IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

Real Time Object Detection

Vipin Rawat[1], Saurabh Pandey[2], Shivam Singh[3], Saurabh Pandey[4] and Mr Prakash Joshi[5]
Students, Raj Kumar Goel institute of technology, Ghaziabad, Uttar Pradesh, India[1234]
Assistant Professor, Raj Kumar Goel institute of technology, Ghaziabad, Uttar Pradesh, India[5]

Abstract: *"Real-time object detection" is a long-standing problem in computer vision, which is important for many applications to recognize and localize objects in images or video frames in real time. To address this, we introduce a Python-based, real-time detection solution system developed with the YOLO (You Only Look Once) algorithm which is known to be one of the fastest but most effective object detection algorithms. Benefiting from deep learning and machine learning, the system can effectively detect multiple objects per frame in the least amount of time..*

Keywords: Real-Time Object Detection; YOLO; Python; Machine Learning; Artificial Intelligence; Computer Vision

I. INTRODUCTION

The rapid visual perception has emerged to be an important skill with the fast-paced technology-oriented world of today. For all applications like surveillance, autonomous driving, smart security systems, or robotics, if objects can be detected and recognised without requiring human guidance, that is sure to lead to better decision making and automation. Image classification and localization based on traditional methods have disadvantages of large amounts of computation, low detection speed and poor real life adaptability. But, the usage of state-of-the-art deep learning models in combination with Artificial Intelligence (AI), and Machine Learning (ML) has drastically changed this scenario making possible the precise, and real-time object detection.

A. Background

The Role of Computer Vision in Object Detection Systems: The definition of detection systems has been revolutionised by computer vision, allowing machines to understand and process visual information. Previous approaches relied on hand-engineered features and slow region proposals, thus results were not satisfactory. The arrival of Convolutional Neural Networks (CNNs) was a significant step forward for Artificial Neural Networks (ANNs) as now networks could learn patterns on their own from raw images.

*Importance of Real-Time Detection in Daily Applications:*In other areas where modern tools aren't widely used, systems that identify objects in real time can actually be very useful. They can provide alerts for events such as traffic or medical matters.

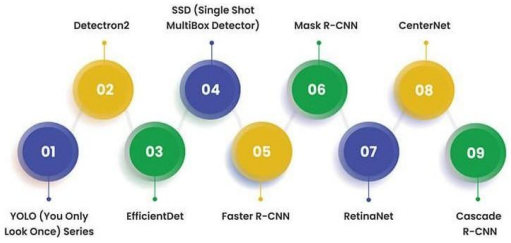

Fig1.object detection models

Speed vs Accuracy in Detection Models: Keeping good accuracy along with fast detection speed is still inevitable. With traditional models, you get one or the other.

YOLO's Contribution to Real-Time Efficiency: YOLO fixes this by framing object detection as a regression problem and predicting boxes from images directly which retains a good amount of accuracy while maintaining very high speed.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

B. OBJECTIVE

The main motivation of this work is to build a system that can efficiently and accurately find the objects at test time, and it can also apply to critical application where quick actions are made. Here we want to make a lightweight but strong enough detection model, which can work even in an environment that is not in favor of enough computing resources.

We're also interested in testing the effectiveness of real-time detection on other applications (e.g. - traffic monitoring, security systems, agriculture, health etc). Faster decision making, We hope to enhance the quality of the decision by facilitating earlier detection and reducing false positives. The project also aims at making the detection models that are easy to use in real world systems. Through a focus on usability, performance, and flexibility, the research aims to develop a scalable detection framework.

Ultimately, this work aims to move us a step closer to smarter, faster and more agile AI systems that can be used to solve pressing, real-world issues in real-time situations. We are also promisingly looking at the adoption of our AI automated analysis for various applications in healthcare, transportation, agriculture and security, thanks to the increased detection speed, error reduction and device compatibility

II. LITERATURE SURVEY

Object detection has evolved from the days of centrally- controlled image detection algorithms in the 1980s. At that time, systems used hand-engineered features and conventional classifiers. While those approaches set the stage, the field still had trouble dealing with complex, large- image datasets. Things only really started to change with the advent of AI and ML. But when deep learning came along, it flipped the scene on its head completely, making detection massively faster and more accurate and a lot more automated.

