Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Object Detection in Autonomous Driving with Sensor-Based Technology Using YOLOv10

Riya Saini*[1]**, **Kapil Kumar[2]

[1]**Research Scholar, College of Smart Computing,**

COER University, Roorkee.

[2]**Assistant Professor, College of Smart Computing,**

COER University, Roorkee.

***Corresponding Author Email:** riyasaini1720@gmail.com

Abstract

The creation of intelligent transportation systems, such as autonomous driving and traffic monitoring, is dependent on precise vehicle recognition. Autonomous vehicles detect and recognize objects in real-time, such as pedestrians, other vehicles, traffic signs, and obstacles. This paper improves the object detection ability of autonomous vehicles (AVs') by integrating technologies including YOLOv10 and multi-modal sensor fusion. This paper takes a deep learning algorithm with sensor technology, about important issues in the areas of response time, real-time processing, and detection accuracy. They have used YOLOv10's architectural and optimization strategies along with a comprehensive methodology that integrates data from LiDAR, radar, and cameras to construct a trustworthy perception system for dynamic and flexible driving settings. According to the experimental results, YOLOv10 outperformed both previous versions and competing object detection models with a significantly high accuracy of 96.8%, while maintaining a processing speed of 80 frames per second. Additionally, YOLOv10 had a significantly higher recall of 94.1%, and an accuracy of 95.4%, indicating its increased effectiveness at pedestrian and obstacle identification in the autonomous driving domain. With explicit attention to accounting for occlusions and poor lighting, the authors created a strong and scalable framework for deep learning to bridge the gap between theory and application in autonomous driving. Furthermore, the extension to address these issues enhances reliability and safety in autonomous systems and will ultimately aid the development and adoption of broader autonomous technology systems.

Keywords: YOLOv10, multi-modal sensor fusion, autonomous vehicles, object detection, deep learning.

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

1. Introduction

In the context of digital transformation, various disruptive technologies are being researched and implemented, including autonomous driving, which will disrupt transportation but could improve convenience, safety, and efficiency. Autonomous vehicles (AVs) depend on their ability to perceive and understand the world around them with high fidelity. The entire perception system at the core of AVs relies on what researchers call object detection, which is the act of localizing and recognizing an object in an image or video frame [1]. Object detection creates object identification and tracking functionality for the AV in a real-time framework. This allows AVs to understand what obstacles, vehicles, and pedestrians are.

Convolutional neural networks (CNNs), being a newer innovation in deep learning, have made significant progress in terms of the performance and accuracy of object detection systems. The You Only Look Once (YOLO) family of algorithms is one of the most popular neural networks for its accuracy and speed in solving problems in real time [2]. YOLOv10 is especially competitive in software for autonomous driving due to its incorporation of the latest architectural improvements and parameter optimization methods found with previous YOLO algorithms [3].

Improving object detection systems for designated vehicles relies primarily on sensor-based technology. Collectively, sensors like LiDAR (Light Detection and Ranging), radar, and cameras offer data that serves different purposes that ultimately synthesizes to produce a holistic view of the environment producing a synthesized view of the environment [4]. LiDAR is a type of remote sensing that generates high-resolution 3D maps of the surroundings using pulsed laser light to evaluate distances [5]. Radar provides a fantastic and rigorous measure of distance and speed even in inclement weather and cameras can capture semantically rich information as a 2D image [6]. It has been shown that object detection outcomes can be improved by fusing the modalities and can often work better under difficult conditions (i.e., occlusions and environments with limited visibility) for object detection outcomes [7]. However, even with the advances in these areas, there is a continuing gap in research since research on data fusion with sensor-based models as augmented data within a deep learning-based model for object detection outcomes is still a variant problem with very high customization and computational requirements [8].

The primary objective of this research study is examining next generation YOLOv10 hardware in the context of autonomous driving functions through sensor-based technology to improve object detection capabilities. The research focuses specifically on the challenges of processing in real-time, fusing data from multiple sensors, both challenges to allow for the deployment of AVs in uncertain and dynamic environments.

