The Academic
International Journal of Multidisciplinary Research

An Online Peer Reviewed / Refereed Journal
Volume 3 | Issue 6 | June 2025
ISSN: 2583-973X (Online)
Website: www.theacademic.in

Real-Time Object Detection in Images: A Review of YOLO and Faster R-CNN Approaches

Mrs. Vaibhavi Vashi

Research Scholar, Vidhyadeep University, Anita-Kim (Gujarat), India

E-mail Id: vaibhavi.kathwadiya@gmail.com

DOI : https://doi.org/10.5281/zenodo.15858344

ARTICLE DETAILS

Research Paper

Accepted: 27-06-2025

Published: 10-07-2025

Keywords:

YOLO, Faster R-CNN, Deep Learning, CNN, Image Processing, Computer Vision

ABSTRACT

One of the main tasks in computer vision is object detection which is importance to many real-time applications such as robotics, autonomous driving, surveillance and ultrasound. Two of the most well-known and frequently helpful frameworks among the different object detection algorithms created in recent years are YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network). Compare between YOLO and Faster R-CNN in this review study, with discussion on their architectural distinctions, better performance signals, pros and cons. YOLO is famous for its quick detection speeds, provides a one-stage detection pipeline that allows for processing in real time with a respectable level of accuracy. The precision and general accuracy of the two-stage identify Faster R-CNN on the other side are superior where albeit at the expense of longer computation times. This study also discusses recent improvements to structure; application positions and standards were outputs on popular database such as COCO and PASCAL VOC. By accepting aspects including speed, accuracy and deployment environment. This paper moto to help researchers and joiners extract the optimal structure for real-world object detection tasks.

Introduction:

Searching and identifying objects inside images or video frames is the aim of object detection is a general computer vision problem. Autonomous vehicles, intelligent surveillance, traffic monitoring, facial recognition, industrial automation and human computer interaction are just a little of the real world uses for it. Among the different methods created in the area, object detection models based on deep learning have illustrate exceptional accuracy and processing speed. Because of their complementing advantages, YOLO (You Only Look Once) and Faster R-CNN (Region-based Convolutional Neural Network) have become two of the most popular object recognition frameworks.

As a result of the increase accuracy for real time object detection system there are two stage, Single-stage and two-stage detectors. YOLO is a single-stage detector which reframes object identify as a regression problem. It can overview images at very high speed through center accuracy by directly assuming bounding boxes and class probabilities from complete images in a single assessment [1]. On the other hand, Faster R-CNN is a two-stage detector that gives high precision at the cost of a higher computational charge. It first creates region proposals before classifying them [2].

A paradigm change was brought about by the initial YOLO algorithm, which was presented by Redmon et al. [1] and processed images at 45 frames per second while reaching real-time object detection performance. After that, YOLOv2 and YOLOv3 increased this's accuracy and resilience. YOLOv4 [3] and YOLOv5 [4] included optimization methods such CSPDarkNet, PANet and auto-learning bounding box anchors to feature boost speed and accuracy. The latest version YOLOv8 [5] is one of the greatest real-time detectors with new architectural features and training enhancements.

As per the the Ren et al. [2], Faster R-CNN which combines a Region Proposal Network (RPN) with a convolutional backbone create an end-to-end experiment system with improved localization and classification accuracy. Rater than YOLO the Faster R-CNN is a strong choice for high precision and comprehensive object instance segmentation such as medical picture analysis and document digitization.

To compare the success of YOLO with Faster R-CNN in different aspects, including occlusion, illumination, and object scale, numerous research have been carried out. For example, YOLO carry out better in terms of speed, but Faster R-CNN is better in terms of accuracy, according to Zhao et al.'s thorough study of object recognition frameworks [6]. Practical contrast employing the PASCAL VOC and MS COCO datasets were used in different publications such as by Tzutalin et al. [7] which confirmed the trade-offs between speed and precision.

Because of their reduced latency YOLO based models have been largely access in real time applications such as license plate identification and helmet recognition. In order to identify non-helmet riders and efficiently extract license plate information from live traffic feeds studies like Prajwal et al. [8] and Singh et al. [9] used YOLO with OCR. These uses illustrate how YOLO is useful in edge computing situations where real-time inference is need.

