

# Exploring the Application of Multi-Sensor Fusion Technology in Enhancing the Performance of Autonomous Vehicles

Enyuan Cao <sup>1,\*</sup>, Leyan Chen <sup>2</sup> and Lingfeng Chen <sup>3</sup>

<sup>1</sup> School of Electromechanical Engineering and Automation, Shanghai University, Shanghai, 200000, China

<sup>2</sup> Graduate School of International Studies, Beijing University of Posts and Telecommunications, Beijing, 100000, China

<sup>3</sup> Zhuhai College of Science and Technology, Zhuhai, 519000, China

\* Corresponding Author Email: 2019213397@bupt.edu.cn

**Abstract.** In recent years, single sensors have exposed many limitations in the field of autonomous driving, and various problems have arisen that make it difficult for single sensors to cope with complex environments. Multi-sensor fusion technology makes up for the shortcomings of a single sensor by integrating the advantages of different sensors. It not only improves the stability and accuracy of perception but also enhances the adaptability of the system in changing scenarios. The purpose of this paper is to analyze the characteristics of multiple sensors and their applications in autonomous driving systems, focusing on the fusion method based on camera and LiDAR and common fusion algorithms, such as the Kalman filter and neural network. This paper concludes that the multi-sensor fusion technology can effectively improve the target detection accuracy, enhance the system robustness, and show significant advantages in obstacle sensing and obstacle avoidance strategies. In addition, enhancement schemes such as optimized fusion algorithms, intelligent modal selection, and edge computing are proposed to address challenges such as real-time and computational complexity. This study provides a theoretical foundation and practical reference for the further application of multi-sensor fusion technology in autonomous driving.

**Keywords:** Multi-sensor fusion, autonomous driving, target detection, environment sensing, obstacle avoidance strategies.

## 1. Introduction

With the rapid development of autonomous driving technology, the accurate perception and localisation of the surrounding environment during driving is one of the hotspots in the field of autonomous driving research nowadays. However, it is often difficult for a single sensor to cope with complex and changing environments. For example, the performance of a camera decreases dramatically under low light or complex weather conditions, while LiDAR performs better under these conditions, but it is difficult to independently complete comprehensive environment sensing. To compensate for the shortcomings of a single sensor, multi-sensor fusion technology combines multiple sensors such as cameras, LIDAR, millimeter-wave radar, ultrasonic radar, etc. By fusing the data from these sensors, the sensing accuracy and robustness of the system can be significantly improved.

Currently, multi-sensor fusion technology has made great progress, and in terms of fusion level, three levels of data level fusion, feature level fusion and decision level fusion have been formed [1]. In terms of fusion algorithms, the weighted average method, Kalman filter method, Bayesian estimation method, and neural network method are formed [2]. For the decision level neural network fusion model for camera and LiDAR, Multi-View 3D Object Detection Network for Autonomous Driving (MV3D) model proposed by CHEN et al. in literature is a typical decision fusion model [3]. Jason Ku et al. introduced the Aggregate View Object Detection Aggregate View Object Detection (AVOD) model algorithm [4]. CenterFusion model algorithm presented by Nabati et al. is a detection network model algorithm that uses view cones to correlate camera and radar features [5]. The

Multimodal VoxelNet for 3D Object Detection (MVXNet) model algorithm introduced by Chen Youpeng is also a typical model algorithm that can process both graphics data and LiDAR data [6].

The aim of this paper is to explore the application of multi-sensor fusion technology in enhancing the performance of autonomous driving vehicles. This paper outlines the levels and algorithms of multi-sensor fusion by comparing the common sensors, analyses and summarizes the algorithms of decision-level neural network fusion model based on camera and LiDAR, and lists the examples of multi-sensor fusion algorithms applied in autonomous driving, and looks forward to the direction of the future development of the technology in view of the existing problems.

