Structure-Guided Diffusion Models for High-Fidelity Portrait Shadow Removal — Supplementary Material

This supplementary material accompanies our paper "Structure-Guided Diffusion Models for High-Fidelity Portrait Shadow Removal", including details of the adversarial loss for training SE-Net, visual illustration of SE-Net's network architecture, more portrait shadow removal results, our results on natural images, and additional validation results for the robustness to inaccurate masks.

1. Adversarial Loss for Training SE-Net

The adversarial loss $\mathcal{L}_{GAN}$ in paper is defined as:

$$\begin{aligned}\mathcal{L}_{GAN} =& \mathbb{E}\big[\log\big(D(G_p(I))\big)\big] + \\ & \mathbb{E}\big[\log\big(1 - D(G_s(I_{syn}))\big)\big],\end{aligned} \tag{1}$$

where D is a discriminator for determining whether the structure maps generated by our SE-Net (referred to as G_s) are in the same domain as the output of the trained PDG model [3] (denoted by G_p).

2. Network Architecture of SE-Net

Figure 2 gives the details the network architecture of SE-Net, where we denote the output channel as c, the convolution kernel size as k, stride in a convolution layer as s, and "Norm" as instance normalization.

3. More Portrait Shadow Removal Results

Figures 3-24 provide more visual comparison of portrait shadow removal on the PSM dataset [4] and our collected dataset. As can be seen, for all these cases, our method generates high-quality results and clearly outperforms the others.

4. Our Results on Natural Images

To adapt our method to shadow removal of natural images, we replace the PDG model from [3] with a generic edge detection model from [1], and then train our framework on ISTD dataset [2]. As shown in Figure 25, even without any parameter tuning or model fine-tuning, our trained model produces very competitive results for natural images from the test set of ISTD, which are comparable or even better than current leading natural image shadow removal methods, manifesting the effectiveness of our shadow removal framework.

5. Extended results of mask robustness tests (60%–80% coverage)

As shown in Figure 1, our method is able to produce satisfactory shadow removal results even large inaccurate shadow masks covering 65% and 75% of the entire image are utilized, validating the robustness of our method to inaccurate masks.

Figure 1. **Shadow removal performance under varying mask inaccuracies.**

References

[1] Xavier Soria, Yachun Li, Mohammad Rouhani, and Angel D. Sappa. Tiny and efficient model for the edge detection generalization. In *ICCV*, 2023. 1

[2] 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. 1

[3] Ran Yi, Yong-Jin Liu, Yu-Kun Lai, and Paul L Rosin. Unpaired portrait drawing generation via asymmetric cycle mapping. In *CVPR*, pages 8217–8225, 2020. 1

[4] Xuaner Zhang, Jonathan T Barron, Yun-Ta Tsai, Rohit Pandey, Xiuming Zhang, Ren Ng, and David E Jacobs. Portrait shadow manipulation. *ACM Transactions on Graphics (TOG)*, 39(4):78–1, 2020. 1, 2, 3, 4, 5, 6, 7, 8

