International Journal of Computational and Experimental Science and ENgineering (IJCESEN)

Vol. 11-No.3 (2025) pp. 4809-4825
http://www.ijcesen.com

ISSN: 2149-9144

Research Article

An Intelligent Vehicle Detection and Recognition Framework for Traffic Cyber-Physical Systems

XiulingWei [1,2], Marloun P. Sejera[3*]

[1] School of Electrical, Electronics, and Computer Engineering, Mapua University Manila, Philippines
[2]Xi'an Siyuan University, Xi'an, Shaanxi 710038 China.
Email: xiulingwei@mymail.mapua.edu.ph **- ORCID:** 0009-0005-1926-7834

[3]Department of Mathematics, Bannari Amman Institute Of Technology, Sathyamangalam, Erode, Tamilnadu, Philippines
* **Corresponding Email:** mpsejera@mapua.edu.ph - **ORCID**: 0009-0001-8147-0156

Article Info:

DOI: 10.22399/ijcesen.3180
Received : 11 May 2025
Accepted : 03 July 2025

Keywords

Vehicle Detection
Vehicle Recognition
Traffic Monitoring
Cyber-Physical Systems (CPS)
Intelligent Transportation Systems (ITS)

Abstract:

The rapid growth of urban traffic has necessitated the integration of intelligent systems within Cyber-Physical Systems (CPS) for real-time monitoring and management. This paper presents an efficient approach to vehicle detection and recognition tailored for Traffic Cyber-Physical Systems (CPS), aimed at enhancing situational awareness, safety, and traffic flow optimization. Leveraging a combination of computer vision techniques and deep learning models, the proposed framework accurately detects vehicles in diverse environmental conditions and recognizes their types and license plates. The system incorporates real-time image acquisition from roadside cameras, pre-processing pipelines, and a convolutional neural network (CNN)-based architecture for robust vehicle classification and identification. Experimental results demonstrate high accuracy and reliability across various traffic scenarios, including low-light and high-occlusion conditions. The proposed model is scalable, adaptable to edge-based deployment, and contributes significantly to the development of responsive and intelligent traffic CPS infrastructure. This research lays the foundation for more comprehensive intelligent transportation systems, facilitating autonomous decision-making and efficient traffic control.

1. Introduction

In recent years, with the rapid development of China's road traffic, the urban road network has constantly improved, and traffic infrastructure construction has made great progress. At the same time, with the continuous improvement of people's living standards, cars have gradually become a necessity for some families. As of September 2020, the number of motor vehicles in China had reached 365 million, including 275 million vehicles. The rapid growth of the number of motor vehicles makes the urban road network difficult to deal with; traffic jams are more and more serious, and traffic violations have become frequent [1]. According to statistics, China's annual economic losses caused by traffic congestion reached as high as 250 billion yuan. Traffic congestion has become a constraint on urban development “Stubborn disease”. Traffic congestion has a direct impact on people’s lives - it extends travel time, increases travel costs, and affects the quality of life and work efficiency. A large number of congested vehicles driving at low speeds not only causes extra energy consumption but also increases exhaust emissions and aggravates urban pollution. Congestion increases the rate of road traffic accidents, especially the secondary accident rate on trunk roads such as expressways. The Traffic System is a sub-system of the social economy; congestion will also affect the healthy development of the whole national economy indirectly through the coupling effect.

In 2006, the National Academy of Sciences issued the American Competitiveness Plan, based on a congressional assessment of America's global technological competitiveness and the urgent need to maintain and improve it. Cyber-Physical System (CPS) is listed as an important research project in program [2]. In 2007, the Presidential Advisory Council on Science and Technology (PCAST) made CPS a very important proposal in the field of network technology and information technology. In

2008, the CPS Steering Group was established in the United States, which is an important research area of CPS application technology in agriculture, defense, transportation, energy, medical, large-scale construction facilities, and so on. In addition, the United States and Europe, in large-scale research funding programs, have invested a large number of research funds in related theoretical and applied technology research. In 2008, the National Science Foundation (NSF) and other organizations co-organized a seminar on Transportation System CPS, which made the CPS in the field of transportation widely concerned by academia and industry. Many scholars have carried out a lot of CPS in the application and development of the discussion. With the gradual development of research work in the field of CPS, many scholars begin to pay attention to the application of CPS in the field of transportation. R. Cartwright and A. Chen thinks that the modern transportation system is a typical CPS system, and the technical bottleneck in the development of the transportation system is mainly embodied in the cost, reliability and re-use. R. Sengupta and YFALLAHL argue that the current road traffic control system is a closed system, not a network-based system, and that the future system needs more open control methods. With the development of T-CPS Technology, T-CPS technology can realize the integration and feedback between traffic information system and traffic physical system, can deepen the study of traffic and vehicle movement law. It is expected that T-CPS will become an important direction for the next generation intelligent transportation system.

In the field of transportation, from the perspective of T-CPS, physical entities and information systems have been more deeply integrated, and the interaction between information systems and physical systems has become closer, which is crucial for effective traffic monitoring of the transportation system.

2. Literature Review

2.1 Intelligent Transportation System

2.1.1 Intelligent Transportation System Development

With the development of modern science and technology, the intelligent transportation system (ITS), which makes use of new automatic control technology, information technology and computing technology, is a revolution to the traditional transportation system. Intelligent transportation system (ITS) means the integration of advanced information technology, data communication and transmission technology, sensing technology, control technology and computing technology on top of better infrastructure, and effectively applied to the entire transportation management system, so as to establish a large-scale, all-round play a role in real-time, accurate and efficient integrated transportation and management system. Its makes the interaction relationship among the three main bodies “People, cars and roads” appear in a new way, using modern scientific and technological means to deal with individual traffic behavior, thus improve the level of transport services and service quality, make modern traffic more safe, efficient and reliable. After more than ten years of promotion, trial and development, ITS has been implemented in some cities and expressway systems in developed countries. It has been proved that its development has become an inevitable choice to solve the contradiction between the increasing traffic demand and the lagging development of traffic service.

Intelligent Transportation System (ITS) is an effective means to solve the imbalance between supply and demand of road traffic, and it is the technical support to realize the goal of green transportation system [3]. Through the application of the intelligent transportation system, the transportation infrastructure will be fully utilized, the level of traffic safety will be greatly improved, and the goal of traffic environmental protection and energy conservation will be better realized, the service level of transportation system will be improved continuously [4]. Information is the basis of intelligent transportation system (ITS) . The accuracy, real-time, comprehensiveness and precision of information acquisition determine the depth and breadth of its development.

The acquisition of traffic state information covers the collection of traffic flow data, the generation and transmission of traffic information, and the storage and processing of traffic information. The effect and quality of traffic state information acquisition will affect the validity of traffic control and traffic guidance strategy, the reliability of traffic information release and service, and the rationality of traffic accident cause analysis. How to obtain traffic information effectively becomes the key problem to determine the development quality and application effect of it. Against the backdrop of rapid advances in computing, communications and control, Cyber Physical Systems (CPS) combine monomer with Physical elements, based on the interaction and feedback between the information system and the physical system, the precise cognition and effective control of the physical system can be realized to realize the closer integration and cooperation between the traffic information system and the traffic physical system, therefore, the coordination and intelligentization of the transportation system can be achieved[5]. The application of information physics fusion system in the field of transportation promotes the intelligent transportation system to a new development stage. The new generation of intelligent transportation system (ITS) is clearly layered, and the traffic information perception system at the bottom of the system is the foundation of the whole its, it requires the following features: it can adapt to many application service requirements, it is convenient and reliable to deploy on a large scale, and multi-sensor units can form an organization network for information transmission and processing.

