

[CSDN首页 \(http://www.csdn.net?ref=toolbar\)](http://www.csdn.net?ref=toolbar)[学院 \(http://edu.csdn.net?ref=toolbar\)](http://edu.csdn.net?ref=toolbar)[下载 \(http://download.csdn.net?ref=toolbar\)](http://download.csdn.net?ref=toolbar)

更多 ▼

[下载 CSDN APP \(http://www.csdn.net/app?ref=toolbar\)](http://www.csdn.net/app?ref=toolbar)[写博客 \(http://write.blog.csdn.net/postedit?ref=toolbar\)](http://write.blog.csdn.net/postedit?ref=toolbar)[登录 \(https://passport.csdn.net/account/login?ref=toolbar\)](http://blog.csdn.net/) | [注册 \(http://passport.csdn.net/account/mobile/register?ref=toolbar&action=mobileRegister\)](http://passport.csdn.net/account/mobile/register?ref=toolbar&action=mobileRegister)[首页 \(http://blog.csdn.net/\)](http://blog.csdn.net/)[移动开发 \(http://blog.csdn.net/mobile/newarticle.html\)](http://blog.csdn.net/mobile/newarticle.html)

全部 ▼

CSDN (<http://www.csdn.net>)

目录



喜欢



收藏



评论



分享

DeepLab：深度卷积网络，多孔卷积 和全连接条件随机场 的图像语义分割 Semantic Image Segmentation with Deep Convolutional Nets, Atro

2017年05月04日 19:08:08

979

0

0

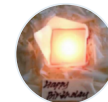
深度卷积网络，多孔卷积 和全连接条件随机场 的图像语义分割

DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs
(<https://arxiv.org/pdf/1606.00915.pdf>)

Taylor Guo (www.qiqitek.com), 2017年5月03日 星期三

摘要

本文的主要任务是深度学习的图像语义分割，主要有3个方面的贡献，有重要的实践价值。首先，用上采样滤波器进行卷积，或“多孔卷积”，是稠密预测任务中的强有力的工具。多孔卷积可以在深度卷积网络中计算特征响应时，明确控制图像分辨率。可以有效地增大滤波器的视野，合成更多内容，而不会增加参数数量或计算量。第

新新大熊 (<http://blog.csdn.net/cicibabe>)

+ 关注

<http://blog.csdn.net/cicibabe>

码云

未开通

原创

粉丝

喜欢

(https://github.com/cicibabe/DeepLab)

14

24

0

他的最新文章

[更多文章 \(http://blog.csdn.net/cicibabe\)](http://blog.csdn.net/cicibabe)

Matterport3D：室内环境RGB-D数据的深度学习
([cicibabe/article/details/78078513](http://blog.csdn.net/cicibabe/article/details/78078513))

React Native 学习笔记（三）---自上而下，直觉理解
([cicibabe/article/details/77947945](http://blog.csdn.net/cicibabe/article/details/77947945))

React Native学习笔记（二）---运行开源App，形成直观感受
([cicibabe/article/details/77921913](http://blog.csdn.net/cicibabe/article/details/77921913))



返回顶部

编辑推荐

最热专栏

二，提出了多孔空间金字塔池化（ASPP），可以在多尺度上鲁棒地分割物体。多孔空间金字塔池化ASPP用多种采样率和视野上的滤波器探测进入的卷积特征层，因此可以在多个尺度上捕捉物体和图像内容。第三，合并深度卷积网络和概率图模型方法，增强对物体边界的定位。通常在卷积神经网络中同时采用最大池化和降采样方法获得不变性，但在定位精度上会有损失。在深度卷积神经网络最后一层带有全连接的条件随机场，合并所有响应结果，克服这个问题，定量地和定性地增强了定位的性能。我们设计的DeepLab系统刷新了PASCAL VOC-2012图像语义分割任务的记录，在测试集上达到79.7% mIOU，在其他3个数据集上的结果也相近：PASCAL-Context，PASCAL-Person-Part，Cityscapes。我们的代码已经开源。



目录



喜欢



收藏



评论



分享

1 简介

深度卷积神经网络将计算机视觉系统的性能推到了新的高度，应用于各种高层应用问题，包括图像分类、物体检测，深度卷积神经网络端到端的训练方式得到了更好的结果，明显优于那些需要手工处理特征的系统。这一成功的关键是深度卷积神经网络内在的对局部图像变换的不变性，可以学习抽象数据表示，如论文13。这种不变性是分类任务所需要的，但可能会妨碍密集的预测任务，比如语义分割，不需要抽取空间信息。

图像语义分割中有3种挑战：（1）特征分辨率减少，（2）不同尺度下的物体的存在状况，（3）由于深度卷积神经网络的不变性造成的定位精度减少。接下来，我们会讨论这些问题，还有Deeplab系统如何克服这些问题。

第一个挑战是由重复最大池化和降采样（步长跨度）造成的，它们作用于用来做图像分类的深度卷积神经网络的一系列层上。深度卷积神经网络采用全卷积方式的时候，会明显降低特征地图的空间分辨率，如论文14。为了解决这个问题，有效地生成更稠密的特征地图，我们从深度卷积神经网络中最后的最大池化层中拿掉降采样操作，将上采样滤波器加入接下来的卷积层中，在更高的采样率上计算特征地图。滤波上采样就是在非零滤波器之间插入空洞。这一技术在信号处理中有悠久的历史，最开始是用于多孔算法中有效计算下采样小波变换，如论文15所示。我们使用了多孔卷积作为上采样滤波卷积运算的简称。这种思想已经在深度卷积神经网络中应用了，如论文3,6,16。在实践中，可以用多孔卷积恢复全分辨率的特征地图，计算更稠密的特征地图，接着在原图像大小和特征响应之间做一个简单的双线性插值。这种算法提供了简单但强有力的方法在稠密预测任务中使用反卷积层，如论文13,14。与通常的大滤波器卷积相比，多孔卷积可以有效地增加滤波器的视野而不会增加参数数量或计算量。

第二个挑战是物体在多尺度图像中的状态造成的。一个处理这个问题的标准方法是给深度神经网络提供这个图像的重新缩放的版本，再合成特征或计算地图分值，如论文6,17,18所示。我们发现这个方法确实可以增加系统的性能，但是需要在所有深度卷积神经网络的层上对输入图像的多尺度版本计算特征响应。受到空间金字塔池化（如

FCIS+code | Fully Convolutional Instan...

DeepLab: Semantic Image Segmentati...

【Deep Learning】DeepLab (/chenyj9...

论文阅读笔记：图像分割方法deeplab以..

语义分割DeepLab v2--DeepLab: Sema...

在线课程



Presto的服务治理与架构在京东的实践与应用
(http://edu.csdn.net/course/series_detail/64?utm_source=blog9)



深入掌握Kubernetes应用实践
(http://edu.csdn.net/course/detail/6080?utm_source=blog9)



返回顶部

论文19,20) 的激发，采用计算效率更好的算法，在多个采样率上重采样特定的特征层来做卷积。这需要用互补有效视野的多个滤波器来探测原图像，因此可以在多个尺度上有用的图像中捕捉物体。与真正对特征重采样不同，我们用不同采样率的多个并行的多孔卷积层做这种映射；我们称这种技术为“多孔空间金字塔池化（ASPP）”。

第三个挑战是物体分类器要求对空间变换具有不变性，内在地限制了深度卷积神经网络的空间精度。减轻这个问题的一个方法是在计算最终的分割结果时，如论文14,21，跳层从多个网络层中提取“超列”特征。需要提出的是，我们采用全连接条件随机场CRF，如论文22，提升了模型的能力可以抓取精细的细节。条件随机场广泛地应用于语义分割中，合并从局部像素和边缘（论文23,24），或超像素（论文25）中获得低阶信息的多路分类器中计算出来的分类分值。构建层次依赖模型（论文26,27,28），和/或分割用的高层依赖信息（论文29,30,31,32,33）的复杂度日益增加，我们使用了论文22提出的全连接配对条件随机场，计算效率更高，可以抓取精细的边缘细节，也适用于较长的依赖项。论文22中的模型增强了基于提升的像素级别分类器的性能。本文工作中，我们展示了，当它与深度卷积神经网络的像素级别分类器耦合的时候，可以取得更好的结果。

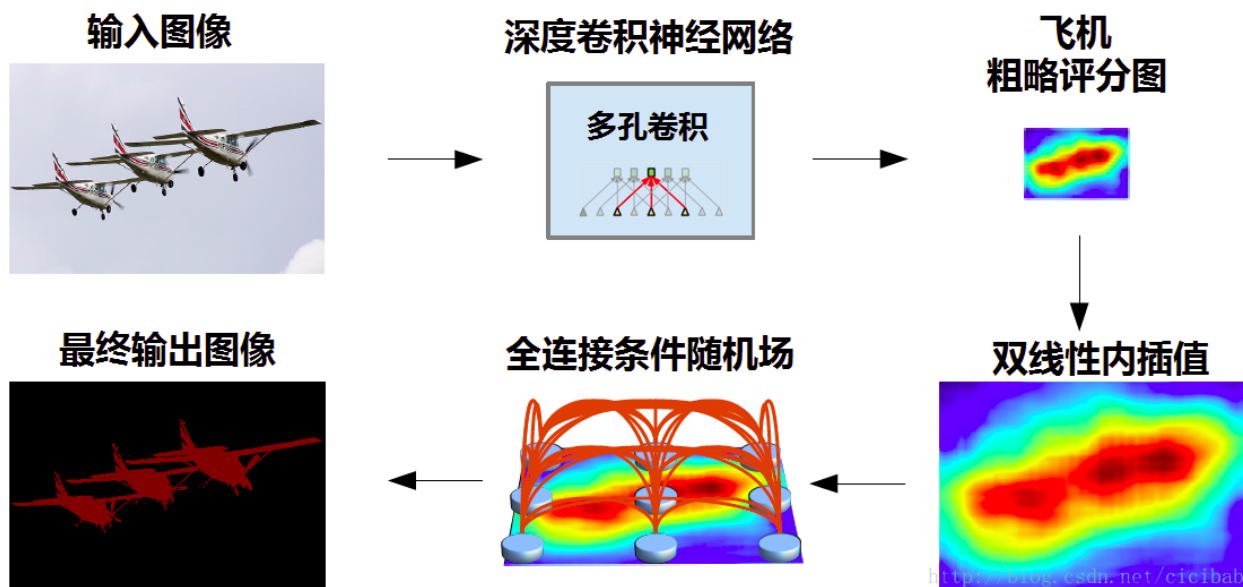


图1：模型演示。一个深度卷积神经网络，比如VGG-16或ResNet-101，采用全卷积的方式，用多孔卷积减少信号降采样的程度（从32x降到8x）。在双线性内插值阶段，增大特征地图到原始图像分辨率。接下来，用条件随机场优化分割结果，更好的抓取物体边缘。