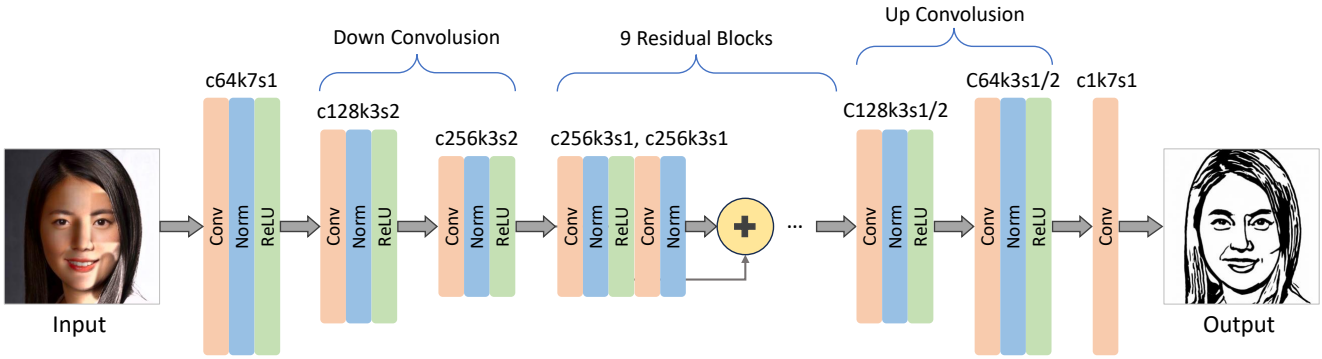

Figure 2. Visual illustration for the network architecture of SE-Net.

Figure 3. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 4. More visual comparison of shadow removal on the dataset provided in [4].

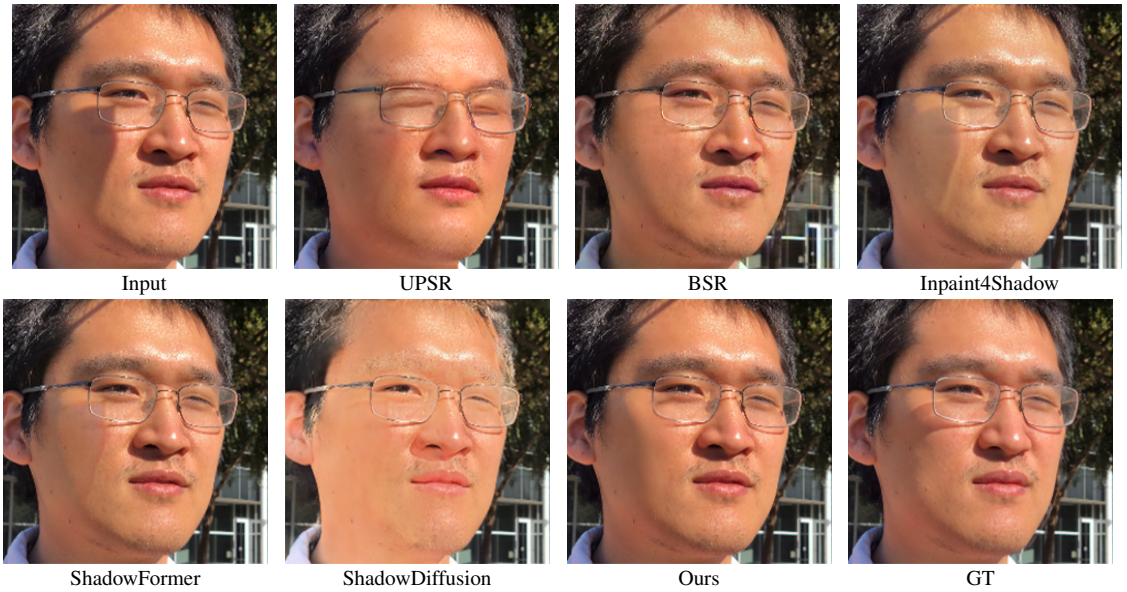

Figure 5. More visual comparison of shadow removal on the dataset provided in [4].

