Self-Supervised Real-World Denoising by Jointly Learning Visible and Invisible Noise Supplementary Material

Shaoyu Wang†
wangshapyu@dlmu.edu.cn
Changze Zhou†
zcz@dlmu.edu.cn
Bolin Song*
bolin_song@dmu.edu.cn
Yiyang Wang
yywerica@dlmu.edu.cn

College of Artificial Intelligence
Dalian Maritime University
Dalian, China

1 Comparison with State-of-the-arts

We also have advantages over other categories of denoising approaches, namely non-learning, supervised learning with synthetic pairs, unpaired learning, and supervised learning with real-world pairs. From the comparisons in Table 1, we can see that our approach surpasses non-learning approaches, unpaired learning methods, and even a few supervised methods with ground truth references such as DnCNN, TNRD, and CBDNet in terms of PSNR.

Dateset	Non-learning		Supervised learning with synthetic pairs		Unpaired learning			Self-supervised learning
	BM3D[8]	WNNM[8]	DnCNN[24]	Zhou *et al.* [24]	GCBD[8]	C2N[8]	D-BSN[24]	Ours
SIDD	25.65/0.685	25.78/0.809	23.66/0.583	34.00/0.898	-	35.35/0.937	-	37.16/0.936
DND	34.51/0.851	34.67/0.865	32.43/0.790	38.40/0.945	35.58/0.922	37.28/0.924	37.93/0.937	38.74/0.943
Dataset	Supervised learning with real-world pairs							Self-supervised learning
	TNRD[8]	CBDNet[8]	RIDNet[8]	AINDNet(R)[24]	VDN[43]	DANet[44]	MIRNet[40]	Ours
SIDD	24.73/0.643	33.28/0.868	38.70/0.950	38.84/0.951	39.26/0.955	39.43/0.956	39.96/0.960	37.16/0.936
DND	33.65/0.831	38.05/0.942	39.24/0.952	39.34/0.952	39.38/0.952	39.58/0.955	39.84/0.957	38.74/0.943

Table 1: Quantitative comparison of PSNR and SSIM on SIDD and DND datasets is conducted, wherein we present the official results as reported in the corresponding paper, which can also be cross-verified from benchmark websites. Red and blue colors are employed to denote the best and second-best outcomes.

†The first two authors contributed equally.
* Corresponding author

2 Ablation on Hyper-parameters

Hyper-parameter λ of TV term. The hyper-parameter λ is employed to control the magnitude of the edge-preserving effect of $\mathbf{x}_i$. Fig. 1 (a) depicts the performance of our model across different values of λ on the SIDD validation dataset. The denoising performance shows a notable decline when the regularization term, represented by $\lambda = 0$, is removed due to the presence of unavoidable image structures in the estimated NV map. Furthermore, due to the degradation of image details caused by PD, the resulting denoised image is significantly influenced by artifacts. When $\lambda \to \infty$, the TV regularization term takes precedence over the loss function, leading to an excessive level of smoothing and consequently resulting in subpar denoising performance. In this study, we select a value of 0.1 for the hyper-parameter λ in order to optimize both qualitative and quantitative performance.

Hyper-parameters α and β of the learning flow for more visible noise. We also investigate the impact of α and β on the denoising and detail enhancement abilities of the entire network. Referring to Sec. 3.2, we aim to determine appropriate values for α and β given these particular conditions: $\alpha > 1$ or $\beta \in (0, 1)$. We optimize the hyper-parameters α and β through grid search, and the optimal combination is found at $\alpha = 1.8$ and $\beta = 0.6$. Thus, we hereby present the curves of α and β as shown in Fig. 1 (b) while maintaining the remaining parameters at their optimal values.

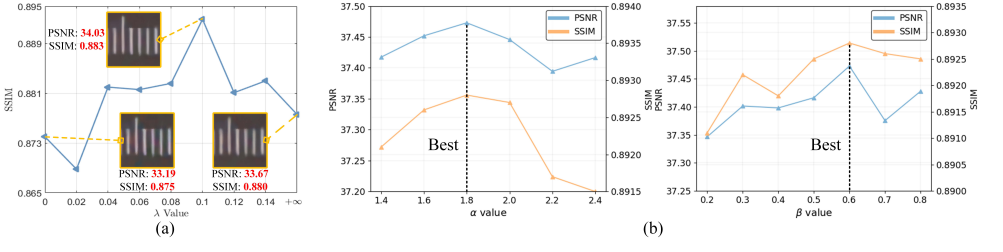

Figure 1: (a) Ablation study of λ of TV regularization term in the flow of learning invisible noise. (b) Ablation study of hyper-parameters α and β in the loss of $\mathcal{L}_v$ for learning more visible noise.

