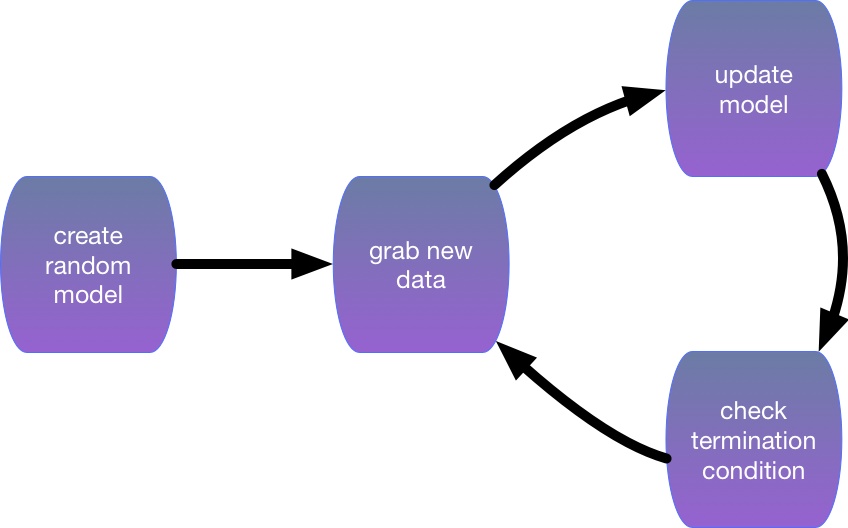
机器学习简单介绍

机器学习是一种数据驱动的学习方式，它的流程如下：



1. 初始化一个几乎什么也不能做的模型；
2. 抓一些有标注的数据集；
3. 修改模型使得它在抓取的数据集上能够更准确执行任务；
4. 重复以上步骤2和3，直到模型看起来不错。

为了更好的解释机器学习的流程，给出一个垃圾邮件分类的例子：

1：分析任务：垃圾邮件分类，属于二分类问题，需要选择一个分类模型，常用的二分类模型有LR、SVM、DNN等，这些模型目前都已经集成只需要调用即可，本例中选择简单的SVM模型；

2：获取标注的数据，垃圾邮件标注为1，正常邮件标注为0

3：将邮件内容作为模型输入，得到模型的输出，并将该输出与输入邮件的标注1或者0想对比，调整模型的各种参数使得模型的输出与标注值的差值在允许误差范围内。

邮件

模型

1（垃圾邮件）

0（正常邮件）

机器学习的核心任务就是获得一个能够解决各种实际问题的模型。

机器学习的两个基本应用：

1：回归分析：回归分析也许是机器学习里最简单的一类任务。在该项任务里，输入是任意离散或连续的、单一或多个的变量，而输出是连续的数值。例如我们可以把本月公司财报数据抽取出若干特征，如营收总额、支出总额以及是否有负面报道，利用回归分析预测下个月该公司股票价格。

2：分类，当模型的输出是离散的类别时，这个机器学习任务就叫做分类。分类在我们日常生活中很常见。例如我们可以把本月公司财报数据抽取出若干特征，如营收总额、支出总额以及是否有负面报道，利用分类预测下个月该公司的CEO是否会离职。在计算机视觉领域，把一张图片识别成众多物品类别中的某一类，例如猫、狗等。

计算机视觉深度模型介绍

一：CNN Architectures

1、卷积层

CNN卷积神经网络是由人工神经网络（ANN）演变而来的。如果直接用ANN对图像进行分类，会将(W,H)大小的三维数组图像扁平化为一维数组作为输入。这样做存在的问题是，扁平化后无法捕捉到图像的局部结构。为了解决这一问题，LeCun等人[13]根据图像处理中卷积操作的思想设计了卷积层。在机器学习中卷积就是局部特征乘以对应的权重，然后再累加起来即可。卷积层的输入是一个二维数组，它使用一个可上下滑动的卷积核对输入做特征变换，最终输出一个二维数组。其转换过程如下：

