International Research Journal of Multidisciplinary Scope (IRJMS), 2025; 6(3):1611-1621

Original Article | ISSN (O): 2582-631X DOI: 10.47857/irjms.2025.v06i03.04661

YOLO Vs Faster RCNN for Object Detection and Recognition

Ranjana Shende[1]*, Sarika Khandelwal[2]

[1]G H Raisoni University, Amravati, GHRCEM, Nagpur, Maharashtra, India, [2]G H Raisoni College of Engineering, Nagpur, Maharashtra, India. *Corresponding Author's Email: ranjana.shende@gmail.com

Abstract

Object detection has become the cornerstone of various real-world applications, ranging from autonomous driving where it helps detect road objects to medical imaging for early disease prediction and gesture recognition systems. Due to their ability to operate continuously and cost-effectively, computer vision models are increasingly being employed in surveillance and monitoring tasks. Unlike humans, who can intuitively recognize and understand objects within images, machines require advanced algorithms to mimic such capabilities. A key challenge in computer vision lies in recognizing and tracking objects in real-time with accuracy and reliability. Human visual systems perform these tasks naturally and swiftly, enabling complex activities like driving. For computers to match such performance, efficient object detection models along with additional hardware such as sensors are essential. This study presents a comparative analysis of two popular deep learning-based object detection algorithms You Only Look Once (YOLO) and Faster Region-based Convolution Neural Network (Faster RCNN). Both models are evaluated based on their detection accuracy, speed, and performance in real-time scenarios. The paper aims to highlight the strengths and limitations of each model, offering insights into their suitability for different applications. The findings suggest that while both algorithms have their merits, the choice between them depends on the specific requirements of a task, such as the trade-off between detection speed and precision. Comparative study shows that Faster RCNN outperforms its accuracy compared with YOLO.

Keywords: Deep Learning, Faster RCNN, Object Detection, Object Recognition, YOLO.

Introduction

Object recognition and classification represent two of image processing's primary uses in computer vision. A subset of computer vision tasks is called object recognition, and it includes things like identifying objects in digital photos. For image classification to be done, it is necessary to predict the class of a single item within a frame. Object detection, that identifies one or more objects in an image, combines these tasks. Algorithms for object recognition frequently leverage deep learning or machine learning to provide pertinent results. When we look at images or movies, we can easily recognize and pinpoint interesting details. Using a computer that imitates this intelligence is the aim of object detection (1). The one of the more significant challenges in computer vision is real-time object detection, so it's often an essential part of systems for computer vision. Real-time object identification models are applied in several fields such as medical image analysis, robotics, video analytics, autonomous cars, multi-object tracking, and object counting. When compared to straighter object detection techniques, because of its superior feature learning and feature representation capabilities, deep learning for detection of objects is one of its most powerful applications for technology. Creating a system or algorithm that can precisely identify and categorize things in pictures or videos is the goal of object detection and recognition. Numerous industries, including computer vision, robotics, autonomous cars, security, and more, heavily rely on this technology. There are two types of object recognition methodologies single stage detector and two stage detectors. Both have benefits over the other in terms of speed and correctness. RCNN model attempts to use deep learning to solve the object recognition problem. This model detected and localized objects in images by combining convolutional neural networks (CNNs) with region proposal techniques (2). The article begins by introducing the traditional approaches to object detection and then contrasts them with other deep learning approaches. The core objective of object detection and recognition is to develop an algorithm or systems that can accurately identify objects in images or in real time and classify them. This study leverages a multidisciplinary approach

(Received 26th March 2025; Accepted 14th July 2025; Published 30th July 2025)

combining image processing, pattern recognition, machine learning, and artificial intelligence to achieve its objectives. The successive is the study's aim, which informs the formulation of the research objectives which aims to obtain high accuracy with real time processing and detect multiple objects.

A comprehensive review was carried out with an emphasis on object detection security through machine learning. To address a lot of research problems, this entails gathering a sizable collection of research articles on a particular subject from several eras. The intention is to give the reader a summary of the current state of the subject or topic and indicate what further research is needed. They demonstrate whether there are any metrics specific to this area of study by focusing on object detection and machine learning measures. The author talks about the different ways in which the works are applied and where they have been used. The top image-based object detection techniques were specifically recommended for video object detection. Frame-by-frame processing techniques don't work in this case since the things have deteriorated appearances such asymmetrical posture and blurry motion (3). The author provided a thorough discussion of a learning object detection method that manages blocking and low-resolution pictures, using varying degrees of RCNN modification. Generic object detection pipelines are covered first in the topic, and they form the foundation for all other tasks that follow. More additional tasks are then briefly covered, including face and conspicuous object detection (4). Author conducted a survey on a variety of cutting-edge object detection algorithms and the underlying theories. They categorized these techniques into three primary categories transformer based, anchor free and anchor-based detectors. These methods differ in how they recognize items in the picture (5).

There are various object detection techniques available as given in Figure 1. We will now talk about some of the most significant deep learning models and approaches from both the past and present that are used in the computer vision field.

