Proceedings of the 2nd International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/75/20240528

Deep learning in autonomous driving: Advantages, limitations, and innovative solutions

Siming You

Minjiang University No.200, Xiyuan Gong Road, Shangjie Town, Minhou County, Fuzhou City, Fujian Province, 35010, China

Louisyou111@outlook.com

Abstract. With the rapid development of autonomous driving technology, deep learning has become a core driver for innovation in testing autonomous driving scenarios. This review paper delves into the critical role of deep learning in autonomous driving technology. The paper will describe how deep learning is at the center of driving innovation. The paper thoroughly explores the application of deep learning in obstacle detection, scene classification and understanding, and image segmentation, emphasizing the significant benefits in perception and decision-making while pointing out the challenges and innovative solutions adopted. The innovative solutions section proposes multimodal fusion and joint learning, new methods for 3D semantic segmentation, etc., aiming to improve image segmentation's accuracy and generalization ability. Overall, deep learning has great potential in automated driving technology, and by innovating and solving challenges, it will advance the system and provide reliable, intelligent, and efficient solutions for future transportation systems.

Keywords: Deep Learning, Autonomous Driving Technology, Image Segmentation, Innovative Solutions.

1. Introduction

With the rise of autonomous driving technology, complex and changing environments make fully autonomous driving difficult. To ensure robust environment perception, cars are equipped with various sensors, among which camera and LiDAR fusion have become a research focus. [1] As one of its key enablers, deep learning leads to technological innovation in autonomous driving. In this context, this thesis aims to analyze in depth the complex challenges faced by current autonomous driving scenarios. It focuses on innovative applications of deep learning in addressing these challenges and identifying gaps in the field.

1.1. Background and Motivation

RISE OF AUTOMATIC DRIVING TECHNOLOGY: In recent years, autonomous driving technology has witnessed a boom and become a cutting-edge transportation innovation.

Deep learning-based automated driving technology is a popular research topic in artificial intelligence, which can benefit automatic driving control based on different methods. [2] Its application potential is believed to improve transportation efficiency, reduce accident rates, and bring greater

DOI: 10.54254/2755-2721/75/20240528

convenience to travel. The rise of this technology has triggered the need for in-depth research on testing methods for autonomous driving scenarios.

The critical role of deep learning in autonomous driving: In autonomous driving systems, deep learning enables efficient processing and decision-making in complex scenarios by learning from large amounts of data. Self-driving cars are built according to specific algorithms and are not intelligent. As the world's top automobile technology continues to evolve, so do the algorithms that manage and control its movements learn to progress. Self-driving technology has become one of the many examples of the power of machine learning in the world. [3] From obstacle detection to lane keeping, deep learning techniques have empowered autonomous driving systems with the intelligence to perceive and make decisions. However, they still face many challenges when faced with complex test scenarios.

1.2. Research Objectives

ANALYZING EXISTING CHALLENGES IN AUTOMATIC DRIVING SCENARIO TESTING: In response to the complexity and diversity that exists in the testing of automated driving scenarios, this research aims to systematically analyze the current testing challenges, including, but not limited to, the complexity of multi-source data simulation, real-time deep learning inference system design, and interactive reinforcement learning test environments.

EXPLORING INNOVATIVE APPLICATIONS OF DEEP LEARNING IN BRIDGING THESE CHALLENGES: Based on the understanding of the existing challenges, this dissertation will delve deeper into how to address the various challenges of testing autonomous driving scenarios by innovatively applying deep learning techniques, especially in the areas of multi-source data fusion, synergistic application of reinforcement learning and deep learning, and real-time model uncertainty estimation.

Through in-depth analysis of existing challenges and exploration of innovative applications of deep learning, this thesis aims to provide valuable insights for the further development of autonomous driving technology.

