

夺吻99次：校草大人，太腹黑！ 顶级老公，太嚣张！ 误惹大神：落跑萌妻你等着 独宠娇妃：王爷翻墙窥香 情定现代：皇上本无罪 宠妻上瘾：狼性邪帝，停！ 霸道老公，撩不停！ 独家占有：穆先生，宠不停！ 重生之龙套姐姐的逆袭 爱在灰烬里重燃 厉害了我的金箍棒 时空胖商人 娇妻甜蜜蜜：恶魔总裁霸道宠 穿越之惹火军嫂 日久成婚 你在心上，爱情那么长 闪婚总裁契约妻 修仙老师在都市 穿入西游 道长求罩：约会僵尸新娘 顶级盛宠：男神索爱100天 花都玄医 萌妻翻身：老公送上门 末日工厂 萌妻要不够：狼性总裁，抱紧爱 喜孕少奶奶：总裁大人，又饿了 重生非洲当酋长 邪王追爱：殿下独宠三小姐 腹黑专宠：快穿女配有点萌 大剑神 大圣 那年我们同过窗 重生之黄金宝鉴 心头肉 重生傻妻向前冲 天穹王座 女神的贴身司机 有种别缠我 乡村朋友圈 功夫兵王 重生之寒门长嫂 破烂王 至尊神魔 青春有约 我的美女特工老婆 至尊重生 我的老婆是总裁 千金归来：追妻365天 嗨，亲爱的9点不见不散 灵武帝尊 天域神座 修真归来 渣男必须死 迷失的青春期 食味娇娘 锦堂娇 至尊掠夺系统 我的无限历程 火影之最强全能系统 守神记 仙凡变 三生三世枕上书 我是你的噩梦 抗战时空倒爷 总裁太冷血：蜜宠娇妻不要逃 枕上萌妻：老公，别靠近 重生之万界主宰 重生完美男神 娘娘带球跑了！ 霸道鬼夫好难缠 金丹老祖在现代 网游重生之毒奶神坑 穿越八零之军妻养成计划 盖世小农民 明星班主任 超能重工 我在洪荒有块田 海贼王之终极分身 连上星际互联网 仙神菜园 洪荒之功德天榜 穿梭诸天 鸿蒙第一掌门 我在洪荒当食神 神奇养殖场 重生之乡村养猪 抗日之逆天英雄 都市之精神狂丘 土豪系统在都市 和外星人打游戏 都市最后一个神仙 wifi修仙 超级吞噬系统 领主之兵伐天下 万界圣师 异界霸主在都市 妖怪公寓 帝仙 巫师纪元 参天 随身仙园 空间 愿无深情共余生 腹黑老公别太坏 我的似水年华 造化玉碟 超级狂少 网游末日之从零开始 美女的超级高手 我的傲娇大小姐 超级司机 全能小村医 重走未来路 我的24小时贴身保镖 女神的超级保镖 这样想着多喜欢 撩婿 最强升级系统 万古余帝 懵懂青春 天命神相 绝世战魂 东陵宝藏之谜 段誉，我要跟你抢老婆 绝世无双 我的师父是神兽 地府交流群 从仙界归来的厨神 从仙界归来 重生之富二代 傲气凌神

新随笔  
联系  
订阅  
管理



<	2017年10月							>
日	一	二	三	四	五	六		
24	25	26	27	28	29	30		
1	2	3	4	5	6	7		
8	9	10	11	12	13	14		
15	16	17	18	19	20	21		
22	23	24	25	26	27	28		
29	30	31	1	2	3	4		

### 公告

昵称：二毛子  
园龄：7个月  
粉丝：2  
关注：0  
+加关注

### 搜索

找找看

谷歌搜索

## 笔记 | 图像语义分割 (FCN、CRF、MRF)、论文延伸 (Pixel Objectness、)

图像语义分割的意思就是机器自动分割并识别出图像中的内容，我的理解是抠图...

之前在Faster R-CNN中借用了RPN(region proposal network)选择候选框，但是仅仅是候选框，那么我想提取候选框里面的内容，就是图像语义分割了。

简单的理解就是，图像的“分词技术”。

参考文献：

- 1、知乎，困兽，[关于图像语义分割的总结和感悟](#)
- 2、微信公众号，沈MM的小喇叭，[十分钟看懂图像语义分割技术](#)

## 一、FCN全卷积：Fully Convolutional Networks

一些简单的名词，下采样=卷积+池化（像素缩小），

上采样=反卷积（像素放大）Caffe和Kera里叫Deconvolution，而tensorflow里叫conv\_transpose。CS231n这门课中说，叫conv\_transpose更为合适。

