

卷积神经网络CNN完全指南终极版（二）



沉迷学习的糕糕
要成为最可爱的产品经理（握拳！

关注她

253 人赞同了该文章

阅读前请保证已阅读上一篇文章

卷积神经网络CNN完全指南终极版（一）

【5】全连接层

上一篇文章我们说到卷积神经网络的前三层，现在开始讲解它的最后一层全连接层。

全连接层的形式和前馈神经网络（feedforward neural network）的形式一样，或者称为多层感知机（multilayer perceptron，MLP），纠结了很久要不要再具体介绍前馈神经网络，最终决定还是先假设读者朋友都已经有了这方面相关的知识，如果有需要的话我再单独写一篇文章介绍FNN。

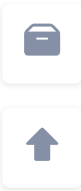
还是回到上一篇中识别图片X的栗子。



原图片尺寸为9X9，在一系列的卷积、relu、池化操作后，得到尺寸被压缩为2X2的三张特征图。



从9X9，到2X2，看起来我们已经做了很多事情了哎！似乎胜利就在前方啦(ง •̀_•́)ง



类并没有什么毛线关系啊。

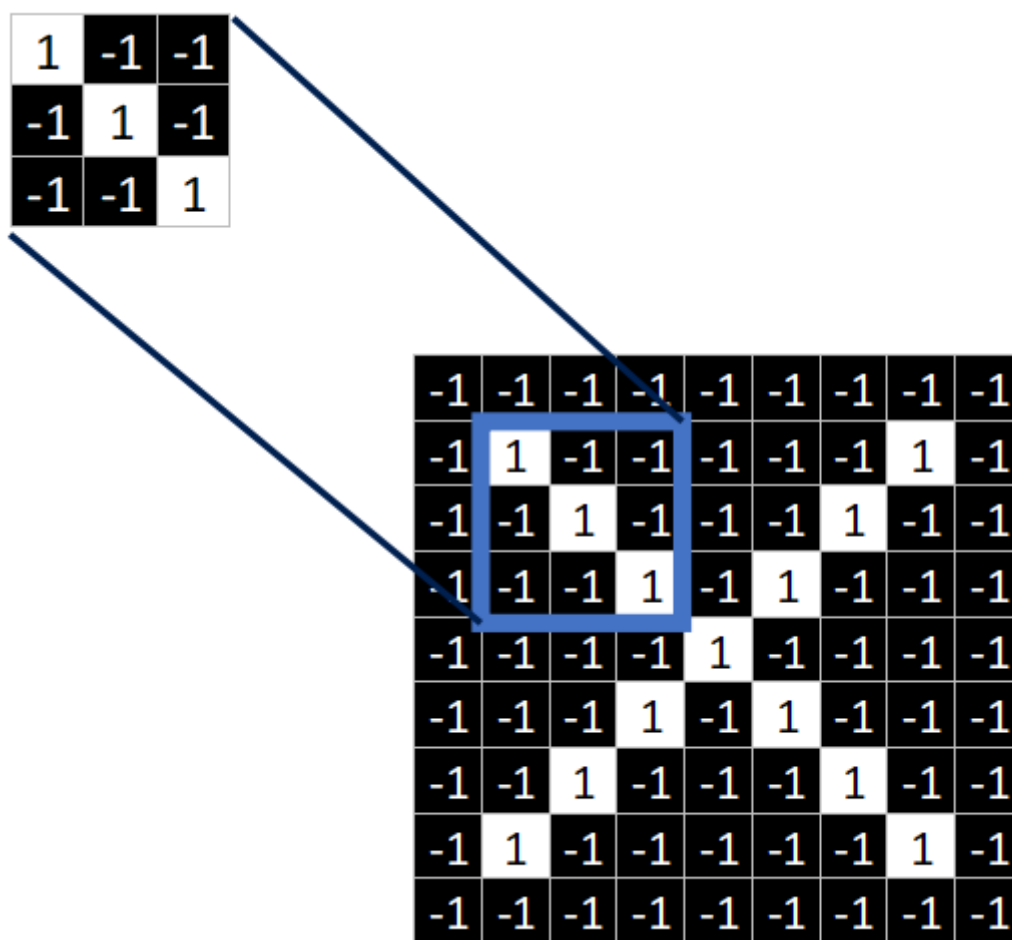
别急。

全连接层要做的，就是对之前的所有操作进行一个总结，给我们一个最终的结果。

它最大的目的是对特征图进行维度上的改变，来得到每个分类类别对应的概率值。

全连接层，顾名思义就是全部都连接起来，让我们把它与卷积层对比起来看。

这么说来的话前面的卷积层肯定就不是全连接了，没错，卷积层采用的是“**局部连接**”的思想，回忆一下卷积层的操作，是用一个3X3的图与原图进行连接操作，很明显原图中只有一个3X3的窗口能够与它连接起来。



