

CNN卷积神经网络

Convolutional Neural Networks

王生生

CNN

✓ 背景

✓ 定义

✓ 问题分析

✓ 基本概念

局部感知

权值共享

池化

✓ 综合

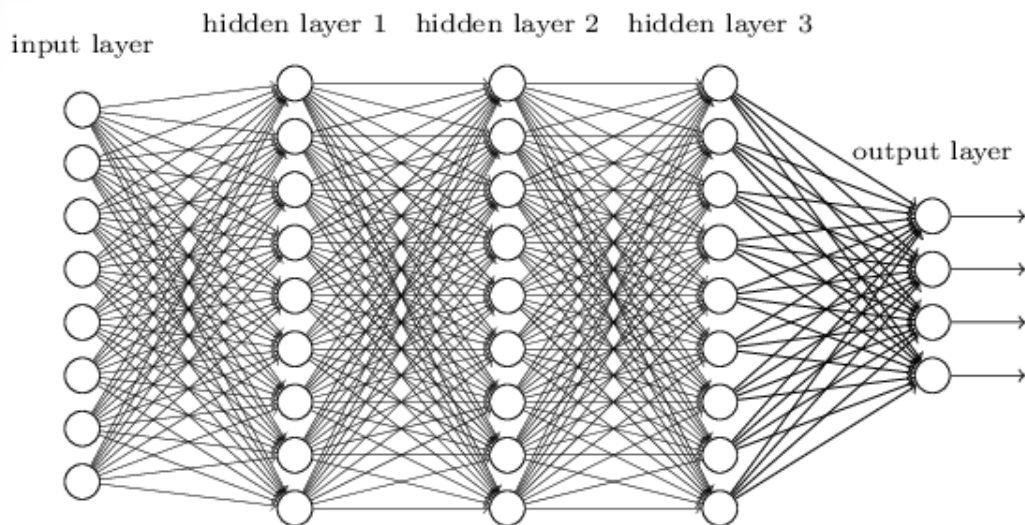
CNN产生背景

输入层

隐藏层/中间层

输出层

传统神经网络



全连接中的每一层网络都和相邻层全部连接。如此并没有考虑到图片中像素的空间分布，不管两个像素的距离远近都一视同仁，显然这样做是不合理的。

因此**卷积神经网络**出现了，它考虑到了输入值（像素）的空间分布，再加上一些人工设定的特性（例如共享权重等），使得它非常容易训练。从而可以做出更深层的网络结构，拥有更好的识别效果。

CNN 定义

卷积神经网络（Convolutional Neural Networks, CNN）是一种前馈神经网络，它的人工神经元可以响应一部分覆盖范围内的周围单元，对于大型图像处理、图像分类有出色表现。

