第五讲 卷积神经网络

2019年4月26日　16:58

1.历史与应用

1.1.历史简介

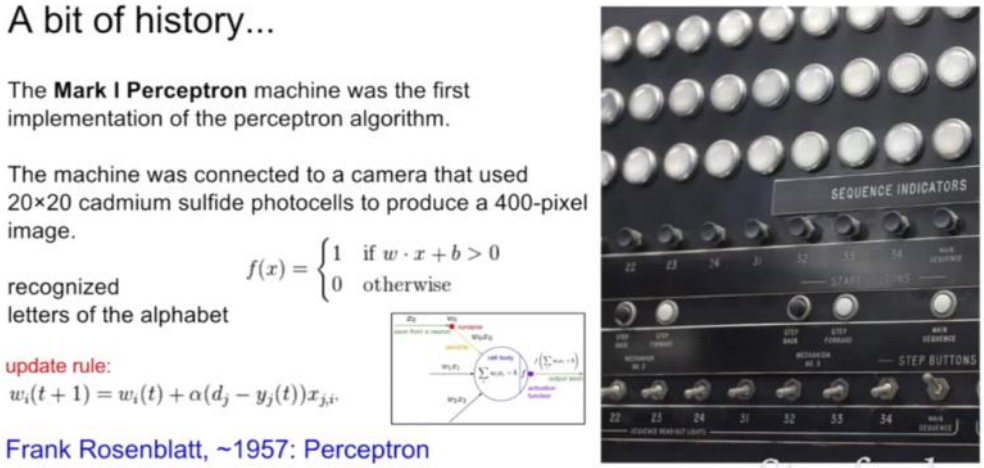


图5.1.1 1957年第一台模拟神经网络的机器

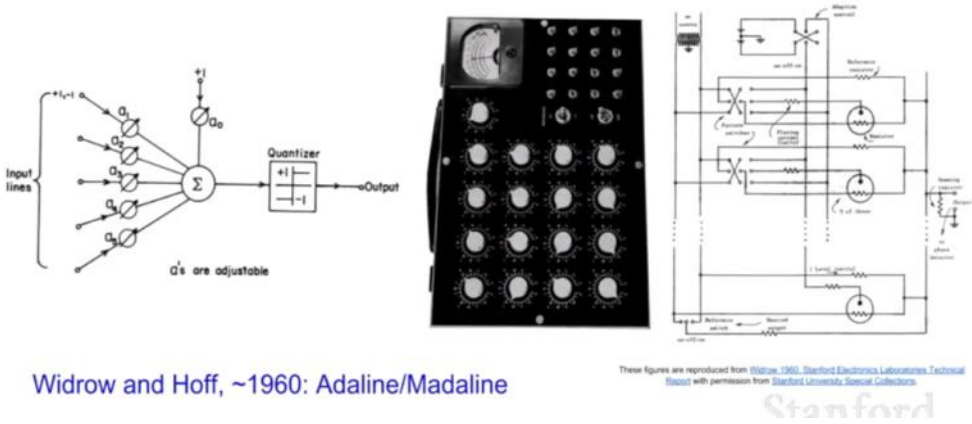


图5.1.2 1960年很接近现代神经网络的设计出现

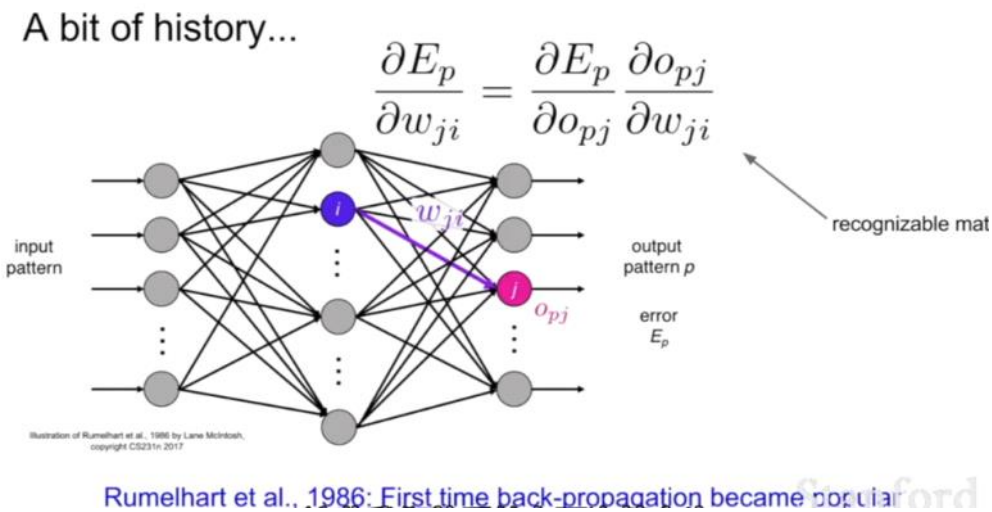


