

冠军的试炼

悟已往之不谏，知来者之可追

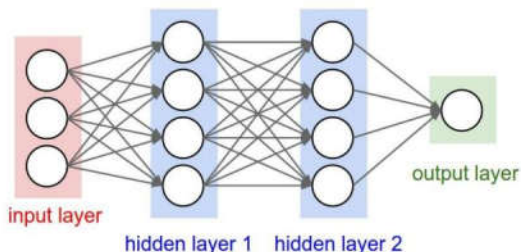
[博客园](#)[首页](#)[新随笔](#)[联系](#)[订阅](#)[管理](#)

随笔 - 69 文章 - 0 评论 - 817

卷积神经网络CNN总结

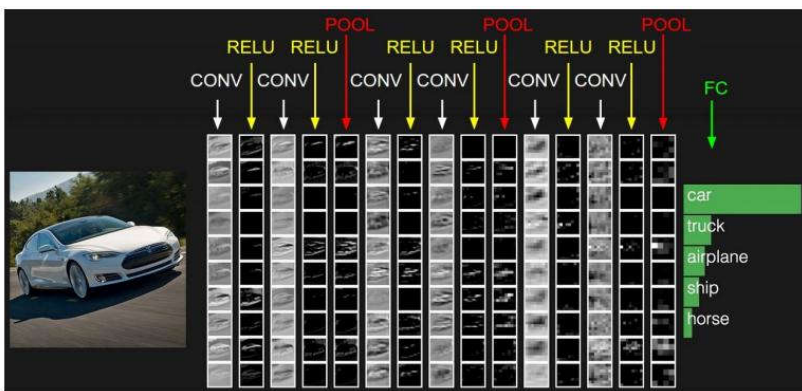
从神经网络到卷积神经网络（CNN）

我们知道神经网络的结构是这样的：



那卷积神经网络跟它是什么关系呢？

其实卷积神经网络依旧是层级网络，只是层的功能和形式做了变化，可以说是传统神经网络的一个改进。比如下图中就多了许多传统神经网络没有的层次。



卷积神经网络的层级结构

- 数据输入层 / Input layer
- 卷积计算层 / CONV layer
- ReLU激励层 / ReLU layer
- 池化层 / Pooling layer
- 全连接层 / FC layer

1.数据输入层

该层要做的处理主要是对原始图像数据进行预处理，其中包括：

- 去均值：把输入数据各个维度都中心化化为0，如下图所示，其目的就是把样本的中心拉回到坐标系原点上。
- 归一化：幅度归一化到同样的范围，如下所示，即减少各维度数据取值范围的差异而带来的干扰，比如，我们有两个维度的特征A和B，A范围是0到10，而B范围是0到10000，如果直接使用这两个特征是有问题的，好的做法就是归一化，即A和B的数据都变为0到1的范围。

公告

昵称: Madcola
园龄: 2年7个月
粉丝: 1334
关注: 30
[+加关注](#)

2019年8月						
日	一	二	三	四	五	六
28	29	30	31	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	5	6	7

搜索

常用链接

[我的随笔](#)
[我的评论](#)
[我的参与](#)
[最新评论](#)
[我的标签](#)

随笔分类⁽⁶⁹⁾

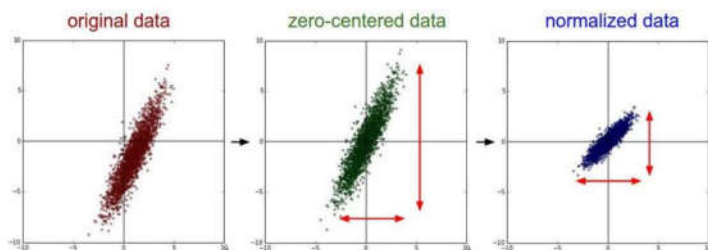
[C++\(1\)](#)
[CUDA\(1\)](#)
[Linux编程\(12\)](#)
[OCR系列\(8\)](#)
[opencv探索\(28\)](#)
[STL\(2\)](#)
[波折岁月\(4\)](#)
[工具技巧\(1\)](#)
[机器学习之旅\(5\)](#)
[深度学习\(4\)](#)
[数字图像处理\(3\)](#)

随笔档案⁽⁶⁹⁾

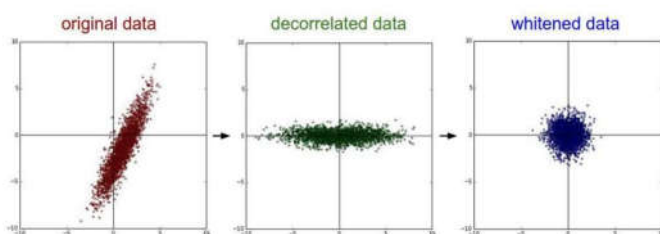
[2019年2月 \(1\)](#)
[2019年1月 \(1\)](#)

