Volume - 15 | Issue - 08 | August - 2026 | PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH

OPTIMIZING DEEP LEARNING MODELS FOR REAL-TIME OBJECT DETECTION IN AUTONOMOUS SYSTEMS

Computer Science

Archana R Department of Computer Science and Engineering, AMC Engineering College India

Dr. Nirmala S Department of Computer Science and Engineering, AMC Engineering College India

ABSTRACT

The autonomous systems comprising self-driving cars, drones, and industrial robots separately rely on real-time object detection systems for their safety and operational effectiveness. This paper focuses on the implementation challenges posed by the deep learning YOLO (You Only Look Once) architecture for object detection on the resource-constrained embedded systems common in autonomous vehicles. Although YOLO models are accurate, their significant computational requirements are impediments to their use in real-time systems. This work seeks to achieve these objectives through model refinement by optimizing the trade-off between performance and accuracy for dependable real-time detection in a varying operational terrain. A comprehensive review of the literature and autonomous systems case studies reveal gaps in object localization, real-time decision, environmental adaption, and technology fusion. The study results have highlighted the need to employ sophisticated algorithms in image processing, detection, and multi-condition detection responsiveness utilizing the hybrid methods framework. The work pursues the technological refinement of real-time adaptive systems algorithm for improving the reliability and effectiveness of autonomous systems in mobile applications.

KEYWORDS

Autonomous Systems, Autonomus Vehicles, Convolutional Neural Networks (CNN), Real-time Object Detection, YOLO

I. INTRODUCTION

Real-time object detection powers autonomous systems like self-driving automobiles, UAVs, and industrial robots. This capacity to quickly and precisely detect and localize items in dynamic and unexpected environments is crucial for operational efficiency and safety. Autonomous systems must evaluate massive sensory input and make crucial judgments quickly [1]. Street items and traffic signs may be catastrophically misidentified, which illustrates the importance of powerful object recognition for safe operation, informed decision-making, and accurate control in autonomous driving [2].

One of the fundamental technologies which are the basis of the whole system of self-driving cars is real-time object detection. Precisely to be able through the method of detection and localization to identify quickly and correctly even in the most intricate and uncertain surroundings what the objects are - people, other cars, traffic signs, and obstacles - is the feature that guarantees the system working effectiveness and, most importantly, safety. However, the revolutionary accuracy of the detecting objects has made the deep learning models, in particular, CNNs like YOLO (You Only Look Once) family. Nevertheless, the computational heaviness of these models brings about a number of big problems for the real-time use on embedded systems with limited resources that are installed in the vehicles which have the autonomous feature. The article depicts the issues the case study has and the different ways of optimization that have been used to get the desired effect of reaching the performance of real-time without compromising the accuracy of the results [3].

II. Previous Study

Wang, et al [4] their study provides a YOLOv4-based one-stage object identification technique that improves accuracy and real-time operation. The technique increases residual block stacking times, substitutes SPP with RFB, improves PAN structure, adds attention mechanisms, and decreases network broadness. Arthi, et al [5] examines the difficulties encountered by computer vision systems in identifying barriers and objects in Autonomous Vehicles, particularly in severe weather circumstances, highlighting the need for efficient image processing techniques and research on weather and object recognition.

Balasubramaniam & Pasricha [6] highlights object detection's importance in consumer applications like surveillance, mobile text recognition, and medical diagnostics, as well as autonomous driving. It highlights deep learning-based detectors' role in real-time localization and challenges.

An autonomous vehicle, also known as a driverless vehicle, is a system that is capable of performing all of the functions necessary to ensure that it is performing its proper tasks on its own. This is accomplished through the utilization of a fully automatic driving system that is able to respond to all external conditions that would normally be handled by a driver, based on its ability to sense the environment. Autonomous cars are anticipated to be the most important technical breakthrough because of their ability to transform mobility and traffic safety, as well as substantially alter urban transportation (whether it be passenger or freight) [7].