2.1.2 Research Progress and Application of CPS

Information Physics fusion systems integrate Computation, Control and Communication (3C) technologies, fusing information systems with the physical world, by making computing and physical processes interact closely and repeatedly, finally, it realizes the accurate, safe and reliable cognition and control of the physical world. CPS has a very wide range of extensions, including embedded systems, wireless sensor networks, information fusion, machine learning and many other fields [6]. Since it was first proposed in 2005, CPS has become a hot research direction, which has attracted the attention of many research institutions and scholars and achieved a lot of results [7].

CPS is a new research field, which emphasizes the deep fusion of the information world and the physical world, and regards computing, communication and control as equal status. In the United States, the American Competitiveness Plan issued in February 2006 lists CPS as an important research project, in July 2007, the U.S. Presidential Council of Advisors on science and technology (PCAST) listed eight key information technologies in a report entitled “Leadership in A Challenge: Information Technology R & D in a competitive world.” CPS ranked first [8]. At the same time, CPS research has been supported by NSF, DOD/DARPA, DOE, NASA, HSARPA and other agencies in the United States, and many international conferences and seminars related to CPS have been held, the basic theory of CPS, the application of CPS, the performance of CPS and the safety of CPS are discussed, CPS technology is also widely used in intelligent transportation, smart grid, aerospace, military exercises, disaster early warning, logistics and supply chain optimization and other fields. The Massachusetts Institute of Technology (MIT) has designed a distributed intelligent robot garden based on mobile robots[9], the autonomous modeling and control of CPS for Dynamic Environment Awareness, multi-node coordinated communication and autonomous task acquisition and execution are studied, in order to further improve the autonomous interaction and efficient real-time communication between CPS nodes, the Carnegie Mellon University (CMU) applies the Support vector machine prediction model and so on to the modeling and optimization of CPS in Smart Grid[10], it realizes the optimal coordination and allocation of energy in wind power generation, and the related methods are expected to be applicable to the scheduling management of Distributed New Energy CPS in the future. In addition to the United States, Europe, Japan and Korea are also engaged in similar research, many scholars in Europe on the CPS framework and modeling new methods of discussion and research [11], the EU has done a lot of work on the integration of intelligent electronic systems and complex systems with multiple components, starting Artemis and other projects. From the perspective of automation research and development, the Korean Academy of Science and Technology and other research institutes are concerned with the cross-platform research of the integration of computing devices, communication networks and embedded objects, the University of Tokyo, Japan, and other applications of CPS technology in Intelligent Medical Devices and robot development, such as the application of great scientific research efforts[12].

Some scholars have Exploratory research on CPS's pervasive network environment, summarizing its requirements or defining its key technologies in the light of technologies such as the Internet of things and cloud computing. However, from the perspective of relevant research, CPS research is still in the stage of defining the concept and attributes, discussing the scope of application, and discussing the related key technologies, even very few of the CPS from the concept to the actual application of the systematic elaboration.

In the research of CPS, it is very important for the overall research of system model design and architecture technology. Dragos Repta et al. [13] introduced a model design method for building an information physics fusion system and verified it by taking the development of Tunneling Ball Device (TBD) as an example, the design process of a small real-time information physical fusion system is completed step by step. Rahul Mangharam and Miroslav Pajic [14] studied the key technologies of distributed information physical fusion systems and introduced the embedded virtual machine method to abstract the control task; the controller can be decoupled from the physical background and migrated in real time. The wireless transmission network itself is also included in the category of distributed controller, and the computing process and rules of the controller are implemented in a fully distributed way in the network. Samy El-Tawab et al. [15], based on CPS theory and key technologies, constructed a small-scale information physical fusion system for freeway accident detection and information release. The system is divided into a node detection layer and road test unit layer. Based on Bayesian cognitive science, the fault detection and recognition algorithm is designed. The system network and detection algorithm are simulated and evaluated using ONE simulation environment. Alexandru Stefanov et al. [16] based on the CPS system design idea, a new model and simulation platform for the SCADA (supervisory control and data acquisition) system is proposed to evaluate the performance and stability of its systems and methods, it includes information physical model, information attack model and information physical security evaluation. In the concrete research of CPS model and System Architecture, based on the CPS design idea, the system model design is completed according to the concrete application scenario, embedded Computing, distributed network and pattern design are integrated to solve many specific problems.

The high level and complexity of CPS in terms of structure and behavior poses a great challenge to the realization of researchers' vision in many fields, requiring the development of a fundamental theoretical framework for a convergent approach to address dynamic CPS. Some scholars focus on the overall level of CPS system and key technology research, in the macro-system research results of CPS. Tan y et al. [17] studied the traditional embedded system architecture, it is pointed out that global time synchronization, human physical activity driving, inter-system component fusion and information distribution are important problems that restrict the performance and efficiency of the old system architecture, a new CPS definition and prototype architecture is proposed, which includes key technologies such as global time synchronization, event/information driven, distribution mechanism, control semantic rules and new generation network technology.

CPS involves a wide range of technology, with the in-depth study of CPS scholars, People's understanding of the specific structure of CPS gradually clear, the corresponding hierarchical structure is proposed for its application in various industries. In the light of the industry 4.0 trend in manufacturing, Lee J et al. [18] proposes a CPS architecture for Industry 4.0 manufacturing systems, as shown in Figure 1, that is, intelligent access layer, data to information conversion layer, information layer, cognitive layer and configuration layer; It is proposed that CPS is composed of two basic elements:

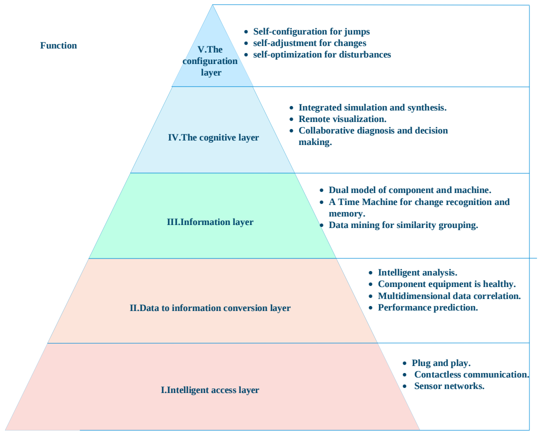


Figure 1. *Industry 4.0 CPS architecture*

1) to ensure the high-level interconnection of real-time data sampling from the physical world and information feedback from the information space.

2) to build the intelligent data management, analysis and computing capabilities needed for the information space.

In order to solve the problem of insufficient coordination and optimization caused by the lack of extensive connectivity and interoperability in traditional transportation systems, Sun et al[19]. introduced the key technology of CPS into the field of intelligent traffic control, taking full advantage of the effective interaction and feedback between its information system and physical system, this paper studies the traffic information physical fusion system (T-CPS) [20].

