# 深度卷积网络：实例探究

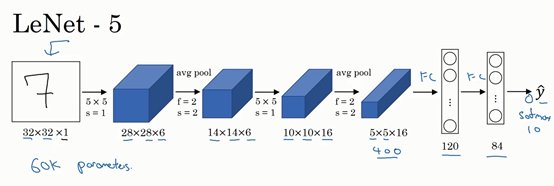
## 为什么要进行实例探究

过去几年的很多研究都是在探究如何用基本构件来搭建有效的网络结构，而我们学习卷积神经网络，最直观的方式之一就是看看他们的做法，学习他们的结构和搭建方式。而在一项任务中表现优异的网络结构，很可能也适用于其他任务。这对我们也有参考价值。

这周我们要介绍一些经典的网络结构和比较新颖的网络结构。经典网络结构如LeNet-5，AlexNet，VGG-16，比较新的结构有resNet和inception Net。

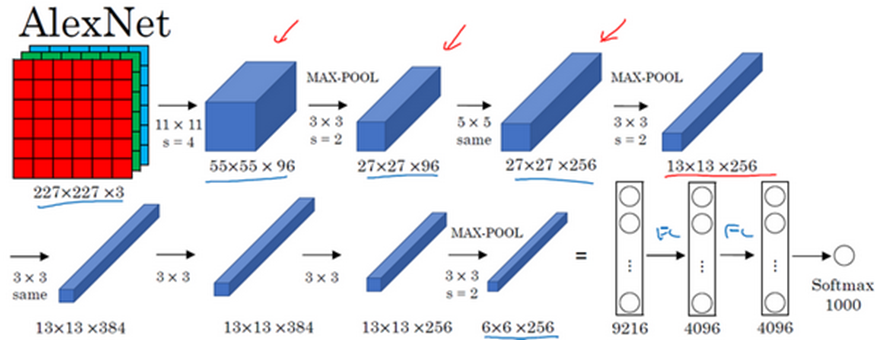
## 经典网络

首先我们来看LeNet-5，这个网络是Lecun在1994年的论文中提出的，目的是为了进行手写数字识别。图像是灰度图像。结构如图所示



在那个论文的年代，当时的人们更习惯于用平均池化而不是最大池化，当时的人们也不使用padding，所以每次卷积过后图像也在变小。分类器现在使用softmax，当时则是用其他分类方法。这个神经网络只有6w个参数，随着网络的加深，通道数量越来越多，而图像尺寸越来越小。这个搭建神经网络的观念一直延续到现在。还有一点就是，多个或一个卷积层之后就会跟上池化层，这个观念也是一直被采纳的。过去，人们使用sigmoid函数或者是tanh函数而不是relu，而且池化后还使用了sigmoid函数。

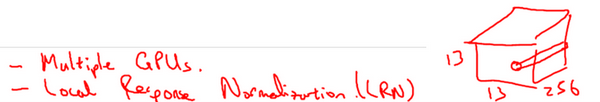
具体lenet的代码请看复现代码。



接着我们来看AlexNet，是以论文的第一作者Alex Krizhevsky的名字命名的，另外两个作者是是Ilya Sutskever和Geoffery Hinton。大概结构如图所示。

实际上，AlexNet和LeNet有很多相似之处，但AlexNet更大，功能也更强，含有大概6000w个参数。

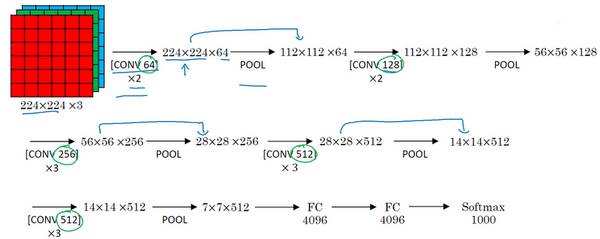
同时，AlexNet采用了relu激活函数，大大加快了反向传播的速度。

AlexNet的论文中还提到了局部相应归一化这个方法， 如图所示，

相当于取图像中的某个像素点，对所有通道同时进行归一化操作，但是后来，很多研究者发现LRN起不到太大作用，所以我们并不适用LRN来进行网络训练。最后两个全连接层使用了dropout方法，有效防止了过拟合。

同时，论文中还提出了一系列的图像增强算法。

AlexNet在图像识别领域取得了突破性进展，是里程碑之一。

接下来介绍VGG net，vgg结构简单，层次多，没有像alexnet一样那么多的超参数。结构如图所示，卷积层+全连接层一共16层，卷积层特征数的增加和尺寸的减小都有规律可循，可以说是结构简单。

## 残差网络(resNet)

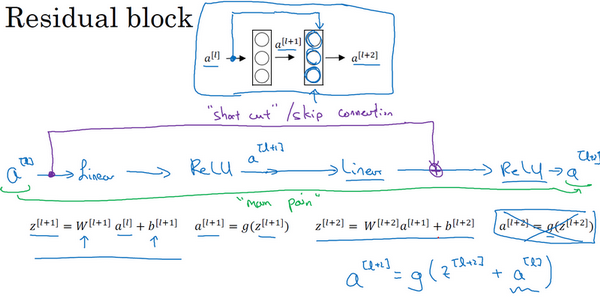
按常理来说，越深的神经网络，参数越多，拟合能力越强，然而我们会发现，当网络大到一定程度时，再增加网络的深度，反而会让网络变得难以训练。

