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 4, June 2025
Impact Factor: 7.67

Real Time Object Detection using OpenCV and Yolo-4

Rane Deepraj Santosh, Patel Sana Shakil, Pardeshi Shreya Dhiraj, Prof. Khemnar Kavita
Department of Artificial Intelligence & Machine Learning
Sahyadri Valley College of Engineering & Technology, Rajuri, Pune.

Abstract: *Real-time object detection is a critical component in various applications such as surveillance, autonomous vehicles, and augmented reality. This paper presents an implementation of real-time object detection using OpenCV and the YOLO (You Only Look Once) framework. YOLO is a stateof-the-art, single-stage detection algorithm known for its speed and accuracy, making it suitable for real-time applications. We discuss the architecture of YOLO, its training process, and the integration of OpenCV for video processing and visualization. Our implementation demonstrates the capability of YOLO to detect multiple objects in real time, achieving high frame rates while maintaining precision. Performance metrics such as mean Average Precision (mAP) and inference time are evaluated, and results are compared against other object detection methods. The findings indicate that our approach effectively balances speed and accuracy, providing a robust solution for real-time object detection in diverse environments. This research contributes to advancing the field of computer vision and enhances the practical applications of object detection technologies.*

Keywords: Real-time object detection- OpenCV- YOLO(You Only Look Once)- Computer vision-Video processing- Autonomous systems- Object classification and localization

I. INTRODUCTION

Object detection is a critical area in computer vision that enables machines to identify and locate objects within images or video streams. With the rapid advancements in artificial intelligence and deep learning, the demand for efficient and accurate object detection systems has surged across various industries, including security, automotive, robotics, and augmented reality. TheYou Only Look Once (YOLO) algorithm represents a significant breakthrough in real-time object detection. Unlike traditional methods that apply classification and localization as separate processes, YOLO treats object detection as a single regression problem. This approach allows it to predict bounding boxes and class probabilities simultaneously, resulting in impressive speed and accuracy. Coupled with OpenCV, a powerful open-source computer vision library, developers can create robust applications that leverage YOLO's capabilities for real time processing. This project aims to develop a real-time object detection system using OpenCV and YOLO, addressing common challenges such as latency and performance in dynamic environments. By providing immediate feedback and accurate results, this system can facilitate a wide range of applications—from surveillance and traffic monitoring to interactive experiences in augmented reality. Through this work, we seek to demonstrate the effectiveness of modern object detection techniques and their potential to transform how machines perceive and interact with the world. Real-time object detection is a pivotal technology in various domains, including autonomous vehicles, security, and assistive technologies. YOLO (You Only Look Once) and its improved version YOLOv3 are leading algorithms in this field. YOLO is renowned for its speed, while YOLOv3 enhances accuracy and robustness. Integrating these models with audio feedback can significantly enhance usability, particularly for visually impaired users or in applications where visual monitoring is impractical. This synopsis outlines the hardware and software requirements and detailed methodology for comparing YOLO andYOLOv3 in real-time object detection with audio feedback.

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 4, June 2025

Impact Factor: 7.67

II. LITERATURE SURVEY

Object detection, a vital task in computer vision, has garnered significant interest due to its diverse applications in fields such as robotics, surveillance, and autonomous systems. Researchers have explored a variety of methods and techniques to advance this area. Historically, object detection relied on machine learning algorithms and handcrafted features. Techniques such as Haar cascades and Histogram of Oriented Gradients (HOG) were commonly employed; however, these methods struggled with complex and varied object appearances. The advent of deep learning has led to a significant paradigm shift in the field: Convolutional Neural Networks (CNNs) have revolutionized object detection since their introduction. Key architectures such as VGG, ResNet, and Inception serve as the foundation for CNNs. Modern real-time object detection techniques, including You Only Look Once (YOLO) and Single Shot Detectors (SSD), have emerged as state-of-the art solutions.

III. METHODOLOGY

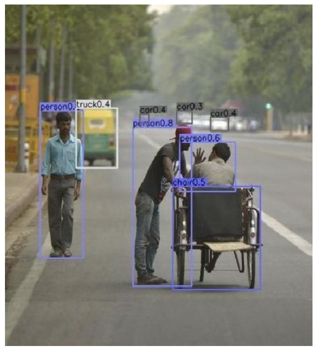
1. Data Collection: Compile a dataset of images that feature the specific objects of interest. These images can be sourced from various platforms, and it is essential that each image is accompanied by annotations or labels that detail the location and type of objects present.
2. Data Preparation: Clean and preprocess the collected images. This may involve resizing the images to a consistent dimension, normalizing pixel values, and applying any necessary transformations to enhance the dataset.
3. Model Development: Implement an object detection model, typically utilizing a single-stage detector such as YOLO (You Only Look Once) or SSD (Single Shot MultiBox Detector), or a region based Convolutional Neural Network (R-CNN). These models are specifically designed for detecting and localizing objects within images.The model acquires the capability to identify objects and predict their bounding box coordinates along with the associated class labels.