2. Application of deep learning knowledge to the testing of autonomous driving scenarios

2.1. Application in Obstacle Detection

In the field of autonomous driving, to ensure safe driving, the Target Detection(TD) performance must be very high to ensure the safety of vehicle traveling, and most autonomous driving uses single-stage TD methods (e.g., SSD methods) to solve this challenge. [4]

Deep learning-based solutions have made significant progress in self-driving cars in recent years. These solutions cover critical tasks such as obstacle recognition, lane detection, and scene recognition, providing more reliable and accurate functionality for automotive AI systems. For example, in places like India, self-driving vehicles face challenges such as roadblocks, potholes, speed bumps, and strolling pedestrians. They utilize VANET to communicate with the car to obtain real-time traffic and road information to improve adaptability and safety. [5] Deep learning techniques enable cars to recognize essential elements such as humans, other vehicles, and road signs by learning on massive images and video datasets.

The introduction of deep learning has significantly improved the predictive modeling capabilities of self-driving cars. Cars can examine sensor information with the help of deep learning algorithms to predict, for example, the probability of a human crossing the road or another vehicle changing lanes quickly. Through deep learning, cars can better understand and judge complex driving scenarios. The wide application of this technology enables autonomous driving systems to have higher perception and decision-making capabilities, laying the foundation for a safe and efficient autonomous driving experience. [6]

Object detection, a core component of the automatic driving system to perceive and understand the surrounding environment, is crucial in automatic driving technology. Its fundamental goal is to enable the system to accurately grasp its surroundings through astute recognition, meticulous classification, and

DOI: 10.54254/2755-2721/75/20240528

precise localization, dotted with various objects that make up the environmental landscape for autonomous entity navigation.

The researchers then conducted a combination of deep learning and traditional image processing frameworks to detect lanes to reduce data collection time and effort while maintaining performance. The detected lanes defined the road, while depth information was utilized to detect obstacles on the road. The experimental results show that their system performs well in a test environment with model cars and is expected to be utilized in a real system. [7]

With the widespread application of deep learning techniques in autonomous driving, deep learning algorithms explicitly designed for object detection tasks have been significantly refined and accelerated for deployment.

Semi-autonomous system techniques have also been derived in recent years and are now being used in a small number of vehicles. [6]

The rapid evolution of these customized deep learning algorithms has driven the performance metrics of object detection systems. Advanced algorithms leverage the neural network infrastructure to extract salient features from image data through complex network structures, thereby significantly improving the accuracy and efficiency of object detection. Such technological advancements have significantly improved the reliability and functionality of autonomous driving systems in real-world, dynamically changing environments. [8]

2.2. Utilization of Scene Classification and Understanding

Deep Learning (DL) has seen remarkable success in the last few years and has become a key component of computer vision applications. Among them, neural networks play a crucial role in solving the problem of object detection and localization in images and videos, providing advanced solutions for many applications. Convolutional neural network (CNN) is a deep learning method suitable for image processing and object detection. It enables object classification and localization by recognizing patterns and features in images. In areas such as self-driving cars, this technique allows vehicles to identify and understand their surroundings, including recognizing lanes, traffic lights, pedestrians, crossings, and traffic signs.

The application of DL allows self-driving cars to perceive and understand complex driving environments more accurately, thereby improving safety and effectiveness. Over time, DL networks can be trained with many images and videos, enabling them to improve their performance in various scenarios gradually.

Another DL model suitable for processing data sequences is the Long Short-Term Memory (LSTM) network, a deep learning model for sequential data that performs exceptionally well when dealing with time-series data. The LSTM network achieves recognition of sequences and patterns through feedback connections and is suitable for remembering and dealing with long-term dependencies. In areas such as autonomous driving, LSTM networks can process time-series data generated by vehicle sensors. One of the key strengths of deep learning (DL) is its ability to analyze and extract meaningful insights from unstructured data, such as images and videos captured by cameras. Deep learning networks are capable of identifying complex patterns and features in data through a process of training, where large amounts of data are used to fine-tune the model's parameters. This training process can be significantly accelerated by utilizing graphics processing units (GPUs), which are designed for parallel computing and can perform complex calculations much faster than central processing units (CPUs).

