Supplementary Material of "Joint Image Super-resolution and Low-light Enhancement in the Dark"

Feihu Zhou[1], Kan Chang[1]✉, Mingyang Ling[2], Hengxin Li[3], and Shucheng Xia[1]

[1] School of Computer and Electronic Information,
Guangxi University, Nanning, China
flyhu2@126.com,kanchang@gxu.edu.cn,xsc@st.gxu.edu.cn,
[2] School of Electrical Engineering, Guangxi University, Nanning, China
lingmy@st.gxu.edu.cn
[3] Shenzhen Xiaomi Communications Co.,Ltd, China
lihengxin@xiaomi.com

In this document, we provide additional details about the DarkSR dataset, more visual comparisons on the DarkSR dataset, as well as the quantitative and qualitative results from training various methods for two specific camera models: the Canon EOS 450D and the Nikon D700.

1 More Details about the DarkSR Dataset

The DarkSR dataset comprises 20 models from four camera brands: *Canon*, *Sony*, *Nikon*, and *Leica*. Detailed information about the models and the corresponding number of training and testing images can be found in Table 1. The testing and training sets comprise 16 and 14 camera models, respectively. Note that to evaluate generalization, the testing set includes 32 images from 6 camera models not present in the training set.

2 More Visual Results Obtained on the DarkSR Dataset

Additional visual results are presented in Figs. 1, 2, 3 and 4. It is important to note that the camera models used in Figs. 1 and 2 are present in both the training and testing sets. However, the camera models used in Figs. 3 and 4 are not included in the training set.

It is obvious that our method produces images with fine details, suitable contrast and no color distortion. Furthermore, the results in Figs. 3 and 4 illustrate that our approach generalizes effectively to unseen camera models, aligning with the findings presented for the *Xiaomi* camera in the main paper.

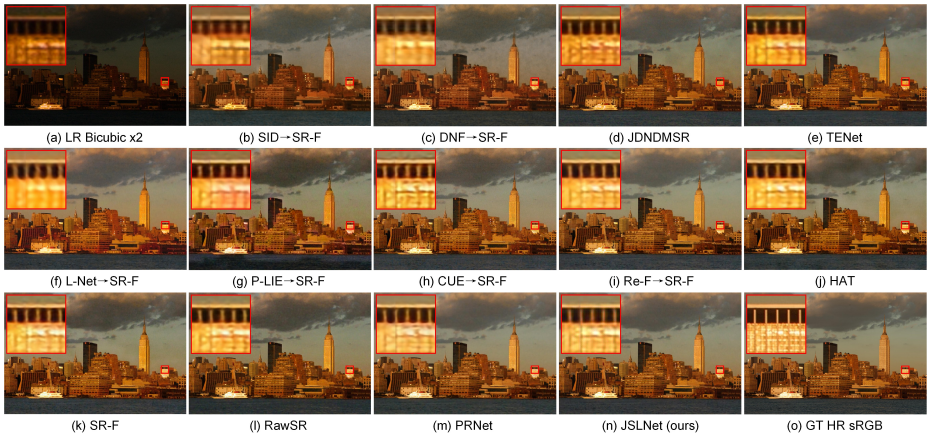

Fig. 1: Visual comparison of different methods on image *a3387*, zoom for best view.

Table 1: Detailed information of the DarkSR dataset. N_{Test} and N_{Train} denote the numbers of testing and training images, respectively.

Camera Brand	Model Name	N_{Test}	N_{Train}	Bayer Pattern
Canon	EOS 5D	29	283	RGGB
	EOS 10D	4	30	RGGB
	EOS 20D	17	170	RGGB
	EOS 30D	0	29	RGGB
	EOS 40D	6	103	RGGB
	EOS D60	4	0	RGGB
	EOS 300D	4	54	RGGB
	EOS 350D	8	0	RGGB
	EOS 400D	1	31	RGGB
	EOS 450D	19	252	RGGB
	1D Mark II	14	141	RGGB
	1D Mark II N	3	0	RGGB
	1D Mark III	8	0	RGGB
	PowerShot G9	0	20	RGGB
SONY	DSLR-A900	10	26	RGGB
LEICA	D-LUX 3	0	24	RGGB
	M8 Digital Camera	0	12	RGGB
Nikon	D200	3	0	RGGB
	D300	6	0	RGGB
	D700	17	220	RGGB

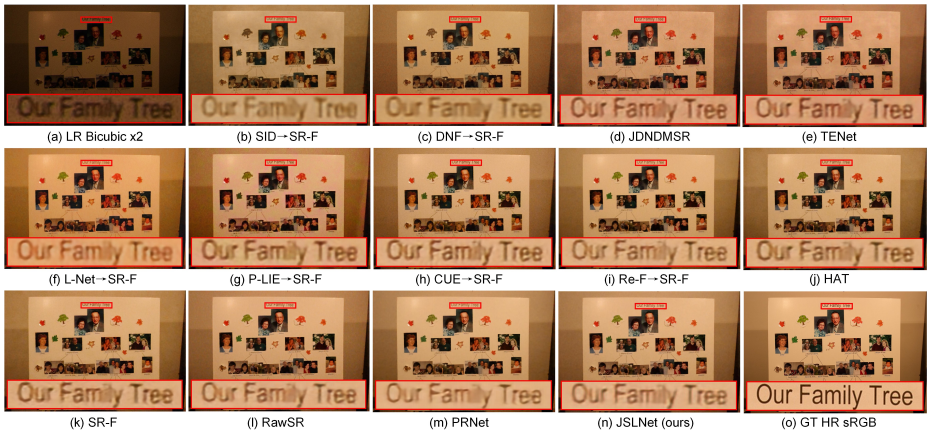

Fig. 2: Visual comparison of different methods on image *a1868*, zoom for best view.

