IJIRIS: International Journal of Innovative Research in Information Security E-ISSN: 2349-7017
Volume 11, Issue 12, December 2025 P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1112.03

Computer Vision for Autonomous Driving

Prof.Suguna A
Computer Science and Engineering
Sri Sairam College of Engineering, Bengaluru, India
Sugunaa.cse@sairamce.edu.in
https://orcid.org/0009-0003-5310-3492
Thamrin.A, Rakshitha.K Srushti B, Shraddha S L, Rukmini.V. M
Computer Science and Engineering
Sri Sairam College of Engineering, Bengaluru, India
sce24cs034@sairamtap.edu.in, sce24cs026@sairamtap.edu.in
sce24cs065@sairamtap.edu.in, sce24cs054@sairamtap.edu.in
sce24cs087@sairamtap.edu.in

Publication History
Manuscript Reference No: IJIRIS/RS/Vol.11/Issue12/DCISI0082
Research Article Open Access| Double-Blind Peer-Reviewed| Article ID: IJIRIS/RS/Vol.11/Issue12/DCIS10082 Received: 28, October 2025, Revised: 05, November 2025, Accepted: 12, November 2025, Published Online: 21, November 2025.
https://www.ijiris.com/volumes/Vol11/iss-12/03.DCIS10082.pdf
Citation:Prof.Suguna,Thamrin,Rakshitha,Shraddha,Rukmini(2025),Computer Vision for Autonomous Driving, IJIRIS: International Journal of Innovative Research in Information Security,Volume 11, Issue 11 of 2025 pages 846-850
Doi:> https://doi.org/10.26562/ijiris.2025.v1112.03
BibTeX Key: Prof.Suguna@Computer
IJIRIS papers should be cited as IJIRIS (International Journal of Innovative Research in Information Security, AM Publications, India 2025, ISSN 2349-7017, https://doi.org/10.26562/ijiris.2025.v1112.03 The journal's official abbreviation is IJIRIS. Orcid: https://orcid.org/0009-0004-9398-7488

Abstract: Autonomous driving has emerged as one of the most transformative innovations in intelligent transportation systems. Modern self-driving vehicles rely heavily on computer vision techniques to understand, interpret, and react to their surroundings in real time. The complexity of the driving environment characterized by dynamic obstacles, traffic signs, lane markings, pedestrians, and varying lighting conditions demands robust perception systems. Computer vision enables autonomous vehicles to detect objects, segment scenes, estimate depth, track motion, and make high-accuracy decisions essential for safe navigation. This paper presents a comprehensive overview of the essential computer vision components used in autonomous driving, including camera systems, sensor fusion, object detection, lane detection, traffic sign recognition, and depth estimation. It also discusses commonly used datasets, algorithms, and deep learning models such as CNNs, YOLO, SSD, HOG, SVM, RANSAC, Canny, and semantic segmentation networks. The work synthesizes how these modules collectively support perception, prediction, and planning stages in autonomous vehicles. Expanded sections detail the design, workflow, implementation considerations, and practical challenges encountered in real-world deployments. The discussion illustrates the critical role of computer vision in enabling vehicles to operate safely and efficiently without human intervention.
Keywords: Autonomous Driving, Computer Vision, Lane Detection, Object Detection, Semantic Segmentation, Traffic Sign Recognition, Depth Estimation, LiDAR, Camera Sensors, YOLO, CNN, ADAS.

I. INTRODUCTION

Autonomous driving technology aims to reduce human driving errors, increase road safety, optimize fuel consumption, and improve traffic efficiency using AI-driven algorithms. Vision-based systems form the backbone of intelligent vehicles because cameras offer rich spatial and semantic information similar to the human visual system. Self-driving vehicles must perform a sequence of tasks: perceive the environment, understand road structures, detect obstacles, predict motion, and execute safe navigation actions. These tasks require real-time processing of data from multiple sensors, including cameras, LiDAR, radar, and ultrasonic devices. Computer vision provides the foundation for object detection, tracking, lane marking interpretation, signage understanding, and depth estimation. While autonomous driving has advanced significantly through deep learning and sensor fusion, perception remains one of the most challenging components. Variations in weather, illumination, occlusion, shadows, and road deterioration introduce complexities. Therefore, robust and scalable computer vision models are essential for reliable autonomous operation. Autonomous vehicles represent a major milestone in modern artificial intelligence, combining advanced perception systems, decision-making algorithms, and real-time control mechanisms. The introduction of high-performance computing hardware, large annotated datasets, and deep learning has significantly accelerated progress in the field. Today, autonomous vehicles are no longer limited to research labs they are being tested on real roads and integrated into smart transportation ecosystems. A key pillar of autonomous driving is computer vision, which enables machines to replicate human-like visual understanding. Unlike humans who rely solely on biological vision, autonomous vehicles must interpret visual information through complex algorithms that process pixel-level data at high speed.

