

Vehicle object detection and improvement methods in specific scenarios

Zihou Guo

NOIC ACADEMY, 50 Featherstone Ave, Markham ON, Canada

17731099560@163.com

Abstract. This paper mainly studies the current mainstream vehicle object detection methods and improvement methods in specific scenarios. The research provides ideas and methods for further enhancing the accuracy and stability of vehicle object detection, which is beneficial in reducing missed and incorrect detections. This can improve the accuracy of target detection in urban areas, construction sites, and other scenarios, thereby ensuring safety. The purpose of this research is to improve vehicle object detection methods and identify directions for enhancement. In this essay, the main way of research is that reading a lot of papers and contrasting the advantage and disadvantage among them. The main finding of this paper is that in different application scenarios, the directions for improvement vary, and the improved algorithms cannot simultaneously accommodate both accuracy and robustness, and the improved methods cannot be widely used. Based on this information, pointing out the future development direction and clear problems.

Keyword: Vehicles, Convolutional Neural Network, Vehicle Object Detection, Deep Learning, Visual Tasks

1. Introduction

With the development of neural networks, autonomous driving is becoming increasingly prevalent in our daily lives. In recent years, autonomous driving has become one of the most popular topics. Vehicle object detection technology has emerged as a major issue in the field of autonomous driving. There is a need to continually refine vehicle object detection technology to enhance the safety of autonomous driving. In practical terms, solving the accuracy problems of vehicle object detection can help us improve the safety issues associated with autonomous driving.

Firstly, vehicle object detection is a key technology in autonomous driving, belonging to the field of object detection research. Object detection is a widely applied AI technology, encompassing areas such as face detection, pedestrian detection, and remote sensing, with vast application potential. Object detection technology can be subdivided into specific task types such as classification, localization, detection, and segmentation, depending on functional requirements. The classification, localization, size, and shape of the detection objects are core issues in object detection. Among these, vehicle object detection, as one of the most common applications in daily life, also faces special challenges such as target overlap and occlusion. Currently, China has over 400 million motor vehicles, and manually solving vehicle object detection issues would require significant manpower and resources, resulting in inefficiency. Advanced and high-quality vehicle object detection algorithms can greatly reduce the time

cost of vehicle object detection, significantly improving detection efficiency. These algorithms can be applied to emerging fields such as new energy vehicles and autonomous vehicles, laying a technical foundation for the long-term development and popularization of autonomous driving. Additionally, there are already mature and advanced deep learning methods for vehicle object detection, such as R-CNN, Fast R-CNN, Faster R-CNN, and RPN. In-depth research and promotion of these methods can effectively advance the application and popularization of vehicle object detection technology.

This paper is divided into five main sections. The second section reviews classical literature. The third section focuses on introducing the algorithm used in this study. The fourth section discusses and analyzes the advantages and disadvantages of the algorithm. Finally, the fifth section summarizes the research, discusses its limitations, and outlines future research directions.

2. Literature review

Cao et al. found that in traditional vehicle detection problems, the need to design manual features reduced detection efficiency. By constructing visual tasks, using deep convolutional neural networks to extract convolutional features of visual task sample images, and performing feature normalization and parallel regression calculations based on Fast R-CNN, they ultimately obtained a vehicle object detection model related to visual tasks, avoiding the design of manual features. However, the initial sample extraction process was time-consuming [1].

Han et al. proposed a vehicle detection algorithm based on deep learning to address the inability of traditional vehicle detection algorithms to adapt to complex scene changes and extract corresponding features. This algorithm, combining the Faster R-CNN open-source framework and LocNet network algorithm, offers higher vehicle object detection and localization accuracy while avoiding redundant information in vehicle detection frames. However, it suffers from missed detections and is not advantageous for detecting small objects [2].

Wang and Zhang proposed a model to address the inefficiencies and poor generalization capabilities of traditional machine learning in vehicle detection applications, which are easily affected by factors such as lighting, target scale, and image quality. This method, based on the Faster R-CNN model, extracts vehicle features through convolution and pooling operations on input images, combines multi-scale training and hard negative mining strategies to reduce the impact of complex environments, and can automatically extract vehicle features, solving the time-consuming and laborious feature extraction problem of traditional methods. It also improves vehicle detection accuracy, with good generalization ability and applicability, but the speed is slightly slow [3].

Shi et al. discovered that using the Fast R-CNN method for feature extraction had issues with long processing time and low detection accuracy. They proposed an improved front vehicle detection method, combining the Faster R-CNN front vehicle detection model with three different sizes of convolutional neural networks, improving the accuracy and robustness of vehicle detection with a certain degree of generalization ability. However, missed detection and sample annotation issues need to be considered [4].

