Proceedings of CONF-SEML 2026 Symposium: Importance of Machine Learning Methods and Analysis in Engineering
DOI: 10.54254/2755-2721/2026.AS32421

Traffic Signal Intelligent Control and System Optimization—Intelligent Traffic Signal Design Based on Machine Learning and Signal Sensors

Mingrui Yang

Rutgers University, New Brunswick, USA
m17301622585@163.com

Abstract. Fixed-time and conventional actuated sig- nal control often perform poorly under unbalanced de- mand, leading to long queues, wasted green time, and spillback risk. This paper proposes an intelligent traffic-signal control framework that adapts phase selection and green splits using multi-source intersection data. The system combines signal sensors with vision-based perception (YOLOv11n) to estimate traffic states for vehicles and pedestrians, and can incorporate GPS data from connected terminals to improve observability under occlu- sion. All measurements are transmitted through a wireless network and fused for real-time decision making. A state- aware controller ranks phase priorities and computes adaptive timing plans that target lower average delay and higher throughput while avoiding inefficient all-red or idle-green intervals. We evaluate the approach in simulation on representative Shanghai intersection geometries under normal, warning, and emergency demand levels, and val- idate feasibility with an ESP32-based prototype. Results show reduced average vehicle delay and shorter queues compared with fixed-time baselines, indicating improved intersection efficiency and robustness.

Keywords: intelligent traffic signal control, multi-source data fusion, YOLO, adaptive signal timing

1. Introduction

With the sustained growth in automobile ownership, road vehicle density continues to increase, and urban traffic congestion has become increasingly severe, emerging as a key factor affecting urban operational efficiency and residents' quality of life. In this context, the limitations of traditional fixed-time traffic signal control have become more pronounced: its fixed timing plans cannot adapt to real-time variations in traffic flow, often resulting in long queues on arterial roads while green phases on minor approaches are underutilized, thereby reducing intersection throughput and exacerbating congestion.

Therefore, to address urban traffic congestion, intelli- gent traffic signal systems have emerged. Such systems acquire real-time data on traffic volume from each approach, queue length, vehicle speed, and vehicle position at intersections, perform online computation based on predefined optimization objectives, and determine a coordinated scheme for the priority phase and its associated

DOI: 10.54254/2755-2721/2026.AS32421

phases, generating an optimal phase sequence and signal timing strategy under the current traffic state. The system perceives traffic conditions through cameras deployed at intersections, collecting data such as approach traffic volume, vehicle speed, queue length, and the numbers of pedestrians and non-motorized vehicles. In addition, the system can integrate GPS positioning; by connecting intersection devices with in-vehicle terminals, it can obtain vehicle location and trajectory information. The above data are transmitted via a wireless communication network to a computing unit. Subsequently, intelligent algorithms such as machine learning and deep learning analyze the data, dynamically generating an optimal signal timing plan with objectives such as maximizing traffic efficiency and minimizing average delay. Finally, optimization commands are issued to the intersection signal controller, enabling adaptive and dynamic control of traffic signals and thereby significantly improving the operational efficiency of intersections and even the regional road network.

The traffic signal control system is an essential com- ponent of modern urban traffic management. Its core function is to guide vehicles and pedestrians to pass in an orderly manner through reasonable allocation of right- of-way, thereby ensuring traffic safety, improving traffic efficiency, and alleviating congestion. Traffic signal control technology has undergone multiple iterations. The earliest approach was fixed-time control, which originated from the classical theory proposed by Webster in 1958 [1]. This method designs a fixed timing plan based on historical traffic flow data and uses mathematical models to compute the optimal signal cycle length and green split in order to minimize average vehicle delay. Although it has the advantages of simple implementation and low cost, it cannot adapt to real-time changes in traffic flow; when traffic demand fluctuates significantly, it can easily lead to congestion and underutilized green time, resulting in low efficiency.

To overcome the limitations of fixed-time control, actu- ated control technology gradually emerged. The SCOOT (Split, Cycle, and Offset Optimization Technique) system proposed by Robertson and Bretherton in 1991 [2] is a representative technology of this stage. This system uses detectors embedded in the roadway (e.g., inductive loop detectors) to measure traffic flow parameters (vehicle counts, occupancy, etc.) in real time, and dynamically optimizes the green split, cycle length, and offsets of signal timing. Actuated control can respond dynamically to real- time traffic demand, and its efficiency is significantly higher than that of fixed-time control. However, this approach still suffers from limitations such as restricted detection coverage (only at detector installation points), susceptibility to environmental interference, and difficulty in identifying detailed information such as specific vehicle types and pedestrians; moreover, multi-sensor data fusion also faces major challenges.

With the development of artificial intelligence technolo- gies, traffic signal control has entered the stage of intelli- gent algorithm-based control. Recent review studies have shown that reinforcement-learning-based and adaptive traffic signal control methods have become major research directions in this field [3]. In 2011, Wang proposed a design method for a traffic signal controller based on a BP fuzzy neural network [4], which combines fuzzy logic with neural networks: fuzzy logic is effective in handling uncertainty and nonlinearity in traffic flow, while neural networks can learn and optimize fuzzy rules and mem- bership functions. This method outperforms traditional approaches in dealing with uncertainty, but it also has drawbacks such as high computational complexity and relatively poor real-time performance. In addition, the design of fuzzy rules and membership functions largely relies on expert knowledge and is difficult to optimize.

In recent years, with breakthroughs in computer vision, intelligent vision-based control has become a frontier direction in traffic signal control research. In 2019, Wang and Zhou proposed a traffic signal recognition method based on high dynamic range imaging and deep learning [5]. By leveraging object detection algorithms such as YOLO, SSD, and Faster R-CNN, cameras are used to detect in real time the positions, speeds, and counts of ve- hicles, pedestrians, and non-motorized traffic participants. Vision technology provides richer and more accurate input information for dynamic signal timing, can distinguish among different types of traffic participants, and offers high accuracy and strong scalability. Nevertheless, its performance is easily affected by adverse weather conditions (rain, snow, fog) and illumination variations (glare, nighttime), it imposes high computational requirements, and model training requires a large amount of high-quality annotated data.

To address the existing challenges in multi-sensor fusion as well as the environmental constraints of vision-based techniques and their reliance on large-scale annotated data, this study introduces targeted improvements. We train models on our own dataset, with a particular focus on the characteristics of traffic scenarios in Shanghai; and we innovatively introduce GPS positioning technology. By connecting GPS devices installed at intersections with invehicle terminals, we assist cameras in obtaining more accurate vehicle location and trajectory information under extreme weather conditions. The structure of this paper is organized as follows: Chapter 2 introduces the overall design framework of the proposed adaptive traffic signal control algorithm; Chapter 3 elaborates on the YOLO-based video traffic perception method and the traffic-state modeling process; Chapters 4 and 5 present the training results of the object detection model and the operational performance of the complete intelligent signal control system; Chapter 6 introduces the experiments and simulation demonstrations based on pygame; Chapter 7 describes the ESP32-based hardware implementation scheme and its experimental demonstrations; finally, Chapter 8 summarizes the entire paper and discusses future research directions.

2. Algorithm model description

This study proposes a video-perception-based adaptive traffic signal control method, aiming to dynamically adjust signal timing by leveraging real-time traffic state information to alleviate traffic congestion at urban intersections, with particular emphasis on optimizing the commonly observed "tailing" behavior or queue spillback phenomena at signalized intersections.