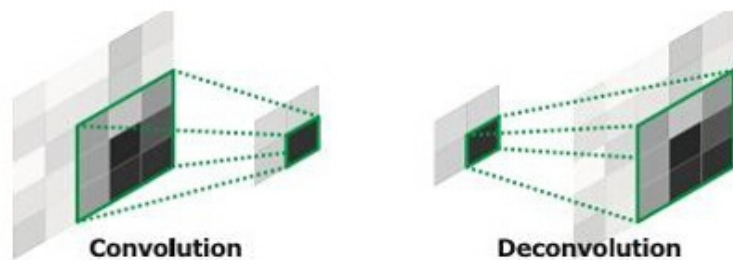
- 卷积与反卷积过程：

## 常用链接

我的随笔  
我的评论  
我的参与  
最新评论  
我的标签

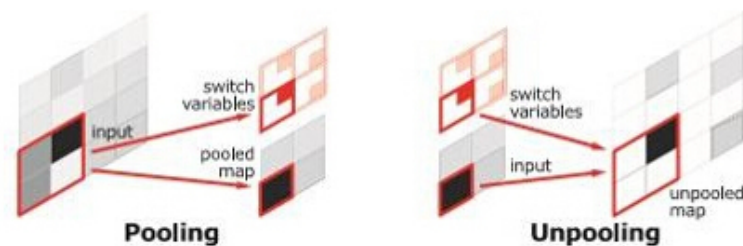
## 随笔档案

2017年2月 (64)  
2017年1月 (43)  
2016年12月 (27)  
2016年11月 (32)  
2016年10月 (37)  
2016年9月 (20)  
2016年8月 (32)  
2016年7月 (41)  
2016年6月 (76)  
2016年5月 (81)  
2016年4月 (85)  
2016年3月 (23)  
2016年2月 (14)  
2016年1月 (5)  
2015年12月 (10)  
2015年11月 (12)  
2015年10月 (5)  
2015年9月 (14)  
2015年8月 (34)  
2015年7月 (93)  
2015年6月 (83)  
2015年5月 (73)  
2015年4月 (92)  
2015年3月 (120)  
2015年2月 (21)  
2015年1月 (86)  
2014年12月 (48)  
2014年11月 (39)  
2014年10月 (151)  
2014年9月 (48)  
2014年8月 (50)  
2014年7月 (142)



[http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)

- 池化与上池化：



[http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)

**池化层作用：**池化操作能够减少计算量，同时也能防止计算结果过拟合

**上池化作用：**上池化的实现主要在于池化时记住输出值的位置，在上池化时再将这个值填回原来的位置，其他位置填0即OK。

## 1、反卷积

2014年6月 (130)  
 2014年5月 (113)  
 2014年4月 (73)  
 2014年3月 (128)  
 2014年2月 (88)  
 2014年1月 (53)  
 2013年12月 (30)  
 2013年11月 (32)  
 2013年10月 (24)  
 2013年9月 (46)  
 2013年8月 (51)  
 2013年7月 (135)  
 2012年4月 (1)  
 2009年7月 (1)  
 2008年11月 (2)  
 2008年10月 (2)  
 2008年7月 (2)  
 2008年6月 (1)  
 2008年4月 (1)

## 最新评论

1. Re:购物网站的推荐算法-  
 个性化推荐算法中如何处理  
 买了还推  
 你好 我想麻烦请教一下 这篇  
 文章的出处是哪啊  
 从哪里转的 啊!!!  
 多谢了啊 '

--KDDA

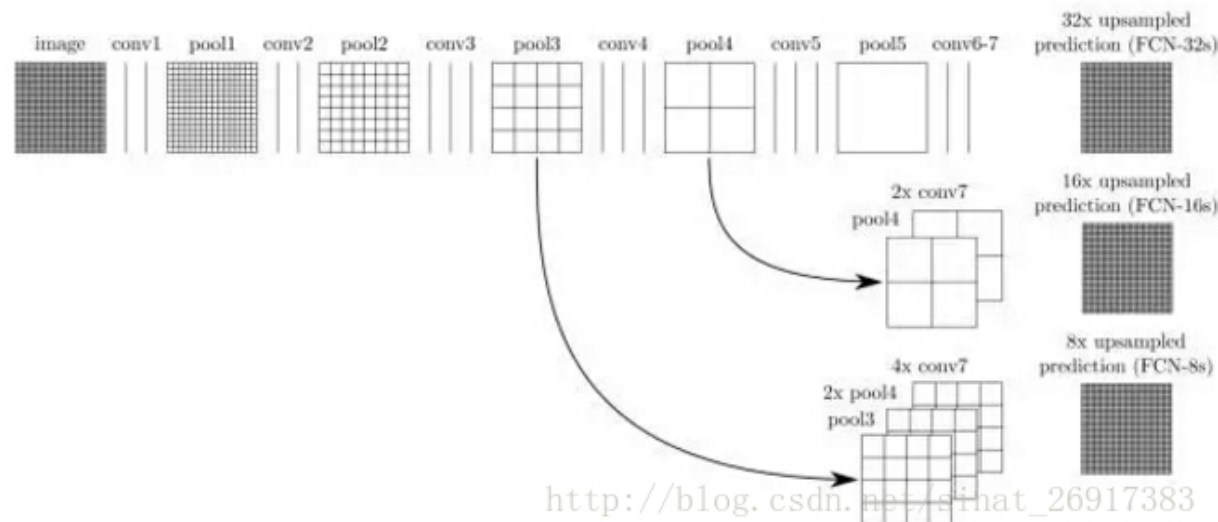
2. Re:区块链技术原理  
 您好, 我们正在寻找区块链  
 人才加入我们团队, 您考虑  
 做技术顾问吗, 如果考虑请  
 联系张先生, 电话:  
 13699209198