Since the next generation of CPS will operate on a larger scale and in a more open environment, its control methods should be extended to a broader context. High-level decision-making algorithms and theories are needed to satisfy the reliability, efficiency, security and robustness of the entire CPS system for information obtained from different time depths and spatial breadth. The information physics fusion system is still a brand-new research field, and various research achievements are continuously enriching and perfecting its definition and content, Table 1 lists the development and evolution of the information physical fusion system concept. The research process of the integrated information physical fusion system can be known. The information physical fusion system emphasizes the deep fusion of the information system and the physical system; it focuses on the deep interaction

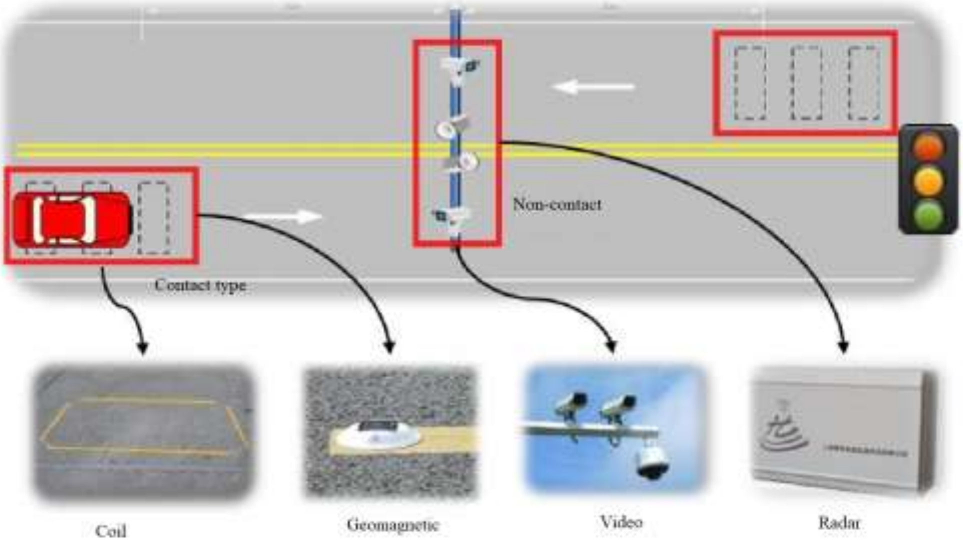
Table 1. *Development of information physics fusion system concept*

Research Unit/scholar	Year	Definition
NSF	2006	CPS is a system in which computational and physical elements are closely integrated and coordinated [21].
Shankar Sastry	2008	The information physical fusion system integrates computing, communication and storage capabilities, has the ability to monitor and control the physical world entities, and can guarantee independence, security, efficiency and real-time [22].
Krogh	2008	The information physical fusion system is composed of computing equipment, communication network, sensor equipment and physical equipment. All equipment cooperates and influences each other, through the deep fusion of physical process and computing process, thus determining the function of the system and behavioral characteristics [23].
R.Poovendran	2010	Information physical fusion system embeds information capability deeply in the physical world and carries out information interaction in the physical world of human, architecture and platform [24].
Lee	2011	The information physical fusion system is a process of integrating physics and computing, and it is a deep interaction between physics and information system, not a simple union [25].
KJ Park	2012	The information physical fusion system is the close integration of computation and physical process, which integrates embedded computing, network monitoring and physical process control, usually including the interaction and feedback between physical process and computational process [26].
Repta D	2014	The concept of information physical fusion system covers the close fusion of information unit and physical unit. The information unit refers to computation, communication and control A physical unit is a system of the natural and man-made world that follows the laws of physics and the continuous properties of time [27].

between the physical elements of the real world and the computing elements of the information space. Information physical fusion system will surpass traditional intelligent system in efficiency, security, reliability, robustness and adaptability, that is to achieve a timely response, accurate control, safe and reliable and efficient complex system.

2.1.3 Traffic information perception technology

Traffic information refers to the circulation of available information in the field of traffic and transport [28], traditional traffic information generally refers to traffic flow parameters, traffic management planning information, traffic facilities information, traffic incident information and traffic environment information. The physical basis of traffic information perception is traffic information detector, which is mainly composed of sensors and supporting systems, high reliable traffic information detector is the key means to ensure the information (speed, density, flow, queue length, etc.) to be collected comprehensively and accurately. So far, many scholars have made good foundation and great progress in the field of traffic information perception, and the main traffic information perception detectors include contact and non-contact. Among them, the contact traffic information detector mainly includes geomagnetic, Coil, etc. The non-contact traffic information detector mainly includes radar, video, infrared and ultrasonic. In order to establish a reliable traffic information perception model, it is necessary to study and analyze the existing traffic information perception technology and system at home and abroad.The development of intelligent transportation system (ITS) makes the content of“Traffic information” more and more extensive [29], it integrates traffic facility information, traffic flow information, parking lot information, traffic incident information, traffic environment information, and sometimes even traffic network information and traffic management information. This paper focuses on the narrow sense of traffic information, that is, the perception of vehicle traffic flow information. Traffic Information Awareness System is the foundation of intelligent transportation system, which is an important sense of intelligent transportation, at present, the main traffic state detection technology is divided into mobile traffic state acquisition technology and fixed installation detection technology. Among them, there are two technologies of mobile traffic state acquisition, floating car sampling and mobile terminal location information service. The fixed-mounted detection technology is widely used in the traffic information sensing system and has the characteristics of diversification. A large number of different kinds of sensors are applied in different scenes, they include induction coils, geomagnetic sensors, pneumatic road pipe detectors, passive Infrared detectors, active Infrared detectors, Doppler radar detectors, vehicle microwave radar detectors, ultrasonic detectors, passive acoustic detectors and video image detectors [30]. Some scholars have done extensive research and induction on the application of fixed-mount detection technology in developed countries abroad, literature [31] summarizes the object perception, accuracy and application cost of various sensors used in traffic information detection engineering and makes a detailed comparison and performance evaluation. Table 2 shows the detection performance of representative products in various types of sensors.



Figures 2. *schematic of common traffic information detectors*

Table 2. *Function list of all kinds of sensors and their typical products*

Sensor type	Objects of perception				
	The vehicle exists	Traffic	Speed	Model	Occupancy
Induction coil.	√	√	√	√	√
Geomagnetic sensor					
3M Microloop	√	√	√	√	√
SPVD	√	√	√	×	√
Pneumatic road pipe detectors.	×	√	√	√	×
Passive Infrared detector					
ASIM IR 254	√	√	√	√	√
Eltec Model 842	√	√	×	×	×
Sienmens PIR-1	√	√	×	×	√
Active Infrared detector					
Autosense II	×	√	√	√	×
Doppler radar detector					
TDN-30	×	√	√	×	×
Loren	√	√	√	√	√
The vehicle has a microwave radar detector.					
Accuwave 150LX	√	√	×	×	×
RTMS	√	√	√	√	√
Ultrasonic detector.					
TC-30	√	√	×	×	×
Lane King	√	√	×	×	×
Passive acoustic detector.					
SmarTek SAS-I	√	√	√	√	√
Video Image Detector					
Autoscope	√	√	√	√	√
VideoTrak	√	√	√	√	√
Traffic Vision	√	√	√	√	√

2.2 Application of radar in traffic

Radar sensors are superior to visual sensors in detecting target distance, direction and speed, and are widely used in the fields of automatic driving assistance systems, traffic detection and security. Scholars have conducted extensive research and application on them.

Radar is a kind of special electromagnetic wave, whose frequency range is between 10GHz and 200GHz, between radio wave, infrared ray and visible light. The range of radar is mainly by transmitting millimeter-wave, then receiving the Echo, using the time difference of the reflected wave of the obstacle to determine the relative distance between the obstacle and the radar, the relative velocity between them is determined by the offset of the reflected wave frequency. At present, 24GHz and 77GHz are the main frequency bands of millimeter wave radars at home

Figure 3. *Radar and application scenarios*

and abroad, and 24GHz is mainly used for short and medium range radars (measuring range is about 15-30m), the 77GHz is mainly used for long-range radars (measuring distances of about 100-200 meters). Because of the powerful and comprehensive information perception ability of MMW radar, the research on MMW radar has been carried out at an early stage in foreign countries. As early as the 1990s, some n, Japanese and German giant enterprises have already started systematic research and gradually applied radar to cars through practice. After 2010, companies such as OTOLAF, Robert Bosch GmbH, Continental, Delfaut, Hela, Denso, Fujitsu, Tianhe, and Farreo all had significant shares of the radar market, compared with foreign countries, domestic millimeter wave radars are still slightly inferior to foreign manufacturers in various detection parameters due to the lack of research on the technology of late-start observation, etc. . At present, domestic millimeter wave radars are mainly used at 24 GHz, the 77GHz millimeter wave radar is still in its initial application stage, and its technology and practical experience are not yet mature.