那除窗口之外的、未连接的部分怎么办呢？我们都知道，采用的是将窗口滑动起来的方法后续进行连接。这个方法的思想就是“**参数共享**”，参数指的就是filter，用滑动窗口的方式，将这个filter值共享给原图中的每一块区域连接进行卷积运算。

敲一下黑板：局部连接与参数共享是卷积神经网络最重要的两个性质！

那么接下来再来看全连接神经网络。

还是拿9X9的输入原图做栗子，要进行全连接的话，那权值参数矩阵应该也是9x9才对，保证每一个值都有对应的权值参数来运算。（二者坐标直接一一对应）

1	-1	-1	1	-1	1	1	-1	1
-1	1	-1	-1	1	-1	-1	1	-1
-1	-1	1	1	-1	1	1	-1	1
-1	-1	1	1	-1	-1	-1	-1	1
-1	1	-1	-1	1	-1	-1	1	-1
1	-1	-1	-1	-1	1	1	-1	-1
1	-1	1	1	-1	1	1	-1	-1
-1	1	-1	-1	1	-1	-1	1	-1
1	-1	1	1	-1	1	-1	-1	1

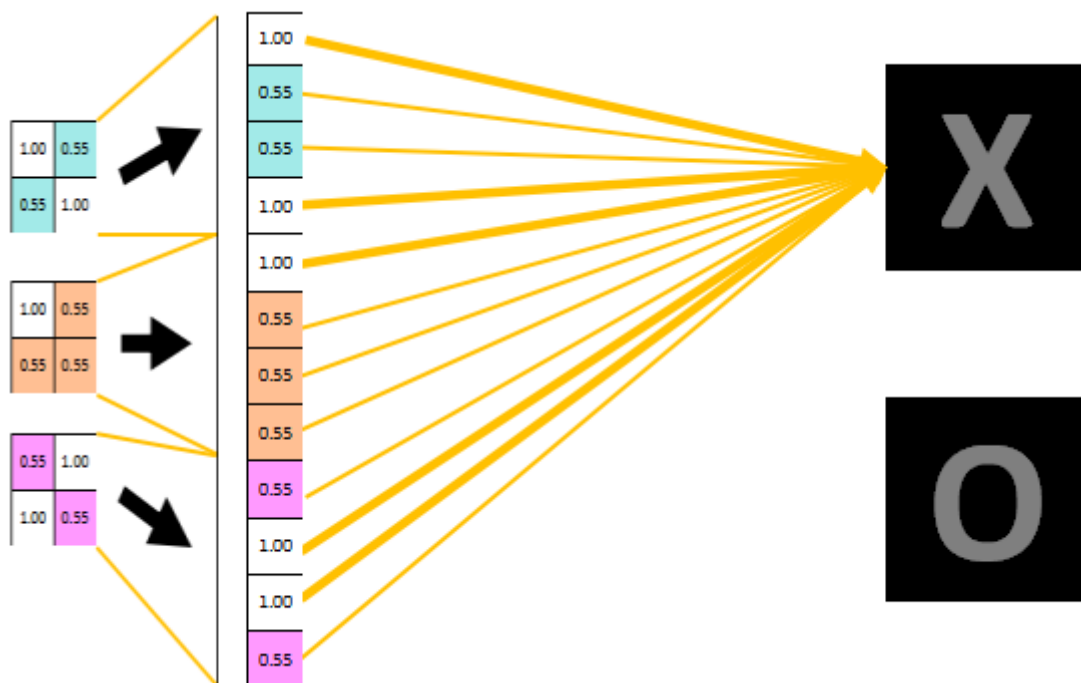
-1	-1	-1	-1	-1	-1	-1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	-1	-1	1	-1	-1	-1	-1
-1	-1	-1	1	-1	1	-1	-1	-1
-1	-1	1	-1	-1	-1	1	-1	-1
-1	1	-1	-1	-1	-1	-1	1	-1
-1	-1	-1	-1	-1	-1	-1	-1	-1

