

# 图像信号基本处理

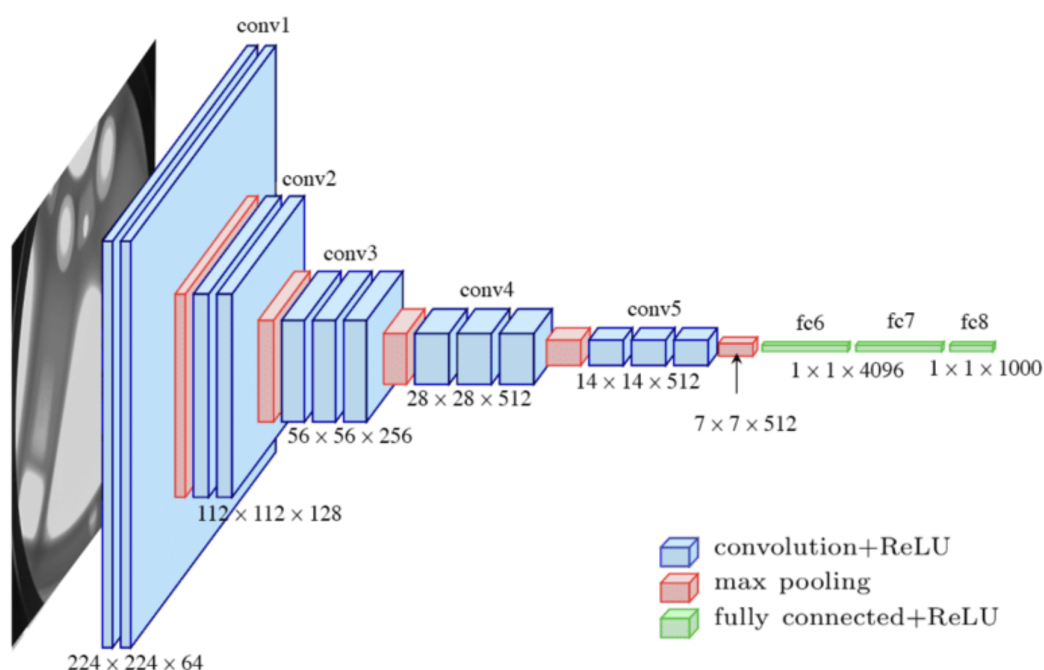
在图像处理的过程中，经常会经历一些操作需要对图像的尺寸等性质进行处理，最常见的就是放大 (zoom in)和缩小 (zoom out)；在这些情况下，需要用到诸如上采样与下采样等操作。

- 缩小图像（或称为下采样或降采样）的主要目的有两个：1、使得图像符合显示区域的大小；2、生成对应图像的缩略图。
- 放大图像（或称为上采样或图像插值）的主要目的是放大原图像,从而可以显示在更高分辨率的显示设备上。

对于一幅图像尺寸为  $M * N$ ，对其进行  $s$  倍下采样，即得到  $\frac{M}{s} * \frac{N}{s}$  尺寸的分辨率图像，当然此时  $s$  应该是  $M$  和  $N$  的公约数，如果考虑的是矩阵形式的图像，就是把原始图像  $s*s$  窗口内的图像变成一个像素，这个像素点的值就是窗口内所有像素的均值。

而图像放大几乎都是采用内插值方法，即在原有图像像素的基础上在像素点之间采用合适的插值算法插入新的元素。

而所谓的池化操作，就是一种非常典型的下采样操作；



同样，在做语义分割或者目标检测的时候，我们用到了相当多的上采样，或者转置卷积，这一系列操作包括卷积核的选择等问题都会直接影响着整个模型的各项性能。

在这一系列过程中，会涉及到各个种类卷积核的相应作用，对大盒小核的选取，也将会进一步地影响包括感受野在内的一系列性质，最终对该系统产生深层影响。

一般而言，对该信号进行上采样有多种方式进行，例如**Nearest Neighbors**，**Bi-Linear Interpolation**等操作，其基本内容如下：

