World Journal of Advanced Research and Reviews

eISSN: 2581-9615 CODEN (USA): WJARAI

Cross Ref DOI: 10.30574/wjarr

Journal homepage: https://wjarr.com/

(Research Article)

Check for updates

Exploring the landscape: Applications and comparative analysis of object detection algorithms

Harshita Mathur * and Vaibhav Gupta

Faculty of Computer Science, Lachoo Memorial College of Science and Technology (Autonomous), Jodhpur, Rajasthan, India.

World Journal of Advanced Research and Reviews, 2024, 24(02), 2157–2163

Publication history: Received on 14 October 2024; revised on 21 November 2024; accepted on 23 November 2024

Article DOI: https://doi.org/10.30574/wjarr.2024.24.2.3584

Abstract

The review provides an in-depth overview of object detection algorithms in computer vision, addressing a broad spectrum of methods and techniques. It thoroughly examines key algorithms such as HOG, R-CNN, SSD, and YOLO, highlighting their respective strengths and limitations. By assessing the performance of each algorithm in various scenarios, the review offers insights into their practical applications across diverse fields, such as surveillance, healthcare, and autonomous systems. Additionally, it discusses current challenges faced in object detection, such as balancing speed and accuracy, and outlines potential research directions aimed at enhancing robustness, efficiency, and adaptability in future applications.

Keywords: Computer vision; Deep learning; DPM; HOG; Object detection; R-CNN; SSD; YOLO.

1. Introduction

Object detection, a crucial task in computer vision, involves identifying and localizing objects in images. The field has evolved substantially, from early techniques like HOG and DPM to modern deep learning models like R-CNN, SSD, and YOLO. This technology has wide-ranging applications in various fields like medical imaging, retail, surveillance, and augmented reality, driving improvements in fields like diagnostic accuracy, inventory management, security, and immersive experiences. This paper provides a study of the development of object detection algorithms, discussing their strengths, limitations, and future directions.

2. Overview of Object Detection Algorithm

Object detection is a key task in computer vision, focused on identifying and locating objects within an image. Over time, numerous algorithms have been developed, and categorized into traditional methods, deep learning-based approaches, and recent advancements.

2.1. Traditional Approaches

2.1.1. Histogram of Oriented Gradients (HOG)

As one of the earliest object detection techniques, HOG captures gradient orientations in localized regions of an image using feature descriptors. By encoding the shape and structure of objects, HOG enables detection based on geometric properties. However, it struggles in complex backgrounds or when objects appear in different poses.

*Corresponding author: Harshita Mathur

2.1.2. Deformable Part Models (DPM)

DPM represents objects as a set of parts, each with its own spatial arrangement and appearance model. By allowing parts to deform while preserving spatial relationships, DPM can handle variations in object orientation and articulation, making it robust to pose changes. However, it is computationally intensive and has been largely replaced by deep learning methods.

2.2. Deep Learning-Based Approaches

2.2.1. Region-Based Convolutional Neural Networks (R-CNN)

R-CNN introduced the concept of region proposals, identifying potential object regions first and then classifying them using a CNN. This method improved accuracy but was slow due to the separate steps of region proposal and classification.

Fast R-CNN and Faster R-CNN

Fast R-CNN improved upon R-CNN by utilizing region of interest (RoI) pooling, which allowed the CNN to process the entire image in one pass. Faster R-CNN further enhanced this by introducing a Region Proposal Network (RPN), which shares computation with the detection network, making object detection faster and more efficient.

Figure 1 Faster R-CNN

2.2.2. Single Shot MultiBox Detector (SSD)

Figure 2 Single Shot MultiBox Detector (SSD)

SSD provided a unified framework that eliminated the need for region proposals, predicting bounding boxes and class scores directly from feature maps at multiple scales. It allows for the detection of objects of varying sizes and aspect ratios in a single pass, resulting in faster detection.

2.2.3. You Only Look Once (YOLO)

YOLO revolutionized object detection by treating it as a single regression problem and predicting both bounding boxes and class probabilities simultaneously. Its end-to-end approach enables the process of real-time object detection with high accuracy, reducing computational complexity and speeding up the process.

