ADVANCES IN TRANSPORTATION STUDIES
An International Journal

Editor in Chief: Alessandro Calvi

Vol. LXVIII April 2026

Contents

D. Mu, Z. Wei, Z. Li, D. Wang	3	Lightweight road foreign object detection algorithm based on improved YOLOv8
S. Gao, Z.P. Fu, B. Li, B. Li, T. Zhu, J. He	19	Driving behavior and risk analysis in urban tunnels with super-large cross-sections
E.M. Choueiri	39	Driven by intelligence: How AI is reshaping the way we move
A. Osama, M. Telima, Y. El Qattan, A. Bayomi, M. Attwa, M.H. Tawfeek	57	Evaluating the impact of speed enforcement in urban areas using probe vehicle data and traffic conflicts
Y. Wang	75	Intelligent automobile traffic sign detection based on context feature aggregation
A.D. Lidbe, M. Dutta, M. Cheshire, E.K. Adanu, X. Li, P. Penmetsa	89	Investigating the socio-demographic determinants of medical travel burden in the U.S.: insights from the 2017 National Household Travel Survey
K. Aziz, F. Chen, I. Ullah	111	Improving transparency and predictability of intersection crash severity analysis through Explainable Boosting Machines
X. Wang, W. Liu, Y. Ji, Q. Wan, W. Hao	131	Modeling and capacity analysis of mixed CAV-HV traffic incorporating variable safety distance in a cellular automaton model
A. Stedmon, D. McKenzie, M. Langham, K. McKechnie, R. Perry, S. Geddes, S. Wilson, M. Mackay, V. Tait	147	Vulnerable road users: investigating the influence of new road markings on motorcycle rider behaviour
T. Xu, L. Li, Y. Han	167	Solving the combined modal split and traffic assignment problem considering the equilibrium on transit networks

aracne

Z. Yang, G.H. Feng	179	An enhanced YOLOv5 model with multi-strategy improvements for detecting helmet-wearing on electric vehicle riders in complex scenes
F. Kasubi, O.M. Mdimi, J.H. Salum, E. Mwambeleko, P. Alluri	193	Understanding lane choice: the role of express lanes on traffic distribution in general-purpose lanes
X.L. Tang, C. Li, L.Y. Wu, L. Zeng	211	Design of intelligent warning method for highway traffic congestion under extreme weather conditions
R. Vazquez, F. Masson	223	Viability to predict driver intentions in unmarked roundabouts using vehicle and contextual information
X.H. Xin, J.Y. Zhang, R. Zhao, M.Y. Zuo, K.D. Liu, S.F. Wang	237	Analysis of accident causes and severity association rules at intersections based on an improved Apriori algorithm
C.X. Chen, H.P. Zhang	257	Passenger boarding behavior assignment approach to mitigate unreliable operation of bus lines
Y. Meng, J. Zhang, Y. Zhang, Z. Liu, G. Qing	277	Computation model of safety index for mountainous expressway accident-prone sections
E.M. Choueiri	295	The future of transportation: trends, technologies, and strategic preparation
F.X. Shi, M. Jia, Y.W. Li	317	Control method for emergency lane change behavior of autonomous vehicles in the internet of vehicles environment
C.F. Shi, P. Wang, B.L. Chen, M.J. Wang, L. Wang, Y. He	329	An obstacle avoidance trajectory selection for autonomous vehicles based on improved artificial potential field method
O. Oluwatomini, D. Eustace, P. Brewick, K.Q. Walsh	341	Modeling the impacts of physical vertical deflection traffic calming measures on speed, traffic volume, noise, emissions, and delay - A systematic review
L.M. Zhang, Y. Ma, T.R. Li, G.W. Li	367	Lateral stability control of autonomous vehicles with joint connected internet of vehicles technology
X. Zhou, C. Alecsandru, S. Bashbaghi, Y. Jeong, Y. Chen	381	Turn-Aware LSTM model for vehicle trajectory forecasting
J. Wang, Q.S. Liu, Y. Wen, S.K. Alias	397	An intelligent charging control algorithm for improving the efficiency and safety of new energy vehicle batteries

Advances in Transportation Studies: an international Journal 68 (2026)

ISBN 979-12-218-2541-1 – ISSN 1824-5463 – DOI 10.53136/97912218254111 – pag. 3-18

Lightweight road foreign object detection algorithm based on improved YOLOv8

D. Mu[1,2] Z. Wei[1,2*] Z. Li[1,2] D. Wang[1,2]

[1]*School of Physics and Electronic Information, Huaibei Normal University, Huaibei, Anhui, 235000, China*
[2]*Anhui Province Key Laboratory of Intelligent Computing and Applications, Huaibei Normal University, Huaibei, Anhui, 235000, China*
[*]*corresponding author; email: 15952240274@163.com*

subm. 8th May 2025 *approv. after rev. 16th August 2025*

Abstract

In recent years, traffic accidents caused by foreign objects on expressways have been on the rise year by year, which have brought serious damage to vehicles or cargo. In view of the unsatisfactory detection accuracy of current road object detection algorithms, and the difficulty of deploying detection models in practical application scenarios with limited computing resources due to the large number of model parameters and high computational complexity, this paper proposes a lightweight road foreign object detection algorithm based on improved YOLOv8n. First, to enhance the detection model's accuracy for small road debris, a generalized building module CONTAINER integrated with multiple contexts was introduced to enhance the detection model's ability to extract local features of small-scale road foreign objects and accelerate convergence. Secondly, the C2f-Faster module is integrated into the detection model's backbone and neck to enhance accuracy with fewer parameters and lower complexity. Finally, in view of the limitations of immutability of border scale and weak generalization ability of CloU, Inner-IoU is used to improve the detection layer, and the scale factor ratio is added to boost boundary box regression accuracy and accelerate convergence. Through the experimental verification of ablation experiment and comparison experiment, the results indicate the proposed algorithm's average accuracy surpasses the traditional method, the enhanced algorithm model boasts an average accuracy improvement of 3%, rising from 96.0% to 99.0%, the calculation amount of the model is reduced by 1.8 from 8.2 to 6.4, and the parameter count has been decreased by 0.7, from 3.0 to 2.3. The algorithm enhances road debris detection accuracy and streamlines the calculation model, which can provide certain reference value for highway inspection and maintenance management.

Keywords – road foreign body, context aggregation, C2f-Faster, inner-IoU

1. Introduction

With rapid economic growth and increasing traffic, road safety has become a major challenge in China. Detecting foreign objects on roads is crucial, as it directly impacts people's safety, vehicle operation, and road flow [1]. Foreign objects on highways refer to non-inherent objects that may pose risks to traffic safety, mainly including: (1) Vehicle debris (tire fragments, cargo remnants, etc.), (2) Natural falling objects (rocks, branches, etc.), (3) Man-made discarded items (obstacles, waste, etc.).These objects can come from traffic accidents, spilled cargo, or intentional placement, posing significant safety risks. The goal of road foreign object detection technology is to monitor road environments in real-time, quickly identifying and addressing potential hazards to ensure safe

and smooth traffic [2]. Common foreign objects are illustrated in Figure 1. Currently, road maintenance departments use two main solutions for dealing with highway foreign objects: The first one is manual inspection, where staff are assigned to patrol and clean specific road sections. While this method can identify and remove objects, it is inefficient, potentially dangerous, and can lead to neglect. The second one is to use high-definition cameras, which are installed at key locations for real-time monitoring. However, these cameras cannot cover entire highways, and due to the unpredictable nature of foreign objects, they may not ensure timely removal.

