Article

Road Car Image Target Detection and Recognition Based on YOLOv8 Deep Learning Algorithm

Hao Wang [*] , Zhengyu Li , Jianwei Li

Article

Road Car Image Target Detection and Recognition Based on YOLOv8 Deep Learning Algorithm

Hao Wang [1,*], Zhengyu Li [2] and Jianwei Li [3]

[1] Beijing Yuandian Technology Inc., Beijing, 102200, China.

[2] Nio Inc., Shanghai, 201805, China.

[3] Independent Researcher, Beijing, 102200, China.

* Correspondence: whdawn1006@gmail.com.

Abstract: In this paper, target detection of car images in roads is performed based on the YOLOv8 model of YOLO family of models, which improves the accuracy and generalisation of the target detection task by combining multi-scale prediction, CSPNet structure and optimisation techniques such as BoF and BoS. The input images contain five types of vehicles such as Ambulance, Bus, Car, Motorcycle and Truck, which are analysed and learnt to have a classification accuracy of 75.4% on Ambulance, 53.5% on Bus, 55.1% on Car, 51.1% on Motorcycle and 42.5% on Truck. Despite the gap in specific classification accuracy, the YOLOv8 model can detect 100% of vehicles on the road, demonstrating good target detection capability. This research is of great significance for improving road traffic safety, intelligent traffic management, and the development of future autonomous driving technology. By optimising the deep learning model to achieve more accurate and efficient vehicle target detection, it can help to improve road safety and traffic efficiency, and promote the progress of intelligent transportation systems.

Keywords: YOLOv8; Deep learning

1. Introduction

In this paper, experiments are conducted using a car target inspection dataset that contains images of cars in various roads and even images of cars in paper content and advertisements, with the labels of the car parts recorded and saved in each car image. In addition, besides cars, the dataset also includes images of various types such as motorbikes and buses, which are used for the inspection process to achieve the classification of car types. The dataset can be used to train a deep learning model for target detection and for validating the detection effect of the model, we selected four images from the dataset and the results are shown in Figure 1.
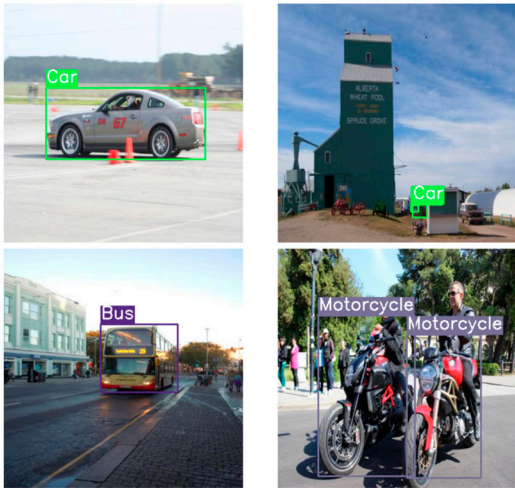

Figure 1. Data Set Presentation. (Photo credit : Original)**.**

2. Related Work on the YOLO Model

YOLO (You Only Look Once) is a popular real-time target detection algorithm that has evolved through several versions.

YOLOv1 was proposed in 2015 by Joseph Redmon et al. It is the first YOLO version.YOLOv1 uses a single neural network for end-to-end target detection, dividing the image into grids and predicting bounding box and category probabilities in each grid. Although fast and simple, it suffered from poor localisation accuracy and difficulty in detecting small targets.

YOLOv2 was introduced in 2016, introducing improvements such as multi-scale prediction and Batch Normalisation to improve performance and support detection of more categories. This version also introduced the Darknet-19 network structure.

YOLOv3 was released in 2018 and improves target detection performance by using a deeper and broader Darknet-53 network structure as well as Feature Pyramid Networks (FPNs). In addition, target detection at different levels allows the model to predict objects of different sizes simultaneously.

YOLOv4, released by Alexey Bochkovskiy in 2020, introduced the CSPDarknet53 backbone network, the Mish activation function, and various technical optimisations to improve accuracy and speed. This version also added Bag of Freebies and Bag of Specials methods to improve training results.

