



专栏

目录视图 摘要视图 阅读

个人资料



shenxiaolu1984

关注 发私信

访问： 241191次

积分： 2697

等级：

排名： 第12223名

【活动】2017 CSDN博客专栏评选 【评论送书】SQL优化、深度学习、数据科学家 CSDN日报20170526 ——《论程序员的时代焦虑与焦虑的缓解》 CSDN日报 | 4.19-5.19 上榜作者排行出炉

【目标检测】RCNN算法详解

标签： 目标检测 深度学习 RCNN detection

2016-04-05 23:10 17645人阅读 评论(4) 收藏 举报

分类： 论文解读 (20)

关闭

版权声明：本文为博主原创文章，未经博主允许不得转载。

目录(?) [+]

Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." Proceedings of the IEEE conference on computer vision and pattern recognition. 2014.

Region CNN(RCNN)可以说是利用深度学习进行目标检测的开山之作。作者Ross Girshick多次在PASCAL VOC比赛中得终身成就奖，如今供职于Facebook旗下的FAIR。



泰国买房





DL框架 (13)
会议概览 (7)
数学 (9)

文章存档

2017年05月 (2)
2017年04月 (1)
2017年02月 (4)
2016年12月 (1)
2016年11月 (8)

展开

阅读排行

【目标检测】Faster RCNN算法.. (32335)
【目标检测】Fast RCNN算法... (27383)
【深度学习】生成对抗网络Ge... (18478)
【目标检测】RCNN算法详解 (17645)
【目标跟踪】KCF高速跟踪详解 (17193)
循环矩阵傅里叶对角化 (9706)
【人体姿态】Convolutional Po... (7644)
【推荐系统算法】PMF(Probab... (7369)
欧拉角、四元数和旋转矩阵 (6239)

这篇文章思路简洁，在DPM方法多年平台期后，效果提高显著。包括本文在内的一系列目标检测**算法**：[RCNN](#), [Fast RCNN](#), [Faster RCNN](#)代表当下目标检测的前沿水平，在github都给出了基于Caffe的源码。

思想

本文解决了目标检测中的两个关键问题。

问题一：速度

经典的目标检测算法使用滑动窗法依次判断所有可能的区域。本文则预先提取一系列较可能是物体的**候选区域**，之后仅在这些候选区域上提取特征，进行判断。

问题二：训练集

经典的目标检测算法在区域中提取人工设定的特征（Haar，HOG）。本文则需要训练深度网络进行**特征提取**。可供使用的有两个**数据库**：

一个较大的**识别库**（ImageNet ILSVC 2012）：标定每张图片中物体的类别。一千万图像，1000类。

一个较小的**检测库**（PASCAL VOC 2007）：标定每张图片中，物体的类别和位置。一万图像，20类。

本文使用识别库进行预训练，而后用检测库调优参数。最后在检测库上评测。

流程

RCNN算法分为4个步骤

- 一张图像生成1K~2K个**候选区域**
- 对每个候选区域，使用深度网络**提取特征**
- 特征送入每一类的SVM **分类器**，判别是否属于该类

关闭



泰国买房





【目标检测】Fast RCNN算法... (7)

【深度学习】生成对抗网络Ge... (7)

【增强学习】Recurrent Visual... (7)

人脸关键点定位.Face Alignmen... (6)

推荐文章

- * 5月书讯：流畅的Python，终于等到你！
- * 【新收录】CSDN日报—— Kotlin 专场
- * Android中带你开发一款自动爆破签名校验工具kstools
- * Android图片加载框架最全解析——深入探究Glide的缓存机制
- * Android 热修复 Tinker Gradle Plugin解析
- * Unity Shader-死亡溶解效果

最新评论

【深度学习】卷积层提速Factorized Convo...
shenxiaolu1984 : @yaogan5984:Tensorflow的tf.nn.separable_conv2d()就实现...

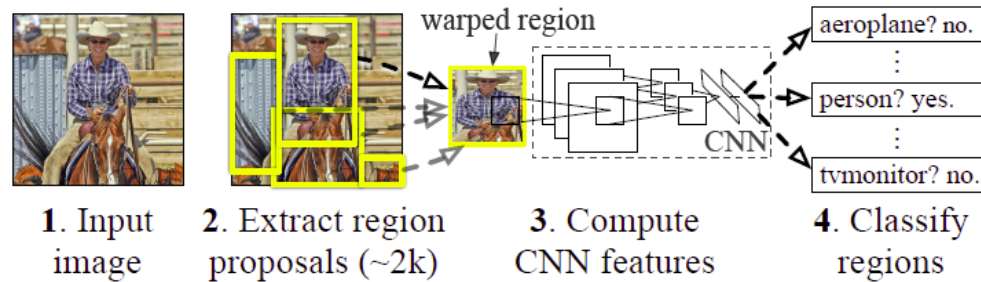
【深度学习】卷积层提速Factorized Convo...
yaogan5984 : 博主，请问下，这个有没有开源的prototxt啊？

循环矩阵傅里叶对角化
kangqi5602 : 楼主，DFT离散Fourier变换矩阵前面为何要除以根号K。还有，这里进行离散fourier变换默认...

