Research on the Application of Recognition and Detection Technology in Automatic Driving

Wanli He [1], Zhengzhe Liu [2] and Shunquan Wang[3, *]

[1] Department Electrical Engineering and Electronic, University of Liverpool, Liverpool, United Kingdom

[2] Dujiangyan Senior High School of Sichuan, Sichuan Province, China

[3] Henry Samueli School of Engineering, University of California, Irvine, California, United States of America

*Corresponding author: shunquaw@uci.edu

Abstract. Object detection is the cornerstone of autonomous driving systems. In autonomous driving systems, vehicle detection is critical to maintaining traffic flow and avoiding collisions. The use of deep learning technology enables the system to skillfully and accurately identify vehicles on the road. Deep learning algorithms can provide autonomous vehicles with precise environmental awareness to enable informed driving decisions. Deep learning algorithms can achieve pedestrian detection. Such algorithms can improve the safety of vehicle operations. Through deep learning models, the system is able to identify and track pedestrians, ensuring appropriate responses in different traffic scenarios. The application of detection technology enables vehicles to promptly identify and interpret road signs and traffic lights in real-time. This recognition ability helps to comprehensively understand traffic conditions and improve the ability to adapt to different road conditions. This study provides an in-depth exploration of the importance and application of deep learning technology in the field of autonomous driving. Special attention is paid to the key role played by object detection methods. and propose future development directions. Deep learning algorithms can expand the application of learning technology to handle real-time traffic conditions and achieve rapid response. Achieve the fusion of multiple sensory information by improving the local processing capabilities of edge computing. This research aims to further improve the level of intelligence inherent in autonomous driving systems.

Keywords: Autonomous driving; object detection, pedestrian detection, vehicle detection, deep learning algorithms.

1. Introduction

With the development of the automotive industry over the past century, mechanical excellence in vehicles has been achieved to a very high level. At present, the core proposition of the automotive industry has come to intelligence with the development of The Times, at this time, it is inseparable from automatic driving technology. Particularly crucial within autonomous driving is the technology of recognition and detection. It is a vital component of autonomous driving systems, analyzing the surrounding environmental images to identify and understand information such as roads, traffic signs, and obstacles. This enables the vehicle to make intelligent driving decisions.

The development of recognition and detection technology in autonomous driving is not only directly related to enhancing the safety performance of vehicles but also impacts the improvement of driving experience and traffic efficiency. In-depth research in this field can drive the rapid advancement of autonomous driving technology, reducing the risk of accidents and elevating the overall intelligence of the transportation system.

This study focuses on recognition and detection technology in autonomous driving, exploring how advanced computer vision and deep learning techniques can be utilized to efficiently perceive complex traffic environments. This includes precise identification of targets such as vehicles, pedestrians, and obstacles, and real-time monitoring of their motion states. Recognition utilizes image analysis with the aid of deep learning algorithms, enhancing the accuracy and robustness of target detection.

Through extensive training and optimization with a large amount of scene data, the system is designed to operate stably in various complex traffic situations. This aims to enhance the perception and decision-making capabilities of autonomous driving systems in complex traffic environments, achieving efficient and accurate recognition and detection of various targets. Through technological innovation, the goal is to improve the safety of autonomous driving while raising the overall intelligence of the transportation system, contributing to the future development of intelligent transportation.

2. Pedestrian Detection in AI Autonomous Driving

Pedestrian detection poses severe challenges to the field of autonomous driving technology. To achieve autonomous driving, the vision system equipped with the vehicle must be able to cope with various complex scenarios. These scenes include different viewing angles, light fluctuations, different poses, and potential ambient occlusion scenarios. These scenarios are set up to accurately identify pedestrians [1]. The use of computer vision for pedestrian detection is developing rapidly in both the computer science and intelligent vehicle fields. This monitoring technology relies on cameras installed in the vehicle to detect nearby pedestrians to assist the driving system. The process aims to assess potential risks and subsequently implement measures to ensure pedestrian safety [2].