On the other hand, Faster R-CNN has been successfully used in surroundings where correctness is valued and computational resources are less. For instance, the accuracy gives benefits of Faster R-CNN surpass its slower carry out in high-resolution satellite image processing and medical diagnostics [10].

The distinction between high-accuracy and real-time detectors is becoming lower as the field develops. Both YOLO and Faster R-CNN are able to good between speed and accuracy glade to hybrid models and improvements in backbone designs, loss functions and training methodologies.

The goal of this review is to illustrate a comprehensive comparison of the Faster R-CNN and YOLO frameworks for real-time object detection. to help researchers and developers get the most valuable model information for certain use cases. It examines upcoming research issues, architectural variations, performance benchmarks and application sectors.

LITURATURE REVIEW

Due to its versatility over various area, object detection has emerged is a key area of study in computer vision in continue years. Manually features and traditional machine learning algorithms were the pillar of early approaches like the Viola-Jones detector. The control underwent a revolt with the introduction of deep learning as CNN-based detectors showed significant gains in speed and accuracy.

Much research and development has gone into the YOLO family of detectors. YOLO was said by Redmon et al. [1] as a different field model for real-time detection and it performed remarkably faster than conventional techniques. By improving feature extraction and bounding box prediction YOLOv2 and YOLOv3 improved the structure. To increased detection speed and accuracy YOLOv4 [3] and YOLOv5 [4] introduced features with Cross Stage Partial Networks (CSPNet), spatial pyramid pooling and weighted residual connections.

The past R-CNN and Fast R-CNN designs were increased by the Faster R-CNN model which was put forth by Ren et al. [2] and present the idea of Region Proposal Networks (RPN). By integrating object classification and presentation creation into a single framework which makes end-to-end practices possible. It is not faster than YOLO but it is more accurate, particularly when it comes to recognizing small things.

Comparative studies, as the ones by Zhao et al. [6] indicate that Faster R-CNN performs exceptionally good in terms of accuracy and especially for complex datasets while YOLO is more considerable for real-time applications because of its single-pass design. These outputs are confirmed by Tzutalin et al. [7] and moreover, benchmark analyses on datasets such as MS COCO and PASCAL VOC.

A multitude of domain certain applications illustrate these models' practically. The way of detect non-helmet riders and retrieve license plates, Prajwal et al. [8] created a YOLOv2 based system that demonstrate successful in real time traffic monitoring. This way was increased by Singh et al. [9] with YOLOv3 for increased accuracy in dynamic settings. In the meantime, Faster R-CNN's information detection skills continue to be pros for applications in industrial inspection, satellite data analysis and medical imaging [10].

To capitalize on the benefits of both frameworks firstly, new research is concentrating on transfer learning and secondly, hybrid models. For instance, employing focus structure to increase the feature extraction or combining RPN with YOLO's detection channel has illustrate possibility in enhancing performance over a number of parameters.

In general, the literature shows a definite trade-off between the accuracy of Faster R-CNN and the speed of YOLO. The need of the application, including latency limitations, computational power and object complexity have a significant impact on the model selection.