YOLOv5, released in 2020 by Ultralytics, is implemented in PyTorch and provides an easy-to-use API interface.YOLOv5 further improves speed and accuracy by optimising the model structure and training strategy while maintaining high performance.

YOLOv6, released by Alexey Bochkovskiy in 2021, continues to optimise model structure and training techniques, further improving performance and adding support for new features.

YOLOv7 was released in 2023, followed by YOLOv8 in 2024.

3. Method

YOLOv8 is a deep learning model based on target detection.The YOLOv8 model is continuously optimised by Joseph Redmon and Ali Farhadi et al. on the basis of the YOLO family of models, aiming at faster and more accurate target detection [7].The network structure of the YOLOv8 model is shown in Figure 2.
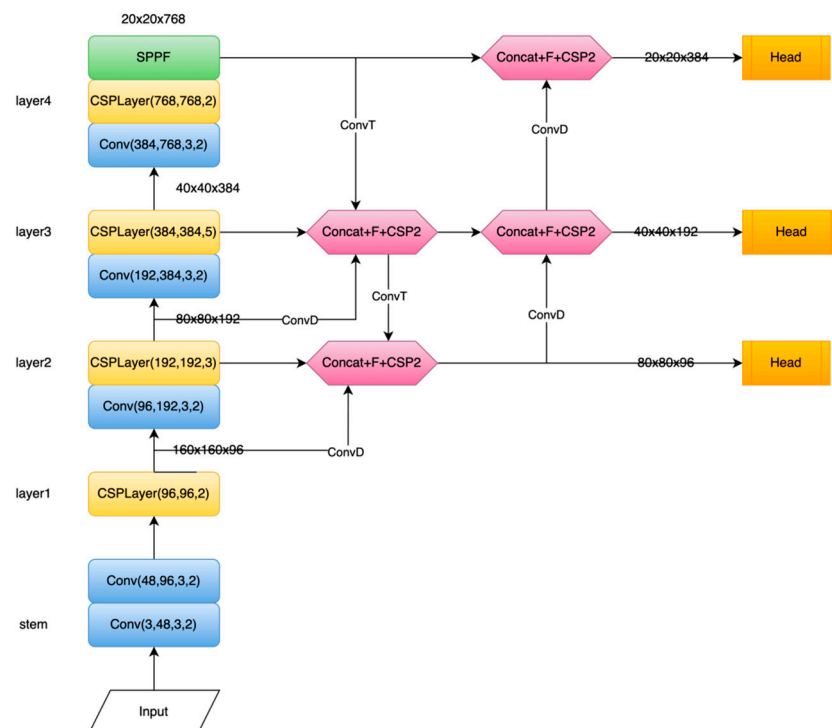

Figure 2. The network structure of the YOLOv8 model. （Photo credit : Original）.

Firstly, YOLOv8 adopts the backbone network Darknet-53 as a feature extractor.Darknet-53 is a deep neural network containing 53 convolutional layers with strong feature extraction capability. With the Darknet-53 network, YOLOv8 can extract rich feature information from the input image, which lays the foundation for the subsequent target detection task [8].

Second, YOLOv8 adopts a multi-scale prediction strategy. This strategy consists of downsampling operations on the input image at different scales and predicting the target frame and category information at each scale. This multi-scale prediction approach enables YOLOv8 to maintain high accuracy in detecting both small-size and large-size targets, effectively improving the generalisation ability of the model [9].

In addition, YOLOv8 introduces the Cross Stage Partial Network (CSPNet) structure to enhance feature fusion and information transfer.The CSPNet structure divides the feature maps in the middle layer of the backbone network into two parts and achieves feature fusion through cross-stage connections, which improves the model's efficiency of utilising feature information of different scales, and alleviates the gradient disappearance problem to a certain extent [10]. the gradient vanishing problem [10].

In addition, YOLOv8 also adopts Bag of Freebies (BoF) and Bag of Specials (BoS) techniques to further optimise the model performance.BoF techniques include conventional training techniques such as data augmentation and learning rate tuning, while BoS techniques include innovative methods such as SAM optimiser and PANet attention mechanism, which help to improve model performance on target detection tasks.