This includes detecting vehicles, pedestrians, traffic lights, lane boundaries, and dynamic obstacles under varying lighting and environmental conditions. Furthermore, autonomous driving systems must operate in complex, unpredictable environments, such as congested city streets, highways, rural roads, and adverse weather. This requires robust algorithms that can generalize well beyond training data. The integration of multiple sensors such as cameras, LiDAR, radar, GPS, and inertial measurement units enhances reliability through redundancy and complementarity. Autonomous vehicles have the potential to transform transportation by reducing accidents caused by human error, improving mobility for elderly or disabled individuals, and optimizing fuel usage and traffic flow. However, achieving full autonomy involves addressing challenges related to perception accuracy, extreme edge cases, real-time processing, and system safety. This motivates research into better computer vision architectures, improved sensor fusion techniques, and more efficient deep learning models.

II. LITERATURE SURVEY

Recent advancements in autonomous driving research emphasize deep learning and robust feature extraction. Convolutional Neural Networks (CNNs) have replaced traditional handcrafted feature approaches, enabling end-to-end scene understanding. Popular models such as YOLO, SSD, and Faster R-CNN deliver real-time object detection suitable for embedded automotive systems. Studies show that computer vision combined with LiDAR improves distance estimation, obstacle detection, and scene reconstruction accuracy. Research on road lane detection highlights the effectiveness of Hough Transform, RANSAC, and deep learning segmentation models in handling lane curvature, shadows, and worn markings. Traffic sign recognition literature demonstrates the use of hybrid CNN-SVM models for high-accuracy classification, even under motion blur. Depth estimation research explores monocular depth networks and stereo matching techniques for 3D perception. The literature indicates that integrating multiple computer vision pipelines increases reliability, making perception more adaptable to real-world conditions. Extensive research has been conducted over the past decade to improve perception systems for autonomous vehicles. Early literature focused on traditional image-processing techniques such as Sobel filters, Canny edge detection, Hough Transform, and template matching. These were effective for basic lane detection and feature extraction but struggled with real-world variability. With the advent of deep learning, especially Convolutional Neural Networks (CNNs), perception capabilities improved dramatically. Several landmark studies introduced real-time object detection models:

- YOLO (You Only Look once) demonstrated end-to-end object detection at high speed, suitable for autonomous platforms.
- SSD (Single Shot Detector) improved multi-scale detection, enabling robust recognition of objects at different distances.
- Faster R-CNN introduced Region Proposal Networks (RPN), significantly increasing accuracy for small or partially occluded objects.

Studies have also explored semantic segmentation models such as UNet, SegNet, and DeepLab, which enable pixel level understanding crucial for lane marking identification, drivable area segmentation, and pedestrian boundary detection. Research on LaneNet, SCNN (Spatial CNN), and RESA has further improved lane detection stability under shadows, rain, and night-time conditions. Additionally, literature highlights the importance of sensor fusion, combining LiDAR and camera features to overcome limitations of single-sensor systems. Works such as KITTI Vision Benchmark and Cityscapes dataset have provided high quality annotated data for evaluating perception models in diverse conditions. Traffic sign recognition research has evolved from classical feature extraction (HOG, SVM) to deep neural networks capable of classifying thousands of sign categories with high accuracy. Recent studies also explore monocular depth estimation using deep networks, reducing the dependency on costly stereo or LiDAR setups. Overall, the literature indicates a clear shift from handcrafted feature methods to deep learning–based end-to-end models, driven by the need for high accuracy, robustness, and generalization in autonomous driving environments.

III. PROPOSED WORK

The proposed work focuses on developing a multi-module autonomous driving perception pipeline using computer vision techniques. The framework includesthe Camera Calibration &Pre-processingwith Correction of radial distortion, noise reduction, and color normalization. The Lane Detection Module Which uses Canny edge detection, Gaussian filtering, region-of-interest masking, and Hough Transform or polynomial fitting for curved lanes. The Object Detection Module which employs deep learning-based detectors (YOLO, SSD, Faster R-CNN) for detecting vehicles, pedestrians, cyclists, and road obstacles. The Traffic Sign Recognition which involves image segmentation, bounding box extraction, feature computation, and classification using CNN models. The Depth Estimation that Uses stereo vision or monocular depth networks to estimate distances for collision avoidance. Finally, the Environment Perception Layer that Integrates lane, object, and traffic sign data for a unified understanding of the scene.This end-to-end system forms the core perception pipeline for a self-driving vehicle.