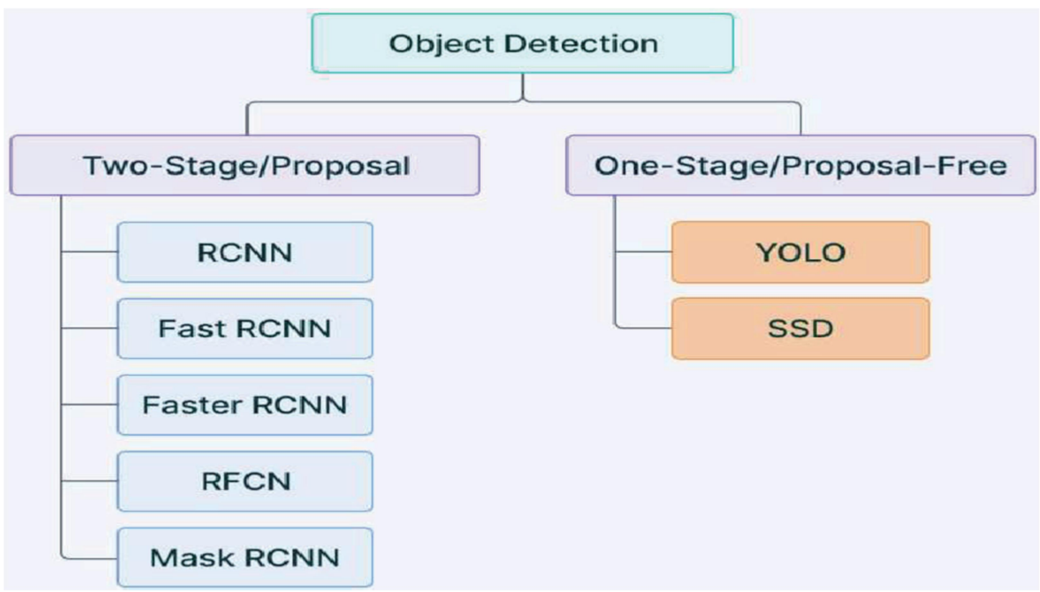

Figure 1. Object Detection Proposal

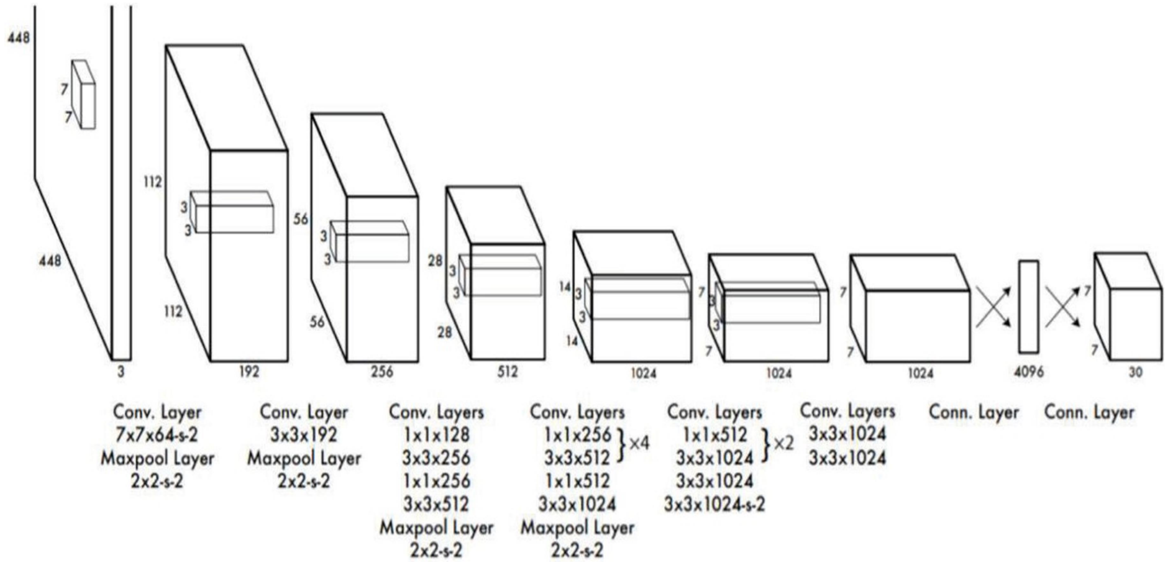

Figure 2 Yolo architecture.

Figure 2: The deep convolutional neural network was use in the original YOLOv1 model for real-time object detection is depicted in the image. The architecture is made by a chain of convolutional layers and fully connected layers. Here is an explanation run down per layer:

CONCLUSION:

Two well-known but different methods in the object detection space are YOLO and Faster R-CNN. Because of its real-time processing optimized model YOLO is good for time sensitive applications with embedded vision systems, traffic monitoring and video surveillance. Its arrangements on edge devices with constrained computational ability is made manageable by its quick inference and comparatively light weight. Faster R-CNN is on the other hand, gives better detection accuracy, particularly when recognizing small and overlapping objects or in complicated surroundings. Because of this, it is better fitted for applications where precision is essential and there is enough processing power.

It is clear from the literature that selecting an object detection architecture needs critical consideration of the trade-off among speed and accuracy. New backbone architectures and training methods are being added to Faster R-CNN structure to low down inference time while YOLO continues to advance toward better accuracy with iterations such as YOLOv5 and YOLOv8.

It is predicted that future developments in object identification will focusing on reducing the imbalance between speed and precision by using hybrid models, increase feature extraction methods and successful model optimization for cross-platform deployment. Furthermore, integrating semi-supervised and unsupervised techniques of investigation with neural techniques for attention and transformers which increase object detection even further.

