PE-YOLO: Pyramid Enhancement Network for Dark Object Detection

Xiangchen Yin[1,2]($^{(\boxtimes)}$), Zhenda Yu[2,3], Zetao Fei[4], Wenjun Lv[1], and Xin Gao[5]

[1] University of Science and Technology of China, Hefei, China
[2] Institute of Artificial Intelligence, Hefei Comprehensive National Science Center,
Hefei, China
`yinxiangchen@mail.ustc.edu.cn, wlv@ustc.edu.cn`
[3] Anhui University, Hefei, China
`wa22201140@stu.ahu.edu.cn`
[4] Qufu Normal University, Qufu, China
`feizetao@163.com`
[5] China University of Mining & Technology, Beijing, China
`bqt2000405024@student.cumtb.edu.cn`

Abstract. Current object detection models have achieved good results on many benchmark datasets, detecting objects in dark conditions remains a large challenge. To address this issue, we propose a pyramid enhanced network (PENet) and joint it with YOLOv3 to build a dark object detection framework named PE-YOLO. Firstly, PENet decomposes the image into four components of different resolutions using the Laplacian pyramid. Specifically we propose a detail processing module (DPM) to enhance the detail of images, which consists of context branch and edge branch. In addition, we propose a low-frequency enhancement filter (LEF) to capture low-frequency semantics and prevent high-frequency noise. PE-YOLO adopts an end-to-end joint training approach and only uses normal detection loss to simplify the training process. We conduct experiments on the low-light object detection dataset ExDark to demonstrate the effectiveness of ours. The results indicate that compared with other dark detectors and low-light enhancement models, PE-YOLO achieves the advanced results, achieving 78.0% in mAP and 53.6 in FPS, respectively, which can adapt to object detection under different low-light conditions. The code is available at https://github.com/XiangchenYin/PE-YOLO.

Keywords: Object detection · Low-light perception · Pyramid enhancement ·

1 Introduction

In recent years, the emergence of convolutional neural networks (CNNs) has promoted the development of object detection. A large number of detectors have been proposed, and the performance of the benchmark datasets is getting

enjoyable results [1,8,14,19]. However, most of the existing detectors are studied in high-quality images under normal conditions. In the real environment, there are often many bad lighting conditions such as night, dark light, and exposure, so that the quality of the image decreases affects the performance of the detector. The visual perception model enables the automatic system to understand the environment and lay the foundation for subsequent tasks such as trajectory planning, which requires a robust object detection or semantic segmentation model. Fig. 1 is an example of dark object detection. It can be found that if the image is appropriately enhanced and restores more potential information of the original fuzzy object according to environmental conditions, the object detection model is adapted to different low-light conditions, which is also a great challenge in the practical application of the model.
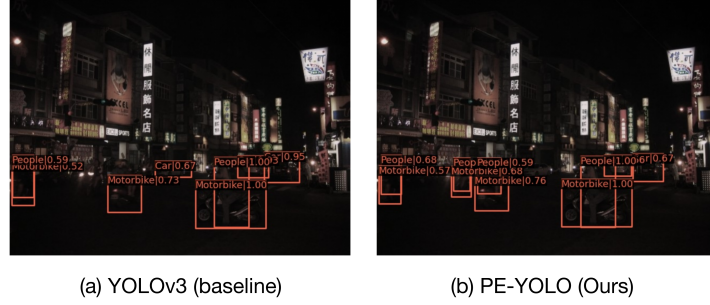

(a) YOLOv3 (baseline)   (b) PE-YOLO (Ours)

Fig. 1. Example of dark object detection. In dark conditions PE-YOLO can recover more potential information of object to get better detection results.

Currently, many methods have been proposed to solve the robustness problem in the dark scenes. Many low-light enhancement models [7,10,24,26] have been proposed to restore image details and reduce the impact of poor lighting conditions. However, the structure of the low light enhancement model is complex, which is not conducive to the real-time performance of the detector after image enhancement. Most of these methods cannot be end-to-end trained with the detector, and supervised learning is required for paired low-light images and normal images. Object detection under low-light conditions can also be seen as a domain adaptation problem. Some researchers [4,13,21] have used adversarial learning to transfer the model from normal light to dark light. But they focus on matching data distribution and overlook the potential information contained in low-light images. In the past few years, some researchers [11,15] have proposed the method of using differentiable image processing (DIP) modules to enhance images and train detectors on in an end-to-end manner. However, DIP are traditional methods such as white balance, which have limited enhancement effects on images.

