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 12, April 2025
Impact Factor: 7.67

Real Time Object Detection using Artificial Intelligence

Dr. Monika Rokade[1], Dr. Sunil S. Khatal[2], Nikita Avhad[3], Suyog Kute[4], Dhanashri Wadekar[5]
[1,2,3,4,5] Department of Computer Engineering
Sharadchandra Pawar College of Engineering, Otur (Dumbarwadi), Junnar, Pune

Abstract: *A wide, active, and challenging field of computer vision is real-time object detection. Image localization refers to the process of finding a single object in an image, while object detection refers to the process of finding several objects in an image. This recognizes a class of semantic items in digital photos and movies. Real-time object detection has a variety of uses, including object tracking, video surveillance, people counting, pedestrian detection, self-driving automobiles, facial recognition, sports ball tracking, and more. When detecting objects with the aid of OpenCV a library of programming functions primarily geared toward real-time computer vision, Convolution Neural Networks is a representative technique of deep learning.*

Keywords: Object detection, OpenCV, CNN, Realtime processing

I. INTRODUCTION

Detecting objects remains one of computer vision and image understanding applications' most fundamental and challenging aspects. Significant advances in object detection have been achieved through improved object representation and the use of deep neural network models. This paper examines more closely how object detection has evolved in the era of deep learning over the past years. We present a literature review on various state-of-the-art object detection algorithms and the underlying concepts behind these methods. We classify these methods into three main groups: anchor-based, anchor-free, and transformer based detectors. Those approaches are distinct in the way they identify objects in the image. We discuss the insights behind these algorithms and experimental analyses to compare quality metrics, speed/accuracy tradeoffs, and training methodologies. The survey compares the major convolutional neural networks for object detection. It also covers the strengths and limitations of each object detector model and draws significant conclusions. We provide simple graphical illustrations summarizing the development of object detection methods under deep learning. Finally, we identify where future research will be conducted. Object detection is a fundamental problem in the field of computer vision, with applications spanning a wide range of industries, from autonomous vehicles and surveillance systems to healthcare and robotics. The objective of object detection is to identify and classify multiple objects within an image or video frame, while simultaneously determining their locations (typically in the form of bounding boxes). The advancement of artificial intelligence (AI), particularly deep learning techniques, has significantly enhanced the accuracy, speed, and robustness of object detection systems. real-time object detection refers to the ability to detect and classify objects in video streams or live camera feeds as they are captured, with minimal delay. Achieving real-time performance is essential for applications where quick decision-making is crucial, such as in autonomous driving, security monitoring, and robotic navigation. This project aims to develop a real-time object detection system using deep learning techniques. The approach focuses on leveraging state-of-the-art AI models, such as YOLO (You Only Look Once), SSD (Single Shot Multibox Detector), or Faster RCNN, to identify and classify objects in real-time. These models are trained on large annotated datasets and optimized for high-performance inference on video data.

OBJECTIVE

- The objective of this project is To identify and locate one or more effective targets from still image.
- To Save Time And Cost .

- To Identify and count Multiple Object in one Image Easily.etc
- To achieve a high degree of detection accuracy while maintaining efficiency in terms of computational resources, making the system scalable to devices with limited processing power such as edge devices and mobile platforms.

II. LITERATURE SURVEY

Paper 1: Real-time Object Detection Using Deep Learning (2023) Authors: K. Vaishnavi a , G. Pranay Reddy

Abstract: As technology improved, object detection, which is connected to video and image analysis, caught researchers' interest. Earlier object recognition techniques are based on hand-crafted features and imprecise architectures and trainable algorithms. One of the main issues with many object detection systems is that they rely on other computer vision methods to support their deep learning-based methodology, which leads to slow and subpar performance. In this article, we present an end-to-end solution to the object detection problem using a deep learning based method. The single shot detector (SSD) technique is the quickest method for object detection from an image using a single layer of a convolution network. Our research's primary goal is to enhance accuracy of SSD method.