In the domain of pedestrian detection, detection methods can be broadly categorized into early traditional approaches and the more prevalent feature-based detection methods currently in use. Firstly, the discussion revolves around traditional detection methods, which typically exhibit well-established theoretical foundations and find widespread application in practical domains such as video surveillance. Even though these techniques are less computationally complicated, they frequently run into problems when used for pedestrian detection, which requires a greater level of detecting precision. Using optical flow as an example, classical detection methods have several drawbacks, including low tracking rates, poor real-time performance, and simple loss of tracked targets. Nevertheless, these methods have seen continuous improvements in recent years, exemplified by motion target tracking and detection methods based on improved optical flow features [3]. Certainly, these methods often combine two or more approaches. This is because traditional detection methods have inherent limitations in accuracy. While combining multiple methods can enhance precision, it also concurrently escalates the hardware requirements of detection devices. Consequently, such methods typically face limitations in their application areas, which contradicts the objectives of our autonomous driving detection system.

The feature-based detection methods differ from traditional detection approaches, and they have undergone considerable advancements in recent years. Earlier feature-based detection methods such as HOG (Histogram of Oriented Gradients) initially faced challenges related to unclear features, low computational complexity, and slow detection speeds. However, these issues have been mitigated to a certain extent through enhancements, as observed in the enhanced HOG pedestrian detection algorithm.

The algorithm preprocesses raw images, extracts their HOG features, and enhances them to generate an improved HOG, experimental results show a recognition rate of 95.49%, significantly boosting pedestrian detection performance [4].In recent years, prevalent detection methods such as YOLO4, R-CNN, and SSD have gained traction. Frameworks like YOLO and SSD predominantly adopt a candidate-box-based approach for pedestrian detection. For instance, the enhanced Tiny-YOLOv3 model exhibited elevated accuracy and quicker real-time detection, attaining an accuracy of 81.13%, a recall rate of 76%, an average IoU of 83.76%, and operated at 62 frames per second (FPS) [5].

Methods of this kind often generate a considerable number of candidate boxes, similar to traditional detection methods that combine multiple detection approaches. Consequently, these methods demand higher hardware specifications; otherwise, they may experience decreased detection speed. In scenarios with less capable hardware, meeting real-time detection requirements becomes challenging. another form of detection relies on regression-based object detection, which usually involves predicting targets in advance to enhance overall real-time performance, such as using CNNs. For pedestrian detection in surveillance scenarios, a novel approach integrating an enhanced Support Vector Machine and Convolutional Neural Network is proposed [6]. Given that CNNs typically prioritize real-time processing over accuracy, enhancing CNNs is especially crucial for improving accuracy.

3. Vehicle Detection in AI Autonomous Driving

In the evolution of autonomous driving technology, vehicle detection plays a significant role. Its essential task is to identify and process vehicles on the road through advanced algorithms to ensure driving safety. Deep learning object detection algorithms dominate in this field, which includes some key methods.

In recent years, vehicle detection has played a significant role in the development of autonomous driving. Its essential task is to process and identify vehicles on the road by utilizing advanced algorithms, and meanwhile ensuring driving safety. Deep learning object detection is the dominant algorithm that is used in this field, and it includes some key detection methods. Faster R-CNN is one of the methods since it has powerful capabilities in processing complex scenes and small objects. Furthermore, Fast R-CNN can accurately locate the position of target objects while providing higher detection accuracy for autonomous driving systems.

Accurate vehicle detection is important in preventing complex traffic collisions. The application of deep learning algorithms in autonomous drive systems can be reflected in vehicle tracking and recognition, because the autonomous driving system first understands the surrounding environment, and then makes appropriate driving decisions by identifying and tracking the surrounding vehicles. [7] .Whether to use YOLO or Faster R-CNN for vehicle detection is often dependent on some specific application and performance requirements, because each deep learning-based monitoring algorithm has its unique advantage in certain scenarios. YOLO has better real-time performance and faster response capabilities, and due to its design. It allows the detection of the entire image in a single forward pass, which enables high processing speeds. Quick response is always a crucial concern regarding autonomous driving systems. In situations such as timely driving decisions need to be made, the rapid response capability of the system is more important. If there are higher requirements for detection accuracy in a complex traffic environment, Faster R-CNN is more suitable. The reason is that it has a two-stage detection process, and it can accurately locate the target object and adapt to scenes with small and overlapping objects. The ability to accurately detect target objects is critical to prevent collisions in different scenarios. The choice of algorithms is sometimes affected by the hardware resources and the actual test results. In practical applications, it is necessary to combine the advantages of both flexible design and adaptive systems [8]. The Kalman filter is excellent for fusing sensor measurements and system model predictions because it provides an improved method for vehicle detection systems. By applying a Kalman filter, it is possible to create a system that fuses both outputs of YOLO and Faster R-CNN to estimate the vehicle's position and motion state with higher accuracy. Furthermore, the Kalman filter not only effectively handles noise and uncertainty in sensor measurements, but also better off the accuracy of tracking vehicles. The advantage of this fusion method is that it combines the real-time capabilities of YOLO and the high-precision detection capabilities of Faster R-CNN to improve the adaptability and robustness of the autonomous driving system. By achieving this goal, the Kalman filter establishes an effective balance between algorithm outputs, especially in complex and unpredictable traffic circumstances. There is no doubt that this fusion approach opens up a new direction for the development of autonomous driving technology by combining deep learning with traditional signal processing [9].