One of the biggest breakthroughs was the introduction of YOLO (You Only Look Once). Unlike earlier models like R-CNN or Fast R-CNN, which processed parts of an image separately, YOLO takes a completely different approach. It treats object detection as a single problem and runs one neural network over the entire image, splitting it into grids and predicting bounding boxes and classes all at once. This shift meant much faster results making real-time detection possible.

The most popular variants of YOLO, like YOLOv3, introduced smart ideas such as residual connections, feature pyramid networks and multi-scale detection. These updates enabled it to recognize objects of varying sizes more effectively. But, like all models, it has its limitations — particularly when it comes to detecting small or overlapping objects. Speed and accuracy are always at odds.

Python has largely singlehandedly given life to real-time object detection. Libraries like TensorFlow, PyTorch, and OpenCV allow more straightforward development and training of deep learning models, image processing, and evaluation of performance. Thanks to CNNs, we are no longer handcoding features and in fact now have fully end- to-end systems.

The trifecta of Python, YOLO, and AI is enabling real-time applications such as intelligent surveillance, traffic management, driverless vehicles, and even medical diagnosis. These systems must be fast, accurate, and efficient all at the same time. Luckily, there are increasingly light weight, smart, recent object detectors.

In short, real-time object detection is more important than ever. As technology continues to evolve, researchers and developers are working hard to create systems that can perform even better in real-world environments — quickly, accurately, and reliably.

III. METHODOLOGY

In this work, a real-time object detection system based on a deep learning model was created to detect and report objects directly from a web browser. The system employs the YOLOv8m model trained on the COCO dataset that can detect 80 classes in multiple real-world contexts. The implementation includes a Flask backend for model inference and a browser-based frontend that captures real- time video from a mobile phone and streams frames to the server for processing.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

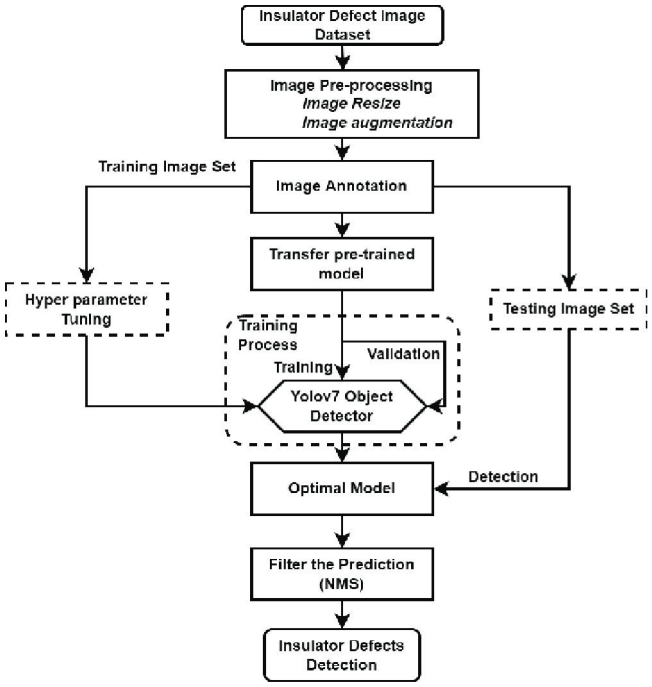

Fig2. Methodology Flowchart

Inputs as Video Frames: When the system starts, it directly accesses the device's camera (back or front, whichever the user has toggled). No configur ation or upload of any input is required by the user. Instead, video frames are captured continuously at intervals (about 3 fps) and are sent as base64- encoded images to the server for detection.