3 More Results of NV Map

We present more results of NV maps acquired through the blind estimation module of our framework. The estimated results selected from the DND benchmark [16] are provided in Fig. 2. In this figure, the pairs of noisy images and corresponding NV maps are presented, where the left image represents the noisy input and the right image displays the estimated NV map. Besides, we also present the outcomes associated with the SIDD benchmark in Fig. 3. From the results provided in Fig. 2 and Fig. 3, it is evident that our estimated NV maps exhibit a clear consistency with the visual effects of the noisy images, even without utilizing ground truth or references for noise level assessment.

In addition, since the SIDD validation dataset provides clean references that are regarded as the ground truth of the noisy images, we employ them to generate the “NV map reference"

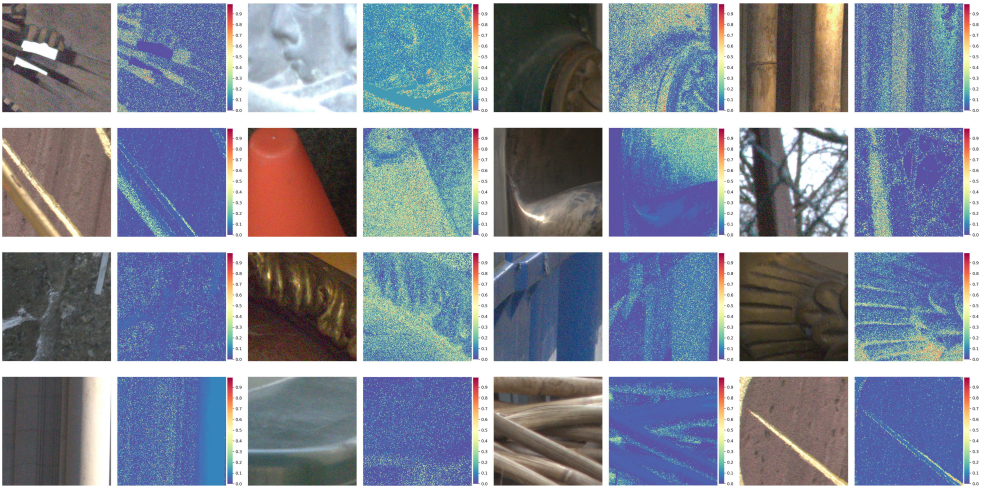

Figure 2: The noisy images and their corresponding NV maps obtained by our blind estimation module. The images are selected from the DND benchmark [16].

