ERUP-YOLO: Enhancing Object Detection Robustness for Adverse Weather Condition by Unified Image-Adaptive Processing

Supplementary Material

7. Visual Comparison with Prior Methods

We present a visual comparison of detection results between our proposed method and prior approaches (YOLOv3 and GDIP-YOLO) across various adverse weather conditions in Section 5.4 of the main paper. We use the models evaluated in Table 3. Figure 9 shows enhanced input image and detection results overlaid on input images for low-light, foggy, and rain conditions from ExDark and DAWN datasets. Figure 10 extends this comparison to snow and sand conditions.

Our proposed method demonstrates a significant image transformation effect compared to GDIP-YOLO and notable improvements in detection performance, particularly in low-light and snow conditions (first row in Figure 9 and first row in Figure 10). However, we observe limitations in foggy and sand conditions, particularly for distant objects obscured by scattering effects. In these cases (Figure 9, second row for foggy, and Figure 10, second and third row for sand), our method tends to overemphasize the brightness of distant objects affected by scattering, sometimes leading to missed detections in the background of the image.

8. Visual Results of Ablation Study

We present a visual comparison of detection results from our Ablation Study in Section 5.5 of the main paper. Figure 11 presents detection results for low-light and foggy conditions using different filter configurations. BPW+KBL combination consistently improves the overall visibility of the scenes across both low-light and foggy conditions. As evident in the second row of Figure 11, BPW+KBL configuration significantly enhances the detection accuracy of objects in extremely dark regions. However, the second and third rows also reveal a limitation of BPW+KBL combination. The global brightness increase tends to oversaturate areas that are already bright due to illumination sources or distant scattering effects in foggy conditions.

9. Evaluation of Image Filters

To quantitatively assess the capabilities of our proposed filters, we conduct four sets of experiments on VOC2007_test dataset. We evaluate the expressiveness of our BPW filter in approximating sequential prior pixel-wise filters (white balance, gamma, contrast, and tone filters). We apply these pixel-wise filters sequentially with random parameters to the test images to generate degraded target images. Then, we optimize the parameters of our BPW filter using the Adam optimizer for 50 iterations with a learning rate of 0.01, minimizing the combined mean squared error (MSE) and structural similarity (SSIM) loss between the filtered images and the target degraded images.

Table 4. PSNR between the degraded and optimized images when using each filter. Each column presents the initial degraded image (initial), the result after applying the optimization filter (final), and the improvement from initial to final (del).

Image filters		PSNR		
Degrade	Optimize	Initial	Final	Det
Pixel Wise	BPW(ours)	13.90	22.29	8.38
BPW(ours)	Pixel Wise	17.62	27.46	**9.84**
Sharp+DeFog	KBL(ours)	12.57	20.09	**7.52**
KBL(ours)	Sharp+DeFog	-8.17	-8.00	0.17

We also conduct the reverse experiment, where our BPW filter randomly degrades the test images, and we measure how well the PW filters can recover the original images. For our KBL filter, we follow a similar procedure. We degrade the test images by applying the combination of prior defog and sharpen filters with random parameters, creating the target degraded images. We also perform the reverse experiment between the defog and sharpen filters and KBL filter.

Table 4 shows the PSNR between the degraded and optimized images when using each filter for degradation and optimization, respectively. Since the degree of degradation from the original image varies depending on the degradation filter, we report the mean of the initial PSNR between the original and degraded images, the final PSNR between the optimized and degraded images, and the increase in PSNR (det) from initial to final. The results demonstrate that our BPW filter can mimic the effects of prior pixel-wise filters with comparable performance, while our KBL filter exhibits superior expressiveness in approximating the combination of prior sharpness and defog filters.

