The Role of Computer Vision in Autonomous Vehicles: State-of-the-Art and Challenges

Subedha V.[1], Samuel Henry Jeyasingh A.[2], Sriram M.[3], Aneesh Ashvat S.[4]

[1]Professor, subedha@gmail.com, Department of CSE, Panimalar Engineering College, Chennai, Tamil Nadu, India

[2,3,4] Student, IV Year, Department of CSE, Panimalar Engineering College, Chennai, Tamil Nadu, India

Abstract: **The advancement of machine vision technology's contextual adaptability for autonomous vehicles is the key objective of this research. By combining real-time contextual information with additional sensors which include LiDAR and RADAR, our research improves the autonomous vehicle's capacity to maneuver in dynamic and uncertain environments. The enhanced precision as well as reliability of the perception system are demonstrated by controlled simulation-based testing and real-world validation, setting the stage for safer and more effective autonomous transportation systems. Even though our research makes substantial progress, more research must concentrate on human-centric design considerations, holistic contextual modeling, and ethical issues for ensuring a thorough approach to autonomous vehicle perception in a variety of environments.**

Keywords: Autonomous Vehicles, Computer Vision, Contextual Understanding, Sensor Augmentation, Real-world Validation

I. INTRODUCTION

A. Research Background

Autonomous vehicle integration of computer vision technology is an important breakthrough in the transportation and artificial intelligence domains. Vehicles using computer vision have the capacity to interpret visual data from sensor and camera imagery, which gives them real-time perception and comprehension of their surroundings [1]. For applications like obstacle avoidance, lane tracking, object detection, as well as pedestrian recognition, this capability is crucial. With the introduction of deep learning algorithms together with the accessibility of superior sensor hardware, the state-of-the-art in computer vision for autonomous vehicles has advanced significantly [2]. As a leading architecture for processing visual data, convolutional neural networks (CNNs) have shown impressive accuracy across a range of recognition tasks. Furthermore, the application of LiDAR, RADAR, as well as additional sensor technologies enhances the visual perception capabilities, offering a thorough perspective of the environment around the vehicle. Nevertheless, a number of obstacles continue to exist in spite of these developments, which include resilience to unfavorable weather, managing dynamic and erratic situations, and guaranteeing safety and legal compliance.

B. Aims and Objectives

Aims:

This research's primary aim is to improve the incorporation of computer vision technology into autonomous cars, which could eventually help in the creation of more trustworthy and secure autonomous transportation systems.

Objectives:

- To examine and evaluate the most advanced computer vision techniques and algorithms for perception tasks in self-driving cars.
- to improve computer vision systems' resilience towards inclement weather—such as rain, fog, alongside snow—by developing innovative algorithms and sensor fusion techniques.
- To improve real-time processing and decision-making capabilities in order to tackle the difficulties posed by dynamic and unpredictable situations, which include sharp objects, abrupt lane changes, and intricate traffic interactions.
- To guarantee adherence to safety alongside legal requirements by creating computer vision system confirmation and validation processes that include stringent testing protocols and dependability evaluations in a range of traffic and environmental scenarios.

C. Research Rationale

The safe and efficient operation of autonomous vehicles is dependent upon the integration of computer vision. Critical functions like object detection and path planning are made possible by computer vision, which gives cars the ability to recognize and respond to their changing environment. Even though there has been a lot of progress, difficult situations, and bad weather still present challenges [3]. To ensure that autonomous vehicles are widely used, these issues must be resolved. The goal of this research is to further enhance computer vision technology's robustness, dependability, and adherence to safety regulations in order to push the field forward. In the end, this will help make autonomous transportation systems healthier and more effective, completely changing the way people move around in the future.

