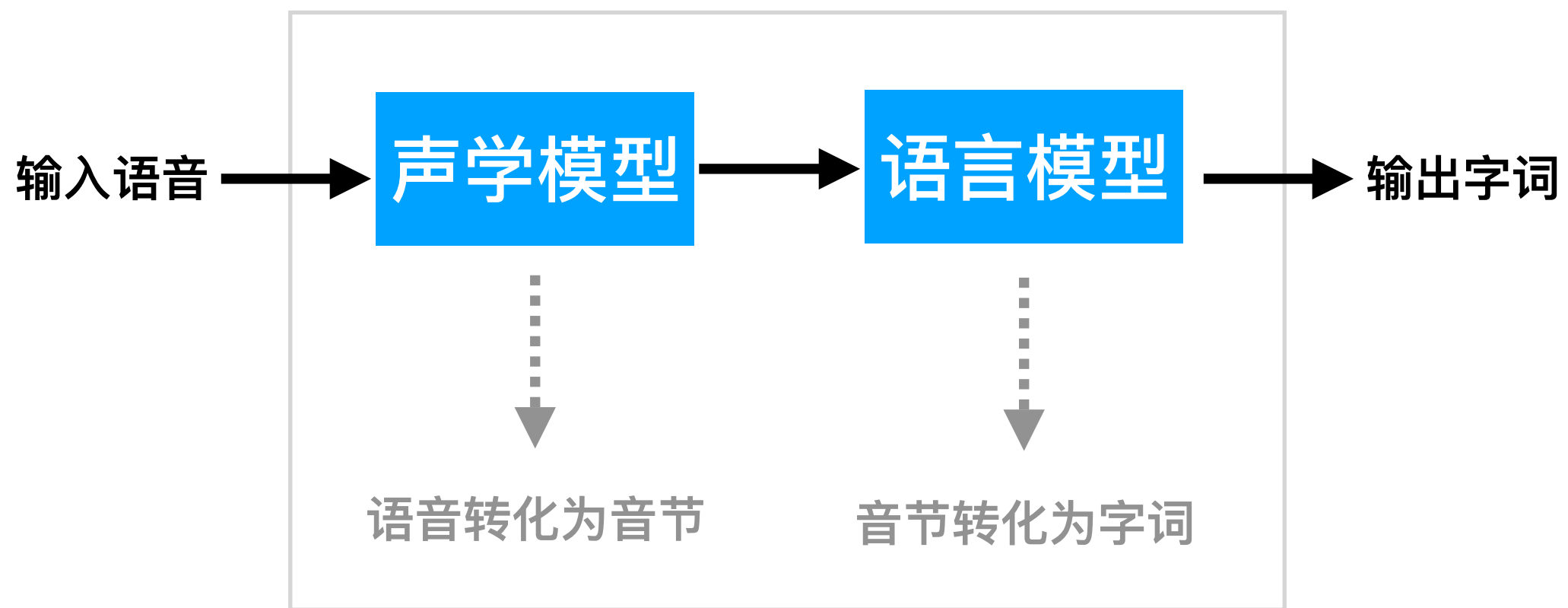


深度学习

卷积神经网络

传统语音识别系统构成

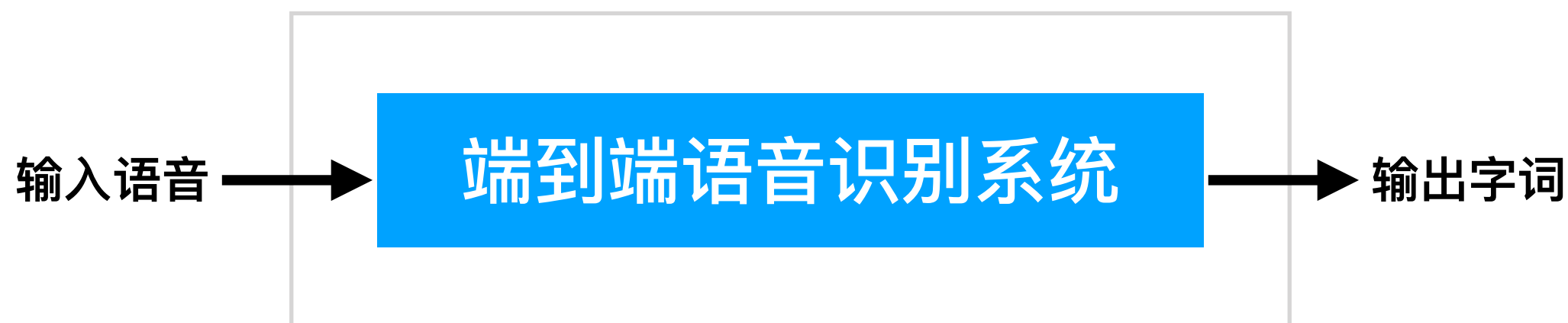


独立训练两个系统。致使模型整体的训练目标与两个系统的训练目标有偏差。

模型训练复杂

模型性能较差

端到端语音识别系统构成



端到端模型仅有一个系统就可以完成整个任务。

训练方便

模型性能好

端到端模型

端到端模型(End-to-end models)是指完成机器学习任务时可以直接由输入样本得到结果的模型。其特点是一个模型可以完成一项复杂任务。端到端模型的代价函数只有一个，在性能上往往优于其他结构的模型。端到端模型在训练时是端到端训练。ANN算法在做语音识别、图像识别等任务时均可以使用端到端模型，是最常用的端到端算法模型。

小练习

以下哪些模型是端到端模型？

- A. 现拥有北京房价与房屋面积的数据，训练线性回归模型做房价预测。其中输入模型的是房屋面积，模型输出房价预测结果。
- B. 使用神经网络识别照片中的红绿灯。神经网络接收照片作为输入，输出是红绿黄三色灯的概率。
- C. 语音识别任务中，将训练好的声学模型与语言模型连接在一起作为一个整体再进行训练，得到最终的模型。
- D. 现有西瓜照片与好坏瓜的标记。通过人工提取西瓜照片中瓜的颜色、纹理形状等特征制作成为代表西瓜照片的样本，然后使用决策树对西瓜照片进行分类。