IV. EXISTING SYSTEM

Existing autonomous vehicle systems primarily rely on a combination of sensors and rule-based algorithms to support driver assistance functions such as adaptive cruise control, lane keeping, collision avoidance, and automatic braking. These systems, commonly referred to as Advanced Driver Assistance Systems (ADAS), provide partial autonomy but still depend heavily on human supervision. Lane detection systems in today's vehicles use simple camera-based edge detection, which struggles in low visibility or worn-out lane markings.

Object detection modules rely on radar for distance measurement, but radar alone lacks semantic understanding of objects. Traffic sign detection is often limited to speed signs and fails under occlusion or high-speed motion blur. Most existing systems lack deep semantic understanding of the environment and cannot predict the future movement of surrounding vehicles or pedestrians. Additionally, current systems rarely integrate advanced deep learning models or robust multi-sensor fusion pipelines. These limitations highlight the need for more advanced perception models and fully automated driving architectures.

V. PROPOSED SYSTEM

The proposed autonomous driving system architecture consists of multiple interconnected layers designed to handle sensing, perception, prediction, and decision-making. The Data Acquisition Layer captures raw input from cameras, LiDAR, radar, and GPS sensors. The Processing and Analytics Layer handles sensor fusion, image preprocessing, feature extraction, and real-time analysis using deep learning models. The Application Layer consists of perception modules responsible for lane detection, object detection, traffic sign recognition, and semantic segmentation. Additionally, motion prediction algorithms analyze the trajectories of surrounding objects. The Decision and Planning Layer converts perception data into actionable driving commands by generating safe paths, controlling speed, and avoiding obstacles. Finally, the Actuation Layer sends commands to the vehicle's steering, acceleration, and braking systems. This modular architecture ensures scalability, fault tolerance, and the ability to incorporate more advanced perception models as technology evolves.

Fig 1. Proposed System

VI. DATA MODEL AND INTERFACES

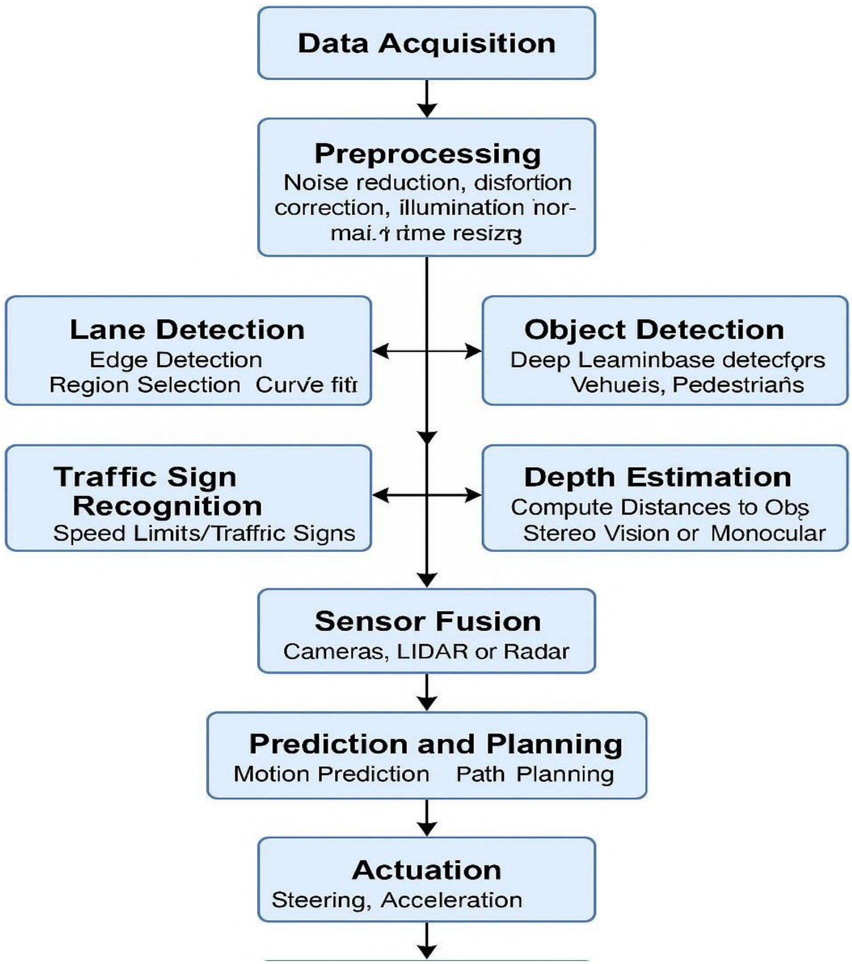
The data model for an autonomous vehicle focuses on structured and time-synchronized sensory information. Visual frames from cameras are stored along with sensor calibration data, timestamps, and positional metadata. LiDAR point clouds are represented as 3D spatial coordinates with reflectance values that help in identifying nearby objects and road structures. Radar data contains range, angle, and velocity information useful for estimating motion. The system's perception dataset integrates all sensory inputs through a sensor fusion schema that aligns information spatially and temporally. Interfaces are developed as modular APIs allowing communication between the perception, prediction, and planning components. These interfaces ensure smooth data flow between detection modules, semantic segmentation outputs, traffic sign classification layers, and the central decision-making system. Such a structured data model enables the vehicle to maintain an accurate and continuously updated understanding of its surroundings.

