

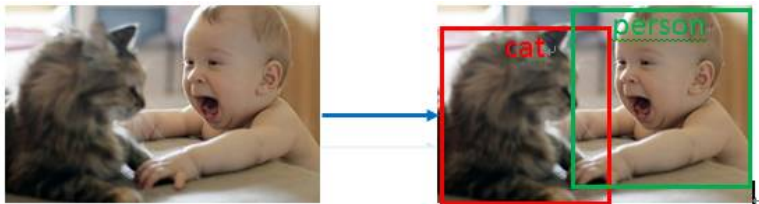
基于深度学习的目标检测研究进展

2016-05-30 深度学习大讲堂 深度学习大讲堂

点击上方“公众号”可订阅哦！

深度学习大讲堂致力于推送人工智能，深度学习方面的最新技术，产品以及活动。

前言

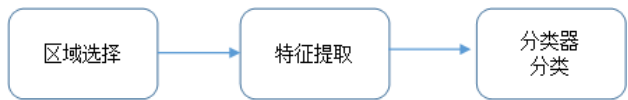


开始本文内容之前，我们先来看一下上边左侧的这张图，从图中你看到了**什么物体**？他们在**什么位置**？这还不简单，图中有一个猫和一个人，具体的位置就是上图右侧图像两个边框(bounding-box)所在的位置。其实刚刚的这个过程就是目标检测，目标检测就是“给定一张图像或者视频帧，找出其中所有目标的位置，并给出每个目标的具体类别”。

目标检测对于人来说是在再简单不过的任务，但是对于计算机来说，它看到的是一些值为0~255的数组，因而很难直接得到图像中有人或者猫这种高层语义概念，也不清楚目标出现在图像中哪个区域。图像中的目标可能出现在任何位置，目标的形态可能存在各种各样的变化，图像的背景千差万别……，这些因素导致目标检测并不是一个容易解决的任务。

得益于深度学习——主要是卷积神经网络(convolution neural network: CNN)和候选区域(region proposal)算法，从2014年开始，目标检测取得了巨大的突破。本文主要对基于深度学习的目标检测算法进行剖析和总结，文章分为四个部分：第一部分大体介绍下传统目标检测的流程，第二部分介绍以R-CNN为代表的结合region proposal和CNN分类的目标检测框架(R-CNN, SPP-NET, Fast R-CNN, Faster R-CNN); 第三部分介绍以YOLO为代表的将目标检测转换为回归问题的目标检测框架(YOLO, SSD); 第四部分介绍一些可以提高目标检测性能的技巧和方法。

一. 传统目标检测方法



如上图所示，传统目标检测的方法一般分为三个阶段：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。

(1) 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）

(2) 特征提取

传送门 for iOS 下载

深度学习大讲堂 微信二维码



分享这篇文章

数据

阅读 10277

点赞 56

更新 6月1日 1:13

海外充值支付

支持Paypal 信用卡 国际

海外充值支付宝

专为海外华人留学生充值支付宝红包，
免费送10美金

UAKA.com

深度学习大讲堂 最新头条文章

见微知著：语义分割中的弱监督学习 2017-12-06

实录：余凯、颜水成、梅涛、张兆翔、山世光同台讨论“深度学习的能与不能” 2017-11-07

章国锋：黑暗中的前行--复杂环境下的鲁棒SfM与SLAM | VALSE2017之十五 2017-11-01

何晖光：“深度学习类脑吗？”--- 基于视觉信息编解码的深度学习类脑机制研究 | VALSE2017之十四 2017-10-25

山世光：我的Face Zero之梦，写在AlphaGo Zero出世之际 2017-10-23

孙剑：如何在公司做好计算机视觉的研究 | VALSE2017之十三 2017-10-18

梅涛：深度学习敲敲视觉理解中的“钉子” 2017-10-12

韩琥：深度学习让机器给人脸“贴标签” 2017-10-09

为国庆献礼：无人智能系统，守护祖国海空长城 2017-10-02

梅涛：“看图说话”——人类走开，我AI来！ | VALSE2017之十二 2017-09-25

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）

(3) 分类器

主要有SVM, Adaboost等。

总结：传统目标检测存在的两个主要问题：一个是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；二是手工设计的特征对于多样性的变化并没有很好的鲁棒性。