An Hengyue et al. [3]. enhanced the accuracy of foreign object detection on roads by incorporating the SE (Squeeze-and-Excitation) attention mechanism and improved the backbone network using Receptive Field Attention (RFA) convolution. Although the modified algorithm effectively improved target detection accuracy, it failed to achieve model lightweighting. Guo Junmei et al. [4]. designed a novel network architecture module by replacing the C2f module with C3D and proposed a new feature fusion method called SFPN (Shallow-Deep Feature Pyramid Network). This method integrates shallow and deep-level information while preserving more original features. The improved model demonstrated significant enhancements in both precision and lightweighting performance. However, the complexity of the modified model still requires further optimization. Wang et al. [5] optimized the feature fusion part of deep learning obstacle detection by replacing the detection head network with a Dyhead dynamic detection head to improve the extraction of occluded features. He also redesigned the loss function using Focaler-IoU. Although this improved average accuracy, it did not achieve true data lightweighting. Zhu Xiaofeng et al. [6] used a lightweight neural network as the backbone for real-time detection of small foreign objects on highways, employing a multi-weight balancing strategy to aid detection. However, the model's computational load and parameter count remain high and need further reduction. In Zou Junyi's research [7] on road bump obstacle recognition method for intelligent vehicles, The SimAm module is integrated into the feature extraction layer to boost the network's feature map perception, and WIoU is used as a positioning loss function to improve the network convergence speed and robustness. However, the improved accuracy needs to be further improved.

Given the current challenges in road object detection algorithms, which include unsatisfactory detection accuracy and difficulty deploying models in resource-limited scenarios due to large parameter counts and high computational complexity, this paper proposes an improved lightweight road foreign object detection algorithm based on YOLOv8. It introduces the CONTAINER module with multi-head context integration, the C2f-Faster module, and Inner-IoU.

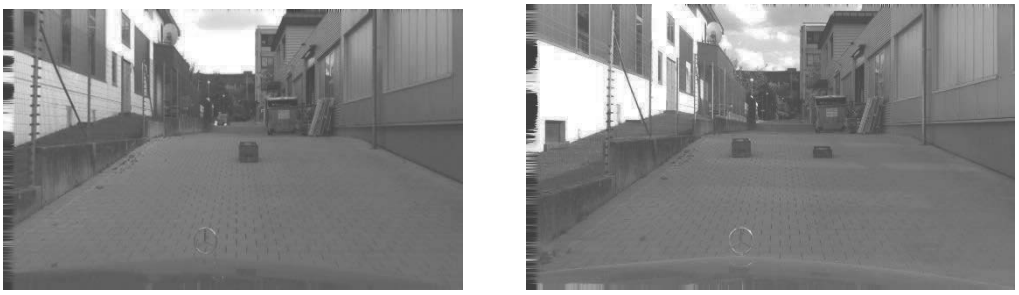


Fig. 1 - Common foreign objects

2. YOLOv8 algorithm

YOLOv8 is a state-of-the-art object detection algorithm encompassing functionalities such as object detection, keypoint identification, and instance segmentation, along with real-time tracking. It uses the C2f (CSP Bottleneck with 2 convolutions) module to maintain lightweight performance while enhancing gradient flow information. The channel number adjusts according to model size, significantly boosting performance. The algorithm employs an anchor-free detection method, predicting the center and aspect ratio of objects directly, which improves detection speed and accuracy. The decoupled head structure separates classification and detection, reducing parameters and computational complexity. Two concurrent pathways are utilized to extract category and location data, enhancing model robustness and better linking position and category.

YOLOv8 has four main modules: input [8], backbone [9], neck [10], and output [11]. During data processing, the input stage uses Mosaic data augmentation, adaptive image scaling, and grayscale padding to enhance data diversity and model generalization. The backbone employs Conv, C2f, and SPPF for deep feature extraction. The neck part features a PAN (Path Aggregation Network) structure, using upsampling, downsampling, and feature map concatenation for effective multi-scale feature fusion. In the output stage, a decoupled head structure is used for classification and regression tasks. The TaskAligned-Assigner method weights classification and regression scores to ensure accurate matching with positive samples. Fig. 2 shows the YOLOv8 network framework [12].

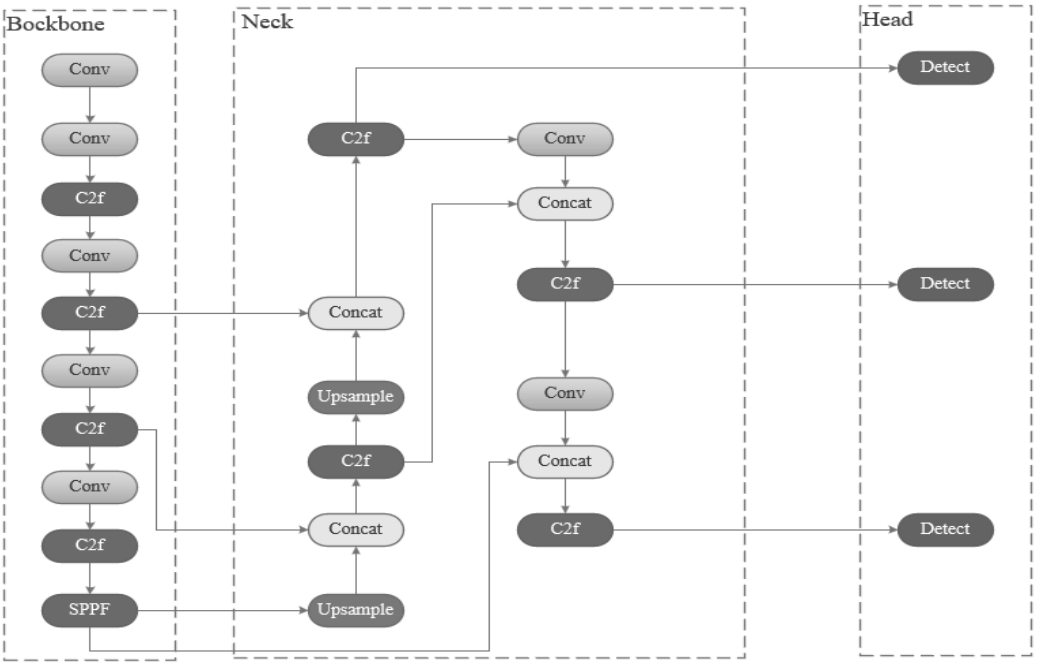


Fig. 2 - YOLOv8 network framework

3. Improved YOLOv8 network

YOLOv8n was selected as the baseline model for improvements. The enhanced network structure is depicted in Fig. 3.

3.1. Introduction of the context aggregation module

Traditional CNNs are effective at extracting local features but have limitations in capturing long-range dependencies in images [13]. Transformers excel at handling long-distance interactions but require more computational resources and training time [14]. MLP-Mixer [15] uses fully connected layers to mix features from different areas, but it may not capture local patterns as effectively as CNNs.

The CONTAINER (Context Aggregation Network) [16] architecture integrates the strengths of these three approaches.Take advantage of long-distance interactions like Transformer, while maintaining CNN's advantages in fast convergence and efficient local feature extraction. In the CONTAINER model, static affinity matrices are learned and visualized at different layers (Layer 1 and Layer 12), as shown in Fig. 4. These matrices highlight the static affinities at each position, with local affinities (similar to convolution operations) enhanced near source pixels, while the affinity value of the source pixel itself is very low.

