INTERNATIONAL JOURNAL OF NOVEL RESEARCH AND DEVELOPMENT (IJNRD)
© 2026 IJNRD | Volume 11, Issue 5, May 2026 | ISSN: 2456-4184 | IJNRD.ORG

Object Detection and Recognition in Real-Time Video Streams

1st ***D.PRIYANKA***

Department Of Advanced Computing And Analytics Vels Institute Of Science, Technology & Advanced Studies Chennai, Tamil Nadu

2nd DR.MUTHUKUMARA

Department Of Advanced Computing And Analytics Vels Institute OF Science, Technology & Advanced Studies Chennai, Tamil Nadu

Abstract - The ability to identify objects in real-time video streams is among the most significant tasks in computer vision and artificial intelligence. It is commonly employed in video surveillance, business monitoring, autonomous vehicles, robotics, and smart megacity operations. The target of this prototype is to characterize and recognize human, vehicle, and other moving objects from live video streams in high-speed and high-accuracy settings.

Traditional image processing methods do not work for dealing with real-time detection because of constraints in speed and accuracy. Thus, this design uses deep learning-based object detection models (similar to YOLO You Only Look Once) and Convolutional Neural Networks (CNN) for the performance of accurate real-time detection. The system takes each frame of a video stream with the corresponding objects, recognizes them, draws a bounding box, and labels accordingly.
Our new system exhibits high detection accuracy, low latency, and high performance. Through its application and experimental findings, it appears that deep learning-based methods have a significant impact in enhancing object detection performance relative to traditional methods. This design contributes to the real-time intelligent monitoring environment at work and the findings of smart automation.

I. INTRODUCTION

Object detection is a computer vision algorithm for spotting objects from images and video frames. The first task is to visualize the object with a camera, search for object detection in the image or video files, and detect any suspicious motion, such as movement. The detection should occur in real time in video streams with speed and accuracy without delay. This makes the task harder: the system must process numerous frames per second.

As artificial intelligence and deep learning have developed, object detection processes are more effective and efficient. Such systems are involved in real-time object detection and control applications for CCTV, traffic analysis systems, airport security, smart city systems, and autonomous driving. The main objective of this work is to design a real-time object detection/recognition process that can analyze live video and correctly recognize a set of objects. That system captures video input, processes frames using deep learning models, detects objects, and displays results in bounding boxes and labels.

II. LITERATURE SURVEY

Object detection has been the subject of many researches. Traditional image processing methods as well as Haar Cascade were used for object detection before.

Nevertheless, these methods were not able to cope with complex backgrounds and several objects. Later deep learning techniques, like R-CNN, Fast R-CNN and Faster R-CNN, improved detection accuracy but it also required high computational demands. Adding YOLO (You Only Look Once), a network can detect objects in a single pass through the network: a more efficient solution. SSD (Single Shot Detector) was another feature with real-time detection capabilities. Research Gap:

Most existing systems had been either focused on accuracy or speed. Achieving both high accuracy and real-time performance is challenging at present. Proposed Solution:

This is a project to deploy a deep learning-based YOLO model to get high accuracy and work in real time.

III. PROPOSED METHODOLOGY

The proposed a systematic workflow:

1. **Video Capture** – Capture live video using webcam or load recorded video.

2. **Frame Extraction** – Convert video stream into individual frames.
3. **Preprocessing** – Resize and normalize frames for model compatibility.
4. **Model Loading** – Load pretrained YOLO model.
5. **Object Detection** – Detect objects using neural network predictions.
6. **Bounding Box Generation** – Draw bounding boxes around detected objects.
7. **Labeling & Confidence Score** – Display object name with probability score.
8. **Output Display** – Show processed frames in real-time.

This methodology continuous frame processing without delay.

IV. ARCHITECTURE DIAGRAM the system consists of multiple modules working together:

Video Input → Frame Processing → Preprocessing → YOLO Model → Object Detection → Bounding Box & Labeling → Real-Time Display

This efficient memory management and optimized frame processing speed.

V. INPUT

the following inputs:

- Live webcam video stream
- Pre-recorded video files
- Image frames for testing

the deep learning model for detection and recognition.

VI. PSEUDOCODE

the library and loading the YOLO model that has been pretrained in system starting. The webcam/video file is turned on to record live video frames. Every single frame is resized and reconstructed into a suitable format that gets fed into the object detection model.

It predicts object categories and bounding box positions. The object is valid if its prediction confidence is greater than or equal to a set threshold. Bounding boxes and labels are made for the frame. The processed frame is shown on the screen. This process continues until the user exits the application.

VII. OUTPUT

The system output is a video stream in real time with detected objects marked. Every object is enclosed inside a bounding box and labeled with its name and confidence percentage. The system can pick up several objects within the same frame. The result looks smooth, video playback as it is without much buffering, hence effective real-time monitoring.

VIII. RESULT AND DISCUSSION

In testing the system was performed indoors and outdoors using various video inputs. The YOLO model is able to accurately detect different frames in a real-time system. Hardware configuration also controls speed of detection, especially hardware equipped with GPU support. The proposed deep learning-based process exhibited significantly faster and more accurate detection results than conventional object detection techniques. Performance was however slightly slower in low-light and highly crowded scenes. In general the system performed very well for surveillance and smart monitoring.