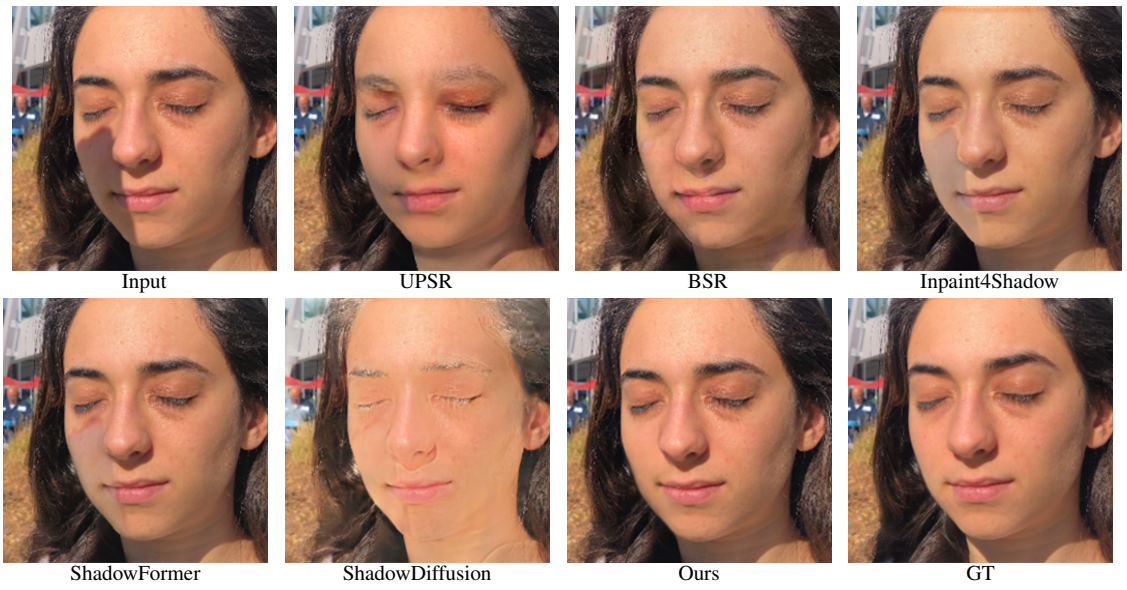

Figure 6. More visual comparison of portrait shadow removal on the PSM dataset [4].

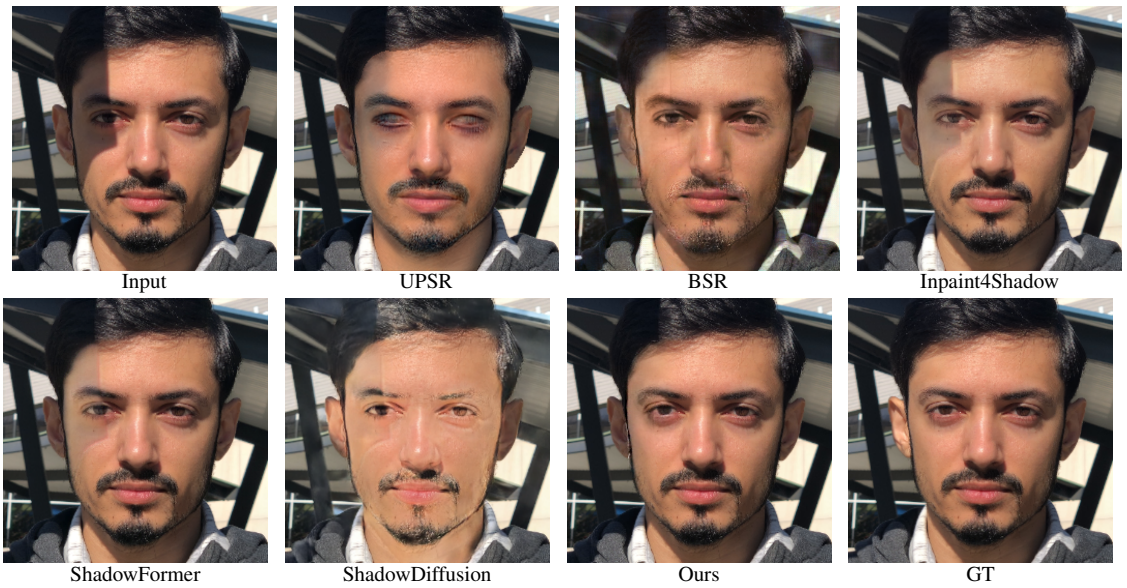

Figure 7. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 8. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 9. More visual comparison of shadow removal on the dataset provided in [4].

Figure 10. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 11. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 12. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 13. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 14. More visual comparison of portrait shadow removal on the PSM dataset [4].

