Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 02 – Issue No: 02 – June 2026

A Comprehensive Survey on Deep Learning-Based Pattern Recognition and Object Detection for Autonomous Vehicles

Ravilla Srinivasulu[1], Theluri Praveen Reddy[2], Are Rishik Reddy[3], Sudipeddolla Hruthik[4], G. Manoj[5] , Mounika Resu[6]

[1,2,3,4,5]Department of CSE, Guru Nanak Institute of Technology, Hyderabad, Telangana, India

[6]Department of Computer science and Engineering, RK College of Engineering (A), Kethanakonda(V), Vijayawada, India

[1]ravillasrinivas4@gmail.com, [2]theluripraveenreddy@gmail.com, [3]rishikreddy0510@gmail.com, [4]hruthikpeddolla@gmail.com, [5]manoj.csegnit@gniindia.org, [6]mounika75.resu@gmail.com

Corresponding author: mounika75.resu@gmail.com

Abstract: The main concept of autonomous cars is founded on well-developed perception systems to read and properly provide feedback to the complex driving conditions. The most important among them, deep learning-based pattern recognition and object detection are now one of the major technologies that allow to reliably identify the road users, traffic signs and obstacles. The paper includes a comprehensive review of the latest advances in the field of deep learning methods of pattern recognition and object detection in autonomous vehicles. The study has thoroughly reviewed the 2D detection methods and 3D, CNN-based model, transformer model, and multi-modal fusion models. Standard, accuracy, real time performance, robustness and efficiency of computation are compared and analyzed. The questionnaire throws some light on the virtues and vices of existing approaches and some of the major issues like negative sensitivity to weather, cost of computation and the poor cross-domain generalization. Finally, possible research directions are discussed in order to lead the next-generation autonomous vehicles in the development of efficient, robust, and real-time perception systems.

Keywords: Autonomous Vehicles; Object Detection; Pattern Recognition; Deep Learning; Multimodal Fusion; 3D Object Detection.

Submitted on: 30 April 2026 Accepted on: 19 June 2026 Published on: 24 June 2026

1. INTRODUCTION

Autonomous vehicles (AVs) heavily use powerful perception systems that are supposed to understand and interact effectively, in a safety-bound manner, with complex driving environments. Among the core elements of autonomous perception, pattern recognition and object detection play a crucial part for detection of road users, traffic signs, lane, obstacles, etc. in real-time. With the fast development of methods in the fields of deep learning, especially convolutional neural networks (CNNs), transformers, and multi-sensor fusion, large progress has been made in the field of visual understanding of self-driving systems. Traditional computer vision approaches that relied on handcrafted features had difficulties in dealing with varying

DoI: doi.org/10.63949/crinfo.v2i2.004

attributes such as illumination, occlusion, scale and weather. Deep learning-based approaches have been able to overcome most of these limitations by automatically learning hierarchical feature representations from large scale datasets. Recent research efforts have focused on both the 2D and 3D object detection, which integrated data from cameras, LiDAR and radar, to increase the perception accuracy and robustness. Despite significant advances, there are still significant challenges in terms of achieving reliable performance in adverse weather, real-time constraints, lack of domain-shift and complex urban scenarios. Therefore, a thorough survey about deep learning-based pattern recognition and object detection techniques for autonomous vehicles is needed for the consolidation of the existing knowledge, analysis of current trends and for the selection of open research problems. This paper overviews the recent architectures, methodologies, applications, and challenges of the autonomous vehicle perception and presents future research directions.