Overall, the YOLOv8 model achieves improved accuracy and generalisation for the target detection task while maintaining the speed advantage by combining multi-scale prediction, CSPNet structure, and optimisation techniques such as BoF and BoS.

4. Result

The input images contain a total of five types of vehicles, namely ambulance, bus, car, motorbike and truck, each type contains 250 images each for training and testing.The training and testing sets are divided according to the ratio of 7:3 and the output data is used for training. The YOLOv8 summary is output during the training process and the results are shown in Table 1.

Table 1. The YOLOv8 summary.

Class	Images	Instances	P	R	mAP50	mAP50-95
Ambulance	250	64	0.754	0.67	0.742	0.603
Bus	250	46	0.535	0.652	0.586	0.461
Car	250	238	0.551	0.345	0.35	0.235
Motorcycle	250	46	0.511	0.454	0.461	0.297
Truck	250	60	0.425	0.3	0.341	0.202

From the results, it can be seen that YOLOv8 has a classification accuracy of 75.4% on Ambulance, 53.5% on Bus, 55.1% on Car, 51.1% on Motorcycle, and 42.5% on Truck, although the classification on vehicle type is low , but YOLOv8 can detect 100% of vehicles on the road and can achieve good target detection.

A randomly selected test set is tested to test the detection effect of YOLOv8, and the results are shown in Figures 3–5.

Figure 3. Test Set Results. （Photo credit : Original）**.**

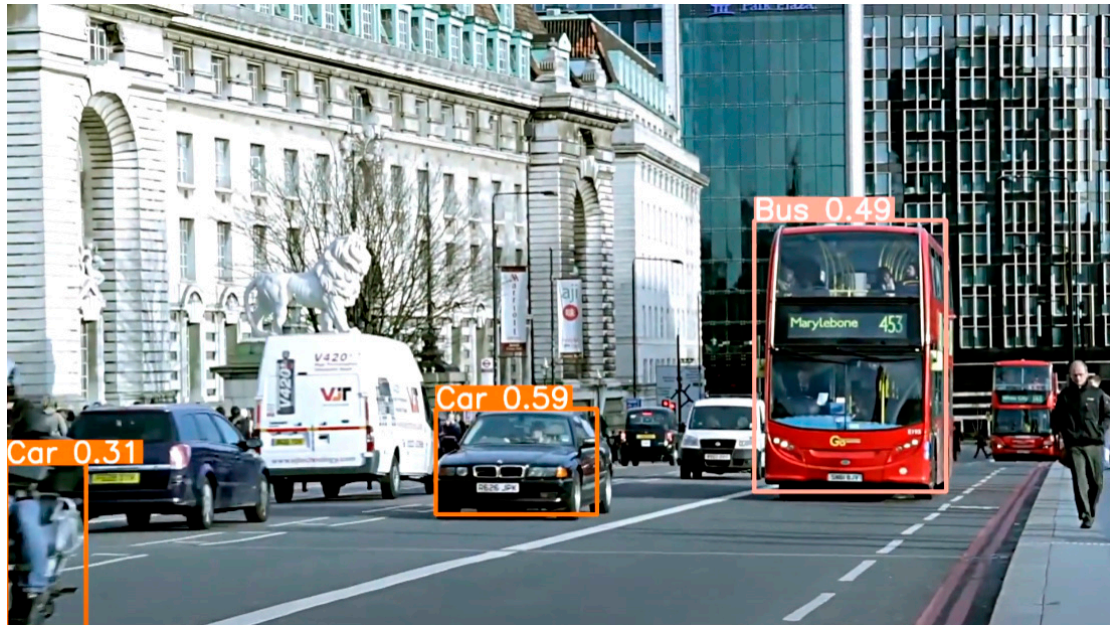**Figure 4.** Test Set Results. （Photo credit : Original）**.**

Figure 5. Test Set Results. （Photo credit : Original）**.**

5. Conclusion

This paper is based on the YOLOv8 model for target detection of car images in roads. By combining multi-scale prediction, CSPNet structure, and optimisation techniques such as BoF and BoS, the model improves the accuracy and generalisation of the target detection task while maintaining the speed advantage. In our experiments, we input a total of 1250 images of five types of vehicles including ambulances, buses, cars, motorbikes and trucks, and analysed the detection results.