目录
喜欢
收藏
评论
分享

返回顶部

DeepLab模型的高层示意图，如图1所示。一个深度卷积神经网络（本文用VGG-16（论文4），或ResNet-101（论文11））训练用于图像分类任务的，可以修改用于语义分割：（1）将所有全连接层改为卷积层（比如，论文14的全卷积网络），（2）通过多孔卷积增加特征分辨率，就可以计算每8个像素的特征响应而不是原始网络中的每32个像素。再使用双线性内插值，用因子8上采样分地图，以达到原始图像的分辨率，得到从输入到一个全连接的条件随机场，用于优化分割结果。

从实践观点看，DeepLab系统有3个主要的优点：（1）速度：由于多孔卷积的优点，稠密深度卷积神经网络在nVidia Titan X GPU上的运行速度是8FPS，全连接条件随机场的平均场推导在CPU上需要0.5秒。（2）精度：在几个数据集中获得了最好的结果，包括PASCAL VOC-2012语义分割对比（论文34），PASCAL-Context（论文35），PASCAL-Person-Part（论文36）和Cityscapes（论文37）。（3）简洁：系统串联了两个构建好的模块，深度卷积神经网络和条件随机场。

本文的DeepLab系统与在ICLR会议上发布的第一版相比有一些改进。新版本通过多尺度输入处理（论文17,39,40）或者多孔空间金字塔池化，可以更好地分割物体。我们构建了DeepLab的剩余净变体，DeepLab采用最新的ResNet图像分类深度卷积神经网络构建，与原来基于VGG-16的网络相比，取得了更好的语义分割性能。最后，我们做了全面实验评估多个模型变量和一些数据集评测，除了PASCAL VOC 2012之外，还包括其他数据集。用Caffe框架来实现这些方法。在DeepLab主页 (<http://liangchiehchen.com/projects/DeepLab.html>)上分享了代码和模型：<http://liangchiehchen.com/projects/DeepLab.html> (<http://liangchiehchen.com/projects/DeepLab.html>)。

2 相关工作

之前开发的大多数成功的语义分割系统都是采用手工特征的单调分类器，比如，提升方法（论文24），随机森林（论文43），或支持向量机（论文44）。通过整合更丰富的信息，内容（论文45）和结构化的预测技术（论文22,26,27,46），取得了实质性的进展，但这些系统的性能受限于特征的表达能力。过去几年里，深度学习在图像分类上取得了突破性的进展，迅速转移到语义分割中。既然这个任务包含了分割和分类，那么一个中心问题是如何合并这两个任务呢。

第一类深度卷积神经网络的语义分割主要采用自下而上的串联图像分割，再串联深度卷积神经网络区域分类。比如，论文7，49中使用了论文47,48的边界框指示和区域遮罩，作为输入提供给深度卷积神经网络，整合形状信息提供给分类流程。类似的，论文50使用超像素表示。尽管这些方法可以从良好的分割产生的形状边缘受益，但它们无法从错误中恢复。

第二类方法用卷积计算的深度卷积神经网络特征用于密集图像标注，将独立获取的分割结合在一起。在第一阶段中论文39在多个图像分辨率上使用深度卷积神经网络，再使用分割树平滑预测结果。最近，论文21建议使用跳层，并联结深度神经网络中计算的中间特征地图用于像素分类。还有，论文51建议用局部方案池化中间特征地图。这些工作仍然是采用的从深度卷积分类器的结果中解耦合的分割算法，因此有提早做决策的风险。

第三类工作使用深度卷积神经网络直接提供类别级像素标注，甚至可以放弃分割。论文14，52的免分割方法直接用全卷积方法将深度卷积神经网络应用到整幅图像上，将深度卷积神经网络的最后全连接层转换成卷积层。为了处理前面简介中提出的空间定位问题，论文14在中间特征地图上上进行上采样，连接分值，论文52，将生成粗略的结果放入另外一个深度卷积神经网络上，从粗到细优化预测结果。在这些工作基础上，对特征分辨率进行控制，引入多尺度池化技术，在深度卷积神经网络的最上层集成稠密连接的条件随机场（论文22）。我们发现这会产生明显好的分割结果，尤其是在物体边缘。深度卷积神经网络和条件随机场当然不是最新的，但之前的工作只是尝试了局部连接条件随机场模型。具体来说，论文53用条件随机场作为基于深度卷积神经网络的重排序系统，论文39将超像素作为结点用于局部配对条件随机场，用图切割用于离散推理。这样，它们的模型受限于超像素计算的误差，或者忽略了长距离依赖。我们的方法将每个像素看做是条件随机场的结点，用于接收卷积神经网络的一元势能。关键是，论文22中的全连接条件随机场中的高斯条件随机场势能可以抓取远程依赖，同时模型服从快速平均场推理。我们注意到平均场推理用于传统图像分割任务，论文54，55,56，但这些老的模型受限于短程连接。在一些独立的工作中，论文57使用了一个非常相似的密集连接条件随机场模型优化深度卷积神经网络的结果，用于物料分类。然而，论文57的深度卷积神经网络模块是用稀疏点监督上训练的，而不是每个像素点的稠密监督。

自从第一版发布后，语义分割领域进展很快。多个组都取得了重要进展，提升了PASCAL VOC-2012语义分割的基准，PASCAL性能评估 (<http://host.robots.ox.ac.uk:8080/leaderboard/displaylb.php?challengeid=11&compid=6>)上比较活跃，如论文17，40,58,59,60,61,62,63。

<http://host.robots.ox.ac.uk:8080/leaderboard/displaylb.php?challengeid=11&compid=6>

(<http://host.robots.ox.ac.uk:8080/leaderboard/displaylb.php?challengeid=11&compid=6>)

有趣的是，大部分性能优秀的方法都在用了DeepLab系统的一个或两个关键模块：高效的稠密特征提取的多孔卷积，和全连接条件随机场的原始深度卷积神经网络分值的优化。我们列出了最重要和有趣的进展。

结构化预测的端到端训练，已经应用于最近的工作中了。我们将条件随机场用于后端处理方法，论文40,59,62,64,65，已经成功地采用深度卷积神经网络和条件随机场。具体来说，论文59,65，将条件随机场平均场推导步骤展开，将整个系统转换成一个端到端的可训练的前馈网络，论文62,用卷积滤波器近似估计稠密条件随机场平均场推理的一次迭代。另外一个论文40,66研究的成果丰富的方向，是学习成对的深度卷积神经网络的条件随机

☰
目录

♥
喜欢

🔖
收藏

💬
评论

🔗
分享

↑
返回顶部

场，明显地增强了性能但计算代价巨大。在不同的方向上，论文63用更快速的域变换模块（论文67）取代平均场推理中的双向滤波模块，加快了速度，降低了整个系统的内存要求，论文18,68，将语义分割和边缘检测合并在一起。

弱监督在很多论文里面都有，在整个训练集上都有像素级别的语义标注，如论文58，69,70,71，结果明显优于弱监督的(预训练)深度卷积网络，如论文72。另外一条研究线，论文49,73，采用了实体分割，同时处理物体识别和语义分割。



