



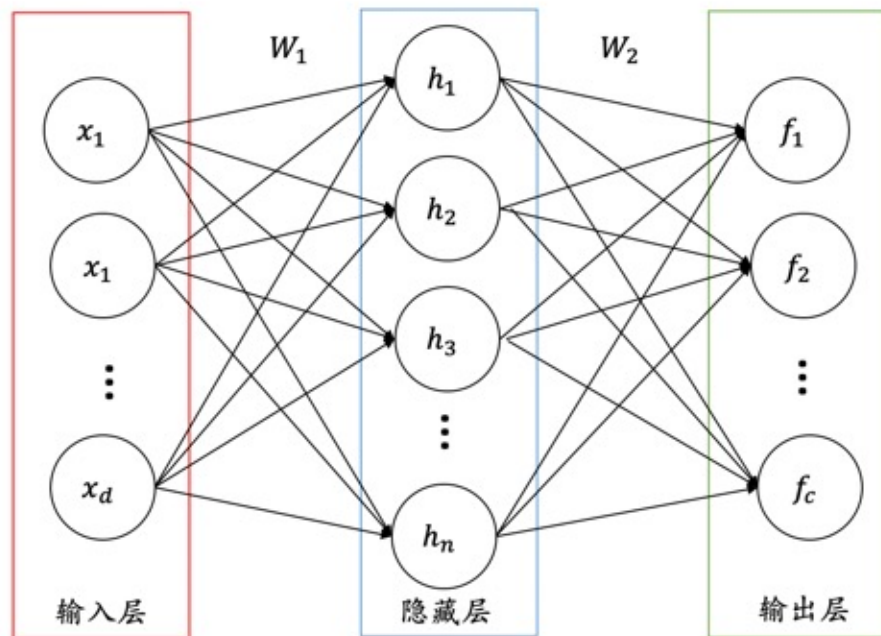
卷积神经网络 (1)

叶山 中国地质大学 (北京)