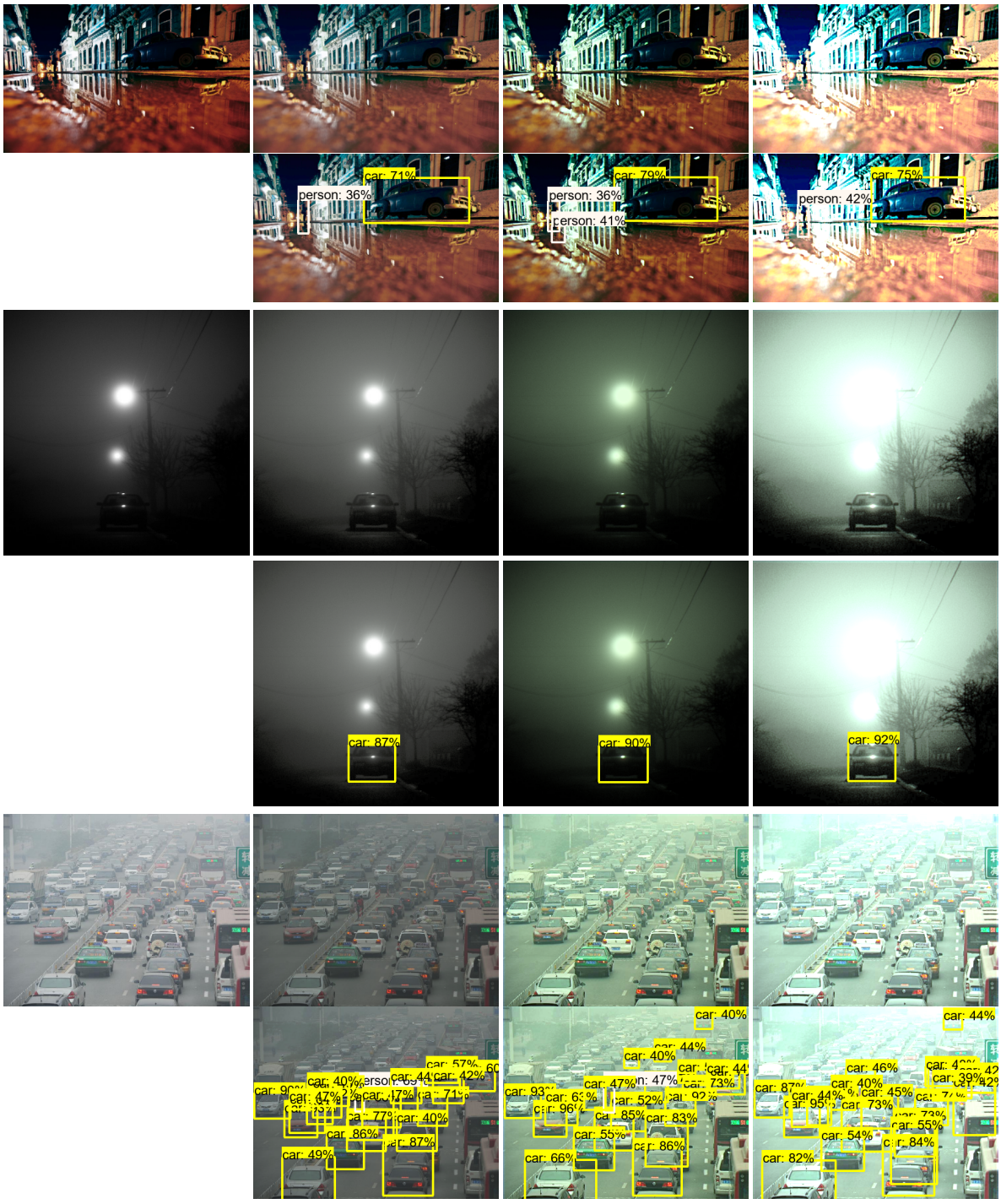
The pixel-wise filters shows a slightly higher mimicking capability compared to our BPW filter. This can be attributed to the fact that the pixel-wise filters comprises a sequential process of four distinct pixel-wise filters, allowing for a high degree of expressiveness through their combination. In contrast, our method achieves nearly equivalent expressiveness using only a single module, which is an improved the conventional tone filter. Despite this minor difference, our approach offers significant advantages. As illustrated in Figure 9, 10, 11, our proposed filter demonstrates high expressiveness in image transformation.

(a)YOLOv3 (b)GDIP (c)ERUP(ours)

Figure 9. Visual comparison of enhanced images and detection results using (a)YOLOv3, (b)GDIP and (c)ERUP(ours). First row shows the low-light results from ExDark, second row shows the fog results from DAWN, and third row shows the rain results from DAWN.

(a)YOLOv3 (b)GDIP (c)ERUP(ours)

Figure 10. Visual comparison of enhanced images and detection results using (a)YOLOv3, (b)GDIP and (c)ERUP(ours). First row shows the snow results from DAWN and second and third row shows the sand results from DAWN.



(a)Original (b)BPW (c)KBL (d)BPW+KBL

Figure 11. Visual comparison of (a)Original input images, enhanced images and detection results using (b)BPW, (c)KBL and (d)BPW+KBL(proposed). First and second row shows the low-light results from ExDark, third row shows the fog results from RTTS.