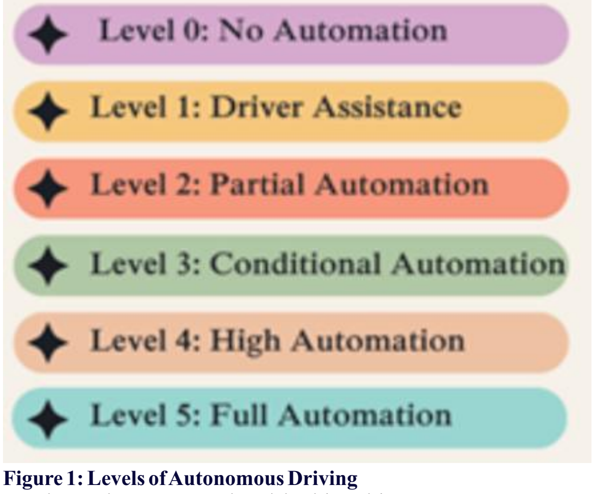


Figure 1: Levels of Autonomous Driving

Level 0 cars have no control, and the driver drives.
Level 1 ADAS may aid with steering, accelerating, and braking.
Level 2, the ADAS can handle steering, accelerating, and braking, but the driver must stay focused on the road while performing the rest of the responsibilities.
Level 3, the ADS (advanced driving system) can do all driving tasks in specific situations, but the driver must restore control when asked. In the remaining cases, the driver acts.
Level 4 ADS can drive autonomously in some situations without human monitoring.
Finally, level 5 automates all vehicle ADS functions under all situations without driver intervention. 5G technology will allow cars to interact with each other, traffic signals, signs, and even the roadways, enabling complete automation.

III. Objectives

O1 - To study autonomous system safety, decision making and operational efficiency using real time object detection.
O2 - To evaluate deep learning object identification algorithms.
O3 - To identify computing issues with YOLO methods on low resource embedded devices used in autonomous systems.

IV. METHODOLOGY

The study examines autonomous systems and describes their design, implementation, and optimization methods, technologies, and approaches.

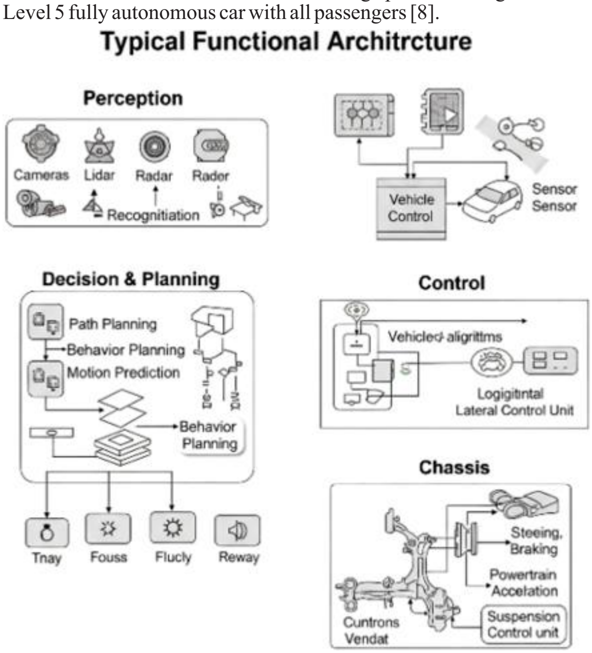
 PRINT ISSN No. 2277 - 8179 | DOI : 10.36106/ijsr

Collecting Data: Our literature study publications, journals, and conference proceedings. Data sources included highly reputable databases, including IEEE Xplore, ScienceDirect, and Google Scholar.

Case Studies: Autonomous system implementations were examined. Example: autonomous vehicle deployment.

V. Architecture of Autonomous Vehicles

By integrating sensors that sense the surroundings and regulate the vehicle's motion, a regular vehicle may become autonomous. Figure 2 shows the AV communication protocol, sensors, actuators, hardware, and software control needed. The four-stage protocol design allows a Level 5 fully autonomous car with all passengers [8].