Overall, the proposed method consists of three core modules: (A) a deep-learning-based video traffic per- ception module; (B) a robust, second-level traffic state fusion module; and (C) a traffic-state quantification and representation module for signal control. In the design pro- cess, the method emphasizes robustness, interpretability, and engineering practicality, avoiding reliance on complex prediction models or black-box control algorithms.

2.1. Video-based traffic perception module

Traffic state information is extracted from monocular traffic video. This paper adopts a self-trained YOLO-series object detection model to perform real-time detection and tracking of traffic participants in the video. The model can recognize six categories of targets: Car, Bus, Truck, Motorcycle, Bicycle, and Pedestrian.

For each detected target i in a frame, the center point of its bounding box is defined in Equation (1):

DOI: 10.54254/2755-2721/2026.AS32421

$$\left(c_x(i,t), c_y(i,t)\right) = \left(\frac{x_1+x_2}{2}, \frac{y_1+y_2}{2}\right) \quad (1)$$

where (x_1, y_1, x_2, y_2) denote the coordinates of the upper-left and lower-right corners of the target bounding box.

To distinguish different traffic behaviors and demands, this paper partitions multiple regions of interest (Region of Interest, ROI) in the video plane, including:

1) Tail Zone, used to identify the risk of queue spillback upstream of the intersection;

2) Bicycle lane region, used to count vulnerable traffic participants such as bicycles and motorcycles;

3) Pedestrian waiting region, used to estimate the number of pedestrians waiting to cross;

4) Left-turn lane region, used to identify left-turn traffic demand.

This paper adopts the following inside/outside decision rule: if the target center point falls within the Tail Zone, the target is regarded as being outside the stop line (outside); otherwise, it is regarded as being inside the stop line (inside). This definition explicitly distinguishes normally queued vehicles from spillback queues that may cause network blockage.

2.2. Real-time speed estimation and outlier handling

For vehicle targets that are successfully tracked, the instantaneous speed is estimated based on the displacement of the center point between two consecutive frames, as shown in Equation (2):

$$v\left(i,t\right) = \frac{\sqrt{(c_x(i,t)-c_x(i,t-1))^2+(c_y(i,t)-c_y(i,t-1))^2}\cdot\alpha}{\Delta t} \quad (2)$$

where α denotes the conversion factor from pixels to realworld distance, and Δt denotes the time interval between frames.

Because the video detection and tracking process is inevitably affected by occlusion, detection jitter, or ID switches, the instantaneous speed may contain obviously unreasonable outliers. To this end, this paper designs a three-layer robust filtering mechanism:

1) Physical plausibility constraint: directly remove speed estimates that exceed a reasonable range for urban roads;

2) Speed-jump constraint: restrict unreasonable abrupt speed changes for the same vehicle within a short time;

3) Time-window outlier rejection: when only a small number of abnormal speeds occur within a short time window, treat them as detection noise and ignore them.

This procedure effectively suppresses speed outliers caused by detection errors while preserving the low-speed char- acteristics that genuinely reflect traffic congestion.

2.3. Second-level traffic state fusion

To reduce the impact of frame-level detection noise on signal control, this paper does not directly use frame- level results; instead, it aggregates traffic states into stable observation data at a resolution of one row per second. Let s denote a one-second time window; different variables adopt different fusion strategies:

• The number of vehicles inside the stop line is com- puted using the median, to reduce the impact of occasional missed detections and false detections;

DOI: 10.54254/2755-2721/2026.AS32421

• The number of vehicles in the Tail Zone is computed using the 90% quantile, to enhance sensitivity to congestion risk;

• The number of waiting pedestrians and the left-turn demand are computed using the maximum, to avoid underestimating peak demand;

• Vehicle speed, after outlier filtering, is represented by the mean as the representative speed for that second.

Through the above robust fusion strategy, the system can significantly improve the stability and reliability of traffic-state inputs while maintaining real-time performance.

3. Adaptive signal control strategy

3.1. Score-based traffic state quantification

To provide a unified representation of the impacts of different vehicle types on intersection capacity, this paper conducts score-based modeling only for three categories of motor vehicles: cars, buses, and trucks. Vehicles inside the stop line and those within the Tail Zone are assigned different weights to reflect their different levels of influence on intersection operations, as defined in Equations (3) and (4):

$$W_{\text{in}}(s) = \sum_k w_k^{\text{in}} N_k^{\text{in}}(s), \tag{3}$$

$$W_{tail}(s) = \sum w_k^{tail} \cdot N_k^{tail}(s) \tag{4}$$

where $N_k^{\text{in}}(s)$ and $N_k^{tail}(s)$ denote the numbers of classk vehicles inside the stop line and within the Tail Zone, respectively, during time window s, and w_k^{in} and w_k^{tail} are the corresponding vehicle-class weights.

To further reflect traffic flow operational efficiency, a speed-related coefficient is introduced to modulate the score of vehicle counts. A lower average vehicle speed indicates a higher degree of impeded traffic operation and a larger control pressure on signal timing. Accordingly, this paper defines the following two core traffic-state indicators in Equations (5) and (6):

$$S_{through}(s) = K_{\text{in}}(s) \cdot W_{\text{in}}(s), \tag{5}$$

$$S_{tail}(s) = K_{tail}(s) \cdot W_{tail}(s) \tag{6}$$

where $S_{through}(s)$ denotes the through-movement traffic pressure indicator inside the stop line, and $S_{tail}(s)$ denotes the spillback risk indicator; $K_{\text{in}}(s)$ and $K_{tail}(s)$ are modulation coefficients associated with average vehicle speed, used to amplify traffic pressure under low-speed operating conditions.

3.2. Spillback-priority state machine design

The spillback risk indicator $S_{tail}(s)$ is mapped into three discrete operating states: normal, warning, and emergency. This state partitioning is used to characterize the potential threat posed by upstream queues to overall traffic operations.

When the system enters the emergency state, the signal control strategy prioritizes clearing queued vehicles in the Tail Zone to prevent congestion from propagating further upstream and

DOI: 10.54254/2755-2721/2026.AS32421

triggering network-level blockage.

To avoid frequent phase switching caused by short- term traffic fluctuations, this paper introduces a hysteresis mechanism during state transitions, thereby improving the stability and executability of the control strategy in practical operations.

3.3. Signal phase structure and duration allocation

This paper adopts a signal control structure consisting of two basic phases:

1) Protected left-turn phase: only left-turn vehicles are permitted to pass, while the pedestrian signal remains red;

2) Concurrent phase: through vehicles and pedestrians are released simultaneously.

The duration of the concurrent phase is jointly determined by through-vehicle demand and pedestrian crossing demand, as defined in Equation (7):

$$T_C(s) = \max(T_{C,veh}(S_{through}(s)), T_{C,ped}(N_{ped}(s))) \tag{7}$$

where $T_{C,veh}(\cdot)$ denotes the minimum service time requirement based on through-movement traffic pressure, and $T_{C,ped}(\cdot)$ denotes the minimum green time required to ensure safe pedestrian crossing.