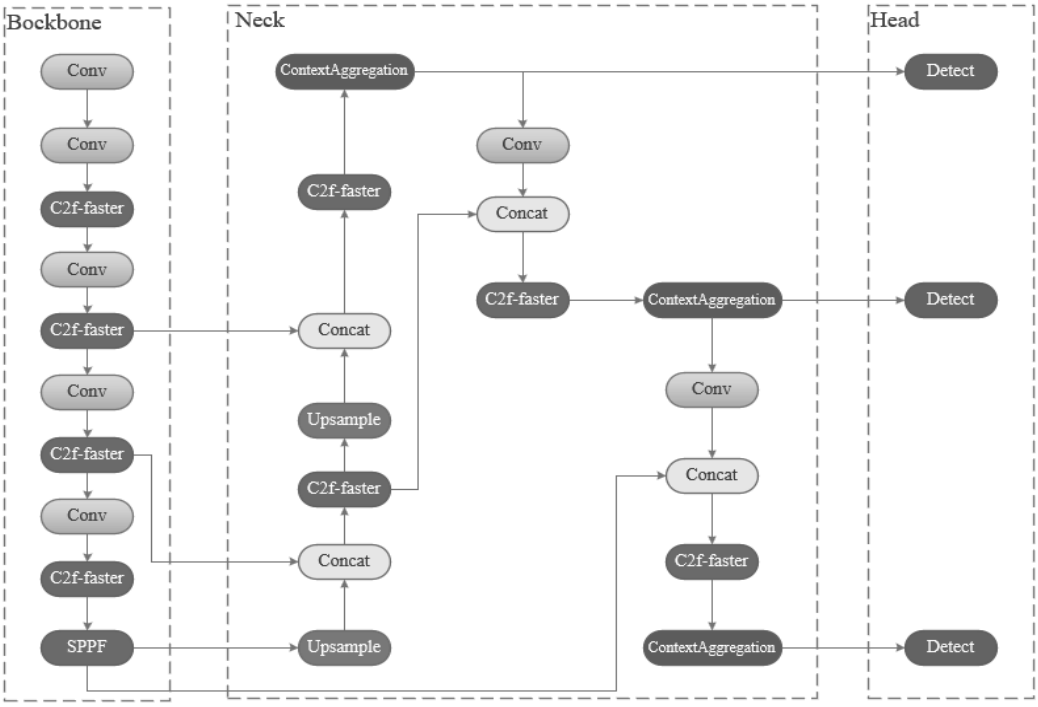


Fig. 3 - Improved YOLOv8n model structure

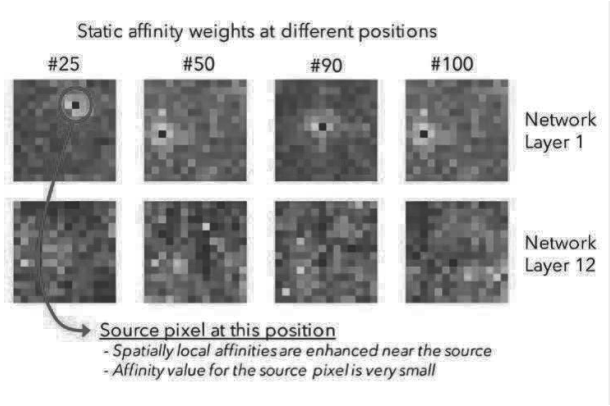


Fig. 4 - Visualization of the Layer 1 and Layer 12

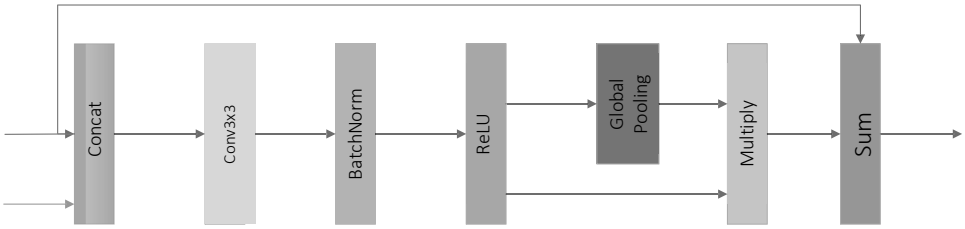
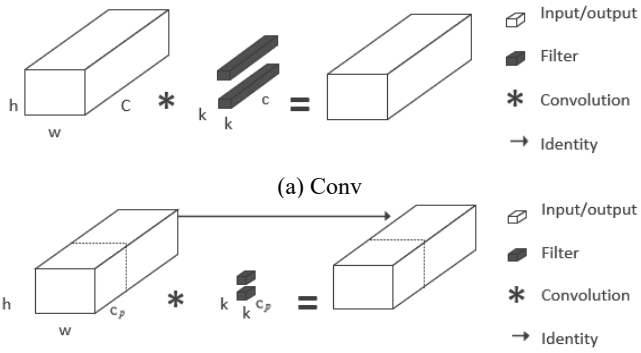


Fig. 5 - The structural block diagram of the CONTAINER module

The structural block diagram of the CONTAINER module is shown in Fig 5. This indicates that during context aggregation, the weight of the current feature is reduced, likely due to residual connections minimizing the need for source features. By combining static and dynamic affinity matrices through learnable parameters and a responsive architecture, CONTAINER has shown strong results in image classification tasks. The efficient extension, CONTAINER-LIGHT, significantly improves performance in detection and segmentation. The attention mechanism Context Aggregation enhances the model's adaptability to complex backgrounds by aggregating global and local information. By preserving the original PAN-FPN [17] multi-scale fusion process and adding a lightweight CONTAINER branch, it generates global contextual features. The output of CONTAINER is then fused with the original feature maps through element-wise addition or channel concatenation, significantly improving small object detection capabilities.

3.2. Introduction of the C2f-Faster module

In the original YOLOv8n backbone network, C2f uses a series of standard convolution and bottleneck modules for feature extraction. While this provides good capability, it results in higher computational complexity and parameter count, making it less suitable for applications requiring fast inference. To minimize computational redundancy and optimize the backbone's parameters, we drew inspiration from FasterNet [18], which achieves higher speeds on various visual tasks without sacrificing accuracy. To create a faster network and reduce floating-point operations (FLOPs), we integrated Partial Convolution (PConv) with C2f. In this paper, FasterBlock in FasterNet is used to replace the Bottleneck module in C2f, which cuts down on model parameters and complexity.



(a) Conv

(b) PConv

Fig. 6 - Structure comparison between Conv and PConv

The FasterBlock module improves upon the partial convolution (PConv) concept and traditional layered Transformer ideas. PConv is a convolution operation that enhances computational efficiency by performing convolutions on only a part of the input feature map, rather than applying them comprehensively as in traditional methods. This reduces unnecessary calculations and memory access by ignoring redundant parts of the input. Because it ignores parts of the input that are considered redundant. This approach is particularly suitable for running deep learning models on devices with limited resources, as it can significantly reduce the computational requirements without sacrificing too much performance. A comparison of Conv and PConv structures is shown in Fig. 6. PConv enables fast and efficient feature extraction by applying filters to only a small portion of the input channel while leaving the rest unchanged [19].

The computation for standard convolution is given by:

$$h * w * k^2 * c^2 \tag{1}$$

For Partial Convolution (PConv), the computation is:

$$h * w * k^2 * {c_p}^2 \tag{2}$$

With a typical offset ratio of $r = \frac{c_p}{c} = \frac{1}{4}$, the FLOPs for PConv are only one-sixteenth of the full convolution, significantly reducing computational overhead [20] and improving efficiency. By selecting the first or last channel as a representative of the entire feature map, memory access frequency is reduced.