Compared with MMW radar, video has the ability of recognizing the visual information such as object scale, color, texture and so on. The vision sensor can distinguish and recognize various kinds of rich traffic information, such as traffic signs, marking lines, license plates, vehicle types, etc., at present, there are mainly target detection methods based on traditional image features, target detection methods based on traditional features combined with machine learning, and target detection methods based on depth learning. At present, the popular traditional target detection methods include frame difference method[32], background subtraction method[33], network and optical flow method[34], for example, Zhuq et al. [35] studied vehicle detection in tunnel scenes, pre-processing video sequences with Histogram equalization, and then modeling the Background using the classic ViBe (visual Background Extractor) algorithm, on this basis, the morphological filter is applied to eliminate some noise interference, and finally the moving traffic target is extracted. However, traffic scenes in practical applications are complex and changeable, and a large amount of environmental noise brings great difficulties to background extraction. Therefore, Unzueta et al. [36] proposed a background subtraction method based on adaptive multimodal, the robustness of the detection algorithm is improved, and the vehicle counting can be carried out better.

Besides, an. Rumaksari et al. [37] also implemented an adaptive shadow filter based on background subtraction, which can adapt to dynamic illumination changes in the environment and be better applied to surveillance video. In recent years, deep learning has been developing rapidly, and target detection technology has received a major breakthrough. In 2014, after SermanetP et al. proposed the R-CNN framework [38], scholars have carried out a lot of research on this basis, it indicates that the target detection technology has entered a brand-new stage. Faster R-CNN [39] is an efficient two-stage target detection algorithm, and FanQ et al. Further utilized Faster R-CNN in the field of vehicle detection, better results were obtained on the Kitti dataset by adjusting parameters and training strategies [40]. Gaoy et al modified the region scale and used more suitable full-image convolution features for vehicle detection, improving the detection effect of small targets [41].

Arunesh et al. [42]. studied the multi-target speed measurement of Doppler radar. To solve the current single-target speed measurement problem of cameras, they used the optical flow method to extract the vehicle contour and fused the radar data with the video to achieve the speed detection of multiple targets. Jiexin Ren et al. [43]. proposed a vehicle speed measurement method based on the fusion of digital camera and radar to solve the shortcomings of the traffic monitoring system. The YOLOv3 algorithm has good vehicle detection performance. It not only has high accuracy but also has the ability to detect vehicles in real time. The experiment verified that the radar could measure the distance, azimuth, pitch angle and speed of the vehicle, and the error is within a reasonable range.

At present, scholars have done a lot of research on the application of radar and video in traffic, mainly focusing on radar. The research mainly focuses on speed detection, traffic flow detection, and azimuth detection, supplemented by video. However, there is little research on the detection of various driving behaviors of vehicles.

2.3 Video Detection of Traffic Objects

In the field of intelligent transportation systems, target detection algorithms are mainly divided into two categories: one is the classic background modeling algorithm [44], and the other is the method of using neural networks for model training. The background modeling algorithm must first obtain the current background frame image and then perform a differential operation on the current video frame image and the background frame image to obtain the shape and position information of the moving target in the current frame, so that the target can be detected. NC Mithun [45] et al. proposed a new detection method that uses multiple spatiotemporal images to provide potential opportunities for identifying vehicles, which can reduce the dependence of static and moving objects on pixels and improve detection accuracy. A Jazayeri et al. proposed a method for locating target vehicles in videos. This method projects the geometric features extracted from the video onto a one-dimensional contour and continuously tracks it. It relies on features and motion behaviors to identify vehicles and uses a hidden Markov model (HMM) to separate the target vehicle from the background and track it [46].

Although the traditional target detection algorithm has a simple implementation principle and low computational complexity, it can only detect moving targets. Especially in today's urban roads, where lanes are densely populated with vehicles, the traditional detection algorithm cannot completely detect some targets, which is prone to false detection and missed detection. Therefore, it is not suitable for use in the construction of intelligent transportation systems. With the development of the computer field, its computing power has made important breakthroughs, and the application field of computer vision using a deep learning framework has developed rapidly. Regarding the detection of traffic targets, since both the target and the background have diverse characteristics, the algorithm that can adapt to this complex environment is the convolutional neural network, which has the ability to handle target detection in complex scenes. Krizhevsky et al. used a deep convolutional neural network (DCNN) structure for target detection, which greatly improved the detection accuracy [47]. Since then, many researchers have conducted research on improving the performance of convolutional neural network models. There are currently two main types of target detection: candidate region based, and single stage based. Girshick R et al. proposed the R-CNN detection network model, which can obtain the region of interest through the candidate region method and use the selective search method to improve the detection efficiency. This method has been tested on the VOC2012 dataset, and its mAP is 53.3%. This algorithm uses a large number of candidate regions to improve accuracy and has a large amount of calculation. Girshick R proposed Fast RCNN [48] and Faster RCNN [49] in 2015. Fast RCNN first extracts the features of the entire image and then directly applies the method of creating candidate regions to the feature map. The mAP of this method is 65.7% through testing. In Faster RCNN, the region proposal network (RPN) is used instead of the candidate region method, which has higher detection efficiency. The methods that use single-shot object detectors mainly include SSD and YOLO series. In the SSD detection algorithm, the VGG19 network is used as a single-shot detector of the feature extractor. It uses the deeper layers in the convolutional network to detect objects, which makes it more difficult to detect objects with smaller pixels [50]. The YOLO series [51] uses DarkNet for feature detection after the convolutional layer, uses an end-to-end detection framework, improves the detection speed, and has better detection performance.

2.4 Research on radar and video information fusion algorithm

At present, in the research of radar video information fusion algorithms, domestic and foreign scholars have put forward a variety of basic theories and solutions. FI.A.T, from Italy, uses radar and long-range infrared cameras to detect and track obstacles in front of vehicles, roads, etc. through centralized data fusion [52]. Toyota, a Japanese company, fuses a radar with a camera. First, images are selected, and regions of interest (rois) are built using radar sensing information. Then, a neural network is used to identify traffic objects in the rois, however, the detection accuracy and efficiency of this method are relatively low [53]. Israel-based Mobileye has developed an advanced Système d'aide à la conduite, à l'exploitation et à la maintenance called EyeQ that supports ADAS based on vision sensors, in the fourth and fifth generation products can also be connected to millimeter wave radar, LIDAR and other multi-sensor signals, further support for semi-autonomous and fully autonomous driving [54]. At the same time, the new-energy-vehicle giant Tesla has teamed up with Nvidia to develop a second-generation Autopilot, complete with autonomous driving. Autopilot can capture and manage the information around the vehicle through a variety of sensors and will be different degrees of hazard of obstacles to distinguish between different colors, providing a“Virtual driver indication” function [55]. On the part of individual and academic researchers, Giancarlo Alessandretti et al. proposed a vehicle detection method based on the fusion of radar and visual data to locate the region of interest (Roi) of the target on the image, then vehicle detection is realized by vehicle vertical symmetry [56]. Huangl et al used Lidar and video camera to detect the vehicle in front, and the experiment verified that the method provided enough information to deal with the multi-view problem of the vehicle and improved the robustness of the system [57].

Multi-sensor Information Fusion (MSIF), as the name implies, processes data from a large number of sensors, which can be of the same or different types.

The data from these sensors are effectively fused to produce a consistent description of the detected environment, and the fusion output is used to make decisions to obtain more complete information than the combination of each part. The basic principle of multi-sensor fusion is similar to the processing method of the human brain. It enables different sensors to form the best combination at multiple levels and spaces, complement each other, and finally provide a unified description result.

In this design, the camera has its own output data, and the radar will also generate its own output data. The output data of all sensors are sent to the fusion system for fusion. Figure 10 shows the post-fusion system structure.