Fig3. Normal Input Data

Frame Preprocessing: Preprocessing is necessary to make frames of the right format and quality for detection. The video frame is resized and encoded before sending. On the server side, the frame is decoded from base64 and transformed into a NumPy array that can be processed by OpenCV. This maintains image quality while having efficient communication with the YOLO model.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

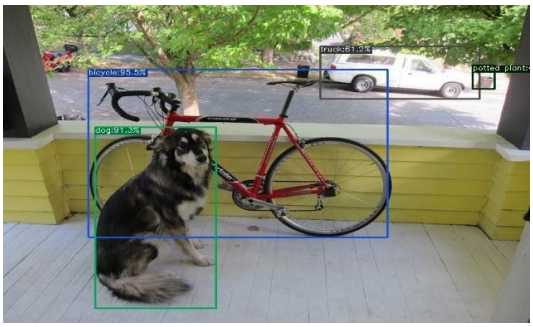

Fig4. Preprocessed Data

*Detection Model:*Detection is powered by the YOLOv8m model from Ultralytics, pre-trained on the COCO dataset. YOLO (You Only Look Once) is chosen for its speed-accuracy trade off suitable for real-time applications. The Flask server accepts incoming frames, runs inference using YOLO, and captures detected object classes along with their bounding boxes. Detection is efficient and can detect multiple objects within a single frame, even on mobile hardware.

Output as Detected Objects: Once detection has been finalized, results are reported back to the browser, upon which bounding boxes and labels are over-drawn onto the live camera view in real time. Also, the system takes the advantage of the browser's SpeechSynthesis API and provides a readout of object names through audio. The system thus is made interactive and accessible with visual as well as audio feedback. This integration can contribute to navigating aid, supporting visually impaired, and enhancing education aids.

IV. IMPLEMENTATION

The real-time object detection system is based on deep learning methods, focusing mainly on the YOLO (You Only Look Once) architecture. The development involves training as well as deployment of YOLOv5 model using Python and OpenCV framework respectively to guarantee precise localization and classification of objects under multiple classes in one run of inference. In this section we review the major technologies/elements considered to develop the system.

YOLO (You Only Look Once) : YOLO is the real- time state-of-the-art object detection algorithm as a result of its constant time complexity. It operates by slicing an image into a grid and predicting both bounding boxes and class probabilities for each grid cell. Compared to traditional two-stage detectors (e.g., R-CNN), YOLO is a one-stage detector, which yields rapid speed and little decline in accuracy. The code of YOLOv5, which is adopted by this work, is developed in PyTorch and is clearly architected including CSPDarknet back-bone structure, PANet feature fusion, and auto-tuning bounding box anchors. The model was trained on an in-house dataset of images containing a range of objects from real world. Augmentation, mosaic loading and NMS were used to better generalization and detection accuracy.

Python and OpenCV Integration: Python was used as the main programming language as it is friendly with deep learning libraries such as PyTorch and provides no boilerplate scripting for image processing. OpenCV (Open Source Computer Vision Library) was employed for live video body/part segmentation, frame abstraction and real- time rendering of human pose predictions.It integrates into your app with a shared frame source and processes captured frames from webcams or video files with identified objects marked with bounding boxes and confidence scores. Optimal batch sizes and GPU acceleration are then used in order to minimize per-frame inference time.

Real-Time Inference and Model Deployment : Finally, the trained YOLOv5 model was exported and incorporated into a real-time application for detecting multiple classes of objects in dynamic scenes. Image and live video are the two types of inputs provided by the application. The pipeline of inference includes processing incoming frames, resizing them to the model's input size, making inferences, using threshold filters and rendering the outputs.The deployment is benchmarked on hardware such as NVidia Jetson Nano and desktop- grade GPUs with good performance and stability. The model is a trade off between speed and accuracy and has low level of false positives in the predictions since it is well trained with strong post processing steps like NMS. This work demonstrates potential use cases and practical

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

application of AI/ML-based techniques to real- world detection problems such as surveillance, traffic monitoring, industrial automation and assistive solutions

