Underexposed Photo Enhancement using Deep Illumination Estimation

Ruixing Wang, Qing Zhang, Chi-Wing Fu, Xiaoyong Shen, Wei-Shi Zheng, Jiaya Jia

CVPR 2019

2020.10.27 윤주열

Underexposed Photo Enhancement

- Requires expert touch, since highly non-linear and subjective

- Segmentation, classification tend to fail for underexposed photos

(a) Input (b) Auto-Enhance on iPhone (c) Auto-Tone in Lightroom (d) Our result

Related Work

- Deep Photo Enhancer : GAN based, unpaired setting
- Distort and Recover : RL based, limited set of actions
- HDRNet : Fast & Light, 20ms on mobile

Method

- Model architecture

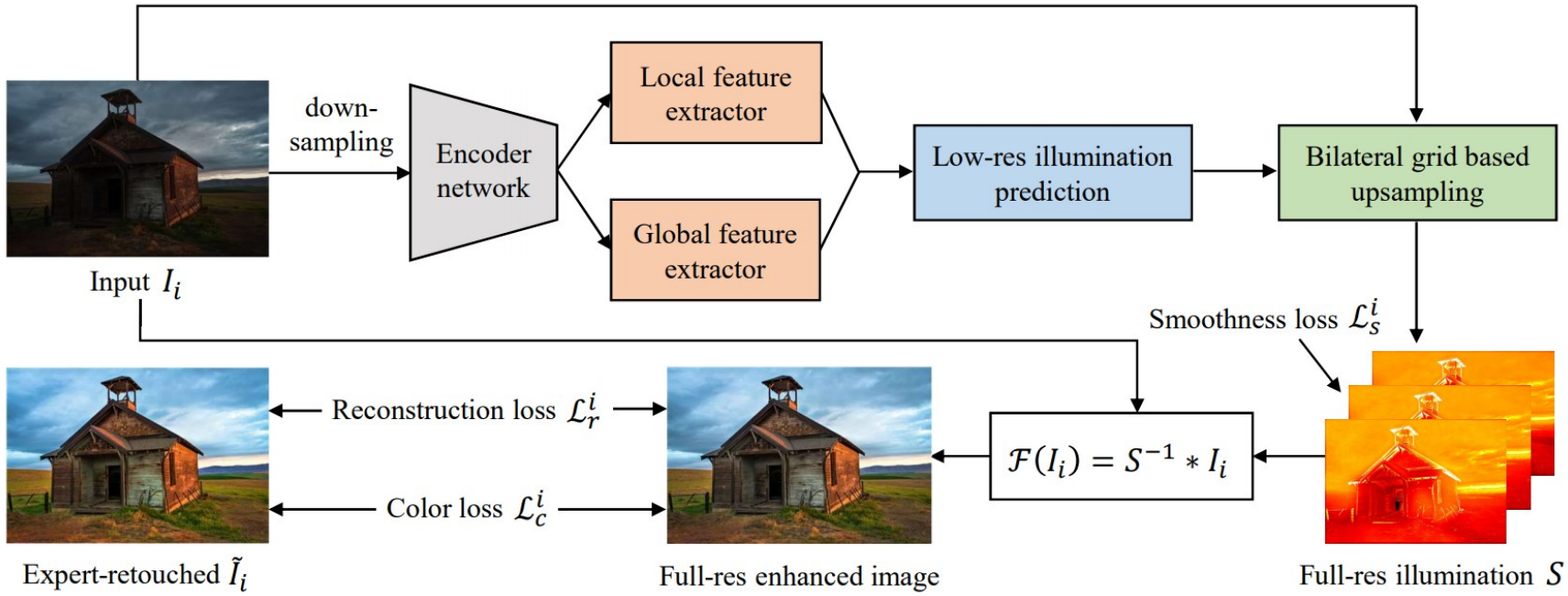
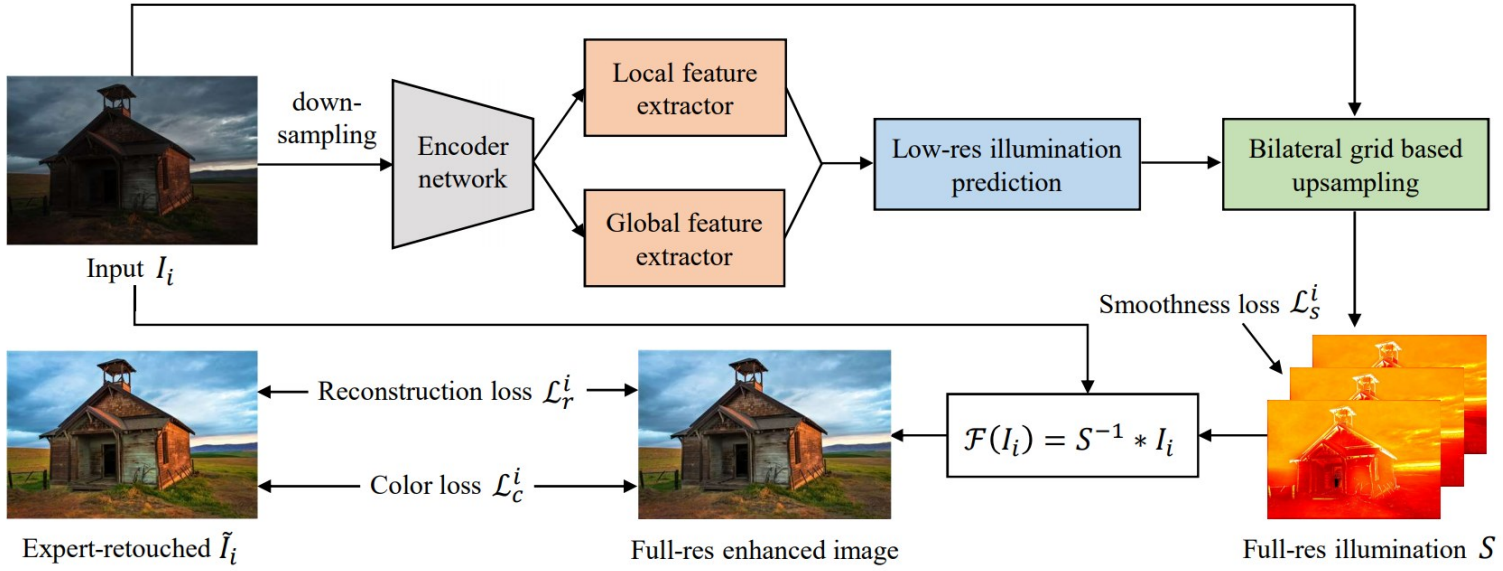
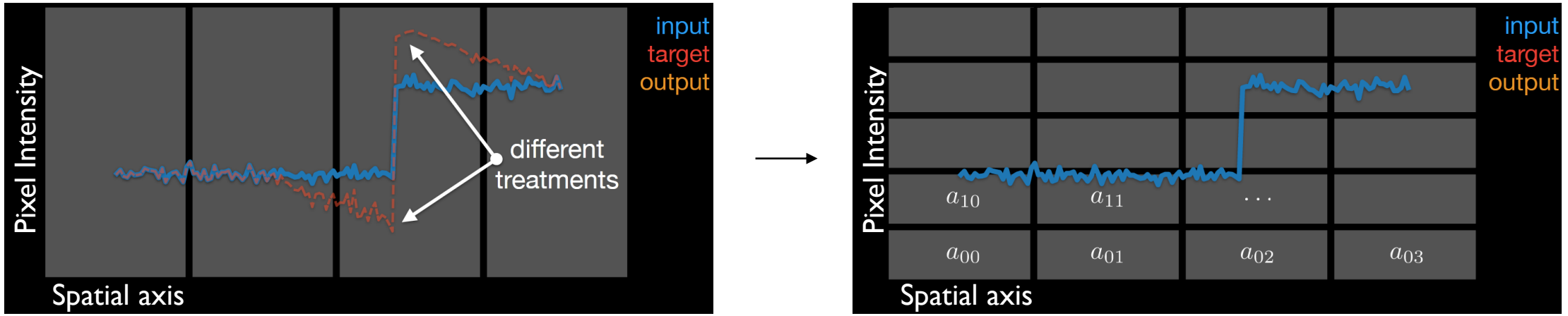


