ScannerNet: A Deep Network for Scanner-Quality Document Images under Complex Illumination

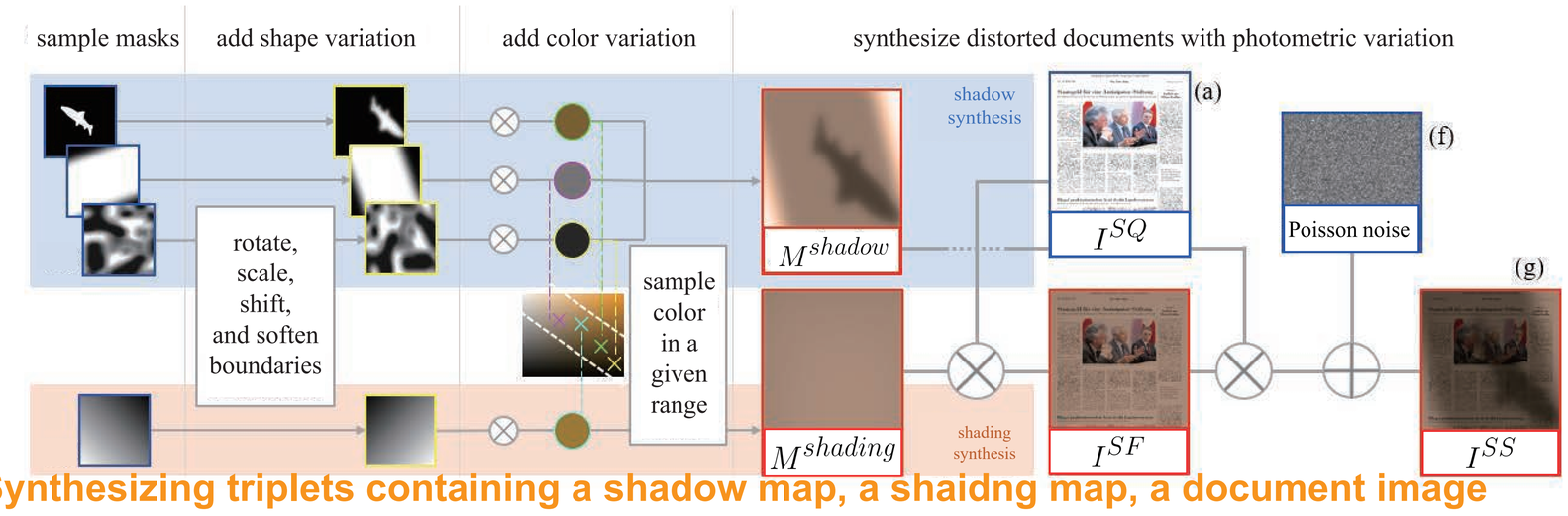
Chih-Jou Hsu[1], Yu-Ting Wu[2], Ming-Sui Lee[1], Yung-Yu Chuang[1]
1 National Taiwan University, Taipei, Taiwan,
2 National Taipei University, New Taipei City, Taiwan

Problem Description

Document images captured by mobile devices are often subject to photometric distortions, including shadows, non-uniform shading, and color shift due to the imperfect white balance of sensors, which significantly reduces legibility and visual quality. Despite the fact that real photographs often contain a mixture of these distortions, the majority of existing approaches to document illumination correction concentrate on only a small subset of these distortions.

Proposed Method

The proposed data synthesis process.