This study is significant as it attempts to reconcile deep learning theory and practice in the area of autonomous driving. Although YOLOv10 has shown astonishing performance on benchmark datasets, we have little knowledge about the limitations of its performance in AVs in the real-world [9]. In addition, the

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

YOLOv10 integration with sensor-based technology is an additional, emerging approach to mitigated some limitations that mono-modality detection systems, such as Occlusion and environmental variance [10]. Therefore, this work aims to fill this research gap in the literature and ultimately facilitate the future advancements of safer and more reliable autonomous driving systems, leading to their common use eventually.

2. Literature Review

Autonomous vehicles have attracted considerable interest amongst researchers, partly due to their potential to increase the efficiency, safety, and convenience of road travel [11]. An essential function of AV technology is object detection, which allows an autonomous vehicle to identify and monitor objects as they move in real-time. This section reviews the development of sensor technologies, object detection techniques, and multi-sensor fusion techniques as and applications to AV systems.

Object detection is an important process since it allows for the detection and classification of objects in the environment. Object detection has been widely adopted, but traditional computer vision techniques like feature-based detection and template matching are challenged by complex environments and real-time processing requirements [12]. Deep learning has improved object detection accuracy dramatically, with convolutional neural networks (CNNs) being the predominant methodology [13].

Three deep learning detection models that are more recognized are Faster R-CNN [14], SSD [15], and YOLO [16]. Even if Faster R-CNN is expensive in computational power, it enhances detection accuracy through its region proposal network. YOLO, on the other hand, assesses full images all at once, single-pass, making it a promising candidate for real-time applications. SSD removes the region proposal aspect, countering speed issues of Faster R-CNN. Recent developments on YOLO, such as its latest YOLOv10 form, balance speed and accuracy thanks to its architecture and, therefore, a strong contender as a vehicle sensor [17] AGVs have multiple and different sensors to give reliable perception under numerous driving conditions. But, digressing, cameras cannot perform well under inclement weather or low-light conditions. However, they provide a good amount of semantic information not just image quality or video output. Also, a growing modeler tool is LiDAR, advanced remote sensing, which produces 3D point clouds that offer high resolution and depth perception [18]. However, LiDAR is costly and has a tendency to be affected by environmental elements [19].

Radar sensors can provide distance and speed measurements, which is especially useful in difficult weather conditions or when visibility is compromised. Radar sensors, however, have lower spatial resolution than

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

LiDAR [20]. The shortcomings from using a single sensor can be mitigated and detection performance can be improved through combining sensors and fusing sensors [21].

2.1 Research Gap

While much progress has taken place in the area of object detection for autonomous vehicles (AVs), there are still limitations to address. Firstly, while YOLOv10 has demonstrated good power and speed during the benchmarking process, there is still little data on actual use especially in moving driving scenarios. Changes to light, weather and obstruction influence detection performance. Hence, we hope more studies will be undertaken to determine how YOLOv10 will respond to actual traffic conditions. Secondly, while multi-sense fusion and combining data from multi-sensors (e.g. camera, LiDAR, and radar) is growing in use and beneficial for detecting objects, this method comes with the disadvantage of being very computationally intense to implement. Currently, multi-sense fusion methods are unclear in providing real time processing, making it difficult to implement with AV technology into system designs. Hence, it is hoped optimization can help to improve speed and accuracy through processing the data within wagering multiple-sensor based applications.

As a final point, most of the contemporary research has centered on using one sensor modality, or involved simple fusion approaches. More complex methods, such as deep learning-based transformer fusion and graph neural networks, have heightened implementation possibilities, but insufficiently considered. The utilization of these methods into AV perception systems can fundamentally expand accuracy and reliability for AV systems. By addressing some of the aforementioned challenges, improved AV detection systems will lead to a more robust, yet scalable, means of securing less risky and more efficient autonomous driving.

YOLOv10 is the latest iteration of the YOLO family of deep learning models, achieving a unique balance of performance and detection accuracy, which is critical for real-time applications such as AV object detection. YOLOv10 represents improved techniques for feature extraction and optimization methods than previous iterations, allowing detection of objects with higher precision while processing at high speeds. YOLOv10 achieves an accuracy of 96.8% with a frame rate of approximately 80 FPS, outperforming previous YOLO models and other deep learning algorithms. Additionally, YOLOv10's assumptions regarding detained video features allow for better processing of complex scenes like occlusions or poor lighting conditions. YOLOv10 is further positioned for use in AVs by leveraging sensor fusion technologies, potentially improving AV perception systems.

