Improved YOLOv8 Object Detection Algorithm for Nighttime Driving Scenarios

Rong Zhang, Song Gao, Pengwei Wang, Binbin Sun, Yuying Fang

Abstract—**To address the challenges in detecting small and occluded targets caused by the decline in visual sensor imaging quality during nighttime driving scenarios, a low-light dynamic object detection algorithm based on YOLOv8 is proposed. First, the Deep Retinex-Net is employed to enhance the original image quality, thereby accentuating the image features. Second, the PAN-FPN module in the YOLO neck network is replaced with a BiFPN fusion module, reducing the loss of feature information in intermediate layers and improving small target detection performance. Finally, a multiframe image fusion module is designed to facilitate feature fusion between historical and current frames, enhancing the detection of occluded targets. To validate the proposed algorithm, both dataset and real vehicle tests are conducted. The results demonstrate that, in nighttime driving scenarios using the Waymo dataset, the mAP of the improved model is 77.3%, marking a 2.8% improvement over the baseline model. In real-world nighttime driving scenarios, the detection accuracy of the improved model is 91.49%, reflecting a 5.78% improvement over the baseline. The proposed algorithm exhibits strong target detection performance and robustness, making it highly suitable for target detection tasks in intelligent connected vehicles.**

Index Terms—**YOLOv8, Multiframe image fusion, Object detection, Nighttime driving scenarios**

I. Introduction

OBJECT detection is a core technology in environment perception, providing accurate information for vehicles to ensure driving safety [1]. However, lighting conditions significantly impact detection accuracy in practical applications [2]. While high detection accuracy can be achieved under good daytime lighting scenarios, low light and noise in nighttime driving scenarios result in dark images with indistinguishable details. Feature extraction becomes especially challenging for occluded and small targets, reducing detection accuracy and impacting driving safety [3], [4]. In driving scenarios, both traditional and deep learning-based object detection methods are employed to identify and locate targets of interest in images or videos [5], [6]. The traditional object detection methods [7], [8] are computationally simple and offer good interpretability, making them suitable for target detection tasks in simpler environments. However, these methods suffer from slow detection speeds and rely on manually designed feature extraction, which limits their performance in complex scenarios. In contrast, deep learning-based object detection methods [9], [10] offer superior feature learning capabilities, robustness and adaptability, making them the mainstream approach in modern object detection.

Currently, deep learning-based object detection methods can be broadly categorized into two-stage and single-stage detection algorithms [11]. The two-stage detection algorithms, exemplified by the R-CNN series [12], [13], offer high detection accuracy. However, the training and detection speeds are relatively slow, making them challenging to implement in real-time vehicle applications. In contrast, single-stage detection algorithms, such as the YOLO series [14], [15] are widely adopted in object detection tasks due to their higher detection accuracy and faster detection speeds. As object detection technology continues to evolve, researchers have increasingly focused on addressing key challenges in real-world scenarios, including image enhancement, small target detection, and occluded target detection. Overcoming these challenges will further enhance the accuracy, robustness, and adaptability of object detection algorithms across diverse driving scenarios.

Existing studies tend to lose the texture and edge details during the image enhancement process, which results in blurred target features and amplified background noise. Consequently, the detection accuracy of the algorithm is compromised. Wang et al. [16] proposed an image adaptive enhancement network, which improves the quality of low-light images and the algorithm's detection accuracy by incorporating the channel attention mechanism SE-Net and constructing a feature enhancement network. However, the method's detection ability for small targets remains limited. Miao et al. [17] applied the optimal MSR algorithm to enhance the original nighttime images for driving scenarios.

Manuscript received March 29, 2025; revised May 26, 2025.

This work was supported by National Natural Science Foundation of China (52102465), Shandong Province Major Science and Technology Innovation Project (2023CXGC010111), Natural Science Foundation of Shandong Province (ZR2022MF230), Shandong Youth Fund (ZR2021QF039), Small and Medium-sized Enterprise Innovation Capability Improvement Project (2022TSGC2277), General programs of National Natural Science Foundation of China (52475269) and General program of Natural Science Foundation of Shandong Province (ZR2024ME179).

Rong Zhang is a postgraduate student of School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, Shandong 255000, China. (e-mail: zhangrong200010@163.com).

Song Gao is a professor of School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, Shandong 255000, China. (e-mail: gaosong@sdut.edu.cn).

Pengwei Wang is an associate professor of School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, Shandong 255000, China. (corresponding author to provide phone: 13287825788; e-mail: wpwk16@163.com).

Binbin Sun is a professor of School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, Shandong 255000, China. (e-mail: sunbin_sdut@126.com).

Yuying Fang is a postgraduate student of School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, Shandong 255000, China. (e-mail: 1592911361@qq.com).

Although the model's AP is improved, the problem of feature information loss in the model's intermediate layers persists. Jiang et al. [18] introduced a self-correcting illumination module to enhance low-light image quality and proposed a dynamic feature extraction module to capture global context information. The approach effectively reduces false detections and missed detections, but the model struggles to balance attention between image details and global context information. Additionally, the network's generalization ability requires improvement. Huang et al. [19] presented a method that combines a multiscale image feature enhancement module with a target feature enhancement module, effectively improving object detection accuracy under low-light conditions. However, the dataset used lacks sufficient representation of rea-world driving scenarios, and the model's generalization ability needs enhancement. Zhang et al. [20] utilized the Retinex algorithm for image enhancement and applied void convolution to reduce network parameters, improving the accuracy and speed of vehicle detection at night. Despite these improvements, the model still struggles with detecting occluded targets. Due to insufficient feature representation for small targets and limited handling of occluded targets, existing algorithms perform poorly in target detection tasks. To further improve detection performance, various strategies have been proposed. Cao et al. [21] reconstructed the backbone network, introduced an attention mechanism, and optimized the loss function, improving small target detection. However, the algorithm's performance on occluded targets has not been significantly enhanced. Wu et al. [22] designed a multiframe feature fusion module to propagate historical frame features to the current frame for fusion, thereby improving detection accuracy for occluded targets. However, the method is not optimized for nighttime driving scenarios, and small target detection remains inadequate. Yang et al. [23] proposed a feature sensing field fusion module and added a P2 detection head with a shallower feature layer, enhancing the model's detection capabilities for both small and occluded targets. However, the model has a large number of parameters, which hampers its detection speed and makes real-time application challenging.

In summary, several challenges still exist in the object detection algorithms for nighttime driving scenarios. Low lighting significantly degrades the quality of captured images. Existing image enhancement techniques struggle to adaptively improve image quality based on real-time conditions. Moreover, current YOLO models often lack sufficient feature information for effective object detection. Due to weak feature representation and partial loss of important features, these models face difficulties in accurately detecting small and occluded targets. Therefore, the detection accuracy of the model requires further improvement.

