

High-Resolution Document Shadow Removal via A Large-Scale Real-World Dataset and A Frequency-Aware Shadow Erasing Net (Supplementary Materials)

Zinuo Li* Xuhang Chen* Chi-Man Pun† Xiaodong Cun†
University of Macau

1. The Experiments on the Resolution of 512

We also evaluate the performance on three datasets: Jung [5], Kligler [6], and the proposed SD7K, in a relatively low resolution of 512×512 where the divisions between training and testing are shown in Table 1. We train each dataset separately to satisfy the settings of the previous works. The batch size is set to 1 for the other two datasets due to the limited samples.

Datasets	# of Training	# of Testing
Jung [5]	67	20
Kligler [6]	272	28
SD7K	6479	760

Table 1. The settings of the training and testing.

We give the quantitative evaluation in Table 2 and the visual results are available in Figure 2 and Figure 3. As shown in Table 2, our proposed FSENet outperforms all other methods under the low-resolution data setting. In the Figure 2 and Figure 3, we can clearly observe that Wang *et al.* [15], SP+M+I Net [8], SG-ShadowNet [12] and ShadowFormer [3] exhibit the phenomenon of incomplete shadow removal. Meanwhile, [15] shows a large difference from the target in white balance. Despite BEDSR-Net [9] performing relatively well in both white balance and shadow removal, a close examination of the finer details reveals that BEDSR-Net yields blurry results with missing texture details. Simultaneously, BEDSR-Net fails in Figure 3 in terms of shadow removal. In comparison, our result achieves the best performance in white balance, shadow removal, and detail restoration.

2. More High-Resolution Results on SD7K

We give more visual comparisons on the high-resolution inputs in Figure 4. The results in the first two rows come from Kligler dataset, and the results in the last two rows

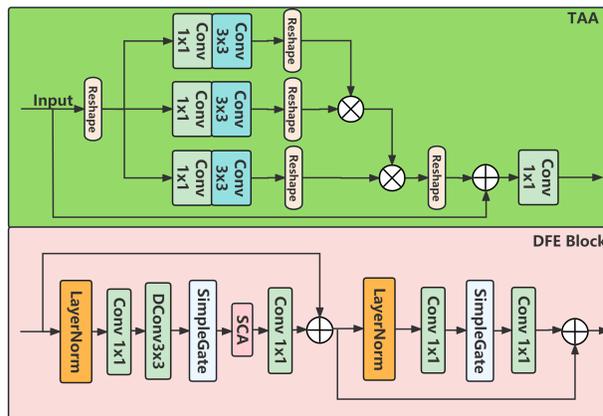


Figure 1. The network structure of our DFE and TAA.

come from SD7K. It can be observed that the shadow removal results from Jung *et al.* [5] are unstable, and the white balance often differs from the target. Meanwhile, shadow removal with MaskShadowGAN [4] is always incomplete, with the original shadow regions prone to leaving stains and causing texts to become unreadable. In contrast, our method performs relatively well in both white balancing and shadow removal.

2.1. The Detailed Structure of DFE and TAA Blocks

In Sec. 4 of the main paper, we show the introduction of the TAA and DFE by text description, here, we give a detailed network structure to better understand our method as shown in Figure 1.

