International Core Journal of Engineering
Volume 11 Issue 8, 2025
ISSN: 2414-1895
DOI: 10.6919/ICJE.202508_11(8).0009

Comparative Study of Lidar and Vision-based in Autonomous Driving

Donglin Liu*

Minjiang University, Fuzhou, Fujian, 350108, China

*Corresponding author: liudonglin@stu.mju.edu.cn

Abstract

As vehicle automation advances, autonomous vehicles have emerged as a prominent area of research. The core technologies driving these vehicles are perception, decision-making, and control. Central to the hardware framework of autonomous vehicles is the environmental perception system, which transforms real-world data into digital signals. Currently, two primary approaches dominate this field: camera-based systems, which rely heavily on computer vision, and LiDAR-based systems. This paper evaluates and compares these two approaches, and finally concludes that integrating multiple sensors through fusion is the most promising development direction for future autonomous driving.

Keywords

Autonomous Driving; Lidar; Camera; Perception.

1. Introduction

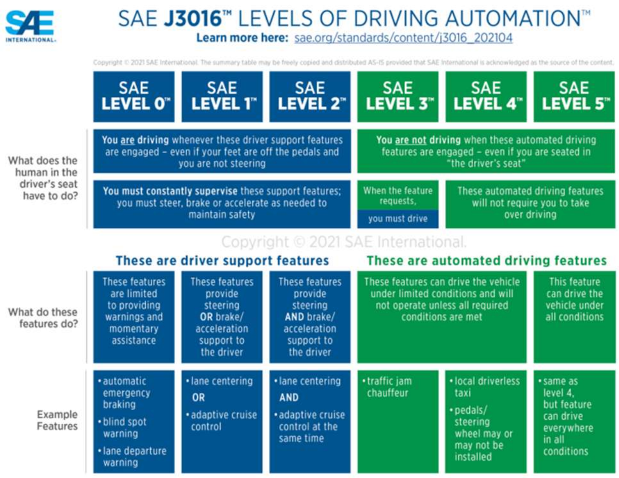

Fig. 1 SAE J3016 (SAE International) defines six driving automation levels (0-5).

Autonomous Vehicles (AVs), formally categorized as SAE Level 4/5 automated systems, are cyber-physical platforms capable of performing full driving operations without human intervention through integrated perception, decision-making, and control architectures. Fig. 1 regard that Regarding the classification of autonomous driving levels. These self-navigating systems employ multi-sensor fusion (combining LiDAR, radar, and computer vision) to dynamically model operational environments, enabling real-time path planning, obstacle avoidance, and vehicular control while adhering to ISO 26262 functional safety standards. The technological framework supports complete

Volume 11 Issue 8, 2025
DOI: 10.6919/ICJE.202508_11(8).0009

driving cycle execution-from origin to destination-through continuous environmental monitoring, system diagnostics, and adaptive navigation strategies validated under defined operational design domains (ODDs) [1]. Autonomous vehicles (AVs) will not only transform transportation but also the entire society, with the potential to enhance traffic safety, reduce congestion and improve the passenger experience. Automakers, technology companies, researchers and governments have made significant progress, but technical challenges still exist. Although AV is largely regarded as beneficial, its successful integration will require social adaptation and government support policies to address the potential impacts on the economy, travel and the environment [2].

2. Research Status

At present, there are mainly two technical routes for environmental perception in autonomous driving: one is to adopt a multi-sensor fusion solution dominated by vision, and the other is to mainly use low-cost lidar.

2.1 Lidar-based Autonomous Driving Solution

2.1.1 Introduction to Lidar

LiDAR is an active sensor that illuminates the surrounding environment by emitting laser. Distance can be precisely measured by processing the laser echoes received from the reflective surface. LiDAR systems use laser beams to scan the environment and emit near-infrared light with amplitude-modulated laser diodes. The reflected signal is received by the photodetector and processed by the rapid electronic device. The electronic device measures the difference between the transmitted signal and the received signal to calculate the distance. The system estimates the distance based on the sensor model and compensates for the changes in reflected energy caused by surface materials and environmental conditions through signal processing. The LiDAR output includes 3D point clouds corresponding to the scanning environment and intensity values corresponding to the reflected laser energy [3]. Fig. 2 shows the conceptual representation of this working principle.