Fig. 3: Visual comparison of different methods on image *a0475*, zoom for best view.

3 Results of Training Separate CNN Models for Two Specific Camera models

Note that training JSLNet for each camera model is unnecessary, because JSLNet can learn ISP pipeline information from the sRGB input. In contrast, Raw-based methods are camera-specific and lack generalization. In this section, for each compared method, we provide the results of training separate CNN models for two specific camera models, Canon EOS 450D and Nikon D700, leveraging the larger number of training images available for these models in the DarkSR dataset. For each camera model, to avoid overfitting, we crop each training image into four sub-images, and then use the sub-images for training. All the other experimental settings are kept the same as that in the main paper.

Fig. 4: Visual comparison of different methods on image *a3104*, zoom for best view.

Table 2: Quantitative results on two specific camera models in the DarkSR dataset. The best results are marked in red.

Methods	Input Type	Canon EOS 450D			Nikon D700		
		CPSNR ↑	SSIM ↑	ΔE ↓	CPSNR ↑	SSIM ↑	ΔE ↓
DNF→SR-F	RAW	20.846	0.7743	13.1976	21.442	0.7648	11.8165
JDnDmSR	RAW	26.941	0.8476	8.1792	22.868	0.8078	9.5585
Re-F→SR-F	sRGB	30.274	0.8552	5.9878	29.295	0.8570	6.2037
SRFormer	sRGB	30.273	0.8553	5.8817	29.391	0.8576	5.9402
PRNet	Dual-input	30.458	0.8621	5.8750	29.450	0.8639	5.8884
JSLNet(ours)	Dual-input	30.558	0.8560	5.8119	29.652	0.8662	5.8385

The quantitative results are shown in Table 2, and the visual comparisons can be found in Figs. 5 and 6, respectively. As shown, JSLNet consistently outperforms other methods across most metrics, demonstrating its superior performance. Furthermore, the visual results produced by JSLNet are remarkably close to the GT images. Despite retraining separate CNN models for each camera model, RAW-based methods still exhibit the poorest performance due to their lack of ISP pipeline information.

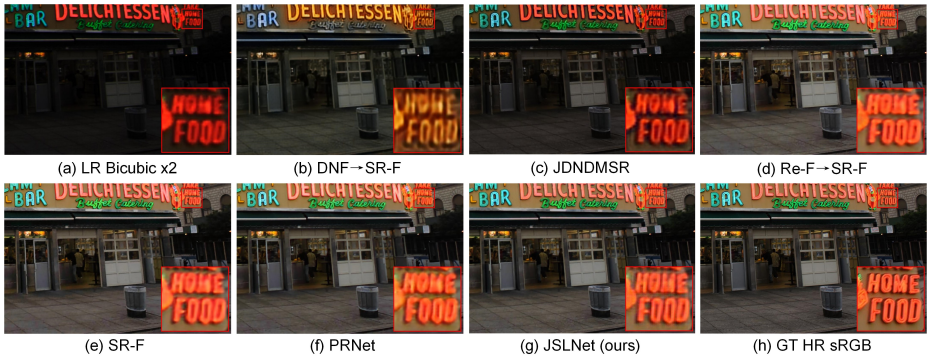

Fig. 5: Visual comparison of different methods on image *a1476* (camera model: Canon EOS 450D), zoom for best view.

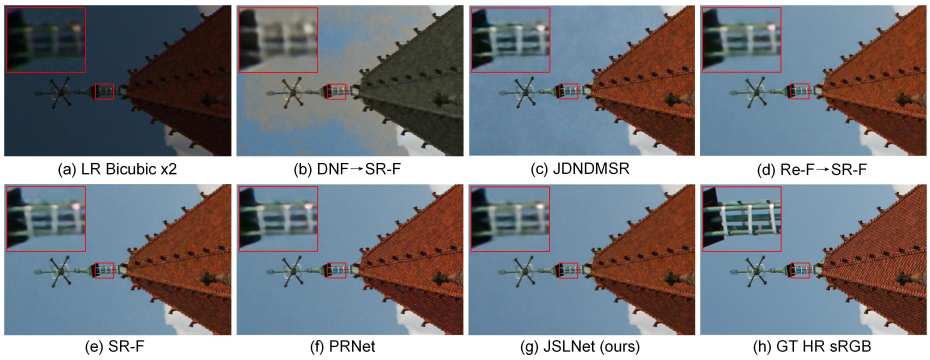

Fig. 6: Visual comparison of different methods on image *a4049* (camera model: Nikon D700), zoom for best view.