ANN端到端做图像识别的问题

- 直接输入样本会使输入层规模巨大，带来维度灾难问题。
- 无先验知识，导致特征提取效率低下。
- 样本数量需求巨大。

亟待解决的问题：

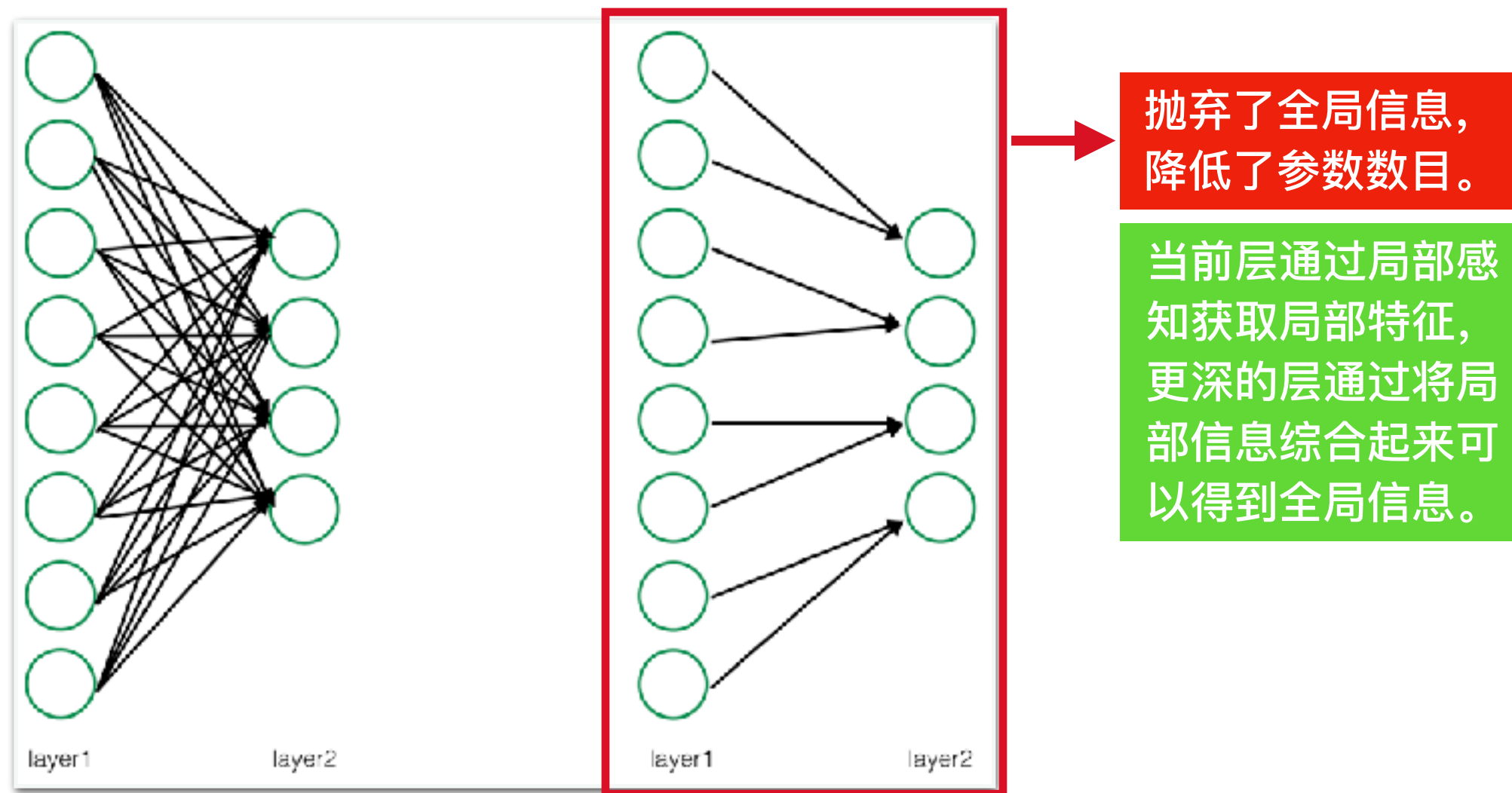
1. 降低全连接层带来的数据维度，以避免维度灾难问题。

2. 提高特征提取效率，加快神经网络训练速度。

卷积神经网络

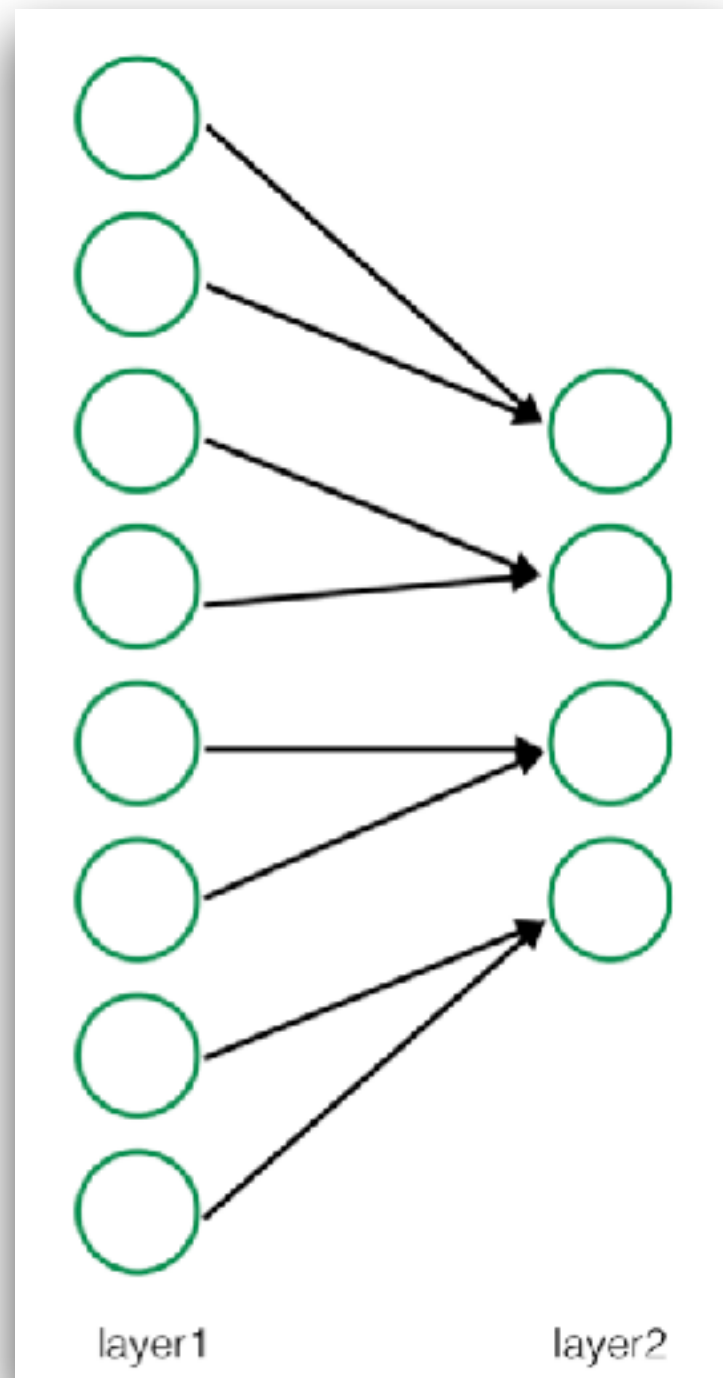
卷积神经网络(Convolutional Neural Network, CNN)是一种经典的前馈神经网络，主要受生物学中的**感受野(receptive field)**的概念提出。感受野在生物体中广泛存在，一个感受野连接多个感受器细胞，这些感受器细胞共同决定了感受野是否兴奋。通过感受野的机制，生物体传入的信号数量会大大降低，同时也能很好的对输入信号进行特征提取。卷积神经网络有两大工具来降低参数数目：卷积与池化。