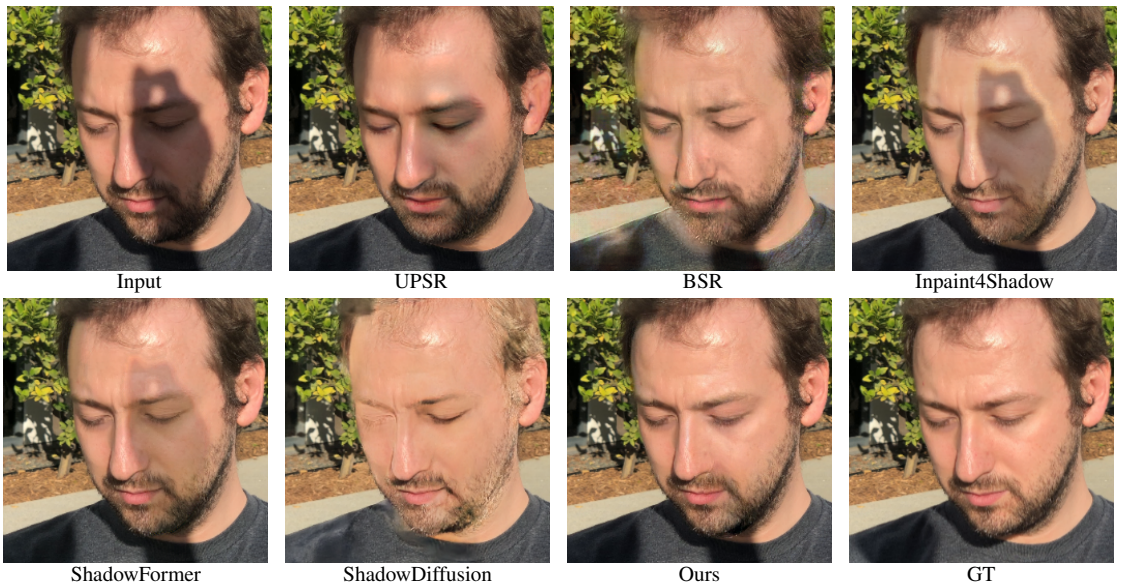

Figure 15. More visual comparison of portrait shadow removal on the PSM dataset [4].

Figure 16. More visual comparison of portrait shadow removal on the PSM dataset [4].

