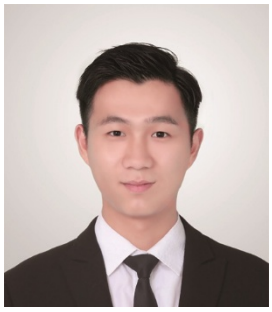
УДК 004.021

VEHICLE AND LABEL DETECTION METHOD BASED ON *YOLOV11*



Guo Qicheng
PhD student at the department of Infocommunication Technologies BSUIR
2733064287@qq.com

Guo Qicheng
PhD student at the Department of Infocommunication Technologies, BSUIR. The field of scientific interest is related to the study of computer vision, image processing, and SLAM (Simultaneous Localization and Mapping).

Abstract. This paper proposes a vehicle and label detection method based on the *YOLOv11* algorithm. By constructing a custom dataset of vehicles and labels, the *YOLOv11* model is trained to achieve precise detection of vehicles and their rear-mounted labels. This paper details the dataset creation process, label design, model training procedure, and the selection of optimal training parameters. Finally, the algorithm's performance is evaluated using a validation set. Experimental results demonstrate that the *YOLOv11*-based vehicle and label detection method exhibits strong performance in terms of accuracy and real-time capability, meeting the requirements of practical applications.

Keywords: Object Detection, *YOLOv11*, Vehicle Detection, Label Recognition.

Introduction. With the rapid advancement of computer vision technology, object detection has been widely applied in fields such as autonomous driving, robotics, and surveillance. In these applications, accurately identifying moving objects and their associated labels is crucial. To achieve precise detection of vehicles and their rear labels, this paper proposes a vehicle and label detection method based on the *YOLOv11* (You Only Look Once Version 11) algorithm.

The *YOLO* series of algorithms [1] has achieved remarkable success in object detection, particularly in real-time detection and efficiency, making it widely adopted. Compared to earlier versions, *YOLOv11* [2] has demonstrated significant improvements in detection speed, accuracy, and scalability. Therefore, this algorithm is selected as the core detection framework for this study.

Dataset Construction. To train the *YOLOv11* model, this study first requires the creation of a custom dataset containing images of small vehicles and their corresponding labels. The dataset construction process involves the following steps:

1)Image Acquisition of Vehicles and Labels

The dataset is built using a four-wheeled experimental vehicle model as the primary subject. Photographs of the vehicle and its rear-mounted labels were captured. During the image acquisition process, care was taken to ensure diversity in the dataset by varying perspectives and distances. The label design incorporates easily recognizable visual features, with red rectangular frames selected as the label markers. This study constructed a dataset comprising 660 images. Figure 1 shows a subset of the dataset images featuring the vehicle and its labels.

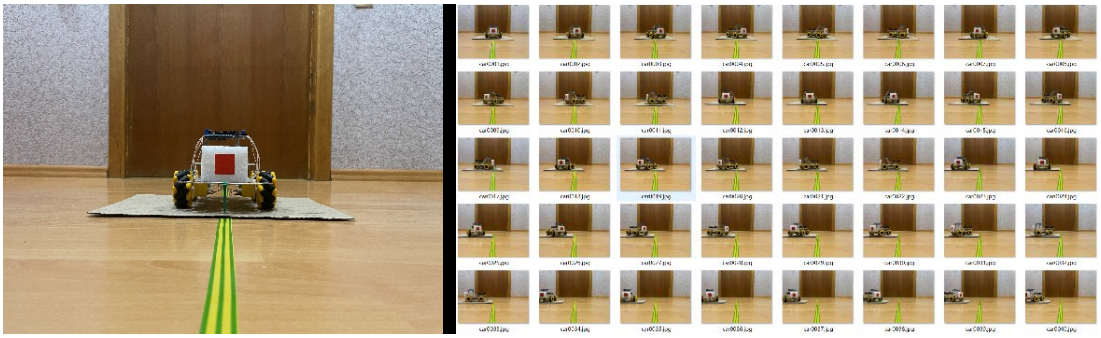


Figure 1. Partial dataset images of the vehicle and its corresponding labels.

2)Label Annotation

The captured images were manually annotated using the labeling tool (LabelImg). In each image, the bounding boxes of the vehicles and the positions of the rear labels were annotated. Each label category was associated with its corresponding vehicle to ensure the accuracy of the data annotation. After the annotation process, *YOLO*-format label files were generated, which contain the category information of each object and the coordinates of the bounding boxes.

3) Dataset Partition

The annotated dataset needs to be divided into training, validation, and test sets. Typically, the training set accounts for 70%-80% of the dataset, the validation set for 10%-15%, and the test set for 10%-15% [3]. The partitioning of the dataset should ensure a uniform distribution of samples across all categories to avoid overfitting or underfitting of the model. In this study, the dataset was partitioned in an 8:1:1 ratio.