To address the aforementioned challenges, an improved YOLOv8 object detection algorithm for nighttime driving scenarios is proposed. First, Deep Retinex-Net is employed to enhance the feature information in low-light images. The approach achieves complex nonlinear modeling through convolutional neural networks, which makes the reflection component and illumination component of the image be accurately separated. During the image enhancement process, a rationality constraint is applied to separated components, effectively suppressing background noise while preserving detailed information, thereby producing more naturally enhanced images. Second, the PAN-FPN module in the YOLOv8 neck network is replaced with the BiFPN fusion module to mitigate the loss of feature information in the intermediate layers of the model. The replacement improves the model's detection performance on small targets. Additionally, redundant node connections are eliminated to streamline the network structure and reduce model parameters. The bidirectional feature fusion mechanism facilitates full interaction between high-level semantic information and low-level detail, enhancing the feature representation of small targets. Furthermore, a learnable weighting mechanism is introduced to dynamically adjust the weights of different feature layers, optimizing the fusion process. Finally, a multiframe image fusion module is proposed, which leverages target consistency and background difference between frames to achieve feature fusion. By integrating historical and current frames, the representation of target feature information is strengthened, enabling enhanced perception of dynamic target behavior and improving the detection of occluded targets. Compared to the baseline YOLOv8 model, the improved algorithmic model significantly enhances the quality of images captured in nighttime driving scenarios. The introduction of the BiFPN fusion module and the multiframe image fusion module addresses the limitations of previous models, such as poor adaptability to nighttime driving scenarios and reliance on single-frame detection. The detection performance of small target and occluded target is further improved. Experimental results demonstrate that the improved algorithmic model achieves superior detection accuracy.

II. The YOLOv8 Principle

You Only Look Once (YOLO) series of algorithms offer high real-time performance and detection accuracy, making them widely applicable in modern autonomous driving scenarios. The algorithms are frequently deployed in real vehicles to carry out object detection tasks. YOLOv8 primarily consists of three parts: Backbone, Neck and Head, as shown in Fig. 1.

Compared to previous versions, the backbone network of YOLOv8 adopts an enhanced structure based on CSPNet (Cross Stage Partial Network). The design improves the model's representation ability and inference speed, while enhancing its ability to extract features from multiscale targets in complex scenarios. The model also demonstrates more stable performance on simulation datasets. The neck network of YOLOv8 incorporates PAN-FPN module, enhancing the model's capacity to capture contextual information. It enables effective identification and localization of targets at various scales, reducing both false and missed detections. The detection head adopts a Decoupled-Head structure. It separates classification and regression tasks into independent branches, minimizing mutual interference.

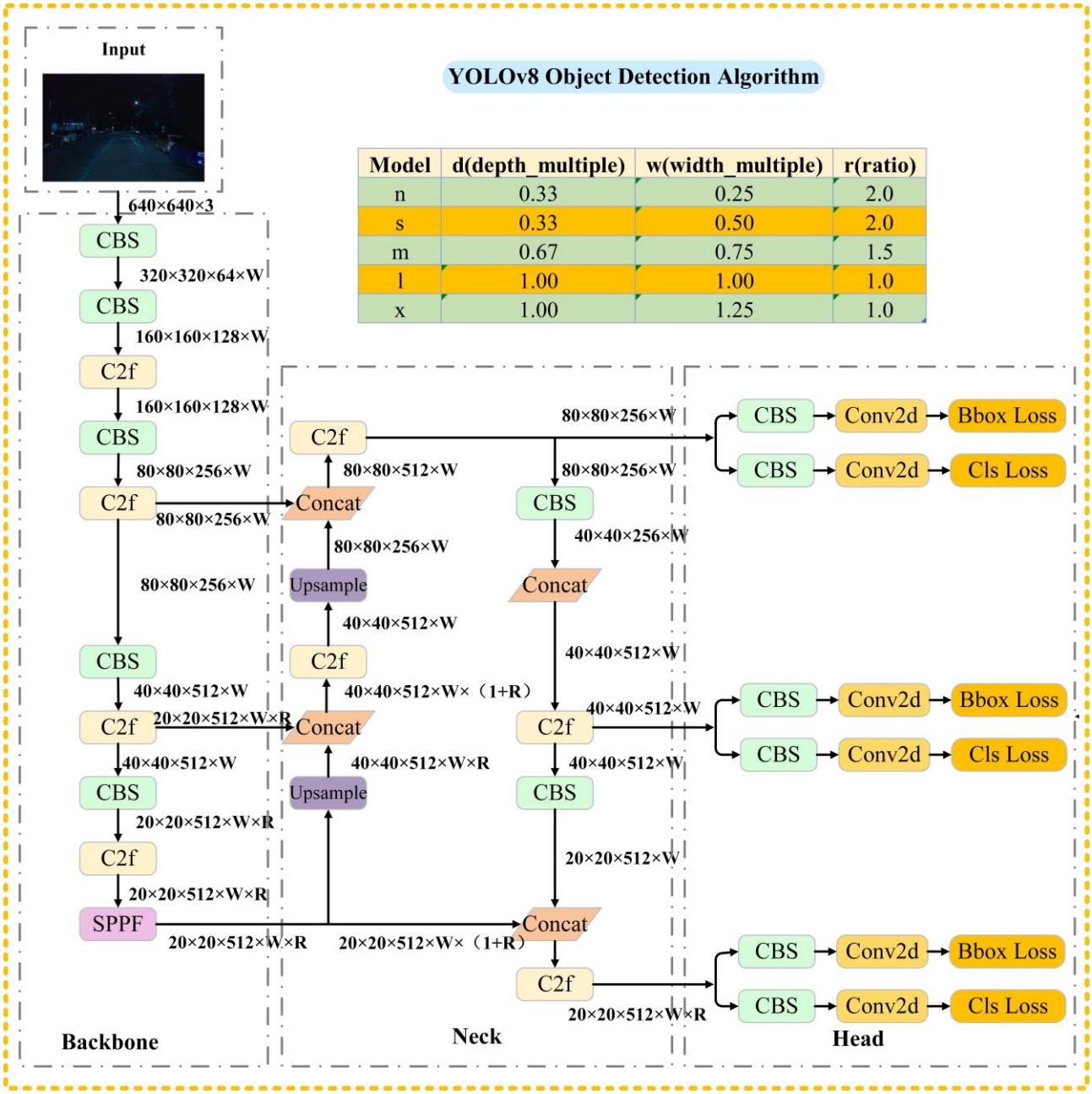


Fig.1 YOLOv8 Object Detection Algorithm

This structure enhances the model's adaptability to diverse driving scenarios. It also enables fast and accurate object detection in complex environments. Meanwhile, YOLOv8 offers five model scales, N, S, M, L, and X, based on the scaling factors. The appropriate scale can be flexibly selected based on available hardware resources. This flexibility reduces the difficulty of deploying the model in real vehicle experiments.

III. Improved YOLOv8 Object Detection Algorithm

YOLOv8 demonstrates strong object detection performance in most autonomous driving scenarios. However, low light conditions and image noise degrade image quality in the nighttime driving scenarios. As a result, object detection algorithms struggle to extract effective features. Additionally, detail loss directly impacts the recall rate and accuracy of small object detection. Furthermore, the visible area of occluded objects becomes even smaller. The phenomenon of missed and false detection is increased. To address these challenges, an improved YOLOv8 object detection algorithm is proposed. The proposed improvements consist of three parts, as shown in Fig. 2. Deep Retinex-Net is introduced to improve the quality of images captured by the camera. BiFPN fusion module and multiframe image fusion module are incorporated to enhance detection performance for small and occluded objects.

