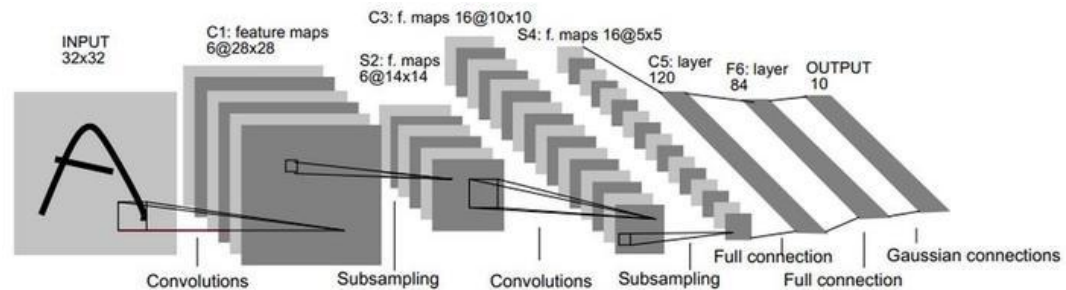


Lenet-5

LeNet5 这个网络虽然很小，但是它包含了深度学习的基本模块：卷积层，池化层，全连接层。是其他深度学习模型的基础，这里对 LeNet5 进行深入分析。



LeNet-5 共有 7 层，不包含输入，每层都包含可训练参数；每个层有多个 Feature Map，每个 FeatureMap 通过一种卷积滤波器提取输入的一种特征，然后每个 FeatureMap 有多个神经元。

各层参数详解

1、输入层

首先是数据 INPUT 层，输入图像的尺寸统一归一化为 32*32。

2、卷积层 C1

输入图片：32*32

卷积核大小：5*5

卷积核种类：6

输出 featuremap 大小：28*28 $(32-5+1)=28$

神经元数量：28*28*6

可训练参数： $(5*5+1) * 6$ （每个滤波器 5*5=25 个 unit 参数和一个 bias 参数，一共 6 个滤波器）

连接数： $(5*5+1) * 6 * 28 * 28 = 122304$

详细说明：对输入图像进行第一次卷积运算（使用 6 个大小为 5*5 的卷积核），得到 6 个 C1 特征图（6 个大小为 28*28 的 feature maps, $32-5+1=28$ ）。我们再来看看需要多少个参数，卷积核的大小为 5*5，总共就有 $6 * (5*5+1)=156$ 个参数，其中+1 是表示一个核有一个 bias。对于卷积层 C1，C1 内的每个像素都与输入图像中的 5*5 个像素和 1 个 bias 有连接，所以总共有 $156 * 28 * 28 = 122304$ 个连接（connection）。有 122304 个连接，但是我们只需要学习 156 个参数，主要是通过权值共享实现的。

3、池化层 S2

输入：28*28

采样区域：2*2

采样方式：4 个输入相加，乘以一个可训练参数，再加上一个可训练偏置。结果通过 sigmoid

采样种类：6

输出 featureMap 大小：14*14 (28/2)

神经元数量：14*14*6

连接数：(2*2+1)*6*14*14

S2 中每个特征图的大小是 C1 中特征图大小的 1/4。

详细说明：第一次卷积之后紧接着就是池化运算，使用 2*2 核 进行池化，于是得到了 S2，6 个 14*14 的特征图 (28/2=14)。S2 这个 pooling 层是对 C1 中的 2*2 区域内的像素求和乘以一个权值系数再加上一个偏置，然后将这个结果再做一次映射。同时有 $5 \times 14 \times 14 \times 6 = 5880$ 个连接。

4、卷积层 C3

输入：S2 中所有 6 个或者几个特征 map 组合

卷积核大小：5*5

卷积核种类：16

输出 featureMap 大小：10*10 (14-5+1)=10

C3 中的每个特征 map 是连接到 S2 中的所有 6 个或者几个特征 map 的，表示本层的特征 map 是上一层提取到的特征 map 的不同组合

存在的一个方式是：C3 的前 6 个特征图以 S2 中 3 个相邻的特征图子集为输入。接下来 6 个特征图以 S2 中 4 个相邻特征图子集为输入。然后的 3 个以不相邻的 4 个特征图子集为输入。最后一个将 S2 中所有特征图作为输入。