The results show that YOLOv8 achieves 75.4% accuracy on ambulance classification, 53.5% on bus classification, 55.1% on car classification, 51.1% on motorbike classification and 42.5% on truck classification. Although the classification accuracy on specific vehicle types is lacking, it is worth noting that YOLOv8 is able to detect 100% of all types of vehicles on the road, providing strong support for road traffic monitoring and intelligent driving systems.

Taken together, the results show that the YOLOv8 model performs well in the overall target detection task, although there is still room for improvement in segmented category recognition. Its efficiency and comprehensiveness make it an important tool for road safety management, traffic flow monitoring, and other fields. In the future, the model's performance on category-specific recognition can be further improved by increasing the amount of training data, adjusting the network structure, or introducing more advanced techniques to better meet the needs of practical applications.

In conclusion, this study provides a useful reference for research and practice in related fields by providing an in-depth analysis of the application of the YOLOv8 model in road vehicle target detection. The model not only shows good detection capability, but also lays a solid foundation for the development of future intelligent transport systems.

References

1. Soylu, Emel, and Tuncay Soylu. "A performance comparison of YOLOv8 models for traffic sign detection in the Robotaxi-full scale autonomous vehicle competition." Multimedia Tools and Applications 83.8 (2024): 25005-25035.
2. Shyaa, Tahreer Abdulridha, and Ahmed A. Hashim. "Superior Use of YOLOv8 to Enhance Car License Plates Detection Speed and Accuracy." Revue d'Intelligence Artificielle 38.1 (2024).
3. Chopade, Rohan, et al. "Automatic Number Plate Recognition: A Deep Dive into YOLOv8 and ResNet-50 Integration." 2024 International Conference on Integrated Circuits and Communication Systems (ICICACS). IEEE, 2024.

4. Li, Wei, Mahmud Iwan Solihin, and Hanung Adi Nugroho. "RCA: YOLOv8-Based Surface Defects Detection on the Inner Wall of Cylindrical High-Precision Parts." Arabian Journal for Science and Engineering (2024): 1-19.

5. Safran, Mejdl, Abdulmalik Alajmi, and Sultan Alfarhood. "Efficient Multistage License Plate Detection and Recognition Using YOLOv8 and CNN for Smart Parking Systems." Journal of Sensors 2024 (2024).

6. Bakirci, Murat. "UTILIZING YOLOv8 for ENHANCED TRAFFIC MONITORING in INTELLIGENT TRANSPORTATION SYSTEMS (ITS) APPLICATIONS." Digital Signal Processing (2024): 104594.

7. Singh, Ripunjay, et al. "Evaluating the Performance of Ensembled YOLOv8 Variants in Smart Parking Applications for Vehicle Detection and License Plate Recognition under Varying Lighting Conditions." (2024).

8. Khan, Israt Jahan, et al. "Vehicle Number Plate Detection and Encryption in Digital Images Using YOLOv8 and Chaotic-Based Encryption Scheme." 2024 6th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT). IEEE, 2024.

9. Li, Shaojie, et al. Utilizing the LightGBM Algorithm for Operator User Credit Assessment Research. 2024.

10. Ma, D.; Li, S.; Dang, B.; Zang, H.; Dong, X. Fostc3net:A Lightweight YOLOv5 Based On the Network Structure Optimization; 2024.

11. Wang, Hao, Jianwei Li, and Zhengyu Li. "AI-Generated Text Detection and Classification Based on BERT Deep Learning Algorithm." arXiv preprint arXiv:2405.16422 (2024).

12. Lai, Songning, Ninghui Feng, Haochen Sui, Ze Ma, Hao Wang, Zichen Song, Hang Zhao, and Yutao Yue. "FTS: A Framework to Find a Faithful TimeSieve." arXiv preprint arXiv:2405.19647 (2024).

13. Chen, Yuyan, et al. "Emotionqueen: A benchmark for evaluating empathy of large language models." arXiv preprint arXiv:2409.13359 (2024).