

Article

Not peer-reviewed version

Algorithm Improvement for Mobile Event Detection with Intelligent Tunnel Robots

Li WAN , Zhenjiang LI , [Changan ZHANG](#) , Guangyong CHEN , [Panming ZHAO](#) ^{*} , [Kewei WU](#)

Posted Date: 1 August 2024

doi: 10.20944/preprints202408.0009.v1

Keywords: YOLOv9 + DeepSORT; Edge Computing; Mobile Event Detection; Smart Tunnels



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Algorithm Improvement for Mobile Event Detection with Intelligent Tunnel Robots

Li Wan ¹, Zhenjiang Li ¹, Changan Zhang ¹, Guangyong Chen ¹, Panming Zhao ^{2,*} and Kewei Wu ²

¹ Shandong Provincial Communications Planning and Design Institute, Shandong Jinan, 250101, China;

² Beijing Zhuoshi Zhitong Technology Co., Ltd., Beijing 100080, China

* Correspondence: zhaopanming_2023@163.com

Abstract: Mobile inspections conducted by intelligent tunnel robots are instrumental in broadening the inspection reach, economizing on inspection expenditures, and augmenting the operational efficiency of inspections. Despite differences from fixed surveillance, mobile-captured traffic videos have complex backgrounds and device conditions that interfere with accurate traffic event identification, warranting more research. This paper proposes an improved algorithm based on YOLOv9 and DeepSORT for intelligent event detection in edge computing mobile device using the intelligent tunnel robot. The enhancements comprise the integration of the Temporal Shift Module to boost temporal feature recognition and the establishment of logical rules for identifying diverse traffic incidents in mobile video imagery. Experimental results show that our fused algorithm achieves a 93.25% accuracy rate, an improvement of 1.75% over the baseline. The algorithm is also applicable to inspection vehicles, drones, and autonomous vehicles, effectively enhancing the detection of traffic events and improving traffic safety.

Keywords: YOLOv9 + DeepSORT; edge computing; mobile event detection; smart tunnels

1. Introduction

China leads globally in the number and scale of highway tunnels, with 21,316 in total, featuring 1,394 extra-long tunnels over 10 km and 5,541 long tunnels from 3 to 10 km. Tunnels have higher accident rates due to their enclosed, narrow design and lack of emergency lanes. Current manual inspection and reactive management lead to issues like poor safety awareness, delayed repairs, inefficient rescues, and lack of real-time traffic monitoring.[1–3]. Traditional tunnel monitoring methods like inspection vehicles, fixed sensors, and manual checks have issues with high costs, incomplete coverage, low efficiency, and risk of oversight. This struggles to keep up with increasing operational demands. Thus, integrating AI into mobile inspection robots is crucial to smarten up tunnel inspections effectively.

Additionally, common deep learning techniques like CNNs, R-CNN, Fast R-CNN, Faster R-CNN, and YOLO effectively extract image features.[4–7]. In [8], the vehicle detection and tracking model employs the YOLOv5 object detector in conjunction with the DeepSORT tracker to identify and monitor vehicle movements, assigning a unique identification number to each vehicle for tracking purposes. In [9], YOLOv8 has been enhanced for highway anomaly detection using the CBAM attention mechanism, MobileNetV3 architecture, and Focal-EIoU Loss.

Target recognition and tracking algorithms for tunnel robots, essential for traffic safety and incident inspection, have been researched and implemented globally in recent years [10–12]. Among them, the method based on deep learning is more advantageous [13,14], Papageorgiou C P et al. proposed a general trainable framework suitable for complex scenarios, in which the system performs feature learning from examples, and does not rely on prior (artificial labeling) models [15]; Yang B et al. [16] proposed an improved convolutional neural network based on Faster RCNN, which enhances the detection performance of small vehicles by fine-tuning the model using vehicle samples specifically for vehicle detection. Lee W J et al. [17] developed a selective multi-stage feature detection

framework based on convolutional neural network, which fully extracts the feature information of vehicle images and exhibits strong robustness to noise. Li Linhui et al. [18] considered the underbody shadow as the candidate area and designed edge enhancement and adaptive feature segmentation algorithms for vehicle detection, which effectively distinguish between underbody shadow and non-underbody shadow interferences, thereby improving the accuracy and reliability of vehicle detection [19].

In summary, the current application of tunnel robots primarily focuses on firefighting, rescue operations, and crack detection [20–22]. However, there are limitations in event detection for tunnel robots [23], as well as insufficient consideration of event characteristics from the perspective of mobile devices. The video background and target pixel position change during robot movement, increasing the computational complexity of the detection process, and if the fixed detection equipment algorithms are still employed[24,25], it would be challenging to ensure the accuracy and real-time nature of event analysis[26–28]. This research aims to enable smart event detection in mobile devices by integrating target tracking and temporal feature extraction. The YOLOv9+DeepSORT algorithm is used for this purpose, along with revised abnormal event detection rules tailored for mobile scenarios. The aim is to improve real-time detection and response to events like tunnel fires, accidents, and pedestrian incursions, thus boosting the intelligence of smart tunnel systems.

2. Methodology

The system’s workflow includes the following steps for mobile sensing: High-definition cameras, infrared cameras, and gas sensors mounted on tunnel inspection robots transmit data to the robots’ edge computing devices. Algorithms deployed on these devices use AI to identify traffic events and send the event information to an AI-based event warning platform. For fixed surveillance equipment, fixed edge computing devices recognize traffic events in the video feeds and relay the output to the AI event warning platform. The AI event warning platform then assesses the events and responds by either using the tunnel robot to issue on-site announcements or dispatching rescue personnel for on-site handling. As shown in Figure 1.