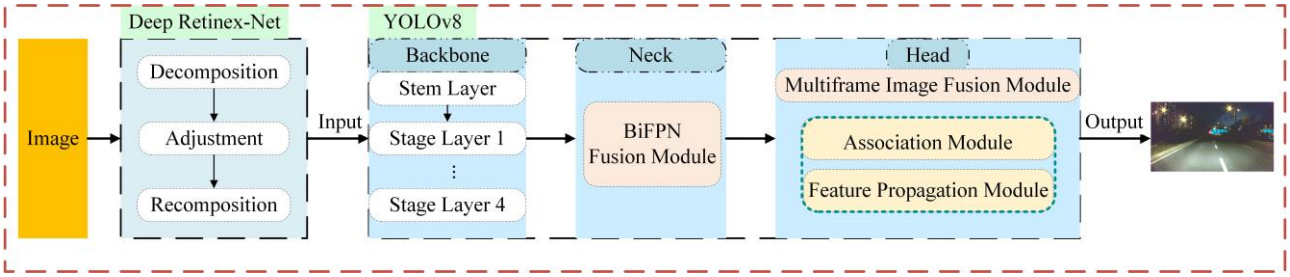


Fig. 2 Improved YOLOv8 Object Detection Algorithm

A. *Deep Retinex-Net Image Enhancement Network*

The image information is influenced by the reflection component *R* and the illumination component *I*. In nighttime driving scenarios, it is difficult for the Retinex algorithm to define the appropriate image decomposition constraints. Moreover, manually designed parameters limit the capability of the Retinex algorithm due to model constraints. The limitations of the Retinex algorithm contribute to false and missed detection in the YOLOv8 object detection process. The principle of the Retinex algorithm is shown in Fig. 3. To address these limitations, Deep Retinex-Net image enhancement network is introduced in this study. The image enhancement process is divided into three stages including decomposition, adjustment and reconstruction as shown in Fig. 4. A data-driven approach is employed to solve the problem of the Retinex algorithm in which the reflection component *R* and the illumination component *I* are fixed. The enhanced images are more suitable for object detection in the nighttime driving scenarios.

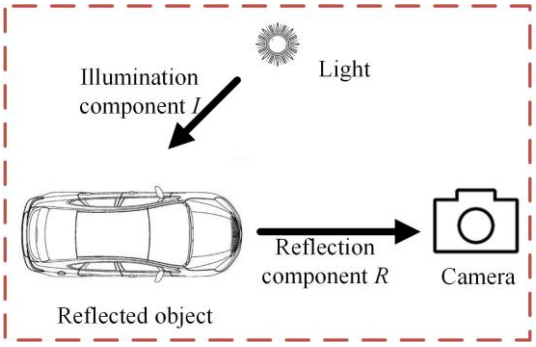


Fig. 3 Retinex algorithm schematic diagram

In the decomposition stage, the initial features of the low-light image are extracted through a 3 × 3 convolutional layer. Then, the initial separation of the illumination component *R* and the reflection component *I* is achieved by nonlinear transformations using multiple 3 × 3 convolutional layers and ReLU activation function. Finally, a 3 × 3 convolutional layer and Sigmoid function are used to normalize the image features. In the adjustment stage, the reflection component is first denoised to further optimize detail performance, and the optimized reflection component $\hat{R}_{low}$ is generated. Meanwhile, local feature information is extracted by a 3 × 3 convolutional layer and ReLU activation function. Multistage down-sampling and up-sampling modules are designed to capture local illumination information and reconstruct the features. Then, channel compression is performed using a 1 × 1 convolution layer to reduce the computational complexity. Finally, features are reconstructed by a 3 × 3 convolutional layer to generate the enhanced illumination components $\hat{I}_{low}$. In the reconstruction stage, the optimized reflection component $\hat{R}_{low}$ and enhanced illumination component $\hat{I}_{low}$ are combined by element level multiplication operation to reconstruct the image. Deep Retinex-Net automatically learn the characteristics of illumination and reflection components of images through convolutional neural network, which no longer relies on artificial assumptions. Through multilevel feature extraction, Deep Retinex-Net can effectively avoid image distortion caused by inaccurate illumination estimation. In nighttime driving scenarios, the traditional Retinex algorithm cannot cope with the actual changes in lighting conditions. And the cumbersome image processing steps will also affect the real-time performance of the system. Therefore, the Deep Retinex-net image enhancement network is proposed to improve image quality by deep learning. The robustness and real-time performance of the system are guaranteed. Deep Retinex-Net has been verified on the Waymo, LIME, LOL (Low-Light paired) and other data sets. The reconstructed images exhibit significant improvements in both brightness and detail. It not only retains the inherent characteristics of the nighttime driving scenarios, but also enhances the lighting conditions of the image. Table I shows the algorithm description of Deep Retinex-Net image enhancement network.

Table I
DEEP RETINEX-NET IMAGE ENHANCEMENT NETWORK

Deep Retinex-Net image enhancement network
Input: Low light image S_{low}
Step 1: Image decomposition
R_{low}, I_{low} = Decom_Net (S_{low})
Decompose low light image into reflection component and illumination component
Step 2: Image adjustment
$\hat{R}_{low}$ = Denoising_Operation (R_{low})
Denoise the reflected component
$\hat{I}_{low}$ = Enhance-Net (I_{low})
Enhanced the illumination component
Step 3: Image reconstruction
$\hat{S}_{low} = \hat{R}_{low} * \hat{I}_{low}$
#Multiply the illumination component and reflection component at element level
Output: Enhanced image $\hat{S}_{low}$

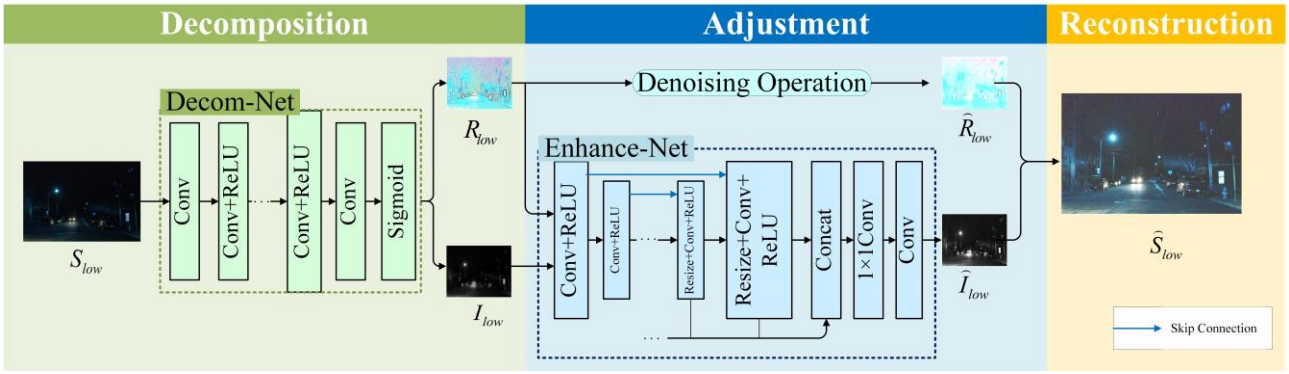


Fig. 4 Deep Retinex-Net image enhancement network

B. *BiFPN Fusion Module*

The neck network of YOLOv8 adopts a PAN-FPN module as shown in Fig. 5. However, the PAN-FPN module relies on a unidirectional flow of information to transfer features, which results in insufficient fusion between local and global semantic information. Moreover, the fusion weight cannot be dynamically adjusted according to the lighting characteristics of the nighttime driving scenarios. It results in inadequate reuse and selective attention to small target features. To address these limitations, BiFPN fusion module is introduced in this study. The bidirectional feature cross scale connection structure and adaptive feature fusion module are used to improve the feature fusion capability of the model. The feature information of small targets is accurately captured and expressed. As shown in Fig. 6, the BiFPN fusion module facilitates comprehensive fusion of top-down and bottom-up feature information. The dynamic weighting mechanism enables the model to adapt to the characteristics of small targets in nighttime driving scenarios. The detection accuracy of the model is improved.