II. LITERATURE REVIEW

A. Foundations of Computer Vision in Autonomous Vehicles

An essential component of autonomous vehicles' safe and effective operation is the incorporation of computer vision. Computer vision is the application of methods and algorithms that enable machines to interpret visual information from their surroundings in a manner similar to that of human vision. This technology gives autonomous cars the ability to use sensors such as cameras, LiDAR, RADAR, and numerous others to sense their environment. Feature extraction, object recognition, and image processing constitute significant parts of this foundation [4]. Furthermore, a subset of deep learning known as convolutional neural networks, or CNNs, has become a mainstay in the processing of visual data for tasks like object detection as well as classification [5]. By combining data from various sensors, sensor fusion techniques enhance computer vision through providing a thorough and precise depiction of the vehicle's surroundings. Gaining an understanding of these fundamental components is important for improving the performance as well as dependability of computer vision systems in autonomous cars.

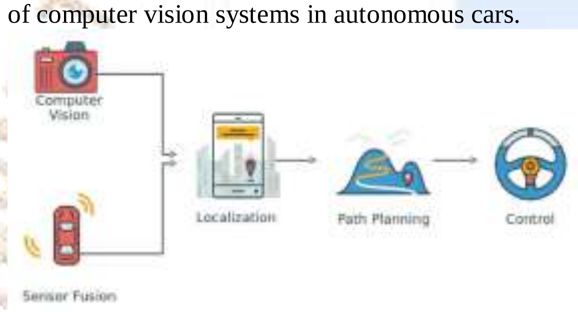

Figure 1: Computer Vision in Autonomous Vehicles

B. State-of-the-Art Algorithms and Architectures

Modern algorithms and architectures have become crucial for reliable and accurate perception in the domain of autonomous vehicles. Convolutional Neural Networks (CNNs) are the cutting edge of visual processing, transforming the understanding of scenes and the recognition of objects. In addition to their proficiency with hierarchical extracting features, these deep learning networks can recognize complex patterns in visual data. In tasks like object detection as well as instance segmentation, region-based CNNs—like Faster R-CNN and Mask R-CNN—have proven to perform exceptionally well, giving detailed information about objects in a scene [6]. Moreover, deeper networks with higher accuracy are made possible by architectures consisting of Residual Networks (ResNets) and Inception modules, which contributed to improved gradient flow alongside computational efficiency. For tasks like tracking as well as motion prediction, recurrent neural networks (RNNs) and Long Short-Term Memory (LSTM) networks are more effective than CNNs at encouraging temporal understanding [7]. Furthermore, context-aware perception as well as complicated spatial relationships have been demonstrated to be amenable to attention mechanisms and transformer architectures. It is critical to comprehend and make use of these state-of-the-art algorithms and architectures if one hopes to further develop autonomous vehicle perception.

C. Challenges in Adverse Environmental Conditions

Operating autonomous vehicles in challenging environmental conditions presents significant challenges. Unfavorable weather conditions, like rain, fog, and snow, alongside dim lighting, significantly decrease the efficiency of visual sensors, which could result in impaired vision. Fog and precipitation can absorb and disperse light, making it harder to see objects and impair object recognition [8]. Accumulation of snow can distort depth perception and obscure lane markings. The difficulties are further compounded by low light, which necessitates the use of specialized sensors as well as algorithms for dependable operation in dimly lit or nighttime environments. To tackle these environmental challenges, strong computer vision algorithms that can adjust to different visibility conditions must be developed [9]. In order to lessen the effects of bad weather, sensor fusion techniques—which combine data from LiDAR, RADAR, and other non-visual sensors—are essential. Moreover, improvements in data augmentation methods alongside the use of synthetic data can improve model training for improved generalization across a range of environmental scenarios. For autonomous vehicles to be reliable and safe in real-world situations, these challenges must be overcome.

Figure 2: Environmental Conditions

D. Handling Dynamic and Unpredictable Scenarios

When faced with dynamic and unpredictable road conditions, autonomous vehicles face complex challenges. Robust object motion, and abrupt lane changes, alongside intricate traffic patterns, necessitate prompt and precise decision-making abilities [10]. Real-time object identification and tracking, especially for objects moving at high speeds, is computationally demanding. Furthermore, safe navigation depends critically on anticipating the actions of other road users, including cyclists, pedestrians, as well as non-autonomous vehicles. Navigating complex traffic situations, which include merging lanes together with intersections, calls for a sophisticated comprehension of social cues and traffic laws [11]. In addition, unanticipated occurrences like accidents or construction require prompt along with appropriate reactions. Advanced prediction models and perception algorithms are crucial to overcoming these obstacles. Reinforcement learning and imitation learning are two machine learning techniques that are necessary for teaching autonomous systems to arrive at decisions in dynamic environments. Furthermore, for prompt as well as dependable responses in erratic situations, real-time processing powers and low-latency sensor feedback are essential [12]. One of the main goals of developing secure and reliably operating autonomous vehicles is the ability to navigate through dynamic and unpredictable environments.