局部感知



图像等数据其空间联系的规律是：局部像素关联紧密，较远距离的像素则相关性较弱。局部像素可以构成微观特征。

局部连接



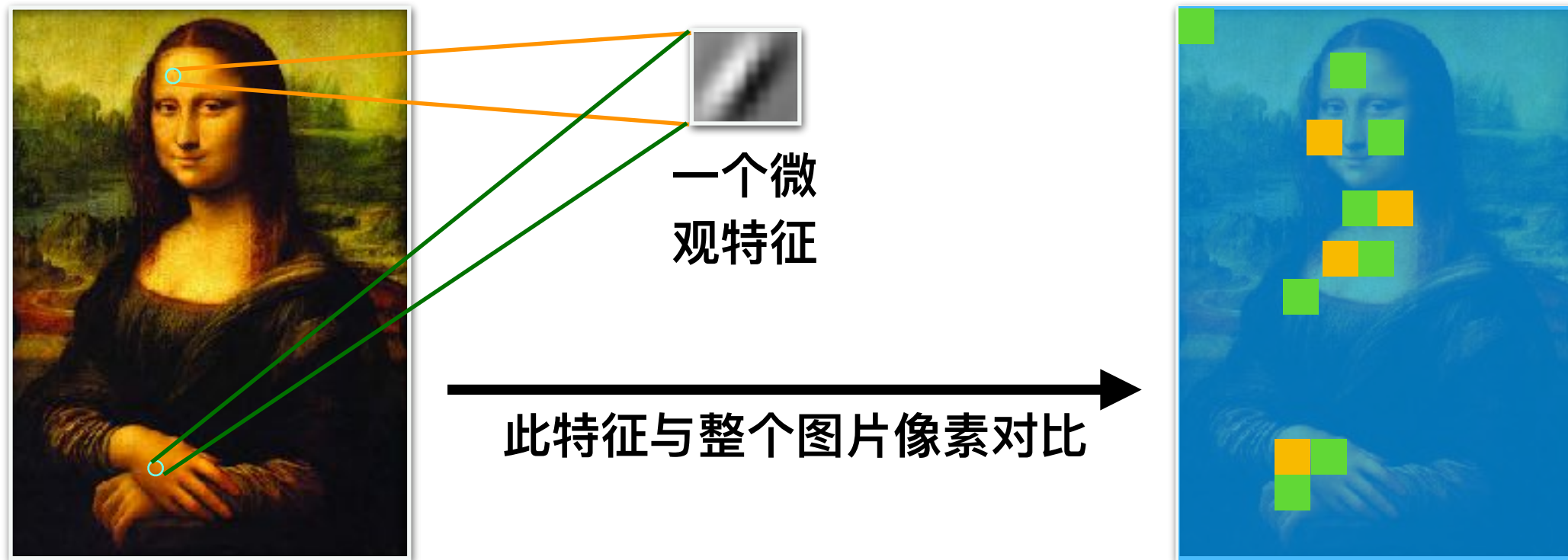
思考：直接使用局部连接是否可行？

1. 当输入一个样本时，局部相关性不清楚，无法设置感知范围。
2. 不同样本的局部相关性不同，无法设置固定的感知范围。

我们知道一定数量的微观特征可以表征任意图片。局部感知的本质是将一个个微观特征连接到了下一层神经网络。

局部感知的应用——卷积

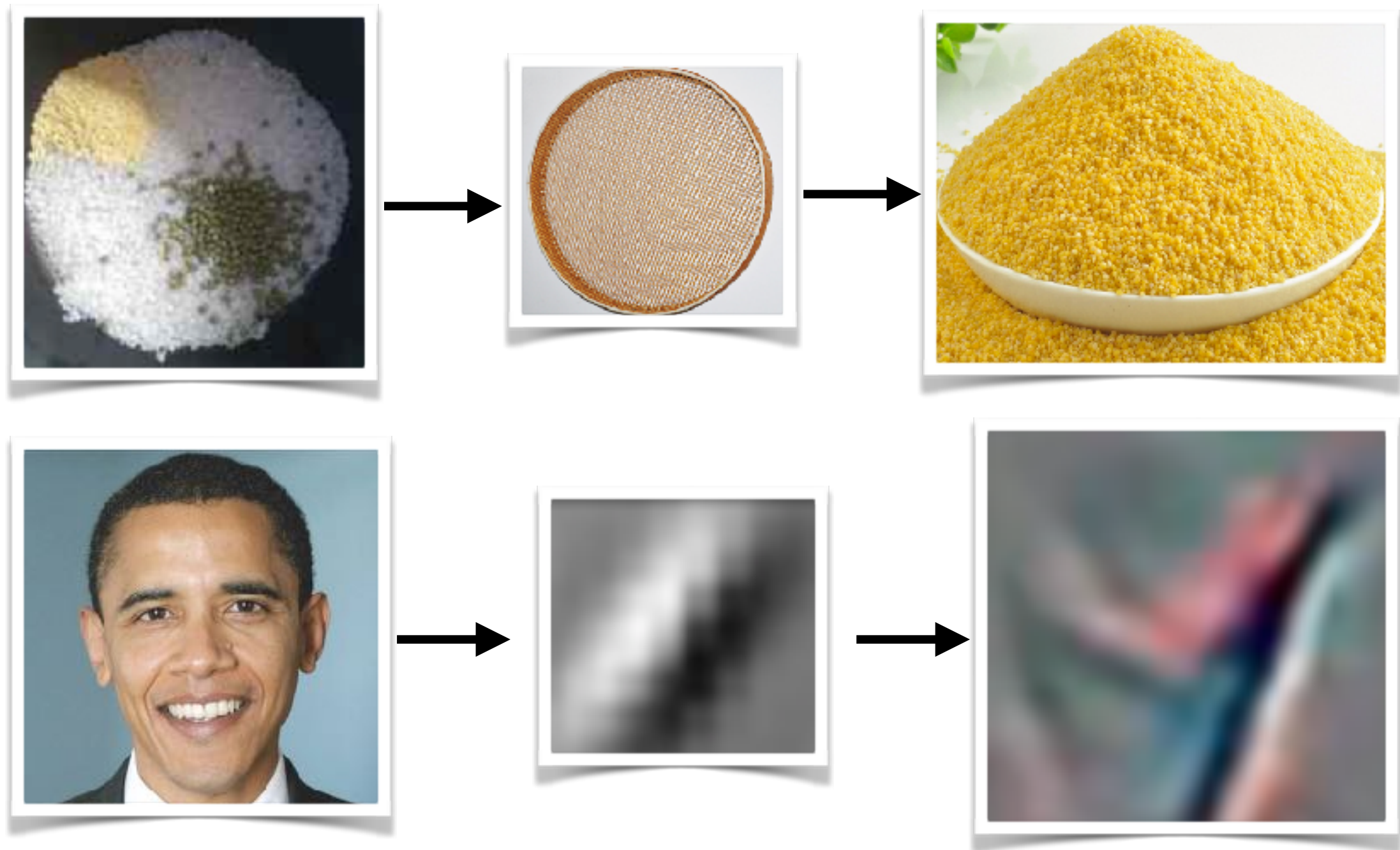
假如已知一个微观特征的形式，我们可以通过比对整个输入数据，找到于此匹配的部分。



图片

绿色部分表示此处是此特征，黄色部分表示此处与此特征比较相近，其它部分表示完全不符合此特征。此图是一个特征分布图。

特征筛选



卷积

1 _{x1}	1 _{x0}	1 _{x1}	0	0
0 _{x0}	1 _{x1}	1 _{x0}	1	0
0 _{x1}	0 _{x0}	1 _{x1}	1	1
0	0	1	1	0
0	1	1	0	0

Image

4		

Convolved
Feature

