Raw-Guided Enhancing Reprocess of Low-Light Image via Deep Exposure Adjustment (supplemental material)

Haofeng Huang, Wenhan Yang, Yueyu Hu, and Jiaying Liu

Peking University, Beijing, P.R.China {huang6013,yangwenhan,huyy,liujiaying}@pku.edu.cn

Abstract. This supplemental material provides a specific network configuration of the proposed RAW-Guiding exposure time adjustment Network (RGNET), subjective evaluations of the Ablation Study and additional evaluation and analysis of RAW-based methods. We also provide extra visual results of ours and other methods for detailed comparison.

1 Specific Network Configuration

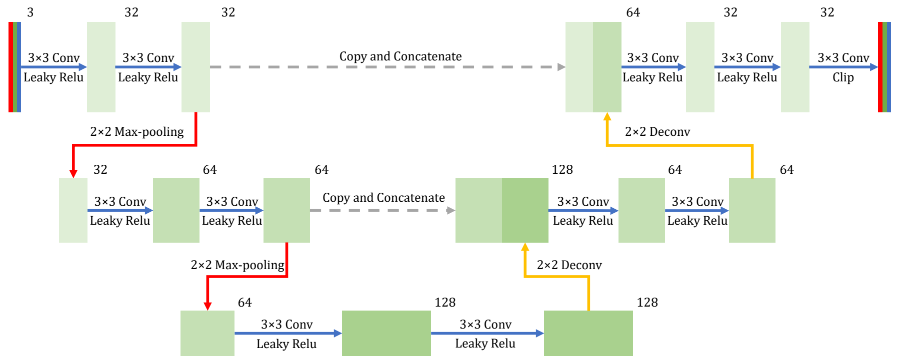
As mentioned in the paper, we construct RGNET with three U-Nets. Based on the analysis that the reconstruction is still the most challenging task, we adopt two identical three-layer U-Nets for RGNET-I and RGNET-III, and a five-layer U-Net for RGNET-II. To simulate the clipping effect, we replace the Sigmoid activation with a simple function to clip images into [0,1]. The specific configuration is shown as Fig. 1

In the training phase, RGNET-I converges after 1500 epochs, note that the L_1 loss is calculated after multiplying the ratio, RGNET-II converges after 2400 epochs and RGNET-III after 2000 epochs, that approximately denotes the difficulty of three tasks to some extent.

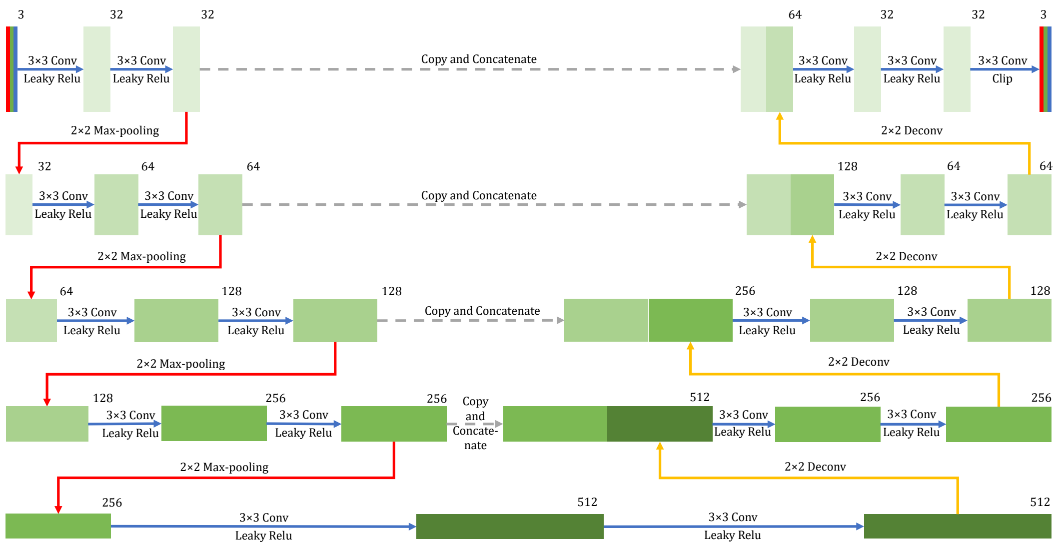
2 Subjective Evaluation in Ablation Study

