Autonomous Vehicles: Exploring the Advancements in Computer Visison Technology

Amina Hoor Zoi
18040967
ec24937@qmul.ac.uk

I. Introduction

Autonomous vehicles (AVs) have seen a surge in investment due to the growing demand for safer and more efficient transportation. Human error is a major cause of traffic accidents, accounting for about 57% of incidents (Ribeiro et al., 2023). This emphasizes the necessity of trustworthy systems with autonomous navigation and decision-making capabilities. Advanced technologies are used by autonomous cars to help solve this problem.

Research shows that computer vision plays a key role in helping AVs understand their surroundings. Studies by Dong et al. (2024) and Chuprov et al. (2023) emphasize improvements in areas like lane detection, object recognition, and localization. However, challenges remain. Much of the current research focuses on isolated tasks, ignoring how different computer vision systems interact. Additionally, many existing systems struggle in difficult environments. The high cost of advanced sensors like LiDAR (Chuprov et al., 2023) and public concerns about safety (Dong et al., 2024) also hinder widespread use.

II. Problem Definition

Accurately observing the surroundings, analysing sensory information, and performing safe and effective driving movements are all part of the challenge. Success in this area is crucial for improving road safety, reducing traffic congestion, and enhancing fuel efficiency. Notably, human error is a leading cause of accidents (Ribeiro et al., 2023), underscoring the importance of AV technology.

The study in this paper is established on a number of statements. Although external conditions may occasionally alter data quality, it is expected that there is enough and trustworthy sensor data accessible for algorithmic processing, such as camera images and LiDAR outputs. Furthermore, it is assumed that sufficient computational resources are available, prioritizing conceptual and algorithmic developments over hardware constraints. Assuming that ethical frameworks will change in tandem with the technology, ethical considerations are crucial in the development of autonomous vehicles.

III. Key Works

In the area of lane detection and object recognition, Ribeiro et al. (2023) present an end-to-end approach for autonomous driving, emphasizing supervised learning methods and dataset creation. Their work integrates multiple functionalities into a unified system, addressing the challenges of generating diverse and realistic datasets essential for training robust models. Dong et al. (2024) provide a comprehensive overview of computer vision applications in autonomous vehicles, examining current techniques and their limitations. Their analysis highlights ongoing challenges in object detection and emphasizes the need for more efficient algorithms. Additionally, Chachlake et al. (2021) explore intelligent lane detection and tracking using deep learning. Their focus on leveraging advanced techniques demonstrates improvements in accuracy and reliability across varying driving conditions. In the domain of localisation and environmental robustness, Chuprov et al. (2023) tackle the difficulties of achieving reliable localisation under adverse environmental conditions. Their research focuses on developing algorithms capable of maintaining performance in challenging scenarios, such as Arctic environments, underscoring the importance of robustness in autonomous systems.

Finally, Unar et al. (2023) investigate the potential of image retrieval techniques for enhancing semantic location identification in navigation systems. This innovative approach represents a newer research direction, highlighting opportunities to improve efficiency and the overall autonomous driving experience.

IV. Evaluation Criteria

Evaluating the performance of computer vision systems for autonomous vehicles requires a thorough approach that considers both the datasets used for training and testing and the evaluation metrics applied. These factors significantly influence the interpretation of results and the overall assessment of system effectiveness.

The success of computer vision models largely depends on the quality and diversity of training data. Effective datasets must be large and cover various scenarios, including different weather, lighting, and object types, to create robust models capable of managing real-world challenges. Including rare and difficult cases is also crucial for handling unexpected conditions. Ribeiro et al. (2023) highlight that a lack of diverse data is still a major challenge in this field.

The ultimate test of a computer vision system's performance is its employment in real-world scenarios, even though simulated data might be helpful in enhancing real-world data. Datasets must include data collected under realistic driving scenarios to ensure their relevance. Existing datasets, such as KITTI and the Waymo Open Dataset, are widely used but often lack sufficient diversity or coverage of niche scenarios. Addressing these gaps is a key area of ongoing research (Ribeiro et al., 2023).

Evaluation metrics should be in line with the computer vision system's particular tasks. In the context of lane detection, recall evaluates the proportion of accurate detections to all real detections, precision evaluates the proportion of correct detections to total detections, and accuracy measures correctly detected lane markings, and the F1-score combines precision and recall. Robustness under different weather and lighting conditions is also important.For object detection, common metrics include precision, recall, and F1-score. Intersection over Union (IoU) measures the overlap between predicted and true bounding boxes, while Mean Average Precision (mAP) averages precision across recall levels. Systems should also be evaluated for detecting objects at various distances and under occlusion.

While orientation accuracy evaluates the difference in angles, positional accuracy calculates the difference between estimated and actual locations for localisation. Additionally crucial is robustness against environmental changes and sensor noise.

The accuracy with which pertinent images are recovered is measured by precision, recall, and F1-score in image retrieval. Computational cost and retrieval speed are other factors that can be used to assess efficiency.