卷积隐含的原则是：输入（此处是图像）的一部分的统计特性与其他部分是一样的。

我们把3*3的筛子叫做卷积核 (Convolution kernel)。右边的图是卷积核筛选特征的结果，我们把它叫做特征图 (Feature map)。特征图反映了某个特征在某个输入（此处是图片）上的激活值。

小练习

1. 对 $10\text{px} \times 10\text{px}$ 的图片使用 $3\text{px} \times 3\text{px}$ 的卷积核做卷积得到的特征图大小是多少?
2. 对 $10\text{px} \times 10\text{px}$ 的图片使用 $2\text{px} \times 3\text{px}$ 的卷积核做卷积得到的特征图大小是多少?

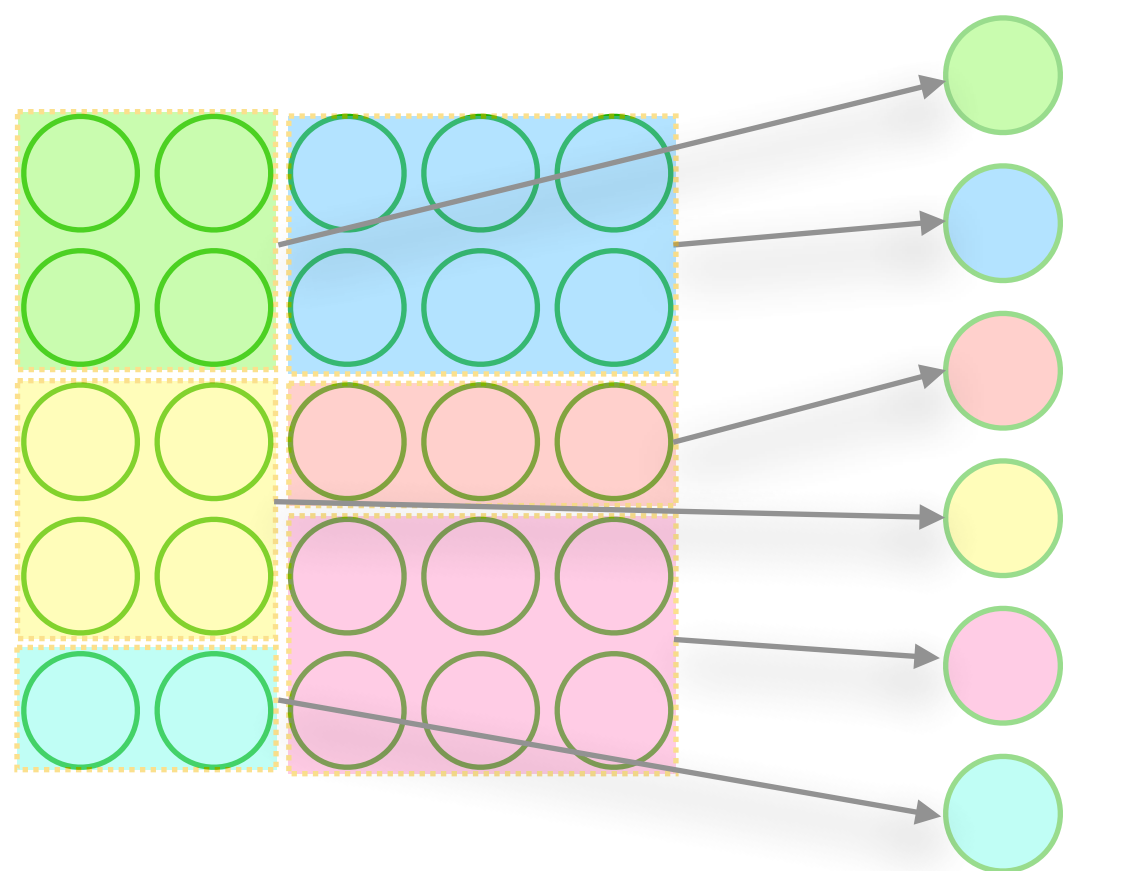
卷积

- 一个卷积核代表一个特征筛选器。卷积就是**使卷积核逐步滑过整个图像**并与对应部分图像进行计算。卷积的结果可以理解为当前特征在各处的激活值组成的矩阵，我们称之为特征图。
- 卷积通常需要逐步滑动卷积核进行运算，其目的是测试所有可能的像素组合。这样就解决了局部感知中无法确定范围的问题了。
- 特征图反映一个特征的图像上的某个位置的可能性。