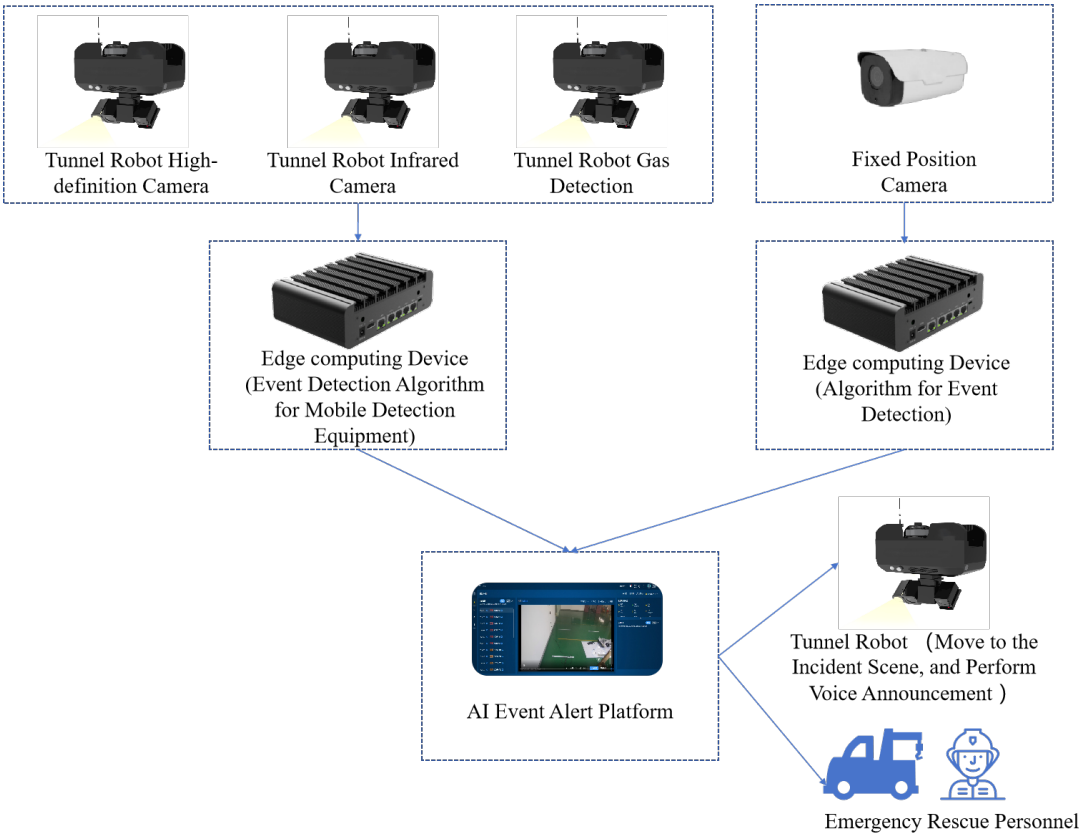


Figure 1. Methodological workflow of the proposed system.

2.1. Intelligent Tunnel Robot

The intelligent tunnel robot has two main parts: the robot itself and the carrying system. The carrying system includes cameras, sensors, and an AI analyzer. The robot has a 5G antenna, wheels, radar, and other features. Its main jobs are smart monitoring, inspection, control, and emergency communication, As shown in Figure 2. Edge Computing Specifications: Processor ARM Cortex-A53 8-core; AI Computing Power 17.6 TOPS INT8.

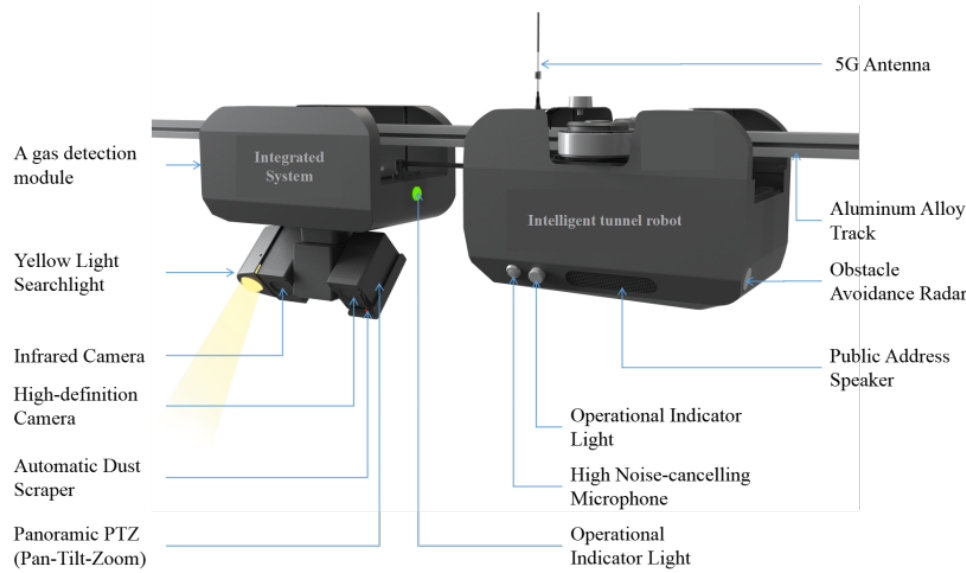


Figure 2. Intelligent Tunnel Robot Module Composition.

Intelligent monitoring uses AI for real-time detection and surveillance of issues like abnormal parking, pedestrian and trespassing. It gathers data on brightness and environmental conditions, including temperature, humidity, gases, and smoke levels. The system provides 360-degree coverage to eliminate blind spots and improve awareness of the tunnel’s status.

Intelligent inspection offers daily and emergency modes. In daily mode, the robot autonomously patrols at 1-1.5m/s, identifying issues like abnormal vehicles and unauthorized entries. In emergency mode, it can quickly reach incidents by speeding up to 7m/s. It uses a light projector to identify and alert fire and deploys a sound and light warning system for effective communication, which aids in preventing additional accidents.

2.2. Designing Algorithms

To achieve high accuracy and efficiency in traffic event detection for mobile device states, considering the limited computing power of edge devices, a lightweight model is required. This paper employs YOLOv9 for object detection in extracted video frames. Detected objects undergo convolution, self-attention mechanisms, and temporal sequence movement to enhance the extraction of channel, spatial, and temporal features, resulting in feature maps. DeepSORT is then used for tracking targets in the feature maps, employing a Kalman filter prediction model and association metrics that use Mahalanobis distance and a Deep Appearance Descriptor with Hungarian Assignment to track trajectories of different objects. Logical rules are applied to determine the occurrence of traffic events, including the identification of six types of traffic events: abnormal parking, pedestrians, wrong-way driving, fires, and road flooding. The event detection results are then outputted.

