深度学习平台与应用

第五讲：卷积神经网络

范琦

fanqi@nju.edu.cn

2024年10月9日

大　　纲

- 卷积神经网络-引言
- 卷积神经网络

课程回顾

南京大学 NANJING UNIVERSITY

- 图像分类

假设有一个标签集合
{人，狗，猫，汽车，…}

模型的分类预测 → 猫

课程回顾：线性分类器

输入图像

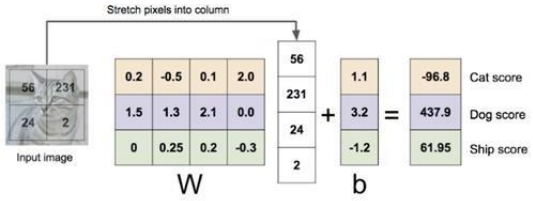
$$f(x,W) = Wx + b$$

$$f(x,W) \longrightarrow$$

571.3	猫
1.25	狗
-132.2	船

Algebraic Viewpoint

$$f(x,W) = Wx$$

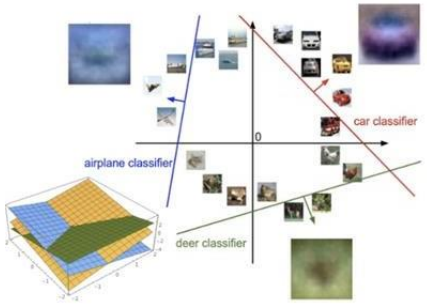


Visual Viewpoint

One template per class

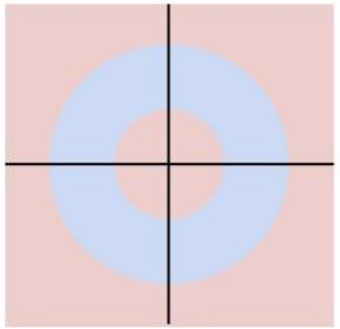
Geometric Viewpoint

Hyperplanes cutting up space



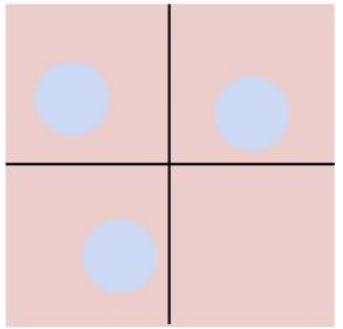
Class 1:
1 <= L2 norm <= 2

Class 2:
Everything else



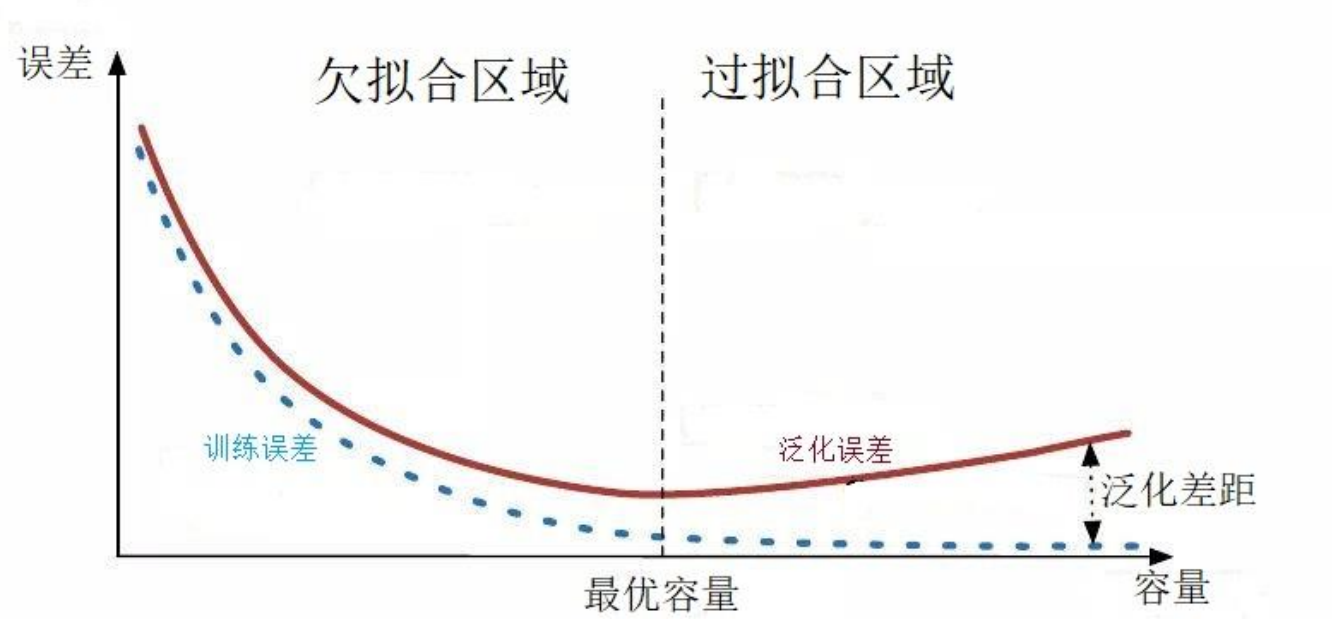
Class 1:
Three modes

Class 2:
Everything else



课程回顾：损失函数与优化

- 我们有训练数据：$\{(x_i, y_i)\}_{i=1}^{N}$
- 我们有线性分类预测：$s = f(x; W) = Wx$
- 我们有损失函数：$L = \frac{1}{N}\sum_{i=1}^{N} L_i + R(W)$
- 优化模型来降低模型在数据上的损失：SGD, SGD+Momentum, RMSProp, Adam



