Technology Application of Autonomous Vehicle in Machine Learning

Fang Shang
College of Letters and Science, University of California Davis,1 Shields Ave, Davis, CA, United States

Keywords: Autonomous Vehicles, Sensor Fusion, Machine Learning, Object Detection, Multi-Modal Data

Abstract: The development of autonomous vehicle technology has significantly promoted innovation in Sensor Fusion and Object Detection methods based on Machine Learning, especially in single-modal target detection and Multi-Modal Data fusion. This research explores object detection techniques based on RGB images and point cloud data, such as Faster-RCNN and PointNet, to improve performance in complex scenes. These methods have significant advantages in distinguishing objects in complex scenes, thus enhancing the perception of vehicles. In addition, research has focused on Multi-Modal Data Fusion technologies, such as fusing images with point clouds and radar data, to enable autonomous driving systems to better cope with severe weather and complex environments. By integrating multiple sensor data, these Machine Learning methods improve the perception and decision reliability of the system. However, challenges in data quality, model generalization, and robustness remain. To solve these problems, it is necessary to optimize the sensor fusion algorithm and further enhance the reliability and security of the model. Future research will focus on improving these sensor fusion strategies to ensure that Autonomous Vehicles achieve reliable and efficient perception under a variety of real-world conditions, supporting safer and more intelligent autonomous driving systems.

1. INTRODUCTION

Nowadays, the vigorous development of artificial intelligence is more and more common in our daily life, and some intelligent technologies have also entered people's society. In recent years, autonomous vehicles have slowly entered the market of society. Over the past decade, autonomous vehicle technology has come a long way. The new capabilities of autonomous vehicle technology will have a profound impact globally and could significantly change society (Daily, Medasani, Behringer, and Trivedi, 2017). In simple terms, a self-driving car is a vehicle that replaces humans with machines and does not require humans to drive the car. Cars drive themselves through computer systems and sensor technology. Autonomous driving requires multiple functions, including positioning, perception, planning, control, and management, and information acquisition isa condition for positioning and perception (Asif, Faisal, He, and Liu, 2019). Indeed, autonomous vehicles achieve autonomous driving, visual computing, radar, monitoring equipment, global positioning system and other technologies through computer systems to automatically and safely cooperate and operate with motor vehicles. These vehicles use sensors to sense their surroundings and make driving decisions based on the perceived data, eliminating the need for an operator. However, the booming development of autonomous vehicles has also caused some technical problems in machine learning, such as autonomous driving relies on sensor data, including cameras, lidar, radar, etc. However, the quality and accuracy of the data directly affect the performance of the model, especially in complex and dynamic traffic environments. And machine learning models often rely on specific scenarios or data distributions during training. However, in practical applications, autonomous vehicles will encounter different environments and extreme situations, such as bad weather, and how to ensure that the model can generalize and operate safely under different conditions, which is a problem worth challenging.

2. SINGLE MODALITY TARGET DETECTION

Autonomous vehicles rely on precise environment-sensing technology to ensure they can safely navigate

Shang and F.
Technology Application of Autonomous Vehicle in Machine Learning.
DOI: 10.5220/0013525800004619
In *Proceedings of the 2nd International Conference on Data Analysis and Machine Learning (DAML 2024)*, pages 446-450
ISBN: 978-989-758-754-2

complex road environments. At present, the object detection methods in automatic driving are mainly divided into two categories, among which the method based on RGB image, and the method based on point cloud data. Methods based on RGB images, such as Faster-RCNN (Zhou, Wen, Wang, Mu, and Richard, 2021) and SSD (Simhambhatla, Okiah, Kuchkula, and Slater, 2019), use traditional cameras to capture visual information and identify and classify different objects through deep learning models. Point cloud-based detection methods, such as PointNet (Paigwar, Erkent, Wolf, and Laugier, 2019) and PointR-CNN (Oliveira et al., 2023), use three-dimensional point cloud data generated by radar sensors for spatial modeling, which can better detect complex structure and distance information.