YOLOv4 version of YOLO introduced improvements in backbone networks, data augmentation, and loss functions, achieving higher accuracy and speed. These are highly effective for real-time detection across various applications.

Figure 3 You Only Look Once (YOLOv4**)**

YOLOv4 includes four primary components:

- Backbone**:** Increases accuracy by using interconnected layers, deepening the network, and expanding the receptive field.
- Neck: Adds layers between the backbone and head, combining spatial and semantic information to improve the prediction head's inputs.
- Head (Dense Prediction): Predicts bounding box coordinates and class confidence scores, dividing the image into a grid and using anchor boxes to detect objects.
- Sparse Prediction: This is used in two-stage detection algorithms like Faster R-CNN, but not in YOLOv4.

3. Comparative Analysis

This section demonstrates a comparative evaluation of YOLO, SSD, and Faster R-CNN based on performance metrics such as mean Average Precision (mAP), inference time, frames per second (FPS), and computational resource usage.

3.1. Mean Average Precision (mAP)

Table 1 Performance Metrics: Mean Average Precision (mAP)

SN.	Algorithm	mAP@0.5	mAP@0.75
1	YOLOv4	54.30%	32.10%
2	SSD (MobileNet)	56.80%	34.50%
3	Faster R-CNN (ResNet-50)	61.20%	37.80%

MAP scores at IoU thresholds of 0.5 and 0.75 are mentioned in Table 1. Among the three algorithms, Faster R-CNN consistently achieves the highest accuracy, with SSD performing second best, and YOLO falling behind in both metrics.

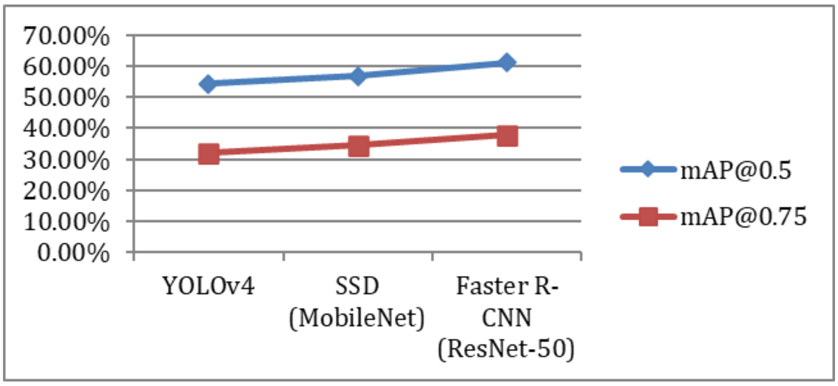

Figure 4 Performance Metrics: Mean Average Precision (mAP)

Faster R-CNN's superior accuracy is greater due to its two-stage detection process, which refines region proposals before classification. While SSD strikes a good balance between speed and precision, it still lags behind Faster R-CNN in accuracy. YOLO's single-stage architecture, although much faster, compromises accuracy, particularly at higher IoU thresholds.

3.2. Inference Time and Frames per Second:

Table 2 Performance Metrics: Inference Time and FPS

SN.	Algorithm	Inference Time(ms)	FPS
1	YOLOv4	25	40
2	SSD (MobileNet)	45	22
3	Faster R-CNN (ResNet-50)	120	8

Table 2 summarizes the inference time per image and FPS. YOLOv4 demonstrates the fastest inference time and highest FPS, followed by SSD, while Faster R-CNN is the slowest.

