



FCN(3)——DenseCRF



冯超

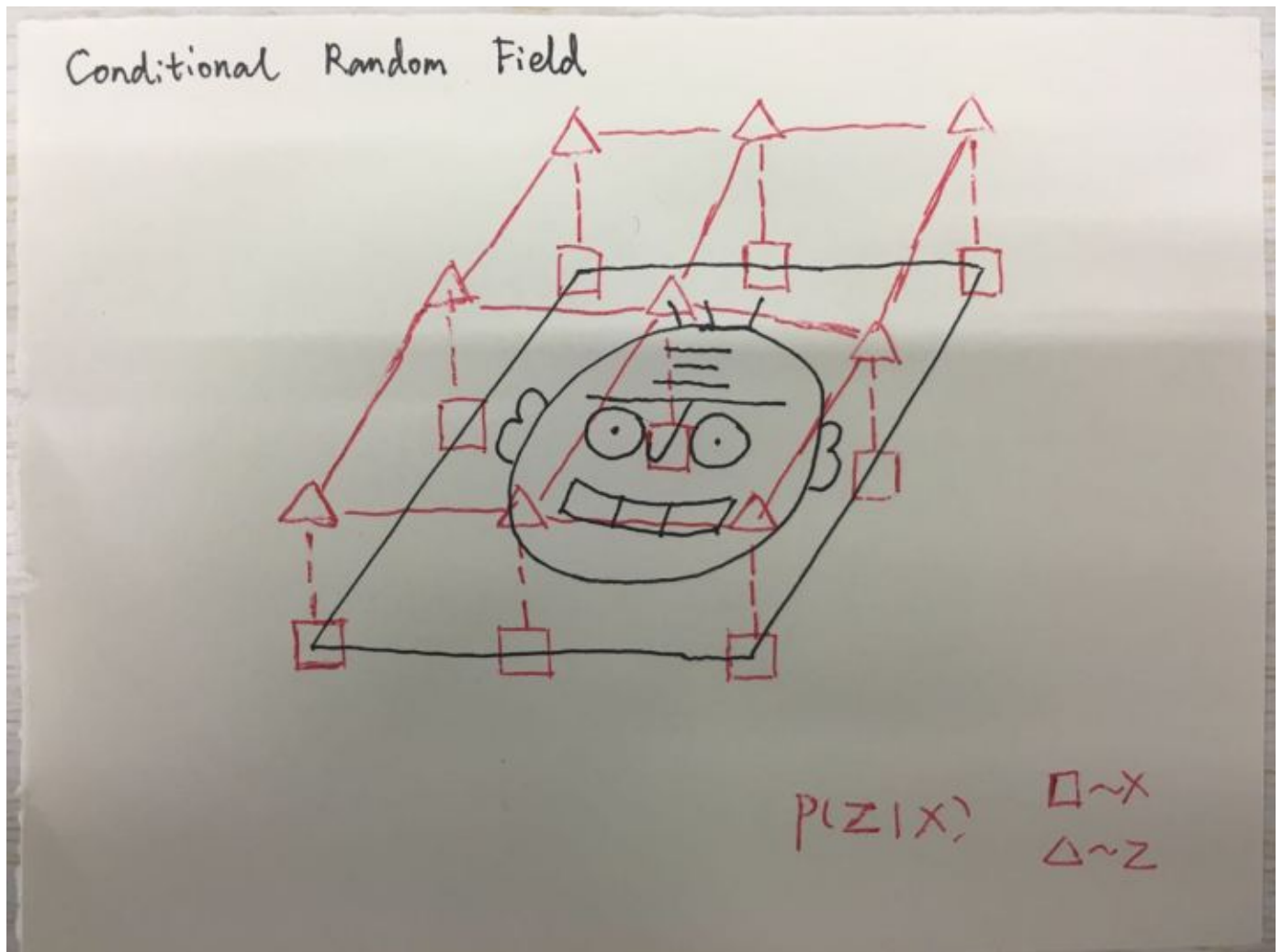
《强化学习精要》《深度学习轻松学》作者,咨询请值乎

48 人赞了该文章

本文收录在[无痛的机器学习第一季](#)。

上一回我们简单介绍了无向图模型和CRF的基本概念，下面我们来看看CRF在图像分割问题上的具体应用。我们简单回忆一下CRF中的两个关键变量，这时我们需要换一下变量的名称——我们用 I 表示图像的像素信息，也就是观察变量，我们用 X 表示图像中每一个像素的类别，也就是我们想要的信息。不过在更多的文档中，大家喜欢用 X 表示图像上的像素，用 Z 表示我们想求的label。

