Autonomous Driving Vehicle Detection Methods under Different Low-Light Scenes

Jiacheng Fan

School of Electronic Engineering, Xidian University, Xi'an, Shaanxi, China
22022100066@stu.xidian.edu.cn

Abstract. Vehicle detection is a core task in environmental perception for autonomous driving, and its performance directly affects driving safety. However, low-light scenes, due to issues such as low light levels, atmospheric scattering, and sensor occlusion, result in image distortion, difficulty in edge detection, and impaired signal transmission. These problems seriously undermine the accuracy and robustness of traditional detection methods, thus emerging as a critical bottleneck in environmental perception for autonomous driving. This paper presents a review of autonomous vehicle detection under low-light scenes, with a focus on the impacts of three typical low-light scenes—nighttime, foggy, and rainy weather—on visual perception, as well as the targeted technical solutions for vehicle detection in these scenes. It combs through existing technical frameworks and provides a systematic organization and introduction. In addition, this paper introduces mainstream vehicle detection datasets, identifies current research bottlenecks, and discusses future development directions. The aim of this paper is to comprehensively understand the influence of low-light scenes on vehicle detection, provide effective references for autonomous driving vehicle detection methods under different low-light scenes, and lay a foundation for the future development and practical application of autonomous driving technology.

Keywords: Low-light scenes, Autonomous driving, Vehicle detection, Generatibe Adversial Network s(GAN), You Only Look Once (YOLO)

1 Introduction

With the rapid development of intelligent connectivity technology, autonomous driving has moved from laboratories to large-scale road tests and is gradually becoming the core of the intelligent transportation system. Its safe operation relies on accurate environmental perception, which helps to detect and recognize various targets in real-time and accurately in complex dynamic road scenarios. As a fundamental task of environmental perception, vehicle detection needs to real-time detect information such as the position, motion state, and shape of surrounding vehicles, providing key inputs for path planning, obstacle avoidance decision-making, and other processes.

However, when vehicles operate in low-light scenes such as nighttime, rainy days, and foggy days, the drastic changes in environmental characteristics pose significant

K. Subramanian (ed.), *Proceedings of the International Workshop on Advances in Deep Learning for Image Analysis and Computer Vision (IWADIC 2025)*, Advances in Computer Science Research 128,
https://doi.org/10.2991/978-94-6239-648-7_18

challenges to detection algorithms: under nighttime scenes, low light increases the difficulty of detecting vehicles, high-intensity glare from oncoming vehicles’ high beams causes "overexposure" of the camera, resulting in unclear vehicle outlines and further detection difficulties; in rainy scenes, surface water accumulation on the lens distorts the captured images and causes occlusion effects. Meanwhile, water droplets in the air lead to light path fluctuations at different positions, which affect the characteristics of the captured targets, such as color, contrast, and reflectivity; in foggy scenes, the atmospheric scattering effect reduces visibility and decreases image contrast, increasing the difficulty of edge detection and impairing the performance of subsequent image processing and computer vision applications. These issues directly lead to a substantial decline in the real-time performance, accuracy, and robustness of traditional object detection methods, along with a significant increase in the miss detection rate and false detection rate. This has become an important bottleneck that restricts the development of autonomous driving from Level 4 above and all-weather safe operation.

This paper take three typical low-light scenes in practice as examples, namely, nighttime, foggy and rainy scenes. This paper propose vehicle detection technical solution for different scenes respectively. Finally, this paper conclude existing research hotspots such as weak generalization ability of cross-scenario, and prospect future research direction.

The research on vehicle detection technology in low-light scenes can summarize the scattered research results in the field of low-light scenes, and build a technical system of low-light vehicle detection. It can take the pain points of detection in low-light scene as an research target, and can directly provide technical support for reducing the accident rate of autonomous driving, and then promote the intelligent transportation to achieve the goal of all-scenario and all-weather safe operation. It will become an important link to improve travel efficiency and people's livelihood safety.

2 Technical Solutions for Different Low-Light Scenes

2.1 Targeted Technical Solutions for Nighttime

The key challenges of nighttime scene are as follows. On the one hand, the low illumination environment in lightless area will bring low image noise and affect the accuracy of vehicle detection. Traditional algorithm depends on the edge detection and feature extraction. On the other hand, local intense light glare formed by high-intensity streetlights, high beams of oncoming vehicles, etc., will cause overexposure of the camera, obscure the intrinsic features of vehicles, increase the difficulty of detection, and raise the false detection rate.