***YOLOv11* Algorithm Overview.** *YOLOv11* (You Only Look Once Version 11) is the latest iteration in the *YOLO* series of algorithms. It achieves object detection through a single forward pass, simultaneously performing category prediction and location regression, demonstrating highly efficient detection capabilities. *YOLOv11* incorporates state-of-the-art deep learning techniques, such as efficient convolutional networks and more advanced loss functions, significantly improving speed while maintaining detection accuracy.

The *YOLOv11* network architecture primarily consists of three components: the backbone network, the detection head, and the output layer. The backbone network is responsible for feature extraction, while the detection head generates category predictions and bounding box regression values through convolutional operations. By performing a single forward pass on the input image, *YOLOv11* can simultaneously output object categories and locations, greatly enhancing detection efficiency. *YOLOv11* has the following advantages:

– efficiency – *YOLOv11* achieves object detection in a single forward pass, making it significantly faster compared to traditional two-stage detection methods;

– accuracy – By integrating various advanced techniques, *YOLOv11* improves the model's accuracy and robustness;

– scalability – *YOLOv11* is adaptable to a wide range of hardware environments, enabling efficient operation across different devices.

Experiments and Results Analysis. The *YOLOv11* framework was employed to train the model for vehicle and label detection. During the training process, hyperparameters such as learning rate, batch size, and weight decay were adjusted, and data augmentation techniques (e.g., rotation, flipping, brightness adjustment) were applied to further enhance the model's generalization capability. Through iterative experiments, appropriate hyperparameters were selected to optimize the training performance. The trained model was evaluated using validation and test sets. Evaluation metrics included mean Average Precision (mAP), recall, and F1-score. By comparing the performance of models under different training parameters, the optimal training strategy was determined.

After multiple rounds of training, the final model achieved a mAP of over 87% on the validation set, with excellent performance in both recall and precision. Figure 2 illustrates the training results. The results on the test set also demonstrated that the *YOLOv11* model can accurately detect vehicles and their rear labels, with detection speed meeting real-time requirements. Figure 3 presents the prediction results.

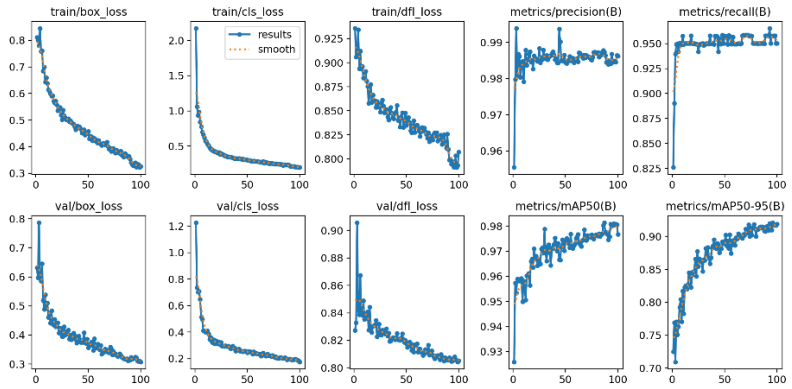


Figure 2. Training results

Due to the diversity of the dataset and the data augmentation techniques employed during the training process, the *YOLOv11* model has demonstrated strong robustness, enabling it to effectively handle vehicle detection tasks under varying lighting conditions and perspectives.

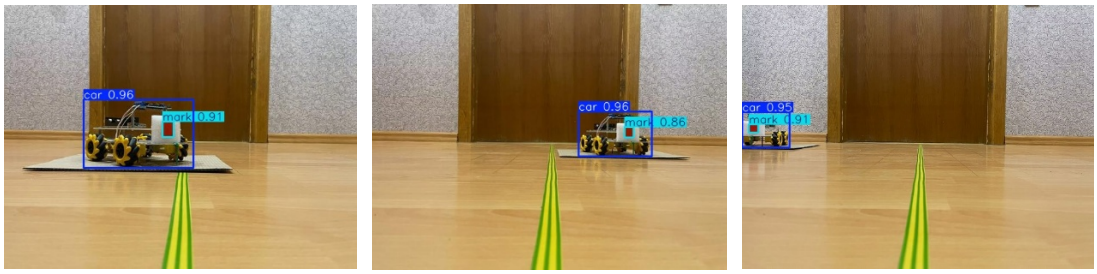


Figure 3. Prediction results

Conclusion. The proposed method for vehicle and label detection based on *YOLOv11*, through training and optimization on a custom dataset, has achieved promising experimental results. The efficiency and accuracy of *YOLOv11* in object detection enable this method to be effectively applied to real-world vehicle detection tasks. Future work will focus on further optimizing model performance and expanding the dataset to enhance adaptability in more complex environments.

References

[1] Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 779-788.

[2] Ultralytics. (2025). YOLOv11 Documentation. Retrieved from https://docs.ultralytics.com/zh/models/yolo11/

[3] Chalapathy, Raghavendra, and Sanjay Chawla. Deep learning for anomaly detection: A survey. IEEE Transactions on Neural Networks and Learning Systems 30.9 (2019): 1-18.