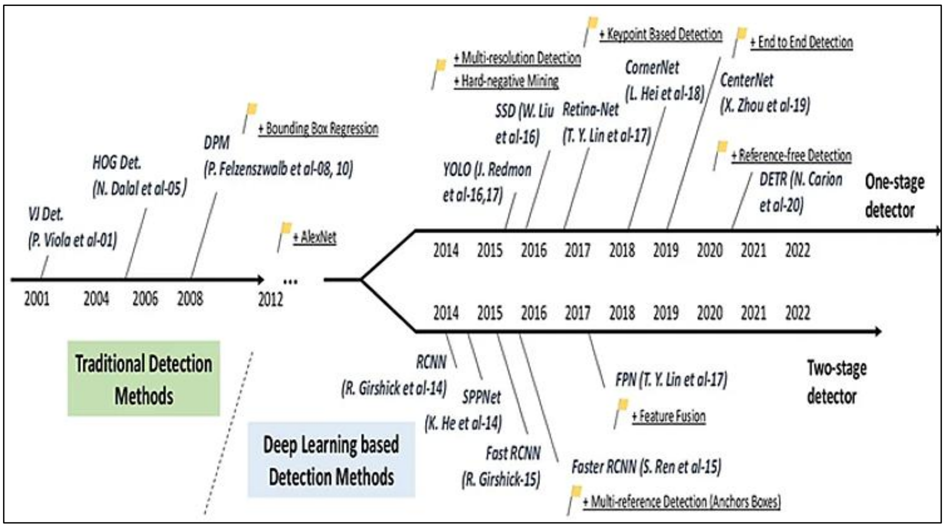

Figure 1: Roadmap and Evolution of Object Detection

Methodology

Generally, object identification algorithms are divided into two groups according to the number of times an image input is transmitted over a network. Two types of object detection methodology are discussed below as given in Figure 2.

Single Shot Object Detection

Single shot object detection algorithms analyze an image in a single pass to identify and locate objects, making them computationally efficient. While they excel in processing entire images quickly, their accuracy diminishes when detecting smaller objects, compared to alternative methods. Nevertheless, their effectiveness makes them suitable for real time object recognition applications in resource constrained environments.

Two Shot Object Detection

This identification of object involves a two-stage process, where the input image is analyzed twice to detect and locate objects. The first stage generates probable object locations, while the second stage refines these proposals to produce specific predictions. Although more accurate than single shot detection, two shot detection requires increased computational resources. The choice between single shot and two shot object detection depends on the specific application requirements and limitation. Commonly single shot detection is suitable for real time applications whereas two shot detection is preferred when precision is paramount.

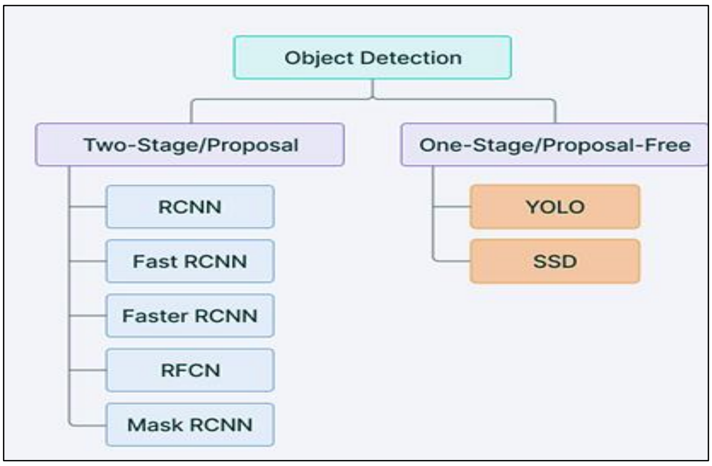

Figure 2: Object Detector

YOLO

The You Only Look Once (YOLO) approach is a one-shot object detection methodology which utilizes a fully convolutional neural network (CNN) to practice images (6). This approach divided an image into a cell grid, applying a single CNN to the entire image. Each cell predicts bounding box matches and confidence scores which represent an object being present in the predicted bounding box. In one forward pass, YOLO models predict bounding boxes and class probabilities immediately, in contrast to two stage detectors that first create region recommendations. When an image is input, YOLO employs a fundamental deep convolution neural system with a foundational architecture to recognize objects exactly and competently.

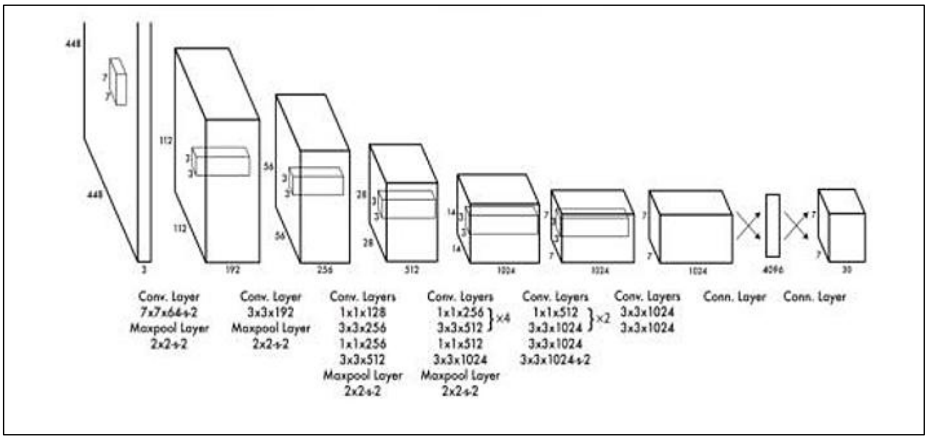

Figure 3: Architecture of YOLO

The architecture functions as follows (as shown in Figure 3):

- An input image is scaled to 448x448 pixels before being fed into the convolutional network.
- A 1x1 convolution is applied, followed by a 3x3 convolution to decrease the number of channels and produce a cuboidal output.
- With the exception of the last layer, which employs a linear activation function, the whole network uses the ReLU activation function.

