引言：

已知函数 p(t)如下图所示，当 t 在$(-\frac{\Delta}{2},\frac{\Delta}{2})$内时取值$\frac{1}{\Delta}$，其余区间取值为 0。现有函数 f(t)如下图所示，当 t 在$(-\frac{\Delta}{2},\frac{\Delta}{2})$内时取值$\boldsymbol{A}$，其余区间取值为 0。

问：f(t)是 p(t)的多少倍？

答：f(t)是 p(t)的 $\frac{A}{\frac{1}{\Delta}}$ 倍。即 f(t) = $\frac{A}{\frac{1}{\Delta}}$ p(t)。

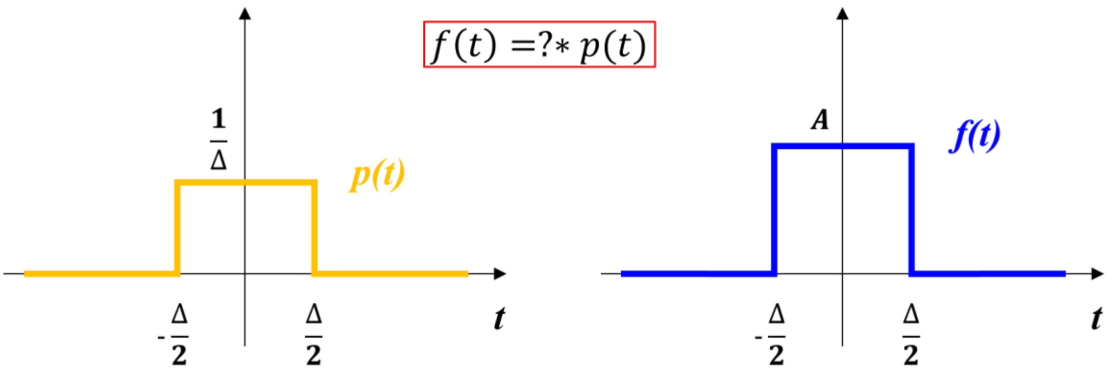

FIGURE. 示意图 1-信号缩放原理

卷积积分运算：

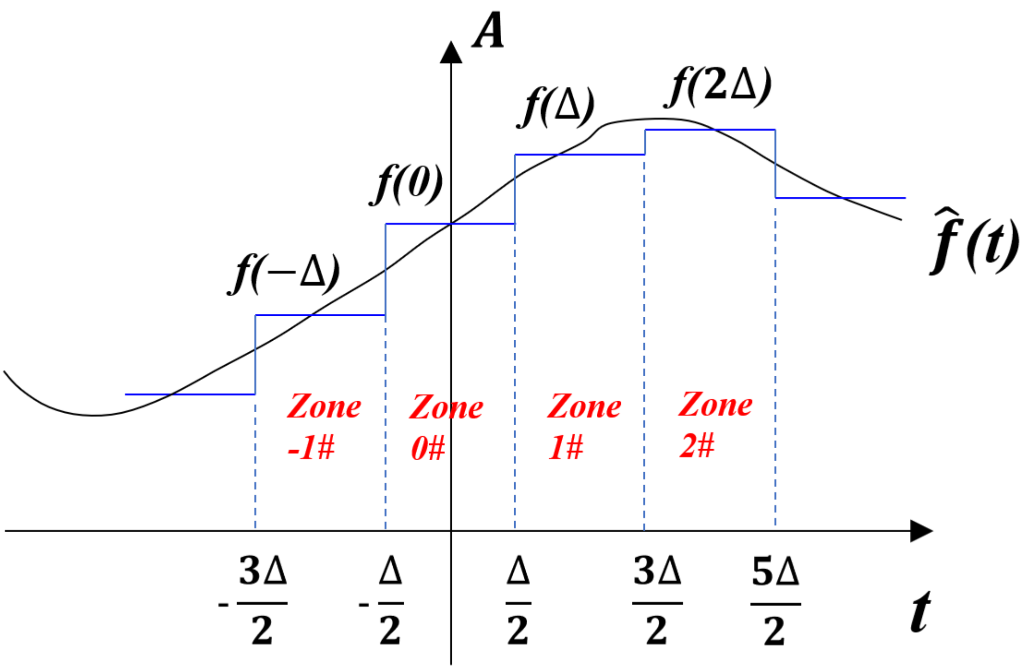

FIGURE. 示意图 2-信号分解原理（缩放和平移）

1.如示意图 2 所示，对于任意函数 $f(t)$，可以使用 $\hat{f}(t)$ 对它进行逼近，可以通过类似上图

中取值为 f(t)的门函数的叠加来获得。比如下图中黑色曲线是函数 $\hat{f}(t)$，是由很多蓝色门函数的叠加获得的。门函数的宽度是 delta，0 号区域上的门函数取值为 f(0)，1 号区域上的门函数取值为 f(delta)，2 号区域上的门函数取值为 f(2*delta)，-1 号区域上的门函数取值为 f(-delta)。
使用引言中的方法，可以得到：

f(0) = f(0)/(1/delta) * p(t) = delta * f(0) * p(t);
f(delta) = f(delta) / (1/delta) * p(t-delta) = delta * f(delta) * p(t-delta);
f(2*delta) = f(2*delta) / (1/delta) * p(t-2*delta) = delta * f(2*delta) * p(t-2*delta);
f(-delta) = f(-delta) / (1/delta) * p(t + delta) = delta * f(-delta) * p(t + delta);
...

所以，$\boldsymbol{f(t)} \approx \hat{\boldsymbol{f}}(\boldsymbol{t}) = \sum_{\boldsymbol{n}=-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{n}\Delta)\Delta \boldsymbol{p}(\boldsymbol{t}-\boldsymbol{n}\Delta)$

2.当 delta 足够小的时候，上述等式取等号。

