International Journal of Science and Engineering Applications
Volume 12-Issue 03, 40 – 42, 2023, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1203.1015

A Car Target Detection Method based on YOLOv7-tiny

Wenlong Wang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Bowen Shi
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Xiaoyuan Wang
College of Electromechanical Engineering
Qingdao University of Science and Technology
Qingdao, China

Abstract: If the driver can discover the safety hazards during the driving and brake the vehicle in time, it will effectively reduce the occurrence of traffic accidents and reduce losses. In order to accurately and quickly detect the vehicle in front, this article proposes a car target detection method based on the YOLOv7-tiny algorithm. This method can use the machine vision to detect the vehicle in front. Provide security assistance. The experimental results show that the average accuracy rate of the methods proposed in this article on six common vehicle recognition reached 80.8%, and the model is more lighter.

Keywords: computer vision; deep learning; target detection; image identification

1.INTRODUCTION

With the rapid development of economy, society and science and technology, gradually improvement of people's living standard, the total quantity of vehicles in our country has increased rapidly. In comparison, the highway mileage in our country has increased slowly in recent years, which has led to the traffic contradiction between supply and demand becoming increasingly serious, and the aggravating contradiction between supply and demand leads to a large number of traffic accidents. According to the China Statistical Yearbook, a total of 244,674 traffic accidents occurred nationwide in 2020, resulting in 61,703 deaths, 250,723 injuries and direct property losses of 131,361 million yuan.

Aiming at the serious traffic safety problem, people hope to solve it effectively through the intelligent automobile. If the driver can find the hidden trouble in advance and timely brake the vehicle, it will effectively reduce the occurrence of traffic accidents and reduce losses. Therefore, it has become a research hotspot of intelligent transportation to enhance the perception of the external environment of vehicles through artificial intelligence technologies such as machine vision. Target detection technology can quickly and accurately detect the vehicle in front and provide feedback to the detection frame, which is of great significance to alleviate traffic safety problems.

In 2015, the YOLOv1 algorithm proposed by J Redmon et al. [1] used neural networks to segment the images to be detected and classify them in each region, reducing the average detection time and improving the detection accuracy at the same time. However, the grid design of the YOLOv1 algorithm also resulted in low detection effect for small targets and multi-targets. To address these shortcomings, researchers improved the YOLOv1 algorithm and successively proposed v2-v5[2-4] versions. This series of improvement overcomes many deficiencies of YOLO algorithm and pushes target detection to a new height. The algorithm also begins to be widely used in various fields. ZHANG et al realized rapid detection of vehicles by improving the YOLOv3 algorithm. In addition to the YOLO series algorithm[5], other single-level detection algorithm performance is also very good. In 2015, Liu et al. proposed the algorithm SSD(Single Shot MultiBox Detector)[6], which uses multiple feature layers. Different resolution feature layers greatly enhance the algorithm's ability to deal with small targets. Corner Net algorithm[7], which predicts target coordinates according to corner points, is also a representative in single-level detection. This method attaches more importance to the information of target edges and greatly improves the detection accuracy. Fan et al. proposed an infrared vehicle target detection algorithm based on improved YOLOv5 to solve such problems as false detection, missed detection and insufficient detection ability of infrared vehicle images[8]. Based on the Yolov5 model, Zhan et al. proposed four methods to improve the detection accuracy of small objects[9]. Ye et al. proposed an AV end-to-end adaptive neural network control turning Angle prediction method based on YOLOv5[10].

In this paper, aiming at the problems of low detection accuracy, slow detection speed, and lack of complete detection of various vehicles in the current vehicle detection model, the original data set was expanded and several common vehicles were added. A vehicle target detection method based on YOLOv7-tiny algorithm is proposed. This method can detect the vehicle in front by using machine vision, and provide safety assistance for vehicle driving through the intuitive feedback image. The model parameters are smaller than those of YOLOv5s, which can run perfectly on the on-board computing platform and has good detection effect on all kinds of vehicles.

2.METHOD

2.1Self-organizing datasets

The self-organizing data set used in this study mainly includes six categories of vehicle tags: Car, Bus, Van, Truck, Cyclist and Motorbike. The data comes from the vehicle tags and self-labeling pictures of the open source data sets KITTI[11], VOC[12] and COCO[13]. The self-organizing data set consists of 33,089 images.

Firstly, four categories of Car, Van, Truck and Cyclist were extracted from KITTI data set as basic data sets. Secondly, the label screening code was used to screen out four categories of bus, motorbike, car and cyclist from VOC data set, and five categories of bicycle, car, motorbike, bus and truck from COCO data set. Modify the code with tag names to change the

International Journal of Science and Engineering Applications
Volume 12-Issue 03, 40 – 42, 2023, ISSN:- 2319 - 7560
DOI: 10.7753/IJSEA1203.1015

above data to the same category name as the KITTI dataset. labellmg annotation tool is used to annotate the unannotated images, and the generated annotation file contains the category name, label category and the coordinates of the boundary box of the target object. Annotated images and files in the self-annotated data set are shown in Figure 1.

Figure 1. Annotate images and annotate files

Finally, the images selected from the open source data set are added to the Kitti-based self-organizing data set as supplements. The vast majority of the images were labeled positive samples, while a small number of unlabeled images were negative samples. According to the principle of random divided into 8:2 ratio, respectively training set and test set. At this point, the data set construction of the whole study was completed. Figure 2 shows the picture in the data set.