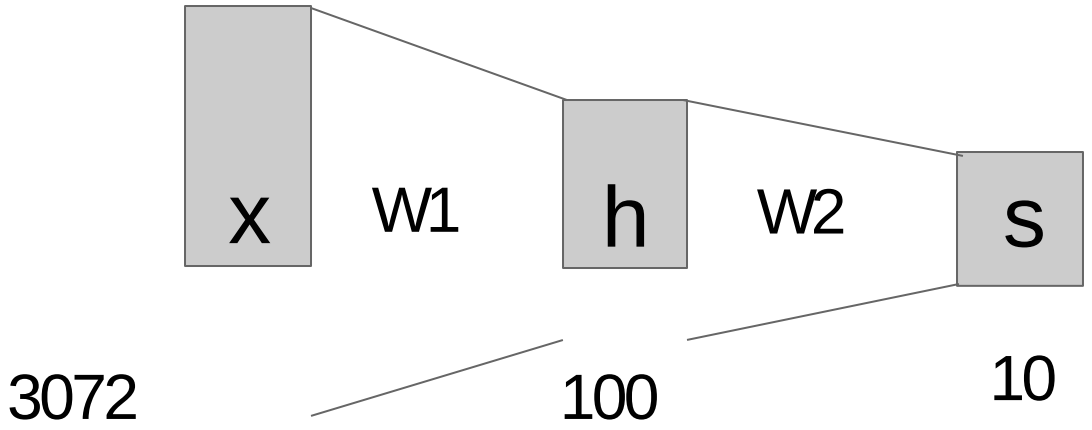
课程回顾：神经网络

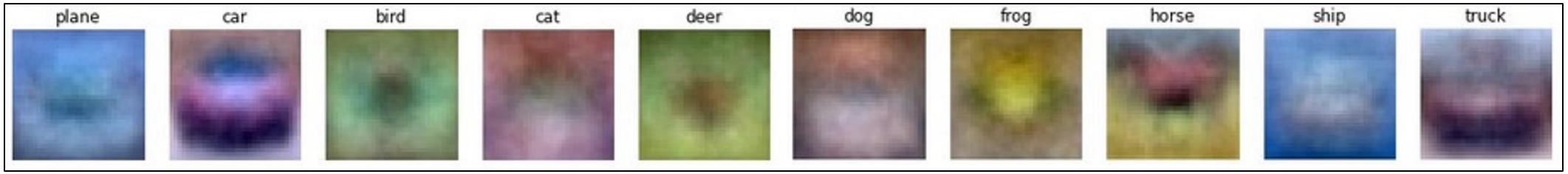
- 线性分类器

$$f = Wx$$

- 2 层神经网络

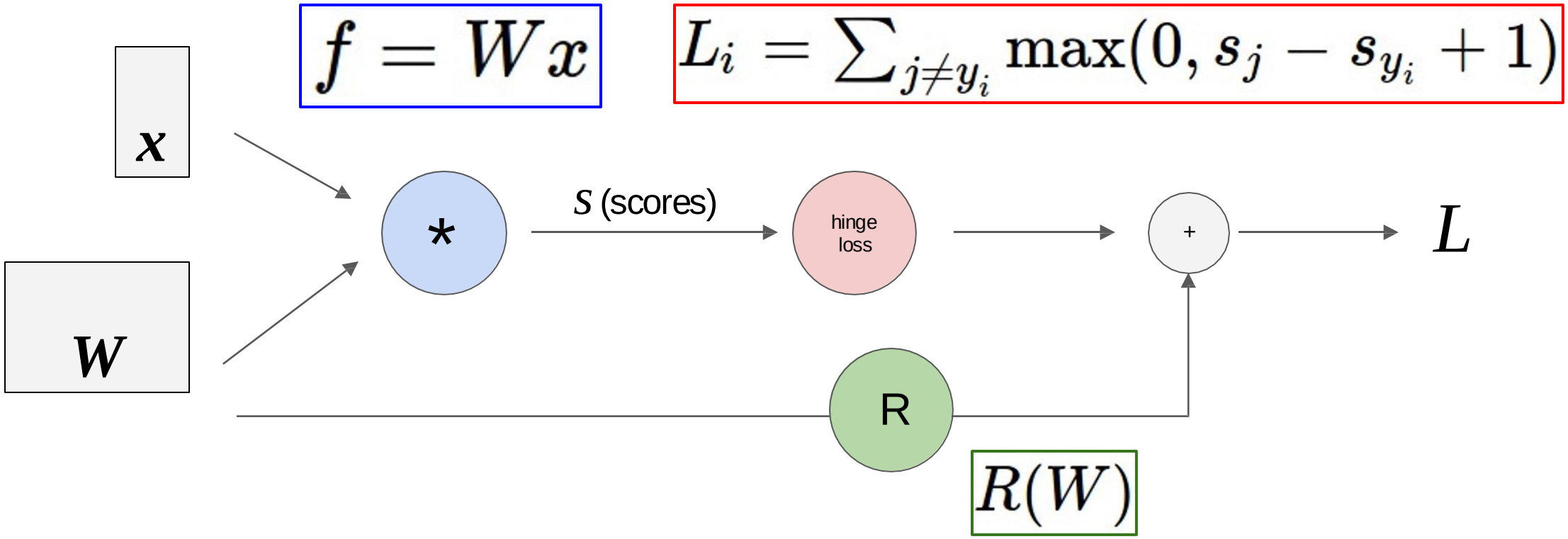
$$f = W_2 \max(0, W_1 x)$$





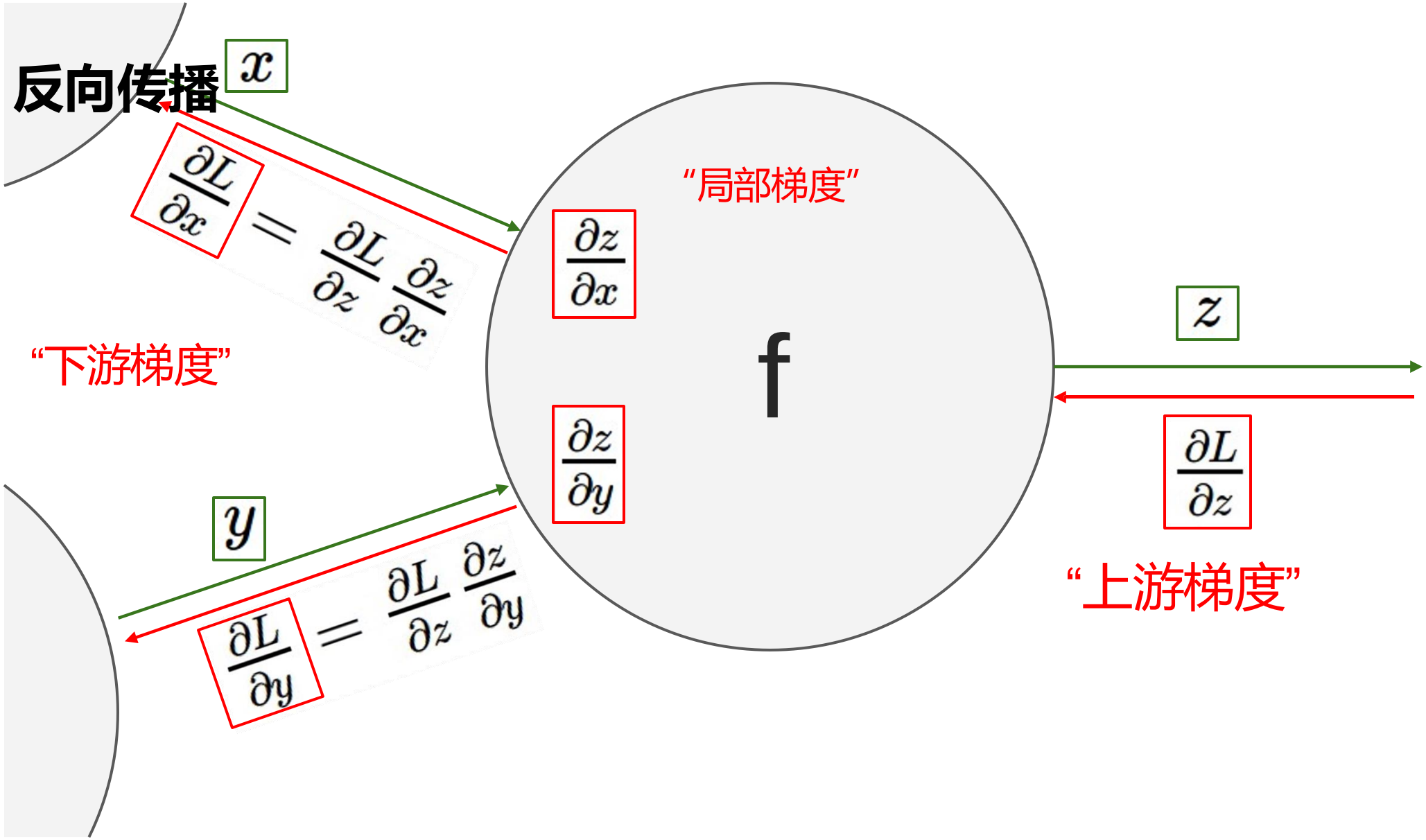
课程回顾：神经网络

- 计算图+反向传播



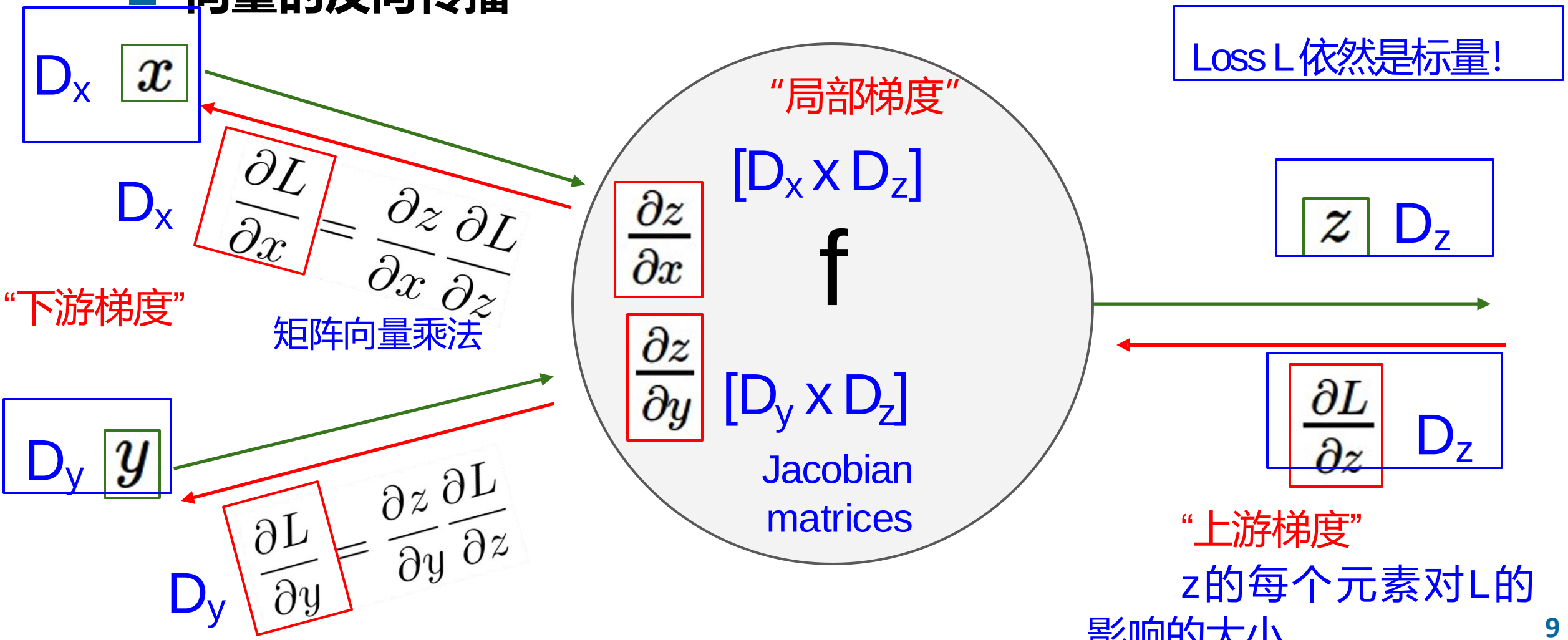
课程回顾：神经网络

- 反向传播



课程回顾：神经网络

向量的反向传播

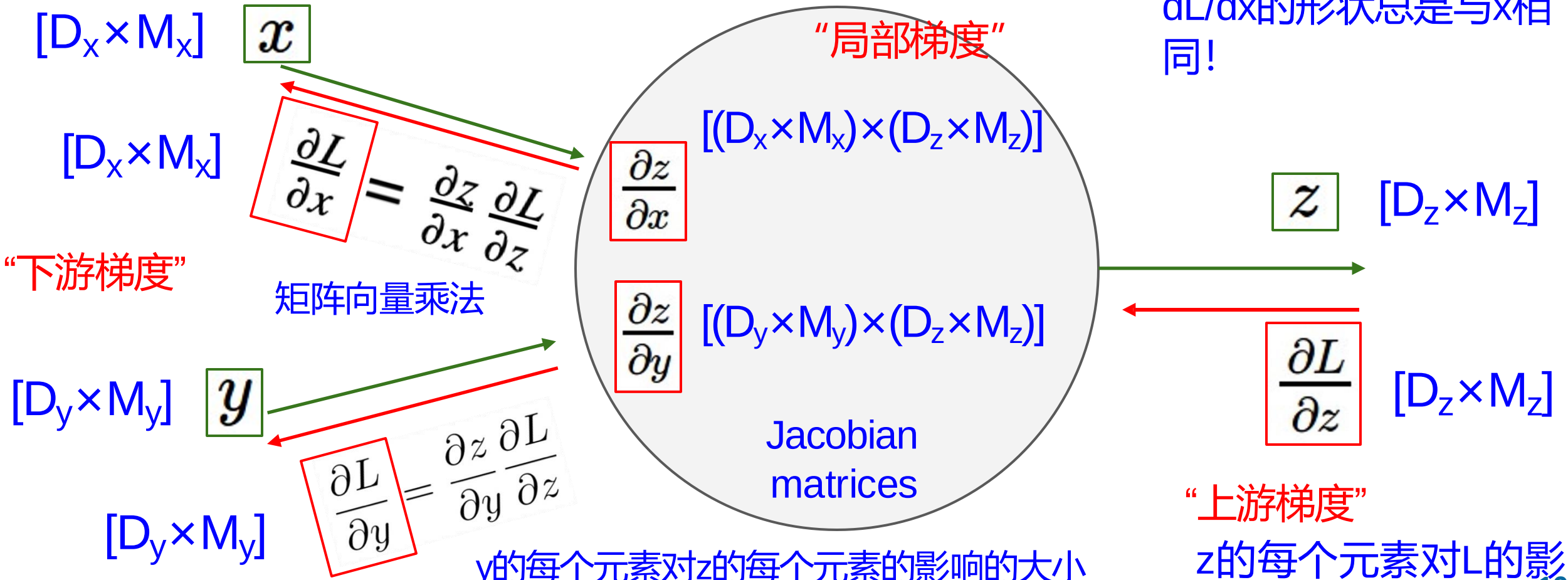


课程回顾：神经网络

- **矩阵的反向传播**

Loss L 依然是标量！

dL/dx的形状总是与x相同！



图像分类

- 图像分类是**最核心的计算机视觉任务**

假设有一个标签集合
{人，狗，猫，汽车，…}

模型的分类预测

Class Scores

图像分类

- **基于图像像素**的图像分类

假设有一个标签集合
{飞机，狗，猫，汽车，…}

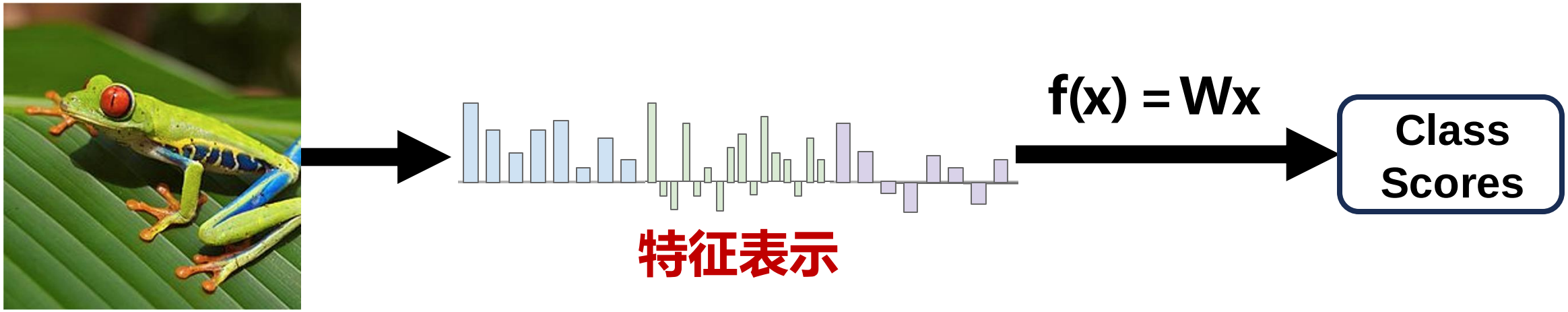
$f(x) = Wx$

Class Scores

图像分类

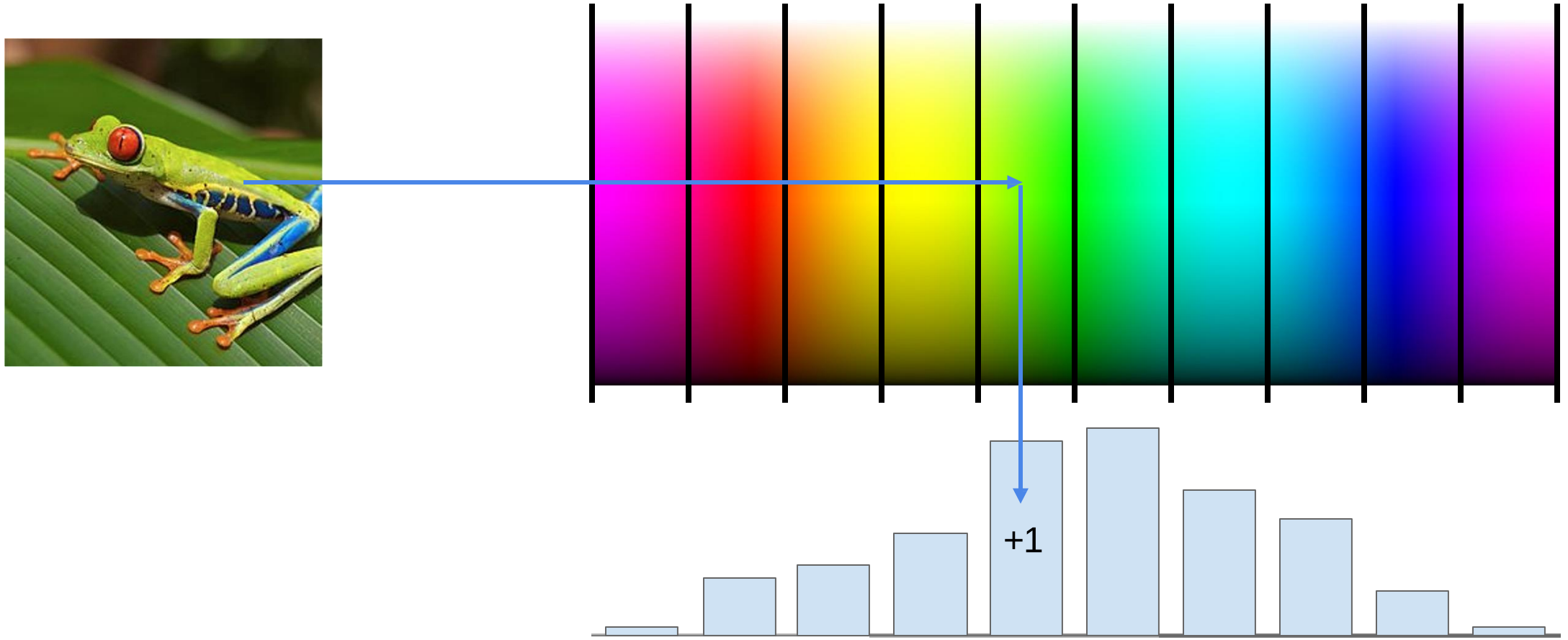
- **基于图像特征**的图像分类

假设有一个标签集合
{飞机，狗，猫，汽车，…}



图像分类

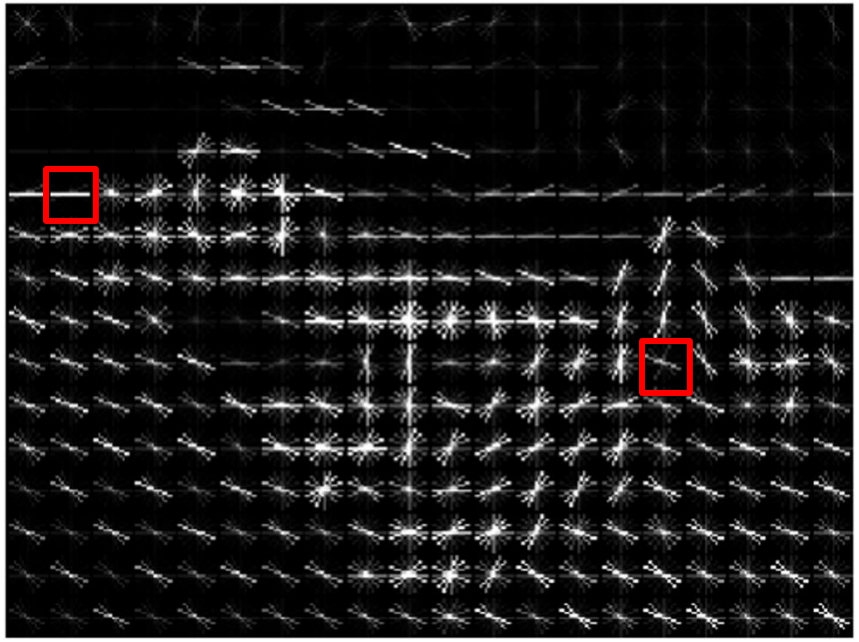
- 特征提取：颜色直方图（Color Histogram）



图像分类

- 方向梯度直方图（Histogram of Oriented Gradients, HoG）

HoG：将图像以8x8为基础区域进行划分。在每个基础区域内，将边缘方向量化为9个数字



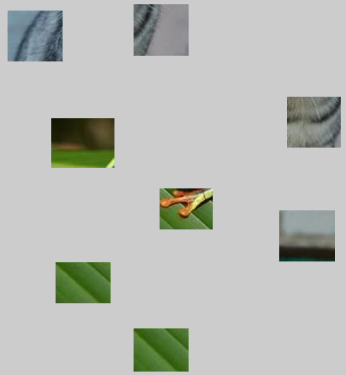
例子：320x240的图像被划分为40x30的区域；每个区域中有9个数字，因此特征向量有30*40*9=10800个数字

南京大学 NANJING UNIVERSITY

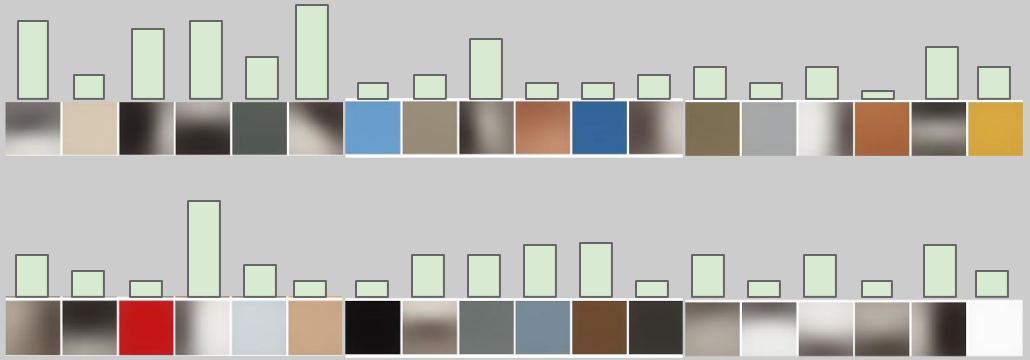
词袋（Bag of Words）

Step 1: 构建 codebook

提取随机图像块



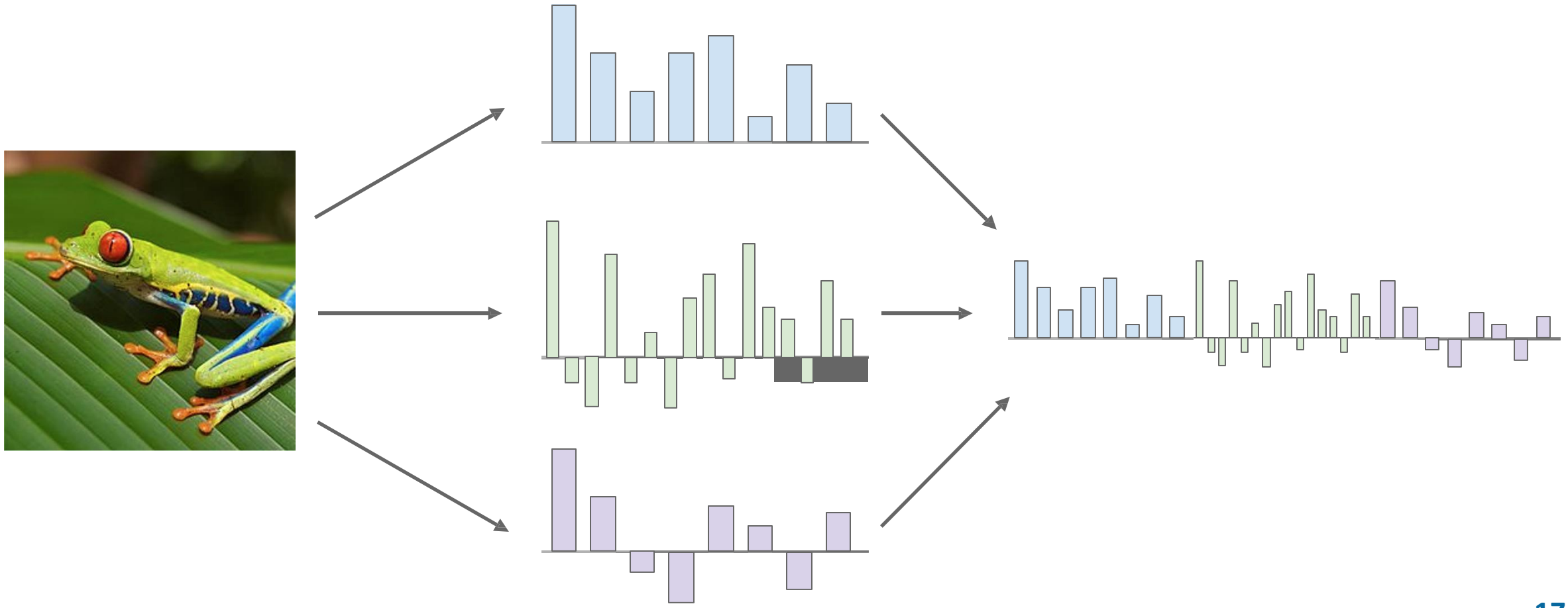
聚类图像块，形成“visual words”的codebook



图像分类

南京大学 NANJING UNIVERSITY

- 图像特征



图像分类

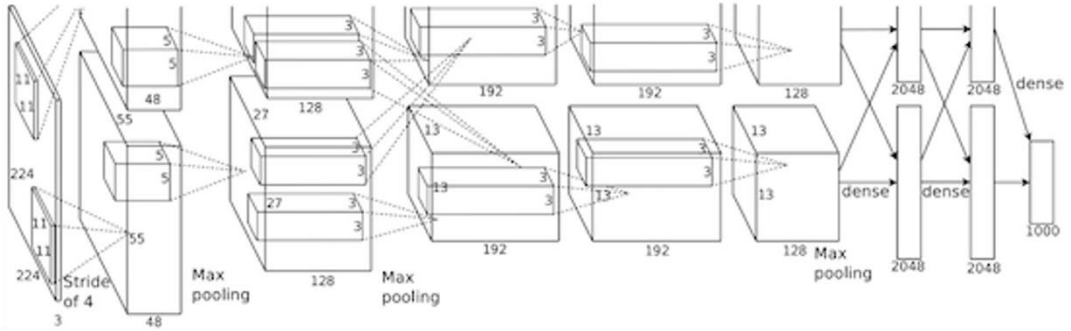
手工图像特征与卷积神经网络

Feature Extraction

f

分类得分

training



Krizhevsky, Sutskever, and Hinton, “Imagenet classification with deep convolutional neural networks”, NIPS 2012. Figure copyright Krizhevsky, Sutskever, and Hinton, 2012. Reproduced with permission.

分类得分

training

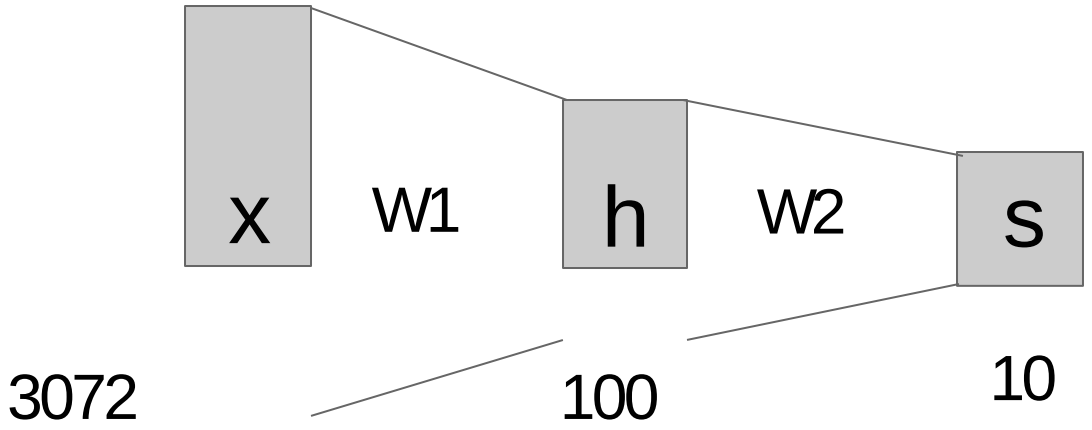
神经网络

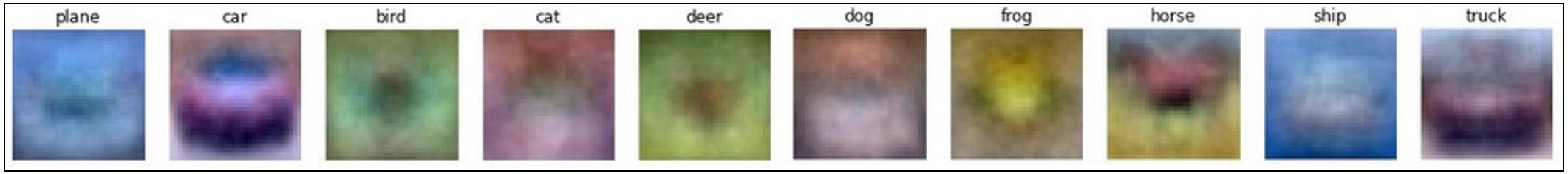
- 线性分类器

$$f = Wx$$

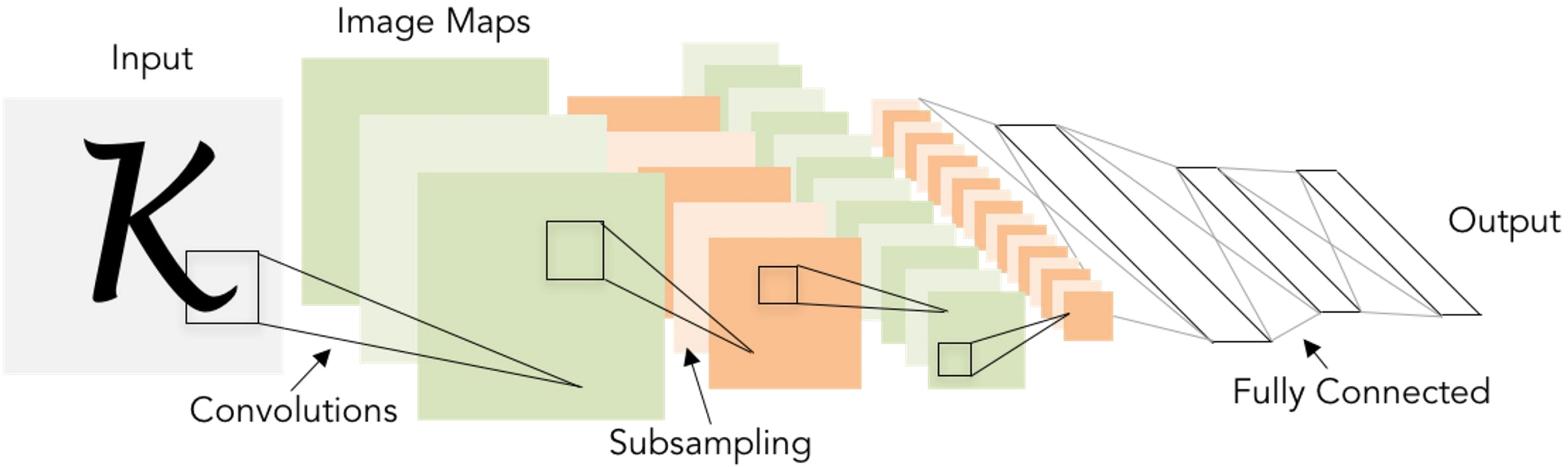
- 2 层神经网络

$$f = W_2 \max(0, W_1 x)$$





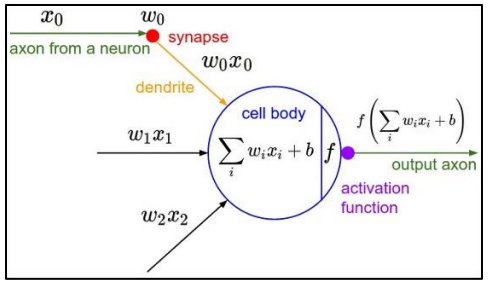
卷积神经网络


Input
Image Maps
Output
Convolutions
Subsampling
Fully Connected

卷积神经网络

- 神经网络的发展历史
- 第一个感知机：Mark I 感知机
- 识别字母：

$$f(x) = \begin{cases} 1 & \text{if } w \cdot x + b > 0 \\ 0 & \text{otherwise} \end{cases}$$



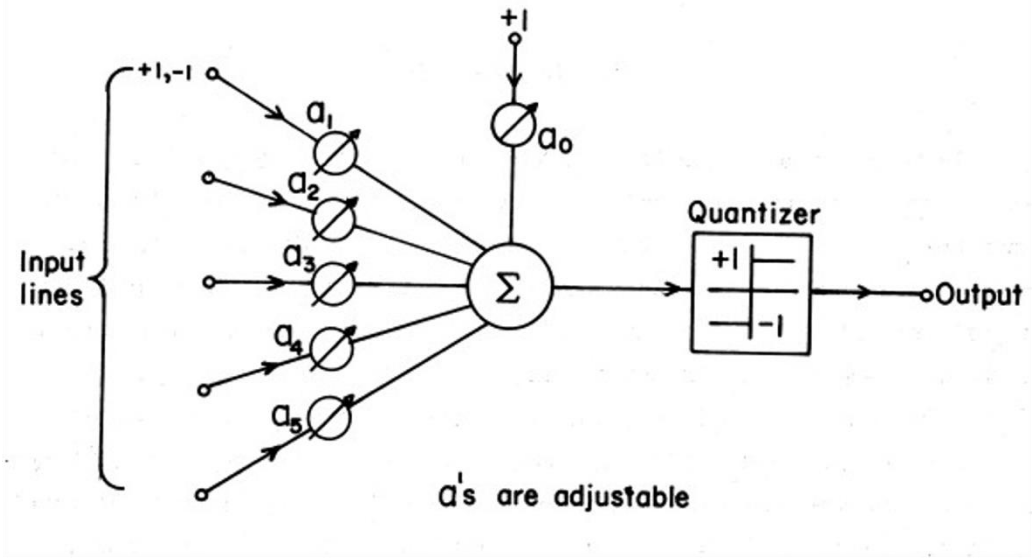
- 网络权重更新规则：

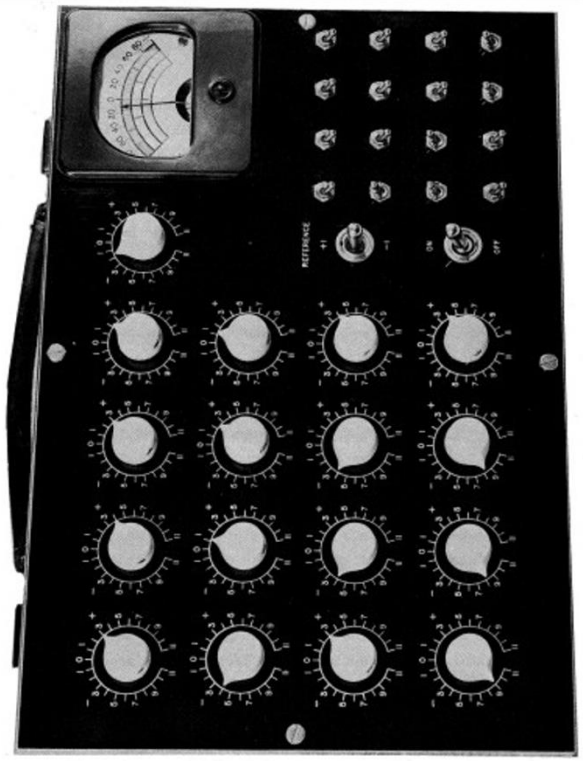
$$w_i(t+1) = w_i(t) + \alpha(d_j - y_j(t))x_{j,i},$$