卷积神经网络是目前广泛使用的一种深度学习网络。测试10000张MNIST图像数据集，CNN的正确率能够达到99.67%。

CNN 问题分析



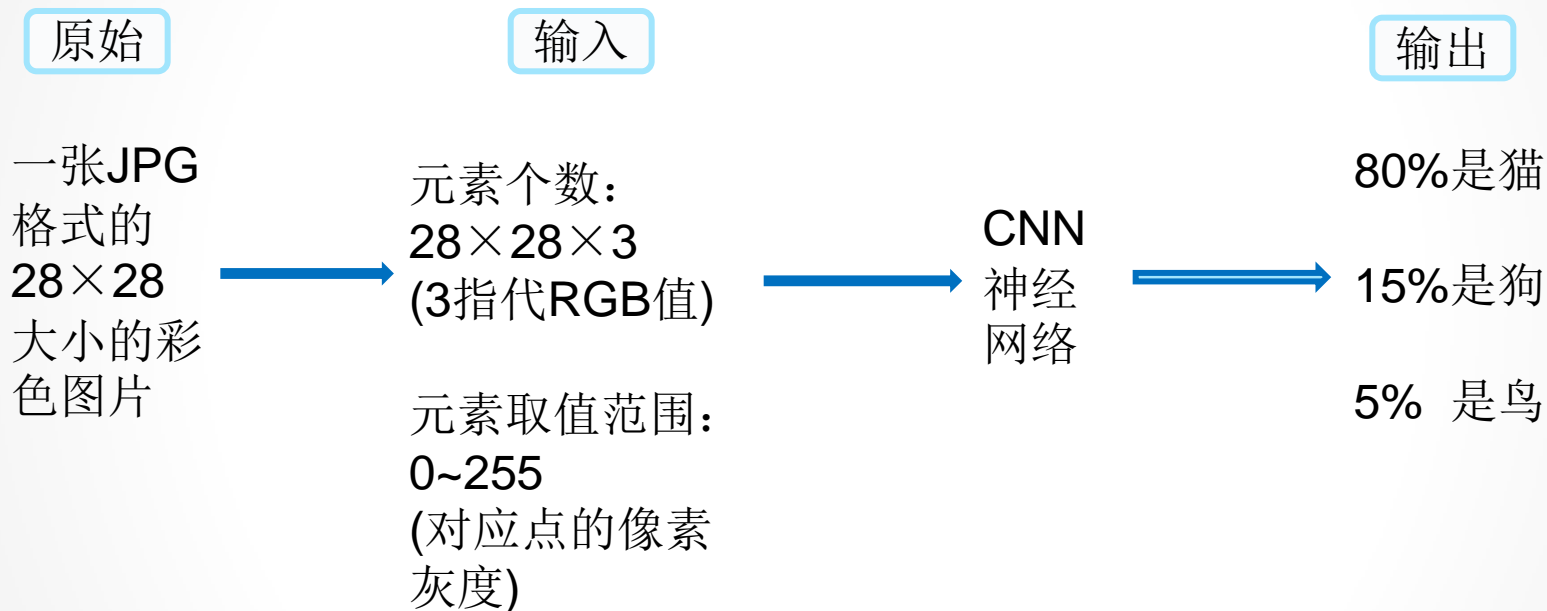
人看到的

```
08 02 22 97 38 15 00 40 00 75 04 05 07 78 82 12 50 77 91 08
49 49 99 40 17 81 18 57 60 87 17 40 98 43 69 48 04 56 62 00
81 49 31 73 55 79 14 29 93 71 40 67 53 88 30 03 49 13 36 65
52 70 95 23 04 60 11 42 69 24 68 56 01 32 56 71 37 02 36 91
22 31 16 71 51 67 63 89 41 92 36 58 22 40 40 28 66 33 13 60
24 47 32 60 99 03 45 02 44 75 33 53 78 36 84 20 35 17 12 50
32 98 81 28 64 23 67 10 26 38 40 67 39 54 70 66 18 38 64 70
67 26 20 68 02 62 12 20 35 63 94 39 63 08 40 91 66 49 94 21
24 55 38 05 66 73 99 26 97 17 78 78 86 83 14 88 34 89 63 72
21 36 29 09 78 00 76 44 20 45 35 14 00 61 33 97 34 31 33 95
78 17 53 28 22 75 31 67 15 94 03 80 04 62 16 14 09 53 56 92
16 39 05 42 96 35 31 47 55 58 88 28 00 17 54 24 36 29 85 57
86 56 00 48 35 71 89 07 03 44 44 37 44 60 21 58 51 54 17 58
19 80 81 68 05 94 47 89 28 73 92 13 86 52 17 77 04 89 55 40
04 52 08 83 97 35 99 16 07 97 57 32 16 26 26 79 33 27 98 66
88 36 48 87 57 62 20 72 03 46 33 67 46 55 12 32 63 93 53 69
04 42 16 73 38 25 39 11 24 94 72 18 08 46 29 32 40 62 76 36
20 69 36 41 72 30 23 88 34 62 99 69 82 67 59 85 74 04 36 16
20 73 35 29 78 31 90 01 74 31 49 71 48 86 81 16 23 57 05 54
01 70 84 71 83 61 94 65 16 92 33 48 61 43 52 01 89 19 67 48
```