二. 基于Region Proposal的深度学习目标检测算法

对于传统目标检测任务存在的两个主要问题，我们该如何解决呢？

对于滑动窗口存在的问题，region proposal提供了很好的解决方案。region proposal（候选区域）是预先找出图中目标可能出现的位置。但由于region proposal利用了图像中的纹理、边缘、颜色等信息，可以保证在选取较少窗口（几千个甚至几百个）的情况下保持较高的召回率。这大大降低了后续操作的时间复杂度，并且获取的候选窗口要比滑动窗口的质量更高（滑动窗口固定长宽比）。比较常用的region proposal算法有selective Search和edge Boxes，如果想具体了解region proposal可以看一下PAMI2015的“What makes for effective detection proposals？”

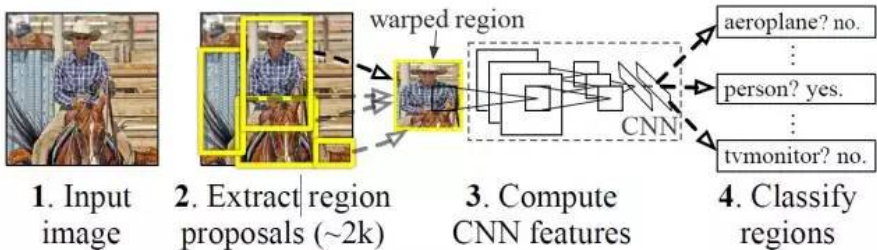
有了候选区域，剩下的工作实际就是对候选区域进行图像分类的工作（特征提取+分类）。对于图像分类，不得不提的是2012年ImageNet大规模视觉识别挑战赛（ILSVRC）上，机器学习泰斗Geoffrey Hinton教授带领学生Krizhevsky使用卷积神经网络将ILSVRC分类任务的Top-5 error降低到了15.3%，而使用传统方法的第二名top-5 error高达26.2%。此后，卷积神经网络占据了图像分类任务的绝对统治地位，微软最新的ResNet和谷歌的Inception V4模型的top-5 error降到了4%以内多，这已经超越人在这个特定任务上的能力。所以目标检测得到候选区域后使用CNN对其进行图像分类是一个不错的选择。

2014年，RBG（Ross B. Girshick）大神使用region proposal+CNN代替传统目标检测使用的滑动窗口+手工设计特征，设计了R-CNN框架，使得目标检测取得巨大突破，并开启了基于深度学习目标检测的热潮。

2 1. R-CNN (CVPR2014, TPAMI2015)

(Region-based Convolution Networks for Accurate Object detection and Segmentation)

R-CNN: Region-based Convolutional Network



上面的框架图清晰的给出了R-CNN的目标检测流程：

(1) 输入测试图像

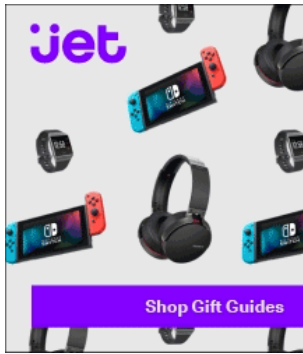
(2) 利用selective search算法在图像中提取2000个左右的region proposal。

(3) 将每个region proposal缩放（warp）成227x227的大小并输入到CNN，将CNN的fc7层的输出作为特征。

(4) 将每个region proposal提取到的CNN特征输入到SVM进行分类。

针对上面的框架给出几点解释：

* 上面的框架图是测试的流程图，要进行测试我们首先要训练好提取特征的CNN模型，以及用于分类的SVM：使用在ImageNet上预训练的模型（AlexNet/VGG16）进行微调得到用于特征提取的CNN模型，然后利用CNN模型对训练集提特征训练SVM。