## 2. Multi-sensor fusion and algorithms

### 2.1. Different sensor characteristics

Different sensors have different positioning accuracies and there are different application and failure scenarios. Fig. 1 shows the performance of four sensors, LiDAR, millimeter wave radar, ultrasonic radar, and camera, in different aspects. Through comparative analysis, it can be seen that LiDAR is suitable for target detection over long distances and has a wide range of detection angles, but it is more costly and performs poorly in bad weather. Millimeter-wave radar speed measurement can directly measure the speed information of the target, which is especially suitable for medium and long-distance target detection, and is extremely adaptable to the environment, but the accuracy is not as good as that of LiDAR, and it is easily affected by the surrounding electromagnetic environment. Ultrasonic radar is only suitable for close-range target detection, and its advantages are low cost and strong adaptability to the environment. The camera has strong temperature stability and a wide range of detection angles, but its accuracy is low and its performance is poor in bad weather. Different sensors can play their own advantages in different use scenarios, and it is difficult to replace each other. Overall, no single sensor can cover all the application scenarios and requirements in autonomous driving. Therefore, multi-sensor fusion technology is used to integrate and process the data obtained from different sensors, thus improving the reliability and robustness of the system and solving the problems faced by the traditional single sensor.

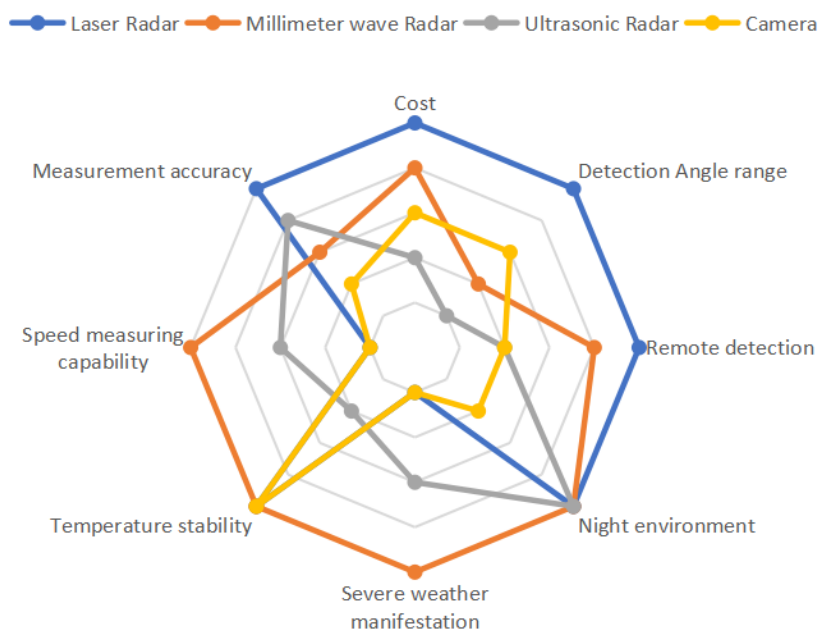


Figure 1. Multi-sensor performance comparison

### 2.2. Layers of sensor data fusion

Data fusion can usually be categorized into three tiers depending on the level of the processing stage. Data level fusion, which is derived from the information obtained by the sensors without any

processing, extracts features based on the fusion results and analyses them for decision-making. Feature level fusion, which first processes the data obtained from the sensors and derives the feature vectors of each key information for fusion. Decision-level fusion, which is characterized by the decision-making process of each sensor's own data, transmits the processed information to the fusion centre to complete the corresponding fusion process[7].

As can be seen from the comparative analysis diagram in Fig. 2, the advantage of data-level fusion is that there is basically no loss of the original data, so the accuracy is high, but because of the huge amount of data, which leads to poor fault-tolerance and anti-jamming ability in the fusion process, and the algorithm is complex. Compared with data-level fusion, feature-level fusion can reduce the amount of information processing, the complexity of the algorithm is lower, and the fault tolerance and anti-interference ability become stronger. While the decision-level fusion method has strong fault tolerance and anti-interference ability and simple algorithms, its accuracy is relatively low and the amount of information loss is large. Of course, feature-level fusion also has problems such as information loss, but considering the real-time and arithmetic requirements of autonomous driving scenarios, the feature-level fusion method has a moderate amount of computation, and it can better balance detection accuracy and computational efficiency. The feature information extracted from different sensors can be fused to improve the accuracy and robustness of the detection under the condition of high error tolerance.