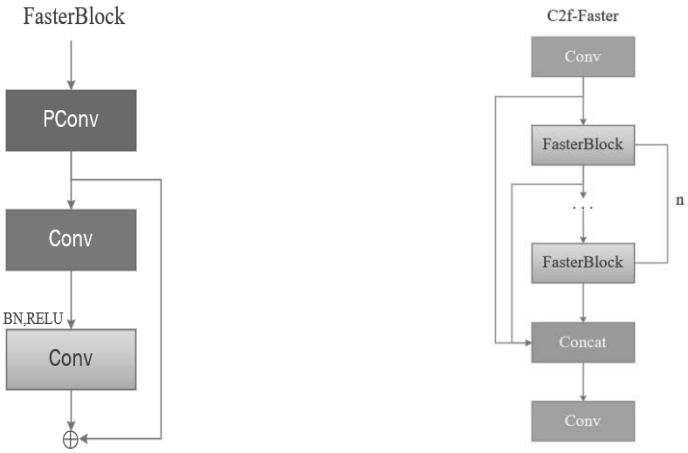
The memory access frequency for standard convolution is:

$$h * w * 2c + k^2 * c^2 \approx h * w * 2c \tag{3}$$

For PConv, it is:

$$h * w * 2c_p + k^2 * c_p^2 \approx h * w * 2c_p \tag{4}$$

With the offset ratio unchanged, PConv accesses memory only 1/4 as often as standard convolution. In C2f-Faster [21], the introduction of PConv technology leverages parallel computing capabilities, accelerating training and inference speeds and improving overall efficiency [22]. Fig. 7 illustrates FasterBlock and C2f-Faster structures.



(a) FasterBlock (b) C2f-Faster

Fig. 7 - FasterBlock and C2f-Faster structure diagram

3.3. Inner-IoU loss function

In object detection, IoU [23] measures the overlap of bounding boxes. During training, it helps evaluate the similarity between anchor boxes and object bounding boxes, assigning positive or negative samples to each anchor point. This process improves the model's ability to predict object boundaries by increasing the similarity between predicted and actual boxes. A higher IoU shows more overlap between predicted and actual boxes, leading to more accurate model predictions. YOLOv8 uses Complete IoU (CIoU) as the loss function for bounding box prediction

The CIoU formula is given by:

$$L_{CIoU} = 1 - IoU + \frac{\rho^2(b,b^{gt})}{c^2} + \alpha v \tag{5}$$

To address the limitations of standard IoU loss functions, the Inner-IoU [24] method adds auxiliary bounding boxes to improve generalization and accelerate box regression. By incorporating a scale factor ratio, Inner-IoU allows adjusting auxiliary box sizes for various datasets and detectors, enhancing generalization capabilities. The concept of auxiliary bounding boxes is illustrated in Fig. 8. The ground truth box is B^{gt}, and the anchor box is B. The center points of the GT box and the anchor box are (x_c^{gt}, y_c^{gt}) and (x_c, y_c). The width and height boxes for the GT, representing its dimensions, are denoted by the variables e^{gt} and h^{gt}, respectively. The anchor box's width (w) and height (h) are used [25]. "Ratio" is the scale factor, usually between 0.5 and 1.5. Inner-IoU defines the boundaries of the auxiliary boxes as follows:

Equations (6) and (7) compute the boundaries of the auxiliary ground truth box:

$$b_l^{gt} = x_c^{gt} - \frac{w^{gt}*ratio}{2}, b_r^{gt} = x_c^{gt} + \frac{w^{gt}*ratio}{2} \tag{6}$$

$$b_t^{gt} = y_c^{gt} - \frac{h^{gt}*ratio}{2}, b_b^{gt} = y_c^{gt} + \frac{h^{gt}*ratio}{2} \tag{7}$$

Equations (8) and (9) determine the edges of the predicted auxiliary box:

$$b_l = x_c - \frac{w*ratio}{2}, b_r = x_c + \frac{w*ration}{2} \tag{8}$$

$$b_t = y_c - \frac{h*ratio}{2}, b_b = y_c + \frac{h*ration}{2} \tag{9}$$

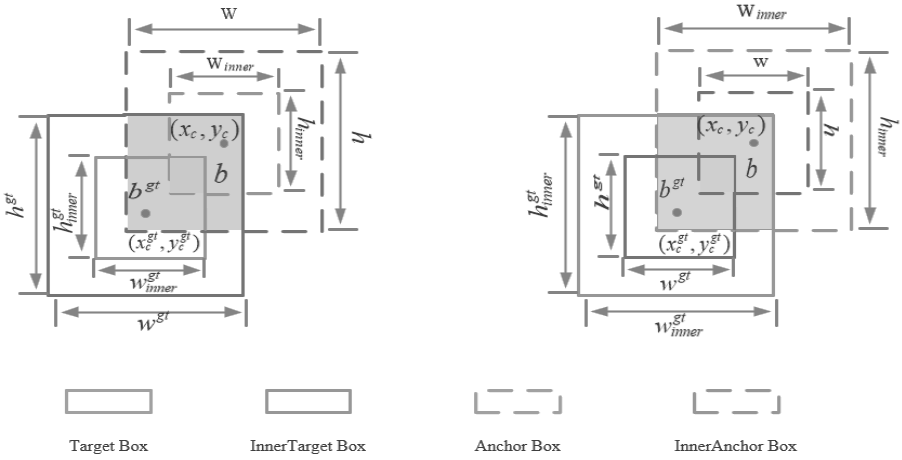


Fig. 8 - Inner-IoU working diagram

Equations (10) and (11) calculate the intersection and union areas of the auxiliary boxes:

$$inter = \left(min(b_r^{gt}, b_r\right) - max\ (b_l^{gt}, b_l)) * (min\left(b_b^{gt}, b_b\right) - max\ (b_t^{gt}, b_t) \tag{10}$$

$$union = (w^{gt} * h^{gt}) * (ration)^2 + (w * h) * (ration)^2 - inter \tag{11}$$

Equation (12) computes the IoU for the auxiliary boxes:

$$IoU^{inner} = \frac{inter}{union} \tag{12}$$

4. Experimental results and analysis

4.1. Dataset introduction

The experiments in this study utilized the "Lost and Found" dataset, which was first introduced in the 2016 IROS conference paper titled 'Lost and Found: Detecting Small Road Hazards for Self-Driving Vehicles [26]'. This benchmark dataset contains 2,104 annotated images covering various small obstacles including tires, boxes, and plastic bags. The distribution of different types of foreign objects is shown in Tab. 1. Each obstacle is carefully labeled for training and testing object detection algorithms. Detecting small obstacles is challenging because they may occupy only a small part of the image and can be affected by lighting, occlusion, and perspective changes. However, using datasets like Lost and Found, researchers can develop and refine algorithms to tackle these challenges, enhancing the safety and reliability of autonomous driving systems in real-world scenarios. The data was randomly split into a training subset and a validation subset, with a proportion of 9 parts for training and 1 part for validation.