$$\boldsymbol{f(t)} = \hat{\boldsymbol{f}}(\boldsymbol{t}) = \lim_{\Delta\to 0}\sum_{\boldsymbol{n}=-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{n}\Delta)\Delta \boldsymbol{p}(\boldsymbol{t}-\boldsymbol{n}\Delta)$$

3.当取极限时，可以改写上述离散求和表达式：

$\Delta \to \mathbf{0}$：

$\Delta = \mathbf{d\tau}$

$\mathrm{n}\Delta = \boldsymbol{\tau}$

$\sum_{\boldsymbol{n}=-\infty}^{+\infty} \quad = \int_{-\infty}^{+\infty}$

故： $\boldsymbol{f(t)} = \hat{\boldsymbol{f}}(\boldsymbol{t}) = \lim_{\Delta\to 0}\sum_{\boldsymbol{n}=-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{n}\Delta)\Delta \boldsymbol{p}(\boldsymbol{t}-\boldsymbol{n}\Delta)$

$$\boldsymbol{f(t)} = \int_{-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{p}(\boldsymbol{t}-\boldsymbol{\tau})\, \boldsymbol{d\tau}$$

（sigma 是离散和，积分是连续和。当取极限时，小矩形的两端聚到一个点上，离散变为连续。）

4.当 $\Delta \to \mathbf{0}$ 时，p(t)变成宽度无穷小，高度无穷大的门函数，就是 δ(t)。所以上式可以写作：

$\boldsymbol{f(t)} = \int_{-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{\delta}(\boldsymbol{t}-\boldsymbol{\tau})\, \boldsymbol{d\tau}$

5.引入这种新的运算，并命名为卷积积分运算，记作：

$$\boldsymbol{f(t)} = \int_{-\infty}^{+\infty} \boldsymbol{f}(\boldsymbol{\tau})\boldsymbol{\delta}(\boldsymbol{t}-\boldsymbol{\tau})\, \boldsymbol{d\tau} = \boldsymbol{f(t)} * \boldsymbol{\delta(t)}$$