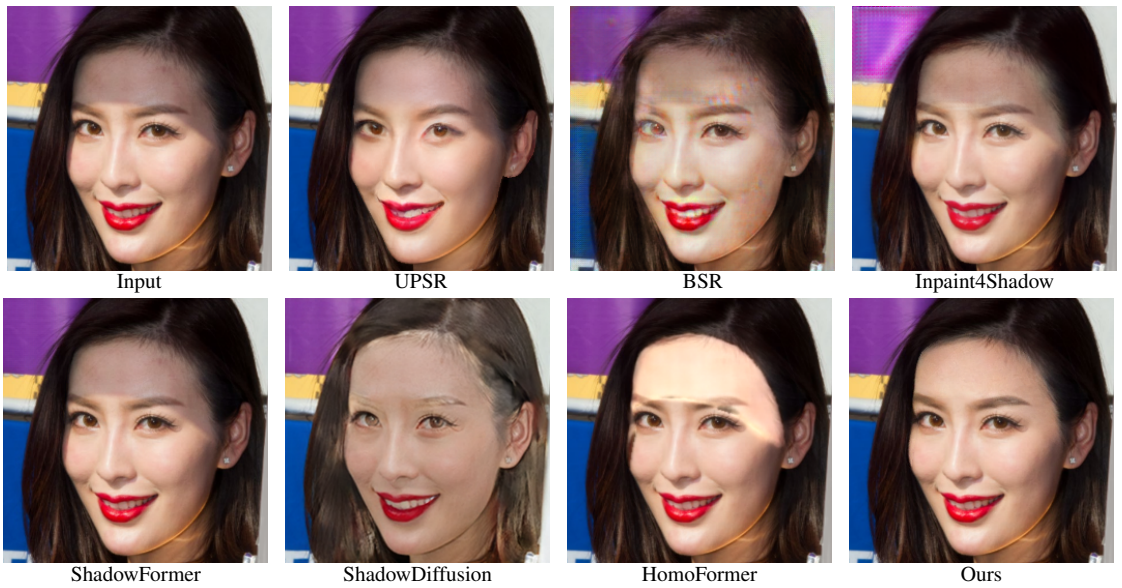

Figure 17. More visual comparison of portrait shadow removal on our collected dataset.

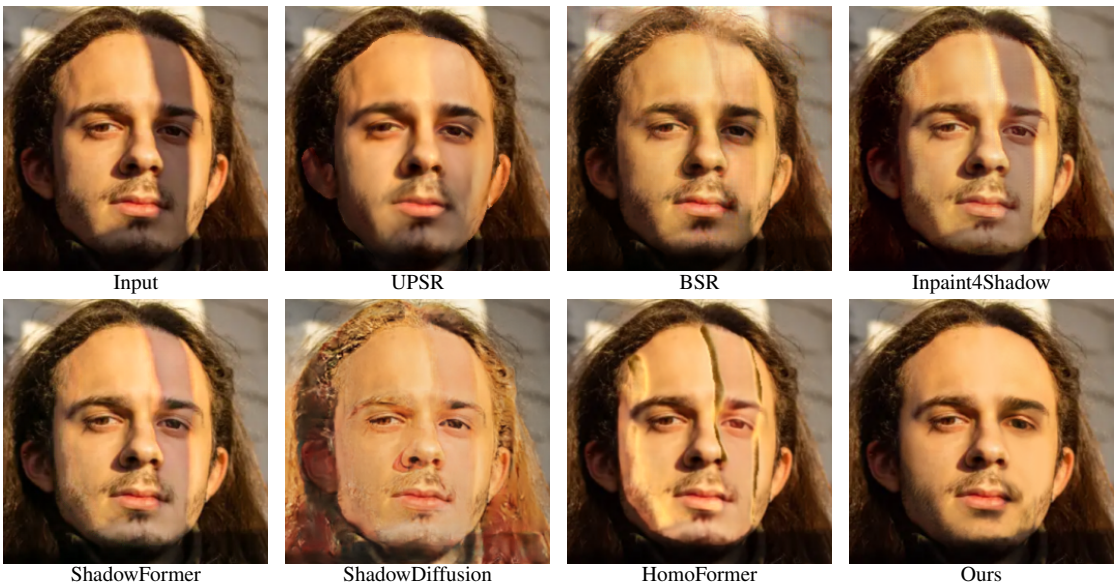

Figure 18. More visual comparison of portrait shadow removal on our collected dataset.

Figure 19. More visual comparison of portrait shadow removal on our collected dataset.