Compared to the PAN-FPN module, the BiFPN fusion module eliminates input nodes that contribute minimally to feature fusion. Additionally, an input path is additionally added to the same hierarchy of input and output nodes, which enhances feature fusion without significantly increasing computational cost. At the same time, each bidirectional path is regarded as a feature network layer. By stacking these layers multiple times, the model achieves deeper and more refined feature fusion. The weighting formula for each output layer feature is defined as follows:

$$P_7^{out} = Conv\left(\frac{P_7^{in} \cdot W_{71} + Resize(P_6^{out}) \cdot W_{72}}{W_{71} + W_{72} + \epsilon}\right) \tag{1}$$

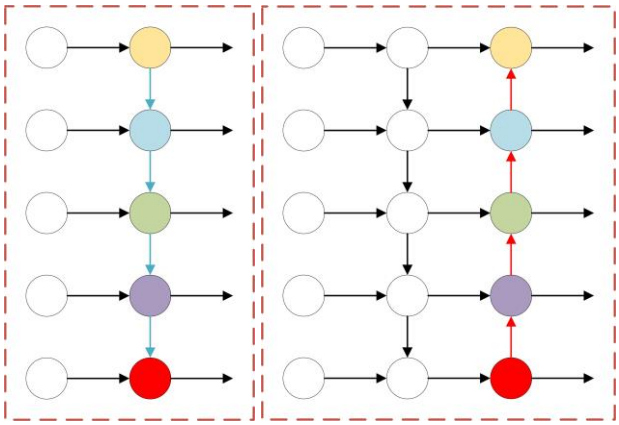
$$P_6^{out} = Conv\left(\frac{P_6^{in} \cdot W_{63} + P_6^{td} \cdot W_{64} + Resize(P_5^{out}) \cdot W_{65}}{W_{63} + W_{64} + W_{65} + \epsilon}\right) \tag{2}$$

$$P_5^{out} = Conv\left(\frac{P_5^{in} \cdot W_{53} + P_5^{td} \cdot W_{54} + Resize(P_4^{out}) \cdot W_{55}}{W_{53} + W_{54} + W_{55} + \epsilon}\right) \tag{3}$$

$$P_4^{out} = Conv\left(\frac{P_4^{in} \cdot W_{43} + P_4^{td} \cdot W_{44} + Resize(P_3^{out}) \cdot W_{45}}{W_{43} + W_{44} + W_{45} + \epsilon}\right) \tag{4}$$

$$P_3^{out} = Conv\left(\frac{P_3^{in} \cdot W_{31} + Resize(P_4^{td}) \cdot W_{32}}{W_{31} + W_{32} + \epsilon}\right) \tag{5}$$

where P_i^{out} denotes the output feature of layer i, $Conv$ denotes the convolution operation for feature processing, P_i^{in} denotes the input feature of layer i, W_{ij} denotes the learnable fusion weight, P_i^{td} denotes the intermediate feature of layer i in the top-down path, and *Resize* denotes the up-sampling or down-sampling in the resolution matching process. $\epsilon = 0.0001$.



(a)PAN (b)FPN

Fig. 5 PAN-FPN fusion module

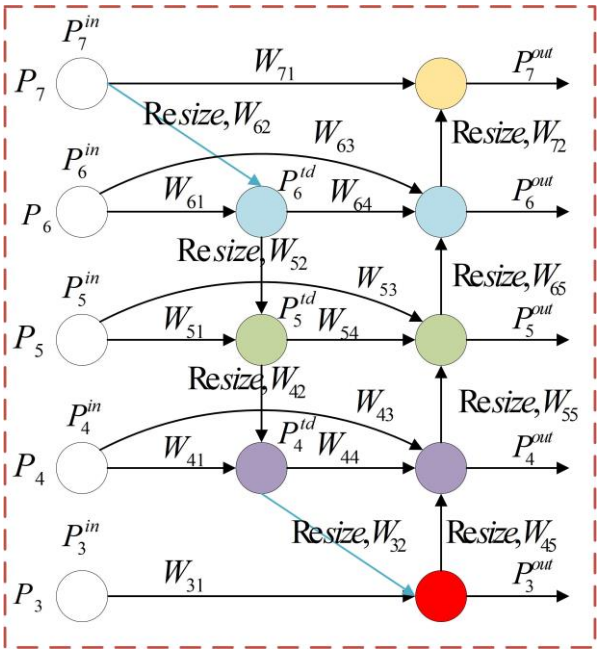


Fig. 6 BiFPN fusion module

The BiFPN fusion module preserves the advantages of multiscale feature fusion. The computational complexity is reduced by the design of bidirectional feature flow and path optimization. It maintains a balance between high-level

semantic information and low-level detail information. Moreover, the fusion weights of different feature layers can be dynamically adjusted. It performs better in the task of small target detection in the nighttime driving scenarios. Therefore, BiFPN fusion module is introduced in this study to further improve the feature fusion efficiency and multiscale target processing capability. It can also achieve higher target detection accuracy. Table II shows the algorithm description of BiFPN fusion module.

Table II
BIFPN FUSION MODULE

BiFPN fusion module

Input: Feature map x
Initialize: Weight parameter W, Splice dimension d,
$\epsilon = 0.0001$
Step 1: Normalize weights
weight = w / (sum(w)+ ϵ)
Step 2：Weighted fusion feature map
x[i] = weight[i] * x[i]
# Multiply the normalized weights with each input feature map one by one. 	
Step 3：Splice feature
# Splice weights along the specified dimension d 	
Output = concat (x, d)

C. Multiframe Image Fusion Module

In nighttime driving scenarios, images captured by the camera become blurred due to the target motion, and the features of the occluded targets are further weakened. Models that rely on single-frame images for target detection are prone to false detection and missed detection, leading to reduced detection accuracy. To address these problems, a multiframe image fusion module is proposed as shown in Fig. 7. The multiframe image fusion module consists of an association module and a feature propagation module, which work together to fuse features from both historical and current frames. By incorporating temporal information, the module mitigates the motion blur problem and captures the motion trajectory of targets. The multiframe image fusion module effectively reduces missed detections caused by occlusions, thereby improving the overall accuracy of the detection model.