深度学习大讲堂 热门头条文章

SeetaFace开源人脸识别引擎介绍 19746阅读

基于深度学习的目标检测研究进展 10277阅读

技术揭秘：海康威视PASCAL VOC2012目标检测权威评测夺冠之道 10031阅读

深度学习框架Caffe源码解析 7983阅读

深度学习中的激活函数导引 5280阅读

基于深度学习的视觉实例搜索研究进展 4923阅读

【Valse首发】CNN的近期进展与实用技巧（上） 4541阅读

【Valse大会首发】领域自适应及其在人脸识别中的应用 2051阅读

【神文】用文言文英语和文言文三语评述机器学习之产生式模型 2011阅读

深度学习大讲堂新年贺词 1891阅读



* 对每个region proposal缩放到同一尺度是因为CNN全连接层输入需要保证维度固定。

* 上图少画了一个过程——对于SVM分好类的region proposal做边框回归 (bounding-box regression)，边框回归是对region proposal进行纠正的线性回归算法，为了让region proposal提取到的窗口跟目标真实窗口更吻合。因为region proposal提取到的窗口不可能跟人工标记那么准，如果region proposal跟目标位置偏移较大，即便是分类正确了，但是由于IoU(region proposal与Ground Truth的窗口的交集比并集的比值)低于0.5，那么相当于目标还是没有检测到。

小结：R-CNN在PASCAL VOC2007上的检测结果从DPM HSC的34.3%直接提升到了66%(mAP)。如此大的提升使我们看到了region proposal+CNN的巨大优势。

但是R-CNN框架也存在着很多问题：

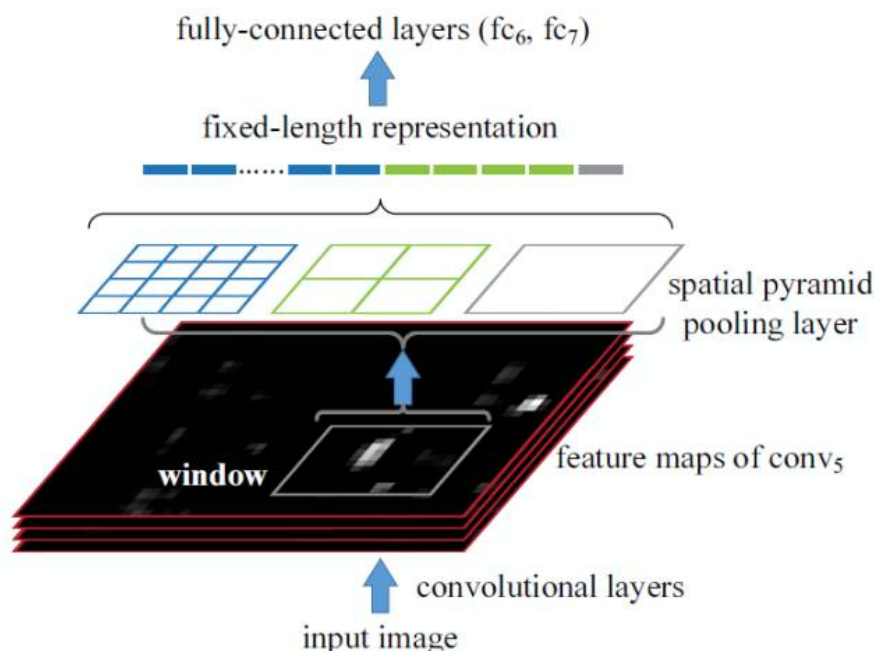
- (1) 训练分为多个阶段，步骤繁琐：微调网络+训练SVM+训练边框回归器
 - (2) 训练耗时，占用磁盘空间大：5000张图像产生几百G的特征文件
 - (3) 速度慢：使用GPU，VGG16模型处理一张图像需要47s。
- 针对速度慢的这个问题，SPP-NET给出了很好的解决方案。

3 1. SPP-NET (ECCV2014, TPAMI2015)

(Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition)

先看一下R-CNN为什么检测速度这么慢，一张图都需要47s！仔细看下R-CNN框架发现，对图像提完region proposal (2000个左右) 之后将每个proposal当成一张图像进行后续处理(CNN提特征+SVM分类)，实际上对一张图像进行了2000次提特征和分类的过程！

有没有方法提速呢？好像是有的，这2000个region proposal不都是图像的一部分吗，那么我们完全可以对图像提一次卷积层特征，然后只需要将region proposal在原图的位置映射到卷积层特征图上，这样对于一张图像我们只需要提一次卷积层特征，然后将每个region proposal的卷积层特征输入到全连接层做后续操作。（对于CNN来说，大部分运算都耗在卷积操作上，这样做可以节省大量时间）。现在的问题是每个region proposal的尺度不一样，直接这样输入全连接层肯定是不行的，因为全连接层输入必须是固定的长度。SPP-NET恰好可以解决这个问题：



