Supplemental Material:
PIDSR: Complementary Polarized Image Demosaicing and Super-Resolution

Shuangfan Zhou[1#] Chu Zhou[2#] Youwei Lyu[1] Heng Guo[1*] Zhanyu Ma[1] Boxin Shi[3,4] Imari Sato[2]

[1] School of Artificial Intelligence, Beijing University of Posts and Telecommunications

[2] National Institute of Informatics

[3] State Key Laboratory for Multimedia Information Processing, School of Computer Science, Peking University

[4] National Engineering Research Center of Visual Technology, School of Computer Science, Peking University

{zhoushuangfan, youweilv, guoheng, mazhanyu}@bupt.edu.cn   zhou_chu@hotmail.com

shiboxin@pku.edu.cn   imarik@nii.ac.jp

A. Details about our synthetic dataset

Since existing public datasets are insufficient for the setting of our PIDSR, we propose a synthetic dataset. It contians 138 different scenes (108 for training and 30 for test), and we perform data augmentation (including random flip and rotation) on each scene, so that in total the training and test sets contain 98496 and 27360 different images respectively.

Fig. A shows the detailed procedure of generating our dataset. For each scene, we place a linear polarizer in front of a conventional RGB camera (Hikrobot MV-CA050-12UC with an F/2.8 lens), and rotate the polarizer to four specific polarizer angles $\alpha_{1,2,3,4} = 0°, 45°, 90°, 135°$ in succession. At each polarizer angle $\alpha_i(i = 1, 2, 3, 4)$, we capture 100 polarized images and average them to serve as the HR ground truth polarized image $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_i}(i = 1, 2, 3, 4)$ (the ground truth of PISR task). Then, we downsample $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_{1,2,3,4}}$ to obtain the ground truth polarized images $\mathbf{I}^{\mathrm{gt}}_{\alpha_{1,2,3,4}}$ (the ground truth of PID task). After that, according to Eq. (4) of the main paper, we manually add mosaic to $\mathbf{I}^{\mathrm{gt}}_{\alpha_{1,2,3,4}}$ to synthesize the corresponding CPFA raw image $\mathbf{R}$. Note that the size of $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_{1,2,3,4}}$ is set to be 1024×1224 for $2\times$ SR experiments and 2048×2448 for $4\times$ SR experiments respectively, while the size of $\mathbf{I}^{\mathrm{gt}}_{\alpha_{1,2,3,4}}$ and $\mathbf{R}$ is fixed to 512×612.

B. Computational complexity analysis

To evaluate the computational complexity, we compare the Params, FLOPs, and inference time on our synthetic test dataset using a single NVIDIA A800 GPU among our PIDSR, three state-of-the-art PID methods (Polanalyser [3], IGRI2 [5], and TCPDNet [6]), and the only existing two PISR methods (PSRNet [1] and CPSRNet [9]), as shown in Tab. A. Here, the size of the input CPFA raw images is

---

# Equal contribution.   * Corresponding author.

Table A. Computational complexity analysis.

Demosaicing	Params[M]	FLOPs[G]	Inference time[s]
Polanalyser [3]	-	-	0.047
IGRI2 [5]	-	-	29.466
TCPDNet [6]	34.522	767.779	0.089
PIDSR	9.162	247.202	0.187

Super-resolution	Params[M]	FLOPs[T]	Inference time[s]
PSRNet ($2\times$) [1]	2.329	7.147	0.624
CPSRNet ($2\times$) [9]	17.110	6.086	2.596
PIDSR ($2\times$)	9.162	1.182	0.435
PSRNet ($4\times$) [1]	2.477	7.358	0.644
CPSRNet ($4\times$) [9]	16.962	5.321	2.709
PIDSR ($4\times$)	9.162	5.910	0.890

512×612 pixels for measuring FLOPs, and the inference time is the time taken to reconstruct a single scene. Note that CPSRNet [9] generates a single polarized image corresponding to one polarizer angle per run. Therefore, to obtain the complete set of outputs for comparison, the model must be executed four times, each with a different polarizer angle. Besides, since Polanalyser [3] and IGRI2 [5] are not learning-based methods, we can neither evaluate their Params nor FLOPs.