Liang Deng et al. proposed a nighttime vehicle detection method that integrates three key technologies: first, performing image registration on near-infrared and visible light images based on SURF features; second, conducting pixel-level fusion via YUV channels after wavelet transform; and third, implementing transfer training for the YOLOv4 model. By virtue of this integrated strategy, the method ultimately achieved higher image entropy and greatly enhanced the accuracy of vehicle detection in nighttime scenes [1].

Ze Wu et al. proposed a novel deep learning-based fusion algorithm. By fusing visible light and infrared images, the algorithm generates new images with rich texture details and clear contours. Subsequently, the fused images are fed into the YOLOv5 network based on the PyTorch framework for real-time detection, which significantly enhances the object detection capability of autonomous driving in nighttime scenes [2].

Che-Tsung Lin et al. proposed a GAN-based data augmentor, AugGAN. By converting images collected during driving to the required domain, it significantly reduces artifacts in the converted images, thereby greatly improving the performance of nighttime vehicle detection [3].

Jia Rushi proposed a CME-YOLO nighttime vehicle detection algorithm improved based on the YOLOv8 architecture. By optimizing the MobileViT network architecture, CMF-MobileViT was obtained to replace the C3 component in the original CSPDarknet53 backbone network. Subsequently, an efficient multi-scale attention mechanism (EMA) was incorporated into the corresponding target layers at the Neck. Finally, to address the problem of a high missed detection rate for overlapping vehicles, MPDIoU was adopted to replace CIoU, which accelerates convergence while further improving the recall rate of the model [4].

U.N. Nisha et al. proposed a nighttime vehicle detection method, which is based on the FR-CNN and YOLOv5 algorithms with GAN as the underlying network. This method can convert the collected nighttime images into daytime images for output, achieving a detection accuracy of 96.75% [5].

2.2 Targeted Technical Solutions for Rainy Weather

The challenges in rainy weather scenes mainly arise from three aspects: First, raindrops falling on cameras lead to water accumulation on the lens surface, which creates an occlusion effect. This distorts the captured images and degrades the quality of detected data, such as that of vehicles, lane markings, and traffic lights. Second, raindrops can be regarded as approximately spherical lenses, causing fluctuations in light paths at different spatial positions. These fluctuations severely interfere with the quality of scenes captured by cameras—for example, they may alter the behavior and shape of targets, affect properties like color, contrast, and reflectivity, and induce "rainfall attenuation effect" and "backscattering effect," thereby reducing the received power of useful signals. Meanwhile, the specular reflection of water puddles on the road surface reflects the light and shadows of street lamps and buildings into "vehicle-like contours," resulting in numerous false detections by the detection model, such as misclassifying road reflections as compact cars.

Sun Zaiming proposed a vehicle detection algorithm under rainy and foggy conditions with multi-level feature fusion. Firstly, a model is designed that can simultaneously remove both rain and fog without switching between a single model and training data. This model can effectively capture information beneficial for removing rain and fog from images and restoring the details of degraded images. Meanwhile, a decoder with learnable weather type queries is introduced. Finally, by integrating the image rain and fog removal model, an overall framework of the vehicle target detection model with enhanced local perception is proposed. The model is tested using datasets under mixed weather conditions and single weather conditions respectively [6].

Hao Sun et al. convolutional neural network takes the rainy image as input, can effectively suppress the rain stripes and atmospheric veiling effect induced by the accumulation of distant rain stripes, and recover a clean image. Firstly, in order to train and evaluate the method, a synthetic dataset containing urban street view images in various scenes, i.e., different rain intensity, rain stripe direction and haze level, is constructed. Secondly, the method is evaluated by quantitative and qualitative experiments based on synthetic data. Experimental results have shown that the method can significantly improve semantic segmentation and object detection for autonomous driving in rainy condition [7].

Hemant Kumar et al. proposed an optimized Faster R-CNN architecture, which uses MobileNetV3 as its backbone and is supplemented by the Feature Pyramid Network (FPN). The Self-Supervised Thermal Network (SSTN) was integrated into this architecture to aid with vehicle detection, and this optimized architecture was trained and evaluated using the DAWN dataset. The SSTN-enhanced Faster R-CNN achieves an accuracy of 87.15%, a mean average precision (mAP) of 85.20%, and a recall rate of 81.37%, demonstrating noticeably better performance than a number of current techniques. The system continues to retain excellent accuracy even in difficult weather situations like rain, snow, and sandstorms [8].

Jian-Gang Wang et al. proposed a Vehicle-Aware Generative Adversarial Networks (VAGAN) to enhance vehicle detection in rainy scenes. This approach leverages the fact that vehicle taillights are typically activated in rainy conditions to detect image features. Images generated by this method are immune to image translation noise, enabling straightforward vehicle localization through color segmentation [9].