This design ensures that, under the premise of satisfying pedestrian safety, through traffic will not be persistently constrained due to a single demand factor. The left-turn phase duration is determined by left-turn demand, but it must not encroach upon the minimum service time required by the concurrent phase. Under the spillback emergency state, the left-turn phase can be shortened or temporarily cancelled to prioritize restoring throughmovement capacity and accelerating congestion dissipation.

The specific algorithm implementation and parameter settings are provided in the traffic signal optimization algorithm code in appendix

4. Machine learning

YOLO (You Only Look Once) is a classical one-stage object detection algorithm whose core idea is to simultane- ously perform object localization and category prediction within a single forward pass. Compared with traditional two-stage object detection methods, YOLO achieves faster inference speed while maintaining relatively high detection accuracy, and is therefore widely used in real-time object detection tasks.

This study selects YOLOv11n as the object detection model. YOLOv11n is a lightweight model in the YOLOv11 family; while significantly reducing the number of param- eters and computational complexity, it can still maintain relatively stable detection performance, making it suitable for practical application scenarios with limited resources or strict real-time requirements. This characteristic is highly consistent with the system requirements of real-time traffic signal control in this paper.

The dataset constructed in this study contains approxi- mately 6,650 images, with 6,650 corresponding annotation files. All images are manually annotated in the YOLO annotation format, and each annotation file contains the target class index and the corresponding normalized bounding-box coordinates.

To analyze the impact of dataset size on the generaliza- tion ability of the model, this paper conducts preliminary experiments using only a small number of samples at the early stage of training. Due to the limited number of train- ing samples, the model exhibits pronounced overfitting during training, and both detection accuracy and gener- alization performance remain low. To alleviate this issue, more traffic-scene-related images are gradually introduced from public datasets, and the dataset configuration file (YAML file) is correspondingly expanded and adjusted.

As the dataset scale continues to expand, the model achieves significant improvements in both detection accuracy and stability. When the dataset size increases to approximately more than 6,600 images, the overall detection performance on the validation set tends to stabilize, enabling the model to reliably perform multiclass traffic object detection.

Model training is completed based on the PyTorch deep learning framework and is conducted in a CUDA-enabled GPU environment. The main training parameter settings are as follows:

- Model architecture: YOLOv11n;
- Number of epochs: 100;
- Batch size: 32;
- Input image size: 640 × 640;
- Optimizer: automatically selected (optimizer = auto);
- Initial learning rate (lr0): 0.0001;
- Final learning rate ratio (lrf): 0.1;
- Dropout: 0.1;
- Random seed (Seed): 0;
- Early stopping patience (Patience): 50.

The above parameter settings ensure stable training while effectively improving the convergence speed of the model. In experiments, they achieve satisfactory detection performance and provide a reliable data foundation for subsequent traffic-state perception and signal control.

5. AI training results

After completing 100 training epochs, the model performance was systematically evaluated on the validation set. The experimental results show that the trained YOLOv11n model exhibits relatively stable detection performance for multi-class traffic object detection. The model achieves a Precision of 0.759 and a Recall of 0.734 on the validation set, indicating that it is able to detect most true targets while effectively controlling the false positive rate.

In terms of overall detection accuracy, the model reaches 0.774 on the mAP@0.5 metric, and achieves 0.600 under the more stringent mAP@0.5:0.95 metric. These results indicate that the model maintains relatively consistent detection capability under different IoU thresholds and demonstrates a certain degree of generalization perfor- mance.

During training, the classification loss (Classification Loss), bounding-box regression loss (Box Loss), and distri- bution focal loss (DFL Loss) were continuously recorded. The corresponding training-loss curves and performance- metric curves are shown in Figures 1. The results show that, as the number of training epochs increases, all loss functions exhibit an overall decreasing trend and gradually stabilize, indicating that the model achieves good convergence under the current parameter settings.

DOI: 10.54254/2755-2721/2026.AS32421

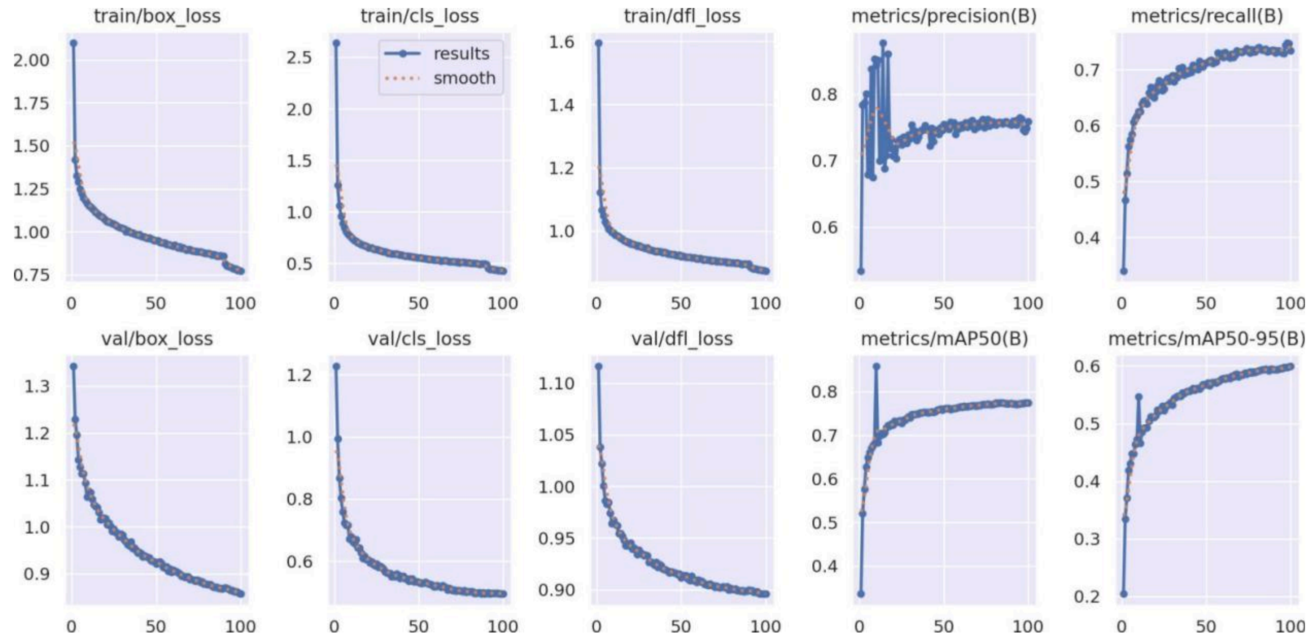

Figure 1. Curves of major loss functions and performance metrics during YOLOv11n model training, including classification loss, bounding-box regression loss, and related detection metrics, used to illustrate the convergence behavior of the model during training

Further inspection of the validation losses in the later stage of training reveals no obvious oscillation or rebound, suggesting that under the current dataset scale and training strategy, the model does not suffer from severe overfitting, and the overall training process remains stable. To further analyze the detection performance under different confidence thresholds, this paper conducts sta- tistical analyses of the Precision–Recall (PR) curve and the F1 curve, with the corresponding results shown in Figures 2 and 3. The PR curve indicates that the model maintains high precision across different recall levels; meanwhile, the F1 curve reaches its peak in the medium-to-high confidence range, suggesting that the model achieves a relatively reasonable balance between precision and recall.