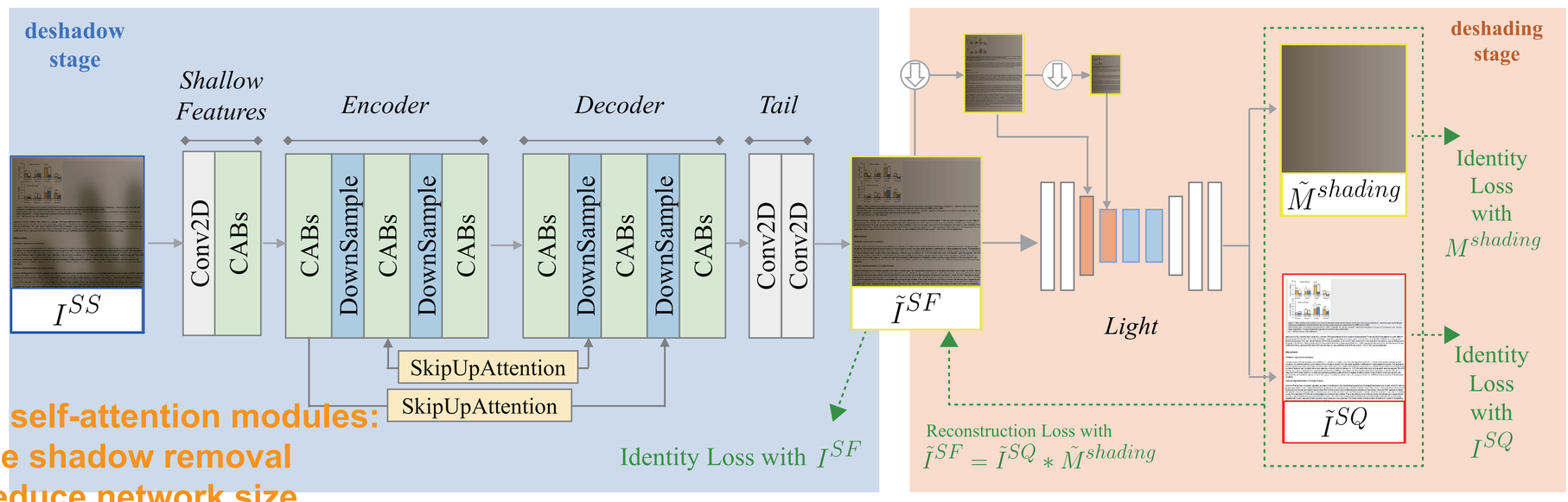
Synthesizing triplets containing a shadow map, a shaidng map, a document image with shading, a document image with shadows and a scanner-quality document image.

Two examples of our SSQD.

With the data synthesis process, we generate 7000 triplets for training our network, 1000 for validation, and 2000 for testing. We call our dataset *Synthetic Scanner-Quality Dataset (SSQD).*

The proposed ScannerNet.

Identity Loss: Providing grouth truth from synthesizing our shadow and shading maps.



Network with self-attention modules: enhancing the shadow removal ablitliy and reduce network size.

Two stage design: exploit the shadows and shading respectively.

Quantitative Results

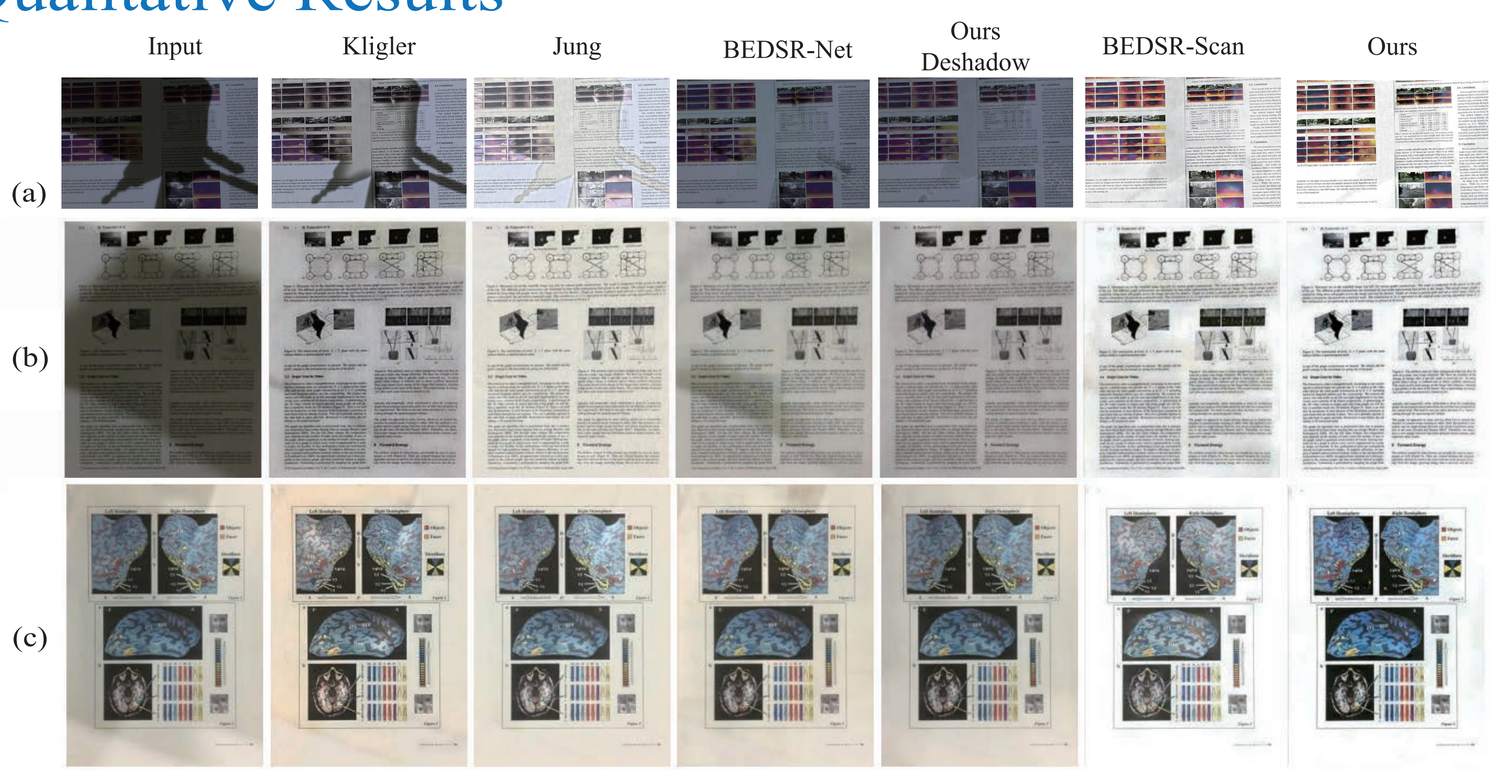
Evaluation for Shadow Removal on Real Images