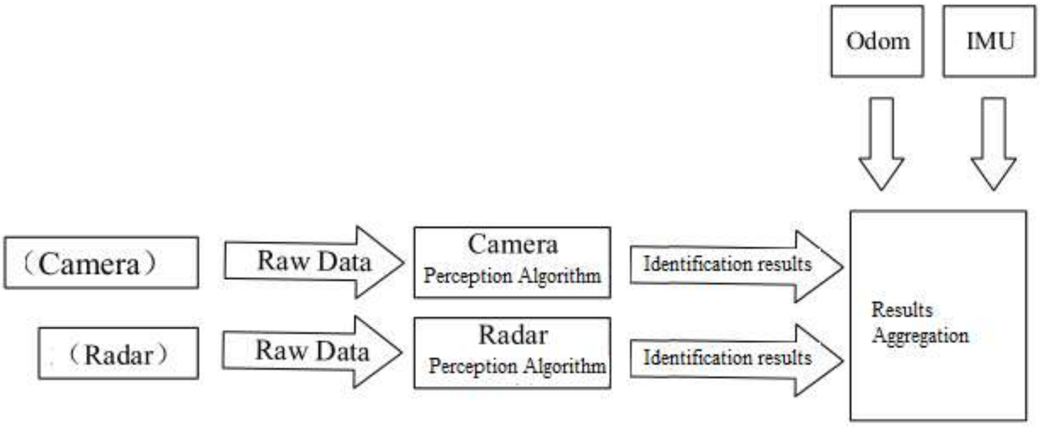


Figure 4. *Fusion System Structure*

3. Methodology

3.1 Radar-based traffic object detection
3.1.1 The Basic Principle of Radar

A radar system consists of a transmitter that emits a specific waveform and a receiver that receives an echo signal. When it works, it will send waves to the front through the antenna. The surface of the object will scatter and reflect the signal when it hits the millimeter wave. The millimeter wave radar system receives the echo signal through several receiving antennas and amplifies down-conversion and signal processing the signal, so as to obtain the velocity, azimuth, range and scattering cross-section of the target. A typical MMW radar model is shown in figure 5 Liner Frequency Mochilated Continious Wave (LFMCW) is a special millimeter Wave radar technology. The radar system will continuously transmit LFM signal, the frequency of the signal will increase linearly with time in one FM period, and the receiving antenna of the radar will capture the reflected signal, the reflected signal is delayed and Doppler effect because of the target's radial range and velocity relative to the radar. The range and velocity of the target are extracted by continuously transmitting multiple sets of the above-mentioned LFM signals, mixing them with the reflected signals, and low pass filter them to obtain multiple sets of difference frequency signals

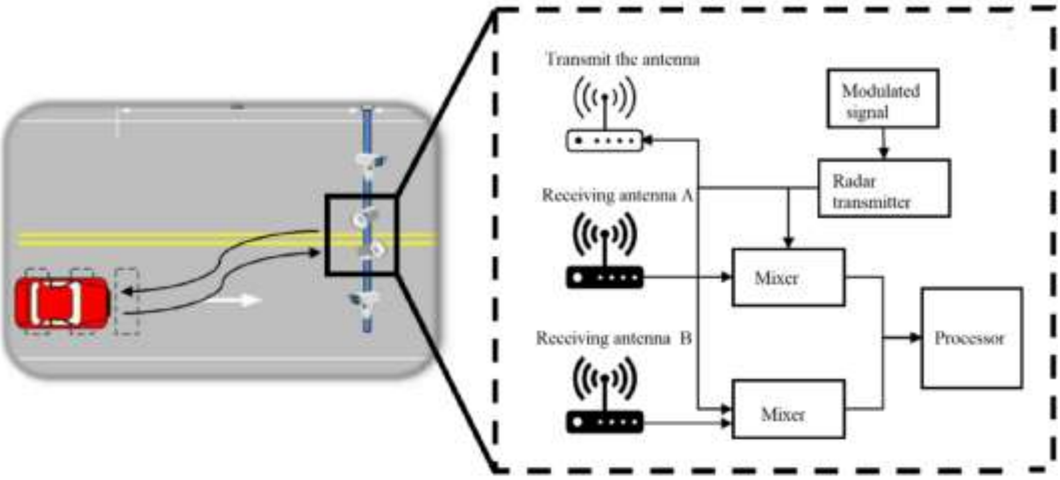


Figure 5. *Schematic diagram of radar working principle*

[58]. Among them, the LFMCW radar can also estimate the target's azimuth by using the range difference of the reflected signals received by multiple antennas.

3.1.2 Radar data acquisition

When processing single sensor data, the main purpose is to verify the multi-target tracking algorithm of radar and video and provide a basis for the subsequent fusion algorithm. Therefore, according to the traffic application scenario of this article, the overpass test scenario is selected to collect radar and video data and analyze them. The radar and camera collect data from a bird's-eye view, and monitor traffic targets such as vehicles, non-motor vehicles, and pedestrians on the traffic road diagonally downward.
The radar development platform used in this article is shown in Figure 6.

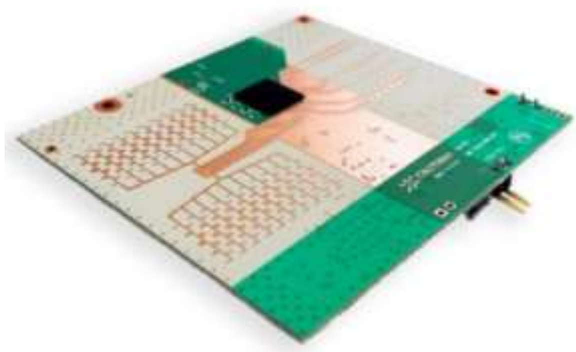


Figure 6. *Radar development platform*

Its parameter format is shown in Table 3. According to the number of targets that the radar can measure and the maximum measurement distance and other parameters, the radar detection range is suitable for traffic application scenarios.

Table 3. *Radar parameters*

Number of targets	≥100
Maximum measurement distance	210m
Distance resolution	0.7m
Maximum measurement speed	31m/s
Speed resolution	0.2m/s
Data rate	50ms
Environment	All-weather，-10~50℃

3.2.3 Radar data processing

The radar data processing process mainly includes the following parts:

(1) Raw data preprocessing After the radar development board collects the original hexadecimal number, it parses the data into decimal target information according to the data parsing format, mainly including frame number, radial distance, speed, angle. According to the radial distance and angle, the target data is converted to (x, y) coordinates and displayed in the two-dimensional rectangular radar coordinate system.

(2) Point trace aggregation During the radar data acquisition process, the same target may have multiple reflection points. Therefore, the aggregation threshold value is set to condense the points that meet the threshold in the same frame. This method reduces the complexity of the data and reduces the subsequent unnecessary processing process. The condensation result is output as the point trace of this target.

(3) Track initiation After the data is condensed, there are multiple target points in each frame. In order to track the target later, the track initiation is required first. If 6 or more of the 8 adjacent frames are marked with the same track number, the track is considered to be stable at the beginning; if 3 or less of the 8 adjacent frames are marked with the same track number, the track is

considered to be a false track and the track is deleted.

(4) Track association Track association mainly matches the started track with the new input data. By evaluating and scoring the track, the quality and status of the track are judged. The new frame data is preferentially associated with the track with a high track quality score for the next update.

(5) Adaptive α-βfilter target tracking After the track starts, the target is tracked by the adaptive α-βfilter algorithm, and the vehicle target ID is marked as the input of the decision-level fusion algorithm. In order to verify the single-sensor target tracking algorithm, this paper collects targets on the traffic road in the overpass test scenario. The data results after tracking and processing the collected radar test data are shown in Figure 8.