图5.1.3 1986年反向传播算法首次出现

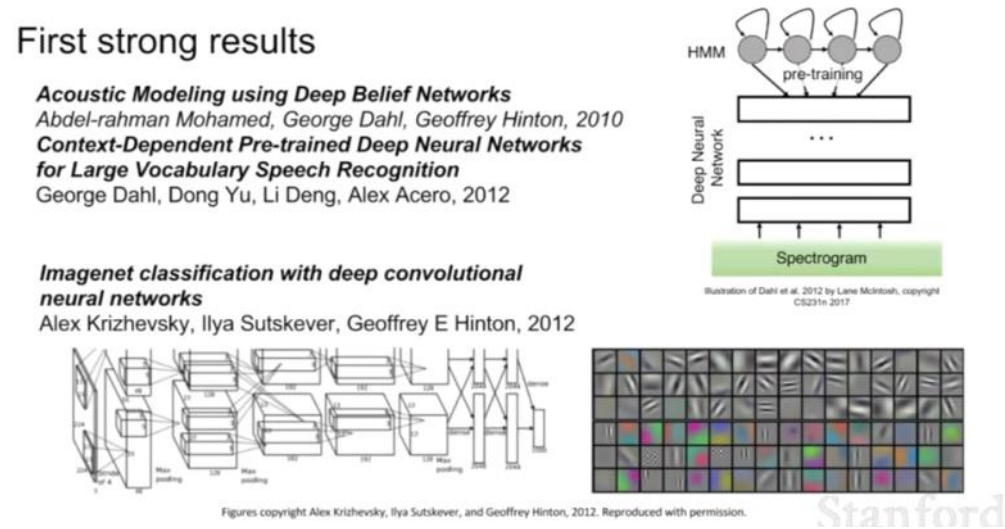


图5.1.4 2012年卷积神经网络在imagenet上大放异彩

cnn五大经典模型

Lenet:

https://github.com/BVLC/caffe/blob/master/examples/mnist/lenet_train_test.prototxt

alexnet:

