




 **AoDeLuo** [关注](#)


 32



 33

 0

 分享

 打赏

R-CNN详解

原创 于 2025-10-21 22:45:00 发布 · 1.5k 阅读 · 32 · 33 · CC 4.0 BY-SA版权

文章标签: #神经网络

38 篇文章 订阅

Yolo-v8.3 Yolo

YOLO (You Only Look Once) 是一种流行的物体检测和图像分割模型，由华盛顿大学的Joseph Redmon 和Ali Farhadi 开发。YOLO 于2015 年推出，因其高速和... 一键部署

文章目录

- R-CNN: Regions + CNN
 - 创新点
 - 网络结构
- 网络结构详解
 - 1.特征区域抽取(Extract region proposal)
 - 2.计算CNN特征(Compute CNN features)
 - 3.区域分类(Classify regions)
 - 4.非极大值抑制(Non-maximum suppression)
 - 5.边界框回归(Bounding box regression)

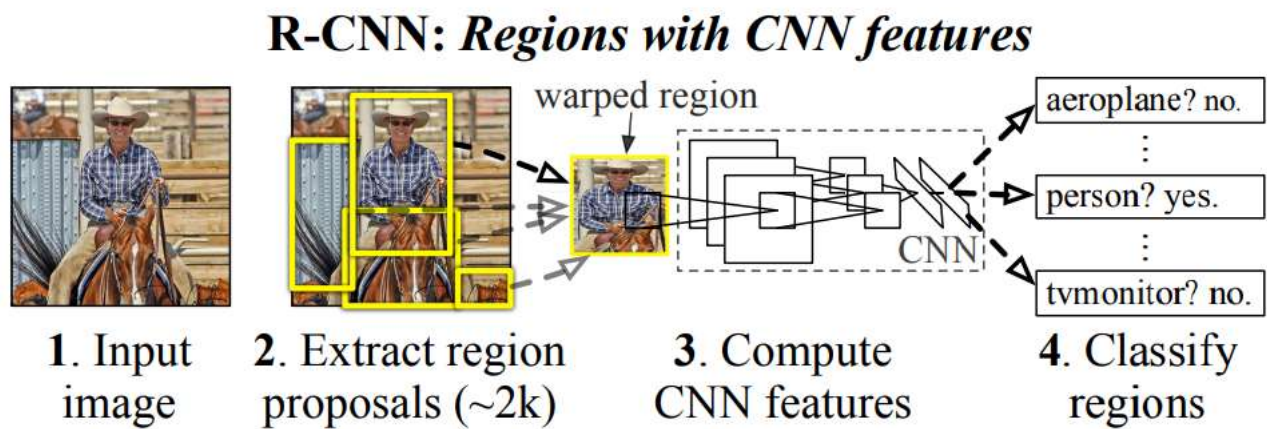
R-CNN : Regions + CNN

《Rich feature hierarchies for accurate object detection and semantic segmentation》
论文地址: <https://arxiv.org/pdf/1311.2524v3.pdf>
代码地址: <https://github.com/rbgirshick/r>

创新点

- R-CNN采用Selective Search算法预先提取出可能包含目标的 Region Proposal (候选区域) , 替代传统算法 (DPM等) 的滑动窗口法, 提高速度。
- 使用CNN提取Region Proposal的特征。从经验驱动特征 (SIFT、HOG) 到数据驱动特征 (CNN feature map) , 提高特征对样本的表示能力。
- 采用大数据集下 (ImageNet ILSVC 2012) 有监督预训练和小数据集下 (PASCAL VOC 2007) 微调 (fine-tuning) 的方法解决小样本难以训练甚至过问题。

网络结构



过程

1. Extract Region Proposal, 使用Sele



AoDeLuo

关注

32 33 0 分享 打赏

2. Compute CNN features, 使用CNN网络计算每个Region Proposal的feature map
3. Classify regions, 将提取到的feature输入到SVM中进行分类
4. Non-maximum suppression, 去除掉重复的box
5. Bounding box regression, 位置精修, 使用回归器精细修正候选框的位置