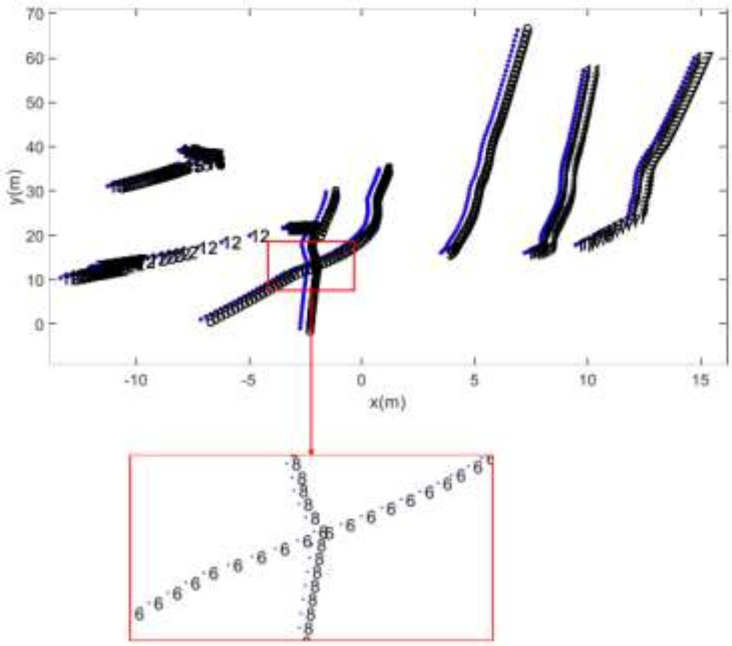


Figure 7. *Radar tracking data*

As shown in Figure 7 after the radar traces are processed, the target ID number is marked on the right. The output result of radar data processing contains the following target information: frame number, ID number, x-coordinate, y-coordinate, velocity. The measured data after tracking processing is shown in Table 4

Table 4. *Radar target tracking data after processing*

frame number	ID number	x(m)	y (m)	velocity (m/s)
4	1	-25.4902	32.2681	-0.01112
4	6	0.75709	34.937	-9.45
5	1	-25.584	32.2035	-0.01112
5	6	0.740645	34.6493	-9.45
6	1	-25.6798	32.1376	-0.01112
6	6	0.719973	34.3277	-9.45
7	1	-25.7777	32.0702	-0.01112
7	6	0.698607	33.9876	-9.45
8	1	-25.8775	32.0015	-0.01112
8	6	0.668752	33.611	-9.45

Table 4 captures some of the measured data results. In frames 4 to 8, two targets appear with ID numbers 1 and 6 respectively, and the position and speed of each target in this frame are also known through radar data analysis.

3.2 Video Detection of Traffic Objects

3.2.1 Video data acquisition

Detect and identify the traffic information of vehicles through video images. It is a method of analyzing the vehicle video monitoring signal to obtain the traffic data information of the vehicle on the road. The application focuses on video image processing. This method has high technical requirements and overcomes the angle requirements of ordinary speed measurement methods. However, video detection technology needs to be further improved. Video detection will definitely become the mainstream technology in the future.

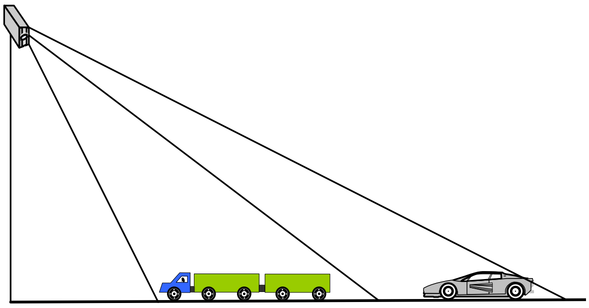


Figure 8. *Video detection side view*

Figure 9. *DJI Osmo Action*

Image sensor	1/2.3” CMOS有效像素:12M
Video format	4K（16:9）（4:3）： 60,50,48,30,25,24fps 2.7K（16:9）（4:3）： 60,50,48,30,25,24fps 1080p“”240,200,120,100,60,50,48,30,25 ,24fps 720p：240,200fps
Horizontal viewing angle	145 degrees
Video format	MOV, MP4 (H.264)

3.2.2 YOLO Series Network Nodel Principle

YOLO, called You Only Look Once: Unified, Real-Time Object Detection, is an end-to-end Object Detection method proposed by Joseph Redmon et al. [59]. It is a kind of convolutional neural network based on the idea of regression, which takes the whole image as the input of the network and can return the boundary frame and the category of the target directly on the image, thus realizing the end-to-end target detection and recognition. The most advantage of the network is its fast speed, and because of the direct selection of the whole graph training model, full use of global information, can better distinguish between the target and background areas. The YOLO convolutional neural network was introduced in 2015 and is now in its fifth generation. The first three generations of models were proposed by Joseph Redmon[60], while the last two generations of models were improved by Alexey Bochkovskiy [61] and Glenn Jocher, respectively, which introduced some excellent optimization strategies in the field of CNN in recent years, from data processing, backbone network, network training, activation function, loss function and other aspects have different degrees of optimization, the performance of their model has reached the most advanced level in the field of image target detection.

In general, the YOLO models follow in the same vein, and the principles of each model are derived from YOLOv [62]. Therefore, here's a brief introduction to the YOLOvl model:

Step 1: Polo models adjust the input image to a fixed 448× 448-pixel size and divide it into S×S grids, if the center of the object in the image is within a grid, the grid is responsible for detecting the object.

Step 2: for each grid, you need to predict borders B, each containing five parameters, the coordinates (x, y), width and height (w, h) and confidence level (c) of the border are respectively, the coordinates (x, Y) represent the center of the prediction border relative to the boundary of the grid. The coordinates (W, H) represent the ratio of the width and height of the predicted border to the width and height of the entire image. Confidence is the IOU value of the predicted border and the real border (IOU in training and a confidence value in reasoning), it reflects the possibility of the existence of the target and the accuracy of the target position prediction based on the current model prediction frame.

Step 3: each grid also predicts the probability of C conditional categories, that is, the probability that a grid belongs to a certain category if it contains a target, then the final dimension of the feature graph used for prediction is S×S× (B×5 + c);

Step 4: multiplying the confidence of each predicted border and the conditional category probability to obtain its confidence score, applying non-maximum suppression to all candidate boxes, excluding the less likely candidate boxes, and outputting the final remaining border, achieve the target positioning and classification.

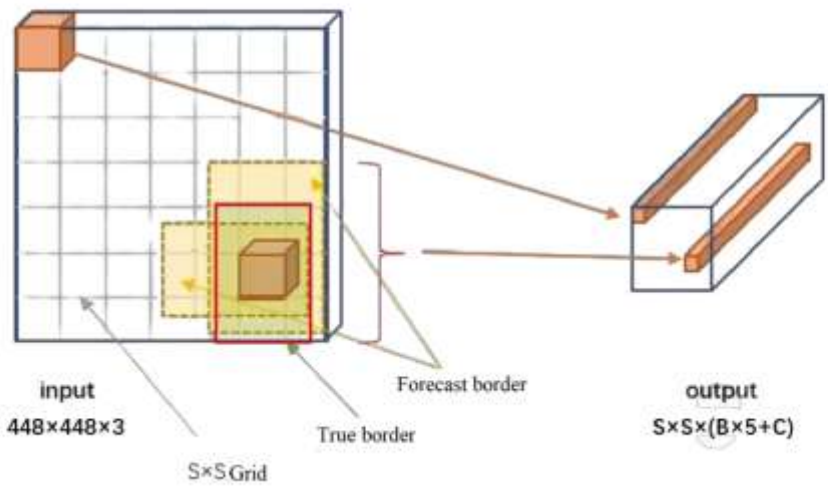


Figure 10. *Schematic diagram of YOLO networks*

Through the above algorithm description, we can see that in YOLOVL model, each grid can only predict B borders, and they share a category, so the model is not good for small target detection, and the recall rate is not high. Therefore, in YOLOV2 and YOLOv3, the idea of RPN is used for reference, and the mechanism of anchor frame is introduced to predict the target, the feature map extracted from the backbone network is directly transformed into a predictive feature map whose scale is S × s and the number of channels is B×(5+C) by convolution of 1×1, where S is the scale of the feature map and B is the number of anchor frames, 5 refers to the above 5 border parameters, c is the number of predicted categories. In addition, YOLOV2 and YOLOv3 fused the feature maps of different scales in order to improve the detection effect of small targets. In YOLOV3, three prediction branches were directly derived, with the scales of 13×13,26×26 and 52×52, respectively, there are three anchor frames on each scale, which greatly improves the detection accuracy of the targets with different scales. So far, the POLO model frame has been basically formed, and subsequent YOLOV4 and YOLOv7 have only optimized the various modules of the network on this basis, with no major changes in principle.