However, the vehicle detection field still faces challenges, such as detection accuracy and reliability under varying weather and traffic conditions. Additionally, the requirement for high real-time performance places a higher demand on the efficiency of algorithms, which requires fast response and processing.

In the future, the development of vehicle detection technology may move in the direction of enhancing learning techniques and the collaboration between humans and machines. So, by utilizing deep learning, autonomous driving systems can continuously learn and optimize on the actual roads to better handle unexpected situations. Moreover, by understanding the habits of drivers, the collaboration between human drivers and autonomous driving systems will become efficient, and further improve the safety of driving [10].

4. Traffic Sign Detection

Vision-based vehicle guidance systems for artificially driven cars operating in road environments have three primary functions: road detection, sign recognition, and obstacle detection. Among these, lane detection is a useful technology in modern autonomous vehicle systems, aiding accurate self-positioning based on detected road lines. Traditional methods utilize edge detection and algorithms based on the Hough transform to draw lines along the detected lanes. In ongoing developments, an improved double-threshold Canny operator based on the Canny edge detection algorithm is proposed for lane edge detection [11]. The use of the Accumulated Probability Hough Transform algorithm has enhanced accuracy and real-time performance. However, previous deep learning-based algorithms treated lane detection as pixel-level lane segmentation, limiting the ability to detect a variable number of lanes. To address this, a two-stage CNN architecture for lane detection is proposed by Ge Zhang, Chaokun Yan, and Jianlin Wang, incorporating a binary lane mask for effective handling of a variable number of lanes[12]. Furthermore, recognizing that lane lines are semantic objects with strong shape priors but weak appearance coherence, a network structure specific to lane line detection is designed to improve the network's ability to capture pixel spatial relationships in rows and columns of images. The structured road traffic marking detection task is an improvement on the LaneNet lane detection algorithm. It adopts the VGG16 model inspired by the FCN semantic segmentation algorithm, replacing the underperforming ENet network structure in the original algorithm. This improves clustering accuracy and algorithm real-time performance [13].

Future improvements in lane detection require enhancing the model's ability to detect varying lane quantities and lane changes. Additionally, obstacle detection models need improved accuracy and adaptability to complex weather conditions, necessitating the establishment of large-scale scene databases for continuous iterative training. Robustness to different weather conditions needs enhancement. Exploring network structures dedicated to lane line detection to improve the network's ability to capture pixel spatial relationships in rows and columns of images is also crucial. It is crucial to handle the conflict between input resolution and semantic or instance segmentation accuracy at the same time as algorithm real-time performance. Techniques for encoding real-world road photographs in high definition need to be investigated. Improving lane detection algorithms to adapt to situations without lane markings and enhancing model robustness in special scenarios are also important considerations.