Xueyang Fu et al. proposed a new deep network architecture based on deep convolutional neural networks (CNNs) for removing rain streaks from a single image [10].

2.3 Targeted Technical Solutions for Foggy Weather

Fog is composed of aerosol particles formed by suspended substances in the atmosphere. It strongly scatters light, reducing visibility and image contrast. This results in images that are generally pale with a sharp drop in contrast, making edge detection challenging and undermining the performance of subsequent image processing and computer vision applications. Furthermore, fog droplets interfere with non-visual sensors such as LiDAR: they block the reflected signals of LiDAR, leading to the loss of vehicle detection information.

Hai Wang et al. proposed a foggy weather vehicle detection network based on the improved YOLOv5, which takes the ResNeXt model modified by structural reparameterization as the backbone. Aiming at the problem of insufficient features in images under foggy scenes, they constructed a new Feature Enhancement Module (FEM) and utilized the attention mechanism to help the detection network recognize the features of images in foggy scenes more accurately [11].

Hasan Abbasi et al. proposed a fog-aware adaptive YOLO algorithm for object detection in foggy scenes. This algorithm mainly consists of three modules: Synthetic Data Generator, Fog Level Evaluator, and Object Detection Block. Before conducting object detection, evaluate the haze of the image. If the haze does not exceed the threshold, defogging is not necessary and YOLOv3 can be used; If the threshold is exceeded, use image-adaptive YOLOv3 for defogging and then perform object detection [12].

Josué Manuel Rivera Velàzquez et al. put forward an approach that employs a pre-trained YOLOv5 model for object detection in foggy scenes, utilizing thermal cameras with different angles of view [13].

Zhang Siyuan proposed a one-stage object detection framework that integrates the feature-supervised defogging algorithm with domain adaptation, comprising an image style transfer module, a feature extraction module, and a detection module. First, style transfer between image domains is achieved through the feature-supervised defogging algorithm and its inverse transformation. Second, a domain classifier is introduced into the high and low-dimensional feature extraction layers, and adversarial training is employed to reduce the feature discrepancy across domains [14].

Wang Dan proposed combining the improved all-in-one dehazing network (AOD-Net) with the improved vehicle detection algorithm based on the YOLOv7 network model, and introduced a foggy image discrimination module, which uses grayscale histogram features to identify foggy images [15].

Sun Zaiming proposed a vehicle detection algorithm for foggy scenes based on an improved U-Net. First, by referring to the network structure of U-Net, an end-to-end defogging network was designed through the introduction of a Transformer. Second, the cross-attention skip connection mechanism was introduced to fuse multi-scale features, thereby enhancing the learning capability of the network to restore images with clean features after defogging. Meanwhile, a vehicle detection model for foggy conditions was constructed by integrating defogging and detection processes [6].

3 Datasets in Low-Light Scenes

3.1 KITTI Dataset

The KITTI dataset was jointly developed by the Karlsruhe Institute of Technology (Germany) and the Toyota Technological Institute (United States). It is a dataset applicable to fields such as mobile robotics and autonomous driving. Through multiple sensor modalities, the dataset records a total of 6 hours of traffic scenarios, covering real-world traffic conditions including highways, rural areas, and complex urban environments. Using 3D bounding box trajectories, the dataset's annotations are displayed for categories like "car," "van," and "truck." The two primary components of this dataset are a 3D visual odometry/SLAM dataset with 22 stereo sequences totaling 39.2 km in length, and a benchmark dataset for stereo matching and optical flow estimates with 194 pairs of training photos and 195 pairs of test images. As a well used standard dataset, it is regularly used to assess how well algorithms perform tasks like vehicle identification, semantic segmentation, and stereo vision in computer vision and autonomous driving [16].

3.2 BDD100K Dataset

The BDD100K dataset is a large-scale driving video dataset jointly developed by teams from the University of California, Berkeley, Cornell University, and Element Inc. It comprises over 100K distinct video clips, covering various scene types such as urban streets, residential areas, and highways, under varying weather conditions and at different times of the day. The dataset encompasses ten tasks, namely image tagging, lane detection, road image detection,etc. Serving as an evaluation benchmark for computer vision research in autonomous driving, this dataset can act as a reference for heterogeneous multi-task learning and baseline studies, thereby facilitating future research endeavors [17].