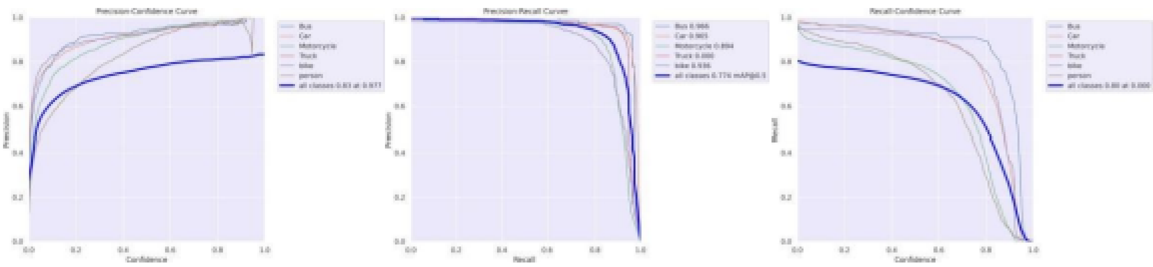

Figure 2. Precision, recall, and precision–recall curves of the YOLOv11n model on the validation set, used to analyze performance under different recall levels

DOI: 10.54254/2755-2721/2026.AS32421

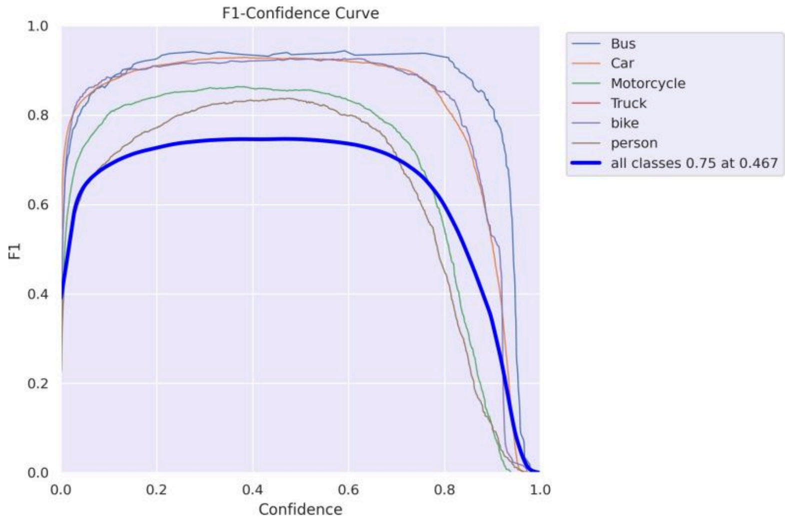

Figure 3. F1 curve of the YOLOv11n model under different confidence thresholds, used to evaluate the trade-off between precision and recall

In terms of class discrimination capability, Figure 4 presents the confusion matrix on the validation set and its normalized results. The confusion-matrix analysis shows that most traffic targets can be correctly classified, and errors are mainly concentrated between categories with similar appearance features, as well as a small number of missed detections. The normalized results further indicate that the overall classification accuracy is high, and the proportion of erroneous predictions between categories is relatively low.

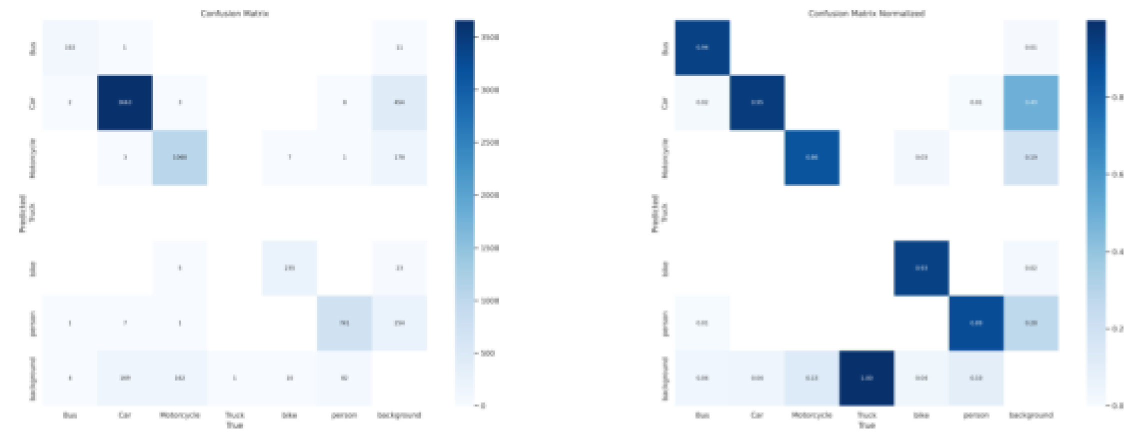

Figure 4. Confusion matrix results of the YOLOv11n model on the validation set, including the raw and normalized confusion matrices, used to present classification outcomes and relative error distribution among traffic target categories

To intuitively verify the detection performance in real traffic scenarios, Figures 5, 6, 7, 8, and 9 also present example detection results on validation-set images. It can be observed that the model can localize traffic targets with relatively high accuracy and generate reasonable bounding boxes, and the predictions are highly consistent with manual annotations, qualitatively validat- ing the applicability of the model to real traffic videos.

DOI: 10.54254/2755-2721/2026.AS32421

Figure 5. Examples of YOLOv11n detection results on validation images, illustrating predicted bounding boxes and categories for traffic targets such as vehicles and pedestrians.

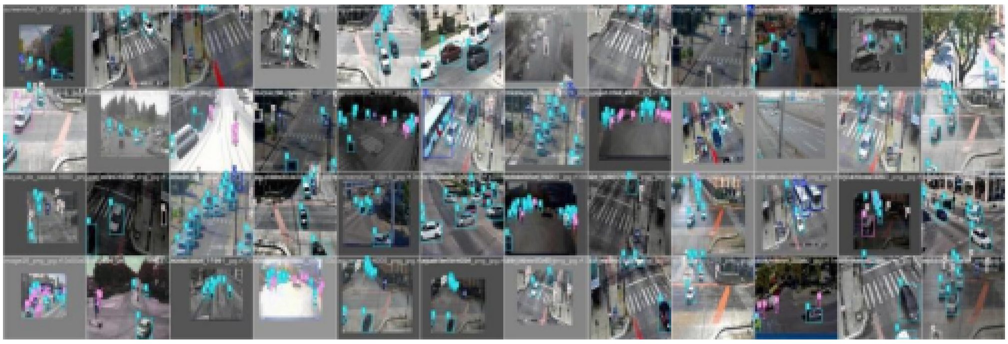

Figure 6. Additional examples of YOLOv11n detection results on validation images

Figure 7. Predicted detections and ground-truth labels for a validation batch sample

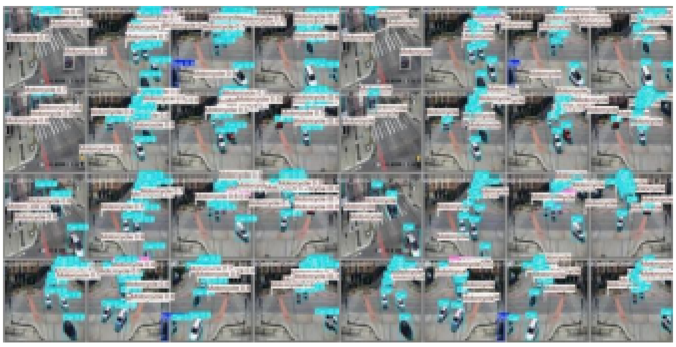

Figure 8. Predicted detections and ground-truth labels for another validation batch sample