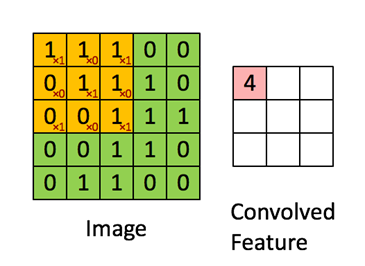


图2-1 卷积层

a. 一个(k, k)大小的卷积核在大小为(W,H)的二维数组中从左到右，从上到下滑动。

b. 每次滑动时将卷积核内输入元素x进行加权求和，再通过一个激活函数得到一个输出y，即y=f(wx+b)。

c. 根据卷积核滑动的位置将其输出合并在一起，得到一个(W-k+1, H-k+1)大小的输出。

为了达到特征变换的一致性，卷积核的权重不会随卷积核的位置而变化。为了学习到更多的变换，每个卷积层可以有多个卷积核，此时输出大小会变成(W-k+1, H-k+1, n)，n为卷积核数。另外输入也可以是三维的，卷积核仍在前两维中滑动，但卷积核的输入则相应变成三维。

2、池化层

在多个卷积层后，可以增加一个池化层进行降维，以达到防止过拟合的效果。池化层也是以卷积核的形式进行转换，但与卷积层不一样的是：

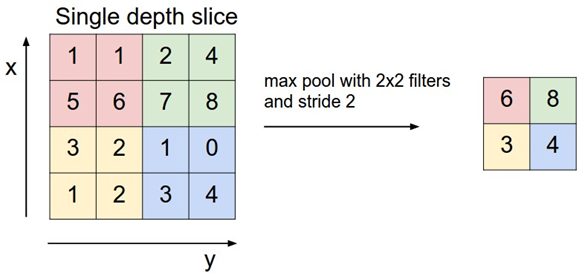


图2-2 池化层

a. 卷积核滑动间距=卷积核大小，滑动的卷积核不重叠；

b. 特征转换形式固定，多为取最大值或平均值；

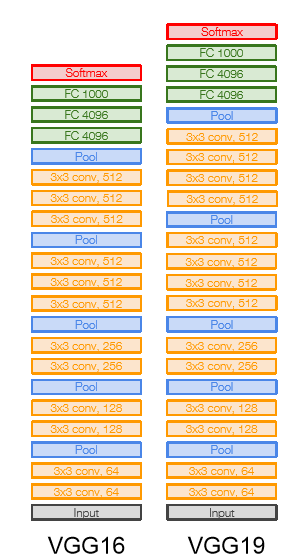
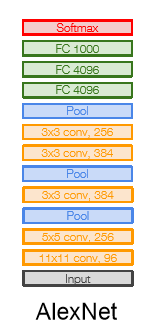
c. 如果输入是三维的话，输出也是三维，第三维个数与输入一致。