The association module calculates the *IoU* between the detection frames $\alpha = \{a_1, a_2 \ldots, a_m\}$ and $\beta = \{b_1, b_2 \ldots, b_n\}$ of the history frames and the current frames. The history and current frames are matched based on the similarity between detected targets. The *IoU* is computed using the following formula:

$$IoU = \frac{|A \cap B|}{|A \cup B|} \tag{6}$$

where A denotes the detection region of the history frame and B denotes the detection region of the current frame.

If there is no intersection between A and B, $A \cap B = 0$, $IoU = 0$. It does not account for the distance between A and B. At the same time, it is not sensitive enough to positional deviation of the history and current frames. It is possible that the *IOU* is equal but the overlap degree is different. The condition is shown in Fig. 8.

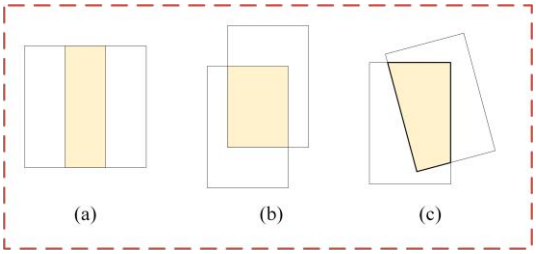


Fig. 8 Equal *IoU* but different overlap

GIoU is introduced to address the low sensitivity of *IoU* to the offset between the history and current frames. It better reflects the overlap degree of history and current frames. The *GIoU* is computed using the following formula:

$$GIoU = IoU - \frac{|C - (A \cup B)|}{|C|} \tag{7}$$

where C denotes the smallest rectangular region containing A and B.

The *GIoU* threshold is set to 0.5. According to the detection region A and B of the history frame and the current frame, the similarity of the target detection frame can be calculated.

$$M = \{(\alpha, \beta) | GIoU(\alpha, \beta) \geq 0.5\} \tag{8}$$

Finally, the nearest matching method is used to traverse and compare the similarity between each feature in the history frame and the current frame. The most similar feature is selected to complete the matching process.

The feature propagation module is designed to propagate features from history frames to the current frame. It calculates the motion offsets of the target prediction frames for both the history frame and the current frame based on feature matching. Weighted propagation is used by combining the similarity *M* calculated in the association module. The feature representation of the current frame is enhanced by fusing the features.

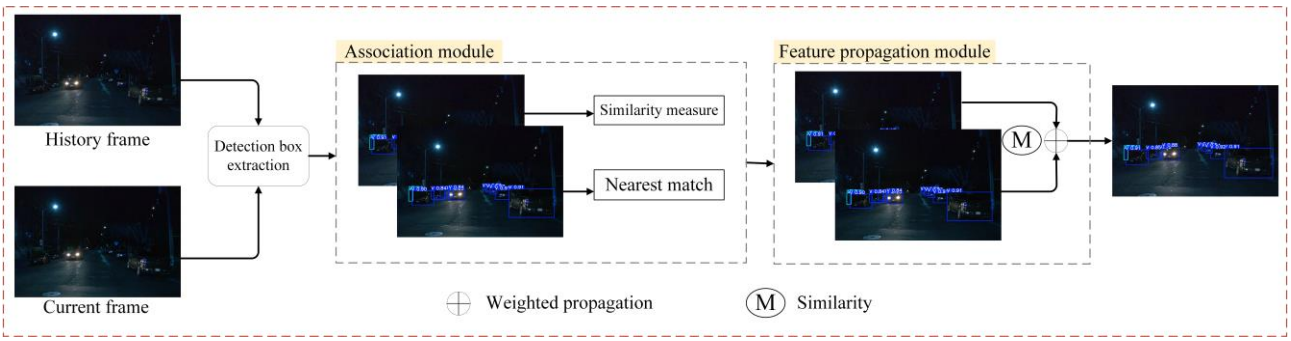


Fig. 7 Multiframe image fusion module

$$\alpha' = \lambda \cdot \alpha + (1-\lambda)\beta \tag{9}$$

where α' denotes the fused target detection frame. λ denotes the weight parameter, which is determined based on the similarity between the target detection frames.

In this study, YOLOv8 object detection algorithm is improved in three aspects. First, Deep Retinex-Net is employed to enhance low-quality images, thereby the quality of the images captured by the camera in the nighttime driving scenarios is improved. It meets the quality requirements of the model for the input image. Second, BiFPN fusion module is introduced into the YOLOv8 object detection algorithm, significantly enhancing the detection of small targets. The module incorporates bidirectional feature fusion and adaptive weighting mechanism, which strengthens the recognition accuracy of small targets in complex backgrounds. Finally, a multiframe image fusion module is proposed. It is able to fuse the features of history frames and current frames to reduce the impact of image blurring on the target detection task. As a result, the detection accuracy for occluded targets is improved, while the occurrences of false and missed detections are reduced.

IV. Experiments

To verify the effectiveness and applicability of the improved YOLOv8 object detection algorithm in nighttime driving scenarios, both simulation experiments and real vehicle experiments were conducted. The simulation experiments include comparative experiments and ablation experiments, which are used to evaluate the object detection performance of the algorithm. The real vehicle experiments assess the practical applicability and robustness of the proposed algorithm in an actual driving environment. By combining both types of experiments, the overall performance of the improved YOLOv8 object detection algorithm is thoroughly validated.

A. Simulation Experiments

Selection and Use of Datasets

Conventional datasets are limited in terms of scale and diversity for nighttime driving scenarios. In this study, the Waymo dataset proposed by Sun et al. [24] is utilized for both training and inference. The dataset contains a large number of nighttime driving scenarios, which can meet the experimental requirements. Meanwhile, the duration of each scene is 20 seconds, providing continuous image sequences of historical and current frames for the multiframe image fusion module. It ensures the feasibility of optimized operation. The camera specifications of Waymo dataset are shown in Table III.

Table III
Waymo Camera Specifications

	F	FL, FR	SL, SR
Size	1920×1280	1920×1280	1920×1040
HFOV	±25.2°	±25.2°	±25.2°

During the training process, vehicle is labelled as ‘V’, pedestrian as ‘P’ and cyclist as ‘C’. A total of 10116 images are selected from the dataset for model training. There are 8133 images in the training set, 991 images in the validation set and 992 images in the test set.

Experimental Environment

The simulation experiments are conducted on the Ubuntu 20.04 operating system using the PyTorch deep learning framework. The specific configuration is shown in Table IV. In the training process, YOLOv8n is adopted as the base model for training. The input image resolution is set to 640×640, and the SGD optimiser is used. The initial learning rate is 0.01, the weight attenuation coefficient is 0.005, the momentum size is 0.9, and the batchsize is 16. A total of 300 epochs are iterated.