yes@cugb.edu.cn

全连接神经网络的局限

全连接神经网络的局限

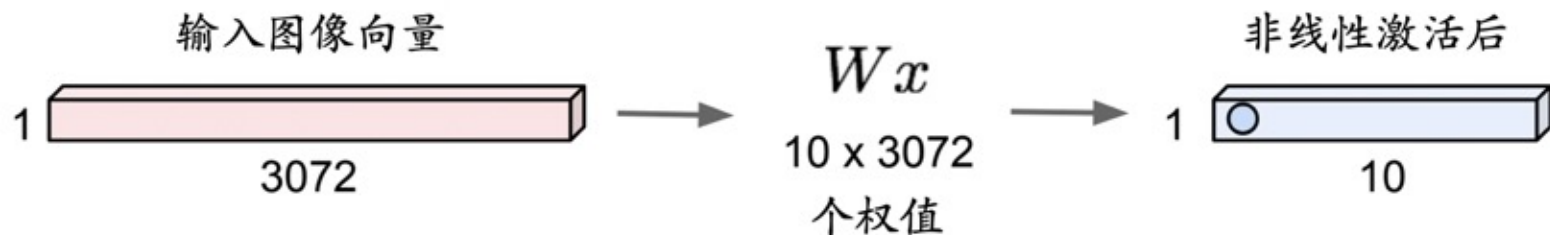


在CIFAR-10图像分类任务中，图片被flatten为长度为3072的向量，即输入层有3072个神经元。

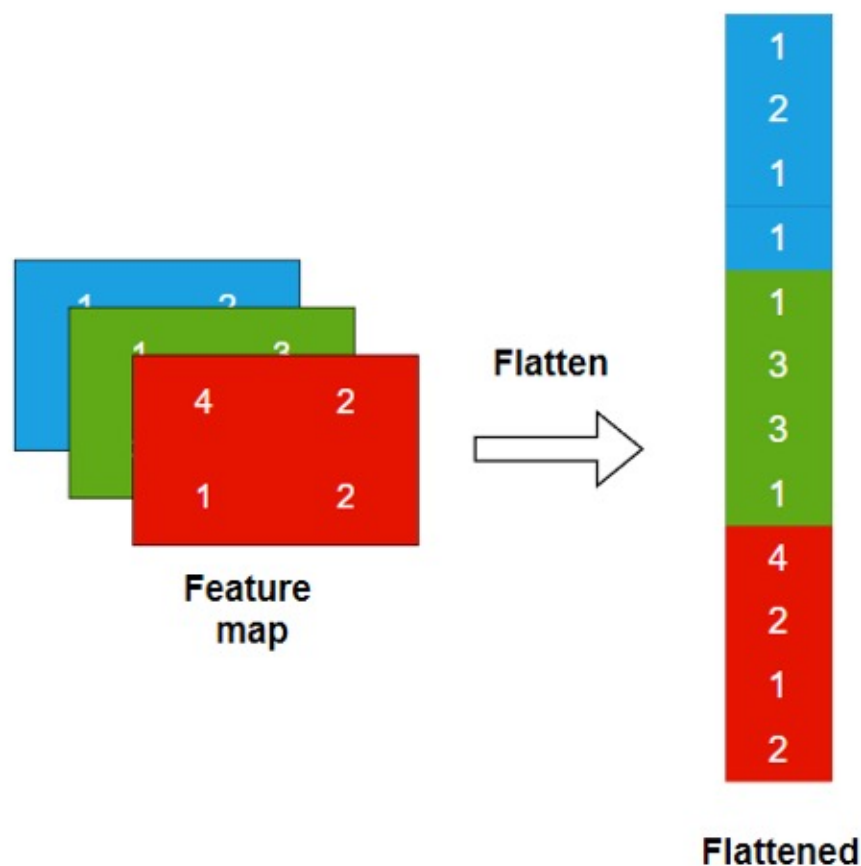
CIFAR-10的类别数为10，因此输出层有10个神经元。

在左图所示的全连接神经网络中，隐藏层的每一个神经元需要学习30720个权值路径。

权值多、计算量大、模型复杂