--海与博

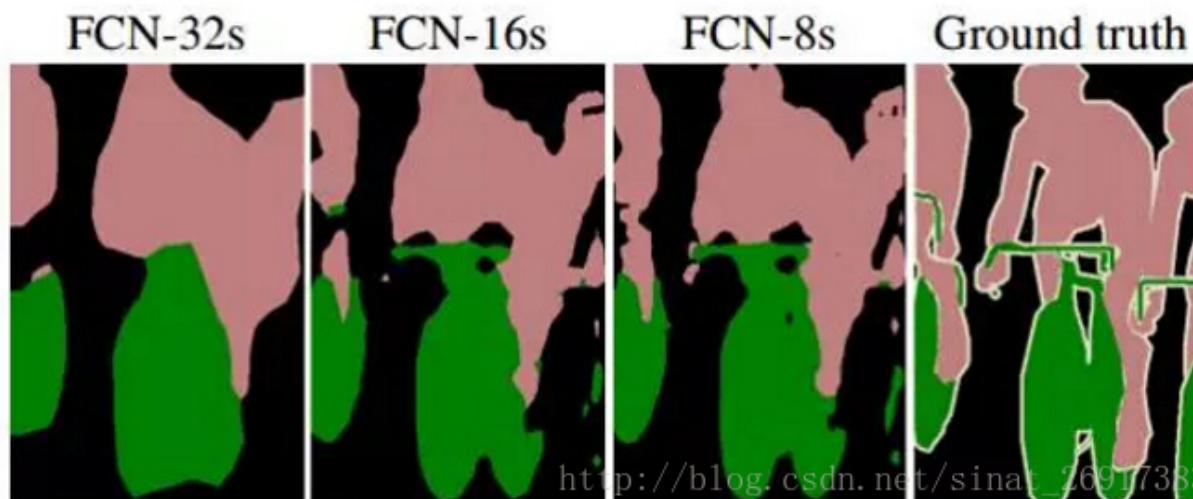
## 阅读排行榜

1. caffe | 深度学习参数调优  
 杂记+caffe训练时的问题  
 +dropout/batch  
 Normalization(5494)

那么反卷积是一种认识“卷积”是什么的好办法。反卷积可以在任何卷积层上进行。那么得到的图像就是一些缩略图形式。



每一种反卷积的样子：



可以发现卷积程度越高，32x图像越模糊，8x跟ground truth还是挺接近的。这里就是感受野（receptive field），8x感受野较小，适合感受细节；32x感受野较大，适合感受宏观。

·  
·

2. NLP | 高级词向量表达  
(一)——GloVe (理论、相关测评结果、R&python实现、相关应用) (767)
3. mongodb去除重复的数据  
(二) (641)
4. NLP+句法结构 (三) | 中文句法结构 (CIPS2016、依存句法、文法) (527)
5. C#利用 string.Join 泛型集合快速转换拼接字符串(495)