- Regularization methods such as dropout and batch normalization are hired to prevent over fitting and ensure the model generalizes well (7).

YOLO v2: The YOLO algorithm was upgraded in 2016 with the introduction of YOLO v2, also known as YOLO9000. This updated version aimed to be faster and more accurate than its predecessor, with the ability to detect a inclusive range of object classes. Key improvements include:

- Darknet 19, a more sophisticated CNN backbone that is a VGGNet architecture variation featuring streamlined convolutional layers and pooling layers.
- The introduction of anchor boxes, a set of predefined bounding boxes with varying scales and aspect ratios.
- YOLO v2 predicts bounding boxes by combining anchor boxes with projected offsets, resulting in more accurate detections.
- Overall, YOLO v2 offers significant enhancements over the original YOLO algorithm, making it a more powerful and versatile object detection tool (8).

YOLO v3: 2018 marked the availability of YOLO v3 the third kind of the YOLO object detection algorithm which aims to outperform YOLO v2 in terms of speed and accuracy. Key enhancements include:

- Darknet-53, a new CNN architecture with 53 convolutional layers, tailored for object detection tasks and achieving state-of-the-art performance across multiple benchmarks.
- Feature Pyramid Networks (FPN), which enable the detection of objects at various scales by constructing a feature map pyramid and processing objects at different scales (9).
- Better small item detection performance as a result of the model's multi scale object viewing capabilities.
- Enhanced stability and precision in comparison to earlier iterations of YOLO, capable of managing a more extensive array of object dimensions and aspect ratios.

In general, YOLO v3 offers improved performance and dependability marking a substantial increase in object detecting capabilities (10).

YOLO v4: The adoption of a revolutionary CNN architecture known as CSPNet (Cross Stage Partial Network) given in Figure 4 is the primary enhancement over YOLO v3. A version of the ResNet architecture created especially for object detection tasks is called CSPNet. CSPNet is quite shallow with only 54 convolutional layers but it performs remarkably well in object recognition. Its novel architecture raises the bar for accuracy and performance in object identification capabilities by allowing for notable improvements.

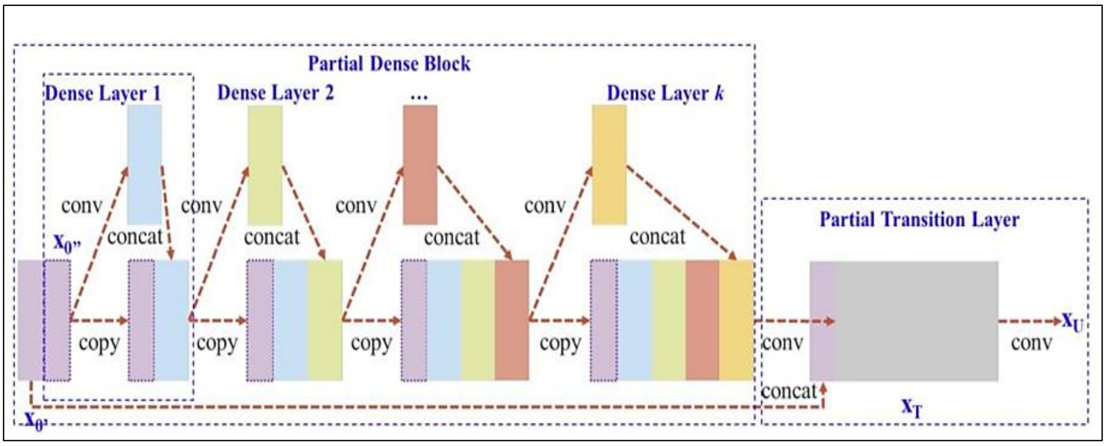

Figure 4: Architecture of CSPNet

To enhance the size and shape of detected objects, YOLO v3 and v4 make use of anchor boxes with varying aspect ratios and sizes. YOLO v4 presents a new method for creating anchor boxes k-means clustering. This method clusters bounding boxes and creates anchor boxes from the cluster centroids, allowing for more accurate representations of object dimensions and shapes. Although YOLO v4 and v3 share similar loss functions during training. In order to enhance

performance on balanced datasets, version 4 presents a novel idea termed GHM loss, a variation of the focal loss function. Additionally, YOLO v4 adopts enhanced Feature Pyramid Network (FPN) designs from YOLO v3, further boosting its object detection capabilities (11).
A comparison between existing state-of-the-art object detectors and the projected YOLOv4 is shown in Figure 5 which highlights notable improvements in speed, accuracy, and computational efficiency. YOLO v4, integrates advanced techniques such as CSP Darknet53, weighted residual connections, and spatial pyramid pooling, resulting in a balanced architecture that achieves real-time detection with high accuracy.