网络结构详解

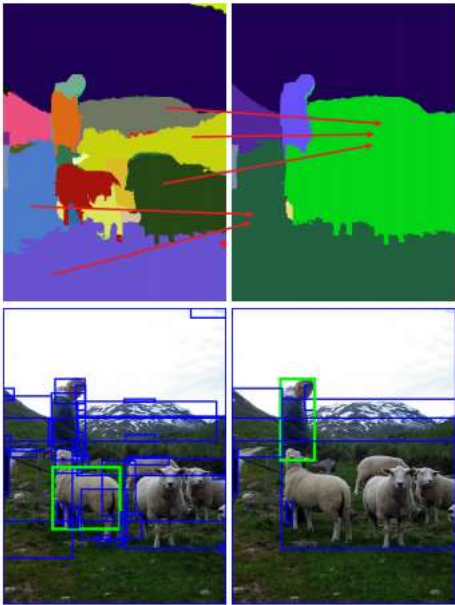
1.特征区域抽取(Extract region proposal)

AI写代码

登录

- 1 首先使用Selective Search算法从输入图像中提取2000个Region Proposal, Selective Search算法主要步骤:
- 2
- 3 1. 使用一种过分割手段, 将图像分割成小区域 (1k~2k 个)
- 4 2. 计算所有邻近区域之间的相似性, 包括颜色、纹理、尺度等
- 5 3. 将相似度比较高的区域合并到一起
- 6 4. 计算合并区域和临近区域的相似度
- 7 5. 重复3、4过程, 直到整个图片变成一个区域
- 8
- 9 在每次迭代中, 形成更大的区域并将其添加到区域提议列表中。这种自下而上的方式可以创建从小到大的不同scale的Region Proposal, 如图所示。

收起 ^



2.计算CNN特征(Compute CNN features)

2.1 缩放区域(Warp region)

由于文中使用的CNN中包含有 全连接层 , 这就需要输入神经网络的图片有相同的size, 但是Selective Search提取的Region Proposal都是不同size的, 需要对每个Region Proposal都缩放到固定的大小 (227*227) 。paper试验了两种不同的处理方法:

(1)各向异性缩放

这种方法比较简单暴力, 不考虑图片的长宽比例, 不考虑图片是否扭曲, 直接缩放到CNN输入的大小227*227。不过这种方法容易导致图片中目标发生严重变, 如下图D所示。

(2)各向同性缩放

1. 先扩充后裁剪: 直接在原始图片中, 把bounding box的边界进行扩展延伸成正方形, 然后再进行裁剪。如果已经延伸到了原始图片的外边界, 那么就 bounding box中的颜色均值填充。如上图(B)所示。



AoDeLuo

关注

32



33

0

分享

打赏



2. 先裁剪后扩充：先把bounding box图片裁剪出来，然后用固定的背景颜色填充成正方形图片(背景颜色也是采用bounding box的像素颜色均值),如上图所示。

注：对于上面的异性、同性缩放，文献还有个padding处理，上面的示意图中第1、3行就是结合了padding=0,第2、4行结果图采用padding=16的结果。最后的试验，作者发现采用各向异性缩放、padding=16的精度最高。

2.2 CNN网络 训练 (Train CNN)

利用Selective Search提取Region Proposal并resize后，接下来使用CNN (AlexNet、VGG) 从每个Region Proposal提取特征。本文训练CNN的方法，包括以下两步：

(1) Pre-training阶段：由于物体标签训练数据少，如果要直接采用随机初始化CNN参数的方法是不足以从零开始训练出一个好的CNN模型。基于此，本文是有监督的预训练，使用一个大的数据集 (ImageNet ILSVC 2012) 来训练AlexNet，得到一个分类的预训练 (Pre-trained) 模型。

(2) Fine-tuning阶段：使用Region Proposal (PASCAL VOC) 对Pre-trained模型进行fine-tuning。首先将原来预训练模型最后的全连接层 (层) 换成21-way的分类层 (20类物体+背景)，然后计算每个region proposal和ground truth的IoU，对于IoU>0.5的region proposal被视为正样本，否则为负样本 (即背景)。在每次迭代的过程中，选取32个正样本和96个负样本组成一个mini-batch (128，正负比：1: 3)。我们使用0.001的学习率和SGD来进行训练 (备注：如果不针对特定任务进行fine-tuning，而是把CNN当做特征提取器，卷积层所学到的特征其实就是基础的共享特征提取层，就类似于SIFT算法——以用于提取各种图片的特征，而f6、f7所学习到的特征是针对特定任务的特征。打个比方：对于人脸性别识别来说，一个CNN模型前面的卷积层所学到的特征就类似于学习人脸共性特征，然后全连接层所学到的特征就是针对性别分类的特征了)