Traffic sign detection and recognition are major research areas in autonomous driving and traffic sign detection. In the context of color model-based traffic sign detection algorithms, an improved threshold segmentation method based on the HSV color space is proposed. After converting the RGB color space of the test image to the HSV color space, adaptive threshold segmentation is applied to the H, S, and V components to complete traffic sign detection. This method not only effectively segments traffic signs but also removes more interference noise. Experimental accuracy and average time consumption are reported at 88.6%, demonstrating the method's effectiveness in eliminating background interference and significantly improving detection speed and accuracy[1]. However, in most cases, color space-based image segmentation methods for traffic sign detection have low accuracy in environments with graffiti contamination, excessive exposure, low visibility, and severe dust blur, as these disturbances seriously disrupt the color characteristics of traffic signs, leading to reduced detection rates. If the traffic sign detection module commonly uses an adaptive color segmentation based on local neighborhood average saturation and circular Hough transform to locate traffic signs in input images, adaptive color thresholds perform better in segmenting images with uneven illumination or very high or very low contrast levels for traffic signs compared to global thresholds. For example, using the GTSRB dataset to develop and train ten convolutional neural networks (CNNs) for the twelve categories of traffic signs, a top-level network achieved a validation accuracy of 99.65% on similar images in the GTSRB dataset. The average time for classifying one input image (60x60 pixels) was 0.54 milliseconds, making it suitable for real-time applications [14]. The application of traffic sign detection technology, such as lane markings and traffic signals, in autonomous driving is discussed in this section. It covers the application of traffic sign detection technology in autonomous driving, technological advancements, potential issues, and future development directions

5. Conclusion

Pedestrian detection is crucial in computer vision, involving finding out the location and size of pedestrians in images or videos, similar to object detection, but facing challenges such as attitude and viewpoint changes. vehicle detection technology plays a key role in the field of autonomous driving, involving algorithmic innovation, environmental adaptability, and exploration of future technologies. With continual technological advancement, these systems are expected to become more intelligent and reliable. The reduction of accidents and enhancement of road safety remain critical challenges in autonomous driving. That unintended lane departure is a major cause of global vehicle collisions, making lane detection the most promising yet challenging task in autonomous driving. many organizations are integrating deep learning technologies with computer vision to address autonomous driving challenges. The future development direction should enhance the learning technology so that the autonomous driving system can continue to learn and optimize on the actual road.

Authors Contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Yang, L. A review of deep learning-based pedestrian re-identification.China Water Transport,2023.
[2] Hui-Xing, and ZHANG YuJin. "A Survey of Computer Vision-Based Pedestrian Detection for Driver Assistance Systems", Acta Automatica Sinica 33.01 (2007): 84-90.
[3] Yang, Y., Yang, Y., & Liu, H. Motion Object Tracking Based on Improved Optical Flow Features. Computer and Modernization,2021.
[4] Li, W.-S., Han, Y., Ruan, M.-H., & Wang, Z.-X. Improved Pedestrian Detection Algorithm Based on Enhanced HOG. Computer Systems & Applications,2020.
[5] Li, X. H., & Song, X. L. Pedestrian Object Detection Research based on Tiny-YOLOv3.Software Guide,2021.
[6] Yang, J., & Yang, Z. Research on Pedestrian Detection Algorithm Based on Improved CNN Multilevel Classification. Computer Applications and Software,2017.
[7] M. Maity, S. Banerjee and S. Sinha Chaudhuri, Faster R-CNN and YOLO based Vehicle detection: A Survey,2021 5th International Conference on Computing Methodologies and Communication,2021:1442-1447.

[8] J. Fan, J. Lee, I. Jung, and Y. Lee, Improvement of Object Detection Based on Faster R-CNN and YOLO," 2021 36th International Technical Conference on Circuits/Systems, Computers and Communications , 2021:1-4,.

[9] Benjdira B, Khursheed T, Koubaa A,et al.Car Detection using Unmanned Aerial Vehicles: Comparison between Faster R-CNN and YOLOv3[C]2019 1st International Conference on Unmanned Vehicle Systems-Oman,2019.

[10] Ahangar, M. Nadeem, Qasim Z. Ahmed, Fahd A. Khan, and Maryam Hafeez. 2021. A Survey of Autonomous Vehicles: Enabling Communication Technologies and Challenges" Sensors,2021, 21(3): 706.

[11] Xueqi Yang. Research on Lane Line Detection and Traffic sign Recognition of outdoor Intelligent Vehicle based on machine vision [D]. Jilin: Changchun University of Technology, 2022.

[12] Zhang, Ge, Chaokun Yan, and Jianlin Wang. Quality-guided lane detection by deeply modeling sophisticated traffic context. Signal Processing: Image Communication 2020,84: 115811.

[13] Yue Guan. Research on intelligent vehicle environment perception method perception method integrating binocular vision [D]. Beijing Institute of Technology, 2020.

[14] Oza, Rudri Mahesh, Angelina Geisen, and Taehyung Wang. "Traffic sign detection and recognition using deep learning,IEEE, 2021.