Tab. 1 - Statistical table of foreign object samples by category

Category	Specific sub-class	Number of examples	proportion
Small objects	Stone, Plastic bags,etc.	2,215	52.5%
Mid-scale objects	Package, Tire,etc.	1,576	37.3%
Large objects	Wooden crate, Bicycle,etc.	428	10.2%
Total	37 subcategories	4219	100%

4.2. Experimental environment

Experiments were run on Windows 11 with PyTorch. Tab. 2 presents the detailed parameters of the experimental setup. To ensure effective training and performance optimization, the study standardized the experimental parameters: the learning rate was set to 0.001, setting the batch size to 4, training epochs to 150, and worker threads to 4. The image input dimension was configured to 640 pixels, with other parameters set to the default values of the YOLOv8n model.

Tab. 2 - Experimental platform parameters

Parameter	Configuration
Operating System	Windows 11
GPU	NVIDIA RTX4060
GPU Memory	16 GB
GPU Acceleration	CUDA12
Network Framework	PyTorch
Programming Language	Python

Precision is based on the prediction results, specifically the proportion of correctly predicted positive samples [27]. Predictions classified as positive can either be true positives (TP) or false positives (FP). This can be expressed with the formula:

$$\text{precision} = \frac{\text{TP}}{\text{TP+FP}} \tag{13}$$

Recall is based on actual samples. It quantifies the percentage of true positive instances that are accurately identified. Among the actual positive samples, they are either correctly predicted as TP or incorrectly as FN. The formula is:

$$\text{recall} = \frac{\text{TP}}{\text{TP+FN}} \tag{14}$$

Average Precision (AP) measures the detection performance of a model for a specific object category [28]. In practical applications, to evaluate a model's performance across multiple categories, the mean Average Precision (mAP) averages category APs, offering a model's overall performance gauge. The formulas are:

$$\text{AP} = \int_0^1 \text{P(R)dR} \tag{15}$$

$$\text{mAP} = \frac{\sum_{\text{j=1}}^{\text{C}} \text{AP}_\text{j}}{\text{C}} \tag{16}$$

4.3. Ablation experiment

To assess the enhancements from the Context Aggregation, C2f-Faster, and Inner-IoU loss in YOLOv8n, ablation studies were performed on the Lost and Found Dataset. Each improvement was added step-by-step: introducing the Context Aggregation module (method A), adding the C2f-Faster module (method B), and using the Inner-IoU loss function (method C). The results are in Tab. 3. From the table, it is observed that the base YOLOv8n model has an mAP@0.5 of 96.0%. By introducing the Context Aggregation module, C2f-Faster module, and Inner-IoU loss function, the model's mAP@0.5 increased by 2.2%, 1.5%, and 1.2%, respectively. The Context Aggregation module contributed the most to performance improvement. Introducing the C2f-Faster attention mechanism in the backbone and neck parts helped the model focus better on important information, optimizing feature fusion [29].

Tab. 3 - Results of ablation experiment

Baseline model	Improvement Method	P/%	R/%	mAP @0.5/%	GFLOPS/G	Params/M	FPS/(f·s^{-1})
YOLOv8n	——	95.7	94.3	96.0	8.2	3.0	73.3
YOLOv8n	+A	96.1	93.2	98.2	8.2	3.0	64.3
YOLOv8n	+B	98.2	89.2	97.5	6.4	2.3	98.6
YOLOv8n	+C	96.0	95.3	97.2	8.2	3.0	85.2
YOLOv8n	+A+B	96.2	92.5	97.4	6.4	2.3	93.6
YOLOv8n	+A+C	97.8	96.2	98.4	8.2	3.0	72.5
YOLOv8n	+B+C	96.2	94.9	98.2	6.4	2.3	94.5
YOLOv8n	+A+B+C	97.7	96.2	99.0	6.4	2.3	92.7

The model's GFLOPS and parameters were reduced by 1.8 and 0.7, respectively, compared to the original network. Switching to Inner-IoU loss and adding a scale factor raised mAP@0.5 by 1.2%. In terms of frame rate, C2f-Faster demonstrates the most significant improvement, achieving a 25.3 FPS increase compared to YOLOv8n, albeit with slightly lower accuracy. Our proposed algorithm, which integrates all three modules, elevates the frame rate from 73.3 FPS to 92.7 FPS. Combining all three modules yielded the best results, increasing mAP@0.5 by 3%, a 19.4 FPS increase in frame rate, with GFLOPS and parameters reduced by 1.8 and 0.7, respectively.

4.4. Comparative experiment

4.4.1. Comparison of different small target detection modules

To highlight the superiority and effectiveness of introducing the Context Aggregation network into the original model, we compared it with SPDConv [30] (composed of SPD layers and Non-strided-Conv layers), AkConv (Adaptive Kernel Convolution), and other popular small target detection networks. The findings are presented in Tab. 4.

The table shows that compared to SPDConv, AkConv, and other popular small target detection networks, the Context Aggregation network achieves the best average precision and computational efficiency. Although AkConv and additional small target detection layers have fewer parameters, their average precision decreases by 1.7 and 0.5 percentage points, respectively, and their computational cost increases by 0.5 and 4.4. Overall, adding Context Aggregation is more beneficial for improving model performance [31].

Tab. 4 - Detection network results of different small targets

Baseline model	P/%	R/%	mAP @0.5/%	GFLOPS/G	Params/M
+ContextAggrefation	96.1	93.2	98.2	8.2	3.0
+SPDConv	95.4	92.8	97.9	54.2	3.6
+AKConv	93.9	94.0	96.5	8.7	2.9
+MobileViTAttention	93.7	93.9	97.7	11.2	3.9
+SEAttention	94.5	92.2	97.5	8.9	3.2
Adding small target detection layer	95.8	94.3	97.8	12.6	2.9

4.4.2. Comparison of different attention mechanisms

To clearly demonstrate the performance advantages of the C2f-Faster module [32], we conducted comparative experiments with the C2f-Attention [33], C2f-MSDA [34], C2f-CA [35], and GhostConv [36] attention modules. The experimental results are presented in Tab .5.

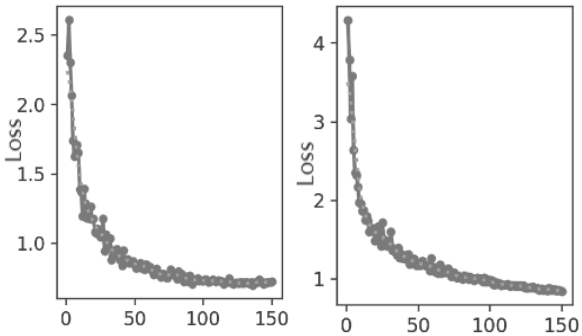
Tab. 5 - Comparative experimental results

Baseline model	P/%	R/%	mAP @0.5/%	GFLOPS/G	Params/M
+C2f-Faster	98.2	89.2	97.5	6.4	2.3
+ C2f-Attention	97.1	88.6	96.7	9.1	3.2
+C2f-MSDA	90.6	90.0	96.4	7.1	2.4
+C2f-CA	95.4	92.8	97.0	43.3	11.6
+GhostConv	94.3.	95.8	96.5	5.8	2.1

Tab. 5 indicates that The GhostConv module exhibits advantages in both parameter count and computational load over the C2f-Faster module, yet it underperforms in terms of accuracy. The C2f-Faster module excels over other attention mechanisms in average precision, parameter count, and computational cost. The reduction in computational cost is particularly significant, with decreases of 2.7, 0.7, and 36.9 compared to the C2f-Attention, C2f-MSDA, and C2f-CA modules, respectively.