Gao et al. addressed the inability of existing convolutional neural network-based vehicle object detection algorithms to effectively adapt to target scale changes, self-deformation, and complex backgrounds. They proposed a vehicle object detection algorithm that integrates multi-scale contextual convolutional features. The algorithm uses a feature pyramid network to obtain feature maps at multiple scales and employs a region proposal network to locate candidate target regions at each scale. It then integrates contextual information of candidate target regions with extracted multi-scale features and finally predicts the vehicle target position and type through multi-task learning. This algorithm has stronger robustness and accuracy [5].

Chen et al. addressed the problem of excessive dependence on pre-trained weights for infrared scene target detection under data scarcity conditions by integrating attention modules to mitigate the detection performance degradation caused by not pre-training. Based on the YOLO v3 algorithm, they incorporated SE and CBAM modules that mimic human attention mechanisms into the network structure to re-calibrate the extracted features at the channel and spatial levels [6].

Li and Zhang proposed an automatic vehicle object detection method for urban street areas in satellite optical remote sensing images with complex backgrounds. The method suppresses vegetation background using multispectral bands, suppresses buildings using a panchromatic band combined with binary morphological methods, and employs the RX algorithm for vehicle object detection. This method is robust, efficient, and does not require manual assistance, making it suitable for automatic vehicle object detection in urban street areas [7].

Yang et al. proposed an improved real-time vehicle detection algorithm based on Faster R-CNN to address issues such as occlusion between vehicles, changes in lighting, shadows, tree branch movements, and the movement of fixed objects in the background affecting detection and recognition accuracy. The algorithm obtains suitable aspect ratios to adapt to significantly different vehicle sizes, improves the region proposal network, reduces computational load, and optimizes the network structure, basically meeting the real-time monitoring needs of vehicles. However, factors like vehicle density and severe occlusion significantly affect detection results [8].

Yuan et al. proposed an image processing method based on secondary transfer learning and the Retinex algorithm for effective identification of vehicles in nighttime aerial images. They used a deep learning algorithm based on Faster R-CNN for quick vehicle detection. This method can train an effective classifier using a small-scale aerial dataset, outperforming traditional machine learning methods with higher recognition accuracy and meeting the needs for rapid detection [9].

Xu et al. proposed PVDNet, a pedestrian and vehicle detection network based on deep learning, to solve the target detection problem in autonomous driving environments. They introduced multi-level skip connections (MLSC) and designed a multi-layer feature fusion method (MLFF) and a one-dimensional convolution method (ODC) to improve accuracy [10].

Jin and Hu proposed a vehicle occupant number detection method based on multispectral infrared images and an improved Faster R-CNN to address the reliability and accuracy issues of current high occupancy vehicle lane detection methods based on radar and infrared thermal imaging technology. The method uses a multispectral infrared imaging system to obtain images of the car interior and combines the Faster R-CNN deep learning algorithm for occupant number detection. Enhancements include a fully convolutional network structure, multi-scale feature prediction, and using ROI-Align instead of ROI-Pooling to improve network generalization. However, the accuracy does not meet industry standards, and there is a significant gap in detection speed compared to YOLOv3 [11].

3. Discussion

Traditional vehicle object detection methods face several challenges, including dependence on manual feature extraction, the need for pre-training, and issues with detection speed and accuracy. The RCNN method selects 2000 candidate regions from the input image and uses AlexNet to extract features from each region. According to the literature, Fast RCNN extracts features from the entire image and then maps each candidate region onto the feature map [9]. Faster RCNN's first stage generates anchor boxes for objects to be detected in the image, while the second stage classifies the objects within the anchor boxes. The YOLO algorithm, a single neural network object detection system, improves detection accuracy and speed with YOLO2. YOLO employs a standalone CNN model for end-to-end object detection. According to the literature, one-stage methods like the YOLO series are computationally simpler and more efficient than two-stage methods like the RCNN series, though the latter typically offer higher accuracy [4].

In the practical application of regional object detection methods, in addition to the basic model framework and concepts, it is also necessary to perform targeted optimization and adaptation based on the specific context of the real-world problem.

For example, according to the literature, regional object detection in practice may encounter issues of target overlap and occlusion, which significantly affect the model's performance [5]. By designing background suppression algorithms tailored to the physical properties of different interferences, the problem of complex background interference in urban areas can be effectively overcome. This approach

offers the advantages of simplicity, high operational efficiency, and no need for manual assistance, making it suitable for detecting vehicle targets in urban areas.

Similarly, according to the literature, small target detection issues are frequently encountered in regional object detection, with many small targets being missed on streets [4]. To address this problem, an improved YOLOv5s-based vehicle object detection method was proposed, incorporating a multi-head self-attention mechanism module and drawing on the fusion computation method of a bidirectional feature pyramid to enhance small target detection. The improved method reduced detection speed but met real-time detection requirements, effectively addressing the issue of missing small vehicle targets, although some missed detections still occur.

