JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

# Comparative Analysis of Real Time Object Detection Methods Using Machine Learning

**Nishita Rathore*1** **B.Tech Student**
**Sarika Chaudhary[2]** **Associate Professor**
**Poonam Sharma[3]** **Associate Professor**
**Richa Nehra[4]** **Assistant Professor**
**[1,2,3,4] Dept. of Computer Science & Engineering**
**DPG Institute of Technology & Management, Gurugram**

## Abstract

This literature review provides a comparative analysis of 12 research papers on real-time object detection. The focus is on understanding developments, strategies, and trends in the field. Specialized techniques such as CNNs, YOLO variants, and hybrid models are explored, focusing on their strengths, limitations, and applications in areas such as autonomous systems, safety management, and robotic vision. The review identifies innovative approaches for handling challenges like occlusions, motion blur, and dynamic environments. This study provides a collective view of the state of real-time product visibility, which forms the basis for future research in this area.

**Keywords**: Convolutional Neural Networks (CNNs), Deep Learning, Object Tracking, Faster R-CNN, OpenCV, Real-Time Object Detection, Single Shot Detector (SSD), YOLO (You Only Look Once)

## 1. INTRODUCTION

Real-time object recognition is a technology that instantly or nearly instantaneously identifies objects in a video or image. This capability is critical for applications that require rapid decision-making, such as self-driving cars, security systems, augmented reality, robotics, and interactive devices [1-3] . The importance of real-time object detection lies in the possibility of situations where delays can be dangerous or harmful [4]. Self-driving cars, for example, rely on detecting pedestrians, traffic signs, other vehicles, and obstacles, enabling them to pass safely and quickly [5], alerting security systems early simply by detecting suspicious activity or objects in real-time, enhancing public safety [1,3]. Similarly, in healthcare, physicians and radiologists contribute to timely treatment decisions by early detection of abnormalities in medical images [6-8]. Real-time object recognition enables robots to better interact with their environment, supporting industrial tasks and personal assistance [9-11].

Deep learning plays an important role in the real-time discovery process. Techniques such as CNNs have revolutionized object recognition by improving accuracy and speed [3]. Unlike traditional methods, deep learning models can learn directly from data, optimizing object recognition even in difficult situations such as poor lighting, unusual views, and chaotic backgrounds [4]. Deep learning models are also more efficient and scalable, with architectures such as YOLO detecting objects in a single step, reducing processing time and enabling real-time processing [9]. This functionality is further improved by advanced hardware such as the GPU and TPU, which can use this model on low-power devices [10].

Real-time object recognition with deep learning is now used in a variety of industries. It recognizes pedestrians, vehicles, and road signs to ensure safe passage through autonomous vehicles [2,7]. In augmented reality, it overlays digital information from known real-world objects in real time, enhancing the user experience [4]. Real-time detection improves security by instantly detecting suspicious activities, unauthorized access, or abandoned items [1]. In retail, it analyses consumer behaviour and product interactions to support personalized marketing and better store design [5,6]. In healthcare, real-time detection helps diagnose diseases earlier and helps during surgery by quickly analysing imaging data [8,11].

This review will focus on recent advances in real-time object detection, analysis, and comparison of different techniques, objectives, and limitations. It will explore the balance between model complexity and the ability to process information in real time, especially in complex situations such as motion blur, blockage, and changing lighting [9,11].

## 2. Literature Review

Zhihao Jin et. at. [1] investigates real-time object tracking by combining CNNs and OpenCV. CNNs are beneficial for feature learning on large amount of data while OpenCV simplifies the computer vision processes. Even though CNNs enhance the accuracy of the tracked object, it is computer intensive. Integrating TensorFlow with OpenCV and other deep learning tools can take the performance a notch higher. The study investigates ought to be directed towards the constraints of high cost of computing making the method more useful within such fields as video surveillance and autonomous driving.