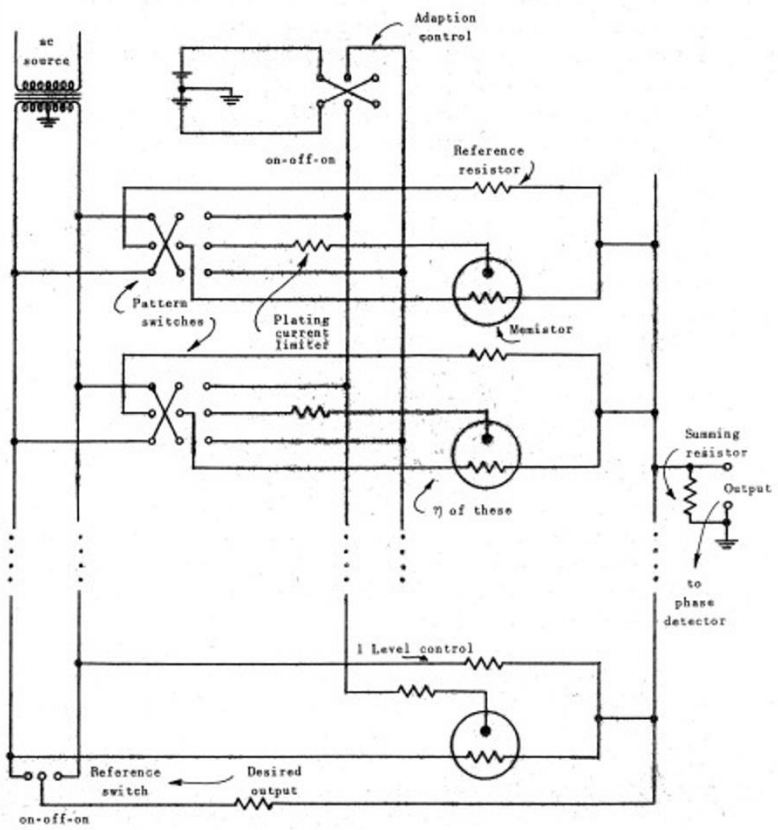
- Frank Rosenblatt, ~1957: Perceptron

卷积神经网络

- 神经网络的发展历史



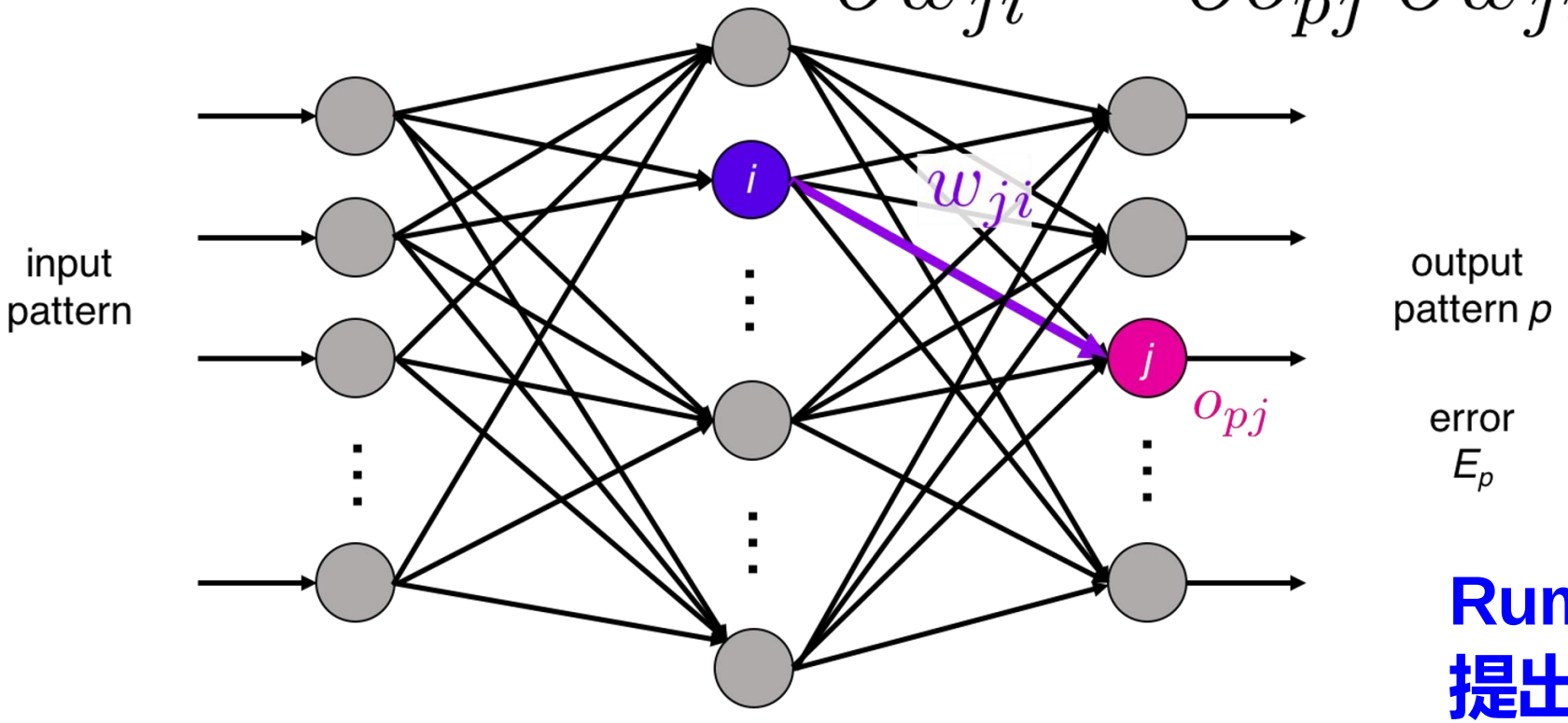




Widrow and Hoff, ~1960 : Adaline/Madaline

- 神经网络的发展历史

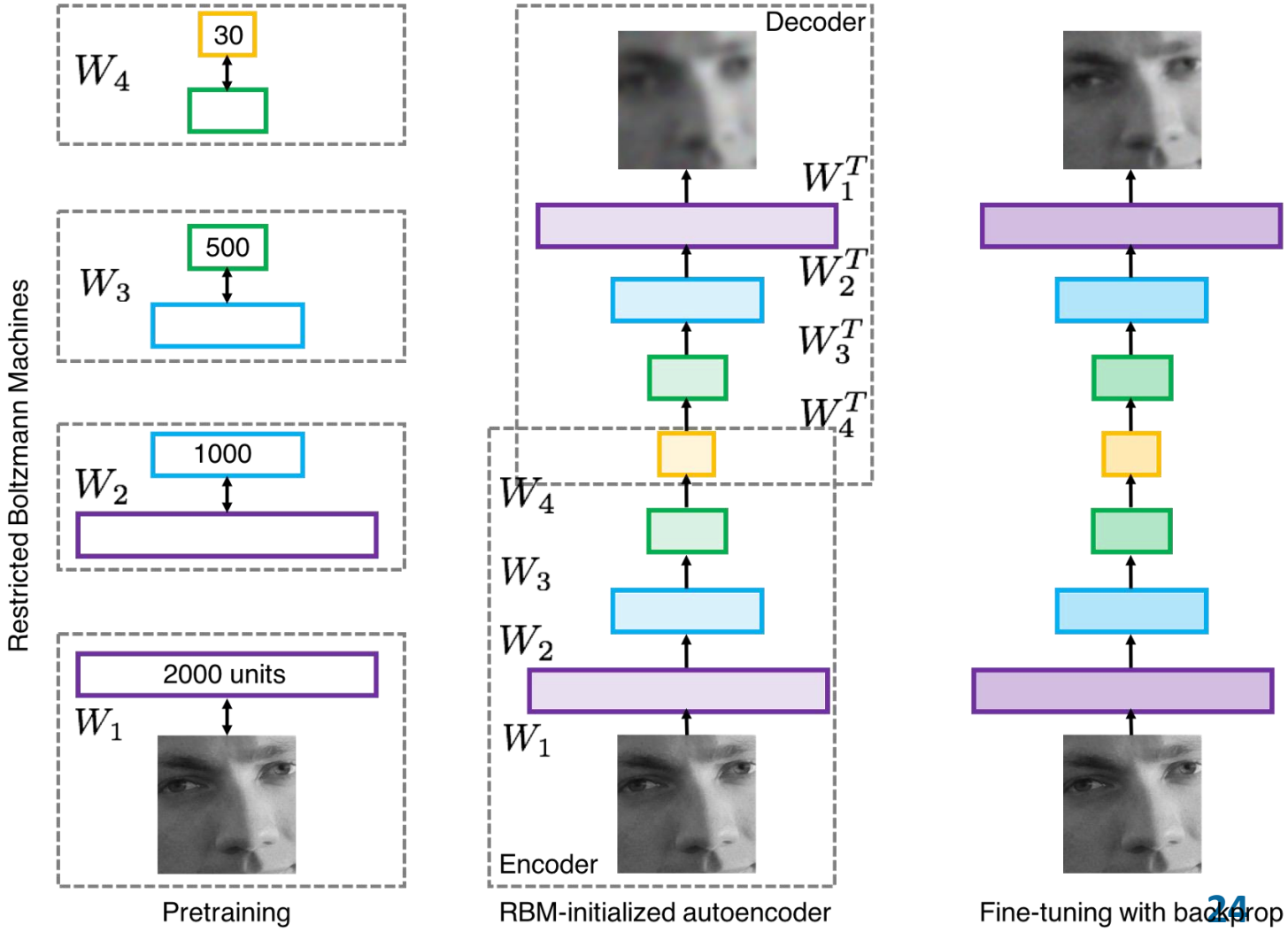
$$\frac{\partial E_p}{\partial w_{ji}} = \frac{\partial E_p}{\partial o_{pj}} \frac{\partial o_{pj}}{\partial w_{ji}}$$



Rumelhart et al., 1986: 提出反向传播

卷积神经网络

- 神经网络的发展历史
- [Hinton and Salakhutdinov 2006]
- 受限玻尔兹曼机
- 重振深度学习研究

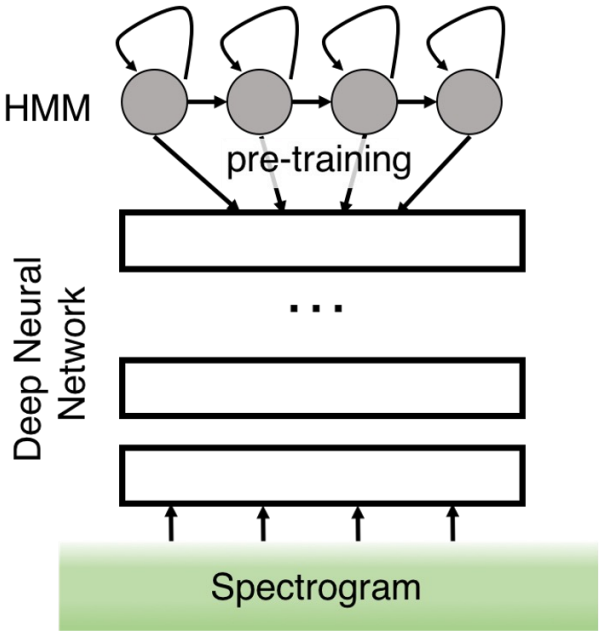


卷积神经网络

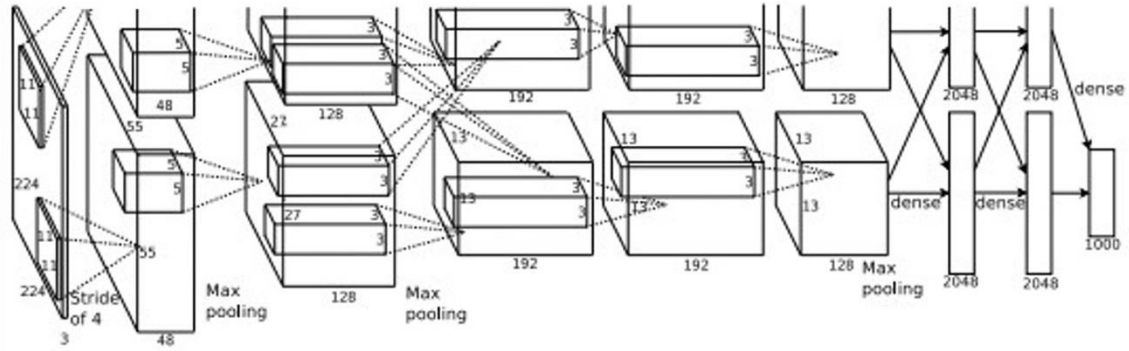
南京大学 NANJING UNIVERSITY

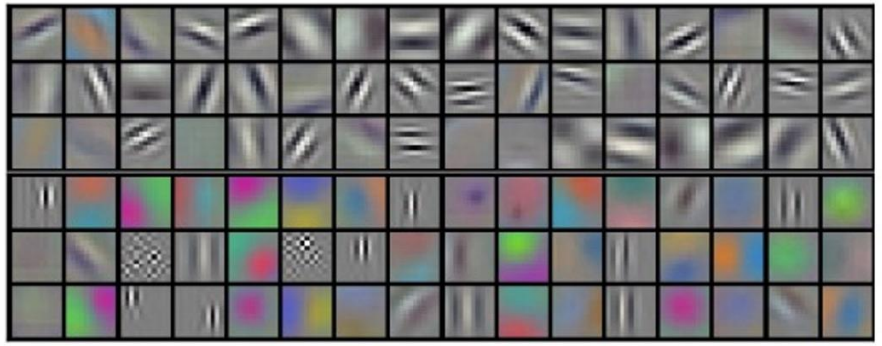
神经网络的发展历史

Acoustic Modeling using Deep Belief Networks
Abdel-rahman Mohamed, George Dahl, Geoffrey Hinton, 2010
Context-Dependent Pre-trained Deep Neural Networks
for Large Vocabulary Speech Recognition
George Dahl, Dong Yu, Li Deng, Alex Acero, 2012



Imagenet classification with deep convolutional
neural networks
Alex Krizhevsky, Ilya Sutskever, Geoffrey E Hinton, 2012





卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的发展历史

Hubel & Wiesel,

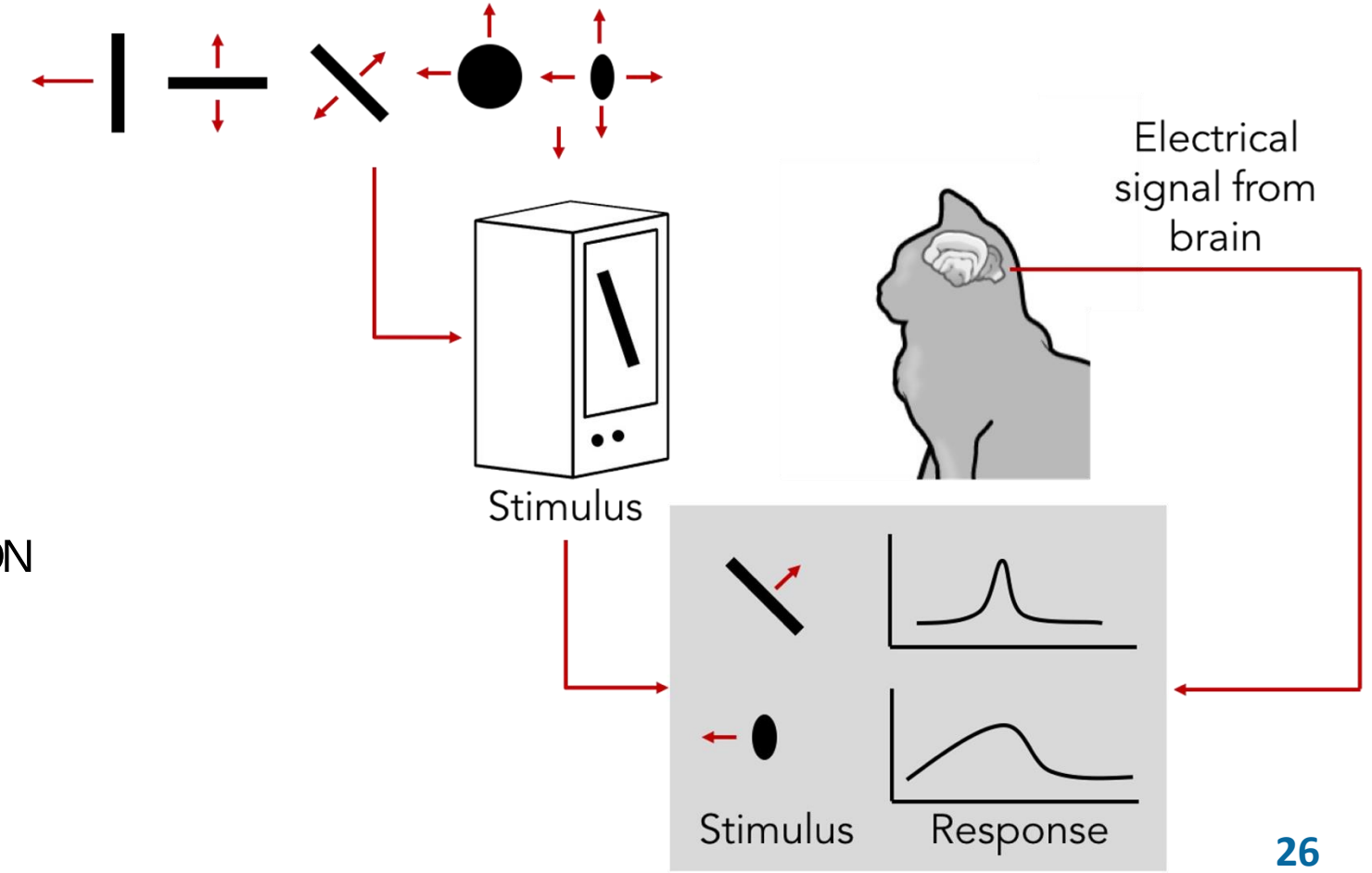
1959

RECEPTIVE FIELDS OF SINGLE NEURONES IN THE CAT'S STRIATE CORTEX

1962

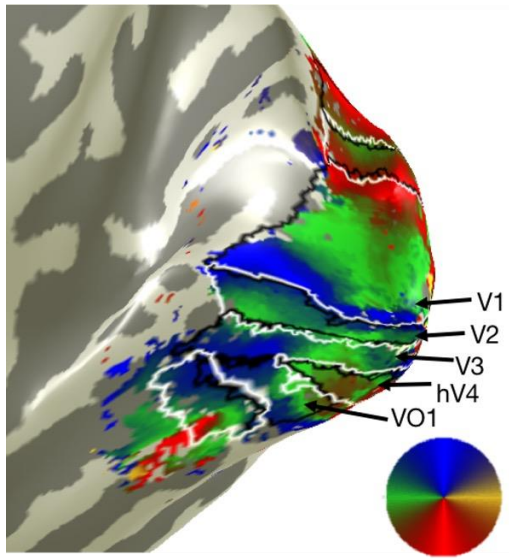
RECEPTIVE FIELDS, BINOCULAR INTERACTION AND FUNCTIONAL ARCHITECTURE IN THE CAT'S VISUAL CORTEX

1968...