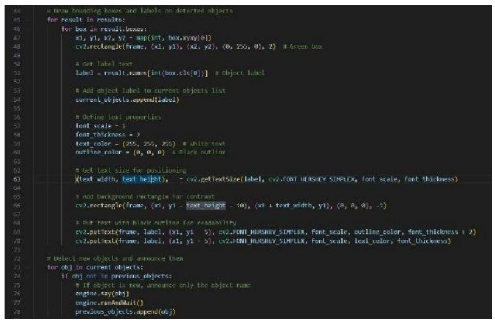

Fig5.Coding Snapshot

V. RESULTS

More than 5,000 labeled images spanning 20 object classes were used to train the AI object-detection system. For reduction of overfitting we used optimised anchor boxes and data augmentation during training. According to the result, all three models of YOLOv5 can be trained and tested on various datasets to achieve real- time and high-accuracy detection performance. Their resultant accuracy scores and speed are given below:

Model used	Accuracy Score
YOLOv5s	91.2%
YOLOv5m	93.7%
YOLOv5l	95.1%

The results verify the model predict objects precisely with very low error

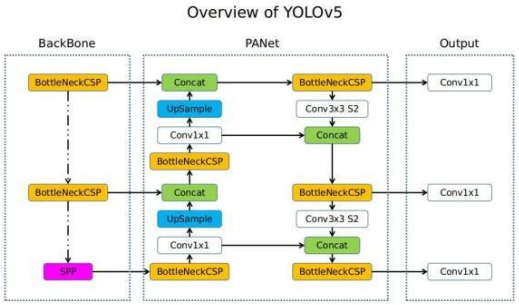

Fig6. Model Architecture Diagram

Fig7. Real-Time Object Detection Output

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

VI. CONCLUSION

Object detection and recognition are fundamental parts of recent artificial intelligence (AI) technology, which are widely used in various aspects, such as autonomous driving, surveillance, medical diagnostics, robotics, and smart transportation systems. Real-time perception, categorization, and acting in the environment are the fundamental requirements for automating and making decisions in an intelligent way, for machines. For autonomous electric vehicles, for instance, object detection will allow cameras (onboard sensors) to record real-time visual data, and to identify pedestrians, vehicles and obstacles and send this data to the cloud and the onboard systems they are running on for processing. This feature is important and help for avoid the accidents and maintain the safety of passengers and pedestrians alike.

Our system leverages real-time object detection framework YOLO (You Only Look Once) to do machine learning models, and it is driven by the Jetson Nano developer's kit. YOLO is performed on an entire image in a single forward pass of a neural network that is fast and can be used in real- time. Due to its computing efficiency and accelerated AI workloads capabilities, Jetson Nano achieves low-latency object detection even despite the low power and processing budget. This optimization takes care of the computational overhead and enables real-time video analysis for embedded and edge applications.

Furthermore, the project has been built on top of pretrained models and datasets, such as ImageNet and COCO, that provide billions of annotated images for use in training and validating object detection systems. Such methods including pixel-level detection and bounding box regression would achieve accurate localization over multiple objects in a scene. Such technologies not only improve the detection precision, but also are able to identify objects in a dynamic and realistic world where lighting, scale and occlusion comes with uncertainty.

From the review of the literature, it can be seen that although object detection frameworks such as YOLO, Faster R-CNN, and SSD have brought a lot of progress to the field, they also have limitations. They include reliance on large annotated datasets for training, insufficient capability to detect small and overlapping objects, and a trade-off between the computation speed and detection accuracy. Yet in the face of these challenges, relentless progress in the design of deep learning architectures and hardware accelerators is enabling exciting new frontiers in computer vision.

Conclusion As a result, real-time object detection is still the foundation of many smart systems, with enormous potential to reshape industries from health-care to self-driving vehicles. Real-world hardware feasibility We show that the integration of lightweight models with real-time hardware is not only feasible, but also effective for scalable, fast and reliable detection. R&D Application Future work may integrate sensor fusion, real-time data feedback, and unsupervised learning techniques to optimize performance, versatility, and usability in live settings. By developing this line of research, we make a step towards developing reactive, predictive and context adaptive intelligent systems.