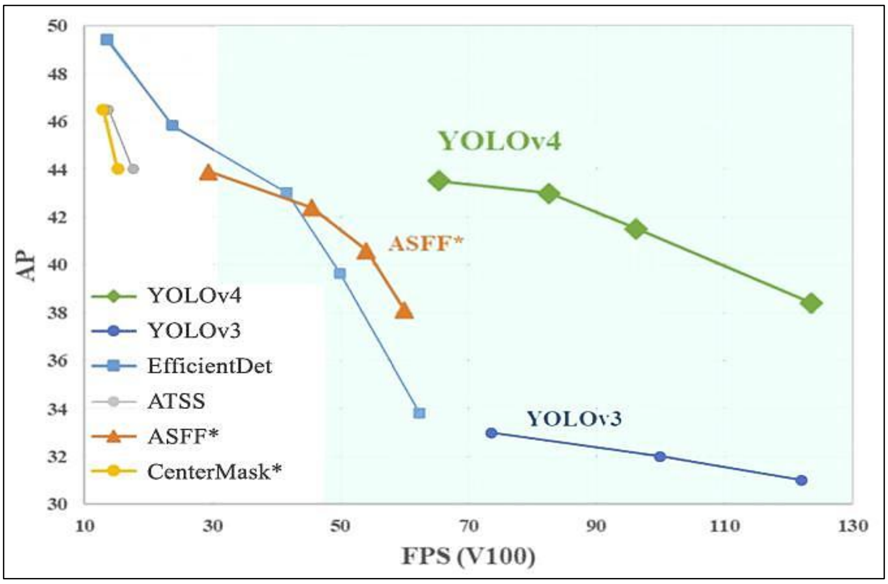

Figure 5: A Comparison between Existing Cutting-Edge Object Detectors and the Projected Yolov4

YOLO v5: The first set of data used to train Yolo was the 20-object category PASCAL VOC dataset. If the D5 dataset, which includes 600 different, object categories, is smaller and less varied, and then YOLO v5 was trained on it. YOLO v5 presents dynamic anchor boxes, generated using a clustering method to match the sizes and shapes of detected objects more accurately. Additionally, YOLO v5 employs Spatial Pyramid Pooling (SPP) to improve detection performance on minor objects by allowing the model to observe objects in various scales (12). SPP is still present in YOLO v5, it does so more efficiently. Furthermore, YOLO v5 introduces IoU loss, a variant of the IoU loss function, aimed at improving performance on unbalanced datasets. Overall, YOLO v5 boasts significant architectural enhancements and improved training data, leading to more accurate and robust object detection capabilities (13).
YOLO v6: YOLO v6 released in 2022 build upon the enhancement over the previous versions. The CNN architecture of YOLO v5 and YOLO v6 is a major distinction. In contrast to Efficient Det, which was employed in YOLO v5, EfficientNet-L2, a modified form of the Efficient Net design, delivers better computing efficiency and fewer parameters in YOLO v6. YOLO v6 makes a number of advancements, such as

- Reduced number of false positives: YOLO v6 is more accurate and less prone to detecting non-existent objects, resulting in fewer false positives.
- Improved anchor box generation: The dense anchor box method in YOLO v6 leads to more accurate bounding box predictions.
- Enhanced computational efficiency: The EfficientNet-L2 architecture in YOLO v6 reduces the number of parameters and improves processing speed.
- Better object detection performance: YOLO v6 incorporates various cutting-edge techniques, resulting in improved detection accuracy and reliability (9).

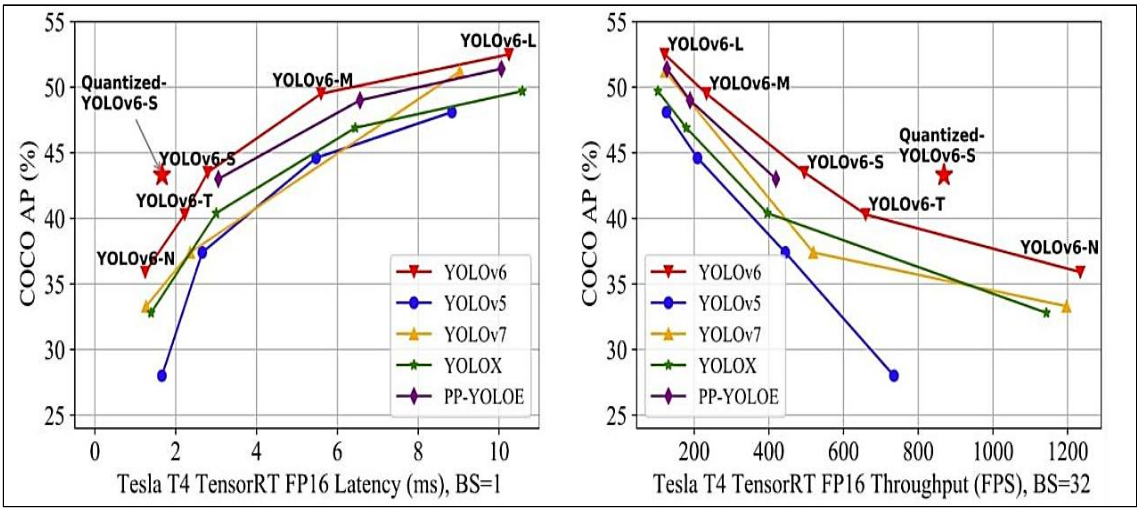

Figure 6: A Comparison of the Most Advanced and Effective Object Detectors

As given in Figure 6 Overall, YOLO v6 offers a more robust and efficient object detection solution compared to its predecessors.