## 评论排行榜

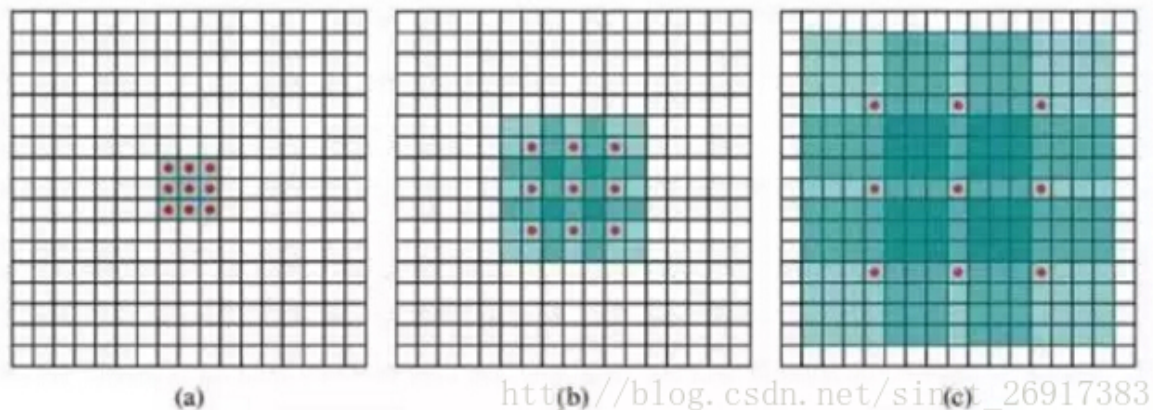
1. 区块链技术原理(1)
2. 购物网站的推荐算法-个性化推荐算法中如何处理买了还推(1)

## 2、“带hole”的卷积——Dilated Convolutions

简单的卷积过程有一些问题：精度问题，对细节不敏感，以及像素与像素之间的关系，忽略空间的一致性等问题。那么新卷积方式hole卷积，用hole卷积核取代池化。

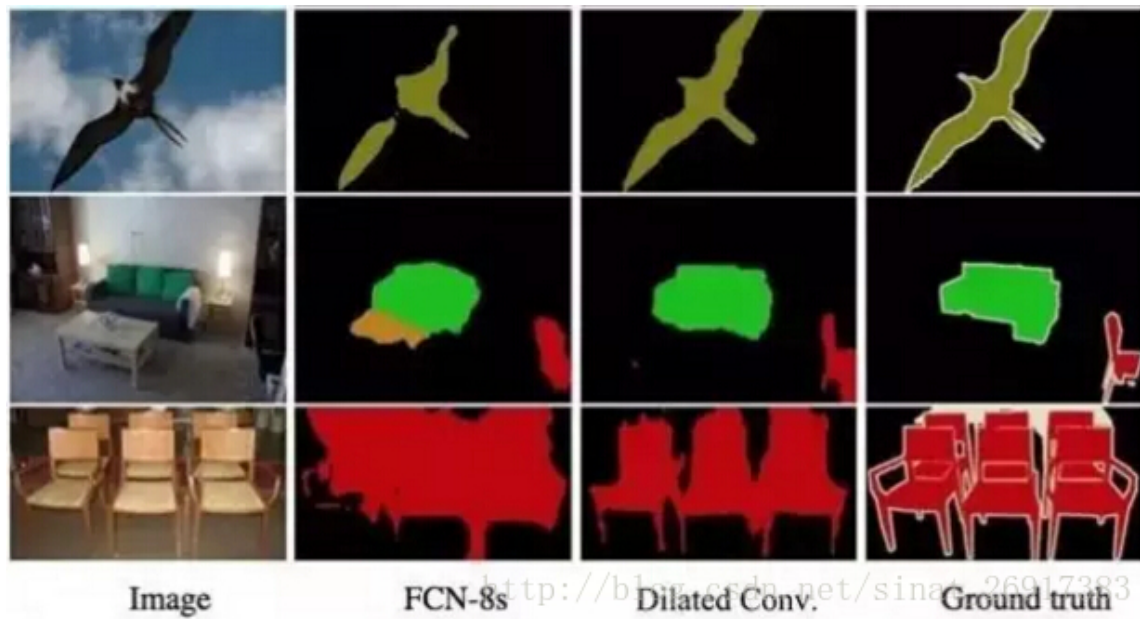
Dilated Convolutions：没有了池化感受野就不会变大，“疏松的卷积核”来处理卷积，可以达到在不增加计算量的情况下增加感受域，弥补不进行池化处理后的精度问题。

操作方式：人为加大了卷积核内部元素之间的距离：



这是水平X轴方向上的扩展，在Y轴上也会有同样的扩展，感受域在没有增加计算（相对于池化操作后）的情况下增大了，并且保留了足够多的细节信息，对图像还原后的精度有明显的提升。

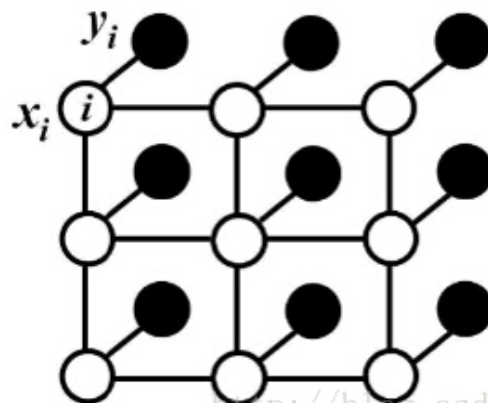




### 3、DenseCRF Conditional Random Field：全连接条件随机场(DenseCRF)

每个像素点作为节点，像素与像素间的关系作为边，即构成了一个条件随机场。

对于每个像素  $i$  具有类别标签  $x_i$  还有对应的观测值  $y_i$ ，这样每个像素点作为节点，像素与像素间的关系作为边，即构成了一个条件随机场。而且我们通过观测变量  $y_i$  来推测像素  $i$  对应的类别标签  $x_i$ 。条件随机场如下：