2.3 Save features

虽然文中训练了CNN网络对region proposal进行分类，但是实际中，这个CNN的作用只是提取每个region proposal的feature。因此，我们输入region proposal行前向传播，然后保存AlexNet的FC7层features，以供后续的SVM分类使用。

3.区域分类(Classify regions)

本文使用SVM进行分类对于每一类都会训练一个SVM分类器，所以共有N (21) 个分类器，我们来看一下是如何训练和使用SVM分类器的。

3.1 训练(Training)

如下图所示，在训练过程中，SVM的输入包括两部分：

(1) CNN feature：这个便是CNN网络为每个region proposal提取的feature，共2000*4096。

(2) Ground truth labels：在训练时，会为每个region proposal附上一个label (标注好的labels称为Ground truth labels)。

在SVM分类过程中，当IoU<0.3时，为负样本，正样本便是ground truth box。然后SVM分类器也会输出一个预测的labels，然后用labels和ground truth计算loss，然后训练SVM。

3.2 测试(Testing)

Testing的过程就是输入经过之前的步骤得到test image的Region Proposal的feature，然后输出对2000个proposal的类别预测值。

4.非极大值抑制(Non-maximum suppression)

经过SVM之后，我们会得到2000个region proposal的class probability，然后我们可以根据‘有无物体’这一类过滤掉一大批region proposal，然后如果某个proposal的最大class probability<阈值，那也可以过滤掉这些region proposal，那剩下的可能如下左图所示，就是有多多个box相互重叠，但是我们目标检测的目标是有一个box即可，那这个时候就需要用到非极大值抑制 (NMS) 了，经过NMS之后，最终的检测结果如下右图所示：

5.边界框回归(Bounding box regression)



AoDeLuo

关注

32

1

33

0

分享

打赏

目标检测问题的衡量标准是重叠面积：许多看似准确的检测结果，往往因为候选框不够准确，重叠面积很小。故需要Bounding box regression步骤。如下色的框表示Ground Truth Box, 红色的框为我们预测得到的region proposal 。那么即使红色的框被分类器识别为飞机，但是由于红色的框定位不准(IoU\$<\$那么这张图相当于没有正确的检测出飞机。所以需要红色的框进行微调，使得经过微调后的窗口跟Ground Truth Box更接近，这样就可以更准确的定位。Bounding box regression算法参考这里。

【R-CNN不足】

- 1. 训练时间长：主要原因是分阶段多次训练，而且对于每个region proposal都要单独计算一次feature map，导致整体的时间变长。
- 2. 占用空间大：主要原因是每个region proposal的feature map都要写入硬盘中保存，以供后续的步骤使用。

multi-stage：文章中提出的模型包括多个模块，每个模块都是相互独立的，训练也是分开的。

3. 测试时间长，由于不共享计算，所以对于test image，也要为每个proposal单独计算一次feature map，因此测试时间也很长。

您可能感兴趣的与本文相关的镜像



Yolo-v8.3

Yolo

YOLO (You Only Look Once) 是一种流行的物体检测和图像分割模型，由华盛顿大学的Joseph Redmon 和Ali Farhadi 开发。YOLO 于2015 年推出，因其高速和...

一键部署

R-CNN、Fast RCNN和Faster RCNN网络介绍

chase

R-CNN、Fast RCNN和Faster RCNN网络介绍

大话目标检测经典模型 (RCNN、Fast RCNN、Faster RCNN)

qq_33287871的博客

目标检测是深度学习的一个重要应用，就是在图片中要将里面的物体识别出来，并标出物体的位置，一般需要经过两个步骤： 1、分类，识别物体是什么 2、定位，找出物体在

rcnn代码—train.py_rcnn 最初代码

1.函数入口在train.py model=crnn.get_crnn(config) AI写代码python 运行 1 通过get_rcnn函数与配置信息(config)构建基础模型 2.get_rcnn函数 defget_crnn(config):model=CRNI