3、全连接层

最后CNN需要学习特征与类别的关系，所以需要将三维的特征打散成1维特征，输入上述的全连接网络，最后得到相应的分类结果。

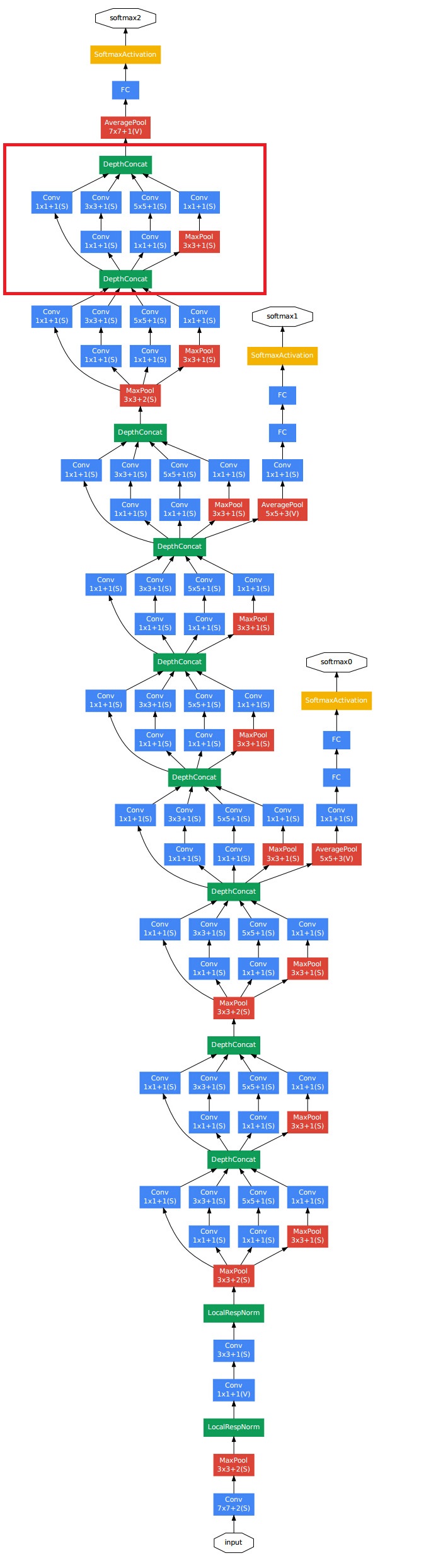
二、CNN相应的结构

AlexNet：



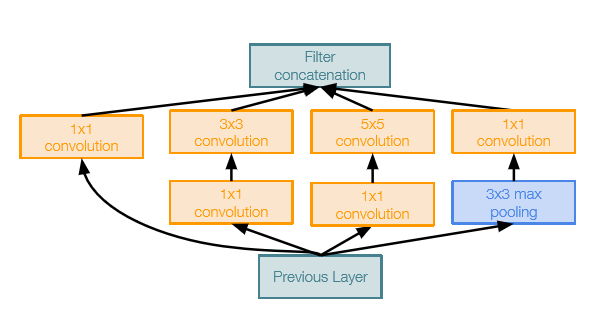
VGG系列是应用最为广泛的深度学习网络，它与AlexNet网络都是由卷积层、池化层和全连接层构成的，不同的是VGG系列深度更深，提取的图像的特征更加全面，因此效果也更优。

GoogLeNet



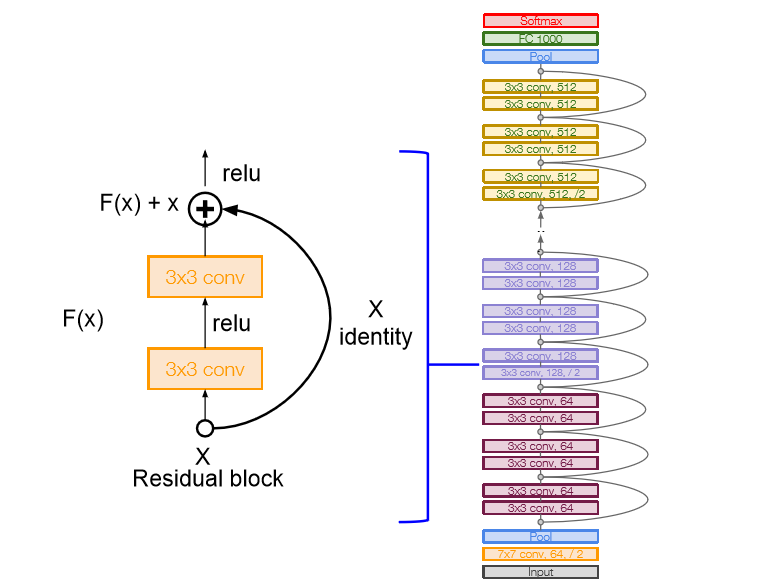
左图是GoogLeNet的框架，也是由卷积层、池化层、全连接层构成的，但是与VGG不同的是，GoogLeNet中有并行的的卷积层，而且左图中红色矩形中的结构是GoogLeNet基本单元，具体构成如下图所示。