【TensorFlow动手玩】数据导入2
无奈的小心酸 : 受教了！感谢

【优化】梯度下降 收敛性 证明

- 使用回归器精细修正候选框位置



候选区域生成

使用了Selective Search¹方法从一张图像生成约2000-3000个候选区域。基本思路如下：

- 使用一种过分割手段，将图像分割成小区域
- 查看现有小区域，合并**可能性最高**的两个区域。重复直到整张图像合并成一个区域位置
- 输出所有曾经存在过的区域，所谓候选区域

候选区域生成和后续步骤相对独立，实际可以使用任意算法进行。

合并规则

优先合并以下四种区域：

- 颜色（颜色直方图）相近的
- 纹理（梯度直方图）相近的
- 合并后总面积小的
- 合并后，总面积在其BBOX中所占比例大的

第三条，保证合并操作的尺度较为均匀，避免一个大区域陆续“吃掉”其他小区域。

例：设有区域a-b-c-d-e-f-g-h。较好的合并方式是：ab-cd-ef-gh -> abcd-efgh -> abcdefgh。
不好的合并方法是：ab-c-d-e-f-g-h -> abcd-e-f-g-h -> abcdef-gh -> abcdefgh。

关闭



泰国买房



 百度云

云计算新用户
注册送520元代金券



立即领取

验证了... 一维推导的结果确实可以直接...

循环矩阵傅里叶对角化
baidu_38400801 : 赞同, 博主的二维推导根本就是错的, 我也验证过了。但是经过matlab验证, 一维推导的结果确实可以直接...

循环矩阵傅里叶对角化
baidu_38400801 : 赞同, 博主的二维推导根本就是错的, 我也验证过了。但是经过matlab验证, 一维推导的结果确实可以直接...

第四条，保证合并后形状规则。

例：左图适于合并，右图不适于合并。



上述四条规则只涉及区域的颜色直方图、纹理直方图、面积和位置。合并后的区域特征可以直接由子区域特征计算而来，速度较快。

多样化与后处理

为尽可能不遗漏候选区域，上述操作在多个颜色空间中同时进行（RGB,HSV,Lab等）。在一个颜色空间中，使用上述四条规则的不同组合进行合并。所有颜色空间与所有规则的全部结果，在去除重复后，都作为候选区域输出。

作者提供了Selective Search的[源码](#)，内含较多.p文件和.mex文件，难以细查具体实现。

特征提取

预处理

使用深度网络提取特征之前，首先把候选区域归一化成同一尺寸227×227。

此处有一些细节可做变化：外扩的尺寸大小，形变时是否保持原比例，对框外区域直接截取还是补灰。会轻微

预训练

网络结构

基本借鉴Hinton 2012年在Image Net上的分类网络2，略作简化3。

关闭



泰国买房





此网络提取的特征为4096维，之后送入一个4096->1000的全连接(fc)层进行分类。

学习率0.01。

训练数据

使用ILVCR 2012的全部数据进行训练，输入一张图片，输出1000维的类别标号。

调优训练

网络结构

同样使用上述网络，最后一层换成4096->21的全连接网络。

学习率0.001，每一个batch包含32个正样本（属于20类）和96个背景。

训练数据

使用PASCAL VOC 2007的训练集，输入一张图片，输出21维的类别标号，表示20类+背景。

考察一个候选框和当前图像上所有标定框重叠面积最大的一个。如果重叠比例大于0.5，则认为此候选框为此标

类别判断

分类器

对每一类目标，使用一个线性SVM二类分类器进行判别。输入为深度网络输出的4096维特征，输出是否属于此

由于负样本很多，使用hard negative mining方法。

正样本



泰国买房





本类的真值标定框。

负样本

考察每一个候选框，如果和本类所有标定框的重叠都小于0.3，认定其为负样本

位置精修

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。故需要一个位置精修步骤。

回归器

对每一类目标，使用一个线性脊回归器进行精修。正则项 $\lambda = 10000$ 。

输入为深度网络pool5层的4096维特征，输出为xy方向的缩放和平移。

训练样本

判定为本类的候选框中，和真值重叠面积大于0.6的候选框。

结果

论文发表的2014年，DPM已经进入瓶颈期，即使使用复杂的特征和结构得到的提升也十分有限。本文将深度学习引入检测领域，一举将PASCAL VOC上的检测率从35.1%提升到53.7%。