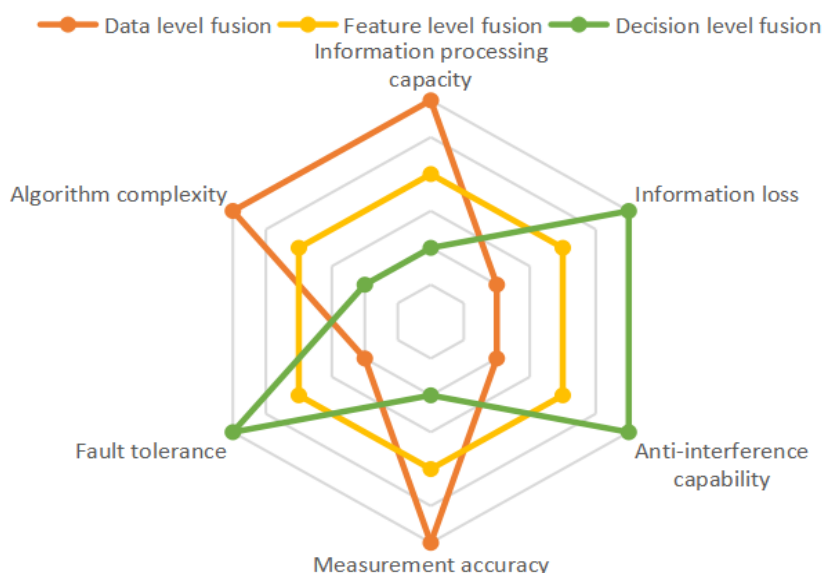


Figure 2. Multi-sensor fusion level performance comparison

### 2.3. Common information fusion methods

In the multi-sensor data fusion system, a large number of data fusion methods need to be used to effectively process various types of sensor data in the system. According to different application background requirements, there are various multi-sensor data fusion algorithms, and the main fusion methods are divided into two categories: classical methods and modern methods, of which the classical methods include a weighted average method, Kalman filter method, Bayesian estimation method, etc., while the modern methods mainly include neural network method.

The weighted average method is the simplest data fusion algorithm, which weights the sensor data to obtain the fusion result. Kalman filtering is also a widely used data fusion method, which is able to obtain a statistically unique optimal estimate in the case of a linear system with a Gaussian white noise model of the error. The Bayesian estimation method optimises the fused information by using a probability density function to represent the source information so that the sensor information is combined in a probabilistic relationship and its measurement uncertainty is expressed in terms of conditional probabilities. The neural network method converts data information into a knowledge base by using a large number of neurons with nonlinear mapping relationships, which can

automatically perform knowledge acquisition and associative reasoning, and is able to fuse uncertain and complex data into data that can be processed by the system through learning and reasoning [8].

The advantages, disadvantages and applicable scenarios of the four common fusion algorithms are shown in Table 1.

**Table 1.** Fusion Algorithm Comparison

Fusion Algorithms	Advantages	Disadvantages	Applicable Scenarios
Weighted arithmetic average	Simple to implement and low computationally, suitable for sensor data fusion with little difference in accuracy	Inability to exploit data time correlation and lack of robustness	Dynamic, linear systems
Kalman filtering	Suitable for dynamic systems, good real-time performance, provides estimation accuracy, computationally efficient	Poor performance on non-linear problems	Static, low precision demand systems
Bayesian estimation	Adaptable for non-linear systems, can handle partial data missing cases	Complex calculations and time inefficiencies	High precision, non-linear systems
Neural network method	Handling complex non-linear systems with high adaptability and generalisation capabilities	Large amount of training data and training time, complex network structure and parameter design	Complex non-linear, data-driven systems

#### 2.4. Decision-level neural network fusion algorithm based on camera and LiDAR

Sensor fusion algorithms are used to fuse the data obtained from multiple sensors to improve the accuracy and reliability of the sensing system. Sensor fusion algorithms are one of the very important techniques in the field of autonomous driving, robotics, drones, etc.