讨论：

上述过程中，卷积的定义使用了一个前提：p(t)是仅在 x=0 左右两侧非常小的区间内取值的函数。**在卷积运算中，p(t)并非直接作用在 t 的所有范围上，而是通过 p(t)的缩放和平移作用在 t 的变化范围中除了 p(t)的门函数取值区间外的其他位置。**当对 Δ 取极限时，虽然$\boldsymbol{p(t)} \to \infty$，但$\boldsymbol{p(t)}\Delta = \mathbf{1}$。也就是说，从 t 的空间中看，当 t 趋于 0 时，p(t)是无穷，但在 p(t)的空间中，p(t)其实是$\boldsymbol{p(t)}\Delta$产生的 1。当对 Δ 取极限时，在 t 的变化范围中除了 p(t)的门函数取值区间外的其他位置，将变为 t 的变化范围中的所有位置。

参考正交性原理。当取极限时，p(t)其实是在与 x 空间垂直的空间上的一个以 p(t)方式变化的函数。当参加卷积运算的函数不是 p(t)时，其实是在 p(t)所在的维度对 p(t)进行缩放。

需要研究的对象有两个维度，卷积积分运算是在一个维度上，用另一个维度的函数进行的叠加运算。示意图 3 是信号的时域波形和频域波形拆解，对照示意图 2，考虑卷积$\boldsymbol{f(t) * g(t)}$。f(t)给定了在 t 的不同取值时，取多少个另外一个维度上的特征，用于平移到在当前 t 值下进行叠加。g(t)给定了在这个 t 值下，另外一个维度上的特征是该空间上的单位值 p(t)的多少倍。

卷积积分是站在当地 t 视角得到的，因为在一个确定的 t 值下，另一个维度上的特征是确定值。傅里叶级数是站在全局 t 视角得到的，整个周期内，由多少个另外一个维度上的特征的叠加产生了在 t 维度上出现的变化。

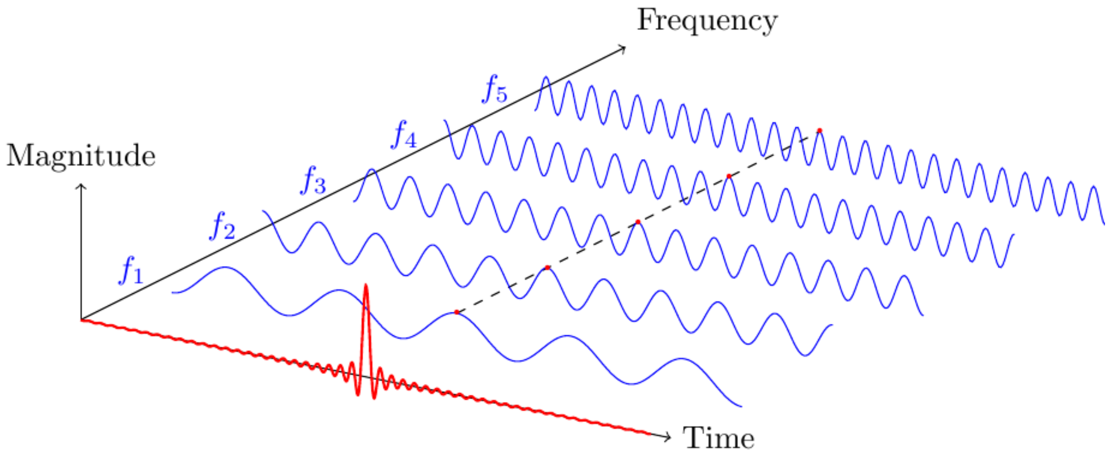

FIGURE. 示意图 3-傅里叶级数将时域信息展开为频域信息

卷积积分运算的定义：

根据以上讨论，参与卷积运算的两个函数可以是任意函数。卷积积分运算的定义为：
已知定义在区间(−∞,+∞)上的两个函数$\boldsymbol{f_1(t)}$和$\boldsymbol{f_2(t)}$，则定义积分：