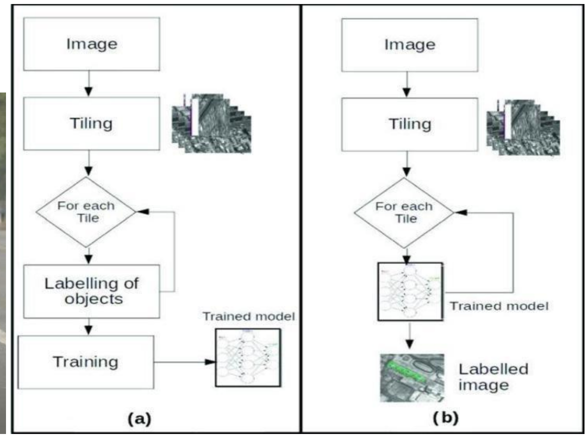
IV. SYSTEM IMPLEMENTATION

Implementing a system based on the YOU ONLY LOOKONCE(YOLO)algorithm for object detection in images involves several key processes. Below is a high-level overview of the steps involved:

1. Compile a comprehensive dataset of images containing potential objects for detection. Each image should be annotated with the locations (bounding boxes) and class labels for the objects present.
2. Resize all images in the dataset to a uniform size that is compatible with YOLO. Typically, square images yield optimal results with this algorithm. Additionally, standardize pixel values and eliminate any noise or outliers from the dataset.
3. Incorporate bounding boxes for each detected object in the images as part of the dataset annotations. Each bounding box should include the object's class label and the corresponding coordinates of the box.
4. Configure the appropriate architecture and hyperparameters for the YOLO model. Choose a specific version of YOLO that aligns with your application requirements (e.g., YOLOv3, YOLOv4). Use the pre- processed and annotated dataset to train the YOLO model, enabling it to detect, classify, and predict the bounding boxes of objects within images.
5. Assess the performance of the YOLO model using a distinct set of images that were not part of the training process. Key evaluation metrics to consider include mean Average Precision (mAP), precision, recall, and intersection over union (IoU). Evaluate the model's accuracy, speed, and overall robustness
6. Develop a software application or system that integrates the trained YOLO model, allowing it to identify the presence and location of objects in new input images.
7. Deploy the YOLO-based object detection system in the intended application environment. Ensure consistent object detection performance by regularly monitoring the system's efficiency and accuracy, making necessary adjustments or updates as required.

DOI: 10.48175/568





V. FUTURE SCOPE

The future of object detection using YOLO is promising, driven by continuous advancements in computer vision and deep learning technologies. It is expected that YOLO will evolve significantly in several key areas. Primarily, there will be an emphasis on enhancing both speed and accuracy, which will make YOLOincreasingly suitable for high precision tasks and real-time applications. With the growing accessibility of customization and transfer learning, organizations will have the opportunity to tailor YOLO to their specific requirements, thereby expanding its use across various sectors.

Additionally, a significant future trend is the integration of semantic segmentation, which will allow for pixel-level object identification alongside traditional object detection. This advancement will facilitate the development of more sophisticated applications in robotics, autonomous vehicles, and medical imaging. Furthermore, the optimization of edge computing is on the horizon, which will support these innovations.

VI. CONCLUSION

To summarize, object detection systems utilizing the YOLO(YouOnly Look Once) framework represent a state-of-the-art and versatile approach within the realm of computer vision. YOLO has achieved significant popularity across various sectors, including autonomous driving, security, healthcare, and beyond, due to its ability to simultaneously detect and classify objects in real-time. Despite its numerous advantages, it is essential to acknowledge certain limitations, such as difficulties in identifying small objects and handling complex scenes. However, ongoing research and development are driving the continuous improvement of YOLO to address these challenges and enhance both accuracy and processing speed. The system's functionality is further enriched through customization, the incorporation of semantic segmentation, and advancements in edge computing.

REFERENCES

[1]. Bruckner, Daniel. Ml-o-scope: a diagnostic visualization system for deep machine learning pipelines.

[2]. No.UCB/EECS-2014 99.CALIFORNIA UNIV BERKELEY DEPTOFELECTRICAL ENGINEERING ANDCOMPUTER SCIENCES, 2014.

[3]. KSaleh, Imad, Mehdi Ammi, and Samuel Szoniecky, eds. Challenges of the Internet of Things: Technique, Use, Ethics. John Wiley & Sons, 2018.

[4]. Petrov, Yordan. Improving object detection by exploiting semantic relations between objects.MS thesis.UniversitatPolitècnica de Catalunya, 2017.

[5]. Nikouei, SeyedYahya, et al. "Intelligent Surveillance as an Edge Network Service: from HarrCascade, SVMtoaLightweight CNN." arXiv preprint arXiv:1805.00331 (2018)