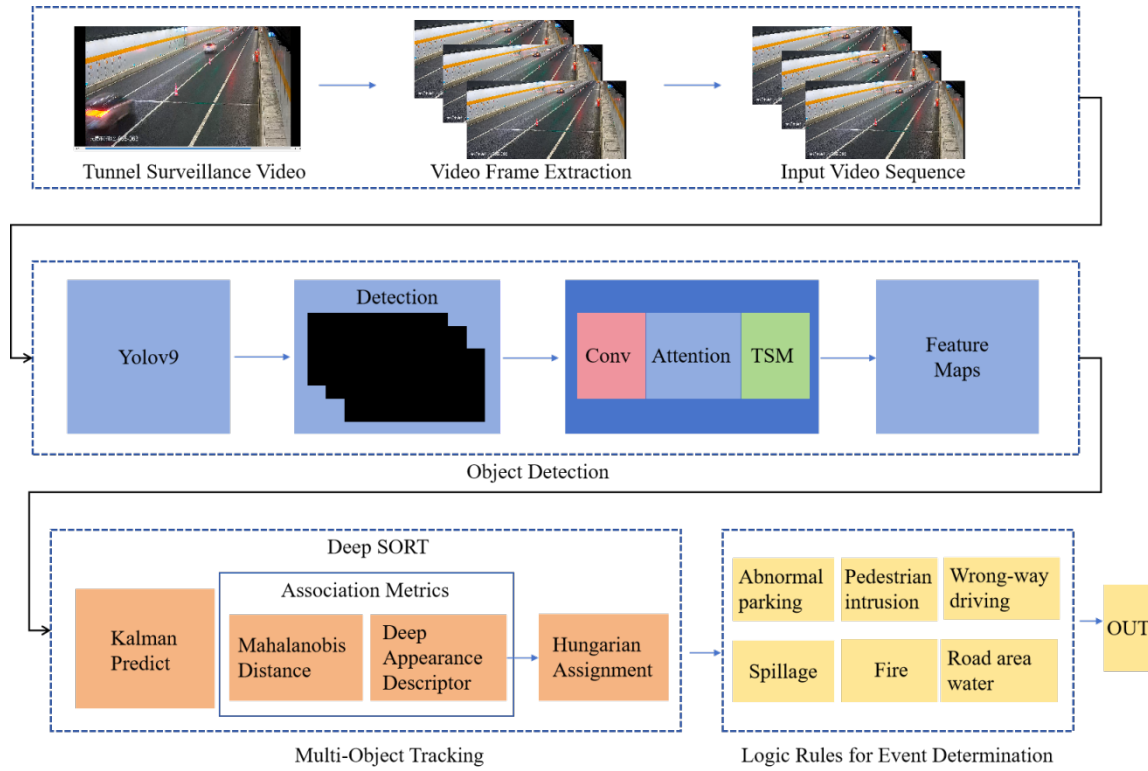


Figure 3. Comprehensive Algorithm Framework Diagram.

2.2.1. Object Detection Algorithms

Traditional deep learning methods often suffer from significant information loss during the feature extraction and spatial transformation processes as the input data propagates through the layers of deep networks. To address the issue of information loss during data transmission, YOLOv9 proposes two solutions: (1) It introduces the concept of Programmable Gradient Information (PGI) to cope with the various transformations required for a deep network to achieve multiple objectives. (2) It designs a novel lightweight network structure based on gradient path planning—Generalized Efficient Layer Aggregation Network (GELAN). Validation on the MS COCO dataset for object detection demonstrates that GELAN can achieve better parameter utilization using only conventional convolutional operators, and PGI can be applied to models of various sizes, from lightweight to large-scale. Moreover, it can be used to obtain comprehensive information, enabling models trained from scratch to achieve better results than state-of-the-art models pretrained on large datasets[29].

2.2.2. Temporal Shift Module

This paper further enhances the algorithm's sensitivity to timing features in traffic videos by introducing weighted time-shifted operations on video frame features. By partially shifting the feature map along the time dimension, information exchange and fusion between adjacent video frames are facilitated. This improves the expression capacity of features in capturing temporal relationships within the time series. The features of the video frame are divided into 4 parts (c_1, c_2, c_3, c_4) in the channel dimension. The features of each channel c_1 are then shifted one time block backwards along the time dimension, while keeping the characteristics of the c_2, c_3 channels unchanged. Additionally, the features of the c_4 channel are shifted one time block forward along the time dimension, and the weights (w_1, w_2, w_3, w_4) are assigned to the movement of each channel. The relationship between the time-shifted feature $c'_i (i=1,2,3,4)$ and the initial feature $c_i (i=1,2,3,4)$ is

expressed as shown in the formula. Among them, T^{-1}, T^0, T^{+1} are the previous moment, the current moment, and the next moment, respectively. The formula is as follows:

$$\begin{aligned} c_1' &= w_1 c_1^{T^{-1}}, c_2' = w_2 c_2^{T^0}, \\ c_3' &= w_3 c_3^{T^0}, c_4' = w_4 c_4^{T^{+1}} \end{aligned} \quad (1)$$

2.2.3. DeepSORT Multi-Object Tracking Algorithms

The DeepSORT algorithm builds upon the SORT algorithm by introducing Matching Cascade and confirmation of new trajectories. It integrates the Kalman filter to estimate the mean and covariance of the track. By utilizing a gate matrix, it constrains the cost matrix to limit the distance between the state distribution of the Kalman filter and the measured value, thereby reducing matching errors. Additionally, the algorithm incorporates the image depth feature within the detection frame range of the YOLOv9 target detection algorithm. To extract features, it leverages the ReID domain model, reducing the occurrence of tracking target ID switches and enhancing target matching accuracy [32].

It is evident that the overall process unfolds as follows:

- (1) Use the Kalman filter to predict where objects will move;
- (2) Apply the Hungarian algorithm to match these predictions with actual objects seen in the current frame, leverage Mahalanobis distance for measuring the similarity between detected and tracked objects;
- (3) Utilize the Deep Appearance Descriptor for the extraction of target appearance features, re-identification of targets, and association of targets, thereby optimizing the matching performance.

