图像分割算法综述

快速浏览：全文分成了两个部分，一个是深度学习方面进行图像分割，一个是传统的图像处理方面的分割。

📅时间：2020年12月7日

📕最后更新：2020年12月9日

# 传统的图像处理方式

我发现目前可能都是基于深度学习的分割论文比较多，然后传统图像处理的方式的算法新的会比较少。

**基于阈值法**

阈值法的基本思想是基于图像的灰度特征来计算一个或多个灰度阈值，并将图像中每个像素的灰度值与阈值作比较，最后将像素根据比较结果分到合适的类别中。因此，该方法最为关键的一步就是按照某个准则函数来求解最佳灰度阈值。

**区域分裂合并以及区域生长**

区域生长是从一组代表不同生长区域的种子像素开始，接下来将种子像素邻域里符合条件的像素合并到种子像素所代表的生长区域中，并将新添加的像素作为新的种子像素继续合并过程，知道找不到符合条件的新像素为止。区域生长是从某个或者某些像素点出发，最终得到整个区域，进而实现目标的提取。而分裂合并可以说是区域生长的逆过程，从整幅图像出发，不断的分裂得到各个子区域，然后再把前景区域合并，得到需要分割的前景目标，进而实现目标的提取。

**分水岭：**

其基本思想是把图像看作是测地学上的拓扑地貌，图像中每一点像素的灰度值表示该点的海拔高度，每一个局部极小值及其影响区域称为集水盆，而集水盆的边界则形成分水岭。分水岭的概念和形成可以通过模拟浸入过程来说明。在每一个局部极小值表面，刺穿一个小孔，然后把整个模型慢慢浸入水中，随着浸入的加深，每一个局部极小值的影响域慢慢向外扩展，在两个集水盆汇合处构筑大坝，即形成分水岭。

**基于边缘检测的分割算法：**

各种算子，通过检测包含不同区域的边缘来解决分割问题，通常不同区域的边界上像素的灰度值变化比较剧烈，如果将图片从空间域通过傅里叶变换到频率域，边缘就对应着高频部分，这是一种非常简单的边缘检测算法。



**基于小波分析和小波变换的分割方法：**

二进小波变换具有检测二元函数的局部突变能力，因此可作为图像边缘检测工具。图像的边缘出现在图像局部灰度不连续处，对应于二进小波变换的模极大值点。通过检测小波变换模极大值点可以确定图像的边缘小波变换位于各个尺度上，而每个尺度上的小波变换都能提供一定的边缘信息，因此可进行多尺度边缘检测来得到比较理想的图像边缘。

**基于遗传算法的图像分割：**

模拟由一些基因串控制的生物群体的进化过程，把该过程的原理应用到搜索算法中，以提高寻优的速度和质量。此算法的搜索过程不直接作用在变量上，而是在参数集进行了编码的个体，这使得遗传算法可直接对结构对象（图像）进行操作。整个搜索过程是从一组解迭代到另一组解，采用同时处理群体中多个个体的方法，降低了陷入局部最优解的可能性，并易于并行化。搜索过程采用概率的变迁规则来指导搜索方向，而不采用确定性搜索规则，而且对搜索空间没有任何特殊要求（如连通性、凸性等），只利用适应性信息，不需要导数等其他辅助信息，适应范围广。

# 基于深度学习的算法分类

全部写出来篇幅过长，我写了一部分了解的算法在后文（也只是一部分），将其他算法总结成一个表，进行了小结，见下面的excel文件，从2013~2019年的主要基于深度学习领域的

在类别上主要分为语义分割（Semantic Segmentation）、实例分割（Instance Segmentation）

语义分割：



## CNN卷积神经网络

采用CNN方法进行图像处理的灵感来源于医学领域，我们发现人类处理视觉的过程是封层处理的，即首先瞳孔摄入像素，大脑皮层检测边缘，最后大脑进一步抽象。我们将卷积与神经网络结合，正好也可以分层处理，主要由三个部分组成：卷积层、池化层、全连接层。卷积层负责提取局部特征，池化层降维，全连接层处理最后输出。

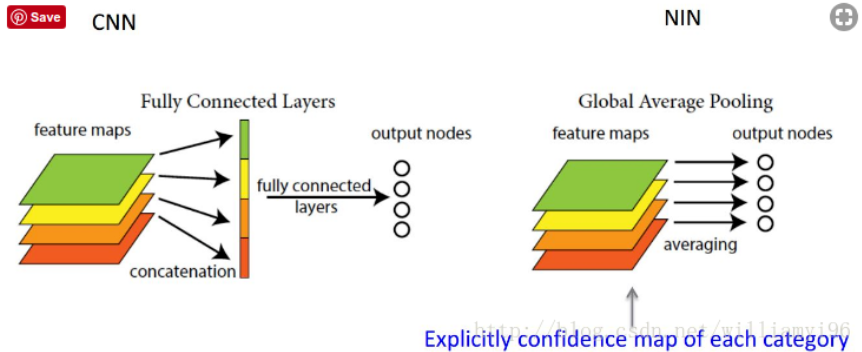
**发展：**

**使用全卷积代替全连接：**

有时候我们发现全连接层也许会破坏图像的空间结构，而且参数量比较大，又或者我们全连接层输入的维度或者图像大小是不确定的，那么原本的全连接层可能无法工作，但是全卷积能解决这个问题，对前层的是全连接的全连接层可以转化程卷积核为1\*1的卷积，而前层是卷积层的全连接层可以转化为卷积核为h\*w的全局卷积，h和w分别为前层卷积结果的高和宽。