DOI: 10.54254/2755-2721/2026.AS32421

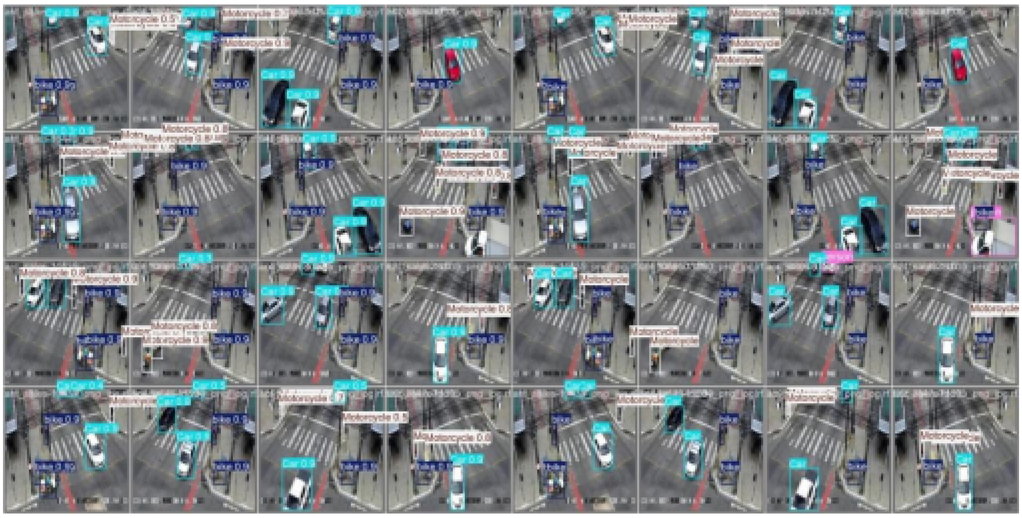

Figure 9. Predicted detections and ground-truth labels for an additional validation batch sample

In summary, after expanding the dataset scale and optimizing parameters, the trained YOLOv11n model demonstrates reliable detection performance in both quan- titative metrics and qualitative results, providing a stable data-input foundation for subsequent video-based traffic- state modeling and the adaptive signal control strategy.

6. Shanghai real-world traffic simulation

To systematically verify the applicability of the pro- posed adaptive traffic signal control method under differ- ent road structures and different traffic operating condi- tions, this paper constructs and simulates multiple types of intersection traffic scenarios based on typical urban road morphologies in Shanghai using Pygame. Three representative intersection types are selected, denoted as A, B, and C. For each intersection type, three traffic operating states—normal, warning, and emergency—are constructed, thereby forming nine typical traffic scenarios in total. The simulation schematic results of each intersec- tion type under different operating states are uniformly presented in the appendix.

Without relying on real traffic detection data, this modeling approach can cover the primary operating pat- terns of different road structures under different traffic load conditions, providing a unified experimental basis for subsequent logical analysis and illustrative verification of the adaptive signal control strategy under multiple operating conditions

6.1. Intersection Type A: typical four-leg inter- section: intersection Type A corresponds to a typical four-leg intersection scenario

For this type, this paper constructs three traffic operating states—normal, warning, and emergency—to simulate the process in which traffic operations evolve gradually from a smooth condition to the emergence of spillback risk and eventually to pronounced congestion. The simulation results of the three states are shown in Figure 10.

DOI: 10.54254/2755-2721/2026.AS32421

Figure 10. Simulation scenarios of a typical Shanghai four-leg intersection (Type A) under normal, warning, and emergency traffic states, illustrating changes in vehicle distribution under different traffic load levels

A-1 Normal State (Normal): Under the normal state, the number of vehicles inside the stop line is at a moderate level, and vehicles mainly queue inside the stop line while waiting for signal release. No obvious vehicle aggregation is observed outside the stop line, the average vehicle speed is high, and overall traffic operation is relatively smooth.

A-2 Warning State (Warning): Under the warning state, a small number of vehicles begin to aggregate outside the stop line, and vehicle speed decreases, but persistent queue spillback has not yet formed. This state is used to characterize a transitional operating condition in which traffic load increases but still remains controllable.

A-3 Emergency State (Emergency): Under the emer- gency state, the number of vehicles outside the stop line increases significantly, and the vehicle speed drops markedly, forming a typical spillback phenomenon. With- out timely signal intervention, congestion may propagate to upstream road segments.

6.2. Intersection Type B: five-leg intersection: in- tersection Type B corresponds to a five-leg intersection scenario

For this type, this paper also constructs three traffic operating states—normal, warning, and emergency—to simulate changes in vehicle distribution and operating con- ditions under different traffic load levels. The geometric structure of this intersection type references a real five- leg intersection in Shanghai (the five-leg intersection is shown in Figure 11), and both the corre- sponding real-road schematic and the simulation results are uniformly presented in Figure 12.

Figure 11. A special five-leg intersection in Shanghai

DOI: 10.54254/2755-2721/2026.AS32421

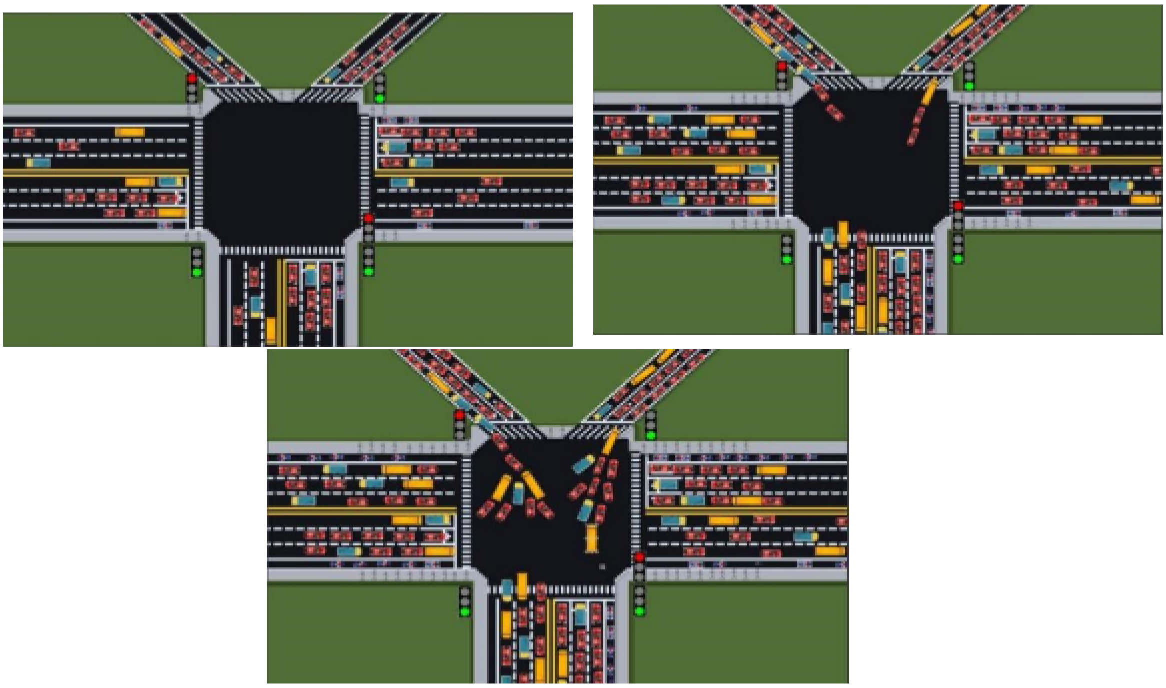

Figure 12. Simulation scenarios of a Shanghai five-leg intersection (Type B) under normal, warning, and emergency traffic states

2-1 Normal State (Normal): Under the normal state, vehicles inside the stop line remain within a controllable range, and there is no obvious aggregation of low-speed vehicles outside the stop line; overall traffic operational efficiency is high.