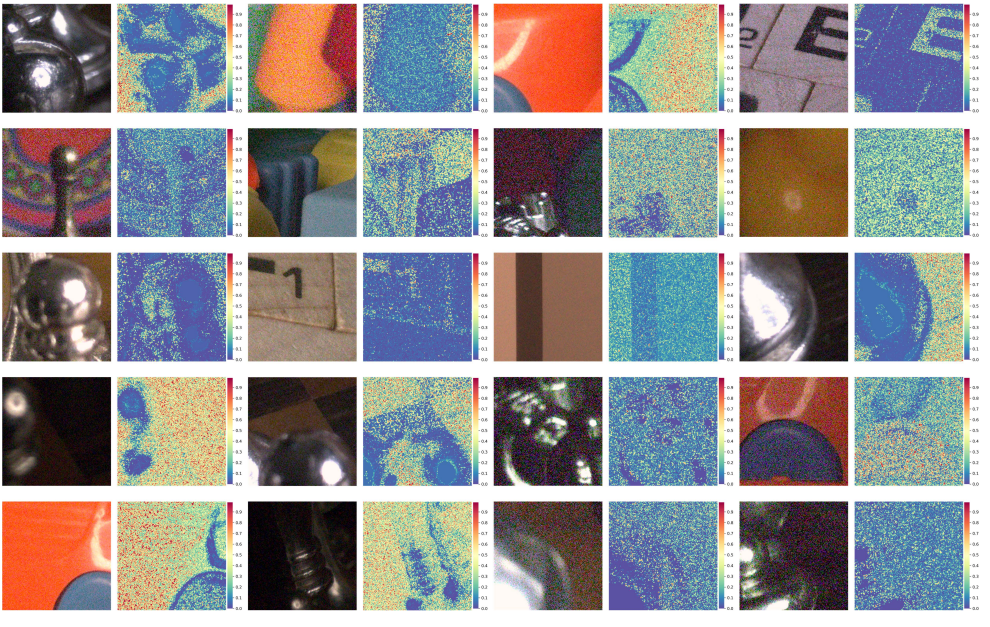

Figure 3: The noisy images and their corresponding NV maps obtained by our blind estimation module. The images are selected from the SIDD benchmark [1].

for validating the reliability of our estimated NV maps. To be specific, the “NV map reference" is computed by normalizing the absolute value of the difference between the noisy image and the clean reference. Then, we provide the results in Fig. 4, from where the top left image shows the noisy image, and the bottom left image represents the clean reference provided in the SIDD validation dataset. The top right image shows the NV map obtained by our blind estimation module, and finally, the bottom right image illustrates the “NV map

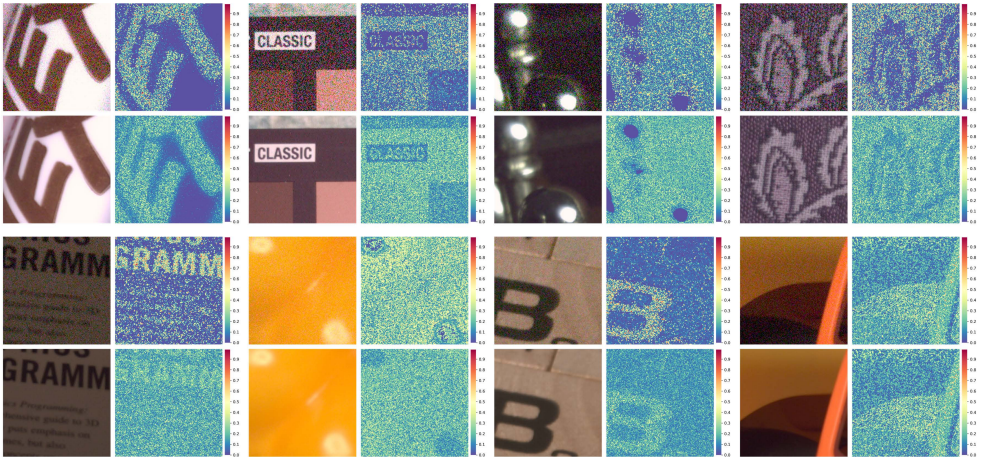

Figure 4: NV map comparisons with the “NV map reference" obtained by the noisy image and the clean reference provided within the SIDD validation dataset [1]. The top left image shows the noisy image, and the bottom left image represents the clean reference provided in the SIDD validation dataset. The top right image shows the NV map obtained by our blind estimation module, and finally, the bottom right image illustrates the “NV map reference".

reference". From the comparisons, it is evident that our estimated NV map exhibits a high degree of similarity with the “NV map reference", thereby confirming the effectiveness of our estimation module.

4 More Visual Comparisons with SOTAs

We present more visual comparisons on SIDD dataset and DND dataset, as depicted in Fig. 5 and Fig. 6, respectively. We compare our approach with other self-supervised methods, including CVF-SID[14], AP-BSN[12], SDAP[15], SS-BSN[8], C-BSN[10] and SASL[13]. For the compared approaches except for CVF-SID[14] and SASL[13], we utilize the official code provided from the author’s Github and train the model according to the configurations stated in each respective paper. While for CVF-SID and SASL, we utilize the code with the pre-trained model released from the authors.

For the DND dataset, since some approaches lack official code or encounter memory limitations, we solely compare our method with CVF-SID, AP-BSN, SDAP, C-BSN and SASL. All the comparisons have verified the superiority of our proposed approach.

References

[1] Abdelrahman Abdelhamed, Stephen Lin, and Michael S. Brown. A high-quality denoising dataset for smartphone cameras. In *CVPR*, 2018.

[2] Saeed Anwar and Nick Barnes. Real image denoising with feature attention. In *ICCV*, 2019.