2. REVIEW OF LITERATURE

Aziz et al. (2020) made an extensive review of deep learning architectures for generic object detection, mentioning two-stage and one-stage detectors and focusing on the evolution from traditional machine learning to deep neural networks. Their effort focused on the importance of feature extraction backbones and detection frameworks but more on generic than autonomous driving contexts. In an article published by Boukerche and Ma in (2021), the authors investigated vision-based autonomous vehicle recognition, where they have found deep learning to be one of the key enablers to intelligent transportation systems. The authors commented on some of the challenges including real-time processing and the variability of the environment, and the authors said that robust perception is an open problem for autonomous systems. Pavel et al. (2022) published a systematic review of the literature in the field of vision-based autonomous vehicle systems based on deep learning. Their study classified perception functions into different categories, and also stressed the rising trend of CNN-based models in improving the detection accuracy. Similarly, R. Nidhya et al. (2024) explored facial recognition technologies using deep learning-based pattern recognition techniques, demonstrating the effectiveness of automated visual recognition systems in real-world applications. However, the focus of the review was mainly on camera-based systems, and the review contained little coverage of multimodal sensing. Alaba and Ball (2022) performed a survey on LiDAR-based 3D object detection methods in autonomous driving, which introduced the benefits of point cloud processing to obtain spatial understanding. The authors discussed voxel-based methods, point-based methods, and a combination of the two, but noted issues of computational complexity and the cost of sensors. Ahmed et al. (2021) investigated the deep learning methods of object detection under challenging environments, and showed that environmental conditions like fog, rain and low light are contributing to the poor performance in object detection. Their results highlighted the need for better and weather independent models. Alaba et al. (2022) also referred to multi-sensor fusion technologies to detect 3D objects and give an explanation of how the use of camera and LiDAR data can enhance the reliability and accuracy. Nevertheless, they had issues with synchronization, calibration as well as increased computational load. Pravallika et al., (2024) published their results with their findings in terms of the new tendencies in transformer-based structures and BEV (birds eye view) representations but acknowledged that the problems

DoI: doi.org/10.63949/crinfo.v2i2.004

Frontiers in Engineering and Informatics
ISSN: 3049-3412 (Online)
Volume No: 02 – Issue No: 02 – June 2026

of their implementation in real-time were still present. In Tahir et al., the detection of objects in case of bad weather conditions is promoted by traditional and deep learning methods. The research made a conclusion that in spite of the methods of deep learning having better performances than the classical techniques, the problem of performance deterioration in severe weather is still an important one. Aher et al. (2024) conducted a review of deep learning-based object detection and tracking for sentinel cars and stressed using temporal modeling and associating tracking. They speculated that the future systems have to make latency and scalability issues. Trigka and Dritsas (2025) have given a wide overview of machine learning models for the object detection task, covering both the classical and deep learning paradigms. While exhaustive, their review was not a special designed to discuss autonomous vehicle perception pipelines, though. Li et al., Object Detector for Autonomous Driving in Complex Environment, Hybrid Deep learning framework for object detection and recognition in autonomous driving scenario. Their work showed the improvement of accuracy with the model alliance, but has not fully considered real-time deployment limitations.

3. RESEARCH GAP

The reviewed studies highlighted several research gaps have been identified in deep learning-based pattern recognition and object detection for autonomous vehicles. First, there is a lack of limited unified surveys, as many studies focus either on generic object detection techniques or on specific object detection modalities such as camera-based or LiDAR-based systems, without providing a unified view of pattern recognition and object detection for self-driving vehicles. Furthermore, although many deep learning models have achieved high detection accuracy, few models achieve sufficient real-time focus, as their computational requirements are often high, making them costly and inadequate for real-time embedded functionality in autonomous motor vehicles. While areas of substantial or complete agreement exist among different models regarding the effectiveness of deep learning for perception tasks, interpretations of agreement and disagreement in reported findings still require further investigation. Beyond that, there has been a lack of study on the framework agreement, which is an explicit or implicit agreement about the nature and conditions of risks that should be identified and measured. Another significant challenge is adverse weather robustness, as current detection systems continue to struggle under challenging environmental conditions such as fog, rain, snow, and low-light scenarios, highlighting the need for smarter and more adaptable models capable of maintaining reliable performance across varying weather conditions. Furthermore, there is an open research question regarding domain adaptation and/or generalisation because existing models, when applied to new domain environments that were not present in their training sets, tend to suffer from domain shift, meaning that there is a need for better standardisation and generalisation. While interpretability and safety validation of deep learning-based perception systems are critical for real-world implementation of self - governing systems, there are relatively few publications on this topic. Finally, edge efficiency is a huge consideration, as the architectures are less available and efficient enough to provide high accuracy and reliability of detection and still work on low-power automotive hardware.

4. PROBLEM STATEMENT

DoI: doi.org/10.63949/crinfo.v2i2.004

Autonomous vehicles rely on the correct recognition of patterns and detection of objects in order to see dynamic events in a road in real time. Although deep learning has enhanced the performance of detection, current models still have some problems, including poor performance under adverse weather, occlusion, illumination change, and processing time. Additionally, the use of multimodal sensor fusion brings extra challenges such as high computational cost, synchronization complexity and difficulty of deployment on an embedded automotive hardware. Therefore, a proper review is required to investigate the existing methods based on deep learning, to see those limitations and to point out the research streams to build powerful, efficient and real-time perception systems for autonomous vehicles.