Paper 2: A Comprehensive Survey on Object Detection Using Deep Learning (2023) Authors: Bhagyashri More, Snehal Bhosale

Abstract: One of the common and difficult issues in com puter vision is to detect the object. Researchers have widely experimented and contributed to the performance improvement of object detection and associated tasks including object classification, localization, and segmentation over the way of the last decade of deep learning's rapid evolution. Object detectors can be broadly categorized into two groups: two stage and single stage object detectors. Two stage detectors primarily focus on selected region proposals via sophisticated architecture whereas single stage detectors concentrate on all feasible spatial region proposals for object detection via relatively easier architecture in one go

Paper 3: Recent Object Detection Techniques: A Survey (2022) Authors: Diwakar, Deepa Raj

Abstract: In the field of computer vision, object detection is the fundamental most widely used and challenging problem. Last several decades, great effort has been made by computer scientists or researchers to handle the object detection problem. Object detection is basically, used for detecting the object from image/video. At the beginning of the 21st century, a lot of work has been done in this field such as HOG, SIFT, SURF etc. are performing well but can't be efficiently used for Real-time detection with speed and accuracy. Furthermore, in the deep learning era Convolution Neural Network made a rapid change and leads to a new pathway and a lot of excelSPCOE, Department of Computer Engineering 2024-25 7 lent work has been done till dated such as region-based convolution network YOLO, SSD, retina NET etc.

Paper 4: Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments (2021)

Authors: Muhammad Ahmed 1,2,† , Khurram Azeem Hashmi

Abstract: Recent progress in deep learning has led to accurate and efficient generic object detection networks. Training of highly reliable models depends on large datasets with

highly textured and rich images. However, in real-world scenarios, the performance of the generic object detection system decreases when (i) occlusions hide the objects, (ii) objects are present in low-light images, or (iii) they are merged with background information. In this paper, we refer to all these situations as challenging environments. With the recent rapid development in generic object detection algorithms, notable progress has been observed in the field of deep learning-based object detection in challenging environments. However, there is no consolidated reference to cover the state of the art in this domain. To the best of our knowledge, this paper presents the first comprehensive overview, covering

Paper 5: A Survey Paper On Object Detection And Localization Methods In Image Processing (2024) Authors: Katroth Balakrishna Maruthiram,2 Dr.G.Venkata Rami Reddy

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 12, April 2025
Impact Factor: 7.67

Abstract: over the past few decades, advances in information technology have significantly transformed how we manage data and information, particularly in terms of data acquisition, processing, and predictions. This interdisciplinary effort has involved the use of image processing techniques and the implementation of AI-based models. The latest technological innovations have made it possible for researchers to conduct computational experiments that would have been impractical using conventional methods.

III. WORKING OF SYSTEMS

The system contains following modules:
Registration User can register here.

- Login: If registration is done then user can login with username and password. Training : 80percent training model, with YOLOv3 algorithm for object .
- Testing : 20percent is testing part.Testing will be done with by giving input as object and predict the output type of object.
- Algorithm : For training and classification YOLOv3 algorithm is used. Output: Detect the object.

