10.18686/frim.v2i6.4686

# Research on Multi-Object Classification and Recognition Methods for Intelligent Vehicles Based on Deep Learning

**Ling Wu, Qiong Wu**

**College of Information Engineering, East University of Heilongjiang, Harbin, Heilongjiang 150066**

***Abstract:*** With the development of intelligent transportation systems, multi-object classification and recognition for intelligent vehicles have become a key research direction in the fields of autonomous driving and intelligent transportation. Deep learning-based methods have become the mainstream approach to solving this problem. This paper reviews the deep learning-based multi-object classification and recognition techniques, analyzes the advantages and disadvantages of current mainstream methods, and discusses future research directions.

***Keywords:*** Deep learning; Intelligent vehicles; Multi-object classification; Autonomous driving; Convolutional neural networks

## 1. Introduction

In recent years, intelligent vehicles and autonomous driving technologies have made significant progress, becoming an important component of modern transportation systems. In intelligent vehicle systems, perception of the surrounding environment is fundamental for achieving autonomous driving, and multi-object classification and recognition are among the most critical tasks. The core of multi-object classification and recognition technology is accurately identifying various objects such as vehicles, pedestrians, and traffic signs and processing them in real time to ensure driving safety. With the rapid development of deep learning technology, methods based on deep learning have shown great potential in multi-object classification and recognition.

This paper aims to systematically review the methods for multi-object classification and recognition in intelligent vehicles based on deep learning, analyze the strengths and weaknesses of existing methods, and propose future research directions.

## 2. Application of Deep Learning in Multi-Object Classification and Recognition

### 2.1 Overview of Deep Learning

Deep learning is a machine learning method based on artificial neural networks, with the ability to automatically learn data features. Compared with traditional machine learning methods, deep learning does not require complex feature engineering and can automatically learn effective feature representations through large amounts of data and computational resources. In multi-object classification and recognition tasks, deep learning methods mainly adopt architectures such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs).

### 2.2 Application of Convolutional Neural Networks in Multi-Object Classification

Convolutional Neural Networks (CNNs) are the mainstream architectures for image classification and recognition tasks. They can progressively extract hierarchical features of images through convolutional layers and pooling layers, thus achieving complex image recognition tasks. In multi-object classification tasks, CNNs can automatically learn features ranging from simple to complex, making them suitable for different types of objects. For example, deep convolutional neural network models such as AlexNet, VGGNet, ResNet, and InceptionNet have achieved significant classification results on large datasets like ImageNet and are widely applied in tasks such as vehicle recognition, pedestrian detection, and traffic sign recognition.

### 2.3 Application of Recurrent Neural Networks and Their Variants in Time-Series Data Processing

For multi-object recognition in intelligent vehicles, environmental data include not only static images but also video data. Recurrent Neural Networks (RNNs) and their variants, such as Long Short-Term Memory (LSTM) networks and Gated Recurrent Units (GRUs), have shown good performance in processing time-series data. In video-based multi-object recognition, RNNs can handle sequential dependencies, improving the temporal correlation of recognition.

## 3. Deep Learning-Based Multi-Object Classification and Recognition Methods for Intelligent Vehicles

### 3.1 Current Status of Object Detection and Classification

Object detection and classification are two key steps in multi-object recognition. Before the rise of deep learning methods, traditional object detection methods relied heavily on sliding windows, HOG features, and SVM classifiers, but these methods performed poorly in complex scenarios. The emergence of deep learning-based methods such as Faster R-CNN, YOLO (You Only Look Once), and SSD (Single Shot MultiBox Detector) has greatly improved the accuracy and speed of object detection.

### 3.2 Single-Stage and Two-Stage Object Detection Methods

Deep learning-based object detection methods are generally divided into single-stage and two-stage methods. Single-stage methods (such as YOLO and SSD) directly predict the location and category of objects, achieving good real-time detection performance; while two-stage methods (such as Faster R-CNN) generate candidate regions and perform classification and regression on these regions, generally providing better detection accuracy.

### 3.3 Application of Multi-Task Learning and Ensemble Learning in Multi-Object Recognition

To further improve the accuracy and robustness of multi-object recognition, researchers have proposed multi-task learning and ensemble learning methods. In a multi-task learning framework, the model is trained not only to learn object detection and classification tasks but also to simultaneously learn other related tasks (such as segmentation and tracking), thereby enhancing the comprehensive understanding of the environment. Ensemble learning combines the predictions of multiple models to reduce potential errors from a single model.

## 4. Challenges and Countermeasures in Intelligent Vehicle Environments

### 4.1 Scene Complexity and Dataset Bias Issues

Multi-object classification and recognition tasks for intelligent vehicles in real environments face numerous challenges, among which scene complexity and dataset bias are two key factors affecting model performance. These issues directly impact the generalization ability and robustness of recognition systems, limiting the effectiveness of intelligent vehicles in complex and variable environments.