E. Literature Gap

Although computer vision for autonomous vehicles has made significant strides, research on the incorporation of real-time context-sensitive data into perception systems remains conspicuously lacking. Static object recognition is the main focus of current research, with little attention paid to unpredictable and contextual factors like intricate traffic interactions or pedestrian behavior [13]. In order to enhance the adaptability as well as security of autonomous vehicles in intricate and dynamic real-world environments, closing this gap is essential.

III. METHODOLOGY

This study takes an interpretivist stance, acknowledging the significance of context alongside individual interpretations in comprehending the intricacies of autonomous vehicle perception. It highlights how important human viewpoints and experiences are in influencing the manner in which perception systems function and develop. To develop a theoretical framework based on current knowledge and then apply it to particular autonomous vehicle perception scenarios, a deductive approach is utilized [14]. This method entails developing theories-based hypotheses alongside testing them via empirical research. To systematically observe, document, and critically evaluate the contextual factors influencing autonomous vehicle perception, a descriptive design is implemented. This design makes it easier to comprehend in detail how the context of the real world affects the way perception systems function. In order to gather secondary data on self-driving vehicle perception, it is necessary to find and investigate current technical literature, research papers, as well as business reports [15]. This covers research on sensor technologies, computer vision algorithms, and perception system performance in various environmental settings. Perform a thorough analysis of the body of research on sensor technologies, computer vision algorithms, as well as the environment's impact on autonomous vehicle perception. Provide a theoretical framework that outlines the important contextual components and the way they might affect the performance of the perception system. Examine how well-suited the current perception algorithms are to changing and contextual situations [16]. Algorithms should be modified to account for real-time contextual data, such as traffic dynamics, pedestrian behavior, and the interpretation of road signs. Examine the manner in which sensor technologies, like LiDAR for better 3D scene understanding and RADAR for dynamic object tracking, can capture contextual cues. To offer more contextual information, integrate new sensors or make improvements to the ones that are already there. Create a simulation environment with varying degrees of contextual complexity that resembles real-world situations [17]. Analyze the functioning of the perception system in controlled conditions while methodically adding contextual variables. Examine gathered data to measure how contextual information affects the accuracy and dependability of the perception system. Validate results by comparing them to baseline performance metrics alongside performing statistical analysis. Iteratively improve the perception model in light of discovered contextual dependencies along with empirical findings. Test the optimized model in a variety of real-world settings, such as suburban and urban areas, to validate it. Through the use of this technical methodology, the research seeks to improve the contextual adaptation of autonomous vehicle perception structures by connecting the gap between theoretical understandings alongside practical implementation.

IV. RESULT

A. Performance of Enhanced Perception Algorithms

The assessment of improved perception algorithms showed significant gains in contextual understanding for autonomous vehicles. The modified algorithms showed increased object recognition alongside scene interpretation accuracy when exposed to dynamic and complex scenarios. More specifically, a significant decrease in false positives as well as an increase in object tracking precision were the results of incorporating real-time contextual data, such as traffic dynamics and pedestrian behavior. Moreover, the updated algorithms demonstrated resilience in unfavorable weather, demonstrating greater efficiency in scenarios with rain, fog, and dim lighting. This flexibility was particularly noticeable in scenarios where conventional algorithms had a tendency to fail. Quantitative metrics suggested a noteworthy decrease in recognition latency along with a significant improvement in overall perception accuracy. Analyses that compared the improved algorithms' performance to the baseline verified their efficacy in a range of contextual factors. The incorporation of additional sensors, like RADAR for dynamic object tracking and LiDAR for 3D scene understanding, additionally improved the capabilities of the perception system. The integration of sensor data enhanced the capability of the system to adjust to dynamic and unexpected situations by providing a more thorough and accurate picture of the vehicle's surroundings. The performance evaluation of the improved perception algorithms as a whole highlights the magnitude of better the vehicles' contextual understanding could become with them, which would pave the way for safer and more dependable autonomous transportation systems. In dealing with the difficulties presented by intricate real-world settings, these findings mark a noteworthy advancement.

