Single Shot Multi Box Detectors (SSD)

Friday, December 7, 2018 4:39 PM

What is SSD?
It is a machine learning algorithm used for the computer vision applications.

What is the purpose?
This algorithm is used to draw bounding boxes around the different objects of interest in an image with vector of values representing probabilities of different class of objects of interest in any particular box.

What is the advantage of SSD?
It is one of the real-time object detection algorithms. It can be run on embedded systems with decent computation capabilities in real time

What is the common pipeline of previous approaches?
Previous approaches involve the pipeline of hypothesizing the bounding boxes, resample pixels or features for each box, and apply the high quality classifier. (e.g. Faster R-CNN)

What is the drawback of the previous approaches that led to the approaches like SSD?
The previous approaches are too computationally intensive for embedded systems, even using high-end hardware, too slow for real-time applications. (e.g. Faster R-CNN operates only at 7 frames per second (FPS))

How SSD increases the prediction speed?
Improvement in speed comes from elimination of bounding box proposals and pixel or feature resampling. But this approach is also used by YOLO algorithm.

How it is different in the context of neural network architecture from YOLO ?
SSD discretizes the output space of bounding boxes into a set of default boxes over different aspect ratios and scales per feature (from different depth of convolutional neural network) map location.

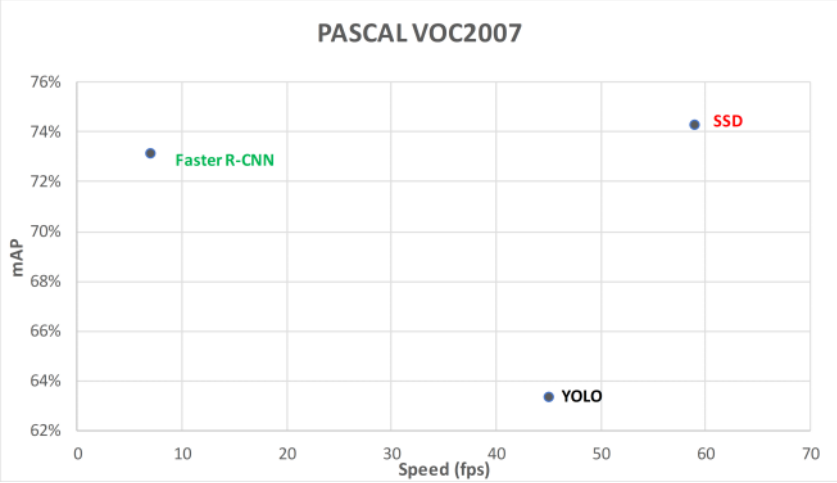
How SSD improves the prediction accuracy when compared to similar approaches?
SSD improves the prediction accuracy by using following tweaks

1. Uses small convolutional filter to predict object categories and offsets in bounding box locations, using separate predictors / filters for different aspect ratio detections.
2. Applying these filters to multiple feature maps from the later stages of a

network in order to perform detections at multiple scales

What is the evidence that shows the performance of SSD is better than YOLO and Faster R-CNN?

All the three neural networks were tested on PASCAL VOC2007 dataset and the mean average precision (mAP) is plotted below. The algorithm which produces the point on the top right region (*High Processing Speed and High Accuracy*) is desirable and SSD lies exactly in that region.



Reference: https://arxiv.org/abs/1512.02325