Yangqing Ye et. al. [2] solves an important problem by designing a real-time object detection system which is aimed at domestic robots that assist individuals in their environment. This system employs special methods which deals with issues of motion blur and partly occluded objects. The system combines DA-Multi-DCGAN for the task of rendering the blurry images crystal clear and AT-LI-YOLO for efficient small object detection so it performs well on less complicated hardware. In this context, the system was trained on the Grasp-17 dataset of a household item and out-competed other models such as YOLOv3 and Faster R-CNN in terms of realisation time as well as precision of the operation. This production has great potential for applications of assistive robots in indoor environments.

The paper "Comparative Analysis of Deep Learning Image Detection Algorithms"[3]focuses on comparing three object detection models namely SSD, Faster R-CNN and YOLO-v3 using Microsoft COCO dataset for the models and reviews the speed, precision, and the model's applicability for real-time like SSD for Quick and performs well in real life applications but is not effective for smaller objects. , Faster R-CNN for Most accurate it's just slower, so it has limited real-time application. ,YOLO-v3 for Strong for real time application as it possesses the best trade-off between speed and accuracy.

This study proposes a framework for detection and tracking in security systems[4]. It combines techniques such as background densification, variable mean detection, and CNNs for near real- time accuracy valuable for environmental monitoring MOT datasets were used to verify the effectiveness and performance of the system and maintenance was achieved without loss of productivity. It is built in Python and C# and is easy to use to control. In the future, the plan is to enhance the system by adding YOLO and other graphics to make work easier in a crowded environment.

Kompalli et. at.[5] Studied how to run the YOLO object recognition model quickly. The key features of a new model, called Faster R-CNN, were combined to make it more accurate. Their method is faster than other models because it uses a special loss function that helps the model make better predictions. It also works well with popular datasets like COCO. This advanced model can be applied to real-life applications such as self-driving cars, safety systems and robotics. The study also showed that the model could be replicated for use in healthcare, smart cities and commercial markets.

This research [6] has been successful in increasing the speed and precision of the SSD (Single Shot Detector) model for conducting real-time object detection. The model has been modified for small object recognition and the model has improved and surpassed over 80% accuracy score on both still frames and videos. This makes the model ideal for applications where mobility, speed and accuracy are the prerequisites such as in image-based security and attendance systems.

Fazal Wahab et. al.[7] developed a real-time system of detection based on SSD and OpenCV for the purpose of detecting the moving and still objects. The SSD_MobileNet_v1 model obtaining 97% accuracy was tested on MS COCO, PASCAL VOC, and Kitti datasets. The study scouted space for the wider application for the detection of object action to make it more applied oriented to monitoring and surveillance for traffic management and security fetching.

P.Devaki et. al. [8] made the real-time object detection device which seeks to assist the visually impaired people in detecting objects around them. Such device comprises the raspberry Pi with open CV and uses SSD with mobilenet to ensure quicker and more effective identification. The prototype delivers messages through voice or vibrating during which time visually impaired users will become aware of close objects.

Parduman Tomar et.al. [9] provide a survey on the use of deep learning for detection of objects in motion and methods of this motion detection by utilizing the CNN, R-CNN and YOLO models. An improved version of the YOLO model is proposed which makes it possible to apply the Darknet53 model making it possible to better detect objects with various sizes of models because the required areas of the model cover several layers of the network. This model also uses k-means clustering to ameliorate detection. Thanks to its high speed which is one of the most attractive features of the YOLO today, it is very useful in real-time applications. This study makes recommendations of incorporating some elements of the technique like edge detection in the future to improve performance further.

D. Pavan et. al.[10] delved into real-time object detection with the use of CNNs and OpenCV. The system is able to identify and follow different objects in a video feed making it useful for applications such as security and counting pedestrians. OpenCV and CNNs interactively contribute into feature extraction and image classification to enhance image recognition and awareness for different use cases such as self-driving cars and helping blind people see.

Malcolm Tivelius's [11] thesis examines real-time miniaturization with Retina Net and YOLOv3 Retina Net was more accurate but took longer to process, while YOLOv3 was faster but struggled with miniatures. Studies show that upgrading hardware and adding preprocessing can significantly improve detection.