思考：为什么卷积核每次滑动二步、三步可以吗？

局部连接与卷积对比

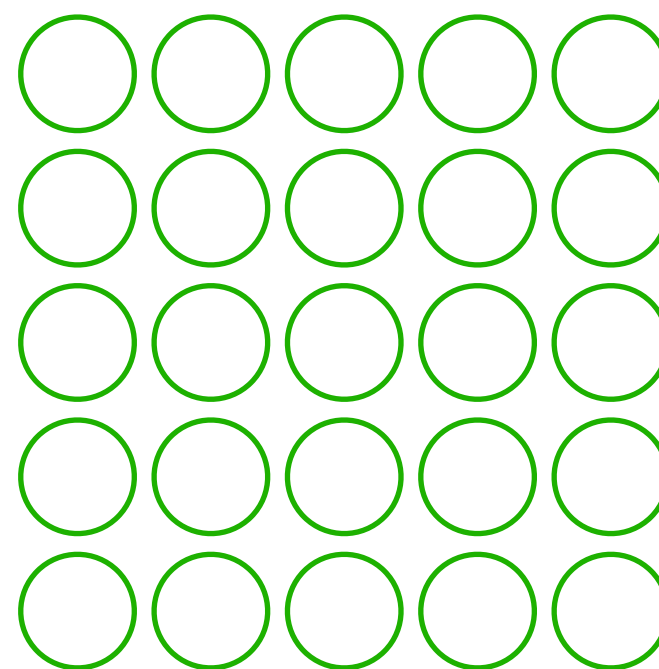
局部连接



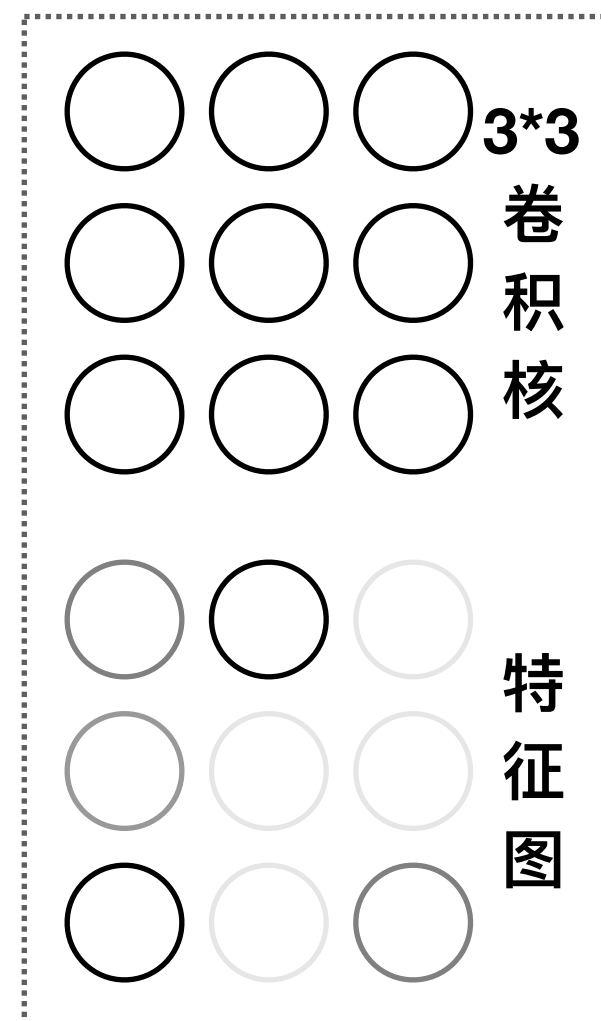
5px*5px的图片

第一个隐藏层

卷积



5px*5px的图片



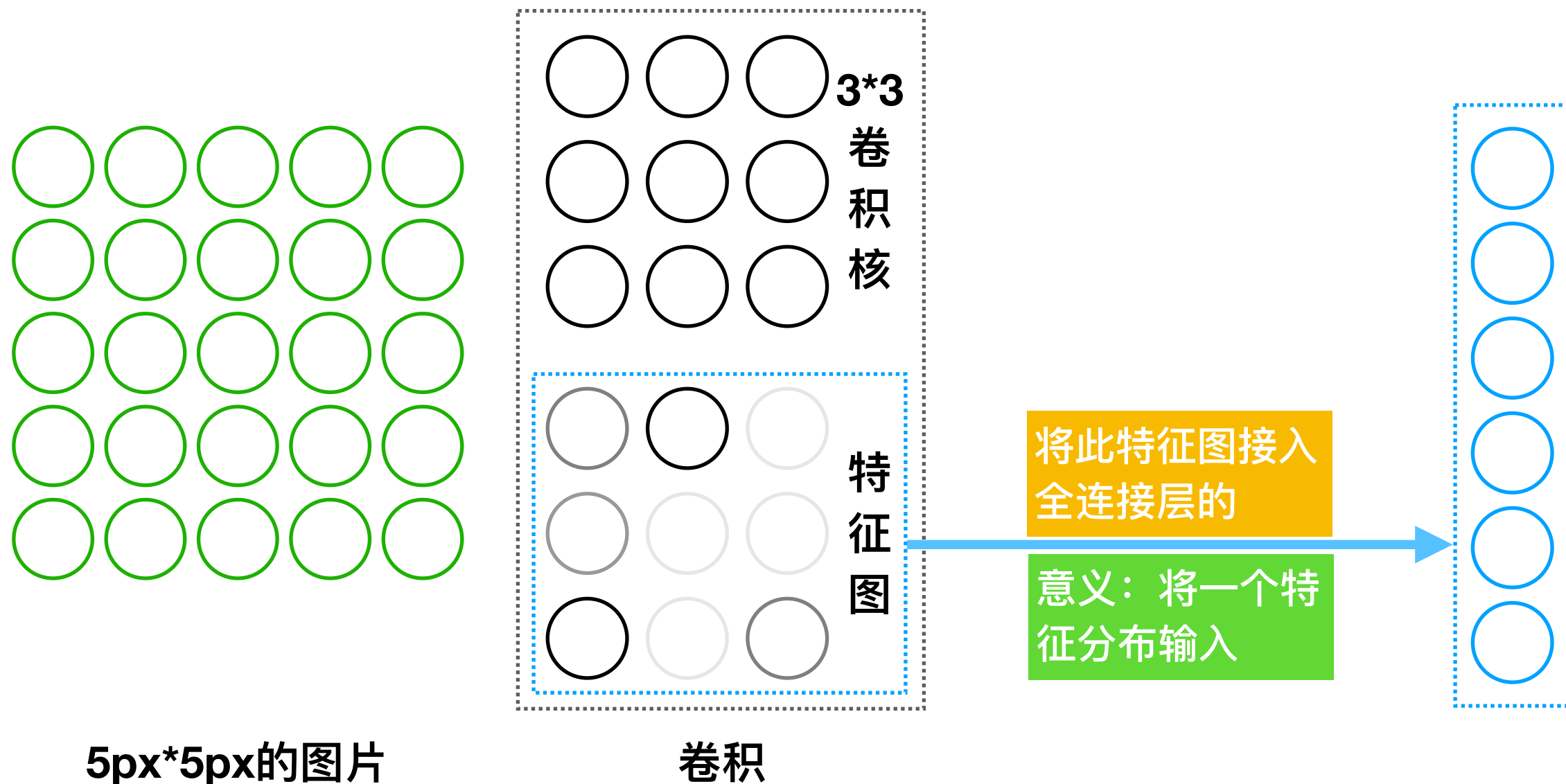
卷积

Author: WangQi

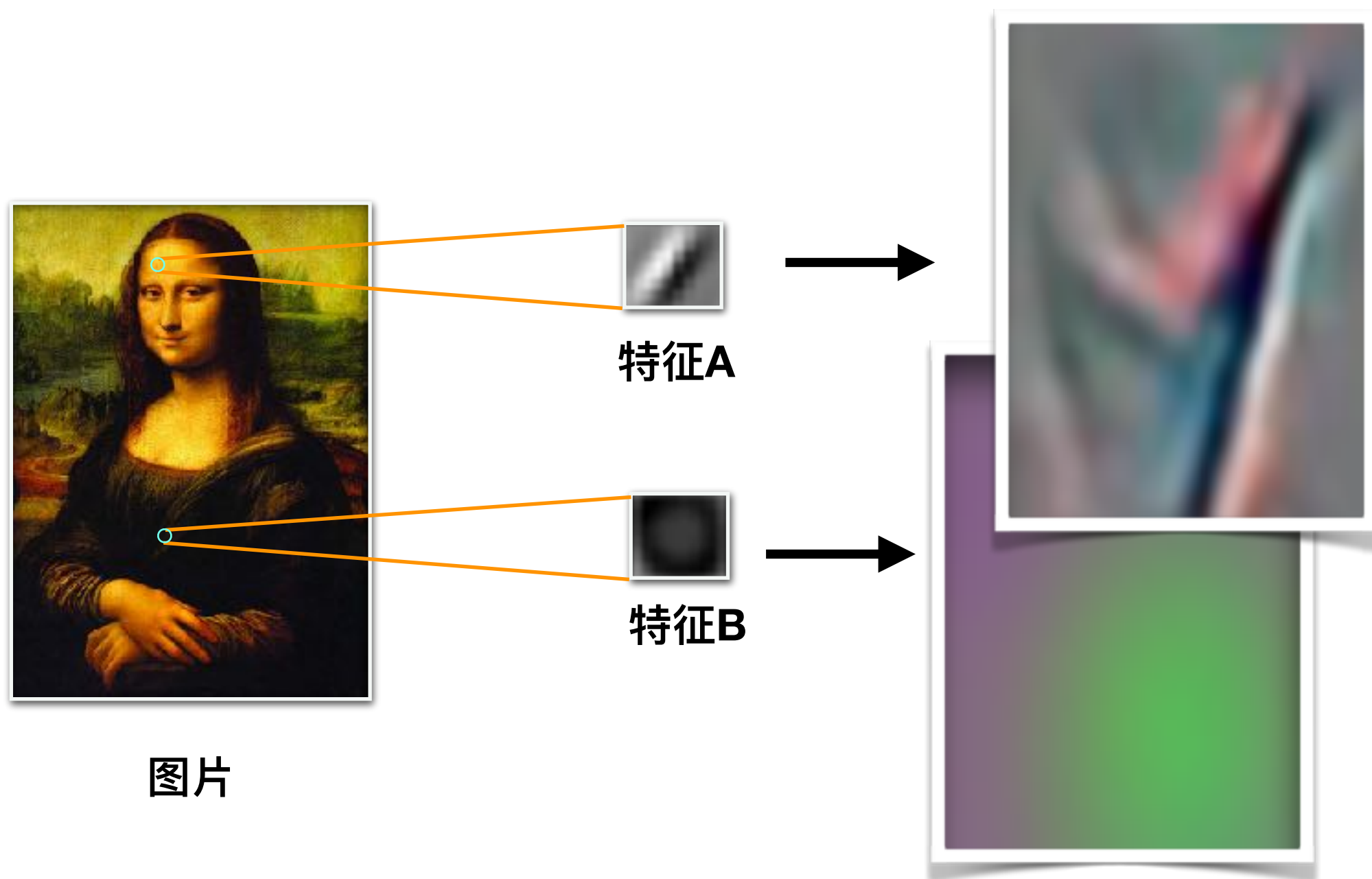
Email: wangqikaixin@gmail.com