2-2 Warning State (Warning): Under the warning state, short-term vehicle aggregation begins to appear outside the stop line, and the average speed decreases, indicating that spillback risk is emerging but has not yet developed into persistent queue spillback.

2-3 Emergency State (Emergency): Under the emer- gency state, a large number of vehicles aggregate outside the stop line and the operating speed decreases signifi- cantly, forming an obvious risk of queue spillback, which is used to simulate extreme operating conditions of a five- leg intersection under high traffic volumes.

6.3. Intersection Type C: roundabout: intersection Type C corresponds to a roundabout scenario

For this type, this paper constructs three traffic operating states— normal, warning, and emergency —to characterize changes in roundabout operating features under different traffic load conditions. The geometric structure references a real roundabout in Shanghai (the real roundabout is shown in Figure 13), and the corresponding real-road schematic and simulation results are uniformly presented in Figure 14.

DOI: 10.54254/2755-2721/2026.AS32421

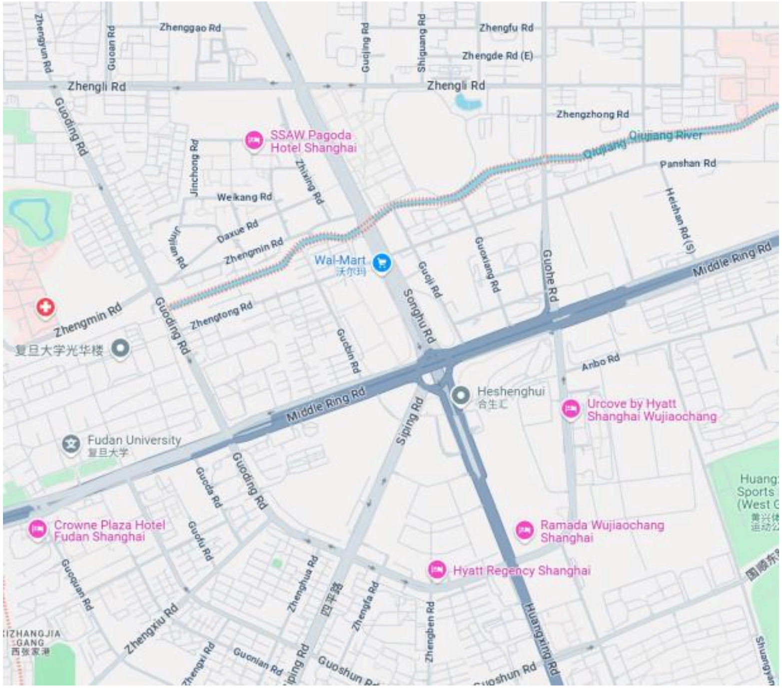

Figure 13. A special roundabout intersection in Shanghai

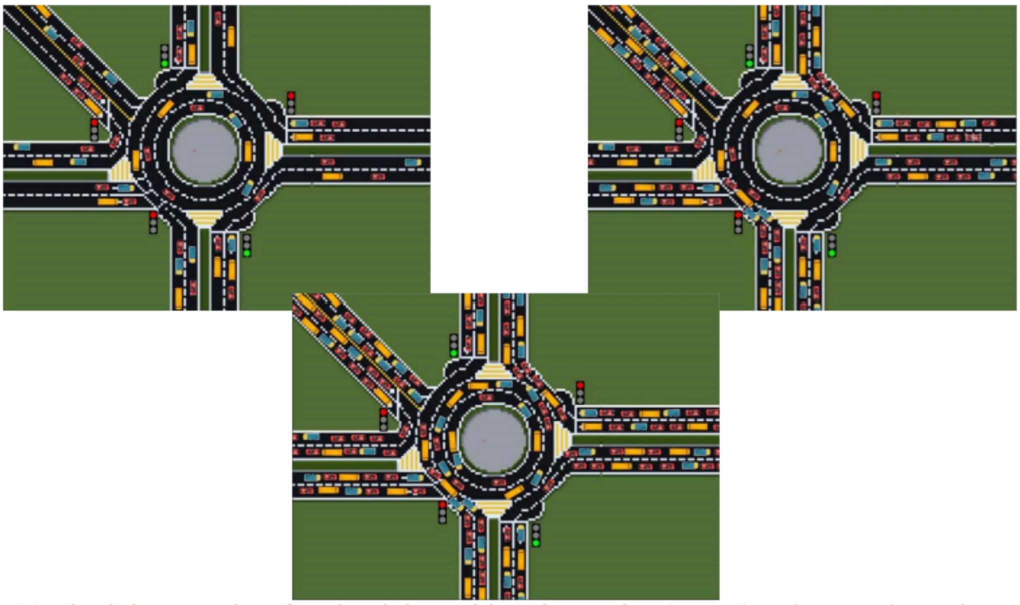

Figure 14. Simulation scenarios of a Shanghai roundabout intersection (Type C) under normal, warning, and emergency traffic states

3-1 Normal State (Normal): Under the normal state, vehicle speed is relatively high, and the distributions of vehicles inside and outside the stop line are relatively clear, with no obvious queue spillback phenomenon.

3-2 Warning State (Warning): Under the warning state, vehicles inside the stop line continue to accumulate, and low-speed vehicle aggregation begins to appear out- side the stop line. Traffic operational efficiency decreases significantly, and the system enters a high-risk operating range.

DOI: 10.54254/2755-2721/2026.AS32421

3-3 Emergency State (Emergency): Under the emer- gency state, vehicles inside and outside the stop line are highly dense, and vehicle speed decreases significantly or approaches standstill, resulting in severe spillback.

By systematically modeling three representative types of Shanghai urban intersections—typical four-leg intersections, five-leg intersections, and roundabouts—under normal, warning, and emergency operating states, this paper establishes a multi-state traffic scenario frame- work that covers different road structures and traffic saturation levels. This framework provides a clear and unified experimental basis for subsequent logical analysis and illustrative verification of the adaptive signal control strategy under different operating conditions.

7. Hardware

To emulate real traffic operating scenarios, this paper builds a traffic signal control simulation system. The system uses the Espressif ESP32 as the core control unit. ESP32 is a highly integrated microcontroller, whose advantages are mainly reflected in its high level of wireless communication integration, prominent low-power characteristics, and progressively enhanced edge AI computing capability. Compared with earlier-generation Raspberry Pi models, ESP32 can enable more professional and efficient system deployment at lower cost in Internet of Things (IoT), embedded control, and lightweight terminal intelligence application scenarios.