IX. SCREENSHOT

Fig 1: Input video

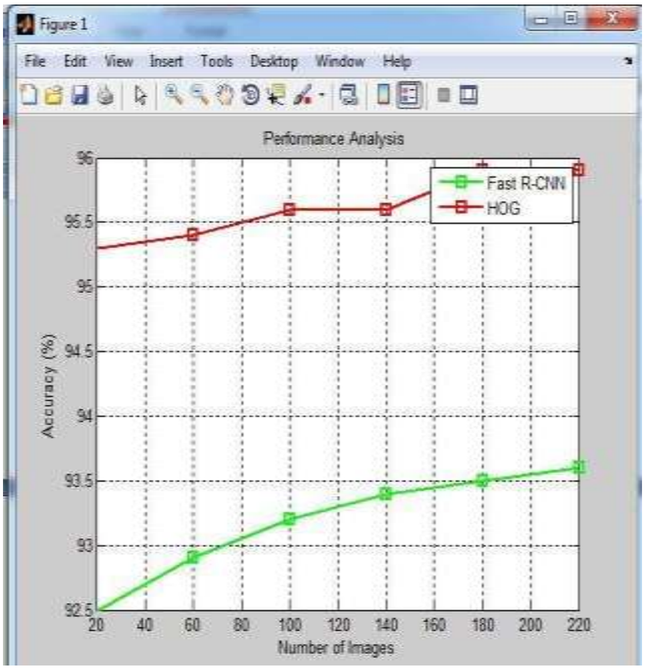

Fig 3: Accuracy Graph

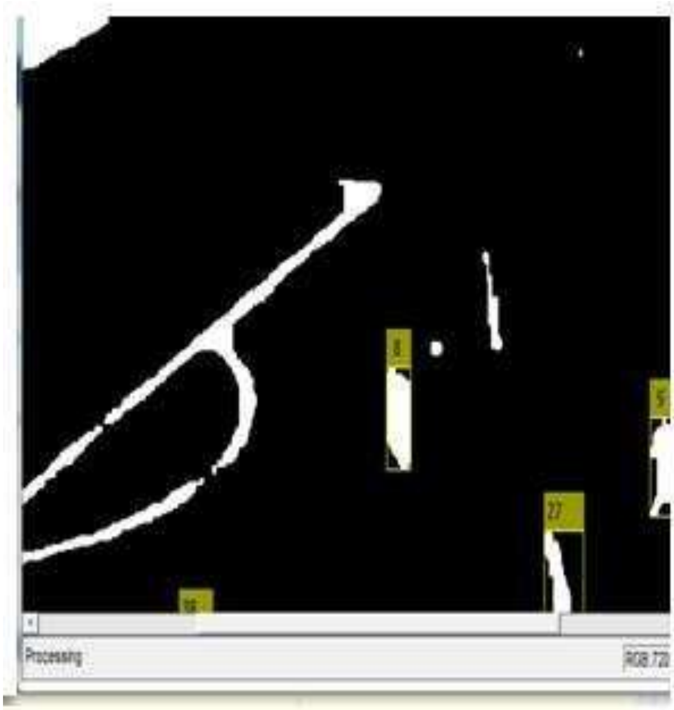

Fig 2: Output video

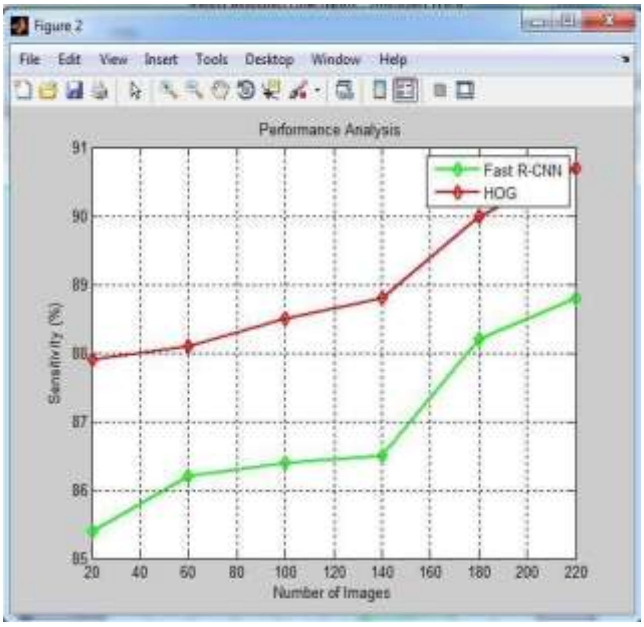

Fig 4: Sensitivity Graph

X. CONCLUSION

This work gives a framework for the development of a deep learning-based real-time object detection and recognition system. The YOLO model guarantees fast detection with good accuracy. The system is applicable to surveillance, traffic monitoring, industrial automation, and security systems. Upcoming improvements could be object tracking, face recognition, and integration with cloud-based monitoring systems.

XI. REFERENCES

[1] Redmon, J., et al., "YOLO: You Only Look Once," 2016.
[2] Girshick, R., "R-CNN," 2014.
[3] Liu, W., et al., "SSD: Single Shot Detector," 2016.
[4] OpenCV Documentation.
[5] Goodfellow, I., Bengio, Y., Courville, A., "Deep Learning," MIT Press, 2016.