Figure 2. Dataset picture

2.2 YOLOv7-tiny Algorithm

In There are many algorithms for target detection. YOLOv7 algorithm is a typical representative of one-stage target detection algorithm. It runs fast and can be used for real-time detection. As a lightweight model of YOLOv7, YOLOv7-tiny can be better arranged on the vehicle platform, and the detection effect is also excellent. YOLOv7-tiny is mainly composed of backbone and head. Different from previous YOLO versions, Yolov7-tiny proposes an ELAN efficient network architecture, and proposes an auxiliary head method to increase training costs and improve accuracy without affecting reasoning time. Algorithm flow figure 3 is shown.

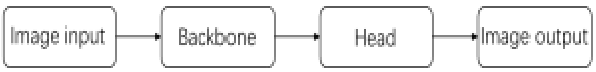

Figure 3. Algorithm flowchart

3.EXPERIMENT

3.1 Model training

In this paper, the idea of transfer learning is adopted and the pre-training model of COCO data set is used as the feature extraction device. The training parameters are set as follows: the size of the input image is 640×640, the batch sent to the network is 32 each time, and the depth multiple and width multiple are set to 0.33 and 0.55. The warm-up strategy was adopted to first warm up the model, set the initial learning rate as 0.01, the number of warm-up learning rounds as 3, the initial momentum as 0.8, the paranoid learning rate as 0.1, and the epoch as 50.

3.2 Model validation and analysis

Model testing performance was evaluated based on accuracy (P) and class average accuracy (mAP). The mAP value is defined as the mean of the average accuracy of each class (average accuracy, AP), and the AP value corresponds to the area under some type of P-R curve. mAP is a performance indicator of comprehensive evaluation. It represents the Average Precision (AP) of all categories in the data set. The specific calculation formula is as follows:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

$$AP = \int_0^1 P(R)d(R) \quad (2)$$

$$mAP = \frac{1}{m}\sum_{i=0}^{m} AP_i \quad (3)$$

Among them, TP is the true class, indicating that the correct prediction is correct, FP is the false positive class, indicating that the wrong prediction is correct, and FN is the false negative class, indicating that the correct prediction is wrong.

The test results of the test set are shown in Table 1:

Table 1. Model performance

category	P/%	AP/%
All	80.8	77.2
Car	80.9	79.1
Van	90.6	98.5
Truck	65.5	63
Bus	85.9	83
Cyclist	79.6	65.6
Motorbike	82.6	73.8

The effect diagram of the real vehicle test is shown in Figure 4

Figure 4. Part of the results

4. CONCLUSION

In this paper, an automobile object detection method based on Yolov7-Tiny is proposed. At the present stage, the identification methods for vehicle detection mostly adopt the method of machine vision. Under the same data set, the accuracy difference between different models is not particularly large. In terms of model size, this model is only 11.7M, which is its unique advantage. It has been verified that the method proposed in this paper has an average accuracy of 80.8% and 77.2% for all kinds of vehicle detection.

5. ACKNOWLEDGMENTS

This study was funded by the Qingdao Top Talent Program of Entrepreneurship and Innovation (Grant No.19-3-2-11-zhc), the Natural Science Foundation of Shandong Province (Grant No. ZR2020MF082), the Foundation of Shandong Intelligent Green Manufacturing Technology and Equipment Collaborative Innovation Center (Grant No. IGSD-2020-012), and the National Key Research and Development Project (Grant No.2018YFB1601500).

6. REFERENCES

[1] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[2] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 7263-7271.

[3] REDMON J, FARHADI A. Yolo v3: An incremental improvement[J]. arXiv preprint arXiv: 1804.02767, 2018.

[4] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLO v4: Optimal speed and accuracy of object detection [J]. arXivPreprint arXiv:2004.10934, 2020.

[5] ZHANG Fukai, YANG Feng, LI Ce. Fast vehicle detection method based on improved YOLOv3. Computer Engineering and Applications, 2019, 55（2）：12

[6] W Liu, D Anguelov, D Erhan, C Szegedy, S Reed, C-Y Fu, and A C Berg, "SSD: Single shot multibox detector," In ECCV, 2016: 21-37.

[7] LAW H, DENG J. Cornernet: Detecting objects as paired keypoints [C]// Proceedings of the European Conference on Computer Vision (ECCV), 2018: 734-750.

[8] Fan, Y, Qiu, Q. Hou, S. Li, Y, Xie, J, Qin, M, Chu, F. Application of Improved YOLOv5 in Aerial Photographing Infrared Vehicle Detection. Electronics 2022, 11, 2344. https://doi.org/10.3390/electronics11152344

[9] Zhan, W, Sun, C, Wang, M et al. An improved Yolov5 real-time detection method for small objects captured by UAV.Soft Comput 26, 361–373 (2022). https://doi.org/10.1007/s00500-021-06407-8

[10] Ye C, Wang Y, Wang Y, Tie M. Steering angle prediction YOLOv5-based end-to-end adaptive neural network control for autonomous vehicles. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering.2022;236(9):1991-2011.doi:10.1177/09544070211053677

[11] Geiger, A., Lenz, P., & Urtasun, R. (2012). Are we ready for autonomous driving? The KITTI vision benchmark suite. 2012 IEEE Conference on Computer Vision and Pattern Recognition, 3354-3361.

[12] Everingham, M., Gool, L.V., Williams, C.K., Winn, J.M., & Zisserman, A. (2010). The Pascal Visual Object Classes (VOC) Challenge. International Journal of Computer Vision, 88, 303-338.

[13] Chen, X., Fang, H., Lin, T., Vedantam, R., Gupta, S., Dollár, P., & Zitnick, C.L. (2015). Microsoft COCO Captions: Data Collection and Evaluation Server.ArXiv, abs/1504.00325.