Figure 3: Overview of our network. First, we downsample and encode the input into a feature map, extract local and global features, and concatenate them to predict the low-res illumination via a convolution layer. Then we upsample the result to produce the full-res multi-channel illumination S (hot color map), and take it to recover the full-res enhanced image. We train the end-to-end network to learn S from image pairs $\{I_i, \tilde{I}_i\}$ with three loss components $\{L_r^i, L_s^i, L_c^i\}$.

Method

- Illumination Prediction

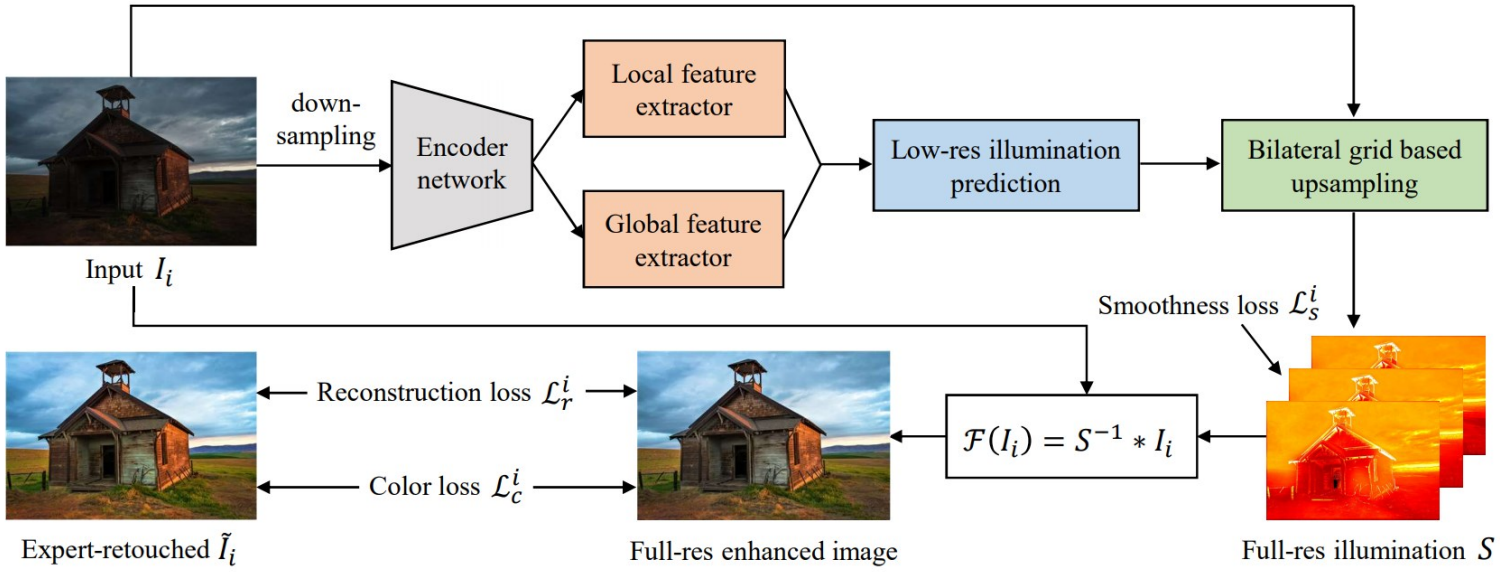




Predict illumination with Bilateral Guided filters

Interpolate filters to fit full resolution

Method



- Losses
  - Reconstruction Loss

$$\mathcal{L}_r^i = \|I_i - S * \tilde{I}_i\|^2,$$
$$s.t. \quad (I_i)_c \leq (S)_c \leq 1 \ , \quad \forall \text{ pixel channel } c$$