3.3 Cityscapes Dataset

The Cityscapes dataset serves as a benchmark suite and large-scale dataset, tailored for training and evaluating pixel-level and instance-level semantic annotation methods. It comprises a large, diverse collection of stereo videos capturing street scenes from 50 distinct cities. Within the dataset, there are 5,000 images annotated with high-quality pixel-level details and 20,000 images with coarse-grained annotations. Additionally, it defines 30 visual categories for annotations, which encompass such types as vehicles, objects, and persons, among others. The primary objective of this dataset is to drive advancements in the semantic understanding of urban scenes [18].

4 Existing Bottlenecks and Future Prospects

4.1 Existing Bottleneck

Despite significant advancements in autonomous driving vehicle detection technology for low-light scenes, several core bottlenecks persist in practical applications, hindering its progression toward higher reliability and stronger robustness. These challenges are manifested in the following aspects:

The relevant datasets for vehicle detection in low-light scenes are both limited in quantity and subpar in quality. On one hand, collecting real-world road condition data for extreme scenarios—such as snowstorms or foggy nights—is highly challenging and costly, leading to an underrepresentation of such samples in training datasets. On the other hand, under low-light conditions, image edges tend to be blurred, making manual annotation susceptible to subjective biases; this, in turn, causes models to learn incorrect feature correlations.

Existing methods are mostly designed for single low-light scenarios (e.g., pure nighttime or pure foggy weather). However, due to their weak cross-scenario generalization capability, their performance drops sharply when applied to complex scenarios (such as rainy nights). For instance, Generative Adversarial Network (GAN)-based methods perform poorly in handling other low-light features during day-night image conversion. Similarly, fog removal models struggle to deal with coexisting vehicle headlight glare, leading to detection errors when these models are deployed in dynamically changing mixed scenarios.

In extreme low-light scenes (such as lightless nights, heavy rainstorms, and dense fog), existing technologies lack sufficient robustness and still fall short of practical requirements. For instance, in heavy rain scenarios, LiDAR is susceptible to the influence of uneven precipitation in the atmosphere. The LiDAR signals are disrupted by the scattering effect of raindrops, leading to detection errors. In fog, water droplets and suspended particles absorb light, reducing light intensity—particularly within the visible spectrum. This impairs the ability of visual sensors to recognize objects and perceive scene depth, resulting in blurred images and reduced contrast. Moreover, the transient and random nature of dynamic interferences (such as sudden glare and dense raindrops) makes it difficult for models to adjust detection strategies in real time, further exacerbating performance fluctuations in extreme scenarios.

4.2 Future Prospects

Dedicated datasets may be constructed to train network models across diverse scenarios, aiming to identify suitable baseline models that advance future research on autonomous driving vehicle detection systems.

An efficient, highly generalizable, and versatile system should be designed, obviating the need to configure different scale groups for varying scenarios. This would enhance domain adaptation capabilities for tasks like object detection, semantic segmentation, and depth estimation, thereby bolstering the robustness of visual perception in rainy and foggy environments.

Research ought to focus on models grounded in weakly supervised or fully unsupervised learning. Fully supervised deep learning models are data-driven, confronting challenges of insufficient training data; moreover, their training depends on annotated data, which is both costly and technically intensive.

Vehicle detection is achieved by integrating computer algorithms with data perceived by intelligent vehicle sensors, yet relying solely on sensors is insufficient for this task. A multi-sensor fusion system may be developed to integrate sensor-acquired data while satisfying real-time performance requirements.

In complex scenarios—including multi-lane, multi-angle, and small vehicle scenarios—false and missed detections are prone to occur owing to difficulties in capturing effective feature information. Consequently, it is imperative to enhance vehicle detection methods via cross-scale feature fusion, improve algorithm robustness, and tackle the issue of inadequate generalization in single-scenario models.

Digital twin technology may be leveraged to provide a test environment that closely approximates real-world conditions for algorithm iteration. By introducing relevant datasets and generating multimodal synchronous data, the efficiency of algorithm verification can be enhanced.

5 Conclusions

Due to the combination of atmospheric scattering, dynamic interferences (like glare and raindrops), and unbalanced illumination, low-light scenes (like nighttime, rainy, and foggy weather) significantly impair the performance of environmental perception systems for autonomous driving. This not only significantly impairs the accuracy and robustness of detection algorithms but also serves as a core bottleneck restricting the safe operation of autonomous driving technology across all scenarios, posing a direct threat to driving safety.

The present paper delivers a thorough review of the domain of vehicle detection applied to autonomous driving in low-light scenes: it analyzes the core challenges in scenarios like nighttime, rainy, and foggy weather, clarifying the commonalities and specific characteristics across different scenarios; sorts out scenario-specific technical solutions based on different scenarios, while elaborating on the scenario coverage value of mainstream datasets; finally, it summarizes current bottlenecks such as weak cross-scenario generalization capability and insufficient robustness in extreme environments, and looks forward to cutting-edge directions including cross-scenario adaptive learning.

