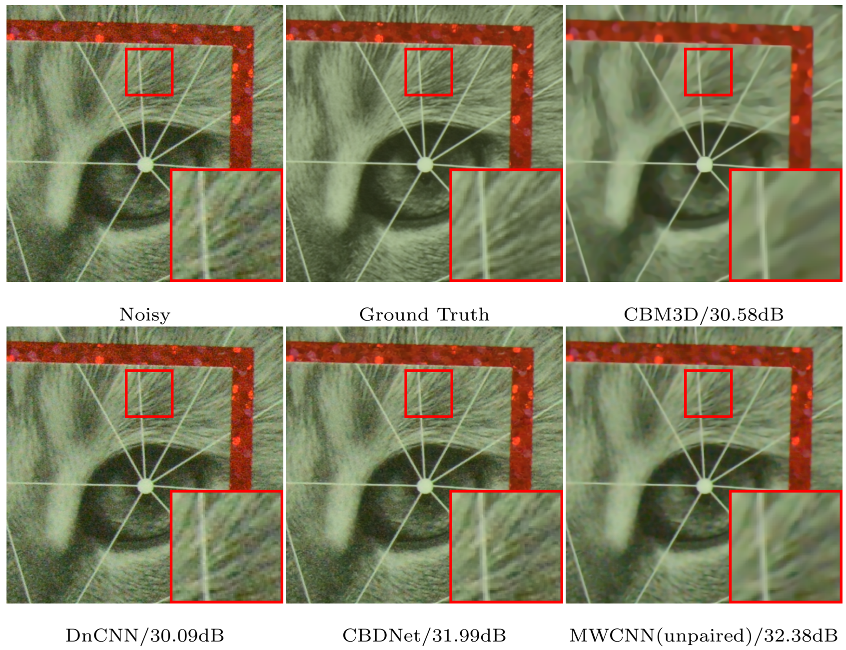


Fig. 9: Denoising results of different methods on real-world images from CC15.

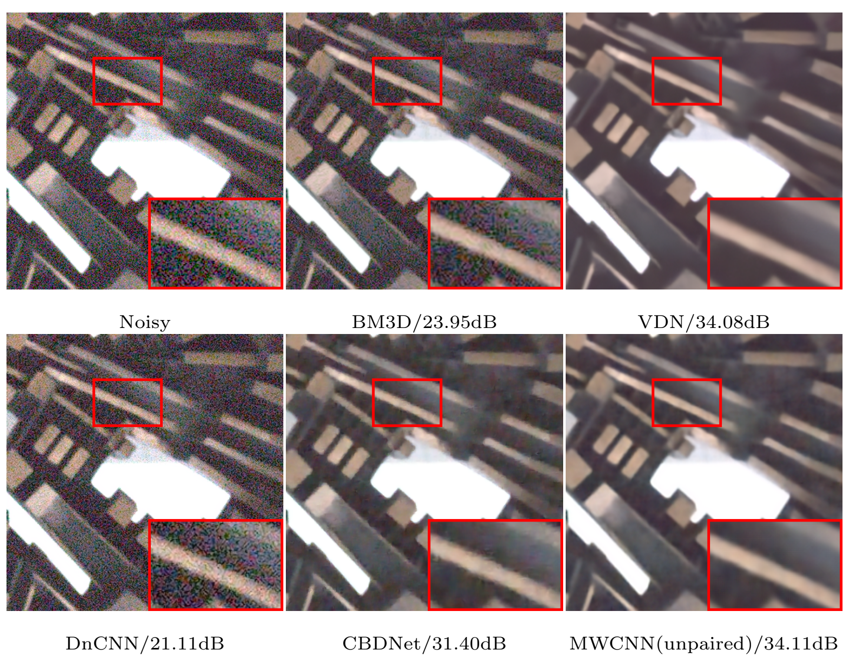


Fig. 10: Denoising results of different methods on real-world images from DND.