5. OBJECTIVES

- To review existing deep learning–based pattern recognition and object detection techniques for autonomous vehicles.
- To analyze the performance of major architectures such as CNNs, transformers, and multimodal fusion models.
- To compare 2D and 3D object detection approaches used in autonomous driving.
- To identify key challenges including real-time processing, adverse weather robustness, and computational efficiency.
- To highlight current research trends and application areas in autonomous vehicle perception.
- To determine research gaps and propose future works on the creation of powerful and effective detection systems.

6. Methodology

This research is done through systematic review and comparative analysis method to analyze deep learning-based pattern recognition and object detection algorithms in autonomous vehicles. The research is aimed at analysing and comparing the existing methods in place of creating a new model.

6.1 Research Design

As a way of comprehending various deep learning models and their performance, a qualitative and quantitative review approach was adopted. The main objective of the study is to locate trends, strengths, and limitations of the existing solutions in autonomous vehicle perception systems.

6.2 Data Collection

The most popular scholarly resources (IEEE Xplore, SpringerLink, ScienceDirect, and Google Scholar) were used to retrieve relevant research papers.The selected papers that were used in this study were published in 2020-2025 and were based on the deep learning methods applied to object detection and pattern recognition in autonomous vehicles.

DoI: doi.org/10.63949/crinfo.v2i2.004

6.3 Classification of Techniques

The gathered studies were divided into four broader categories to be analyzed more easily:

- 2D object detection models (e.g., CNN-based)
- 3D object detection models (mainly LiDAR-based methods)
- Transformer-based models
- Multimodal fusion models (uniting information of several sensors)

This categorization aided in the realization of the effectiveness of various approaches and their areas of success.

6.4 Evaluation Criteria

The models were compared to each other on the basis of performance measures that are commonly used such as:

- Accuracy (in terms of mAP)
- Speed (in FPS)
- Computational efficiency
- Performance under challenging conditions such as bad weather and occlusion

This was compared with the results in standard benchmark datasets such as KITTI and nuScenes.

6.5 Comparative Analysis

To learn the trade-off between accuracy and real-time performance, a thorough comparison of various models was done. Tables and graphs were implemented to clearly demonstrate differences in performance of models. The behavior of models in the actual driving situation was given special attention.

6.6 Identification of Research Gaps

Upon reviewing the current approaches, some gaps in research were found. These are high computation cost, reduced performance in adverse weather conditions, low capability to generalize to new environments and absence of lightweight models to be used in real-time.

7. EXPERIMENTAL RESULTS

To demonstrate the practicability of the deep-learning powered pattern recognition and object detection systems to manage autonomous cars, the comparative analysis of the exemplary models was performed on the basis of the literature. The standard evaluation measures were used to reduce the analysis to the generally accepted benchmarks such as KITTI and nuScenes .

DoI: doi.org/10.63949/crinfo.v2i2.004

7.1 Performance Summary

The performance summary indicates that a camera-based perception system using 2D CNN-based detectors is able to deliver high detection rates with an average mean Average Precision (mAP) in the range of 75% to 90% for detectors such as Faster R-CNN, SSD and YOLO. One of these is a one-stage detector that can make inferences fast, hence ideal for real-time autonomous driving applications. Although 3D object detection using LiDARs is in general more costly in terms of computation, it has better spatial localisation and depth estimation performance in more complex traffic scenarios. Transformer-based models have shown to improve the capability to model the circular context and achieve better detection accuracy in complex scenes, but they utilize substantial amounts of trainable data and computing power. In addition, multisensor fusion models always outperform single-modality systems, as they fuse information from different sensors, making the system more robust against occlusions and different illumination.

7.2 Robustness Analysis

The findings from studies on object detection systems that were reviewed show that fog, rain and snow have a negative impact on the efficiency of the object detection systems. Moreover, the occlusion and small object detection are still complex tasks for existing models. Moreover, when deploying in real-time, large model sizes and high latency are common problems that add to the computational needs and decrease the operational efficiency.

7.3 Comparative Insight

Generally, the method of deep-learning is exquisitely superior to the conventional techniques of vision systems in the application of autonomy driving. Nevertheless, there is a deadly trade-off between accuracy and the performance of the model run calculations - and there are even models which perform well that can be anything but optimized to run on low-power automotive hardware.