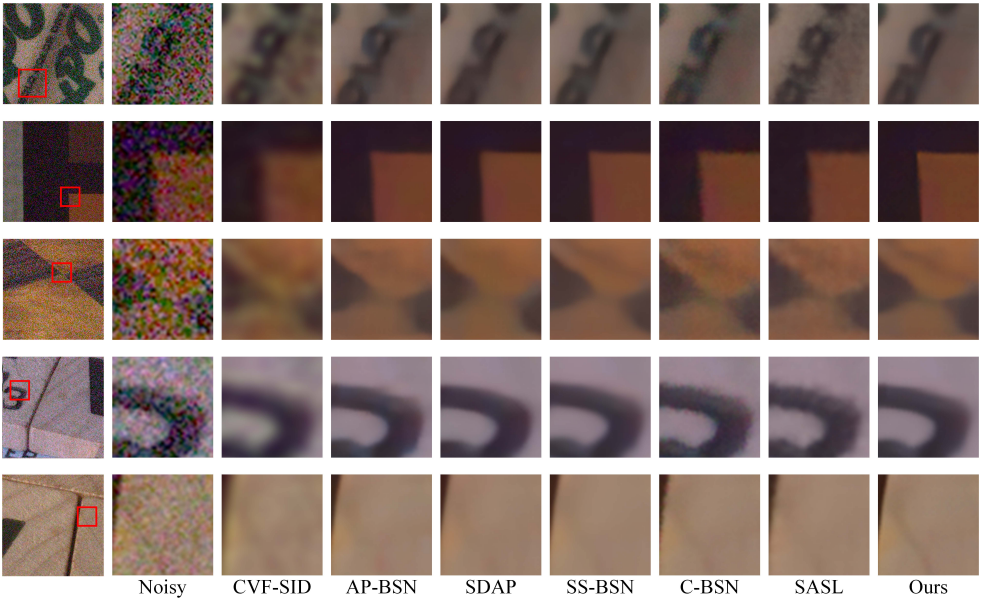

Figure 5: Visual comparisons on SIDD dataset [1] with state-of-the-art self-supervised real-world denoising approaches, including CVF-SID, AP-BSN, SDAP, SS-BSN, C-BSN and SASL.

[3] Jingwen Chen, Jiawei Chen, Hongyang Chao, and Ming Yang. Image blind denoising with generative adversarial network based noise modeling. In *CVPR*, 2018.

[4] Yunjin Chen and Thomas Pock. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6):1256–1272, 2016.

[5] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, 2007.

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

[7] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *CVPR*, 2019.

[8] Young-Joo Han and Ha-Jin Yu. SS-BSN: Attentive blind-spot network for self-supervised denoising with nonlocal self-similarity. In *IJCAI*, 2023.

[9] Geonwoon Jang, Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. C2N: Practical generative noise modeling for real-world denoising. In *ICCV*, 2021.