  - Smoothness Loss

$$\mathcal{L}_s^i = \sum_p \sum_c \omega_{x,c}^p \, (\partial_x S_p)_c^2 + \omega_{y,c}^p (\partial_y S_p)_c^2$$

$$\omega_{x,c}^p = (|\partial_x L_i^p|_c^\theta + \epsilon)^{-1} \text{ and}$$
$$\omega_{y,c}^p = (|\partial_y L_i^p|_c^\theta + \epsilon)^{-1} .$$

  - Color Loss

$$\mathcal{L}_c^i = \sum_p \angle((\mathcal{F}(I_i))_p, (\tilde{I}_i)_p)$$

Results

- Ablation study

Input | Naive Regression | $\mathcal{L}_r^i$ | $\mathcal{L}_r^i + \mathcal{L}_s^i$ | $\mathcal{L}_r^i + \mathcal{L}_s^i + \mathcal{L}_c^i$ | Expert-retouched

Figure 4: Ablation study that demonstrates the effectiveness of each component ($\mathcal{L}_r^i$, $\mathcal{L}_s^i$, and $\mathcal{L}_c^i$) in the loss function.

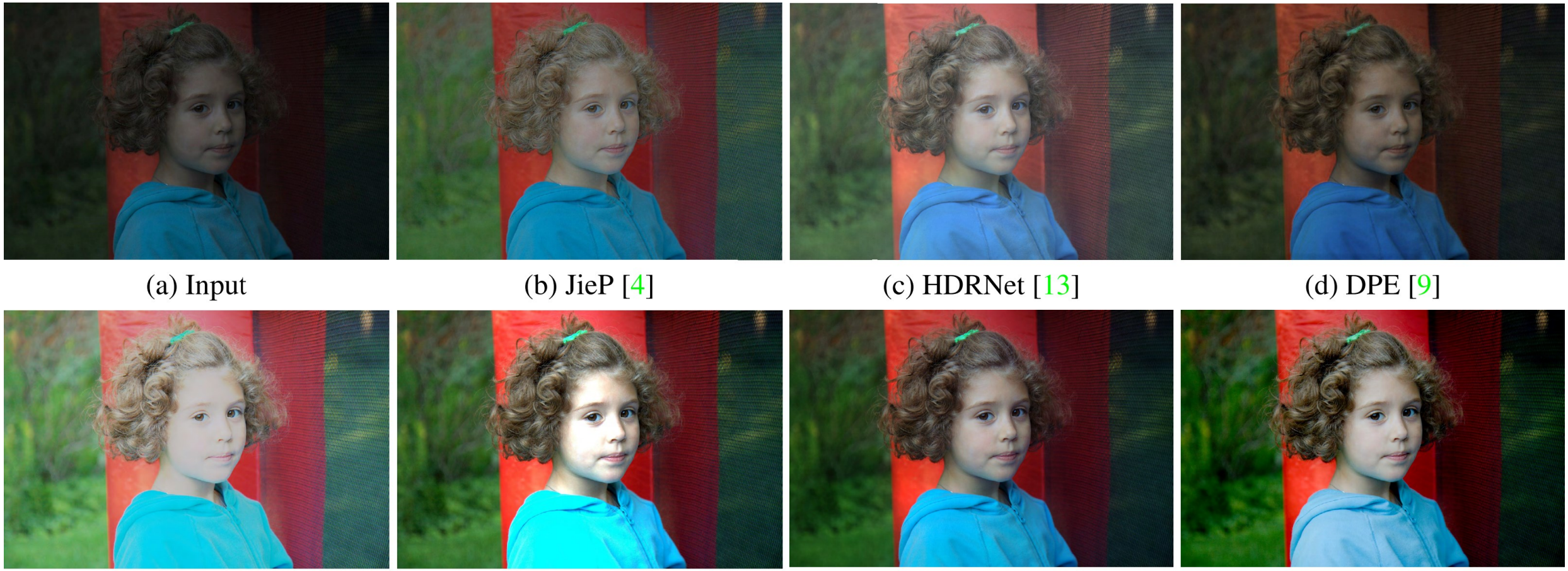
Results

(a) Input (b) JieP [4] (c) HDRNet [13] (d) DPE [9]

(e) White-box [15] (f) Distort-and-Recover [22] (g) Our result (h) Expert-retouched

Figure 6: Visual comparison with state-of-the-art methods on a test image (a) from our dataset.