In conclusion, the selected needs of the application in terms of speed, accuracy and available computational resources should indicate the decision between YOLO and Faster R-CNN. Designing efficient real-time object detection systems that meet a variety of real-world needs can be increased by a full grasp of the pros and cons of both structures.

References

1. Automated License Plate Recognition for Non-Helmeted Motor Riders Using YOLO and OCR arxiv.org+15journals.riverpublishers.com+15e3s-conferences.org+15
2. Real time license plate number extraction of non-helmet person using YOLO algorithm sciencescholar.us
3. Real-time license plate detection for non-helmeted motorcyclist using YOLO sciencedirect.com+1sciencescholar.us+1
4. Detection of License Plate Numbers and Identification of Non-Helmet Riders using YOLO v2 and OCR Method turcomat.org+10discovery.researcher.life+10ijisae.org+10
5. Detection of Non-Helmet Riders and Extraction of License Plate Number using YOLO v2 and OCR Method en.wikipedia.org+3discovery.researcher.life+3e3s-conferences.org+3
6. Detection of non-helmet riders and extraction of license plate number using Yolo v2 and OCR method
7. Prajwal M. J. et al., Detection of non-helmet riders and extraction of license plate number using Yolo v2 and OCR method reddit.com+13ijisae.org+13e3s-conferences.org+13
8. Smart surveillance system for automatic detection of license plate number of motorcyclists without Helmet mdpi.com+7ijisae.org+7journals.riverpublishers.com+7
9. Helmet Detection and Number Plate Recognition for Safety and Surveillance System journal.ijresm.com+1e3s-conferences.org+1
10. Helmet Detection and License Plate Recognition using CNN turcomat.org+1ijisae.org+1
11. Automatic Detection of Helmet and License Plate Recognition using CNN & GAN arxiv.org+3sersc.org+3ijisae.org+3
12. Fast Helmet and License Plate Detection Based on Lightweight YOLOv5 dl.acm.org+2mdpi.com+2arxiv.org+2

13. Enhancing Helmet Violation Detection and License Plate Recognition through Optimization of YOLOV8 Models with Edge Computing Integration itm-conferences.org+15dl.acm.org+15ijisae.org+15
14. Real-Time Helmet Violation Detection Using YOLOv5 and Ensemble Learning arxiv.org
15. An automatic detection of helmeted and non-helmeted motorcyclist with license plate extraction using convolutional neural network discovery.researcher.life+4e3s-conferences.org+4journals.riverpublishers.com+4
16. Motorcyclist's Helmet Wearing Detection Using Image Processing e3s-conferences.org
17. Automatic detection of bike-riders without helmet using surveillance videos in real-time journal.ijresm.com+2e3s-conferences.org+2sersc.org+2
18. Automatic helmet detection on public roads e3s-conferences.org
19. Safety helmet wearing detection based on image processing and machine learning e3s-conferences.org+1sciencescholar.us+1
20. Improved OCR based automatic vehicle number plate recognition using features trained neural network e3s-conferences.org+1sciencescholar.us+1
21. Single line license plate detection using opencv and tesseract e3s-conferences.org
22. Hybrid Approach for Detecting the Traffic Violations Based on Deep Learning Using the Real-Time Data dl.acm.org
23. Machine Learning Based Approach for Traffic Rule Violation Detection dl.acm.org+1sciencescholar.us+1
24. YOLOV5 Based A Real Time Automatic Number Plate And Helmet Recognition System dl.acm.org
25. YOLOv8-Based Helmet and Vest Detection System for Safety Assessment turcomat.org+15dl.acm.org+15ijisae.org+15
26. Smart Cloud-Edge Video Surveillance System dl.acm.org+1sersc.org+1
27. An Intelligent Video Surveillance System using Edge Computing based Deep Learning Model dl.acm.org+1arxiv.org+1
28. Helmet use detection of tracked motorcycles using CNN-Based Multi-Task Learning itm-conferences.org+4sciencescholar.us+4ijisae.org+4
29. Helmet Use Detection of Tracked Motorcycles Using CNN-Based Multi-Task Learning sciencescholar.us
30. A Survey on Helmet Detection by CNN Algorithm arxiv.org+3itm-conferences.org+3sciencedirect.com+3