"Object Recognition by Deep Learning: A Review" by Zhong-Q Zhao et al. [12]. It explores the transition from traditional recognition methods, which rely on hand features, to deep

learning models that simply learn complex models Reviews key models such as R-CNN, Fast R-CNN , Faster R-CNN, and YOLO, and focuses on their improvements in accuracy and efficiency. The paper also discusses key tasks such as face and pedestrian detection, and suggests future research directions to improve model performance in complex environments

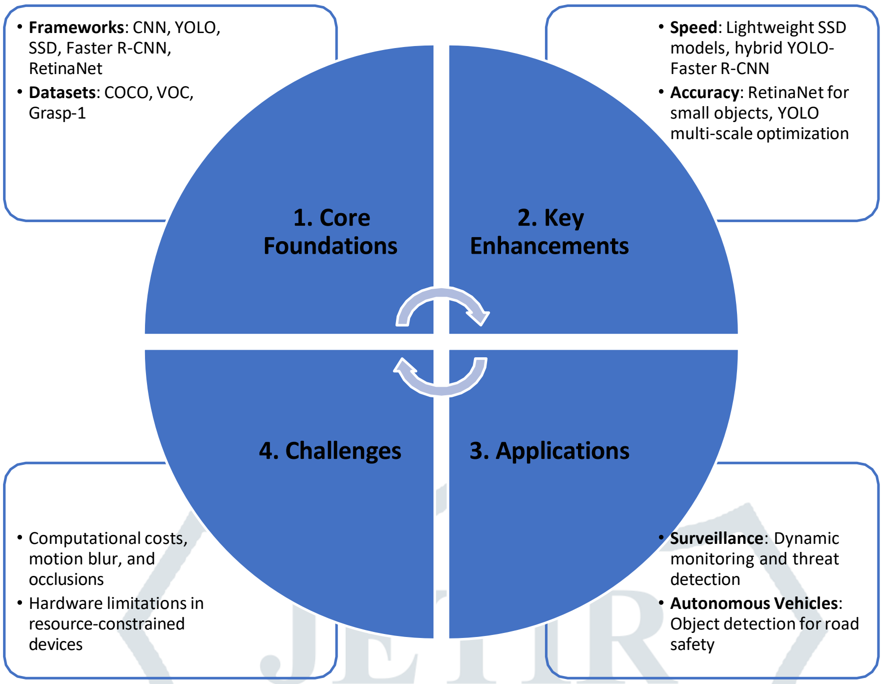


**Fig:1 Hierarchical Model for Real-Time Object Detection**

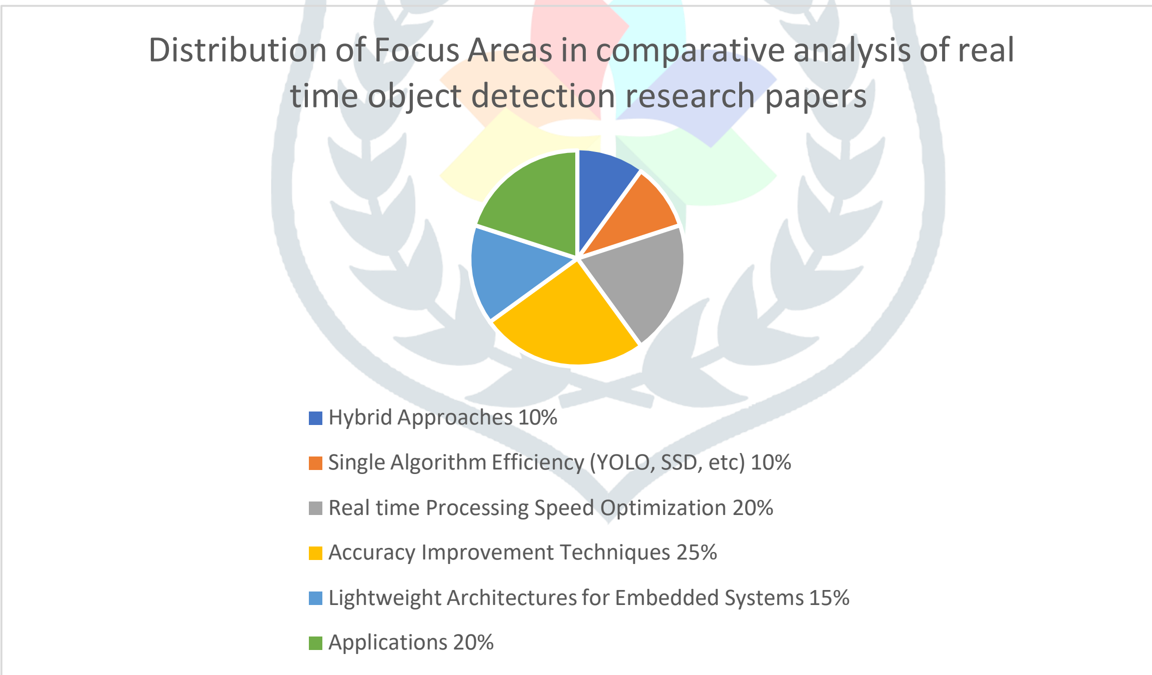


**Fig:2 Distribution of focus areas**

Pie chart in fig.2 represents the distribution of focus areas in comparative analysis of real- time object recognition research papers. This visual summary shows how the papers are

categorized into major topics such as hybrid methods, single algorithm efficiency, real-time optimization, accuracy improvement, small architecture, and applications This helps to highlight trends and areas of focused research effort in the field. Comparative analysis of all the methods is shown in Table 1.

**Table 1. Analysis of different methods**

| Paper | Goal | Tech Used | Method | Key Insights | Challenges | Applications |
|---|---|---|---|---|---|---|
| **Jin et al. (2024)** | Improve real-time object tracking for surveillance. | CNNs, OpenCV, SVMs | Sequentially tracks objects using CNNs in video frames. | Effective tracking but needs more lightweight models for faster speeds. | High data and computing power needs. | Surveillance, autonomous systems. |