2.1 Based on RGB Images

Faster-RCNN algorithm. Faster-RCNN is a common target detection algorithm, which generates candidate boxes through area suggestion network (RPN) and combines with reinforcement neural network (CNN) for target classification and location (Zhou et al., 2021). In automatic driving, it is very important to detect vehicles, pedestrians, road signs and other targets, and the complexity of the scene puts forward requirements for detection algorithms. In the Faster-RCNN algorithm, they use regional proposal network (RPN) instead of selective search to recommend candidate proposals, which greatly improves the quality of candidate proposals, reduces the amount of computation, and realizes end-to-end training classification. On the other hand, in the proposed model, the target detection problem in the automatic driving scenario, the Faster RCNN network is optimized and the residual connection module based on spatial attention is added (Ni, Shen, Chen, Cao, and Yang, 2022). This module helps the network to better focus on important details in the input image, enhance the detection ability of small targets and blocked targets, and effectively retain discriminant features to reduce information loss, thus improving the accuracy of target detection, and is more suitable for complex automatic driving environments.

Different from before, the working principle of an improved Faster-RCNN model. Faster-RCNN not only retains the basic structure of RPN (Zhou et al., 2021), but also adds residual connection module and spatial attention mechanism to perform better, especially when dealing with challenges such as occlusion and small targets in automatic driving. These improvements enable the autonomous vehicle model to better detect fine targets, improving the accuracy and effect of detection. Meanwhile, similar to the image-based method of Faster-RCNN, some methods for detecting SSDS in autonomous vehicles (Simhambhatla et al., 2019). SSD is an object detection algorithm based on feedforward form neural network, which can directly predict fixed-size bounding boxes. The detection performance of SSD is achieved by simultaneous target detection with multi-scale feature maps. The detection results are efficiently processed with non-maximum suppression to achieve the maximum boundary box and finally generate accurate target detection results. CARLA simulator to generate synthetic data, combined with SSD and Faster RCNN model for training and testing, and evaluated its target detection performance in automatic driving by analyzing various performance parameters, aiming to improve the accuracy and practicability of the model (Niranjan, VinayKarthik, and Mohana, 2021). An explanation of how the SSD algorithm works and its advantages in object detection, especially in the application of autonomous vehicles. Specifically, it illustrates how SSDS can simultaneously perform object location and classification tasks with a single forward pass, avoiding complex steps such as Faster-RCNN that require a regional proposal network, thereby increasing detection speed. And the primary role of SSDS is to enable real-time object detection through their simple, efficient architecture.

2.2 Point Cloud Based

Compared to methods based on RGB images, point cloud data has a natural advantage in processing complex three-dimensional structure and depth information, especially when detecting distant objects and dealing with occlusions. Scenario description is how to deal with the challenges of selective attention and data accuracy in deep learning. Since these operations are non-differentiable, the network cannot be trained by backpropagation. To solve this problem, the Spatial Transformation Network (STN) module (Paigwar et al., 2019). The STN module allows the network to perform an explicit spatial transformation of the input data internally. In addition, a series of Point Net methods. The application of the Point Net family in autonomous vehicles improves perception by processing 3D point cloud data (Wang and Goldluecke, 2021). By introducing new key-point sampling algorithms and integrating dynamic occupancy heat maps, PointPainting enhances the understanding of complex scenes, improves the accuracy of target detection and recognition, and contributes to the safety and decision-making capabilities of autonomous driving systems. The application of visual attention mechanisms and

PointNet and its improved methods in autonomous driving, especially when processing 3D point cloud data. Visual attention mechanisms help the network focus on the important parts, while methods such as PointPainting improve detection efficiency and accuracy by improving key point sampling and data fusion. These technologies enable autonomous vehicles to better perceive complex three-dimensional environments, especially when detecting multiple objects and processing dynamic scenes.