C. Additional results on synthetic data

We provide additional visual quality comparisons on synthetic data among our PIDSR, three state-of-the-art PID methods (Polanalyser [3], IGRI2 [5], and TCPDNet [6]), and the only existing two PISR methods (PSRNet [1] and CPSRNet [9]), as shown in Fig. B.
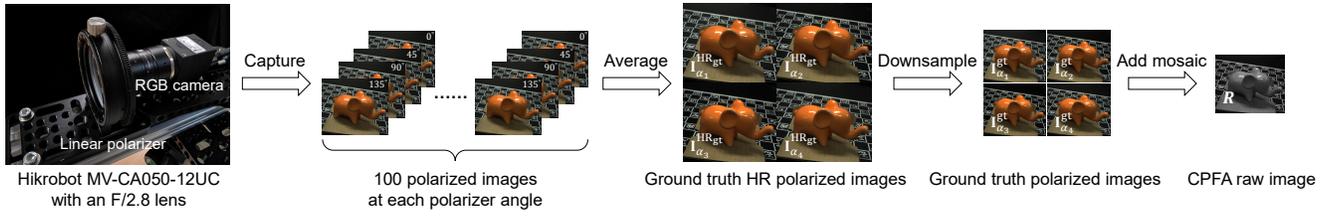
Figure A. The detailed procedure of generating our dataset.
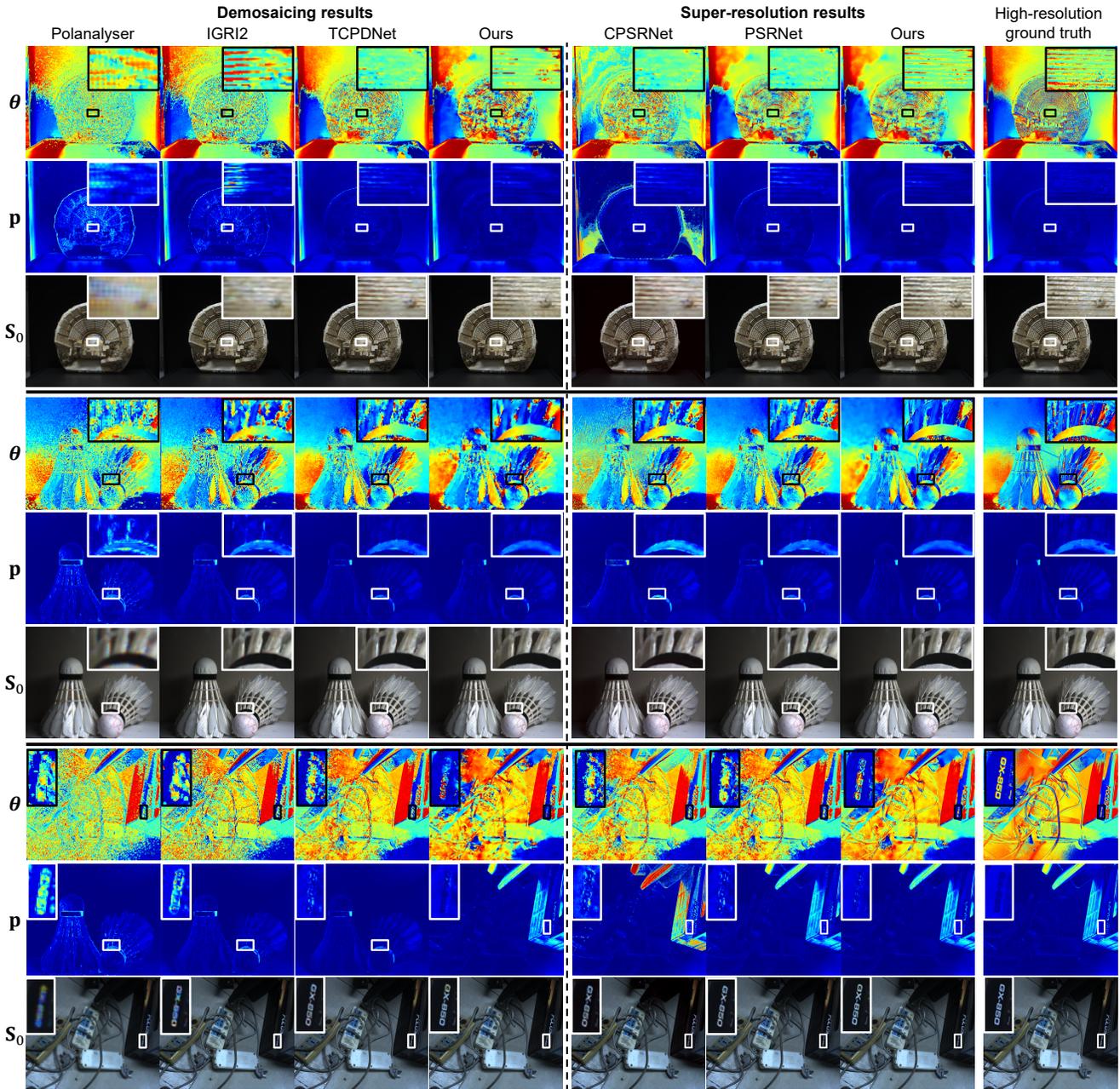
Figure B. Qualitative comparisons on synthetic data of both demosaicing and $2\times$ SR tasks.

