

## lanyuxuan100的博客

目录视图

摘要视图

RSS 订阅

## 个人资料



lanyuxuan100

关注

发私信

访问：48831次

积分：1248

等级：BLOG &gt; 4

排名：千里之外

原创：78篇

转载：1篇

译文：0篇

评论：29条

异步赠书：Kotlin领衔10本好书 SDCC 2017之区块链技术实战线上峰会 程序员9月书讯 每周荐书：Java编程（评论送书）

## Pyramid Scene Parsing Network ( PSP NET )

2017-04-11 17:09

357人阅读

评论(0)

收藏

分享

分类：

Deep Learning ( 42 )

Computer Vision ( 8 )

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(?)

[+]

## 简介

在这篇文章中，我们通过金字塔池化与金字塔场景解析网络来实现不同区域上下文信息的整合，从而探索全局上下文信息对场景理解能有多大的帮助。

## 文章搜索

## 文章分类

- linux (17)
- Tensorflow (32)
- Python (18)
- Deep Learning (43)
- Computer Vision (9)
- caffe (8)
- cuda安装 (4)
- cudnn (4)
- segmentati (2)
- instance-s (1)
- mask-rcnn (5)
- detection (2)
- R-CNN (1)
- fast-r-cnn (1)
- faster-rcn (1)
- spp (1)
- spp-net (1)
- 车道线分割 (1)
- cuda (1)
- FCN (4)
- SegNet (1)
- machine-learning (1)
- lifelong学习 (1)

由于PSPNET提供了一个良好的场景解析先验架构，我们的全局先验表达对产生高质量的场景解析任务结果是有效的。

本文提出的方法在不同的数据集上取得了不错的效果，2016年12月在imagenet，voc2012及cityscapes上都取得了第一名。

这里VOC上的mIOU达到了85.4%，cityscapes上80.2%。

## 背景

在计算机视觉中，基于语义分割的场景解析是一项基础任务。目标就是对图像中的每个像素类。场景解析提供对一个场景的比较全面的理解，预测标签，坐标，及每个元素的形状等。这项研究在自动驾驶，机器人场景理解，安防等等好多领域都有应用，这里就不一一列举了。

而场景解析的难度，与标签及场景的复杂程度都有关；

datasets review

早些时候的场景解析任务是对2668张图片分出33个场景，LMO dataset；

稍微近一些的是PASCAL VOC数据集上，在相似的物体上提供了更加详细的标签；比如椅子和沙发，马和牛等；

最近的ADE20K dataset是最有挑战性的一个数据集，提供了一个尺度大且词汇内容丰富且类别更多。

[关闭](#)

[darknet](#) (1)[YOLO](#) (1)[SSD](#) (1)

### 文章存档

[2017年05月](#) (29)[2017年04月](#) (50)

### 阅读排行

[如何查看tensorflow的版本与安...](#) (3151)[Tensorflow中遇到OOM when al...](#) (2741)[tensorflow版SSD使用经验](#) (2543)[Tensorflow报错：AttributeError...](#) (2068)[Tensorflow使用pip安装后没有...](#) (1974)[Tensorflow报错ImportError: lib...](#) (1920)[Tensorflow指定使用哪个GPU](#) (1615)[语义分割与实例分割的区别](#) (1449)[Check failed: registry.count\(type...](#) (1448)[caffe报错Check failed: outer\\_nu...](#) (1411)

### 评论排行

[caffe报错Check failed: outer\\_nu...](#) (6)[sift-flow数据集](#) (4)[Caffe FCN Test | Check failed: e...](#) (3)[Tensorflow训练Kitti道路分割数...](#) (3)[Tensorflow训练CNN网络识别m...](#) (2)

目前大多数实时的分割都是基于FCN的，虽然深度卷积网络CNN提高了动态物体理解的能力，但是当面对多样的场景与未知的类别时容易产生问题。

例如在图2 中第一行，程序将boat错认为是car。

所以为了获得精确的场景感知，自然而然地就需要场景上下文信息来作为先验知识。对应典型的复杂场景理解，事先获得整张图像的全局特征，空间金字塔池化被应用，因为其中可以有不同的空间信息来描述整体场景。Spatial pyramid pooling network即SPP net提高了这一能力。