A Point R CNN method (Oliveira et al., 2023). Point R-CNN enables accurate 3D object detection in self-driving cars by processing point cloud data. Its two-stage approach combines raw point data and voxel representation to improve detection efficiency and accuracy, helping the environment perception and decision making of the autonomous driving system. Which indicates that the main role of Point RCNN in autonomous vehicles is target detection and recognition based on three-dimensional point cloud data. Specifically, Point RCNN takes a two-stage approach to take full advantage of raw point cloud data and voxel representation. Point RCNN first takes the voxel representation as input and performs lightweight convolution operations to produce a small number of high-quality initial predictions. In the initial prediction, the attention mechanism effectively combines the coordinate and indexed convolution features of each point, preserving both the accurate positioning of the target object and the context information. In the second stage, the network uses fusion features inside the point cloud to further refine the prediction. The refinement process at this stage can better capture the fine structure characteristics of the object, thus enhancing the accuracy of the final detection results.

3. MULTIMODAL FUSION

In autonomous vehicles, the multi-modal fusion method refers to the integrated processing of data from different sensors, such as LiDAR and cameras, to improve perception, positioning, decision-making and control capabilities. Single-mode methods are simple, low cost, and fast to process, but their performance may be limited in complex environments. Multimodal fusion combines the advantages of different sensors to provide greater perceptual accuracy and safety, suitable for demanding complex tasks such as autonomous driving. Because different sensors have different advantages and limitations, the comprehensive use of multi-modal data can make up for the shortcomings of a single sensor and improve the accuracy of the system.

3.1 Image and Point Cloud Fusion

A new multi-sensor fusion framework for autonomous vehicles (Shahian Jahromi, Tulabandhula, and Cetin, 2019). The framework combines deep learning-based full convolutional neural networks (FCNx) and traditional Extended Kalman filters (EKF) to provide cost-effective, real-time, and robust environment awareness solutions. This shows that the multi-sensor fusion framework of image fusion, combined with a full convolutional neural network and extended Kalman filter, can optimize the perception tasks such as road segmentation and obstacle detection. This method improves the sensing accuracy and can realize real-time performance in embedded systems. It is economical, efficient and robust.

A sensor fusion algorithm that is not directly related to the specific approach of image and point cloud fusion but focuses on enhancing the robustness and security of CAVs against sensor failures and attacks through sensor redundancy and fusion algorithms, while previous approaches focus on image and point cloud fusion to improve perception accuracy (Yang and Lv, 2022). The $H\infty$ controller is also introduced to reduce the influence of estimation error and communication noise on vehicle queue cooperation and expand the application field. In addition, Farag's Computer Vision-based Vehicle Detection and Tracking Method (RT_VDT) is used for autonomous driving or Advanced Driver assistance systems (ADAS). This method mainly relies on RGB image processing and generates the bounding box of the vehicle through a series of computer vision algorithms, which is significantly different from the previous image and point cloud fusion methods. The RT_VDT method relies on RGB images to realize real-time vehicle detection and tracking through computer vision algorithms, emphasizing speed and accuracy, and low computing resource requirements.

3.2 Image and radar data fusion

In addition to image and point cloud fusion in multimodal, image and radar data fusion can enhance detection, and how can analysis radar data supplement image data, especially in severe weather conditions. A dynamic Gaussian method, whose process performs occupation mapping, to determine the drivable area of the autonomous vehicle within the radar field of view (Hussain, Azam, Rafique, Sheri, and Jeon, 2022). Different from the deep learning technology that relies on a large amount of data, the Gaussian process method can still work effectively in the case of sparse data and the combination of image data and radar data, and the dynamic Gaussian process can combine the environmental information provided by the radar and the detailed information provided by the image sensor to enhance the perception and path planning capabilities of the automatic driving system under different environmental conditions. The application of multi-radar data fusion technology in the environment perception of autonomous vehicles, especially the weight calculation method in the case of sensor data deviation (Ren, He, and Zhou, 2021). This study uses correlation changes between sensors, sensor consistency, and stability to evaluate sensor reliability, thereby optimizing the data fusion process. Different from the dynamic Gaussian process, this method focuses on the data fusion of multiple radar sensors, especially dealing with the sensor data deviation problem, and optimizing the weight by evaluating the correlation between sensors. Compared with the previous Gaussian process method, the dynamic evaluation of sensor performance is emphasized.