(手残.....图画的好丑.....意会到了就好，不要在意外表.....大概就是这个意思，但好像这个例子又举得不恰当，还是推荐大家直接去了解全连接神经网络。)

还是回来看接下来的操作，得到了2X2的特征图后，对其应用全连接网络，再全连接层中有一个非常重要的函数-----Softmax，它是一个分类函数，输出的是每个对应类别的概率值。比如：

【0.5, 0.03, 0.89, 0.97, 0.42, 0.15】就表示有6个类别，并且属于第四个类别的概率值0.89最大，因此判定属于第四个类别。

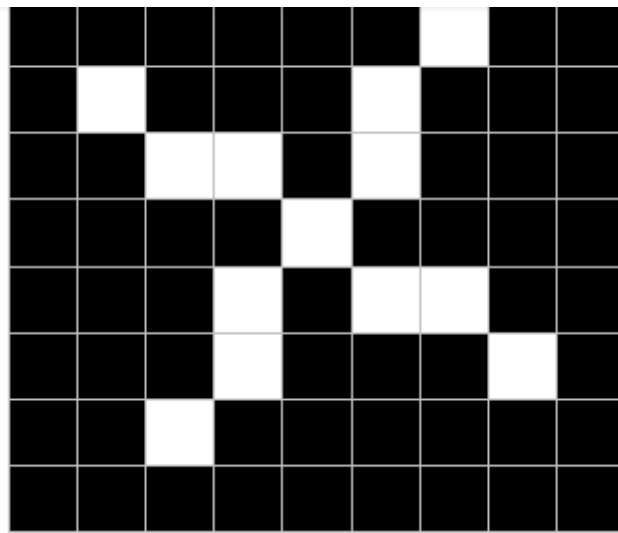
注意:本例中因为只有两个类别X和O，而且数据量到此已经非常少了，因此直接将三个特征图改变维度直接变成一维的数据。（相当于全连接层的每个参数均为1）