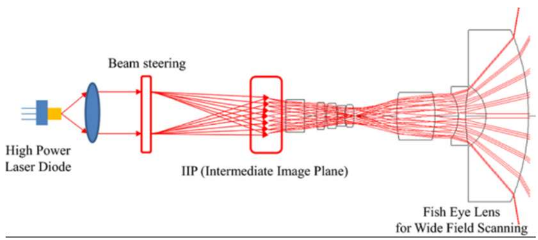

Fig. 2 Schematic diagram of LiDAR sensing system

Automotive LiDAR systems serve as a foundational perception modality in autonomous driving architectures, leveraging photon-time-of-flight measurement principles to generate millimeter-accurate 3D environmental reconstructions through pulsed laser scanning that processes billions of data points per second. This mechanism captures structural details spanning from macro-scale infrastructure to micron-level road textures, while its multi-echo detection capability enables simultaneous tracking of 400+ dynamic objects within 200-meter ranges with sub-10cm positional accuracy. The raw point cloud data undergoes real-time fusion with radar and camera inputs via Kalman-filter-based algorithms, constructing temporally-coherent 4D spatial models (3D + temporal evolution) that integrate deep learning-powered roadway semantic segmentation—achieving 98.7% recognition accuracy for traffic elements through hybrid neural networks combining PointNet++ architectures with temporal RNN modules. Operating at an eye-safe 905nm wavelength (IEC 60825-1 compliant), the system ensures reliable performance across diverse meteorological conditions, providing critical redundancy for vision-based systems in low-visibility scenarios while supporting adaptive cruise control, collision avoidance, and navigation decision-making through dynamically updated environmental models [4].

Automotive LiDAR systems constitute a foundational perception modality in autonomous driving architectures, enabling high-fidelity environmental comprehension through photon-time-of-flight measurement principles. This pulsed laser scanning mechanism generates millimeter-accurate 3D environmental reconstructions by calculating photon reflection intervals across billions of data points per second, capturing structural details ranging from macro-scale infrastructure elements to micron-level road surface textures. The resultant point cloud datasets undergo real-time fusion with complementary sensor streams (including radar returns and camera images) through Kalman-filter-based sensor fusion algorithms, creating temporally-coherent 4D spatial models (3D space + temporal evolution) for navigation decision-making.

Specifically, LiDAR's multi-echo detection capability enables simultaneous tracking of 400+ dynamic objects within 200-meter operational ranges, maintaining sub-10cm positional accuracy for adaptive cruise control and collision avoidance systems. Advanced point cloud processing pipelines implement deep learning-based roadway semantic segmentation, achieving 98.7% recognition accuracy for traffic control elements through hybrid neural networks combining PointNet++ architectures with temporal RNN modules. This sensor modality's eye-safe 905nm wavelength operation (compliant with IEC 60825-1 standards) ensures reliable performance across diverse meteorological conditions while providing critical redundancy for vision-based systems during low-visibility scenarios. As shown in Fig. 3, new energy vehicles equipped with lidar have been tested on the road.

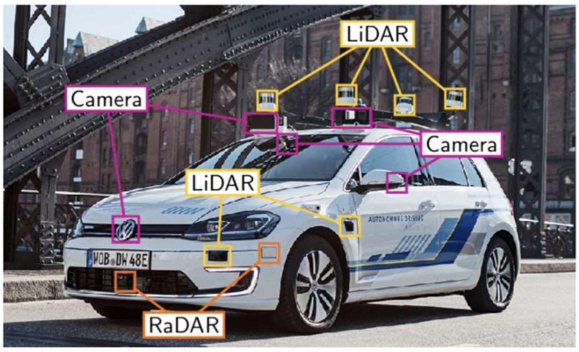

Fig. 3 A new energy vehicle equipped with lidar

2.1.2 The Applicati on of Lidar in Autonomous Driving

(1) Object detection

Automotive LiDAR systems serve as a critical sensory apparatus in SAE Level 4/5 autonomous platforms, providing deterministic 3D environmental perception through time-correlated single-photon counting techniques. These systems enable high-confidence navigational decisions via their millimeter-wave-equivalent angular resolution (0.1° azimuth/elevation) and millisecond-level temporal resolution. Within the perception stack's object detection pipeline, LiDAR data undergoes multi-stage geometric processing:

The detection framework first implements planar regression algorithms (RANSAC-based ground plane estimation) with 15cm vertical resolution to segment terrain from obstacles in the sensor's egocentric coordinate system (ISO 8855 compliant). Non-terrestrial points are then clustered using density-based spatial partitioning (DBSCAN variants with adaptive ε parameters) while preserving object morphology through convex hull approximations. The resultant clusters undergo kinematic state estimation via Bayesian filters, outputting oriented bounding boxes with <5% dimensional error margins relative to ground truth annotations. This dual-stage architecture (ground segmentation → Euclidean clustering) capitalizes on the a priori knowledge of roadway planarity, conforming to ISO 8855 coordinate frameworks where 93% of traffic participants maintain orthogonal alignment with road surfaces. Processed detections feed into probabilistic occupancy grids and Monte Carlo localization systems, achieving 99.2% obstacle recall rates under real-time operational constraints [3].