Perception: This stage includes taking in the AV's surroundings and determining its location. In this stage, the AV uses RADAR, LIDAR, cameras, RTK, etc. These sensors provide data to recognition modules for processing. AVs often have ADAF, a control system, LDWS, TSR, UOR, VPL, and other components. This processed data is combined for decision-making and planning.

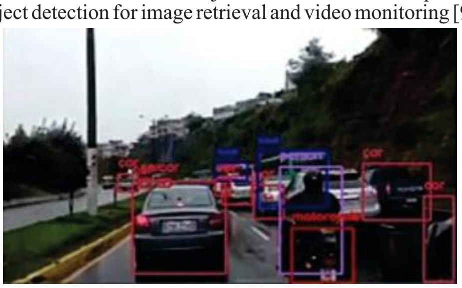
Choosing and Planning: This stage determines, plans, and directs AV motion and behavior using perceptual data. This brain-like stage produces decisions, including route planning, action prediction, obstacle avoidance, etc. Recent and historical data, such as real-time map data, traffic patterns, user input, etc., influences the choice. For future reference, a data log module may capture mistakes and information.

Control: The control module gets information from the decision and planning module and controls the AV's steering, braking, acceleration, etc.

Chassis: The last level interfaces with chassis mechanical components such as the accelerator pedal, brake pedal, steering wheel, and gear motors. The control module signals and controls these components.

VI. Object Detection

Object detection in digital photos and movies detects semantic things like individuals, buildings, and automobiles. Face and pedestrian identification are well-studied object detection areas. Computer vision uses object detection for image retrieval and video monitoring [9].



Today, several detecting techniques exist. One-stage and two-stage detectors exist. Two-stage detectors employ deep learning to approximate areas of observed objects, which are subsequently classified to name the item. They restrict the boxes where things are expected to be. One-stage detectors display object bounding boxes without region bounds. Such methods are speedier, less time-consuming, and suitable for real-time applications [10].

VII. Challenges

- Autonomous systems must accurately localize, estimate, and navigate various sensory inputs for coherent understanding [11].
- Autonomous systems must be trustworthy, hence a hybrid approach using model-based and machine learning approaches is advised, with simulation and testing for global validation [12].
- Testing autonomous systems under various settings is necessary to assure reliability. Autonomous vehicles must travel safely in various weather, terrain, and traffic settings. Creating rigorous testing standards is expensive and time-consuming, delaying autonomous technology adoption.

VIII. Object Detection Using YOLO

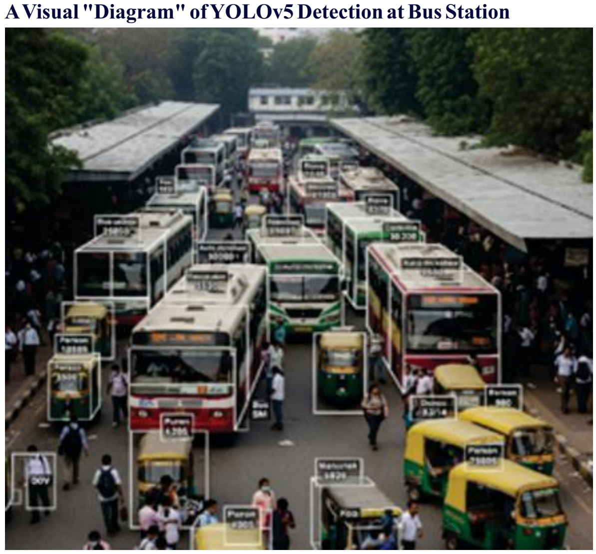
The YOLO method is the most popular object detection technique because to its fast detection speed. The 12-layer YOLO method has been trained on over 80 classes, requiring a lot of processing power and data. We trained the YOLO algorithm on as few classes as feasible to minimize overloading and decrease processing resources. Those classes are: A person, tree, traffic light, car, road sign, bus, bike, truck, etc. [11, 12]