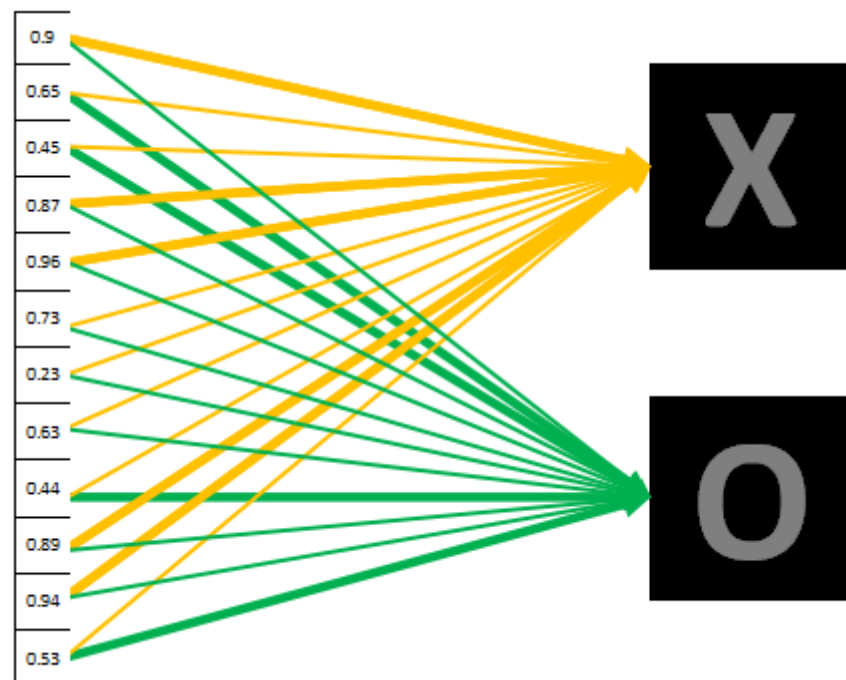
展开的数据即为属于类别X的概率值，值大小也在对应X的线条粗细中表现出来了。

以上所有的操作都是对标准的原图X来进行的，因此最终分类显示即为X毋庸置疑。

假设对一张看起来并不标准的图进行分类。如下



对于进行一系列操作后，假设得到的概率值如下所示：



0.9表示极大可能是X，因此对应到X的黄色线条比对应到O的绿色线条要粗很多很多。

我们对结果进行统计分析后可判断这张图片里的字母为X。



【6】神经网络的训练与优化

前面说了那么多，其实只是一个大致的框架的设计而已，里面的参数具体是多少则是需要训练的。

那么神经网络到底需要训练什么呢？

训练的就是那些卷积核（filter）。

针对这个识别X的例子，我们可以人为定义三个3X3的卷积核，便可实现对X的特征提取。但是在实际运用中，比如识别手写字母，几乎不可能存在标准的写法，每个人的字迹都完全不同，因此原来的那三个标准的卷积核就变得不再适用了，为了提高CNN模型的通用性（机器学习中的“**泛化能力**”），就需要对卷积核进行改写。经过成千上万的训练集来训练，每一次加入新的数据，都有可能对卷积核里的值造成影响。

那么具体的训练方法是什么呢？

在训练时，我们采用的训练数据一般都是带有标签label的图片。如果图片中的字母是X，则label=x，如果图片中的字母是A，则label=A。 标签能直观地反映图片。

在最开始，训练前，我们定义一个大小为3X3的卷积核，那么里面具体的值是多少，我们都不知道，但又不能为0吧，所以就用随机初始化法来进行赋值，卷积核获取到了一个随机值，便可以开始工作。