4. Results discussion

In this study, the proposed intelligent vehicle detection and recognition framework was experimentally validated for performance in various traffic scenarios. The experimental results show that the framework demonstrates high accuracy and reliability in vehicle detection and recognition, performing stably under different lighting conditions, weather conditions, and varying degrees of vehicle occlusion. Specifically, the vehicle detection accuracy reached 63 % in low-light conditions, and the accuracy remained above 64% even under high occlusion conditions. This indicates that the framework has strong adaptability to complex traffic environments.

Moreover, compared to traditional single-sensor detection methods, the multi-sensor detection method that integrates radar and video information proposed in this study demonstrates significant advantages in terms of detection accuracy and robustness. For instance, when measuring vehicle speed, using only video sensors results in a larger measurement error. However, by integrating radar data, the error can be reduced to within 65%. This demonstrates that by fusing information from multiple sensors, the strengths of different sensors can be fully utilized, compensating for the limitations of a single sensor, thereby enhancing the accuracy and reliability of traffic information perception.

However, despite the achievements of this study, several limitations remain. For instance, the current framework may affect the system's real-time performance when handling large-scale traffic flows. In high-density traffic scenarios, the speed of vehicle detection and recognition might slow down, failing to meet the demands of real-time traffic monitoring. Moreover, the accuracy of detecting and recognizing certain special types of vehicles, such as motorcycles and tricycles, is relatively low, necessitating further algorithm optimization to enhance the detection capabilities for these vehicles.

5. Conclusion

This study proposes a vehicle detection and recognition method based on a traffic information physical fusion system. By integrating radar and video data, the method achieves accurate vehicle detection and recognition. Experimental results show that this method demonstrates high accuracy and reliability in various complex traffic scenarios, providing robust support for intelligent transportation systems. Compared to traditional single-sensor detection methods, the multi-sensor information fusion approach offers significant improvements in detection accuracy and robustness, better meeting the demands of modern traffic monitoring.

However, this study also has some limitations. When dealing with large-scale traffic flows, the system's real-time performance needs further improvement; for certain special types of vehicles, the accuracy of detection and recognition requires further optimization. Future research will focus on addressing these issues to enhance the vehicle detection and recognition framework. Specifically, more efficient algorithms and data processing methods will be explored to improve the system's real-time performance; at the same time, the ability to detect and recognize different types of vehicles will be enhanced, improving the algorithm's generalization and adaptability. Additionally, the depth and breadth of multi-sensor information fusion will be further studied, exploring more sensor combinations and fusion strategies to enhance the accuracy and reliability of traffic information perception, providing stronger technical support for the development of intelligent transportation systems.

Author Statements:

- **Ethical approval:** The conducted research is not related to either human or animal use.
- **Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper
- **Acknowledgement:** The authors declare that they have nobody or no-company to acknowledge.

- **Author contributions:** The authors declare that they have equal right on this paper.
- **Funding information:** The authors declare that there is no funding to be acknowledged.
- **Data availability statement:** The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

References

[1] Ministry of Transport of the People's Republic of China. (2023). *Statistical Bulletin on the development of the 2023 transport sector*. Beijing: Ministry of Transport of the People's Republic of China.

[2] Lu, H., & Li, R. (2014). The development status and trend of urban intelligent transportation system. *Engineering Research-Engineering from an Interdisciplinary Perspective*, (01).

[3] Zhao, N., Yuan, J., & Xu, H. (2014). Overview of intelligent transportation systems. *Computer Science, 41*(11), 7-11.

[4] Zhou, Z. M., Zhao, M., & Sun, D. H. (2019). An improved mean-field lattice hydrodynamic model with consideration of the average effect of multi-lattice interaction. *IEEE Access*, (6), 168798-168804.

[5] Hehenberger, P., Vogel-Heuser, B., Bradley, D., et al. (2016). Design, modelling, simulation and integration of cyber physical systems: Methods and applications. *Computers in Industry, 82*(C), 273-289.

[6] Zhang, N. (2021). A cloud-based platform for big data-driven CPS modeling of robots. *IEEE Access, 9*, 34667-34680.

[7] Chen, X., Zhu, Y., Zhao, Y., et al. (2021). Hybrid modeling and model transformation of AADL for verifying the properties of CPS space-time compositions. *IEEE Access, 9*, 99539-99551.

[8] CPS Steering Group. (2011, December 18). *Report: Cyber Physical Systems Summit*. http://cesg.tamu.edu/icCPS2013/_doc/CPS_Submmit_Report.pdf

[9] Wang, Z., & Xie, L. (2011). A review of information physical fusion systems. *Acta Automatica, 37*(10), 1157-1166.

[10] Krogh, B. H., Lee, E., Lee, I., et al. (2008). *Cyber-physical systems: Executive summary*. CPS Steer Group, Wash. DC.

[11] Li, Z., Zhang, T., & Zhang, J. (2011). A review of Information Physical Fusion Systems (CPS). *Computer Science, 38*(9), 25-31.

[12] Zhou, B., Wang, Y. P., Yu, G., & Wu, X. K. (2017). A lane-change trajectory model from drivers' vision view. *Transportation Research Part C, 85*, 609-627.

[13] Repta, D., Stanescu, A. M., Moisescu, M. A., et al. (2014). A cyber-physical systems approach to develop a generic enterprise architecture. In *Engineering, Technology and Innovation (ICE), 2014 International ICE Conference on* (pp. 1-6). IEEE.

[14] Mangharam, R., & Pajic, M. (2013). Distributed control for cyber-physical systems. *Journal of the Indian Institute of Science, 93*(3), 353-387.

[15] Olariu, S., & El-Tawab, S. S. (2012). Friend: A cyber-physical system for traffic flow related information aggregation and dissemination. In *World of Wireless, Mobile and Multimedia Networks* (pp. 1-6).

[16] Stefanov, A., Liu, C. C., Govindarasu, M., et al. (2015). SCADA modeling for performance and vulnerability assessment of integrated cyber–physical systems. *International Transactions on Electrical Energy Systems, 25*(3), 498-519.

[17] Tan, Y., Goddard, S., & Perez, L. C. (2008). A prototype architecture for cyber-physical systems. *ACM Sigbed Review, 5*(1), 26.

[18] Lee, J., Bagheri, B., & Kao, H. A. (2015). A cyber-physical systems architecture for industry 4.0-based manufacturing systems. *Manufacturing Letters, 3*, 18-23.

[19] Sun, D., Li, Y., Liu, W., et al. (2013). A review of traffic information physical fusion system and its key technologies. *Chinese Journal of Highways, 26*(1), 144-155.