目录



喜欢



收藏



评论



分享

多孔卷积最初是用于高效计算论文15中的多孔算法的非抽取小波变换。有兴趣的读者可以阅读论文74了解小波。

多孔卷积在信号处理中也指“noble identities”（是多采样率系统中用于将抽取器或插值器移动到合适位置），构建

在输入信号和滤波采样率相互作用基础上的，如论文75。我们在论文6中使用了多孔卷积这个词。相同的操作，在

论文76中称为空洞卷积，常规卷积进行上采样（或论文15中的空洞加宽）滤波。之前很多作者都在深度卷积神经

网络中使用相同的操作用于稠密特征提取，如论文3,6,16。处理分辨率的增强外，多孔卷积还可以增大滤波器的感

受野，获得更多的内容，论文38中所示，受益很多。这种方法后来论文76也采用了，使用了一系列的采样率逐步

增大的多孔卷积层集合多尺度内容。多孔空间金字塔池化算法抓取多尺度上的物体和内容，也采用了多个不同采

样率的多个多孔卷积层，是并行布局而不是串行。有趣的是，多孔卷积技术也应用于其他任务，比如物体检测(论

文12,77)，实体分割(论文78)，视觉问答(论文79)，和光流法(论文80)。

如我们预期的，将更高级的图像分类卷积神经网络整合进DeepLab，比如，论文11的深度残差网络，可以生成更好的结果。论文81也发现了这个结论。

3 方法

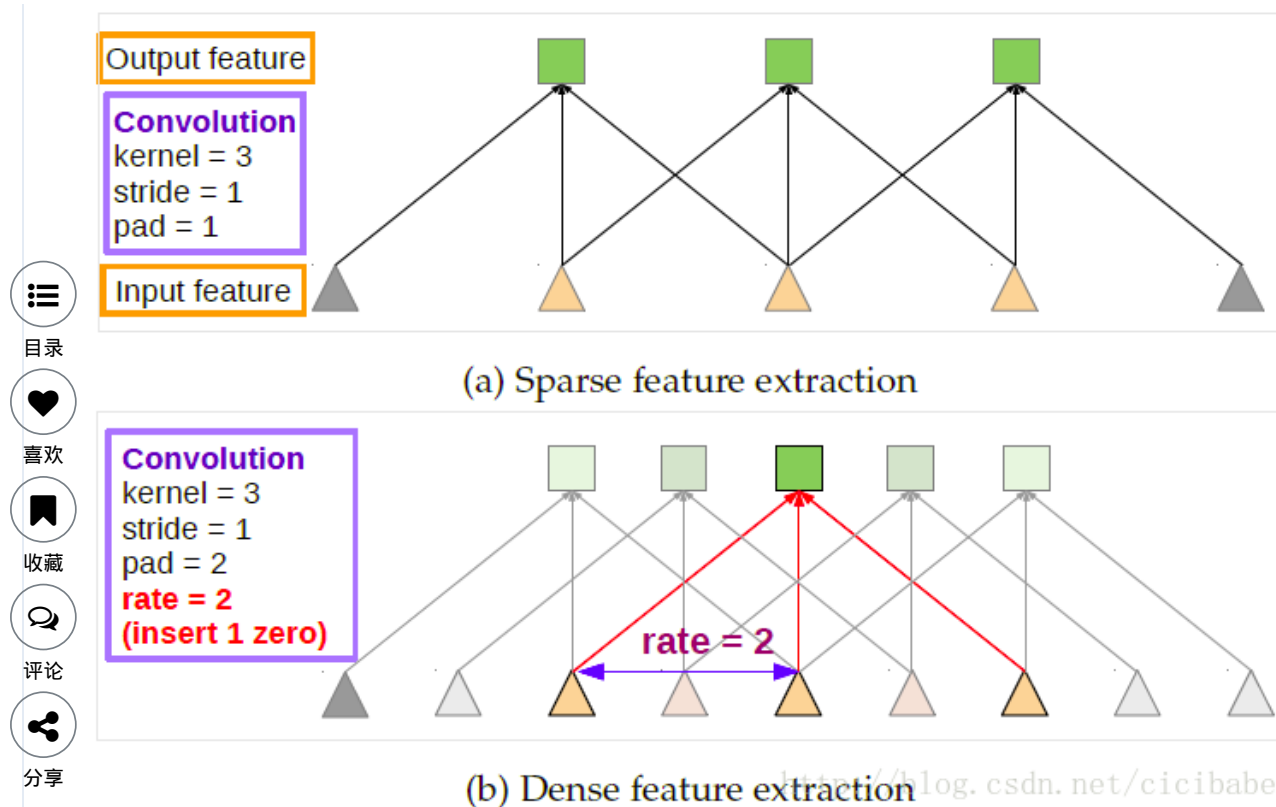
3.1 用于稠密特征提取和视野增大的多孔卷积

深度卷积神经网络的语义分割或其他稠密预测任务，简洁成功地应用了全卷积模式，如论文3,14。然而，这些网络连续层上重复合并最大池化和跨步明显减少了生成特征地图的空间分辨率，典型的最新深度卷积神经网络上每个方向的因子是32。论文14的补救方法是用去卷积层，需要更多的内存和时间。

我们使用多孔卷积，最开始是在多孔算法中提高非抽取小波变换的计算效率。之前也应用于深度卷积神经网络，如论文3,6,16。这个算法使我们可以在任何期望的分辨率上计算任何层的响应。网络训练好后，也可以无缝与训练过程整合。



返回顶部



(b) Dense feature extraction <http://blog.csdn.net/cicibabe>

图2：多孔卷积的一维展示。（a）低分辨率输入特征地图上采用标准卷积的稀疏特征提取。（b）应用于高分辨率输入特征地图，多孔卷积比例 $r=2$ 的稠密特征提取。

先考虑一维信号处理的情况，一维输入信号 $x[i]$ ，滤波器为 $w[k]$ ，长度为 k 。多孔卷积的输出为 $y[i]$ ：

$$y[i] = \sum_{k=1}^K x[i + r \cdot k] w[k] \quad (1)$$

比例参数 r 对应输入信号的采样步长。标准卷积是 $r=1$ 的特例。如图2所示。

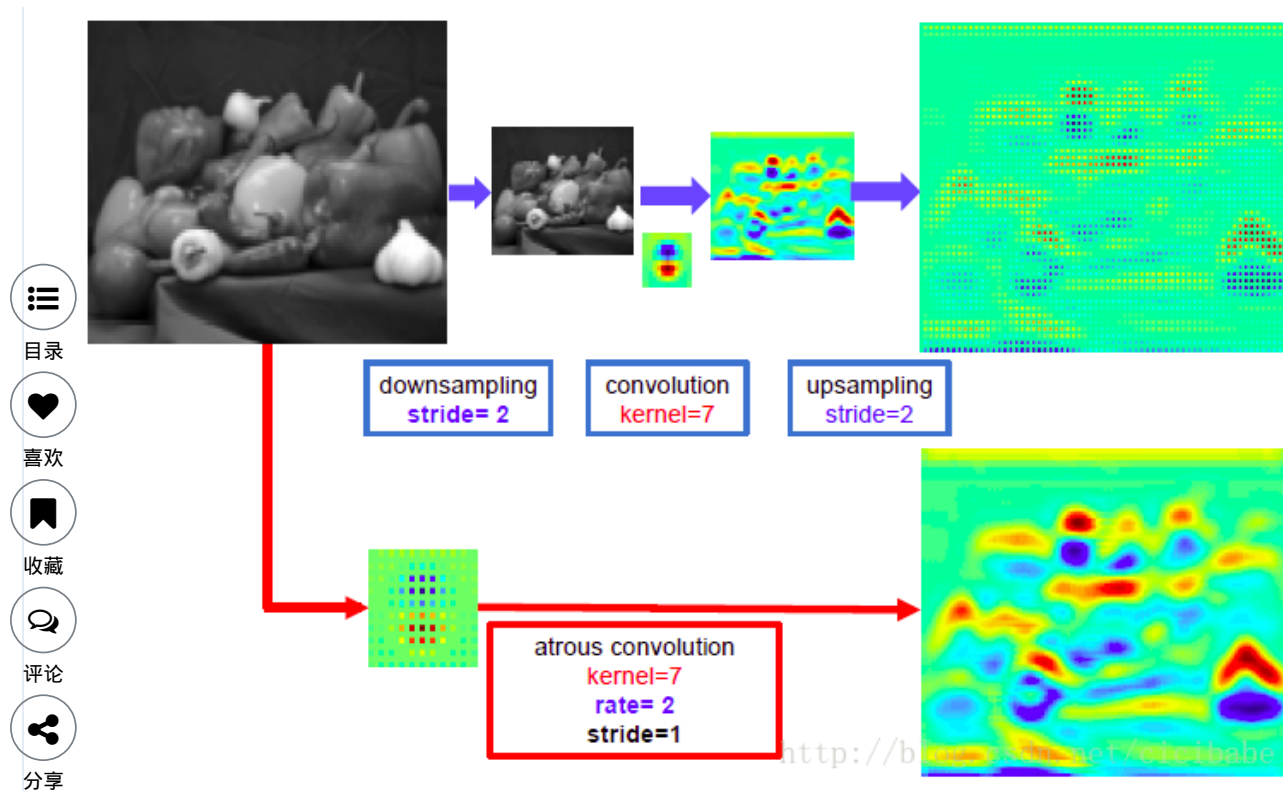


图3：二维多孔卷积示意图。上排：低分辨率输入特征地图的标准卷积稀疏特征提取。下排：高分率输入特征地图多孔卷积比例 $r=2$ 的稠密特征提取。

图3展示了2D算法操作的一个简单例子：给定一幅图像，先用因子为2的降采样减小分辨率，再用一个核卷积运算，这里是垂直高斯导数（滤波器）。如果将生成的特征地图植入原始图像坐标中，只能在图像位置的1/4处得到响应。相反，如果用空滤波器对全分辨率图像做卷积运算，就可以计算所有图像位置的响应，可以用因子2上采样原始滤波器，在滤波器的值之间填入0。尽管有效滤波器的大小增加了，只需要处理非零滤波值就可以了，因此滤波器参数和每个位置的操作数量保持不变。生成流程可以简单明确地控制神经网络特征响应的空间分辨率。

在深度卷积神经网络中可以使用层链中的多孔卷积，可以在任意高的分辨率上计算最后的深度卷积神经网络响应。比如，为了对VGG-16或ResNet-101网络中计算出来的特征响应的空间密度翻倍，我们发现最后的池化或卷积层会减小分辨率（分别是'pool5'或'conv5_1'），设置步长为1以避免信号抽取，用比例 $r=2$ 的多孔卷积层取代所有

后续的卷积层。这个方法应用于整个网络，就可以在原始图像分辨率上计算特征响应，但这样成本比较高。我们可以采用混合方法，达到一个比较好的效率/精度平衡，用因子为4的多孔卷积增加计算的特征地图和密度，用另外一个因子为8的快速双线性插值在原始图像分辨率上恢复特征地图。双线性插值在这样的配置中足够了，因为分类得分值地图（对应数概率）非常平滑，如图5所示。与论文14的去卷积方法不同，本方法将图像分类网络转换成稠密特征提取器，不需要学习更多的参数，得到更快的深度卷积神经网络训练。