4.4.3. Comparison of loss function effects

To demonstrate the effectiveness of the Inner-IoU loss function, experiments were conducted using the CIoU loss function on the YOLOv8 model. As shown in Fig. 9, Fig. 9(a) shows results from the Inner-IoU loss, contrasting with Fig. 9(b)'s CIoU loss results. From 0 to 50 epochs, the Inner-IoU loss function converges significantly faster than the CIoU loss function and quickly stabilizes between 50 and 100 epochs [37]. The conducted experiments thoroughly validate the effectiveness of the Inner-IoU loss function, demonstrating its superiority in enhancing the performance of the detection model.



(a) Inner-IoU (b) CIoU

Fig. 9 - Comparison of loss functions

4.4.4. Comparison of different model performance

In order to more accurately assess the performance of our enhanced model, we have conducted a series of comprehensive evaluations and comparisons with various metrics and methodologies, We evaluated it against multiple object detection models, such as Faster R-CNN , SSD [38], YOLOv5n,NanoDet,YOLOv6 [39], YOLOv7 [40], YOLOv5-Lite, EfficientDet-Lite [41] and YOLOv8n. Additionally, we validated the algorithm's performance on TinyPerson [42], a

specialized dataset for small object detection, which is equally suitable for our proposed method. As shown in Tab .6, our improved YOLOv8n achieves higher average precision than the other models, outperforming Faster-RCNN, SSD, YOLOv5n, YOLOv6, YOLOv7, NanoDet, YOLOv5-Lite, EfficientDet-Lite and the original YOLOv8n by 16.4, 18.7, 2.2, 3.1, 3.3, 3.7,1.5,9.5and 3 percentage points, respectively. Additionally, our model has relatively fewer parameters. Faster-RCNN and SSD have larger computational and parameter requirements but lower precision. Although YOLOv5n, NanoDet, YOLOv5-Lite, and EfficientDet-Lite require fewer computational resources and parameters than our model, their accuracy is lower by 1.8, 3.7, 1.5, and 9.5 percentage points, respectively. YOLOv5s, YOLOv6, YOLOv7, and YOLOv8n all have higher parameter and computational demands but lower precision compared to our algorithm. Our improved YOLOv8n achieves a higher frame rate than other models, with an increase of 19.4 FPS compared to the original version. Overall, the improved YOLOv8 model performs better, especially for detecting highway foreign objects [43].

Tab. 6 - Comparing experimental results across models in the lost and found dataset

Baseline model	P/%	R/%	mAP @0.5/%	GFLOPS/G	Params/M	FPS/(f·s^{-1})
Faster-RCNN	78.7	74.3	82.6	132.2	41.3	6.94
SSD	73.2	70.6	80.3	86.3	40.3	3.76
EfficientDet-Lite	80.6	85.2	89.5	3.9	0.8	65.3
YOLOv5n	96.2	97.1	97.2	4.2	1.8	70.2
YOLOv5-Lite	95.6	96.3	97.5	4.0	1.6	3.65
YOLOv6	93.5	88.7	95.9	11.8	3.6	79.6
YOLOv7	94.1	91.0	95.7	103.2	36.5	69.5
YOLOv8n	95.7	94.3	96.0	8.2	3.0	73.3
NanoDet	93.6	94.2	95.3	0.72	0.95	32.6
TinyPerson Dataset	96.4	97.3	98.6	6.4	2.3	91.2
Ours	97.7	96.2	99.0	6.4	2.3	92.7

4.5. Visualization analysis

To clearly show the performance differences before and after our algorithm improvements, we conducted image-by-image detection on the test set. We selected representative data from complex scenes with large scale variations, dense distributions, and high background similarity for visualization experiments. By comparing YOLOv5n, YOLOv8n, and our improved model, we can clearly see the enhanced performance of the improved model in handling these complex backgrounds, as shown in Fig. 10. Higher confidence in the detection boxes indicates a greater likelihood of the target being present [44], showing that the model has learned more detailed information about the targets. The results from four sets of experiments reveal that the basic YOLOv8n and YOLOv5n models have lower detection rates for highway foreign objects in complex backgrounds and tend to miss detections. In contrast, our improved algorithm significantly enhances small object detection, greatly reducing missed detections and effectively improving the accuracy of highway foreign object detection. As shown in Figure e, tests conducted under simulated elevated surveillance camera perspectives typical of highway monitoring systems - demonstrate that our algorithm maintains exceptional detection performance even under high-angle, long-distance imaging conditions. These results validate the model's robustness to multi-scale perspective transformations. The P-R curves for the YOLOv8n algorithm and our improved

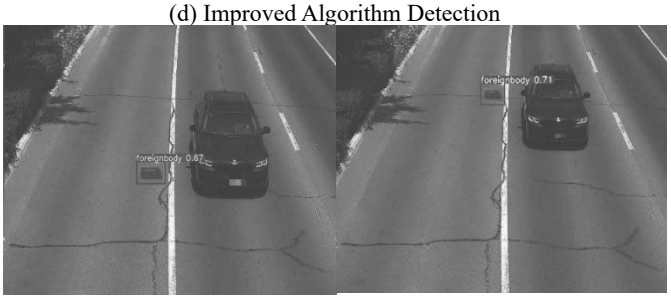
algorithm are shown in Fig. 11, demonstrating a significant increase in average precision after the improvements. The missed detection rates (MDR) and average confidence scores for foreign object detection in complex scenarios across different models are presented in Table 7.

(a) Original Image

(b) YOLOv5n Detection

(c) YOLOv8n Detection

(d) Improved Algorithm Detection

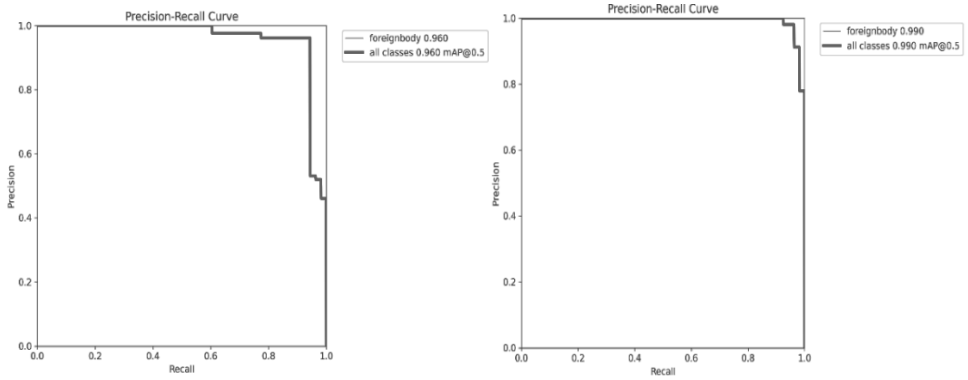


(e) Emulate high-speed camera detection performance

Fig. 10 - Comparison of detection effect before and after algorithm improvement

Tab. 7 - Statistical table of detection performance in complex scenarios

Model	MDR (%)	Avg Confidence
YOLOv5n	9.3	0.83
YOLOv8n	12.7	0.81
Ours	5.3	0.89



(a)YOLOv8n (b) Detection of Our Algorithm

Fig. 11 - Comparison of P-R curve between YOLOv8n algorithm and this algorithm

5. Conclusion

This paper addresses the challenges in highway foreign object detection by proposing a lightweight detection algorithm based on an improved YOLOv8n. First, by adding a generalized module for multi-head context integration (Context Aggregation), the model achieves optimal average precision while minimizing computational load, as compared to other small object detection algorithms. Secondly, the introduction of the C2f-Faster module shows that, compared to other popular attention mechanisms, it achieves optimal results in terms of average precision, computational load, and parameter count. Lastly, replacing YOLOv8's CIoU loss with Inner-IoU speeds up the model's convergence. In practical applications, the system can help vehicles quickly identify foreign objects at long distances, providing drivers with sufficient reaction time to take timely evasive actions. Experimental results indicate that the improved YOLOv8 algorithm offers high detection accuracy with lower computational resource requirements, making it well-suited for road foreign object detection. Looking ahead, three promising research directions emerge from this work: (1) Multi-modal sensor fusion incorporating LiDAR and thermal imaging could enhance detection reliability in low-visibility conditions, particularly for non-metallic objects, (2) On-device deployment optimization for edge computing platforms would facilitate real-time processing in resource-constrained vehicle systems, (3) Semi-supervised learning approaches may reduce annotation costs while maintaining detection accuracy for rare object categories. These extensions would further bridge the gap between laboratory performance and real-world deployment requirements.