YOLO v7: YOLO v7 works on nine anchor boxes In contrast to previous iterations it can identify a greater variety of item sizes and shapes, which helps lower the amount of false positives. By using the novel loss function known as focal loss YOLO v7 has made significant progress. Figure 7 represent Performance and inference speed comparison of YOLO v7. Ii can identify tiny objects and has a higher overall accuracy thanks to its higher resolution. One of the main benefits of YOLO v7 is its speed (9). It scans images at a rate of 155 frames per second, which is far quicker than current the most advanced object detection algorithms. Up to 45 frames per second might be processed by even the initial basic YOLO model. Because increased processing rates are essential, this makes it suited for sensitive real-time applications like self-driving cars and surveillance.

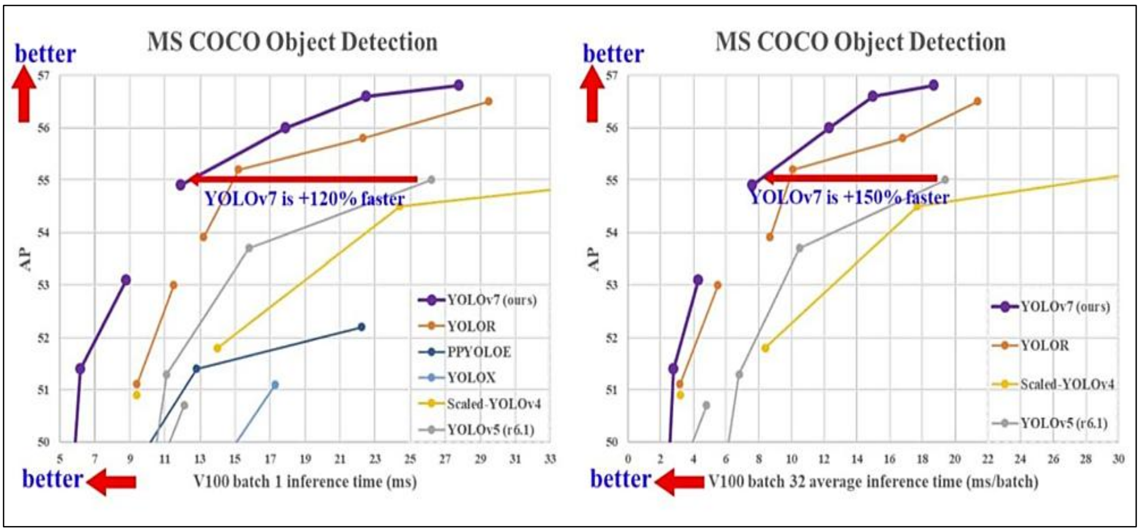

Figure7: Performance and Inference Speed Comparison of YOLO v7

The precision of YOLO v7 is superior to that of other object recognition methods. It acquires an average correctness of 37.2% at an intersection over union threshold of 0.5 on the famous COCO dataset which is in line to other contemporary object detection techniques. The numerical performance is compared below shown in Figure 8.

Model	#Param.	FLOPs	Size	AP^{val}	AP^{val}_{50}	AP^{val}_{75}	AP^{val}_{S}	AP^{val}_{M}	AP^{val}_{L}
YOLOv4	64.4M	142.8G	640	49.7%	68.2%	54.3%	32.9%	54.8%	63.7%
YOLOR-u5 (r6.1)	46.5M	109.1G	640	50.2%	68.7%	54.6%	33.2%	55.5%	63.7%
YOLOv4-CSP	52.9M	120.4G	640	50.3%	68.6%	54.9%	34.2%	55.6%	65.1%
YOLOR-CSP	52.9M	120.4G	640	50.8%	69.5%	55.3%	33.7%	56.0%	65.4%
YOLOv7	36.9M	104.7G	640	51.2%	69.7%	55.5%	35.2%	56.0%	66.7%
improvement	-43%	-15%	-	+0.4	+0.2	+0.2	+1.5	=	+1.3
YOLOR-CSP-X	96.9M	226.8G	640	52.7%	71.3%	57.4%	36.3%	57.5%	68.3%
YOLOv7-X	71.3M	189.9G	640	52.9%	71.1%	57.5%	36.9%	57.7%	68.6%
improvement	-36%	-19%	-	+0.2	-0.2	+0.1	+0.6	+0.2	+0.3
YOLOv4-tiny	6.1	6.9	416	24.9%	42.1%	25.7%	8.7%	28.4%	39.2%
YOLOv7-tiny	6.2	5.8	416	35.2%	52.8%	37.3%	15.7%	38.0%	53.4%
improvement	+2%	-19%	-	+10.3	+10.7	+11.6	+7.0	+9.6	+14.2
YOLOv4-tiny-3l	8.7	5.2	320	30.8%	47.3%	32.2%	10.9%	31.9%	51.5%
YOLOv7-tiny	6.2	3.5	320	30.8%	47.3%	32.2%	10.0%	31.9%	52.2%
improvement	-39%	-49%	-	=	=	=	-0.9	=	+0.7
YOLOR-E6	115.8M	683.2G	1280	55.7%	73.2%	60.7%	40.1%	60.4%	69.2%
YOLOv7-E6	97.2M	515.2G	1280	55.9%	73.5%	61.1%	40.6%	60.3%	70.0%
improvement	-19%	-33%	-	+0.2	+0.3	+0.4	+0.5	-0.1	+0.8
YOLOR-D6	151.7M	935.6G	1280	56.1%	73.9%	61.2%	42.4%	60.5%	69.9%
YOLOv7-D6	154.7M	806.8G	1280	56.3%	73.8%	61.4%	41.3%	60.6%	70.1%
YOLOv7-E6E	151.7M	843.2G	1280	56.8%	74.4%	62.1%	40.8%	62.1%	70.6%
improvement	=	-11%	-	+0.7	+0.5	+0.9	-1.6	+1.6	+0.7