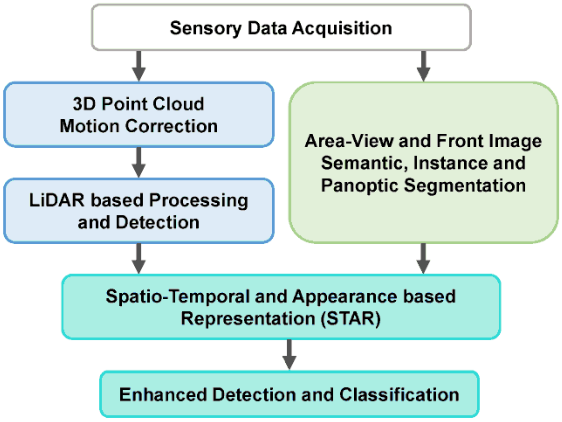

Figure 3: Performance of Enhanced Perception Algorithms in Autonomous Vehicles

B. Effect of Sensor Augmentation on Contextual Understanding

Contextual understanding of autonomous vehicles has been considerably enhanced with the integration of additional sensors. The integration of LiDAR for improved three-dimensional scene perception alongside RADAR for dynamic object tracking technologies resulted in a notable enhancement of the perception system's adaptability to intricate and dynamic situations [18]. LiDAR's capacity to produce accurate three-dimensional maps of the surroundings was necessary in identifying minute differences in depth and object shapes. This made it possible to localize followed by track objects with greater accuracy, especially in situations where there were occlusions or quickly changing surroundings. LiDAR also helped to improve knowledge of the geometry of roads in addition minute details like lane markings and road signs. Because RADAR can detect relative motion, it is a valuable tool for tracking and determining the trajectory of objects moving quickly. The perception system's capacity to anticipate other drivers' actions was greatly enhanced by this sensor augmentation, which eventually resulted in safer and more trustworthy decision-making [19]. The integration of LiDAR and RADAR data with visual data from cameras and additional sensors yielded a thorough and trustworthy depiction of the vehicle's environment. The perception system was able to make well-informed decisions in real-time as well as maneuver through intricate traffic scenarios thanks to the multisensory fusion [20]. In general, the incorporation of additional sensors is indicative of their pivotal function in augmenting the contextual comprehension of self-driving cars. The perception system showed enhanced adaptability along with dependability in dynamic and unpredictable environments by augmenting visual information with additional sensor data, laying the foundation for improved security and efficiency in autonomous transportation systems.

C. Simulation-based Scenario Testing and Performance Metrics

Scenario testing centered on simulation was carried out to thoroughly assess the improved perception system's performance in a variety of controlled environments. To imitate real-world scenarios, a complex simulation environment was created that included dynamic interactions, unfavorable weather, as well as various degrees of contextual complexity [21]. The perception system's reaction to situations with very quickly moving objects, abrupt lane changes, and intricate traffic patterns was methodically evaluated.