| **Ye et al. (2023)** | Recognize blurred and occluded objects in robotic vision. | AT-LI-YOLO, DA-Multi-DCGAN | Uses attention-based mechanisms to detect objects under motion blur. | Strong performance with occluded objects in dynamic settings. | Issues with complex scenes and severe blur. | Home robotics, object recognition. |
| **Srivastav et al. (2021)** | Compare SSD, Faster R-CNN, and YOLO-v3 for performance and accuracy. | SSD, Faster R-CNN, YOLO-v3 | Tests each model on COCO dataset for real-time detection. | YOLO-v3 performs best for speed and balance; Faster R-CNN has highest accuracy. | SSD struggles with small objects; Faster R-CNN slower but more accurate. | General object detection, real-time applications. |
| **Abba et al. (2024)** | Enable real-time detection for security monitoring. | CNNs, Background Subtraction, Labelling | Combines traditional methods with deep learning for dynamic surveillance. | Reliable accuracy in varied settings. | Performance affected by lighting, weather changes. | Security, real-time monitoring. |
| **Kompalli et al. (2023)** | Blend YOLO's speed with Faster R-CNN's precision. | YOLO, Faster R-CNN | Hybrid model to balance accuracy and speed. | Optimized real-time detection through model fusion. | High model complexity, requires more computation. | Robotics, security, autonomous vehicles. |
| **Vaishnavi et al. (2023)** | Enhance SSD for real-time efficiency on low-power devices. | SSD with depth-wise convolutions | Multi-layer SSD to speed up detection, particularly for low-res images. | High accuracy (80%) with low processing requirements. | Small objects in complex environments remain a challenge. | Surveillance, automation. |