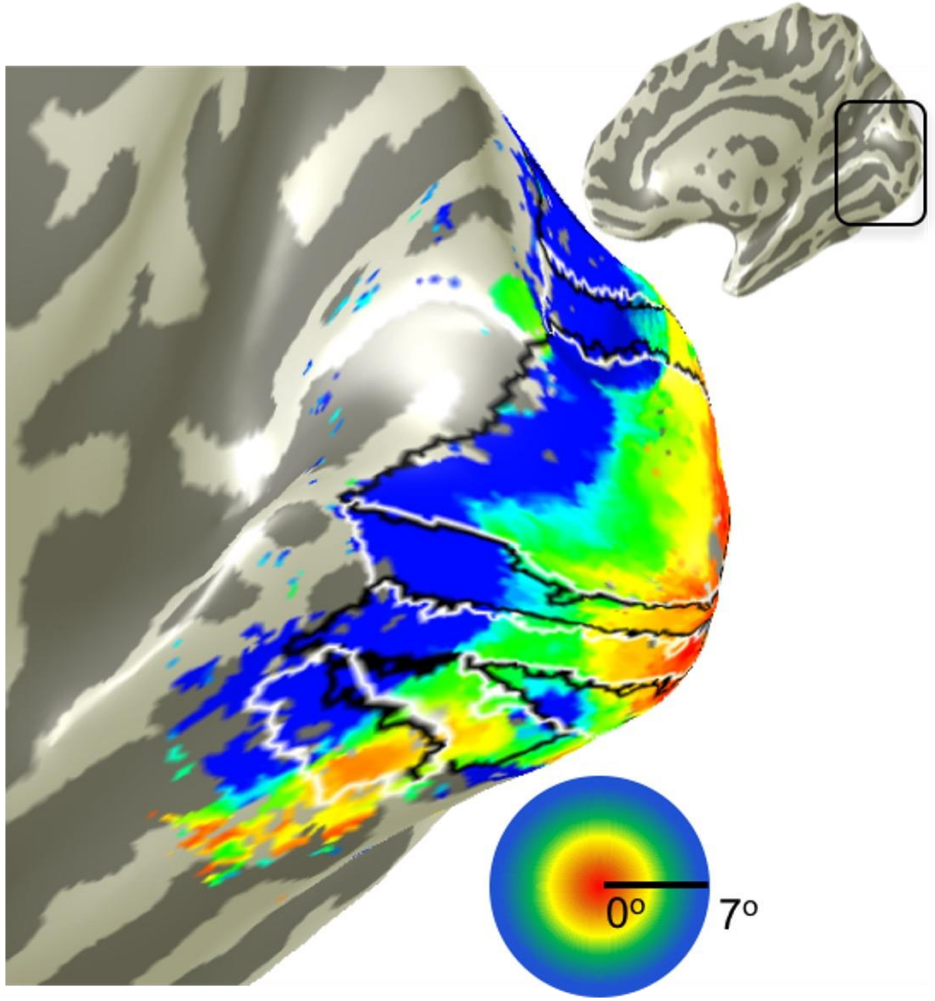


卷积神经网络

- 卷积神经网络的发展历史
- 皮层地形图：大脑皮层中的相邻细胞被视野中的相邻区域激活

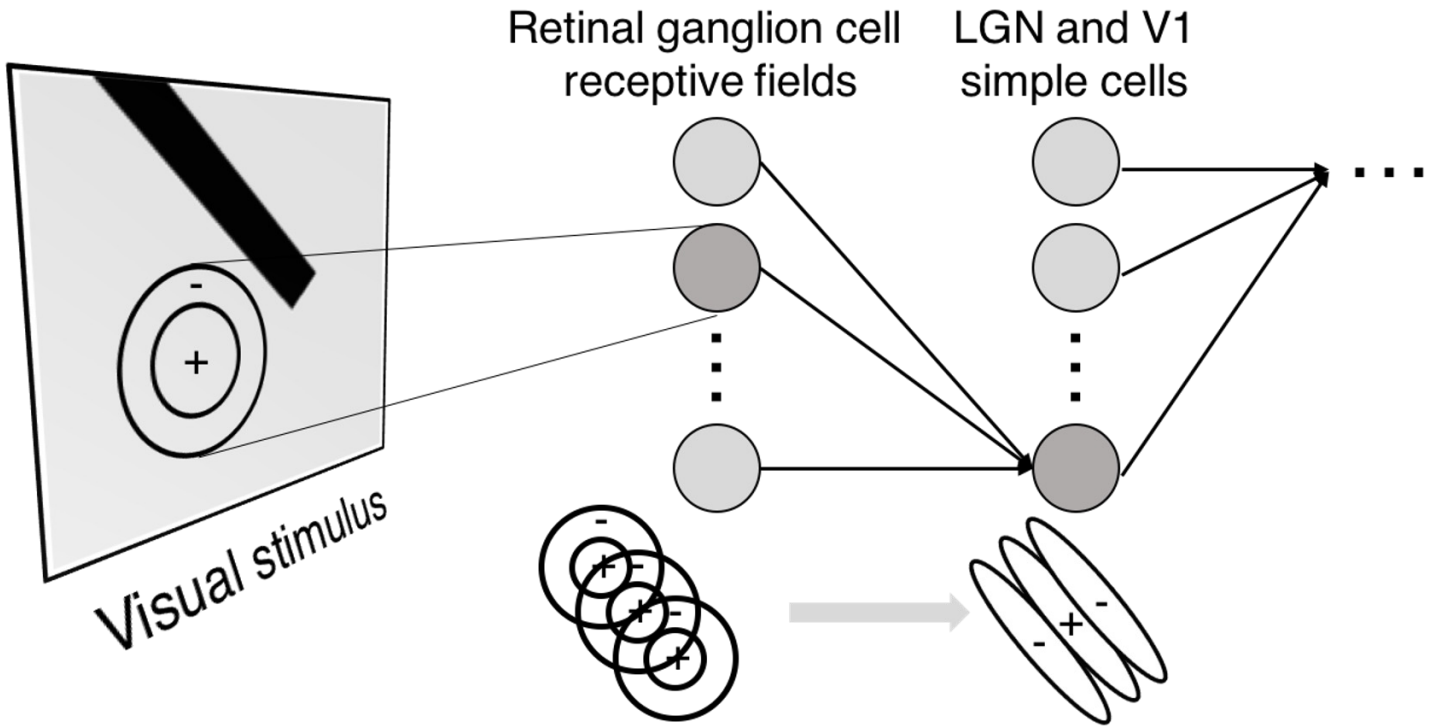


人类大脑



卷积神经网络

- 卷积神经网络的发展历史
- 视觉分级感知



Simple cells:
Response to light orientation

Complex cells:
Response to light orientation and movement

Hypercomplex cells:
response to movement with an end point

No response

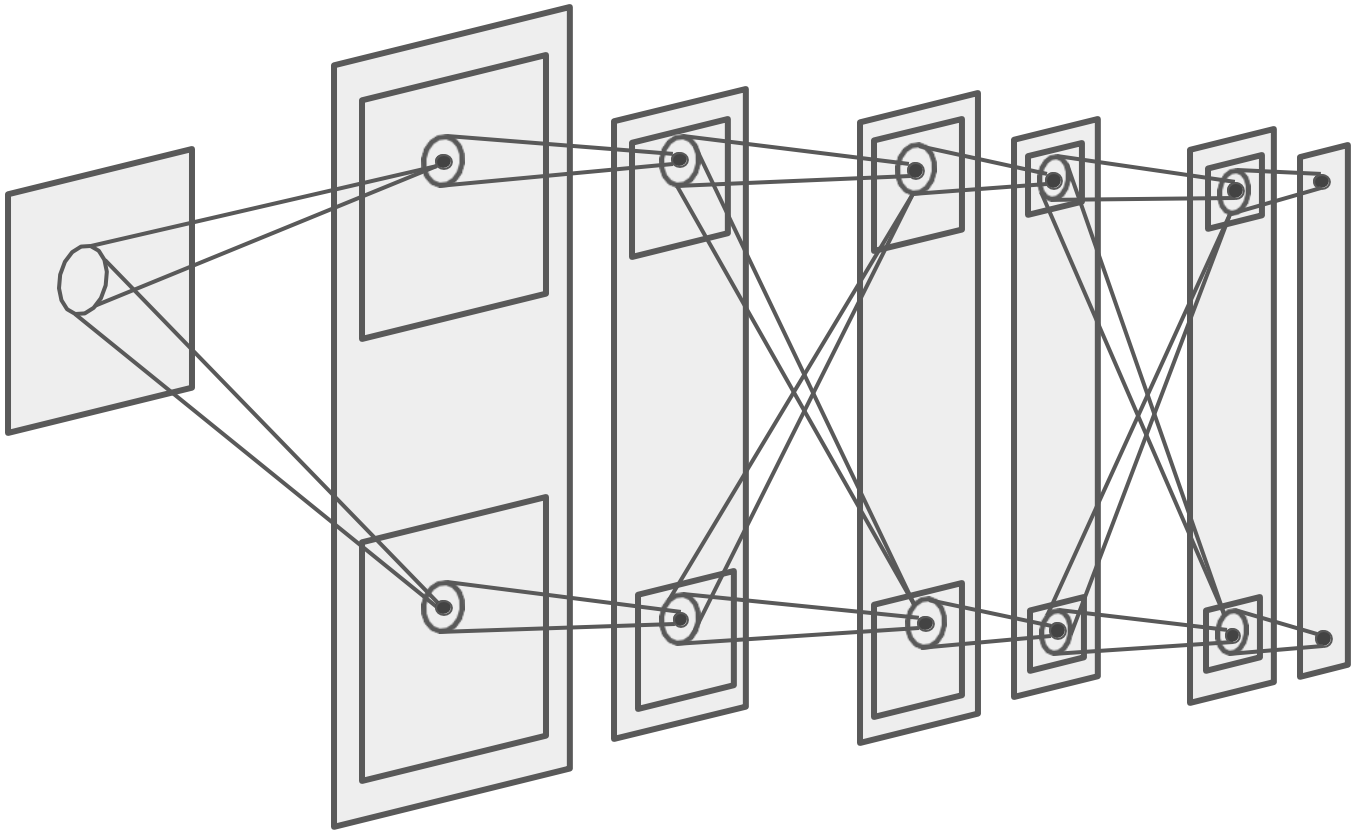
Response (end point)

卷积神经网络

南京大学 NANJING UNIVERSITY

- **卷积神经网络的发展历史**

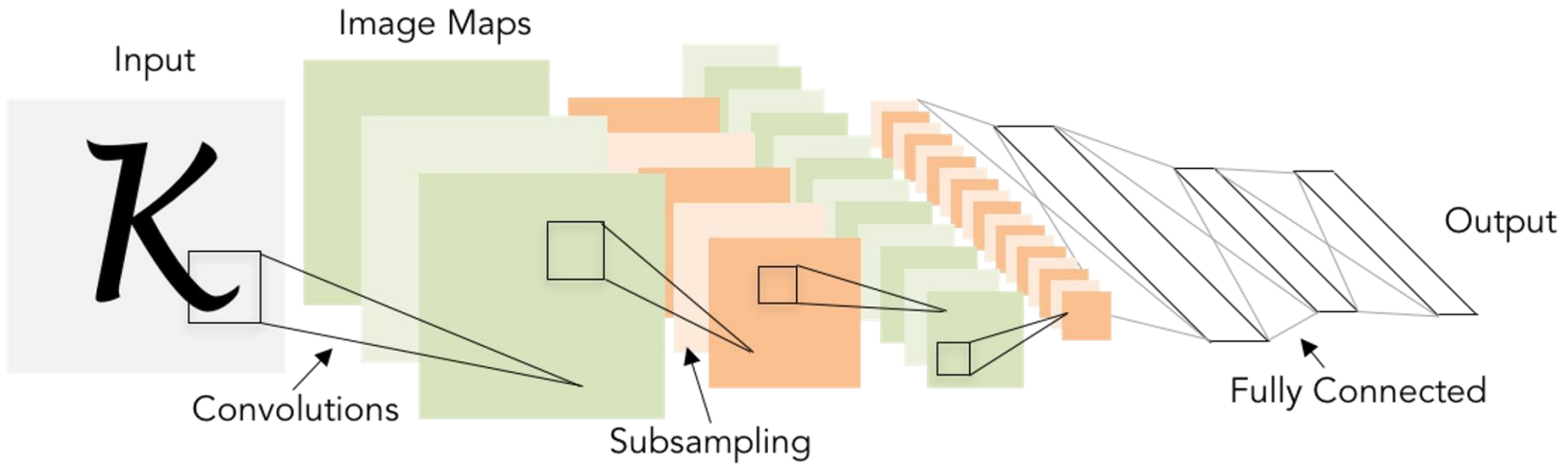
Neocognitron
[Fukushima 1980]



三明治结构 (SCSCSC…)
simple cells: modifiable parameters
complex cells: perform pooling

卷积神经网络

- 卷积神经网络的发展历史
- 将基于梯度优化的卷积网络应用到文档识别
- [LeCun, Bottou, Bengio, Haffner 1998]



卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的发展历史

ImageNet Classification with Deep Convolutional Neural Networks [Krizhevsky, Sutskever, Hinton, 2012]

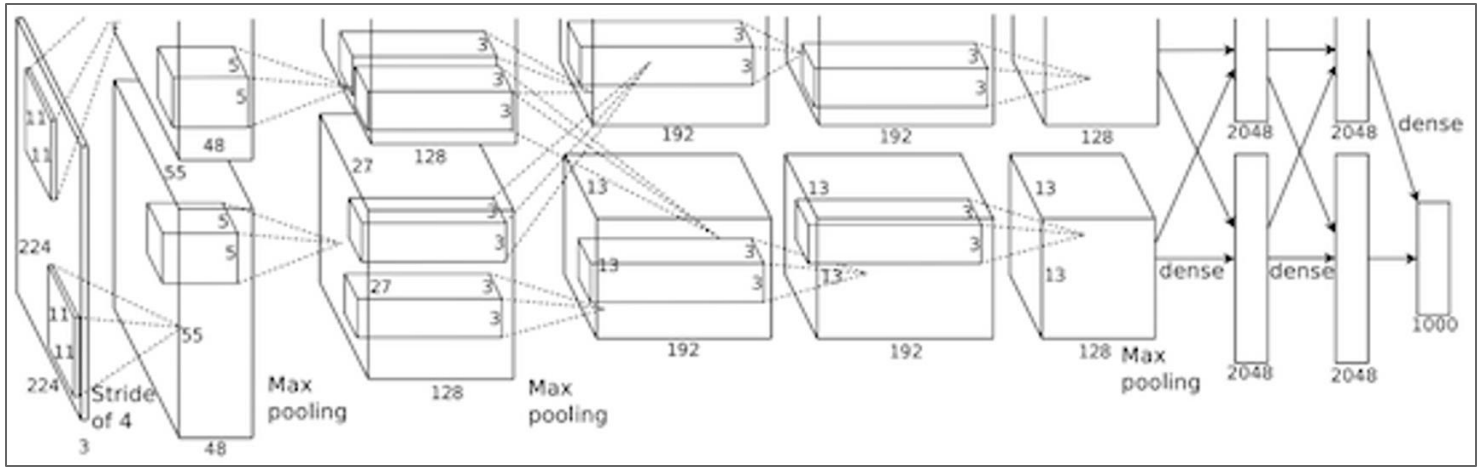


Figure copyright Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton, 2012. Reproduced with permission.

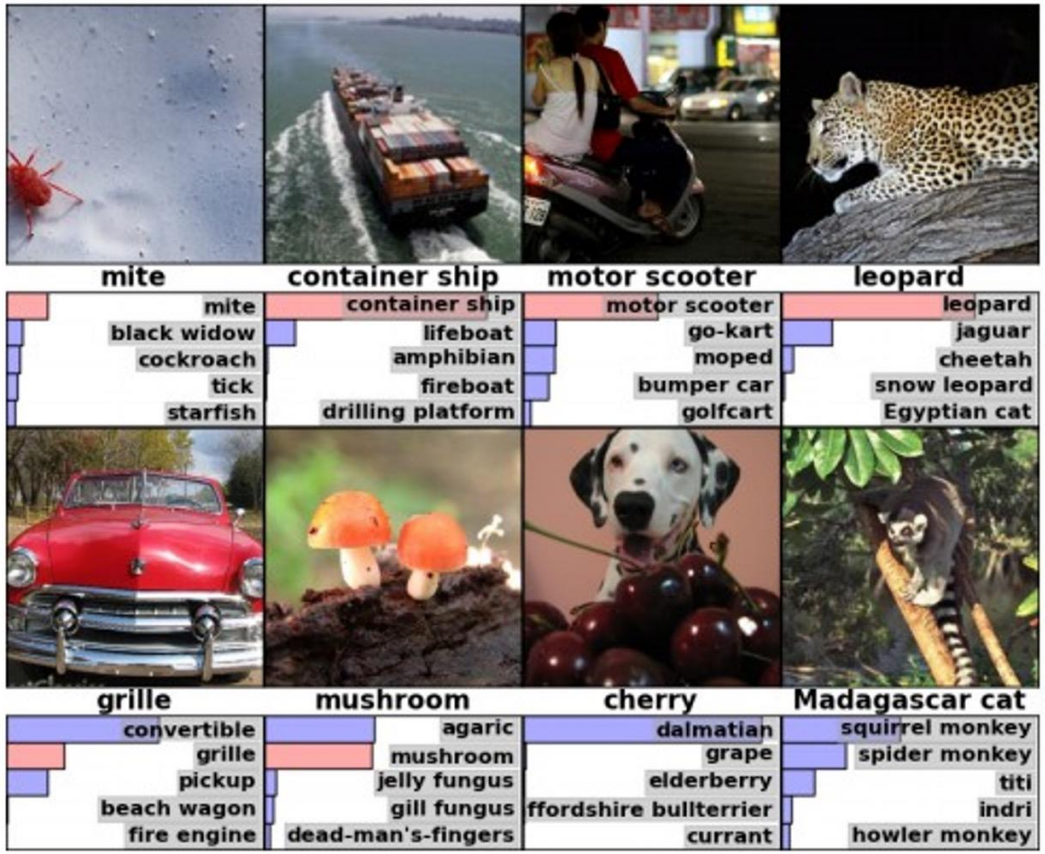
"AlexNet"

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用

Classification



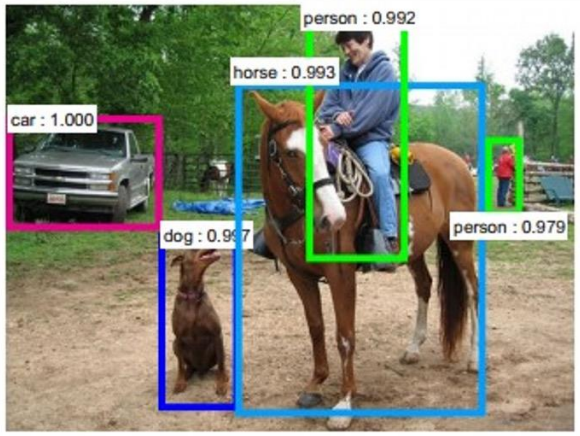
Retrieval

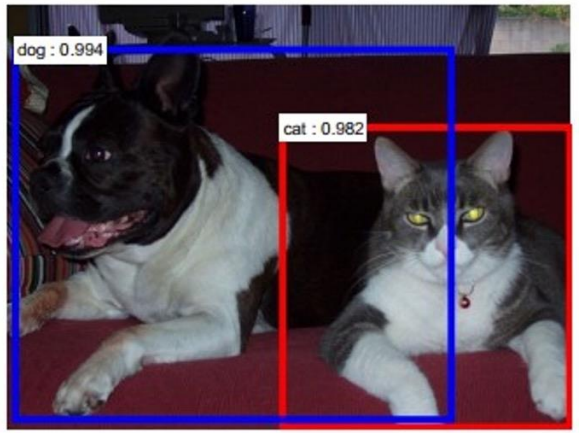
卷积神经网络

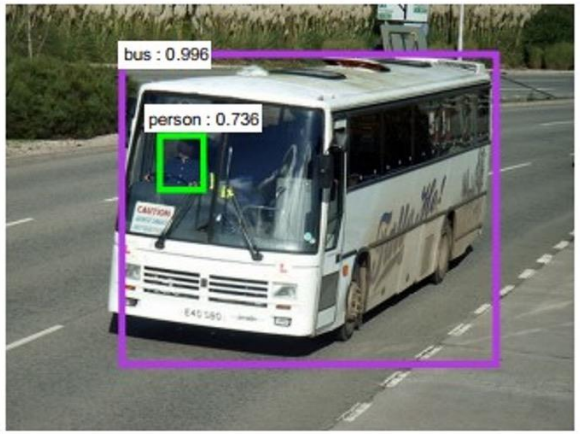
南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用