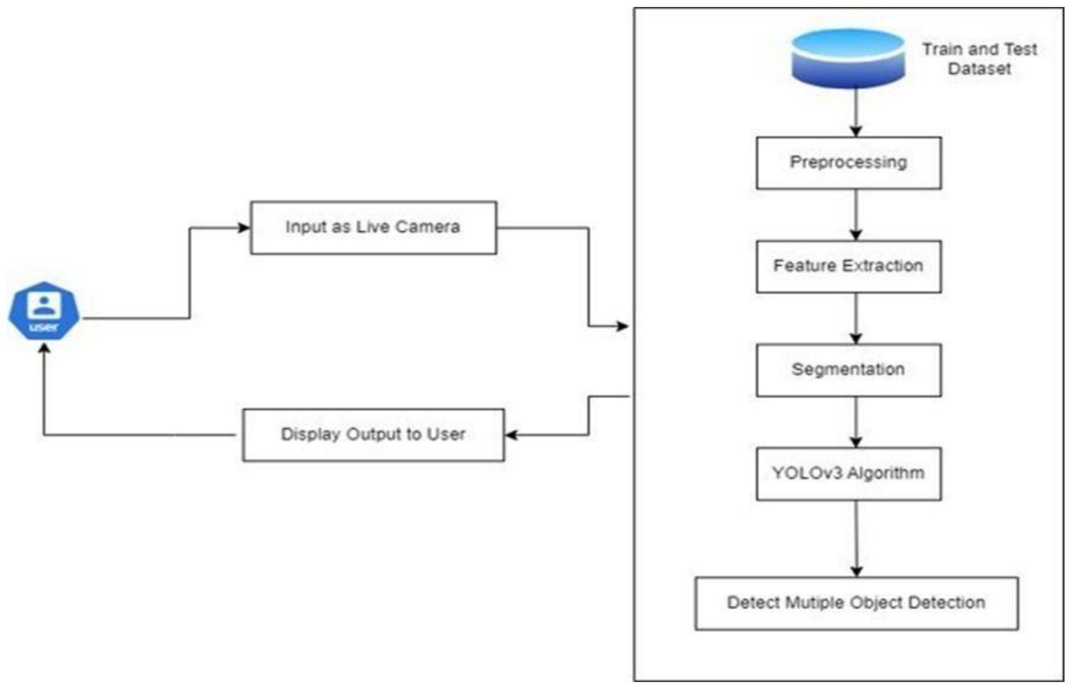

Fig.1 System Architecture

Real-time object detection using Artificial Intelligence (AI) leverages deep learning techniques to identify and classify objects in images or video streams instantly. These systems use various architectures and models to extract features from input data, and each approach varies in how it balances accuracy and speed to meet real-time constraints.

Convolutional Neural Networks (CNNs) for Feature Extraction

At the core of most real-time object detection systems is a Convolutional Neural Network (CNN). CNNs are designed to automatically learn spatial hierarchies of features from input images through convolutional layers. These layers identify low-level features such as edges and textures, progressing to more complex patterns like shapes and objects as the data passes through deeper layers. For object detection, CNNs act as the feature extractor, transforming raw image pixels into high-level, spatially aware representations that are critical for detecting objects of various sizes and shapes.

Single-Stage vs. Two-Stage Detectors

Object detection models can be broadly classified into two categories: single-stage and two-stage detectors. Two-stage detectors, like Faster R-CNN, first generate region proposals and then classify these regions to detect objects. The first

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 12, April 2025 Impact Factor: 7.67

stage identifies regions in the image where objects are likely to be located, and the second stage refines these proposals and classifies them. While two-stage detectors tend to be more accurate, they are slower due to their multi-step processes. In contrast, single-stage detectors like YOLO and SSD simultaneously predict bounding boxes and class labels in one pass through the network, making them faster and more suitable for real-time applications. The system thus balances accuracy and speed, depending on the specific use case.

Anchor Boxes and Proposal Generation

A key feature in most object detection systems is the use of anchor boxes. Anchor boxes are predefined bounding box shapes and sizes that help detect objects of various dimensions. These anchor boxes are generated by considering different aspect ratios and scales that might be present in the input image. In two-stage detectors like Faster R-CNN, a Region Proposal Network (RPN) is used to generate potential object proposals. In single-stage models like YOLO, the network directly predicts offsets from these anchor boxes to generate bounding boxes around detected objects. The use of anchor boxes accelerates the detection process by reducing the search space and simplifying the task of identifying object boundaries.

Non-Maximum Suppression (NMS) for Refining Predictions

After detecting potential objects, object detection systems often face the issue of multiple overlapping predictions for the same object. Non-Maximum Suppression (NMS) is a technique used to eliminate redundant bounding boxes by keeping only the one with the highest confidence score. This is crucial for ensuring that the system does not produce multiple detections for a single object, which would lead to false positives. NMS works by sorting the bounding boxes according to their confidence scores and progressively eliminating boxes that overlap with others above a certain threshold. This step is essential in streamlining object detection results and ensuring the accuracy of the system in realtime applications.

Integration of Transformer Models for Enhanced Detection

The emergence of transformer models, such as DETR (Detection Transformer), has introduced a new paradigm in object detection systems. Unlike traditional CNN-based methods, transformer models are designed to capture longrange dependencies between different parts of the image. They treat the detection task as a direct set prediction problem, where the transformer processes the entire image in parallel and produces object predictions. By applying selfattention mechanisms, transformers can better understand contextual relationships and improve the accuracy of detecting objects in complex scenes, especially in cases with occlusions or clutter. The integration of transformer-based models into realtime systems offers significant improvements, although these models are typically slower and require more computational resources than their CNN-based counterparts.