- PCA/白化: 用PCA降维; 白化是对数据各个特征轴上的幅度归一化

去均值与归一化效果图:



去相关与白化效果图:



2. 卷积计算层

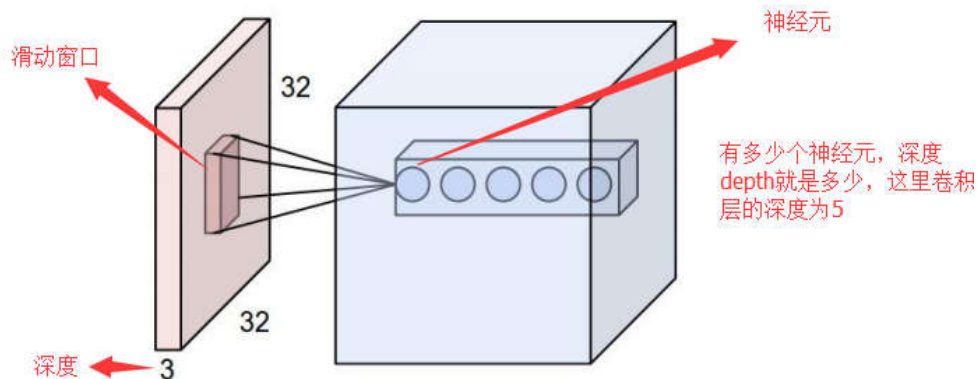
这一层就是卷积神经网络最重要的一个层次, 也是“卷积神经网络”的名字来源。

在这个卷积层, 有两个关键操作:

- 局部关联。每个神经元看做一个滤波器(filter)
- 窗口(receptive field)滑动, filter对局部数据计算

先介绍卷积层遇到的几个名词:

- 深度/depth (解释见下图)
- 步长/stride (窗口一次滑动的长度)
- 填充值/zero-padding



填充值是什么呢? 以下图为例, 比如有这么一个5*5的图片 (一个格子一个像素), 我们滑动窗口取2*2, 步长取2, 那么我们发现还剩下1个像素没法滑完, 那怎么办呢?

2018年12月 (2)
2018年10月 (1)
2018年9月 (3)
2018年5月 (1)
2018年4月 (2)
2018年2月 (6)
2018年1月 (3)
2017年12月 (4)
2017年11月 (3)
2017年10月 (1)
2017年9月 (4)
2017年8月 (3)
2017年7月 (5)
2017年6月 (4)
2017年5月 (17)
2017年4月 (1)
2017年2月 (2)
2017年1月 (5)

积分与排名

积分 - 213285
排名 - 1747

最新评论

1. Re: 【OCR技术系列之四】基于深度学习的文字识别 (3755个汉字)
感谢博主的无私奉献

--寒冬夜行人lee

2. Re: 我的2018: OCR、实习和秋招
学长好厉害! 我实习内容也是在做OCR, 每天照着你的好多博客看...我是18级研究生 2020年毕业后最近也在边实习边准备秋招 (但是我好菜), 可不可以加学长微信交流交流呀, 我的微信是Dreaminice.....

--Cocoalate

3. Re: 【Keras】基于SegNet和U-Net的遥感图像语义分割
@wenny-bell我也有这样的问题, 请问您解决了么? 可以加下qq讨论下, 2724858160...

--嗯哼!!!

4. Re: 【Keras】基于SegNet和U-Net的遥感图像语义分割
@wenny-bell我也有这样的问题, 请问您解决了么? 可以加下qq讨论下, 2724858160...

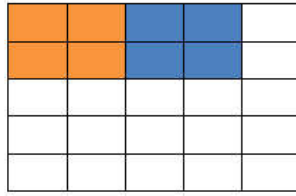
--嗯哼!!!

5. Re: OpenCV探索之路 (十一): 轮廓查找和多边形包围轮廓
请教下, `int thresh_size = (100 / 4) * 2 + 1;` //自适应二值化阈值这个阈值的定义是怎么给出的呢? 用你设定的参数确实得到了很好的效果, 但是不知道为什么这样设置. 谢谢~.....