With the paper's constraints of space, we show visual results of controlled experiments here. In Fig. 2 and Fig. 3 :

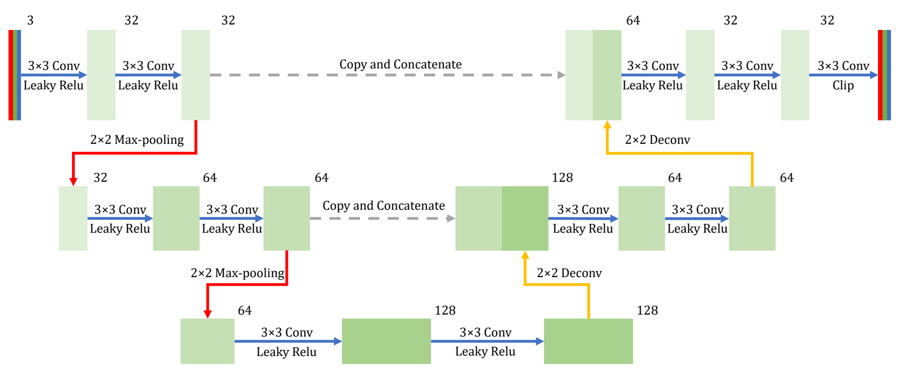
- "w/o RAW" denotes an end-to-end trained U-Net with low/normal-light processed RGB image pairs.
- "with MS-SSIM" denotes the identical architecture with "w/o RAW", multi-scale SSIM loss added, whose weight is set to 0.5 .
- "with mean adjusted" denotes that inputs' values multiply by the ratio of intensities of image pairs, and then the U-Net follows.
- "with original RAW" means the linear process is not adopted by the RGNET, *i.e.* the RGNET-I and RGNET-III target to invert and simulate the whole processing pipeline.
- The inputs' and targets' exposure time is shown left.