上图对应的就是SPP-NET的网络结构图，任意给一张图像输入到CNN，经过卷积操作我们可以得到卷积特征（比如VGG16最后的卷积层为conv5_3，共产生512张特征图）。图中的window就是原图一个region proposal对应到特征图的区域，只需要将这些不同大小window的特征映射到同样的维度，将其作为全连接的输入，就能保证只对图像提取一次卷积层特征。SPP-NET使用了空间金字塔采样 (spatial pyramid pooling)：将每个window划分为4*4, 2*2, 1*1的块，然后每个块使用max-

pooling下采样，这样对于每个window经过SPP层之后都得到了一个长度为 $(4*4+2*2+1)*512$ 维度的特征向量，将这个作为全连接层的输入进行后续操作。

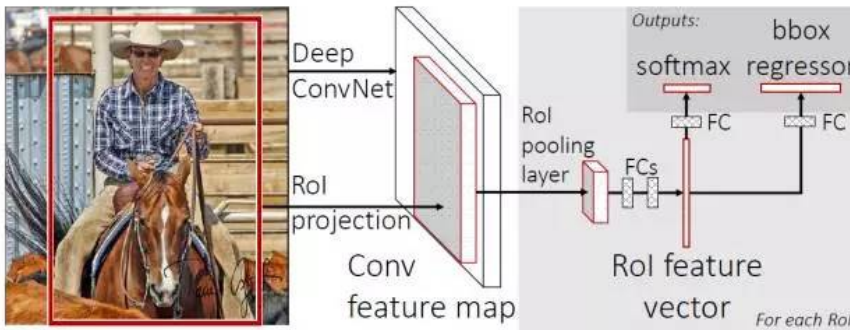
小结：使用SPP-NET相比于R-CNN可以大大加快目标检测的速度，但是依然存在着很多问题：

- (1) 训练分为多个阶段，步骤繁琐：微调网络+训练SVM+训练训练边框回归器
- (2) SPP-NET在微调网络的时候固定了卷积层，只对全连接层进行微调，而对于一个新的任务，有必要对卷积层也进行微调。（分类的模型提取的特征更注重高层语义，而目标检测任务除了语义信息还需要目标的位置信息）

针对这两个问题，RBG又提出Fast R-CNN，一个精简而快速的目标检测框架。

4 1. Fast R-CNN(ICC2015)

有了前边R-CNN和SPP-NET的介绍，我们直接看Fast R-CNN的框架图：



与R-CNN框架图对比，可以发现主要有两处不同：一是最后一个卷积层后加了一个ROI pooling layer，二是损失函数使用了多任务损失函数(multi-task loss)，将边框回归直接加入到CNN网络中训练。

- (1) ROI pooling layer实际上是SPP-NET的一个精简版，SPP-NET对每个proposal使用了不同大小的金字塔映射，而ROI pooling layer只需要下采样到一个7x7的特征图。对于VGG16网络conv5_3有512个特征图，这样所有region proposal对应了一个7*7*512维度的特征向量作为全连接层的输入。

- (2) R-CNN训练过程分为了三个阶段，而Fast R-CNN直接使用softmax替代SVM分类，同时利用多任务损失函数边框回归也加入到了网络中，这样整个的训练过程是端到端的(除去region proposal提取阶段)。

- (3) Fast R-CNN在网络微调的过程中，将部分卷积层也进行了微调，取得了更好的检测效果。

小结：Fast R-CNN融合了R-CNN和SPP-NET的精髓，并且引入多任务损失函数，使整个网络的训练和测试变得十分方便。在Pascal VOC2007训练集上训练，在VOC2007测试的结果为66.9%(mAP)，如果使用VOC2007+2012训练集训练，在VOC2007上测试结果为70%（数据集的扩充能大幅提高目标检测性能）。使用VGG16每张图像总共需要3s左右。

缺点：region proposal的提取使用selective search，目标检测时间大多消耗在这上面（提region proposal 2~3s，而提特征分类只需0.32s），无法满足实时应用，而且并没有实现真正意义上的端到端训练测试（region proposal使用selective search先提取出来）。那么有没有可能直接使用CNN直接产生region proposal并对其分类？