To address the above issues, we propose a pyramid enhancement network (PENet) that enhance low-light images and capture potential information about objects. We have joint PENet with YOLOv3 to construct an end-to-end dark object detection framework called PE-YOLO. In PENet, we first decompose the image into multiple components of different resolutions using the Laplacian pyramid. In each scale of the pyramid, we propose detail processing module (DPM) and low-frequency enhancement filter (LEF) to enhance the components. DPM consists of context branch and edge branch, which context branch globally enhance components by capturing long-range dependencies and edge branch enhance the texture of components. LEF uses a dynamic low-pass filter to obtain low-frequency semantics and prevent high-frequency noise to enrich feature information. We only use normal detection loss during model training to simplify the training process, without the need for clear ground truth of the image. We validated the effectiveness of our method in the low-light object detection dataset ExDark [16], the results show that compared with other dark detectors and low-light enhancement models, PE-YOLO achieved the advanced results, reaching 78.0% in mAP and 53.6 in FPS respectively, which can adapt to object detection in dark conditions.

Our contribution could be summarised as follow:

- We build a pyramid enhancement network (PENet) to enhance different low-light images. We propose a detail processing module (DPM) and a low-frequency enhancement filter (LEF) to enhance the components.

- By jointing PENet with YOLOv3, we propose an end-to-end trained dark object detection framework PE-YOLO to adapt the dark conditions. During training, we only use normal detection loss.

- Compared with other dark detectors and low light enhancement models, our PE-YOLO achieved advanced results in ExDark dataset, achieving enjoyable accuracy and speed.

2   Related Work

2.1   Object Detection

Object detection models are divided into three categories: one-stage models, two-stage models, and anchor-free-based models. Faster RCNN [20] does not obtain region recommendations through selective search, but rather through a region proposal network (RPN). It enables candidate region proposals, feature extraction, classification, and regression to be trained end-to-end within the same network. Cai et al. propose Cascade RCNN [2], which cascades multiple detection heads, and the current level will refine the regression and classification results of the previous level. YOLOv3 [19] proposed the new feature extraction network DarkNet-53. Inspiring from the feature pyramid nework (FPN), YOLOv3 adopts multi-scale feature fusion. In addition, recently anchor-free-based detectors [12, 23] have appeared, they abandoned anchor and changed it to key point-based detection.

2.2   Low-light Enhancement

The goal of low-light enhancement tasks is to improve human visual perception by restoring image details and correcting color distortion and to provide high-quality images for high-level visual tasks such as object detection. Zhang et al. [26] propose Kind, it can be trained through paired images with different levels of illumination, without the need for ground truth. Guo et al. [9] propose Zero DCE, which transforms low-light enhancement tasks into image-specific curve estimation problems. Lv et al. [17] propose a multi-branch low light enhancement network (MBLLEN), which extracts features at different levels and generates output images through multi branch fusion. Cui et al. [5] propose an Illumination Adaptive Transformer (IAT), through dynamic query learning to construct an end-to-end Transformer. After the low-light enhancement model restores the details of the image, the effect of the detector is improved. However, most low-light enhancement models are complex and have a great impact on the real-time performance of the detector.

2.3   Object Detection in Adverse Condition

Object detection under adverse conditions is crucial for the robust perception of robots, and robust object detection models have emerged for some adverse conditions. Some people transfer detectors from the source domain to the target domain through unsupervised domain adaptation [4,13,21], adapting the model to harsh environments. Liu et al. [15]propose IA-YOLO, which adaptively enhances each image to improve detection performance. They propose a differentiable image processing (DIP) module for harsh weather and used a small convolutional neural network (CNN-PP) to adjust the parameters of DIP. On the basis of IA-YOLO, Kalwar et al. [11] propose GDIP-YOLO. GDIP proposes a gating mechanism that allows multiple DIPs to operate in parallel. Qin et al. [18] propose detection-driven enhancement network (DENet) is used for object detection in adverse weather condition. Cui et al. [6] propose a multi-task automatic encoding transform (MAET) for dark object detection, exploring the potential space behind lighting conversion.