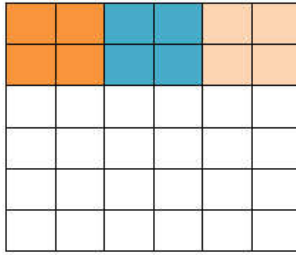
--foxlucia

阅读排行榜

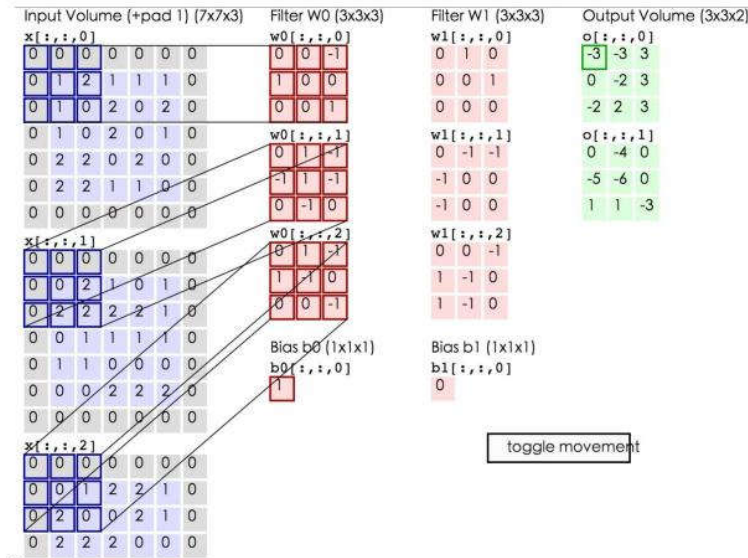
1. 卷积神经网络CNN总结(219292)
2. 基于深度学习的目标检测技术演进: R-CNN、Fast R-CNN、Faster



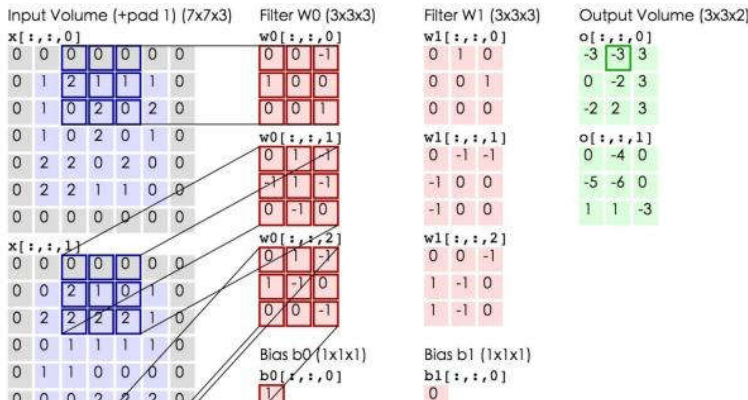
那我们在原先的矩阵加了一层填充值，使得变成6*6的矩阵，那么窗口就可以刚好把所有像素遍历完。这就是填充值的作用。



卷积的计算（注意，下面蓝色矩阵周围有一圈灰色的框，那些就是上面所说的填充值）



这里的蓝色矩阵就是输入的图像，粉色矩阵就是卷积层的神经元，这里表示了有两个神经元（w0,w1）。绿色矩阵就是经过卷积运算后的输出矩阵，这里的步长设置为2。



R-CNN(206008)

3. OpenCV探索之路（二十四）图像拼接和图像融合技术(88663)

4. CNN网络架构演进：从LeNet到

DenseNet(58616)

5. OpenCV探索之路（二十三）：特征检测和特征匹配方法汇总(54127)

6. 【OCR技术系列之四】基于深度学习的文字识别（3755个汉字）(51233)

7. C++ STL快速入门(45935)

8. OpenCV探索之路（六）：边缘检测（canny、sobel、laplacian）(45154)

9. Linux编程之UDP SOCKET全攻略(41919)

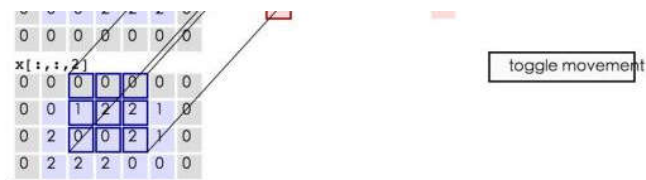
10. OpenCV探索之路（十四）：绘制点、直线、几何图形(37784)

评论排行榜

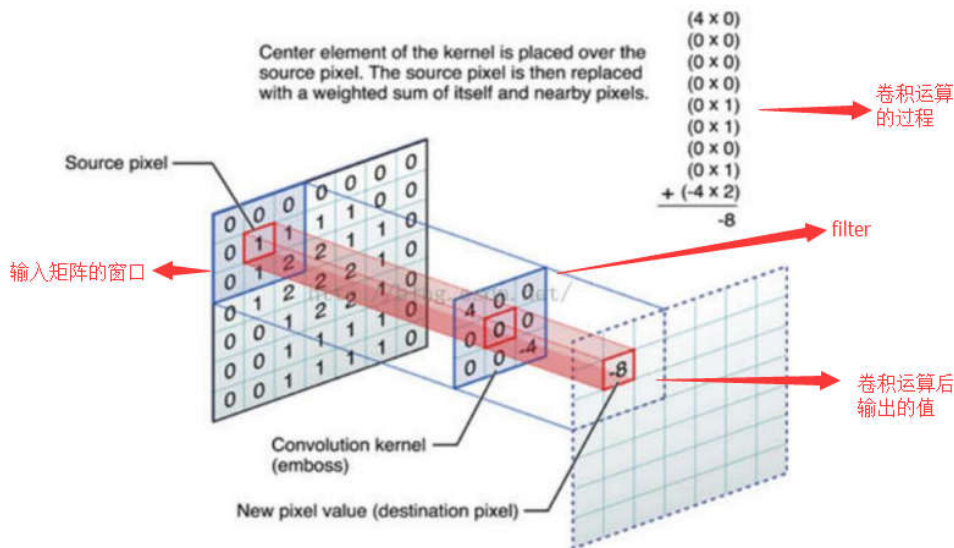
1. 【OCR技术系列之四】基于深度学习的文字识别（3755个汉字）(82)
2. 【Keras】基于SegNet和U-Net的遥感图像语义分割(73)
3. OpenCV探索之路（二十四）图像拼接和图像融合技术(67)
4. 【OCR技术系列之八】端到端不定长文本识别CRNN代码实现(59)
5. 【Keras】从两个实际任务掌握图像分类(33)