The application of radar-based environmental perception to high automation and the transformation of radar and camera information fusion methods (Dickmann et al. ,2016) and (Liu, Zhang, He, and Zhao, 2022). It emphasizes the need for fast situation update, dynamic object motion prediction, position estimation and semantic information in the driverless phase, and requires radar signal processing to provide revolutionary solutions in these new fields. And to solve the inherent defects of a single sensor in bad weather conditions. The scheme uses the radar as the main hardware and the camera as the auxiliary hardware to match the observed values of the target sequence by using the Mahalanobis distance and performs data fusion based on the joint probability function. Both approaches emphasize the use of radar as the primary sensor, with cameras as auxiliary, to optimize data fusion through Mahalanobis distance matching and joint probability functions, with a particular focus on performance improvements in adverse weather conditions. The previous approach may be broader and not refer specifically to bad weather or specific matching algorithms.

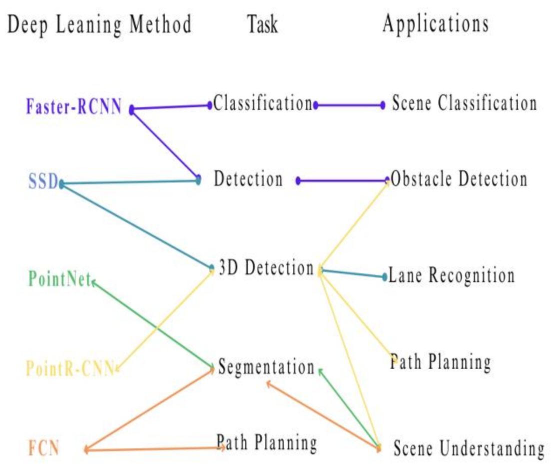

Figure 1: Deep learning tasks and applications

See the picture above. Figure 1 shows the relationship between tasks and applications of deep learning methods in autonomous driving. The application of autonomous vehicles in machine learning covers a variety of deep learning methods and specific tasks. These methods include Faster-RCNN, SSD, PointNet, PointR-CNN, and FCN, each of which plays a critical role in different tasks and application scenarios. Faster-RCNN is mainly used for classification and object detection, such as scene classification and obstacle detection, to help autonomous vehicles identify objects in the environment. SSDS are also used for object detection and are particularly suitable for real-time obstacle detection. PointNet and PointR-CNN are better at 3D object detection and segmentation for tasks such as lane recognition and scene understanding, helping cars better sense and navigate in complex 3D environments. FCN is used for path planning and scene understanding to ensure that the vehicle can plan a safe driving path by splitting the various elements of the environment. The tasks and applications of each deep learning approach complement each other, working together to enable the perception, decision making, and control functions of autonomous driving systems, thereby improving the safety and reliability of driving. This fully demonstrates the widespread application of machine learning in autonomous driving and its indispensable value.

4. CONCLUSIONS

However, sensor data quality and model generalization ability in complex environments are still urgent problems to be solved. At present, automatic driving target detection methods mainly include Faster-RCNN (Zhou et al., 2021) based on RGB images and PointNet based on point cloud data. The RGB image method focuses on visual recognition and point cloud data is more suitable for processing three-dimensional structures. In addition, multi-modal fusion technology combines data from different sensors to effectively improve perception, positioning and decision-making capabilities. Future improvements and research will make breakthroughs in improving sensor data quality, optimizing multi-modal fusion algorithms, and enhancing model generalization ability. First, improving the data quality of the sensors is the basis for ensuring the safe operation of the autonomous driving system in complex road conditions, especially the accuracy of data acquisition in bad weather is crucial. Second, the optimization of multimodal fusion technology will enable autonomous vehicles to better combine data from cameras, LiDAR and radar to improve environmental awareness and decision-making. In addition, enhancing the generalization ability of the model will be key, especially in response to different geographical environments and uncertainties, and the robustness of the model needs to be significantly improved. These future studies will advance the safety and reliability of autonomous driving systems in the real world, preparing them for wider deployment.