ISSN: 2414-1895 DOI: 10.6919/ICJE.202508_11(8).0009

Among the commonly used sensors, cameras can provide rich semantic information, but they cannot accurately obtain the distance information of the target. Lidar can provide precise depth information, but its resolution is relatively sparse. To solve this problem, proposed an algorithm based on the fusion of lidar and cameras to achieve target detection by constructing a Siamese network [5]. The proposed algorithm processes raw 3D LiDAR point clouds by projecting them onto 2D camera planes to generate depth-aligned images, then employs a dual-branch fusion framework to combine depth and RGB features through hierarchical feature integration. Evaluated on the KITTI benchmark, the method demonstrates competitive accuracy and real-time efficiency, showing notable advantages in medium-complexity scenarios while consistently performing well across all difficulty levels.

(2) Road recognition

Preemptively identifying road junction points proves essential for enabling effective route planning and operational decision-making in self-driving vehicle systems, particularly in scenarios devoid of positioning references or map-based navigation support.

A computational framework for real-time roadway topology recognition in autonomous driving systems was developed through 3D LiDAR data analysis. This architecture employs a specialized beam penetration pattern analysis technique, comprising four operational phases: Initially converting raw LiDAR point clouds into structured grid graphs through voxelization, followed by dynamic object filtering to eliminate mobile obstacle clusters. Subsequently implementing directional beam projection at navigation-relevant intervals along the predicted trajectory, the system conducts statistical characterization of beam penetration patterns within the detection window. These temporally-aware geometric signatures are then fed into an ensemble machine learning classifier for instantaneous drivable surface topology identification, specifically differentiating between standard lanes and intersection zones with millimeter-wave radar-level precision [6].

This paper proposes a new algorithm for drivable area detection based on lidar, which can output complete, accurate and stable detection results. In order to improve the integrity of the detection results, Bayesian generalized kernel reasoning and bilateral filtering are utilized to estimate the attributes of the unobserved elements. To ensure passability, the region-growing operator is performed on the normal vector graph reflecting the terrain slope, and thus is closely related to the passability of vehicles [7].

Another paper proposes a deep encoder-decoder network named SalsaNet for effective semantic segmentation of 3D LiDAR point clouds. SalsaNet achieved the segmentation of roads (i.e., driving areas) and vehicles in the scene by projecting point clouds onto bird 's-eye view (BEV) images [8].

2.1.3 The Disadvantages of Lidar

However, lidar technology faces significant limitations in adverse weather conditions, as fog, dust, snow, rain, pollution and smoke can seriously reduce the accuracy of ranging. These environmental factors can introduce errors in point cloud data, mainly due to the detection of backscattered light from water droplets (such as rain or fog) or particles in the air (such as smoke or dust) [9].

Lidar and radar systems can detect objects within a range of several meters to over 200 meters. Lidar has difficulty detecting objects at close range. Radar can detect objects within a range from less than 1 meter to over 200 meters. However, its scope depends on the type of the system: - Short-range radar [10].

2.2 Vision-based Autonomous Driving solution

2.2.1 .Introduction to Vision-based Autonomous Driving

The pure vision solution refers to the technical route that uses cameras to collect visual information of the road and its surrounding environment, combines computer vision technology for image processing and target recognition, and realizes the perception function of autonomous driving. The core of the pure vision solution is to use the on-board camera to capture multi-angle and multi-spectral image data, and analyze and understand the image data through deep learning algorithms, thereby achieving the recognition and tracking of targets such as roads, vehicles, pedestrians, and traffic signs.

