Table 1: Detection accuracy by a different number of training images. VGG16 is used for the base network.

# training images	mAP @0.5 IoU	AVE
5,000	0.540	3.182
10,000	0.892	2.882
15,000	0.900	2.658
20,000	0.884	2.620
25,000	0.887	2.570
30,000	0.887	2.600