Table IV
Experimental Environment

Category	Version
CPU	Intel® Core™ i7-14700KF CPU @ 5.6 GHz × 28
RAM	32G
GPU	NVIDIA GeForce RTX 4070 Ti Super
VRAM	16GB
Python	Python 3.8.20
Pytorch	Pytorch 1.11.0
CUDA	CUDA 11.3
cuDNN	cuDNN 8.9.5

Evaluation Indicators

In this study, mAP (mean Average Precision) and FPS are adopted to validate the model performance. Where mAP is used to evaluate the model detection accuracy, and FPS is used to evaluate the model detection speed. The calculation formulas for the evaluation metrics are as follows:

$$P = \frac{TP}{TP + FP} \tag{10}$$

$$R = \frac{TP}{TP + FN} \tag{11}$$

where *TP* denotes the number of positive samples that were correctly identified, *FP* denotes the number of negative samples that were incorrectly identified as positive samples, *FN* denotes the number of positive samples that were incorrectly identified as negative samples, *P* denotes the precision rate, and *R* denotes the recall rate.

$$AP = \int_0^1 Pd(R) \tag{12}$$

$$mAP = \frac{1}{3}\sum_{i=1}^{3} AP_i \tag{13}$$

where $AP_i\,(i = 1,2,3)$ represents the accuracy rate of single category prediction for vehicle, pedestrian and cyclist respectively. mAP represents the average of the three categories of AP for vehicle, pedestrian and cyclist. It can reflect the detection accuracy of the model.

Deep Retinex-Net Comparison Experiments

In this study, Deep Retinex-Net is applied to process the images captured by the camera in nighttime driving scenarios. The image quality is improved, and feature information is enriched, resulting in an enhanced image that meets the input requirements of the YOLOv8 algorithm.

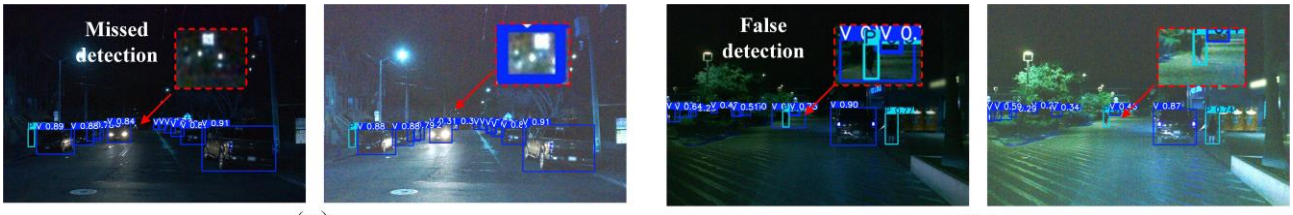


Fig. 9 Comparison of model detection before and after Deep Retinex-Net improvement

In the experiment, the improved YOLOv8 object detection algorithm is applied to both the original and enhanced images. The detection results, shown in Fig. 9, are categorized into two groups: (a) original images on the left and (b) enhanced images on the right. The enhanced images avoid overexposure and exhibit significantly improved illumination. Detail loss in highlighted areas is mitigated, and target edges in dark, reflective regions appear sharper. Deep Retinex-Net effectively separates the reflective component of the image, resulting in a more balanced color representation. Meanwhile, noise suppression makes the enhanced image clean, and the contrast in the target area is significantly improved.

In group (a), the original image suffers from poor lighting conditions, resulting in a missed detection of the farthest vehicle in the left lane. After enhancement with Deep Retinex-Net, the improved YOLOv8 object detection algorithm successfully detects the farthest vehicle in the left lane. In group (b), a false detection occurs in the original image, where the improved YOLOv8 object detection algorithm erroneously identifies a single vehicle in the distance as two separate vehicles. After the original image is enhanced by Deep Retinex-Net, the improved YOLOv8 object detection algorithm eliminates the false detection. These results demonstrate that Deep Retinex-Net significantly improves image quality in challenging nighttime driving scenarios. By enhancing visual clarity and dynamic range, it helps reduce both missed and false detections that are commonly caused by insufficient illumination and poor image fidelity. Consequently, the integration of Deep Retinex-Net with the object detection algorithm enhances the reliability and accuracy of object detection in low-light environments.

Comparative Experiments of Different Object Detection Models

In this study, the proposed algorithmic model is evaluated alongside other object detection models in the nighttime driving scenario to better assess the performance of the improved YOLOv8 object detection algorithm. The results of the tests are presented in Table V.

Table V
COMPARATIVE EXPERIMENTS OF DIFFERENT OBJECT DETECTION MODELS

Model	Parameter/M	mAP0.5/%	FPS
YOLOv5s	7.0	67.1	91
YOLOv8-DEL	2.8	68.0	112
YOLOv8s	11.1	75.0	84
FE-YOLOv8	11.8	75.6	82
ours	5.2	77.3	98

First, the proposed model in this study has 5.2 million parameters, which is higher than the YOLOv8-DEL model. It indicates that the improved model has stronger learning and feature expression capabilities. The number of parameters in the proposed model is smaller than that of the YOLOv5s, YOLOv8s and FE-YOLOv8 models, indicating that the improved model is less likely to overfit during training. As a result, its generalization ability is enhanced. Second, the mAP of the proposed model is 77.3%, which is higher than that of the YOLOv5s, YOLOv8-DEL, YOLOv8s, and FE-YOLOv8 models. It demonstrates that the detection accuracy of the improved model is improved. The algorithm performs particularly well in nighttime driving scenarios, effectively reducing false and missed detections. Finally, the FPS of the proposed model is 98. Although it is lower than YOLOv8-DEL model, it still meets the practical requirements for real-time detection. The improved YOLOv8 object detection algorithm strikes a usable balance between accuracy and real-time performance. Therefore, the improved model demonstrates superior detection performance compared to other models. The effectiveness of the proposed algorithm is validated.

Ablation Experiments

In this study, ablation experiments are conducted to further validate the effectiveness of the BiFPN fusion module and the multiframe image fusion module. The experimental results are presented in Table VI, where √ indicates the inclusion of a module.

Table VI
ABLATION EXPERIMENT

Base model	BiFPN fusion module	Multiframe image fusion module	Parameter /M	mAP0.5 /%	FPS
YOLOv8n			3.2	74.5	103
	√		4.2	76.4	108
		√	4.0	76.8	97
	√	√	5.2	77.3	98

Table VI demonstrates that both improved modules can enhance the performance of the object detection algorithm. In the YOLOv8n base model, the PAN-FPN module in the neck network is replaced with a BiFPN fusion module. The loss of feature information in the middle layer is reduced. The mAP of the model is increased by 1.9%. When a multiframe image fusion module is added to the YOLOv8n base model, it fuses current and historical frames to address the problem of feature information loss. The mAP of the model is increased by 2.3%. Adding both the BiFPN fusion module and the multiframe image fusion module simultaneously increases the mAP by 2.8%, while also reducing the occurrence of false and missed detections. Although the FPS of the improved model decreases, it still meets the real-time object detection requirements.

