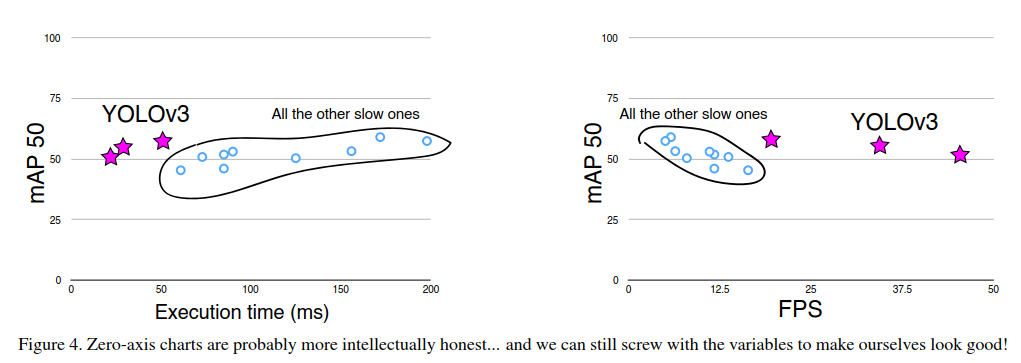
1.YOLOv3达到的**显著指标**：

It achieves 57.9 AP50 in 51 ms on a Titan X, compared to 57.5 AP50 in 198 ms by RetinaNet, similar performance but 3.8× faster



2. During YOLO9000 training we use sum of squared error loss, YOLOv3 predicts an objectness score for each bounding box using logistic regression. If a bounding box prior is not assigned to a ground truth object it incurs no loss for coordinate or class predictions, only objectness.

3.Each box predicts the classes the bounding box may contain using multilabel classiﬁcation. We do not use a softmax as we have found it is unnecessary for good performance,instead we simply use independent logistic classiﬁers. During training we use binary cross-entropy loss for the class predictions

作者任务softmax多分类对好的分类效果没有太大作用，改成了多个独立的logistics回归分类器。进行多分类。

This formulation helps when we move to more complex domains like the Open Images Dataset [7]. In this dataset there are many overlapping labels (i.e. Woman and Person). Using a softmax imposes the assumption that each box has exactly one class which is often not the case

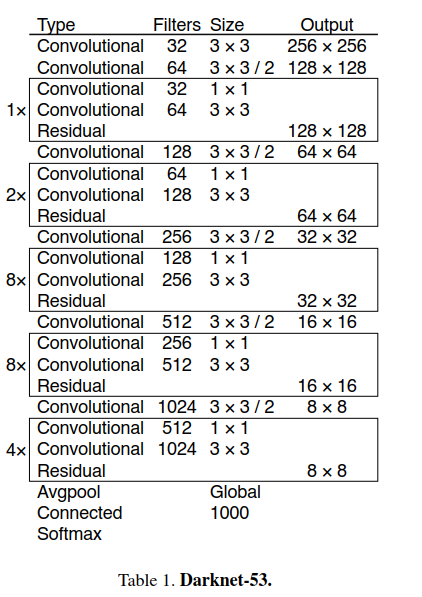
能够使用更多的图像数据集合，并且标签变得可以重叠。

4.关于多尺度目标的问题：

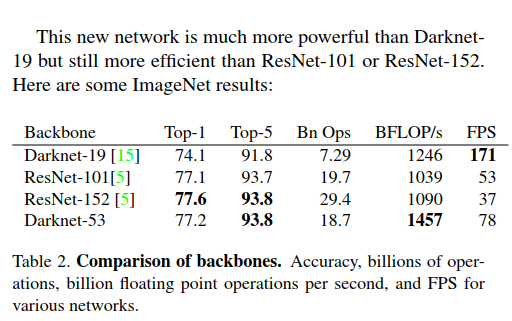
①选用了类似FPN的结构，做了三层预测，YOLOv3 predicts boxes at 3 different scales. Our system extracts features from those scales using a similar concept to feature pyramid networks。

②依然使用k-means聚类的方法选取anchor的比例。We still use k-means clustering to determine our bounding box priors. We just sort of chose 9 clusters and 3 scales arbitrarily and then divide up the clusters evenly across scales. On the COCO dataset the 9 clusters were:(10 × 13), (16 × 30), (33 × 23), (30 × 61), (62 × 45), (59 × 119), (116 × 90), (156 × 198), (373 × 326).

