International Journal Research of Leading Publication (IJLRP)
E-ISSN: 2582-8010 ● Website: www.ijlrp.com ● Email: editor@ijlrp.com

Real Time Object Detection System by using YOLO Framework

Awadhesh Singh Chouhan[1], Shivam Rathore[2], Roshan Dhomane[3], Girdhar Kalamdhad[4], Prof. Nilesh Mishra[5]

[1, 2, 3, 4]B.Tech scholar, [5]Assistant Professor
Department of Artificial intelligence & Data Science
Shri Balaji Institute of technology and Management
Betul, RGPV University M.P., India

Abstract

Real-time object detection has emerged as a crucial task in computer vision, with applications spanning autonomous vehicles, surveillance, and robotics. This paper provides a comprehensive analysis of state-of-the-art object detection models, particularly focusing on YOLO variants (YOLOv3, YOLOv7) and their comparative performance. Various training methodologies, including step-by-step training, Faster R-CNN, and hybrid deep learning models, are explored to highlight advancements in accuracy and computational efficiency[1]. Experimental results demonstrate that YOLOv7 surpasses YOLOv3 in accuracy, while achieves state-of-the-art performance without additional pretraining datasets. This study contributes to the ongoing development of real-time detection systems by analyzing the trade-offs between speed, precision, and practical implementation.

Keywords: Real-Time Object Detection, YOLO, Faster R-CNN, Deep Learning

1. INTRODUCTION

Object detection is a fundamental aspect of computer vision that enables the localization and classification of multiple objects in images or videos. Traditional methods like sliding window detection have been replaced by deep learning-based approaches such as Convolutional Neural Networks (CNNs), Single Shot Detectors (SSD), Faster R-CNN, and the You Only Look Once (YOLO) series.Recent advancements have focused on improving detection speed while maintaining accuracy, particularly for real-time applications such as autonomous vehicles and surveillance. This paper explores and compares multiple deep learning-based object detection models, including YOLOv3, YOLOv7, and to determine their suitability for real-world implementations.

2. LITERATURE REVIEW

2.1 Real-Time Object Detection Techniques

Object detection algorithms have evolved from traditional handcrafted feature extraction methods to deep learning models. The most widely used deep learning architectures for object detection include:

- **YOLO (You Only Look Once):** A one-stage detector known for its high speed and reasonable accuracy.
- **Faster R-CNN:** A two-stage detector that offers high accuracy but is computationally expensive.

According to Alsharabi (2023), real-time object detection relies on Convolutional Neural Networks (CNNs), hardware accelerators (GPUs, TPUs), and optimization techniques such as non-maximum suppression and feature pyramid networks.[1]

2.2 Step-by-Step Training in Object Detection

Ouyang et al. (2023) propose a novel training method called step-by-step training in which eliminates the reliance on additional datasets for pretraining. By leveraging a pre-trained YOLO model to initialize the backbone and encoder, achieves state-of-the-art performance while reducing computational costs.

2.3 Comparative Performance of YOLO Variants

Jadhav et al. (2023) evaluate YOLOv3, Tiny-YOLOv3, and Faster R-CNN for real-time object detection in construction vehicles. Their results indicate that YOLOv3 outperforms other models in both speed and accuracy. Similarly, Ramana et al. (2023) propose a hybrid model combining YOLO and Faster R-CNN, optimizing real-time detection accuracy while maintaining high inference speed.

A comparative study between YOLOv3 and YOLOv7 highlights that YOLOv7 improves upon YOLOv3 in precision and robustness, making it more suitable for real-time applications.

3. SYSTEM ARCHITECTURE AND DESIGN

3.1 Functional requirements

1. Input Module (Data Acquisition)

The Input Module is the first stage in a real-time object detection system, responsible for acquiring, preprocessing, and managing data before it is passed to the detection model. This module ensures efficient data ingestion while maintaining real-time processing capabilities.

Source: Live camera feeds, video stream, captured images.

2. Feature Extraction (Backbone Network)

The Feature Extraction Module is a crucial part of a real-time object detection system. It transforms raw image/video input into meaningful representations, allowing the detection model to accurately identify and classify objects. This module leverages deep learning models, mainly Convolutional Neural Networks (CNNs), to extract hierarchical features at different levels.

