A. Display of Object Detection Results in Real-World Situation

In this section we would give more examples on real data sets. Fig.A1 shows a case in UG^2+ DARK FACE [11] dataset. Fig.A2 shows the detection results of ExDark [8] dataset. Our MAET has shown better detection performance when facing actual dark light detection tasks.

B. Ablation Study

B.1. Other Low-Light Image Synthesis Methods

Although the purpose of this paper is not for a perfect low-light image synthesis pipeline, we have also compared with existing low light image synthesis methods (from normal light sRGB to low light sRGB) and evaluate their impact on low-light object detection tasks.

The work in [6] proposed to use the Retinex model [5] to generate low light counterpart by normal light images:

$$I(x) = R(x) \cdot L, \tag{1}$$

in this equation, $R(x)$ is the clear normal-lit images (same as x in Eq.15), $I(x)$ is the generated low-lit counterpart (same as $t_{deg}(x)$ in Eq.15) and L is the random fixed illumination value, here L, same as our parameter k's range.

The work in [9, 10, 13] proposed to use an invert gamma correction with additional noises to generate low light degraded image from normal light counterpart, the equation shown as follow:

$$t_{deg}(x) = x^{\gamma} + n, \tag{2}$$

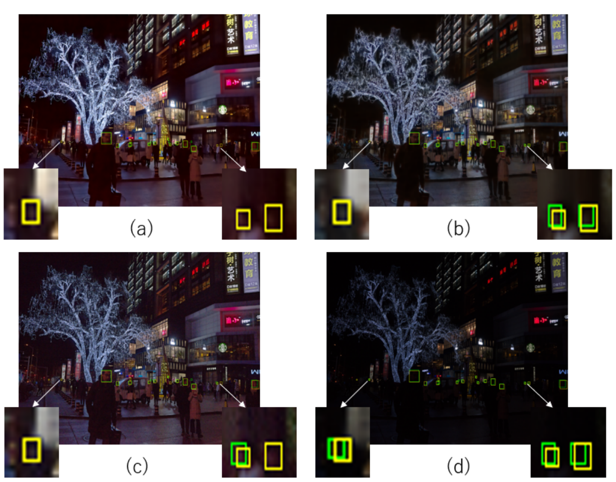

Figure A1. Detection results of UG^2+ DARK FACE dataset [11]. (a)/(b)/(c) is the detection result of YOLO based on the dataset preprocessed by MBLLEN [10]/KIND [12]/Zero-DCE [3] and (d) is the is the detection result of MAET-YOLO model on original dataset, yellow and green boxes are ground truth boxes and prediction results, respectively.

		ExDark	DARK FACE
YOLO (L)	L_R	0.698	0.511
	L_G	0.709	0.528
	L_{GP}	0.712	0.532
	L_{Gmix}	0.713	0.535
	$\boldsymbol{L_{ours+m}}$	0.706	0.530
	$\boldsymbol{L_{ours}}$	**0.716**	**0.540**

Table B1. The experiment results on ExDark [8] dataset and UG^{2+} DARK FACE [11] datasets by using different kinds of synthetic low light COCO [7] dataset. The detection results verify the reliability of our synthesis method.

here γ is the gamma curve parameter and n is the additional noise (Poisson noise or Gaussian-Poisson noise model).

B.2. Demosacing's Influence

Demosacing is an essential part in camera image signal processing pipeline [1, 4, 2], which aims to recover intermediate gray-scale image to the R/G/B value by interpolating the missing values in Bayer pattern. Unlike the previous work [2], we ignore this step for simplicity in our ISP procedure. In supplementary material, we show an example after adding mosaicing process after invert WB process ((d) in Fig.3) and demosacing process after WB process ((f) in Fig.3).

To evaluate effects of different low light data generation methods: Retinex based generation method L_R, inverse gamma curve L_G, inverse gamma curve with additional possion noise L_{GP}, inverse gamma curve with additional mixed Gaussian-Poisson noise model L_{Gmix}, our proposed data generation method in Sec.3.3 L_{ours}, and our proposed data generation method with mosaicing and demosacing process L_{ours+m}. We measured the performance of using different dark light data on the real world datasets [8, 11], shown as YOLO (L) in Table B1, the training configurations and strategies are same as Sec.4, it could be seen that our synthetic method is of greatest help in improving the detection performance of real datasets [8], [11].

References

[1] Chapter 5 - comparison of color demosaicing methods. In Peter W. Hawkes, editor, *Advances in Imaging and electron Physics*, volume 162 of *Advances in Imaging and Electron Physics*, pages 173–265. Elsevier, 2010.

[2] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T Barron. Unprocessing images for learned raw denoising. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 11036–11045, 2019.

[3] C. Guo, C. Li, J. Guo, C. C. Loy, J. Hou, S. Kwong, and R. Cong. Zero-reference deep curve estimation for low-light image enhancement. In *2020 IEEE/CVF Conference*