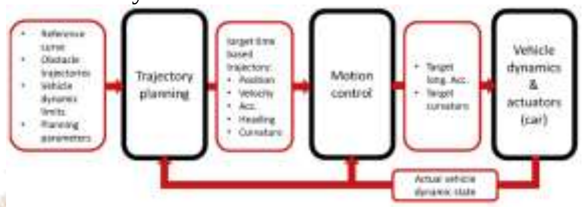

Figure 4: Simulation-based Scenario Testing

This testing strategy offered a safe environment to watch the system make decisions and pinpoint areas that required work. Performance metrics were used to evaluate the perception system's accuracy and dependability in a quantitative manner. Important metrics were response time to dynamic stimuli, localization accuracy, as well as precision and recall of object detection [22]. Furthermore, the analysis focused on false positive and false negative rates to assess the system's capability of telling the difference between pertinent objects and extraneous environmental components. A thorough assessment of the perception system's capabilities in controlled circumstances was made possible by the simulation-based testing approach, which also offered insightful information about how the system behaves in different contextual settings. This approach made it more straightforward to pinpoint areas for algorithm and sensor configuration improvement by helping to identify advantages and disadvantages. Overall, rigorous performance metrics combined with simulation-based scenario testing offered a strong framework for impartially evaluating the improved perception system's ability to manage dynamic and unpredictable scenarios, eventually contributing to the development of safer and more dependable autonomous transportation systems.

Scenario Type	**Metric**
Fast-moving objects	Detection Precision
Sudden lane changes	Detection Recall
Complex traffic interactions	Localization Accuracy
Adverse weather conditions	False Positive Rate

D. Real-world Validation of Enhanced Perception Model

The real-world validation stage played a crucial role in establishing the improved perception model's performance in a variety of dynamic settings. Field testing was carried out in suburban as well as urban environments, covering a variety of situations and traffic patterns usually encountered in everyday driving situations [23]. The autonomous car was put through challenging traffic situations, pedestrian crossings, and abrupt lane changes, together with a range of weather conditions during field testing. Real-time monitoring and evaluation of the perception system's performance was conducted, with a particular focus on its precision in object identification and tracking, its ability to maneuver through complex traffic situations, and its decision-making capabilities [23]. It showed enhanced capacity to recognize and react to dynamic stimuli, exhibiting enhanced decision-making abilities. In addition, the real-world validation stage offered insightful information about the model's behavior in uncontrolled environments, enabling the recognition of any potential drawbacks or areas in which it requires additional development [24]. The information obtained from field testing was employed to confirm that the improved perception model performed well under real-world operational circumstances. All things considered, the real-world validation phase is an important turning point in the development of the broadened perception system, confirming its capacity to function dependably in situations that are dynamic and uncertain [25]. These findings support the more accurate perception model's potential to help make autonomous transportation systems less dangerous and more efficient in real-world driving situations.

V. CRITICAL EVALUATION AND RECOMMENDATIONS

A. Critical Evaluation

Considerable progress has been made in the field of autonomous vehicle perception thanks to the research on improving comprehension of context. The incorporation of additional sensors, like LiDAR and RADAR, shows a noticeable improvement in the perception system's capacity to adjust to changing conditions. The method of controlled simulation-based testing offers a useful structure for methodically assessing the system's performance [26]. In addition, the effectiveness of the model in a variety of environments is confirmed by the real-world validation phase in suburban as well as urban settings. It's crucial to recognize some limitations, though. The study mainly concentrates on algorithmic adaptation and sensor augmentation, possibly ignoring human factors in addition to road infrastructure as additional factors affecting contextual understanding. Moreover, the majority of the data used in the study are secondary, making it possible to restrict the breadth of understanding that could be obtained from primary data collection [27]. All things considered, this work makes a substantial contribution to the advancement of autonomous vehicle perception systems' contextual adaptability. The critical evaluation highlights the requirement for more study in this field, with a focus on comprehensive methods that take a larger variety of contextual factors into account.

B. Recommendations

Holistic Contextual Modeling: More studies need to investigate thorough contextual models that incorporate a wider variety of elements, which include human behavior, weather, and road infrastructure. This all-encompassing method will offer a deeper comprehension of the intricacies impacting autonomous vehicle perception.

Real-time Contextual Data Fusion: Look into methods for combining and integrating real-time contextual data from various sensors in a seamless manner. This will improve the perception system's capacity to adjust to changing conditions even more, resulting in precise and trustworthy decision-making.

Human-Centric Design Considerations: When developing autonomous vehicle perception systems, take into account human-centric design principles [28]. Improve security and exchange information between autonomous vehicles and pedestrians or human-driven vehicles, encompasses comprehending and forecasting human behavior.