3   Method

Dark images have poor visibility due to low-light interference, which affects the performance of the detector. To address this issue, we propose a pyramid enhanced network (PENet) and joint YOLOv3 to construct a dark object detection framework PE-YOLO. The overview of the framework of PE-YOLO is shown in Fig. 2.

3.1   Overview of PE-YOLO

PENet decomposes the image into components of different resolutions through the Laplacian pyramid. In PENet, we enhance the components of each scale

through proposed detail processing module (DPM) and low-frequency enhancement filter (LEF).

Define the image $I \in R^{h \times w \times 3}$ as input, we obtain sub images of different resolutions using a Gaussian pyramid.

$$G(x) = Down(Gaussian(x)) \tag{1}$$

where **Down** represents downsampling, **Gaussian** represents Gaussian filter, and the size of Gaussian kernel is 5×5. After each Gaussian pyramid operation, the width and height of the image are halved, which means the resolution is the original $\frac{1}{4}$. Obviously, the downsampling operation of the Gaussian pyramid is irreversible. In order to recover the original high-resolution image after upsampling, the lost information is required, and the lost information constitutes the components of the Laplacian pyramid. The definition of the Laplacian pyramid is

$$L_i = G_i - Up(G_{i+1}) \tag{2}$$

among which L_i is the ith layer of the Laplacian pyramid, G_i represents the ith layer of the Laplacian pyramid, and **Up** represents bilinear upsampling operation. When reconstructing the image, we only need to perform the reverse operation of (2) to restore the high-resolution image.
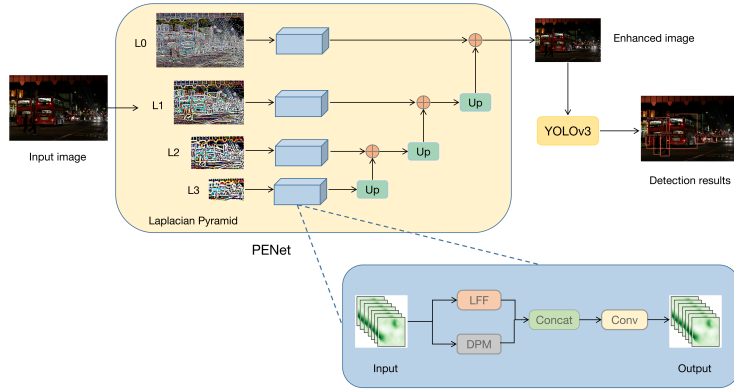

Fig. 2. Overview of PE-YOLO. We use detail processing module (DPM) and low-frequency enhancement filter (LEF) to enhance the images.

We obtained four components of different scales through the Laplace pyramid, as shown in Fig. 3. We found that the Laplacian pyramid pays more attention to global information from bottom to top, while on the contrary it pays more attention to local details. They are all information lost during the image downsampling process, which is also the object of our PENet enhancement. We enhance the components through detail processing module (DPM) and low-frequency enhancement filter (LEF), and the operations of DPM and LEF are

parallel. We will provide introduction to DPM and LEF later in next section. By decomposing and reconstructing the Laplacian pyramid, PENet can be made lightweight and effective, which helps to improve the performance of PE-YOLO.
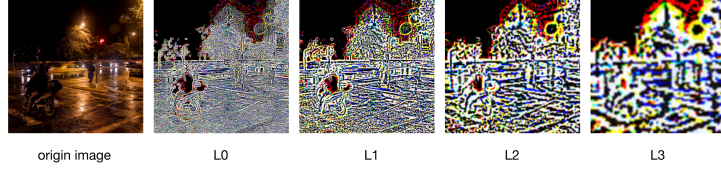

origin image   L0   L1   L2   L3

Fig. 3. Visualization of each layer in the Laplacian pyramid. The second to fourth column is the component of the Laplacian pyramid, and the resolution is reduced from left to right.

3.2   Detail Enhancement

We propose a detail processing module (DPM) to enhance the components in the Laplacian pyramid, which is divided into contextual branch and edge branch. The details of DPM are shown in Fig. 4. Context branch obtains contextual information by capturing remote dependencies, and globally enhances components. The edge branch uses two Sobel operators in different directions to calculate image gradients to obtain edges and enhance the texture of the components.