*Equal Contribution

†Corresponding Authors

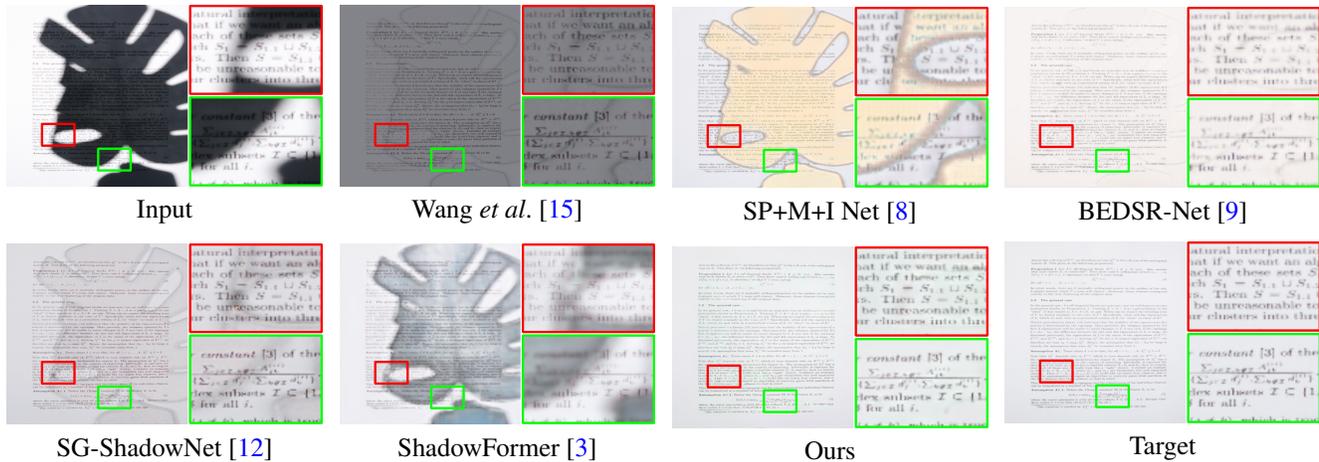


Figure 2. Qualitative results of the methods comparison in low-resolution training samples on SD7K dataset.

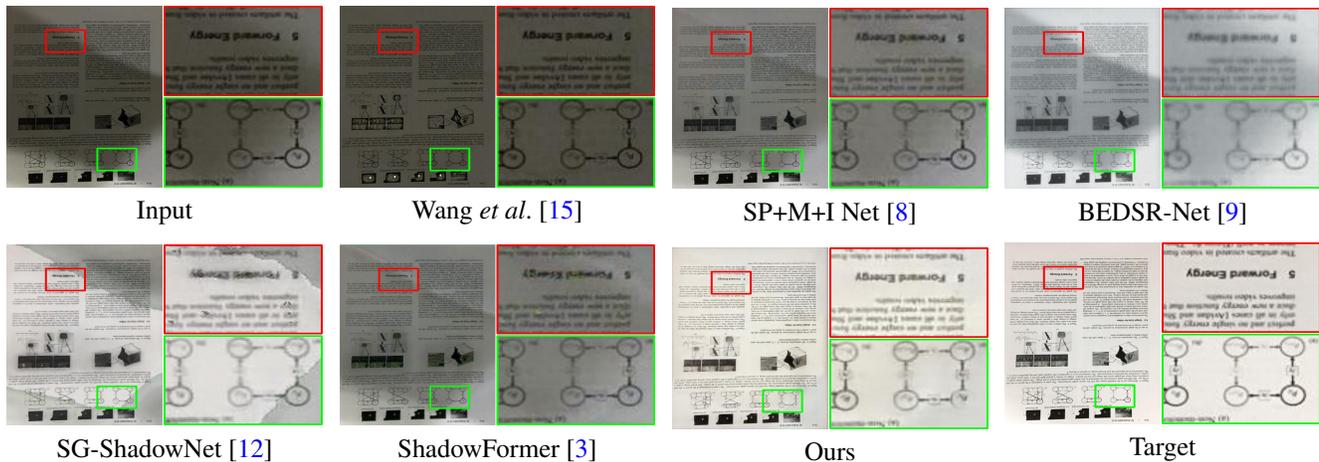


Figure 3. Qualitative results of the methods comparison in low resolution on Jung dataset.

Method	Jung (512 × 512)			Kligler (512 × 512)			SD7K (512 × 512)		
	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓	PSNR↑	SSIM↑	RMSE↓
Input	13.01	0.82	60.85	13.26	0.80	56.73	15.95	0.89	44.09
Wang et al. [15]	11.17	0.78	73.27	15.73	0.82	44.04	15.31	0.82	47.88
Wang et al. [13]	9.11	0.71	90.99	15.38	0.72	48.03	13.32	0.68	67.48
Shah et al. [11]	14.69	0.80	47.97	8.36	0.70	97.86	9.89	0.71	86.35
Jung et al. [5]	22.77	0.88	19.13	14.30	0.84	49.91	19.86	0.92	26.76
AEFNet [2]	23.52	0.85	19.44	19.53	0.89	27.72	24.18	0.95	16.83
BEDSR-Net [9]	21.51	0.85	22.58	22.31	0.75	20.86	21.50	0.90	30.52
DHAN [1]	20.58	0.82	25.95	25.66	0.84	15.49	25.61	0.85	14.27
LG-ShadowNet [10]	19.99	0.84	27.69	26.29	0.87	14.38	24.88	0.86	16.77
Mask-ShadowGAN [4]	19.41	0.82	29.16	25.79	0.87	15.01	24.82	0.87	15.43
SG-ShadowNet [12]	22.90	0.86	19.20	25.34	0.91	14.96	28.22	0.96	10.61
SP+M Net [7]	20.04	0.84	29.13	18.85	0.86	29.57	18.84	0.91	30.94
ST-CGAN [14]	13.93	0.33	52.51	12.17	0.44	63.22	12.87	0.32	61.17
ShadowFormer [3]	20.61	0.85	26.96	17.26	0.77	36.31	23.71	0.90	17.54
BMNet [16]	23.12	0.80	18.19	26.15	0.84	14.17	24.86	0.80	15.59
Ours	24.36	0.88	16.04	29.11	0.94	10.49	28.67	0.96	10.00