[http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)

其中： $y_i$ ：卷积后的值； $i$ ：像素； $x_i$ ：类别标签

条件随机场符合吉布斯分布。

在全链接的 CRF 模型中，有一个对应的能量函数：

$$E(\mathbf{x}) = \sum_i \psi_u(x_i) + \sum_{i < j} \psi_p(x_i, x_j)$$

那么  $E(x)$  由两个部分组成，可以简单理解为：

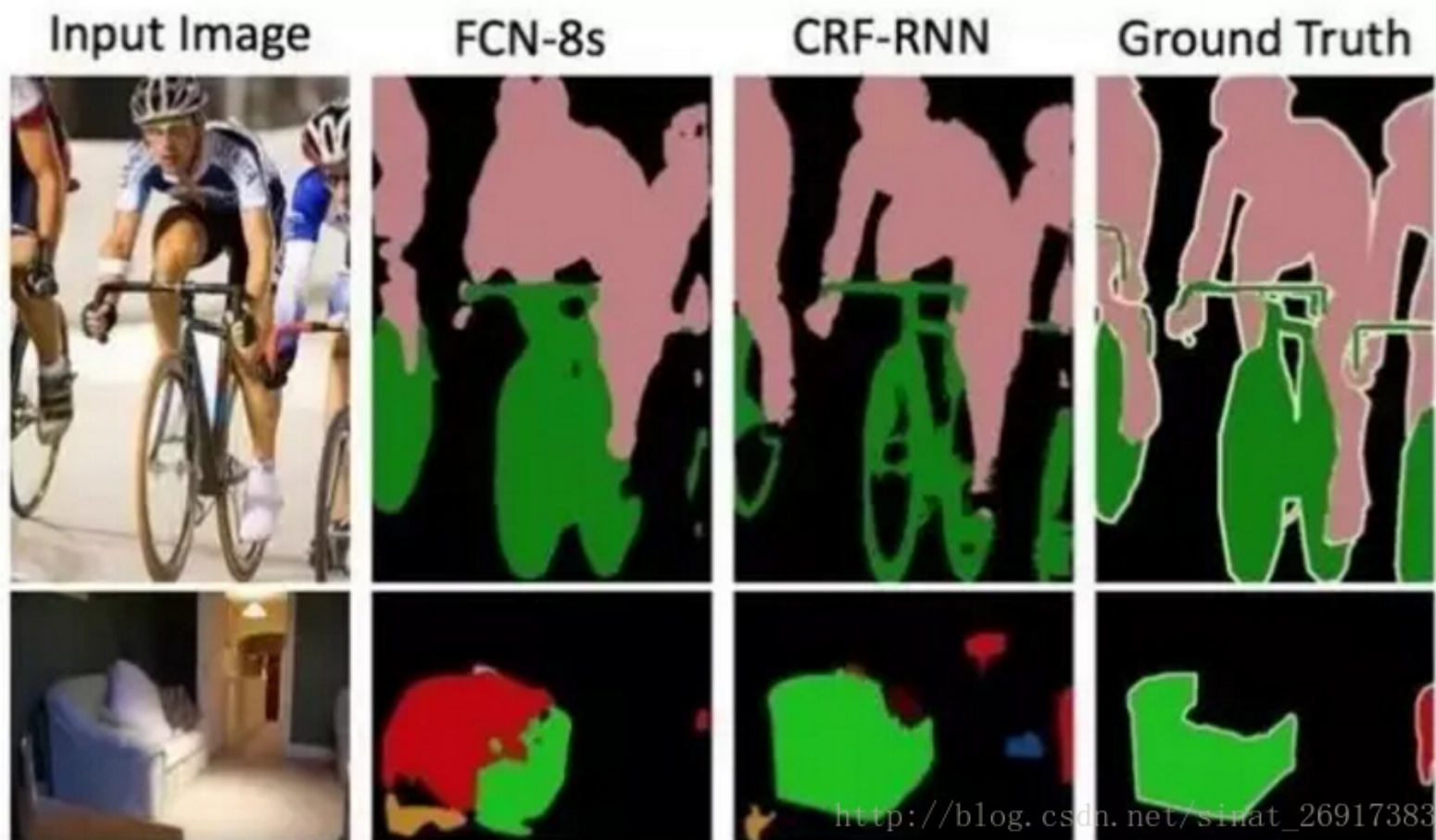
$E(x)$  = 一元函数 + 二元函数

一元函数：来自于前端 FCN 的输出

二元函数：是描述像素点与像素点之间的关系，鼓励相似像素分配相同的标签，而相差较大的像素分配不同标签，而这个“距离”的定义与颜色值和实际相对距离有关。所以这样 CRF 能够使图片尽量在边界处分割。