Context branch. We use a residual block to process features in before and after obtaining remote dependencies, and residual learning allows rich low-frequency information to be transmitted through skip connections. The first residual block changes the channel of the feature from 3 to 32, and the second residual block changes the channel of the feature from 32 to 3. Capturing global information in the scene has been proven to be beneficial for low-level visual tasks such as low-light enhancement. The structure of the context branch is described in Fig. 4, which is defined as

$$CB(x) = x + \gamma(F_1(\hat{x})) \tag{3}$$

where $\hat{x} = \sigma(F_2(x)) \cdot x$, F is the convolutional layer with kernel 3×3, γ is Leaky ReLU, and σ is the Softmax function.

Edge branch. Sobel operator is a discrete operator that uses both Gaussian filter and differential derivation. It can find edges by calculating gradient approximation. We use Sobel operators in both the horizontal and vertical directions to re-extract edge information through convolutional filters and use residuals to enhance the flow of information. This process is represented as

$$EB(x) = F_3(Sobel_h(x) + Sobel_w(x)) + x \tag{4}$$

where $Sobel_h$ and $Sobel_w$ represents Sobel operations in the vertical and horizontal directions, respectively.

3.3 Low-Frequency Enhancement Filter

In each scale component, the low-frequency component has most of the semantic information in the image, and they are the key information for the detector prediction. To enrich the semantics of the reconstructed image, we propose low-frequency enhancement filter (LEF) to capture low-frequency information in the components. The details of LEF are shown in Fig. 5. Assuming the component $f \in R^{h \times w \times 3}$, we first transform it into $f \in R^{h \times w \times 32}$ through a convolutional layer. We use a dynamic low-pass filter to capture low-frequency information, and we use average pooling for feature filtering, which only allows information below the cutoff frequency to pass through. The low-frequency thresholds for different semantics are different. Considering the multi-scale structure of Inception [22], we used adaptive average pooling with sizes of 1×1, 2×2, 3×3, 6×6, and used upsampling at the end of each scale to restore the original size of the features. A low-pass filter is formed under average pooling of different kernel sizes. We divide f into four parts through channel separation, namely $\{f_1, f_2, f_3, f_4\}$. Each part is processed using different sizes of pooling, which is described as

$$Filter(f_i) = Up(\beta_s(f_i)) \tag{5}$$

where f_i is part of f divided on the channel, **Up** is bilinear interpolation sampling, β_s is an adaptive average pooling of different sizes of $s \times s$. Finally, after tensor splicing each $\{f_i, i = 1, 2, 3, 4\}$, we restore them to $f \in R^{h \times w \times 3}$.

4 Experiments

4.1 Dataset and Implementation Details

Dataset: We use the ExDark dataset to validate the effectiveness of our PE-YOLO. ExDark is a low-light object detection dataset used for research on object detection and image enhancement. It collected a total of 7363 images under 10 different lighting conditions, from extremely low light to dusk, with 12 bounding box annotations of objects in the images. We divided ExDark into 80% for training and 20% for testing, and the specific division is consistent with IAT [5] and MAET [6].

Details: All trained and tested images are resized to 608×608, and data augmentation methods such as random cropping, flipping, and multi-scale resizing are used during training. Batch-size is set to 8, the optimizer uses SGD, the initial learning rate is set to 0.001, and the weight decay is set to 0.0005. Train PE-YOLO for 30 epochs and run our model on a single RTX 3090 GPU. The deep learning framework is Pytorch, and we use mmdetection [3] to achieve our model.
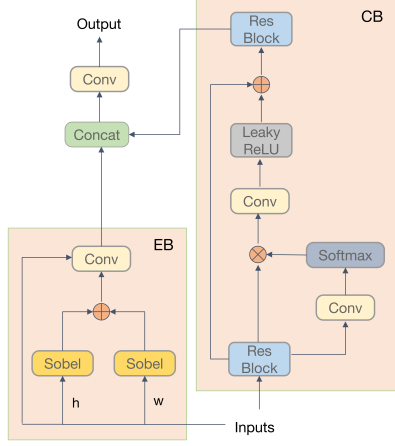
Fig. 4. Detail of DPM, contains context branch (CB) and edge branch (EB).
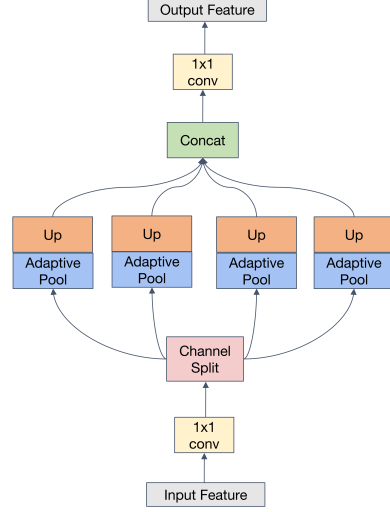
Fig. 5. Details of low-frequency enhancement filter (LEF). LEF is composed of adaptive averge pooling in different sizes to intercept low-frequency components.