好了，为了更好地展示我们提到的模型，我们给出一个简单的图形：

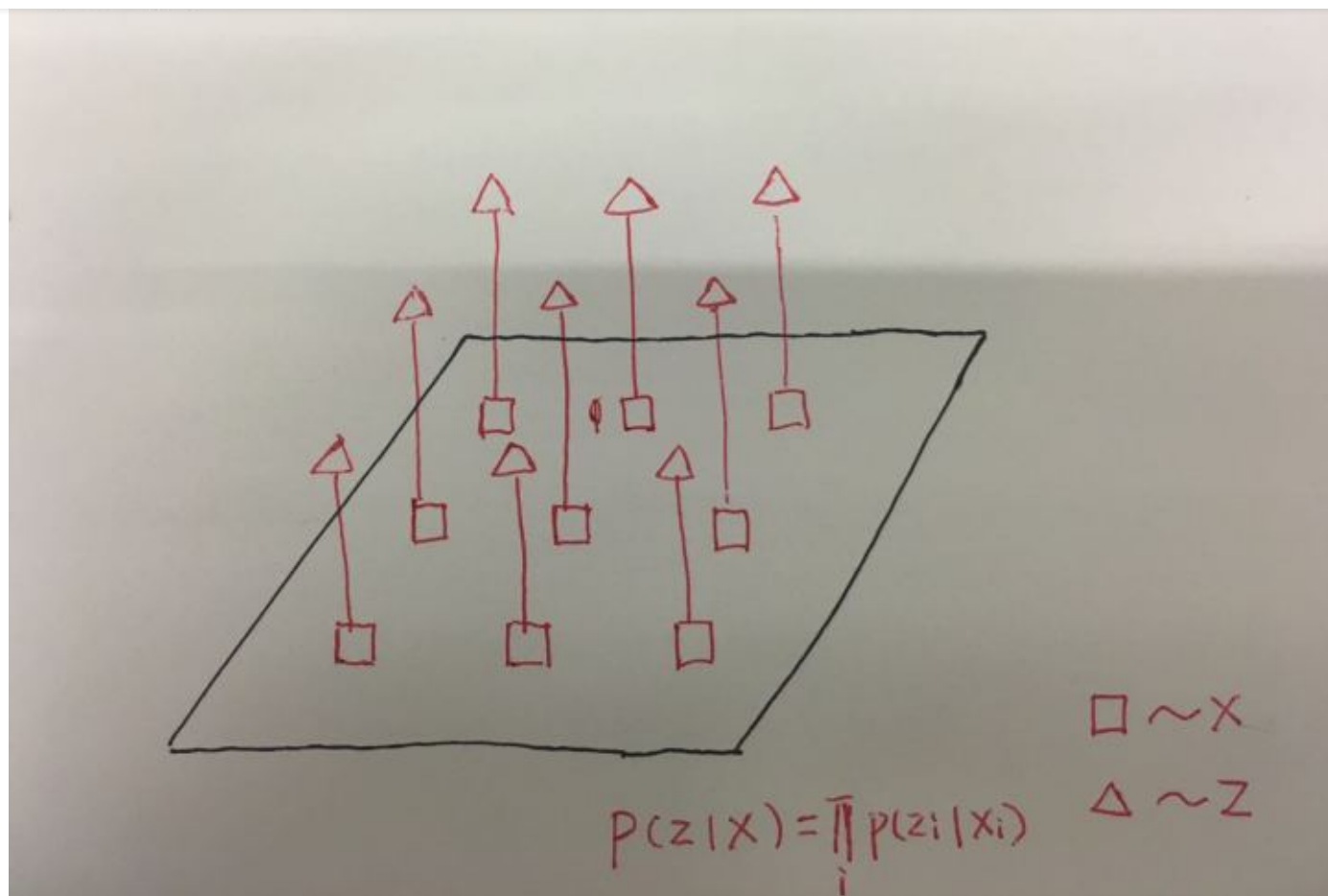


下面的就是我们建模的过程了，最简单的建模方式就是每一个像素的类别只和自己所在的像素有关。我们可以用一个简单的图来表示：