Figure 8: Quantitative Comparison of the Performance YOLO v7

Working of Faster RCNN

The suggested study uses a camera to take images and videos, which are pre-processed and subsequently analyzed to identify objects. Following pre-processing, a quicker RCNN approach is used for object detection. The three parts of Faster RCNN are separated (14).

Convolution Layers

In order to extract appropriate data from the images, filters are developed in the convolution layers. If we train these filters to retrieve say, the key elements of a face, they will identify the unique colors and contours of the human face. Convolution network are composed of three layers the fully connected layer, which is used for classification and detection tasks, the pooling layer, and the convolution layer (15). Sliding a filter along the length of the original image yields the location of the pooling layer. A two-dimensional feature map is generated as a matrix.

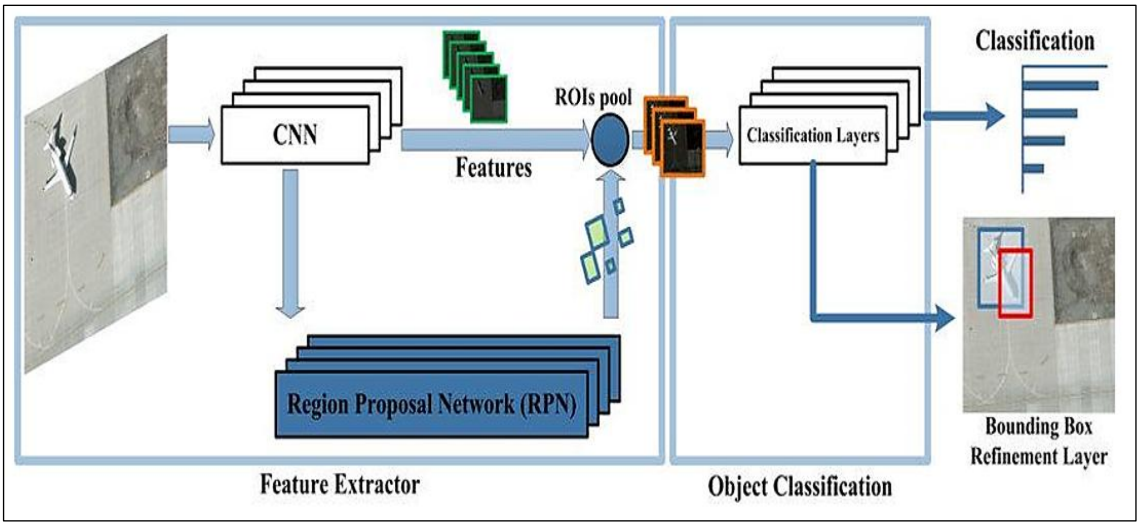

Figure 9: Architecture of Faster RCNN

Region Proposal Network (RPN)

In order to identify whether an object and its bounding box are present in the image, the multilayer perceptron referred to as the region proposal network repeats across the convolution layers' feature map (16). The feature map of the final shared convolution layer is routed via n × n rectangular sliding window. Concepts for K regions are produced for each window. An anchor box is used as the reference point while determining each proposal (17).

Classes and Bounding Boxes Prediction

RPN is used as an input to a fully intersected neural network which is utilized to predict bounding box and object class (18). Figure 9 illustrates generation of region proposals where an object might be located. Anchor boxes come in a variety of sizes to accommodate objects of all sizes. Every region proposed in the image has a fixed-length feature path removed by the ROI pooling layer. ROI aims to reduce the feature map overall size to the same level. Regression and classifiers are our last options. Classifier for figuring out if an image has an object in it or just background. Regressions provide a bounding box for the object being categorized.