Table 2. Quantitative results of the models' comparison on three datasets in low-resolution. The best result is highlighted in red and bold.



Figure 4. Qualitative results of the methods comparison in high resolution, the first two rows represent the results on the Kligler dataset and the last two rows depict the results on the SD7K dataset.

References

- [1] Xiaodong Cun, Chi-Man Pun, and Cheng Shi. Towards ghost-free shadow removal via dual hierarchical aggregation network and shadow matting gan. In *AAAI*, pages 10680–10687, 2020. [2](#)
- [2] Lan Fu, Changqing Zhou, Qing Guo, Felix Juefei-Xu, Hongkai Yu, Wei Feng, Yang Liu, and Song Wang. Auto-exposure fusion for single-image shadow removal. In *CVPR*, pages 10571–10580, 2021. [2](#)
- [3] Lanqing Guo, Siyu Huang, Ding Liu, Hao Cheng, and Bihan Wen. Shadowformer: Global context helps image shadow removal. In *AAAI*, 2023. [1](#), [2](#)
- [4] Xiaowei Hu, Yitong Jiang, Chi-Wing Fu, and Pheng-Ann Heng. Mask-shadowgan: Learning to remove shadows from unpaired data. In *ICCV*, pages 2472–2481, 2019. [1](#), [2](#), [3](#)
- [5] Seungjun Jung, Muhammad Abul Hasan, and Changick Kim. Water-filling: An efficient algorithm for digitized document shadow removal. In *Asian Conference on Computer Vision*, pages 398–414. Springer, 2018. [1](#), [2](#), [3](#)
- [6] Netanel Kligler, Sagi Katz, and Ayellet Tal. Document enhancement using visibility detection. In *CVPR*, pages 2374–2382, 2018. [1](#)
- [7] Hieu Le and Dimitris Samaras. From shadow segmentation to shadow removal. In *ECCV*, August 2020. [2](#)
- [8] Hieu Le and Dimitris Samaras. Physics-based shadow image decomposition for shadow removal. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 44(12):9088–9101, 2021. [1](#), [2](#)
- [9] Yun-Hsuan Lin, Wen-Chin Chen, and Yung-Yu Chuang. Bedsr-net: A deep shadow removal network from a single document image. In *CVPR*, pages 12905–12914, 2020. [1](#), [2](#)
- [10] Zhihao Liu, Hui Yin, Yang Mi, Mengyang Pu, and Song Wang. Shadow removal by a lightness-guided network with training on unpaired data. *IEEE Transactions on Image Processing*, 30:1853–1865, 2021. [2](#)
- [11] Vatsal Shah and Vineet Gandhi. An iterative approach for shadow removal in document images. *IEEE ICASSP*, pages 1892–1896, 2018. [2](#)
- [12] Jin Wan, Hui Yin, Zhenyao Wu, Xinyi Wu, Yanting Liu, and Song Wang. Style-guided shadow removal. In *ECCV*, pages 361–378. Springer, 2022. [1](#), [2](#)
- [13] Bingshu Wang and CL Philip Chen. An effective background estimation method for shadows removal of document images. In *IEEE International Conference on Image Processing (ICIP)*, pages 3611–3615. IEEE, 2019. [2](#)
- [14] Jifeng Wang, Xiang Li, and Jian Yang. Stacked conditional generative adversarial networks for jointly learning shadow detection and shadow removal. In *CVPR*, pages 1788–1797, 2018. [2](#)
- [15] Jian-Ren Wang and Yung-Yu Chuang. Shadow removal of text document images by estimating local and global background colors. *IEEE ICASSP*, pages 1534–1538, 2020. [1](#), [2](#)
- [16] Yurui Zhu, Jie Huang, Xueyang Fu, Feng Zhao, Qibin Sun, and Zheng-Jun Zha. Bijective mapping network for shadow removal. In *CVPR*, pages 5627–5636, June 2022. [2](#)