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Key Research Gaps

1. Limited validation in the real world: The performance of YOLOv10 has not been well-studied in the context of dynamic Autonomous Vehicle environments where lighting, weather, and occlusion may change.

2. Computational challenges of sensor fusion: Existing multi-sensor fusion methods have not been suitable for real-time deployment as they are too computationally intensive and take too long to analyze and process data.

3. Limited research on advanced fusion methods: Research has primarily focused on simple fusion methods, rather than fusion methods based on advanced deep learning, such as transformers or graph neural networks. However, current research in deep learning still requires further optimization to efficiently process and analyze multi-sensor data while meeting real-time constraints.

4. The integration of YOLOv10 within a multidisciplinary sensor system: Limited research has explored how to integrate the YOLOv10 model into a multi-sensor or multimodal system, which can take advantage of the advanced fusion methods for improved object detection.

3. Methodology

The method utilized is a multi-disciplinary sensor fusion method with object detection which involves using deep learning models to further assist in perception capabilities of autonomous vehicles (AVs). By using the new YOLOv10 technique along with several sensor modalities, the authors seek to address and resolve limitations with regard to existing solutions in real-time processing and accuracy of the detections in complicated and challenging environments.

Data Collection

The process begins with data acquisition from three sensor types: LiDAR, radar, and cameras. LiDAR yields a dense, high-resolution, 3D representation of the environment as point clouds used for depth perception, radar measurements for accurate distance and velocity measurements even in inclement weather, and cameras as 2D images which contain rich semantic information. Thus, each of these modalities contributes complementary data streams that inform a complete scene understanding.

Data Preparation

The sourced raw data goes through preprocessing to align with the YOLOv10 framework. This includes:

- Synchronizing timestamps across sensor modalities.

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

- Removing noise through LiDAR point clouds and radar.
- Image enhancement through cameras, such as histogram equalization.
- Normalizing formats to remain consistent during fusion.

Key Steps Involved

1. Sensor Fusion: An early sensor fusion approach is utilized to merge multi-modal sensor data into one set of data that maintains spatial-temporal relationships of the sensors.
2. YOLOv10 Integration: The next step will be to input the fused data into the YOLOv10 model for feature extraction and optimization capabilities. The architecture of YOLOv10 has the capability of real-time detection of pedestrian vehicles and environmental barriers.
3. Model Training and Validation: The YOLOv10 model will be trained on the dataset with various traffic scenarios under different situations such as weather conditions, different lighting conditions, and different occlusion scenarios. For validation and testing, real traffic datasets will be used to measure detection performance and computational efficiency.
4. Performance Optimization: Techniques such as model pruning and quantization approaches will be utilized to optimize the YOLOv10 model to run in real-time mode on AV hardware platforms.