| | | | | | | |
|---|---|---|---|---|---|---|
| **Wahab et al. (2022)** | Real-time detection of static and moving objects. | SSD_MobileNet, OpenCV, COCO, VOC | SSD_MobileNet with OpenCV tested on multiple datasets. | 97% accuracy and fast processing for real-time use. | Limited by processing power on low-cost devices (Raspberry Pi). | Traffic monitoring, surveillance, general object detection. |
| **Devaki et al. (2019)** | Build a detection system for assistive tech on Raspberry Pi. | SSD, MobileNet, OpenCV, HOG, R-CNN | Implements SSD and Mobile Net for auditory feedback on Raspberry Pi. | Efficient detection and response on low-power setup. | Environment affects accuracy (e.g., lighting). | Assistive devices, smart home systems, autonomous robots. |
| **Tomar et al. (2022)** | Detect objects of various sizes in real-time. | YOLO, Darknet-53, clustering | Uses YOLO with k-means clustering to boost multi-scale accuracy. | Improved accuracy across different object sizes. | Edge detection could further improve, but at higher computation cost. | Surveillance, video analysis, autonomous driving. |
| **Pavan et al. (2022)** | Multi-object detection for better situational awareness. | CNNs, OpenCV, R-CNN, SSD, YOLO, MobileNet | Real-time multi-object detection using CNNs in video feeds. | Enhances navigation and situational awareness. | Needs substantial datasets and processing power. | People counting, assistive tech, navigation systems. |
| **Tivelius (2020)** | Small object detection for quality control. | RetinaNet, YOLOv3 | Compares RetinaNet (accuracy) with YOLOv3 (speed). | RetinaNet better for small objects, YOLOv3 faster for real-time use. | RetinaNet is more accurate but slower; YOLOv3 is faster but less precise. | Manufacturing quality control. |
| **Zhao (2021)** | Review of deep learning models for object detection. | CNNs, R-CNN, YOLO, SSD, Mask R-CNN, FPN | Compares models on strengths, applications, and trade-offs. | YOLO and SSD are fast; Faster R-CNN and Mask R-CNN offer higher accuracy. | Balancing speed and accuracy remains a challenge. | Surveillance, facial recognition, autonomous driving. |

## 3. Identified gaps

Following gaps were identified while studying different research methods.

1. Increased spending on computers: Many models, such as CNN and advanced algorithms such as RetinaNet [11] or Faster R-CNN [3], provide high accuracy but are computationally expensive, limiting their application to powerful devices learn or in an embedded system.

2. Microscopic detectors: The detection of particles in turbulent environments remains a challenge, as noted in studies such as [6,11]. YOLO, although efficient, struggles with small features [11], while images such as

RetinaNet perform well but at slower speeds.

3. How to handle occlusions and motion blur: The intensive use of occlusions and motion blur is less investigated. [11] introduced DA-Multi-DCGAN to address this, but only for specific applications such as home service robots.

4. Simple environment: Real-time feature recognition on low-power hardware such as the Raspberry Pi [8] faces limitations in accuracy and processing speed.

5. Adaptability to applications: Current models are optimized for specific applications (e.g. security, robotics) but lack optimal frameworks for different applications such as healthcare, smart cities and retail

6. Unseen real-time action: Although object recognition is well studied, few papers deal with real-time action or behaviour recognition [7].

# 4. Results and discussions

Diverse Methodologies: Each paper tends to use different methods, architectures, datasets, or performance metrics, making it difficult to directly compare findings. This variation requires careful analysis to extract meaningful comparisons across studies.

Wide Range of Applications: The papers span various applications, from surveillance and autonomous driving to robotics and assistive technology. Evaluating the relevance and potential impact of each approach across such different fields can be challenging without specific knowledge of each area.

Inconsistent Reporting Standards: Not all papers report information in the same way. Details like specific metrics, testing environments, or hardware setups are often omitted or varied, which makes it harder to accurately assess each model's performance, particularly for real- time applications.