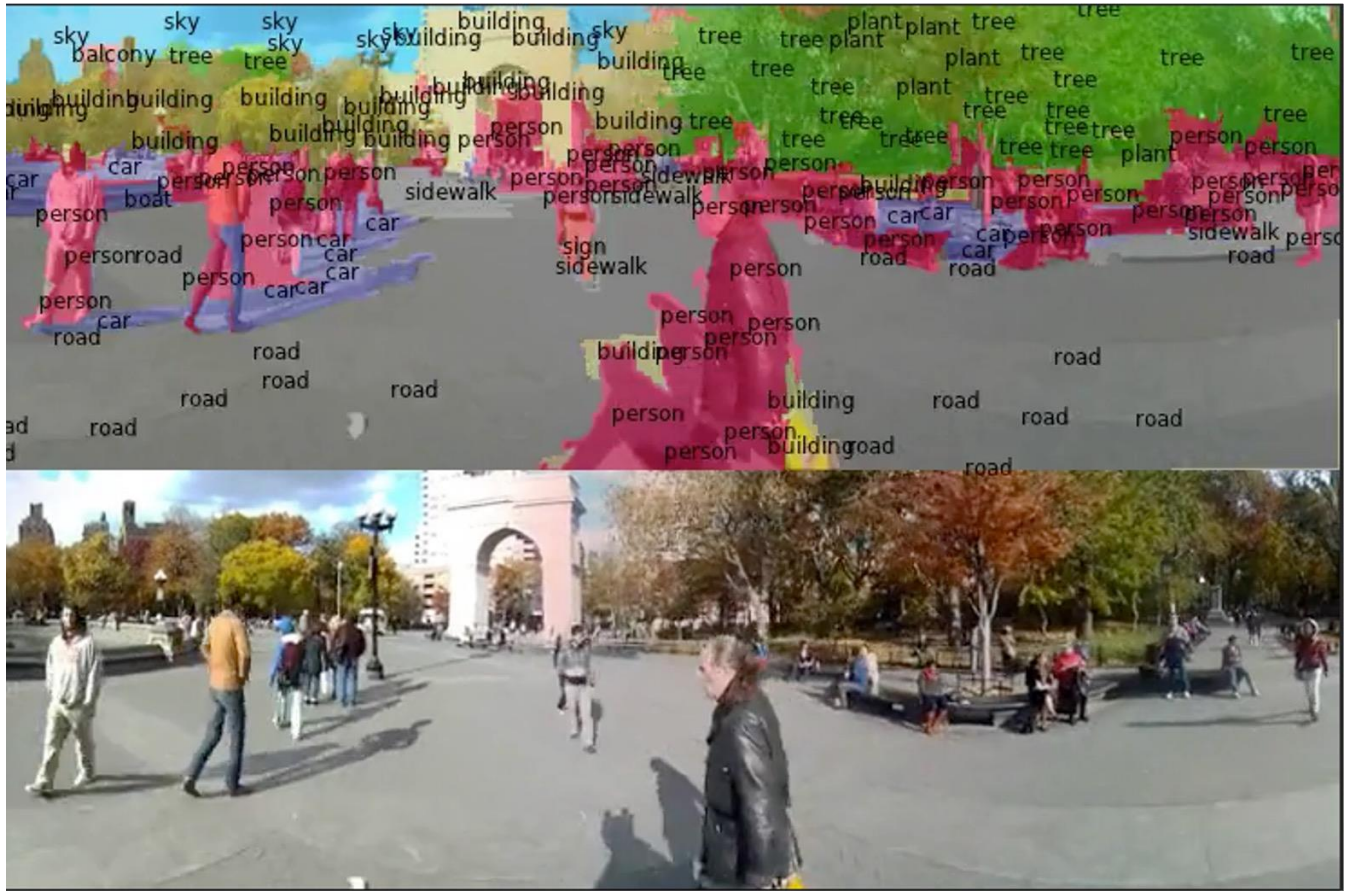
Detection







Segmentation



卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用

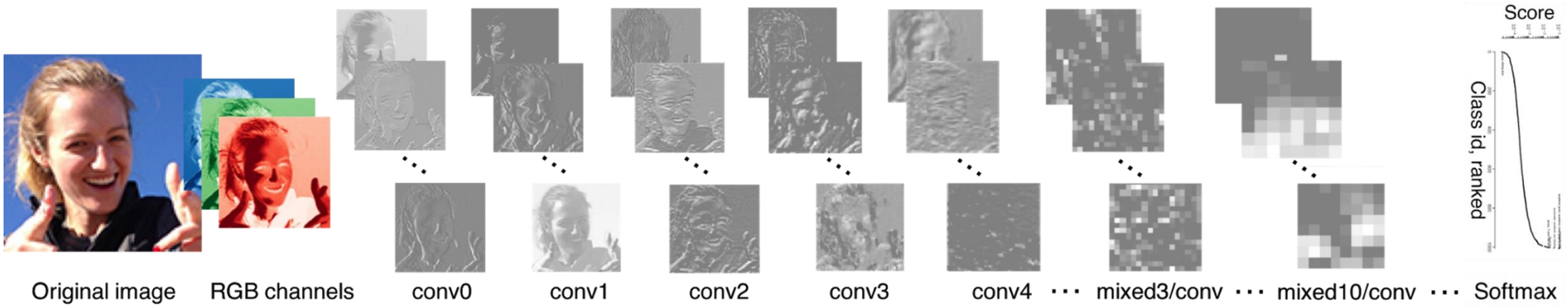
自动驾驶汽车

NVIDIA Tesla 系列显卡

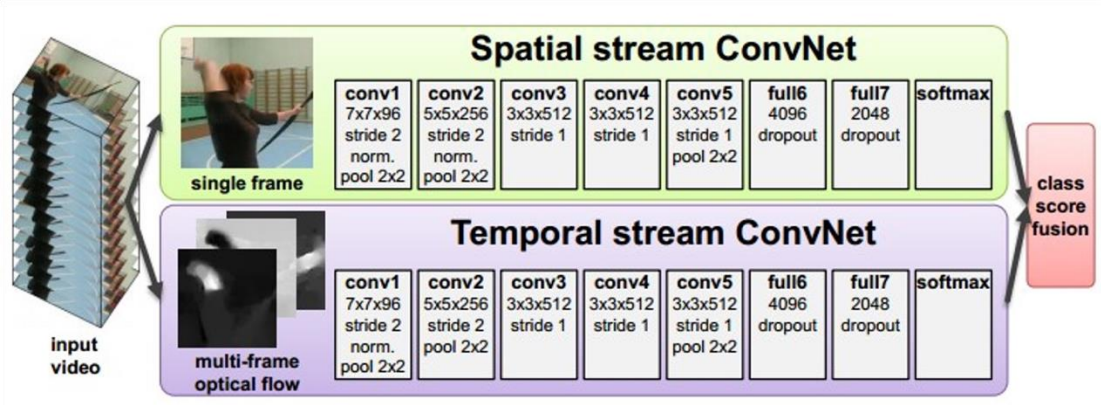
卷积神经网络

南京大学 NANJING UNIVERSITY

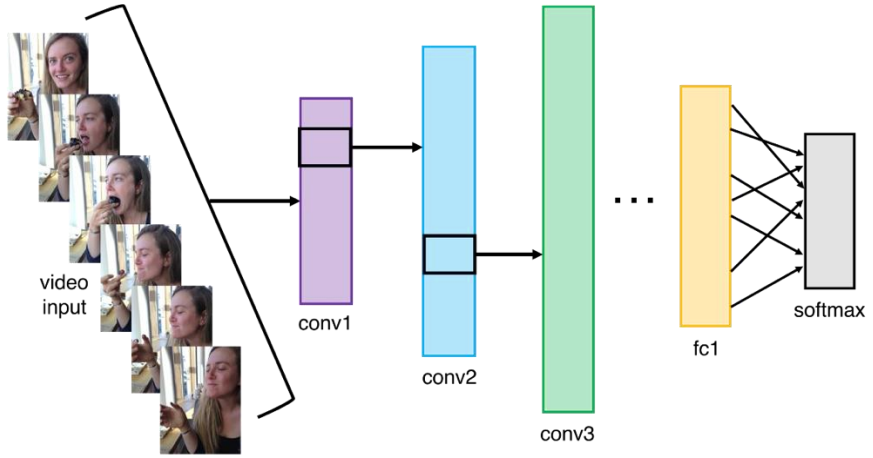
卷积神经网络的广泛应用



[Taigman et al. 2014]



[Simonyan et al. 2014]

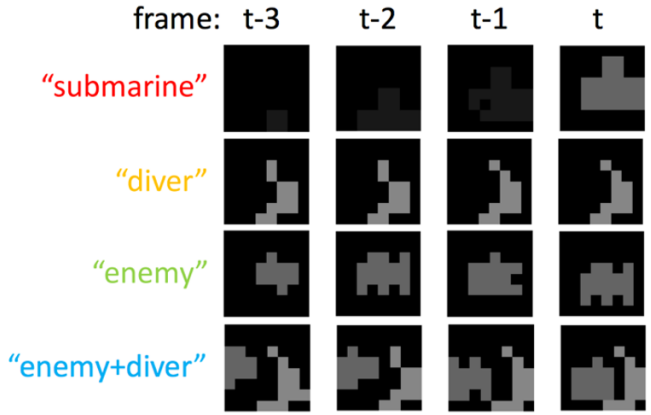


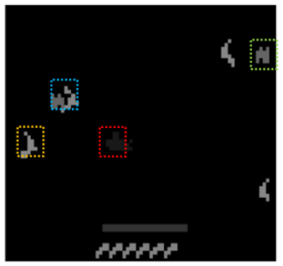
卷积神经网络

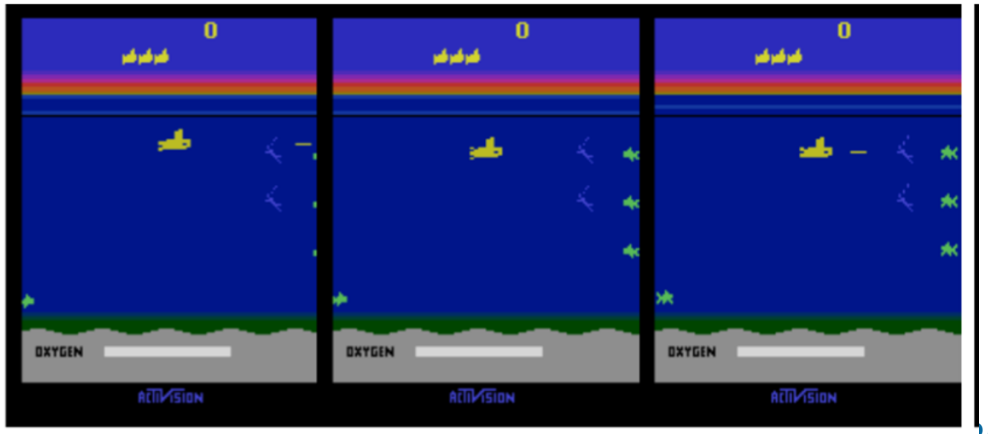
南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用

[Toshev, Szegedy 2014]





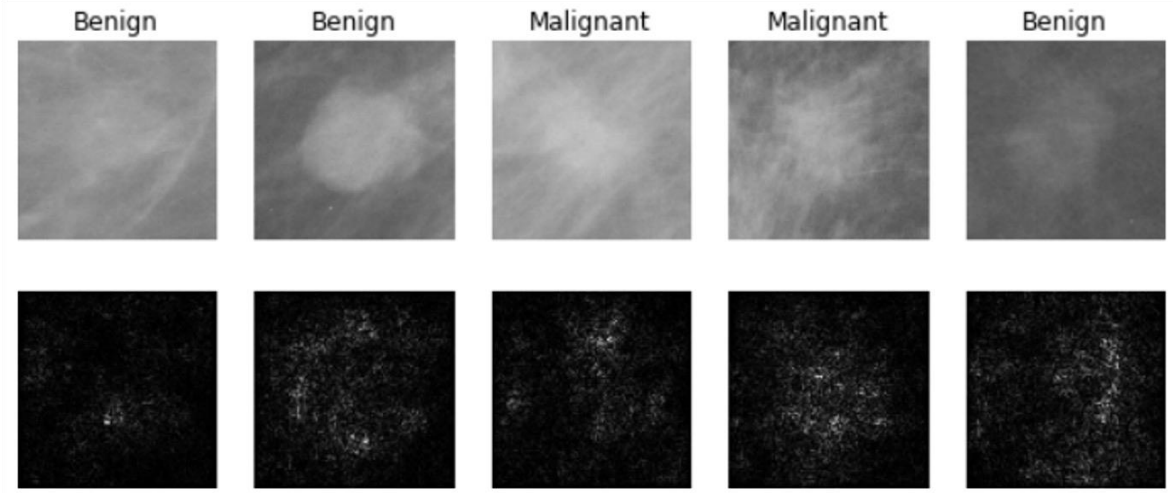


[Guo et al. 2014]

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用



[Levy et al. 2016]

[Dieleman et al. 2014]

[Sermanet et al. 2011]
[Ciresan et al.]

卷积神经网络

南京大学 NANJING UNIVERSITY

- 卷积神经网络的广泛应用

Kaggle Challenge

Mnih and Hinton, 2010

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用

No errors

A white teddy bear sitting in the grass

A man riding a wave on top of a surfboard

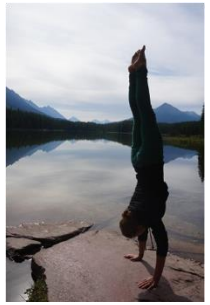
Minor errors

A man in a baseball uniform throwing a ball

A cat sitting on a suitcase on the floor

Somewhat related

A woman is holding a cat in her hand



A woman standing on a beach holding a surfboard

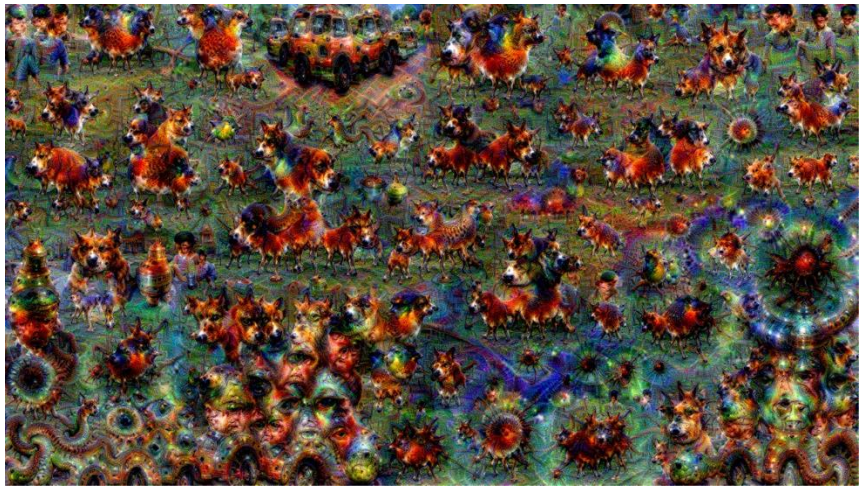
图像描述

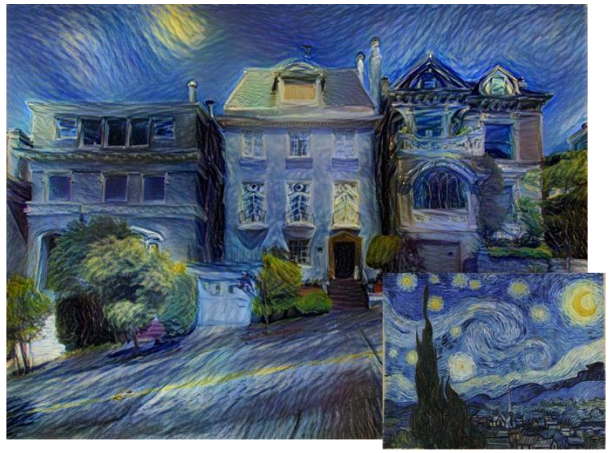
[Vinyals et al., 2015]
[Karpathy and Fei-Fei, 2015]