In this system, the ESP32 serves as the main control unit, analyzes and processes the data collected by ultra- sonic sensors, and maps the traffic optimization decisions to an LED traffic light module. Due to limitations in experimental materials, this paper uses ultrasonic sensors as a substitute for the GPS positioning system to detect the spatial occupancy of vehicles at the intersection. The specific hardware wiring connections of the system are shown in Figure 15.

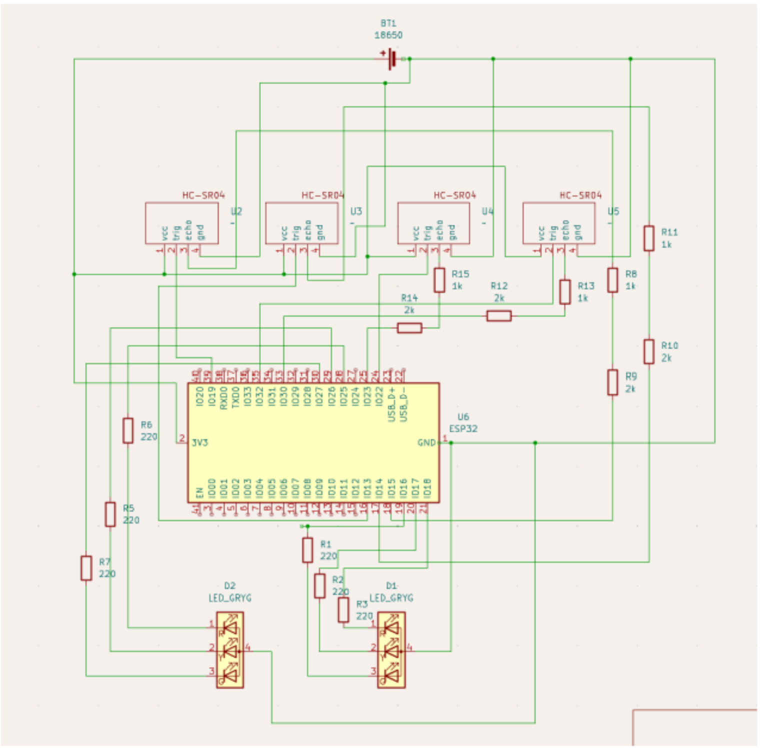

Figure 15. Overall schematic of the ESP32-based traffic signal control simulation system, including the control unit, sensor modules, and LED traffic light display module

This paper further demonstrates the system perfor- mance under different traffic states. Because the front and rear traffic lights are started simultaneously during the experiments, while the lane controlled by the rear traffic light exhibits no change in traffic flow and thus maintains an unchanged signal timing, the timing adjustment of the front traffic light can be judged intuitively by comparing the state changes of the front and rear traffic lights. The overall system prototype and the ESP32 physical board are shown in Figure 16.

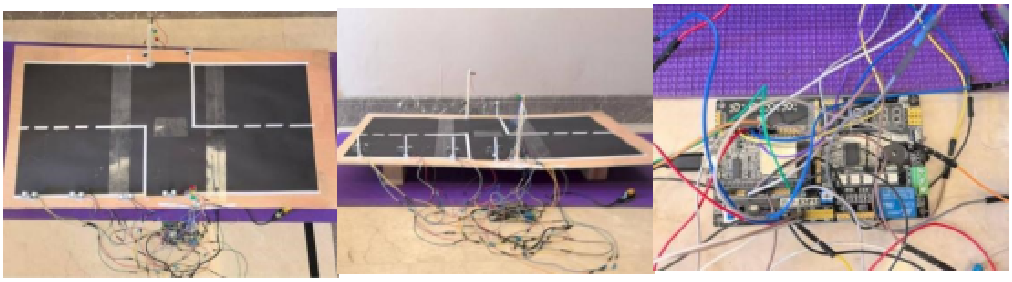

Figure 16. System overview: simulated opposing traffic flow and traffic scene (left), side view of the model (middle), and ESP32 circuit details (right)

In Case 1, when the road traffic flow is determined to be sparse, the signal durations are set to 20 s for red and 10 s for green; the corresponding red and green states are shown in Figures 17 and 18, respectively. In Case 2, when the road traffic flow is determined to be congested, the signal durations are adjusted to 10 s for red and 20 s for green; the corresponding red and green states are shown in Figure 19. At this time, a clear timing difference between the front and rear traffic lights can be observed, i.e., when the front red is on the rear green is on, or when the front green is on the rear red is on.

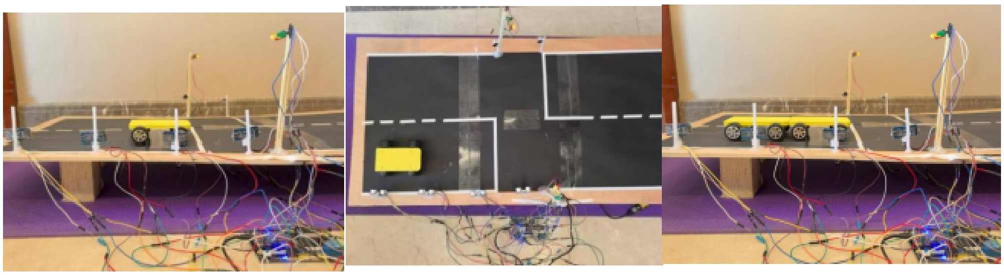

Figure 17. Case 1 (red): traffic signal operation under sparse traffic flow, where the timing of the two traffic lights remains consistent

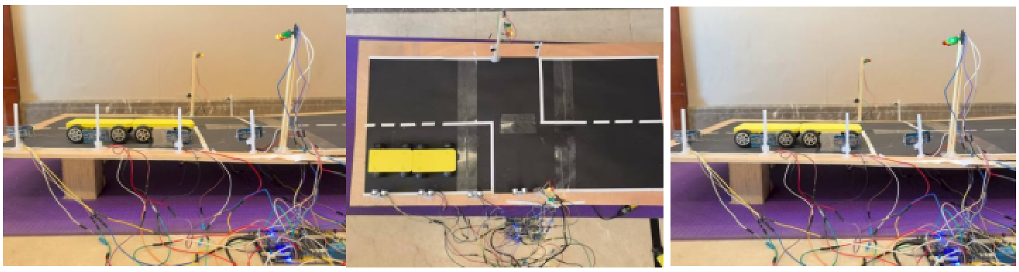

Figure 18. Case 1 (green): traffic signal operation under sparse traffic flow, where the timing of the two traffic lights remains consistent

DOI: 10.54254/2755-2721/2026.AS32421

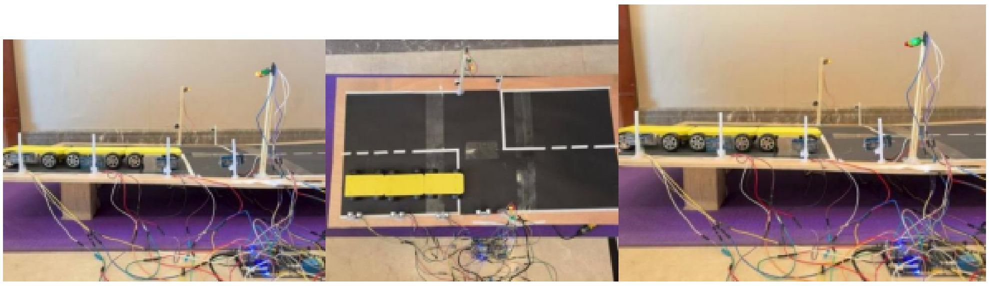

Figure 19. Case 2: traffic signal operation under congested traffic flow, where signal timing changes and the two sides no longer turn green simultaneously

