**引言**

在深度学习出现之前，传统的目标检测方法大概分为区域选择（滑窗）、特征提取（SIFT、HOG等）、\*\*分类器（SVM、Adaboost等）\*\*三个部分，其主要问题有两方面：一方面滑窗选择策略没有针对性、时间复杂度高，窗口冗余；另一方面手工设计的特征鲁棒性较差。自深度学习出现之后，目标检测取得了巨大的突破，最瞩目的两个方向有：

（1）以RCNN为代表的基于Region Proposal的深度学习目标检测算法（RCNN，SPP-NET，Fast-RCNN，Faster-RCNN等），它们是two-stage的，需要先使用启发式方法（selective search）或者CNN网络（RPN）产生Region Proposal，然后再在Region Proposal上做分类与回归。

（2）以YOLO为代表的基于回归方法的深度学习目标检测算法（YOLO，SSD等）,其仅仅使用一个CNN网络直接预测不同目标的类别与位置。

**什么是传统目标检测？**

首先我们先来了解一下什么是目标检测？简单来说就是把存在的目标从图片中找到并识别出来。我们发现这对于我们人来说十分简单，但对于计算机而言，它是怎么做到的呢？

传统目标检测方法分为三部分：区域选择 → 特征提取 → 分类器

即首先在给定的图像上选择一些候选的区域，然后对这些区域提取特征，最后使用训练的分类器进行分类。下面我们对这三个阶段分别进行介绍。

区域选择

这一步是为了对目标的位置进行定位。由于目标可能出现在图像的任何位置，而且目标的大小、长宽比例也不确定，所以最初采用滑动窗口的策略对整幅图像进行遍历，而且需要设置不同的尺度，不同的长宽比。这种穷举的策略虽然包含了目标所有可能出现的位置，但是缺点也是显而易见的：时间复杂度太高，产生冗余窗口太多，这也严重影响后续特征提取和分类的速度和性能。（实际上由于受到时间复杂度的问题，滑动窗口的长宽比一般都是固定的设置几个，所以对于长宽比浮动较大的多类别目标检测，即便是滑动窗口遍历也不能得到很好的区域）

特征提取

由于目标的形态多样性，光照变化多样性，背景多样性等因素使得设计一个鲁棒的特征并不是那么容易。然而提取特征的好坏直接影响到分类的准确性。（这个阶段常用的特征有SIFT、HOG等）