Results



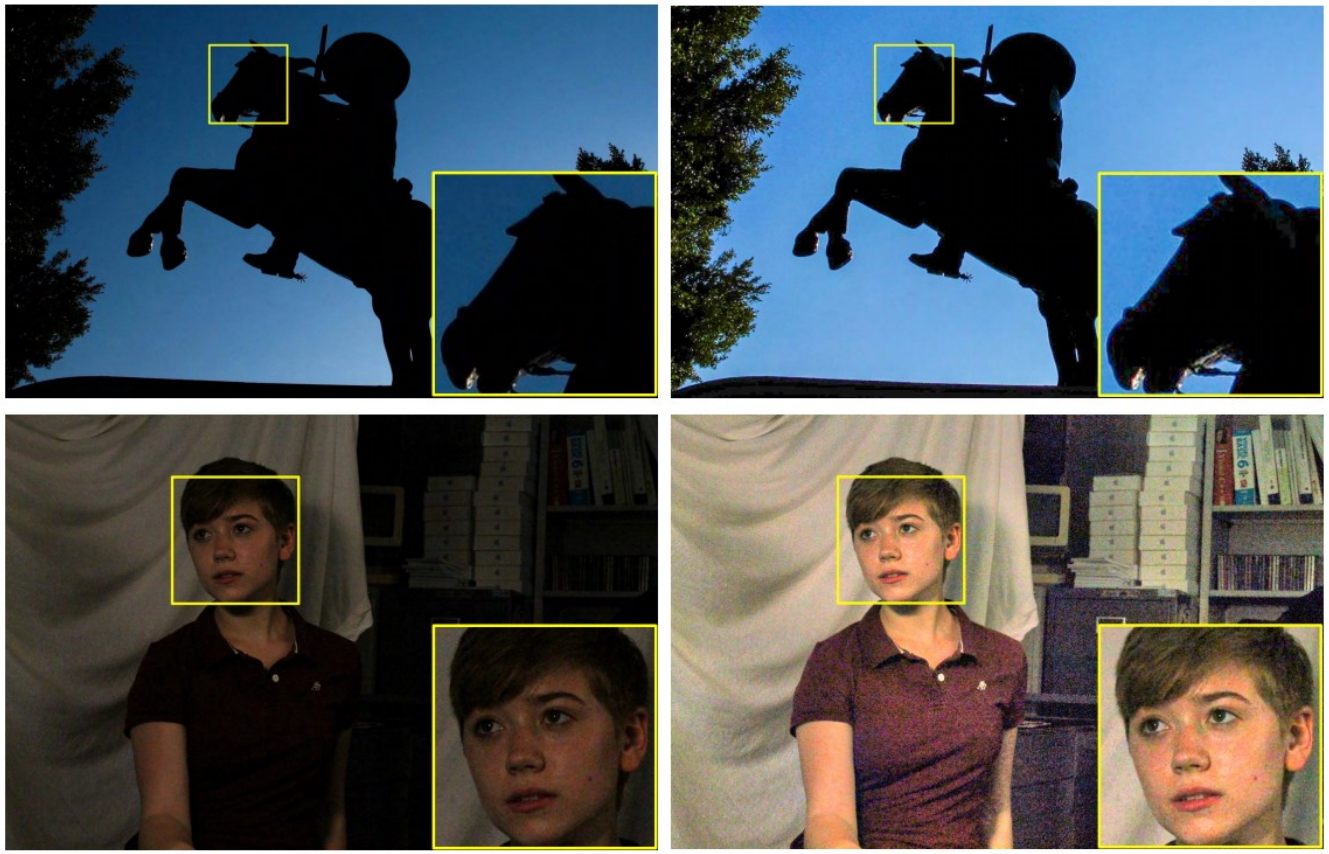
(a) Input (b) JieP [4] (c) HDRNet [13] (d) DPE [9]

(e) White-box [15] (f) Distort-and-Recover [22] (g) Our result (h) Expert-retouched

Figure 7: Visual comparison with state-of-the-art methods on a test image (a) from the MIT-Adobe FiveK [3] dataset.