与原来的方法不同，为了适当的使用全局上下文信息，本文提出金字塔场景解析网络。在FCN的基础之上，将像素级别的特征扩展到全局金字塔池化的特征上。

局部与全局线索的结合，使得最终的预测更加可靠。

这篇文章也提出一种优化策略，使用深度监督loss。

这篇文章给出了所有的实现细节，以便能够按照文章所述达到很好的效果，并且公开了训练

关闭

- Tensorflow使用pip安装后没有... (2)
- 语义分割与实例分割的区别 (2)
- Check failed: registry.count(type... (2)
- Tensorflow报错ImportError: lib... (1)
- Tensorflow中遇到OOM when al... (1)

### 推荐文章

- \* CSDN日报20170828——《4个方法快速打造你的阅读清单》
- \* Android检查更新下载安装
- \* 动手打造史上最简单的 Recycleview 侧滑菜单
- \* TCP网络通讯如何解决分包粘包问题
- \* SDCC 2017之区块链技术实战线上峰会
- \* 快速集成一个视频直播功能

### 最新评论

Tensorflow报错：AttributeError: 'module' ob...  
dwnsxz : 有用!!!

tensorflow版SSD使用经验  
Mundane\_World : 你这个是linux系统吗

caffe报错Check failed: outer\_num\_ \* inner\_n...  
Darlewo : @hongzhiyang0804:遇到同样的问题，不知您是怎么解决的

caffe报错Check failed: outer\_num\_ \* inner\_n...  
hongzhiyang0804 : 楼主你好，我用的是lmdb文件做的标签，图片像素也已经进行了归一化了，可是依然出现这个问题Check...

Caffe FCN Test | Check failed: error == cuda...  
adruinoZ : @lanyuxuan100:请问怎么解决的，我训练也out of memory

Caffe FCN Test | Check failed: error == cuda...  
lanyuxuan100 : 解决了

码。

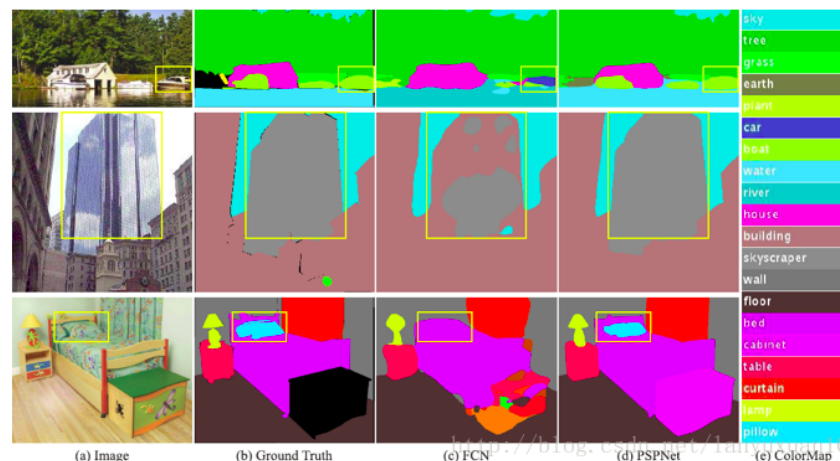
FCN等分割网络的感受野实际上没有理想中的大，只能得到局部上下文信息。

一些尺寸比较大的物体往往会超过FCN的感受野。

Mismatched Relationship

Confusion Categories

Inconspicuous Classes



### (3) 贡献

a.提出一种金字塔场景解析网络结构，为了将困难场景的上下文信息嵌入到以FCN为基础的络上。

b.提出一种深度ResNet的优化方法，使用深度loss；

c.建立一套语义分割与场景解析系统，并给出所有的实施细节；

## 相关工作

下面让我们来回顾一下场景解析与语义分割任务在近年来的研究进展：

首先，由于使用卷积层代替全连接层，受此启发，很多网络在场景解析与语义分割方面都取得了长足的

关闭

Caffe FCN Test | Check failed: error == cuda...

hukongtao1994 : 楼主最后问题解决了吗？

Check failed: registry.count(type) == 1 (0 vs. ...

qq\_38335390 : @m0\_37407756:我也遇到了这个问题，请问你解决了吗？

Check failed: registry.count(type) == 1 (0 vs. ...

向阳- : 你好，我遇到了和您一样的错误，可是我按照您做的操作了依旧报错，我在Makefile.config里 ...

caffe报错Check failed: outer\_num\_ \* inner\_n...

IGlli : 我在训练fcn时，我用的voc里的VOCS egDataLayer，也是报这个错误，我的label图是...