Mask RCNN 源代码解析 (1) - 整体思路_mask rcnn代码

MaskRCNN 属于 RCNN这一系列的应该还是比较最终版本,融合多种算法的思想,这里对Mask RCNN从源代码进行解析,主要写几篇文章,一个总结大的思路,其他文章整理细节。

R-CNN史上最全讲解 热门推荐

weixin_43702653的博客

R-CNN史上最全讲解

R-CNN(目标检测算法)介绍

Sea_c的博客

因为一个框可以用四个值来表示 (x,y,w,h) 分别是框的中心点横坐标, 中心点纵坐标, 宽, 高。由于物体标签训练数据少, 如果要直接采用随机初始化CNN参数的方法,目前的

目标检测Faster_r_cnn代码的使用(python+caffe版本)

AI写代码 1 2 faster_rcnn模型链接: https://pan.baidu.com/s/1miDWCEc 密码: xq4y 2:运行模型 \$ cd \$FRCN_ROOT \$./tools/demo.py AI写代码 1 2 结果如下: 训练自己的数据集

cascade rcnn检测代码_cascade r-cnn代码

cascade rcnn检测代码 本文介绍了一种基于Cascade-RCNN的目标检测方法实现过程,包括网络配置、数据集准备及检测配置等关键步骤,并展示了如何通过MATLAB进行模型部

RCNN详解

panfengblog

RCNN详解 RCNN即region proposals (候选区域) + CNN, 是将CNN引入目标检测领域的开山之作 (2014年), 大大提高了目标检测的效果, 在其后也是出现了更优异的变体

【目标检测算法】R-CNN (详解)

风口IT猪的成长录

目标检测算法之R-CNN学习目标1. 目标检测-Overfeat模型1.1 滑动窗口1.2 Overfeat模型总结2. 目标检测-R-CNN模型2.1 完整R-CNN结构2.2 候选区域 (Region of Interest) 得

faster rcnn 代码解析10_ml.m..m.lm

---# Fast/r R-CNN# Licensed under The MIT License [see LICENSE for details]# Written by Bharath Hariharan# ---importxml.etree.ElementTreeasETimportosimportPickle

【目标检测经典算法】R-CNN、Fast R-CNN和Faster R-CNN详解系列三: Faster R-CNN图文详解

CV在读

在图像上预设好的不同大小, 不同长宽比的参照框。论文中设定每个区域可以生成k个anchor box。可以看到。

目标检测系列—Cascade R-CNN 详解

Azperk的博客

是由和于 2018 年提出的一种改进型方法, 旨在通过提升目标检测的精度。传统的单阶段目标检测器通常会在精度上遇到瓶颈, 特别是对于难以检测的目标。而 Cascade R-CNN

算法之Mask R-CNN 详解 最新发布

分享技术应用、实践场景等, 评论区开放技术难题讨...

Mask R-CNN 作为计算机视觉领域的重要算法, 在多个领域都有着广泛的应用前景。无论是新手还是成手, 都可以通过深入学习和实践, 不断挖掘其潜力, 为解决实际问题提供

【目标检测经典算法】R-CNN、Fast R-CNN和Faster R-CNN详解系列一: R-CNN图文详解

CV在读

是一种常用于目标检测的候选区域生成方法。在传统的目标检测算法中, 需要对图像中的每个可能包含目标的区域进行检测, 但这样做会导致计算量巨大, 尤其是在图像具有大

目标检测系列—Mask R-CNN 详解

Azperk的博客

是由等人于 2017 年提出的, 它在此基础上进行了扩



AoDeLuo

关注

32



33

0

分享

打赏

- RCNN

原文理解与代码分析

weixin_45595421的博客
- RCNN

算法步骤详解与代码
- 深度学习之目标检测R-CNN模型算法流程

详解说明（超详细理论篇）

qq_55433305的博客
- RCNN

算法分为4个步骤:获取候选区域：对于一张输入的图像，首先使用selective search算法获取2000个左右的候选区域。获取图像特征：将图像输入到卷积神经网络中获取图
- R-CNN