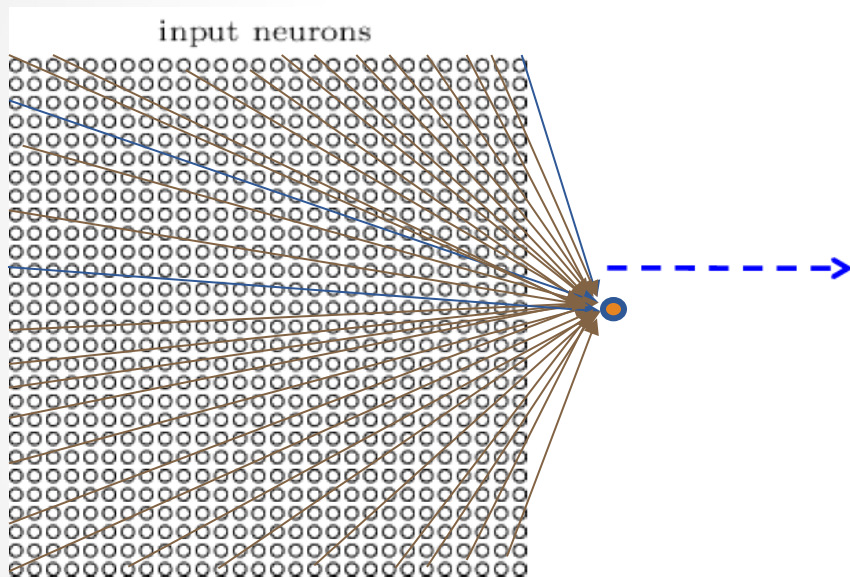
计算机看到的

CNN 问题分析



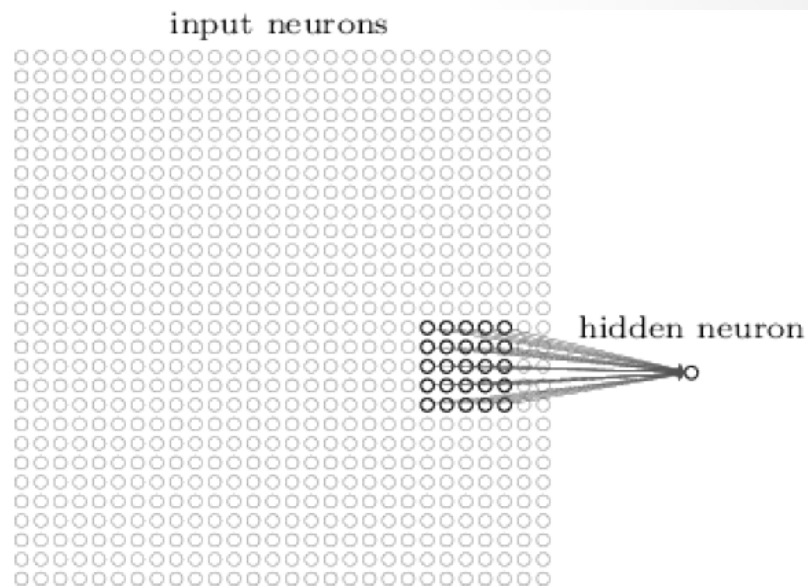
局部感知

全连接



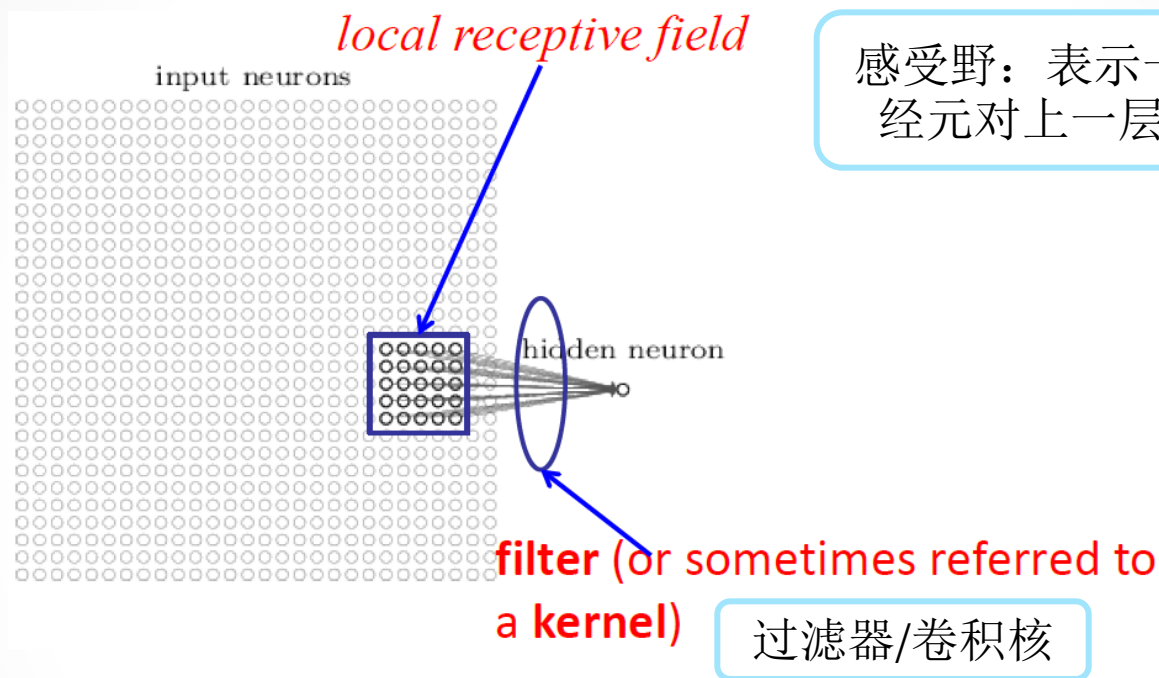
一个隐藏神经元
与上一层
全部
的神经元连接

CNN



一个隐藏神经元
与上一层
一小块区域
的神经元连接

局部感知



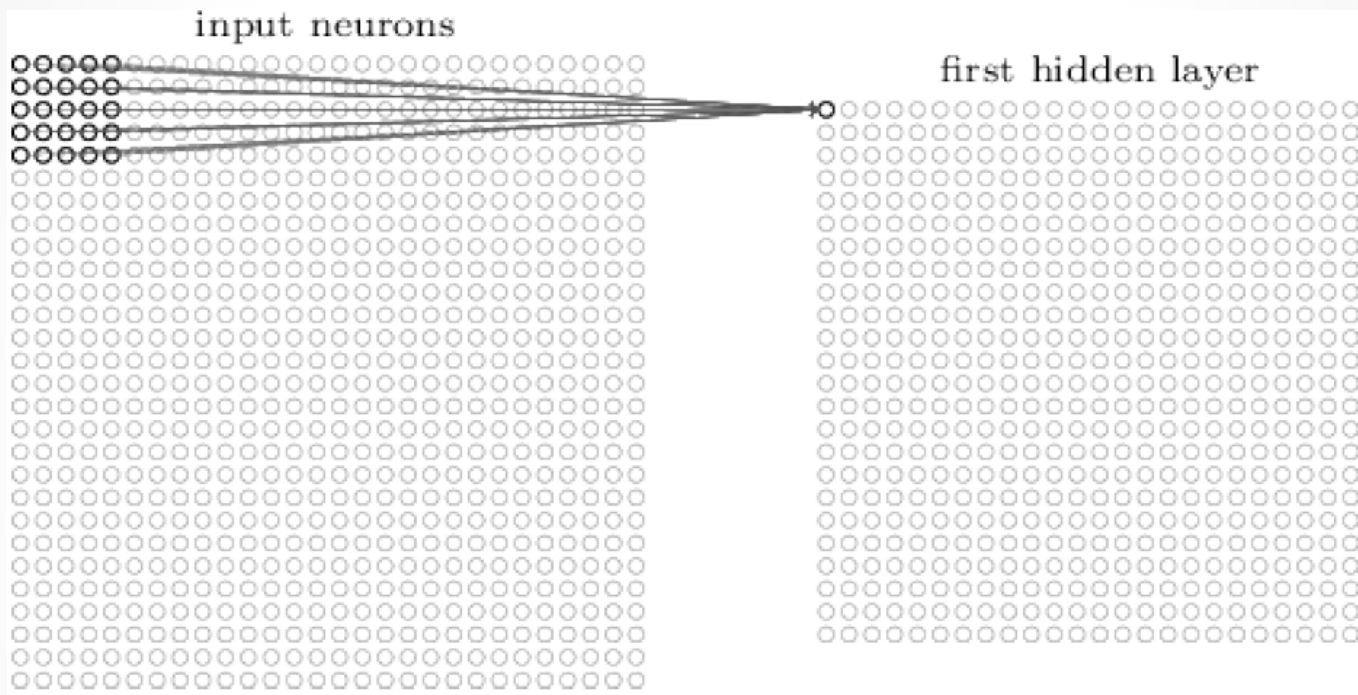
感受野：表示一个隐藏层神经元对上一层的感受区域

以输入 $28 \times 28 \times 3$ 的像素数组， $5 \times 5 \times 3$ 的过滤器数组为例：

过滤器与输入数组对应位置的值相乘再相加后得到的数值，可表示过滤器在当前感受野的过滤情况。

注：过滤器的深度必须与输入数组的深度相同。

局部感知



将过滤器滑过整个输入数组后将得到一个 $24 \times 24 \times 1$ 的数组，称之为**激活映射**（activation map）或**特征映射**（feature map）。

局部感知示例1

输入层

过滤器

特征映射

0	1	1	1	0	0	0
0	0	1	1	1	0	0
0	0	0	1	1	1	0
0	0	0	1	1	0	0
0	0	1	1	0	0	0
0	1	1	0	0	0	0
1	1	0	0	0	0	0

1	0	1
0	1	0
1	0	1

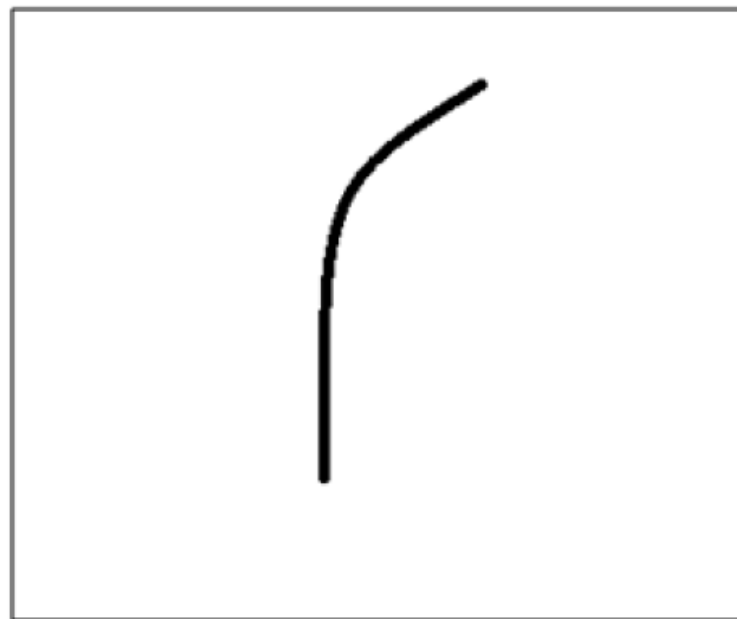
1	4	3	4	1
1	2	4	3	3
1	2	3	4	1
1	3	3	1	1
3	3	1	1	0

局部感知示例2

过滤器（像素表示）

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

过滤器（图像表示）



每个过滤器可以被看成是特征标识符。特征指直边缘、曲线之类。

假设过滤器是 $7 \times 7 \times 3$ 的曲线检测器。

为了易于分析，暂且忽略该过滤器的深度为 3 个单元，只考虑过滤器和图像的顶层层面。

作为曲线过滤器，它将有一个像素结构，在曲线形状附近会产生高数值。

局部感知示例2



Original image



Visualization of the filter on the image



Visualization of the receptive field

0	0	0	0	0	0	30
0	0	0	0	50	50	50
0	0	0	20	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0
0	0	0	50	50	0	0

Pixel representation of the receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Multiplication and Summation = $(50*30)+(50*30)+(50*30)+(20*30)+(50*30) = 6600$ (A large number!)