Based on the analysis in Section 2.1, by fusing two common complementary sensors, a camera and LiDAR, the shortcomings of a single sensor in different scenarios can be effectively dealt with, so as to improve the robustness and accuracy of target detection. Comparing the three different data fusion levels in Section 2.2, this paper chooses feature-level fusion as the final solution. The feature-level fusion method can fuse the feature information extracted from different sensors under the guarantee of high error tolerance, so as to improve the accuracy and robustness of detection. Comparing and analysing several information fusion methods mentioned in Section 2.3, this paper chooses the neural network method as the final solution, which can deal with the analysis of complex nonlinear systems with high adaptability and generalization ability. In summary, in this section, an example of a decision-level neural network fusion algorithm based on camera and LiDAR will be analyzed.

MV3D is a multimodal fusion model applied to 3D target detection for autonomous driving, which is able to simultaneously use multi-view images and radar point cloud data for object detection with high accuracy and robustness [9].

AVOD is a two-stage multi-sensor fusion algorithm for 3D object detection, which is capable of using both single image and point cloud data. AVOD algorithm is capable of detecting 3D objects more accurately by fusing the advantages of image and point cloud data, but it may be difficult to satisfy the requirements of smart traffic system's real-time requirements [10].

CenterFusion is an algorithm for target detection that fuses the prediction results of multiple detectors through deep neural networks to improve the accuracy and stability of detection, effectively suppressing false and missed detections and adapting to multi-scenario target detection, but the fixed number of detectors may not be able to fully satisfy the demand of roadside detection [11].

MVXNet is a classical neural network architecture for the simultaneous processing of camera and LiDAR data for more accurate and robust target detection. With parallel inputs and feature fusion,

MVXNet takes full advantage of both sensors. However, MVXNet also has shortcomings, such as its primitive network structure, which makes it difficult to cope with complex scenarios and may result in omission and false detection in scenarios with significant occlusion and illumination changes. In addition, due to the complex feature fusion method, MVXNet has a high demand for computational resources, which increases the cost and complexity of training and detection.

Fig. 3 shows the performance comparison of MV3D, AVOD, CenterFusion, and MVXNet in terms of accuracy robustness, computation amount, and real-time performance. Combined with the above brief introduction of the four algorithms, it can be seen that MV3D and AVOD enhance the detection accuracy by fusing the image and the point cloud data, but both of them have high computation amounts and are limited in complex scenes. CenterFusion, on the other hand, suppresses false and missed detections through multi-detector fusion, and has better real-time performance, while MVXNet achieves more accurate and robust target detection by simultaneously processing camera and LiDAR data through a two-branch structure, but is not the best in terms of real-time performance. In addition, CenterFusion and MVXNet are suitable for small target and long-distance target detection, but there may be some misjudgments for some objects, and there are leakage and misdetection problems in complex situations such as occlusion. Overall, these algorithms can improve the robustness of detection by using multimodal data, but the choice should be weighed against the demand for computational resources and the suitability of specific application scenarios.