https://github.com/BVLC/caffe/blob/master/models/bvlc_alexnet/deploy.prototxt

GoogleNet、VGG、Deep Residual Learning

1.2.应用

2.1 cnn表现强势领域

图5.1.5 图像分类与检索领域

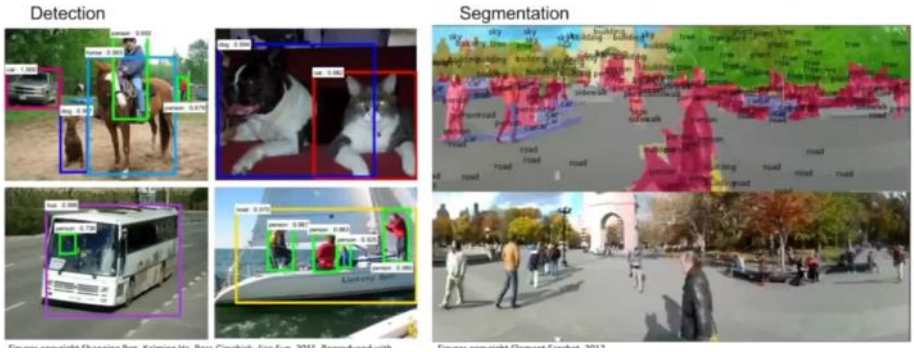


图5.1.6 图像探测与分割

图5.1.7 2d或3d姿势识别

除此之外包括 Image captioning,风格迁移等等

2.卷积和池化

2.1卷积操作

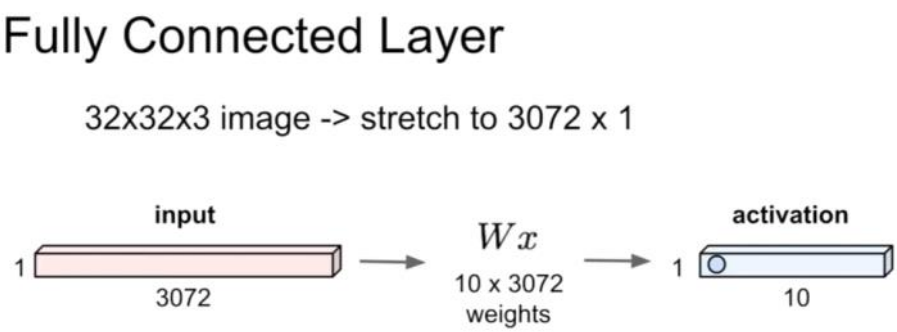


图5.2.1 神经网络全连接层

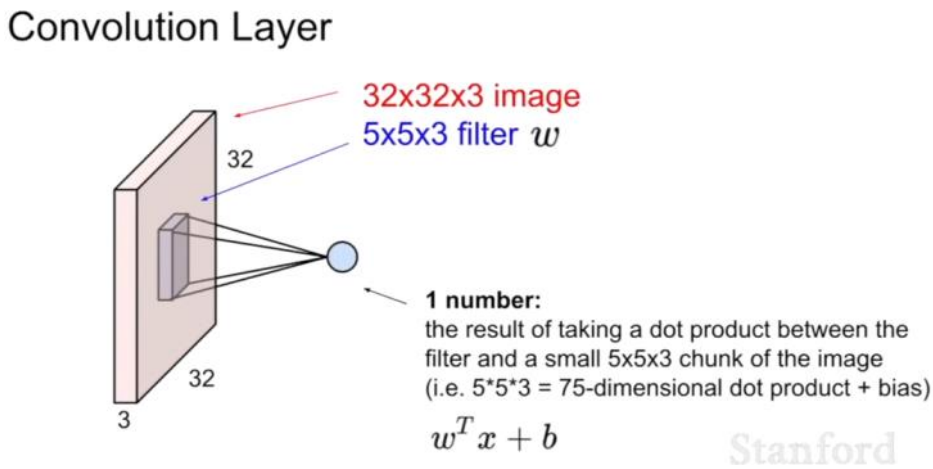


图5.2.2 卷积神经网络

从神经网络到卷积神经网络的变化：之前的神经网络我们将图像的像素以及三通道值组成一个一维的向量，直接输入到神经网络的第一层，经过与参数矩阵W的点积运算到达隐藏层；而卷积神经网络则是使用卷积核遍历图像(卷积操作)之后形成的新的"图像"，作为输入到达下一层。

卷积：图像中不同数据窗口的数据和卷积核（一个滤波矩阵）作内积的操作叫做卷积。其计算过程又称为滤波（filter），本质是提取图像不同频段的特征。

卷积核：也称为滤波器**filter**，带着一组固定权重的神经元，通常是n*m二维的矩阵，n和m也是神经元的**感受野**。n*m 矩阵中存的是对感受野中数据处理的系数。一个卷积核的滤波可以用来提取特定的特征（例如可以提取物体轮廓、颜色深浅等）。通过卷积层从原始数据中提取出新的特征的过程又成为**feature map(特征映射)**。filter_size是指filter的大小，例如3*3； filter_num是指每种filter_size的filter个数，通常是通道个数。