分类器

主要有SVM，Adaboost等。

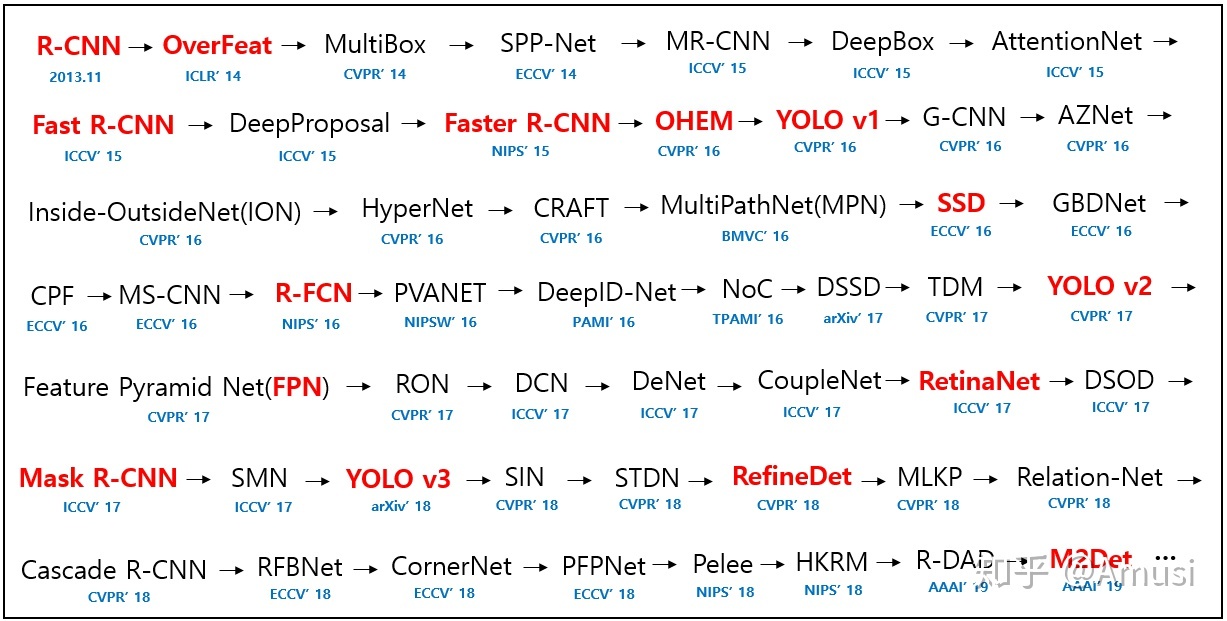
传统目标检测方法不足

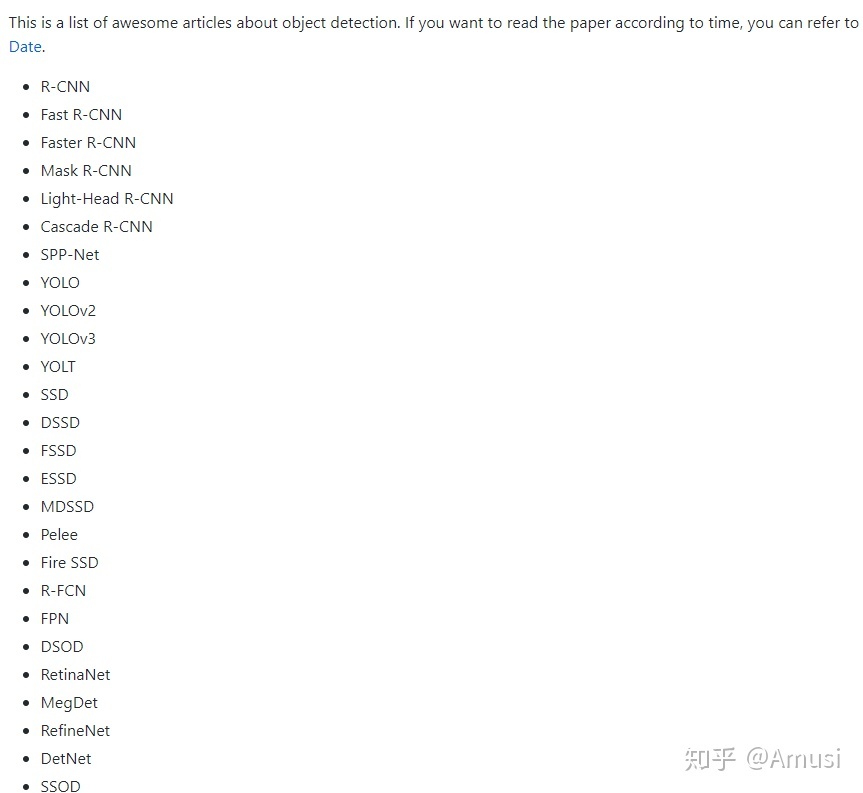
总结一下，传统目标检测存在的两个主要问题：

（1）基于滑动窗口的区域选择策略没有针对性，时间复杂度高，窗口冗余；

（2）手工设计的特征对于多样性的变化并没有很好的鲁棒性。

**基于深度学习的目标检测算法**





**R-CNN**

R-CNN的全称是Region-CNN，是第一个成功将深度学习应用到目标检测上的算法。R-CNN基于[卷积神经网络](https://baike.baidu.com/item/%E5%8D%B7%E7%A7%AF%E7%A5%9E%E7%BB%8F%E7%BD%91%E7%BB%9C/17541100" \t "https://baike.baidu.com/item/R-CNN/_blank)(CNN)，线性回归，和[支持向量机](https://baike.baidu.com/item/%E6%94%AF%E6%8C%81%E5%90%91%E9%87%8F%E6%9C%BA/9683835" \t "https://baike.baidu.com/item/R-CNN/_blank)(SVM)等算法，实现目标检测技术。

R-CNN算法解析参照：<https://www.jianshu.com/p/5056e6143ed5>

**Fast R-CNN与Faster-R-CNN**

Fast R-CNN:2014年R-CNN横空出世，首次将卷积神经网络带入目标检测领域。受SPPnet启发，rbg在15年发表Fast R-CNN，它的构思精巧，流程更为紧凑，大幅提高目标检测速度。

Fast R-CNN算法参照：<https://www.jianshu.com/p/fbbb21e1e390>

Fast R-CNN论文及中文翻译：<https://blog.csdn.net/quincuntial/article/details/79132243>

Fast R-CNN论文解析：<https://blog.csdn.net/tmylzq187/article/details/51441553>

Faster-R-CNN:<https://blog.csdn.net/shenziheng1/article/details/82907663>