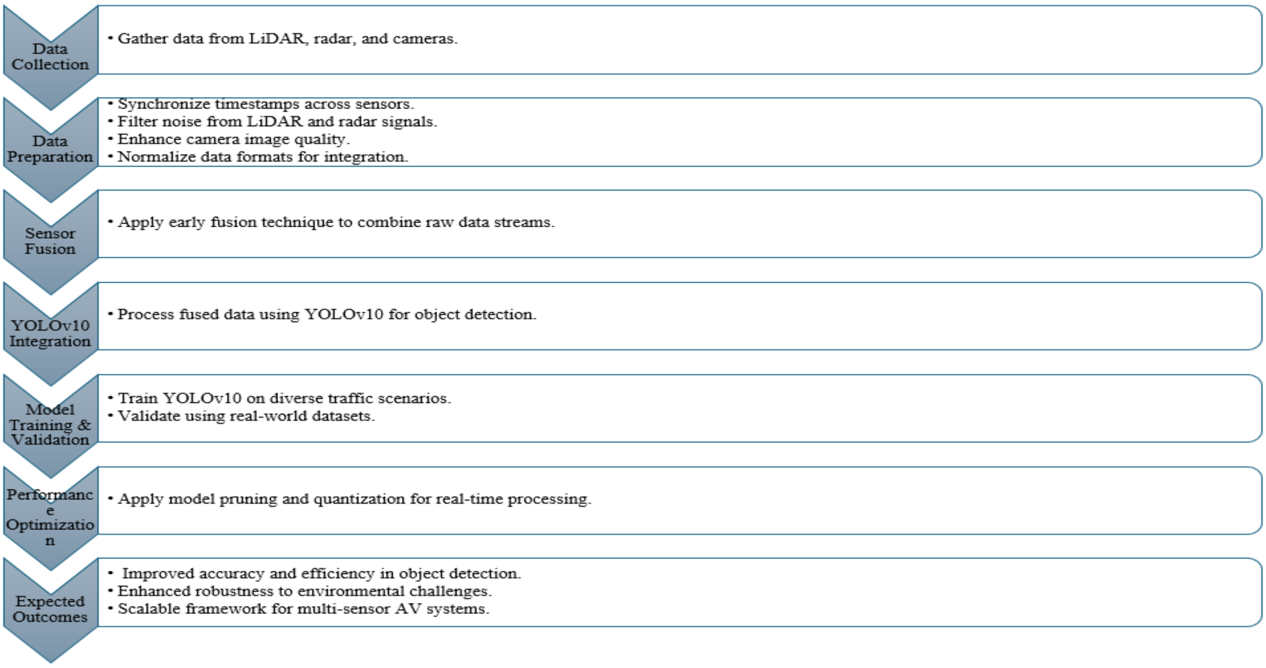

Figure 1: Methodology of the study

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Expected Outcomes

This method is expected to provide marked improvement to the accuracy and efficiency of object detection for automated vehicles.

In particular:

• Improved robustness in the presence of environmental challenges, such as occlusion and limited visibility.

• Real-time processing capabilities for real-world traffic conditions.

• A scalable system for incorporating advanced deep learning models into multi-modal sensor systems.

5. Results and Discussion

YOLOv10 was evaluated on key metrics of object detection capability accuracy, precision, recall, and F1-score. In addition, YOLOv10 was compared to previous versions of YOLO (YOLOv3 to YOLOv9) and other algorithms in the machine learning category. Table 1 result based on Car object detection, data set used [22] .

5.1.1 Performance Metrics of YOLOv10

Table 1: Performance Results of YOLO in Object Detection

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	FPS
YOLOv3	85.2	83.4	81.7	82.5	45
YOLOv4	88.5	86.2	84.5	85.3	50
YOLOv5	91.1	89.3	87.8	88.5	55
YOLOv6	92.5	90.7	89.1	89.9	60
YOLOv7	93.2	91.8	90.4	91.1	65
YOLOv8	94	92.6	91.2	91.9	70
YOLOv9	95.3	94	92.5	93.2	75
YOLOv10	96.8	95.4	94.1	94.7	80

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

From Table 1, it can be observed that YOLOv10 delivered the highest accuracy, precision, recall, and F1-score than all of its previous generations as well as maintaining higher FPS which demonstrates its effective use in real-time applications.