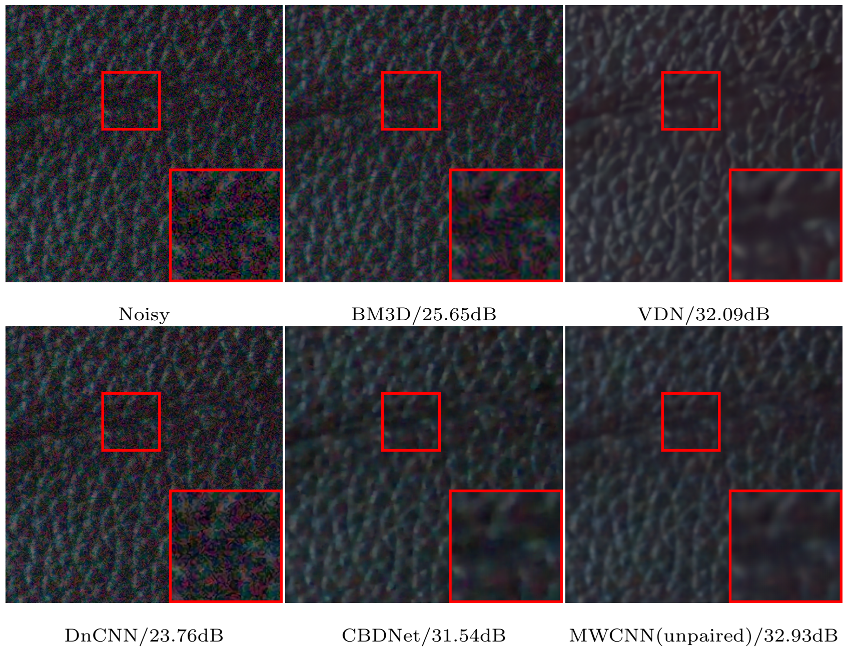


Fig. 11: Denoising results of different methods on real-world images from DND.


Fig. 12: Denoising results of different methods on real-world images from DND.


Fig. 13: Denoising results of different methods on real noisy images from RNI6. From top to bottom: noisy images, denoised images by BM3D [1], denoised images by DnCNN-B [9], denoised images by our MWCNN(unpaired).


Fig. 14: Denoising results of different methods on real noisy images from RNI15. From top to bottom: noisy images, denoised images by CBM3D [1], denoised images by CBDNet [2], denoised images by our MWCNN(unpaired).

## References

1. Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K.: Image denoising by sparse 3-d transform-domain collaborative filtering. TIP **16**(8), 2080–2095 (2007) 1, 2, 6, 7, 11, 12
2. Guo, S., Yan, Z., Zhang, K., Zuo, W., Zhang, L.: Toward convolutional blind denoising of real photographs. In: CVPR. pp. 1712–1722 (2019) 2, 12
3. Laine, S., Karras, T., Lehtinen, J., Aila, T.: High-quality self-supervised deep image denoising. In: NIPS. pp. 6968–6978 (2019) 1, 2, 3, 4, 5
4. Lebrun, M., Colom, M., Morel, J.M.: The noise clinic: a blind image denoising algorithm. Image Processing On Line **5**, 1–54 (2015) 1
5. Liu, P., Zhang, H., Zhang, K., Lin, L., Zuo, W.: Multi-level wavelet-CNN for image restoration. In: CVPR Workshops. pp. 773–782 (2018) 3, 4, 5, 6, 7
6. Nam, S., Hwang, Y., Matsushita, Y., Joo Kim, S.: A holistic approach to cross-channel image noise modeling and its application to image denoising. In: CVPR. pp. 1683–1691 (2016) 1
7. Plotz, T., Roth, S.: Benchmarking denoising algorithms with real photographs. In: CVPR. pp. 1586–1595 (2017) 1
8. Yue, Z., Yong, H., Zhao, Q., Meng, D., Zhang, L.: Variational Denoising Network: Toward blind noise modeling and removal. In: NIPS. pp. 1688–1699 (2019) 2
9. Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L.: Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. TIP **26**(7), 3142–3155 (2017) 2, 11