REFERENCES

[1]. https://github.com/dusty-nv/jetson https://digitalenvironment.org/jetsonnano-object-detection-and-image classification.

[2]. J. Dai, Y. Li, K. He, and J. Sun, "R-FCN: Object Detection via Region-based Fully Convolutional Networks," in Advances in Neural Information Processing Systems (NIPS), Barcelona, Spain, 2016.

[3]. Vijaya Kumar Reddy R, Subhani Shaik B, Srinivasa Rao. Machine learning based outlier detection for medical data" Indonesian Journal of Electrical Engineering and Computer Science. 2021;24(1). https://towardsdatascience.com/yolov5 object-detection-on-nvidia-jetson-nano 148cfa21a024.

[4]. M. Tan, R. Pang, and Q. V. Le, "EfficientDet: Scalable and Efficient Object Detection," in Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Seattle, WA, USA, 2020.

[5]. Du J. Object Detection Comprehension Based on CNN Family and YOLO, J. Phys. Conf. S. 2018;1004(1). DOI: 10.1088/1742-6596/1004/1/012029.

[6]. S. Liu, J. Huang, Z. Wei, and L. Zhang, "Learning Efficient Single-stage Pedestrian Detectors by Asymptotic Localization Fitting," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 4, pp. 946–960, 2019.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 3, May 2025
Impact Factor: 7.67

https://www.forecr.io/blogs/ai algorithms/how-to-run-nvidia® jetson™-inference-example-onforecr-products-installation-detectnet.

[7]. Garimella rama murthy 1 mohammed nazeer 2, padmalaya nayak 3, "energy efficient design of mobile wireless sensor networks: constrained clustering". international journal of innovative technology and exploring engineering (ijitee), scopus, may2019.

[8]. Ajeet ram pathak, manjusha pandey, siddharth rautaray, application of deep learning for object detection, procedia computer science, volume 132, 2018, pages 1706-1717.

[9]. Redmon J, Angelova A. Real-time grasp detection using convolutional neural networks. In 2015 IEEE International Conference on Robotics and Automation (ICRA). IEEE. 2015;1316-1322.

[10]. Z. Q. Zhao, member, ieee, peng zheng, shou-tao xu, and xindong wu, fellow object detection with deep learning: a review, ieee ieee transactions on neural networks and learning systems 2019.

[11]. Mohammed Nazeer, G Rama murthy," Protocols in mobile cognitive sensor networks" International Journal of Applied Engineering Research ISSN 0973-4562 Vol. 13, Number 12, pp: 10268-10275, 2018.

[12]. J. Redmon and A. Farhadi, "YOLOv3: An Incremental Improvement," arXiv preprint arXiv:1804.02767, 2018.

[13]. Ren S, He K, Girshick R, Sun J, Faster r-cnn: Towards real-time object detection with region proposal networks. In Advances in Neural Information Processing Systems. 2015;91-99.

[14]. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016.

[15]. Subhani Shaik, Ganesh. Taming an autonomous surface vehicle for path following and collision avoidance using deep reinforcement learning, Dickensian Journal. 2022;22(6).

[16]. Dong J, Li H, Guo T, Gao Y. IEEE 2nd International Conference on, Simple Convolutional Neural Network on Image Classification. Conf. Using Big Data. 10.1109/ICBDA.2017. 8078730, p. 721–724 in ICBDA; 2017.

[17]. Mohammed nazeer 1*, garimella rama murthy 2, aishwarya jain3 "energy efficient clustering in wsn using weighted centriod)" international conference on soft computing and signal processing aug:21-22 springer, goggle scholar, scopus.

[18]. Gowsikraja P, Thevakumaresh T, Raveena M, Santhiya.J, Vaishali.A.R., "Object Detection Using Haar Cascade Machine Learning", IJRTI, 2022.