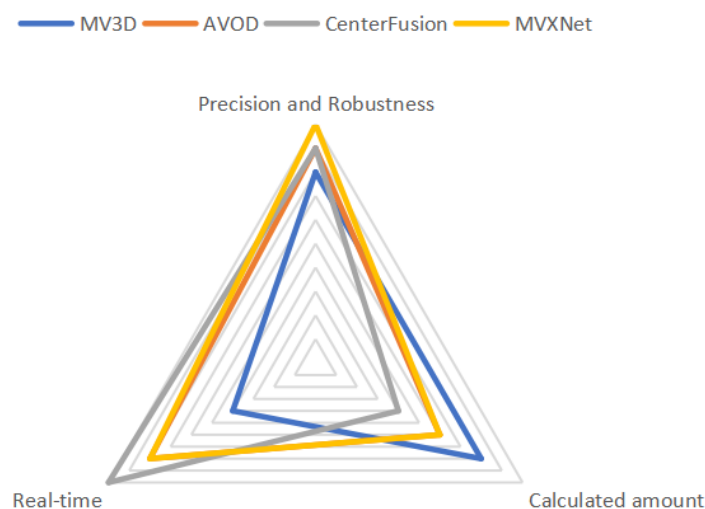


Figure 3. Comparison of neural network fusion algorithms

### 3. Applications of Multi-Sensor Fusion Algorithms in Autonomous Driving

#### 3.1. Obstacle Perception and Avoidance Strategies

Obstacle perception and obstacle avoidance are crucial aspects of autonomous driving technology, and are key capabilities to ensure safe and efficient vehicle operation. Automatic driving vehicles need to recognize and respond to possible static or dynamic obstacles, including pedestrians, vehicles, roadblocks, etc., in real time and accurately in complex road environments. This is not only directly related to driving safety, but also of great significance for improving transportation efficiency and user experience.

However, there are obvious limitations to relying on a single sensor for obstacle sensing. For example, cameras may not be able to provide reliable data under low light or strong light interference; LiDAR, although it has high-precision 3D spatial sensing capability, its performance will be affected under rain and snow; millimeter-wave radar, although it performs well under bad weather, is deficient in close-range obstacle recognition and object classification. These limitations reduce the perception accuracy of the system, which in turn affects the effectiveness of the obstacle avoidance strategy. To overcome the limitations of a single sensor, multi-sensor fusion becomes a key means to achieve

efficient obstacle sensing and obstacle avoidance. By fusing data from multiple sensors such as LIDAR, millimeter-wave radar, and cameras, the system is able to take full advantage of their respective strengths and realize complementary sensing data. For example, the combination of high-precision spatial data from LiDAR and the long-range target detection capability of millimeter-wave radar can significantly improve the accuracy of obstacle detection. At the same time, multi-sensor fusion also improves the robustness of the system so that it can still provide stable and reliable obstacle sensing results in complex and changing environments, leading to more efficient obstacle-avoidance strategies. This fusion technology not only improves the overall performance of the autonomous driving system but also provides a solid foundation for safe decision-making of the vehicle in different scenarios.

In terms of obstacle avoidance strategies, dynamic programming algorithms are commonly used for path planning to ensure smooth and real-time obstacle avoidance paths. Obstacle avoidance strategies based on safety potential field adjust the vehicle trajectory by setting the “potential field” for obstacles to avoid collision. In addition, the cost function obstacle avoidance strategy calculates the distance and dynamic characteristics of obstacles to select the optimal path to ensure smooth obstacle avoidance.

## 3.2. Case study

### 3.2.1. Real-Time Roadside Perception Based on LiDAR and Camera Fusion

Chen Youpeng [6] proposed a multi-sensor fusion method based on camera and LiDAR, aiming to improve the detection accuracy and speed of autonomous driving technology in roadside sensing scenarios. The method is specially designed for the detection of motorized vehicles, non-motorized vehicles and pedestrians in complex road environments, and combines the advantages of two types of sensors, camera, and LiDAR, to address the shortcomings of a single sensor in terms of real-time performance and accuracy. In order to ensure the accurate fusion of the two types of sensor data, Chen Youpeng designed a spatio-temporal synchronization processing algorithm to solve the problem of temporal and spatial desynchronization of camera and LiDAR data in dynamic environments. This spatio-temporal synchronization technique can effectively reduce the errors caused by sensor desynchronization, and improve the accuracy of the data and the stability of the overall system.