D. Additional results on real data

We provide additional visual quality comparisons on real data among our PIDSR, three state-of-the-art PID methods (Polanalyser [3], IGRI2 [5], and TCPDNet [6]), and the only existing two PISR methods (PSRNet [1] and CPSRNet [9]), as shown in Fig. C.

E. Qualitative evaluation of ablation study

We present an example of the qualitative evaluation of the ablation study in Fig. D. As shown, our complete PIDSR produces the results with the best visual quality.

F. Additional results of application

We provide additional results of polarization-based reflection removal (PRR) produced by the following approaches: (1) "PID→PISR→PRR": performing PID and PISR sequentially on the CPFA raw image, then performing reflection removal; (2) "PID→PRR→SISR": performing PID on the CPFA raw image first, then performing reflection removal, and performing single image super-resolution (SISR) in the end; (3) "PIDSR→PRR": performing our PIDSR on the CPFA raw image first, then performing reflection removal, as shown in Fig. E. Here, the SR scale is 2, and the involved PRR, PID, PISR, and SISR methods are selected to be RSP [2], TCPDNet [6], PSRNet [1], and OmniSR [8] respectively.

G. Generalization ability evaluation

To validate the generalization ability of our PIDSR and the compared methods (including three state-of-the-art PID methods (Polanalyser [3], IGRI2 [5], and TCPDNet [6]), and the only existing two PISR methods (PSRNet [1] and CPSRNet [9])), we also evaluate them on two existing datasets (KAUST dataset [7] and Tokyo Tech dataset [4]). The KAUST dataset [7] provides 40 different scenes with a resolution of 1024×1024 pixels, and the Tokyo Tech dataset [4] provides 40 different scenes with a resolution of 1024×768 pixels. Those scenes were also captured by placing a linear polarizer in front of a conventional RGB camera and rotating it to four specific polarizer angles $\alpha_{1,2,3,4} = 0°, 45°, 90°, 135°$ in succession.

Since the KAUST dataset [7] and Tokyo Tech dataset [4] are initially designed for the PID task only, in order to make them applicable for both PID and PISR tasks, we adopt a similar approach as our procedure of dataset capturing. First, we treat the provided polarized images as the ground truth HR polarized images $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_{1,2,3,4}}$ (*i.e.*, the ground truth of PISR task). Then, we downsample $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_{1,2,3,4}}$ to half the resolution, and treat the downsampled ones as the ground truth polarized images $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_{1,2,3,4}}$ (*i.e.*, the ground truth of PID task). After that, we manually add mosaic to $\mathbf{I}^{\mathrm{HR_{gt}}}_{\alpha_{1,2,3,4}}$ to synthesize the corresponding CPFA raw images $\mathbf{R}$.