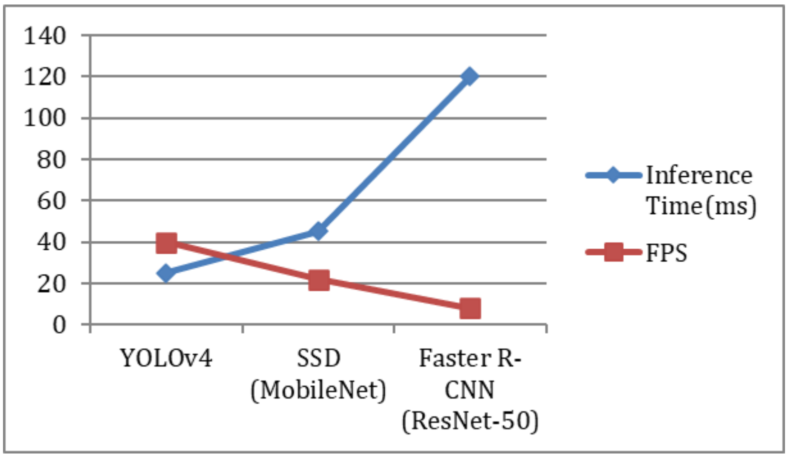

Figure 5 Performance Metrics: Inference Time and FPS

The speed advantage of YOLOv4 stems from its single-pass detection mechanism, making it ideal for real-time applications such as autonomous driving and live video surveillance. SSD delivers a balance between speed and accuracy, making it suitable for various applications. In contrast, Faster R-CNN, while offering high accuracy, has much slower inference times, which limits its usefulness in real-time scenarios.

4. Applications of object detection algorithms

4.1. Autonomous Vehicles

4.1.1. Application

Object detection is crucial for the safe operation of autonomous vehicles. These vehicles must recognize pedestrians, other vehicles, obstacles, traffic signs, and lane markings in real time to make navigation decisions. The goal is to ensure safety by predicting potential collisions or hazards, enabling vehicles to take necessary actions such as braking or changing direction.

4.1.2. Algorithms Used

- YOLO (You Only Look Once): YOLO is known for its fast processing speed, making it ideal for real-time applications. It detects multiple objects in an image with a single pass through the network.
- Faster R-CNN (Region-Based Convolutional Neural Network): This algorithm offers high accuracy in detecting objects. It breaks the detection process into two stages, first identifying regions of interest and then classifying those regions, providing precise detection of various objects.

4.2. Medical Imaging

4.2.1. Application

In medical imaging, object detection algorithms are used to identify anomalies such as tumors, lesions, or other diseases in scans like X-rays, MRIs, and CT scans. These algorithms help in early diagnosis and assist radiologists by reducing the need for manual inspection, improving the overall accuracy and speed of medical assessments.

4.2.2. Algorithms Used

- Faster R-CNN: Its two-stage process is well-suited for detecting complex patterns, such as tumors, in medical images, where high precision is essential.
- Transformers-Based Models: These models are becoming increasingly popular for their ability to comprehend spatial relationships between objects in medical scans, delivering state-of-the-art performance in detecting subtle anomalies.

4.3. Retail and Surveillance

4.3.1. Application

In retail settings, object detection is employed for inventory management, allowing systems to automatically monitor products on shelves, helping businesses maintain stock levels and minimize theft. In surveillance, object detection is utilized to track individuals, detect suspicious activities, and ensure security in real time

4.3.2. Algorithms Used

- SSD (Single Shot Detector): SSD came into the picture for real-time detection with moderate accuracy. It processes images in a single pass, making it efficient for real-time applications like tracking.
- YOLO: Its ability to process frames quickly makes it well-suited for surveillance systems, where timely detection of potential threats or activities is essential.

4.4. Augmented Reality

4.4.1. Application

In augmented reality (AR), object detection enables real-time recognition of real-world objects, allowing virtual content to be seamlessly integrated into the user environment. Applications range from gaming to education, where users interact with digital elements overlaid on physical objects.

4.4.2. Algorithms Used

- YOLO: Its fast processing speed ensures that objects are detected quickly, enhancing the immersive experience of AR by delivering real-time interactions.
- SSD: Its capacity for real-time object detection with minimal computational demands makes it ideal for mobile AR applications, where efficiency is crucial.