局部感知示例2



Visualization of the filter on the image

0	0	0	0	0	0	0
0	40	0	0	0	0	0
40	0	40	0	0	0	0
40	20	0	0	0	0	0
0	50	0	0	0	0	0
0	0	50	0	0	0	0
25	25	0	50	0	0	0

Pixel representation of receptive field

*

0	0	0	0	0	30	0
0	0	0	0	30	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	30	0	0	0
0	0	0	0	0	0	0

Pixel representation of filter

Multiplication and Summation = 0

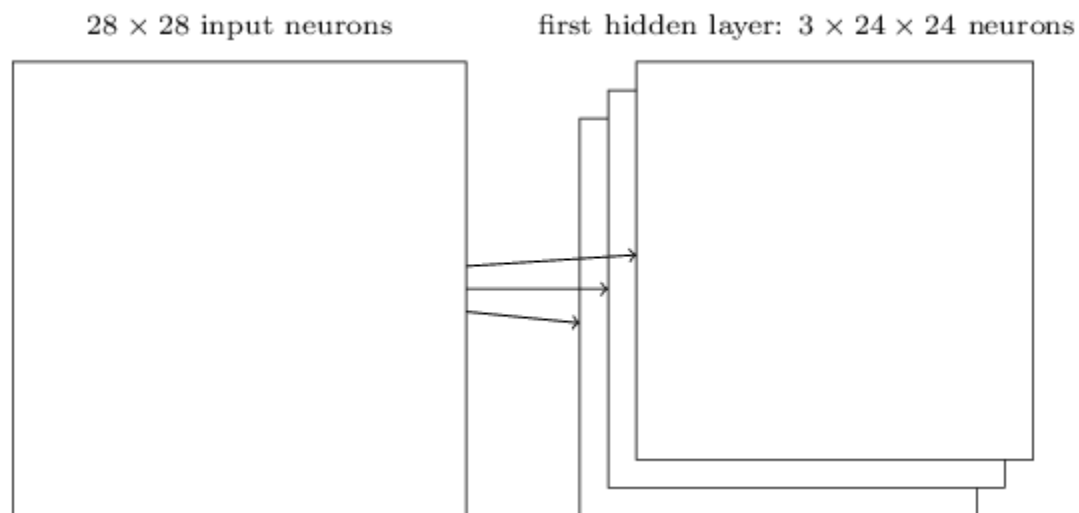
6600: 高数值，曲线激活了过滤器，说明原图左上角中存在曲线。

0: 低数值，没有激活过滤器，说明原图右上角没有曲线。

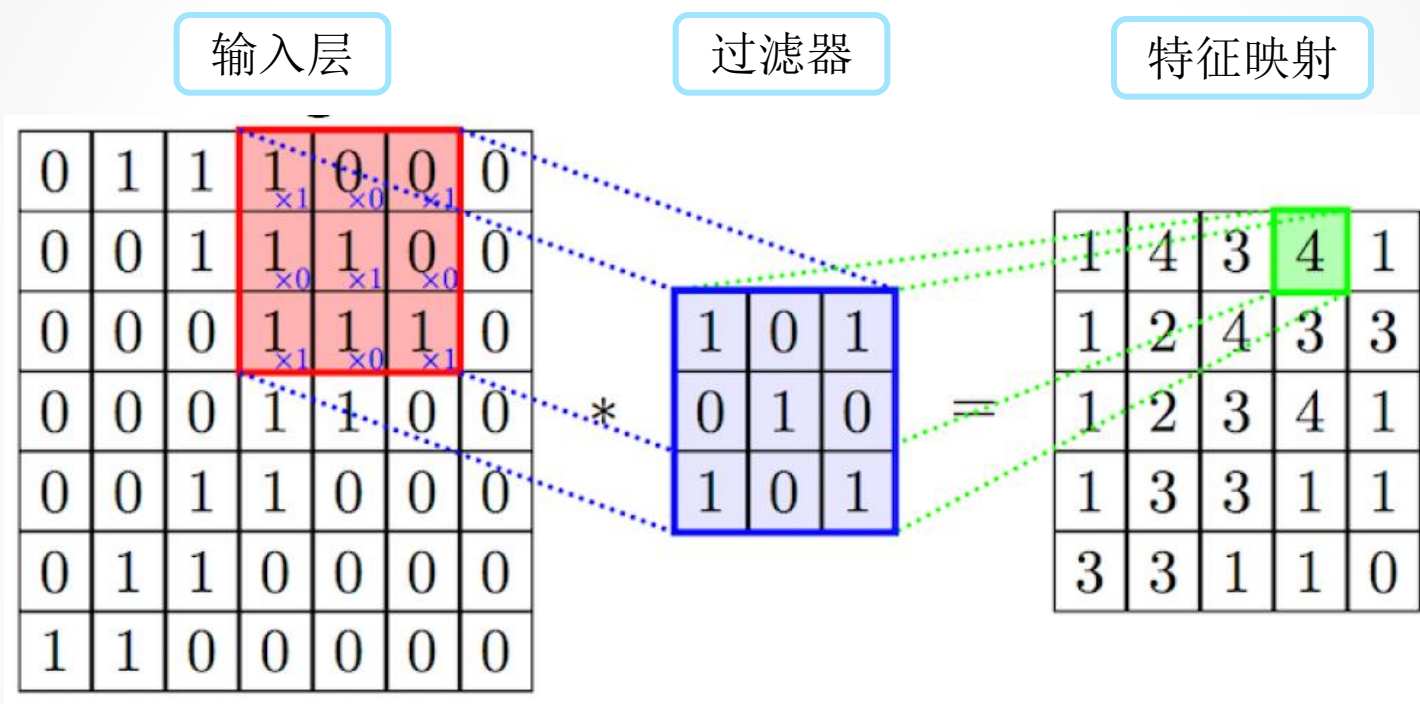
这是检测右弯曲线的过滤器。还有检测左弯曲线或直线边缘的过滤器。过滤器越多，激活映射的深度越大，对输入内容的了解也就越多。