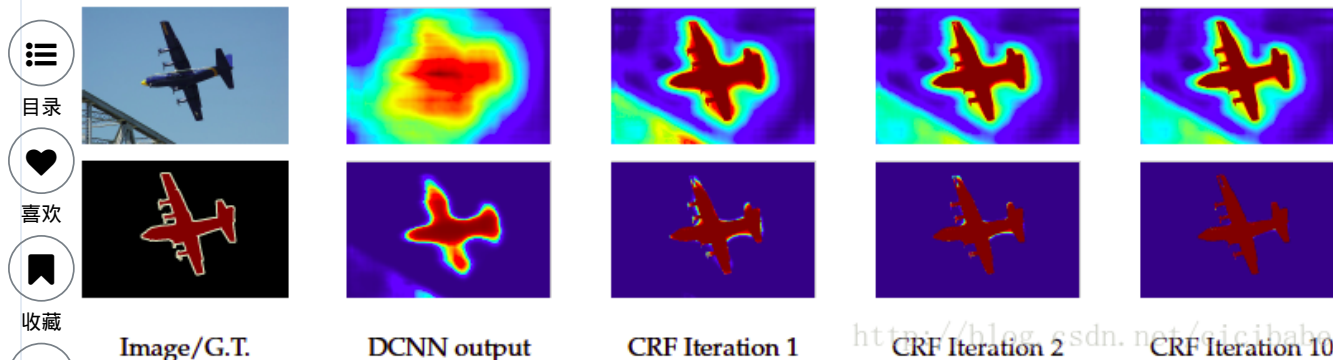


图5：飞机的范围地图（做柔性最大函数之前的输入）和置信地图（柔性最大函数之后的输出）。我们提供了每次平均场迭代后的分类得分地图（第一排）和置信地图（第二排）。卷积神经网络的最后层输出作为平均场推理的输入。

分享 多孔卷积可以在任何深度卷积神经网络层上任意扩大滤波器的视野。最新的深度卷积神经网络采用空间小卷积核（通常是 3×3 ），可以保持计算量和参数数量保持不变。比例为 r 的多孔卷积在连接滤波值之间加入 $r-1$ 个0零值。有效地将一个 $k \times k$ 大小的滤波核扩大到 $k_e = k + (k-1)(r-1)$ 而不会增加参数数量和计算量。因此，提供了一个有效的机制控制视野，在精确定位（小视野）和语境同化（大视野）之间取得最好的平衡。我们成功地做了实验：DeepLab-LargeFOV模型变量（论文38）采用了比例 $r=12$ 的多孔卷积，取得了显著的性能，第4章会详细说明。

在实现上，有两种方法可以有效地执行多孔卷积。第一种是插入空洞（零元素）或者对输入特征地图同等稀疏地采样，来对滤波器进行上采样。如论文15所示，我们在早期工作中，论文6,38中执行了这样的方法，接着是论文76，在Caffe框架内，添加到im2col函数中（它从多通道特征地图中提取了量化区域）稀疏采样内部特征地图。第二种方法，最初由论文82提出，论文3,16中也有应用，用一个与多孔卷积比例 r 相同的因子对输入特征地图进行子采样，对每个 $r \times r$ 可能的偏移，消除隔行扫描生成一个 r^2 降低的分辨率地图。接下来对这些中间地图使用标准卷积，隔行扫描生成原始图像分辨率。将多孔卷积转换成常规卷积，这样我们就可以使用现有优化好的卷积例程。我们已经用TensorFlow框架实现了第二种方法。

3.2 用多孔空间金字塔池化的多尺度图像表示

深度卷积神经网络表现出了卓越的性能可以表示尺度，只需要简单地在包含多个尺度物体的数据集上进行训练就可以了。明确的物体尺度可以增强深度卷积神经网络的处理大小物体的能力（论文6）。

我们实验了两种方法处理语义分割上的尺度变换。第一种方法是标准多尺度处理（论文17,18）。我们用共享相同参数的并行深度卷积神经网络分支，从多个（我们的实验里有3个）原始图像不同尺度版本，提取深度卷积神经网络得分地图。为了生成最后的结果，我们从并行的深度卷积神经网络分支到原始图像分辨率之间对特征地图进行双线性插值，并且融合它们，在不同尺度上获得每个位置的最大响应。在训练和测试时都这样做。多尺度处理明显增强了性能，但是需要对输入图像的多个尺度上在所有深度卷积神经网络层上计算特征响应。

第二种方法受论文20中R-CNN空间金字塔池化方法成功的启发，任意一个尺度上的区域都可以用在这个单一尺度上重采样卷积特征进行精确有效地分类。我们实现了他们方法的一种变化形式，使用多个不同采样率上的多个并行多孔卷积层。每个采样率上提取的特征再用单独的分支处理，融合生成最后的结果。前面提到的“多孔空间金字塔池化”（DeepLab-ASPP）方法泛化了DeepLab-LargeFOV，图4有演示。



目录



喜欢



收藏



评论



分享



返回顶部

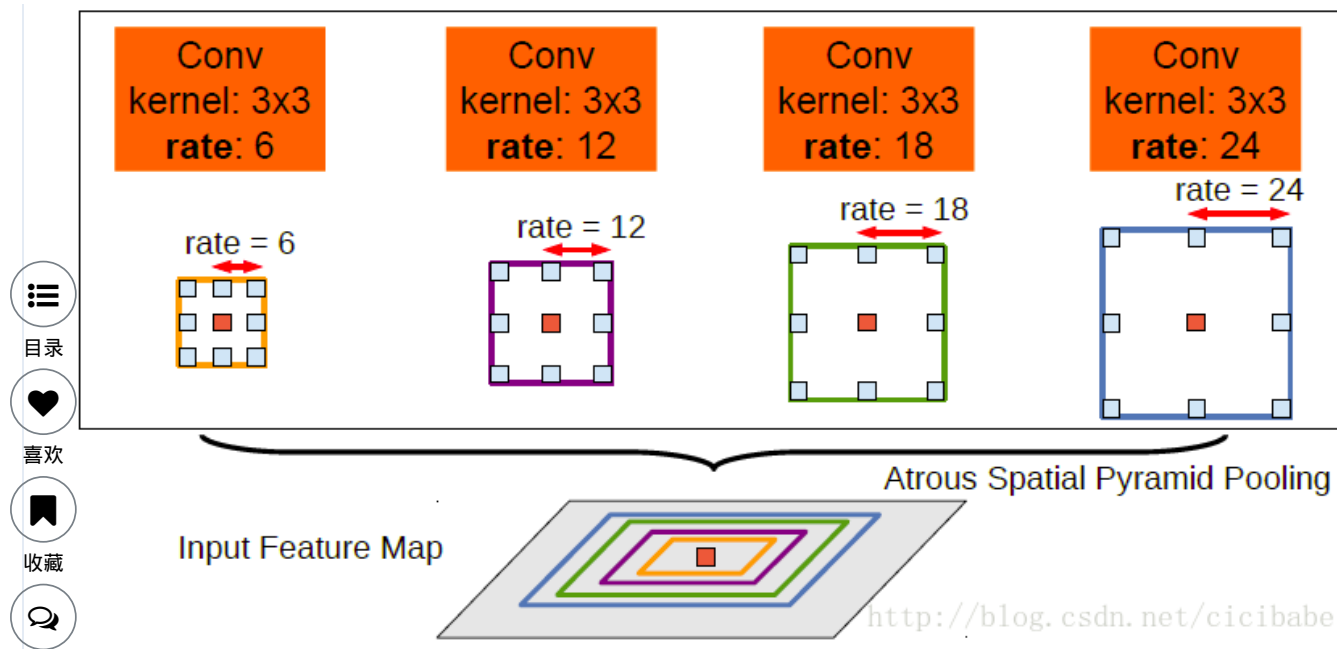


图4：多孔空间金字塔池化（ASPP）。为了分类中间像素（橙色），ASPP用不同采样率的多个并行滤波器开发了多尺度特征。视野有效区用不同的颜色表示。

3.3 用于精确边界恢复的全连接条件随机场的结构预测

精确定位和分类性能好像是深度卷积神经网络中天生的平衡关系：带有最大池化层的更深的模型在分类任务上非常成功，但不变性增加，顶层结点的大感受野只能产生平滑响应。如图5所示，深度卷积神经网络得分地图可以预测是否有物体，以及物体出现的大致位置，但不能真正描绘它们的边界。

之前的工作在两个方面应对这种定位挑战。第一种方法是从卷积网络的多层获取信息，为了更好地估计物体边界，论文14,21,52。第二种方法采用超像素表示，将定位问题表示成一个低阶分割方法，如论文50。

我们将深度卷积神经网络的识别能力和全连接条件随机场优化的定位精度耦合在一起，非常成功地处理定位挑战问题，生成了精确的语义分割结果，在一个详细的层级上恢复物体边界，超越了现有方法。我们的第一版工作论文38发布后，在这个研究方向上有很多后来跟进的论文，如17,40,58,59,60,61,62,63,65。

传统方法中，条件随机场用于平滑带噪声的分割图，论文23,31。通常，这些模型将邻近结点耦合，有利于将相同标记分配给空间上接近的像素。定性的说，这些短程条件随机场主函数会清除构建在局部手动特征上层弱分类器的错误预测。

与这些弱分类器相比，现代深度卷积神经网络架构，比如像我们在本文中用的，生成得分地图和定性不同的语义标记预测。如图5所示，得分地图通常非常平滑，生成相同的分类结果。在这种情形下，用短程条件随机场可能就不好了，我们的目标是要恢复详细的局部结构而不是平滑它。用局部条件随机场关联中的反差灵敏势，可以增强定位，但还是会漏掉细小结构，通常都需要处理离散优化问题。



目录

为了克服短程条件随机场的局限，我们在系统中整合了论文22的全连接随机场模型。模型使用了如下能量函数：



喜欢

$$E(x) = \sum_i \theta_i(x_i) + \sum_{ij} \theta_{ij}(x_i, x_j) \quad (2)$$



收藏

其中 x 是像素分配的标记。我们将其用作单点势能 $\theta_i(x_i) = -\log P(x_i)$ ，其中 $P(x_i)$ 是深度卷积神经网络计算的像素 i 处的标记分布概率。一对势能有相同的形式，可以用全连接图进行推理，比如，