For instance, according to the literature, addressing the problem of convolutional neural networks' excessive reliance on pre-trained weights in object detection algorithms, an integrated design of a convolutional neural network was proposed [3]. Based on the YOLOv3 algorithm, SE and CBAM attention modules were embedded in the original network to increase sensitivity to useful features, mitigating the performance degradation caused by not pre-training. The final adjusted model showed good detection performance for vehicle targets in urban street scenes.

According to the literature, in addressing the problem of designing manual features in traditional vehicle object detection, convolutional features of visual tasks were extracted using deep convolutional neural networks [1]. Features were normalized and parallel regression calculations were performed based on Fast RCNN, effectively avoiding the dependence on manual features. However, the initial sample extraction process was time-consuming and could not form an end-to-end detection process with Fast RCNN.

4. Conclusion

Currently, all research approaches can improve their accuracy and detection speed in specific scenarios. This science becomes more and more mature. Most improved methods combine the based system such as Fast-RCNN, Faster-RCNN, and YOLO, with other algorithm and model. Let them can use in special scenes. The finding of research concludes popular ways to improve the technology of object detection, and it provides the future development direction and way. They address the problem of overlapping and occlusion of detection targets in urban areas and small target detection well. However, these algorithms are constrained by the inability to simultaneously balance speed and accuracy, as well as the over-reliance of convolutional neural networks on pre-trained weights. They limited the efficiency of detection and increasing the risk in our daily life. Pre-trained also takes up a lot of time and money. The issue of manual feature extraction remains, and improvements still cannot form an end-to-end detection process with Fast RCNN. Although these ways can get the detection areas, they reduce the accuracy of detecting objects and increasing the error rate. There is no end-to-end detection. Declining the transmission speed and accuracy. Small target detection still faces missed detections, and regional object detection methods need specific modifications for particular scenarios, limiting their broad applicability. Increasing the potential risk of road detection. Every problem needs a special way to solve, and it is very difficult and inefficient. Future research can further enhance the speed, accuracy, and compatibility of object detection. It will improve the safety and efficiency in human society significantly.

References

- [1] Cao, S. Y., Liu, Y. H., & Li, X. Z. (2017). Vehicle Object Detection Based on Fast R-CNN.
- [2] Han, K., Zhang, H. Y., Wang, Y., & Xu, M. (2017). A Vehicle Detection Algorithm Based on Faster R-CNN. *Journal of Southwest University of Science and Technology*, 32(4), 65-70.
- [3] Wang, L., & Zhang, H. H. (2018). Application of Faster R-CNN Model in Vehicle Detection. *Computer Applications*, 38(3), 666-670.
- [4] Shi, K. J., Bao, H., Xu, B. X., Pan, W. G., & Zheng, Y. (2018). Intelligent Vehicle Detection Method in Front of Road Based on Faster RCNN. *Computer Engineering*, 44(7), 36-41.
- [5] Gao, L., Chen, N. N., & Fan, Y. (2019). Vehicle Object Detection with Fusion Multi-Scale Context Convolution Features. *Opto-Electronic Engineering*, 46(4), 180331-1.

- [6] Chen, H. C., Cheng, Y., & Zhang, C. Y. (2017). Application of Convolutional Neural Network in Fast Detection of Vehicle Object. *Journal of Software*, 28(s1), 107-114.
- [7] Li, Z. H., & Zhang, J. Q. (2014). Automatic Detection Method of Vehicle Targets in Urban Block Starborne Optical Remote Sensing Images. *Infrared and Laser Engineering*, 43(11), 3751-3755.
- [8] Yang, W., Wang, H. Y., Zhang, J., & Zhang, Z. B. (2019). An Improved Algorithm for Real-Time Vehicle Detection Based on Faster-RCNN. *Journal of Nanjing University (Natural Sciences)*, 55(2), 231-237.
- [9] Yuan, G. L., Hou, J., & Yin, K. Y. (2019). Night Aerial Vehicle Recognition Method Based on Transfer Learning and Image Enhancement. *Journal of Computer-Aided Design & Computer Graphics*, 31(3), 467-473.
- [10] Xu, Q., Li, Y., & Wang, G. (2019). Pedestrian and Vehicle Detection Based on Deep Learning. *Journal of Jilin University (Engineering and Technology Edition)*, 49(5), 1661-1667.
- [11] Jin, X., & Hu, Y. (2020). Method of Vehicle Occupant Number Detection Based on Improved Faster R-CNN. *Infrared Technology*, 42(11), 1103-1110.