卷积层：多个卷积核叠加

参数：神经元个数 = 卷积核个数

步长：每次移动卷积核的步数

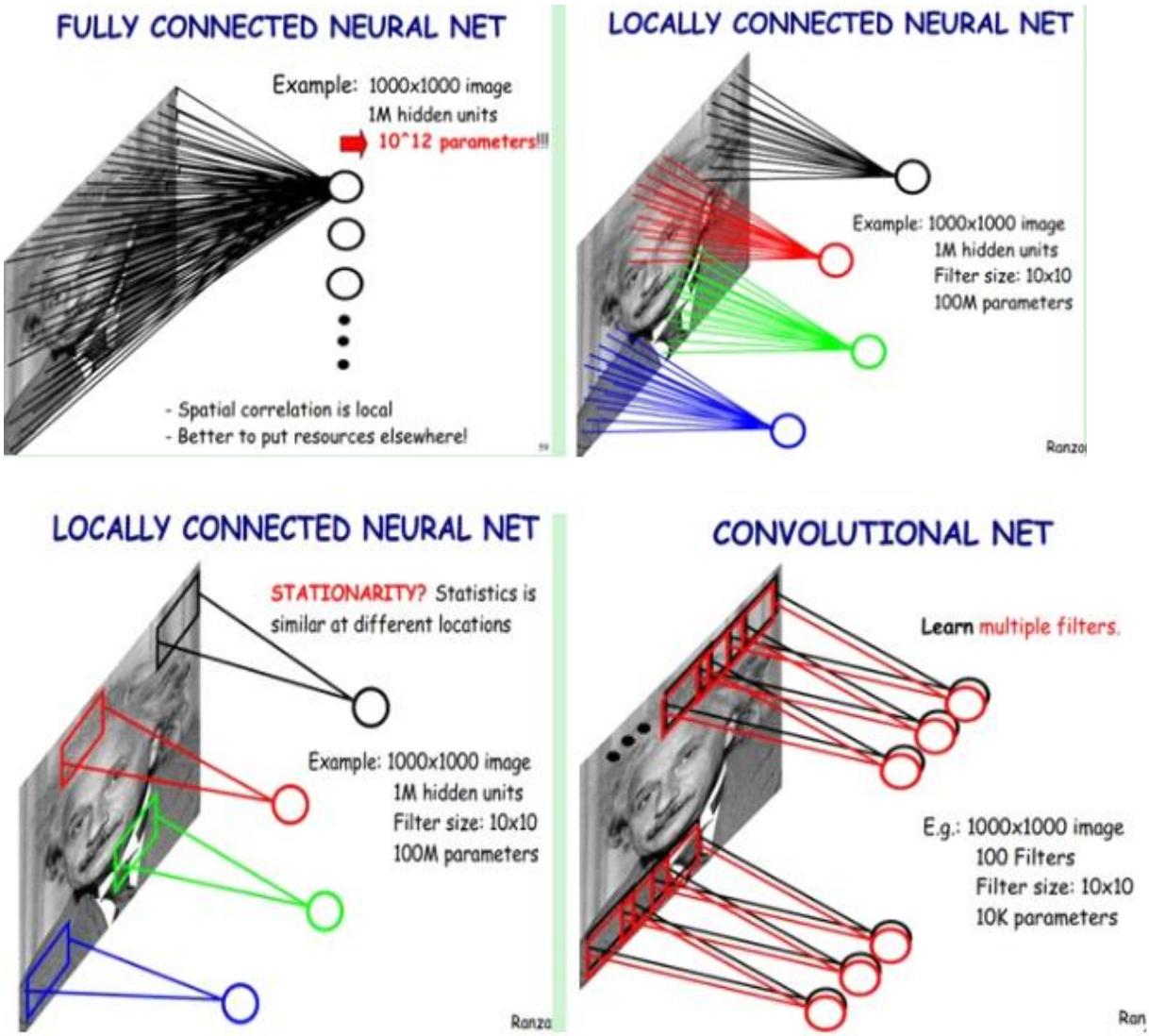


图5.2.3 全连接网络、局部连接网络、卷积网络

全连接与局部连接：在神经网络中，图像的每个像素值对每个神经元都会有影响，但事实上人对图像的识别并不是这样，我们可能看到一个图片的局部特征就能了解图像，而卷积神经网络正是将神经网络每个神经元的全连接状态给断掉，每个神经元只负责维护**某种局部**的图像信息。

一种卷积核提取了一种图像特征，n种卷积核就能提取图像的n种图像特征，一层卷积神经网络就是这样由n种卷积核组成的。

卷积网络的核心思想是将：局部感受野、权值共享以及时间或空间亚采样这三种结构思想结合起来获得了某种程度的位移、尺度、形变不变性。

Input Volume (pad 1) (7x7x3) Filter W0 (3x3x3) Filter W1 (3x3x3) Output Volume (3x3x2)