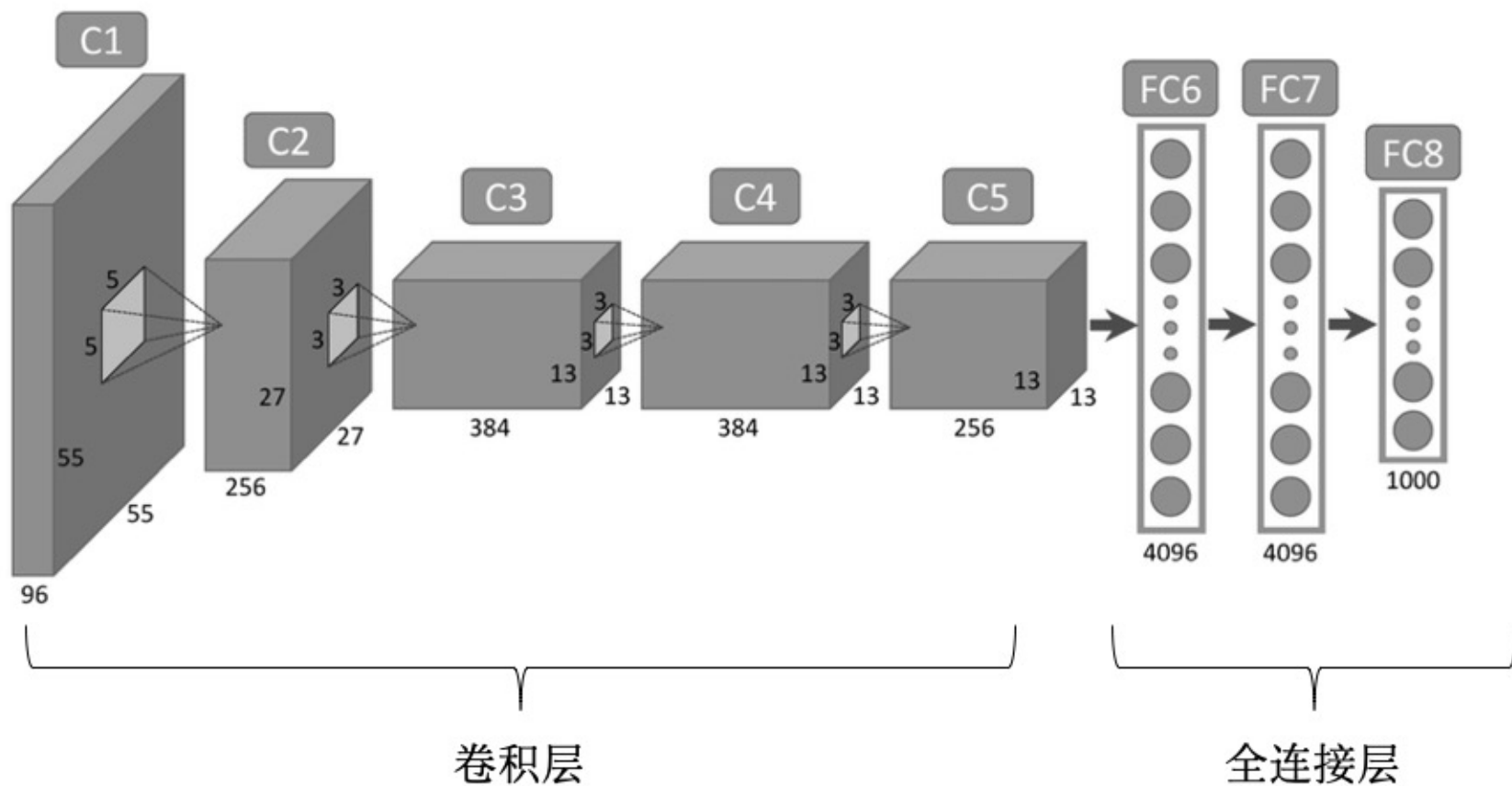


全连接神经网络的局限

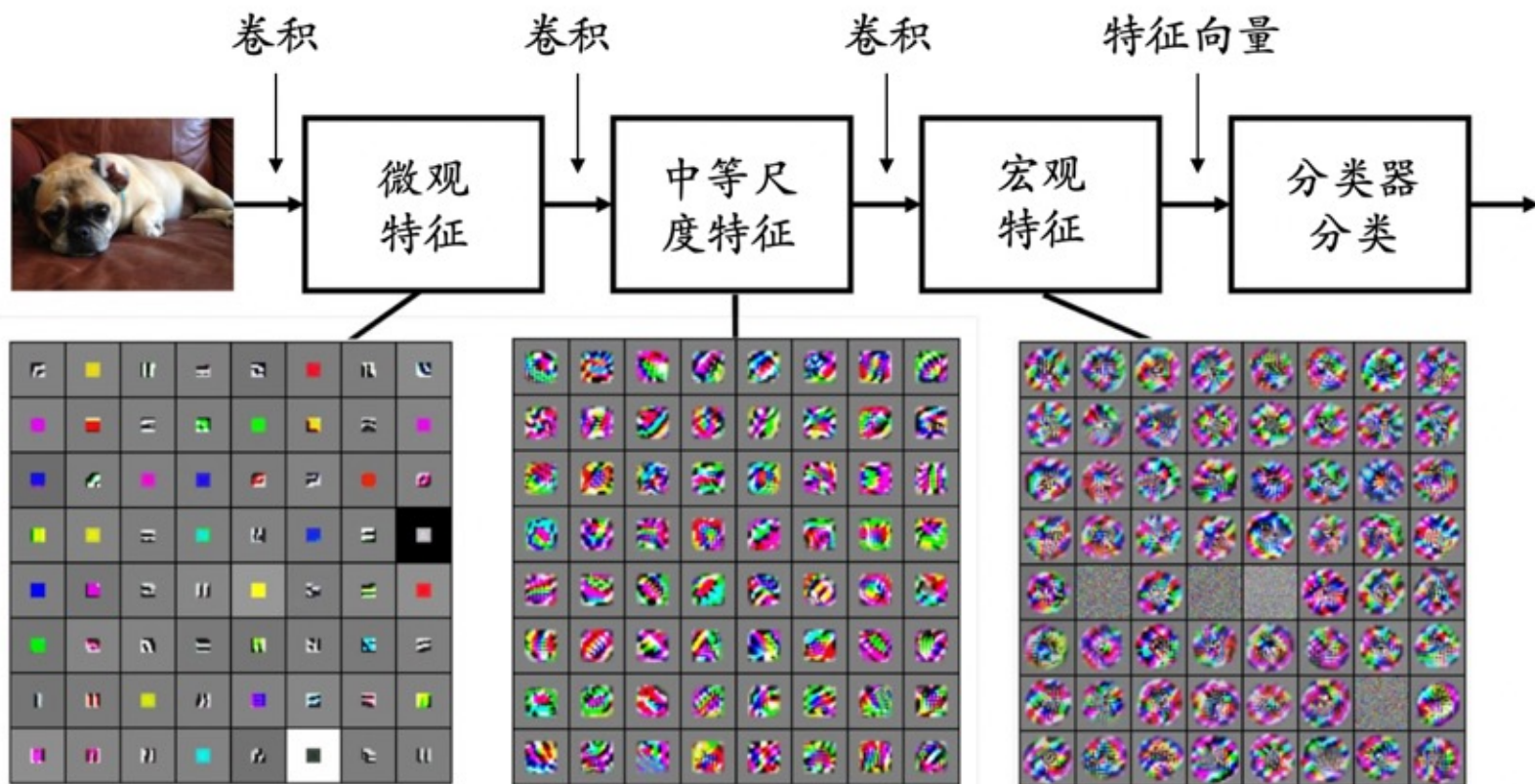


Flatten操作让像素点之间的空间关系信息受损。

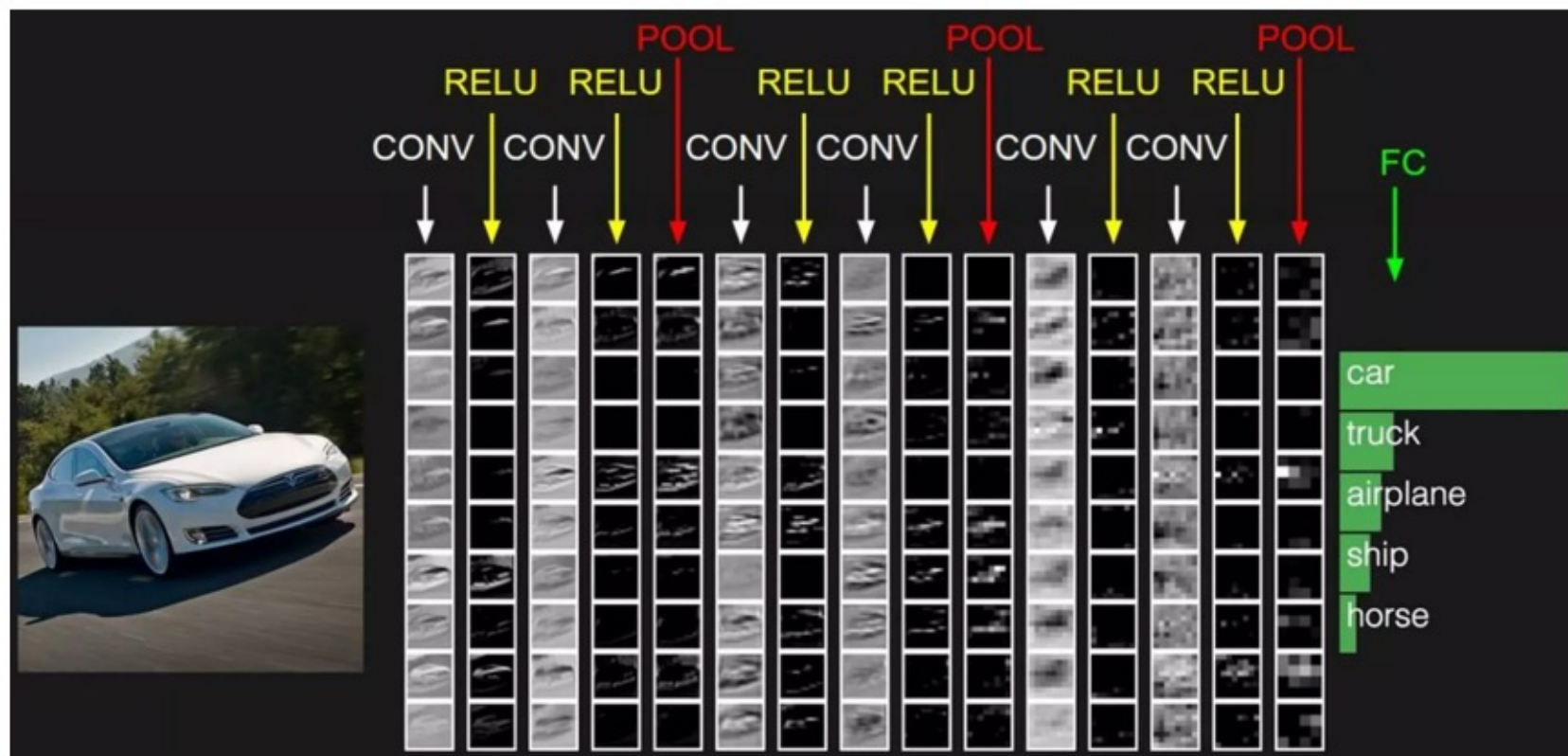
卷积神经网络 CNN



卷积神经网络 CNN



卷积神经网络 CNN



CONV: 卷积层

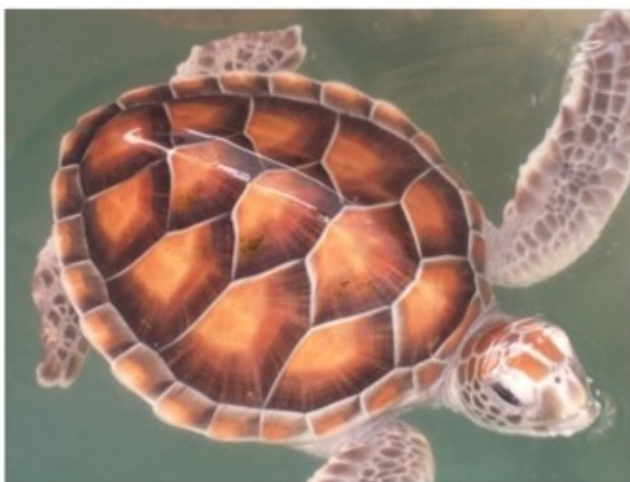
RELU: ReLU层 (激活层)

