经典卷积神经网络的结构一般满足如下表达式：

输出层 -> （卷积层+ -> 池化层？）+  -> 全连接层+

上述公式中，“+”表示一个或者多个，“？”表示一个或者零个，如“卷积层+”表示一个或者多个卷积层，“池化层？”表示一个或者零个池化层。“->”表示 forward 方向。

　　下面将分别介绍 [LeNet-5](https://www.cnblogs.com/wuliytTaotao/p/9544625.html#LeNet-5)、[AlexNet](https://www.cnblogs.com/wuliytTaotao/p/9544625.html#AlexNet) 和 [VGG-16](https://www.cnblogs.com/wuliytTaotao/p/9544625.html#VGG-16) 结构。

1. LeNet-5（modern）

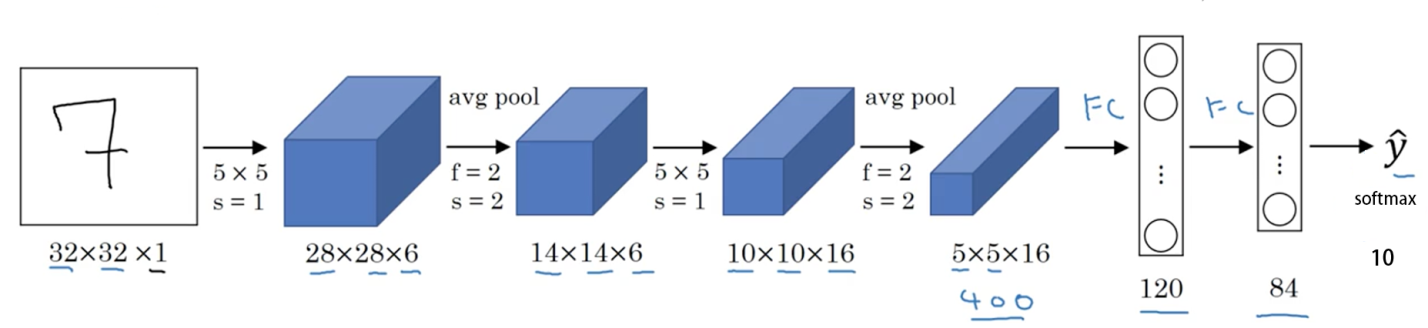


图 1  LeNet-5

1.1 LeNet-5 结构：

* 输入层

图片大小为 32×32×1，其中 1 表示为黑白图像，只有一个 channel。

* 卷积层

filter 大小 5×5，filter 深度（个数）为 6，padding 为 0， 卷积步长 s=1s=1，输出矩阵大小为 28×28×6，其中 6 表示 filter 的个数。

* 池化层

average pooling，filter 大小 2×2（即 f=2f=2），步长 s=2s=2，no padding，输出矩阵大小为 14×14×6。

* 卷积层

filter 大小 5×5，filter 个数为 16，padding 为 0， 卷积步长 s=1s=1，输出矩阵大小为 10×10×16，其中 16 表示 filter 的个数。

* 池化层

average pooling，filter 大小 2×2（即 f=2f=2），步长 s=2s=2，no padding，输出矩阵大小为 5×5×16。注意，在该层结束，需要将 5×5×16 的矩阵flatten 成一个 400 维的向量。

* 全连接层（Fully Connected layer，FC）

neuron 数量为 120。

* 全连接层（Fully Connected layer，FC）

neuron 数量为 84。

* 全连接层，输出层

现在版本的 LeNet-5 输出层一般会采用 softmax 激活函数，在 LeNet-5 提出的论文中使用的激活函数不是 softmax，但其现在不常用。该层神经元数量为 10，代表 0～9 十个数字类别。（图 1 其实少画了一个表示全连接层的方框，而直接用 y^y^ 表示输出层。）