Rapid Advancements: The field of deep learning evolves quickly, so some methods can become outdated as newer models and techniques emerge. This rapid pace requires constant updates, as the relevance of older methods can diminish quickly.

# Conclusion and Future Work

The research will conclude with a comparative analysis of the selected models, emphasizing the practicality and limitations of each model in real-world applications. Recommendations for future research will include exploring advanced architectures like transformers and spatiotemporal networks for improved tracking and detection, especially in dynamic environments with higher occlusion or complex object interactions.

The reviewed studies collectively emphasize that future research should focus on refining these models to improve their robustness and scalability, with particular attention to lightweight architectures for mobile and embedded applications, adaptive learning methods, and action recognition for contextual awareness. The continued evolution of these methods will be essential to meeting the demands of real-time object detection across increasingly complex and dynamic environments, driving impactful applications across industries.

# References


1. Jin, Z., & Yang, H. (2023). *Real-time object tracking using deep learning and OpenCV*. Proceedings of the 2023 International Conference on Machine Learning and Automation. https://doi.org/10.54254/2755-2721/35/20230406

2. Ye, Y.; Ma, X.; Zhou, X.; Bao, G.; Wan, W.; Cai, S. Dynamic and Real-Time Object Detection Based on Deep Learning for Home Service Robots. *Sensors* **2023**, *23*, 9482. https://doi.org/10.3390/s23239482

3. Srivastava, S., Divekar, A.V., Anilkumar, C. *et al.* Comparative analysis of deep learning image detection algorithms. *J Big Data* **8**, 66 (2021). https://doi.org/10.1186/s40537-021-00434-w

4. Sani Abba , Ali Mohammed Bizi , Jeong-A Lee b, Souley Bakouri , Maria Liz Crespo. Real-time object detection, tracking, and monitoring framework for security surveillance systems. DOI:https://doi.org/10.1016/j.heliyon.2024.e34922

5. *Kompalli, V. R., Reddy, S. V. M., Kumar, M. T., & Madala, S. (2023).* Real-time object detection using deep learning. International Journal of Innovative Science and Research Technology, *8(5), 1638-1643. 10.5281/zenodo.7992750*

6. Y. Xu, "Real-time Object Detection Using Deep Learning," ResearchGate, 2023. https://www.researchgate.net/publication/371600210_Real- time_Object_Detection_Using_Deep_Learning

7. Wahab F, Ullah I, Shah A, Khan RA, Choi A and Anwar MS (2022) Design and implementation of real-time object detection system based on single-shoot detector and OpenCV. Front. Psychol. 13:1039645. https://doi:10.3389/fpsyg.2022.1039645

8. Devaki, P., Shivavarsha, S., Bala Kowsalya, G., Manjupavithraa, M., & Vima, E.A. (2019). Real-Time Object Detection using Deep Learning and Open CV. International Journal of Innovative Technology and Exploring Engineering (IJITEE), 8(12S), 411- 41 DOI: 10.35940/ijitee.L1103.10812S19

9. Tomar, P., Sagar, & Haider, S. (2022). *A Study on Real-Time Object Detection using Deep Learning*. *International Journal of Engineering Research & Technology (IJERT)*, 11(5), 465-471. https://doi.org/10.1109/icbda.2017.8078730

10. Pavan, D., Kumar, V. S. A., Sampreeth, J. A. S., & Naveen, K. S. (2020). *Real-time object detection using deep learning* [Bachelor of Technology project report]. Anil Neerukonda Institute of Technology and Sciences.

11. Tivelious, M. (2020). *Real-time Small Object Detection using Deep Neural Networks* (Master’s thesis, KTH Royal Institute of Technology, School of Electrical Engineering and Computer Science, Stockholm, Sweden).

**12.** Z.-Q. Zhao, P. Zheng, S.-T. Xu, and X. Wu, "Object Detection with Deep Learning: A Review," ***IEEE Transactions on Neural Networks and Learning Systems***, vol. XX, no. XX, pp. XX-XX, 2017.