7.4 Comparison of Deep Learning–Based Object Detection Methods for Autonomous Vehicles

Table 1. Performance comparison of representative models reported in the literature.

Method Type	Representative Model	Modality	Strengths	Limitations	Typical mAP (%)	Real-Time (FPS)
2D Detection (Two-Stage)	Faster R-CNN	Camera	High accuracy, strong localization	Slow inference	80–90	5–10
2D Detection (One-Stage)	YOLOv5	Camera	Very fast, suitable for real time	Slightly lower accuracy	75–88	30–60
2D Detection (One-Stage)	SSD	Camera	Good speed–accuracy balance	Weak for small objects	70–85	20–40

DoI: doi.org/10.63949/crinfo.v2i2.004

3D Detection (Voxel-Based)	VoxelNet	LiDAR	Accurate 3D localization	High computation	65–78	5–15
3D Detection (Point-Based)	PointRCNN	LiDAR	Precise object boundaries	Complex pipeline	68–80	8–12
Transformer-Based	DETR	Camera	Global context modeling	Slow convergence	70–85	10–20
Multimodal Fusion	PointPainting	Camera + LiDAR	High robustness, better accuracy	Fusion complexity	75–88	10–18
BEV-Based Fusion	BEVFormer	Multi-sensor	Strong spatial understanding	Heavy computation	78–90	8–15

7.5 Comparison of Object Detection Models: mAP VS FPS

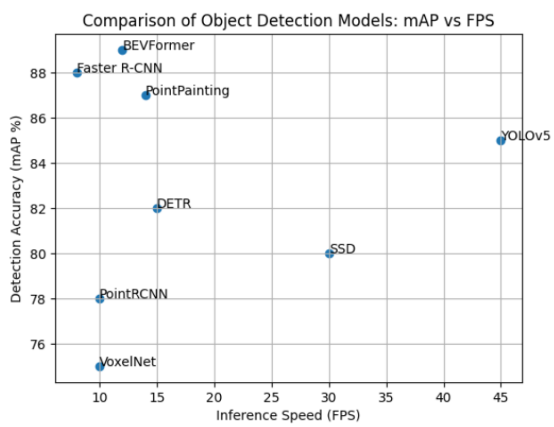

Figure 1. Comparison of Object Detection Models: mAP vs FPS

Figure 1 presents the mAP and FPS comparison of object detection models. YOLOv5 had the fastest inference speed of about 45 FPS among the considered models and had a relatively balanced detection accuracy of around 85% mAP, which is ideal for autonomous driving applications in real time. Faster R-CNN and BEVFormer obtained the highest detection performance in terms of mAP scores (88–89%), but their FPS figures suggest they require more complex computing and are less suitable for real-time systems. SSD achieves a good compromise between processing speed and detection accuracy, making it well- suited for resource-constrained scenarios. VoxelNet and PointRCNN, which are based on 3D LiDAR, on the other hand, had significantly lower frames per second (FPS) because of the heavy computational requirements involved in working with point clouds. Multimodal fusion techniques (e.g., PointPainting) were used to improve accuracy and robustness of the detection, but their use sacrificed real-time performance. Likewise, the transformer-based DETR model demonstrated competitive detection accuracy by improving global

DoI: doi.org/10.63949/crinfo.v2i2.004

context modelling, but it still encountered the problem of real-time processing for autonomous vehicle system applications.

8. CONCLUSION

This survey was targeted to give a detailed overview of deep-learning-based methods of pattern recognition and object detection to autonomous vehicles. In the research paper, solutions 2D CNN based detectors, 3D lidar based solutions, transformer solutions and multimodal fusion frameworks have been discussed as key solutions. The comparative analysis implied that the deep learning models are much superior in array of detection accuracy and strength compared to the traditional vision methodologies. Nevertheless, there is the compromise of accuracy and timely range. Real-time deployment can be performed using light-weight models like the YOLO variants, transformer and fusion-based models can be more accurate but come at the cost of greater computational complexity. However, despite the aforementioned drawbacks like sensitivity to adverse weather, dealing with occlusions, high computational requirements and constrained cross-domain generalization, complete trustworthiness of autonomous perception has not yet been allowed. Overall, the results show that although significant progress has been made, there is still a need for further research to build solid, efficient and scalable perception systems for the future deployment of autonomous vehicles in the real world.