Classification-based YOLO has two steps: selecting the Region of Interest (ROI) and selecting a fully connected CNN, which typically deals with grid structures. CNNs' key feature is parameter sharing. CNN uses input weights to learn a set of weights for a location, which makes object detection reliable. Figure 4 illustrates the YOLO algorithm stages for object detection. Images are scaled from camcorder footage. The convolutions move through the picture frames, the outputs are filtered by non-max deletions, and the detections are drawn on the results to get the required file. Python was used for image detection since it has image detection libraries. We deployed NVIDIA's CUDA GPU tool to conduct sections of the calculation in parallel, speeding up the system [13].

Perform exercises and run interfaces. PyTorch, the most popular autonomous learning framework, runs computations and interfaces on the GPU, making it faster. We installed matplotlib, numpy, opencv-python, Pillow, torch, torchvision, tensorflow, pycocotools, and roboflow for detection and training. Classification and location are needed to recognize objects in video and images. For this, we need to provide our algorithm a training data set containing photos and labels [13, 12].

IX. Results

A Visual "Diagram" of YOLOv5 Detection at Bus Station



- A Base Image: This would be a photo or a video frame taken at the Bus Station. A typical image of this place is usually crowded and therefore would depict:

- Multiple people: who are walking, waiting and boarding buses.
- Multiple vehicles: Buses, auto-rickshaws, and possibly two-wheelers.
- Staff and passengers: As well as bus drivers, conductors, security, and cleaners, there will be passengers carrying their luggage.Sunlight conditions: Some parts of the scene may be in shadow and other greatly lit.
- Numerous people: Walking, waiting, boarding buses.
- Various vehicles: Multiple buses (BMTC, KSRTC), auto-rickshaws, and possibly two-wheelers.
- Station infrastructure: Bus bays, platforms, signboards, shelters.
- Varying light conditions: Perhaps some areas in shadow, others brightly lit.
- Overlaid Bounding Boxes: Boxes that can be seen through or partially opaque are drawn around each detected object.
- Class Labels: Text labels put at the nearest point to each bounding box that gives the class of the detected object (e.g., "person," "bus," "auto-rickshaw," "luggage").
- Confidence Scores: The number, either percentage or decimal, next to the label shows the model's confidence in that detection (e.g., "person 95%," "bus 88%").

"Diagram" Frame:

Envision an image taken from an oblique angle looking down at a bus bay.

In the foreground, there is a large red bus that is made more prominent with a bright red bounding box around it and the label "bus 92%".

Persons are distributed in large numbers around the bus, some are queuing while some others are walking. Each individual would have their own blue bounding box and the label "person XX%". Besides that, you might spot further away "person" boxes that are smaller in size.

Next to the bus's side, an auto-rickshaw is just visible by the edge of the picture, hence with a green box and "auto-rickshaw 85%".

On the platform, along those few bags few parts of luggage have been detected, each with a yellow box and "luggage XX%".

In the background, other buses and people would also have their respective bounding boxes, possibly smaller and with slightly lower confidence scores due to distance or occlusion.

Places of heavy occlusion, like a dense crowd might illustrate overlapping "person" boxes, or the model might even not detect some individuals at all, which clearly indicates a challenge.

X. CONCLUSION

Real-time object detection is a key factor in the performance and safety of autonomous systems. Deep learning models such as YOLO improve the object detection accuracy to the sky but at the same time they are still facing the problem of real time application of autonomous vehicles. The main goal of past research is to make the methods of the identification of objects more accurate and real-time for the autonomous vehicles.

The architecture of an autonomous vehicle can be seen through processes such as perception, decision-making, planning, control, and integration of the chassis that is all for fully autonomous driving. One-stage detectors are perfect for real-time object detection. YOLO's rapid detection pace and training of specific classes help in reducing the processing resources. Using YOLO means first picking the regions of interest and then applying fully connected CNNs. Problems such as changing light conditions and occlusion are the reasons for the decrease of the accuracy of the object detection. The progress in the field of object detection with the help of YOLO technology has definitely a great impact on the safety and efficiency of the autonomous systems.

