**Advances and Applications in Mathematical Sciences**
Volume 18, Issue 8, June 2019, Pages 741-751


# OBJECT DETECTION AND RECOGNITION

**SHAKTI ARORA, HARISH SAINI and ASHISH KUMAR**

Panipat Institute of Engineering and Technology
Samalkha, India
E-mail: Shakti.nagpal@gmail.com
harish.cse@piet.co.in
ashishghn282@gmail.com

## Abstract

Object detection and recognition is a part of computer vision technology. In the field of computer vision object detection and recognition is considered to be one of difficult and challenging task. Many algorithms, models and approaches were proposed with the time, all these algorithms, models and approaches were compared with each other and those were adopted that gives us the accurate and efficient results. Like Exhaustive Search, ESS algorithm, selective search, state of art recognition approaches came with time and the approaches that were introduced in state of art recognition approaches came out to be the best approaches like Faster RCNN. This paper presents a review of various techniques that are used in detection and recognition of objects in various images and done a comparative analysis of all these techniques and proposed the optimal solution for object detection.

## Introduction

Object detection and recognition deals with image processing related to computer vision and handles the detection of instances of objects belonging to certain classes like cars, two wheelers, animals, houses, persons, birds etc. Researchers include face detection and pedestrian detection as object detection.

**Applications of object detection:**

There are many applications of object detection which are very interesting and very useful in today's world. Some of them are as follows:

- Manufacturing industry: object detection can be used in manufacturing

2010 Mathematics Subject Classification: 94A13.
Keywords: Regression, Bounding box, Tensor Flow.
Received February 23, 2019; Accepted March 17, 2019

industries like you want your machine to detect the objects of any kind of shape like square, rectangle, circular etc.

- Video surveillance: object detection can be used for video surveillance. It is a system of monitoring activity in an area or building using a TV. System in which signals are transmitted from a TV camera to the receivers through cables etc.

- Vehicle detection: vehicle detection is also an application of object detection such as object detection can be used in estimating the speed of object and it can also be used to detect the type of ship entering the port.

- Face detection: This application deals with the process of detecting the face of any person. E.g. facebook recognizes your face when you upload any photo on facebook.

- Self driving cars: object detection finds its great application in self driving cars like they can accurately detect vehicles, streets, buildings, pedestrian and road signs easily.

People counting: object detection can be used for counting people counting. This is used for crowd statistics during festivals etc. In digital images and video, every object class has its own special features that help in classifying the class like all circles are round in shape and all circles have perpendicular corners to detect any object.

Object detection is widely used in computer vision task like it is used in tracking objects e.g. tracking a ball during football match or a cricket ball in a cricket match.

Object detection finds its application in many fields like it is used in auto driving cars to detect streets, buildings, road signs and pedestrians so that the car driven safely automatically. And it can be used in face detection also in which the system recognizes the face of a person like you upload a photo on Facebook and it recognizes your photo and also face unlock technique is also based on object detection.

In this case an image will be given as the input and a bounding box along with class of each object in the image will be the output.

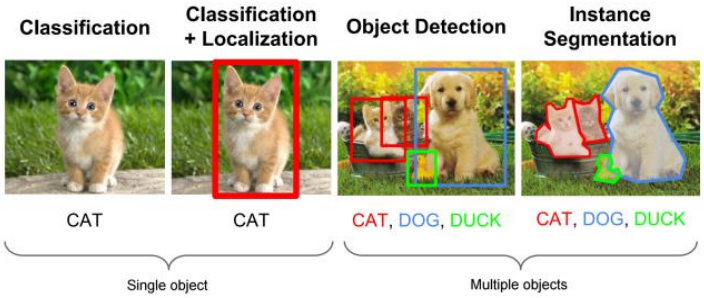


**Figure 1.1.** Object Detection with different instances.

**Problem which can be analyzed:**

Object detection problem: search for the position of the object of interest in the image. It detects all objects in image.

Image classification problem: the whole image is evaluated against a specific type such that best guesses for the image can be evaluated.

Object localization: detect the number of the target objects.

**Literature Review**

• Dilip Kumar Parsad et al. (2010) worked on object detection in real images, in that they proposed a new object detection/ recognition process, in addition to the usual features they proposed to use the geometrical shapes like linear cues, ellipses and quadrilaterals as additional features. They also proposed a new hierarchical codebook that provides good generalization and discriminative property.

• Juan Wu, et al. (2012) worked in the object detection and recognition field and in that they include the description of the existing object detection methods associated with color or shape, important methods developed to tackle the object classification problem and comparison of performance of these techniques.