Table B. Generalization ability evaluation on KAUST dataset[7].

Metric	PSNR↑/SSIM↑		MAE↓
Demosaicing	$\mathbf{S}_0$	$\mathbf{p}$	$\boldsymbol{\theta}$
Polanalyser [3]	38.48/0.9620	31.92/0.8717	7.4131
IGRI2 [5]	44.15/0.9866	35.37/0.9324	5.1473
TCPDNet [6]	45.21/0.9870	39.01/0.9430	4.0684
PIDSR	**46.99/0.9893**	**39.23/0.94933**	**3.7633**
Super Resolution	$\mathbf{S}_0^{\mathrm{HR}}$ (2×)	$\mathbf{p}^{\mathrm{HR}}$ (2×)	$\boldsymbol{\theta}^{\mathrm{HR}}$ (2×)
PSRNet [1]	41.36/0.9648	34.55/0.8917	5.6266
CPSRNet [9]	39.74/0.9573	34.81/0.9011	5.3150
PIDSR	**42.76/0.9720**	**35.29/0.9049**	**5.2796**

Table C. Generalization ability evaluation on Tokyo Tech dataset [4].

Metric	PSNR↑/SSIM↑		MAE↓
Demosaicing	$\mathbf{S}_0$	$\mathbf{p}$	$\boldsymbol{\theta}$
Polanalyser [3]	33.10/0.9042	26.88/0.6983	20.0742
IGRI2 [5]	37.75/0.9634	30.12/0.7985	16.9239
TCPDNet [6]	39.26/0.9718	33.66/0.8497	13.8324
PIDSR	**39.83/0.9759**	**35.14/0.8593**	**12.7000**
Super Resolution	$\mathbf{S}_0^{\mathrm{HR}}$ (2×)	$\mathbf{p}^{\mathrm{HR}}$ (2×)	$\boldsymbol{\theta}^{\mathrm{HR}}$ (2×)
PSRNet [1]	36.25/0.9458	31.10/0.8091	14.4521
CPSRNet [9]	34.46/0.9278	32.03/0.8004	14.2948
PIDSR	**37.97/0.9574**	**32.61/0.8104**	**14.0238**

Results are shown in Tab. B and Tab. C. Note that all learning-based methods involved in the comparisons, including our PIDSR, TCPDNet [6], PSRNet [1], and CPSRNet [9], were trained exclusively on our dataset. For evaluation, these models were tested on the entire KAUST dataset [7] and Tokyo Tech dataset [4], ensuring that the test datasets remain entirely unseen during training. From the results we can see that our PIDSR consistently outperforms the compared methods on all metrics in both demosaicing and SR tasks, which demonstrates that our PIDSR has better generalization ability than all compared methods.

References

[1] Haofeng Hu, Shiyao Yang, Xiaobo Li, Zhenzhou Cheng, Tiegen Liu, and Jingsheng Zhai. Polarized image super-resolution via a deep convolutional neural network. *Optics Express*, 31(5):8535–8547, 2023. 1, 3, 5

[2] Youwei Lyu, Zhaopeng Cui, Si Li, Marc Pollefeys, and Boxin Shi. Reflection separation using a pair of unpolarized and polarized images. In *Proc. of Advances in Neural Information Processing Systems*, 2019. 3