使用卷积特征图方法一

一次卷积得到的特征图可以看作是一个特征的分布。

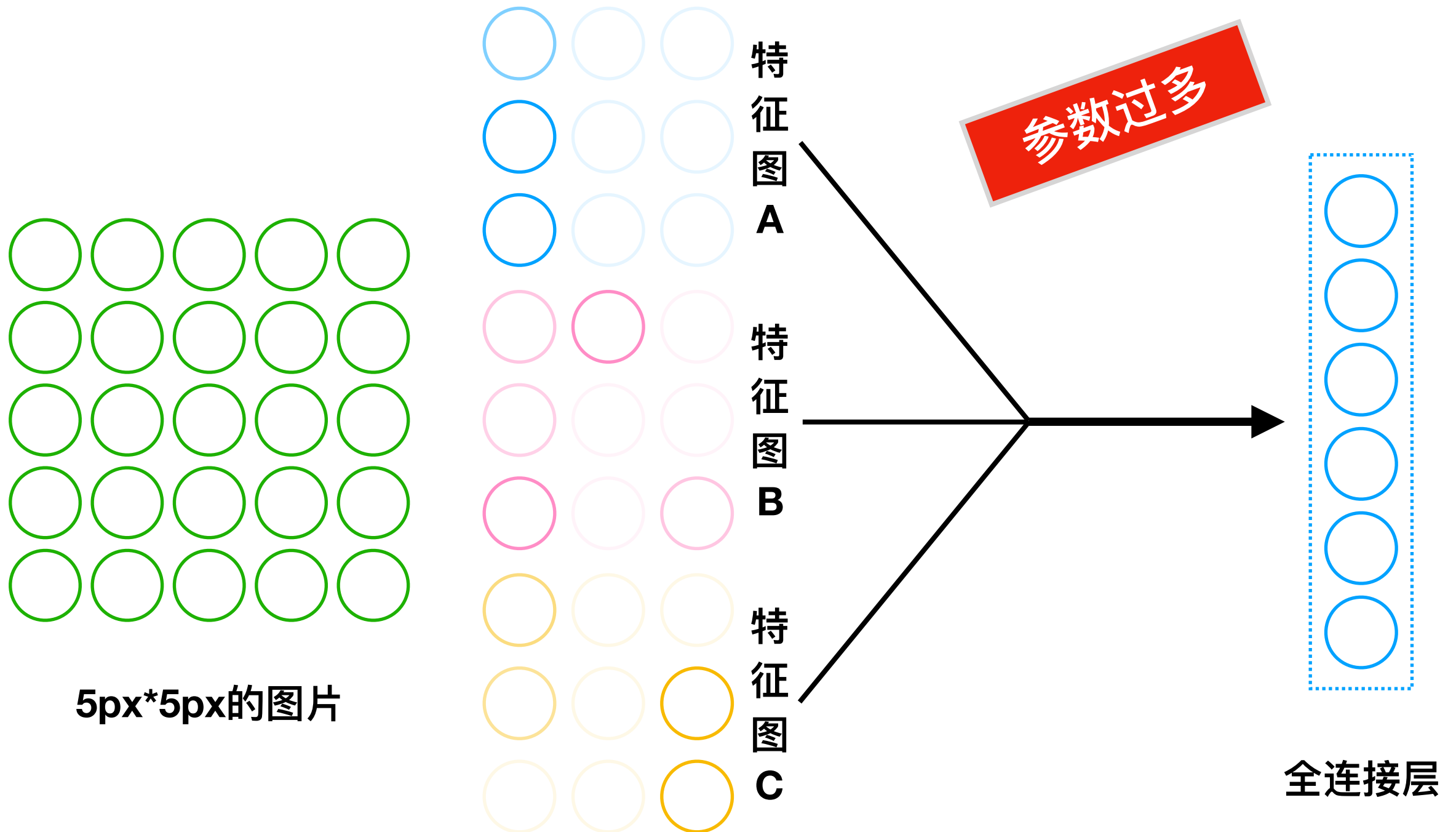


多核卷积



如果总共拥有 n 个微观特征，则需要进行 n 次卷积，得到 n 张特征图。

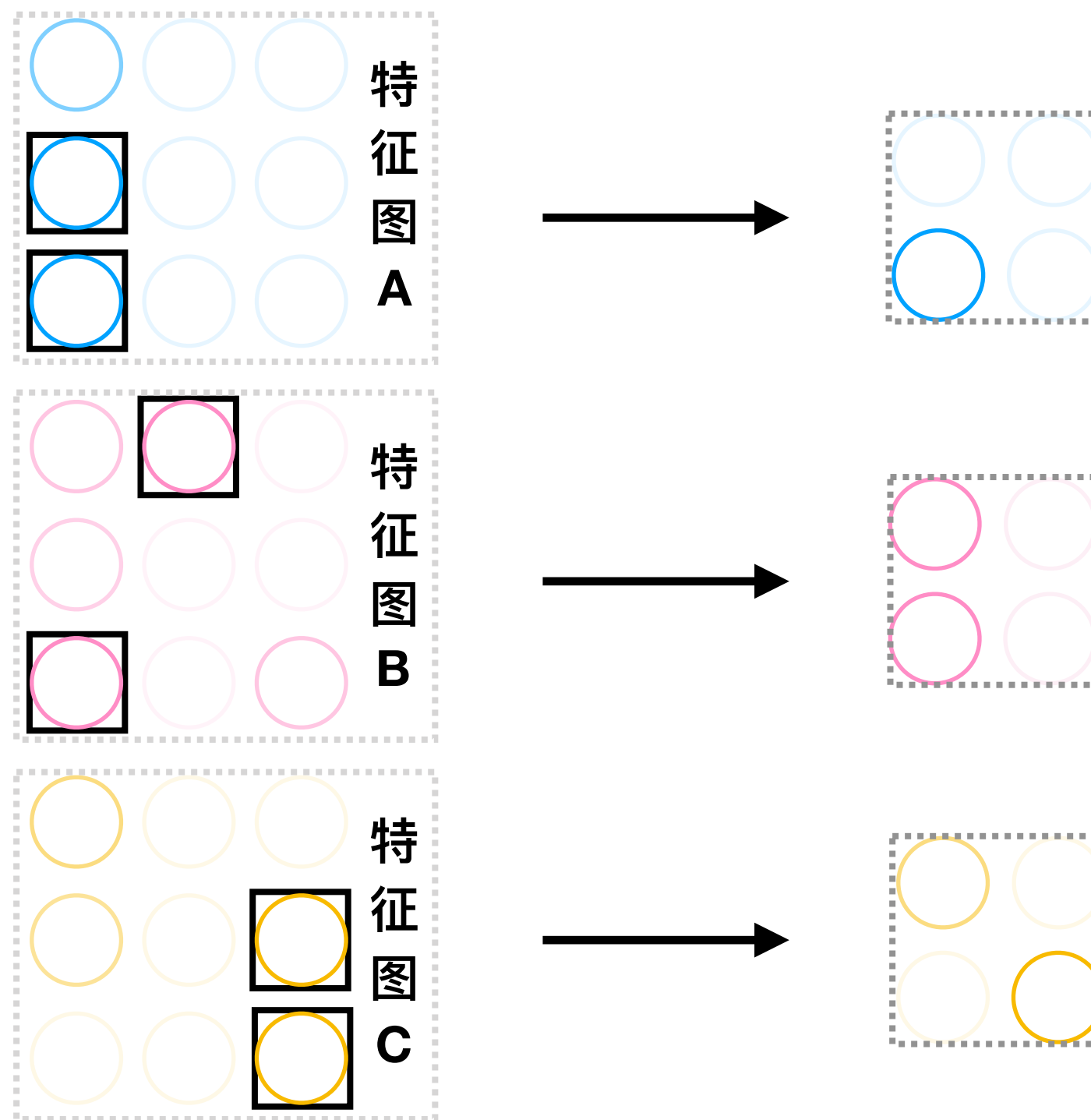
输入多个特征图



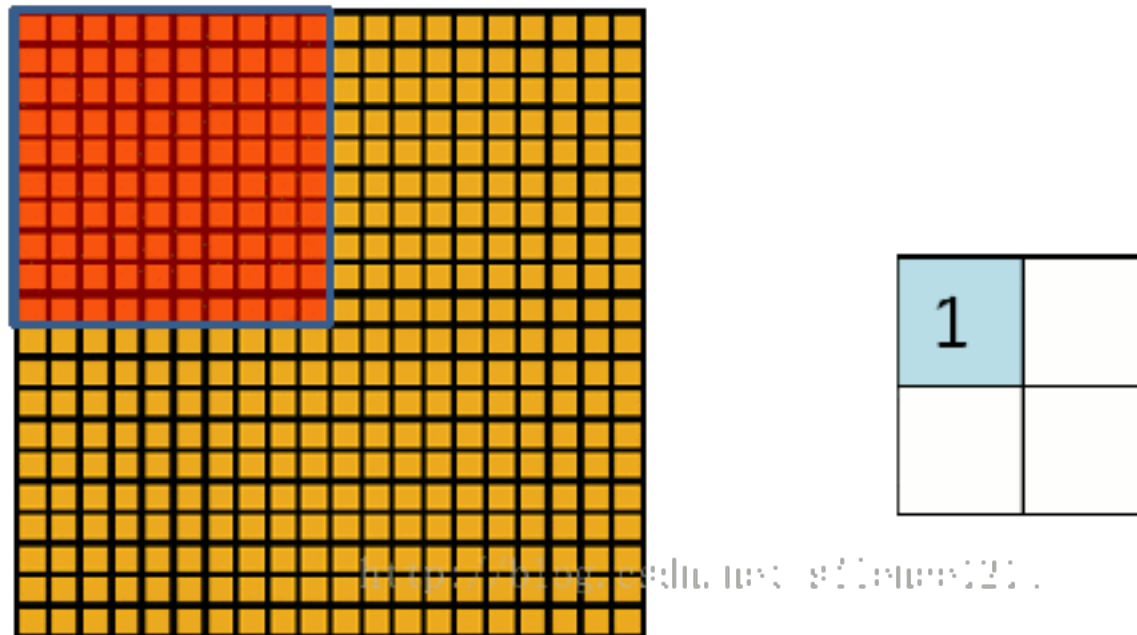
5px*5px的图片

3个卷积核得到的特征图

特征图下采样



最大池化



Convolved
feature

Pooled
feature