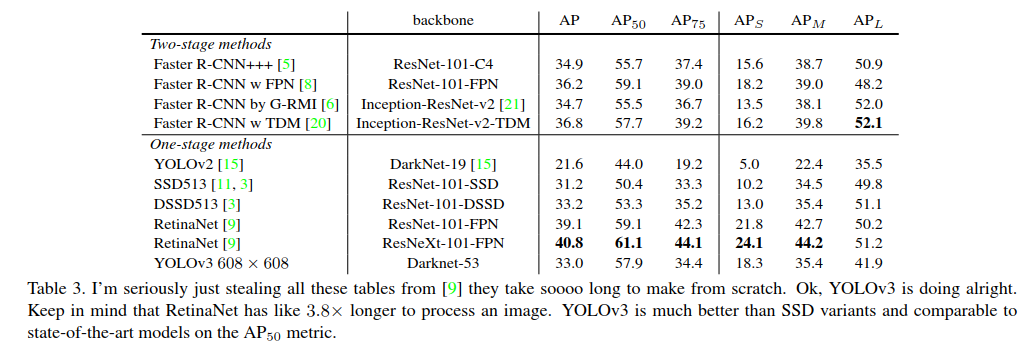
5.新的网络结构：YOLOv2和Darknet-19的组合，融入了最新的残差网络。一共有53层，所以我们取名为Darknet-53。



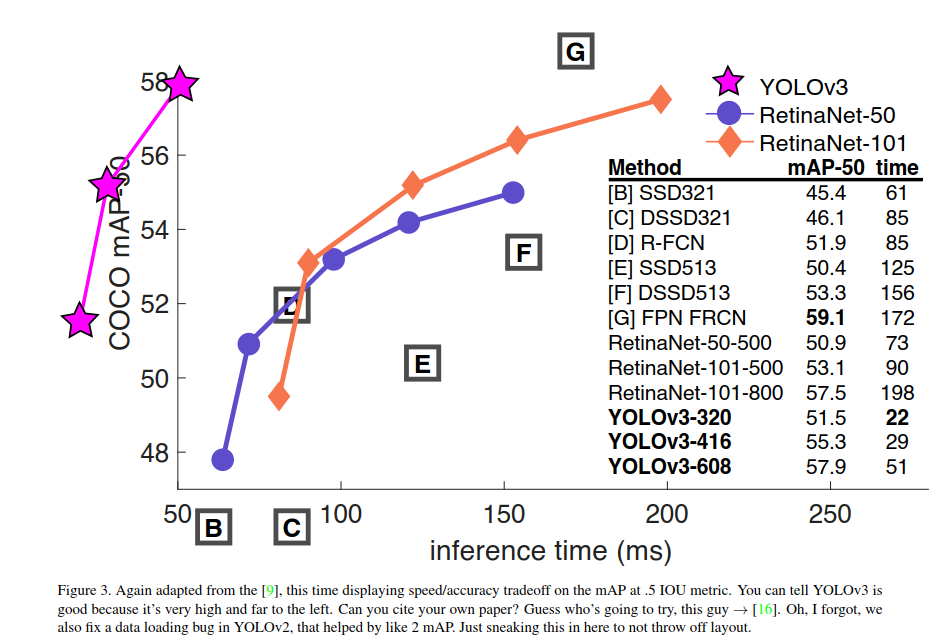
Darknet-53的效果比ResNet-102，ResNet-152要好，如下



与现有方法的比较



mAP 50可以比拟RetinaNet，优势在于，检测速度很快。



Things We Tried That Didn’t Work