POOL: 池化层

FC: 全连接层

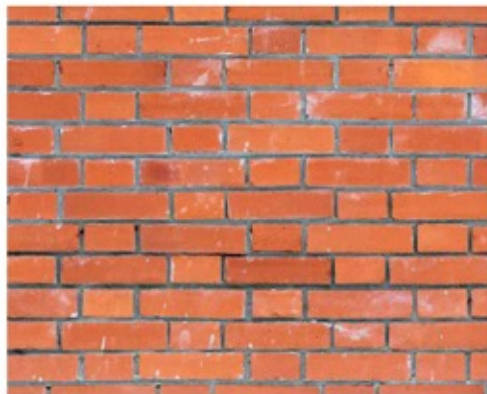
卷积核组

纹理 Texture



纹理 Texture

规则纹理



砖块堆砌

半规则纹理



农田卫星图

随机纹理



纸张褶皱



布料纹路



城市街道



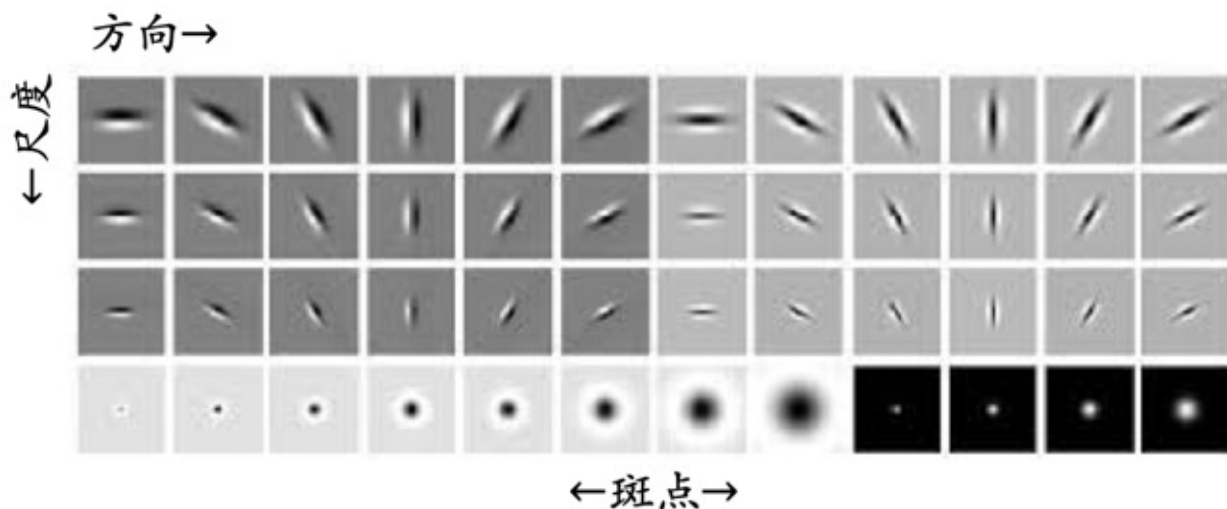
河流流向

基于卷积核组的纹理表示

1. 利用卷积核组提取图像中的纹理基元。

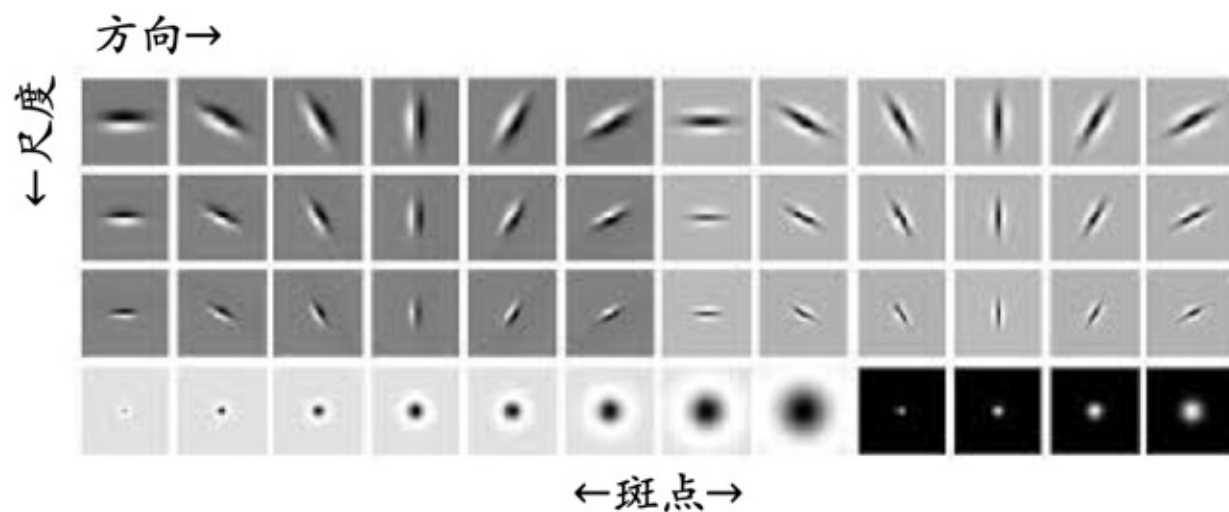
- 纹理基元又称纹理元素，是具有某种色调属性和/或区域属性的像素连续集。它可以通过其平均亮度、最大或最小亮度、尺寸、形状等特征来进行描述。
- 纹理通常是由一些重复的局部子结构（local pattern）组合起来的，这些子结构即纹理基元。