(a) RGNET-I

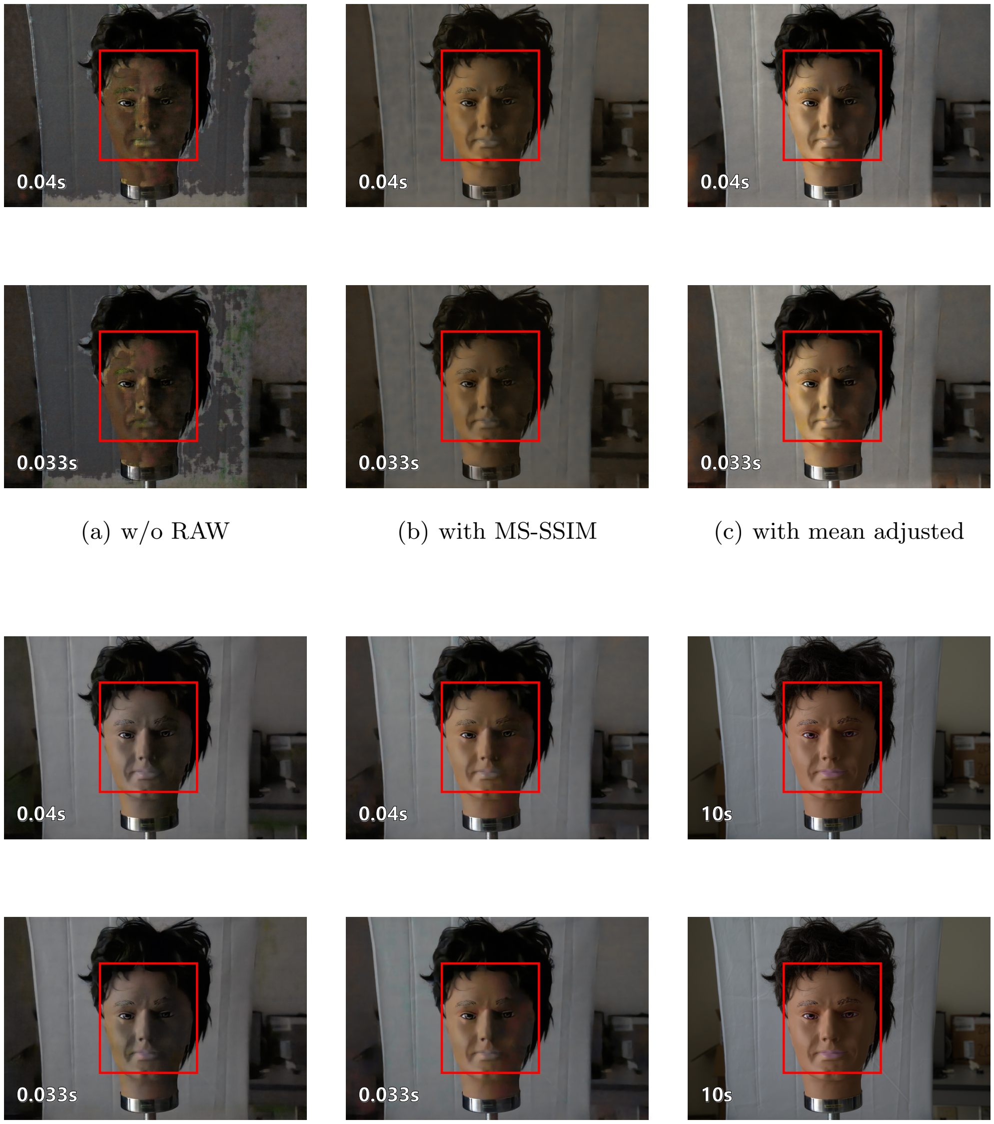


(b) RGNET-II



(c) RGNET-III

Fig. 1. The details of RGNET architecture.

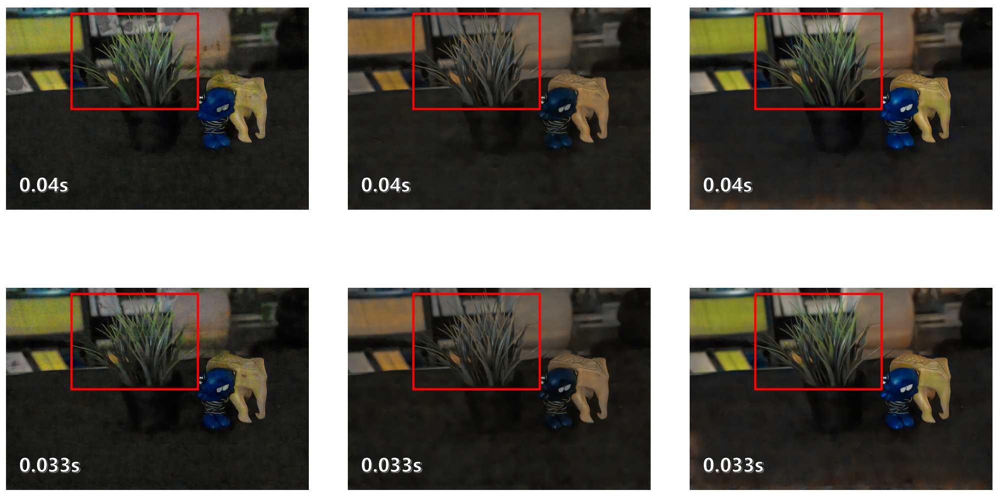


(a) w/o RAW (b) with MS-SSIM (c) with mean adjusted

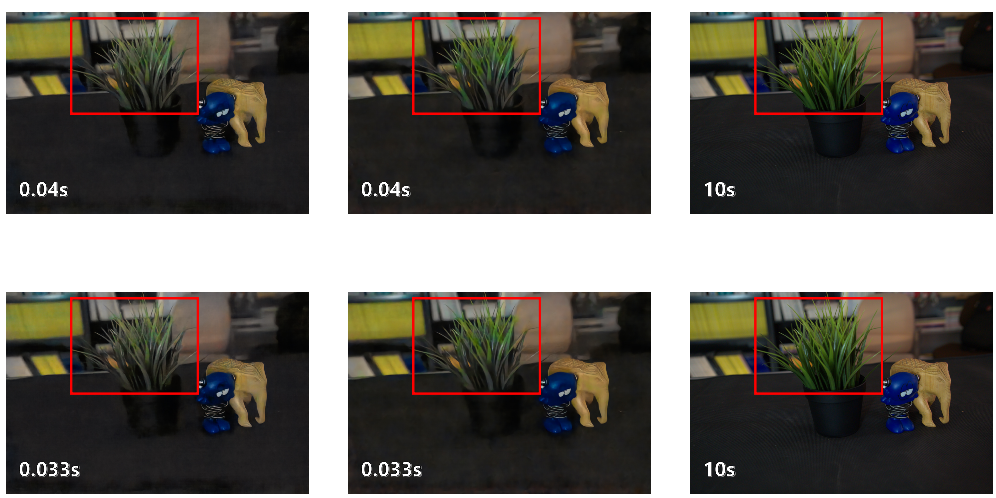
(d) with original RAW (e) Ours (f) Ground Truth

Fig. 2. Qualitative results of Ablation Study.