▲ 赞同 48



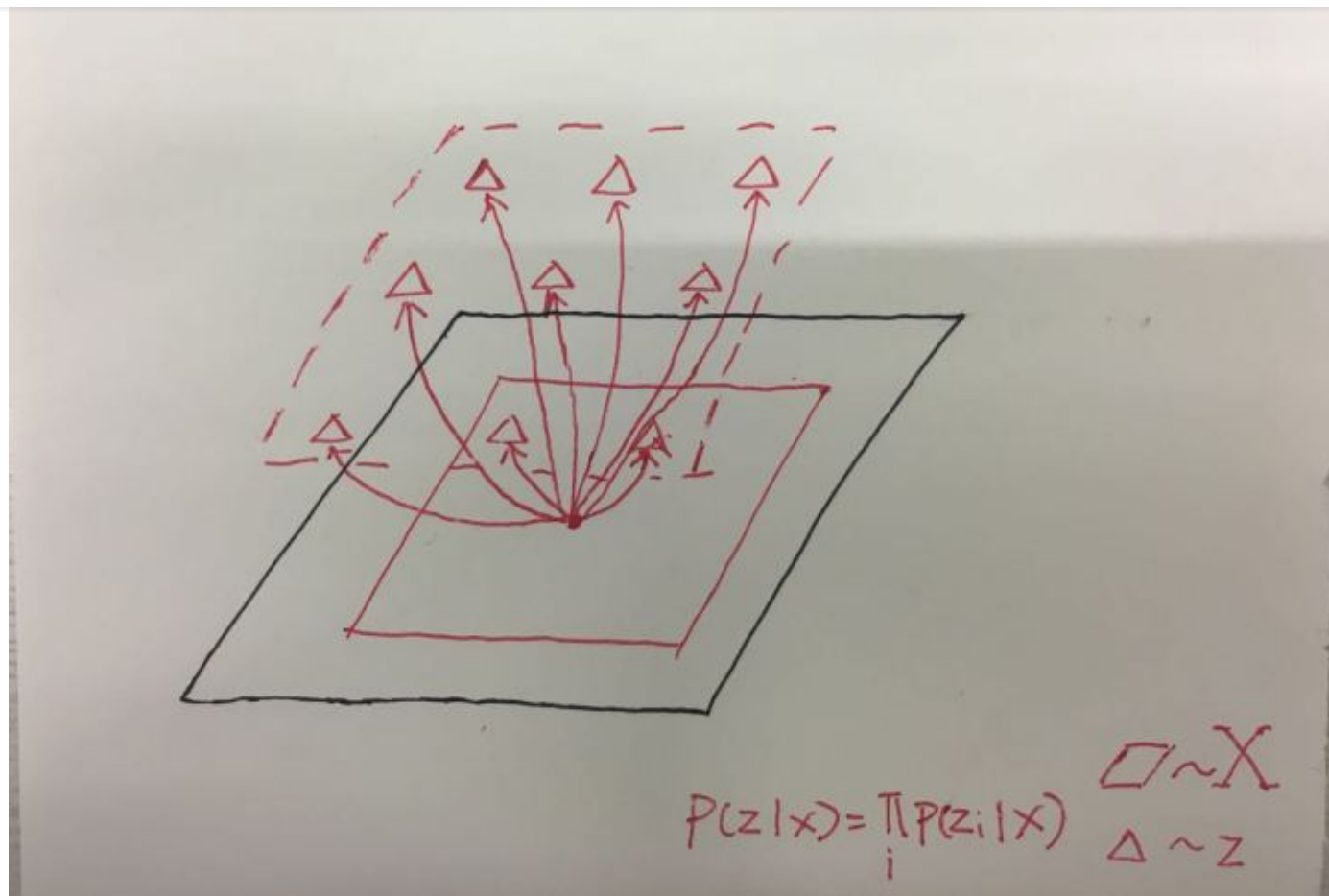
● 9 条评论



这个图看上去比较简单吧，而且计算也比较简单。我们可以建一个模型，输入是一个像素的特征，输出是一个像素所代表的类别。那么这个模型和一般的分类问题没有太多区别。