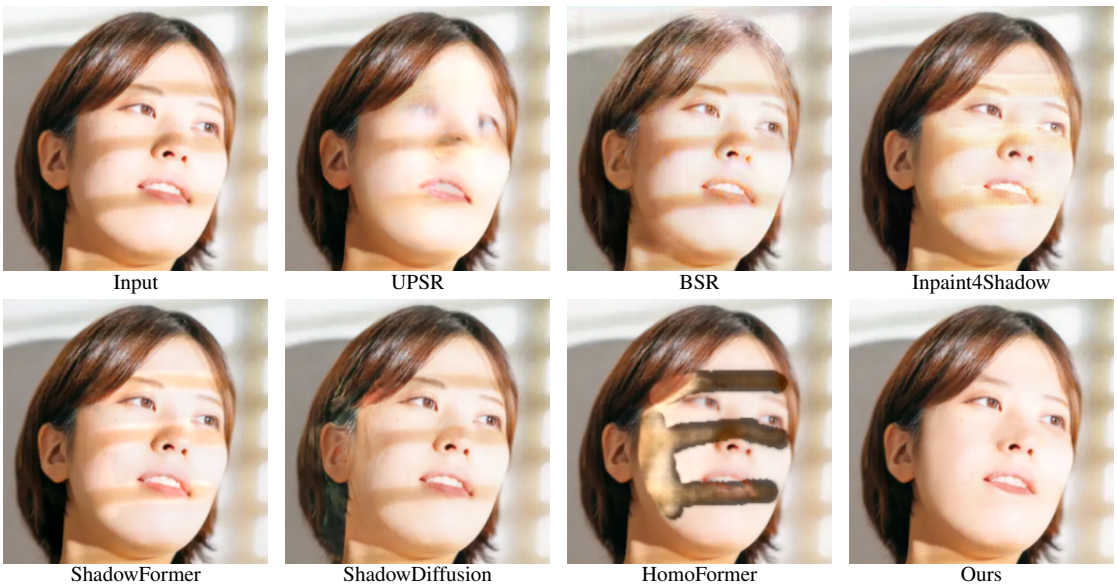

Figure 20. More visual comparison of portrait shadow removal on our collected dataset.

Figure 21. More visual comparison of portrait shadow removal on our collected dataset.

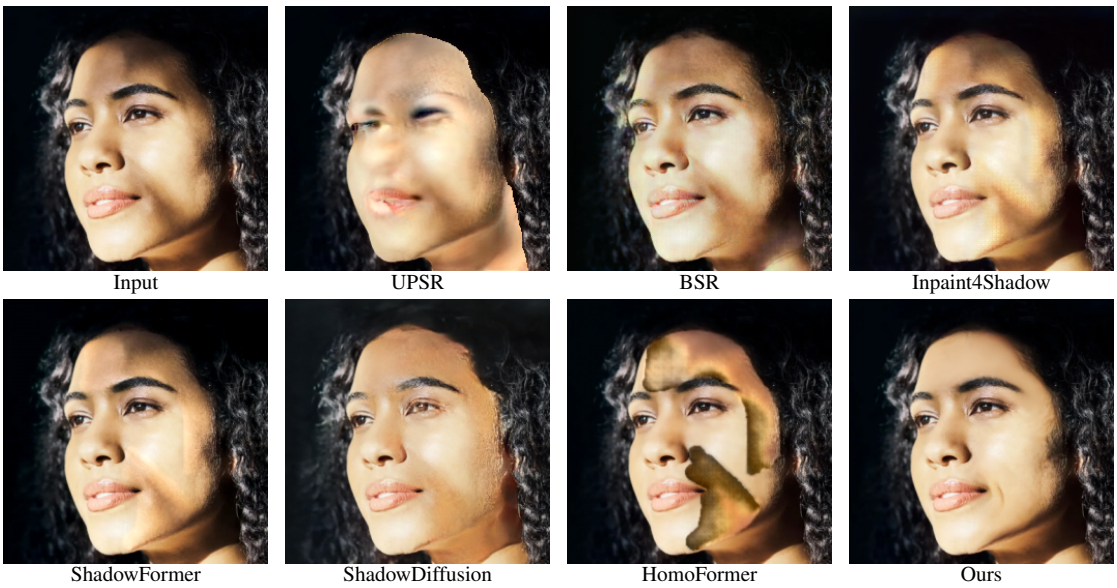

Figure 22. More visual comparison of portrait shadow removal on our collected dataset.