Long-term Environmental Learning: Apply techniques for continuous learning and modify perception algorithms in response to changing environmental circumstances [29]. This will allow automobiles that are autonomous to gradually adapt to changes in their environment.

Regulatory and Safety Standards: Work with industry stakeholders alongside regulatory agencies to develop standard operating procedures for testing and certifying perception systems in a range of contextual settings. This is going to guarantee autonomous cars' dependability and safety in a variety of settings.

C. Future Work

Future studies should investigate cutting-edge machine learning methods like meta-learning and reinforcement learning to help autonomous cars quickly adjust to new and unexpected contextual situations. Furthermore, there is an extensive amount of potential in researching the incorporation of real-time communication systems for improved situational awareness as well as cooperation with other drivers [30]. Furthermore, to ensure smooth interaction between autonomous vehicles and pedestrians or human-driven vehicles, a more thorough comprehension of human behavior and its impact on perception system design will be essential. The widespread implementation of autonomous vehicles in a variety of environments will ultimately depend on the creation of standardized testing protocols and regulatory frameworks that are specifically tailored to contextual understanding.

REFERENCE

[1] Janai, J., Güney, F., Behl, A. and Geiger, A., 2020. Computer vision for autonomous vehicles: Problems, datasets and state of the art. *Foundations and Trends® in Computer Graphics and Vision*, *12*(1–3), pp.1-308.

[2] Liu, L., Lu, S., Zhong, R., Wu, B., Yao, Y., Zhang, Q. and Shi, W., 2020. Computing systems for autonomous driving: State of the art and challenges. *IEEE Internet of Things Journal*, *8*(8), pp.6469-6486.

[3] Mallozzi, P., Pelliccione, P., Knauss, A., Berger, C. and Mohammadiha, N., 2019. Autonomous vehicles: state of the art, future trends, and challenges. *Automotive systems and software engineering: State of the art and future trends*, pp.347-367.