REFERENCES

Daily, M., Medasani, S., Behringer, R., & Trivedi, M., 2017. Self-driving cars. *Computer*, 50(12), pp. 18–23.

Dickmann, J., et al., 2016. Automotive radar: The key technology for autonomous driving: From detection and ranging to environmental understanding. In *2016 IEEE Radar Conference (RadarConf)*, Philadelphia, PA, USA, pp. 1–6.

Faisal, A., et al., 2019. Understanding autonomous vehicles: A systematic literature review on capability, impact, planning, and policy. *Journal of Transport and Land Use*, 12(1), pp. 45–72.

Hussain, M. I., Azam, S., Rafique, M. A., Sheri, A. M., & Jeon, M., 2022. Drivable region estimation for self-driving vehicles using radar. *IEEE Transactions on Vehicular Technology*, 71(6), pp. 5971–5982.

Liu, Z., Zhang, X., He, H., & Zhao, Y., 2022. Robust target recognition and tracking of self-driving cars with radar and camera information fusion under severe weather conditions. *IEEE Transactions on Intelligent Transportation Systems*, 23(7), pp. 6640–6653.

Ni, J., Shen, K., Chen, Y., Cao, W., & Yang, S. X., 2022. An improved deep network-based scene classification method for self-driving cars. *IEEE Transactions on Instrumentation and Measurement*, 71, pp. 1–14.

Niranjan, D. R., VinayKarthik, B. C., & Mohana, 2021. Deep learning-based object detection model for autonomous driving research using CARLA simulator. In *2021 2nd International Conference on Smart Electronics and Communication (ICOSEC)*, Trichy, India, pp. 1251–1258.

Paigwar, A., Erkent, O., Wolf, C., & Laugier, C., 2019. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) Workshops*, pp. 0–0.

Ren, M., He, P., & Zhou, J., 2021. Improved shape-based distance method for correlation analysis of multi-radar data fusion in self-driving vehicles. *IEEE Sensors Journal*, 21(21), pp. 24771–24781.

Shahian Jahromi, B., Tulabandhula, T., & Cetin, S., 2019. Real-time hybrid multi-sensor fusion framework for perception in autonomous vehicles. *Sensors*, 19(19), pp. 4357.

Silva, A. L., Oliveira, P., Duraes, D., Fernandes, D., Nevoa, R., Monteiro, J., Melo-Pinto, P., & Machado, J., 2023. A framework for representing, building, and reusing novel state-of-the-art three-dimensional object detection models in point clouds targeting self-driving applications. *Sensors*, 23(23), pp. 6427.

Simhambhatla, R., Okiah, K., Kuchkula, S., & Slater, R., 2019. Self-driving cars: Evaluation of deep learning techniques for object detection in different driving conditions. *SMU Data Science Review*, 2(1), pp. 1–23.

Wang, L., & Goldluecke, B., 2021. Sparse-PointNet: See further in autonomous vehicles. *IEEE Robotics and Automation Letters*, 6(4), pp. 7049–7056.

Yang, T., & Lv, C., 2022. A secure sensor fusion framework for connected and automated vehicles under sensor attacks. *IEEE Internet of Things Journal*, 9(22), pp. 22357–22365.

Zhou, Y., Wen, S., Wang, D., Mu, J., & Richard, I., 2021. Object detection in autonomous driving scenarios based on an improved Faster-RCNN. *Applied Sciences*, 11(11), pp. 11630.