Acknowledgments

This research was supported by the Open Fund Project of the Anhui Key Laboratory of Intelligent Computing and Application (Project No. AFZNJS2024KF03) and the 2023 Anhui New Era Education Quality Engineering Project (2023lhpysfjd044).

References

1. Wu Hongtao, Liu Liyuan, Meng Ying, et al. Threat Degree Analysis Method for Scattered Road Objects Based on Dynamic Multi-Feature Fusion[J]. Computer Science, 2020, 47(S1): 196-205.
2. Ye Ting, Chen Zhongxian, Tan Rongxian, et al. Analysis of Foreign Object Protection for Power Battery Front-end in M Electric Vehicles[J]. Automobile Test Report, 2023, (08): 152-154.
3. An H, Liu W, **ng Y. Research on foreign object detection on highways based on improved YOLO in harsh weather[C]//Eighth International Conference on Traffic Engineering and Transportation System (ICTETS 2024). SPIE, 2024, 13421: 299-306.
4. Guo J, Lou H, Chen H, et al. A new detection algorithm for alien intrusion on highway[J]. Scientific reports, 2023, 13(1): 10667.
5. Wang Meng, Li Wei, Gao Rong, et al. Foreign Object Debris Detection of Pavement Based on Deep Learning of 3D Point Cloud. Computer Systems & Applications, 2021, 30(02): 165-170.
6. Zhu Xiaofeng, Li Lin, Zhang Dejin, et al. RoadNetv2:Real-Time Algorithm for Highway Weak Abandoned Objects Detection. Computer Engineering and Applications, 2023, 59(01): 317-324.
7. Zou Junyi, Liu Chang, Guo Wenbin, et al. Research on Road Bump Obstacle Recognition for Intelligent Vehicle. China Mechanical Engineering, 2024, 35 (06): 951-961.
8. Chen H, Zhou G, Jiang H. Student behavior detection in the classroom based on improved YOLOv8[J]. Sensors, 2023, 23(20): 8385.
9. Cheng G, Chao P, Yang J, et al. SGST-YOLOv8: An Improved Lightweight YOLOv8 for Real-Time Target Detection for Campus Surveillance[J]. Applied Sciences, 2024, 14(12): 5341.
10. Liu J, Liu X, Chen H, et al. MDD-YOLOv8: A Multi-Scale Object Detection Model Based on YOLOv8 for Synthetic Aperture Radar Images[J]. Applied Sciences (2076-3417), 2025, 15(4).
11. Hu D, Yu M, Wu X, et al. DGW-YOLOv8: A small insulator target detection algorithm based on deformable attention backbone and WIoU loss function[J]. IET image processing, 2024, 18(4): 1096-1108.
12. Varghese R, Sambath M. Yolov8: A novel object detection algorithm with enhanced performance and robustness[C]//2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS). IEEE, 2024: 1-6.
13. Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6): 84-90.
14. Dao T, Fu D, Ermon S, et al. Flashattention: Fast and memory-efficient exact attention with io-awareness[J]. Advances in neural information processing systems, 2022, 35: 16344-16359.
15. Tolstikhin I O, Houlsby N, Kolesnikov A, et al. Mlp-mixer: An all-mlp architecture for vision[J]. Advances in neural information processing systems, 2021, 34: 24261-24272.
16. Gao P, Lu J, Li H, et al. Container: Context aggregation network[J]. arxiv preprint arxiv:2106.01401, 2021.
17. Zhiheng K, Ning L. PyramNet: Point cloud pyramid attention network and graph embedding module for classification and segmentation[J]. arxiv preprint arxiv:1906.03299, 2019.
18. Chen J, Kao S, He H, et al. Run, don't walk: chasing higher FLOPS for faster neural networks[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 12021-12031.
19. Tolstikhin I O, Houlsby N, Kolesnikov A, et al. Mlp-mixer: An all-mlp architecture for vision[J]. Advances in neural information processing systems, 2021, 34: 24261-24272.
20. Ren Yiwei, Zeng Kun. Design of an information monitoring scheme for foreign object intrusion in emergency lanes on long downhill sections of national and provincial mountain highways. Highway, 2022, 67(06): 237-243.
21. Zhu J, Hu T, Zheng L, et al. YOLOv8-C2f-Faster-EMA: an improved underwater trash detection model based on YOLOv8[J]. Sensors, 2024, 24(8): 2483.
22. Xu Jing, Shi Shugang, Li Shunwei, et al. Improved design of mobile clearing facilities for highway pavement obstacles. China Science and Technology Information, 2022, (03): 67-68.
23. Rezatofighi H, Tsoi N, Gwak J Y, et al. Generalized intersection over union: A metric and a loss for bounding box regression[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2019: 658-666.

24. Zhang H, Xu C, Zhang S. Inner-iou: more effective intersection over union loss with auxiliary bounding box[J]. arxiv preprint arxiv:2311.02877, 2023.
25. Ketenci H Ç, Özsoy S, Aydoğdu H İ, et al. Drowning in submerged cars caused by traffic accidents. Turkish Journal of Trauma & Emergency Surgery, 2022, 28(8): 1115.
26. Pinggera P , Ramos S , Gehrig S ,et al.Lost and Found: Detecting Small Road Hazards for Self-Driving Vehicles[J].IEEE, 2016.DOI:10.1109/IROS.2016.7759186.
27. Yang Jiechao, Xu Jiangchun, Lu Wanrong, et al. Study on Embedded Expressway Clearance Intrusion Detection and Tracking. Automation Instrumentation, 2018, 39(12): 70-73+80.
28. Li X, Wang W, Wu L, et al. Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection. Advances in Neural Information Processing Systems, 2020, 33: 21002-21012.
29. Pandey K C, Kumar N, Gupta K B, etal. An Efficient Algorithm for Roadways Obstacles Detection Using Smartphone in Infrastructure-less Environment. International Journal of Recent Technology and Engineering (IJRTE), 2019, 8 (2): 5976-5982.
30. Sunkara R, Luo T. No more strided convolutions or pooling: A new CNN building block for low-resolution images and small objects[C]//Joint European conference on machine learning and knowledge discovery in databases. Cham: Springer Nature Switzerland, 2022: 443-459.
31. Xiao Z, Jun Y, Tong X, et al. Road Obstacle Object Detection Based on Improved YOLO V4. International Journal of Advanced Network, Monitoring and Controls, 2021, 6 (3): 18-25.
32. Balasundaram A, Shaik A, Prasad A, et al. On-road obstacle detection in real time environment using an ensemble deep learning model. Signal, Image and Video Processing, 2024, 18 (6-7): 5387-5400.
33. Rateke T, von Wangenheim A. Passive vision road obstacle detection: a literature map**[J]. International Journal of Computers and Applications, 2022, 44(4): 376-395.
34. Yang Qianyun, Shen Yan. Corn Disease and Pest Image Recognition Based on Improved YOLO v8s Model[J]. Jiangsu Agricultural Sciences, 2025, 53(05): 231-243.
35. Jiao J, Tang Y M, Lin K Y, et al. Dilateformer: Multi-scale dilated transformer for visual recognition[J]. IEEE Transactions on Multimedia, 2023, 25: 8906-8919.
36. Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 13713-13722.
37. Han K, Wang Y, Tian Q, et al. Ghostnet: More features from cheap operations[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 1580-1589.
38. Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C]//Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.
39. Li C, Li L, Jiang H, et al. YOLOv6: A single-stage object detection framework for industrial applications[J]. arxiv preprint arxiv:2209.02976, 2022.
40. Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2023: 7464-7475.
41. Tan M, Pang R, Le Q V. Efficientdet: Scalable and efficient object detection[C]//Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2020: 10781-10790.
42. Yu X, Gong Y, Jiang N, et al. Scale match for tiny person detection[C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. 2020: 1257-1265.
43. Saemi M M, See J, Tan S. Lost and found: Identifying objects in long-term surveillance videos//2015 IEEE International Conference on Signal and Image Processing Applications (ICSIPA). IEEE, 2015: 99-104.
44. Pinggera P, Ramos S, Gehrig S, et al. Lost and found: detecting small road hazards for self-driving vehicles//2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE, 2016: 1099-1106.