Evaluation: We use mAP and FPS to validate the effectiveness of our model. mAP is the average AP of all categories in the detection model, and a larger value indicates a higher accuracy of the model. It is represented as

$$mAP = \frac{\sum_{i=1}^{C} AP_i}{C} \tag{6}$$

where C is the number of categories, and AP is the Averge Precision for each category, calculated by the area of the Precision Recall curve. FPS is the number of image frames detected by the model per second, and a larger FPS indicates a faster model detection speed.

4.2   Experimental Results

To verify the effectiveness of PE-YOLO, we conducted many experiments on the ExDark dataset. Firstly, we compare PE-YOLO with other low-light enhancement models. Due to the lack of detection capability of the low light enhancement model, we will use the same detector as PE-YOLO to experiment on all enhanced images. We set the IoU threshold of mAP to 0.5, and the performance comparison is shown in Table 1. We found that directly using low-light enhanced models before YOLOv3 did not significantly improve detection performance. Our PE-YOLO is 1.2% and 1.1% higher on mAP than MBLLEN and Zero-DCE, respectively, achieving the best results.

Table 1. Performance comparisons between PE-YOLO and low-light enhancement models. It shows mAP and AP in each class. The bold number has the highest score in each column.

Model	Venue	Bicycle	Boat	Bottle	Bus	Car	Cat	Chair	Cup	Dog	Motorbike	People	Table	mAP
YOLOv3 [19]	arXiv	79.8	75.3	78.1	92.3	83.0	68.0	69.0	79.0	78.0	77.3	81.5	55.5	76.4
KIND [26]	MM2019	80.1	77.7	77.2	**93.8**	83.9	66.9	68.7	77.4	79.3	75.3	80.9	53.8	76.3
MBLLEN [17]	BMVC 2018	82.0	77.3	76.5	91.3	**84.0**	67.6	69.1	77.6	**80.4**	75.6	**81.9**	**58.6**	76.8
Zero-DCE [9]	CVPR 2020	84.1	77.6	78.3	93.1	83.7	70.3	69.8	77.6	77.4	76.3	81.0	53.6	76.9
PE-YOLOv3 (Ours)	-	**84.7**	**79.2**	**79.3**	92.5	83.9	**71.5**	**71.7**	**79.7**	79.7	**77.3**	81.8	55.3	**78.0**

We visualized the detection results of different low light enhancement models, as shown in Fig. 6. We found that although MBLLEN and Zero DCE can significantly improve the brightness of the image, they also enlarge the noise in the image. PE-YOLO mainly captures the potential information of objects in low-light images, while suppressing noise in high-frequency components, thus PE-YOLO has better detection performance.
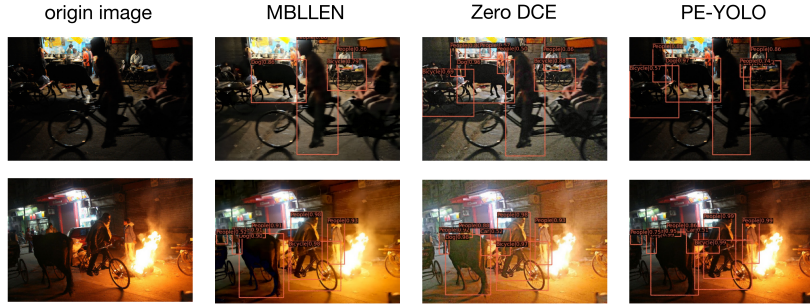

Fig. 6. Detection Results in PE-YOLO and other low-light enhancement models.