Apparently, without a label indicating the extent of brightening and only with L_1 loss, "w/o RAW" brightens different areas nonuniformly to make sure there are always pixels with the right brightness. Multi-scale SSIM loss partly lessens artifacts but still cannot handle low light images with different illumination, *i.e.* shorter-exposure images' outputs are always darker, although they correspond to an identical target. After adjusting images' intensities, the network dynamically outputs images with approximately correct illumination but low contrast. "with original RAW", as well as all these RGB-based architecture, cannot handle the color cast properly because the significant meta-data for white balance is not utilized by these methods.



(a) w/o RAW (b) with MS-SSIM (c) with mean adjusted



(d) with original RAW (e) Ours (f) Ground Truth

Fig. 3. Qualitative results of Ablation Study.

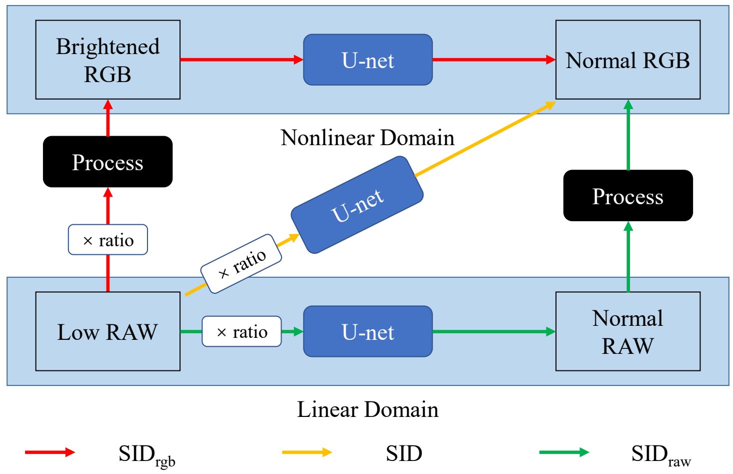


Fig. 4. Pipelines of three discussed RAW-based methods. $\mathrm{SID_{rgb}}$ denotes that the reconstruction stage is in RGB's nonlinear domain and $\mathrm{SID_{raw}}$ in RAW's linear domain.

3 Evaluation in Effect of RAW in Enhancement

In this part we discuss RAW-based methods and take a glance at how RAW images benefit the enhancement. Due to the lack of low light enhancement methods with RAW images as input besides SID, we construct two similar pipelines as competing methods for evaluation as shown in Fig. 4. For fairness, we adopt the identical network structure with SID.

Table 1. Quantitative results of discussed RAW-based methods. "$\mathrm{SID_{raw}}$" denotes that results of the network are processed with meta-data recorded by normal-light RAW files and "$\mathrm{SID_{raw}}$ with low meta-data" means process uses low-light RAW files' inaccurate meta-data. Note that the former pipeline is inaccessible in the application.

Method	PSNR
$\mathrm{SID_{rgb}}$	28.36
SID	28.63
$\mathrm{SID_{raw}}$	29.50
$\mathrm{SID_{raw}}$ with low meta-data	28.20

RAW images possess two significant features we value *i.e.* linearity and recording meta-data for image processing systems *e.g.* weights for white balance and the CCM matrix. With evaluation results shown in Table 1, although $\mathrm{SID_{rgb}}$ conduct process with meta-data, reconstructing in the nonlinear domain drops more than gain from the process, compared with SID. On the other hand, fully utilizing linearity and meta-data, $\mathrm{SID_{raw}}$ outperforms SID. To be strict, during the testing phase meta-data for "Normal RAW" is unknown. To make up, directly using unmatched meta-data for “Low RAW” instead induces biases in process and finally a drop in PSNR as shown in Table 1 and Fig. 5. This gap may be narrowed via learning-based methods but it’s out of our discussion.

