#2758 High-Resolution Document Shadow Removal via A Large-Scale Real-World Dataset and A Frequency-Aware Shadow Erasing Net

Zinuo Li[1], Xuhang Chen[1], Chi-Man Pun[1] and Xiaodong Cun[1]
[1]University of Macau

ICCV23 PARIS

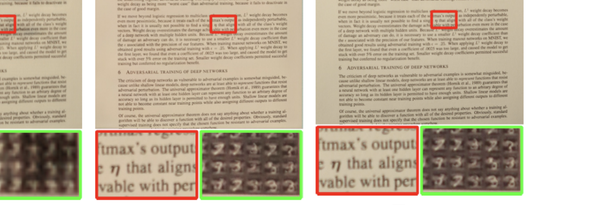
澳門大學 UNIVERSIDADE DE MACAU UNIVERSITY OF MACAU

Introduction

1. We provide SD7K, the only large-scale real-world dataset consisting of high-resolution shadow and the associated shadow-free images under various illumination conditions currently.
2. We propose FSENet, a frequency-aware network with a carefully designed network structure to handle high-resolution document shadows.

Data Distrubution

Data distribution of SD7K and quantitative comparison across all document shadow datasets:



Visual Result

Input | Jung *et al.* | MaskShadowGAN | Ours | Target

Figure 2: Visual comparison on high-resolution of different methods for image enhancement on the SD7K dataset.

Methodology and Quantitative Result

In order to enable multi-frequency optimization, we use a Laplacian pyramid to divide the original image into two high-frequency areas (textures) and one low-frequency area (color), different frequency components are refined by distinct modules. The low frequency is sent to Dimension-Aware Transformer (DAT) and Deep Feature Extraction (DFE), and the high-frequencies are sent to a series of carefully designed convolution with Texture Recover Module (TRM).

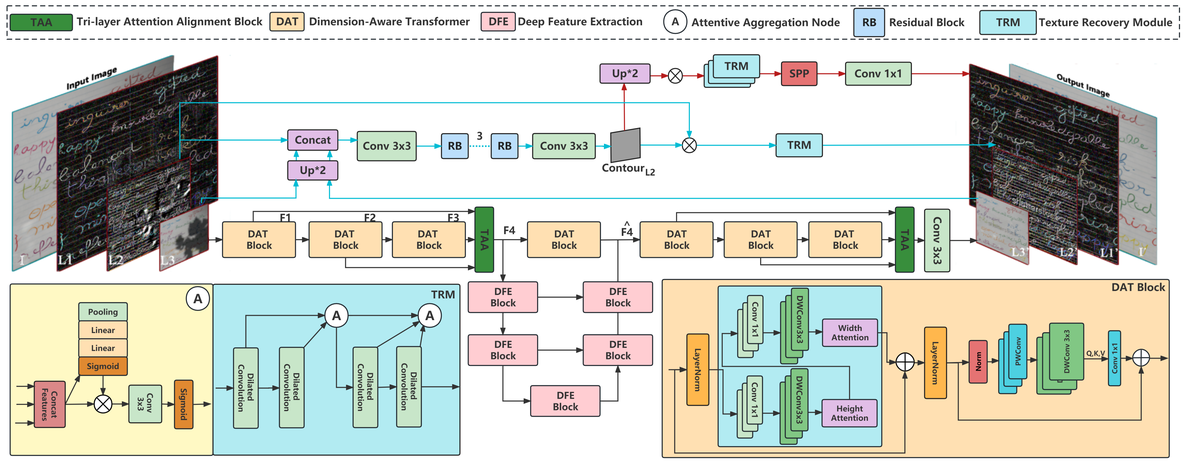


Figure 1: The overall structure of our FSENet. Different frequency domains are effectively learned and optimized.

Method	SD7K												Param (M)
	512 × 512				1024 × 1024				2462 × 3699 (Full Size)				
	PSNR↑	SSIM↑	RMSE↓	Time(s) ↓	PSNR↑	SSIM↑	RMSE↓	Time(s) ↓	PSNR↑	SSIM↑	RMSE↓	Time(s) ↓	
Input	15.95	0.89	44.09	N/A	15.95	0.90	44.09	N/A	15.94	0.91	44.14	N/A	N/A
Wang	15.31	0.82	47.88	> 10	15.31	0.85	47.88	> 30	15.29	0.86	47.95	> 60	N/A
Wang	13.32	0.68	67.48	> 10	13.32	0.71	67.48	> 30	13.26	0.74	68.07	> 60	N/A
Shah	9.89	0.71	86.35	> 10	9.90	0.76	86.35	> 30	9.88	0.79	86.46	> 60	N/A
Jung	19.86	0.92	26.76	> 5	19.86	0.93	26.76	> 10	19.82	0.92	26.86	> 30	N/A
MaskGAN	24.82	0.87	15.43	0.96	24.82	0.85	15.43	2.07	24.67	0.86	15.72	7.59	28.29
DHAN	25.61	0.85	14.27	0.21	25.60*	0.83*	14.27*	N/A	25.42*	0.85*	14.60*	N/A	27.82
AEFNet	24.18	0.95	16.83	0.11	23.00*	0.90*	19.14*	N/A	22.94*	0.90*	19.28*	N/A	391.10
BEDSR-Net	21.50	0.90	30.52	0.85	19.81*	0.85*	33.03*	N/A	19.74*	0.86*	33.18*	N/A	29.44
ShadowFormer	23.71	0.90	17.54	0.15	22.69*	0.84*	19.65*	N/A	22.64*	0.87*	19.79*	N/A	11.35
BMNet	24.86	0.80	15.59	0.18	24.84*	0.79*	15.67*	N/A	24.70*	0.83*	15.96*	N/A	2.11
Ours	28.69	0.97	9.98	0.24	28.68	0.97	9.98	0.75	28.67	0.96	10.00	7.93	29.34

Table 1: Quantitative comparisons on the SD7K dataset of different shadow removal methods. The top result is highlighted in red. Our proposed method significantly outperforms similar methods in terms of image quality metrics, yet its speed does not lag far behind.

References

[1] Jie Liang, Hui Zeng, and Lei Zhang. High-resolution photorealistic image translation in real-time: A laplacian pyramid translation network. In *CVPR*, 2021.

Conclusion

In this paper, we tackle the important problem of removing shadows from high-resolution documents. We gather a high-resolution dataset SD7K, which is comprised of over 7000 triplets of real-world document images with distinct characteristics under various illumination conditions. In the meanwhile, we propose FSENet to remove shadows in higher resolution via frequency decomposition.

Contact Information

Web https://www.cis.um.edu.mo/ cmpun/
Email cmpun@umac.mo
Phone +853 8822 4369