Real-Time Processing and Edge Computing

For real-time object detection, systems must be optimized to process images or video streams with minimal latency. Real-time detection often involves deploying models on specialized hardware, such as Graphics Processing Units (GPUs), Field Programmable Gate Arrays (FPGAs), or dedicated AI accelerators. These hardware platforms enable faster computation by parallelizing operations. Moreover, edge computing has become a critical aspect of real -time object detection, especially for mobile devices, autonomous vehicles, and drones. Edge computing allows data processing to occur locally on the device, reducing the reliance on cloud services and minimizing communication delays. This is particularly important in applications that require immediate decision-making, such as autonomous driving, where fast and accurate object detection is crucial for navigation and safety.

Real-time object detection systems powered by AI combine various techniques, including CNNs, anchor boxes, NMS, and transformers, to detect and classify objects quickly and accurately. The use of efficient hardware and edge computing further enables these systems to function in time-sensitive applications, offering a balance between computational efficiency and detection accuracy. These advances have revolutionized real-time object detection, opening up new possibilities for industries like healthcare, surveillance, autonomous driving, and robotics.

IV. ADVANTAGES

- High Accuracy: Provides precise object detection in complex environments, even with occlusions or varying lighting.

- Speed and Efficiency: Enables real-time detection with fast processing speeds, critical for time-sensitive tasks.
- Scalability: Easily adaptable to detect a wide range of objects across diverse applications.
- Real-time Processing: Utilizes edge computing for local data processing, reducing latency and reliance on cloud servers.
- Reduced Human Error: Minimizes manual intervention, ensuring more consistent and accurate results.

V. DISADVANTAGES

- High Computational Requirements: Requires significant computational power, especially for real-time processing.
- Data Dependency: Performance heavily depends on the quality and quantity of labeled data for training.
- Limited Generalization: Models may struggle with detecting objects outside their training dataset or in unfamiliar conditions.
- Energy Consumption: Real-time object detection systems can consume substantial energy, especially in mobile or edge devices.

VI. FUTURE SCOPE

The future scope of real-time object detection using artificial intelligence lies in enhancing the speed and accuracy of detection models through efficient algorithms and hardware optimization. Researchers are focusing on improving performance in complex environments with occlusions and varying lighting conditions. The integration of multi-modal data, such as combining visual, textual, and audio information, promises to create more robust systems. Additionally, advancements in tiny object detection, 3D object detection, and the use of few-shot learning for training models with limited data are expected to expand the potential applications in fields like autonomous driving, robotics, medical imaging, and augmented reality.

VII. CONCLUSION

In this application, we are detecting object using a live camera. Yolo v3 algorithm is used for detection of object. We can also keep a count of object after the detection. This work proposes an approach towards generic object counting with unsupervised local image information. Moreover, we propose to learn from local image features, and predict global image object counts. this project has successfully demonstrated the implementation of real-time object detection using artificial intelligence (AI) techniques, particularly deep learning algorithms. By leveraging powerful models such as Convolutional Neural Networks (CNNs) and utilizing advanced frameworks like TensorFlow and PyTorch, the system has been trained to identify and classify objects in real-time from live video feeds with high accuracy.

REFERENCES

[1]. K. Vaishnavi a, G. Pranay Reddy , ”Real-time Object Detection Using Deep Learning”, Volume 38, Issue 8, Page 24-32, 2023; Article no.JAMCS.101284 ISSN: 2456-9968

[2]. Bhagyashri More1, Snehal Bhosale, ”A Comprehensive Survey on Object Detection Using Deep Learning”, https://doi.org/10.18280/ria.370217

[3]. 3 Diwakar, Deepa Raj , ”Recent Object Detection Techniques: A Survey”, DOI: 10.5815/ijigsp.2022.02.05

[4]. Muhammad Ahmed 1,2,† , Khurram Azeem Hashmi, ”Survey and Performance Analysis of Deep Learning Based Object Detection in Challenging Environments”, Sensors 2021, 21, 5116. https://doi.org/10.3390/s21155116

[5]. Katroth Balakrishna Maruthiram,2 Dr.G.Venkata Rami Reddy, ”A Survey Paper On Object Detection And Localization Methods In Image Processing”, © 2024 IJCRT — Volume 12, Issue 6 June 2024 — ISSN: 2320-2882

[6]. S. Agarwal and D. Roth, “Learning a sparse representation for object detection,” in Computer Vision—ECCV 2002, A. Heyden, G. Sparr, M. Nielsen, and P. Johansen, Eds. Berlin, Germany: Springer, 2002, pp. 113–127, doi: 10.1007/3-540-47979-18.