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积神经网络的广泛应用



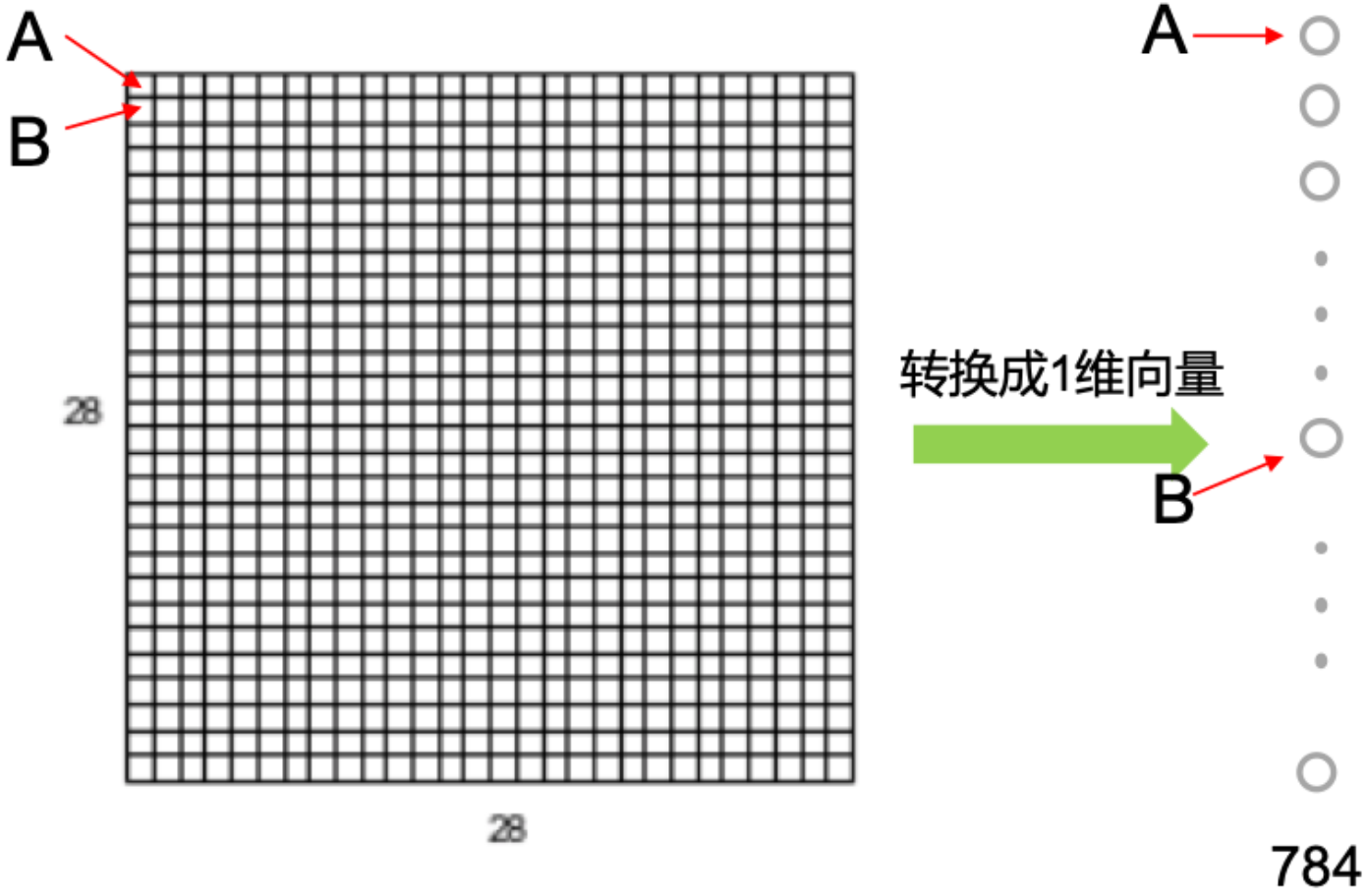


大　　纲

- 卷积神经网络-引言
- 卷积神经网络

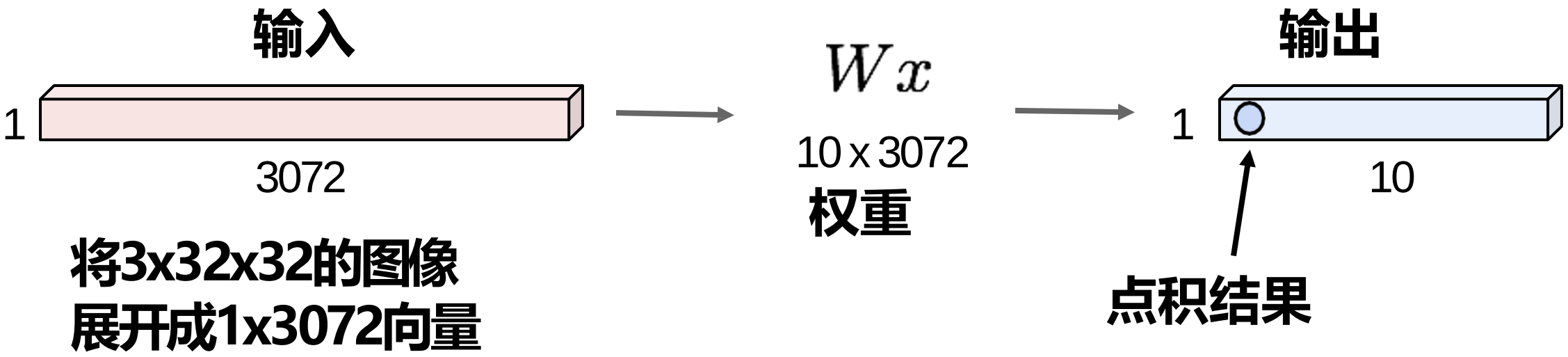
卷积神经网络

- 全连接层（参数量，优化，空间信息）



卷积神经网络

- 全连接层（参数量，优化，空间信息）

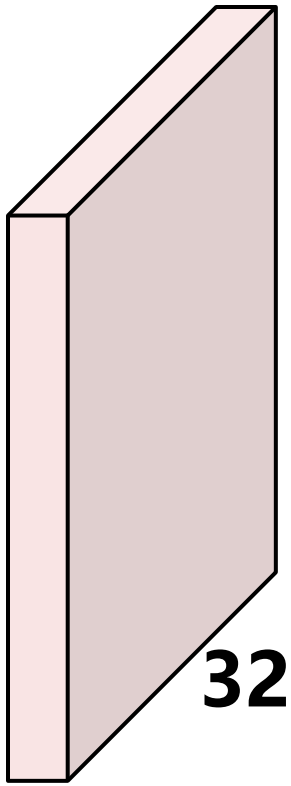


卷积神经网络

南京大学 NANJING UNIVERSITY

- 卷积层

3x32x32图像：保持图像的空间结构



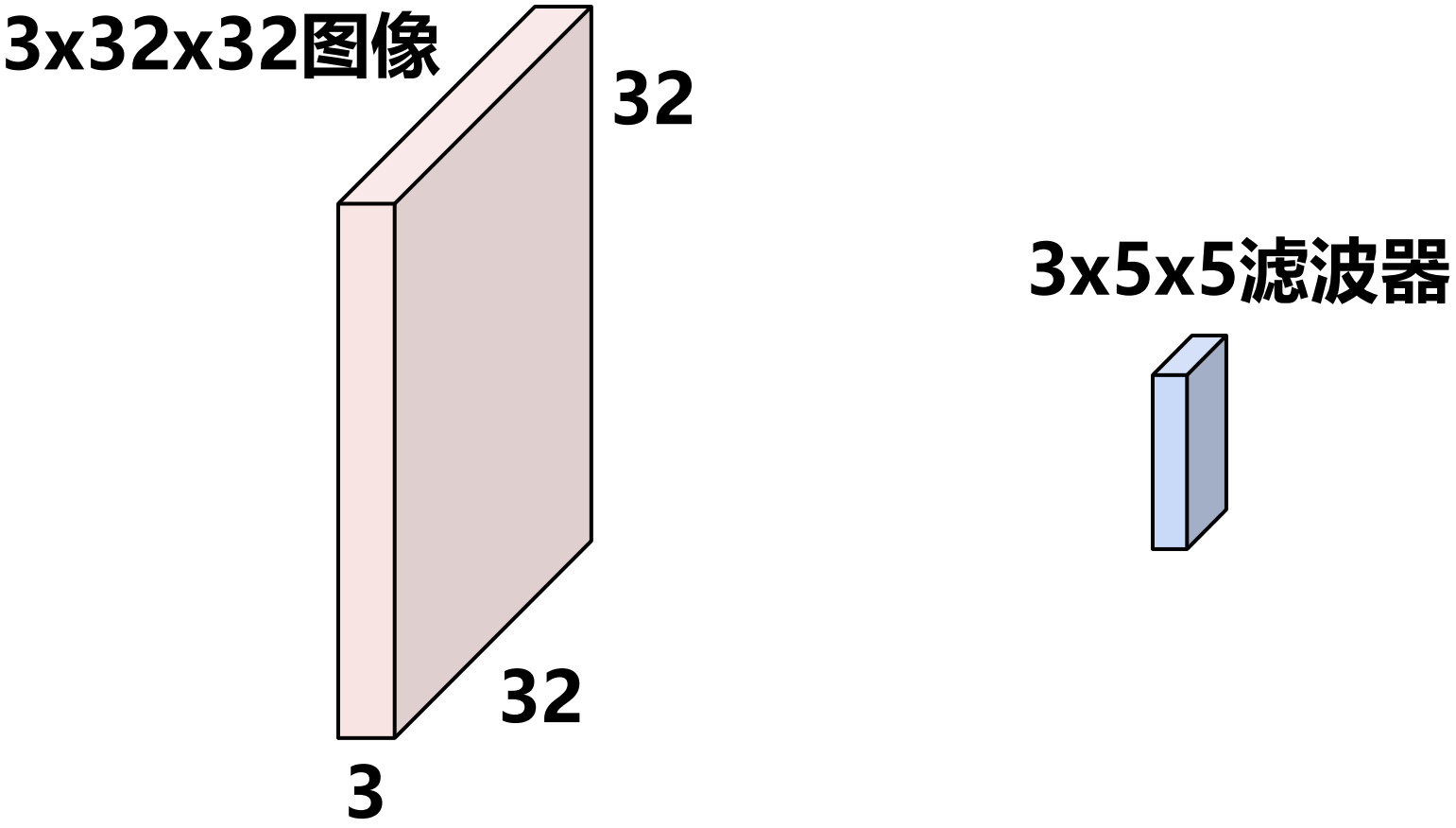
32（高度，height）

32（宽度，width）

3（深度，depth）

卷积神经网络

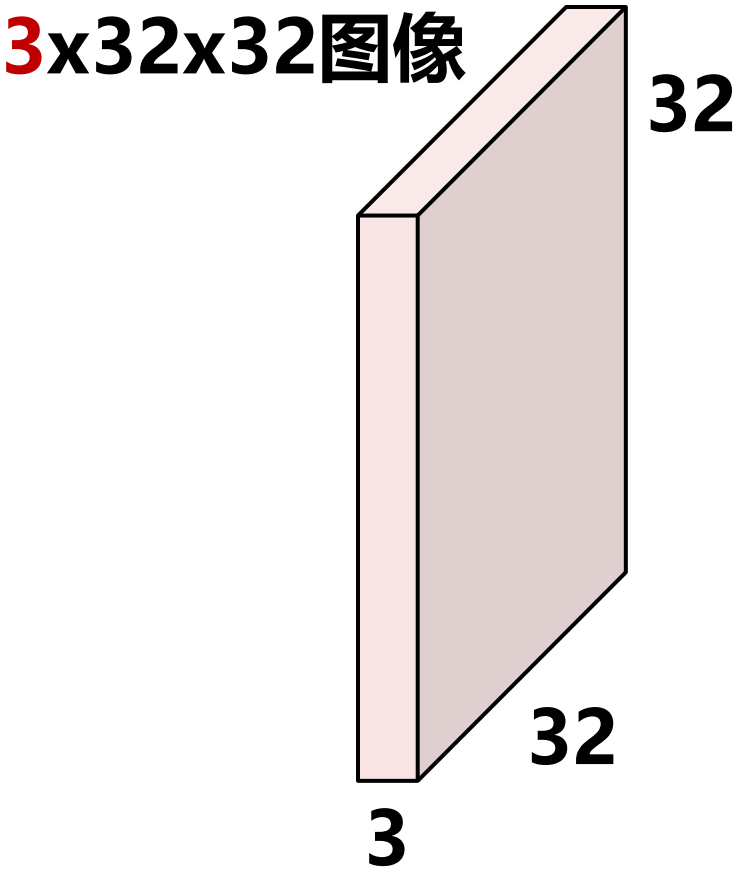
- 卷积层

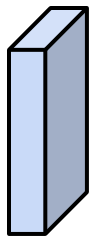


卷积神经网络

- 卷积层

滤波器的深度始终与输入的深度相同

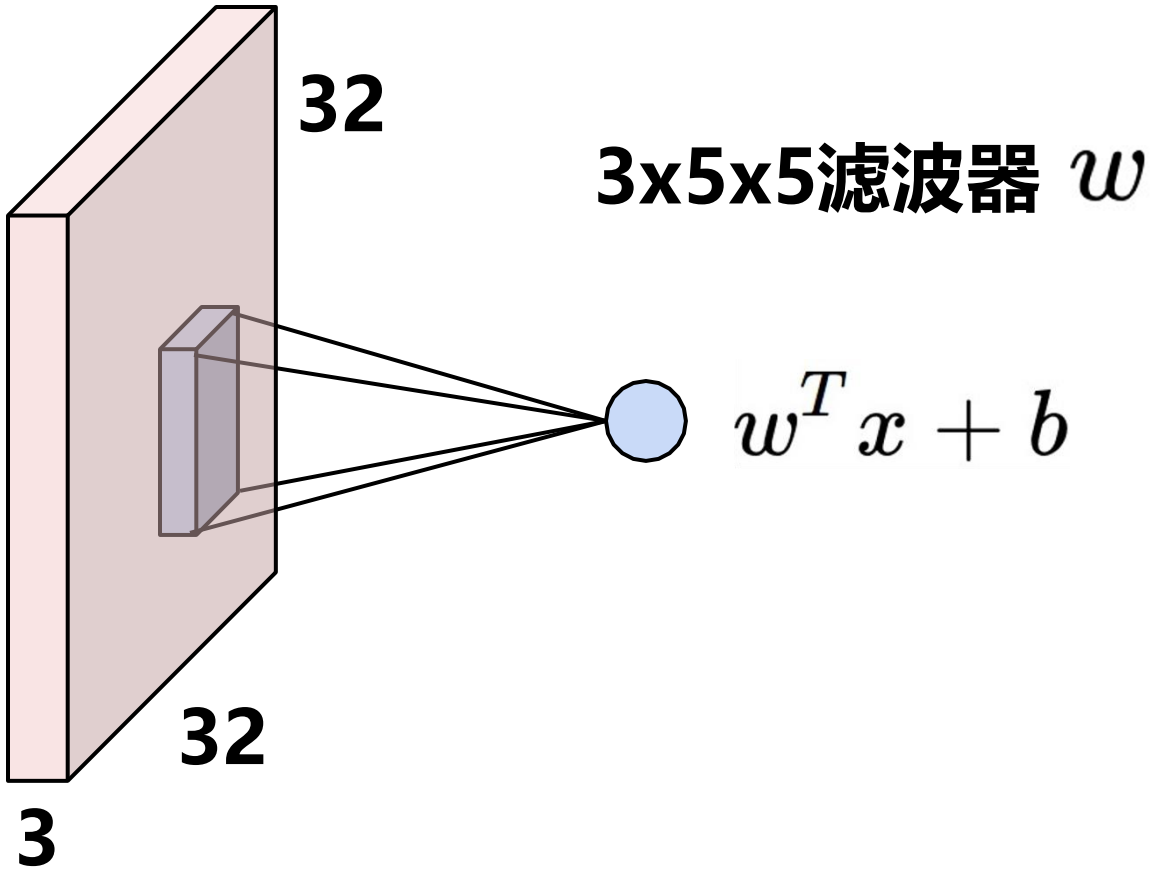




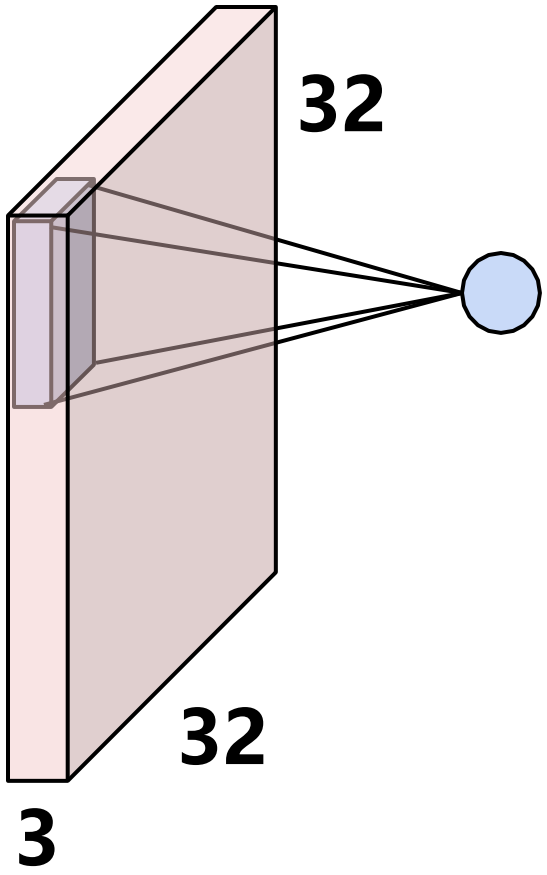
将滤波器与图像卷积，
即“在图像上空间滑动，计算点积”