推荐排行榜

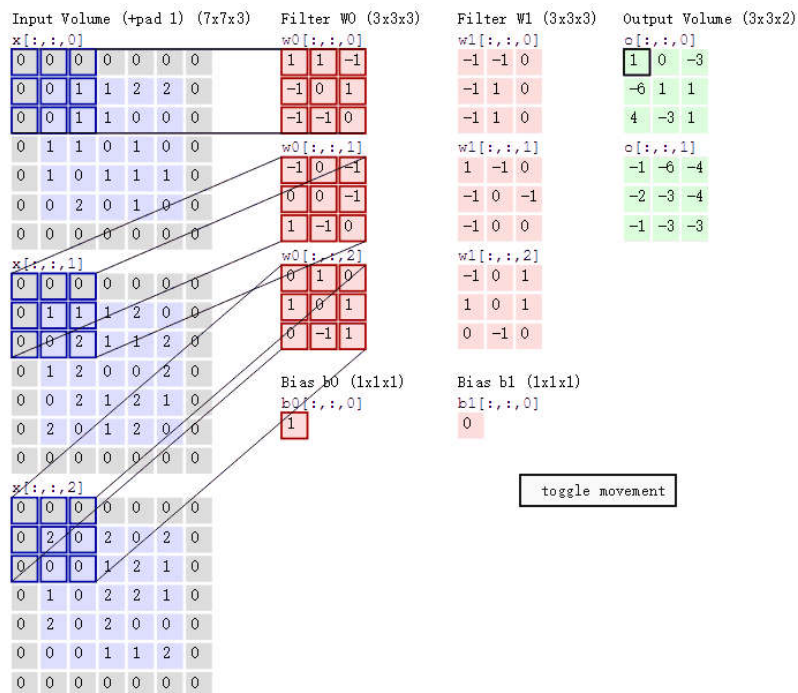
1. 基于深度学习的目标检测技术演进：R-CNN、Fast R-CNN、Faster R-CNN(92)
2. 卷积神经网络CNN总结(61)
3. 我的2018：OCR、实习和秋招(22)
4. 【OCR技术系列之四】基于深度学习的文字识别（3755个汉字）(21)
5. 【Keras】基于SegNet和U-Net的遥感图像语义分割(20)
6. CNN网络架构演进：从LeNet到DenseNet(20)
7. OpenCV探索之路（二十四）图像拼接和图像融合技术(18)
8. 我在北京实习的四个月(15)
9. 【OCR技术系列之一】字符识别技术总览(13)
10. 读研以来的一些感想：名校好在哪里？(13)



蓝色的矩阵(输入图像)对粉色的矩阵 (filter) 进行矩阵内积计算并将三个内积运算的结果与偏置值b相加 (比如上面图的计算: $2 + (-2 + 1 - 2) + (1 - 2 - 2) + 1 = 2 - 3 - 3 + 1 = -3$), 计算后的值就是绿框矩阵的一个元素。



下面的动态图形象地展示了卷积层的计算过程:

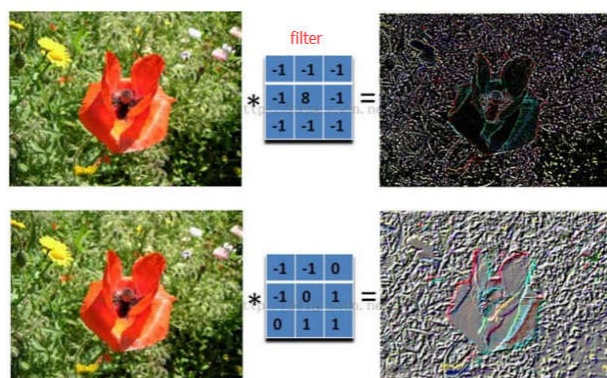


参数共享机制

- 在卷积层中每个神经元连接数据窗的权重是固定的, 每个神经元只关注一个特性。神经元就是图像处理中

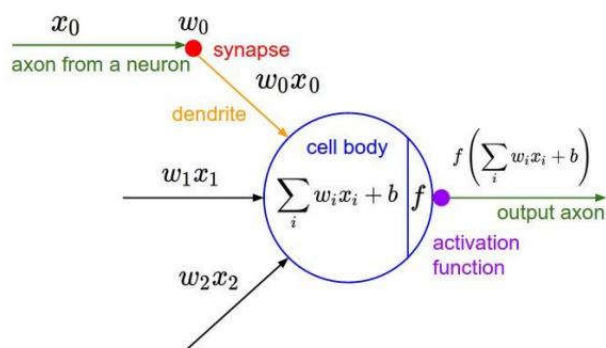
的滤波器，比如边缘检测专用的Sobel滤波器，即卷积层的每个滤波器都会有自己所关注一个图像特征，比如垂直边缘，水平边缘，颜色，纹理等等，这些所有神经元加起来就好比就是整张图像的特征提取器集合。

- 需要估算的权重个数减少: AlexNet 1亿 => 3.5w
- 一组固定的权重和不同窗口内数据做内积: 卷积

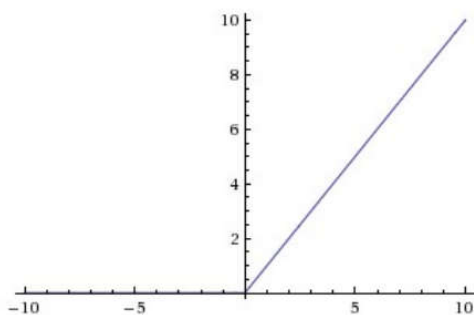


3.激励层

把卷积层输出结果做非线性映射。



CNN采用的激励函数一般为ReLU(The Rectified Linear Unit/修正线性单元)，它的特点是收敛快，求梯度简单，但较脆弱，图像如下。



激励层的实践经验：

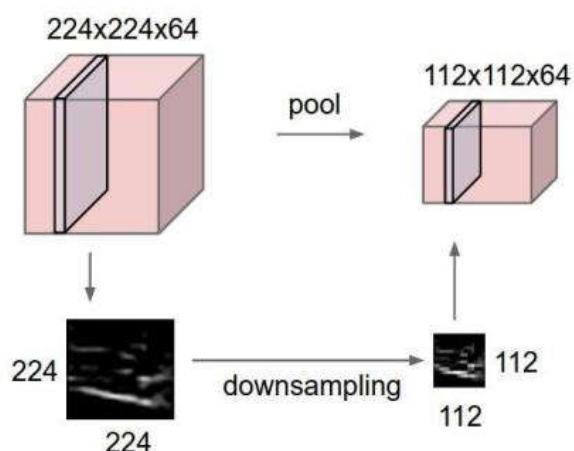
- ① 不要用sigmoid！ 不要用sigmoid！ 不要用sigmoid！
- ② 首先试ReLU，因为快，但要小心点
- ③ 如果2失效，请用Leaky ReLU或者Maxout
- ④ 某些情况下tanh倒是不错的结果，但是很少

4.池化层

池化层夹在连续的卷积层中间，用于压缩数据和参数的量，减小过拟合。
简而言之，如果输入是图像的话，那么池化层的最主要作用就是压缩图像。

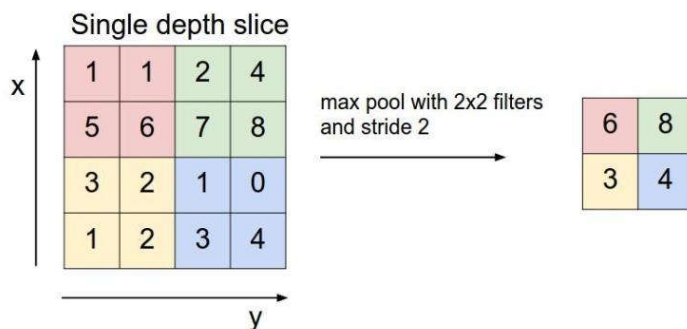
这里再展开叙述池化层的具体作用。

1. 特征不变性，也就是我们在图像处理中经常提到的特征的尺度不变性，池化操作就是图像的resize，平时一张狗的图像被缩小了一倍我们还能认出这是一张狗的照片，这说明这张图像中仍保留着狗最重要的特征，我们一看就能判断图像中画的是只狗，图像压缩时去掉的信息只是一些无关紧要的信息，而留下的信息则是具有尺度不变性的特征，是最能表达图像的特征。
2. 特征降维，我们知道一幅图像含有的信息是很大的，特征也很多，但是有些信息对于我们做图像任务时没有太多用途或者有重复，我们可以把这类冗余信息去除，把最重要的特征抽取出来，这也是池化操作的一大作用。
3. 在一定程度上防止过拟合，更方便优化。