评论

连接图像像素所有配对， i, j 。具体来说，如论文22，我们用如下表达式：



分享

$$\theta_{ij}(x_i, x_j) = \mu(x_i, x_j) \left[w_1 \exp \left(-\frac{\|p_i - p_j\|^2}{2\sigma_\alpha^2} - \frac{\|I_i - I_j\|^2}{2\sigma_\beta^2} \right) + w_2 \exp \left(-\frac{\|p_i - p_j\|^2}{2\sigma_\gamma^2} \right) \right] \quad (3)$$

其中，如果 $x_i \neq x_j$ ， $\mu(x_i, x_j) = 1$ ，或者0；在波茨模型中，只有显著标记的结点才会惩罚。表达式的剩下部分用了两个不同特征空间的高斯核；第一个是像素位置（记为 p ）和RGB颜色（记为 I ）间的双向核，第二个核是像素位置。超参数 $\sigma_\alpha, \sigma_\beta, \sigma_\gamma$ 控制高斯核的尺度。第一个核强制相同和位置的像素具有相同的标记，第二个核在强制平滑时只考虑空间上的接近程度。

关键是，这个模型可以有效近似概率推理。全分解平均场估计 $b(x) = \prod_i b_i(x_i)$ 的信息传递更新可以表示成双边空间下的高斯卷积。高维滤波算法明显地加速了这个计算过程，使算法在实际中非常快，用论文22的实现方法，在PASCAL VOC图像上平均少于0.5秒。



返回顶部

4 实验结果

我们调好了VGG-16和ResNet-101网络的模型权重用于ImageNet预训练，按照论文14的步骤，直接用在语义分割任务中。我们换掉了最后一层ImageNet分类器，使用的语义分类的目标数量（包括背景）。损失函数是深度卷积神经网络输出图（相比原图的8倍子采样）中每个空间位置的交叉熵的和。在整个损失函数中，所有的位置和标记都分配了相同的权重（除了未标记的像素被忽略外）。我们的目标是基准标记（子采样为8）。我们用论文2中标准的随机梯度下降算法优化了所有网络层中相对目标函数的权重。我们解耦合了深度卷积神经网络和条件随机场训练步骤，设置条件随机场参数时假定深度卷积神经网络的一元项是确定的。

我们在4个数据集上评估了模型：PASCAL VOC 2012，PASCAL-Context，PASCAL-Person-Part和Cityscapes。我们先报告第一版在PASCAL VOC2012的主要结果，然后是所有数据集上的最新结果。

4.1 PASCAL VOC 2012

数据集 PASCAL VOC 2012分割对标，论文34，包含20个前景物体类和1个背景类。原始的数据集包含1464（训练集）1449（验证集）和1456（测试集）像素级别的标记图像用于训练，验证和测试。数据集又用论文85的标注进行加强，生成10582（训练集增强）训练图像。性能的测量，用21个类别的平均单位中的像素交叉（IOU）来衡量。

4.1.1 第一版结果

我们使用了VGG-16在ImageNet上预训练，用于3.1节所说的语义分割。用了20个mini-batch，初始化学习率为0.001（最后分类层的是0.01），每2000次迭代学习率乘以0.1。动量是0.9，权重衰减是0.0005。

深度卷积神经网络在训练集增强集上调好后，我们用论文22方法交叉验证了条件随机场参数。我们使用了缺省值 $w_2=3$ ， $\sigma_\gamma=3$ ，在验证集上交叉验证了100个图像，搜索到 w_1 ， σ_α ， σ_β 的最优值。我们使用了从粗到细的搜索方法。参数的最初搜索范围是 $w_1 \in [3:6]$ ， $\sigma_\alpha \in [30:10:100]$ 和 $\sigma_\beta \in [3:6]$ （Matlab记号），再在第一轮最好的值附近优化搜索步长大小。使用了10个平均场迭代。



目录



喜欢



收藏



评论



分享



返回顶部

Kernel	Rate	FOV	Params	Speed	bef/aft CRF
7×7	4	224	134.3M	1.44	64.38 / 67.64
4×4	4	128	65.1M	2.90	59.80 / 63.74
4×4	8	224	65.1M	2.90	63.41 / 67.14
3×3	12	224	20.5M	4.84	62.25 / 67.64



目录



喜欢



收藏



评论



分享

表1：调整fc6层的核大小和多孔采样率r后的视野的效果。展示了条件随机场前后的模型参数数量，训练速度（图像数量每秒），和验证集平均IOU。DeepLab-LargeFOV（核大小 3×3 ， $r=12$ ）取得了最好的平衡。

视角和条件随机场：表1中，我们报告了用不同视角大小的DeepLab模型变量的实验，如3.1节所述，在'fc6'层调节内核大小和多孔采样率r。我们开始直接使用VGG-16网络，用原始 7×7 核大小， $r=4$ （最后两个最大池化层没有用步长）。这个模型经过条件随机场后，性能为67.64%，但比较慢（训练时每秒1.44个图像）。将核大小变为 4×4 后模型速度增加到2.9个图像每秒。测试了两个网络变量小（ $r=4$ ）和大（ $r=8$ ）视野大小，后者性能更好。最后，核大小改为 3×3 ，更大的多孔采样率（ $r=12$ ），在'fc6'和'fc7'层中保留4096个滤波器中的一个包含1024的随机子集，使网络更薄。生成的模型，DeepLab-CRF-LargeFOV，匹配VGG-16的性能（ 7×7 核， $r=4$ ）。同时，DeepLab-LargeFOV速度是3.36倍更快，更少的参数数量（20.5M而不是134.3M）。

条件随机场提升了所有模型变体的性能，平均IOU增加了3-5%。

测试集验证：DeepLab-CRF-LargeFOV在PASCAL VOC 2012测试集上进行了评估。平均性能70.3% IOU。

4.1.2 第一版发布后的改进

第一版发布后，我们对模型做了3个主要改进：（1）训练时的不同学习策略，（2）多孔空间金字塔池化，（3）更深的网络和多尺度处理。

学习率：训练DeepLab-LargeFOV时，我们使用了不同的学习率策略。与论文86类似，我们也发现采用“多核细

胞”学习率策略（学习率乘以 $\left(1 - \frac{iter}{max_iter}\right)^{power}$ ）比“步

幅”学习率（用固定步幅大小减小学习率）。



返回顶部

Learning policy	Batch size	Iteration	mean IOU
step	30	6K	62.25
poly	30	6K	63.42
poly	30	10K	64.90
poly	10	10K	64.71
poly	10	20K	65.88

<http://blog.csdn.net/cicibabe>

表2：PASCAL VOC 2012 验证集结果（%）（在条件随机场之前）作为不同的学习超参数变化。训练DeepLab-LargeFOV时，“多核细胞”学习策略比“步幅”更有效率。

如表2所示，使用“多核细胞”（power=0.9），相同的样本数据大小，相同训练迭代次数，性能会比“步幅”策略优1.17%。固定样本数量大小，增加训练迭代次数到10K，可以将性能增强到64.90%（1.48%增加）；然而，由于更多的训练迭代次数增加总的训练时间会增加。将样本数量减少到10，性能保持不变（64.90%比64.71%）。最后，用样本数量=10，迭代20K次，为了获得与“步幅”策略相似的训练时间。令人意外的是，在验证集上性能是65.88%（与“步幅”相比有3.63%的改进），在测试集上是67.7%，是在原来没有条件随机场之前的DeepLab-LargeFOV的“步幅”配置上。后面，用“多核细胞”学习率做剩下的实验。



目录



喜欢



收藏



评论



分享



返回顶部

多孔空间金字塔池化：我们用3.1节描述的多孔空间金字塔池化ASPP流程做了实验。

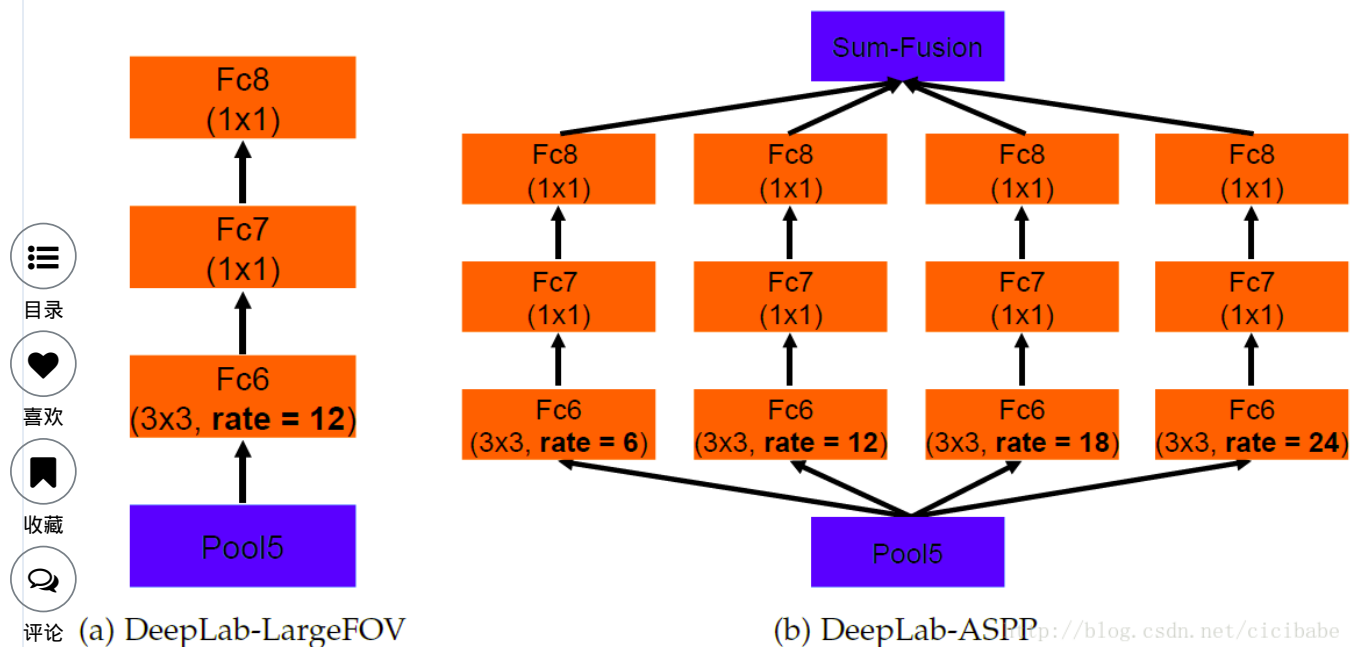


图7：DeepLab-ASPP用不同比例采用多个滤波器在不同尺度上抓取物体和内容。

如图7所示，VGG-16中ASPP使用了几个并行fc6-fc7-fc8分支。他们都使用了3×3的核但在'fc6'上是不同的多孔比例r，为了抓取不同大小的物体。

Method	before CRF	after CRF
LargeFOV	65.76	69.84
ASPP-S	66.98	69.73
ASPP-L	68.96	71.57