Results

- Failure cases



(a) Inputs (b) Our results

Figure 10: Failure cases. Input images with mostly black regions (*top row*) and with noise (*bottom row*).

Results

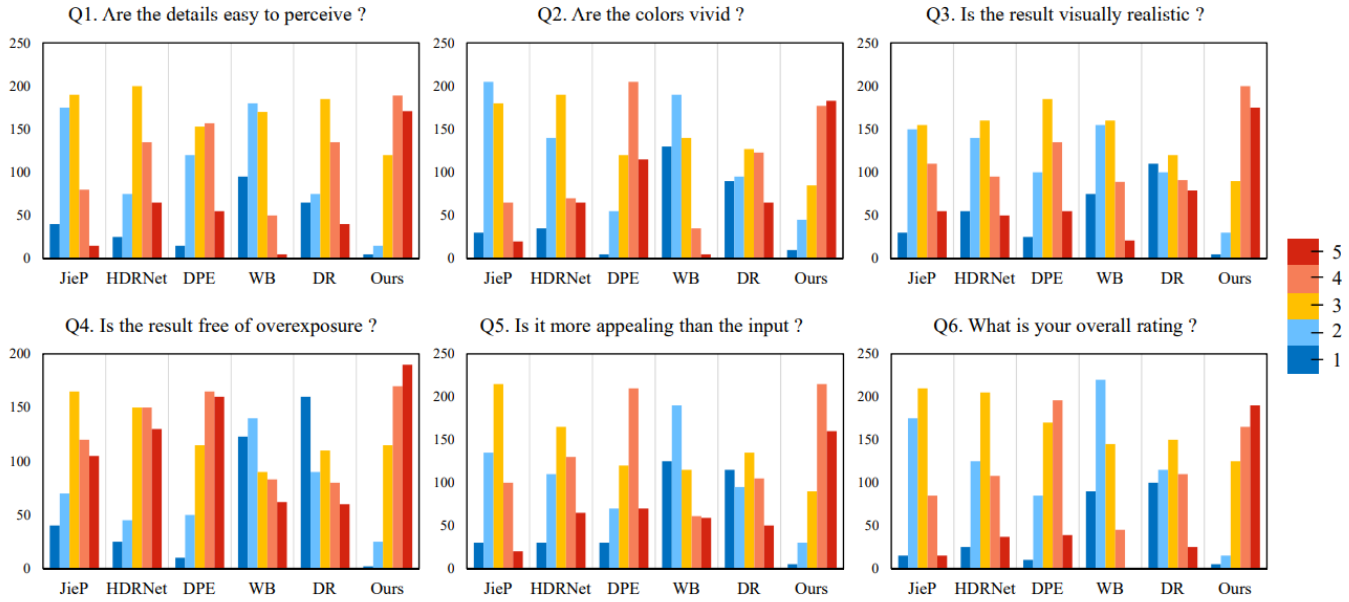


Figure 8: Rating distributions for various methods on the six questions in the user study. The ordinate axis shows the rating frequency received by the methods from the participants. WB and DR mean White-Box [15] and Distort-and-Recover [22].

Table 2: Quantitative comparison between our method and state-of-the-arts on the MIT-Adobe FiveK dataset.

Method	PSNR	SSIM
HDRNet [2]	21.96	0.866
DPE [1]	22.15	0.850
White-Box [3]	18.57	0.701
Distort-and-Recover [4]	20.97	0.841
Ours w/o $\mathcal{L}_r$, w/o $\mathcal{L}_s$, w/o $\mathcal{L}_c$	21.97	0.867
Ours with $\mathcal{L}_r$, w/o $\mathcal{L}_s$, w/o $\mathcal{L}_c$	22.31	0.871
Ours with $\mathcal{L}_r$, with $\mathcal{L}_s$, w/o $\mathcal{L}_c$	22.89	0.884
Ours	**23.04**	**0.893**