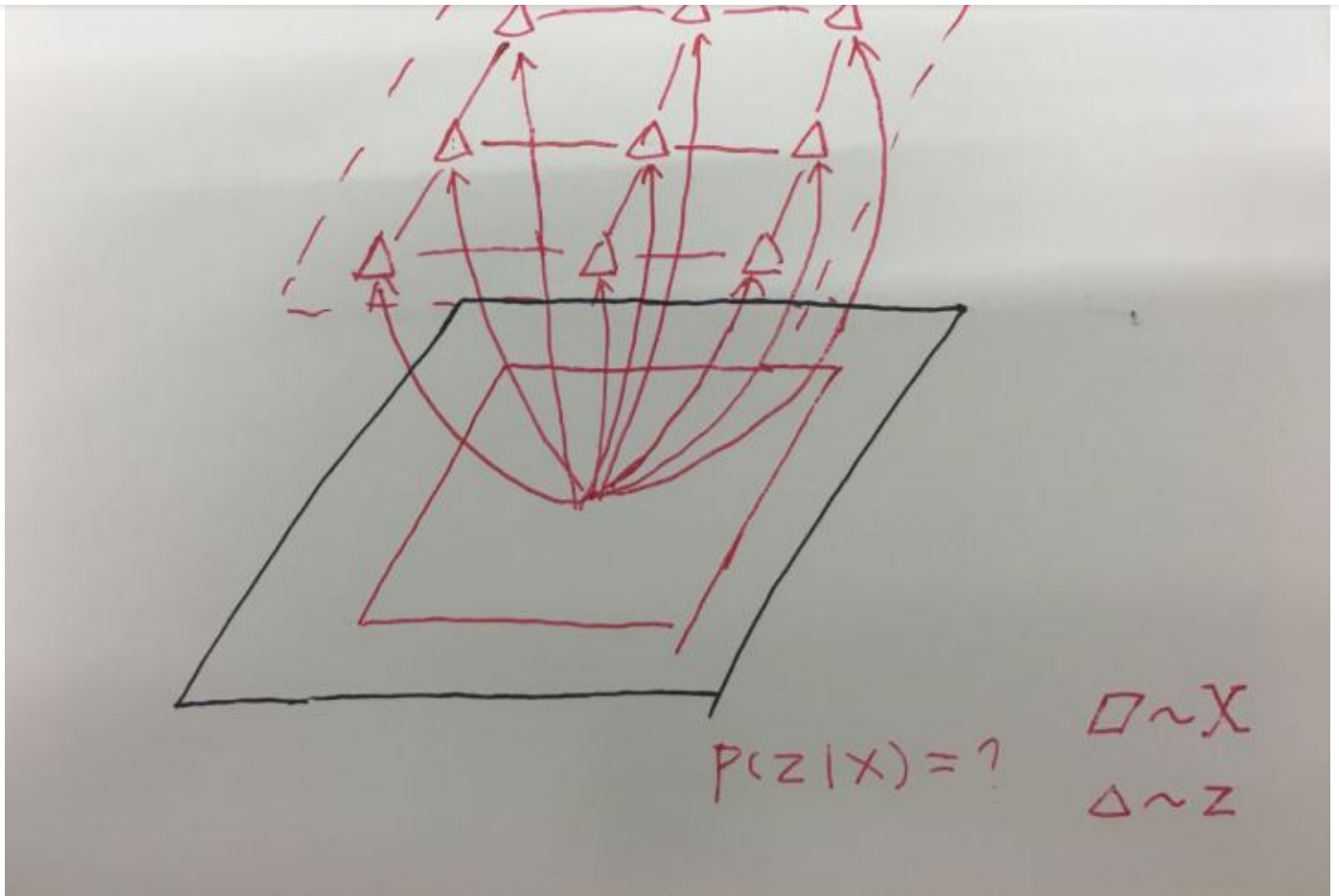
不过这样的效果显然不好，谁能只通过一个像素就知道这个地方是什么类别呢？这个模型虽然简单好解，但是我们一眼就可以看出，它并不是一个高级的模型，所以这个模型pass了。不要问我什么理由.....