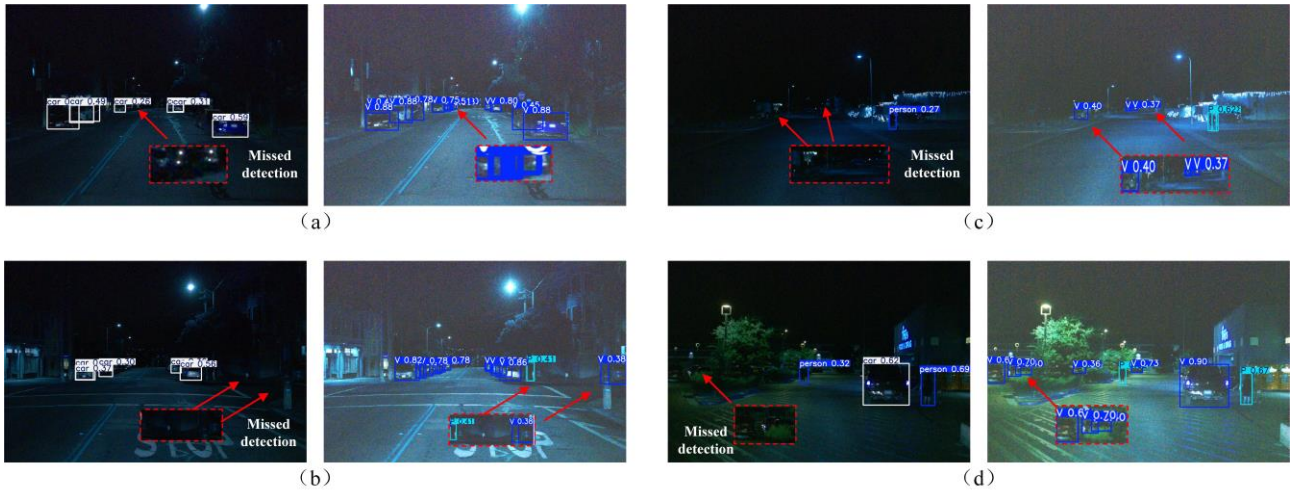


Fig. 10 Comparison of detection results before and after model improvement

Visualization Analysis

The improved YOLOv8 object detection model and the base model were used to perform visual detection analysis in the nighttime driving scenarios of Waymo dataset. The detection results are shown in Fig. 10. In the group (a), (b), (c), and (d), the detection results of the base model are shown on the left, while the detection results of the improved model are shown on the right. The lighting conditions of the nighttime driving scenarios are poor, resulting in low-quality images captured by the camera. The base model is highly susceptible to miss detection. However, the improved YOLOv8 object detection model utilizes Deep Retinex-Net to enhance the images. These high-quality images provide richer feature information for the detection model, significantly reducing the occurrence of missed detections. Meanwhile, the improved YOLOv8 object detection model shows greater improvement in detecting small and occluded targets compared to the base model. The object detection performance in the nighttime driving scenarios is effectively improved.

B. *Real Vehicle Experiments*

Real vehicle experiments are performed to evaluate the performance of the improved YOLOv8 object detection algorithm in a nighttime driving scenario in Zhangdian District, Zibo City. The Haval H7 autonomous vehicle serves as the test platform, which is equipped with a RERVISION camera (with a resolution of 1920×1080), a Velodyne 32-line LiDAR, a combined GPS-IMU navigation system, and an industrial-grade control computer. The corresponding Ubuntu operating system is equipped. These components are capable of meeting the requirements for autonomous driving algorithm validation. The RERVISION camera is calibrated to ensure the geometric accuracy of the images. An 8×10 checkerboard calibration board is used with a single grid length of 66 mm. A total of 20 checkerboard images are selected. The real vehicle experiment is shown in Fig. 11.

Fig. 11 Real vehicle experiment

Fig. 13 Comparison diagram of object detection results in real vehicle experiments

As illustrated in Fig. 12, the overall mean reprojection error of the camera calibration is 0.15, with the maximum mean reprojection error not exceeding 0.21. These results indicate that the calibration is reliable and meets the accuracy requirements for subsequent visual processing tasks.

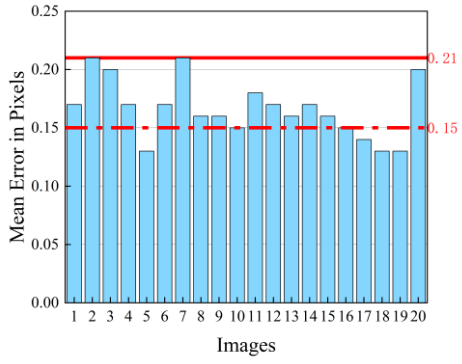


Fig. 12 Mean error diagram

The internal matrix of the camera is:

$$K = \begin{bmatrix} 961.94 & 0.50 & 326.79 \\ 0 & 959.26 & 189.16 \\ 0 & 0 & 1 \end{bmatrix} \quad (14)$$

The radial distortion factor of the camera is:

$$\begin{bmatrix} k_1 & k_2 & k_3 \end{bmatrix} = \begin{bmatrix} -0.4171 & -0.1231 & 4.4841 \end{bmatrix} \quad (15)$$

The tangential distortion factor of the camera is:

$$\begin{bmatrix} p_1 & p_2 \end{bmatrix} = \begin{bmatrix} 0.0044 & 0.0026 \end{bmatrix} \quad (16)$$

The video data of the nighttime driving scenarios in Zhangdian District is recorded by the industrial-grade control computer onboard the test platform. Using previously obtained calibration parameters, distortion correction is applied to the original images to ensure geometric accuracy. The final detection results are presented in Fig. 13. Compared with the base model, the improved model leverages Deep Retinex-Net to extract richer image information and employs the BiFPN and multiframe fusion modules to enhance the feature representation of small and occluded targets. In the nighttime driving scenario, the improved model demonstrates significantly enhanced performance compared to the base model, particularly in detecting occluded and small targets. The base model achieves a target detection accuracy of 85.71%, while the improved model attains an accuracy of 91.49%, representing a 5.78% improvement. The enhancement demonstrates the improved model's superior reliability and robustness in nighttime driving scenarios.

V. Conclusion

In this study, an improved YOLOv8 object detection algorithm is proposed to address the challenges of low detection accuracy for small and occluded targets in nighttime driving scenarios, caused by the poor image quality of cameras. First, low-quality images are enhanced using Deep Retinex-Net, which meets the algorithm's image quality requirements. Second, the PAN-FPN module in YOLOv8 neck network is replaced with a BiFPN fusion module, which reduces the loss of target detail information and improves the detection performance for small targets. Finally, a multiframe image fusion module is designed to perform feature fusion between historical and current frames, enhancing the feature representation of occluded targets and reducing false and missed detections. Experimental results demonstrate that the improved YOLOv8 object detection model achieves a 2.8% increase in mAP compared to the baseline model in the nighttime driving scenario of the Waymo dataset.

Furthermore, the mAP of the proposed model outperforms other state-of-the-art models. Simulation experiments validate the accuracy and generalizability of the improved YOLOv8 object detection algorithm. In real vehicle experiments, the detection accuracy of the improved model is enhanced by 5.78% compared to the baseline model. The results confirm that the improved algorithm not only performs well in the simulation environment but also exhibits strong adaptability and robustness in real-world applications. While the proposed algorithm improves detection accuracy, it increases the number of model parameters and the demand for training resources, which can slow down detection performance. In static scenes or scenes with minimal frame differences, the multiframe image fusion module provides limited enhancement the multiframe image fusion module provides limited enhancement and may introduce redundant computations. In future work, the algorithm's lightweight operation will be explored to enhance efficiency. Additionally, the fusion strategy will be dynamically adjusted based on scene characteristics to prevent unnecessary complexity. The improved YOLOv8 object detection algorithm will further improve the environmental perception ability of intelligent connected vehicles in nighttime driving scenarios.

