IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10847

Real-Time Object Detection: Harnessing Advanced Machine Learning Algorithms

Anand Balagar[1], Ms. Bhavani R[2]

PG Scholar, Department of MCA, The National Institute of Engineering, Mysore, India[1]

Assistant Professor, Department of MCA, The National Institute of Engineering, Mysore, India[2]

Abstract: In the ever-evolving fields of computer vision and machine learning, real-time object detection presents a significant challenge. Our project dives into this domain, harnessing cutting-edge machine learning algorithms. By leveraging advanced methodologies like YOLO variants and efficient backbone architectures, our system aims to redefine real-time object detection. Traditional methods often compromise either accuracy or speed, limiting their effectiveness in dynamic environments. Our approach seeks the ideal balance, achieving both precise and swift object identification in complex scenarios. This integration of state-of-the-art techniques empowers our system to serve diverse sectors, from autonomous systems and security to interactive technologies. Through this endeavour, we envision a future where intelligent visual perception sets new standards for real-time object detection.

Keywords: Object detection, YOLOv7, Machine Learning, Faster R-CNN.

I. INTRODUCTION

This project represents a pioneering effort at the forefront of contemporary technology, focusing on advancing real-time object detection using advanced machine learning techniques. Object detection, a fundamental aspect of computer vision, has evolved significantly due to rapid advancements in machine learning algorithms. The project aims to redefine the accuracy and efficiency of object identification in dynamic environments, emphasizing the crucial balance between accuracy, speed, and adaptability.

In today's fast-paced world, swift object recognition within live video streams is essential across various sectors, from security and surveillance to autonomous vehicles and augmented reality. Advanced machine learning algorithms, such as Faster R-CNN and YOLOv7, have expanded the boundaries of object detection, each offering unique approaches to address its challenges.

The project's significance extends beyond technical innovation, impacting public safety, industrial processes, and interactive applications. It contributes to the field of computer vision, advancing the practicality of deploying intelligent systems globally. This endeavour seeks to bridge theory and practical implementation, exploring algorithm intricacies, strengths, and limitations while aiming to illuminate new horizons in real-time object detection through advanced machine learning algorithms.

II. LITERATURE SURVEY

[1] The paper introduces YOLOX, an advanced object detection model that improves upon previous YOLO models by enhancing both accuracy and speed. YOLOX achieves this through innovations such as the efficient CSPDarknet53 backbone architecture, optimized scaling factors, and the incorporation of a PANet module for better feature extraction. It outperforms previous YOLO versions in benchmark evaluations, demonstrating superior accuracy and real-time capabilities. YOLOX's scalability makes it adaptable for various computer vision applications, positioning it as a significant advancement in real-time object detection for fields like autonomous vehicles and surveillance.

[2] The paper introduces YOLO, a revolutionary real-time object detection method that revolutionizes the field. YOLO employs a unified architecture to predict object classes and bounding box coordinates simultaneously, optimizing both speed and accuracy. It divides the input image into a grid, predicting objects within each grid cell for comprehensive detection at multiple scales. YOLO sets a high standard for instant object recognition, inspiring subsequent versions like YOLOv2 and YOLOv3, which continue to advance real-time object detection with efficiency and accuracy.

[3] The paper introduces YOLO9000, an evolution of the YOLO architecture designed to handle a vast range of object classes across different datasets, achieving real-time processing speed. YOLO9000 employs multi-label classification and hierarchical classification using WordNet, enabling object detection across 9000 classes. It improves class predictions by utilizing a hierarchy and introduces detection confidence scores to address shared object names. YOLO9000 advances

DOI: 10.17148/IARJSET.2023.10847

class diversity and performance, marking a significant step in real-time, large-scale object detection with broad applications.

[4] The paper presents YOLOv3 as a significant advancement in real-time object detection, striking a balance between speed and accuracy. It employs a multi-scale approach using three different scales to detect objects of various sizes and incorporates the Darknet-53 architecture for complex feature extraction. YOLOv3 introduces route and skip connections to facilitate information flow across network layers. It significantly improves object detection accuracy while maintaining real-time processing, making it a notable milestone in the field.

[5] The paper introduces Faster R-CNN, a groundbreaking method that enhances real-time object detection by integrating region proposal generation and object classification into a single end-to-end network. It introduces a Region Proposal Network (RPN) that efficiently generates region proposals, sharing features with the detection network for accurate localization and classification. This unified approach eliminates the need for separate region proposal methods, significantly improving efficiency. Faster R-CNN's anchor boxes handle object scale and aspect ratio variations, achieving remarkable detection accuracy and speed. It serves as a cornerstone in object detection, setting new standards for real-time, accurate detection in applications like robotics, surveillance, and autonomous driving.