GoogLeNet的基本单元



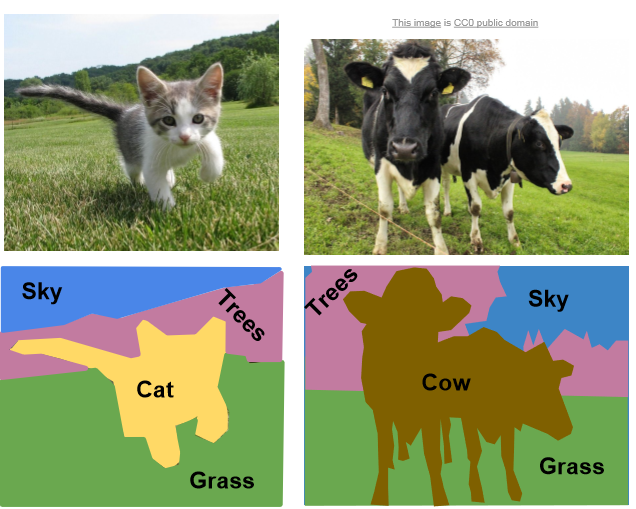
通过更改上述基本单元的结构，谷歌研发人员提出了一系列的网络框架GoogLeNet V1，V2，V3。简单的说就是卷积层和池化层的排列组合。

ResNet



残差网络的出现，完美的解决了深度网络的梯度消失的问题，使得网络的深度可以达到150层以上。从上面的结构图可以看出，网络的基本构成还是卷积层、池化层和全连接层，与之前网络唯一的不同是，层与层之间的连接不再是相邻层链接，而是跨层链接。