<http://blog.csdn.net/cicibabe>



目录



喜欢



收藏



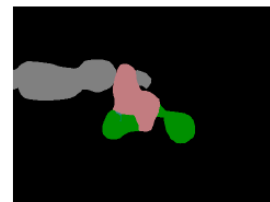
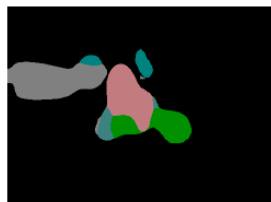
评论



分享

表3：PASCAL VOC 2012验证集上基于VGG-16 DeepLab模型的ASPP效果的性能（平均IOU）。LargeFOV：单分支， $r=12$ 。ASPP-S：4个分支， $r=\{2,4,8,12\}$ 。ASPP-L：4个分支， $r=\{6,12,18,24\}$ 。

表3，用几种不同的设置给出结果：（1）我们的基线LargeFOV模型，单分支 $r=12$ ，（2）ASPP-S，有4个分支，更小的多孔比例（ $r=\{2,4,8,12\}$ ），（3）ASPP-L，有4个分支，更大的比例（ $r=\{6,12,18,24\}$ ）。对每个变量，我们都给出了适用条件随机场前后的结果。如表中所示，ASPP-S在基准LargeFOV使用条件随机场前，有1.22%的改进。然而，结果条件随机场后，LargeFOV和ASPP-S性能相似。另外一方面，ASPP-L在基准LargeFOV使用条件随机场前后，都有一致的改进。我们在测试集上评估了ASPP-L+CRF模型，性能为72.6%。图8展示了不同流程的可视化效果。



(a) Image

(b) LargeFOV

(c) ASPP-S

(d) ASPP-L

图8：ASPP定性分割结果，与基准LargeFOV模型对比。ASPP-L模型使用了多个大视野可以在多个尺度上成功地抓取物体和图像内容。



返回顶部

DeepLab参考资料

DeepLab:

DeepLab主页 (<http://liangchiehchen.com/projects/DeepLab.html>)

DeepLab中文翻译与理解 / 网络架构图 (<http://blog.csdn.net/chenyj92/article/details/53448161>)

DeepLab中文翻译二 (<http://www.dongzhuoyao.com/deeplab-semantic-image-segmentation-with-deep-convolutional-nets-atrous-convolution-and-fully-connected-crfs/>)

深度学习之DeepLab用于语义分割 (<http://blog.csdn.net/yxq5997/article/details/53693869>)

【总结】图像语义分割之FCN和CRF (<https://zhuanlan.zhihu.com/p/22308032>)

图像语义分割之FCN和CRF (<http://blog.csdn.net/u012759136/article/details/52434826>)

DeepLab: Semantic Image Segmentation (<http://blog.csdn.net/tonyyang1995/article/details/51915968>)

diving into deep learning (<http://www.rsipvision.com/ComputerVisionNews-2017May/#16>)



目录



喜欢



收藏



评论



分享

双线性内插值算法:

图像缩放的双线性内插值算法的原理解析 (http://blog.sina.com.cn/s/blog_ab584cac0101h0xy.html)

小波变换:

小波变换教程 (1) : 基本原理 (<http://blog.jobbole.com/101976/>)

小波变换 完美通俗解读 (<http://blog.csdn.net/wanglang3081/article/details/17750565>)

小波变换 完美通俗解读@百度文库 (<https://wenku.baidu.com/view/c7b4ebef7f1922791688e8e3.html>)

小波变换轻松入门 (http://www.360doc.com/content/13/0927/14/10724725_317448978.shtml)

小波变换 (wavelet transform) (<https://www.zhihu.com/topic/19621077/top-answers>)

采样方法 (一) (http://blog.csdn.net/dark_scope/article/details/70992266)

拉普拉斯抠图矩阵:

主页 : A Closed Form Solution to Natural Image Matting (<http://www.wisdom.weizmann.ac.il/~levina/>)

图像处理中的全局优化技术(Global optimization techniques in image processing and computer vision) (二) (<http://blog.csdn.net/mulinb/article/details/9079645>)

A closed form solution to natural image matting。《Computer Vision for Visual Effects》讲解笔记。
(http://blog.csdn.net/d_nel/article/details/44133197)

A Closed Form Solution to Natural Image Matting编程中遇到的问题
(http://blog.sina.com.cn/s/blog_67d185b801017auu.html)



返回顶部

PETA PIXEL GIF (https://giphy.com/gifs/3o7ZeFAS1W6ZmQteJa?utm_source=iframe&utm_medium=embed&utm_campaign=tag_click)

拉普拉斯矩阵/映射/聚类 (<http://blog.csdn.net/yujianmin1990/article/details/48420483>)

Deep Spreadsheets with ExcelNet (<http://www.deepexcel.net/>)

参考文献

目录



喜欢



收藏



评论



分享

[1] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," in Proc. IEEE, 1998.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in NIPS, 2013.

[3] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun, "Overfeat: Integrated recognition, localization and detection using convolutional networks," arXiv:1312.6229, 2013.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in ICLR, 2015.

[5] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," arXiv:1409.4842, 2014.

[6] G. Papandreou, I. Kokkinos, and P.-A. Savalle, "Modeling local and global deformations in deep learning: Epitomic convolution, multiple instance learning, and sliding window detection," in CVPR, 2015.

[7] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in CVPR, 2014.

[8] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov, "Scalable object detection using deep neural networks," in CVPR, 2014.

[9] R. Girshick, "Fast r-cnn," in ICCV, 2015.

[10] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards realtime object detection with region proposal networks," in NIPS, 2015.

[11] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv:1512.03385, 2015.

[12] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, and S. Reed, "SSD: Single shot multibox detector," arXiv:1512.02325, 2015.

[13] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in ECCV, 2014.

[14] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in CVPR, 2015.

- [15] M. Holschneider, R. Kronland-Martinet, J. Morlet, and P. Tchamitchian, "A real-time algorithm for signal analysis with the help of the wavelet transform," in *Wavelets: Time-Frequency Methods and Phase Space*, 1989, pp. 289–297.
- [16] A. Giusti, D. Ciresan, J. Masci, L. Gambardella, and J. Schmidhuber, "Fast image scanning with deep max-pooling convolutional neural networks," in *ICIP*, 2013.
- [17] L.-C. Chen, Y. Yang, J. Wang, W. Xu, and A. L. Yuille, "Attention to scale: Scale-aware semantic image segmentation," in *CVPR*, 2016.
- [18] I. Kokkinos, "Pushing the boundaries of boundary detection using deep learning," in *ICLR*, 2016.
- [19] S. Lazebnik, C. Schmid, and J. Ponce, "Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories," in *CVPR*, 2006.
- [20] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *ECCV*, 2014.
- [21] B. Hariharan, P. Arbel'aez, R. Girshick, and J. Malik, "Hyper columns for object segmentation and fine-grained localization," in *CVPR*, 2015.
- [22] P. Krahenbuhl and V. Koltun, "Efficient inference in fully connected crfs with gaussian edge potentials," in *NIPS*, 2011.
- [23] C. Rother, V. Kolmogorov, and A. Blake, "GrabCut: Interactive foreground extraction using iterated graph cuts," in *SIGGRAPH*, 2004.
- [24] J. Shotton, J. Winn, C. Rother, and A. Criminisi, "Texton boost for image understanding: Multi-class object recognition and segmentation by jointly modeling texture, layout, and context," *IJCV*, 2009.
- [25] A. Lucchi, Y. Li, X. Boix, K. Smith, and P. Fua, "Are spatial and global constraints really necessary for segmentation?" in *ICCV*, 2011.
- [26] X. He, R. S. Zemel, and M. Carreira-Perpinán, "Multiscale conditional random fields for image labeling," in *CVPR*, 2004.
- [27] L. Ladicky, C. Russell, P. Kohli, and P. H. Torr, "Associative hierarchical crfs for object class image segmentation," in *ICCV*, 2009.
- [28] V. Lempitsky, A. Vedaldi, and A. Zisserman, "Pylon model for semantic segmentation," in *NIPS*, 2011.
- [29] A. DeLong, A. Osokin, H. N. Isack, and Y. Boykov, "Fast approximate energy minimization with label costs," *IJCV*, 2012.
- [30] J. M. Gonfau, X. Boix, J. Van de Weijer, A. D. Bagdanov, J. Serrat, and J. Gonzalez, "Harmony potentials for joint classification and segmentation," in *CVPR*, 2010.
- [31] P. Kohli, P. H. Torr et al., "Robust higher order potentials for enforcing label consistency," *IJCV*, vol. 82, no. 3, pp.



目录



喜欢



收藏



评论



分享



返回顶部

302–324, 2009.

[32] L.-C. Chen, G. Papandreou, and A. Yuille, “Learning a dictionary of shape epitomes with applications to image labeling,” in ICCV, 2013.