Results and Discussion

Proposed system identifies objects in images and videos in real time. We set up a few experiments to verify that the suggested method is accurate. The designed solution trains a pre trained RCNN model more quickly using a dataset by utilizing the TensorFlow Object Detection API. Real-time images are captured by the implemented system when the program is run, allowing for object detection and classification. Both YOLO and Faster R-CNN models are initialized with weights pretrained on large benchmark datasets like COCO before being fine-tuned on the target dataset. The system can also recognize many things, such as a person, bicycle and a slipper together. Figure 10 and Figure 11 illustrates the outcome of confident object detection. Bottle displays 66% confidence, Scissor displays 55% confidence, Cup displays 99% confidence and person displays 85% confidence also other smaller tiny objects are detected like slipper.

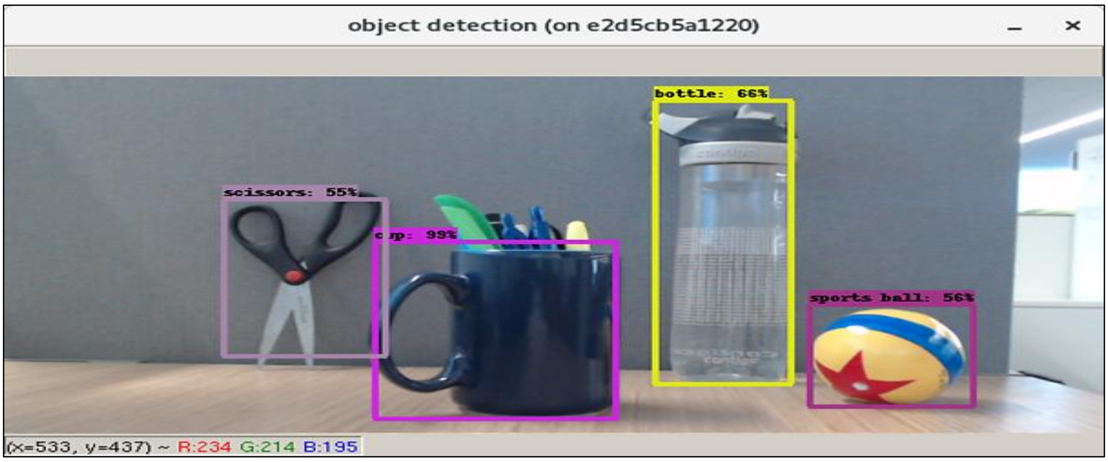

Figure 10: Real Time Object Detection with YOLO

Figure 11: Object Detection with Faster RCNN

Comparative Analysis

Analysis of single stage object detection DETR (Detection Transformer) and SSD (Single Shot MultiBox Detector)/YOLO is described in Table 1. SSD and YOLO is a CNN-based architecture, combining multiple convolution layers for feature extraction and prediction (19). DETR employs a Transformer-based architecture, initially designed for natural language processing. It utilizes self-attention mechanisms to capture global context. DETR/SSD achieves state-of-the-art accuracy in object detection due to its Transformer architecture (20).

Table 1: Analysis of One Stage Detector

Criteria	DETR	SSD/YOLO
Architecture	Transformer-based	CNN-based
Speed	Medium	Fast
Accuracy	Very High	High
Training Speed	Slower	Faster
Object Class Limit	Limited by vocabulary	Limited by anchors
Real-time Inference	No	Yes
Anchor Boxes	Not used	Yes
Box Regression Method	Learnable offsets	Anchor-based

Analysis of two stage object detection RCNN, Fast RCNN and Faster RCNN with Prediction time / image is given in Table 2. RCNN uses selective search method to extract image. In Fast RCNN Feature maps are extracted from each image after it is sent through CNN just once. Faster RCNN replace selective search with Region Proposal Network (RPN). RPN is a multilayer perceptron that moves over the feature of convolution layer to predict weather there is object or not also predict bounding box of that object.

Table 2: Analysis of Two Stage Detector

Algorithm/ techniques	Features	Prediction time / image
R-CNN	Generates areas using selective search around 2000 areas are extracted per image.	40-50 seconds
Fast RCNN	Feature maps are extracted from each image after it is sent through CNN just once. These maps are subjected to selective search in order to produce predictions.	2 seconds
Faster RCNN	Uses a quicker approach to perform selective search instead of the region proposal network (RPN) method.	0.2 seconds

Conclusion

The single shot object detection algorithm is SSD and YOLO. YOLO prioritizes speed while faster RCNN focuses on accuracy. Similar to many object detection algorithms, YOLO v7 has trouble identifying small things and isn't entirely accurate at classifying objects at various scales. Because YOLO v7 can be computationally demanding, it may not be possible to run it in real time on devices with partial resources such as smartphones or other edge devices. While working with two stage a faster R-CNN deep learning network technique, objects in images may be successfully identified and classified.

Through the convolution layer, RPN is used. We ascertain whether or not the bounding box contains an object for each geographic point. Fixed size ROI pooling is used from varying area sizes, and an item is then sorted into distinct categories. It is quicker than a fast RCNN because it requires less computing time because selective search is not needed. Faster RCNN is more robust to different object poses, scales and occlusions. Future scope of Faster RCNN can be extended in 3D and Multimodal Object Detection to handle multi-view or multi-modal inputs.

