Deep Learning & Detection on Smooth Objects

Student: Soh See Kiat Desmond   Supervisor: Assoc Prof Qian Kemao

Deep Learning is used in various real world applications for computer vision like healthcare analysis, security & facial detection. Some images in these applications might contain smooth & unusual objects. The motivation is to compare and contrast various neural network architectures performance on detection and localization of these objects for different scenarios. We define smooth objects as one which is unusual and does not contain a lot of features as compared to other more common objects (eg face). We will base our experiment on a fidget spinner which is smooth & unusual in shape.

| SSD (Single Shot Multibox Detector) with MobileNet	FasterRCNN Resnet
Object localization and classification is done in a single forward pass of the network. Multibox is the technique developed by Szegedy for bounding box regression.	FasterRCNN Resnet has 2 networks. One network is to propose regions (Region Proposal Network) while the other network is to use these proposals to detect objects.

Training Phase

During the training phase for SSD MobileNet, the total loss remains relatively stable at around 3620 steps. The value of loss is 0.924 while mean Average Precision value is 0.854.

During the training phase for FasterRCNN Resnet, the total loss remains relatively stable at around 2822 steps. The value of loss is 0.443 while mean Average Precision value is 0.920.
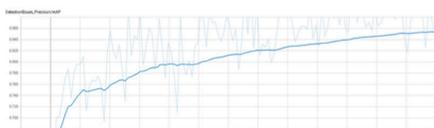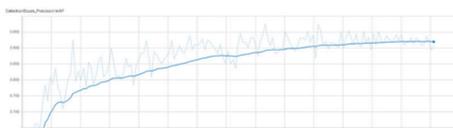

Value	Step	Time	Relative
0.8537	3.620k	Tue Sep 18, 01:40:15	17d 9h 30m 25s

Figure 28 – SSD mean Average Precision over steps
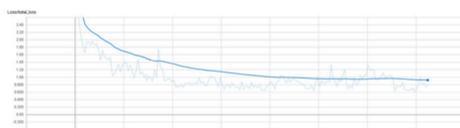
Value	Step	Time	Relative
0.924	3.620k	Tue Sep 18, 01:40:15	17d 9h 30m 25s

Figure 29 – SSD Total Loss over steps

Value	Step	Time	Relative
0.9197	2.822k	Fri Aug 31, 07:52:27	2d 15h 54m

Figure 33 – FasterRCNN mean Average Precision over steps
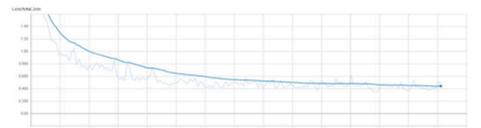
Value	Step	Time	Relative
0.4425	2.822k	Fri Aug 31, 07:52:27	2d 15h 54m

Figure 34 – FasterRCNN Resnet 101 Total Loss over steps

Object Detection

Once training was completed, SSD MobileNet was able to achieve an Intersection Over Union (IOU) value of 0.883 for the following sample image.

SSD IOU where red box is prediction coordinates and green box is ground truth

Similarly, FasterRCNN Resnet was used to detect the object. It is able to achieve an Intersection Over Union (IOU) value of 0.977 for the same sample image.

FasterRCNN IOU where red box is prediction coordinates and green box is ground truth

When it comes to detection time, SSD MobileNet model is able to make predictions on images at a much quicker speed as compared to the FasterRCNN Resnet model. On the other hand, FasterRCNN outperforms SSD MobileNet in terms of confidence level, IOU, Precision, Loss value & Mean Average Precision. A summary of the values collected by running the models on unseen data is shown in the tables on the right.

	Value
Average Confidence	0.91686
Average Detection Time	7.20367
Average IOU (TP - IOU>0.5)	0.83325
Average IOU (TP, FP & FN - IOU any)	0.46689
Precision	0.52381
Total Loss	0.924
Mean Average Precision (mAP)	0.8537

SSD MobileNet on Unseen Data

	Value
Average Confidence	0.916698595
Average Detection Time	43.40775166
Average (TP - IOU>0.5)	0.908067857
Average (TP, FP & FN - IOU any)	0.707880556
Precision	0.823529412
Total Loss	0.4425
Mean Average Precision (mAP)	0.9197

FasterRCNN on Unseen Data