REFERENCES

[1] T. Diwan, G. Anirudh, J. V. Tembhurne. “Object detection using YOLO: Challenges, architectural successors, datasets and applications,” *Multimedia Tools and Applications*, vol. 82, no. 6, pp. 9243-9275, 2023.

[2] X. X. Zhang, B Story, D Rajan. “Night time vehicle detection and tracking by fusing vehicle parts from multiple cameras,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 7, pp. 8136-8156, 2021.

[3] Y. X. Xiao, A. W. Jiang, J. H. Ye, and M. W. Wang. “Making of night vision: Object detection under low-illumination,” *IEEE Access*, vol. 8, pp. 123075-123086, 2020.

[4] H. Y. Li, B. Q. Xu, Z. Y. Zhan and W. F. Wang. "Small Target Detection Method of Optical Remote Sensing Image Based on Multi-scale Information Fusion," *IAENG International Journal of Computer Science*, vol. 51, no. 6, pp. 681-687, 2024.

[5] V. K. Sharma, and R. N. Mir. “A comprehensive and systematic look up into deep learning based object detection techniques: A review,” *Computer Science Review*, vol. 38, pp. 100301, 2020.

[6] X. D. Li and Y. J. Zhang. "Improved Road Damage Detection Algorithm Based on YOLOv8n," *IAENG International Journal of Computer Science*, vol. 51, no. 11, pp. 1720-1730, 2024.

[7] S. S. A. Zaidi, et al. “A survey of modern deep learning based object detection models,” *Digital Signal Processing*, vol. 126, pp. 103514, 2022.

[8] X. D. Dong, S. Yan, and C. Q. Duan. “A lightweight vehicles detection network model based on YOLOv5,” *Engineering Applications of Artificial Intelligence*, vol. 113, pp. 104914, 2022.

[9] L. Aziz, M. S. B. H. Salam, U. U. Sheikh, and S. Ayub. “Exploring deep learning-based architecture, strategies, applications and current trends in generic object detection: A comprehensive review,” *IEEE Access*, vol. 8, pp. 170461-170495, 2020.

[10] J. Deng, et al. “A review of research on object detection based on deep learning,” *Journal of Physics: Conference Series. IOP Publishing*, vol. 1684, no. 1, pp. 012028, 2020.

[11] P. Peng, K. K. Geng, Z. W. Wang, Z. C. Liu and G. D. Yin. “Review on Environmental Perception Methods of Autonomous Vehicles,” *Journal of Mechanical Engineering*, vol. 59, no. 20, pp. 281-303, 2023.

[12] M. Maity, S. Banerjee, S. S. Chaudhuri. “Faster r-cnn and yolo based vehicle detection: A survey,” *2021 5th International Conference on Computing Methodologies and Communication (ICCMC). IEEE*, pp. 1442-1447, 2021.

[13] X. X. Xie, G. Cheng, J. B. Wang, X. W. Yao and J. W. Han. “Oriented R-CNN for object detection,” *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp. 3520-3529, 2021.

[14] Y. Zhang, et al. “Real-time vehicle detection based on improved yolo v5,” *Sustainability*, vol. 14, no. 19, pp. 12274, 2022.

[15] A. S. Geetha, M. A. R. Alif, M. Hussain, and P. Allen. “Comparative Analysis of YOLOv8 and YOLOv10 in Vehicle Detection: Performance Metrics and Model Efficacy,” *Vehicles*, vol. 6, no. 3, pp. 1364-1382, 2024.

[16] F. F. Wang, et al. “Low-illumination Object-Detection Algorithm Based on Image Adaptive Enhancement,” *Computer Engineering*, vol. 50, no. 10, pp. 352-361,2024.

[17] Y. Miao, F. Liu, T. Hou, L. Liu, and Y. Liu. “A nighttime vehicle detection method based on YOLO v3,” *2020 Chinese Automation Congress (CAC). IEEE*, pp. 6617-6621, 2020.

[18] C. J. Jiang, X. Y. He, and J. Xiang. “LOL-YOLO: Low-Light Object Detection Incorporating Multiple Attention Mechanisms,” *Computer Engineering and Applications*, vol. 60, no. 24, pp. 177-187, 2024.

[19] Y. L. Huang, X. L. Zhang. “Low-light object detection algorithm based on image feature enhancement,” *Electronic Measurement Technology*, vol. 47, no. 13, pp. 167-175, 2024.

[20] R. Zhang, S. B. Gao, X. Zhao, and X. L. Hou. “Algorithm on nighttime target detection for unmanned vehicles based on an improved YOLOv5s,” *Electronic Measurement Technology*, vol. 46, no. 17, pp. 87-93, 2023.

[21] J. W. Cao, F. Luo, and W. C. Ding. “BS-YOLO: a novel small target detection algorithm based on BSAM attention mechanism and SCConv,” *Computer Engineering*, pp. 1-10, 2024.

[22] S. B. Wu, J. L. Geng, C. Wu, Z. X. Ran and K. Y. Chen. “MultiSensor Fusion 3D Object Detection Based on MultiFrame Information,” *Transactions of Beijing Institute of Technology*, vol. 43, no. 12, pp. 1282-1289, 2023.

[23] L. Yang, Y. F. Chen, H. M. Li, J. X. Shi, and P An. “Object detection algorithm for autonomous driving scenes based on improved YOLOv8,” *Computer Engineering and Applications*, vol. 61, no. 1, pp. 131-141, 2025.

[24] P. Sun, et al. “Scalability in perception for autonomous driving: Waymo open dataset,” *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 2446-2454, 2020.

Rong Zhang received his B.S. degree in Automotive Service Engineering from Shandong Jiaotong University, Jinan, China, in 2022. He is currently pursuing the M.S. degree in Vehicle Engineering at Shandong University of Technology, Zibo, China. He is also with the Key Laboratory of New Energy Vehicle Integrated Design and Intelligent Technology of Shandong Province. His research direction is environment perception for intelligent vehicles.

Song Gao is currently a Professor at the School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, China. He is also with the Key Laboratory of New Energy Vehicle Integrated Design and Intelligent Technology of Shandong Province. His currently research interests include energy system matching theory and control technology of electric vehicle, intelligent vehicle and intelligent transportation system.

Pengwei Wang received his Ph.D. degree in Mechanical Engineering from the Shandong University of Technology, Zibo, China in 2020. He is currently an Associate Professor at the School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, China. He is also with the Key Laboratory of New Energy Vehicle Integrated Design and Intelligent Technology of Shandong Province. His research interests include intelligent vehicles, vehicle system dynamics, and motion control.

Binbin Sun received his Ph.D. degree in Mechanical Engineering from the Shandong University of Technology, Zibo, China in 2017. He is currently a Professor at the School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo, China. He is also with the Key Laboratory of New Energy Vehicle Integrated Design and Intelligent Technology of Shandong Province. His research interests include intelligent vehicles, new energy vehicle control technology, Intelligent technology of agricultural machinery.

Yuying Fang received her B.S. degree in Transportation from Shandong University of Technology, Zibo, China, in 2023. She is currently pursuing the M.S. degree in Vehicle Engineering at Shandong University of Technology, Zibo, China. She is also with the Key Laboratory of New Energy Vehicle Integrated Design and Intelligent Technology of Shandong Province. Her research interests include trajectory planning and motion control of intelligent vehicles.