Faster R-CNN框架就是符合这样需要的目标检测框架。

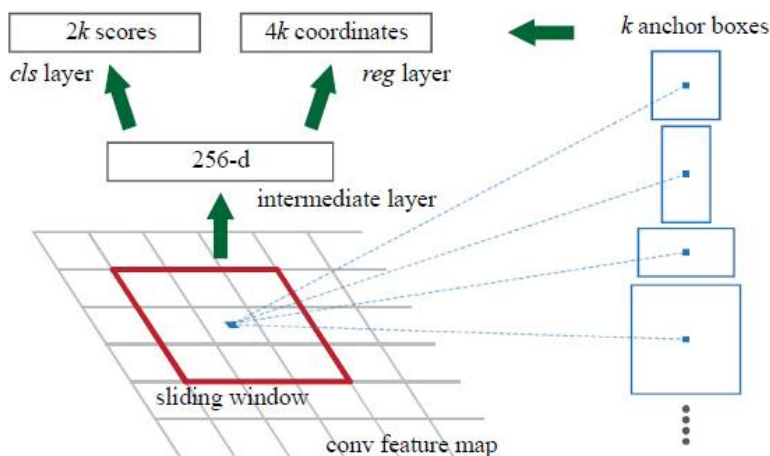
5 1. Faster R-CNN(NIPS2015)

(Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks)

在region proposal + CNN分类的这种目标检测框架中，region proposal质量好坏直接影响到目标检测任务的精度。如果找到一种方法只提取几百个或者更少的高质量的首选窗口，而且召回率很高，这不但能加快目标检测速度，还能提高目标检测的性能（假阳例少）。RPN(Region Proposal Networks)网络应运而生。

RPN的核心思想是使用卷积神经网络直接产生region proposal，使用的方法本质上就是滑动窗口。RPN的设计比较巧妙，RPN只需在最后的卷积层上滑动一遍，因为

anchor机制和边框回归可以得到多尺度多长宽比的region proposal。



我们直接看上边的RPN网络结构图（使用了ZF模型），给定输入图像（假设分辨率为600*1000），经过卷积操作得到最后一层的卷积特征图（大小约为40*60）。在这个特征图上使用3*3的卷积核（滑动窗口）与特征图进行卷积，最后一层卷积层共有256个feature map，那么这个3*3的区域卷积后可以获得一个256维的特征向量，后边接cls layer和reg layer分别用于分类和边框回归（跟Fast R-CNN类似，只不过这里的类别只有目标和背景两个类别）。3*3滑窗对应的每个特征区域同时预测输入图像3种尺度（128,256,512），3种长宽比（1:1,1:2,2:1）的region proposal，这种映射的机制称为anchor。所以对于这个40*60的feature map，总共有约20000(40*60*9)个anchor，也就是预测20000个region proposal。

这样设计的好处是什么呢？虽然现在也是用的滑动窗口策略，但是：滑动窗口操作是在卷积层特征图上进行的，维度较原始图像降低了16*16倍（中间经过了4次2*2的pooling操作）；多尺度采用了9种anchor，对应了三种尺度和三种长宽比，加上后边接了边框回归，所以即便是这9种anchor外的窗口也能得到一个跟目标比较接近的region proposal。

NIPS2015版本的Faster R-CNN使用的检测框架是RPN网络+Fast R-CNN网络分离进行的目标检测，整体流程跟Fast R-CNN一样，只是region proposal现在是用RPN网络提取的（代替原来的selective search）。同时作者为了让RPN的网络和Fast R-CNN网络实现卷积层的权值共享，训练RPN和Fast R-CNN的时候用了4阶段的训练方法：

- (1) 使用在ImageNet上预训练模型初始化网络参数，微调RPN网络；
- (2) 使用(1)中RPN网络提取region proposal训练Fast R-CNN网络；
- (3) 使用(2)的Fast R-CNN网络重新初始化RPN，固定卷积层进行微调；
- (4) 固定(2)中Fast R-CNN的卷积层，使用(3)中RPN提取的region proposal微调网络。

权值共享后的RPN和Fast R-CNN用于目标检测精度会提高一些。

使用训练好的RPN网络，给定测试图像，可以直接得到边缘回归后的region proposal，根据region proposal的类别得分对RPN网络进行排序，并选取前300个窗口作为Fast R-CNN的输入进行目标检测，使用VOC07+12训练集训练，VOC2007测试集测试mAP达到73.2%（selective search + Fast R-CNN是70%），目标检测的速度可以达到每秒5帧（selective search+Fast R-CNN是2~3s一张）。