As technology advances, deep learning is vital in various fields, providing powerful tools for solving complex problems. [9]

2.3. Image Segmentation

Image segmentation algorithms have undergone several development stages, from early histogram bundles to modern methods based on deep learning (DL). Early techniques include thresholding, K-mean clustering, etc., while current algorithms are based on graph cuts, active contours, conditional and Markov random fields, and sparsity-based methods. In recent years, algorithms based on DL models

DOI: 10.54254/2755-2721/75/20240528

have achieved significant results in image segmentation, including CNNs, expanded CNNs, R-CNNs, FCNs, RNNs, generative models, adversarial models, and attention-based models. [10]

With the emergence of semantic segmentation in a unique form, the amount of information obtained from the segmentation of images has been further enhanced. Existing camera-LIDAR fusion methods are available for 2D and 3D semantic segmentation. 2D/3D semantic segmentation aims to predict the category labels for each pixel and point. In 2D semantic segmentation, feature-level fusion is essential, including parallel processing and fusion of autoencoder networks, multilevel feature-level fusion, and applying up-sampled depth images and images. The best-performing models typically employ two parallel CNN branches that fuse dense depth images and image data in the final convolutional layer by skipping connections. At the same time, the field of 3D semantic segmentation has seen the emergence of several innovative approaches. 3DMV networks improve semantic prediction performance in voxelized point clouds by subtly fusing image semantic and point features. The UPF framework efficiently characterizes the learning by combining image features, geometric structure, and context to optimize point cloud semantic label prediction.MVPNet, on the other hand, achieves accurate semantic label speculation for each point by fusing multi-view images and 3D geometric information. In addition, SPLATNet employs sparse bilateral convolution to efficiently process multimodal data, including images and point clouds, to enhance spatially aware representation learning. These approaches provide new directions for 3D semantic segmentation research, optimizing performance and addressing challenges like point cloud voxelization. [11] For example, obtaining accurate environment perception and precise localization in complex dynamic environments are critical requirements for the safe driving of autonomous vehicles. Both require acquiring and processing highly accurate and information-rich data from real-world environments. Multiple sensors, such as LiDAR, must be equipped on a self-driving car or mapping vehicle to collect and extract information about the target environment to achieve this data. However, image data lacks 3D geo-referenced information. Therefore, utilizing dense, geo-referenced, and accurate 3D point cloud data from LiDAR becomes a new alternative. [2] Lidar is a crucial component of the perceptual sensor suite in autonomous driving technology. Its unique features, including high-resolution velocity imaging, detection of objects in occlusion and at long ranges, and robust performance in adverse weather conditions, contribute significantly to data collection. [12]

2.4. Advantages and Limitations of Deep Learning

2.4.1. Advantages

1 In the case of frequent changes or disturbances, parametric methods can lead to prediction errors. Deep learning models are potentially innovative in providing new ways for more accurate traffic motion analysis. [1]

2 For data collected from on-board sensors, such as optical, infrared, and radar images and point cloud data, existing methods for image classification, object detection, and segmentation are well adapted to maritime, offshore environments, which are much simpler scenarios compared to the complex road conditions of self-driving cars. DL techniques have been proven to outperform other traditional methods in terms of accurate and robust detection [13]

3 utilizes deep learning models such as CNN, Fast R-CNN, Faster R-CNN, YOLO, SSD, etc., capable of learning complex feature representations to improve the accuracy of traffic signs. Deep learning methods excel in multi-layer feature extraction and are particularly suitable for processing large-scale image data. Deep learning models can automatically learn key features, improving robustness in complex environments compared to traditional methods. In addition, after real-time consideration, YOLO and SSD perform better in real-time ATSDR, highlighting the advantages of deep learning in handling real-time tasks. [14]

2.4.2. Limitations

1 Most modern machine learning techniques cannot recognize fundamental similarities in traffic flow details, and driverless cars face several challenges, including problems with proving their practical

protectiveness, neural network structure, interpretability, and reliance on large amounts of training data and computational hardware. The goal is to improve the performance of transportation systems through sophisticated devices and sensors. However, challenges still need to be addressed due to the complexity of road network modeling and traffic prediction. [1]

2 Deep Learning (DL) has challenges, such as a lack of supervised learning and adaptation to test distributions. Migration learning, generative adversarial networks, and contrast learning are used to overcome these problems. Contrast learning has promise for pre-training navigation modules in marine environments. Deep reinforcement learning learns control laws directly, but reward function design is complex. The research uses path, obstacle avoidance, and speed to design rewards and propose methods for agents to learn intrinsic reward functions autonomously. [13]

3 The big data challenge involves the use of LiDAR to collect millions to billions of natural scene points in different urban or rural environments, creating difficulties in data storage and processing. Second, the accuracy challenge manifests itself in the importance of accurately sensing road objects, but the diversity of intra- and extra-class objects and the inhomogeneity, sparseness, and missingness of the data challenge perception accuracy. [15]

3. Innovative solutions

3.1. Multimodal Fusion and Joint Learning

Joint learning using multi-sensor data, such as images and LiDAR, to improve the accuracy and robustness of image segmentation.