• J. R. R Uijilings, et al. (2012) submitted their technical report to IJCV and in that paper they the problem of generating possible object location for use in object recognition was addressed. They introduced selective search which combined the strength of both exhaustive search and segmentation. Like segmentation they used the image structure to guide their sampling

process. Like exhaustive search, they aim to capture all possible object locations.

• Dilip K Parsad et al. (2012) had made a review on the different aspects of the object detection and in his paper he included feature selection, learning model, object representation, matching features and object templates and the boosting schemes. He also discussed the merits and demerits of the state of the art research. His main aim was to provide a summary of the state of art techniques in object detection.

• Rodrigo Versace et al. (2015) solar worked in the current and future directions of object detection, they divided the object detection methods in five categories with their merits and demerits and the five categories are coarse to fine boosted classifiers, dictionary based, deformable part based model, deep learning and triangle image processing architectures.

• Divya Patel et al. (2015) worked in this field and they worked on different techniques and methods for detecting and recognizing objects. They worked on object detection techniques like multi-component object detection, multi-class though transform approach, latent though transform and boosted haar cascade technique etc.

• Sukanya C. M., et al. (2016) worked in this field i.e. the field of object detection and recognition, in that paper they presented a survey of different techniques in the field of computer vision and object recognition. In that paper they mainly gave a review and study of different methods of object detection. They discussed background subtraction, optical flow, point detector, frame differencing to detect objects. They also discussed the accuracy and limitation of these methods, and included various approaches that had been used by different researchers for object detection.

• Pratik Kalshetti, et al. (2016) worked in the field of object detection and recognition and their project aimed to incorporate state of the art technique for object detection with the goal of achieving high accuracy with a real time performance. They discussed about bounding boxes, (regression + classification), two stage (RCNN, fast RCNN, Faster RCNN) method and unified methods (YOLO, SSD).

• Sandeep Kumar, et al. (2017) worked in the field of object detection and recognition. In that they make prediction with a single network evaluation.

In that object detection was a regression problem to spatially separated bounding boxes and associated classes probabilities.

• Kartik Umesh Sharma et al. (2017) made a review and an approach in object detection in images. In that paper they divided the process of object detection into five major categories namely sliding window based, contour based, graph based, fuzzy based and context based. Apart from this they also summarized other approaches like shape based detection and steiner tree based approach.

• Zhong qio et al. (2017) worked in object detection with deep learning. In that paper they provide a review on deep learning based object etection frameworks, their review begun with the history of deep learning and its representative toll CNN. Then they focused on typical generic object detection architectures along with some modifications and useful tricks to improve detection performance further, they also briefly surveyed several specific tasks including salient object detection, face detection and pedestrian detection.

• Mukesh Tiwari et al. (2017) presented their paper in which they provided a review on different object detection, tackling, recognition technique, feature descriptor and segmentation method which is based in video frames and various tackling technologies. They have identified and discussed the limitation/ future scope.

• Anupriya George et al. (2017) studied object detection algorithms for computer vision systems and in their paper they presented a survey of the same. They started by introducing the concepts that were behin object recognition on computer vision i.e. how the object is recognized using matched patterns, how the input image is being featured and the major components of the object recognition system. They discussed that how object recognition can be successful for a visible change in the environment.

**Review analysis**

As it can be seen that there are different algorithm and strategies had been developed in the field of object detection and recognition, like exhaustive search, an algorithm which was not used in practical life because of the reason that if we have a $N \times N$ image then if the value of $N$ is increased then

it is impossible to implement it in real life. Then ESS model came which solved the problem of exhaust search, in that branch and bound technique was proposed, this model had complexity of $O(N^2)$ and the worst case of this model was $O(N^4)$. Then an update of this algorithm was proposed which reduced the worst case complexity to $O(N^3)$, and that could be implemented in real life, but that model was not used. Then selective search was introduced which depend upon certain steps that are preprocessed to guess small set of locations where object is there.

After all these algorithms and approaches state of the art recognition approach was introduced as deep learning was increasing at that time and that was classified into two categories (i) two stage detection and (ii) unified detection. The concepts that were used into that were bounding box and regression + classification etc. so these are the approaches and algorithms that were developed with the time.

Faster RCNN is one the best state of the art approaches. It is very fast and accurate. The code is available and you can tune to specific datasets.

**Existing Algorithms:**

A lot of work h in the field of been done by researchers object detection and recognition from earlier times also, but at that time they lack the accuracy because at that time machine learning was not developed upto that extent, Machine learning lack with its techniques at that time.