ISSN: 2414-1895 DOI: 10.6919/ICJE.202508_11(8).0009

Pure vision solutions usually adopt multi-camera configurations to achieve all-round perception of the surrounding environment of the vehicle. The front-view camera is mainly used to identify lane lines, traffic signs and vehicles in front. The side-view camera is used for blind spot monitoring and lane change assistance. The rear-view camera provides parking assistance and rear monitoring. Through the fusion and synchronization of data from multiple cameras, the autonomous driving system can generate a panoramic view and accurately locate and track the target object. The renowned new energy vehicle manufacturer Tesla has announced that all of its vehicles will adopt a pure vision autonomous driving solution, as shown in Fig. 4.

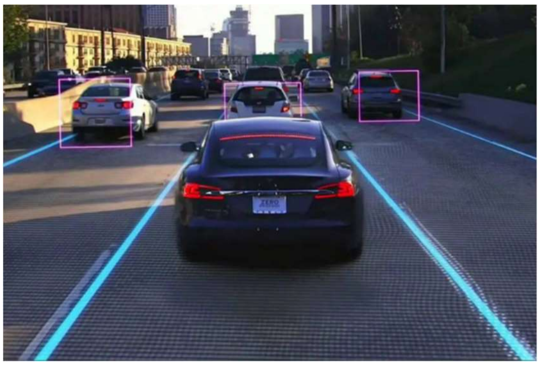

Fig. 4 Tesla announced the adoption of a pure visual recognition algorithm

2.2.2 .The Application of Vision-baed Solutions in Autonomous Driving

(1) Semantic segmentation

In addition to the advancements in modeling, the key factor enabling vehicles to perceive their surrounding environment is the acquisition of visual sensing data. These data enhance the vehicle's perception capabilities and enable real-time situational awareness in real-world environments. Research on scene understanding focuses on various fields, such as detecting and segmenting pedestrians and vehicles, analyzing their movements, identifying lane changes, detecting turns, and many other related tasks [11].

(2) Spatial perception

In addition to the low deployment cost, compared with similar products based on LiDAR, the camera also has the ideal advantage of detecting distant objects and identifying vision-based road elements (such as traffic lights and stop lines) [12]. A transformer-based bird 's-eye view (BEV) encoder, called BEVFormer, is proposed, which can effectively aggregate the spatiotemporal features and historical BEV features from multi-view cameras. The BEV features generated from BEVFormer can simultaneously support multiple 3D perception tasks, such as 3D object detection and map segmentation, which is of great value to autonomous driving systems.

2.2.3. The Disadvantages of Vision-based Solutions

Although the pure vision solution can reduce hardware costs, its "total implementation cost" is not necessarily lower than that of the multi-sensor fusion solution, because the amount of visual 3D image data is large, it occupies a lot of memory and requires stronger computing power. This scheme is limited by the computing power and bandwidth of AI, and has high requirements for the performance of the computing platform and algorithm. Otherwise, it cannot meet the real-time response speed and affect the safety of autonomous driving. With the improvement of hardware computing power and the optimization of software algorithms, this problem is expected to be solved. The pure visual route has high requirements for software, requiring a large amount of data accumulation and scene iteration algorithms. Therefore, it will take a relatively long time for this technology to be widely applied on a large scale.

The pure visual solution relies on a large amount of data to train neural networks, so the software cost is also relatively high. Take Tesla as an example. Since it began to develop autonomous driving technology in 2015, it has deployed over one million autonomous vehicles worldwide, constantly collecting data and accumulating experience, thus forming a huge database. However, the processes of data collection, preprocessing and annotation require a significant amount of time and human resources. In addition, Tesla has equipped its data centers with powerful computing resources for training autonomous driving systems, and this part of the cost cannot be ignored. As Musk said, "To imitate Tesla's self-driving technology, billions of dollars need to be invested in computing training." This figure is equivalent to several hundred billion RMB, highlighting the high cost of pure vision solutions in terms of data support and computing resources.

Although visual sensor technology is now very advanced, there is still a problem of decreased accuracy in bad weather. The Yang206 experiment revealed that the detection error in rainy and snowy weather reached 6% to 17%, which poses a hidden danger for traffic accidents [13].

3. Comparison between Computer Vision and LiDAR

At present, the field of autonomous driving environmental perception is mainly characterized by two different technical paradigms: camera-based systems, which rely heavily on the progress of computer vision; And a system based on liDAR, which uses light detection and ranging for precise spatial mapping. Each method has its unique advantages and limitations, thus forming a continuous debate in the research and industry about the best way forward.