全连接条件随机场的不同就在于，二元势函数描述的是每一个像素与其他所有像素的关系，所以叫“全连接”。

通过对这个能量函数优化求解，把明显不符合事实识别判断剔除，替换成合理的解释，得到对FCN的图像语义预测结果的优化，生成最终的语义分割结果。



#### 4、马尔科夫随机场(MRF)

在Deep Parsing Network中使用的是MRF，它的公式具体的定义和CRF类似，只不过作者对二元势函数进行了修改：

$$\Psi(y_i^u, y_i^v) = \sum_{k=1}^K \lambda_k u_k(i, u, j, v) \sum_{\forall z \in N_j} d(j, z) p_z^v$$

其中，作者加入 $\lambda_k$ 为label context，因为只是定义了两个像素同时出现的频率，而 $\lambda_k$ 可以对一些情况进行惩罚。比如，人可能在桌子旁边，但是在桌子下面的可能性就更小一些。所以这个量可以学习不同情况出现的概率。

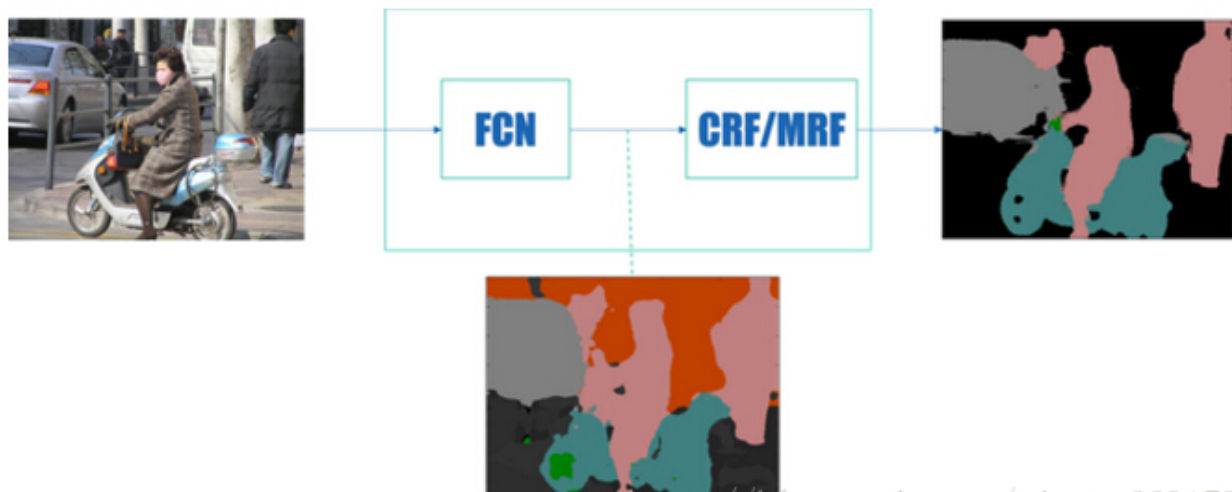
而原来的距离 $d(i,j)$ 只定义了两个像素间的关系，作者在这儿加入了个triple penalty，即还引入了 $j$ 附近的 $z$ ，这样描述三方关系便于得到更充足的局部上下文。

## 5、高斯条件随机场(G-CRF)

这个结构使用CNN分别来学习一元势函数和二元势函数。

## 二、一些成型的分割结构

知乎，困兽，[关于图像语义分割的总结和感悟](#)在文中提到通用框架：



[http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)