池化层用的方法有Max pooling 和 average pooling，而实际用的较多的是Max pooling。

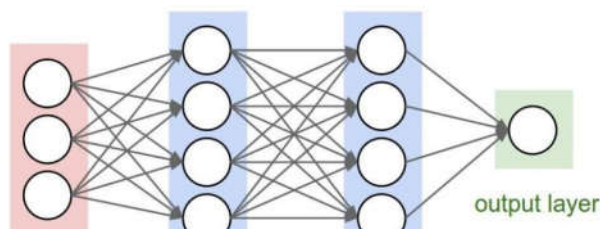
这里就说一下Max pooling，其实思想非常简单。



对于每个2*2的窗口选出最大的数作为输出矩阵的相应元素的值，比如输入矩阵第一个2*2窗口中最大的数是6，那么输出矩阵的第一个元素就是6，如此类推。

5. 全连接层

两层之间所有神经元都有权重连接，通常全连接层在卷积神经网络尾部。也就是跟传统的神经网络神经元的连接方式是一样的：





一般CNN结构依次为

1. INPUT
2. $[[CONV \rightarrow RELU] * N \rightarrow POOL?]*M$
3. $[FC \rightarrow RELU]*K$
4. FC

卷积神经网络之训练算法

1. 同一般机器学习算法，先定义Loss function，衡量和实际结果之间差距。
2. 找到最小化损失函数的W和b，CNN中用的算法是SGD（随机梯度下降）。

卷积神经网络之优缺点

优点

- 共享卷积核，对高维数据处理无压力
- 无需手动选取特征，训练好权重，即得特征分类效果好

缺点

- 需要调参，需要大样本量，训练最好要GPU
- 物理含义不明确（也就说，我们并不知道没个卷积层到底提取到的是什么特征，而且神经网络本身就是一种难以解释的“黑箱模型”）

卷积神经网络之典型CNN

- LeNet，这是最早用于数字识别的CNN
- AlexNet，2012 ILSVRC比赛远超第2名的CNN，比
- LeNet更深，用多层小卷积层叠加替换单大卷积层。
- ZF Net，2013 ILSVRC比赛冠军
- GoogLeNet，2014 ILSVRC比赛冠军
- VGGNet，2014 ILSVRC比赛中的模型，图像识别略差于GoogLeNet，但是在很多图像转化学习问题(比如object detection)上效果奇好

卷积神经网络之 fine-tuning

何谓fine-tuning?

fine-tuning就是使用已用于其他目标、预训练好模型的权重或者部分权重，作为初始值开始训练。

那为什么我们不用随机选取几个数作为权重初始值？原因很简单，第一，自己从头训练卷积神经网络容易出现问题；第二，fine-tuning能很快收敛到一个较理想的状态，省时又省心。

那fine-tuning的具体做法是？

- 复用相同层的权重，新定义层取随机权重初始值
- 调大新定义层的的学习率，调小复用层学习率

卷积神经网络的常用框架

Caffe

- 源于Berkeley的主流CV工具包，支持C++,python,matlab
- Model Zoo中有大量预训练好的模型供使用

Torch

- Facebook用的卷积神经网络工具包
- 通过时域卷积的本地接口，使用非常直观
- 定义新网络层简单

TensorFlow

- Google的深度学习框架
- TensorBoard可视化很方便
- 数据和模型并行化好，速度快

总结

卷积网络在本质上是一种输入到输出的映射，它能够学习大量的输入与输出之间的映射关系，而不需要任何输入和输出之间的精确的数学表达式，只要用已知的模式对卷积网络加以训练，网络就具有输入输出对之间的映射能力。

CNN一个非常重要的特点就是头重脚轻（越往输入权值越小，越往输出权值越多），呈现出一个倒三角的形态，这就很好地避免了BP神经网络中反向传播的时候梯度损失得太快。

卷积神经网络CNN主要用来识别位移、缩放及其他形式扭曲不变性的二维图形。由于CNN的特征检测层通过训练数据进行学习，所以在使用CNN时，避免了显式的特征抽取，而隐式地从训练数据中进行学习；再者由于同一特征映射面上的神经元权值相同，所以网络可以并行学习，这也是卷积网络相对于神经元彼此相连网络的一大优势。卷积神经网络以其局部权值共享的特殊结构在语音识别和图像处理方面有着独特的优越性，其布局更接近于实际的生物神经网络，权值共享降低了网络的复杂性，特别是多维输入向量的图像可以直接输入网络这一特点避免了特征提取和分类过程中数据重建的复杂度。