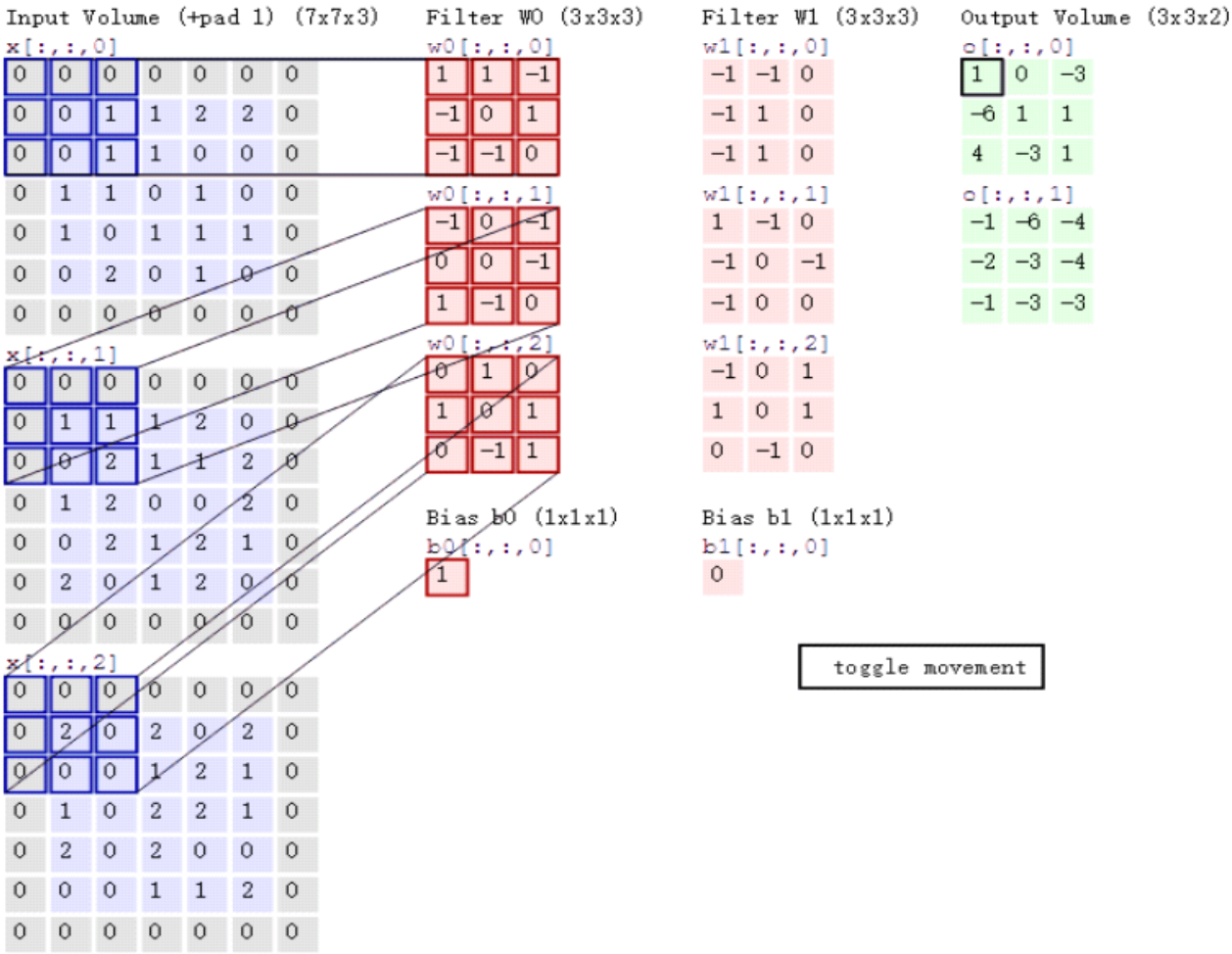


图5.2.4 卷积操作计算过程

0填充(pad1)的7x7x3矩阵，这里有两个卷积核，w0、w1都是3x3x3的卷积核，步长为2.

Output1 = **w0 x input** + bias

w0 x input = 分为三个通道分别计算点积最后相加。

第一个通道和对应权重的结果：0*1+0*1+0*(-1)+0*(-1)+0*0+1*1+0*(-1)+0*(-1)+1*0 = 1
第二个通道和对应权重的结果：0*(-1)+0*0+0*(-1)+0*0+1*0+1*(-1)+0*1+0*(-1)+2*0 = -1
第三个通道和对应权重的结果：0*0+0*1+0*0+0*1+2*0+0*1+0*0+0*(-1)+0*0 = 0
偏执为1
1+（-1）+ 0 + 1 = 1 -> 第一个卷积核输出的第一个值