2. 利用纹理基元的统计信息表示图像中的纹理。

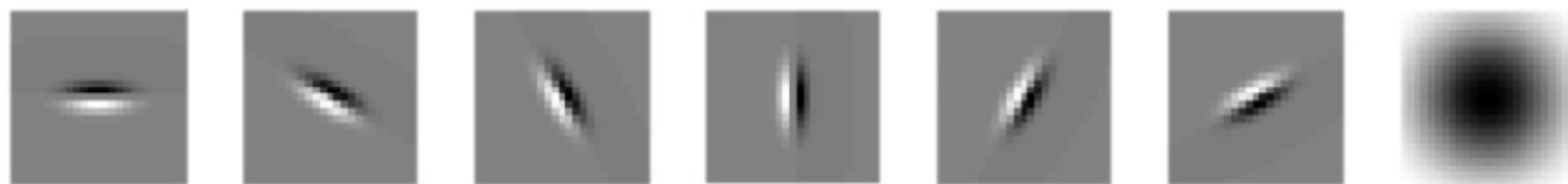


卷积核组（filter bank）：一系列卷积核的集合，通常包含了不同的尺度和方向。

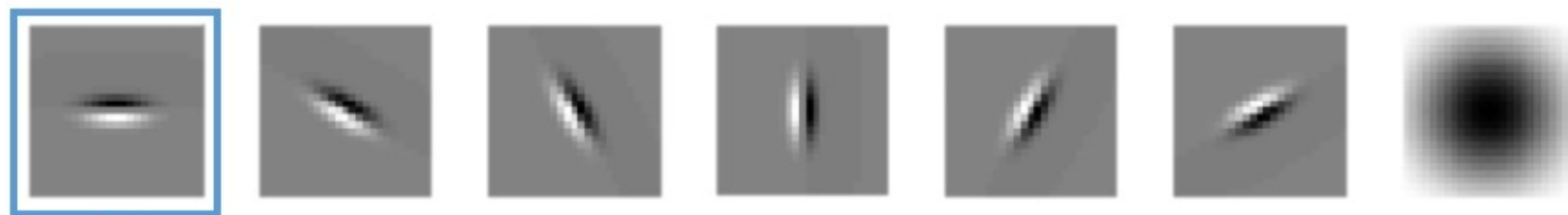
卷积核组



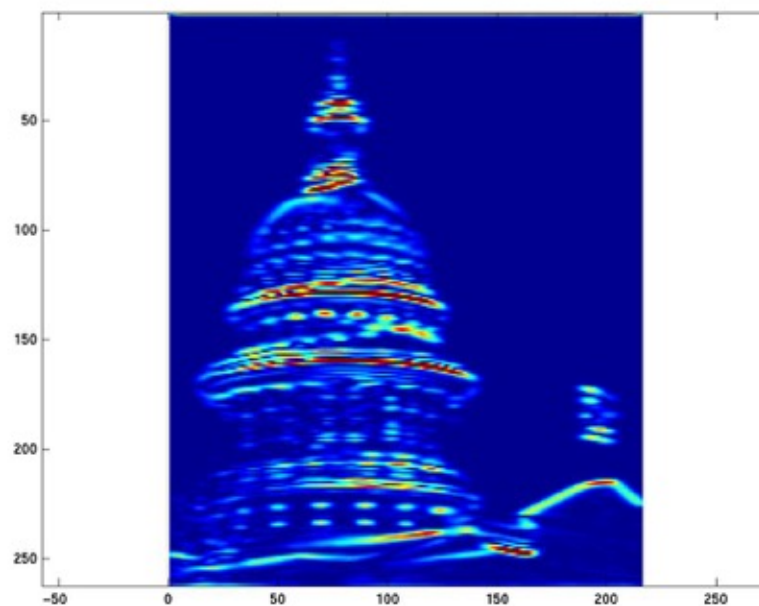
其中7个



卷积核组



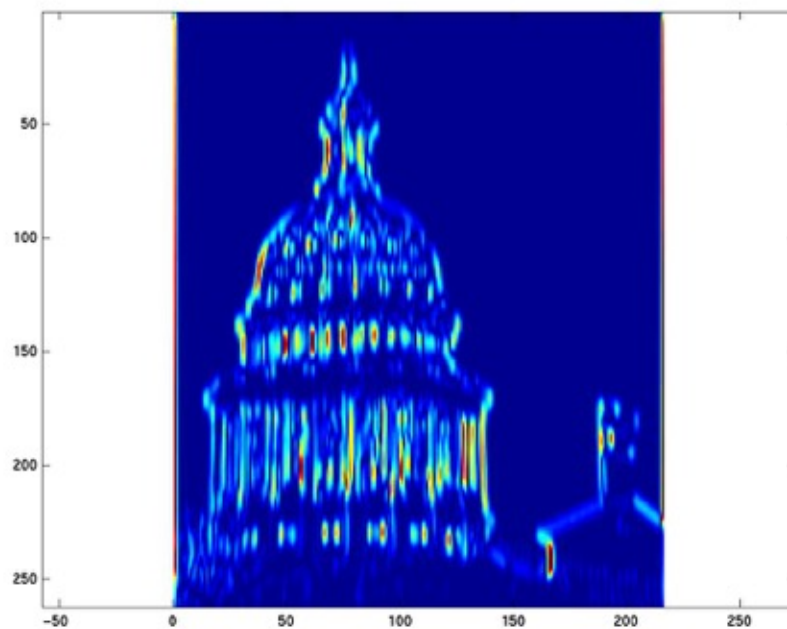
卷积



卷积核组



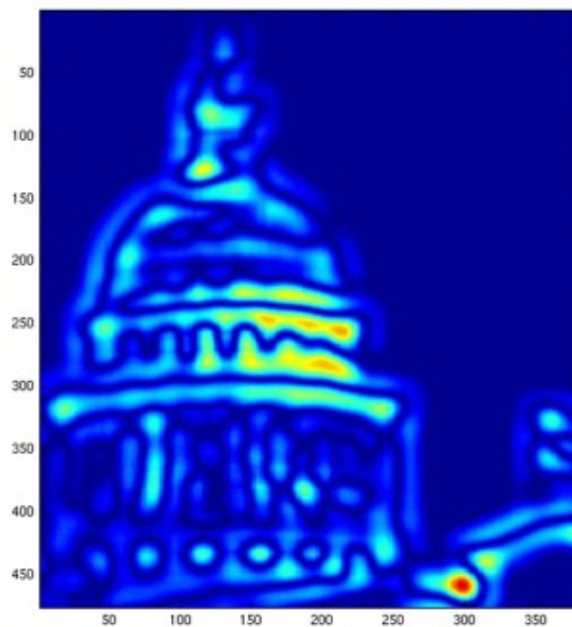
卷积



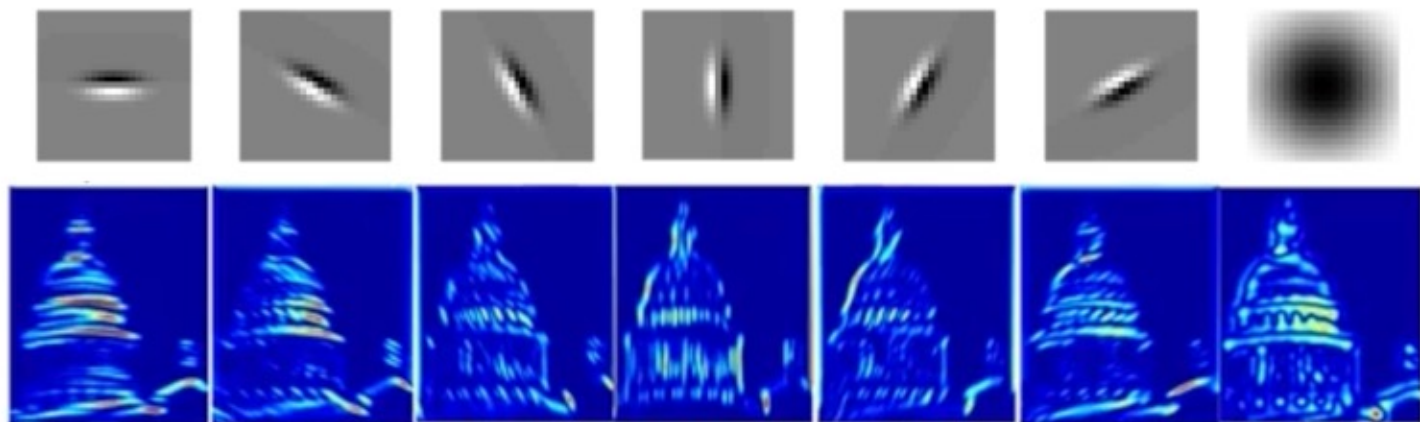
卷积核组



卷积



基于卷积核组的图像表示