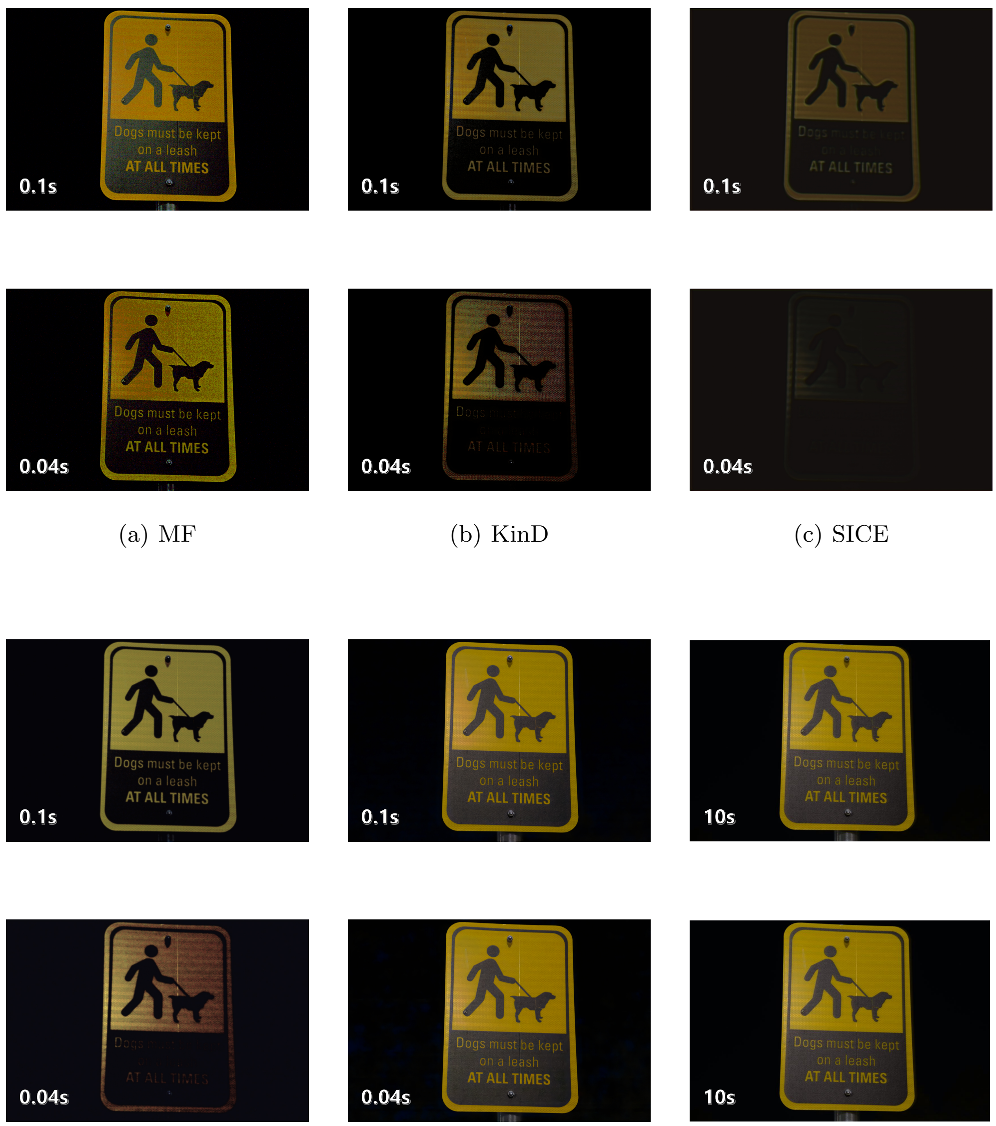
(a) $\mathrm{SID}_{\mathrm{raw}}$ with low meta-data (b) $\mathrm{SID}_{\mathrm{raw}}$ (c) Ground Truth

Fig. 5. Qualitative results to show how inaccurate meta-data affects the image processing system.

These experiments additionally show the significance of linearity and meta-data, and provide extra reasons for the proposed architecture. Note that all these methods are inaccessible without RAW images and show an approximate upper bound for RGB-based methods.

4 Extra Visual Results for Comparison

We choose five methods with the best PSNR scores, including MF, KinD, SICE, LLNet and ours, and provide extra visual results of them for detailed comparison.