R-CNN

m0_51797359的博客
- 目标检测 | R-CNN、Fast R-CNN与Faster R-CNN理论讲解

目标检测：R-CNN、Fast R-CNN与Faster R-CNN理论讲解，根据B站up霹雳吧啦Wz总结。此文章包含R-CNN算法原理讲解、Fast R-CNN算法原理讲解、Faster R-CNN原理讲

houjingchuan的博客
- RCNN

代码简单实现

微凉的博客
- 本文

代码

来自于github(https://github.com/1297rohit/RCNN),可以去给原作者点个????，源代码使用了kears，现在主要使用pytorch实现了一遍。首先看RCNN需要做什么 如上[
- 深度学习-目标检测(一)-R-CNN

目标检测的暴力方法是从左到右、从上到下滑动窗口，利用分类识别目标。为了在不同观察距离处检测不同的目标类型，我们使用不同大小和宽高比的窗口。如下图所示：这样

m0_73426548的博客
- RCNN(R-CNN、Region-CNN)网络，DL目标检测的先驱，Two-Stage代表网络

RCNN，第一个成功将深度学习应用到目标检测上的算法，可以说是利用深度学习进行目标检测的开山之作，是Two-Stage算法的代表，即先生成候选区域，再利用CNN进行识

weixin_43331421的博客
- cascade mask r-cnn

代码
- Cascade Mask R-CNN

是一种用于实例分割任务的深度学习模型，它是基于Mask R-CNN的改进版本。Cascade Mask R-CNN通过级联多个Mask R-CNN模型来逐步提升实例分

关于我们 招贤纳士 商务合作 寻求报道 400-660-0108 kefu@csdn.net 在线客服 工作时间 8:30-22:00

公安备案号11010502030143 京ICP备19004658号 京网文〔2020〕1039-165号 经营性网站备案信息 北京互联网违法和不良信息举报中心 家长监护 网络110报警服务 中国互联网举报中心 Chrome商店下载 账号管理规范 版权与免责声明 版权申诉 出版物许可证 营业执照

©1999-2026北京创新乐知网络技术有限公司



AoDeLuo

博客等级 码龄10年

221

1695

4668

6万+

原创

点赞

收藏

粉丝

关注

私信



Mall4j

Java开源商城

Gitee 15K Star

百万开发者的选择

电商项目必备！Java 开源商城系统
SpringBoot+Vue，功能齐全，全源码交付，可二开

广告


热门文章

海康威视监控摄像头连接电脑教程

44835

python plt 绘图详解 (plt.版本)

39723



AoDeLuo

关注

32

33

0

分享

打赏

CIFAR-10数据集简介 20087

MPU-6050详解 15453

yolov8使用详解 12961

分类专栏

- C

项目管理

1篇
- R

机器人

1篇
- C

日常记录
- C

Cmaker

14篇
- 展开全部

上一篇：官网安装的cuda和conda安装的...

下一篇：LeNet-5详解

大家在看

Linux初学者应掌握的常用命令 514

深度学习图像分类之农作物病虫害智能识别
农业生产中病虫害早期预警系统搭建 农作
物健康监测与精准防治提供数据 棉花小麦
病虫害识别 玉米病虫害检测数据集第10392
期 22

Java 大视界 -- 基于 Java 的大数据可视化
在城市地下管网管理与风险预警中的应用
419

智慧城市治理之无人机河道垃圾巡检 水上
漂浮物检测模型的开发 助力水环境垃圾清
理 水上安全监测等场景 塑料垃圾识别 河道
漂浮物识别第10383期 235

从淘宝推荐到微信搜索：查找算法如何支撑
亿级用户——动画可视化

最新文章

EBOM和PBOM的区别

SCARA机械臂国内外厂商对比

CPM:一个轻量级的 CMake 依赖管理工具

2025年 31篇

2024年 31篇

2023年 120篇

2022年 39篇



Yolo-v8.3

YOLO (You Only Look Once) 是一种流行的物...

Yolo

镜像市场

一键部署

目录

文章目录

R-CNN: Regions + CNN

创新点

网络结构

网络结构详解

- 1.特征区域抽取(Extract region propo...
- 2.计算CNN特征(Compute CNN feat...
- 3.区域分类(Classify regions)
- 4.非极大值抑制(Non-maximum supp...
- 5.边界框回归(Bounding box regressi...

收起 ^



AoDeLuo

关注

32



33

0

分享

打赏