$$\boldsymbol{f(t) = \int_{-\infty}^{+\infty} f_1(\tau) f_1(t-\tau)\, d\tau}$$

为$\boldsymbol{f_1(t)}$和$\boldsymbol{f_2(t)}$的卷积积分运算，简称卷积。记作：

$$\boldsymbol{f(t) = f_1(t) * f_2(t)}$$

$$\text{或 } \boldsymbol{f(t) = f_1(t) \otimes f_2(t)}$$

参考资料：《信号与系统》 https://www.icourse163.org/course/XIDIAN-483006

卷积神经网络的理解：

可以将图像中的不同图案理解为不同的特征，比如圆形状和方形状对应不同的特征。

图像的 RGB 矩阵描述了图像的信息。

1.通过卷积运算，在 RGB 数字维度上，取另一个维度上的特征进行缩放和平移，得到一个以 RGB 维度上信息为参变量的特征表示。可以在 RGB 维度上，将这种特征表示看作是一条函数曲线。但是这条曲线是在描述和上述两个维度都正交的另一个维度中的信息。就像信号中的时域和频域的卷积，得到以时域为参变量的正交于时域和频域的信息一样。

2.将卷积层的矩阵当做另一个维度上的特征，该矩阵依次按照 RGB 层的矩阵上对应位置处的信息描述在 RGB 层进行缩放和平移，当完成 RGB 层上所有位置处描述的缩放和平移操作后，卷积完成，得到的新矩阵是正交于卷积层矩阵和 RGB 矩阵的、以 RGB 层信息为参变量的、另一个维度上的信息表示。

3.可以将上一步中得到的新维度的信息当做类似时域频域中的时域，重复上一步中的卷积操作，又得到正交于第一个卷积层矩阵、第二个卷积层矩阵和 RGB 矩阵的、以 RGB 层信息为参变量的、另一个维度上的信息表示。

4.不断产生新的维度，将会在高维空间中，以高维空间中的一条向量的形式表征 RGB 层中包含的信息。因此，直到我们在高维空间中得到了这样的一条向量，就可以使用这样的一条向量进行分类，即连接神经网络。

以上是卷积神经网络的整体流程。

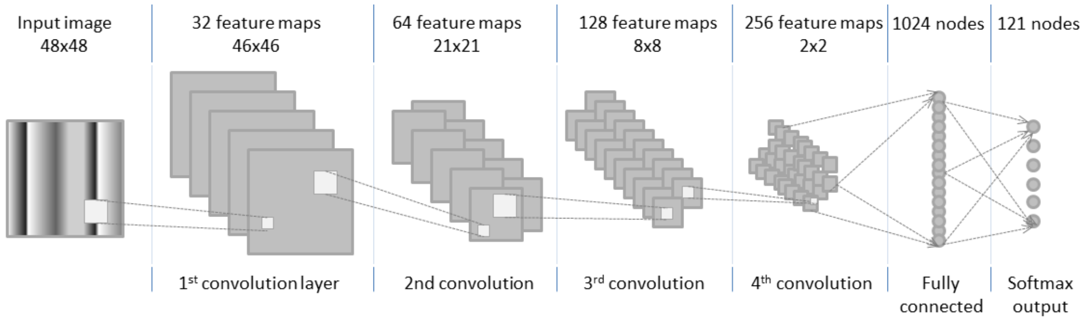

FIGURE. 示意图 4-卷积神经网络的信息处理流程

以下是对整体流程中的一个局部处理方法的讨论。

■ 卷积神经网络中的池化：

1.每层卷积之后得到了新维度矩阵，可以把新维度矩阵划成若干个大块，每个大块中包含了同样的小块。

2.可以认为每个大块中包含了许多信息，就像是高维空间中的，顶点位于空间原点的一个锥体内有长短不一、空间角度略有差异、但同在同一个锥体内的空间向量。
使用池化的方式突出新维度中的突出信息，就像是在锥体中选出最长的那条向量或者最贴近锥体轴指向的那条向量。