2.3. Logic Rules for Event Determination in Mobile Device

This paper focuses on the development of traffic event detection logic for four specific scenarios within tunnel environments, utilizing the unique characteristics of mobile device surveillance footage. The scenarios addressed are Abnormal Parking, Pedestrian Intrusion, Wrong-way Driving, and Flame detection.

The detection process is primarily object recognition-based, aimed at mitigating the impact of the dynamic and complex environment encountered during the movement of the monitoring devices. Additionally, the paper discusses the conversion of pixel coordinates within the video to real-world coordinates, while also accounting for the movement of the device itself. This is crucial to avoid misclassifications, such as incorrectly identifying stationary vehicles as wrong-way drivers due to relative motion between the vehicle and the moving device.

The standard duration for event detection is set at 15 seconds. For flame detection, to enhance the reliability of the recognition, a threshold for identifying high-temperature areas is proposed, with the stipulation that the area must be at least 0.4 meters by 0.4 meters.

The detailed process for event determination is outlined below, using abnormal parking as an illustrative example: the position coordinates of events in different video frame states are deduced from the real-time position of camera motion and pixel coordinates in camera lenses. Furthermore, whether a vehicle is moving or stationary within a certain period of time is determined by calculating the distance between positions in consecutive N frames (representing a specific time interval). As shown in Figure 4.

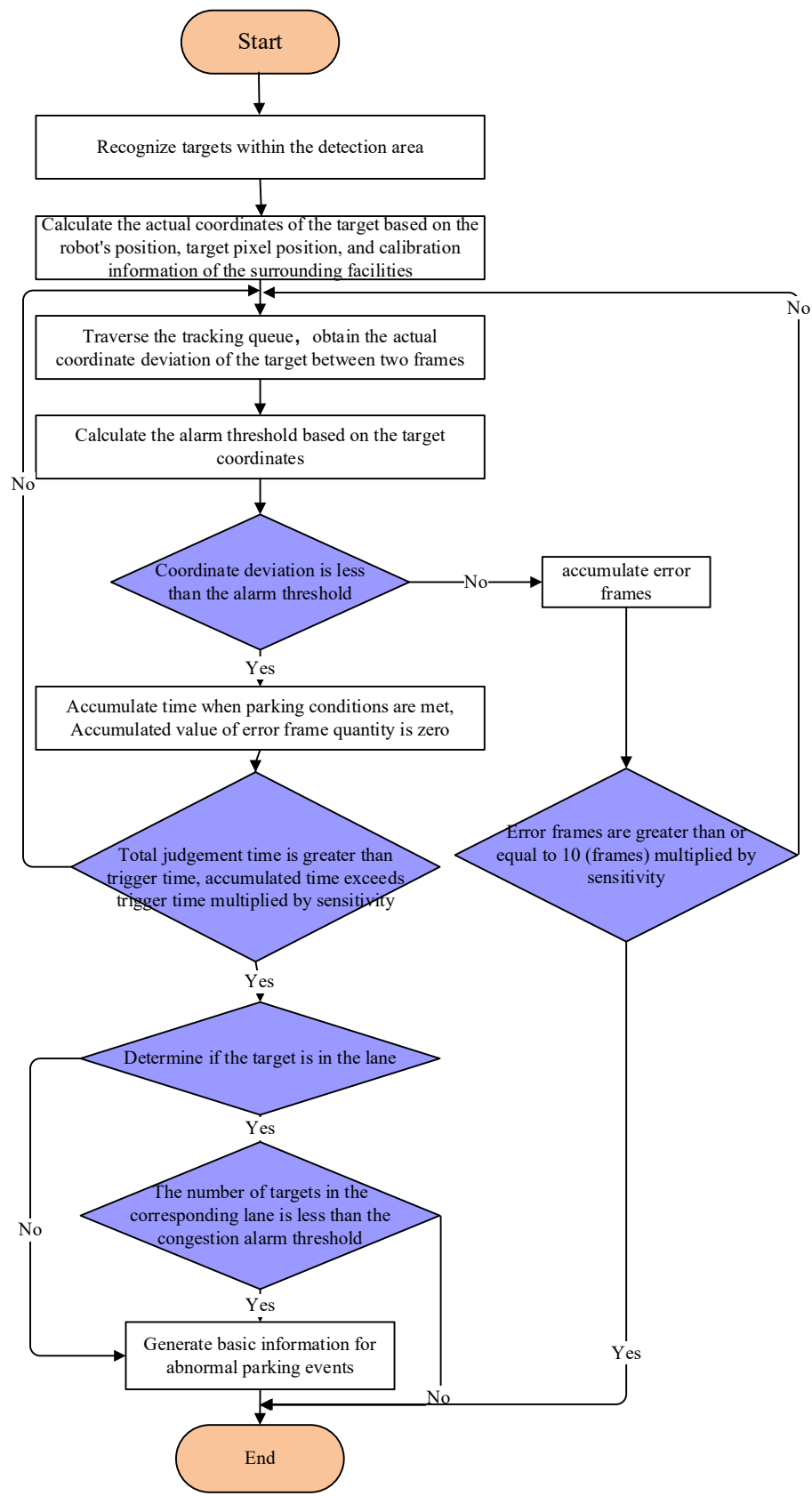


Figure 4. Flowchart of rules for abnormal parking events with mobile cameras.

3. Experimental Setup

3.1. Datasets

This algorithm uses a proprietary training set consisting of 210,000 images, which includes: 55,000 images for abnormal parking detection, 5,0000 images for wrong-way driving detection,, 50,000 images for pedestrian detection, 55,000 images for flame detection.

To validate the experiment’s reliability, the intelligent tunnel robot’s mobile state algorithm underwent testing in two environments: simulated tunnels and mobile video capture. With a priority on replicating real tunnel conditions, the majority of tests were conducted in the simulated tunnel setup, with each event class tested 80 times. In contrast, mobile video testing was performed 20 times per event class. Test event picture as shown in Figure 5.



Figure 5. Experiment site and video scene.

3.2. Ablation Study

This paper conducts ablation studies on YOLOv9, YOLOv9+SORT, YOLOv9+DeepSORT, and YOLOv9+DeepSORT+TSM. The ablation study comprises the following key points:

A preliminary evaluation of the YOLO algorithm suite was performed, encompassing YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, YOLOv8, and YOLOv9, prior to the commencement of ablation studies on the integrated algorithmic framework.