下面我们再来一个模型，每一个像素的类别和所有像素的图像信息有关，这个模型实际上就和FCN的输出相同，我们可以用下面的图表示这个模型：



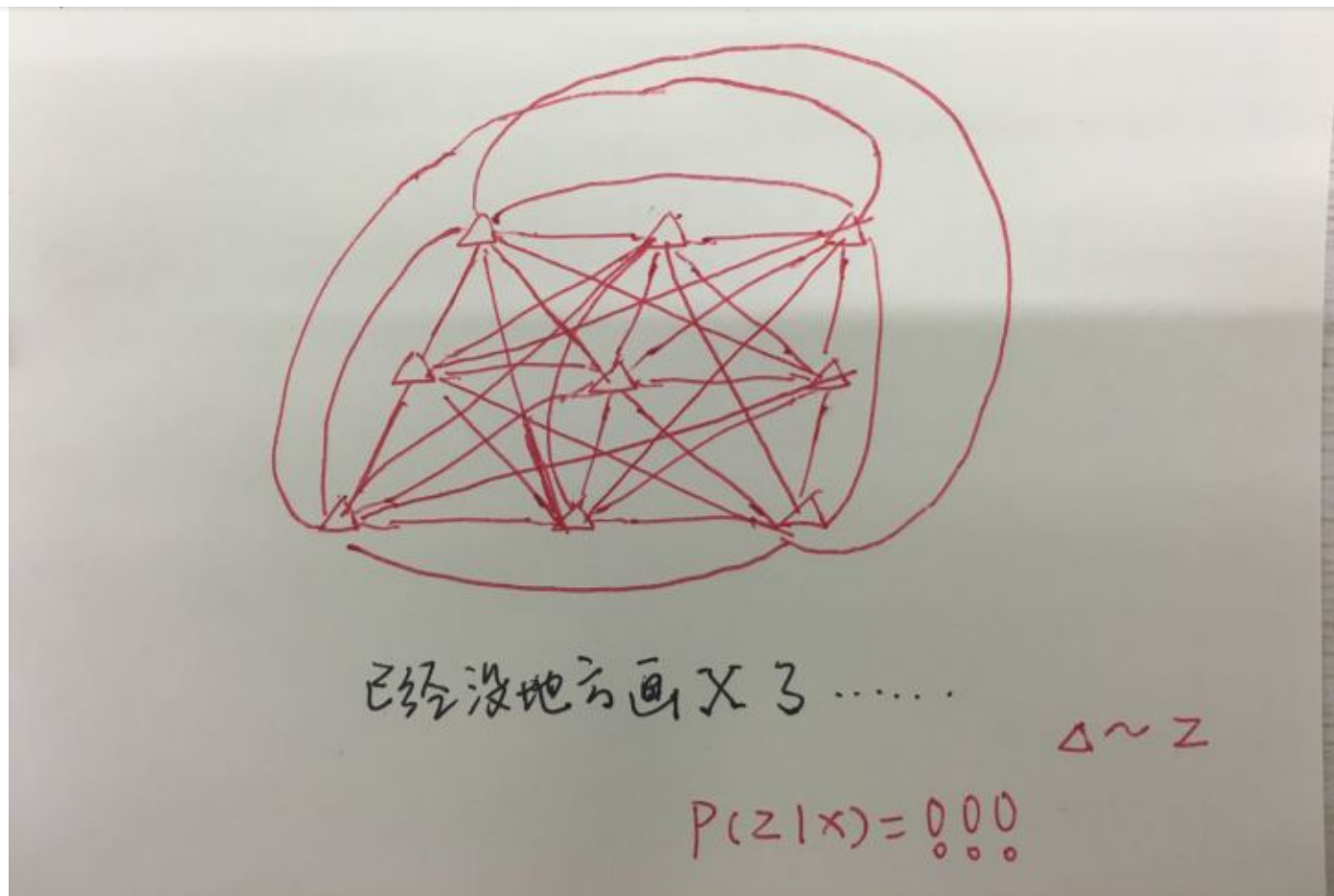
CNN模型虽然很强大，但是CNN模型缺少一个关键点，就是每一个像素点类别之间实际上存在着一定关系，也就是我们常说的图像的平滑性——每一个图像像素点的类别都有可能和临近点的类别相近，这个特性是CNN模型所不具有的。所以我们对这个模型也保留意见。