则：可训练参数： $6 \times (3 \times 5 \times 5 + 1) + 6 \times (4 \times 5 \times 5 + 1) + 3 \times (4 \times 5 \times 5 + 1) + 1 \times (6 \times 5 \times 5 + 1) = 1516$

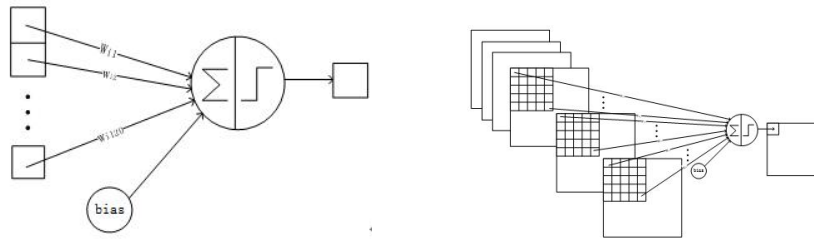
连接数：10*10*1516=151600

详细说明：第一次池化之后是第二次卷积，第二次卷积的输出是 C3，16 个 10x10 的特征图，卷积核大小是 5*5。我们知道 S2 有 6 个 14*14 的特征图，怎么从 6 个特征图得到 16 个特征图了？这里是通过 S2 的特征图特殊组合计算得到的 16 个特征图。具体如下：

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	X				X	X	X			X	X	X	X	X	X	X
1	X	X				X	X	X			X	X	X	X	X	X
2	X	X	X				X	X	X			X		X	X	X
3		X	X	X			X	X	X	X			X		X	X
4			X	X	X			X	X	X	X		X	X		X
5				X	X	X			X	X	X	X		X	X	X

C3 的前 6 个 feature map (对应上图第一个红框的 6 列) 与 S2 层相连的 3 个 feature map 相连接 (上图第一个红框)，后面 6 个 feature map 与 S2 层相连的 4 个 feature map 相连接 (上图第二个红框)，后面 3 个 feature map 与 S2 层部分不相连的 4 个 feature map 相连接，最后一个与 S2 层的所有 feature map 相连。卷积核大小依然为 5*5，所以总共有

$6 \times (3 \times 5 \times 5 + 1) + 6 \times (4 \times 5 \times 5 + 1) + 3 \times (4 \times 5 \times 5 + 1) + 1 \times (6 \times 5 \times 5 + 1) = 1516$ 个参数。而图像大小为 10×10 ，所以共有 151600 个连接。C3 与 S2 中前 3 个图相连的卷积结构如下图所示：



上图对应的参数为 $3 \times 5 \times 5 + 1$ ，一共进行 6 次卷积得到 6 个特征图，所以有 $6 \times (3 \times 5 \times 5 + 1)$ 参数。为采用上述这样的组合有两个原因：1) 减少参数，2) 这种不对称的组合连接的方式有利于提取多种组合特征。

5、池化层 S4

输入： 10×10

采样区域： 2×2

采样方式：4 个输入相加，乘以一个可训练参数，再加上一个可训练偏置。结果通过 sigmoid

采样种类：16

输出 featureMap 大小： 5×5 ($10/2$)

神经元数量： $5 \times 5 \times 16 = 400$

连接数： $16 \times (2 \times 2 + 1) \times 5 \times 5 = 2000$

S4 中每个特征图的大小是 C3 中特征图大小的 $1/4$

详细说明：S4 是 pooling 层，窗口大小仍然是 2×2 ，共计 16 个 feature map，C3 层的 16 个 10×10 的图分别进行以 2×2 为单位的池化得到 16 个 5×5 的特征图。有 $5 \times 5 \times 5 \times 16 = 2000$ 个连接。连接的方式与 S2 层类似。

6、卷积层 C5

输入：S4 层的全部 16 个单元特征 map（与 s4 全相连）

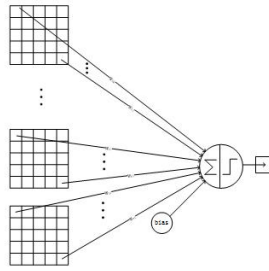
卷积核大小： 5×5

卷积核种类：120

输出 featureMap 大小： 1×1 ($5 - 5 + 1$)

可训练参数/连接： $120 \times (16 \times 5 \times 5 + 1) = 48120$

详细说明：C5 层是一个卷积层。由于 S4 层的 16 个图的大小为 5×5 ，与卷积核的大小相同，所以卷积后形成的图的大小为 1×1 。这里形成 120 个卷积结果。每个都与上一层的 16 个图相连。所以共有 $(5 \times 5 \times 16 + 1) \times 120 = 48120$ 个参数，同样有 48120 个连接。C5 层的网络结构如下：