本文的前两个步骤（候选区域提取+特征提取）与待检测类别无关，可以在不同类之间共用。这两步在GPU上约需13秒。

同时检测多类时，需要倍增的只有后两步骤（判别+精修），都是简单的线性运算，速度很快。这两步对于1000个类别，约需100秒。

以本论文为基础，后续的fast RCNN⁴（参看[这篇博客](#)）和faster RCNN⁵（参看[这篇博客](#)）在速度上有突飞猛进。RCNN³在PASCAL VOC上解决了多类检测问题。

1. J. Uijlings, K. van de Sande, T. Gevers, and A. Smeulders. Selective search for object recognition. IJCV, 2013. [↩](#)
2. A. Krizhevsky, I. Sutskever, and G. Hinton. ImageNet classification with deep convolutional neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2012. [↩](#)
3. 所有层都是串行的。relu层为in-place操作，偏左绘制。 [↩](#)



泰国买房





4. Girshick, Ross. “Fast r-cnn.” Proceedings of the IEEE International Conference on Computer Vision. 2015. [↩](#)

5. Ren, Shaoqing, et al. “Faster R-CNN: Towards real-time object detection with region proposal networks.” Advances in Neural Information Processing Systems. 2015. [↩](#)

顶
8

踩
0

- [上一篇](#) [【优化】logistic和softmax代价函数](#)
- [下一篇](#) [【目标检测】Fast RCNN算法详解](#)

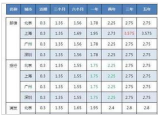
相关文章推荐

- [【目标检测】Fast RCNN算法详解](#)
 - [【目标检测】RCNN算法详解](#)
 - [【目标检测】Faster RCNN算法详解](#)
 - [【目标检测】Faster RCNN算法详解](#)
 - [【目标检测】Faster RCNN算法详解](#)
- [【目标检测】RCNN算法详解](#)
 - [【目标检测】Fast RCNN算法详解](#)
 - [【目标检测】Faster RCNN算法详解](#)
 - [【目标检测】Fast RCNN算法详解](#)
 - [【目标检测】Fast RCNN算法详解](#)

关闭



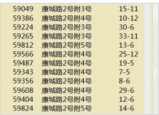
澳洲房价



存款利率



一年拿本科文



公租房摇号



oa系统



人大在职研究生



扫地



泰国买房



我的更多文章



云计算新用户
注册送520元代金券



【目标检测】Faster RCNN算法详解 (2016-04-21 15:08:06)

【目标检测】Fast RCNN算法详解 (2016-04-12 13:05:42)

参考知识库



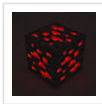
深度学习知识库

13029 关注 | 578 收录



MySQL知识库

22558 关注 | 1471 收录



算法与数据结构知识库

16473 关注 | 2320 收录

猜你在找

Oracle数据库基础入门培训课程视...

精通memcached数据库管理深度讲...

SQL Server 2008数据库基础及应用

Winform数据库编程:ADO.NET入门

搜狗郭理勇：小而美-Sogou数据库..

数据结构基础系列(1)：数据结构...

SQL Server数据库设计和开发基础...

SQLServer数据库基础

mySQL数据库深度讲解和设计SQ...

零基础学编程-数据库基础



一年拿本科文凭



文档管理系统




存款利率



无地址注册公司



日语学



山水文园二手房



美国洛杉矶房产

查看评论



betacoding

有一点没有明白，关于神经网络，是直接已经将训练好的神经网络结构和模型，针对最后的输出层进行修改即可是吗？那么这应该是所谓的模型调优吧！



will_notebook

讲解的都是干货，解惑了



竹斋

很好的文章，三篇看来不错



泰国买房



关闭



Mudwall79

1楼 2016-05-02 19:03发表

楼主这篇是我看过的最清晰的一篇RCNN介绍

您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

* 以上用户言论只代表其个人观点，不代表CSDN网站的观点或立场

核心技术类目

- 全部主题
- Hadoop
- AWS
- 移动游戏
- Java
- Android
- iOS
- Swift
- 智能硬件
- Docker
- OpenStack
- VPN
- Spark
- ERP
- IE10
- Eclipse
- CRM
- JavaScript
- 数据库
- Ubuntu
- NFC
- WAP
- jQuery
- BI
- HTML5
- Spring
- Apache
- .NET
- API
- ExtJS
- SDK
- IIS
- Fedora
- XML
- LBS
- Unity
- Splashtop
- UML
- components
- Windows Mobile
- Rails
- QEMU
- KDE
- Cas
- CloudStack
- FTC
- coremail
- OPhone
- CouchBase
- 云计算
- iOS6
- Rackspace
- Web App
- SpringSide
- Maemo
- Compuware
- aptech
- Perl
- Tornado
- Ruby
- Hibernate
- ThinkPHP
- HBase
- Pure
- Solr
- Angular
- Cloud Foundry
- Redis
- Scala
- Django
- Bootstrap



泰国买房

