Object Detection in Computer Vision

Neelam Somai, Faizan Sayyad, Kona Sireesha, Jaymeet Mehta
Ms. Archana das

Abstract

Probabilistic Object Detection in computer vision helps to locate and identify objects in video or image. It is widely used in computer vision tasks such as face recognition, face detection , autonomous self driving cars and many more. The aim of project is to demonstrate object detection using a single shot multi-box detector (SSD) and you look only once(YOLO) and to improve accuracy of model. The comparison has been developed between these two algorithms in order to understand which algorithm is faster and accurate.

Introduction

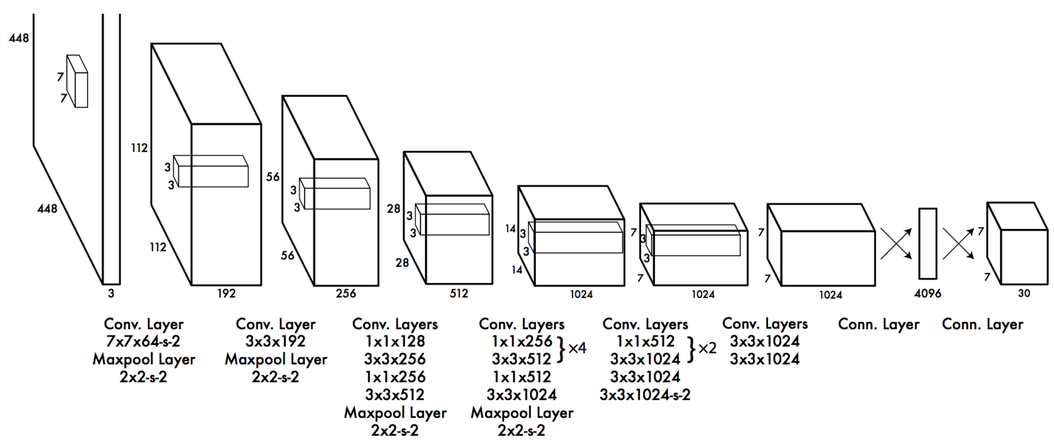
- Humans are very fast and accurate at identifying multiple objects and performing multiple tasks. So, humans can easily train the computers to classify and detect multiple objects within an image with high accuracy by having large amounts of data, faster GPUs, and appropriate algorithms. Humans train the computer for three tasks. They are image classification, Object localization and object detection.
- Here our aim is to detect object from video and audio.
- To do this we can use Deep learning based on computer vision for videos and images
- In this project we have used different technologies to achieve it

Proposed Method

- After some analysis of different technologies, we decided to apply YOLOv3 and SSD for our project

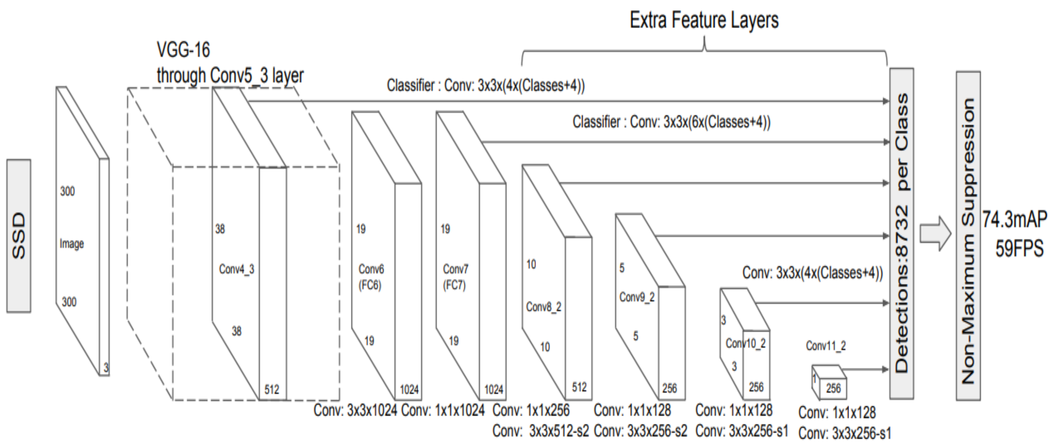
- The dataset used for training the models is Microsoft Common Objects in Context (MS COCO).COCO dataset is an excellent object detection dataset that has 80 different classes, 80,000 training images and 40,000 validation images available.

1. YOLOv3
   - You only look once (YOLO) is an object detection system targeted for real-time processing. YOLO divides the input image into an S×S grid. Each grid cell predicts an object.