VII. METHODOLOGY

The methodology of the autonomous driving system is based on a multi-step perception pipeline. First, sensory inputs undergo preprocessing to remove noise, correct distortions, and normalize illumination. Camera frames are fed to deep learning-based object detection networks such as YOLO or Faster RCNN, which identify vehicles, pedestrians, cyclists, and road obstacles. For lane detection, the system applies a combination of edge extraction, region masking, and polynomial fitting or deep segmentation networks to accurately track lane boundaries even in curved or shadowed roads. Traffic sign recognition is performed through CNN-based classifiers trained on large sign datasets. Depth estimation is achieved either through stereo matching or monocular depth networks, providing essential 3D information for collision avoidance. All perception outputs are combined through sensor fusion algorithms that integrate visual, LiDAR, and radar data for improved reliability. The fused data is then passed to prediction and planning algorithms, completing the full perception-to-action pipeline. The methodology for the autonomous driving perception system follows a structured and sequential process that enables the vehicle to understand its surroundings and make safe navigation decisions. The system begins with data acquisition, where visual information from cameras and additional sensing inputs from LiDAR and radar are captured continuously. This raw data undergoes preprocessing steps such as noise reduction, distortion correction, illumination normalization, and frame resizing to ensure that the subsequent perception modules receive clean and consistent input.

IJIRIS: International Journal of Innovative Research in Information Security E-ISSN: 2349-7017
Volume 11, Issue 12, December 2025 P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1112.03

Once preprocessing is completed, the lane detection component processes the visual frames to identify road boundaries. This is achieved through edge extraction techniques, region-of-interest selection, and polynomial curve fitting, or in some cases deep-learning-based segmentation models that provide more reliable detection under challenging environmental conditions such as shadows, faded markings, or night-time driving. Simultaneously, the object detection module analyzes each frame using deep learning algorithms like YOLO or Faster R-CNN to detect vehicles, pedestrians, cyclists, and other road obstacles. Detected objects are localized using bounding boxes and classified according to their types, helping the system understand potential hazards around the vehicle. In parallel, a traffic sign recognition module leverages color information, region extraction, and convolutional neural networks to accurately identify speed limits, warning signs, and mandatory traffic instructions. Depth estimation is another critical component of the methodology and is performed either through stereo vision, which calculates depth from disparities between left and right camera images, or through monocular depth prediction networks that infer 3D distance information from a single image. This depth data helps the vehicle assess how far various objects are located in real-world space. To ensure higher reliability and robustness, the system incorporates sensor fusion, where data from cameras, LiDAR, and radar are combined using filtering and alignment algorithms. Sensor fusion mitigates individual sensor weaknesses and enables more consistent perception, especially in environments with poor visibility or high traffic density. After perception data is fully processed, it is passed to the prediction and planning layer, which analyzes the motion patterns of other road users and forecasts their future positions.
Based on this understanding, the planning module formulates a safe driving path by evaluating multiple factors such as lane structure, object distance, collision risk, and traffic rules. Finally, the actuation module converts these planned decisions into control commands that regulate the vehicle's steering, acceleration, and braking, ensuring smooth and safe autonomous movement. This complete methodology forms an end-to-end pipeline that transforms raw sensor input into intelligent driving actions.