7、全连接层 F6

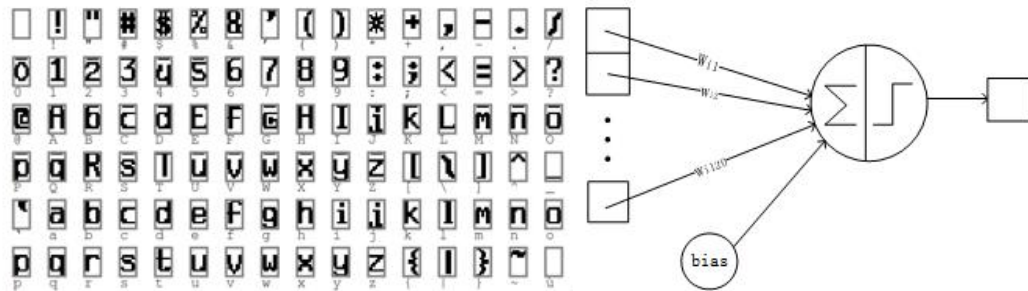
输入：c5 120 维向量

计算方式：计算输入向量和权重向量之间的点积，再加上一个偏置，结果通过 sigmoid 函数输出。

可训练参数： $84 \times (120 + 1) = 10164$

详细说明：6 层是全连接层。F6 层有 84 个节点，对应于一个 7x12 的比特图，-1 表示白色，1 表示黑色，这样每个符号的比特图的黑白色就对应于一个编码。该层的训练参数和连接数是 $(120 + 1) \times 84 = 10164$ 。ASCII 编码图如下：

F6 层的连接方式如下：

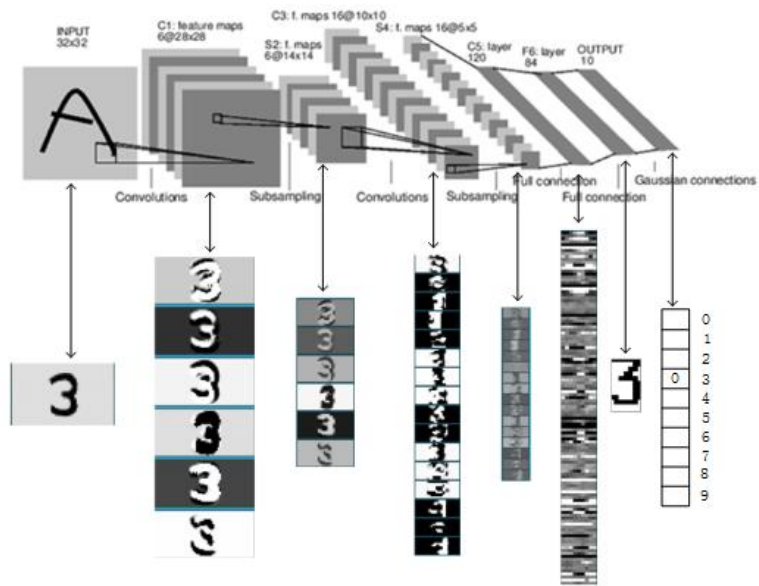


8、输出层（全连接层）

Output 层也是全连接层，共有 10 个节点，分别代表数字 0 到 9，且如果节点 i 的值为 0，则网络识别的结果是数字 i 。采用的是径向基函数（RBF）的网络连接方式。假设 x 是上一层的输入， y 是 RBF 的输出，则 RBF 输出的计算方式是：

$$y_i = \sum_j (x_j - w_{ij})^2$$

上式 w_{ij} 的值由 i 的比特图编码确定， i 从 0 到 9， j 取值从 0 到 $7 \times 12 - 1$ 。RBF 输出的值越接近于 0，则越接近于 i ，即越接近于 i 的 ASCII 编码图，表示当前网络输入的识别结果是字符 i 。该层有 $84 \times 10 = 840$ 个参数和连接。下图是 LeNet-5 识别数字 3 的过程。



1. LeNet-5 是一种用于手写体字符识别的非常高效的卷积神经网络。
2. 卷积神经网络能够很好的利用图像的结构信息。
3. 卷积层的参数较少，这也是由卷积层的主要特性即局部连接和共享权重所决定。