卷积神经网络

卷积层



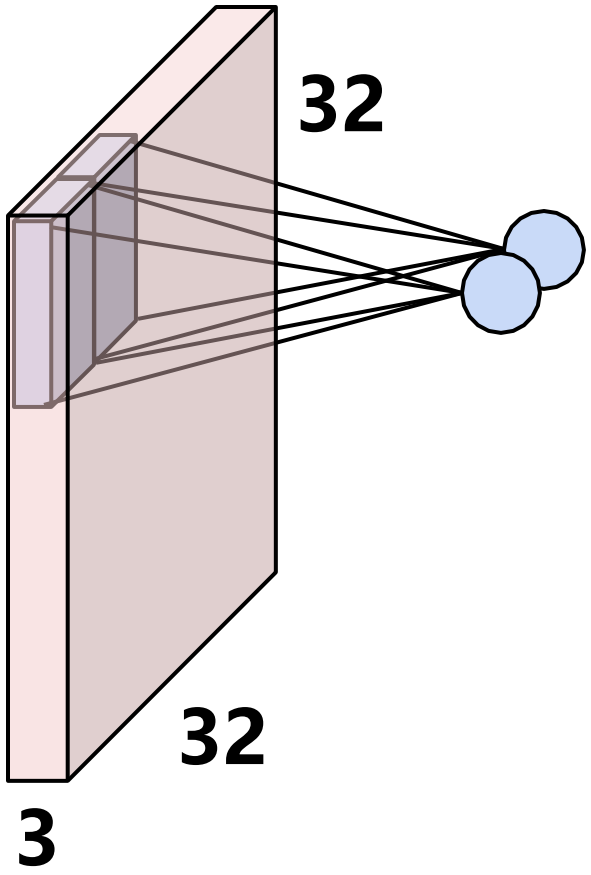
卷积层



卷积神经网络

南京大学 NANJING UNIVERSITY

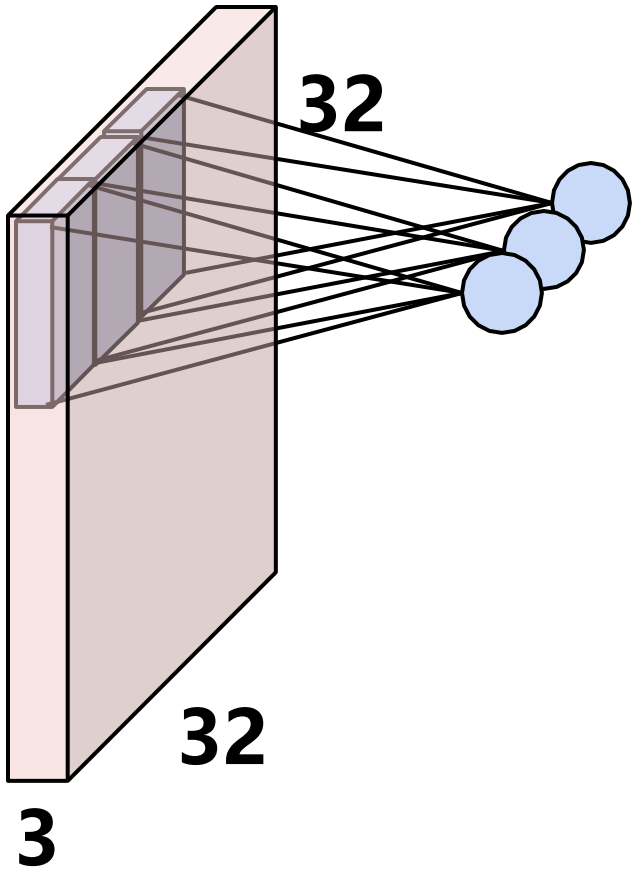
卷积层



卷积神经网络

南京大学 NANJING UNIVERSITY

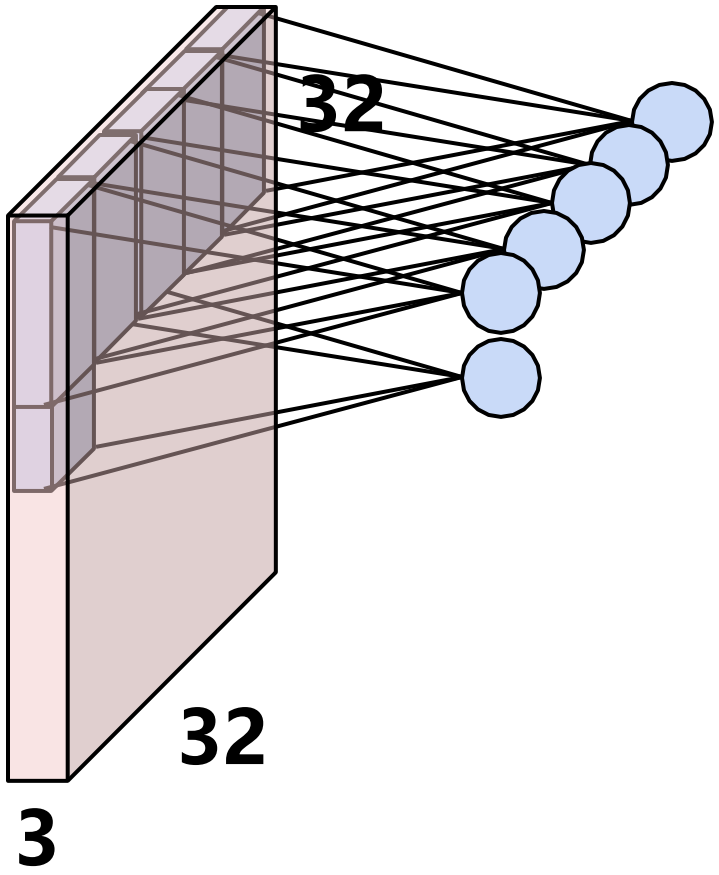
卷积层



卷积神经网络

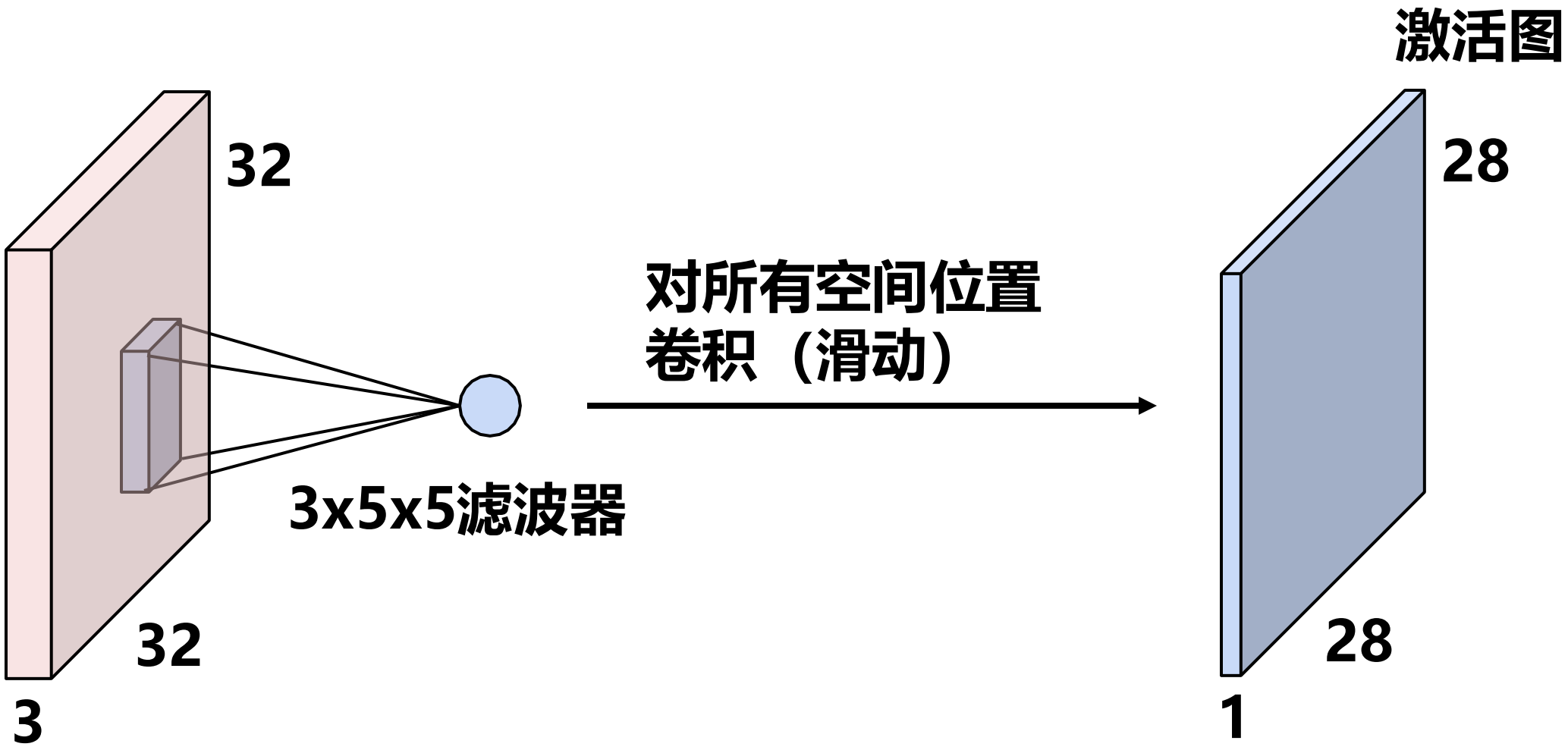
南京大学 NANJING UNIVERSITY

卷积层



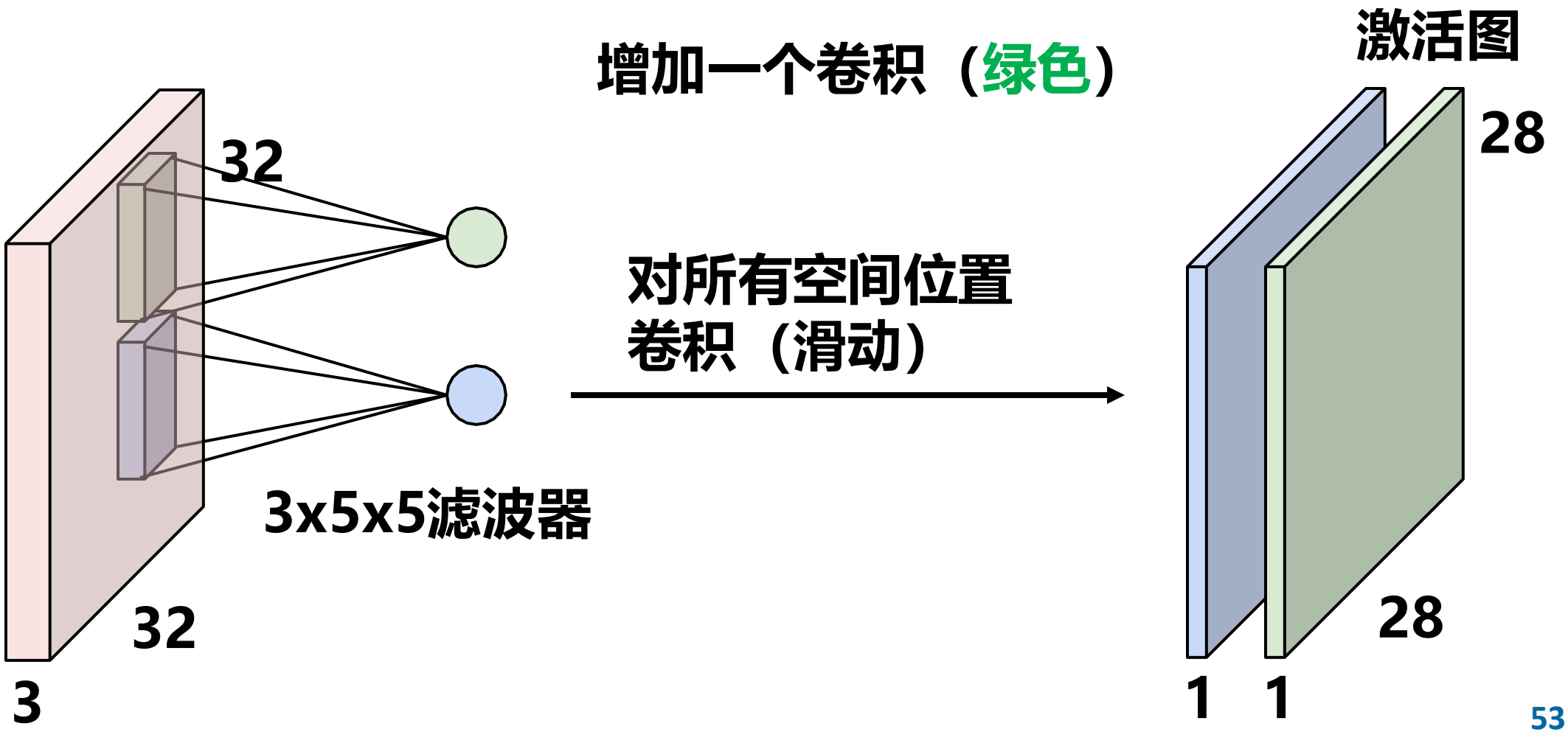
卷积神经网络

- 卷积层



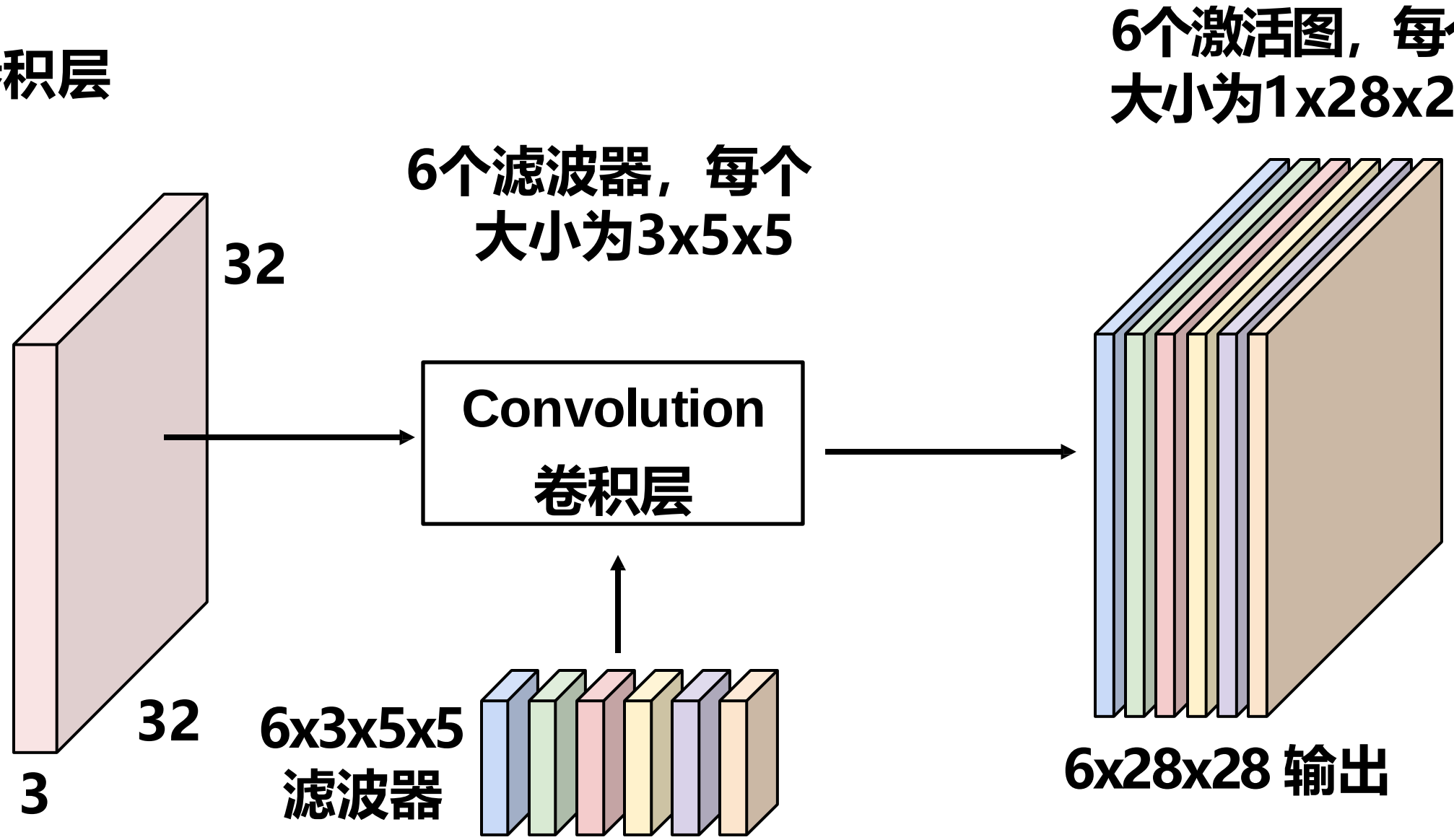
卷积神经网络

- 卷积层



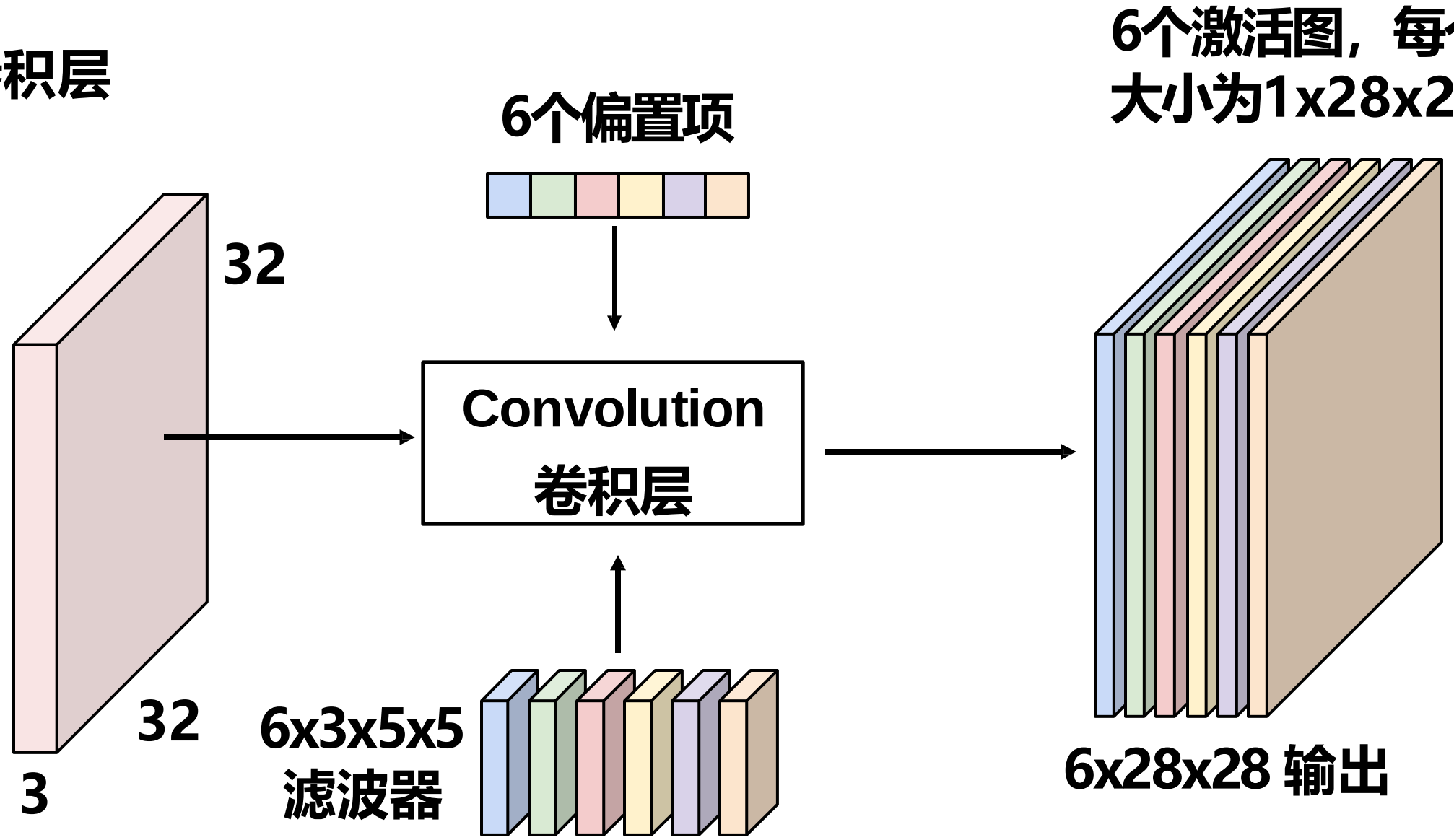
卷积神经网络

- 卷积层


6个激活图，每个
大小为1x28x28
6个滤波器，每个
大小为3x5x5
32
32
3
Convolution
卷积层
6x3x5x5
滤波器
6x28x28 输出

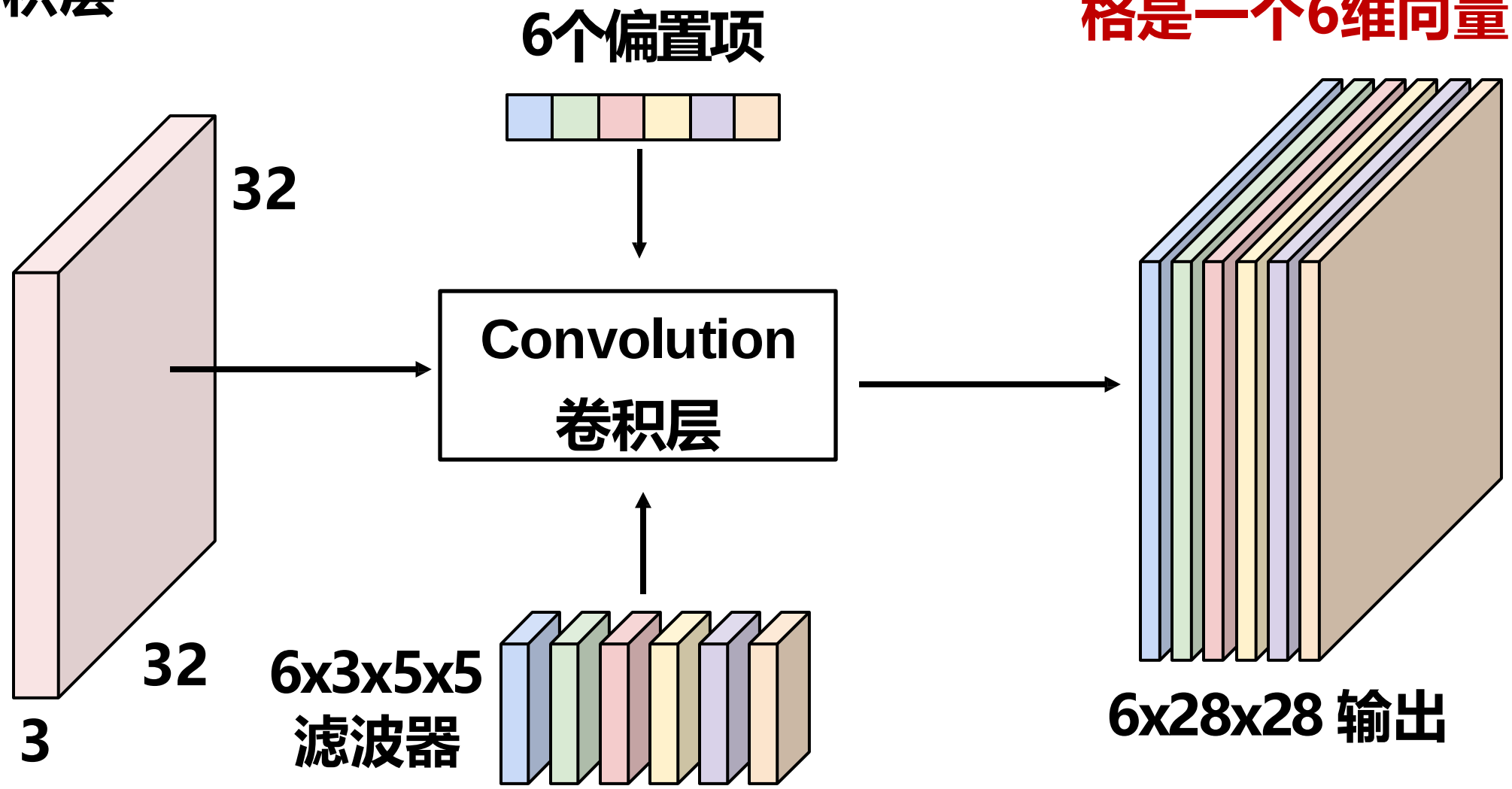
卷积神经网络

- 卷积层


6个偏置项
6个激活图，每个大小为1x28x28
32
32
3
Convolution
卷积层
6x3x5x5
滤波器
6x28x28 输出

卷积神经网络

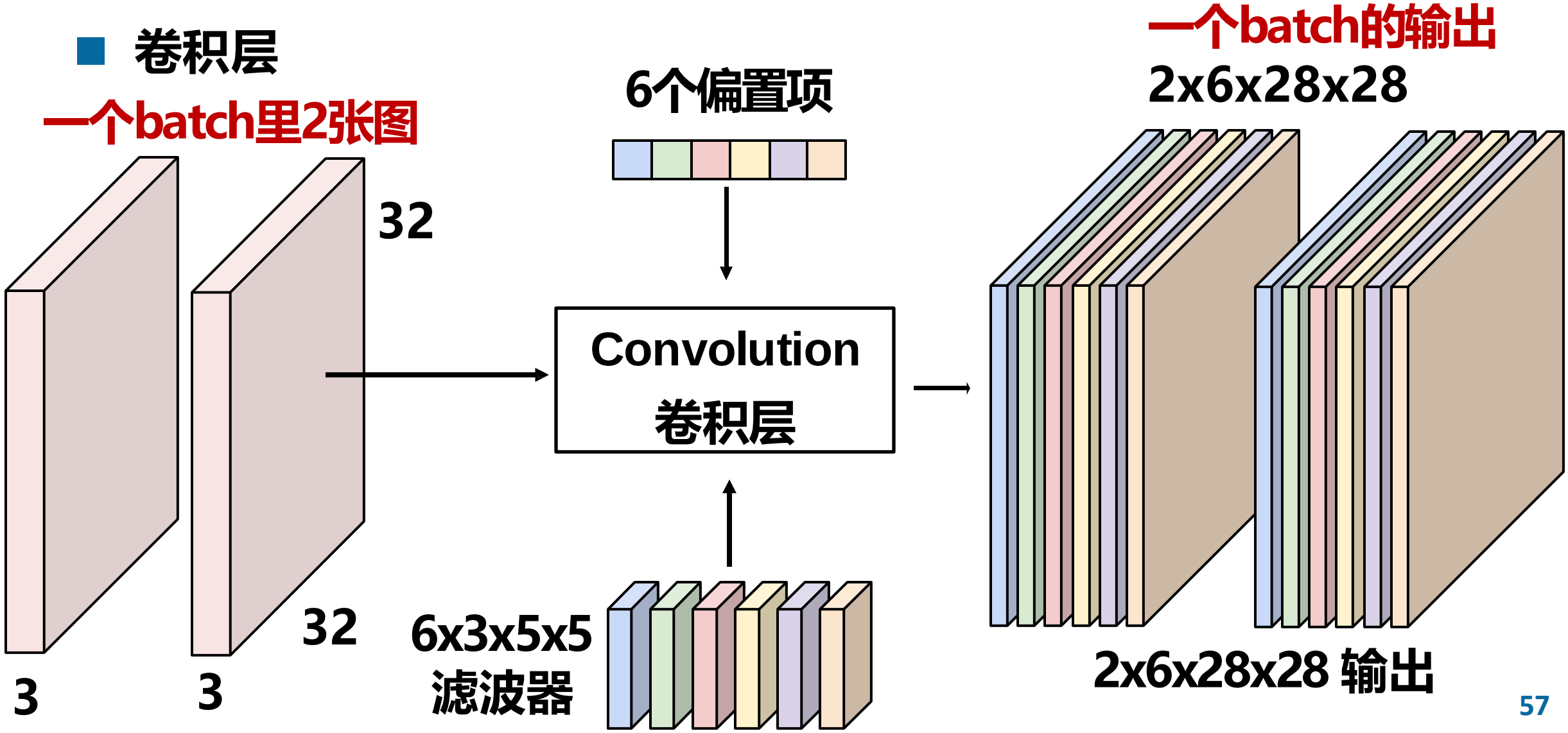
- 卷积层



28x28网格，每个网格是一个6维向量

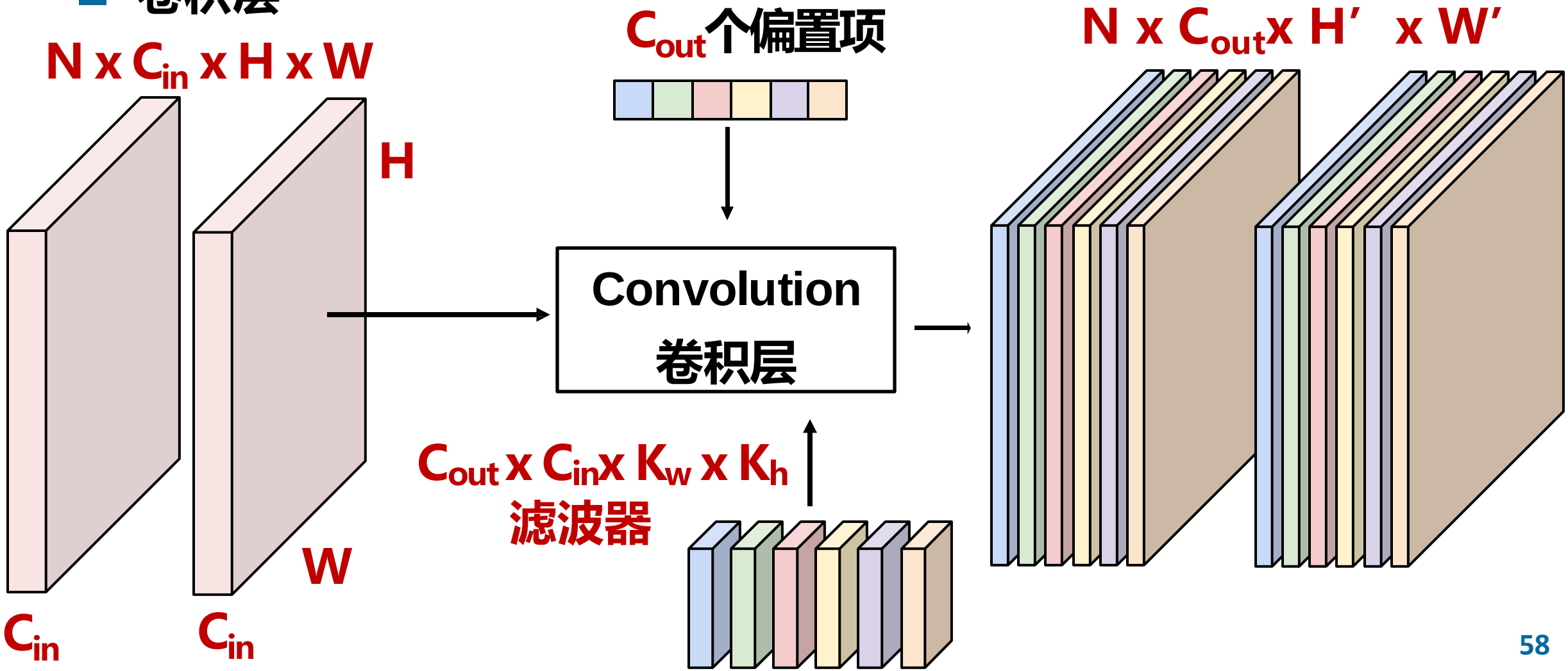
卷积神经网络

- 卷积层


一个batch里2张图
32
32
3
3
6个偏置项
Convolution
卷积层
6x3x5x5
滤波器
一个batch的输出
2x6x28x28
2x6x28x28 输出