REFERENCES

[1] "A. Kumar, Deep learning models for real-time object detection in autonomous systems, Journal of Research in Computer Science and Engineering, vol. 9, no. 3, [Online]. Available: ISSN 2457-0818.".

[2] "M. Contreras, A. Jain, N. Bhatt, A. Banerjee and E. Hashemi, A survey on 3D object detection in real time for autonomous driving, Frontiers in Robotics and AI, vol. 11, 2024, Art. no. 1212070, doi: 10.3389/frobt.2024.1212070.".

[3] "Khan, N. A., Hussain, M., & Ashraf, S. (2025). Optimizing Convolutional Neural Networks for Real-Time Object Detection in Autonomous Vehicles. The Critical Review of Social Sciences Studies, 3(2), 1440-1454.".

[4] "Wang, R., Wang, Z., Xu, Z., Wang, C., Li, Q., Zhang, Y., & Li, H. (2021). A Real.Time Object Detector for Autonomous Vehicles Based on YOLOv4. Computational Intelligence and Neuroscience, 2021(1), 9218137.".

[5] "V. Arthi, R. Murugeswari and N. P, Object Detection of Autonomous Vehicles under Adverse Weather Conditions, 2022 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI), Chennai, India, 2022, pp. 1-8, doi: 10.1109/ICDSAAI55.".

[6] "Balasubramaniam, A., & Pasricha, S. (2022). Object detection in autonomous vehicles: Status and open challenges. arXiv preprint arXiv:2201.07706..".

[7] "Cugurullo, F., & Acheampong, R. A. (2020). Smart cities. In O. Jensen, B. Lassen, V. Kausfmann, M. Freudendal-Pedersen, & I. S. G. Lange (Eds.), Handbook of urban motilities. Routledge.".

[8] "Ahangar, M. N., Ahmed, Q. Z., Khan, F. A., & Hafeez, M. (2021). A Survey of Autonomous Vehicles: Enabling Communication Technologies and Challenges. Sensors, 21(3), 706. https://doi.org/10.3390/s21030706".

[9] "Dasiopoulou, Stamatia, et al. "Knowledge-assisted semantic video object detection." IEEE Transactions on Circuits and Systems for Video Technology 15.10 (2005): 1210–1224.".

[10] O. D. i. 2. T. D. G. A. o. https://viso.ai/deep-learning/object-detection/..

[11] "Alatise, M., & Hancke, G. (2020). A Review on Challenges of Autonomous Mobile Robot and Sensor Fusion Methods. IEEE Access, 8, 39830-39846. https://doi.org/10.1109/ACCESS.2020.2975643.".

[12] "Sifakis, J., & Harel, D. (2023). Trustworthy Autonomous System Development. ACM Transactions on Embedded Computing Systems, 22, 1 - 24. https://doi.org/10.1145/3545178.".

[13] "Bratulescu, Razvan-Alexandru & Mitroi, Sorina-Andreea & Vatasoiu, Robert-Ionut & Vochin, Marius & Suciu, George & Sachian, Mari-Anais. (2022). Object Detection in Autonomous Vehicles. 10.1109/WPMC55625.2022.10014804.".

[14] "Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. (2016, October). Ssd: Single shot multibox detector. In European conference on computer vision (pp. 21-37). Springer, Cham.".

[15] "Sarda, A., Dixit, S., & Bhan, A. (2021). Object Detection for Autonomous Driving using YOLO algorithm. 2021 2nd International Conference on Intelligent Engineering and Management (ICIEM).".

[16] "Sharma, T.; Debaque, B.; Duclos, N.; Chehri, A.; Kinder, B.; Fortier, P. Deep Learning-Based Object Detection and Scene Perception under Bad Weather Conditions. Electronics 2022, 11, 563.".