We compared the performance of PE-YOLO with other dark detectors, as shown in Table 2. In addition, we visualized the detection results of the dark detector and PE-YOLO, as shown in Fig. 7, which clearly showed that PE-YOLO was more accurate in object detection. PE-YOLO is 0.7% and 0.2% higher in mAP compared to DENet and IAT-YOLO pre-trained with the LOL dataset, and our PE-YOLO is also almost the highest on FPS. The above data indicate that PE-YOLO is more suitable for detecting objects in dark conditions.

4.3 Ablation Study

To analyze the effectiveness of each component in PE-YOLO, we conducted ablation studies, and the results are shown in Table 3. After adopting context branching, PE-YOLO increased from 76.4% to 77.0% in mAP, indicating that

Table 2. Performance comparisons between PE-YOLO and dark detectors. The bold number has the highest score in each column.

Model	Venue	Bicycle	Boat	Bottle	Bus	Car	Cat	Chair	Cup	Dog	Motorbike	People	Table	mAP	FPS
YOLOv3 [19]	arXiv	79.8	75.3	78.1	92.3	83.0	68.0	69.0	79.0	78.0	77.3	81.5	55.5	76.4	54.0
MAET [6]	ICCV 2021	83.1	78.5	75.6	**92.9**	83.1	73.4	71.3	79.0	**79.8**	77.2	81.1	57.0	77.7	-
IAT-YOLOV3 (LOL pretrain) [5]	BMVC2022	79.8	76.9	78.6	92.5	83.8	**73.6**	**72.4**	78.6	79.0	**79.0**	81.1	**57.7**	77.8	52.6
DENet [18]	ACCV2022	80.4	**79.7**	77.9	91.2	82.7	72.8	69.9	**80.1**	77.2	76.7	**82.0**	57.2	77.3	**54.8**
PE-YOLOv3 (Ours)	-	**84.7**	79.2	**79.3**	92.5	**83.9**	71.5	71.7	79.7	79.7	77.3	81.8	55.3	**78.0**	53.6

capturing remote dependencies is effective for enhancement. After adopting edge branch, the mAP increased from 77.0% to 77.6%, indicating that edge branch can enhance the texture of components and enhance the details of the enhanced image. After adopting LEF, the mAP increased from 77.6% to 78.0%, indicating that capturing low-frequency components is beneficial for obtaining potential information in the image. In the end, our model improved from 76.4% to 78.0% on mAP and only decreased by 0.4 on FPS.

Fig. 7. Detection Results in PE-YOLO and other dark detectors.

Table 3. Ablation study on PE-YOLO. "CB" represents context branch, "EB" represents edge branch, and LEF represents low-frequency enhancement filter.

DPM		LEF	mAP	FPS
CB	EB			
			76.4	54.0
✓			77.0	53.9
✓	✓		77.6	53.8
✓	✓	✓	78.0	53.6

5   Conclusion

To achieve more robust dark object detection, we propose a pyramid enhancement network (PENet) that performs detail restoration and captures potential information. By combining PENet and YOLOv3, we build a dark object detection framework named PE-YOLO. We first use the Laplacian pyramid to decompose the image into four components with different resolutions, and propose a detail processing module (DPM) and a low-frequency enhancement filter (LEF) for component enhancement. In addition, PE-YOLO trains in an end-to-end way, without additional loss function. We conducted experiments in the ExDark dataset, and the experimental results show that compared with the low-light enhancement models and the dark detectors, PE-YOLO achieves the best results and can effectively detect objects in dark conditions. However, our model should be studied on more detectors and further improve performance while maintaining lightweight.