[20] National Science Foundation of the United States. (2011, June 27). *Cyber-Physical Systems (CPS)*. http://www.nsf.gov/pubs/2010/nsf10515/nsf10515.htm?org=NSF

[21] Cardenas, A. A., Amin, S., & Sastry, S. (2008). Secure control: Towards survivable cyber-physical systems. In *The 28th International Conference on Distributed Computing Systems Workshops* (pp. 495-500). IEEE.

[22] CPS Steering Group. (2008). *Cyber-physical systems, Executive Summary* (Technical Report). Available: http://varma.ece.cmu.edu/Summit/

[23] Krogh, B. H., Lee, E. A., Mok, A., et al. (2010, July 10). *Cyber-physical systems executive summary*. http://varma.ece.cmu.edu/Submmit/CPS-Executive-Summary.pdf

[24] Poovendran, R. (2010). Cyber-Physical Systems: Close encounters between two parallel worlds. *Proceedings of the IEEE, 98*(8), 1363-1366.

[25] Lee, E. A., & Seshia, S. A. (2011). *Introduction to embedded systems: A cyber-physical systems approach*. Lee & Seshia.

[26] Park, K. J., Zheng, R., & Liu, X. (2012). Cyber-physical systems: Milestones and research challenges. *Computer Communications, 36*(1), 1-7.

[27] Repta, D., Stanescu, A. M., Moisescu, M. A., et al. (2014). A cyber-physical systems approach to develop a generic enterprise architecture. In *Engineering, Technology and Innovation (ICE), 2014 International ICE Conference on* (pp. 1-6). IEEE.

[28] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*. arXiv preprint arXiv:1804.02767.

[29] De Fabritiis, C., Ragona, R., & Valenti, G. (2008). Traffic estimation and prediction based on real time

floating car data. In *Intelligent Transportation Systems, 2008. ITSC 2008. 11th International IEEE Conference on* (pp. 197-203). IEEE.

[30] White, J., Quick, J., & Philippou, P. (2004). The use of mobile phone location data for traffic information. In *Road Transport Information and Control, 2004. RTIC 2004. 12th IEE International Conference on* (pp. 321-325). IET.

[31] Herrera, J. C., Work, D. B., Herring, R., et al. (2010). Evaluation of traffic data obtained via GPS-enabled mobile phones: The Mobile Century field experiment. *Transportation Research Part C: Emerging Technologies, 18*(4), 568-583.

[32] Lipton, A. J., Fujiyoshi, H., & Patil, R. S. (1998). Moving Target Classification and Tracking from Real-time Video. In *1998 IEEE Workshop on Applications of Computer Vision* (pp. 8-14). IEEE Computer Society.

[33] McKenna, S. J., Jabri, S., Duric, Z., et al. (2000). Tracking groups of people. *Computer Vision & Image Understanding, 80*(1), 42-56.

[34] Meyer, D., & Denzler, J. (1997). Model based extraction of articulated objects in image sequences for gait analysis. In *1997 International Conference on Image Processing* (p. 78). IEEE Computer Society.

[35] Zhu, Q., & Ma, C. Q. (2017). A vehicle detection method in tunnel video based on ViBe algorithm. In *2017 IEEE 2nd Advanced Information Technology, Electronic and Automation Control Conference* (pp. 1545-1548). IEEE.

[36] Unzueta, L., Nieto, M., Cortes, A., et al. (2012). Adaptive multicue background subtractio for robust vehicle counting and classification. *IEEE Transactions on Intelligent Transportation Systems, 13*(2), 527-540.

[37] Rumaksari, A. N., Sumpeno, S., & Wibawa, A. D. (2017). Background subtraction using spatial mixture of Gaussian model with dynamic shadow filtering. In *2017 International Seminar on Intelligent Technology and Its Applications* (pp. 296-301). ISITIA.

[38] Sermanet, P., Kavukcuoglu, K., Chintala, S., et al. (2013). Pedestrian detection with unsupervised multi-stage feature learning. In *2013 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3626-3633). IEEE Computer Society.

[39] Taigman, Y., Yang, M., Ranzato, M., et al. (2014). DeepFace: Closing the gap to human-level performance in face verification. In *2014 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 1701-1708). IEEE Computer Society.

[40] Fan, Q., Brown, L., & Smith, J. (2016). A closer look at Faster R-CNN for vehicle detection. In *2016 Intelligent Vehicles Symposium* (pp. 124-129). IV.

[41] Gao, Y., Guo, S., Huang, K., et al. (2017). Scale optimization for full-image-CNN vehicle detection. In *2017 Intelligent Vehicles Symposium* (pp. 785-791). IV.

[42] Roy, A., Gale, N., & Hong, L. (2009). Fusion of Doppler Radar and video information for automated traffic surveillance. *2009 12th International Conference on Information Fusion*, 1989-1996.

[43] Ren, J., Wang, Y., Han, Y., et al. (2019). Information fusion of digital camera and radar. In *Proceedings of the 2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)* (pp. 2-8).

[44] Xiang Hua Hou, & Hong Hai Liu. (2013). Research of Background Modeling Algorithm Method Based on Multi-Frame Average Method in Moving Target Detection. *Applied Mechanics and Materials, 2617*(380-384).

[45] Mithun, N. C., Rashid, N. U., & Rahman, S. M. M. (2012). Detection and classification of vehicles from video using multiple time-spatial images. *IEEE Transactions on Intelligent Transportation Systems, 13*(3), 1215-1225.

[46] Jazayeri, A., Cai, H., Zheng, J. Y., & Tuceryan, M. (2011). Vehicle detection and tracking in car video based on motion model. *IEEE Transactions on Intelligent Transportation Systems, 12*(3), 583-595.

[47] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems* (pp. 1097-1105).

[48] Girshick, R. (2015). *Fast R-CNN*. Computer Science.

[49] Ren, S., He, K., Girshick, R., et al. (2015). *Faster R-CNN: Towards real-time object detection with region proposal networks*.

[50] Liu, W., Anguelov, D., Erhan, D., et al. (2016). SSD: Single shot multibox detector. In *European Conference on Computer Vision* (pp. 21-37).

[51] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement*.

[52] Polychronopoulos, A., Scheunert, U., & Tango, F. (2004). Centralized data fusion for obstacle and road borders tracking in a collision warning system. *International Conference on Information Fusion*.

[53] Haselhoff, A., Kummert, A., & Schneider, G. (2007). Radar-Vision Fusion with an Application to Car. Following using an Improved AdaBoost Detection Algorithm. In *Intelligent Transportation Systems Conference, 2007. ITSC. IEEE* (pp. 854-858).

[54] Developpement, Y. (2015). Moving closer to automated driving, Mobileye Unveils EyeQ4System-on-Chip with its First Design Win for 2018. *Sensors, 2015*, 483-488.

[55] Wang, C. Y., Liao, H. Y. M., Wu, Y. H., et al. (2020). CSPNet: A new backbone that can enhance learning capability of cnn. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops* (pp. 390-391).

[56] Alessandretti, G., Broggi, A., & Cerri, P. (2007). Vehicle and guard rail detection using radar and vision data fusion. *IEEE Transactions on Intelligent Transportation Systems, 8*(1), 95-105.

[57] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.
[58] Redmon, J., Divvala, S., Girshick, R., et al. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
[59] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2017). Imagenet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), 84-90.
[60] Redmon, J., Divvala, S., Girshick, R., et al. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788).
[61] Alnujaim, I., Oh, D., & Kim, Y. (2020). Generative adversarial networks for classification of micro-doppler signatures of human activity. *IEEE Geoscience and Remote Sensing Letters, 17*(3), 396-400.
[62] Seyfioglu, M. S., Erol, B., Gurbuz, S. Z., & Amin, M. G. (2019). Dnn transfer learning from diversified micro-doppler for motion classification. *IEEE Transactions on Aerospace and Electronic Systems, 55*(5), 2164-2180.