Explore deep learning models that can effectively fuse data generated from different sensors to achieve more comprehensive environment sensing.

3.2. New methods for 3D semantic segmentation

Further, develop the field of 3D semantic segmentation and explore more innovative methods, such as deep learning models based on point clouds, which can more accurately segment the environment.

Consider introducing temporal information to improve understanding of dynamic scenes, such as vehicles, pedestrians, etc., at different time steps.

3.3. Adversarial Learning for Generative Models for Image Segmentation

Consider using adversarial generative models to generate more realistic and detailed image segmentation results, especially when faced with complex scenes.

Explore how Generative Adversarial Networks (GANs) can be used to improve the performance and generalization of image segmentation models.

3.4. Innovative Applications of Migration Learning and Contrastive Learning

Use migration and contrast learning to address the problem of insufficient data for supervised learning in image segmentation and improve the generalization ability of models in new environments.

Investigate the potential of combining migration learning techniques, which allow models to adapt to new environments, with generative adversarial networks (GANs), a class of machine learning algorithms, to develop adaptive image segmentation models that can perform effectively across different environments or datasets.

4. Future Outlook and Research Directions

Future research trends will focus on potential applications of deep learning in autonomous driving and future directions to break through current challenges. Driven by academics and innovation, research will explore the fusion of multimodal data and sensor integration to improve the accuracy of environment perception. Emphasis is placed on the introduction of reinforcement learning and autonomous learning to enable autonomous driving systems to respond more intelligently to changes in the real-time environment and improve adaptation to dynamic traffic scenarios. Meanwhile, innovative research on

3D scene understanding and semantic segmentation is emphasized for a deeper understanding of complex 3D environments and enhanced perception. The application of adversarial learning and generative modeling will be further investigated to generate more realistic and enriched image results and to improve model robustness and adaptability. Finally, optimizing real-time performance and driving edge computing applications will be important directions for future research to meet the demand for low latency and efficient decision-making in autonomous driving systems. These trends will drive the development of autonomous driving technology and provide more reliable, intelligent, and efficient solutions for future transportation systems.

5. Conclusion

Deep learning has shown significant advantages in autonomous driving, especially in adapting to complex environments, multi-sensor fusion, and automatic feature learning. These advantages enable deep learning models to perform more accurate traffic motion analysis, adapt to maritime and offshore environments, and excel in real-time tasks. Innovative solutions include multimodal fusion and joint learning, new methods for 3D semantic segmentation, generative models for adversarial learning for image segmentation, and innovative applications of transfer learning and contrast learning, which aim to improve the accuracy and generalization of image segmentation.

However, deep learning also faces challenges, including difficulties in recognizing traffic flow details, lack of supervised learning, and adaptation to test distributions. Methods such as migration learning, generative adversarial networks, and contrast learning are used to overcome these challenges. Future research directions will focus on improving the accuracy of environment perception, introducing reinforcement learning and autonomous learning, delving into 3D scene understanding and semantic segmentation, optimizing real-time performance, and promoting the use of edge computing.

Overall, deep learning has great potential in autonomous driving technology, and through continuous innovation and challenge-solving, it will be expected to drive the development of autonomous driving systems and provide more reliable, intelligent, and efficient solutions for future transportation systems.

References

[1] A. Kumar and P. Saha, "A review of deep learning models for traffic flow prediction in Autonomous Vehicles," 2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN), Greater Noida, India, 2020, pp. 303-308, doi: 10.1109/ICACCCN51052.2020.9362972.