3.可以认为图像中的某个特征对应于最长的那条向量，也可以认为特征对应于最贴近轴向的那条向量，这需要视情况而定。当以前者为准则时，对应着最大池化方法。当以后者为准则时，对应着平均池化方法。

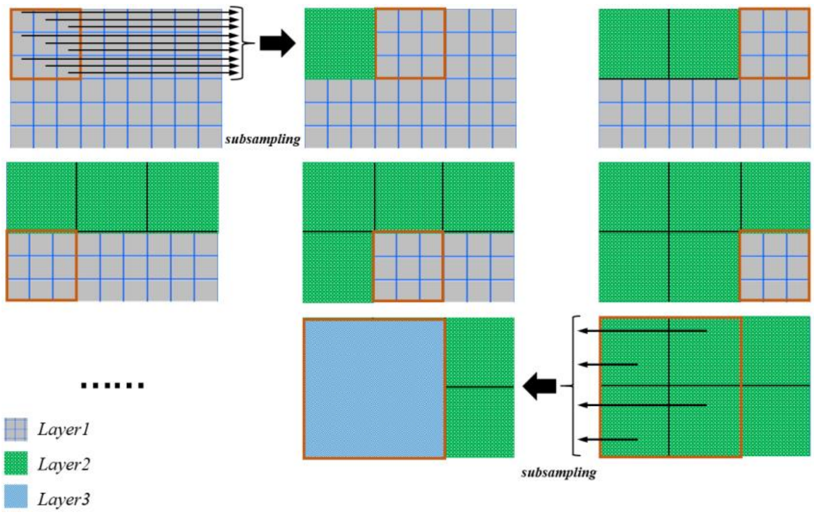

FIGURE. 示意图 5-池化运算的定义

但无论根据什么样的原则选出向量，池化的目的都是根据你指定的原则，过滤出最符合你所指定的原则的信息。**即池化会过滤出每个大块中最强的信息。**池化前后，信息在同一维度。

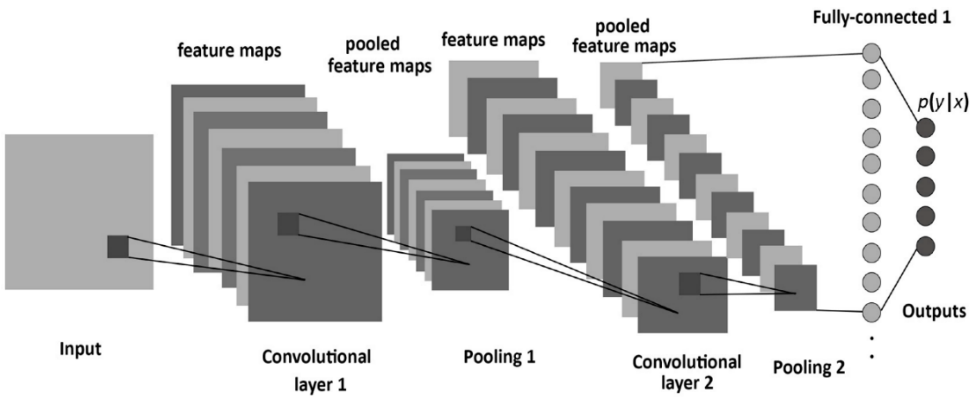

FIGURE. 示意图 6-卷积运算和池化运算的连接

注意：使用卷积神经网络时，**不要让注意力被图像上的图案所分散**。将所有类型的图案**一律视为数字矩阵**，在 RGB 格式下只是多了几层同样维度的矩阵而已。