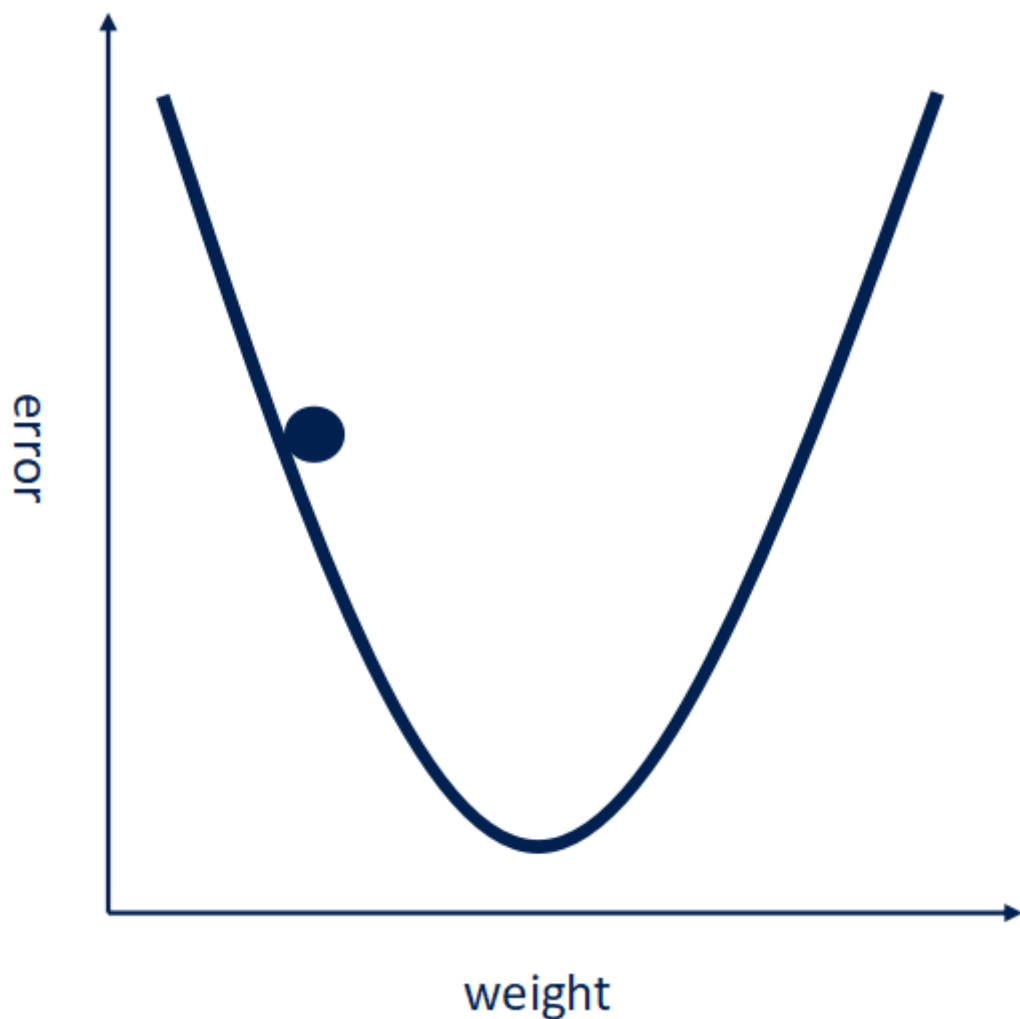
卷积神经网络便可以开始工作了，输入一张带有标签的图片（假设图片内容是字母X）。经网络识别后判断是X的概率为0.3。本来应该是1.0的概率，现在只有0.3，问题就很明显了，存在了很大的误差。

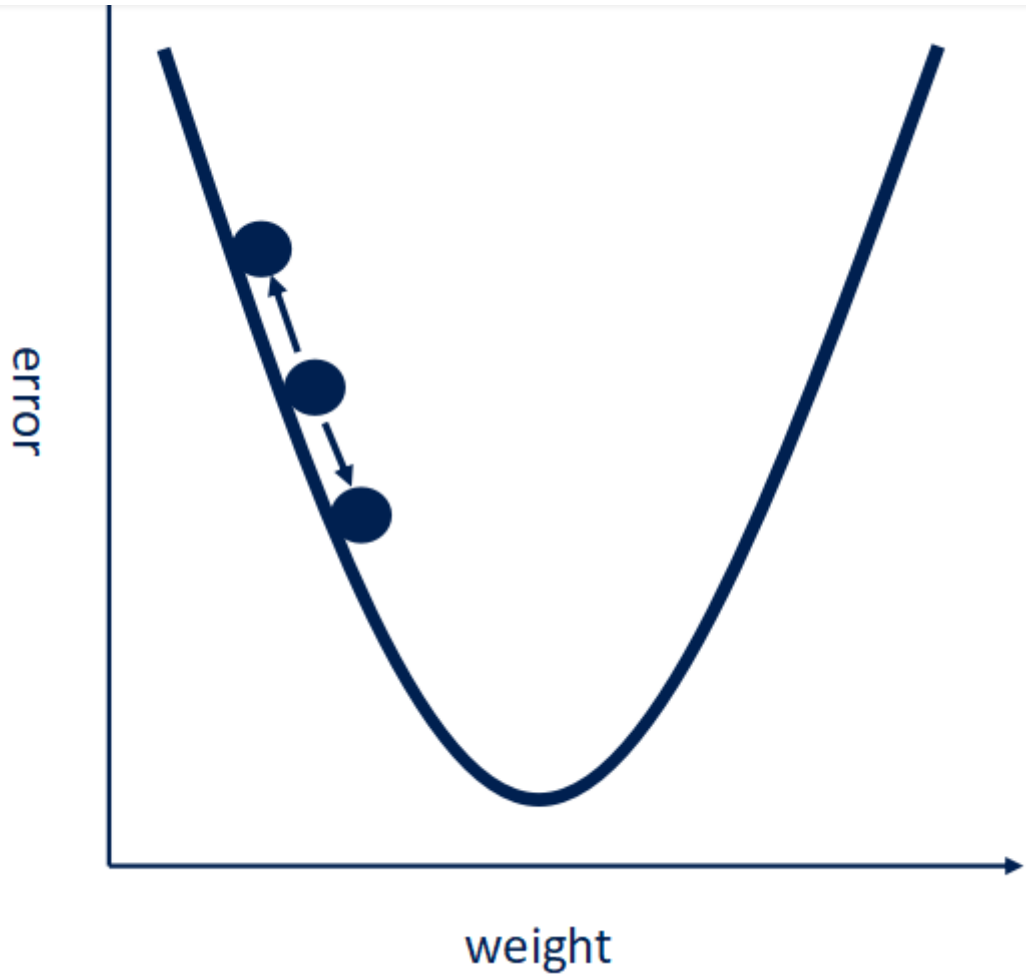
一种简单定义误差error的计算公式为 $error = (result - label)^2$