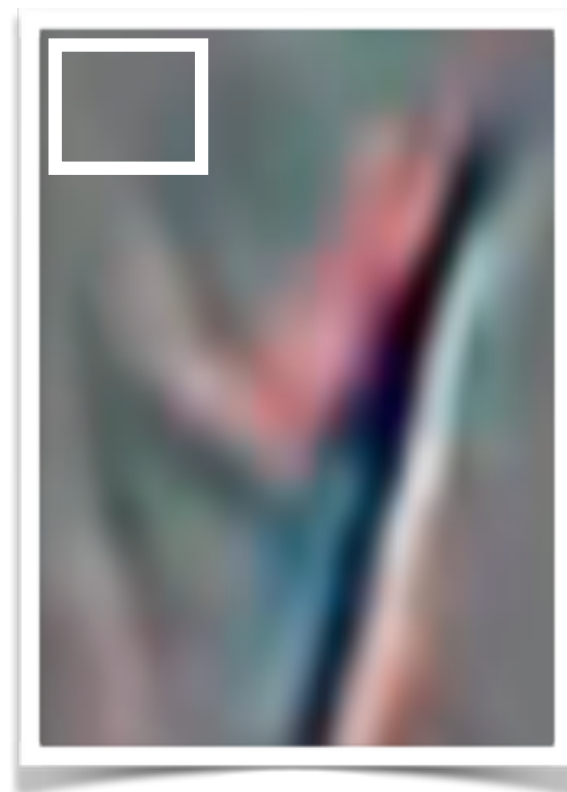
提取局部最大值代表局部范围内的特征激活值。

最大池化

由于特征的稀疏性，一个特征图的局部一般仅存在少量的有效特征，可以使用最大激活值代表当前区域的激活值。

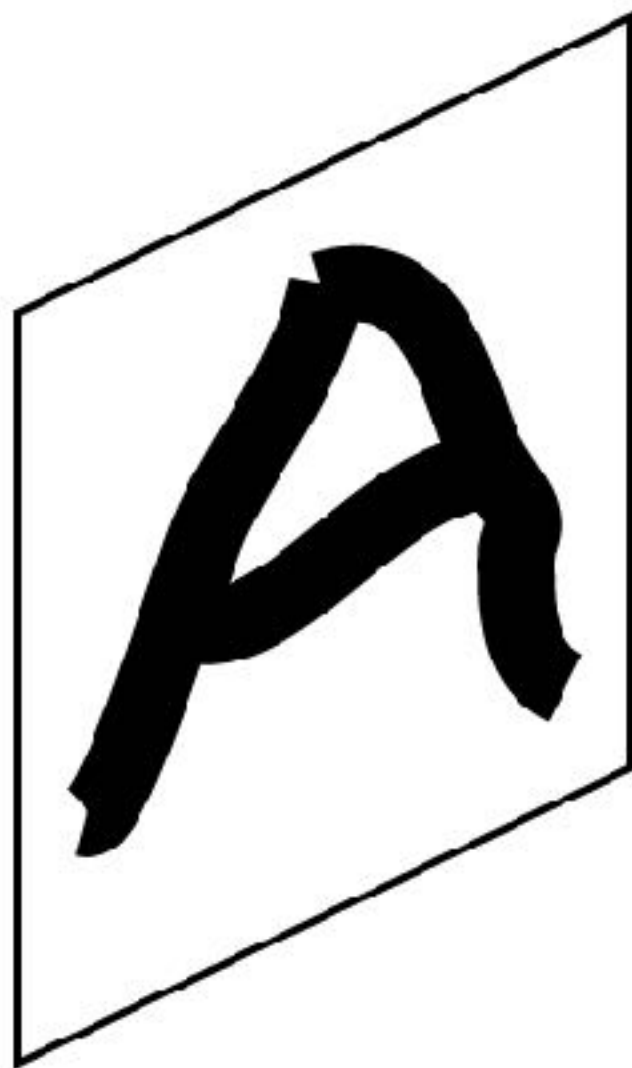


图片

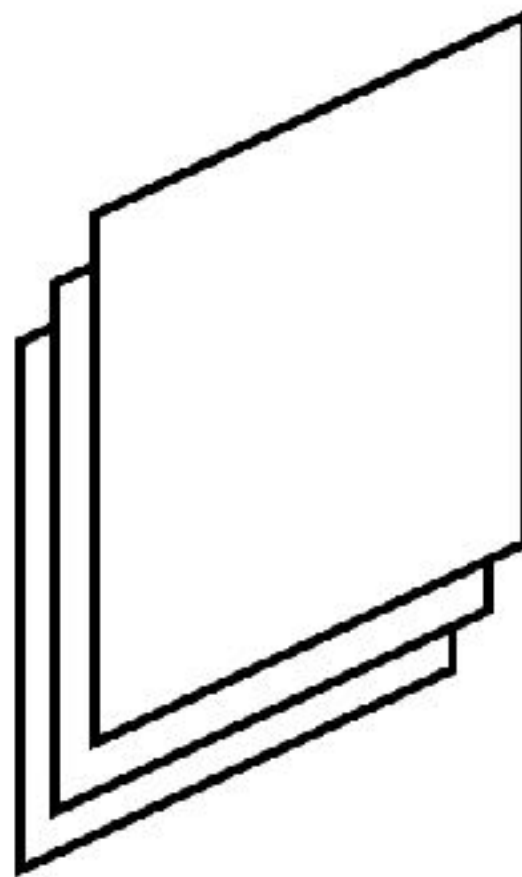


特征图

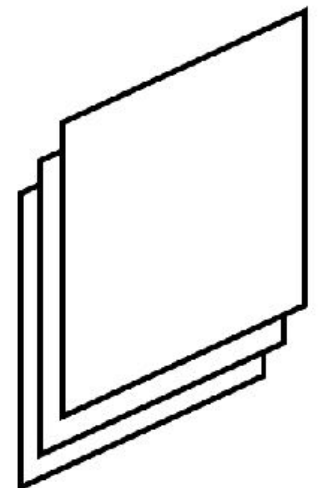
卷积与池化



28*28像素的图像



3个由5*5卷积核得到的
特征图：3*24*24



3个6*6池化核池化后
的特征图：3*4*4

通过卷积与池化，可以有效的提取数据中的特征，并达到降低输入纬度的目的。

小结

- ANN可以用来建立端到端模型，并进行端到端训练。
- 卷积神经网络包含有卷积层与池化层，可以降低输入维度，并有效提取特征。
- 利用数据的局部相关性，可以进行局部特征提取。
- 利用卷积核可以对图像的各个区域进行扫描，可以得到特征分布图。
- 一个卷积核可以得到一个特征图，通常我们需要多个卷积核才能准确提取所有特征。
- 通过下采样可以降低特征图的维度。
- 卷积核与池化核的大小、移动步长等信息不尽相同。

THANKS