于是我们新鲜出炉的新模型又来了——这一次我们让每一个像素点的类别和它的4邻域的分类相关，于是它的模型变成了这个样子：



这个模型瞬间高大上了许多，我们前面提到的平滑性已经得到了一定的满足。但是有时候仅仅是这种程度的关联是不够的，如果临近的几个像素恰好产生了一点波动，不是我们想要的那个样子，那么这个像素点的类别就有可能出问题。我们还需要更深层次的关联。于是乎，传说中的denseCRF就诞生了。

denseCRF的模型是什么样子呢？如下图所示：



好吧，基本上能连在一起的都连在一起的，我们把模型的复杂程度提到了最高的程度，完全成了一团乱麻了.....

Dense CRF

我们看过了它的模型形式，下面直接来看看模型的能量函数表达形式：

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

可以看出每一个像素都有一个unary的函数，也就是说在这个特征函数里w，剩下的参数都出现在图像的像我们只考虑当前像素的类别，而暂时不考虑别的像素的类别信息。

而后面的pairwise函数里，每一个像素的类别都和其他像素计算一个energy feature。于是就有了上面的公式形式。注意这里的能量函数是所有像素点的联合能量和不是某一个像素点的能量，所以后面的pairwise项，我们有 $n(n-1)/2$ 组特征，总的来说，特征数量是像素数量的平方级别，如果我们有一张100万像素的图片，那么我们会建立4950亿组pairwise特征。正是因为这种复杂的形式，所以这个模型被称作Dense CRF。满满的全是dense啊！

关于unary函数的内容，我们可以尽情发挥，在denseCRF中我们并没有对这一部分做详细的限制。因此关于这部分我们就暂时略去不谈了，在后面的文章中我们会重新回来，挖掘unary函数的潜力。下面我们专注于解决pairwise这个大坑。



$$\psi_p(x_i, x_j) = \mu(x_i, x_j) \sum_{m=1}^M w^{(m)} k^{(m)}(f_i, f_j)$$

可以看出，pairwise函数中还是比较复杂的，我们从左往右以此介绍。

首先是 $\mu(x_i, x_j)$ ，这一项被称作label compatibility项，简单来说这里约束了“力”传导的条件，只有相同label条件下，能量才可以相互传导。具体来说，“一个像素可能是飞机”的能量可以和“另一个像素可能是飞机”的能量相互传导，从而增加或者减少后者“可能是飞机”的能量，从而影响“可能是飞机”的概率，而“一个像素可能是飞机”的能量是不能影响“另一个像素是人”的概率的。

文章中也提到了，这个简单地一刀切似乎有点不人性。拿Pascal-VOC中的20类+背景类来说，有些类别之间的相似性是很强的，而另外一些类别则完全不相似，所以作者后面提到了一些学习相关的事情，这里我们就不再深入下去了。