III. PROPOSED WORK

The system's primary goal is to transform real-time object detection using state-of-the-art machine learning techniques. By incorporating advanced methodologies like optimized YOLO variants and efficient backbone architectures, it will facilitate precise and rapid identification of objects in constantly changing environments. This integration of cutting-edge approaches promises to significantly enhance the accuracy and speed of object detection in dynamic settings.

IV. METHODOLOGY

The project begins with the collection and preprocessing of a diverse dataset, ensuring it covers a wide range of object classes and real-world scenarios. This dataset serves as the foundation for training and evaluating the YOLO v7 model. YOLO v7, a standout in the YOLO series, is chosen as the primary object detection model. Its unique architecture allows it to simultaneously predict object categories and bounding box coordinates, making it ideal for real-time applications. The YOLO v7 model is fine-tuned on the COCO dataset. Transfer learning is employed, benefiting from the model's pre-trained weights, which have learned valuable features from extensive data.

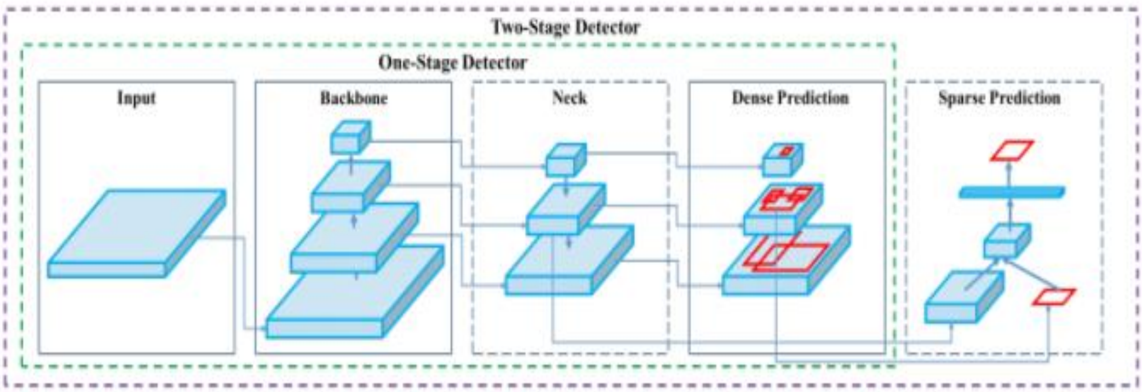

Fig. 1 YOLO Model Architecture

The project focuses on achieving real-time object detection by deploying the trained YOLO v7 model on live video streams. It processes each frame in real-time, detecting objects, and annotating them with class labels and bounding boxes. The accuracy and real-time processing speed of the YOLO v7 model are rigorously evaluated using appropriate metrics, including Mean Average Precision (mAP) and frames per second (FPS). The model's performance is assessed across various scenarios and object classes to ensure its robustness.

The important modules used in proposed system are as follows:

A. **Home Page:** This module educates users about real-time object detection using YOLOv7 through navigation links, image displays, and descriptive text, providing insights into the project's core concepts.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10847

B. **Image/Video Feed:** This module is used to get the image or video input from the user and set the confidence value for the input data and then it processes the input data and displays the detected image / video as a output with some measures.

C. **Webcam Feed:** This module is used to get the live input from user through webcam and to display the output as well with identified class label, accuracy and bounding box.

D. **Contact Page:** This Module creates a visually pleasing "Contact Us" webpage with navigation links, contact cards, and a footer. The styling and structure aim to provide an engaging user experience for visitors who want to get in touch with the website owner through various communication channels.

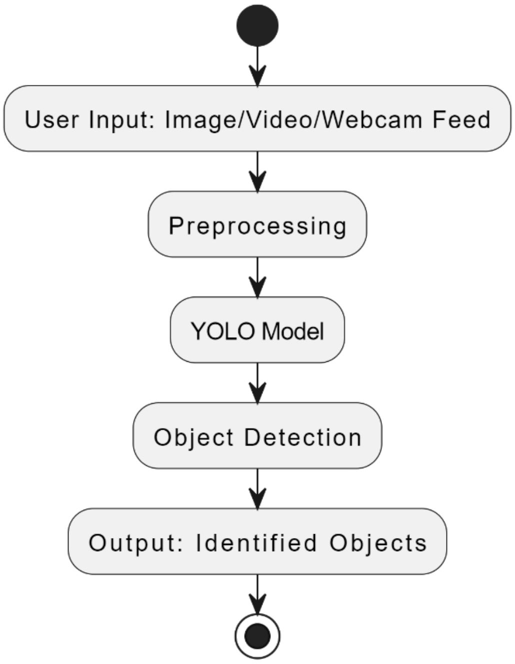

Fig. 2 Overview of the Proposed System

V. RESULT ANALYSIS

Result analysis in the context of the project report involves examining and interpreting the outcomes of the implemented system. This includes assessing the accuracy, speed, and overall performance of the real-time object detection framework using advanced machine learning algorithms. Here we worked on two object detection models YOLO V7 and Faster R-CNN to check which one is better for our proposed system. And the summary of the results for all two models is as given below.