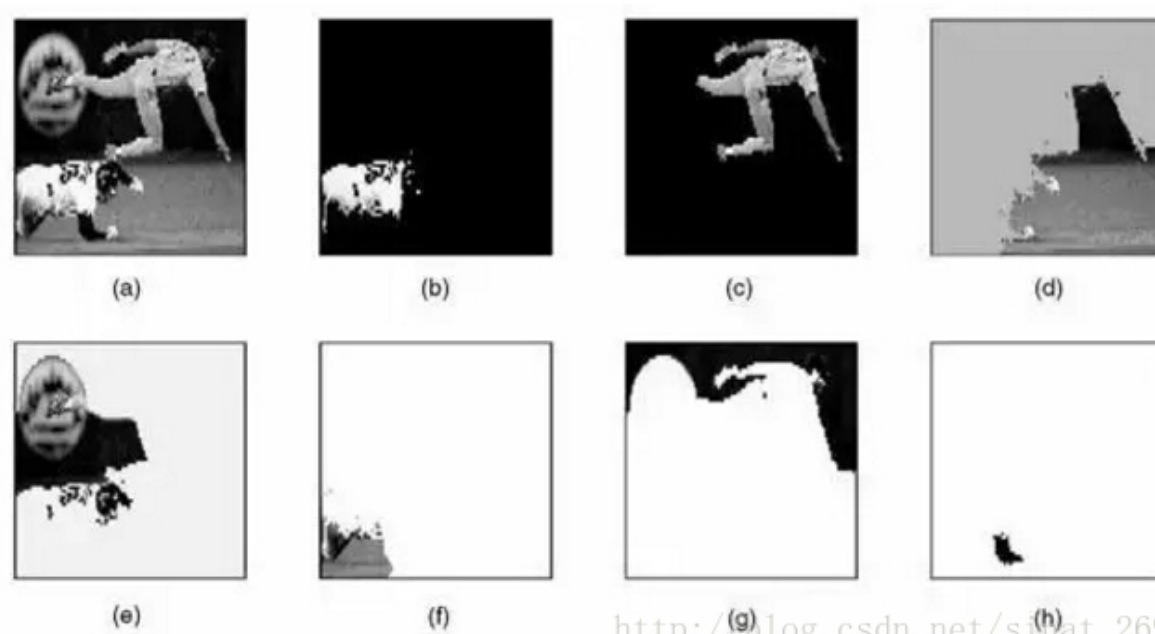
前端使用FCN进行特征粗提取，后端使用CRF/MRF优化前端的输出，最后得到分割图。

### 1、“Normalized cut”的图划分方法

它的思想主要是通过像素和像素之间的关系权重来综合考虑，根据给出的阈值，将图像一分为二。在实际运用中，**每运行一次N-cut，只能切割一次图片**，为了分割出图像上的多个物体，需要多次运行，下图示例了对原图a进行7次N-cut后，每次分割出的结