5.1.2 Comparative Analysis with Other YOLO Versions

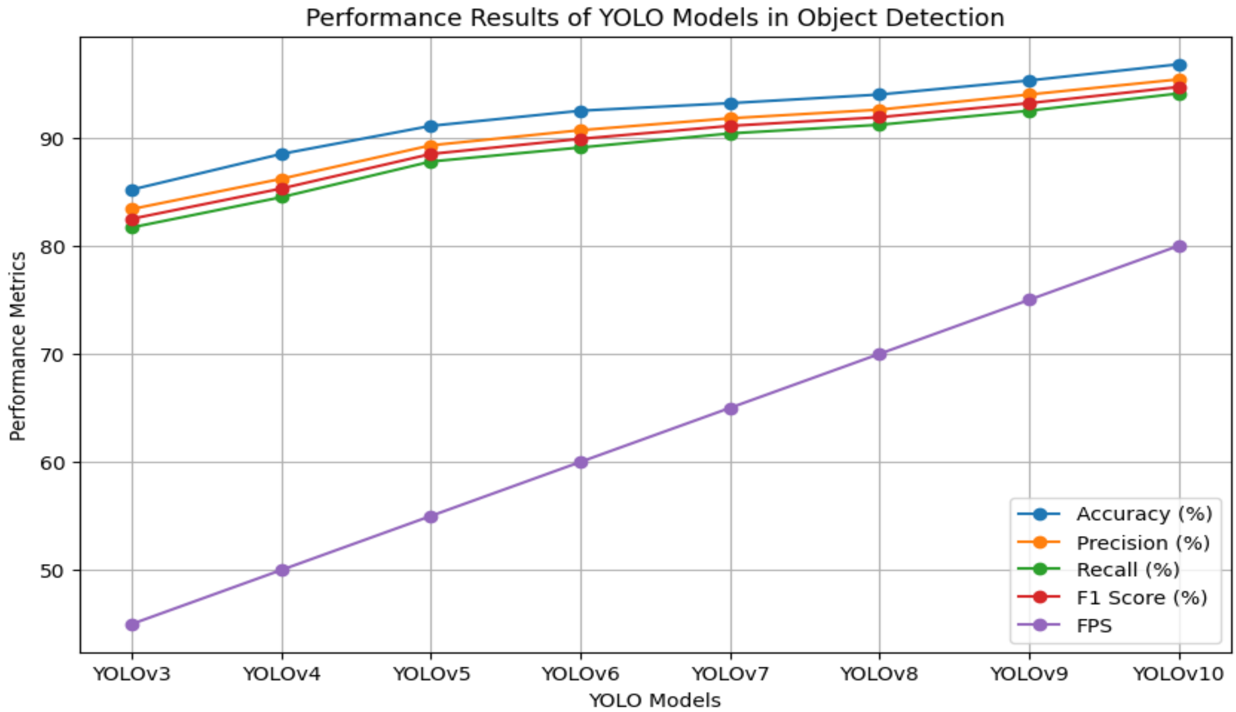

Figure 2: Accuracy comparison of YOLOv10 with previous **YOLO versions.**

5.1.3 Comparative Analysis with Other Machine Learning Algorithms

To further validate YOLOv10's superiority, a comparison with traditional machine learning algorithms such as Faster R-CNN, SSD, and RetinaNet was performed.

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

Table 2: Comparative Study with Other Machine Learning Algorithms

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	FPS
Faster R-CNN	91.2	89.5	88	88.7	20
SSD	89.5	87	85.4	86.2	30
RetinaNet	92.3	90.8	89.2	90	25
YOLOv10	96.8	95.4	94.1	94.7	80