[7]. X. Li, W. Wang, X. Hu, J. Li, J. Tang, and J. Yang, "Generalized focal loss v2: Learning reliable localization quality estimation for dense object detection," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2021, pp. 11627–11636, doi: 0.1109/CVPR46437.2021.01146.

[8]. X. Wang, A. Shrivastava, and A. Gupta, "A-fast-RCNN: Hard positive generation via adversary for object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 3039–3048, doi: 10.1109/CVPR.2017.324.

[9]. X. Li, W. Wang, L. Wu, S. Chen, X. Hu, J. Li, J. Tang, and J. Yang, "Generalized focal loss: Learning qualified and distributed bounding boxes for dense object detection," in Proc. 34th Int. Conf. Neural Inf. Process. Syst. Red Hook, NY, USA: Curran Associates, 2020, pp. 21002–21012.

[10]. K. Aurangzeb, S. Aslam, M. Alhussein, R. A. Naqvi, M. Arsalan, and S. I. Haider, "Contrast enhancement of fundus images by employing modified PSO for improving the performance of deep learning models," IEEE Access, vol. 9, pp. 47930–47945, 2021, doi: 10.1109/ACCESS.2021.3068477.

[11]. Zhang, S., Wu, Y., & Zuo, W. (2017). Object Detection and Segmentation using Dense Region Proposal Networks. IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[12]. Zhang, H., Liu, X., & Sclaroff, S. (2016). Detecting Objects in RGB-D Indoor Scenes. IEEE International Conference on Computer Vision (ICCV).

[13]. Gupta, A., & Dollar, P. (2016). Visualizing and Understanding Convolutional Networks. IEEE International Conference on Computer Vision (ICCV).

[14]. Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. International Conference on Machine Learning (ICML).

[15]. Wang, H., & Zhang, Y. (2017). A Comprehensive Review of Real-Time Object Detection Techniques. Journal of Artificial Intelligence Research.

[16]. Song, Z., & Wang, L. (2018). Real-Time Object Detection for Autonomous Driving: A Survey. IEEE Transactions on Intelligent Transportation Systems.

[17]. Yu, Z., & Yuan, Y. (2017). A Survey of Object Detection and Tracking in Video Surveillance Systems. Journal of Computer Vision.

[18]. Chen, D., & Wei, Y. (2017). Real-Time Object Detection and Tracking Using Convolutional Neural Networks. International Journal of Computer Applications.

[19]. Zhang, X., & Zhong, W. (2020). Object Detection in Real-Time Applications with Deep Learning. Neural Processing Letters.

[20]. Liu, F., & Xu, X. (2020). Object Detection Using Deep Convolutional Networks in Complex Environments. IEEE Access.

[21]. Choi, W., & Lee, Y. (2020). A Survey of Real-Time Object Detection Methods for Autonomous Vehicles. IEEE Transactions on Vehicular Technology.

[22]. Zhang, Y., & Tao, H. (2019). Real-Time Object Detection with Multilayer Convolutional Neural Networks. Journal of Artificial Intelligence and Data Mining.

[23]. Hinton, G., & Salakhutdinov, R. (2006). Reducing the Dimensionality of Data with Neural Networks. Science.

[24]. Zoph, B., & Le, Q. V. (2018). Neural Architecture Search with Reinforcement Learning. International Conference on Learning Representations (ICLR).

[25]. He, K., & Sun, J. (2015). Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[26]. Wei, L., & Zhou, H. (2021). An Efficient Real-Time Object Detection System for Video Surveillance Applications. IEEE Transactions on Cybernetics.

[27]. Zhang, L., & Li, X. (2021). Towards Real-Time Object Detection with Small Scale Networks. International Journal of Computer Vision.

[28]. Zeng, Z., & Xu, Y. (2018). Deep Learning-Based Object Detection for Industrial Automation. Automation Science and Engineering.

[29]. Ren, Z., & Wang, S. (2017). Real-Time Object Detection with YOLOv3 and Deep Learning Techniques.

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 12, April 2025
Impact Factor: 7.67

Proceedings of the International Conference on Computer Vision and Image Processing.
[30]. Li, Y., & Zhang, F. (2019). Fast and Efficient Object Detection Using Lightweight Neural Networks. Journal of Machine Learning and Computer Vision