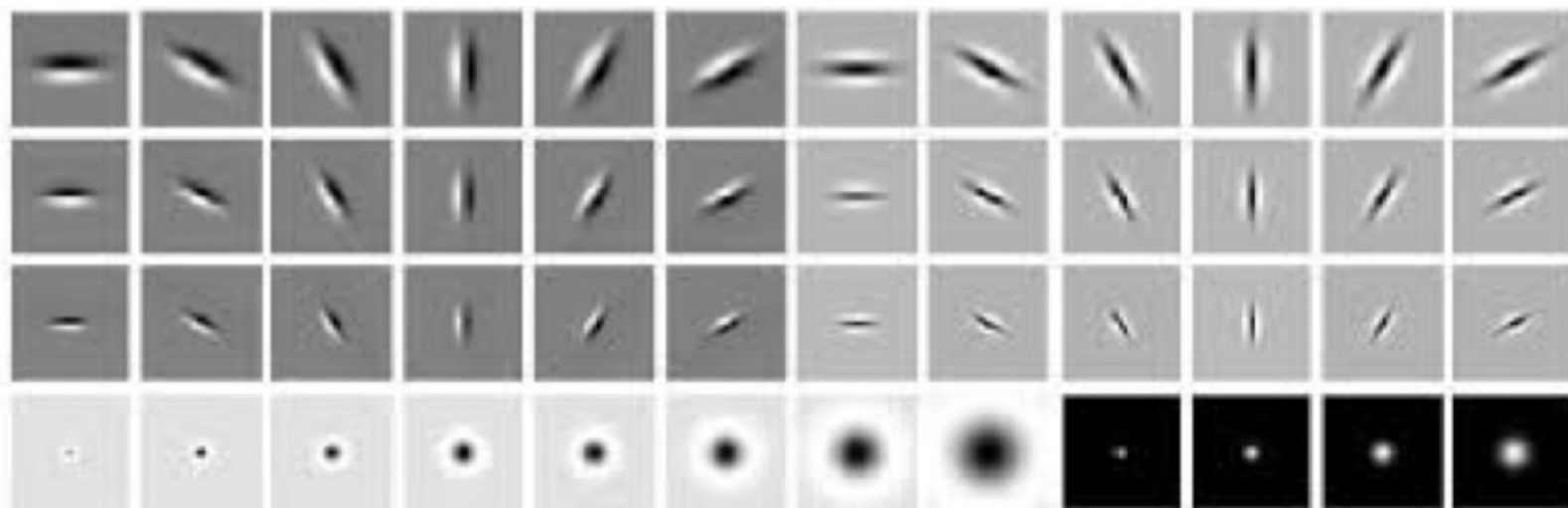


该卷积核组有7个卷积核，因此形成7个特征图。原图中的每一个像素点都会在每一个特征图上形成一个响应值。每个像素可组成特征响应值向量：

$$P_{(x,y)} = [p_1, p_2, p_3, p_4, p_5, p_6, p_7]$$

像素角度

基于卷积核组的图像表示

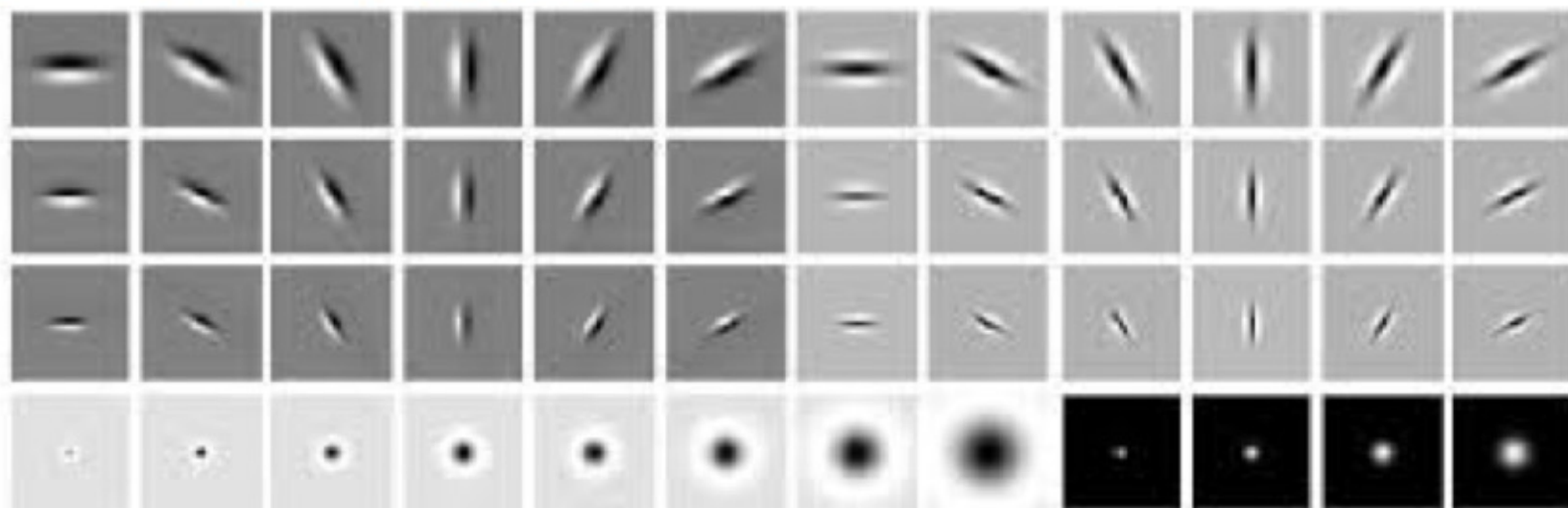


如果使用的是这个卷积核组，
则每个像素点会产生48个响应
值。每个像素可组成特征响
应值向量：

$$P_{(x,y)} = [p_1, p_2, p_3, \dots, p_{48}]$$

像素角度

纹理表示的特点



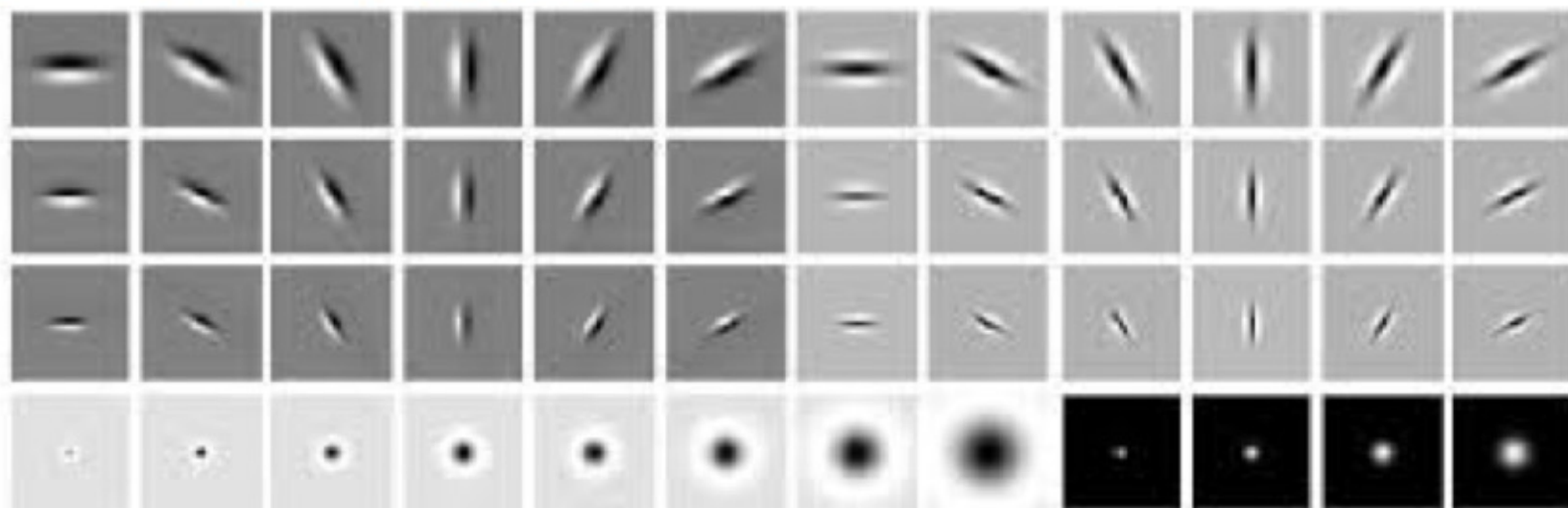
每一个像素点产生48个响应值，用响应值组建一个向量。如果该向量的其中某一个值很大，意味着什么？