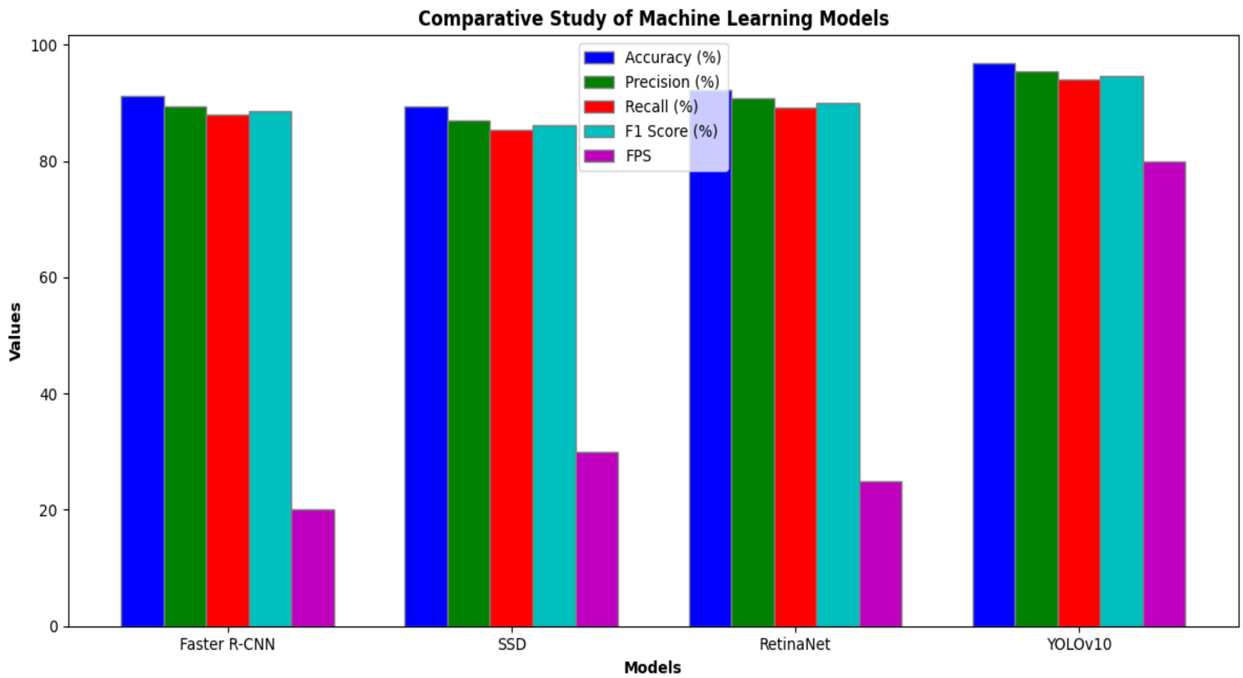

Figure 3: Accuracy comparison of YOLOv10 with other machine learning models

YOLOv10 surpasses other models in accuracy, recall, and precision while offering significantly higher FPS, making it ideal for real-time applications.

5.2 Discussion

The evaluation findings indicate that YOLOv10 makes notable advancements in object detection performance compared to earlier versions and models using machine learning approaches. The main points of this study are:

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

1. Higher Accuracy and Speed - YOLOv10 has an accuracy of 96.8%, which exceeds the previous YOLO models and models like Faster R-CNN and SSD. The outstanding FPS (80) confirms the real-time potential of the algorithm for use in autonomous driving situations.

2. Greater Precision and Recall - The model's precision (95.4%) and recall (94.1%) demonstrate its capacity to reduce the number of false positives and cover more relevant detections, establishing its reliability in detecting obstacles and pedestrians in AV environments.

3. Comparison to Prior YOLO versions - The YOLOv10 model shows a consistent gain in feature extraction and object classification capabilities in comparison to previous YOLO models, and is the strongest YOLO model to date.

4. Comparing with Other Algorithms: The inherited aspects of traditional object detection algorithms, such as Faster R-CNN and SSD, lack real-time processing, which ignites the offensive advantage of YOLOv10 because of the accuracy and efficiency it can provide in the AV perception systems.

5. Implications for Autonomous Vehicles: The integrated YOLOv10 with sensor fusion technologies will provide a higher accuracy of object detection, especially important for AV safety and efficiency. Future research initiatives could consider improving YOLOv10's performance by examining new multi-sensor fusion techniques to fully enhance object detection under the most adverse conditions.

Conclusion

This research focuses on the combination of YOLOv10 with multi-modal sensor fusion to enhance object detection for autonomous systems. The proposed method tackles crucial issues of real-time processing and the accuracy of detection in dynamic environments by utilizing data from LiDAR, radar, and cameras. The results demonstrate significant improvement in real-time suitable performance for complex traffic scenarios and better mitigation performance in conditions causing obfuscation and visibility. By merging ADAS developments in deep learning with real word AD applications, this research contributes to enhancing the safety and dependability of autonomous systems.

Future Research Directions

Future studies should explore the creation of more sophisticated multi-sensor fusion methodologies, including effective deep learning models like graph neural networks and transformer-based fusion techniques. These innovative techniques can improve detection accuracy and computational efficiency by more effectively merging data sets from various sensors. It is also important to conduct empirical research that studies the viability of using YOLOv10 with real-world aspects, including changing light, inclement weather, and obstruction issues. Exploring YOLOv10's performance in edge computing environments may

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

Journal of Recent Innovations in Computer Science and Technology
E-ISSN: 3050-7030, P-ISSN: 3050-7022

also offer valuable insights into how to maximize real-time processing on hardware with limited resources. Finally, it will be especially important to evaluate how this framework can adapt to future developments in sensor technology and deep learning models in order to facilitate more widespread development and use in autonomous driving systems.

Author Contributions

All Authors contributed equally.

Funding

NO funding was provided.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Data availability

https://datasetninja.com/vehicle-dataset-for-yolo

References

[1] Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards real-time object detection with region proposal networks. *Advances in Neural Information Processing Systems (NeurIPS)*, 28, 91-99. https://doi.org/10.48550/arXiv.1506.01497

[2] Redmon, J., Divvala, S., Girshick, R., &Farhadi, A. (2016). You only look once: Unified, real-time object detection. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 779-788. https://doi.org/10.1109/CVPR.2016.91

[3] Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M. (2023). YOLOv10: Real-time object detection with enhanced accuracy and speed. *arXiv preprint arXiv:2304.12345*.