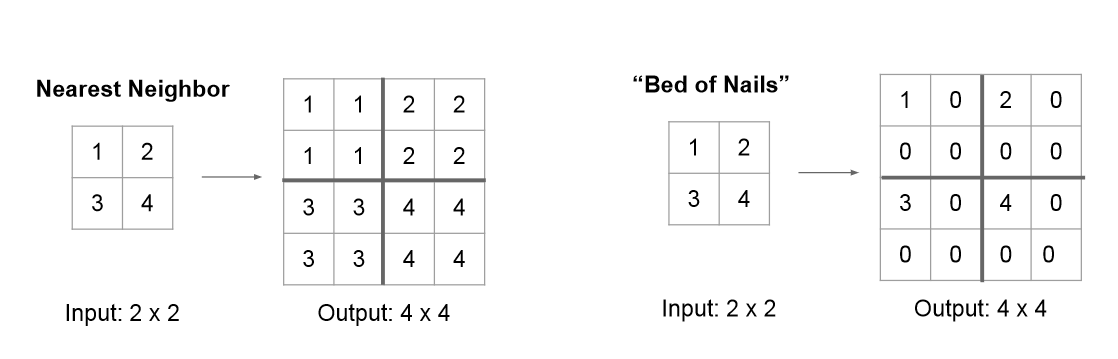
语义分割



输出仍然是图像，基本原理是像素分类。

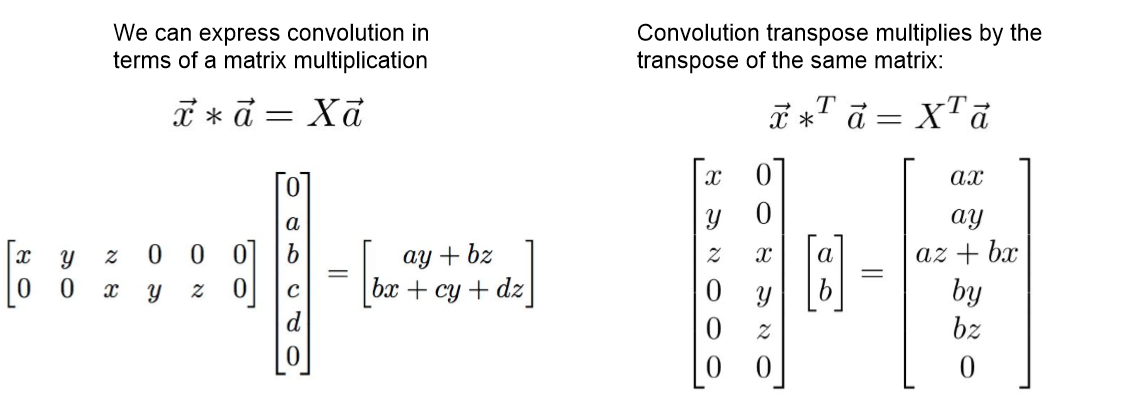
1FCN

上采样



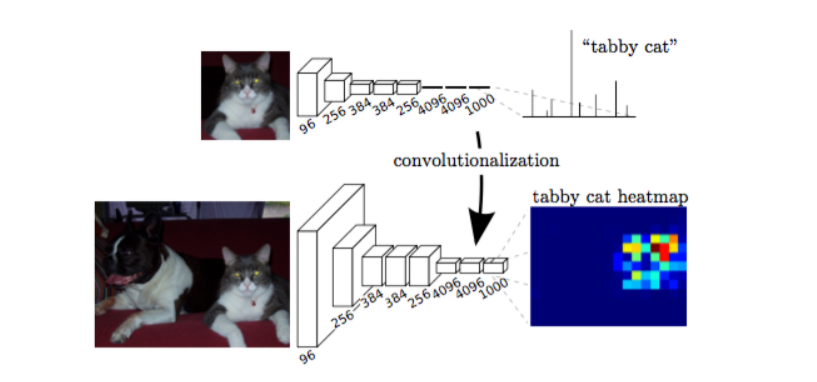
上采样与CNN中的下采样是互逆的，上采样通过填充数据，使得输出的矩阵大于输入的矩阵。

转置卷积



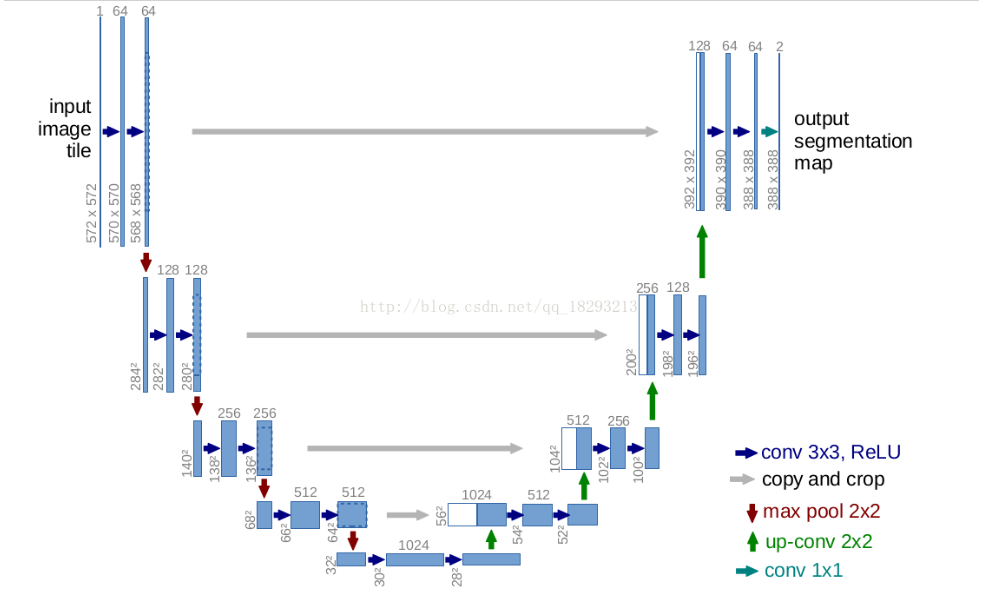
左边是卷积层的数学表达方式，右边是转置卷积的数学表达式，x表示卷积，a为输入矩阵。从上式可以看出，卷积层是压缩输入，而转置卷积层是放大输入。

在FCN网络中因为输出是图像，因此FCN的网络构成是卷积层、下采样层、上采样层以及转置卷积层。具体的FCN的网络构架为：



为了输出图像，之前CNN中的全连接层由转置卷积层和上采样层代替。

FCN的另一种网络结构为U-Net



网络结构如图所示, 蓝色代表卷积和激活函数, 灰色代表复制, 红色代表下采样, 绿色代表上采样然后在卷积, conv 1X1代表核为1X1的卷积操作, 可以看出这个网络没有全连接,只有卷积和下采样. 这也是一个端到端的图像, 即输入是一幅图像, 输出也是一副图像。

上述网络就是计算机视觉中最常用的深度神经网络，基本上可以用来解决计算机视觉中所有的任务，比如图像分类、检测、分割等。

下周任务：

1：机器学习入门PPT

2：搭建深度学习平台，主要是GPU的使用