该像素所在位置的纹理和响应值最大的这个卷积核所对应的纹理最相近。

$$P_{(x,y)} = [p_1, p_2, p_3, \dots, p_{48}]$$

像素角度

纹理表示的特点



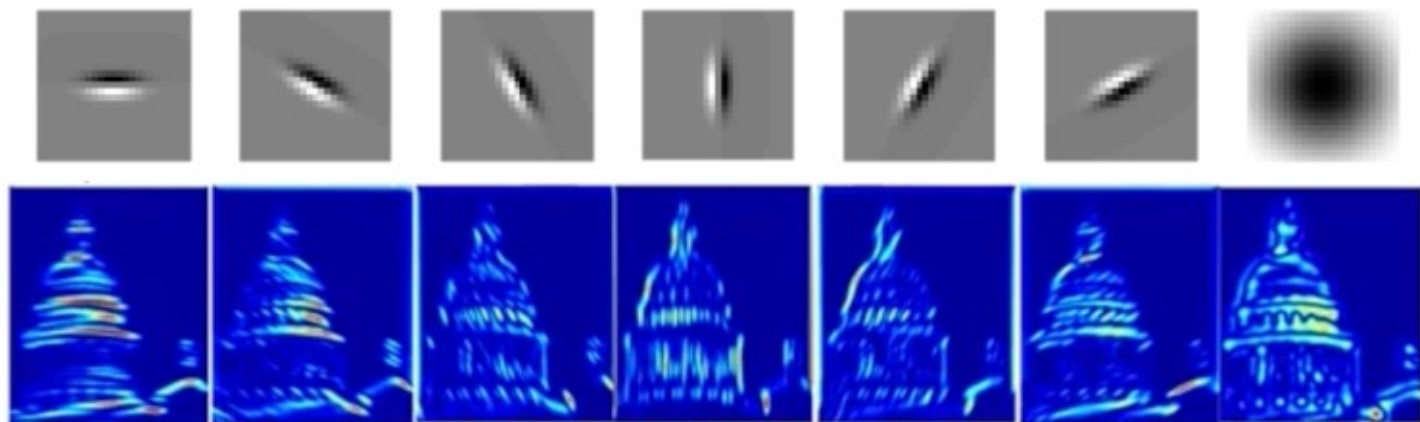
对于同一个像素，这48个响应值是否有可能全都很大？

不可能。该向量应当是一个稀疏向量，只有少数位数（甚至1位）数值较大。

$$P_{(x,y)} = [p_1, p_2, p_3, \dots, p_{48}]$$

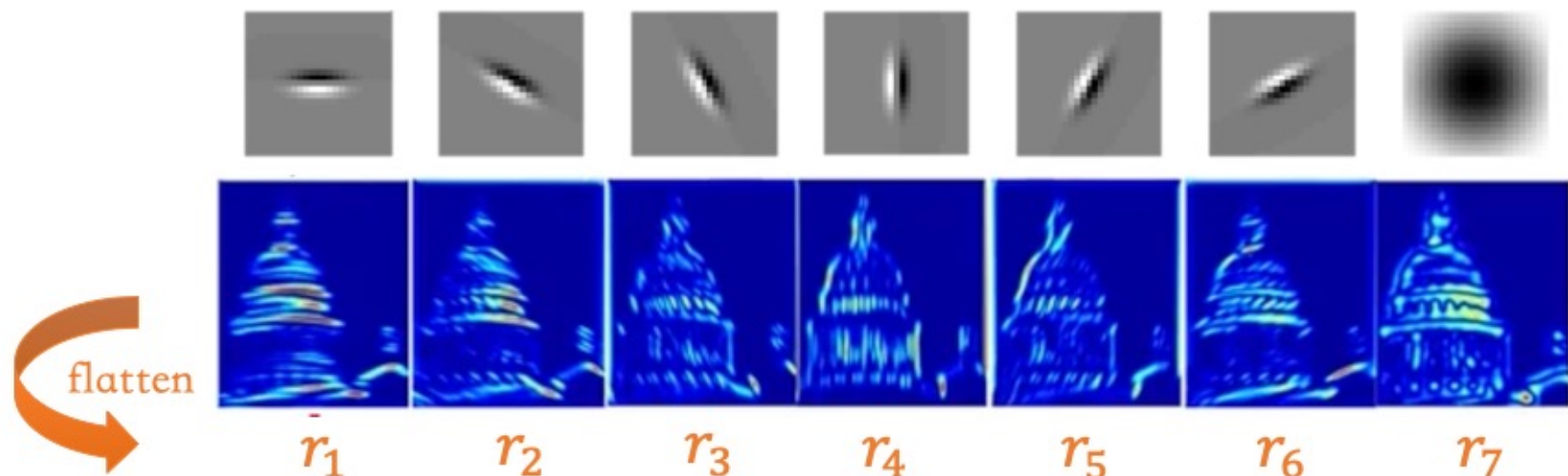
像素角度

基于卷积核组的图像表示



特征响应图的高亮区域，代表图像的相应位置存在和其对应卷积核检测方向相符的纹理基元。

基于卷积核组的图像表示

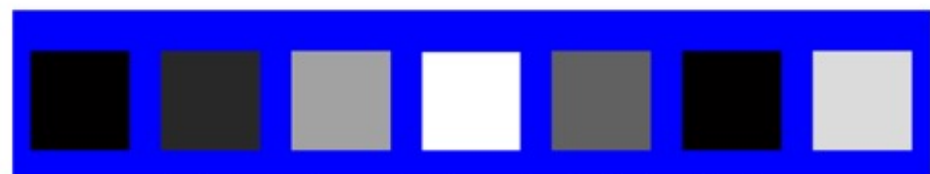
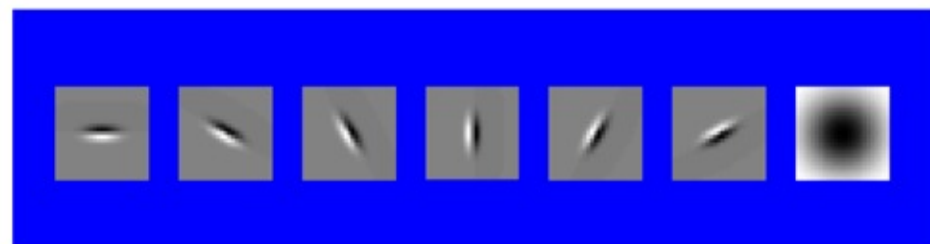


把每一个特征响应图做flatten操作，展开成向量。再把这些向量拼接起来，作为图像的表达。

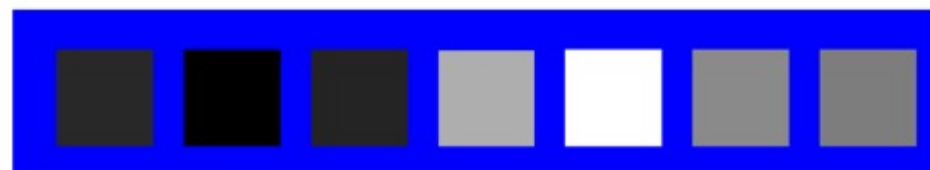
$$R = [r_1, r_2, r_3, r_4, r_5, r_6, r_7]$$

