Real Time Object Detection Using TensorFlow

Veena A[1], Vasanthamma H[2] , Heenakousar M Makandar[3], Uppar Gangadevi[4], Maitra k[5]
[1]Assistant Professor, CSE Department, PDIT, Hosapete,.
[2]Head of Department, CSE, Department, PDIT, Hosapete,
[3]Student of CSE, BE PDIT College, Hosapete.
[4]Student of CSE, BE PDIT College, Hosapete.
[5]Student of CSE, BE PDIT College,

Abstract: *In the progress of computer vision systems, Real time object detection has become an main topic in development. Creating the specific machine learning model which have the ability to recognizing and localizing many objects in a single image has become a main challenge in a Computer vision. By using advanced deep learning, object detection has become more easier than before.*

The aim of project is to detect the object in real time with high accuracy by using Tensorflow. Tensorflow object detection API is the Google open source machine learning frame work that is easy to build, train and place the object detection.

The object is trained using accessible data set. It is used to detect the object in videos as well as specific images. The result is fast and precise, thus helps those applications which demand object detection.

Keywords: Machine learning, deep learning, Tensorflow, coco model.

Introduction: Object Detection is common term used for computer vision techniques for locating and labelling the objects. This can be applied for both static and dynamic images. Object detection is difficult task in computer vision. Detecting an object is major problem because there are many variations in inclination, lighting, position, and background in an image of very similar objects that makes an result uncompleted. Now with the help of Deep learning and neural network, we can finally outfit such problems without approaching various trial and error methods.

In this case we are using open source library i.e Tensorflow Object Detection API which is easy to use and Google Open source framework for machine learning where pre-trained models for detecting objects reached highest accuracy in detection of objects. TensorFlow Object Detection API is taken due to its demand in object detection and its easy to approach. The input to the system will be a image, and the output will be a bounding box correlating to all the objects in the image, along with the class of object in each box.

Related Work:

There has been many research on object detection using conventional computer vision methods. However, it show the deficiency of deep learning based techniques such as Bounding Box, Classification + Regression, Unified Method, SSD approach.

1. Bounding Box

The bounding box is a rectangle drawn on the image which closely fits the object in the image. A bounding box subsists for every occurrences of all objects in the image. In the box, 4 numbers (x, y, width, height) are established. This can be trained using a distance measure between established and flooring truth bounding box. The distance measure is a Jaccard distance which provides convergence above junction between the established and flooring truth boxes as shown in fig 3.

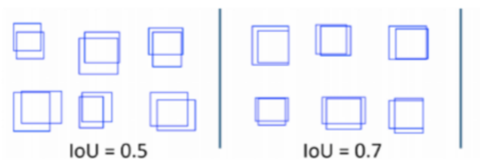


Figure 3: Jaccard distance

2. Classification + Regression:

The bounding box is predicted by using regression and the class within the bounding box is predicted using classification. The architecture is shown in Fig 4

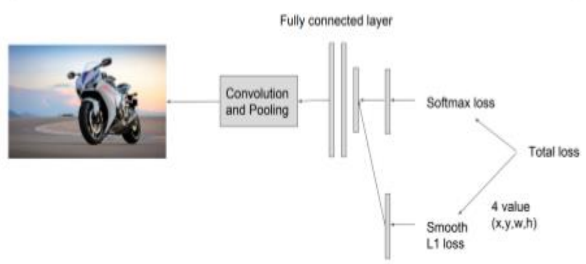


Figure 4: Architecture overview

3.Unified Method

Using complexity feature maps from later layers of the network, used to run another network over these feature maps to analyse class scores and bounding box offsets shown in fig 5. The steps are mentioned below:

1. Train a CNN with regression and classification objective.

2. Get activation from layers to deduce classification and location with a fully connected or complexity layers.

3. While training, use jaccard distance to relate predictions with the flooring truth.

4. During deduction, use non-maxima suppression to sieve multiple boxes around the same object.

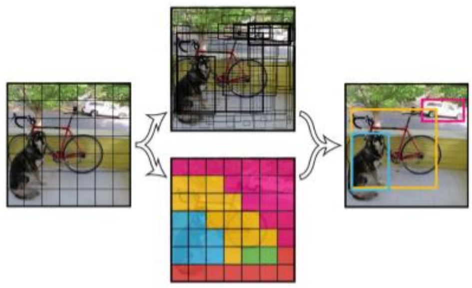
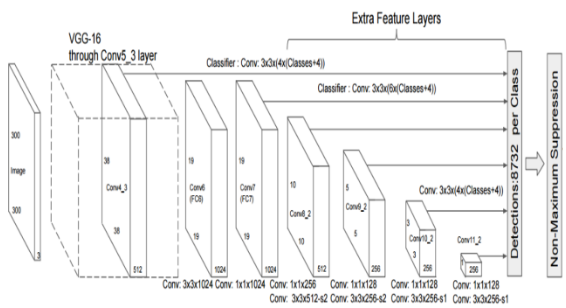


Fig 5:Unified Method

The major technique used: SSD (uses different activation maps for establishing of classes and bounding boxes) Using multiple scales helps to achieve a higher mAP(mean average precision) by being able to detect objects with different sizes on the image better. Thus the technique used in this project is SSD.



All these feature maps are used for predicting bounding boxes at different scales. Thus the overall idea of SSD is shown in Fig. 6. Some of the activations are passed to the sub-network that acts as a classifier and a localizer.

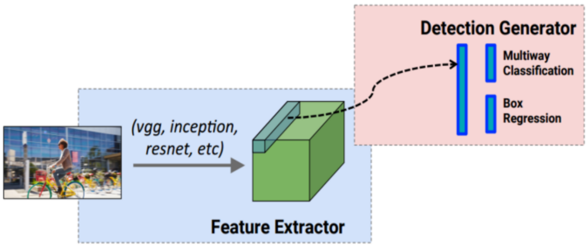


Fig.6: Overall idea of SSD

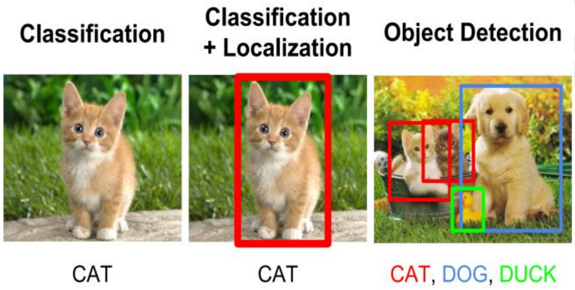
Result:

In today's web log post we learned how to perform real-time object detection using deep learning.

We finished this by combining two separate methods:

1. Object detection with deep learning and OpenCV
2. Efficient, threaded video streams with OpenCV

The end result is a deep learning-based object detection which can process approximately 6-8 FPS depending on the speed of the system.



Conclusion:

An exact and efficient object detection system has been developed. This project uses recent techniques in the field of computer vision and deep learning. Custom dataset was created using labelIing and the evaluation was consistent. This can be used in real-time applications which require object detection for pre-processing in their pipeline. An important scope is to train the system on a video sequence for usage in tracking applications.

References:

[1] Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.

[2] Ross Girshick. Fast R-CNN. In International Conference on Computer Vision (ICCV), 2015.

[3] Shaoqing Ren, Kaiming He, Ross Girshick, and Jian Sun. Faster R-CNN: Towards realtime object detection with region proposal networks. In Advances in Neural Information Processing Systems (NIPS), 2015

[4] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In The IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[5] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, ChengYang Fu, and Alexander C. Berg. SSD: Single shot multibox detector. In ECCV, 2016