In the course of the ablation study, a module incorporating logical rules for event judgment was added to the YOLOv9 + DeepSORT + TSM configuration.

The experimental design involved segmenting the testing area based on distance, with multiple independent sampling tests conducted within each segment to ensure the reliability of the findings, as depicted in Figure 6,the test area is divided into three effective regions: the proximal region (T1), the middle region (T2-T4), and the distal region (T5).

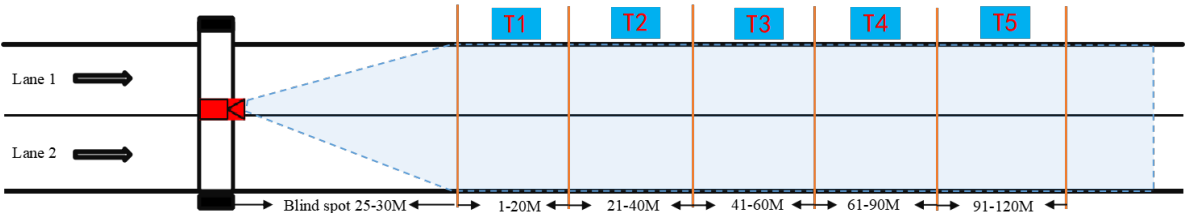


Figure 6. Test viewshed division reference diagram.

The evaluation criteria for the experimental outcomes are delineated by the following metrics: Accuracy, Precision, Recall, MAP(Mean Average Precision), and FPS(Frames Per Second).

For a given sample, there are four possible outcomes of classification predictions: True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN). By aggregating these outcomes across all test samples, one can derive evaluation metrics such as Accuracy, Precision, and Recall.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

$$Precision = \frac{TP}{TP + FP} \quad (3)$$

$$Recall = \frac{TP}{TP + FN} \quad (4)$$

$$AP = \sum_{n=1}^N P_n \cdot \omega_n \quad (5)$$

In the formula, P_n is the precision at the n -th recall level, ω_n is the weight corresponding to the recall level (usually the recall rate itself), N is the number of recall levels.

$$MAP = \frac{1}{N} \sum_{i=1}^N AP_i \quad (6)$$

In the formula, N represents the total number of classes, AP_i denotes the Average Precision for the i -th class.

4. Results

This article compares the performance of YOLOv5s, YOLOv5m, YOLOv5l, YOLOv5x, YOLOv8, and YOLOv9. The outcomes of this comparison are detailed in Table 1 below. Based on Table 1, the YOLOv5s, YOLOv5l, and YOLOv5m algorithms exhibit higher computational efficiency compared to YOLOv8 and YOLOv9. In terms of Recall, YOLOv5s, YOLOv5m, YOLOv5l, and YOLOv8 slightly outperform YOLOv9. However, YOLOv9 surpasses other YOLO algorithms in both MAP and Precision. Considering the high complexity of the tunnel inspection scenario and the limited computing power of edge devices, this study selects YOLOv9 for object detection.

Table 1. Comparison table of YOLO algorithm.

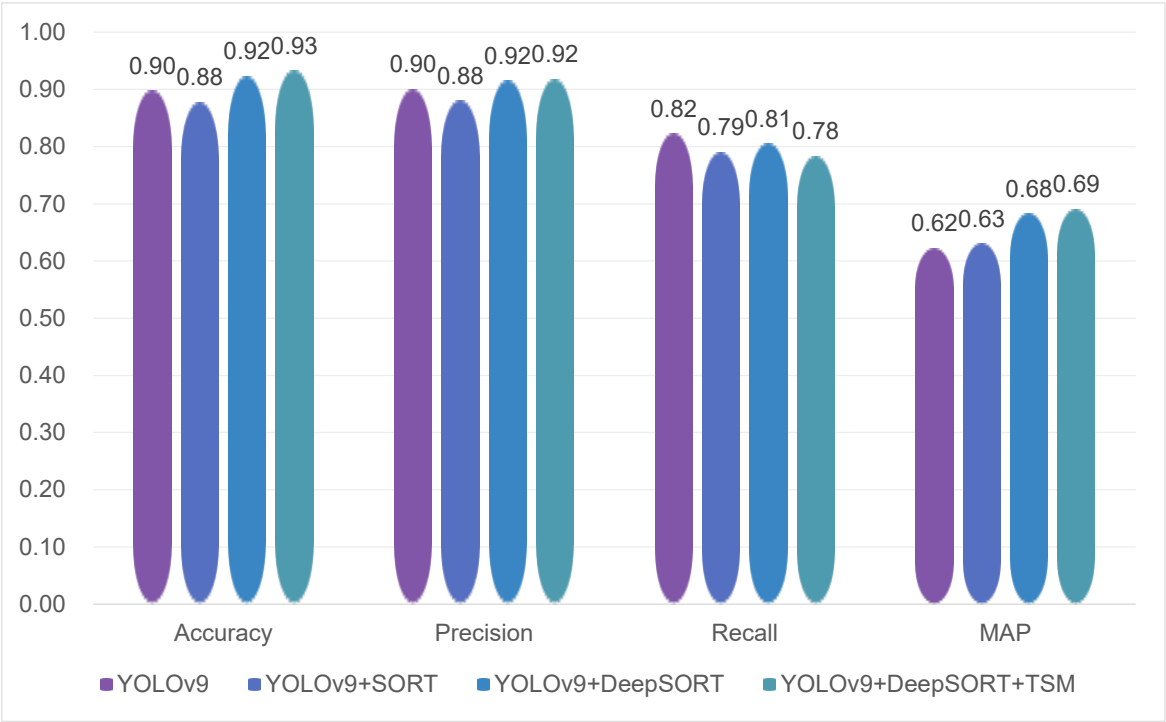
Model name	Epoch/Iteration	Batch Size	MAP	Processing frame rate (fps)	Precision	Recall
YOLOv5s	300epoch	512	57.1	142	84.77%	76.44%
YOLOv5m	300epoch	512	58.7	105	85.71%	75.98%
YOLOv5l	300epoch	512	61.6	68	85.93%	77.29%
YOLOv5x	300epoch	512	61.7	55	89.17%	77.54%
YOLOv8	300epoch	512	61.9	45	89.23%	77.32%
YOLOv9	300epoch	512	62.1	60	90.08%	76.59%

Based on the results of the ablation study in Table 2, the YOLOv9+DeepSORT+TSM algorithm proposed in this paper achieves higher Accuracy, Precision, and MAP, outperforming YOLOv9, YOLOv9+SORT, and YOLOv9+DeepSORT. Specifically, the Accuracy is 3.5% higher than the baseline algorithm YOLOv9, the Precision is 1.75% higher, and the MAP is 6.75% higher. However, in terms of Recall and FPS, YOLOv9+DeepSORT+TSM is slightly inferior to the baseline algorithm, with a 4% lower Recall and 4.75 frames per second (FPS) less, which is not significantly noticeable in practical applications.