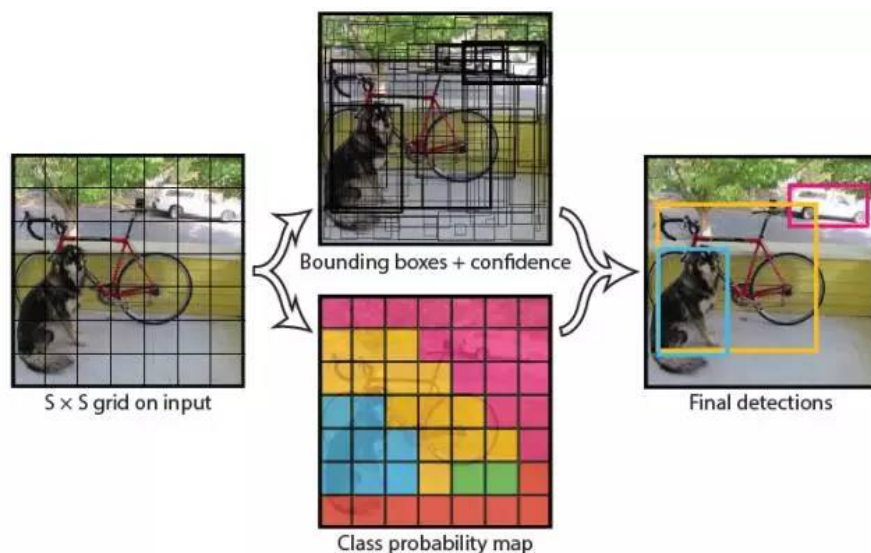
需要注意的是，最新的版本已经将RPN网络和Fast R-CNN网络结合到了一起——将RPN获取到的proposal直接连到ROI pooling层，这才是一个真正意义上的使用一个CNN网络实现端到端目标检测的框架。

小结：Faster R-CNN将一直以来分离的region proposal和CNN分类融合到了一起，使用端到端的网络进行目标检测，无论在速度上还是精度上都得到了不错的提高。然而Faster R-CNN还是达不到实时的目标检测，预先获取region proposal，然后在对每个proposal分类计算量还是比较大。比较幸运的是YOLO这类目标检测方法的出现让实时性也变的成为可能。

三. 基于回归方法的深度学习目标检测算法

1 YOLO (CVPR2016, oral)

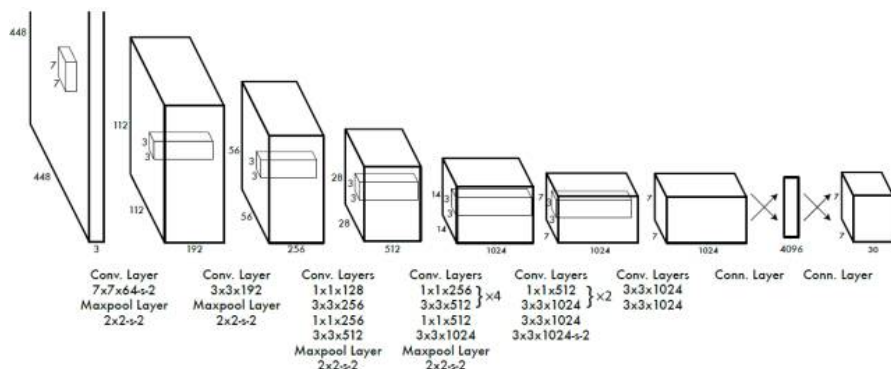
(You Only Look Once: Unified, Real-Time Object Detection)



(1) 给个一个输入图像，首先将图像划分成 7×7 的网格

(3) 根据上一步可以预测出 $7*7*2$ 个目标窗口，然后根据阈值去除可能性比较低的目标窗口，最后NMS去除冗余窗口即可。

可以看到整个过程非常简单，不需要中间的region proposal在找目标，直接回归便完成了位置和类别的判定。



6/10

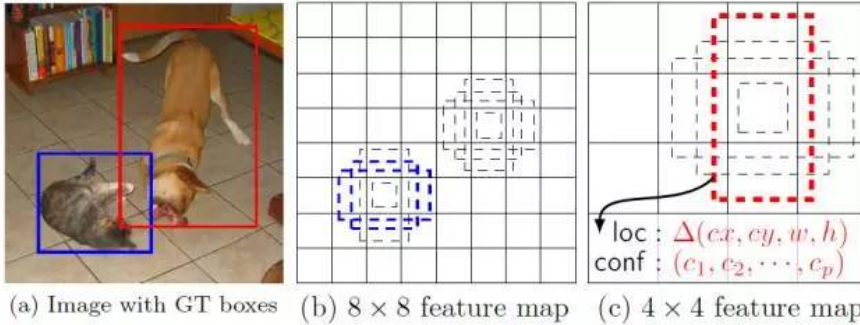
4096维的全图特征直接在每个网格上回归出目标检测需要的信息（边框信息加类别）。