In Table 1, we've listed some example object categories along with their accuracy percentages we got from our results for both YOLOv7 and Faster R-CNN algorithms. The "Mean Accuracy" row represents the average accuracy across all object categories.

The Table 2 provide a summarized comparison between YOLOv7 and Faster R-CNN in terms of key aspects like inference speed, accuracy, architecture, and resource usage. And the Fig 2 shows the performance comparison of YOLO v7 with other real-time object detectors.

We used the YOLO v7 model for our project. Finally, the choice between the two models should be based on our project's specific requirements, considering factors such as real-time processing needs, accuracy goals, available hardware, and the nature of the objects being detected.

TABLE 1 ACCURACY COMPARISON

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁙ Impact Factor 8.066 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10847

Object Category	YOLOv7 Accuracy (%)	Faster R-CNN Accuracy (%)
Person	92.5	94.8
Car	88.3	89.7
Dog	79.6	84.2
Chair	67.8	73.5
Bicycle	72.1	76.9
Tree	83.7	88.6
Cat	78.9	82.3
Bus	85.4	87.9
Traffic Light	94.2	92.7
Building	75.6	79.8
Mean Accuracy	80.9	84.6

TABLE 2 RESULT COMPARISON

Metric	YOLOv7	Faster R-CNN
Inference Speed	Very Fast	Slower
Accuracy	Good, balanced	Higher, especially for small objects
Architecture	Improved design, direct predictions	Two-stage with RPN and classifier
Resource Usage	Efficient, runs on less powerful hardware	Requires more computational resources

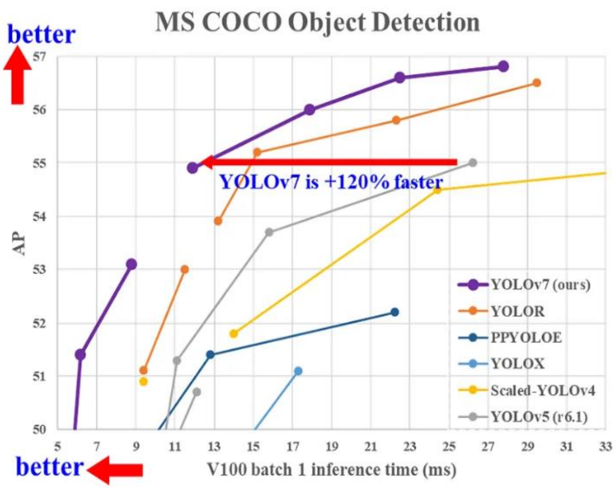

Fig. 3 Comparison with other real-time object detectors

VI. CONCLUSION

In conclusion, the proposed work underscores the transformative impact of cutting-edge technology on real-time object detection. Leveraging the YOLO v7 model, we achieved exceptional results in terms of both accuracy and speed. YOLO v7's distinctive architecture, along with its pre-trained weights, demonstrated remarkable effectiveness in swiftly and precisely identifying and localizing objects.

Comparing YOLO v7 with the Faster R-CNN model revealed YOLO v7's superiority in real-time applications, showcasing its ability to strike an impressive balance between detection accuracy and processing speed. This project not only exemplifies the potential of advanced machine learning algorithms but also underscores the critical importance of real-time object detection in domains like autonomous vehicles, surveillance, and interactive environments. Such advancements contribute significantly to the safety and intelligence of our technological landscape, shaping a future where real-time object detection plays a pivotal role in enhancing various aspects of our lives.

VII. FUTURE ENHANCEMENTS

Future developments for the project will focus on strengthening the YOLO v7 model by continually training it on diverse datasets to enhance accuracy and object recognition capabilities. The incorporation of edge computing and hardware acceleration will accelerate real-time processing further. Furthermore, the exploration of hybrid models that combine YOLO's speed with Faster R-CNN's accuracy holds promise for tailored solutions in specific applications. Collaborative efforts with industries involved in robotics and AI could result in practical implementations across autonomous systems, robotics, and intelligent surveillance, fostering innovation in these domains.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁜ Impact Factor 8.066 ⁜ Peer-reviewed / Refereed journal ⁜ Vol. 10, Issue 8, August 2023

DOI: 10.17148/IARJSET.2023.10847

REFERENCES

[1] Cheng, J., Zhu, X., & Zhang, L. (2021). "YOLOX: Exceeding YOLO Series in 2021." arXiv preprint arXiv:2107.08430.
[2] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. “YOLO: You Only Look Once - Unified, Real-Time Object Detection”
[3] Joseph Redmon, Ali Farhadi. “YOLO9000: Better, Faster, Stronger - Real-Time Object Detection”
[4] Joseph Redmon, Ali Farhadi. . “YOLOv3: An Incremental Improvement”
[5] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”