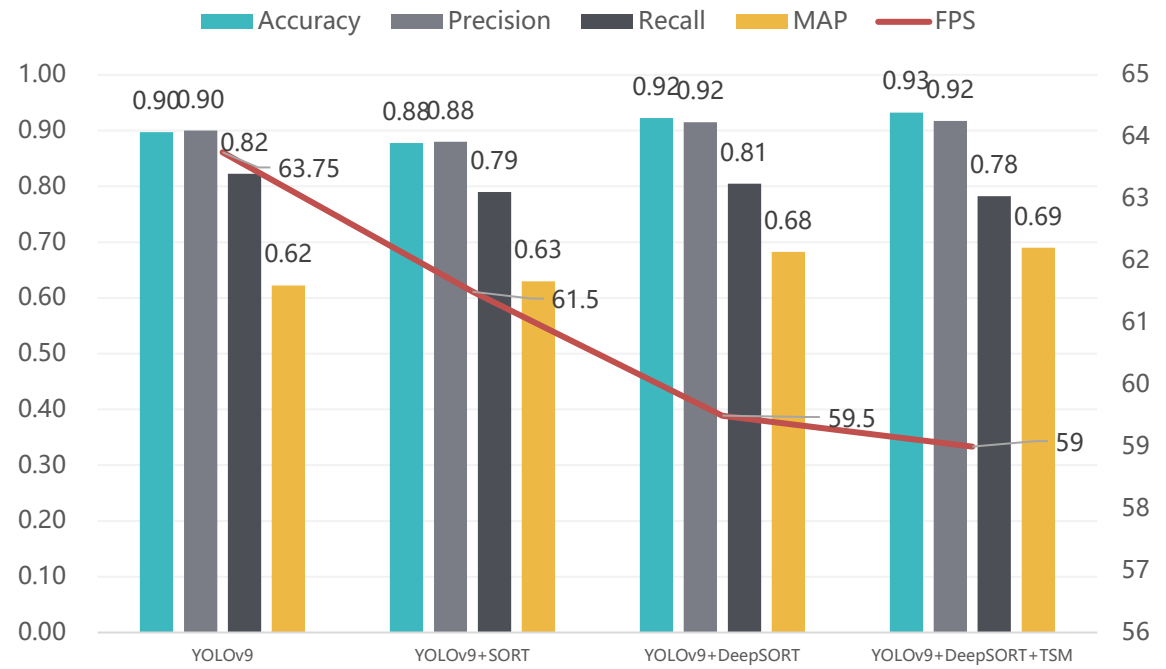
Table 2. The ablation study results.

Algorithms	Traffic events	Accuracy	Precision	Recall	MAP	FPS
YOLOv9	Abnormal parking	0.92	0.90	0.82	0.66	66
	Pedestrians	0.92	0.88	0.80	0.65	64
	Wrong-way driving	0.85	0.92	0.84	0.58	65
	Flame	0.90	0.90	0.83	0.60	60
	Average	0.8975	0.9000	0.8225	0.6225	64
YOLOv9+SORT	Abnormal parking	0.89	0.92	0.86	0.68	66
	Pedestrians	0.90	0.87	0.74	0.60	64
	Wrong-way driving	0.84	0.85	0.84	0.64	58
	Flame	0.88	0.88	0.72	0.60	58
	Average	0.8775	0.8800	0.7900	0.6300	62
YOLOv9+DeepSORT	Abnormal parking	0.92	0.90	0.80	0.70	65
	Pedestrians	0.93	0.94	0.78	0.66	60
	Wrong-way driving	0.94	0.92	0.82	0.68	55
	Flame	0.90	0.90	0.82	0.69	58
	Average	0.9225	0.9150	0.8050	0.6825	60
YOLOv9+DeepSORT+TSM	Abnormal parking	0.94	0.92	0.85	0.72	60
	Pedestrians	0.95	0.94	0.76	0.70	58
	Wrong-way driving	0.92	0.90	0.78	0.66	60
	Flame	0.92	0.91	0.74	0.68	58
	Average	0.9325	0.9175	0.7825	0.6900	59

The experimental results are analyzed and displayed through bar charts and line charts. As shown in Figure 7a, It can be observed that the recognition effect achieved by integrating object recognition algorithms with multiple object tracking algorithms is superior to that obtained by using only object recognition algorithms. Furthermore, the addition of a temporal smoothing module (TSM) to the fusion algorithm further enhances the recognition effect compared to the simple integration of object recognition and multiple object tracking algorithms. In Figure 7b, it is evident that all four algorithm structures based on YOLOv9 have achieved an Accuracy of 90% or higher, indicating that the YOLOv9 algorithm has demonstrated applicable performance in the experiments. Additionally, with the addition of fusion algorithm modules, there is a slight decrease in frames per second (FPS). However, the YOLOv9+DeepSORT+TSM shows a negligible decrease in processing speed compared to YOLOv9+DeepSORT, suggesting that the incorporation of the TSM module has a minimal impact on computational power.



(a)



(b)

Figure 7. (a) Performance Evaluation of Different Fusion Algorithms. (b) Performance Evaluation of Different Fusion Algorithms.