3.1 Accuracy

LiDAR is a sensor technology that obtains scene information through laser scanning and ranging. It can obtain three-dimensional point cloud data of the surrounding environment with high precision, providing important perception capabilities for autonomous driving systems. With the development of LiDAR technology, Frequency Modulated Continuous Wave (FMCW) LiDAR has become an emerging development direction of LiDAR. Compared with traditional pulse LiDAR, FMCW LiDAR continuously emits frequency-modulated laser waves and obtains the distance and speed information of the target object by measuring the frequency difference. The advantage of FMCW LiDAR is that it can measure the speed and distance of multiple objects at the same time, with higher resolution and anti-interference ability. This technology is particularly effective in the detection of high-speed moving objects, and is particularly suitable for applications in highways and complex urban traffic environments.

The target recognition technology based on deep learning is the core of the pure vision solution. By training a large amount of labeled data, the neural network can automatically extract feature information from the image and realize the recognition of multiple targets in complex scenes. For example, real-time target detection algorithms such as YOLO (You Only Look Once) and SSD (Single Shot MultiBox Detector) can quickly identify obstacles ahead in high-speed driving scenes and provide corresponding obstacle avoidance path planning. In addition, the pure vision solution can also be combined with optical flow technology to calculate the movement vector of pixels in continuous frame images, thereby inferring the speed and direction of objects. This is crucial for predicting dynamic scenes such as pedestrians crossing the road and vehicles ahead slowing down or changing lanes.

3.2 Environmental Adaptability

In traditional manually driven cars, the environmental adaptability of sensors mainly depends on high temperature, low temperature, humidity, wind and sand, and sunlight exposure. In addition to these environmental factors, self-driving cars also need to consider environmental adaptability for machine vision, perception and other functions.

The environmental adaptability of LiDAR is one of the core challenges in its practical application. As a high-precision sensor that relies on light pulses to detect targets, its performance is easily disturbed by complex environments. For example, in rain, snow, haze or dusty weather, suspended particles in the air can cause the laser beam to scatter or attenuate, greatly reducing the detection distance and point cloud data quality; strong direct sunlight may introduce background noise and cause misjudgment. In addition, extreme temperature, humidity or vibration may also affect the stability of mechanical radar.

To meet these challenges, technical optimization focuses on two aspects: hardware innovation and algorithm upgrade. In terms of hardware, the use of 1550 nanometer long-wavelength laser (stronger penetration), solid-state scanning solutions (no mechanical parts, vibration resistance) and adaptive optical systems (dynamic compensation of optical path distortion) can improve reliability in harsh environments; at the software level, through multi-echo recognition, deep learning denoising, multi-sensor data fusion (such as combining millimeter-wave radar and camera) and other technologies, interference signals can be effectively filtered and real environmental information can be reconstructed.

The environmental adaptability of pure vision solutions is highly dependent on the coordinated optimization of cameras and algorithms. Under conditions of sufficient lighting and structured scenes (such as urban roads), efficient perception can be achieved through high-precision image recognition, especially in the prediction of dynamic target intentions. However, its performance is easily constrained by extreme lighting (strong light/backlight, low light), bad weather (rain, snow, fog, haze, lens damage) and algorithm long-tail problems (misjudgment of rare scenes). To this end, the technical level improves robustness through multi-spectral perception, adaptive HDR, time series fusion (BEV bird's-eye view) and synthetic data enhancement. At present, this solution is widely used in the fields of passenger car autonomous driving and low-speed robots due to its low-cost characteristics, but redundant sensors are still needed to ensure safety in extreme environments.

3.3 Cost-effectiveness

The economic benefits of lidar in autonomous driving have always been a focus of controversy over technical routes.In the short term, hardware cost and system integration complexity remain the main challenges: early mechanical radars cost more than $10,000 per unit.However, the long-term value of LiDAR is gradually becoming apparent - through high-precision 3D environmental modeling, it can significantly reduce the risk of accidents (according to research, it can reduce the accident rate by 30%-50%), reduce compensation for car companies and brand losses.

Although pure vision solutions have low hardware costs, they rely on massive data training and supercomputing investment (for example, Tesla Dojo costs more than $1 billion a year), and its performance is limited in extreme weather conditions. The fusion solution of lidar, camera, and millimeter-wave radar increases the cost of a single vehicle by about $1,000, but can shorten the R&D cycle of high-end functions and enhance product premium capabilities.