局部感知

使用3个过滤器，检测三种特征：



权值共享

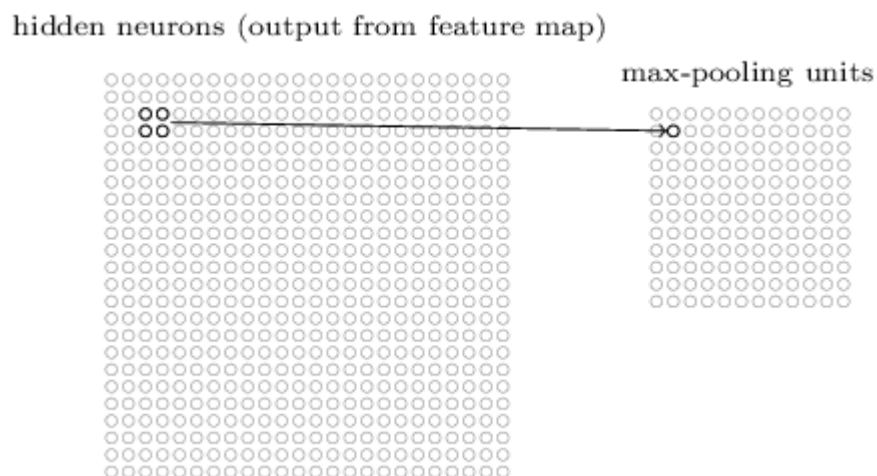


从输入层得到特征映射的过程中使用了同一个过滤器，即权值参数共享。

优点：权值共享大大减少了参数个数

池化

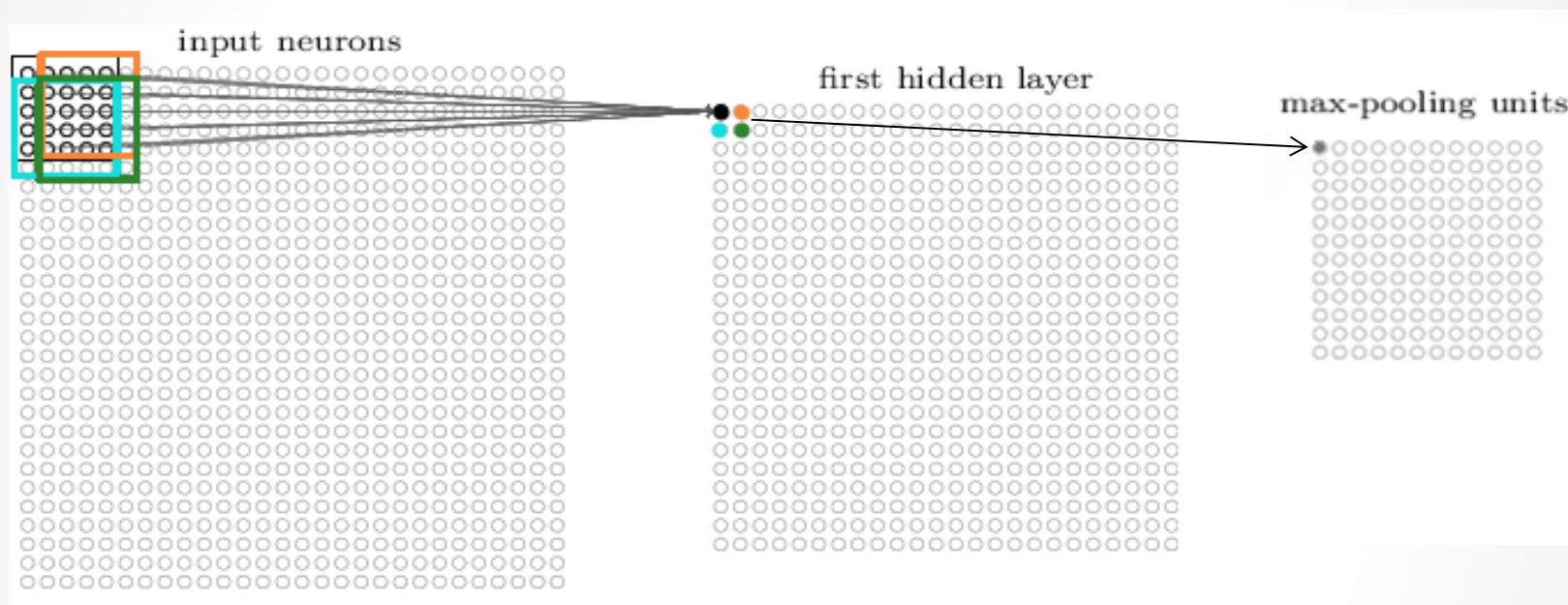
池化层接在卷积层后面，用于简化卷积的输出结果，即把卷积层的输出结果进行压缩：



上图使用 **最大池化法（max-pooling）** 将四个神经元压缩为一个神经元，即在4个相邻的神经元中取最大的值，其它丢弃。

上图 24x24 的特征图经过 2x2的max-pooling之后变成了12x12的特征图。

