DSWEC: A Saliency-Weighted Defense against Camera Exposure Tampering*

Seohyun Kim[1] and Sanghoon Jeon*[2]

[1] Department of Automobile and IT Convergence, Kookmin University, Seoul 02707, Republic of Korea
tjgus4470@kookmin.ac.kr

[2] * Department of Automobile and IT Convergence, Kookmin University, Seoul 02707, Republic of Korea
sh.jeon@kookmin.ac.kr

Abstract

With the acceleration of autonomous driving commercialization, the security vulnerabilities of cameras, the core sensors of vehicles, have emerged as an important issue. This paper examines the impact on camera-based perception performance when the automatic exposure control (AEC) parameters of an image signal processor (ISP) are externally tampered with. As a countermeasure, we propose DSWEC (Dual-Pass Saliency-Weighted Exposure Control), a defense framework that detects attacks in real time and restores lost information through exposure correction weighted for driving-critical objects. Experimental results using the CARLA simulator show that the proposed attack reduced pedestrian detection from 93.3% to 14.3%, while DSWEC restored it to 91.7%.

Keywords: Automatic Exposure Control (AEC), Adversarial Attack Defense, Camera Sensor Security, Autonomous Driving

1 Introduction and Method

The safety of autonomous vehicles depends on the reliability of camera-based perception systems. Recent studies have reported that tampering with the automatic exposure control (AEC) parameters of an image signal processor (ISP) can cause brightness distortion and perception errors [3]. This study simulates such AEC manipulation attacks and proposes a defense framework, DSWEC (Dual-Pass Saliency-Weighted Exposure Control) [2, 5]. The proposed method was verified through driving scenarios based on the CARLA simulator.

An attacker can access the ECU/CAN bus to directly manipulate AEC parameters (exposure time, gain, etc.), intentionally generating overly dark or bright images to interfere with object detection. To mitigate this threat, this study designed the following defense mechanisms:

1. Detection of inconsistency between sensor metadata and image statistics
2. Detection of temporal discontinuity
3. Histogram-based anomaly detection
4. Detection of abnormal situations by integrating the above three signals into an attack determination score, restoring normal AEC parameters, and performing saliency-weighted exposure correction.

2 Evaluation

2.1 Research Question

1. What is the impact of AEC tampering on pedestrian, vehicle, and lane recognition rates?
2. How effectively can DSWEC detect attacks and restore perception performance?
3. How does DSWEC differ in performance from existing image enhancement methods?

*Proceedings of the 9th International Conference on Mobile Internet Security (MobiSec'25), Article No. P-69, December 16-18, 2025, Sapporo, Japan.

2.2 Results and Analysis

2.2.1 Attack Effectiveness Analysis

Figure 1 illustrates the changes in perception performance caused by AEC tampering under normal and attack conditions.
"Avg. Conf." denotes the average confidence score of detected objects obtained from the YOLOv5s perception network.

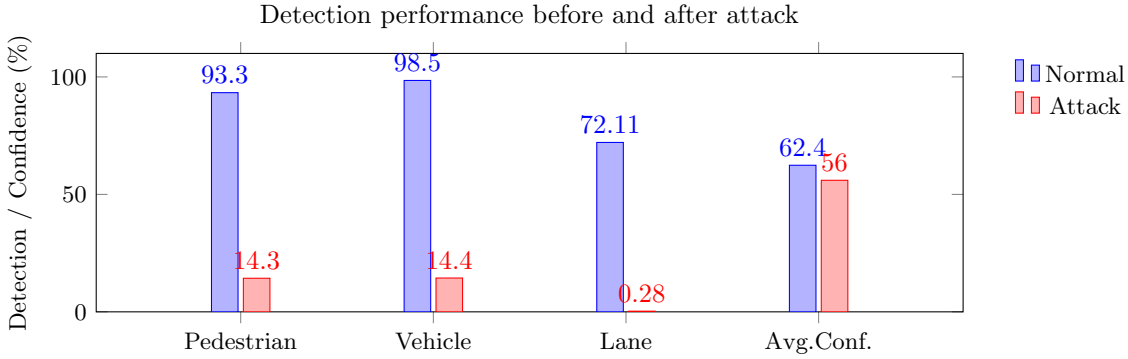

Figure 1: Comparison of detection performance before and after attack.

2.2.2 Defense Framework Performance Evaluation

Table 1 compares the perception recovery performance of several enhancement and defense methods, including the proposed DSWEC. [1, 4]

Table 1: Comparison of detection performance by defense method

Method	Pedestrian (%)	Vehicle (%)	Lane (%)	Avg. Conf.
Attack	14.3	14.4	0.28	0.56
CLAHE	54.8	51.4	38.68	0.51
MSRCR	43.6	46.6	40.48	0.47
Zero-DCE	58.2	81.4	59.61	0.50
DSWEC (Proposed)	91.7	95.7	64.34	0.6295

3 Conclusion

This study confirmed that AEC tampering can severely disrupt perception systems, with pedestrian and vehicle detection decreasing by about 85% and lane recognition by 99.6%. Conventional image enhancement methods show high performance variance under different lighting and parameter conditions, but DSWEC provides stable recovery based on attack awareness. Since this study was conducted in the CARLA simulator environment, actual driving conditions may show differences. Future work will further verify the robustness of DSWEC under various attack conditions and explore real-time implementation feasibility.

References

[1] 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 *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pages 1780–1789, 2020.

[2] I. Jatzkowski, D. Wilke, and M. Maurer. A deep-learning approach for the detection of overexposure in automotive camera images. In *2018 21St international conference on intelligent transportation systems (ITSC)*, pages 2030–2035. IEEE, 2018.

[3] Y. Man, M. Li, and R. Gerdes. Remote perception attacks against camera-based object recognition

systems and countermeasures. *ACM Transactions on Cyber-Physical Systems*, 8(2):1–27, 2024.

[4] T. P. H. Nguyen, Z. Cai, K. Nguyen, S. Keth, N. Shen, and M. Park. Pre-processing image using brightening, clahe and retinex. *arXiv preprint arXiv:2003.10822*, 2020.

[5] D. Ye, C. Chen, C. Liu, H. Wang, and S. Jiang. Detection defense against adversarial attacks with saliency map. *International Journal of Intelligent Systems*, 37(12):10193–10210, 2022.