American Journal of Electronics & Communication

Sudipto Kumar Mondal et. al., American Journal of Electronics & Communication, Vol. II (2), 1-4

# Real-time object detection comparative study

Sudipto Kumar Mondal, Sanhita Dey, Soumyajit Dey

Dept. of CSE, UEM Kolkata

**ABSTRACT:-**

Object detection is a computer technology related to computer vi-sion and image processing that deals with detecting instances of semantic objects of a certain class (such as humans, buildings, or cars) in digital images and videos. In this domain of object detection, it includes many detecting methods , such as face detection and also pedestrian detection. Object detection has applications in huge areas of computer vision, which includes image retrieval, video surveillance.Deep Neural methods in object detection using one-stage processes generally include di erent versions of YOLO and SSD. The paper which we are publishing here we are compar-ing some the image detection algorithms. In this project we are going to develop a system for visually impaired people for assisting them in their daily work and give them a free life.

## 1 INTRODUCTION:-

Object detection is a computer technology of computer vision.For achieving a good understanding of image, it needs to recognize categories of object instances, as in object classi cation, but also need to locate them. This task is commonly called object detection. Object detection is a much harder task than merely classifying images into distinct classes, and there will always be a trade-o between accuracy and e ciency when developing object detecting networks. We also need to notice the accelerating complexity and hurdles which has come from adding additional classes for detecting objects. So for solving this problem, many algorithms are created but the most eminent, powerful, and accurate algorithms approached over the time are Region based Convolutional Neural Networks that is known as RCNN , You Only Look Once which is known as YOLO, SSD and RetinaNet.For extracting features and performing the object detection on the regions, the approach of the solution take many regions of those frames. There are many advantages of neural networks methodology.They have the ability to learn by themselves and produce the result which is not limited to the input provided. Very general representations of objects are generally learnt by YOLO. It executes all other detection methods which also includes DPM and R-CNN, by a wide margin when deducting from natural images to artwork on both of the Picasso Dataset and the People-Art Dataset.

## 2 LITERATURE SURVEY

1.RCNN: RCNN is the abbreviation of Regional Based Convolutional Neural Networks.The rst stage of RCNN pipeline is the generation of region proposals in an image which could belong to any particular image.Many research papers o er di erent models to generate category-independent region proposals like objectness , selective search, category independent object proposals etc., but selective search is widely used because it enables controlled observations with other detection task.

2.Fast RCNN: Fast R-CNN builds on previous work to e ciently classify object proposals using deep convolutional networks.If we compare the previous work then Fast R-CNN employs several innovations for upgrading training and testing speed and also gives a rising graph in accurate detection. For each test region of interest or ROI, which extracts a featured vector from featured map which is provided as input to nal completely connected layers and outputs a class which contains the probability distribution and the set of predicted boundaries.

3.Faster RCNN: Faster R-CNN is a single-stage model that is trained end-to-end. It uses a RPN that is known as novel region network to generate region

proposals, which saves time compared to other algorithms like Selective Search. It uses the Region of Interest Pooling layer for extracting a contant length feature vector from each region proposal.It uses the convolutional network and share it with Object detection network.It deducts the proposed frame to 300 from 2000, So performance is also raised.

4.YOLO: YOLO stands for You Only Look Once. YOLO is developed in the year 2016.YOLO is a real-time object detection system which applies a single neural network to the full input image, hence the name You only look once.The bounding box and class predictions are made after an evaluation of the image given as input. The fastest architecture of YOLO can get 45 FPS and a smaller version of YOLO that is Tiny-YOLO which achieves up to 244 FPS (Tiny YOLOv2 or Tiny YOLO Version 2) on a computer with a GPU. When Tiny-YOLOv2 runs on a non-GPU laptop (Dell XPS13), the model speed reduces drastically from 244 FPS to near about 2.4FPS. With this shrinking, real-time object detection is not easily accessible on some nonGPU devices,like most of the cellphones or laptops.