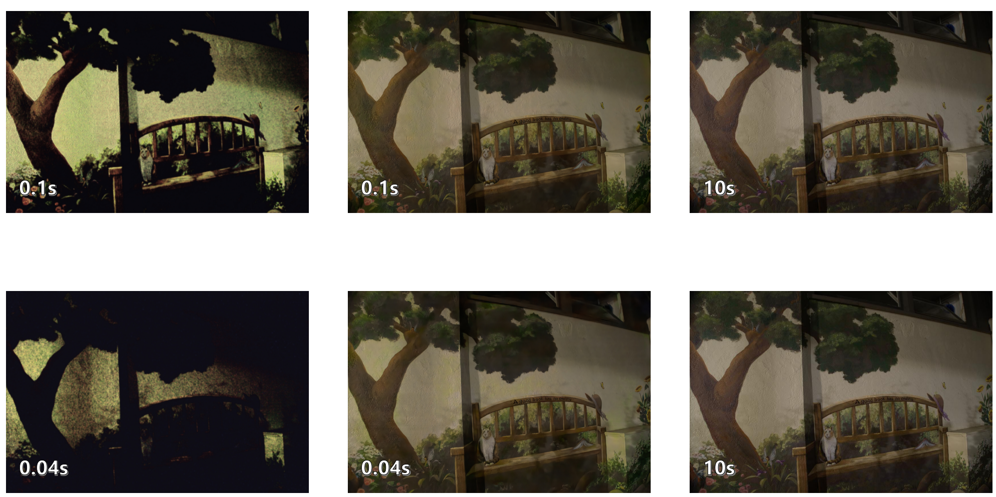


(a) MF (b) KinD (c) SICE

(d) LLNet (e) Ours (f) Ground Truth

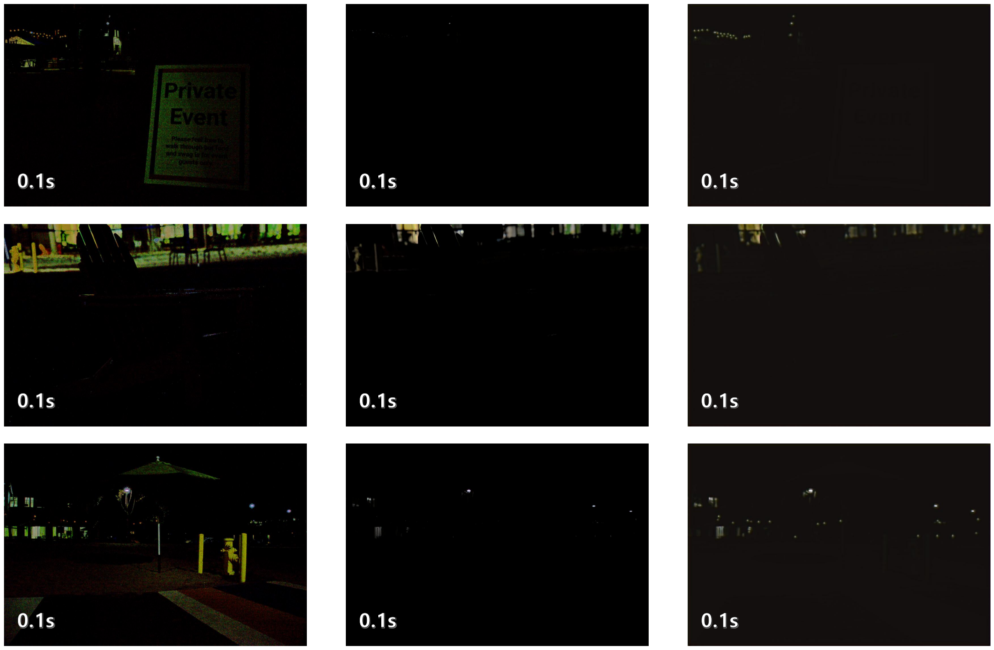
Fig. 6. Qualitative results of chosen methods.

(a) MF (b) KinD (c) SICE



(d) LLNet (e) Ours (f) Ground Truth

Fig. 7. Qualitative results of chosen methods.



(a) MF (b) KinD (c) SICE

(d) LLNet (e) Ours (f) Ground Truth

Fig. 8. Qualitative results of chosen methods.