1.2 LeNet-5 一些性质：

* 如果输入层不算神经网络的层数，那么 LeNet-5 是一个 7 层的网络。（有些地方也可能把 卷积和池化 当作一个 layer）（LeNet-5 名字中的“5”也可以理解为整个网络中含可训练参数的层数为 5。）
* LeNet-5 大约有 60,000 个参数。
* 随着网络越来越深，图像的高度和宽度在缩小，与此同时，图像的 channel 数量一直在增加。
* 现在常用的 LeNet-5 结构和 [Yann LeCun 教授在 1998 年论文](http://yann.lecun.com/exdb/publis/pdf/lecun-01a.pdf)中提出的结构在某些地方有区别，比如激活函数的使用，现在一般使用 ReLU 作为激活函数，输出层一般选择 softmax。

2. AlexNet

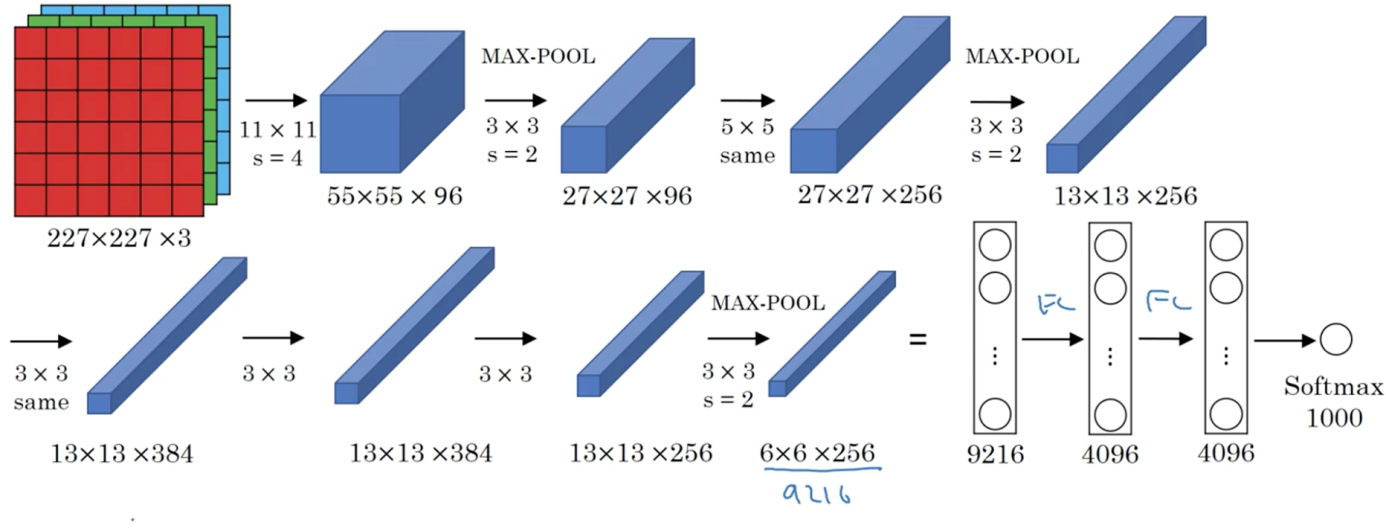


图 2  AlexNet

2.1 AlexNet 结构：

* 输入层：图像大小为 227×227×3，其中 3 表示输入图像的 channel 数（R，G，B）为 3。
* 卷积层：filter 大小 11×11，filter 个数 96，卷积步长 s=4s=4。（filter 大小只列出了宽和高，filter矩阵的 channel 数和输入图片的 channel 数一样，在这里没有列出）
* 池化层：max pooling，filter 大小 3×3，步长 s=2s=2。
* 卷积层：filter 大小 5×5，filter 个数 256，步长 s=1s=1，padding 使用 same convolution，即使得卷积层输出图像和输入图像在宽和高上保持不变。
* 池化层：max pooling，filter 大小 3×3，步长 s=2s=2。
* 卷积层：filter 大小 3×3，filter 个数 384，步长 s=1s=1，padding 使用 same convolution。
* 卷积层：filter 大小 3×3，filter 个数 384，步长 s=1s=1，padding 使用 same convolution。
* 卷积层：filter 大小 3×3，filter 个数 256，步长 s=1s=1，padding 使用 same convolution。
* 池化层：max pooling，filter 大小 3×3，步长 s=2s=2；池化操作结束后，将大小为  6×6×256 的输出矩阵 flatten 成一个 9216 维的向量。
* 全连接层：neuron 数量为 4096。
* 全连接层：neuron 数量为 4096。
* 全连接层，输出层：softmax 激活函数，neuron 数量为 1000，代表 1000 个类别。