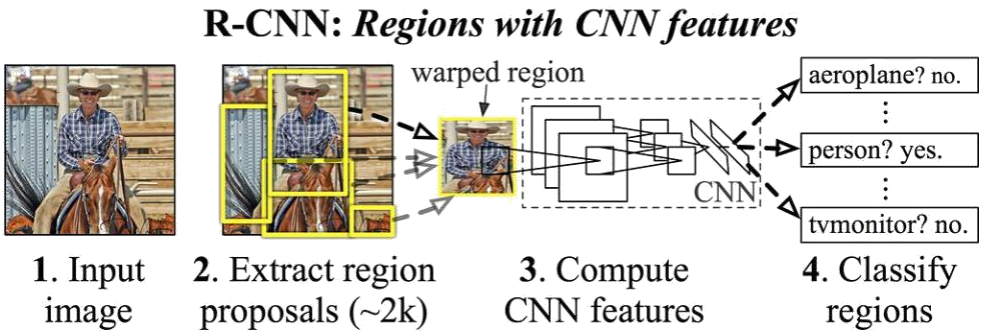


Figure 1: OBJECT DETECTION OVERVIEW WITH RCNN

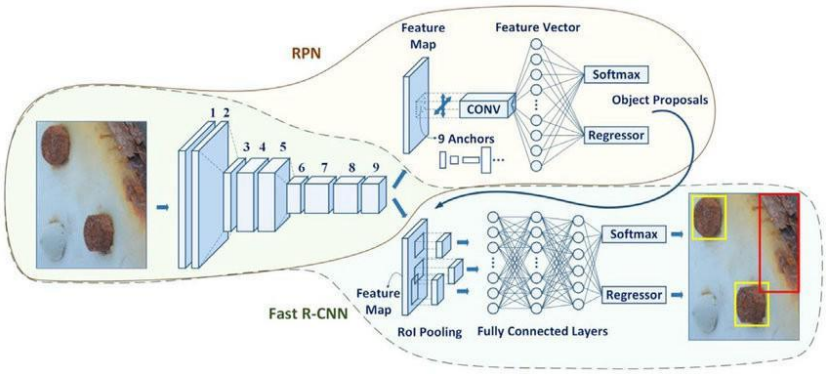


Figure 2: OBJECT DETECTION OVERVIEW WITH FAST RCNN

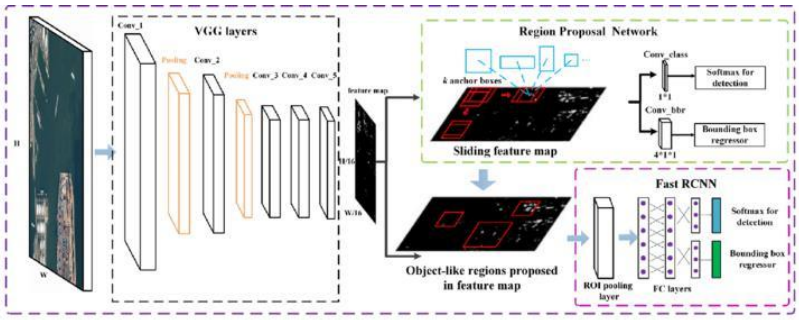


Figure 3: OBJECT DETECTION OVERVIEW WITH FASTER RCNN

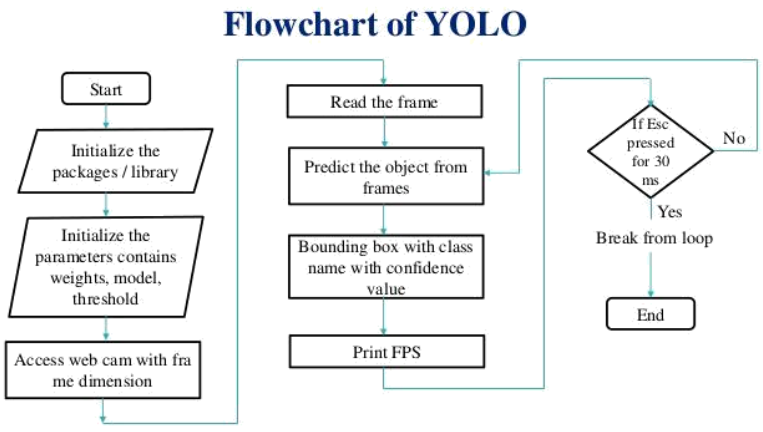


Figure 4: YOLO STRUCTURE

## 3 RESULT ANALYSIS:-

The result comes as the system is based on some of the parameters of the models. The common features are performance, resources, time taken which are needed and accuracy, which when are taken into consideration in almost all the analysis.When they are applied the general feature over the R-CNN method of the object detection proposed by the Ross Girshick, the results show that it is much faster than the old methods based on the classi cation methods. A new method You Only Look Once (YOLO) is proposed for the recognition of the objects. Previous methods use proposed regions to identify the object in the image, it actually never considers the full image.Rather the regions with high probability of having the objects are passed in the system for the object detection. But in YOLO, it has only one Convolutional network and the whole image is analysed by this network. It divides the image into SxS grid and take m bounding boxes. For each box the network outputs a class probability and the classes with chance higher than the threshold value are used to locate the object. This way it has many advantage over to its single convolutional neural network. First of all, it predicts bounding boxes and class probability directly from the whole image in one evaluation. Secondly, the whole detection process is done in a single network; hence it is easy to optimise the network. It is much faster than the RCNN, Fast RCNN and Faster RCNN as it has only one convolutional neural network.

| Model | Latency | Test Time per image (second) | mAP | FPS | RealTime |
|---|---|---|---|---|---|
| R-CNN | High | 50 | ~60 | < 1 | No |
| Fast R-CNN | Medium | 2 | ~70 | < 1 | No |
| Faster R-CNN | Medium | 0.2 | ~70 | 7 | No |
| Yolo | Low | 0.025 | ~60 | 46 | Yes |

Table1: Comparison of di erent models(R-CNN, Fast-RCNN Faster-RCNN and YOLO)

## 4 CONCLUSION

Comparing and looking at all the parameter and results by test and compar-ison and considering speed and accuracy which are most important parameters in object detection application as speed is required to get the result faster and to accuracy is to get the result accurately and it is necessary to have less computation time and also process the input fast to provide the result to the user. By our result YOLO(YOU ONLY LOOK ONCE) is the best compared to R-CNN,Fast R-CNN and Faster R-CNN also YOLO only comes with one convolution network and with its fast computation speed it can deliver result quickly and accurately to the user compared to the other methods and the accuracy of the method can be managed as per the requirement of the system.

## 5 REFERENCE

1. Ross Girshick, Je Donahue, Trevor Darrell, Jitendra Malik, UC Berke-ley, "Rich feature hierarchies for accurate object detection and semantic segmentationTech report (v5)", 22 OCT 2014.

2. Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi, "You Only Look Once: Uni ed, Real-Time Object Detection", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

3. Joseph Redmon, Ali Farhadi,"YOLO9000: Better, Faster, Stronger", IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017

4. Joseph Redmon, Ali Farhadi,"Yolov3: An incremental improvement", arXiv preprint arXiv:1804.02767, 2018