With the passage of time new approaches, algorithms and techniques were developed so that object detection and recognition can be improved. Some of them are as follows:

**Exhaustive search:**

The best ever approach will be applying a brute force. Promising rectangles are needed. Output is all the boxes in rectangle. Our output is all the boxes of given image.

In $N \times N$ image we will have $N^4$ and if the value of $N$ is increased to a large number then it. Because of this reason this is impossible to implement this in real life. This approach is not in use. If we want to apply brute force algorithm to some specific class of problems then we must implement four procedures: first, next, valid, output. These procedures should take as a parameter $P$ for particular instance to solve the problem:

First generates the first candidate solution for $S$, next generate the next candidate for $S$ after current one, valid check that whether the candidate is a solution for $S$, output generate the solution as appropriate to the application.

d← first(S)

while d ≠ Λ do

if valid(S,d) then output(S,d)

d ← next(S,d)

end while

Where $S$ represents the parameter and $d$ represents the current solution for problem.

**ESS Algorithm:**

To solve the problems of Exhaustive search in an efficient manner Branch and Bound technique was proposed. To identify the best windows in order they combined recognition with detection. The complexity of this algorithm is $O(N^2)$, but its worst case is $O(N^4)$. The worst case happens when object that is to be identified does not exist in the image. Later, an update of this algorithm called IESS was proposed that reduced the worst case to $O(N^3)$. This is very fast in real time. In some cases IESS can be used to efficiency purpose. This approach is not in use.

**Selective search:**

This type of approaches is dominant approaches now. These approaches depend upon certain steps that are preprocessed to guess small set of locations where object is there. This approach is heavily used nowadays.

The processes that are followed in selective search algorithm are:

1. It adds all the bounding boxes that correspond to the segmented part to the list of regional proposals.

2. Then, after the first step, it groups the adjacent segments on the basis of their similarity.

3. Go to step 1, after grouping all the segments it produces the final

candidate region proposals.

In this way it provides us the output.

**State of the art recognition approaches:**

As the concept of deep learning was introduced we moved from extract hand craft feature to learn feature. In deep learning the approaches are broadly classified into two categories that are used:

• First one is Two stage detection and it includes RCNN, Fast RCNN, Faster RCNN

• Second one is Unified detection and it includes YOLO, SSD

**The concepts that are used in these techniques are under:**

**Bounding box:**

A rectangle is drawn on the image that tightly fits it's objects. The rectangle that exists is for every instance of every object that is present in the image.

**Classification + Regression:**

Regression is used to predict the bounding of the image and classification is used to predict the class within the bounding box.

**Two Stage Method:**

Object Proposal is extracted using different computer vision techniques and further resized according to input classification network, which is called feature extractor. State vector machine is trained properly to extract the features of object and background of the object, *A* trained bounding box regressor provide the corrective proposal boxes as output. Specified technique is computationally very intensive but provides very accurate results

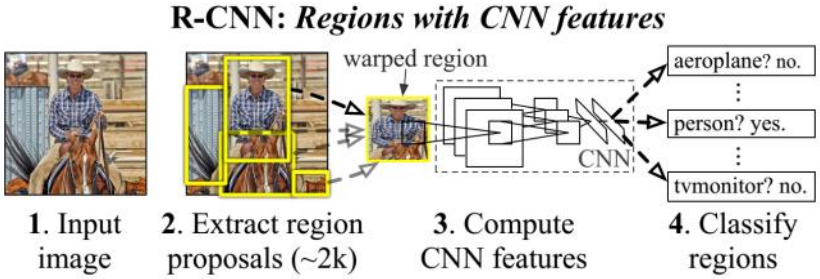

**Figure 1.2.** CNN feature Extraction.

**Unified method:**

Predefined set of boxes are defined to look for a set of objects. To generate and predict the scores of class and the off set values of bounding box, the convolutional feature map from lower layers of network is used, and another network over the lower feature map is implemented to extract more features. Following steps of unified methods are:

1. A CNN is trained with regression and classification objective.

2. Activation is gathered from the later layers so that classification and location can be inferred with fully connected and convolution layers.

3. Jaccard distance is used during training so that prediction and growth truth can relate.

4. Non maxima suppression is used during inference so that multiple boxes around the same object could be filtered.

The major techniques that follow this strategy are:

• SSD: For the prediction of bounding boxes and classes SSD uses different activation maps.

• YOLO: For the prediction of bounding boxes and classes YOLO uses single activation map.