V. Discussion

A. Lane Detection

Lane detection has improved with algorithms like Canny edge detection and Hough line transforms (Ribeiro et al., 2023). However, these methods face challenges in poor lighting, obstructed views, and inconsistent road markings. Deep learning approaches, as noted by Chachlake et al. (2021), show promise for creating more robust systems but struggle with real-time performance in dynamic settings. Their success depends heavily on diverse and well-annotated datasets (Ribeiro et al., 2023).

B. Object Detection

Deep learning models like YOLO and Faster R-CNN (Dong et al., 2024) have greatly improved object detection in autonomous vehicles, enabling accurate identification of pedestrians, vehicles, and bicycles. However, these models are computationally intensive, making real-time use challenging in resource-limited environments. Additionally, achieving reliable detection in harsh weather, such as heavy rain or snow, remains difficult. Developing more efficient and robust algorithms to handle complex scenarios with varying lighting and conditions is crucial.

C. Accurate Localisation

GPS is a reliable source of positional data but struggles in areas with signal blockages, such as urban canyons. Computer vision-based localization, using features from cameras or LiDAR, provides a valuable alternative (Chuprov et al., 2023). These methods show promise, even in extreme environments like the Arctic. However, their accuracy and reliability need improvement, especially in dynamic settings and under varying lighting and weather conditions.

D. Image Retrieval for Semantic Location

The application of image retrieval for semantic location identification represents a promising development (Unar et al., 2023). This approach can enhance user experience and navigational efficiency by allowing users to specify a desired location using an image query. The success of this approach is heavily dependent on the quality of the image dataset and the ability to extract relevant features that effectively capture semantic information of a location. Further research is needed to improve the robustness of this method, particularly regarding variations in lighting, viewpoint, and weather conditions.

.Conclusion

Progress in computer vision, especially in deep learning and image retrieval, has enhanced tasks such as lane detection, object recognition, and location. Nonetheless, obstacles persist, notably the substantial computing requirements of deep learning models and their dependence on high-quality data. Dependable systems require extensive, varied, and meticulously annotated datasets to manage multiple driving circumstances. Ensuring consistent performance in all weather and lighting is crucial for safety, and image retrieval techniques can further enhance navigation and user experience.

In conclusion, even though computer vision has advanced significantly for autonomous vehicles, much more needs to be done. The key to achieving the full potential of safe and extensively used autonomous driving systems is overcoming these obstacles through additional research.

References

[1]Dong, X., & Cappuccio, M. L. (2024). Applications of Computer Vision in Autonomous Vehicles: Methods, Challenges and Future Directions. Multimedia Tools and Applications, 83, 20537–20558. https://doi.org/10.1007/s11042-023-16387-6

[2]Chuprov, S., Viksnin, I., Belyaev, P., Gataullin, R., Reznik, L., & Neverov, E. (2023). Robust Autonomous Vehicle Computer-Vision-Based Localization in Challenging Environmental Conditions. Applied Sciences, 13, 5735. https://doi.org/10.3390/app13095735

[3]Ribeiro, I. A., Ribeiro, T., Lopes, G., & Ribeiro, A. F. (2023). End-to-End Approach for Autonomous Driving: A Supervised Learning Method Using Computer Vision Algorithms for Dataset Creation. Algorithms, 16(411). https://doi.org/10.3390/a16090411

[4]Unar, S., Su, Y., Liu, P., Wang, Y., & Cappuccio, M. L. (2023). Towards applying image retrieval approach for finding semantic locations in autonomous vehicles. Multimedia Tools and Applications, 83, 20537–20558. https://doi.org/10.1007/s11042-023-16387-6

[5]Kshirsagar, A., & Shah, M. (2021). Anatomized study of security solutions for multimedia: deep learning-enabled authentication, cryptography, and information hiding. In Advanced Security Solutions for Multimedia. IOP Publishing. https://doi.org/10.1088/978-0-7503-3735-9.ch7.

[6]Chachlake, A., Lee, J., Kim, S. H., & Kang, K. J. (2021). Intelligent lane detection and tracking method for autonomous driving based on deep learning. Expert Systems with Applications, 177, 114951. https://doi.org/10.1016/j.eswa.2021.114951.

[7] Amiri, Z., Heidari, A., Navimipour, N., & Unal, M. (2022). Resilient and dependability management in distributed environments: a systematic and comprehensive literature review. Cluster Computing, 26(6), 1565–1600. https://doi.org/10.1007/s10586-022-03462-0.

[8] Sahu, B. K., Kumar Sahu, B., Choudhury, J., & Nag, A. (2019). Development of hardware setup of an autonomous robotic vehicle based on computer vision using Raspberry Pi. In 2019 Innovations in Power and Advanced Computing Technologies. https://doi.org/10.1109/i-PACT44901.2019.8960011.

[9] Khokhlov, I., Chuprov, S., Reznik, L., & Manghi, K. (2022). Multi-Modal Sensor Selection with Genetic Algorithms. In 2022 IEEE Sensors. https://doi.org/10.1109/SENSORS55856.2022.10001234.

[10] Tseng, Y.-H., & Jan, S.-S. (2018). Combination of computer vision detection and segmentation for autonomous driving. Proceedings of the 2018 IEEE International Conference on Robotics and Automation (ICRA), 1047-