

实训周报

15331416 赵寒旭

1. 学习内容

阅读两篇论文，并做了简单的阅读报告。

1) 论文名字

Alex Krizhevsky, Ilya Sutskever and Geoffrey E. Hinton: **ImageNet Classification with Deep Convolutional Neural Networks**

Karen Simonyan, Andrew Zisserman: **Very Deep Convolutional Networks for Large-scale Image Recognition**

2) 论文内容简述

	AlexNet	VGGNet
简介	2012 年，AlexNet 以显著优势获得了 ImageNet 的冠军，训练了到 2012 年为止最大的卷积神经网络之一，top-5 错误率降低到了 16.4%，相比于第二名 26.2% 的错误率有了巨大的提升。	研究了卷积网络深度对大型图像识别准确性的影响，主要评估小的卷积核 (3×3) 同等架构下随着网络深度的增加卷积网络的性能变化，随着网络深度达到 16-19 层，网络的性能也有着显著的提升。
主要贡献	① 介绍了完整的 CNN 架构模型和多种训练技巧； ② 实现了多 GPU 上的训练，加速了大型卷积操作；	① 证明了深度在卷积神经网络中的重要性； ② 说明相比于直接使用大的卷积核，多个小卷积核的堆叠可能会有更好的效果；
数据处理	系统要求固定尺寸 256×256 输入。 方法：将图像缩放至短边长为 256 再剪裁中间的 256×256 区域。 预处理：每个像素点减去 RGB 的均值。	输入固定尺寸 224×224 图像。 预处理：每个像素点减去 RGB 的均值。
网络结构	5 个是卷积层，3 个全连接层，最后一个全连接层输出一个 1000 维的 softmax 来表达对于 1000 个类别的预测； 响应归一化层在第一、二个卷积层后面； 最大值池化层跟在响应归一化层和第五个卷积层后面； ReLU 被应用在每个卷积层和全连接层。	所有卷积层的卷积核大小都为 3×3 ，共经过 5 个最大池化层，卷积核数目从 64 开始，每经过一个池化层卷积核数翻倍直到 512 个时停止。 最后跟 3 个全连接层和一个 softmax 层，每个隐含层后面都跟一个 ReLU 激活函数进行非线性处理。
训练细节	① 激活函数采用 ReLU，使学习周期大大缩短，提高速度和效率；	① 激活函数为 ReLU ② 除了一个网络之外，我们所有的网络都不含局部响应归一

	<p>② Local Response Normalization (局部响应归一化) 模拟生物神经元的侧抑制作用——当前神经元的作用受附近神经元作用的抑制;</p> <p>③ 重叠池化: AlexNet 使用的池化层是可重叠的, 池化时每次移动的步长 s 小于池化的边长 z, 效果比 $s = z = 2$ 的传统池化效果要好, 在产生相同维度的输出时分别将 top-1 和 top-5 的错误率降低了 0.4% 和 0.5%。观察还发现, 采用有重叠的池化能让模型更难过拟合</p>	<p>化, 它不能提升在 ILSVRC 数据集上的性能, 但会导致内存消耗和计算时间的增加。</p> <p>③ 采用最大池化: 在 2×2 的窗口进行, 步长为 2</p> <p>④ 使用非常小的 3×3 的滤波器 (是足够捕获一个区域信息的最小的卷积核) 其中一个配置中使用 1×1 的卷积滤波器 (对输入通道的线性变换), 这样的卷积不改变输入通道的维度, 且可以提高模型的学习能力。</p>
方法介绍	<p>① 图像平移和水平镜像: 测试过程中网络会抽取 5 个 (4 角和中间) 224×224 的区块及其水平镜像进行预测, 然后将 softmax 层对这十个区块做出的预测取平均;</p> <p>② RGB 通道的亮度改变: 对训练集中的每张图片都做主成分分析 (PCA), 得到对应的特征向量和特征值。训练过程中, 每当这张样本被重复一次, 我们就对其每一个像素按照公式进行变换, 得到一张“新”样本做下一轮的训练。通过这种方法我们可以获得富有多样性的样本;</p> <p>③ Dropout: 每轮迭代时, 网络中每个隐含层的输出节点都有一半的可能性被置为 0</p>	<p>① 网络权重的初始化很重要, 因为深层网络中的梯度不稳定, 初始化不好可能会导致学习停滞。为此先训练一个深度较浅的网络 A, 再利用已经训练好的 A 网络权值初始化深度较深的网络。(迁移学习)</p> <p>② 输入图像获取: 为了获得固定大小的 224×224 的卷积网络输入图像, 他们从重新缩放的训练集图像中随机裁剪 (每个 SGD 迭代的每张图像进行一次裁剪);</p> <p>(? 这里的操作和 Alex 直接抽取 5 个区块不同, VGG 是先缩放再裁剪一块)</p> <p>③ 数据增强的方法: 同 Alex, 使用随机水平翻转和 RGB 颜色偏移;</p>
观点分歧	<p>在第一、二个卷积层后面跟响应归一化层。</p>	<p>除了一个网络之外, 所有的网络都不含局部响应归一化, 它不能提升在 ILSVRC 数据集上的性能, 但会导致内存消耗和计算时间的增加。</p>

2. 遇到困难以及解决方案, 或者学习收获

很多名词和过程刚接触时不能很快理解, 通过查找网上的博客和一些深度学习的课程资料建立起了相关概念。

我个人是一边看一边查找资料整理学习报告，速度比较慢，因为平时作业也比较多，两篇文章都放在实训时间阅读，但是两个星期也没有看完，展示之前又另外抽了大半天时间才做完了 VGG 的总结。

论文阅读分享的环节其实也没有来得及准备什么，只提供了简单整理的总结做展示，其实也没有做 ppt，之后有机会会努力准备尽量讲清楚一些。

学习收获：基本了解卷积神经网络的结构，但具体细节还需要仔细理解。

3. 下周计划。

继续整理两篇论文的阅读总结。

熟悉 pytorch 框架，并且自己实现 alexnet 的网络结构