小结：YOLO将目标检测任务转换成一个回归问题，大大加快了检测的速度，使得YOLO可以每秒处理45张图像。而且由于每个网络预测目标窗口时使用的是全图信息，使得false positive比例大幅降低（充分的上下文信息）。但是YOLO也存在问题：没有了region proposal机制，只使用7*7的网格回归会使得目标不能非常精准的定位，这也导致了YOLO的检测精度并不是很高。

2 SSD

(SSD: Single Shot MultiBox Detector)

上面分析了YOLO存在的问题，使用整图特征在7*7的粗糙网格内回归对目标的定位并不是很精准。那是不是可以结合region proposal的思想实现精准一些的定位？SSD结合YOLO的回归思想以及Faster R-CNN的anchor机制做到了这点。



上图是SSD的一个框架图，首先SSD获取目标位置和类别的方法跟YOLO一样，都是使用回归，但是YOLO预测某个位置使用的是全图的特征，SSD预测某个位置使用的是这个位置周围的特征（感觉更合理一些）。那么如何建立某个位置和其特征的对对应关系呢？可能你已经想到了，使用Faster R-CNN的anchor机制。如SSD的框架图所示，假如某一层特征图(图b)大小是 8×8 ，那么就使用 3×3 的滑窗提取每个位置的特征，然后这个特征回归得到目标的坐标信息和类别信息(图c)。

不同于Faster R-CNN，这个anchor是在多个feature map上，这样可以利用多层的特征并且自然的达到多尺度（不同层的feature map 3×3 滑窗感受野不同）。

小结：SSD结合了YOLO中的回归思想和Faster R-CNN中的anchor机制，使用全图各个位置的多尺度区域特征进行回归，既保持了YOLO速度快的特性，也保证了窗口预测的跟Faster R-CNN一样比较精准。SSD在VOC2007上mAP可以达到72.1%，速度在GPU上达到58帧每秒。

总结：YOLO的提出给目标检测一个新的思路，SSD的性能则让我们看到了目标检测在实际应用中真正的可能性。

四. 提高目标检测方法

R-CNN系列目标检测框架和YOLO目标检测框架给了我们进行目标检测的两个基本框架。除此之外，研究人员基于这些框架从其他方面入手提出了一系列提高目标检测性能的方法。

(1) 难分样本挖掘 (hard negative mining)

R-CNN在训练SVM分类器时使用了难分样本挖掘的思想，但Fast R-CNN和Faster R-CNN由于使用端到端的训练策略并没有使用难分样本挖掘（只是设置了正负样本的比例并随机抽取）。CVPR2016的Training Region-based Object Detectors with Online Hard Example Mining(oral)将难分样本挖掘(hard example mining)机制嵌入到SGD算法中，使得Fast R-CNN在训练的过程中根据region proposal的损失自动选取合适的region proposal作为正负例训练。实验结果表明使用OHEM (Online Hard Example Mining) 机制可以使得Fast R-CNN算法在VOC2007和VOC2012上mAP提高4%左右。

(2) 多层特征融合

Fast R-CNN和Faster R-CNN都是利用了最后卷积层的特征进行目标检测，而由于高层的卷积层特征已经损失了很多细节信息（pooling操作），所以在定位时不是很精准。HyperNet等一些方法则利用了CNN的多层特征融合进行目标检测，这不仅利用了高层特征的语义信息，还考虑了低层特征的细节纹理信息，使得目标检测定位更精准。

(3) 使用上下文信息

在提取region proposal特征进行目标检测时，结合region proposal上下文信息，检测效果往往会更好一些。（Object detection via a multi-region & semantic segmentation-aware CNN model以及Inside-Outside Net等论文中都使用了上下文信息）