References

1. Bochkovskiy, A., Wang, C.Y., Liao, H.Y.M.: Yolov4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020)
2. Cai, Z., Vasconcelos, N.: Cascade r-cnn: Delving into high quality object detection. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6154–6162 (2018)
3. Chen, K., Wang, J., Pang, J., Cao, Y., Xiong, Y., Li, X., Sun, S., Feng, W., Liu, Z., Xu, J., et al.: Mmdetection: Open mmlab detection toolbox and benchmark. arXiv preprint arXiv:1906.07155 (2019)
4. Chen, Y., Li, W., Sakaridis, C., Dai, D., Van Gool, L.: Domain adaptive faster r-cnn for object detection in the wild. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3339–3348 (2018)
5. Cui, Z., Li, K., Gu, L., Su, S., Gao, P., Jiang, Z., Qiao, Y., Harada, T.: You only need 90k parameters to adapt light: a light weight transformer for image enhancement and exposure correction. In: 33rd British Machine Vision Conference 2022, BMVC 2022, London, UK, November 21-24, 2022. BMVA Press (2022), https://bmvc2022.mpi-inf.mpg.de/0238.pdf
6. Cui, Z., Qi, G.J., Gu, L., You, S., Zhang, Z., Harada, T.: Multitask aet with orthogonal tangent regularity for dark object detection. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 2553–2562 (2021)
7. Dudhane, A., Zamir, S.W., Khan, S., Khan, F.S., Yang, M.H.: Burst image restoration and enhancement. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 5759–5768 (2022)
8. Everingham, M., Van Gool, L., Williams, C.K., Winn, J., Zisserman, A.: The pascal visual object classes (voc) challenge. International journal of computer vision **88**, 303–338 (2010)
9. Guo, C., Li, C., Guo, J., Loy, C.C., Hou, J., Kwong, S., Cong, R.: Zero-reference deep curve estimation for low-light image enhancement. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. pp. 1780–1789 (2020)

10. Guo, X., Li, Y., Ling, H.: Lime: Low-light image enhancement via illumination map estimation. IEEE Transactions on image processing **26**(2), 982–993 (2016)
11. Kalwar, S., Patel, D., Aanegola, A., Konda, K.R., Garg, S., Krishna, K.M.: Gdip: Gated differentiable image processing for object-detection in adverse conditions. arXiv preprint arXiv:2209.14922 (2022)
12. Law, H., Deng, J.: Cornernet: Detecting objects as paired keypoints. In: Proceedings of the European conference on computer vision (ECCV). pp. 734–750 (2018)
13. Li, C., Du, D., Zhang, L., Wen, L., Luo, T., Wu, Y., Zhu, P.: Spatial attention pyramid network for unsupervised domain adaptation. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XIII 16. pp. 481–497. Springer (2020)
14. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context. In: Computer Vision–ECCV 2014: 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part V 13. pp. 740–755. Springer (2014)
15. Liu, W., Ren, G., Yu, R., Guo, S., Zhu, J., Zhang, L.: Image-adaptive yolo for object detection in adverse weather conditions. In: Proceedings of the AAAI Conference on Artificial Intelligence. vol. 36, pp. 1792–1800 (2022)
16. Loh, Y.P., Chan, C.S.: Getting to know low-light images with the exclusively dark dataset. Computer Vision and Image Understanding **178**, 30–42 (2019)
17. Lv, F., Lu, F., Wu, J., Lim, C.: Mbllen: Low-light image/video enhancement using cnns. In: BMVC. vol. 220, p. 4 (2018)
18. Qin, Q., Chang, K., Huang, M., Li, G.: Denet: Detection-driven enhancement network for object detection under adverse weather conditions. In: Proceedings of the Asian Conference on Computer Vision. pp. 2813–2829 (2022)
19. Redmon, J., Farhadi, A.: Yolov3: An incremental improvement. arXiv preprint arXiv:1804.02767 (2018)
20. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems **28** (2015)
21. Sasagawa, Y., Nagahara, H.: Yolo in the dark-domain adaptation method for merging multiple models. In: Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16. pp. 345–359. Springer (2020)
22. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., Rabinovich, A.: Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 1–9 (2015)
23. Tian, Z., Shen, C., Chen, H., He, T.: Fcos: Fully convolutional one-stage object detection. In: Proceedings of the IEEE/CVF international conference on computer vision. pp. 9627–9636 (2019)
24. Wei, C., Wang, W., Yang, W., Liu, J.: Deep retinex decomposition for low-light enhancement. arXiv preprint arXiv:1808.04560 (2018)
25. Wu, A., Han, Y., Zhu, L., Yang, Y.: Instance-invariant domain adaptive object detection via progressive disentanglement. IEEE Transactions on Pattern Analysis and Machine Intelligence **44**(8), 4178–4193 (2021)
26. Zhang, Y., Zhang, J., Guo, X.: Kindling the darkness: A practical low-light image enhancer. In: Proceedings of the 27th ACM international conference on multimedia. pp. 1632–1640 (2019)