9. FUTURE WORK

The future research in this regard should lean towards building lightweight yet high-accuracy deep learning models that can run on the edge automotive hardware in real-time mode. Improving the robustness of perception in adverse weather conditions, under occlusions and in low light conditions is still a key to reliable perception. In addition, multimodal fusion strategies with higher efficiency and optimized transformer or hybrid architectures are required to reduce the accuracy/people cost. Moreover, future work should also focus on domain adaptation for improved generalization across diverse environments as well as explainable AI techniques to make the autonomous vehicle systems safe and trustworthy.

REFERENCES

[1] Dang, N. C., Moreno-García, M. N., & De la Prieta, F. (2020). Sentiment Analysis Based on Deep Learning: A Comparative Study. Electronics, 9(3), 483. https://doi.org/10.3390/electronics9030483

[2] K. S. Yogi, V. Dankan Gowda, D. Sindhu, H. Soni, S. Mukherjee and G. C. Madhu, "Enhancing Accuracy in Social Media Sentiment Analysis through Comparative Studies using Machine Learning Techniques," 2024 International Conference on Knowledge Engineering and Communication Systems (ICKECS), Chikkaballapur, India, 2024, pp. 1-6, doi: 10.1109/ICKECS61492.2024.10616441

DoI: doi.org/10.63949/crinfo.v2i2.004

[3] Gandhar, A., Gandhar, S., Rehalia, A., Priyadarshi, P., & Tiwari, M. (2024). A Comparative Analysis of Machine Learning Algorithms for Sentiment Analysis in Indian Social Media. Library of Progress-Library Science, Information Technology & Computer, 44(3).

[4] Qaiser, S., Yusoff, N., Ali, R., Remli, M. A., &Adli, H. K. (2021). A comparison of machine learning techniques for sentiment analysis. Turkish Journal of Computer and Mathematics Education, 12(3), 1738-1744.

[5] Singh, N.K., Tomar, D.S. &Sangaiah, A.K. Sentiment analysis: a review and comparative analysis over social media. J Ambient Intell Human Comput 11, 97–117 (2020). https://doi.org/10.1007/s12652-018-0862-8

[6] Sudhir, P., & Suresh, V. D. (2021). Comparative study of various approaches, applications and classifiers for sentiment analysis. Global Transitions Proceedings, 2(2), 205-211. https://doi.org/10.1016/j.gltp.2021.08.004

[7] R. Nidhya, D. J. A. Pabi, U. DivyaSree, B. Abhishek, A. Harika and S. A. Alim, "Exploring Facial Recognition Technologies for Classroom Management," 2024 10th International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 2024, pp. 1702-1707, doi: 10.1109/ICCSP60870.2024.10544143.

[8] Kansal, R., &Diwaker, C. (2025). Efficiency Determination of Various Machine Learning Techniques for Sentiment Analysis on Social Media Platforms. Engineering, Technology & Applied Science Research, 15(4), 25584-25589. https://doi.org/10.48084/etasr.11158

[9] Enduri, M. K., Sangi, A. R., Anamalamudi, S., Manikanta, R. C. B., Reddy, K. Y., Yeswanth, P. L., ... & Karthikeya, A. (2023). Comparative study on sentimental analysis using machine learning techniques. Mehran Univ. Res. J. Eng. Technol, 42(1), 207. https://doi.org/10.22581/muet1982.2301.19

[10] ShiXiao, X., Alobaedy, M. M., Goyal, S. B., Singla, S., Kang, S., & Chadha, R. (2023). Real-time sentiment analysis on social networks using meta-model and machine learning techniques. Scalable Computing: Practice and Experience, 24(4), 819-834. https://doi.org/10.12694/scpe.v24i4.2247.

[11] Rodríguez-Ibánez, M., Casánez-Ventura, A., Castejón-Mateos, F., & Cuenca-Jiménez, P. M. (2023). A review on sentiment analysis from social media platforms. Expert Systems with Applications, 223, 119862. https://doi.org/10.1016/j.eswa.2023.119862

[12] Reddy, M. A. S., & Prakash, C. (2024). Sentiment Analysis of Social Media Networking Sites: A Comparative Study on User Engagement and Public Opinion. 10.15680/IJMRSET.2024.0709012

DoI: doi.org/10.63949/crinfo.v2i2.004