By integrating the technological evolution trajectory in the field of autonomous driving vehicle detection, this review accurately identifies research gaps, provides technical references, and promotes the development of autonomous driving towards safety and generalization in complex environments.

References

1. Deng, L., Pan, M., Jin, R., Xie, Z.: Night target detection approach based on near infrared image fusion on vehicles. In: 2022 5th International Conference on Pattern Recognition and Artificial Intelligence (PRAI), Chengdu, China, pp. 755–759. IEEE, Piscataway (2022)
2. Wu, Z., Miao, X., Li, W., Yu, H.: Deep learning based nighttime target enhancement detection algorithm for intelligent vehicles. In: 2022 6th CAA International Conference on Vehicular Control and Intelligence (CVCI), Nanjing, China, pp. 1–6. IEEE, Piscataway (2022)
3. Lin, C.-T., Huang, S.-W., Wu, Y.-Y., Lai, S.-H.: GAN-based day-to-night image style transfer for nighttime vehicle detection. IEEE Trans. Intell. Transp. Syst. 22(2), 951–963 (2021)
4. Jia, R.: Research on vehicle detection and tracking algorithm under nighttime scenarios based on deep learning. Dalian Jiaotong University, Dalian (2025)
5. Nisha, U. N., Ranjani, G.: Deep learning based night time vehicle detection for autonomous cars using generative adversarial network. In: 2022 International Conference on Augmented Intelligence and Sustainable Systems (ICAISS), Trichy, India, pp. 336–341. IEEE, Piscataway (2022)
6. Sun, Z.: Key technologies for vehicle detection in rainy and foggy weather [D]. North China Electric Power University, Beijing (2023)
7. Sun, H., Ang, M. H., Rus, D.: A convolutional network for joint deraining and dehazing from a single image for autonomous driving in rain. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, pp. 962–969. IEEE, Piscataway (2019)
8. Kumar, H., Mamoria, P., Dewangan, D. K.: Improving Faster R-CNN for vehicle detection under varying conditions with domain adaptation technique. In: 2025 Fourth International Conference on Power, Control and Computing Technologies (ICPC2T), Raipur, India, pp. 1–6. IEEE, Piscataway (2025)

9. Wang, J.-G., Wan, K.-W., Yau, W.-Y., Pang, C. H., Lai, F. L.: VAGAN: Vehicle-aware generative adversarial networks for vehicle detection in rain. In: 2020 16th International Conference on Control, Automation, Robotics and Vision (ICARCV), Shenzhen, China, pp. 363–368. IEEE, Piscataway (2020)
10. Fu, X., Huang, J., Zeng, D., Huang, Y., Ding, X., Paisley, J.: Removing rain from single images via a deep detail network. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, pp. 1715–1723. IEEE, Piscataway (2017)
11. Wang, H., et al.: YOLOv5-Fog: A multiobjective visual detection algorithm for fog driving scenes based on improved YOLOv5. IEEE Trans. Instrum. Meas. 71, 1–12 (2022)
12. Abbasi, H., Amini, M., Yu, F. R.: Fog-aware adaptive YOLO for object detection in adverse weather. In: 2023 IEEE Sensors Applications Symposium (SAS), Ottawa, ON, Canada, pp. 1–6. IEEE, Piscataway (2023)
13. Rivera Velázquez, J. M., Khoudour, L., Saint Pierre, G., Duthon, P., Liandrat, S., Bernardin, F., Fiss, S., Ivanov, I., Peleg, R.: Analysis of thermal imaging performance under extreme foggy conditions: Applications to autonomous driving. J. Imaging 8(11), 306 (2022)
14. Zhang, S.: Vehicle detection in foggy environment and domain adaptation research based on generative adversarial network. Chongqing University, Chongqing (2021)
15. Wang, D.: Research on deep learning-based vehicle detection algorithms in foggy weather. Shenyang University of Technology, Shenyang (2024)
16. Geiger, A., Lenz, P., Stiller, C., et al.: Vision meets robotics: The KITTI dataset. Int. J. Robot. Res. 32(11), 1231–1237 (2013)
17. Yu, F., et al.: BDD100K: A diverse driving dataset for heterogeneous multitask learning. In: 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, pp. 2633–2642. IEEE, Piscataway (2020)
18. Cordts, M., et al.: The Cityscapes dataset for semantic urban scene understanding. In: 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, pp. 3213–3223. IEEE, Piscataway (2016).