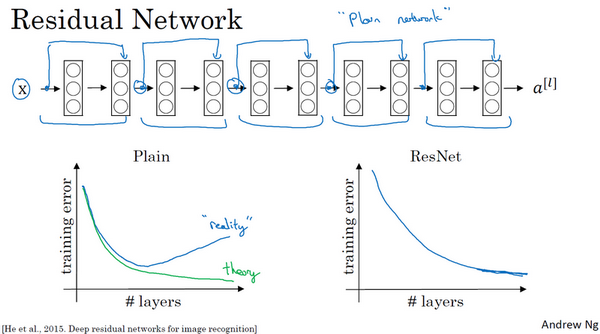
这样考虑，就算比原来再额外增加一层，最坏也就是和输入层次形成相同的映射，为何会导致模型变差呢？其实，网络层次变多后，梯度下降与反向传播变得很困难，没有办法那么容易学到全等映射。

为了解决这个问题，我们可以通过相隔n层，直接把第k层的输出传到k+n层去，不就能直接得到全等映射了吗？而且由于我们的网络还在学习，我们可以得到比全等映射更强的函数，由此一些科学家提出了resNet的结构。

使用resnet，有时层次甚至可以超过100层，resnet使用了跳跃连接的结构。

如图所示，

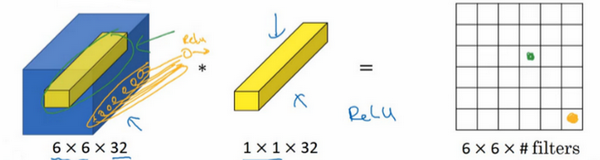
我们得到，想将传播到后两层，我们可以直接通过主路径传达，也可以通过中间的捷径，直接将传到层，。

我们设 = y, 我们原来要学习的映射是y = x，现在我们令y = x+f(x),我们只需学习出一个f(x) = y-x的映射即可，这种映射比全等映射容易学习，就让我们的网络能够更深。

我们在向后传递激活值时，需要保持维度一致，所以残差网络使用了大量的same卷积。如果维度改变，可以通过乘上一个矩阵来更改维度。如果通道改变，可以使用1x1的卷积来改变通道。

## 网络中的网络及1x1卷积

如图所示，我们正在使用一个1x1的卷积核和原图形进行卷积运算



这个1x1x32的滤波器中的32个数字可以这样理解，把原图形作为输入，他的每个输入神经元都和这个32个数字中的某一个数字相乘，得到相应的输出结果。如果有多个滤波器，就相当于结果中有多个神经元。这种方法叫做1x1卷积，也叫做network in network。虽然这个论文中的架构没有被广泛使用，但是1x1卷积这个方法却被广泛使用了。

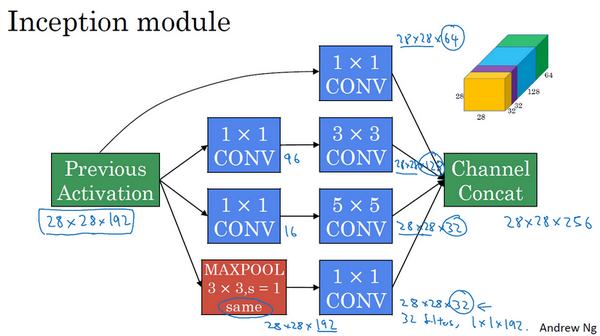
同时，1x1卷积核可以用来进行升维或者降维。

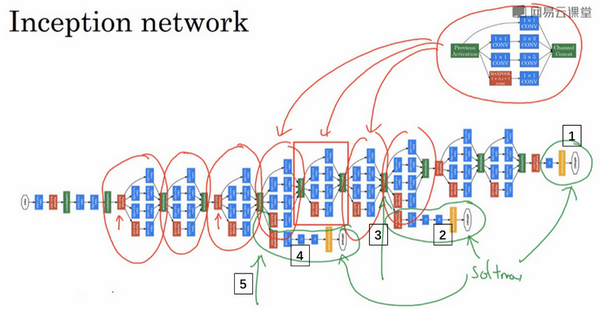
## Google inception网络

在构建网络的时候，你会考虑你用多大的池化层，多大的卷积层，而inception网络就可以替你做出决定。

值得一提的是，inception这个词来自于与盗梦空间，其中有句台词叫we need to go deeper。

如图所示，这是一个inception模块

，而inception网络就是由多个这样的模块构成的，可以由网络自行选择所需要的卷积与池化参数。



通过inception网络，可以让网络变得很深。

## 使用开源的实现方案与迁移学习

当你要实现某项任务时，如果别人已经有过相关经历，你完全可以去网络时使用开源的实现方案，或者是下载预训练模型进行迁移学习，尤其是你的数据量比较小的时候。

## 数据增强

大部分计算机视觉中的的CNN网络都需要喂入大量的数据，而我们的数据很可能不够，这时使用数据增强就是一个比较好的做法。常用的做法有改变通道颜色，截取图像，翻转图像，PCA颜色增强等方法。