该文章属于“深度学习大讲堂”原创，如需要转载，请联系loveholicguoguo。

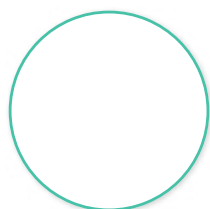
作者简介

王斌

中科院计算所前瞻研究实验室跨媒体计算课题组博士生，导师张勇东研究员。研究方向为基于深度学习的目标检测。2015年作为计算所MCG-ICT-CAS团队成员，参加ImageNet竞赛ILSVRC2015的目标检测任务获得全球第5名。

往期文章精选：

- 【脑洞】Hinton剑桥演讲：大脑神经元的误差反向传播机制
- 【阿里集团卜居深度解析】卷积神经网络的硬件加速
- 全卷积网络：从图像级理解到像素级理解
- 长文干货！走近人脸检测：从VJ到深度学习（下）
- 长文干货！走近人脸检测：从VJ到深度学习（上）
- 人脸识别简史与近期进展
- 面部特征点定位概述及最近研究进展
- 【Valse首发】CNN的近期进展与实用技巧（上）
- 【Caffe实战与小数据深度学习】CNN的近期进展与实用技巧（下）
- 【Valse大会首发】领域自适应及其在人脸识别中的应用



欢迎关注我们！

深度学习大讲堂致力于推送人工智能，深度学习的最新技术，产品和活动！

深度学习大讲堂



觉得不错，分享给更多人看到 微信 QQ空间 新浪微博 腾讯微博 人人 Twitter 豆瓣 百度贴吧

深度学习大讲堂 最新文章:

- 2017-12-06 见微知著：语义分割中的弱监督学习
- 2017-11-07 实录：余凯、颜水成、梅涛、张兆翔、山世光同台讨论“深度学习的能与不能”
- 2017-11-01 章国锋：黑暗中的前行--复杂环境下的鲁棒SfM与SLAM | VALSE2017之十五
- 2017-10-25 何晖光：“深度学习类脑吗？”--- 基于视觉信息编解码的深度学习类脑机制研究 | VALSE2017之十四
- 2017-10-23 山世光：我的Face Zero之梦，写在AlphaGo Zero出世之际
- 2017-10-18 孙剑：如何在公司做好计算机视觉的研究|VALSE2017之十三
- 2017-10-12 梅涛：深度学习敲敲视觉理解中的“钉子”
- 2017-10-09 韩琰：深度学习让机器给人脸“贴标签”
- 2017-10-02 为国庆献礼：无人智能系统，守护祖国海空长城
- 2017-09-25 梅涛：“看图说话”——人类走开，我AI来！|VALSE2017之十二



海外充值支付宝

深度学习大讲堂 热门文章:

- SeetaFace开源人脸识别引擎介绍 阅读/点赞：19746/52
- 基于深度学习的目标检测研究进展 阅读/点赞：10277/56
- 技术揭秘：海康威视PASCAL VOC2012目标检测权威评测夺冠之道 阅读/点赞：10031/41
- 深度学习框架Caffe源码解析 阅读/点赞：7983/38
- 深度学习中的激活函数导引 阅读/点赞：5280/68
- 基于深度学习的视觉实例搜索研究进展 阅读/点赞：4923/55
- 【Valse首发】CNN的近期进展与实用技巧（上） 阅读/点赞：4541/37
- 【Valse大会首发】领域自适应及其在人脸识别中的应用 阅读/点赞：2051/33
- 【神文】用文言文英语和白话文三语评述机器学习之产生式模型 阅读/点赞：2011/73
- 深度学习大讲堂新年贺词 阅读/点赞：1891/43



技术揭秘：海康威视
ASCAL VOC2012目标
检测权威评测夺冠之...



【CVPR2016论文快
讯】细粒度视觉分类的
最新进展_深度学习大...



深度 | 神经网络和深度
学习简史第四部分：深
度学习终迎伟大复兴_...



广告

干货 | 深度学习之CNN
反向传播算法详解_机器
学习算法与Python学习



深度学习入门：投身深
度学习你需要哪些准
备？_机器学习研究会



【珍藏】了解CNN这一
篇就够了：卷积神经网
络技术及发展_新智元



【推荐】(TensorFlow)
深度学习中文分词开源
了！98%准确率的深...



【干...
词袋...
图像...

热榜