以下是我自己在学习CNN的时候遇到的一些困惑，以及查阅一些资料后得到的一些答案。

第一个问题：为什么不用BP神经网络去做呢？

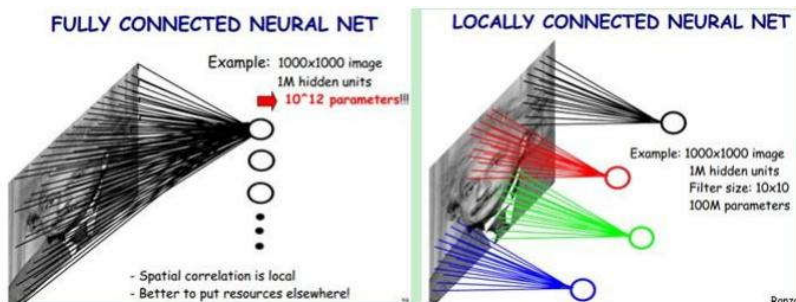
1. 全连接，权值太多，需要很多样本去训练，计算困难

- 应对之道：减少权值的尝试，局部连接，权值共享

卷积神经网络有两种神器可以降低参数数目。

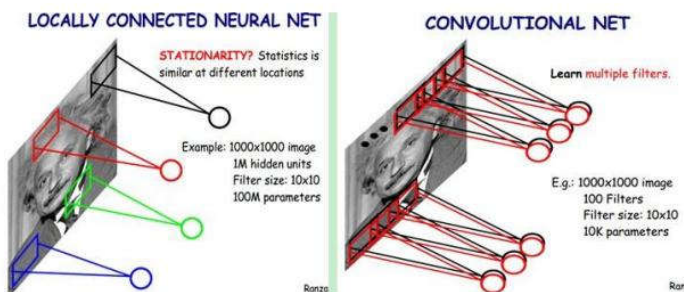
第一种神器叫做局部感知野，一般认为人对外界的认知是从局部到全局的，而图像的空间联系也是局部的像素联系较为紧密，而距离较远的像素相关性则较弱。因而，每个神经元其实没有必要对全局图像进行感知，只需要对局部进行感知，然后在更高层将局部的信息综合起来就得到了全局的信息。

第二级神器，即权值共享。



2. 边缘过渡不平滑

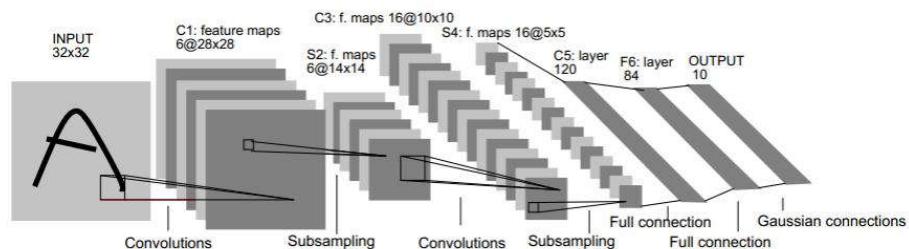
- 应对之道：采样窗口彼此重叠



第二个问题：LeNet里的隐层的神经元个数怎么确定呢？

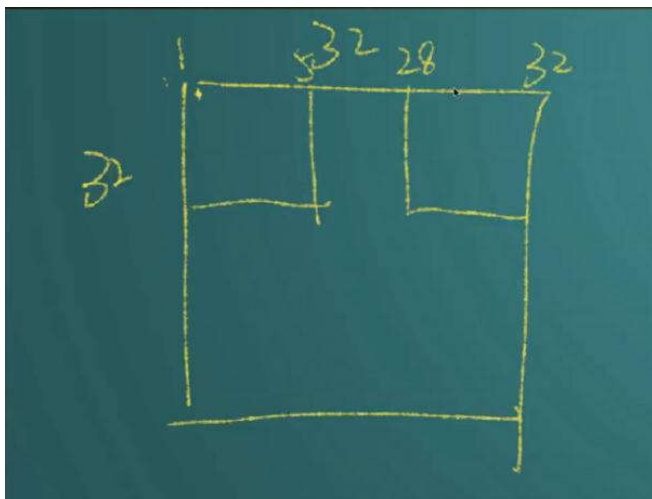
它和原图像，也就是输入的大小（神经元个数）、滤波器的大小和滤波器在图像中的滑动步长都有关！

LeNet-5共有7层，不包含输入，每层都包含可训练参数（连接权重）。输入图像为32*32大小。



例如，我的图像是1000x1000像素，而滤波器大小是10x10，假设滤波器没有重叠，也就是步长为10，这样隐层的神经元个数就是 $(1000 \times 1000) / (10 \times 10) = 100 \times 100$ 个神经元了。

那重叠了怎么算？比如上面图的C2中28*28是如何得来的？这里的步长就是1，窗口大小是5*5，所以窗口滑动肯定发生了重叠。下图解释了28的由来。



第三个问题：S2层是一个下采样层是干嘛用的？为什么是下采样？

也就是上面所说的池化层，只是叫法不同而已。这层利用图像局部相关性的原理，对图像进行子抽样，可以减少数据处理量同时保留有用信息，相当于图像压缩。

分类： 机器学习之旅

好文要顶

关注我

收藏该文



Madcola

关注 - 30

粉丝 - 1334

+加关注

« 上一篇: Linux编程之ICMP洪水攻击

» 下一篇: 基于深度学习的目标检测技术演进: R-CNN、Fast R-CNN、Faster R-CNN

posted @ 2017-04-30 20:10 Madcola 阅读(219295) 评论(21) 编辑 收藏

评论列表

#1楼 2018-05-09 14:41 老游条

实在看不懂，这种东西是不是智商没有140看不懂啊？

支持(4) 反对(26)

