

机器学习练习题答案

通过习题，帮助理解机器学习，当然我们更专注于深度学习。

卷积神经网络

1. 相比全连接深度神经网络（DNN），卷积神经网络（CNN）在图像识别上有什么优势？
- 连续的层部分连接，权值重用，有比全连接的DNN更少的参数，使得训练速度更快，减少了过拟合的风险，并且只需要更少的数据；
 - 当CNN学习到可以检测某个特征的内核后，它可以在图片的各个位置都检测到该特征，相反，当DNN在一个位置学习一个特征时，它只能在特定位置检测到它；因为图片通常具有很强的重复特征，在使用更少的训练样本的前提下，CNN可以能够比DNN更好地推广用于图像处理任务，如分类。
 - 最后，DNN对像素是如何组织的没有先验知识，但CNN的架构嵌入了这个先验知识。较低层通常在图像的小区域内识别特征，而较高层将较低层特征组合为更大特征，这在大多数的自然图像上都取得了很好的效果，让CNN比DNN取得了决定性的领先。
2. 假设构建一个有3层卷积的CNN，网络结构是这样，每层的核为3X3，步长（strides）为2，相同填充（Same Padding），第一层输出100个特征图（feature map），第二层输出200个，第三层输出400个。输入是200X300像素的RGB图像。这个网络的参数数目是多少？如果参数的类型都是32bits的float，那么当网络对单张图片进行学习时，最小需要使用多少内存（RAM）？当处理50张图片时又会占用多大的内存？
- 参数数目
 - 一：输入通道3维，核大小3x3，共3x3x3，100个特征图100x3x3x3，同时加上每个特征图的偏置项，100x3x3x3+100；
 - 二：输入通道100维，核大小3x3，共3x3x3，200个特征图200x3x3x100，同时加上每个特征图的偏置项，200x3x3x100+200；
 - 三:输入通道200维，核大小3x3，共3x3x200，400个特征图400x3x3x200，同时加上每个特征图的偏置项，400x3x3x200+400；
 - 训练一个实例时需要的内存
 - 包括每个特征图所占的内存及参数所占内存；
 - 200X300像素的图像经过这样（每层的核为3X3，步长（strides）为2，相同填充（Same Padding））的卷积层，得到的特征图大小分别为100x150，50x75，25x38，每个数值数据32bits=4B，如表中所示，每个实例在三层卷积过程中占用的内存约为10Mb，加上网络中的参数约为13.479MB；
 - 训练50个实例时需要的内存
 - 当使用反向传播进行训练时，之前的所有计算参数都需要保留，卷积层占内存50x10=500MB，参数3.4MB，输入图片50x200x300x3x4B=34.332MB，共计约537.7MB。

		一	二	三	总数
参数数目		100x(3x3x3+1)=2800	200x(3x3x100+1)=180200	400x(3x3x200+1)=720400	903400
特征图大小		100x150	50x75	25x38	
占内存大小	903400x4B=3.446MB	100x100x150x4B=5.722MB	200x50x75x4B=2.861MB	400x25x38B=1.450MB	13.479

3. 当训练一个CNN网络时，GPU的内存耗尽，可以怎样来解决这个问题？

即寻找可以较少网络训练占内存的方法，由上题，可以通过减小每次处理的输入数量，参数数目，步长等，或者直接扩展硬件条件；

- 减小每个batch size的大小；
- 在一层或多层中使用更大的步长来减小维度；
- 移除一个或多个卷积层；
- 使用16bit类型数据而不是32bits；
- 使用集群来训练CNN。

4. 步长一样，为什么要添加一个最大池化层而不是卷积层？

因为最大池化层是不增加训练参数的，而使用卷积层则会增加很多参数。

5. 什么时候添加一个局部响应标准化层（local response normalization layer）？

local response normalization layer是在同一位置激活最强的神经元会抑制其他特征图中的神经元，这使得每个特征图都能检测到自己独特的特征，并与其他特征图区别开来，并迫使它们去寻找更多的特征，local response normalization layer通常在较低的层使用，以使更高更深层次的特征能建立在更多的底层特征之上。

6. 相比LeNet-5，AlexNet有什么创新？GoogleNet和ResNet呢？(在著名的Imagenet竞赛中，AlexNet, GoogleNet, ResNet分别是2012,2014和2015年的取得最高准确率的方案)

- AlexNet
 - 打破常规，在卷积层上直接叠加卷积层，而不是每个卷积层上必定要加一个池化层。
- GoogleNet
 - 提出inception modules，使得CNN网络可以有比之前所有网络更深的深度，同时参数更少。
- ResNet
 - 提出skip connections，是的网络的层数可以突破100层。

7, 8, 9代码的实现将会[上传到这](#)

7. 构建一个CNN网络，尝试在MNIST数据上取得尽可能高的准确率。

8. 使用Inception v3分类大图像。

9. 大图像分类的迁移学习。