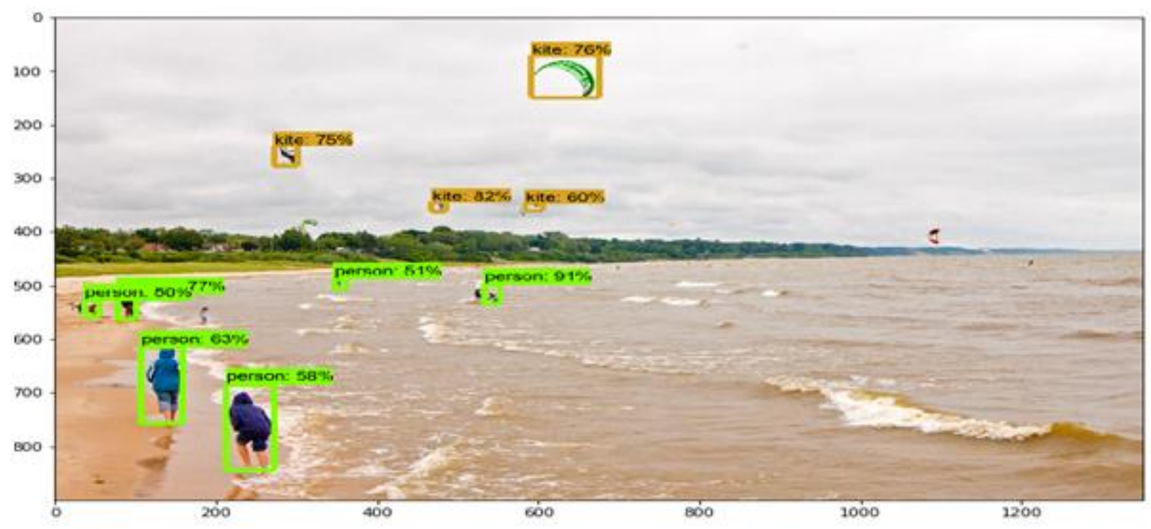
2. SSD
   - SSD is designed for real-time object detection . SSD is faster than Faster R-CNN due to some improvement which includes multi-scale features and default boxes. SSD uses low resolution images which helps to increase the speed further. SSD object detection is made up of two components:
     1. Extracting features maps
     2. 2.Convolution filters to detect objects

     SSD uses VGG16 to extract feature maps.

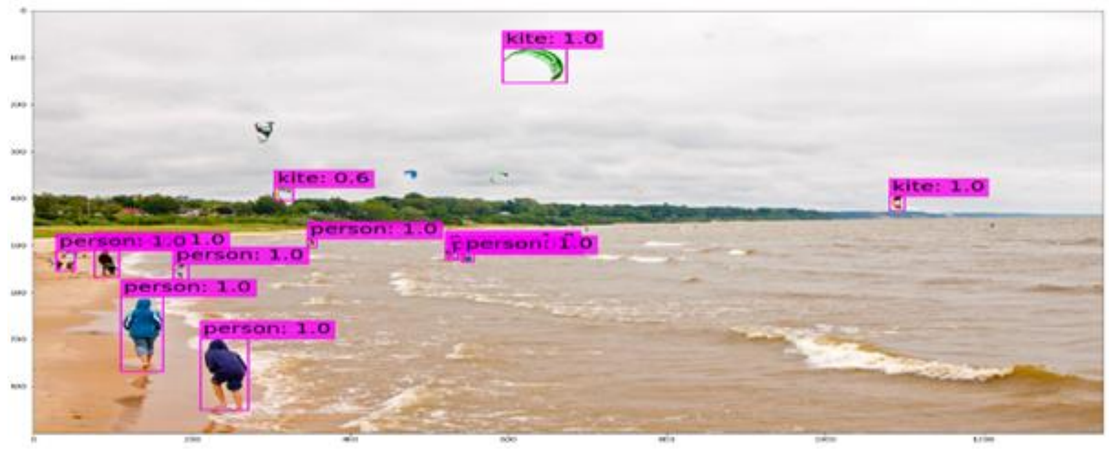


Experimental Results and Discussion

- The results obtained with yolov3 and SSD are shown below:
- YOLOv3 has better results for our test images than SSD. The certainty of objects in the image obtained with YOLOv3 are higher. In terms of speed also YOLOv3 performed better than SSD



Object detection in SSD



Object detection in YOLOv3

Objection detection For Cars and Buses using SSD

Objection detection For Cars and Buses using YOLOv3

- While doing object detection on video we got more FPS(Frames per second) in SSD rather than YOLOv3
- YOLOv3 is slow because of the computation in grids
- SSD is fast in video because of the computation on multi-box

Conclusions

From our project we can conclude that YOLOv3 is more accurate than SSD for images but as it has more computing it is slow for real time which gives low FPS for video where on the other hand SSD is faster than YOLov3 so we can get good fps for real time scenario which is important in current situation

References

- Data references
- [1] Link to the yolo weights and cfg
  https://pjreddie.com/darknet/yolo/
- [2] Link to the detection model zoo
  - https://github.com/tensorflow/models/blob/master/research/object_detection/g3doc/detection_model_zoo.md