[33] P. Wang, X. Shen, Z. Lin, S. Cohen, B. Price, and A. Yuille, “Towards unified depth and semantic prediction from a single image,” in CVPR, 2015.

[34] M. Everingham, S. M. A. Eslami, L. V. Gool, C. K. I. Williams, J. Winn, and A. Zisserma, “The pascal visual object classes challenge a retrospective,” IJCV, 2014.



目录

[35] R. Mottaghi, X. Chen, X. Liu, N.-G. Cho, S.-W. Lee, S. Fidler, R. Urtasun, and A. Yuille, “The role of context for object detection and semantic segmentation in the wild,” in CVPR, 2014.



喜欢

[36] X. Chen, R. Mottaghi, X. Liu, S. Fidler, R. Urtasun, and A. Yuille, “Detect what you can: Detecting and representing objects using holistic models and body parts,” in CVPR, 2014.



收藏

[37] M. Cordts, M. Omran, S. Ramos, T. Rehfeld, M. Enzweiler, R. Benenson, U. Franke, S. Roth, and B. Schiele, “The cityscapes dataset for semantic urban scene understanding,” in CVPR, 2016.



评论

[38] L.-C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, and A. L. Yuille, “Semantic image segmentation with deep convolutional nets and fully connected crfs,” in ICLR, 2015.



分享

[39] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, “Learning hierarchical features for scene labeling,” PAMI, 2013.

[40] G. Lin, C. Shen, I. Reid et al., “Efficient piecewise training of deep structured models for semantic segmentation,” arXiv:1504.01013, 2015.

[41] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, “Caffe: Convolutional architecture for fast feature embedding,” arXiv:1408.5093, 2014.

[42] Z. Tu and X. Bai, “Auto-context and its application to highlevel vision tasks and 3d brain image segmentation,” IEEE Trans. Pattern Anal. Mach. Intell., vol. 32, no. 10, pp. 1744–1757, 2010.

[43] J. Shotton, M. Johnson, and R. Cipolla, “Semantic texton forests for image categorization and segmentation,” in CVPR, 2008.

[44] B. Fulkerson, A. Vedaldi, and S. Soatto, “Class segmentation and object localization with superpixel neighborhoods,” in ICCV, 2009.

[45] J. Carreira, R. Caseiro, J. Batista, and C. Sminchisescu, “Semantic segmentation with second-order pooling,” in ECCV, 2012.

[46] J. Carreira and C. Sminchisescu, “CPMC: Automatic object segmentation using constrained parametric min-cuts,” PAMI, vol. 34, no. 7, pp. 1312–1328, 2012.

[47] P. Arbel´aez, J. Pont-Tuset, J. T. Barron, F. Marques, and J. Malik, “Multiscale combinatorial grouping,” in CVPR,



返回顶部

2014.

[48] J. Uijlings, K. van de Sande, T. Gevers, and A. Smeulders, “Selective search for object recognition,” IJCV, 2013.

[49] B. Hariharan, P. Arbel´aez, R. Girshick, and J. Malik, “Simultaneous detection and segmentation,” in ECCV, 2014.

[50] M. Mostajabi, P. Yadollahpour, and G. Shakhnarovich, “Feedforward semantic segmentation with zoom-out features,” in CVPR, 2015.

[51] J. Dai, K. He, and J. Sun, “Convolutional feature masking for joint object and stuff segmentation,” arXiv:1412.1283,

2014.



目录

[52] D. Eigen and R. Fergus, “Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture,” arXiv:1411.4734, 2014.



喜欢

[53] M. Cogswell, X. Lin, S. Purushwalkam, and D. Batra, “Combining the best of graphical models and convnets for semantic segmentation,” arXiv:1412.4313, 2014.



收藏

[54] D. Geiger and F. Girosi, “Parallel and deterministic algorithms from mrfs: Surface reconstruction,” PAMI, vol. 13, no. 5, pp. 401–412, 1991.



评论

[55] D. Geiger and A. Yuille, “A common framework for image segmentation,” IJCV, vol. 6, no. 3, pp. 227–243, 1991.



分享

[56] I. Kokkinos, R. Deriche, O. Faugeras, and P. Maragos, “Computational analysis and learning for a biologically motivated model of boundary detection,” Neurocomputing, vol. 71, no. 10, pp. 1798– 1812, 2008.

[57] S. Bell, P. Upchurch, N. Snavely, and K. Bala, “Material recognition in the wild with the materials in context database,” arXiv:1412.0623, 2014.

[58] G. Papandreou, L.-C. Chen, K. Murphy, and A. L. Yuille, “Weakly- and semi-supervised learning of a dcnn for semantic image segmentation,” in ICCV, 2015.

[59] S. Zheng, S. Jayasumana, B. Romera-Paredes, V. Vineet, Z. Su, D. Du, C. Huang, and P. Torr, “Conditional random fields as recurrent neural networks,” in ICCV, 2015.

[60] J. Dai, K. He, and J. Sun, “Boxsup: Exploiting bounding boxes to supervise convolutional networks for semantic segmentation,” in ICCV, 2015.

[61] H. Noh, S. Hong, and B. Han, “Learning deconvolution network for semantic segmentation,” in ICCV, 2015.

[62] Z. Liu, X. Li, P. Luo, C. C. Loy, and X. Tang, “Semantic image segmentation via deep parsing network,” in ICCV, 2015.

[63] L.-C. Chen, J. T. Barron, G. Papandreou, K. Murphy, and A. L. Yuille, “Semantic image segmentation with task-specific edge detection using cnns and a discriminatively trained domain transform,” in CVPR, 2016.

[64] L.-C. Chen, A. Schwing, A. Yuille, and R. Urtasun, “Learning deep structured models,” in ICML, 2015.

[65] A. G. Schwing and R. Urtasun, “Fully connected deep structured networks,” arXiv:1503.02351, 2015.



返回顶部

[66] S. Chandra and I. Kokkinos, "Fast, exact and multi-scale inference for semantic image segmentation with deep Gaussian CRFs," arXiv:1603.08358, 2016.

[67] E. S. L. Gastal and M. M. Oliveira, "Domain transform for edge aware image and video processing," in SIGGRAPH, 2011.

[68] G. Bertasius, J. Shi, and L. Torresani, "High-for-low and low-for high: Efficient boundary detection from deep object features and its applications to high-level vision," in ICCV, 2015.

[69] P. O. Pinheiro and R. Collobert, "Weakly supervised semantic segmentation with convolutional networks," arXiv:1411.6228, 2014.



目录

[70] D. Pathak, P. Krahenbuhl, and T. Darrell, "Constrained convolutional neural networks for weakly supervised segmentation," 2015.



喜欢

[71] S. Hong, H. Noh, and B. Han, "Decoupled deep neural network for semi-supervised semantic segmentation," in NIPS, 2015.



收藏

[72] A. Vezhnevets, V. Ferrari, and J. M. Buhmann, "Weakly supervised semantic segmentation with a multi-image model," in ICCV, 2011.



评论

[73] X. Liang, Y. Wei, X. Shen, J. Yang, L. Lin, and S. Yan, "Proposal-free network for instance-level object segmentation," arXiv preprint arXiv:1509.02636, 2015.



分享

[74] J. E. Fowler, "The redundant discrete wavelet transform and additive noise," IEEE Signal Processing Letters, vol. 12, no. 9, pp. 629–632, 2005.

[75] P. P. Vaidyanathan, "Multirate digital filters, filter banks, polyphase networks, and applications: a tutorial," Proceedings of the IEEE, vol. 78, no. 1, pp. 56–93, 1990.

[76] F. Yu and V. Koltun, "Multi-scale context aggregation by dilated convolutions," in ICLR, 2016.

[77] J. Dai, Y. Li, K. He, and J. Sun, "R-fcn: Object detection via region-based fully convolutional networks," arXiv:1605.06409, 2016.

[78] J. Dai, K. He, Y. Li, S. Ren, and J. Sun, "Instance-sensitive fully convolutional networks," arXiv:1603.08678, 2016.

[79] K. Chen, J. Wang, L.-C. Chen, H. Gao, W. Xu, and R. Nevatia, "Abc-cnn: An attention based convolutional neural network for visual question answering," arXiv:1511.05960, 2015.

[80] L. Sevilla-Lara, D. Sun, V. Jampani, and M. J. Black, "Optical flow with semantic segmentation and localized layers," arXiv:1603.03911, 2016.

[81] Z. Wu, C. Shen, and A. van den Hengel, "High-performance semantic segmentation using very deep fully convolutional networks," arXiv:1604.04339, 2016.

[82] M. J. Shensa, "The discrete wavelet transform: wedding the a trous and mallat algorithms," Signal Processing, IEEE

Transactions on, vol. 40, no. 10, pp. 2464–2482, 1992.

[83] M. Abadi, A. Agarwal et al., “Tensorflow: Large-scale machine learning on heterogeneous distributed systems,” arXiv:1603.04467, 2016.

[84] A. Adams, J. Baek, and M. A. Davis, “Fast high-dimensional filtering using the permutohedral lattice,” in Eurographics, 2010.

[85] B. Hariharan, P. Arbel'aez, L. Bourdev, S. Maji, and J. Malik, “Semantic contours from inverse detectors,” in ICCV, 2011.



[86] W. Liu, A. Rabinovich, and A. C. Berg, “Parsenet: Looking wider to see better,” arXiv:1506.04579, 2015.

目录

[87] T.-Y. Lin et al., “Microsoft COCO: Common objects in context,” in ECCV, 2014.



[88] A. Arnab, S. Jayasumana, S. Zheng, and P. Torr, “Higher order potentials in end-to-end trainable conditional random fields,” arXiv:1511.08119, 2015.

喜欢



[89] R. Vemulapalli, O. Tuzel, M.-Y. Liu, and R. Chellappa, “Gaussian conditional random field network for semantic segmentation,” in CVPR, 2016.

