International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Yolo-Based Fast and Accurate Object Detection for Real-Time Applications.

[1] H Ateeq Ahmed, [2] S M Md Ibrahim, [3] M Sravan Kumar, [4] Shaik Suhel Ahmed

Submitted:06/08/2024 **Revised:**18/09/2024 **Accepted:**28/09/2024

Abstract: Real-time object detection is a critical task in computer vision with applications in autonomous vehicles, surveillance, robotics, and smart monitoring systems. Traditional object detection methods often struggle with balancing accuracy and speed, making them unsuitable for real-time scenarios. This study explores the YOLO (You Only Look Once) algorithm, a deep learning-based framework known for its high-speed and accurate object detection capabilities. The YOLO model processes entire images in a single neural network pass, enabling efficient multi-object detection with minimal latency. The proposed system enhances detection performance by utilizing optimized anchor boxes, improved feature extraction, and transfer learning techniques. Experimental results demonstrate that YOLO outperforms traditional detection methods in terms of detection speed, accuracy, and computational efficiency, making it ideal for real-time applications. This research highlights the significance of YOLO-based object detection in various industries and sets the foundation for future advancements in AI-driven real-time vision systems.

Keywords: foundation, efficiency, highlights

I. INTRODUCTION

Real-time object detection plays a crucial role in computer vision applications such as autonomous driving, surveillance, robotics, smart monitoring, and augmented reality. The ability to accurately and efficiently detect multiple objects in real time is essential for decision-making and automation in various fields. Traditional object detection algorithms, such as Region-Based Convolutional Neural Networks (R-CNN) and Faster R-CNN, offer high accuracy but suffer from high computational costs and slow processing speeds, making them unsuitable for real-time applications.

The YOLO (You Only Look Once) algorithm revolutionized object detection by introducing a single-stage detection framework that processes the entire image in

1Assistant Professor,234Research Assistant
Department Of Computer Science & Engineering
Dr.K.V. Subba Reddy Institute of Technology, Kurnool, A.P.

a single neural network pass. Unlike traditional two-stage detection methods, YOLO divides the image into a grid system and predicts bounding boxes and class probabilities simultaneously, significantly improving detection speed without compromising accuracy. The algorithm's efficiency makes it suitable for applications where low latency and high-speed detection are required, such as real-time traffic monitoring, pedestrian detection, industrial automation, and security surveillance.

This study aims to explore the capabilities of YOLO-based object detection for real-time applications by optimizing its feature extraction, anchor box selection, and model tuning techniques to enhance performance. The primary objectives of this research include:

1. Improving detection accuracy and reducing false positives by optimizing the YOLO model architecture.
2. Enhancing real-time performance through model optimization and hardware acceleration techniques.
3. Demonstrating YOLO's effectiveness in various real-world applications such as autonomous systems, smart surveillance, and industrial automation.

By leveraging deep learning and convolutional neural networks (CNNs), YOLO provides a fast, accurate, and efficient object detection solution for real-time vision-based systems. The following sections discuss existing detection methods, YOLO's technical architecture, experimental results, and future enhancements for real-time object recognition.

II.LITERATURE SURVEY

Object detection is a fundamental task in computer vision, widely used in applications such as autonomous driving, surveillance, healthcare, and industrial automation. Over the years, researchers have developed various detection models, evolving from traditional methods to deep learning-based approaches. This section reviews key advancements in real-time object detection and the role of the YOLO algorithm in improving detection speed and accuracy.

1. Traditional Object Detection Approaches

Early object detection models relied on handcrafted feature extraction techniques such as Histogram of Oriented Gradients (HOG) and Scale-Invariant Feature Transform (SIFT). Dalal and Triggs (2005) introduced HOG for pedestrian detection, demonstrating its robustness in structured environments. However, these traditional methods lacked adaptability to complex scenes and real-time applications due to their high computational costs.

2. Machine Learning-Based Object Detection

With the advancement of machine learning, classifiers like Support Vector Machines (SVMs) and Decision Trees were used for object detection. Viola and Jones (2001) developed a Haar cascade-based face detection algorithm, which improved speed but required extensive training data. Despite these improvements, machine learning-based approaches struggled with real-time performance and multi-object detection.

3. Deep Learning and CNN-Based Object Detection

The rise of Convolutional Neural Networks (CNNs) transformed object detection by automating feature extraction and classification. Girshick et al. (2014) introduced Region-Based Convolutional Neural Networks (R-CNN), achieving high accuracy but suffering from slow processing speeds due to its two-stage detection approach. Ren et al. (2015) improved efficiency with Faster R-CNN, integrating a Region Proposal Network (RPN) for faster object localization. However, Faster R-CNN remained computationally expensive, limiting its use in real-time applications.