VIII. CONCLUSION

Computer vision is a fundamental component enabling autonomous vehicles to interpret their environment, detect obstacles, and make intelligent driving decisions. The combination of lane detection, object recognition, traffic sign identification, and depth estimation forms a complete perception stack crucial for safe navigation. While challenges such as poor lighting, occlusion, adverse weather, and real-time processing constraints remain, advancements in deep learning, sensor fusion, and edge processing continue to improve performance.

E-ISSN: 2349-7017
P-ISSN: 2349-7009
https://www.ijiris.com/archives
https://doi.org/10.26562/ijiris.2025.v1112.03

Future work may explore transformer-based models, multimodal perception, and robust domain adaptation techniques to further enhance autonomous vehicle perception capabilities.

REFERENCES

1. Krizhevsky, I.Sutskever, and G.E.Hinton, "ImageNet classification with deep convolutional neural networks," Advances in Neural Information Processing Systems (NeurIPS), 2012.
2. J.Redmon, S.Divvala, R.Girshick, and A. Farhadi, "You Only Look Once: Unified, real-time object detection," Proc. IEEE Conf. Comput.Vis. Pattern Recognit. (CVPR), 2016–2020.
3. S.Ren, K.He, R.Girshick, and J.Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," IEEE Trans. Pattern Anal. Mach. Intell. (TPAMI), vol. 39, no. 6, pp. 1137–1149, 2015.
4. R.Girshick, "Fast R-CNN," Proc. IEEE Int. Conf. Comput. Vis. (ICCV), 2015, pp. 1440–1448.C., V., & S., J. J. T. (2024). Enhanced Breast Cancer Detection in Ultrasound Imaging Using an Attention Based U-Net++ Architecture for Improved Tumor Segmentation. International Journal of Electronics and Communication Engineering, 11(8), 82–89. https://doi.org/10.14445/23488549/ijece-v11i8p109
5. Priya, A.T.R.K.,Rethinakumari, M.,Valarmathi, C., Anuja, R., Devi, R. S., & Kavyalakshmi, A. (2024). Intelligent Routing Protocols for Secure and Efficient Transmission of IoT Medical Data. In 2024 International Conference on Advancement in Renewable Energy and Intelligent Systems (AREIS) (pp. 1–6). 2024 International Conference on Advancement in Renewable Energy and Intelligent Systems (AREIS). IEEE. https://doi.org/10.1109/areis62559.2024.10893613.
6. F. Fleischer, "GTSRB Traffic Sign Dataset," *IEEE ITS*, 2013.
7. KITTI Vision Benchmark Suite, 2012.
8. BDD100K: Berkeley Deep Drive Dataset, 2018.
9. C,V.,&A,S.(2022). UTILIZATION OF UGV IN MINING SECTORS. International Research Journal of Computer Science, 09(07), 160–165. https://doi.org/10.26562/irjcs.2022.v0907.005.
10. C,V.,C,A. R,B.,Gowda, K.B.,& G.R,K.(2025).Sleep Stage Classification using SimCLR Framework. In National level Technical Symposium (Advaya 2k25). National level Technical Symposium (Advaya 2k25). International Journal of Advanced Trends in Engineering and Management. https://doi.org/10.59544/cpta3980/advaya2k25p3.
11. C.Valarmathi. (2024). The Combination of Feature Extraction and Classification by Bag of Visual Words to Detect Breast Cancer for Improved Accuracy. Journal of Electrical Systems, 20(3), 2949–2955. https://doi.org/10.52783/jes.4639.
12. Ramamoorthy, R., Velu, A., Valarmathi, C., & Ananthi, M. (2025). Evaluating Mobility Models for IH-VANETs: A Simulation-Based Analysis. In 2025 International Conference on Computing and Communication Technologies (ICCCT) (pp. 1–5). 2025 International Conference on Computing and Communication Technologies (ICCCT). IEEE. https://doi.org/10.1109/iccct63501.2025.11020005.
13. C,V.,K,D., Sanofar, F., & Keshava, K. (2023). Smart Home Intelligent System for Elder People Using Raspberry Pi. International Research Journal of Computer Science, 10(05), 204–210. https://doi.org/10.26562/irjcs.2023.v1005.18.
14. Valarmathi,C., & Sharanya, S. (2024). Cyberbullying Detection in Instagram Comments. INTERANTIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING AND MANAGEMENT, 08(008), 1–5. https://doi.org/10.55041/ijsrem37403.