In Case 3, when a vehicle is detected in the middle of the roadway, i.e., a spillback phenomenon occurs (due to congestion at the downstream intersection, vehicles from the upstream intersection cannot completely pass through and must remain in the middle of the intersection), to ensure normal passage for vehicles in the other direction, the system keeps the traffic signal continuously in the red state. The corresponding effect is shown in Figure 20, until the spillback vehicle disappears. The simulation implementation of the entire system can also be obtained via the GitHub link. Under conditions where experimental materials and cost allow, different types of traffic signal optimization algorithms can be deployed within this system framework, and the simulation system built in this paper can provide a reference for subsequent related studies.

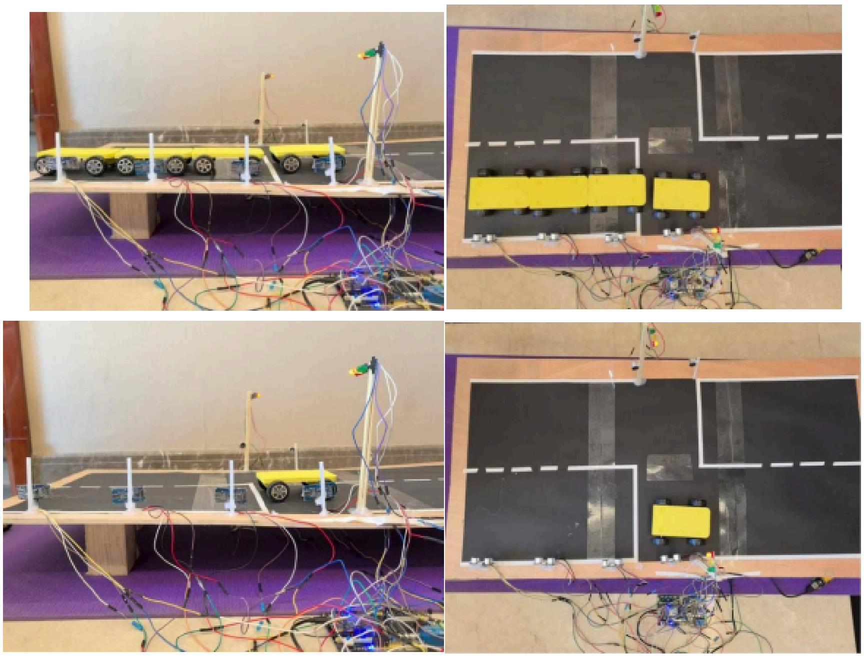

Figure 20. Case 3: traffic signal operation under spillback, where the signal remains red until vehicles clear even if no vehicles are present behind the stop line

To demonstrate the system operating logic under limited experimental conditions, this paper simplifies the traffic signal duration optimization algorithm. In the simulation system, when the detected number of vehicles is greater than 3, the system determines it as a traffic congestion state; when the number of vehicles is less than 3, it de- termines it as a smooth-traffic state. When the ultrasonic sensor detects a vehicle in the middle of the intersection, the system will always maintain the red-light state. The corresponding program code is uniformly presented in appendix in the ESP32 physical simulation system code.

8. Conclusion

This paper proposes a video-perception-based adaptive traffic signal control method, with a particular focus on modeling and optimizing the often-overlooked spillback phenomenon at signalized intersections. The method ex- tracts fundamental traffic-state information such as vehi- cles, pedestrians, and their speeds from real-time traffic video, and adopts a robust second-level traffic state fusion strategy to construct a signal control framework with a clear structure and interpretable logic.

Unlike traditional traffic signal control methods that rely on fixed timing plans or complex prediction models, the proposed control strategy is built upon an explicit score-based traffic-state mechanism and priority rules. At the design level, by distinguishing vehicle states inside and outside the stop line and explicitly modeling low-speed operating conditions, the system is ensured to maintain a basic level of service under different traffic load conditions. When spillback risk is detected, the system increases the control priority of through movements via a state-machine mechanism and constrains the left-turn phase, thereby preventing congestion from further propagating upstream in terms of control logic.

For method validation, this paper constructs multiple intersection models based on typical Shanghai urban road morphologies and conducts systematic scenario-based simulation analyses under different traffic operating states. Meanwhile, by building a simplified hardware simulation system, the executability of the signal control logic in an actual control flow is verified. The experimental and illustrative results indicate that the proposed method exhibits good integrity and implementability in terms of structural design and control logic.

It should be noted that, due to limitations in ex- perimental conditions and materials, this paper has not yet conducted quantitative performance evaluation of the method using real intersections or mature traffic simulation platforms. Therefore, the focus of this paper is to propose an interpretable signal-control modeling approach for spillback phenomena and to verify its logical rationality through multi-scenario examples, rather than to quantitatively analyze the magnitude of performance improvement.

Future work will focus on systematic validation of this method in standard traffic simulation environments or real-road scenarios, including further analysis of metrics such as vehicle delay, queue length variation, and levels of service for different traffic participants, so as to more comprehensively evaluate its practical application effects in complex urban traffic environments.

Acknowledgements

During the research and writing of this paper, guidance and support were provided from multiple sources, and I hereby express my sincere gratitude.

First, I would like to express my heartfelt thanks to my advisor for the meticulous guidance provided in research direction selection, method design, system implementation, and manuscript revision. The advisor's rigorous academic attitude, systematic research thinking, and patient and detailed guidance played a crucial role in the successful completion of this work.

In addition, I would like to thank my family for their understanding, support, and encouragement throughout the research process, which enabled me to remain focused during research and writing

DOI: 10.54254/2755-2721/2026.AS32421

and to successfully complete this study.

References

[1] Webster, F. V. (1958). Traffic Signal Settings. Road Research Laboratory, UK..
[2] Robertson, D. I. Bretherton R. D. (1991). Optimizing networks of traffic signals in real time-The SCOOT method. IEEE Transactions on Vehicular Technology, 40(1), 11-15.
[3] Zhao, H., et al. (2024). A survey on deep reinforcement learning approaches for traffic signal control. Engineering Applications of Artificial Intelligence.
[4] Wang, S. C. (2011). Design of traffic signal controller based on BP fuzzy neural network. Journal of Yunnan University for Nationalities(Natural Science Edition), 20(6), 511-514.
[5] Wang, J. Zhou L. (2019). Traffic light recognition with high dynamic range imaging and deep learning. IEEE Transactions on Intelligent Transportation Systems, 20(4), 1341-1352.

Appendix

This appendix provides the implementation of the core program code of the proposed system, including the processing workflow of traffic-state data and an example implementation of the signal control logic. The code is provided to demonstrate the feasibility of the proposed method at the engineering implementation level and to facilitate understanding of the descriptions of system architecture and control workflow in the main text.