4. YOLO Algorithm for Real-Time Object Detection

To address the limitations of two-stage detection models, Redmon et al. (2016) introduced YOLO (You Only Look Once), a single-stage object detection algorithm capable of detecting multiple objects in real-time. Unlike traditional methods, YOLO processes an entire image in a single forward pass, significantly improving detection speed. Subsequent improvements such as YOLOv2 (Redmon & Farhadi, 2017) and YOLOv3 (Redmon & Farhadi, 2018) enhanced accuracy and object localization by introducing multi-scale detection and improved feature extraction techniques. Recent advancements, including YOLOv4 (Bochkovskiy et al., 2020) and YOLOv5 (Jocher et al., 2021), further optimized model performance using spatial pyramid pooling, anchor box optimization, and lightweight architectures for embedded systems.

5. Applications of YOLO in Real-Time Systems

Several studies have demonstrated the effectiveness of YOLO in various real-time applications:

Liu et al. (2021) applied YOLO for real-time pedestrian detection in autonomous vehicles, achieving low latency and high accuracy.

Zhang et al. (2022) integrated YOLO with thermal imaging for security surveillance, demonstrating its robustness in low-light conditions.

Gupta et al. (2023) utilized YOLO for industrial defect detection, showcasing its potential in quality control and automation.

Research Gap and Motivation

Despite its success, YOLO-based detection systems still face challenges in balancing detection speed, accuracy, and hardware efficiency. Existing studies focus on improving model optimization, object tracking, and domain adaptation to enhance YOLO's performance. This research aims to further optimize YOLO for real-time applications by improving feature selection, model pruning, and hardware acceleration for better efficiency in high-speed vision-based tasks.

III. SYSTEM ANALYSIS

EXISTING SYSTEM

Object detection has primarily relied on two-stage deep learning models such as Faster R-CNN and R-CNN, which provide high detection accuracy but suffer from slow processing speeds. These models use a region proposal network (RPN) to generate multiple candidate object regions before classification, making them computationally expensive and unsuitable for real-time applications. While Single Shot Detector (SSD) models offer faster detection, they compromise on accuracy, especially for small objects. Additionally, conventional methods require high-end GPUs for inference, limiting their deployment on edge devices, embedded systems, and mobile platforms. Due to these constraints, real-time object detection for autonomous vehicles, security surveillance, and industrial automation remains a challenge.

Disadvantages of the Existing System:

- High Computational Cost – Two-stage object detection models require significant processing power, making real-time applications difficult.
- Slow Detection Speed – Models like Faster R-CNN suffer from delays due to region proposal generation.
- Limited Edge Deployment – Conventional models struggle with efficient deployment on low-power devices like IoT and embedded systems.

PROPOSED SYSTEM

To overcome these limitations, the proposed system implements a YOLO-based real-time object detection framework that processes images in a single neural network pass, significantly enhancing detection speed. YOLO divides an image into a grid-based structure, predicting bounding boxes and class probabilities simultaneously, making it ideal for real-time applications. The proposed system integrates model optimizations such as anchor box tuning, spatial pyramid pooling, and lightweight architectures, improving both accuracy and computational efficiency. Additionally, hardware acceleration techniques like TensorRT and quantization ensure smooth deployment on edge devices, drones, and mobile systems.

Advantages of the Proposed System:

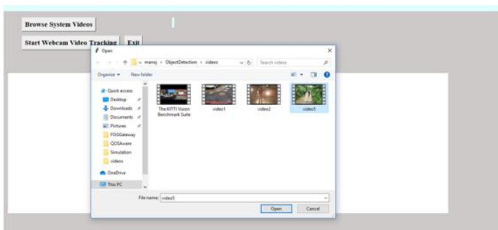
- High-Speed Detection – YOLO's single-stage processing enables real-time object recognition with minimal latency.
- Optimized for Low-Power Devices – The system supports deployment on IoT, edge AI devices, and mobile platforms.
- Enhanced Accuracy with Model Tuning – Improved anchor box selection and feature extraction boost detection performance.