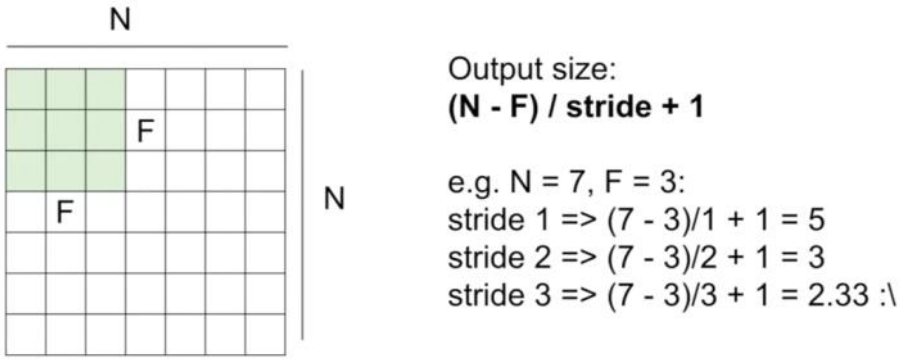


图5.2.5 卷积核、步长和输出尺寸计算公式

图像经过卷积之后输出结果"尺寸"计算公式：(N-F)/stride + 1

N:原图像尺寸N*N

F:卷积核尺寸F*F

Stride:步长

2.2 池化

这节课没讲什么是池化！！

3.视觉之外的卷积神经网络

3.1 池化