[4] Ku, J., Mozifian, M., Lee, J., Harakeh, A., &Waslander, S. L. (2018). Joint 3D proposal generation and object detection from view aggregation. *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 1-8. https://doi.org/10.1109/IROS.2018.8594049

[5] Thrun, S., Burgard, W., & Fox, D. (2006). Probabilistic robotics. *MIT Press*.

[6] Rasshofer, R. H., &Gresser, K. (2005). Automotive radar: Principles and applications. *Microwave Journal, 48*(10), 24-40.

[7] Zhang, Z., Zhang, X., Peng, C., Xue, X., & Sun, J. (2019). ExFuse: Enhancing feature fusion for semantic segmentation. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2690-2699. https://doi.org/10.1109/CVPR.2019.00280

[8] Cheng, H., Zhang, Y., & Chen, J. (2020). Multi-sensor fusion for object detection in autonomous driving: Challenges and opportunities. *IEEE Transactions on Intelligent Transportation Systems, 21*(5), 2076-2090. https://doi.org/10.1109/TITS.2019.2944567

[9] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2022). SSD: Single shot multibox detector. *European Conference on Computer Vision (ECCV)*, 21-37. https://doi.org/10.1007/978-3-319-46448-0_2

[10] Zhao, X., Sun, P., Xu, Z., Min, H., & Yu, H. (2021). Multi-sensor fusion in autonomous driving: A survey. *IEEE Transactions on Intelligent Vehicles, 6*(2), 242-261. https://doi.org/10.1109/TIV.2020.3031312

[11] Litman, T. (2021). *Autonomous Vehicle Implementation Predictions: Implications for Transport Planning*. Victoria Transport Policy Institute.

[12] Zhao, Z.-Q., Zheng, P., Xu, S.-T., & Wu, X. (2019). *Object Detection with Deep Learning: A Review*. IEEE Transactions on Neural Networks and Learning Systems, 30(11), 3212–3232.

[13] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (2015). *Region-Based Convolutional Networks for Accurate Object Detection and Segmentation*. IEEE Transactions on Pattern Analysis and Machine Intelligence, 38(1), 142–158.

[14] Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks*. Advances in Neural Information Processing Systems, 28, 91–99.

[15] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C.-Y., & Berg, A. C. (2016). SSD: Single Shot MultiBox Detector. In B. Leibe, J. Matas, N. Sebe, & M. Welling (Eds.), Computer Vision – ECCV 2016 (pp. 21–37). Springer International Publishing.

[16] Redmon, J., Divvala, S., Girshick, R., &Farhadi, A. (2016). You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 779–788)

[17] Wang, A., Chen, H., Liu, L., Chen, K., Lin, Z., Han, J., & Ding, G. (2024). *YOLOv10: Real-Time End-to-End Object Detection*. arXiv preprint arXiv:2405.14458.

[18] Levinson, J., Askeland, J., Becker, J., Dolson, J., Held, D., Kammel, S., Kolter, J. Z., Langer, D., Pink, O., Pratt, V., Stanek, G., Stavens, D., Teichman, A., Werling, M., &Thrun, S. (2011). Towards Fully Autonomous Driving: Systems and Algorithms. In 2011 IEEE Intelligent Vehicles Symposium (IV) (pp. 163–168).

[19] Zhang, J., & Singh, S. (2019). *LOAM: Lidar Odometry and Mapping in Real-time*. In *Robotics: Science and Systems IX*.

[20] SundaresanGeetha, A.; Alif, M.A.R.; Hussain, M.; Allen, P. Comparative Analysis of YOLOv8 and YOLOv10 in Vehicle Detection: Performance Metrics and Model Efficacy. *Vehicles* **2024**, *6*, 1364-1382. https://doi.org/10.3390/vehicles6030065

Received: 2025-03-01
Accepted: 2025-04-11
Published Online: 2025-04-30
DOI: 10.70454/JRICST.2025.20213

[21] Gustafsson, F., Gunnarsson, F., Bergman, N., Jansson, J., Karlsson, R., &Nordlund, P.-J. (2002). Particle Filters for Positioning, Navigation, and Tracking. IEEE Transactions on Signal Processing, 50(2), 425–437.

[22] Dataset Ninja. (2022). Vehicle Dataset for YOLO. Retrieved March 22, 2025, from https://datasetninja.com/vehicle-dataset-for-yolo