收藏



[90] Z. Yan, H. Zhang, Y. Jia, T. Breuel, and Y. Yu, “Combining the best of convolutional layers and recurrent layers: A hybrid network for semantic segmentation,” arXiv:1603.04871, 2016.

评论



[92] F. Shen and G. Zeng, “Fast semantic image segmentation with high order context and guided filtering,” arXiv:1605.04068, 2016.

分享

[93] Z. Wu, C. Shen, and A. van den Hengel, “Bridging category-level and instance-level semantic image segmentation,” arXiv:1605.06885, 2016.

[94] K. He, X. Zhang, S. Ren, and J. Sun, “Identity mappings in deep residual networks,” arXiv:1603.05027, 2016.

[95] F. Xia, P. Wang, L.-C. Chen, and A. L. Yuille, “Zoom better to see clearer: Human part segmentation with auto zoom net,” arXiv:1511.06881, 2015.

[96] X. Liang, X. Shen, D. Xiang, J. Feng, L. Lin, and S. Yan, “Semantic object parsing with local-global long short-term memory,” arXiv:1511.04510, 2015.

[97] X. Liang, X. Shen, J. Feng, L. Lin, and S. Yan, “Semantic object parsing with graph lstm,” arXiv:1603.07063, 2016.

[98] J. Wang and A. Yuille, “Semantic part segmentation using compositional model combining shape and appearance,” in CVPR, 2015.

[99] P. Wang, X. Shen, Z. Lin, S. Cohen, B. Price, and A. Yuille, “Joint object and part segmentation using deep learned potentials,” in ICCV, 2015.



返回顶部

[100] V. Badrinarayanan, A. Kendall, and R. Cipolla, "Segnet: A deep convolutional encoder-decoder architecture for image segmentation,"arXiv:1511.00561, 2015.

[101] J. Uhrig, M. Cordts, U. Franke, and T. Brox, "Pixel-level encoding and depth layering for instance-level semantic labeling,"arXiv:1604.05096, 2016.



目录



标签：深度学习 (<http://so.csdn.net/so/search/s.do?q=深度学习&t=blog>) /

喜欢



卷积神经网络 (<http://so.csdn.net/so/search/s.do?q=卷积神经网络&t=blog>) /



收藏



评论



分享

举报



相关文章推荐

FCIS+code | Fully Convolutional Instance-aware Semantic Segmentation (/bea_tree/article/details/73088212)

个人觉着RFCN系列的工作还是非常好的，所提出的position sensitive score maps 非常有启发性，包括本文的一些工作都是有很高的价值的。 主要涉及的论文 [1] I...



bea_tree (http://blog.csdn.net/bea_tree) 2017-06-13 18:01 417



返回顶部

DeepLab: Semantic Image Segmentation (/tonyyang1995/article/details/51915968)

DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully C...

 TonyYang1995 (<http://blog.csdn.net/TonyYang1995>) 2016-07-15 11:10  2706



重磅！CSDN自主举办：人工智能峰会来袭！

2017年下旬 AI先行者们正在关注哪些关键技术？又如何从理论跨越到企业创新实践？CSDN邀请阿里巴巴、微软、商汤科技、微博大咖...



喜欢



(http://www.baidu.com/cb.php?c=lgF_pyfqHmsrHTYrjT0IZ0qnfK9ujYzP1D4P1Ts0Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1Ydm1u-uyR1mhP-nWc4nvDY0AwY5HDdnjm4PjD1PjD0IgF_5y9YIZ0IQzq-4Z8mLPbUB48ugfEpZNGXy-jULNzTvRETvNzpyN1gVw-IA7GUatvrjqdIAdxTvqdThP-UvTkn0KzujYk0AFV5H00TZcqn0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1cVrHnz0AqLUWYs0ZK45HcsP6KWThnqnHnkrfj)


【Deep Learning】DeepLab (/chenyj92/article/details/53448161)

【Deep Learning】DeepLab

 chenyj92 (<http://blog.csdn.net/chenyj92>) 2016-12-04 22:22  1602

论文阅读笔记：图像分割方法deeplab以及Hole算法解析 (/tangwei2014/article/details/50453334)

1. deeplab方法概述 2. deeplab对**FCN**更加优雅的处理方式 3. Hole算法 4. 代码

 tangwei2014 (<http://blog.csdn.net/tangwei2014>) 2016-01-04 00:29  15351

语义分割DeepLab v2--DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolut (/zhangjunhit/article/details/72675556)

DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Con...



zhangjunhit (<http://blog.csdn.net/zhangjunhit>) 2017-05-24 16:35 970

语义分割 DeepLabv3--Rethinking Atrous Convolution for Semantic Image Segmentation (/zhangjunhit/article/details/73614317)



Rethinking Atrous Convolution for Semantic Image Segmentation <https://arxiv.org/abs/1706.05587v1>代码还...



zhangjunhit (<http://blog.csdn.net/zhangjunhit>) 2017-06-23 10:49 1690



完语义分割经典论文:DeepLab (/bea_tree/article/details/55096031)



Here we go 本文所提到的是DeepLabv2 : DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, At...



bea_tree (http://blog.csdn.net/bea_tree) 2017-02-14 11:03 1314



Deeplab v3 | Rethinking Atrous Convolution for Semantic Image Segmentation (/bea_tree/article/details/73527732)



鉴于之前写过deeplab v1与2 , 继续读一读新出的v3 整体来说, 这篇文章还是在探究两个大家一直在做的问题: 全局信息于多尺度信息。另外作者还强调了BN的使用。本文在原有的框架下提出了两种框架: ...



bea_tree (http://blog.csdn.net/bea_tree) 2017-06-21 13:04 2473

腾讯公司有举办真情大反馈抽奖活动吗,经理专线 (/cs365315634/article/details/6809642)

腾讯公司总部专线:【400-067+9599】(客户部王经理接听) 腾讯公司总部二线:【400-067+9599】(7×24小时为您服务)

腾讯公司总部专线:【400-067+9599】(7×24...



cs365315634 (<http://blog.csdn.net/cs365315634>) 2011-09-24 01:36 108



返回顶部

论文阅读理解 - (Deeplab-V3)Rethinking Atrous Convolution for Semantic Image Segmentation (/zziahgf/article/details/75314719)

Deeplab-V3Rethinking Atrous Convolution for Semantic Image Segmentation摘要 DeeplabV1&V2 - 带孔卷积(atrous...



oJiMoDeYe12345 (<http://blog.csdn.net/oJiMoDeYe12345>) 2017-07-18 17:32 1100



SegNet: A Deep Convolutional Encoder-Decoder Architecture for Image Segmentation 视频语义分割demo跑通 (/lemianli/article/details/76687508)

目录



先进入到<https://github.com/TimoSaemann/caffe-segnet-cudnn5>, 下载caffe-segnet-cudnn5-master.zip 再进入到https://...

喜欢



lemianli (<http://blog.csdn.net/lemianli>) 2017-08-04 16:23 233



收藏



深度卷积网络CNN与图像语义分割 (/liuxiabing150/article/details/48441969)

评论

原文地址: <http://blog.csdn.net/xiahouzuoxin/article/details/47789361> 转载请注明出处: <http://xiahouzuoxin.github...>



liuxiabing150 (<http://blog.csdn.net/liuxiabing150>) 2015-09-14 17:19 822

分享

深度卷积网络CNN与图像语义分割 (/thesby/article/details/50947437)

说好的要笔耕不辍, 这开始一边实习一边找工作, 还摊上了自己的一点私事困扰, 这几个月的东西都没来得及总结一下。这就来记录一下关于CNN、Caffe、Image Sematic Segmentation相关...



thesby (<http://blog.csdn.net/thesby>) 2016-03-21 16:13 1203

深度卷积网络CNN与图像语义分割 (/u012746763/article/details/51083840)

深度卷积网络CNN与图像语义分割 2015-08-16 / xiahouzuoxin Tags: CNN 转载请注明出处: <http://xiahouzuoxin.github.io/n...>



u012746763 (<http://blog.csdn.net/u012746763>) 2016-04-07 10:49 458



返回顶部

深度卷积网络CNN与图像语义分割 (/xiahouzuoxin/article/details/47789361)

转载请注明出处: <http://xiahouzuoxin.github.io/notes/> 级别1 : DL快速上手级别2 : 从Caffe着手实践级别3 : 读paper , 网络Train起来级别3 : Demo跑起来...



xiahouzuoxin (<http://blog.csdn.net/xiahouzuoxin>) 2015-08-19 22:33 67697

深度卷积网络CNN与图像语义分割 (/zoeyunjj/article/details/51279686)

目录

出处: <http://xiahouzuoxin.github.io/notes/html/深度卷积网络CNN与图像语义分割.html> 级别1 : DL快速上手级别2 : 从Caffe着手实践级别3 : ...



zoeyunjj (<http://blog.csdn.net/zoeyunjj>) 2016-04-29 10:13 605



收藏

深度卷积网络CNN与图像语义分割 (/sherry_gp/article/details/50924781)



转自 : <http://blog.csdn.net/xiahouzuoxin/article/details/47789361> 转载请注明出处: <http://xiahouzuoxin.github.io/>

评论



sherry_gp (http://blog.csdn.net/sherry_gp) 2016-03-18 16:37 758



分享

VGG-大规模图像识别的深度卷积网络 Very Deep Convolutional Networks for Large-Scale Image Recognition (/cicibabe/article/details/71172576)

VGG-大规模图像识别的深度卷积网络 Very Deep Convolutional Networks for Large-Scale Image Recognition 本文讨论大规模图像识别中, 卷...



cicibabe (<http://blog.csdn.net/cicibabe>) 2017-05-04 17:18 727

图像语义分割技术-深度卷积网络图像识别 (/u011001084/article/details/71158994)

http://www.leiphone.com/news/201705/YbRHBVljhqVBP0X5.html?utm_source=tuicool&utm_medium=referral ...