Abbreviations

CNN: Convolutional Neural Network, CSPNet: Cross Stage Partial Network, FPN: Feature Pyramid Networks, R-CNN: Region-based Convolutional Neural Network, SSD: Single Shot MultiBox Detector, VGGNet: Visual Geometry Group Network, YOLO: You Only Look Once.

Acknowledgement

The authors would like to express their sincere gratitude and appreciation to G H Raisoni University Amravati, for providing the necessary resources and facilities. Special thanks to guide Dr. Sarika Khandelwal whose constructive comments and suggestions helped improve the quality of this work.

Author Contributions

Ranjana Shende: Conceptualization, Design, Methodology, Investigation, Writing – Original Draft, Sarika Khandelwal: Formal Analysis, Writing Review, Editing, Supervision.

Conflict of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Ethics Approval

This study did not involve any data collection requiring ethics approval.

Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

References

1. Railkar Y, Nasikkar A, Pawar S, Patil P, Pise R. Object Detection and Recognition System Using Deep Learning Method. In2023 IEEE 8th International Conference for Convergence in Technology (I2CT). IEEE. 2023 Apr 7:1-6. https://ieeexplore.ieee.org/abstract/document/10126316
2. Girshick R. Fast R-CNN. Proc. IEEE International Conference Computing Vision. 2015. https://doi.org/10.48550/arXiv.1504.08083
3. Ray S. A quick review of machine learning algorithms. In 2019 International conference on machine learning, big data, cloud and parallel computing (COMITCon). IEEE. 2019 Feb 14:35-39. https://ieeexplore.ieee.org/abstract/document/8862451
4. Singh YB, Mishra AD, Nand P. Use of machine learning in the area of image analysis and processing. In2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN). IEEE. 2018 Oct 12:117-120. https://ieeexplore.ieee.org/stamp/stamp.jsp?arnumber=8748338&casa_token=vFNP7T_C-vAAAAAA:qDkAk1_BtlD3my_LlLbU_N-PYCaljm7a7vZd-cbnEWcK_4so-nvXBj-QMnpx-LYpR6KAEVBG5hcd&tag=1
5. Amjoud AB, Amrouch M. Object detection using deep learning, CNNs and vision transformers: A review. IEEE Access. 2023 Apr 10;11:35479-516.
6. AlZaabi A, Abi Talib M, Nassif AB, Sajwani A, Einea O. A systematic literature review on machine learning in object detection security. In2020 IEEE 5th International Conference on Computing Communication and Automation (ICCCA). IEEE. 2020 Oct 30:136-139. https://ieeexplore.ieee.org/document/9250836
7. Joiya F. Object detection: yolo vs faster R-CNN. Int Res J Modern Eng Technol Sci. 2022 Sep;9:1911-5.
8. YOLO v2 – Object Detection. 2025. https://www.geeksforgeeks.org/machine-learning/yolo-v2-object-detection/
9. Klingler N. YOLOv4: A Fast and Efficient Object Detection Model. Blog. 2024. https://viso.ai/deep-learning/yolov4/
10. Sharma T, Chehri A, Fofana I, Jadhav S, Khare S, Debaque B, Duclos-Hindie N, Arya D. Deep learning-based object detection and classification for autonomous vehicles in different weather scenarios of Quebec, Canada. IEEE Access. 2024 Jan 15;12:13648-62.
11. Zhao ZQ. Object Detections with Deep Learning: A Review. IEEE Trans Neural Network Learning System. 2019. https://arxiv.org/pdf/1807.05511
12. Asokan A, Anitha J. Machine Learning based Image Processing Techniques for Satellite Image Analysis -A Survey. 2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon), Faridabad, India. 2019:119-124. https://doi.org/10.1109/COMITCon.2019.8862452
13. Santra S, Roy S, Sardar P, Deyasi A. Real-time vehicle detection from captured images. In2019 International Conference on Opto-Electronics and Applied Optics (Optronix). IEEE. 2019 Mar 18:1-4. https://www.ijert.org/real-time-vehicle-detection-from-captured-image

14. Kong X, Li X, Zhu X, Guo Z, Zeng L. Detection model based on improved faster-RCNN in apple orchard environment. Intelligent System Appl. 2024;21: 200325.
15. Ren S, He K, Girshick R, Sun J. Faster R-CNN: Towards real-time object detection with region proposal networks. IEEE transactions on pattern analysis and machine intelligence. 2016 Jun 6;39(6):1137-49.
16. Panda C. Object detection and tracking using faster R-CNN. International Journal of Recent Technology and Engineering. 2019 Sep;8:4894-900.
17. Vijiyakumar K, Govindasamy V, Akila V. An effective object detection and tracking using automated image annotation with inception based faster R-CNN model. International Journal Cognitive Computing Eng. 2024;5:343-356.
18. Vaishnavi K, Reddy GP, Reddy TB. Real-time Object Detection Using Deep Learning. J Advanced Math Computer Sci. 2023;38(8):24–32.
19. Haris M, Glowacz A. Road object detection: A comparative study of deep learning-based algorithms. Electronics. 2021;10(16):1932.
20. Wu Y, Zhang H, Li Y, Yang Y, Yuan D. Video object detection guided by object blur evaluation. IEEE Access. 2020 Nov 18;8:208554-65.