[3] Ryota Maeda. Polanalyser: Polarization image analysis tool, 2019. 1, 3

[4] Miki Morimatsu, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. Monochrome and color polarization de-
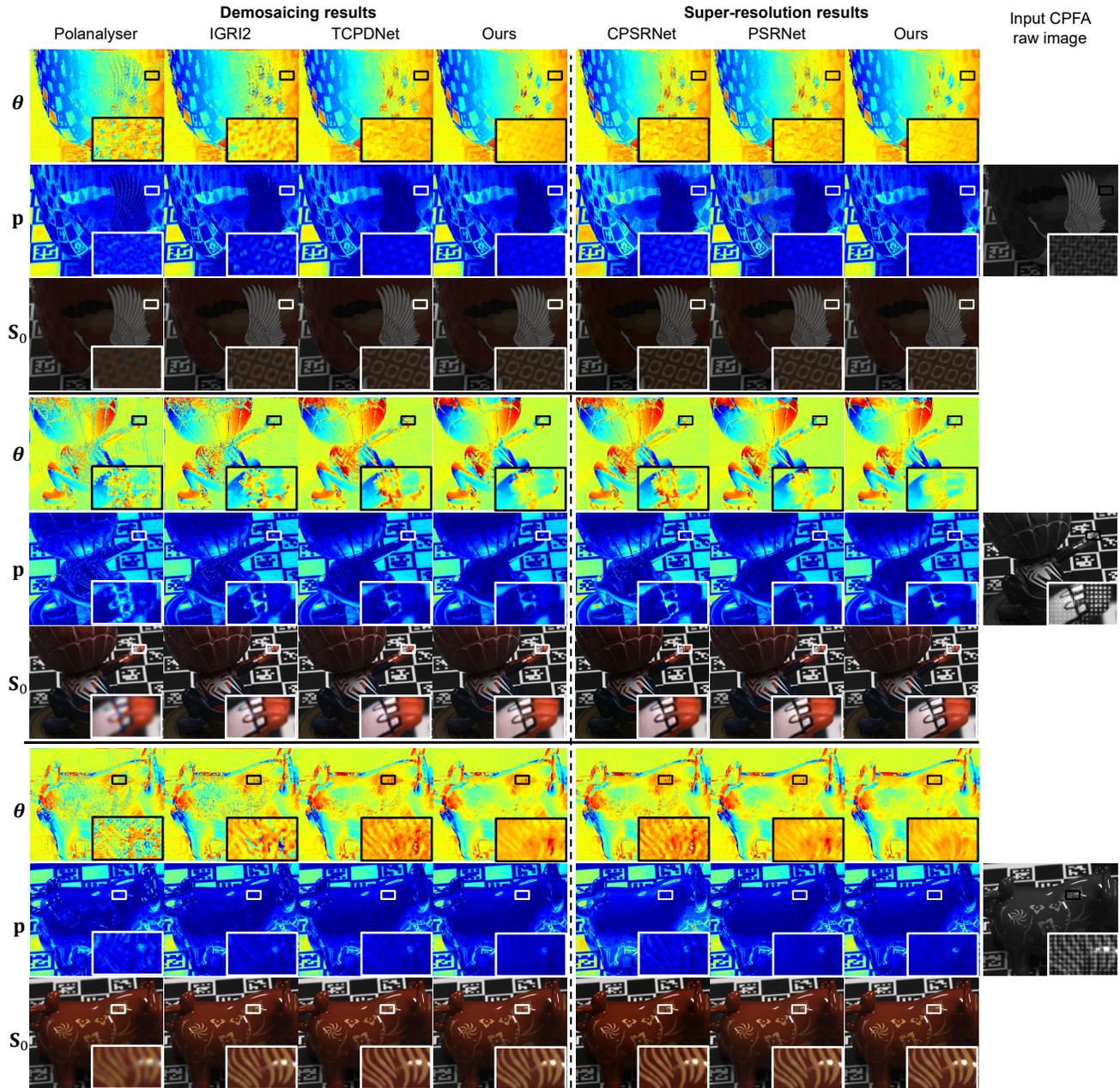
Figure C. Qualitative comparisons on real data of both demosaicing and $4\times$ SR tasks.

mosaicking using edge-aware residual interpolation. In *Proc. of International Conference on Image Processing*, pages 2571–2575, 2020. 3