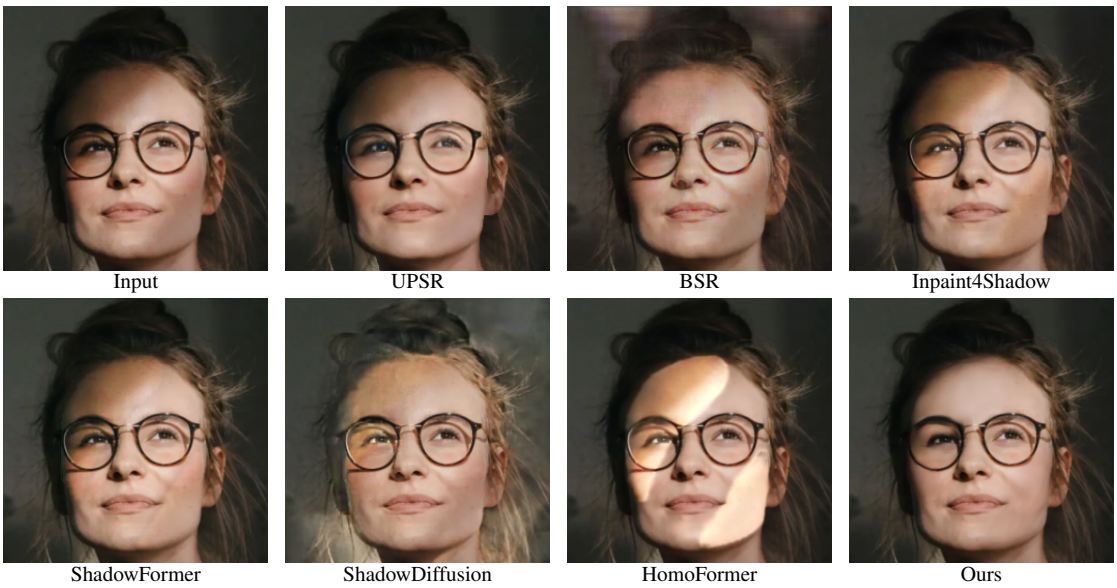

Figure 23. More visual comparison of portrait shadow removal on our collected dataset.

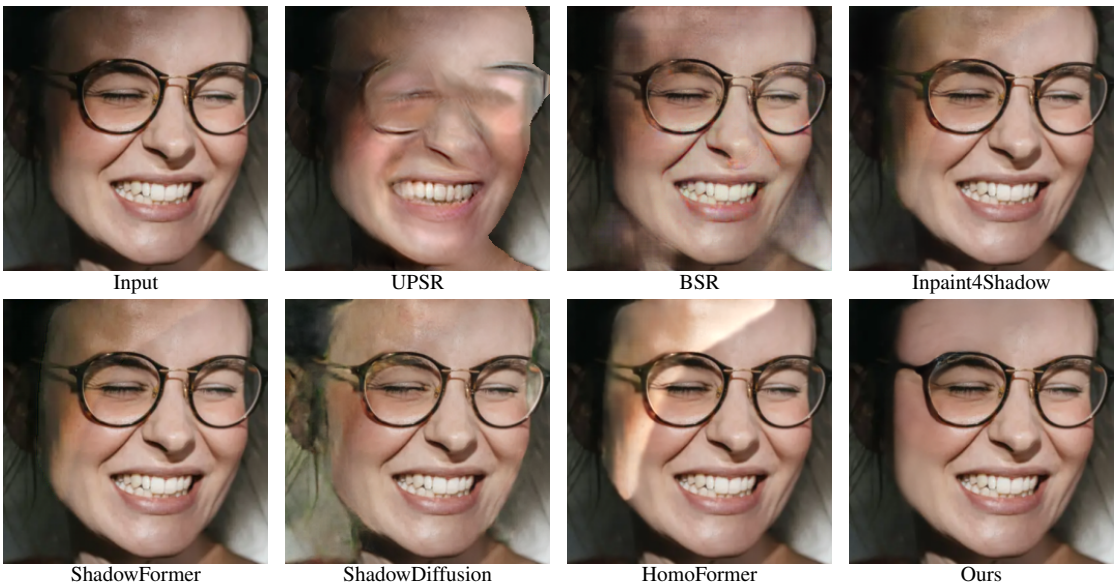

Figure 24. More visual comparison of portrait shadow removal on our collected dataset.