进步。

为了扩大网络的感受野，又使用的带孔的卷积（dilated convolution），之后又有了deconvolution。

其他工作主要沿着两个方向：一个方向是多尺度特征的集成。因为在深度网络中，高层特征包含较多的语义信息，较少的位置信息。

结合多尺度特征能够提升效果。

另外一个方向是结构预测，早期的工作使用条件随机场（CRF）来进行后处理，对分割结果进行refine。

后来一些方法将类似于CRF的结构预测机制放到了网络中；

上面两个方向都使得场景解析的结果变得更好；

为了很好的利用全局先验知识，有些方法先使用传统方法而不是深度学习网络来提取特征，然后和深度网络检测也有类似的方法。

有人证明使用全局均值池化能够提高语义分割的效果；然而在这篇文章中对ADE20K数据进行的实验表明，使用全局均值池化产生的描述子不足以代表这些上下文信息。

关闭

## 3. Pyramid Scene Parsing Network

这篇文章观察与分析了使用FCN进行解析失败的一些典型例子，

### 3.1 Important Observations

场景中存在着可视模式的共现。比如，飞机更可能在天上或者在跑道上，而不是公路上。

许多类别具有高度相似的外表。

场景中包括任意尺寸的物体，小尺寸的物体难以被识别但是有时候对于场景理解很重要

### 3.2 Pyramid Pooling Module

#### Pyramid Pooling Module

在一个深度学习网络中，感受野的大小往往可以近似的看成能够使用的上下文信息；

尽管理论上说ResNet的感受野比原始输入图像大，但是一些论文已经证明实际上的感受野

很多，尤其是在高层特征的时候。

这导致很多网络都没有充分考虑到全局上下文信息，尤其是尺度大的目标，比如道路可通行

易造成分割效果不好。

这里就通过提出一种全局先验的表达方式来解决上述问题。

由于在很多分类任务中被广泛使用，全局均值池化被成功地应用到分割任务上，考虑到ADE20K的复杂性，这一策略不足以覆盖到必要的信息。

这些场景中的像素被标注为很多材料和物体。直接将他们模糊化成一个单一维度的向量会丢失很多空间信息，从而产生二义性。

将局部区域上下文信息与全局上下文信息结合起来的方式，在解决不同类别之间的判别是很有用的。

一个更有效的表达方式是将多级不同尺度不同区域的上下文信息都包含进来；

关闭

为了去除不同子区域之间的上下文信息损失，本文采用层级的全局先验知识，包含不同尺度不同区域的信息，这个结构就叫金字塔池化模块。

## Deep Supervision for ResNet-Based FCN

越深越好的网络能够产生更好的效果。然后，加深网络结构会给优化带来困难，于是ResNet使用skip connect来解决。

辅助loss

通过添加辅助loss来对一些初始化的结果进行监督，之后学习最终loss的残差；

这样一个特别深的网络的优化问题就转换为了子问题，并且每个的学习难度都下降了。

这就是将复杂问题转换为简单问题的方法。

## Experiments

顶 踩  
0 0

- [上一篇](#) [使用Tensorflow测试自己的分割网络模型](#)

[关闭](#)

- [下一篇](#) [图像分割任务中的尺度问题](#)

#### 相关文章推荐

- 图像分割任务中的尺度问题
- [【免费】深入理解Docker内部原理及网络配置--王...](#)
- Pyramid Scene Parsing Network论文解读
- SDCC 2017之区块链技术实战线上峰会--蔡栋
- PSPnet : Pyramid Scene Parsing Network
- php零基础到项目实战
- 论文阅读理解 - Pyramid Scene Parsing Network
- C语言及程序设计入门指导
- 论文实践讨论 - Pyramid Scene Parsing Network
- Android入门实战
- SPP-net论文笔记《Spatial Pyramid Pooling in Deep C...
- 5天搞定深度学习框架Caffe
- Beyond Bags of Features: Spatial Pyramid M...
- Scene Parsing
- 关于2006\_CVPR\_Beyond Bags of Features
- Human Parsing with Contextualized Convoluti...

#### 查看评论

暂无评论

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

关闭



[公司简介](#) | [招贤纳士](#) | [广告服务](#) | [联系方式](#) | [版权声明](#) | [法律顾问](#) | [问题报告](#) | [合作伙伴](#) | [论坛反馈](#)

网站客服

杂志客服

微博客服

[webmaster@csdn.net](mailto:webmaster@csdn.net)

400-660-0108

| 北京创新乐知信息技术有限公司 版权所有 | 江苏知之为计算机有限公司 | 江苏乐知网络技术有限公司

京 ICP 证 09002463 号 | Copyright © 1999-2017, CSDN.NET, All Rights Reserved



关闭