[5] Miki Morimatsu, Yusuke Monno, Masayuki Tanaka, and Masatoshi Okutomi. Monochrome and color polarization demosaicking based on intensity-guided residual interpolation. *IEEE Sensors Journal*, 21(23):26985–26996, 2021. 1, 3

[6] Vy Nguyen, Masayuki Tanaka, Yusuke Monno, and Masatoshi Okutomi. Two-step color-polarization demosaicking network. In *Proc. of International Conference on Image Processing*,

pages 1011–1015, 2022. 1, 3, 5

[7] Simeng Qiu, Qiang Fu, Congli Wang, and Wolfgang Heidrich. Linear polarization demosaicking for monochrome and colour polarization focal plane arrays. In *Computer Graphics Forum*, pages 77–89, 2021. 3

[8] Hang Wang, Xuanhong Chen, Bingbing Ni, Yutian Liu, and Jinfan Liu. Omni aggregation networks for lightweight image super-resolution. In *Proc. of Computer Vision and Pattern Recognition*, pages 22378–22387, 2023. 3

[9] Dabing Yu, Qingwu Li, Zhiliang Zhang, Guanying Huo, Chang
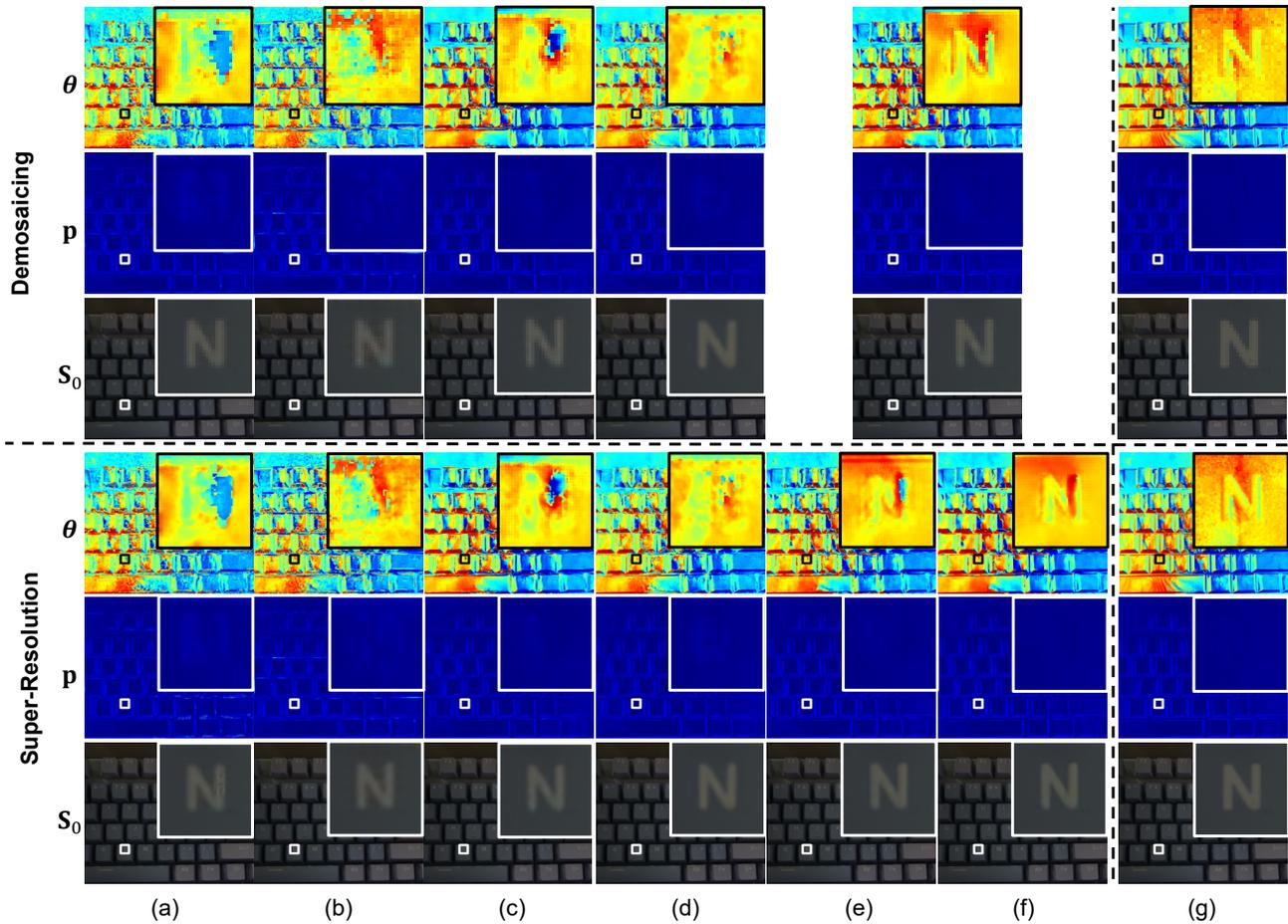
Figure D. Quantitative evaluation results of ablation study for demosaicing and super-resolution task. (a) Sequential $\mathcal{D}$ and $\uparrow$. (b) Single-stage pipeline. (c) Without SFI blocks. (d) TCPDNet [6]$\to$ ours (SR only). (e) Ours (demosaicing only)$\to$PSRNet [1]. (f) Our complete PIDSR. (g) Ground truth. Note that (e) and (f) share the same demosaicing result.