特征图角度

基于卷积核组的图像表示



第4个特征图最亮：竖直特征较多。



第5个特征图最亮：右上-左下的特征较多。



第7个特征图最亮：斑点状特征较多。

\bar{r}_1 \bar{r}_2 \bar{r}_3 \bar{r}_4 \bar{r}_5 \bar{r}_6 \bar{r}_7

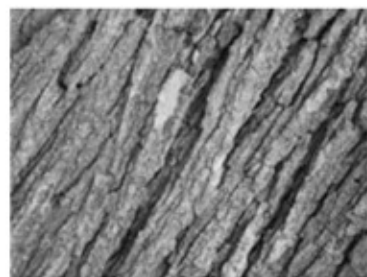
A

B

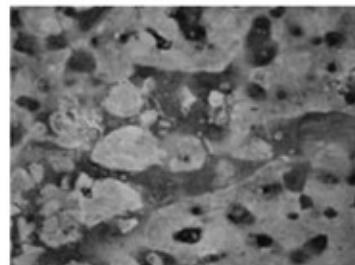
C



1



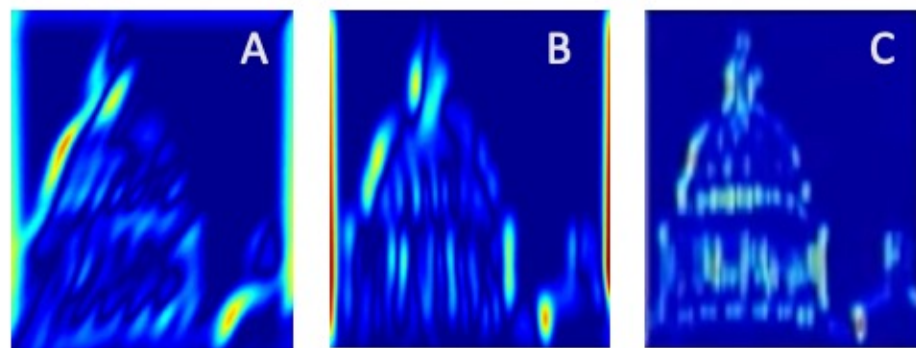
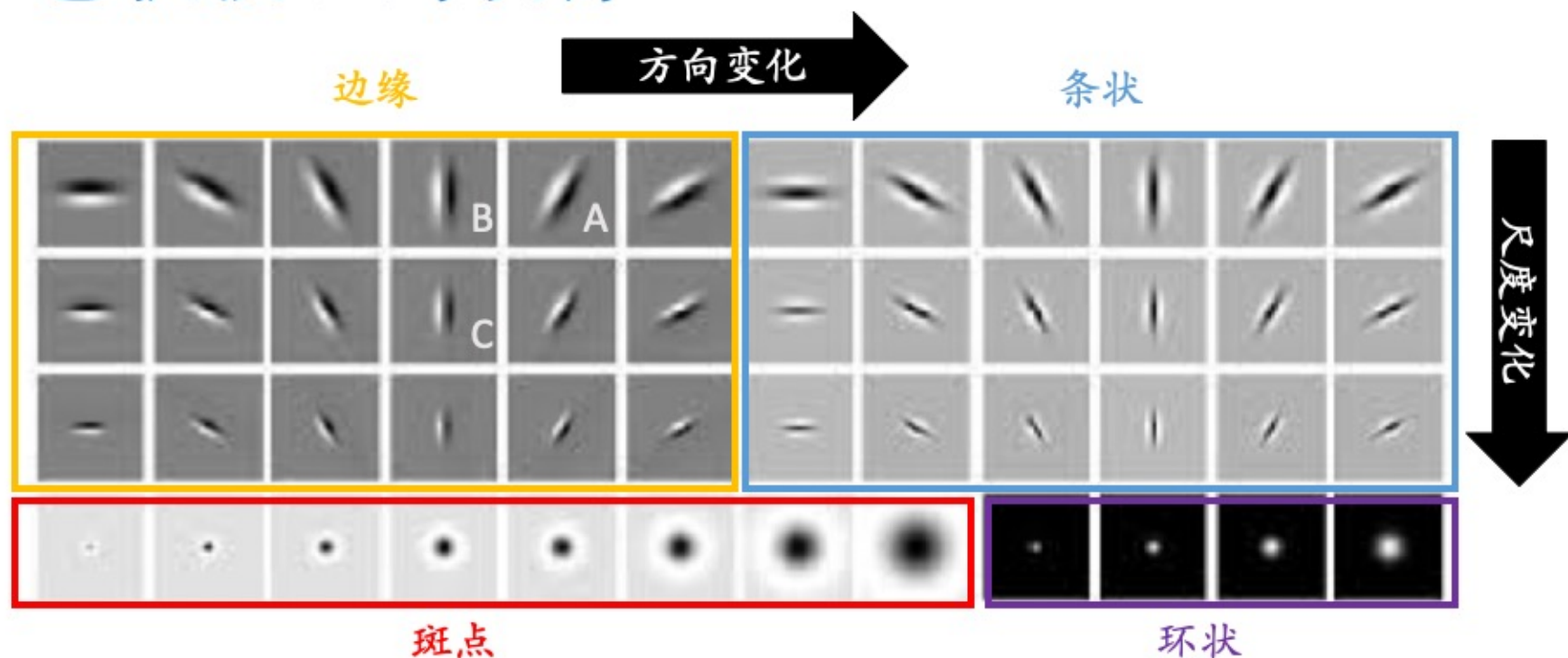
2



3

规律多于噪音
忽略纹理基元的具体位置

卷积核组的设计



$$R = [\bar{r}_1, \bar{r}_2, \bar{r}_3 \cdots \bar{r}_{48}]$$

卷积核组的特征
响应图均值向量

综合考虑三个需求：特征类型、尺度、方向

纹理表示的步骤

1. 设计卷积核组。
2. 对图像实施卷积操作，描述图像中每个位置存在的纹理基元。
3. 用特征响应图（或其统计值，比如均值）表示图像中的纹理特性。

Pop Quiz

1. 以下哪种激活函数通常用于输出层?

A. Softmax | B. Sigmoid | C. ReLU | D. Tanh

2. 衡量模型预测值与真实值之间差距大小的函数被称为_____?

A. 卷积函数 | B. 损失函数 | C. 正则化函数 | D. 优化函数

3. 去掉激活函数的全连接神经网络相当于一个_____?

A. 神经元 | B. 权值矩阵 | C. 线性分类器 | D. 决策树

4. 训练集中的所有样本都被训练过一次, 被称为一个_____?

A. 批次 (batch) | B. 迭代 (iteration) | C. 轮次 (epoch) | D. 优化 (optimization)

5. 在信息论领域, (信息) 熵越大, 表示_____?

A. 信息的不确定性越大 | B. 信息的不确定性越小

C. 数据的取值范围越大 | D. 数据的取值范围越小