加和项里面就是经典的权重*feature的套路了，其中

$$k^{(m)}(f_i, f_j) = w^{(1)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_\alpha^2} - \frac{|I_i - I_j|^2}{2\theta_\beta^2}\right) + w^{(2)} \exp\left(-\frac{|p_i - p_j|^2}{2\theta_\gamma^2}\right)$$

这一项以特征的形式表示了不同像素之前的“亲密度”。前面我们提到了特征不同于tabular形式的factor，我们看不到表格，只能看到公式。上面的公式中，第一项被称作appearance kernel，第二项被称作smooth kernel。这里面有很多变量，我们一一看。

appearance kernel里面的p表示像素的位置——position，我们的图像是2维的，那么position就有2维。I表示图像的像素值——Intensity，我们的图像是彩色的，那么Intensity就有3维。分式下面的两个参数无论从位置还是长相都像高斯分布中的方差，这里的含义也差不多的。

于是乎我们可以看出，如果两个像素距离近且颜色近，那么这个feature特征就会很强，反之则不强。同时分母也控制了强弱性，分母越大越难强起来。其实这一项和图像处理中的bilateral filter很像。我们相当于在一个5维的空间内寻找相近的像素对并给予它们的特征加强。

看完了前面一项，在看后面那一项就不会觉得太难。大家照着上面那一段话自行分析吧。

好了，到这里我们就交代完DenseCRF的表达形式了，下面就要进入Mean field variational inference的环节了。

广告时间

更多精彩尽在 [《深度学习轻松学：核心算法与视觉实践》](#)！

私货时间2

机器学习

深度学习 (Deep Learning)

概率图模型

文章被以下专栏收录



无痛的机器学习

专栏主营业务：让更多人能看的懂的机器学习科普+进阶文章。欢迎各位大神投稿或协...

进入专栏

推荐阅读

FCN(5)——DenseCRF推导

本文收录在无痛的机器学习第一季。经过前两篇文章，我们了解了CRF的基本概念，了解了许多许多的CRF模型，也了解了Mean field variational inference的基本概念，那么这一回我们开始真刀真...

冯超

FCN(2)——CR入门

本文收录在无痛的季。前面我们简单这个将High-Level Level任务的模型。Low并不是我们通和Low，两种任务

冯超

9 条评论

⇌ 切换为时间排序

写下你的评论...



fayefan

1 年前

请问博文里面的“文章中也提到了”，这个文章指的是哪篇文章？谢谢！

 赞



张小迪 回复 fayefan

1 年前



知乎用户

1 年前

分母越小，越难强起来吧。



1



cjh2012

1 年前

论文看了，有个都没说到的地方。也可能是我理解错了。

图的势函数是从最大子团计算的。dense crf的最大团，好像只有一个，所有节点连接在一起的这个子团。

然后吉布斯分布里，指数里面的累加可以去掉。

不知道是否理解正确



2



冯超 (作者) 回复 cjh2012

1 年前

没有太看明白您的意思，不过我感觉是对的



赞



查看对话



CalmDog 回复 cjh2012

1 个月前

是的，因为是dense的所以类别组成的无向图所有点两两连接，所有点构成了一个最大团



赞



查看对话



ann HHHH

1 个月前

请问 说cnn不会考虑临近点 缺少平滑性的理由是什么呢？本人小白 刚刚入门的新手 恳请解答！



赞



吴凡

1 个月前

您好，CNN提取的是一种局部特征，对于相邻的像素，局部特征应该是相似的，从而使得相邻类别相近，不知道这种理解有什么问题？



赞



高压线 回复 ann HHHH

18 天前

我感觉作者叙述不清楚（也有可能是我理解错了），作者这儿的平滑性指大的是label之间的，也就是相邻像素之间的label之间应该有关系，但是planB中的label之间并没有线的连接关系，作者把这个叫做不平滑，但是我感觉CNN是存在平滑性的，CNN的score map存在明显的平滑关系。



赞



查看对话