Input reflection-contaminated CPFA raw image

Reflection-free CPFA raw image (for reference)

Input reflection-contaminated CPFA raw image

Reflection-free CPFA raw image (for reference)

(1) PID$\to$PISR$\to$PRR (2) PID$\to$PRR$\to$SISR (3) PIDSR$\to$PRR (ours) (1) PID$\to$PISR$\to$PRR (2) PID$\to$PRR$\to$SISR (3) PIDSR$\to$PRR (ours)
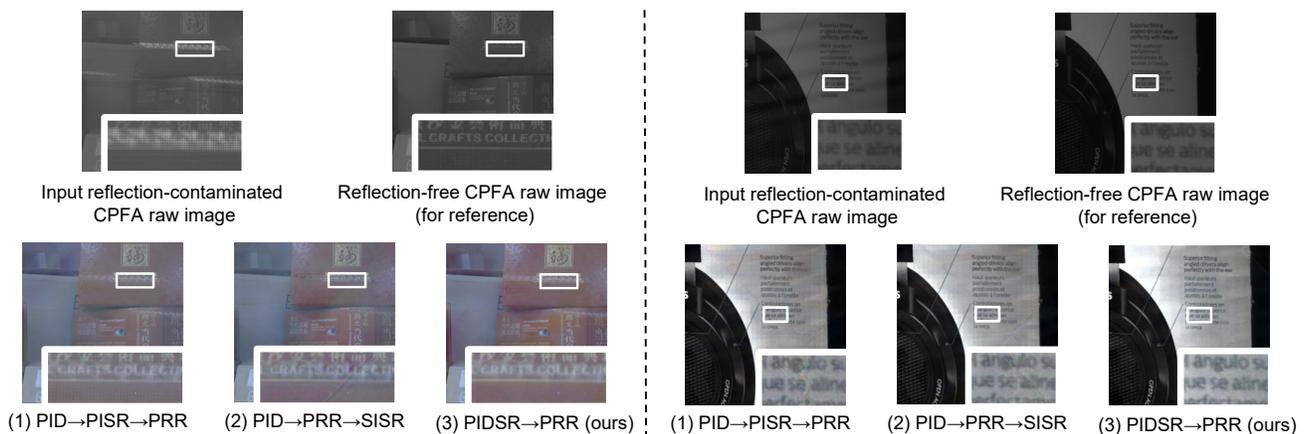
Figure E. Additional results of polarization-based reflection removal.

Xu, and Yaqin Zhou. Color polarization image super-resolution reconstruction via a cross-branch supervised learning strategy. *Optics and Lasers in Engineering*, 165:107469, 2023. 1, 3