Vehicles encounter various scenes during driving, including urban roads, highways, rural roads, and mountainous areas. Each scene has its unique characteristics, such as high-density traffic and mixed pedestrian and non-motorized vehicles in urban roads, high-speed vehicles with fewer pedestrians on highways, and complex terrains and sudden obstacles on rural roads. Moreover, different weather conditions (such as sunny, rainy, snowy, and foggy days) and illumination variations (such as daytime, dusk, and nighttime) can significantly affect the visual characteristics of the environment, directly impacting the visibility of objects and model recognition performance. For example, in foggy weather, distant vehicles and pedestrians may become blurred; under nighttime or strong backlight conditions, objects may be completely hidden in shadows. For intelligent vehicles to maintain stable recognition capabilities in various scenarios, the model must possess strong robustness and adaptability. Existing traffic datasets for training deep learning models, such as KITTI, COCO, and Cityscapes, are often collected under specific environments and conditions. These datasets may cover the characteristics of some scenes and object types but cannot comprehensively cover all possible situations that intelligent vehicles may encounter in the real world. For example, some datasets may primarily contain urban road scenes, lacking samples of rural roads, highways, or mountainous environments; some datasets may be collected under clear daytime conditions, lacking data for extreme weather and lighting conditions such as rainy, snowy, or nighttime scenarios. Due to the limitations of dataset collection, deep learning models trained on them may perform poorly and even fail to work correctly when facing unseen environments and situations. This dataset bias issue leads to poor generalization ability in real-world applications, even if the model performs well on the test set.

### 4.2 Computational Resource Limitations and Real-Time Requirements

In the multi-object classification and recognition system for intelligent vehicles, computational resource limitations and real-time requirements are two closely related challenges. The environmental perception system of an intelligent vehicle needs to detect, classify, and track various objects (such as vehicles, pedestrians, and traffic signs) in the external environment in real time. This requires algorithms to complete complex computational tasks in an extremely short time to ensure driving safety and comfort. However, the hardware computing resources that intelligent vehicles can be equipped with are limited, making it a challenging task to achieve efficient real-time processing with limited computational resources.

The limitations of computational resources in intelligent vehicles are mainly reflected in two aspects: the computing power of processors and the capacity of memory. Unlike data centers or high-performance computers, computing devices in intelligent vehicles are usually constrained by space and power consumption, making it difficult to install large GPUs or TPUs. Meanwhile, the edge computing units in intelligent vehicles need to process large amounts of image and sensor data under controlled power consumption. Such hardware limitations make it difficult to deploy large-scale deep neural network models directly in onboard systems, especially those that need to process high-resolution images and complex scenes. Therefore, how to reduce computational overhead and energy consumption while ensuring model recognition accuracy has become an urgent problem to solve.

### 4.3 Issues of Anomaly Detection and Small Object Detection

Anomaly detection and small object detection are two challenging aspects of multi-object classification and recognition tasks in intelligent vehicles, directly affecting the vehicle's ability to handle complex driving environments and safety. In real driving scenarios, intelligent vehicles need to not only accurately recognize regular traffic objects (such as vehicles, pedestrians, and traffic signs) but also quickly and accurately detect and classify anomalies and small objects to ensure driving safety and stability.

Anomaly detection refers to the ability of intelligent vehicles to quickly and accurately recognize and judge unexpected or uncommon scenarios. For example, on urban roads, situations such as a suddenly appearing pedestrian, an animal crossing the road, dropped cargo, or another vehicle's emergency lane change all fall under anomalies. These situations are often instantaneous and irregular, making them difficult to handle effectively through conventional object detection methods. Anomaly detection requires intelligent vehicles to have strong environmental perception capabilities and rapid response capabilities; otherwise, it may lead to traffic accidents. Traditional object detection algorithms are generally optimized based on training data in regular scenarios. Due to the uncertainty and low frequency of anomalies, it is difficult to fully represent and cover them in training data. Therefore, intelligent vehicles often show poor robustness and response capabilities when facing these "long-tail" events, leading to increased risks of missed or false detections.

## 5. Conclusion

Multi-object classification and recognition methods for intelligent vehicles based on deep learning represent a research field with broad application prospects. Although significant progress has been made with existing methods, challenges remain in addressing complex environments, meeting real-time requirements, and handling computational resource constraints. Future research should focus more on methodological innovation and enhancing practical application performance to further advance autonomous driving technology.

## References


[1] Liu, Hui. Research on Vehicle Type Recognition Algorithm Based on Deep Learning [D]. Qingdao University of Technology, 2018.

[2] Wu, Chen. Multi-Target Vehicle Detection and Tracking Methods Based on Deep Learning [J]. Practical Automotive Technology, 2023, 48(4):14-17.

[3] Feng, Jihao. Research on DCT Vehicle Intelligent Control Considering Driving Behavior and Driving Environment [D]. Chongqing University, 2022.

[4] He, Danni. Research on Multi-Vehicle Detection and Tracking Algorithm Based on Deep Learning [D]. Dalian University of Technology, 2019.


**Acknowledgments:** This work is supported by the project "Research on Target Recognition Methods for Autonomous Driving in Open Environments" (Heilongjiang Provincial Natural Science Foundation, 2023, Project Number: LH2023F041).