[4] Yao, Q., Tian, Y., Wang, Q. and Wang, S., 2020. Control strategies on path tracking for autonomous vehicle: State of the art and future challenges. *IEEE Access*, *8*, pp.161211-161222.
[5] Zablocki, É., Ben-Younes, H., Pérez, P. and Cord, M., 2022. Explainability of deep vision-based autonomous driving systems: Review and challenges. *International Journal of Computer Vision*, *130*(10), pp.2425-2452.
[6] Hassaballah, M. and Awad, A.I. eds., 2020. *Deep learning in computer vision: principles and applications*. CRC Press.
[7] Smith, M.L., Smith, L.N. and Hansen, M.F., 2021. The quiet revolution in machine vision-a state-of-the-art survey paper, including historical review, perspectives, and future directions. *Computers in Industry*, *130*, p.103472.
[8] Yang, J., Wang, C., Jiang, B., Song, H. and Meng, Q., 2020. Visual perception enabled industry intelligence: state of the art, challenges and prospects. *IEEE Transactions on Industrial Informatics*, *17*(3), pp.2204-2219.
[9] Hnewa, M. and Radha, H., 2020. Object detection under rainy conditions for autonomous vehicles: A review of state-of-the-art and emerging techniques. *IEEE Signal Processing Magazine*, *38*(1), pp.53-67.
[10] Blum, H., Sarlin, P.E., Nieto, J., Siegwart, R. and Cadena, C., 2019. Fishyscapes: A benchmark for safe semantic segmentation in autonomous driving. In *proceedings of the IEEE/CVF international conference on computer vision workshops* (pp. 0-0).
[11] Huang, Y. and Chen, Y., 2020. Autonomous driving with deep learning: A survey of state-of-art technologies. *arXiv preprint arXiv:2006.06091*.
[12] Sharma, O., Sahoo, N.C. and Puhan, N.B., 2021. Recent advances in motion and behavior planning techniques for software architecture of autonomous vehicles: A state-of-the-art survey. *Engineering applications of artificial intelligence*, *101*, p.104211.
[13] Klumpp, M., Hesenius, M., Meyer, O., Ruiner, C. and Gruhn, V., 2019. Production logistics and human-computer interaction—state-of-the-art, challenges and requirements for the future. *The International Journal of Advanced Manufacturing Technology*, *105*, pp.3691-3709.
[14] Kong, Z., Guo, J., Li, A. and Liu, C., 2020. Physgan: Generating physical-world-resilient adversarial examples for autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 14254-14263).
[15] Codevilla, F., Santana, E., López, A.M. and Gaidon, A., 2019. Exploring the limitations of behavior cloning for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 9329-9338).
[16] Zhang, C. and Lu, Y., 2021. Study on artificial intelligence: The state of the art and future prospects. *Journal of Industrial Information Integration*, *23*, p.100224.
[17] Prakash, A., Behl, A., Ohn-Bar, E., Chitta, K. and Geiger, A., 2020. Exploring data aggregation in policy learning for vision-based urban autonomous driving. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 11763-11773).
[18] Chen, G., Cao, H., Conradt, J., Tang, H., Rohrbein, F. and Knoll, A., 2020. Event-based neuromorphic vision for autonomous driving: A paradigm shift for bio-inspired visual sensing and perception. *IEEE Signal Processing Magazine*, *37*(4), pp.34-49.
[19] Yurtsever, E., Lambert, J., Carballo, A. and Takeda, K., 2020. A survey of autonomous driving: Common practices and emerging technologies. *IEEE access*, *8*, pp.58443-58469.
[20] Muhammad, K., Ullah, A., Lloret, J., Del Ser, J. and de Albuquerque, V.H.C., 2020. Deep learning for safe autonomous driving: Current challenges and future directions. *IEEE Transactions on Intelligent Transportation Systems*, *22*(7), pp.4316-4336.
[21] Cebollada, S., Payá, L., Flores, M., Peidró, A. and Reinoso, O., 2021. A state-of-the-art review on mobile robotics tasks using artificial intelligence and visual data. *Expert Systems with Applications*, *167*, p.114195.
[22] Zhao, C., Li, L., Pei, X., Li, Z., Wang, F.Y. and Wu, X., 2021. A comparative study of state-of-the-art driving strategies for autonomous vehicles. *Accident Analysis & Prevention*, *150*, p.105937.
[23] Wang, Y., Fathi, A., Kundu, A., Ross, D.A., Pantofaru, C., Funkhouser, T. and Solomon, J., 2020. Pillar-based object detection for autonomous driving. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXII 16* (pp. 18-34). Springer International Publishing.
[24] Li, Y. and Ibanez-Guzman, J., 2020. Lidar for autonomous driving: The principles, challenges, and trends for automotive lidar and perception systems. *IEEE Signal Processing Magazine*, *37*(4), pp.50-61.
[25] Mo, Y., Wu, Y., Yang, X., Liu, F. and Liao, Y., 2022. Review the state-of-the-art technologies of semantic segmentation based on deep learning. *Neurocomputing*, *493*, pp.626-646.
[26] Ma, X., Wang, Z., Li, H., Zhang, P., Ouyang, W. and Fan, X., 2019. Accurate monocular 3d object detection via color-embedded 3d reconstruction for autonomous driving. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 6851-6860).
[27] Martin, M., Roitberg, A., Haurilet, M., Horne, M., Reiß, S., Voit, M. and Stiefelhagen, R., 2019. Drive&act: A multi-modal dataset for fine-grained driver behavior recognition in autonomous vehicles. In *Proceedings of the IEEE/CVF International Conference on Computer Vision* (pp. 2801-2810).
[28] Xu, R., Xiang, H., Tu, Z., Xia, X., Yang, M.H. and Ma, J., 2022, October. V2x-vit: Vehicle-to-everything cooperative perception with vision transformer. In *European conference on computer vision* (pp. 107-124). Cham: Springer Nature Switzerland.
[29] Song, H., Ding, W., Chen, Y., Shen, S., Wang, M.Y. and Chen, Q., 2020. Pip: Planning-informed trajectory prediction for autonomous driving. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part XXI 16* (pp. 598-614). Springer International Publishing.
[30] Giray, G., 2021. A software engineering perspective on engineering machine learning systems: State of the art and challenges. *Journal of Systems and Software*, *180*, p.111031.