#2楼 2018-05-23 14:23 Andre_Ma

好文，顶一个

支持(3) 反对(0)

#3楼 2018-05-29 20:56 草长莺飞二月天

真好

支持(1) 反对(0)

#4楼 2018-06-08 19:11 一桶白开水

排版很舒服点个赞

支持(4) 反对(0)

#5楼 2018-07-04 14:08 Leon_zm

好文！

支持(1) 反对(0)

#6楼 2018-07-08 00:21 人工智能4708

楼主很用心的写了这篇文章，排版清晰，图文并茂，还加入了人自己的理解。让我终于对卷积神经网络有了大致的了解。之前看了花书上讲的东西，表示看不懂。

支持(6) 反对(0)

#7楼 2018-07-22 12:17 神思者123

绝对好文 🍻

支持(2) 反对(0)

#8楼 2018-08-26 09:01 merely06

中间那个动图没有截完整，误导初学者。。。希望博主看到能修改一下

支持(4) 反对(0)

#9楼 2018-10-25 10:45 岳麓笑笑生

@ 老游条

先看吴恩达的课吧，有基础就很好理解了

支持(0) 反对(0)

#10楼 2018-11-03 19:14 维常之华

通俗易懂，一目了然

支持(0) 反对(0)

#11楼 2019-01-02 19:51 fuzhihong0917

@ 老游条

发现自己智商不够140的应该不在少数吧。。。。。。

支持(0) 反对(0)

#12楼 2019-01-10 18:48 我知道你知道我知道

终于有一篇是我看的明明白白的了

支持(1) 反对(0)

#13楼 2019-01-19 20:10 fuzhihong0917

@ merely06

引用

中间那个动图没有截完整，误导初学者。。。希望博主看到能修改一下

中间卷积过程的gif，右边绿色的Output，有错误。希望博主可以修改一下。

支持(0) 反对(0)

#14楼[楼主] 2019-01-20 11:36 Madcola

@ fuzhihong0917

谢谢提醒，已更换新的gif图

支持(1) 反对(0)

#15楼 2019-02-12 16:05 冰封丽丝

这篇文章太棒了 条理特别清楚 讲的也很细致 对于初学卷积神经网络的小白 对卷积神经网络也有了大致的了解 有助于进行后面进一步的学习啊 之前看有关论文觉得太难懂 都快磨灭我的学习热情了。。。

支持(0) 反对(0)

#16楼 2019-02-26 10:04 秋野小姐

好文

支持(0) 反对(0)

#17楼 2019-02-26 16:47 realzero

一般cnn结构为"2. [[CONV -> RELU]*N -> POOL?]*M "; 对这点有疑问，一般每次卷积完都得激励吗，我看Alex模型，好像是在全连接的地方才用激励函数吧，还是说，每次卷积完都得激励，一个十几次卷积的网络模型需要激励十几次？不是很懂，望解答，谢谢

支持(0) 反对(0)

#18楼 2019-03-25 21:13 GIS_HKL

不错啊，学长

支持(0) 反对(0)

#19楼 2019-04-04 08:57 ktxi

写的真好

支持(0) 反对(0)

#20楼 2019-04-23 09:59 珠珠要加油

写的很好，终于有一篇能看懂了，谢谢分享

支持(0) 反对(0)

#21楼 2019-05-10 11:41 xiaohu_01

这篇文章太棒了 条理特别清楚 讲的也很细致，顶啊!!!!!!!!!!!!!!

支持(0) 反对(0)

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】超50万C++/C#源码: 大型实时仿真组态图形源码

【推荐】华为云·云创校园套餐9元起, 小天鹅音箱等你来拿

【推荐】零基础轻松玩转云上产品, 获赠礼加返百元大礼

相关博文:

- [卷积cnn总结](#)
- [从神经网络到卷积神经网络 \(CNN\)](#)
- [CNN卷积神经网络-tensorflow](#)
- [卷积神经网络CNN总结](#)
- [卷积神经网络CNN](#)

最新新闻:

- [一线 | 阿里回应即将上线独立网约车平台: 没有的事](#)
 - [在高通、华为、三星之间寻找“机会”的联发科](#)
 - [美国首部 CRISPR 法律警告不要 DIY 自己的 DNA](#)
 - [知乎周源发融资全员信强调工作效率: 快则生慢则死](#)
 - [到处作恶的台风究竟是怎么来的? 背后竟藏着这些秘密...](#)
- » [更多新闻...](#)