2.2 AlexNet 一些性质：

* 大约 60million 个参数；
* 使用 ReLU 作为激活函数。

3. VGG-16

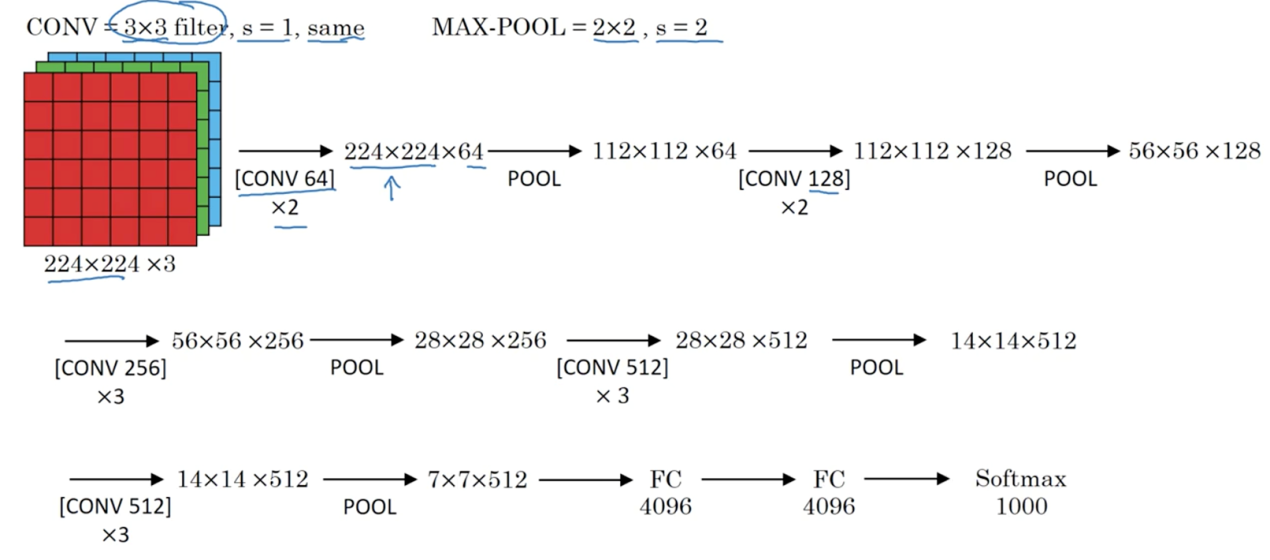


图 3  VGG-16

3.1 VGG-16 结构：

* 输入层
* 卷积层
* 卷积层
* 池化层
* 卷积层
* 卷积层
* 池化层
* 卷积层
* 卷积层
* 卷积层
* 池化层
* 卷积层
* 卷积层
* 卷积层
* 池化层
* 卷积层
* 卷积层
* 卷积层
* 池化层
* 全连接层
* 全连接层
* 全连接层，输出层

3.2 VGG-16 一些性质：

* VGG-16 中的 16 表示整个网络中有 trainable 参数的层数为 16 层。（trainable 参数指的是可以通过 back-propagation 更新的参数）
* VGG-16 大约有 138million 个参数。
* VGG-16 中所有卷积层 filter 宽和高都是 3，步长为 1，padding 都使用 same convolution；所有池化层的 filter 宽和高都是 2，步长都是 2。