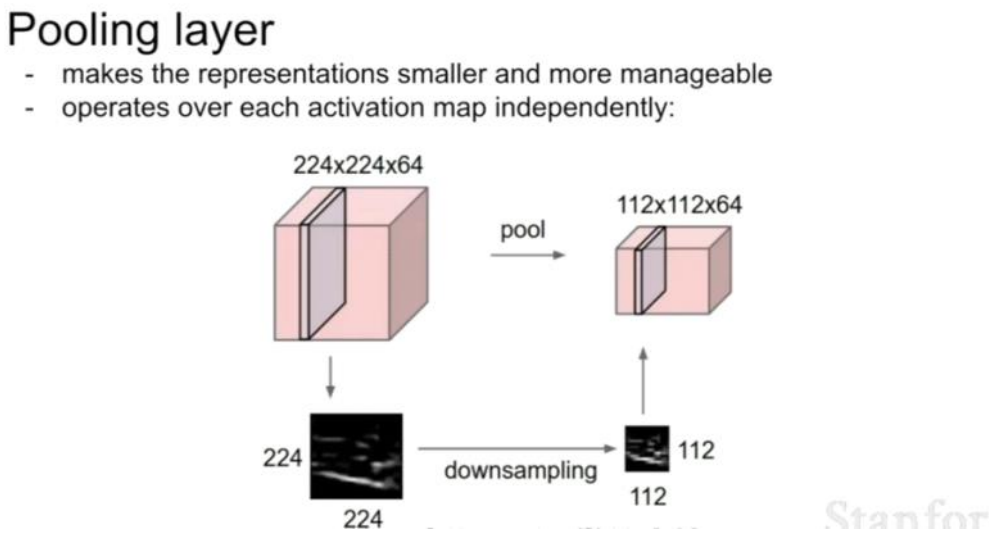
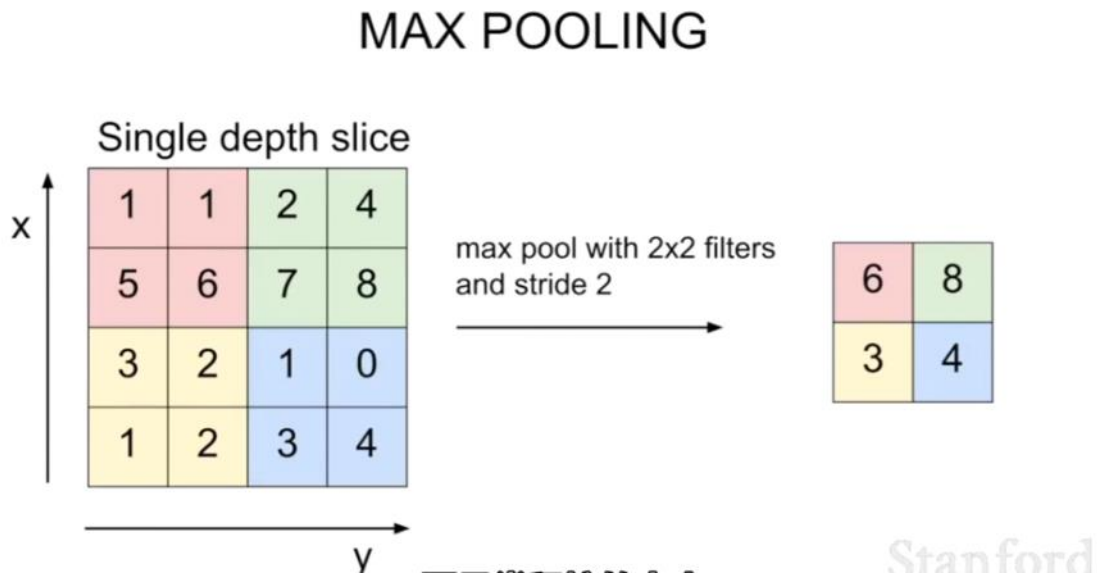


图5.3.1 池化层

池化：对输入的特征图进行压缩，一方面使特征图变小，简化网络计算复杂度；一方面进行特征压缩，提取主要特征。

Max pooling

压缩时取块中的最大值



Avy pooling

与max pooling 类似，不过avy取filter的平均值。