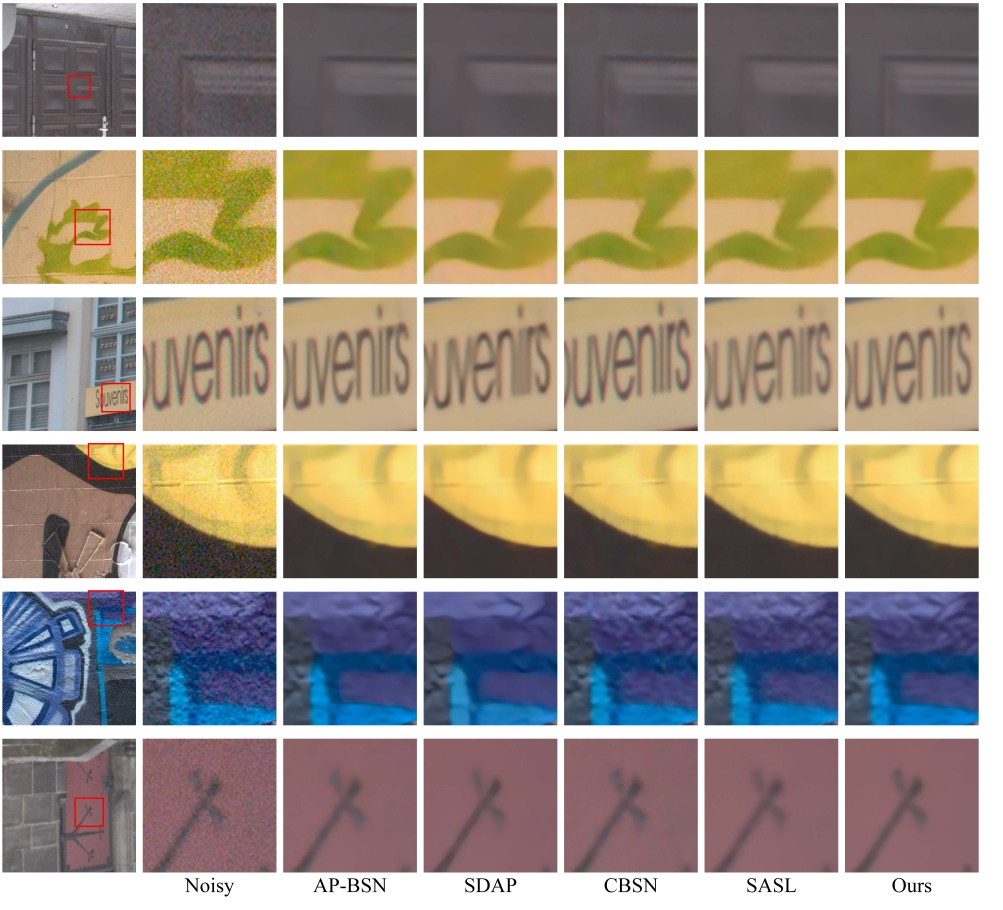

Figure 6: Visual comparisons on DND dataset [16] with state-of-the-art self-supervised real-world denoising methods, including AP-BSN, SDAP, C-BSN and SASL.

[10] Yeong Il Jang, Keuntek Lee, Gu Yong Park, Seyun Kim, and Nam Ik Cho. Self-supervised image denoising with downsampled invariance loss and conditional blind-spot network. In *ICCV*, 2023.

[11] Yoonsik Kim, Jae Woong Soh, Gu Yong Park, and Nam Ik Cho. Transfer learning from synthetic to real-noise denoising with adaptive instance normalization. In *CVPR*, 2020.

[12] Wooseok Lee, Sanghyun Son, and Kyoung Mu Lee. AP-BSN: Self-supervised denoising for real-world images via asymmetric pd and blind-spot network. In *CVPR*, 2022.

[13] Junyi Li, Zhilu Zhang, Xiaoyu Liu, Chaoyu Feng, Xiaotao Wang, Lei Lei, and Wangmeng Zuo. Spatially adaptive self-supervised learning for real-world image denoising. In *CVPR*, 2023.

[14] Reyhaneh Neshatavar, Mohsen Yavartanoo, Sanghyun Son, and Kyoung Mu Lee. CVF-SID: Cyclic multi-variate function for self-supervised image denoising by disentangling noise from image. In *CVPR*, 2022.

[15] Yizhong Pan, Xiao Liu, Xiangyu Liao, Yuanzhouhan Cao, and Chao Ren. Random sub-samples generation for self-supervised real image denoising. In *ICCV*, 2023.

[16] Tobias Plötz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *CVPR*, 2017.

[17] Xiaohe Wu, Ming Liu, Yue Cao, Dongwei Ren, and Wangmeng Zuo. Unpaired learning of deep image denoising. In *ECCV*, 2020.

[18] Zongsheng Yue, Hongwei Yong, Qian Zhao, Lei Zhang, and Deyu Meng. Variational Denoising Network: Toward blind noise modeling and removal. In *NeurIPS*, 2019.

[19] Zongsheng Yue, Qian Zhao, Lei Zhang, and Deyu Meng. Dual adversarial network: Toward real-world noise removal and noise generation. In *ECCV*, 2020.

[20] Syed Waqas Zamir, Aditya Arora, Salman Khan, Munawar Hayat, Fahad Shahbaz Khan, Ming-Hsuan Yang, and Ling Shao. Learning enriched features for real image restoration and enhancement. In *ECCV*, 2020.

[21] 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(7):3142–3155, 2017.

[22] Yuqian Zhou, Jianbo Jiao, Haibin Huang, Yang Wang, Jue Wang, Honghui Shi, and Thomas Huang. When awgn-based denoiser meets real noises, 2019.