In terms of data fusion, Chen Youpeng adopts the channel attention mechanism and feature-level fusion technology to integrate the data from the camera and LIDAR into a multimodal feature input, giving full play to the camera's rich image texture and LIDAR's precise distance sensing capability. To improve the performance of the system under complex conditions such as occlusion or low light, the study also designs an improved target detection network structure, including a feature pyramid-based image detection network and a multi-branch LIDAR point cloud detection network. Experimental validation shows that the method exhibits high detection accuracy and real-time performance in road traffic scenarios, effectively improving the detection of small targets. With this scheme, the intelligent roadside platform is able to sense the road environment in real-time and accurately under multiple weather and lighting conditions, providing high-quality data support for the safety and decision-making of self-driving vehicles.

### 3.2.2. Adaptive Obstacle Avoidance with Multi-Sensor Fusion

The research work of Yue Wang [12] focuses on a multi-sensor fusion adaptive obstacle avoidance strategy that utilizes a combination of cameras, LiDAR, and millimeter wave radar to sense the environment. Based on the different conditions of the environment, the method dynamically adjusts the weights of each sensor and uses fuzzy reasoning to realize the nonlinear mapping between the sensor detection inputs and the fusion result outputs by building a probabilistic spatial knowledge base to achieve adaptive sensing. For example, under good lighting conditions, the camera is weighted higher to provide rich image information, while in weather such as rain, snow and fog, it relies more on LiDAR. The method is effectively validated in emergency obstacle avoidance scenarios, lane changing obstacle avoidance scenarios, and global planning scenarios, which not only improves the

response speed of the system in complex dynamic scenarios, but also enhances the accuracy of the obstacle avoidance decision, and is able to flexibly cope with unexpected obstacles in different scenarios.

### 3.2.3. Fusion of LiDAR and Millimeter-Wave Radar

Mingjie Li [13] et al. proposed an obstacle sensing method based on LiDAR and millimeter-wave radar, aiming to utilize the advantages of both to complement each other. The method deploys millimeter-wave sensor nodes through a uniform linear array and combines the LIDAR signal echo data to achieve time-frequency conversion of the signal and power spectral density feature extraction using matched filtering and wavelet multi-scale decomposition techniques. The signal and noise subspace separation methods, as well as Wigner-Vill distribution detection and multi-resolution feature clustering analysis, improve the response speed and optimize the accuracy of data fusion and obstacle sensing. Specifically, millimeter-wave radar is suitable for long-distance target detection and can provide stable distance information in bad weather such as rain and snow, while LiDAR has advantages in the accurate positioning of obstacles. By fusing data from these two sensors, the system not only improves the detection range of obstacles but also increases the accuracy and reliability of identification. This approach is applicable to a wide range of complex environments, especially under complicated road conditions, and can significantly improve the safety of autonomous driving systems.

## 4. Discussion

### 4.1. Challenge

Challenges faced by autonomous driving systems include data processing complexity, real-time requirements, adaptation to environmental changes, system robustness, and large-scale data processing and storage. In terms of data processing complexity, autonomous driving systems need to integrate data from different sensors, such as point cloud data generated by LIDAR and image data captured by cameras, which come in different formats and frequencies. A typical scenario is that in a busy city road, the system needs to identify multiple targets such as pedestrians, vehicles, traffic signs, etc. simultaneously, and for each of them different algorithms need to be applied to classify, localise and predict them, thus increasing the difficulty of data fusion. In addition, self-driving vehicles must process sensor data in a very short period of time (usually within tens of milliseconds) in order to react in a timely manner, especially when an obstacle is detected ahead. For example, when the vehicle is travelling at high speed on a motorway, the system needs to complete obstacle detection, distance calculation, and route adjustment within tens of milliseconds, which could lead to a collision if the delay is too long. The system also needs to work properly under various weather and light conditions, but fluctuations in sensor performance under different environmental conditions may affect the stability of the system. It is also a challenge to ensure that when a sensor fails, other sensors can supplement the missing information and ensure the safety and reliability of the system. Finally, multi-sensor systems generate a large amount of data, which places higher demands on data storage, transmission and processing capabilities.