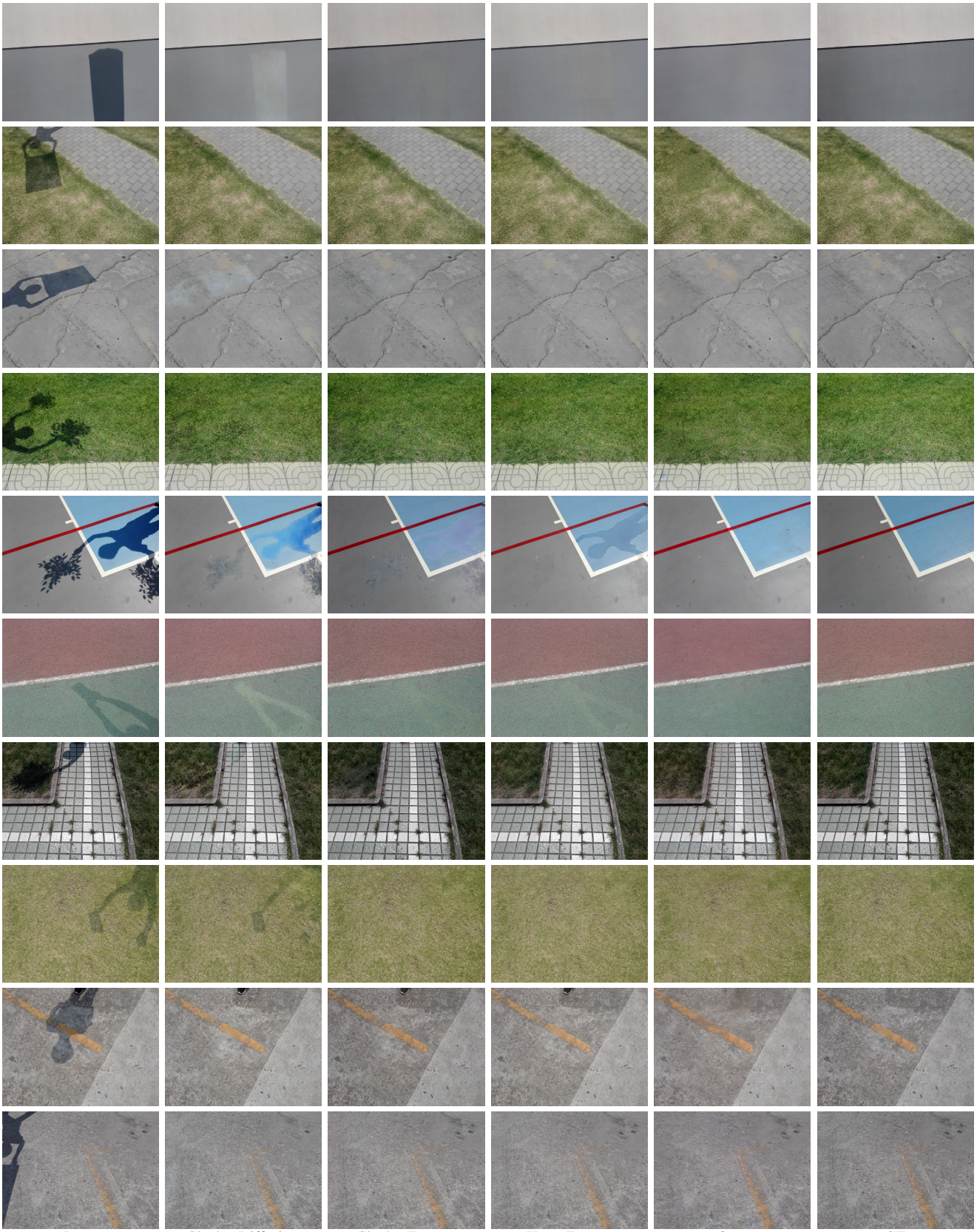

Figure 25. Visual comparison with state-of-the-art natural image shadow removal methods on the ISTD dataset.