在Nearest Neighbors中，直接将离被插入点最近像素的值作为被插入的值：

### Nearest Neighbor

1	2
3	4

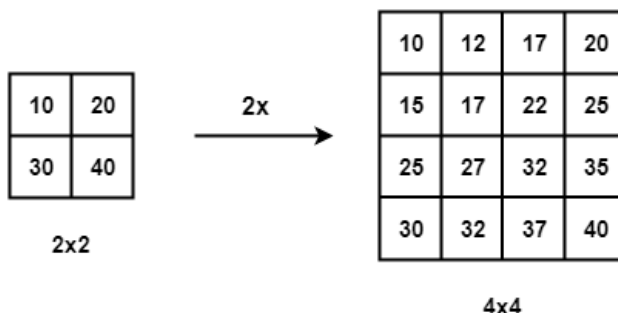


1	1	2	2
1	1	2	2
3	3	4	4
3	3	4	4

Input: 2 x 2

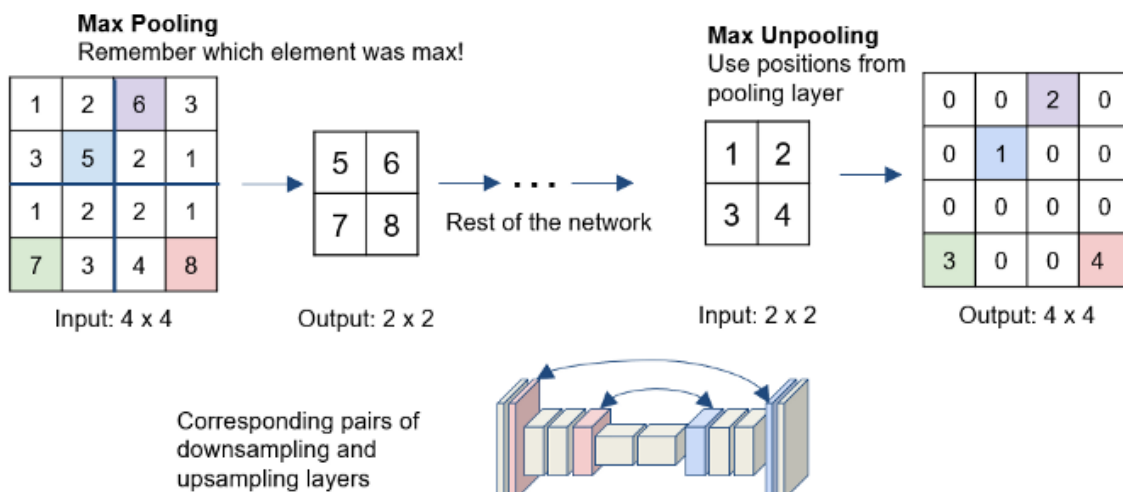
Output: 4 x 4

而双线性插值法则使用最近的多个元素进行回归操作完成赋值：

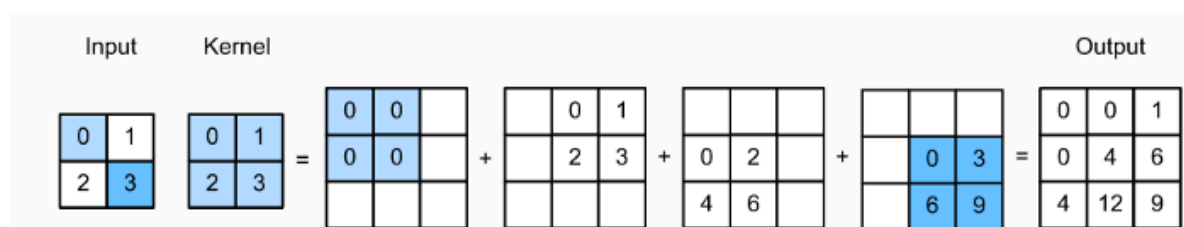


而上池化（Max-Unpooling）操作则完成了对池化操作的逆运算，它利用记录的，存下的位置值完成了操作：

The Max-Pooling layer in CNN takes the maximum among all the values in the kernel. To perform max-unpooling, first, the index of the maximum value is saved for every max-pooling layer during the encoding step. The saved index is then used during the Decoding step where the input pixel is mapped to the saved index, filling zeros everywhere else.



而转置卷积操作，又被称为反卷积操作；实际上，它并不是所谓卷积操作的逆，而是一种与卷积操作运算方法类似的上采样操作，它将会使用输入的矩阵与卷积核进行运算，在预设的stride与padding操作下，可以完成上采样的基本操作：



可以看出，其主要操作即对输入的矩阵分别在对应的位置卷积生成新矩阵后，利用overlapping的性质对矩阵进行叠加操作，得到最终结果。

**此外，如同卷积操作可以使用卷积矩阵进行一样，转置卷积也可以使用其对应的矩阵操作完成计算。**