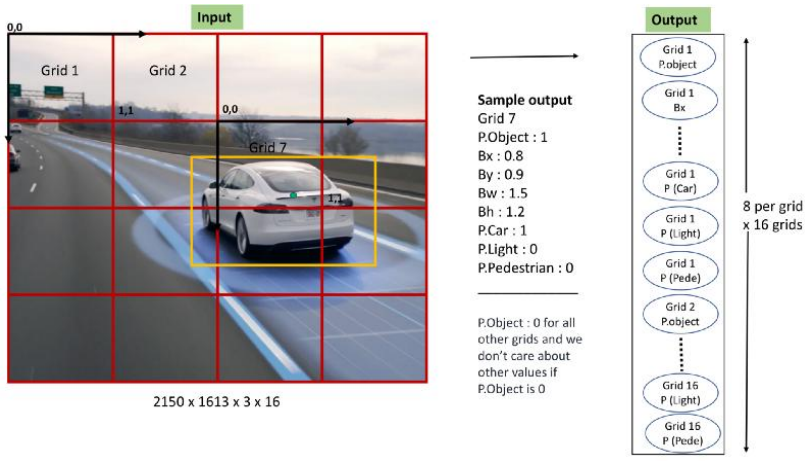


**Figure 1.3.** gridding calculation of distant objects.

**Comparative Analysis:**

**Table 1.** Comparative analysis.

| Algorithm/ Approach | Methodology | Advantage | Limitation |
|---|---|---|---|
| Exhaustive Search | It applies brute force and promising rectangles are needed | | It cannot be used practically. |
| ESS Algorithm | It combined recognition with detection | | It cannot be used in practically. |
| Selective Search | It depend upon certain steps that are preprocessed to guess small set of locations where object is there | It is heavily used. | |
| State of art recognition | Deep learning | Provides accurate and efficient result. | |

**Proposed Approach:**

As we identified that selective search and sliding window techniques are slower and less accurate whereas RCNN is providing the accuracy of objects. The proposed approach is based on the Tensor Flow. The classification of data sets is increased from thousands to million and cost of tagging and maintenance of the classification is reduced. A hierarchal model is adopted to maintain the object classification. A trained Labeled detection mechanism will be designed to train the different objects as well as increase the localization of the objects.

## Conclusion

With the passage of time there are many developments in the field of object detection and recognition and many algorithms and approaches were

proposed and improvements were made in that so that accurate and efficient result can be made. The approaches continued from earlier times and now they include the concept of deep learning to improve the results in the object detection. The fastest algorithm is working with object detection is generating output in 0.2 seconds. Faster RCNN gives higher accuracy. An optimized computation is achieved with SSD. A super-fast approach which is not fast is YOLO.

And the work is still going in this field so that it can be improved upto a great extent.

## References

[1] P. P. Bonissone, Soft computing: the convergence of emerging reasoning technologies, Springer-Verlag 1: 6-18. Object Detection and Classification, in International Conference on Computer and Computing Technologies in Agriculture(ccta-2012), vol -392,ppno-183-188

[2] Khushboo Khurana and Reetu Awasthi, Techniques for Object Recognition in Images and Multi-Object Detection, in International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 2, Issue 4, PP. No. 1383-1388 April 2013.

[3] Xiaofeng Ning, Wen Zhu and Shifeng Chen, Recognition, object detection and segmentation of white background photos based on deep learning, in Youth Academic Annual Conference of Chinese Association of Automation (YAC), china, 2017.

[4] Marcin Woźniak and Dawid Połap, Object detection and recognition via clustered features, in peel defects detection, Neural Networks, Volume 98 (2018), 16-33.

[5] Prince Grover, Evolution of Object Detection and Localization Algorithms.

[6] Katherine Bouman, Image Processing Technology and Applications: Object Recognition and Detection in Natural Images in IEEE signal processing society lecture notes ,nov,2012

[7] Liming Wang, Jianbo Shi and Gang Song and I-fan Shen Object Detection Combining Recognition and Segmentation, In Proc. ACCV, 2007.

[8] S. Belongie, J. Malik and J. Puzicha, Shape Matching and Object Recognition Using Shape Contexts, IEEE Transactions on Pattern Analysis & Machine Intelligence, April 2002, pp. 509-522, vol. 24.

[9] Kartik Umesh Sharma, Nileshsingh V. Thakur, A review and approach for object detection in images, In International Journal of Computational Vision and Robotics, vol-7 ppno-196, 2017

[10] Liming Wang, Jianbo Shi, Gang Song and I-fan Shen, Object Detection Combining Recognition and Segmentation.