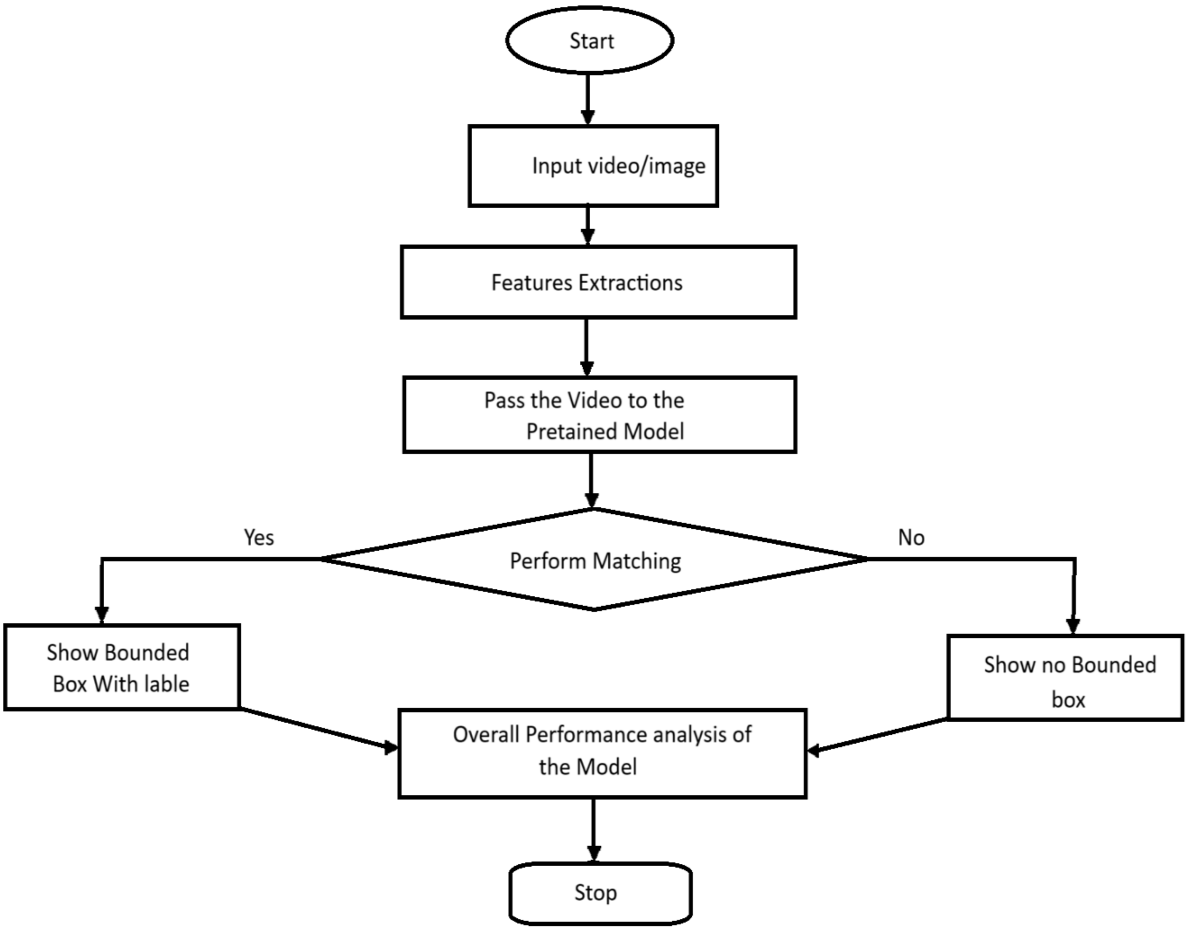
3. Object Detection Model (Detection Head)

The Object Detection Model is the core of a real-time object detection system. It takes processed image data from the Feature Extraction Module and detects objects by predicting bounding boxes and class labels in real-time**.**

Key Functions:

- Localization: Determines where objects are located in an image (bounding box regression).
- Classification: Identifies what each detected object is (labeling).
- Speed Optimization: Ensures real-time inference for applications like surveillance, autonomous driving, and robotics.[5]

3.2 Entity Relationships Diagram



4. METHODOLOGY

4.1 Data collection

For real-time object detection, data collection involves capturing images or video streams, labeling objects with bounding boxes or other annotations, and preparing the dataset for training machine learning models, often using tools like OpenCV and libraries like COCO.

4.2 Preprocessing

In real-time object detection, data preprocessing involves preparing images or video frames for analysis by techniques like resizing, noise reduction, and data augmentation, ensuring efficient and accurate object detection.

4.3 Model Selection

For real-time object detection, model selection hinges on balancing speed and accuracy. Popular choices include YOLO (You Only Look Once) for its speed, and for a good trade-off, while TensorFlow offers a wide range of pre-trained models.

5. RESULT

Detecting an object using the laptop's camera. Through this, we are finding out the time stamp of that object.

6. FUTURE SCOPE

The future Scope of real-time object detection is incredibly promising, with applications spanning diverse industries such as healthcare, autonomous systems, robotics, smart cities, and security. As AI models improve and processing power increases, real-time object detection will become even more seamless and pervasive in everyday life, transforming how we interact with our environments and technology.

7. CONCLUSION

This study highlights the advancements in real-time object detection, particularly comparing YOLO variants and hybrid deep learning techniques. While YOLOv7 provides a balance between speed and accuracy, introduces an innovative training method that eliminates the need for large-scale pretraining. Future research should focus on optimizing these models for low-power devices and real-time embedded systems.

We introduce YOLO, a unified model for object detection. Our model is simple to construct and can be trained directly on full images. Unlike classifier-based approaches, YOLO is trained on a loss function that directly corresponds to detection performance and the entire model is trained jointly. Fast YOLO is the fastest general-purpose object detector in the literature and YOLO pushes the state-of-the-art in real-time object detection. YOLO also generalizes well to new domains making it ideal for applications that rely on fast, robust object detection.

REFERENCES

1. Alsharabi, N. (2023). "Real-Time Object Detection Overview: Advancements, Challenges, and Applications," *Journal of Amran University*.
2. C.-Y. Wang, A. Bochkovskiy, and H.-Y. M. Liao, "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors, "2022, arXiv:2207.02696.
3. Redmon, Joseph; Divvala, Santosh; Girshick, Ross; Farhadi, Ali (2016-05-09). "You Only Look Once: Unified, Real-Time Object Detection".
4. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors Authors: Chien-Yao Wang, Alexey Bochkovskiy, Hong-Yuan Mark Liao Published: July 2022.
5. Hui Zhou , Meiwei Kong , Hexiang Yuan , Yanyan Pan , Xinru Wang , Rong Chen , Weiheng Lu , Ruizhi Wang , Qunhui Yang(2024).” Real-time underwater object detection technology for complex underwater environments based on deep learning. Ecological Informatics Volume 82, September 2024, 102680