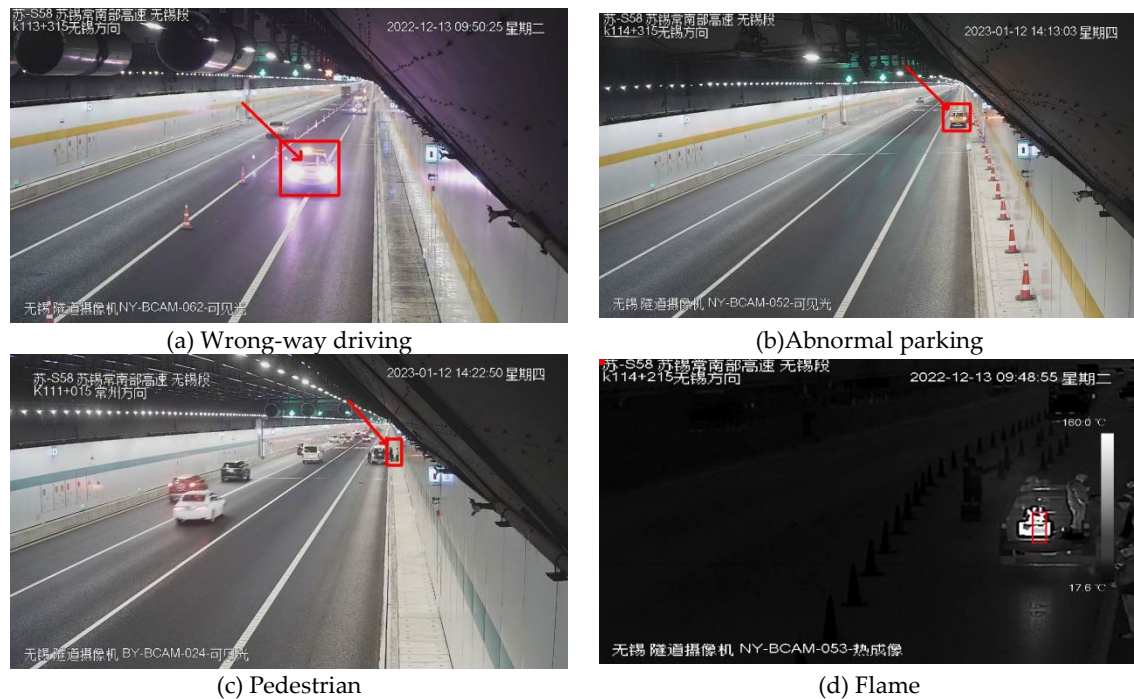


Figure 8. Traffic Event Detection Effectiveness.

5. Conclusions

This research paper aims to meet the demands of daily inspection, emergency handling, and intelligent operation by focusing on event detection in the mobile state. Specifically, it focuses on event detection based on intelligent tunnel robots in smart tunnels operating in a mobile state. The core contributions and conclusions of this study are as follows:

Design of Event Detection Algorithm in Moving State: To effectively detect events in the dynamic state of an intelligent tunnel robot, this paper utilizes YOLOv9 for object detection and identification within tunnels. Through comprehensive improvements in detection speed and recognition accuracy, real-time and accurate detection of tunnel traffic events is achieved. The DeepSORT algorithm is employed to detect image depth features within a specific frame range, while leveraging the ReID domain model for feature extraction. This helps reduce target ID jumps and enhance target matching accuracy. Time Shift Module (TSM) is utilized to model temporal information from video frame features, capturing time series characteristics and enhancing the algorithm's ability to detect time-related features.

Redesign of Event Analysis Rules, such as Abnormal Parking in Mobile Device States. By employing the real-time position of camera motion and pixel coordinates, the position coordinates of events in different video frame states are calibrated and derived. Based on the distance of the position in continuous N-frame pictures over a period of time, calculations determine whether the vehicle is moving or stationary. If the vehicle's coordinate position does not change within the detection frame and time frame, it is considered abnormal parking.

Although the proposed improved method for event detection algorithms in the mobile detection state based on intelligent tunnel robots demonstrates superior performance in recognizing mobile device states, there remains scope for further enhancement. Future studies can explore the implementation of transformer algorithms in the moving state, optimize data acquisition and algorithm training based on actual tunnel scene events, and further improve the real-time accuracy of event detection. These advancements will contribute to enhanced intelligent operation and traffic safety management of key road sections such as tunnels, thereby elevating the construction and operational standards of smart tunnels.

Author Contributions: Conceptualization, Li Wan and Kewei Wu; Methodology, Li Wan and Zhenjiang Li; Software, Kewei Wu; Validation, Changan Zhang, Guangyong Chen and Panming Zhao; Formal Analysis, Changan Zhang, Guangyong Chen and Panming Zhao; Investigation, Changan Zhang, Guangyong Chen and Panming Zhao; Resources, Li Wan; Data Curation, Kewei Wu and Panming Zhao; Writing—Original Draft Preparation, Li Wan, Changan Zhang, and Guangyong Chen; Writing—Review & Editing, Li Wan and Kewei Wu; Visualization, Kewei Wu and Panming Zhao; Supervision, Kewei Wu; Project Administration, Li Wan; Funding Acquisition, Li Wan.

Funding: Shandong Provincial Department of Transportation (a government agency) in China; Project Name: Research on Tunnel Control Technology Based on Improving Traffic Efficiency and Operational Safety (June 2019 to June 2022).

Data Availability Statement: The datasets presented in this article are not readily available because the data are part of an ongoing study. And researchers have entered into a contract with a third party, which includes confidentiality clauses requiring that the relevant data shall not be disclosed within a specified period. Requests to access the datasets should be directed to authors.

Acknowledgements: This work was supported by project “Research on Intelligent Management and Control Technology of Tunnels Based on Improving Traffic Efficiency and Camp Safety”. The authors would like to thank our colleagues of Shandong Provincial Communications Planning and Design Institute and Beijing Zhuoshi Zhitong Technology Co., Ltd. for their excellent technical support and valuable suggestions.

Conflict of Interest: Authors have no conflict of interest relevant to this article.