Advances in Transportation Studies: an international Journal 68 (2026)

ISBN 979-12-218-2541-1 – ISSN 1824-5463 – DOI 10.53136/97912218254112 – pag. 19-38

Driving behavior and risk analysis in urban tunnels with super-large cross-sections

S. Gao[1] Z.P. Fu[2,3] B. Li[1,4*]
B. Li[2,3] T. Zhu[5*] J. He[2,3]

[1]*Shijiazhuang Transportation Investment Development Group Co., Ltd., Shijiazhuang 050000, Hebei, China*
[2]*CCCC First Highway Consultants Co. Ltd., Xi'an 710068, Shaanxi, China*
[3]*State Key Laboratory of Green and Long-Life Road Engineering in Extreme Environment, Xi'an 710065, Shaanxi, China*
[4]*Shijiazhuang Tiedao University, Shijiazhuang 050043, Hebei China*
[5]*College of Transportation Engineering, Chang'an University, Xi'an 710064, Shaanxi, China*

[*]*corresponding authors: Biao Li: hb_libiao@126.com; Tong Zhu: zhutong@chd.edu.cn*

subm. 12nd June 2025 approv. after rev. 19th August 2025

Abstract

With the development of urban roads, super-large cross-section urban tunnels (SCUT) with four or more lanes appear. However, there is a lack of research regarding the impact of super-large cross-section urban tunnels on driver behavior and associated risk augmentation, leading to insufficient evidence for optimizing traffic facility configurations for SCUT. This paper investigates driving behavior and risk factors in SCUT by driving simulation. The virtual environment for variety urban tunnels is constructed. By leveraging the collected driving behavior data, a driving behavior risk index (DBRI) tailored specifically for tunnels is developed. Moreover, machine learning methods are employed to extract the non-linear relationships among the driving behavior risk index, the number of lanes in the tunnel, the illumination luminance, the color temperature of the tunnel, and the width of the tunnel. The results of the study indicate the following: (1) Compared to tunnels with three lanes or fewer, the risk associated with driving in SCUT is reduced. (2) The color temperature of the tunnel environment has a significant impact on driving behavior, followed by tunnel brightness. Maintaining a high-level range of color temperature in the tunnel environment is beneficial for reducing driving risks. (3) Wider lanes may potentially increase the risk of unsafe driving behavior in SCUT. Finally, based on the results, pertinent recommendations are put forward.

Keywords – driving behavior, urban tunnels, risk, driving simulation, super-large cross-section

1. Introduction

With the advancement of urban infrastructure in China, the construction of urban tunnels has been steadily increasing. By the end of 2021, a total of 406 urban road tunnels had been constructed across 82 cities in China [1].

The tunnel environment is confined, with relatively narrow internal space, limited light conditions, and poor air circulation, which can easily cause drivers to experience psychological oppression. This also has a certain impact on driving behavior [2]. In recent years, there has been an emergence of super-large cross-section urban tunnels (SCUT) with four or more lanes [3].

However, the characteristics of driving behaviors within the SCUT environment remain ambiguous, and the distinctions compared to those in conventional tunnels have yet to be thoroughly investigated. This gap in knowledge results in a lack of sufficient guidance and basis for establishing appropriate lighting conditions and traffic engineering measures.

To bridge this gap, this paper first develops a virtual environment for driving simulation tests within a SCUT. A comprehensive set of behavioral indicators is then collected to construct an urban tunnel driving behavior risk index. Furthermore, a quantitative model is established to analyse the relationship between the number of lanes in the tunnel section, lighting conditions, and driving behavior. The ultimate goal of this study is to explore the differences between SCUTs and conventional tunnel environments and provide preliminary recommendations regarding lighting and lane width settings for SCUTs.

2. Literature review

A substantial body of research has delved into the characteristics of driving behavior within urban tunnel environments. The research primarily zeroes in on two key themes: driving behavior and risk indicator analysis, as well as the interplay between design elements and driving behavior.

Research focusing on driving behavior and risk indicators has made significant inroads into understanding how drivers act within tunnel settings and identifying associated risk factors. For instance, Miller et al. conducted a systematic investigation into the critical characteristics of driving speed in tunnel environments [4]. Their work lays down foundational insights into how drivers adjust their speeds in response to tunnel-specific factors such as lighting, geometry, and psychological effects. The lighting conditions within a tunnel can create a unique visual environment, influencing a driver's perception of speed and distance. The tunnel's geometry, including its length, curvature, and slope, also affects driving speed as drivers need to adapt to the changing road conditions. Moreover, the psychological effects, such as the feeling of confinement or anxiety, can lead to variations in speed. Additionally, Calvi et al. utilized driving simulator technology to analyze driving behavior, employing indicators like speed, acceleration, and lateral positioning [5]. Their findings emphasize the necessity for enhanced safety measures at tunnel entrances and exits. At these points, drivers are transitioning between different driving environments, and their behavior may be more erratic. For example, they may accelerate too quickly upon exiting the tunnel or decelerate abruptly upon entering, increasing the risk of accidents. This highlights the importance of considering these risk - related behaviors in tunnel design and management.

Research on the interplay between design elements and driving behavior encompasses several important aspects, including the impact of lighting conditions, traffic design optimization, and the use of guidance systems. Previous literature has extensively evaluated the influence of lighting systems on driver behavior and traffic safety within road tunnels through driving simulator experiments. The results consistently show that improved lighting conditions enhance driver performance. Better lighting facilitates earlier detection of critical situations, such as the presence of obstacles or sudden changes in road conditions. It also improves lateral trajectory control during transitions, enabling drivers to maintain a more stable and safer driving path within tunnels [6-8]. This indicates that lighting is a crucial factor in influencing driving behavior and should be carefully considered in tunnel design. A well-lit tunnel can reduce driver fatigue and improve their ability to react to unexpected events, thereby enhancing overall safety.

Studies have also emphasized the optimization of tunnel facilities, which integrate speed management [9], as an effective approach to reducing accident risks. By optimizing traffic design