Methods	Model Size	Bako's Dataset PSNR	Bako's Dataset SSIM	Jung's Dataset PSNR	Jung's Dataset SSIM	Lin's RDSRD PSNR	Lin's RDSRD SSIM	Kligler's Dataset PSNR	Kligler's Dataset SSIM	Our SSQD PSNR	Our SSQD SSIM
Input Shadow Image		28.45	0.974	20.35	0.885	21.73	0.809	19.31	0.843	18.93	0.692
Bako [8]		35.22	**0.982**	23.70	0.902	28.24	0.866	29.66	0.905	24.39	0.850
Jung [8]		13.88	0.806	**28.49**	0.911	14.45	0.705	19.21	0.872	14.66	0.768
BEDSR-Net [18]	19.8M	35.07	0.981	27.23	0.912	**33.48**	**0.908**	**32.90**	**0.935**	25.37	0.821
DeshadowNet (ours)	**3.6M**	**35.86**	0.981	26.39	**0.914**	33.00	0.905	31.92	0.932	**32.09**	**0.908**
U-Net [8] (SDSRD)	17.3M	32.44	0.979	25.66	0.910	30.10	0.893	30.15	0.929	26.82	0.867
U-Net [8] (SSQD)	17.3M	33.68	0.977	25.82	0.915	30.92	0.895	31.71	0.935	30.27	0.900
DeshadowNet (SDSRD)	3.6M	35.38	0.982	26.28	0.903	31.53	0.902	28.76	0.918	26.55	0.846
DeshadowNet (SSQD)	3.6M	35.86	0.981	26.34	0.914	33.00	0.905	31.92	0.932	32.09	0.908

Evaluation for Scanner Quality on Synthetic Images

Method	SSIM ↑	PSNR ↑
Bako *et al.* [1]	0.86	20.19
Jung *et al.* [10]	0.91	27.3
Li *et al.* [17]	0.37	13.24
Li *et al.* trained on SSQD	0.33	14.08
BEDSR-Net [18]	0.83	22.81
BEDSR-Net trained on SSQD	**0.93**	27.97
Ours	**0.93**	**28.11**

Qualitative Results



We compare our methods with previous approaches.
First, we compare the shadow removal results of Kligler, Jung, BEDSR-Net, and our DeshadowNet.
Second, we compare the scanner-quality results of our method with those of BEDSR-Net trained on our dataset (BEDSR-Scan).

- (a) illustrates an example with extremely dark shadows that can only be removed completely by our method.
- Complicated multi-cast shadows are present in (b), and only Jung and DeshadowNet can handle them successfully.
- Our method better preserves the colors of figures in (a)(c).