### 4.2. Scheme

In order to overcome the challenges faced by multi-sensor fusion techniques and optimise their application, several options can be considered. Firstly, the efficiency and accuracy of data fusion can be improved by adopting advanced fusion algorithms such as Kalman filtering, particle filtering and deep learning. Second, the most appropriate sensor modalities are intelligently selected for data processing according to different environments and task requirements, thus reducing unnecessary data processing burdens. In addition, the establishment of a unified data format standard helps to simplify the fusion process of different sensor data. Edge computing is also an effective solution to improve system real-time and reduce latency by performing preliminary data processing at the sensor end and reducing the burden of data transmission to the central processing unit. Data preprocessing

is equally crucial as feature extraction, where data from various types of sensors are first processed to remove noise and extract key information to improve the efficiency and accuracy of the subsequent fusion process. Finally, through simulation and emulation tests in the virtual environment, the performance of the multi-sensor fusion algorithm in various complex scenarios can be verified before actual deployment to ensure its stability and reliability.

## 5. Conclusion

In this study, high-precision motion information acquisition of vehicle circumferential targets is achieved through the cooperative fusion of multi-source sensors, which provides reliable data support for autonomous driving environment sensing and path planning. In this paper, it is concluded that multi-sensor fusion can significantly improve the accuracy of target detection and tracking, and especially show better robustness under complex road conditions. However, sensor fusion still faces the challenges of real-time and computational complexity. Future research could focus on the development of efficient fusion algorithms to further enhance the real-time response capability of the system. In addition, with the improvement of sensor technology and computational capability, the application of deep learning of multi-source data will become an important development direction of the automatic driving perception system. Through the introduction of machine learning and deep learning techniques, algorithms can be trained on large-scale datasets to improve the robustness of target detection and recognition in uncertain environments.

## Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

## References

- [1] Huang Mangguo, Fan Shangchun, Zheng Dezhi, et al. Research progress of multi-sensor data fusion technology. *Sensors and Microsystems*, 2010, 29 (3): 5 - 8.
- [2] Wang X. Research on multi-sensor data fusion problem [D]. Changchun: Jilin University, 2006.
- [3] Chen XZ Ma HM. Wan J. Multi-view 3D object detection network for autonomous driving [CV// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. New York: IEEE, 2017: 6526 - 6534.
- [4] Jason Ku, Melissa Mozifian, Jungwook Lee. Joint 3D proposal generation and object detection from view aggregation [1. 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2018: 1 - 8. and camera fusion for 3d object.
- [5] Nabati R, Qi H. Centerfusion: Center-based radar detection [C]// Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2021: 1527 - 1536.
- [6] Chen Youpeng. A multi-sensor fusion roadside sensing method based on camera and LiDAR [D]. Beijing Jiaotong University, 2023.
- [7] Gaoxiang Wang. Research on vehicle circumferential multi-source sensing target motion information acquisition for autonomous driving [D]. Jilin University, 2023.
- [8] Zhou P. Research and prospect of multi-sensor data fusion technology. *Internet of Things Technology*, 2015, 5 (05): 23 - 25.
- [9] SHA Junliang, CAO Jingsheng, DONG Yining, et al. Research on improved MV3D target detection model using hybrid fusion. *Modern Electronic Technology*, 2024, 47 (19): 175 - 180.
- [10] Yao Lu, Wu Shaobo. Multi-sensor fusion vehicle detection based on optimised AVOD algorithm. *Journal of Beijing University of Information Science and Technology (Natural Science Edition)*, 2022, 37 (02): 76 - 80.
- [11] J. Huang, J. S. Liu. 3D target detection model for autonomous driving based on improved Centrefusion. *Radio Engineering*, 2024, 54 (02): 507 - 514.

- [12] Wang Yue. Research on Obstacle Avoidance Strategies for Autonomous Driving Vehicles Based on Sensor Fusion [D]. Tianjin University, 2020.
- [13] LI Mingjie, LIU Xiaofei. An obstacle sensing method for autonomous driving based on fusion of millimetre wave sensors and lidar signals. *Laser Journal*, 2023, 44 (6): 199 - 203.