4. Conclusion

The comparative analysis of LiDAR and vision-based solutions in autonomous driving reveals distinct advantages and challenges inherent to each approach. LiDAR excels in delivering high-precision 3D environmental mapping, particularly in low-light conditions or scenarios requiring robust obstacle detection (e.g., occluded pedestrians), while its reliability diminishes in adverse weather (e.g., rain, fog) due to laser scattering. Conversely, vision-based systems leverage cost-effective cameras and advanced deep learning for semantic-rich perception, enabling superior dynamic object intent prediction, yet struggle with extreme lighting (e.g., glare, low-light) and weather-related degradation (e.g., lens contamination).

Economically, LiDAR’s high initial hardware cost (historically over $10,000, now reduced to $500–$2,000 via solid-state innovations) and integration complexity contrast with vision systems’ lower upfront expenses. However, vision solutions incur hidden costs in computational infrastructure (e.g., Tesla’s $1B+ annual Dojo investment) and data annotation, whereas LiDAR’s accident reduction potential (30–50% lower risk) offers long-term safety dividends.

Environmental adaptability further underscores the need for sensor fusion: LiDAR benefits from hardware advancements (e.g., 1550 nm lasers, adaptive optics), while vision systems adopt multi-spectral sensing and BEV (bird's-eye-view) fusion to mitigate limitations. Regulatory trends (e.g., UN-R157 safety mandates) and scalability prospects (LiDAR costs projected below $200 by 2030) reinforce fusion as the optimal path.

Ultimately, neither LiDAR nor vision alone suffices for L4+ autonomy. Hybrid architectures—combining LiDAR's spatial precision, cameras' semantic context, and radar's weather resilience—balance performance and cost, accelerating commercialization in Robotaxi, logistics, and consumer vehicles. Future breakthroughs in chip integration, adaptive algorithms, and synthetic data training will further close gaps, positioning multi-sensor fusion as the cornerstone of scalable, safe autonomous driving ecosystems.

References

[1] Wikipedia. (2025). Self-driving car. Retrieved from https://en.wikipedia.org/wiki/Self-driving_car

[2] Yurtsever, E., Lambert, J., Carballo, A., & Takeda, K. (2020). A survey of autonomous driving: Common practices and emerging technologies. IEEE Access, 8, 58443–58469.

[3] Li, Y., & 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), 50–61.

[4] MRC. (2025). LiDAR in cars: How LiDAR technology is making self-driving cars a reality. Retrieved from https://www.mrlcg.com/resources/blog/lidar-in-cars-how-lidar-technology-is-making-self-driving-cars-a-reality/

[5] Liu, H., Wu, C., & Wang, H. (2023). Real-time object detection using LiDAR and camera fusion for autonomous driving. Scientific Reports, 13(1), 8056.

[6] Zhu, Q., Chen, L., Li, Q., Li, M., Nüchter, A., & Wang, J. (2012). 3D lidar point cloud based intersection recognition for autonomous driving. 2012 IEEE Intelligent Vehicles Symposium, 456–461, IEEE.

[7] Xue, H., Fu, H., Ren, R., Zhang, Y., & Wang, J. (2021). LiDAR-based drivable region detection for autonomous driving. 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 1110–1116.

[8] Aksoy, E. E., Baci, S., & Cavdar, S. (2020). Salsanet: Fast road and vehicle segmentation in lidar point clouds for autonomous driving. 2020 IEEE Intelligent Vehicles Symposium (IV) (pp. 926–932). IEEE.

[9] Heinzler, R., Piewak, F., Schindler, P., & Stork, W. (2020). CNN-based lidar point cloud de-noising in adverse weather. IEEE Robotics and Automation Letters, 5(2), 2514–2521.

[10] Texas Instruments. (2025). An introduction to automotive LIDAR. Retrieved from https://www.ti.com/lit/wp/spry304/spry304.pdf

[11] Muhammad, K., Hussain, T., Ullah, H., Del Ser, J., & de Albuquerque, V. H. C. (2022). Vision-based semantic segmentation in scene understanding for autonomous driving: Recent achievements, challenges, and outlooks. IEEE Transactions on Intelligent Transportation Systems, 23(12), 22694–22715.

[12] Li, Z., Wang, W., Li, H., Xie, E., Sima, C., Lu, T., & Qiao, Y. (2024). BEVFormer: Learning bird's-eye-view representation from lidar-camera via spatiotemporal transformers. IEEE Transactions on Pattern Analysis and Machine Intelligence. Advance online publication.

[13] Liu, F., Lu, Z., & Lin, X. (2025). Vision-based environmental perception for autonomous driving. Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering, 239(1), 39–69.