训练的终极目的就是使得这个误差最小，常用的方法是 **梯度下降法**。

内部设计的具体复杂公式在此不多做叙述。

简单的说可以参照下图，要使得误差error最小，就是让卷积核里的参数w往梯度下降最小的反向改变。





用这种方法来改变卷积核里的参数W使得误差最小。

在现有的各大深度学习框架中，CNN的优化可直接通过定义优化器解决，因此这里只是简单叙述原理以供了解。

写在最后：

CNN的部分到此先告一个段落，毕竟它只是个工具，我想把更多精力放在它的应用上，日后有想到什么更好的描述再另写（感觉有点潦草结尾啊.....）

接下来本专栏将会专注于图像处理相关，很多篇幅的重点写【风格迁移】，并会介绍另一个网络GAN生成式对抗网络。

欢迎各位提建议！

发布于 2017-07-28

深度学习（Deep Learning）

卷积神经网络（CNN）

神经网络

▲ 赞同 253 ▼

● 61 条评论

➤ 分享

★ 收藏

...

文章被以下专栏收录

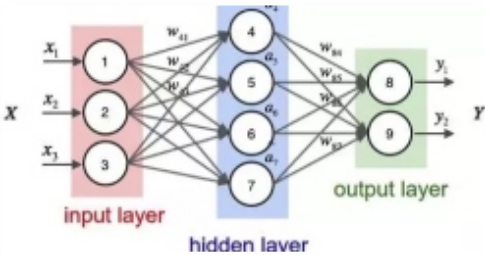


神经网络--从入门到放弃

专注深度学习的视觉领域，分享最新学习的论文

关注专栏

推荐阅读



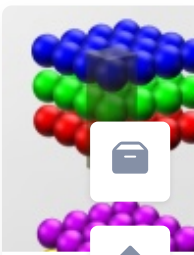
CNN的全面解析（带你简单轻

一个框架看懂优化算法之异同SGD/AdaGrad/Adam

Adam那么棒，为什么还对SGD念念不忘 (1) —— 一个框架看懂优化算法 机器学习界有一群炼丹师，他们每天的日常是：拿来药材（数

CNN经典结构解析（1）——VGG，Inception

前言上一篇讲了CNN的基本结构，主要分析了CNN的设计思路，以及各个基本零部件的功能。入门CNN之后，迎面而来的就是各种经典结



CNN入门么