[2] A. Khanum, C. -Y. Lee and C. -S. Yang, "Involvement of Deep Learning for Vision Sensor-Based Autonomous Driving Control: A Review," in IEEE Sensors Journal, vol. 23, no. 14, pp. 15321-15341, 15 July 15, 2023, doi: 10.1109/JSEN.2023.3280959.

[3] E. Jebessa, K. Olana, K. Getachew, S. Isteefanos and T. K. Mohd, "Analysis of Reinforcement Learning in Autonomous Vehicles," 2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC), Las Vegas, NV, USA, 2022, pp. 0087-0091, doi: 10.1109/CCWC54503.2022.9720883.

[4] G. Tong, J. Cai, L. Huang, B. Shang, J. Liu and M. Chen, "SSD Based Target Detection Algorithm for the Autonomous Driving Vision System," 2021 IEEE 2nd International Conference on Information Technology, Big Data and Artificial Intelligence (ICIBA), Chongqing, China, 2021, pp. 213-217, doi: 10.1109/ICIBA52610.2021.9688170.

[5] N. S. Manikandan, G. Kaliyaperumal and Y. Wang, "Ad Hoc-Obstacle Avoidance-Based Navigation System Using Deep Reinforcement Learning for Self-Driving Vehicles," in IEEE Access, vol. 11, pp. 92285-92297, 2023, doi: 10.1109/ACCESS.2023.3297661.

[6] A. Agrawal, "Autonomous Self Driving System Using Deep Learning Appraoches and Impact," 2023 International Conference on Circuit Power and Computing Technologies (ICCPCT), Kollam, India, 2023, pp. 961-970, doi: 10.1109/ICCPCT58313.2023.10245567.

[7] N. S. Manikandan, G. Kaliyaperumal and Y. Wang, "Ad Hoc-Obstacle Avoidance-Based Navigation System Using Deep Reinforcement Learning for Self-Driving Vehicles," in IEEE Access, vol. 11, pp. 92285-92297, 2023, doi: 10.1109/ACCESS.2023.3297661.

[8] Jingyuan Zhao et al., "Autonomous driving system: A comprehensive survey," in Expert Systems with Applications Volume 242, 15 May 2024, 122836

[9] Badr Ben Elallid et al., "A Comprehensive Survey on the Application of Deep and Reinforcement Learning Approaches in Autonomous Driving," in Journal of King Saud University - Computer and Information Sciences, Volume 34, Issue 9, October 2022, Pages 7366-7390.

[10] N. N. Ali, N. A. Kako and A. S. Abdi, "Review on Image Segmentation Methods Using Deep Learning," 2022 4th International Conference on Advanced Science and Engineering (ICOASE), Zakho, Iraq, 2022, pp. 7-12, doi: 10.1109/ICOASE56293.2022.10075607.

[11] Y. Cui et al., "Deep Learning for Image and Point Cloud Fusion in Autonomous Driving: A Review," in IEEE Transactions on Intelligent Transportation Systems, vol. 23, no. 2, pp. 722-739, Feb. 2022, doi: 10.1109/TITS.2020.3023541.

[12] A. Srivastav and S. Mandal, "Radars for Autonomous Driving: A Review of Deep Learning Methods and Challenges," in IEEE Access, vol. 11, pp. 97147-97168, 2023, doi: 10.1109/ACCESS.2023.3312382.

[13] Y. Qiao, J. Yin, W. Wang, F. Duarte, J. Yang, and C. Ratti, "Survey of Deep Learning for Autonomous Surface Vehicles in Marine Environments," in IEEE Transactions on Intelligent Transportation Systems, vol. 24, no. 4, pp. 3678-3701, April 2023, doi: 10.1109/TITS.2023.3235911.

[14] A. Juyal, S. Sharma and P. Matta, "Traffic Sign Detection using Deep Learning Techniques in Autonomous Vehicles," 2021 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), Chennai, India, 2021, pp. 1-7, doi: 10.1109/ICSES52305.2021.9633959.

[15] Li, Y., Ma, L., Zhong, Z., Liu, F., Chapman, M. A., Cao, D., & Li, J. (2020). Deep learning for lidar point clouds in autonomous driving: A review. IEEE Transactions on Neural Networks and Learning Systems, 32(8), 3412-3432, doi: 10.1109/TNNLS.2020.3015992.