References

1. Li, W. F., C. Yuan, K. Li, H. Ding, Q. Z. Liu and X. H. Jiang, (2021) “Application of Inspection Robot on Operation of Highway Tunnel,” *Technology of Highway and Transport* 37(S1), 35-40.
2. Wu, W. B., (2021) “Automatic Tunnel Maintenance Detection System based on Intelligent Inspection Robot,” *Automation & Information Engineering* 42(1), 46-48.
3. Wang, Z. J., (2020) “Application of robotics in highway operation,” *Henan Science and Technology* (19), 98-100.
4. Lee, C.; Kim, H.; Oh, S.; Doo, I. A Study on Building a “Real-Time Vehicle Accident and Road Obstacle Notification Model” Using AI CCTV. *Appl. Sci.* 2021, 11, 8210.
5. Nancy, P.; Dilli Rao, D.; Babuaravind, G.; Bhanushree, S. Highway Accident Detection and Notification Using machine Learning. *Int. J. Comput. Sci. Mob. Comput.* 2020, 9, 168–176.
6. Ghosh, S.; Sunny, S.J.; Roney, R. Accident Detection Using Convolutional Neural Networks. In *Proceedings of the 2019 International Conference on Data Science and Communication (IconDSC)*, Bangalore, India, 1–2 March 2019.
7. Zhang, X. (2022) Research on Traffic Event Recognition Method Based on Video Classification and Video Description. M. D. Dissertation, ShanDong University, China.
8. Basheer Ahmed, M.I., Zaghdoud, R., Ahmed, M.S., Sendi, R., Alsharif, S., Alabdulkarim, J., Albin Saad, B.A., Alsabt, R., Rahman, A., and Krishnasamy, G., (2023) A Real-Time Computer Vision Based Approach to Detection and Classification of Traffic Incidents. *Incidents. Big Data and Cognitive Computing.* 2023, 7, 22. <https://doi.org/10.3390/bdcc7010022>.
9. Ren, A. H., Y. F. Li, Y. Chen, (2024) Improved Detection of Unusual Highway Traffic Events for YOLOv8. *Proc. of Laser Journal*, <https://link.cnki.net/urlid/50.1085.TN.20240628.0926.002>.
10. Li, Y., (2021) Research on Autonomous Mobile Platform of Intelligent Tunnel Inspection Robot Based on Laser SLAM. M. D. Dissertation, Chang’ an University, China.
11. Manana, M., C. Tu and P. A. Owolawi, (2017) A survey on vehicle detection based on convolution neural networks. *Proc. of 2017 3rd IEEE International Conference on Computer and Communications (ICCC)*. IEEE, Chengdu, China, 1751-1755.
12. Rosso, M. M., M.Giulia, A. Salvatore, A. Aloisio, B. Chiaia and G. C. Marano, (2023) “Convolutional networks and transformers for intelligent road tunnel investigations,” *Computers and Structures*, 275, 106918.
13. Liu, J., Z. Y. Zhao, C. S Lv, Y. Ding, H. Chang, and Q. Xie, (2022) “An image enhancement algorithm to improve road tunnel crack transfer detection,” *Construction and Building Materials*, 348, 128583.
14. Zhang, G. W., J. Y. Yin, P. Deng, Y. Sun, L. Zhou and K. Zhang, (2022) “Achieving Adaptive Visual Multi-Object Tracking with Unscented Kalman Filter” *Sensors*, 22(23), 9106.
15. Papageorgiou, C. P., M. Oren and T Poggio, (1998) A general framework for object detection. *Proc. of Sixth International Conference on Computer Vision (IEEE Cat. No. 98CH36271)*. IEEE, Bombay, India, 555-562.

16. Yang, B., Y. Y. Zhang, J. M. Cao and L. Zou, (2018) On road vehicle detection using an improved faster RCNN framework with small-size region up-scaling strategy. Proc. of In Image and Video Technology: PSIVT 2017 International Workshops, Wuhan, China, Revised Selected Papers 8, Springer International Publishing, 241-253.
17. Lee, W. J., D. S. Pae, D. W. Kim and M. T. Lim, (2017) A vehicle detection using selective multi-stage features in convolutional neural networks. Proc of 17th International Conference on Control, Automation and Systems (ICCAS). IEEE, Jeju, Korea, 1-3.
18. Li, L. H., Z. M. Lun and J. Lian, (2017) "Road Vehicle Detection Method based on Convolutional Neural Network," Journal of Jilin University(Engineering and Technology Edition),47(02), 384-391.
19. Fu, Z. Q. (2020) Research on Key Technology for Tunnel Monitoring of Inspection Robot Based on Image Mosaic. M. D. Dissertation, Ningbo University, China.
20. Ma, H. J., "Design and research of highway tunnel inspection robot system," Western China Communications Science & Technology, 10, 122-124.
21. Tan, S. (2020) Research on Object Detection Method Based on Mobile Robot. M. D. Dissertation, Beijing University of Civil Engineering and Architecture, China.
22. Li, J. Y., "Research on the application of automatic early warning system for safety monitoring of cut-and-cover tunnel foundation pit based on measurement robot technology," Journal of Highway and Transportation Research and Development(Applied Technology Edition), 6(02),267-269.
23. Tian, F, C. L. Meng and Y. C. Liu, "Research on the construction scheme of intelligent tunnel in Jiangsu Taihu Lake for generalized vehicle-road coordination," Journal of Highway and Transportation Research and Development(Applied Technology Edition), 16(04),268-272.
24. Jin, Y., Y. Xu, F. Y. Han, S. Y. He and J. B. Wang, "Pixel-level Recognition Method of Bridge Disease Image Based on Deep Learning Semantic Segmentation," Journal of Highway and Transportation Research and Development(Applied Technology Edition), 16(01),183-188.
25. Song, Z. L. (2021)Intelligent Analysis of Traffic Events based on Software-defined Cameras. M. D. Dissertation, Southwest Jiaotong University, China.
26. Hu, M. (2021)Design and Implementation of Intelligent Video Analysis System based on Deep Learning. M. D. Dissertation, Huazhong University of Science & Technology, China.
27. Zhang, C. S., D. Zhou, "Application of event intelligent detection system based on radar and thermal imaging technology in extra-long tunnels," Journal of Highway and Transportation Research and Development(Applied Technology Edition),16(10),313-316.
28. Dong, M. L.(2022)Research on Highway Traffic Incident Detection Based on Deep Learning. D. M. Dissertation, Xi'an Technological University,China.
29. Wang, C. Y., Yeh, I. H., Liao, H. M, (2024) "YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information." Journal of Computer Science, cs>arXiv:2402.13616.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.