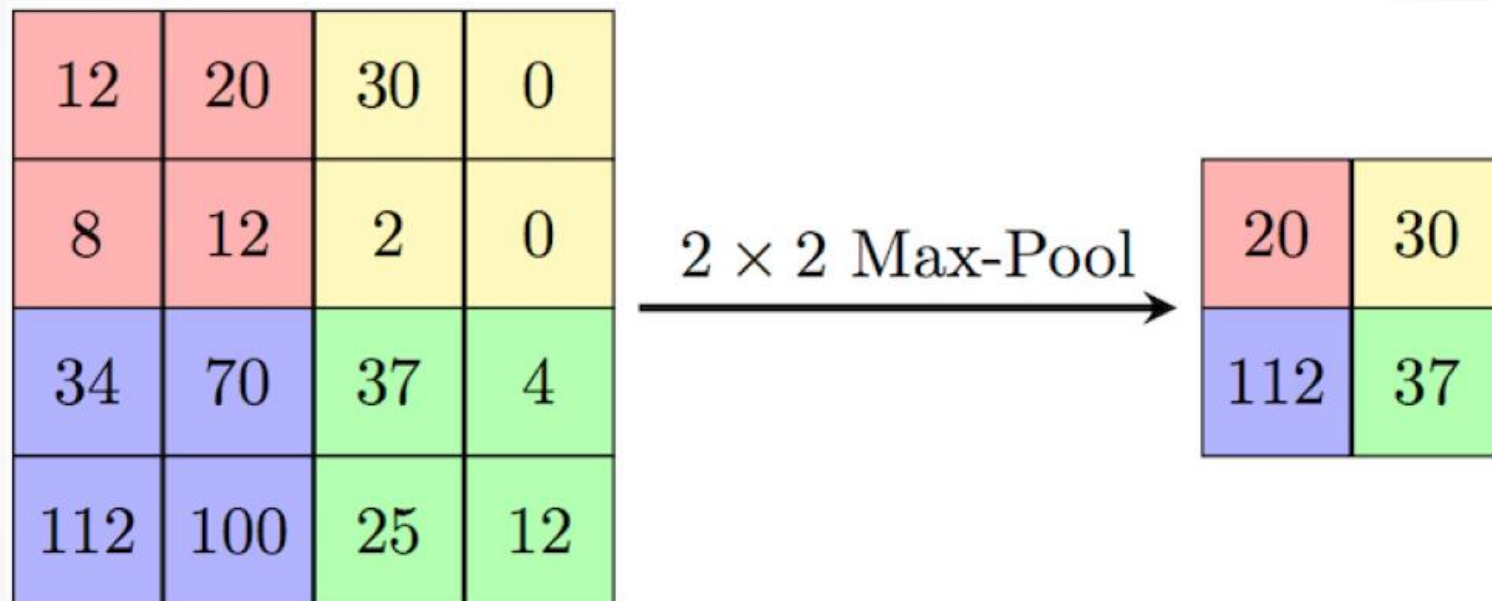
池化



可将 max-pooling 看作去检测某个特征在图片中是否存在，而忽略其精确位置。

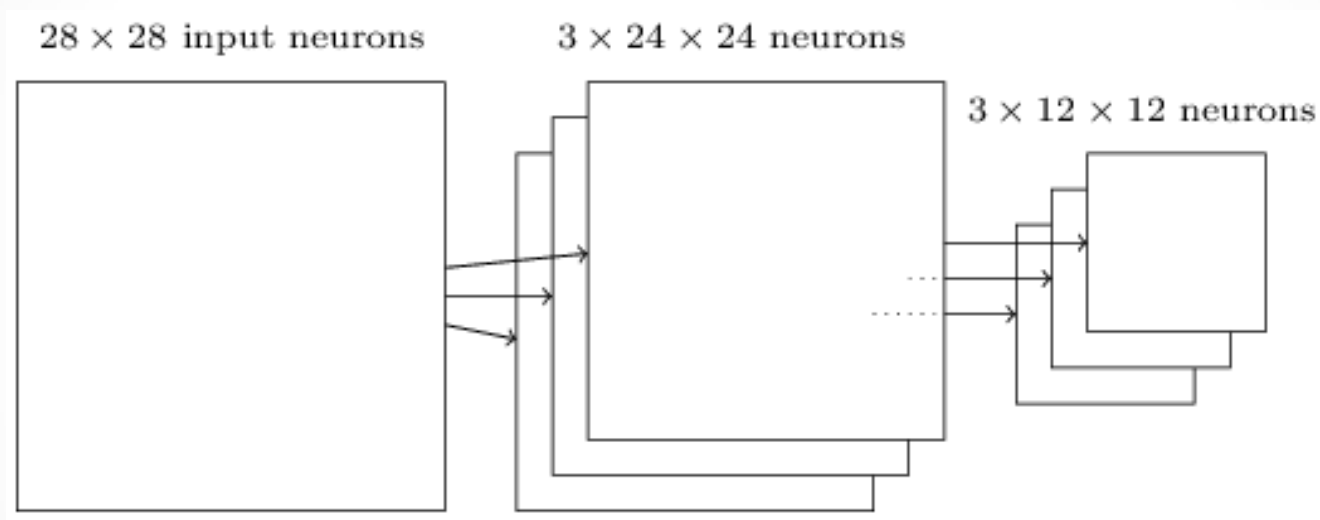
优点：池化后特征数量大大减少

池化示例

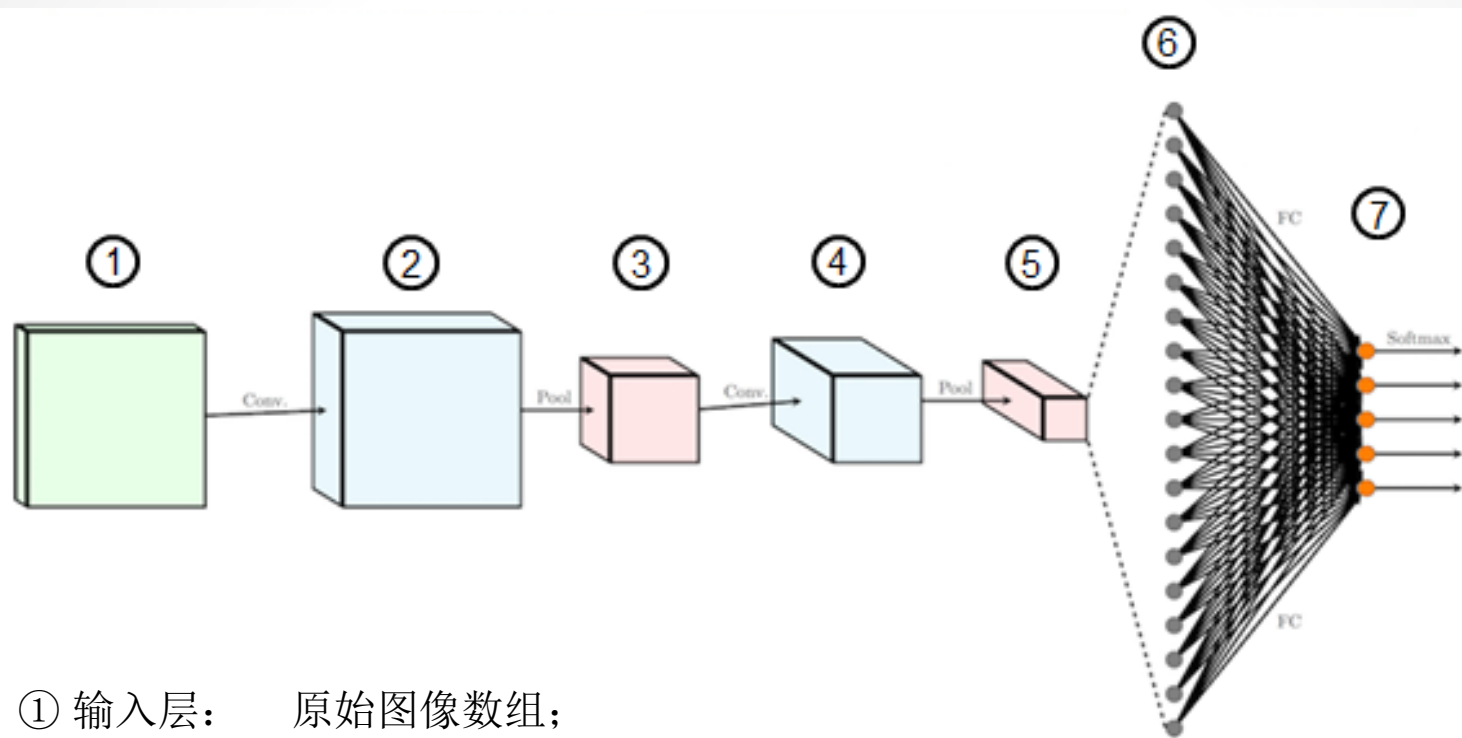


池化

对卷积得到的特征映射分别进行池化：



综合



- ① 输入层：原始图像数组；
- ② 卷积层：采用检测**低级特征**的过滤器，得到激活映射；
- ③ 池化层：压缩数据；
- ④ 卷积层：采用检测**高级特征**的过滤器，得到激活映射。
- ⑤ 池化层；
- ⑥ 全连接层：分类器；
- ⑦ 输出层：**N**维向量，数值表示特定类别的概率。

参考资料

- [1] <https://www.zhihu.com/question/52668301/answer/131573702>
- [2] <https://www.zhihu.com/question/39022858/answer/81026163>
- [3] <https://zhuanlan.zhihu.com/p/23185164>
- [4] <http://blog.csdn.net/zouxy09/article/details/8781543/>
- [5] <http://blog.csdn.net/stdcoutzyx/article/details/41596663/>

谢谢！