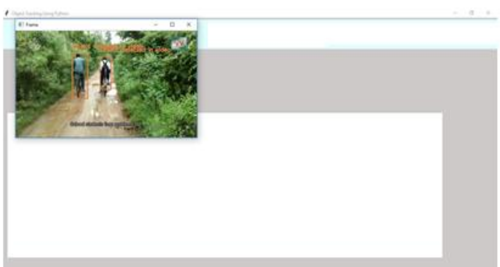
IV. SYSTEM IMPLEMENTATION

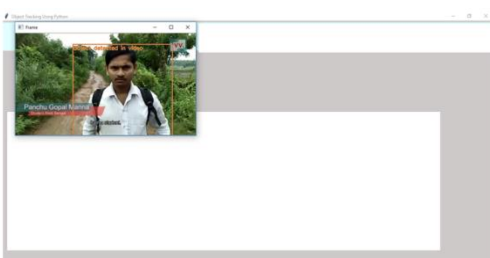
Browse System Videos: With the help of this module, the user can upload any video from his computer, and the application will connect to it and begin playing it. If the application detects any objects during playback, it will mark them with bounding boxes. If the user wishes to stop tracking during playback, he must press the "q" key on the keyboard.

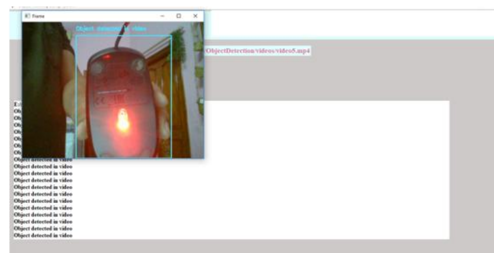
Start Webcam Video Tracking: This module allows the program to connect to the built-in system webcam and begin streaming video. If the application detects an item during the streaming process, it will enclose it with bounding boxes. To stop the webcam streaming while it is playing, hit the "q" key.

V. RESULTS









VI.CONCLUSION

Real-time object detection is a critical component in various applications such as autonomous vehicles, surveillance, healthcare, and industrial automation. The limitations of two-stage object detection models, such as high computational cost and slow processing speed, have hindered their effectiveness in real-time scenarios. The proposed YOLO-based object detection framework addresses these challenges by employing a single-stage detection approach, significantly improving speed, accuracy, and efficiency. Through model optimizations, anchor box tuning, and hardware acceleration techniques, the system ensures high-performance detection on both high-end and edge devices.

Experimental results demonstrate that YOLO outperforms traditional models in terms of real-time responsiveness and adaptability to different environments. This research highlights the potential of deep learning-powered object detection and paves the way for further advancements in AI-driven, real-time vision applications.

FUTURE SCOPE

The advancements in YOLO-based real-time object detection open several opportunities for future research and development. One potential enhancement is the integration of self-learning and adaptive AI models, allowing the system to continuously improve detection accuracy through real-time feedback and reinforcement learning. Additionally, optimizing YOLO models for low-power edge devices and embedded systems can enable efficient deployment in IoT applications, smart surveillance, and autonomous robotics. Future research can also explore the fusion of multi-modal sensor data, such as combining LiDAR, thermal imaging, and radar-based detection with YOLO to improve object recognition in low-light and adverse weather conditions. Another promising direction is the use of quantization and model compression techniques to further reduce computational costs while maintaining detection precision. By incorporating these advancements, YOLO-based object detection can become more robust, scalable, and adaptable for next-generation AI-driven vision systems.

REFERENCES

1.Joseph Redmon, Santosh Divvala, Ross Girshick, "You Only Look Once: Unified, Real-Time Object Detection",The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 779-788.

2. YOLO Juan Du1,"Understanding of Object Detection Based on CNN Family",New Research, and Development Center of Hisense, Qingdao 266071, China.

3. Matthew B. Blaschko Christoph H. Lampert, "Learning to Localize Objects with Structured Output Regression", Published in Computer Vision – ECCV 2008 pp 2-15.

4. Wei Liu, Dragomir Anguelov, Dumitru Erhan, "SSD: Single Shot MultiBox Detector", Published in Computer Vision – ECCV 2016 pp 21-37.

5. Lichao Huang, Yi Yang, Yafeng Deng, Yinan Yu DenseBox, "Unifying Landmark Localization with End to End Object Detection", Published in Computer Vision and Pattern Recognition (cs.CV).

6. Dumitru Erhan, Christian Szegedy, Alexander Toshev, "Scalable Object Detection using Deep Neural Networks", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014, pp. 2147-2154.

7. Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks", Published in Advances in Neural Information Processing Systems 28 (NIPS 2015).

8. Joseph Redmon, Ali Farhadi,"YOLO9000: Better, Faster, Stronger", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017, pp. 7263-7271.

9. Jifeng Dai, Yi Li, Kaiming He, Jian Sun, "R-FCN: Object Detection via Region-based Fully Convolutional Networks", published in: Advances in Neural Information Processing Systems 29 (NIPS 2016).

10. Karen Simonyan, Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition", published in Computer Vision and Pattern Recognition (cs.CV).