**论文读后感**

我最喜欢的论文题目是《YOLO9000:Better, Faster, Stronger》，这是一篇关于目标检测并且发表在计算机视觉顶级盛会--CVPR 2017会议上的优秀文章。我选择的论文方向为：基于计算机视觉的室内导航系统设计与实现。搜索的关键词为：目标识别与检测、室内定位、室内导航。

目标检测系统要求快速，准确以及能识别大范围种类数量。但是，目前基于深度神经网络方法的目前检测系统能识别的物品种类较少。其原因在于：相对于物品分类数据集，目标检测数据集中的物品种类较少。标记目标识别数据集所耗费的精力远大于标记物品分类数据集。物品分类数据集包含成千上万种超过数百万张图片，而目标识别数据集就很小了。

本文中，作者提出了一种结合不同类型数据集的方法。基于该方法，作者提出了一种新的联合训练方法，结合目前物品分类数据集的优点，将其应用于训练目标检测模型。模型可以从目标检测数据集中学会准确定位目标，同时从物品分类数据集中学会识别更多的种类，增强模型的鲁棒性。

采用该方法，作者训练了可以识别超过9000种物品的实时目标检测与识别系统-YOLO9000。首先，作者在YOLO的基础上进行了改进，产生了YOLOv2（获得了state of the art）。然后，作者采用数据集结合方法和联合训练方法，采用ImageNet和COCO数据集训练该模型，使该模型可以识别和检测 超过9000种类别。下面简单介绍如下：

一、Better-更好

YOLO一代有很多缺点，作者希望改进的方向是:改善recall，提升定位的准确度，同时保持分类的准确度。

目前计算机视觉的趋势是更大更深的网络，更好的性能表现通常依赖于训练更大的网络或者把多种model综合到一起。但是YOLO v2则着力于简化网络。

因为使用了限制让数值变得参数化，也让网络更容易学习、更稳定。 采用聚类方法选择boxes维度和直接预测bounding boxes中心位置提高YOLO将近5%准确率

二、Faster-更快

大部分检测框架是基于VGG-16作为特征提取网络，但是VGG-16比较复杂，耗费计算量大；YOLO完成一次前向过程只用8.52 billion 运算，而VGG-16要30.69billion。YOLO框架使用了类似GoogLeNet的网络结构，计算量比VGG-16小，但YOLO精度只是稍低于VGG-16。

三、Stronger-更强健

作者提出了将分类数据和检测数据综合的联合训练机制。该机制使用目标检测标签的数据训练模型学习定位目标和检测部分类别的目标；再使用分类标签的数据取扩展模型对多类别的识别能力。在训练的过程中，混合目标检测和分类的数据集。当网络接受目标检测的训练数据，反馈网络采用YOLOv2 loss函数；当网络接受分类训练数据，反馈网络只更新部分网络参数。

另外，通过读这篇论文，也使我认识和体会到了以下几点：

1、目标检测的算法主要可以分为两大类：two-stage detector和one-stage detector。前者是指类似Faster RCNN，RFCN这样需要region proposal的检测算法，这类算法可以达到很高的准确率，但是速度较慢。虽然可以通过减少proposal的数量或降低输入图像的分辨率等方式达到提速，但是速度并没有质的提升。后者是指类似YOLO，SSD这样不需要region proposal，直接回归的检测算法，这类算法速度很快，但是准确率不如前者。研究发现正负样本极不均衡是主要原因。在今后的研究工作中，需要进一步解决正负样本极不均衡的问题。

2、我对今后研究方向的想法是：基于计算机视觉的室内导航方向。因为目前室内定位和导航技术有很好的前景，大型商场和停车场内GPS定位不准确，因此室内精确导航的意义就很重要了。

3、读过该论文后，使我不仅了解了目标检测领域的一些知识，而且也学习到了一篇经典论文的脉络结构应该如何组织。这篇论文的结构严谨，层次分明，采用了递进式的分析结构，逻辑性强，文笔流畅，表达清晰，重点突出。文章格式相当的符合学术规范，反映了作者很强的科研能力。

总之，正如一句名言所说：读一本好书就像和一个高尚的人说话。我相信站在巨人的肩膀上才能有更高的成就，我以后要多读书，读好书，不断提高科研水平和自身修养，尽量为中国的科研事业做出自己力所能及的贡献。