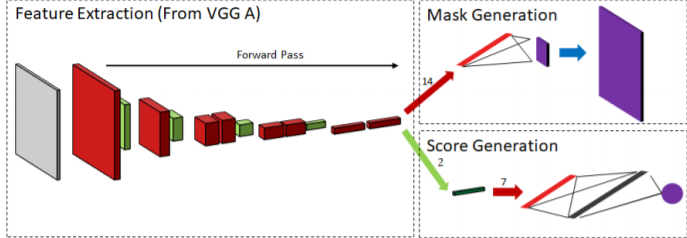
**使用GAP（Global Average Pooling）代替全连接：**

全连接最后要对每个特征图分类，并且之前还要展开向量，GAP将二者合并，最后卷积层输出的图数量对应类别的数量，即在结构上做了正则化防止拟合。



**DeepMask和SharpMask**

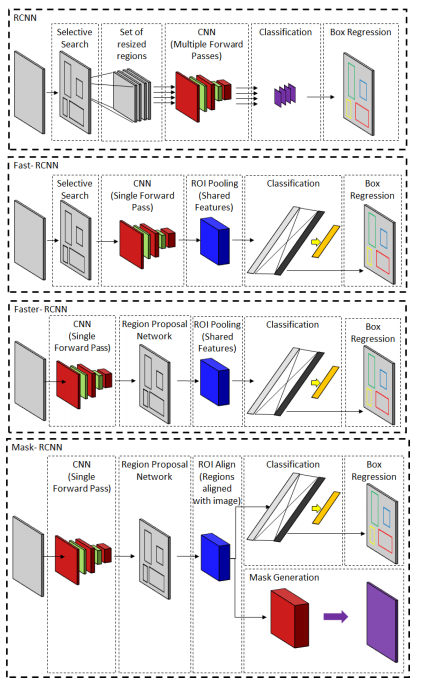
DeepMask也使用了FCN模型，但是最后是按两个分支处理：Mask generation和Score generation分支。Mask部分对图片内部的场景进行识别，由一个1\*1的卷积层接分类层实现，之后是score部分即判断一个图块是否处于正中心附近，然后目标是否完整存于图块当当中。最终输出的目标是一个score，用来显示是否满足之前的要求。



SharpMask是他的升级版，加入了refine的方式，最终生成更高分辨率的mask。

**RCNN (Region-based Convolutional Neural Networks)**

CNN也有许多问题，其中一个就是类似AlexNet的方法不能检测多个物体，RCNN采用了选择搜索算法去寻找可能的分类。选择搜索算法（selective search technique）分析不同的图像特征，例如材质、颜色以及像素在某个实体的集中度。除此之外还有升级版的Fast RCNN、Faster RCNN，有关RCNN的系列算法如下图所示：



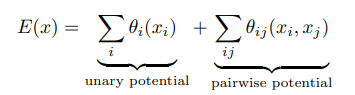
**Deeplab算法**

虽然说像素级别的分割时特别有效的，但是两个因素互相影响着彼此，并且使得性能降低。首先小的卷积核不能去获取上下文的信息，在分类算法中，我们利用池化层去提升感知区域。

但是这会降低分割输出的sharpness，deeplab系列算法用于解决这个问题。Deeplab目前有四个系列的论文，包括v1，v2，v3，v3+。

**使用pixel correlation 提升CNN分割图像的性能**

使用可能性模型例如mrf或者crf主要是基于unary一元和pairwise对偶的激活函数：



CRF经常通过两种方式和深度学习相结合，第一种是使用后训练模型，另一种采用所谓的训练模型在一个端对端网络例如deep parsing network或者spatial propagation networks。

**多层网络**

针对这样的问题，即我们要处理分割的图像大小常常是难以预测的。这时候**PSPNet**和**RefineNet**算法应运而生。**PSPNet**除了传统的空洞卷积FCN对于像素预测，且将像素级特征扩展到专门设计的全局金字塔池化中。局部和全局的线索共同作用以使最终的预测更加可靠。而且设计了相应的优化策略。**RefineNet**它是一种多路径的提炼网络，利用多级抽象特征进行高分辨率的语义分割，通过递归方式提炼低分辨率的特征，生成高分辨率的特征。

**使用自动编码器Convolutional autoencoders：**

思路来自自动编码器。自动编码器基本上由编码器，用于将输入表示形式从原始输入编码为可能的输入低维中间表示和尝试解码的解码器，从中间表示重构原始输入。损失是根据原始输入图像与重建的输出图像。解码器部分的生成性质具有通常会被修改并用于图像分割。不像传统的自动编码器，在分段过程中，损耗是根据重建的像素级类别分布，与所需的像素级类别分布。这种细分方法是与分类方法相比，更多的是生成过程RCNN或DeepLab算法。这样的方法的问题是防止在编码过程中图像过度抽象。这种方法的好处是能够生成更清晰的边界，并且并发症少得多。

与常规卷积相比，使用基于自动编码器的方法的主要优点是特征提取器是选择输入大小的自由。巧妙地使用下采样和上采样操作可以输出像素级与输入图像具有相同分辨率的概率。这个好处已经使具有多尺度特征转发的编码器-解码器体系结构具有对于没有预先确定输入大小的网络变得无处不在，同时也需要与输​​入大小相同的输出。

**采用残差链接的设计**

U-Net：Unet借鉴了FCN网络，其网络结构包括两个对称部分：前面一部分网络与普通卷积网络相同，使用了3x3的卷积和池化下采样，能够抓住图像中的上下文信息（也即像素间的关系）；后面部分网络则是与前面基本对称，使用的是3x3卷积和上采样，以达到输出图像分割的目的。此外，网络中还用到了特征融合，将前面部分下采样网络的特征与后面上采样部分的特征进行了融合以获得更准确的上下文信息，达到更好的分割效果。