5. Conclusion

This paper provided a comparison of several object detection algorithms, emphasizing their strengths and limitations. Despite significant advancements, challenges remain in achieving an optimal balance between accuracy and speed. Future research will likely focus on incorporating AI technologies like natural language processing (NLP) and reinforcement learning to develop more intelligent systems. Enhancing robustness and efficiency will be crucial for broader adoption in sectors such as healthcare and autonomous vehicles. The next wave of innovation aims to make object detection systems more reliable, adaptable, and better equipped to meet the demands of modern AI-driven solutions.

Compliance with ethical standards

Disclosure of conflict of interest

No conflict of interest to be disclosed.

References

[1] Karthikeyan PR. Illumination invariant foreground detection algorithms for real time applications [Master's thesis]. Chennai, India: [University]; 2019.

[2] Maryreeja Y. Object tracking system for traffic surveillance network [Ph.D. dissertation]. Chennai, India: Anna University; 2019.

[3] Sahoo PK, Kanungo P, Mishra S. Slow moving object detection in indoor scene videos [Technical report]. Odisha, India: Department of Electronics and Communication Engineering, Centurion University of Technology and Management; 2019.

[4] Anusha C, Avadhani PS. A novel method for object detection using deep learning [Master's thesis]. Visakhapatnam, India: Department of Computer Science and Systems Engineering, Andhra University; 2020.

[5] Nirmala A, Arivalagan S. A study and development of robust multi object recognition and tracking model using mask region based convolution neural network [Master's thesis]. Annamalai, India: Department of Computer and Information Science, Annamalai University; 2020.

[6] Kaur B. Design of algorithm for robust detection of multi object [Ph.D. dissertation]. Chandigarh, India: Chandigarh University; 2021.

[7] Rawat SS. Developing algorithms for small target detection using infrared images [Master's thesis]. Dehradun, India: Uttarakhand Technical University, Computer Science and Engineering; 2021.

[8] Revathi T. Performance improvement on object detection using deep learning with image enhancement techniques [Master's thesis]. Chennai, India: Anna University; 2021.

[9] Hou BX. Deep learning-based low complexity and high efficiency moving object detection methods [Ph.D. dissertation]. Santa Clara, CA: Santa Clara University; 2022.

[10] Praveen kumar B, Eswaran P. FPGA based acceleration for object detection and tracking applications [Technical report]. Chennai, India: Department of Electronics and Communication Engineering, SRM Institute of Science and Technology; 2022.

[11] Jain P, Meenu. Object recognition using artificial neural network [Master's thesis]. Kurukshetra, India: Department of Mechanical Engineering, National Institute of Technology; 2022.

[12] Francis Alexander Raghu A, Ananth JP. An improved object detection and localization using deep convolutional neural network architecture [Master's thesis]. Chennai, India: Faculty of Information and Communication Engineering, Anna University; 2023.

[13] Vashisht M, Kumar B. Artificial intelligence based approach for object detection in digital image processing [Master's thesis]. Faridabad, India: Department of Computer Science Engineering, Manav Rachna International Institute of Research and Studies; 2023.

[14] Premanand V, Dhananjay Kumar. Deep learning approaches for detection and tracking of multiple moving objects in video [Master's thesis]. Chennai, India: Faculty of Information and Communication engineering, Anna university; 2023.

[15] Bali S, Kumar T, Tyagi SS. Development of object detection technique using machine learning [Master's thesis]. Faridabad, India: Department of Computer Science Engineering, Manav Rachna International Institute of Research and Studies; 2023.

Author’s short biography

Dr Vaibhav Gupta, PhD (Computer Applications) I hold a doctoral degree in Computer Applications and have been actively involved in research in the areas of Cloud Computing and Computer Vision for over 18 years. My research interests lie in exploring the applications of cloud computing and computer vision in real-world problems. With a strong academic background and a passion for innovation, I strive to contribute meaningfully to the field of computer science."	
Mrs. Harshita Mathur (MCA Hons.) I am pursuing PhD in Computer Science and have been actively involved in research in the area of computer vision. My research aims to bridge the gap between computer vision and real-world applications, with a focus on developing innovative solutions.