果。



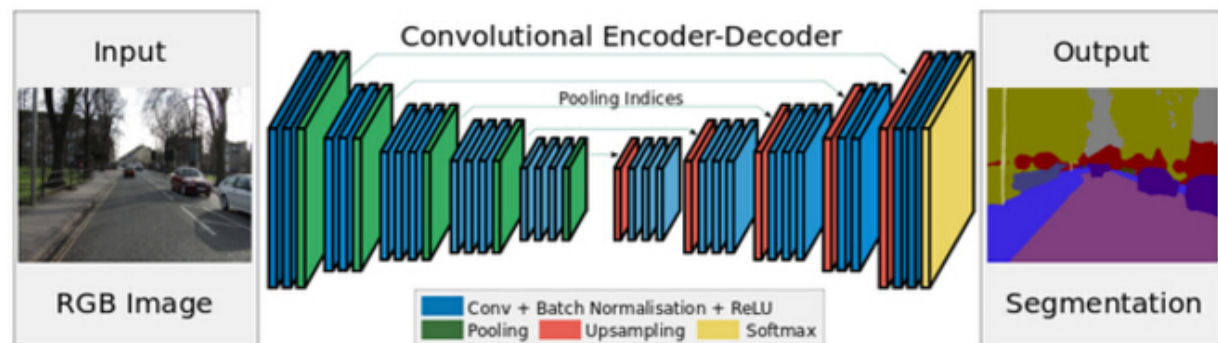
## 2、Grab Cut

增加了人机交互，在分割过程中，需要人工干预参与完成。需要人工选择主体候选框，然后将中部作为主体参考，然后剔除和主体差异较大的部分，留下结果。



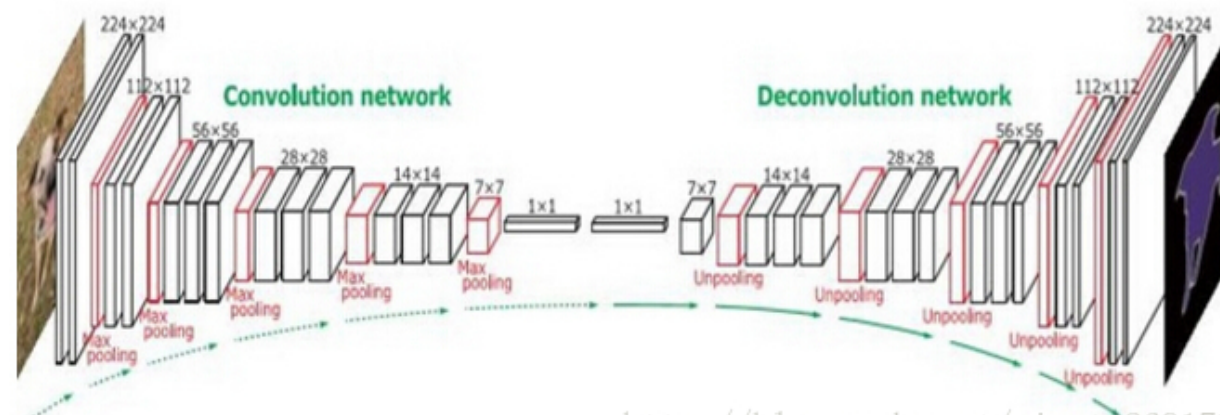
此技术中，抠出来的部分叫“前景”，剔除的部分叫“背景”。缺点也很明显，首先，它同N-cut一样也只能做二类语义分割，说人话就是一次只能分割一类，非黑即白，多个目标图像就要多次运算。其次，它需要人工干预，这个弱点在将来批量化处理和智能时代简直就是死穴。

## 3、segNet



[http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)

#### 4、DeconvNet



[http://blog.csdn.net/sinat\\_26917383](http://blog.csdn.net/sinat_26917383)

这样的对称结构有种自编码器的感觉在里面，先编码再解码。这样的结构主要使用了反卷积和上池化。

#### 5、DeepLab

现在的很多改进是基于这个网络结构的进行的。

为了保证之后输出的尺寸不至于太小，FCN的作者在第一层直接对原图加了100的padding，可想而知，这会引入噪声。

DeepLab非常优雅的做法：将pooling的stride改为1，再加上 1 padding。这样池化后的图片尺寸并未减小，并且依然保留了池化整合特征的特性。

因为池化层变了，后面的卷积的感受野也对应的改变了，这样也不能进行fine-tune了。所以，Deeplab提出了一种新的卷积，带孔的卷积：Atrous Convolution.

## 延伸一：Pixel Objectness —— 更好的自动抠图、图像检索、图像重定向技术

论文《Pixel Objectness》提出了一个用于前景对象分割的端到端学习框架。给定一个单一的新颖图像，我们的方法为所有“像对象”区域 - 即使对于在训练期间从未见过的对象类别，产生像素级掩码。**我们将任务制定为使用深完全卷积网络实现的将前景/背景标签分配给每个像素的结构化预测问题。**

我们的想法的关键是采用训练与图像级对象类别示例，以及采用相对较少的注释的边界级图像合。我们的方法大大改善了ImageNet和MIT数据集上的前景分割的最先进的水平 - 在某些情况下，有19%的绝对改进。此外，在超过100万的图像，我们显示它很好地归纳到用于训练的前景地图中看不见的对象类别。

最后，我们演示了我们的方法如何有利于图像检索和图像重定向，这两种方法在给定的高质量前景图的领域将会有好的效果。

### 论文成果在caffe有成品案例

github：<https://github.com/suyogduttjain/pixelobjectness>

paper：<https://arxiv.org/abs/1701.05349>

官方网址：<http://vision.cs.utexas.edu/projects/pixelobjectness/>

[好文要顶](#)[关注我](#)[收藏该文](#)

二毛子

关注 - 0

粉丝 - 2

[+加关注](#)

0

0

« 上一篇：[Android的ListView异步加载图片时，错位、重复、闪烁问题的分析及解决方法](#)

» 下一篇：[cips2016+学习笔记 | 简述常见的语言表示模型（词嵌入、句表示、篇章表示）](#)

posted on 2017-02-05 17:57 二毛子 阅读(104) 评论(0) 编辑 收藏

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)，[访问网站首页](#)。

【推荐】50万行VC++源码: 大型组态工控、电力仿真CAD与GIS源码库

【推荐】报表开发有捷径：快速设计轻松集成，数据可视化和交互



#### 最新IT新闻:

- 盖茨切换到Android暗示Surface Phone可能无见光之日
  - 丰田马自达成立合资新公司，日系新能源会迎来第二春吗？
  - 紧跟摩拜、ofo，哈罗单车联手支付宝参战“免押金”模式
  - 特斯拉出手：造世界最大锂离子电池 蓄满能不得了
  - 北美首个三星AI实验室落户蒙特利尔大学
- » 更多新闻...



#### 最新知识库文章:

- 实用VPC虚拟私有云设计原则
  - 如何阅读计算机科学类的书
  - Google 及其云智慧
  - 做到这一点，你也可以成为优秀的程序员
  - 写给立志做码农的大学生
- » 更多知识库文章...



Powered by:  
[博客园](#)  
Copyright © 二毛子