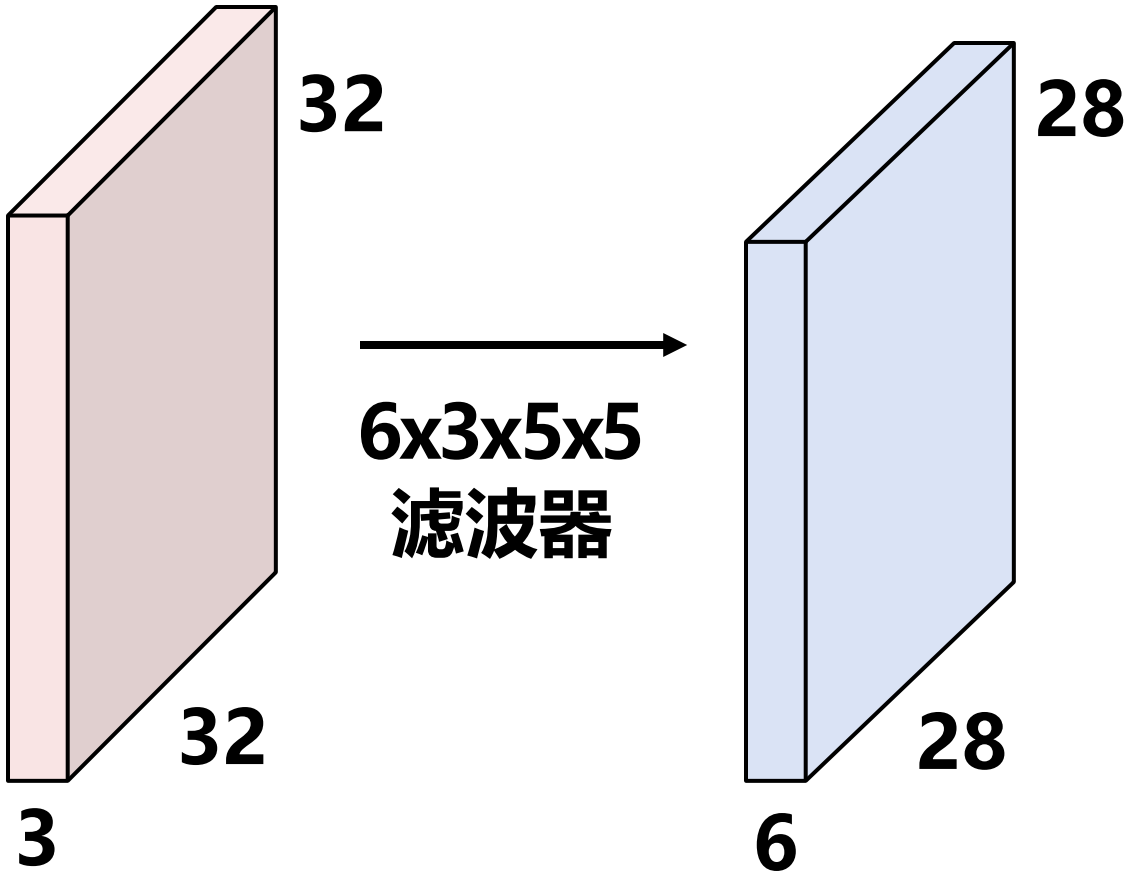
卷积神经网络

- 卷积层


N x C_{in} x H x W
H
W
C_{in}
C_{in}
C_{out}个偏置项
Convolution
卷积层
C_{out} x C_{in}x K_w x K_h
滤波器
N x C_{out}x H′ x W′

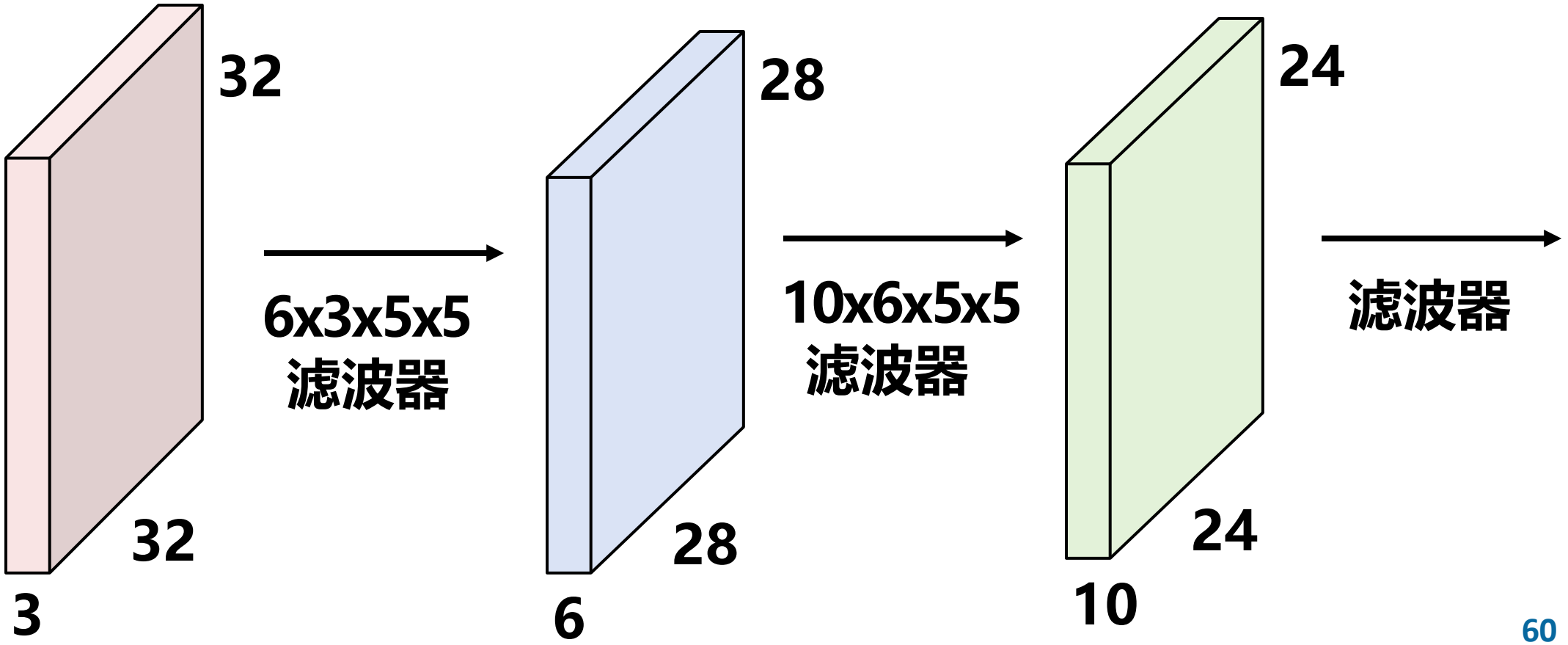
卷积神经网络

- 卷积神经网络由一系列卷积层组成



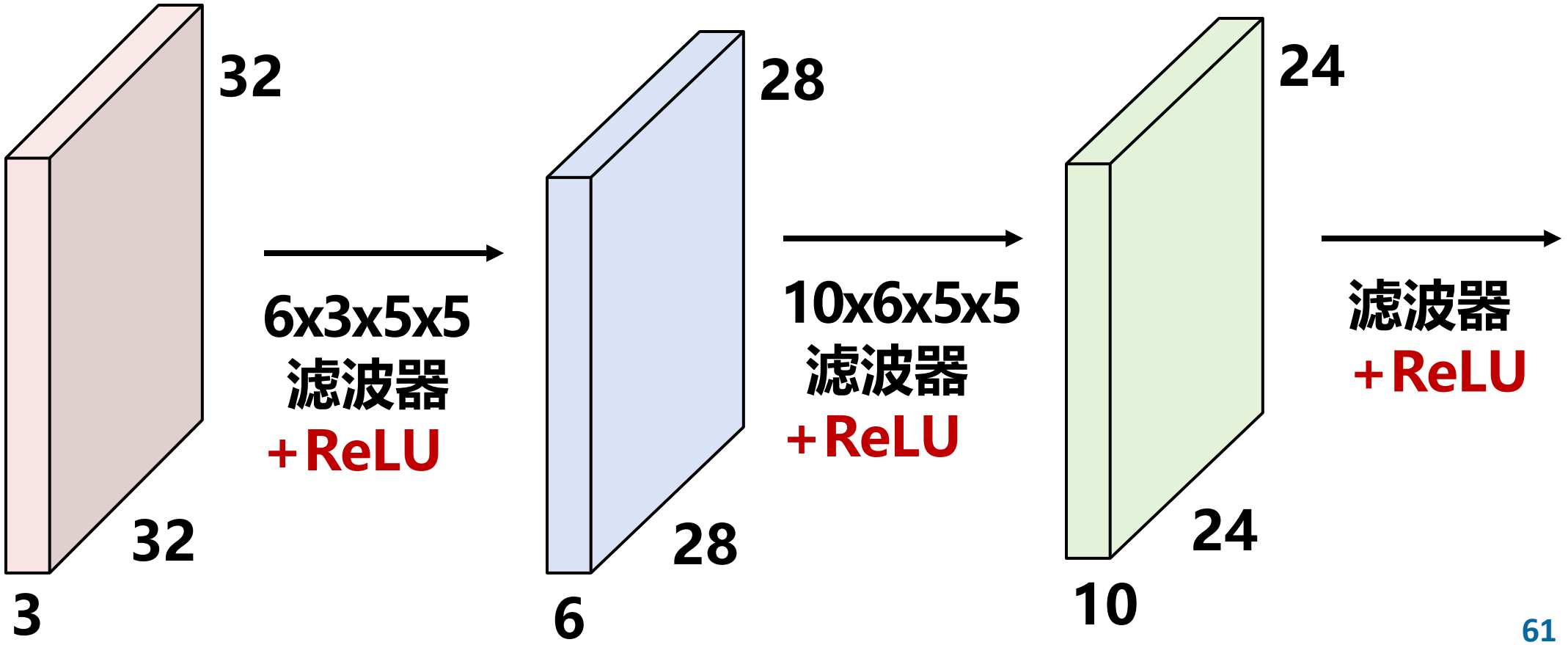
卷积神经网络

- 卷积神经网络由一系列卷积层组成



卷积神经网络

- 卷积神经网络由一系列卷积层组成



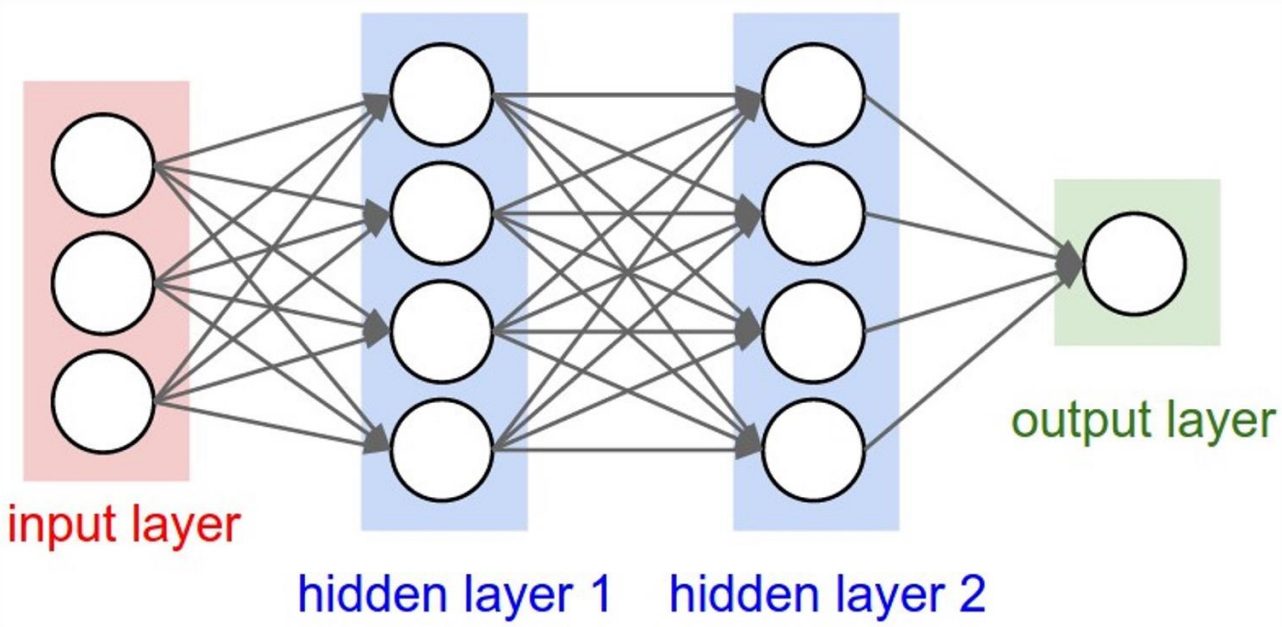
- 线性分类器学到了什么?

$$f(x,W) = Wx + b$$

每个类别的模板

卷积神经网络

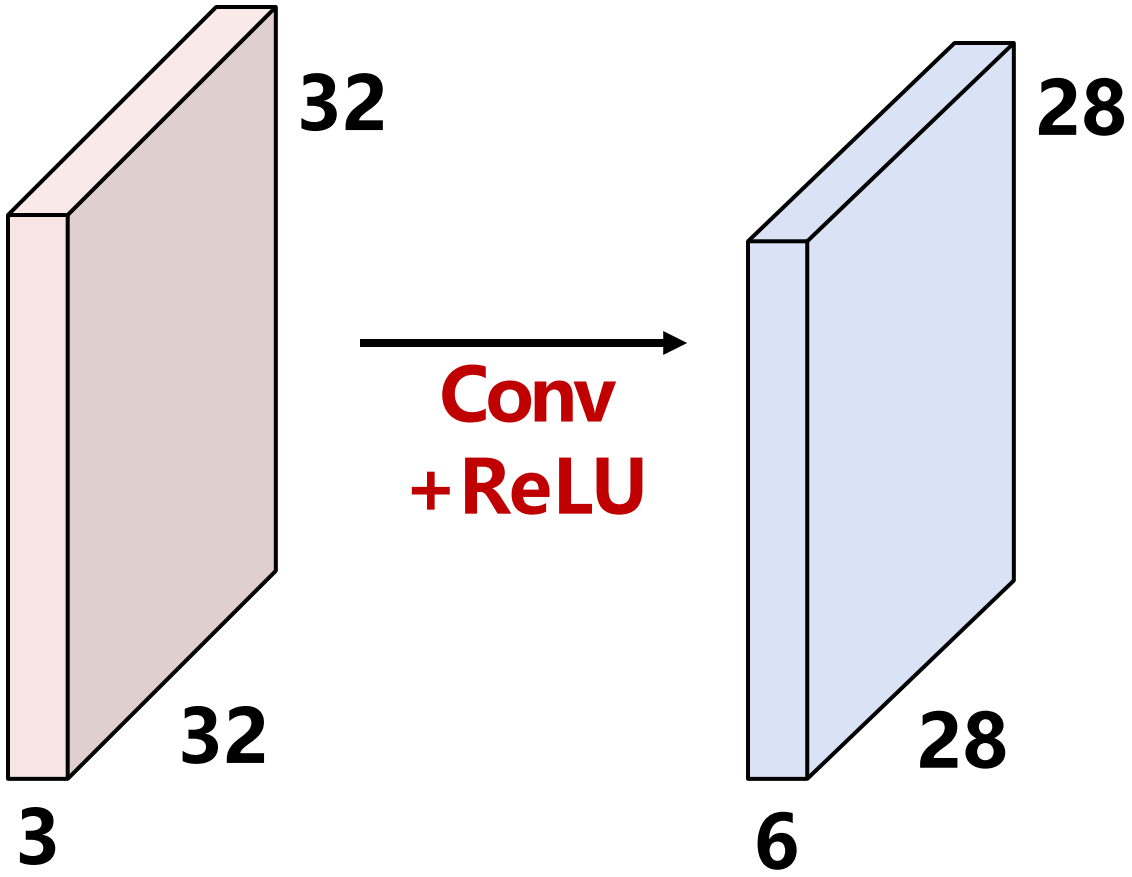
- 全连接神经网络学到了什么？



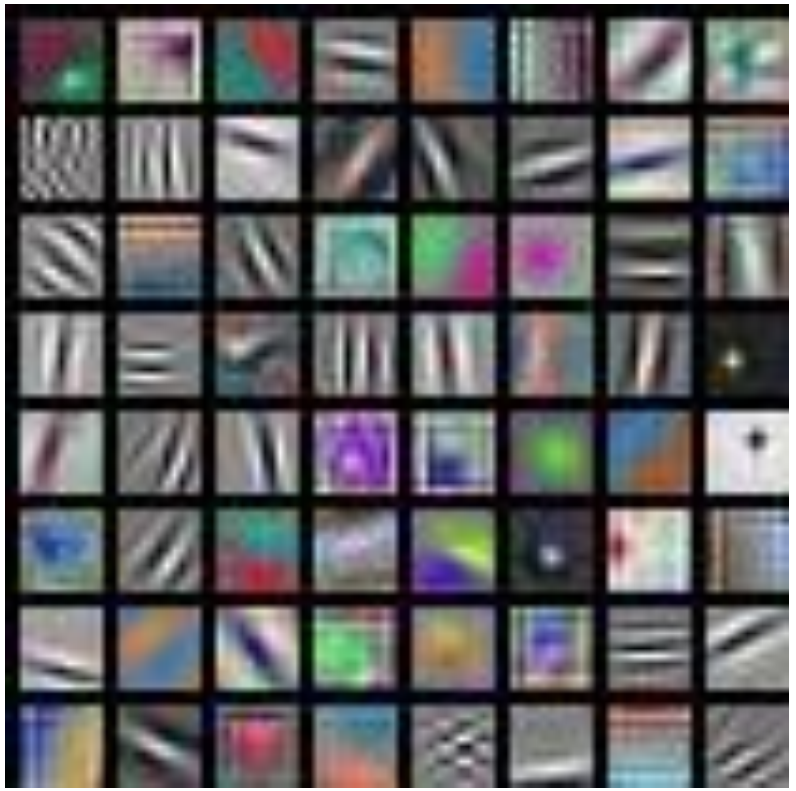
整张图像的的模板库

卷积神经网络

- 滤波器学到了什么?



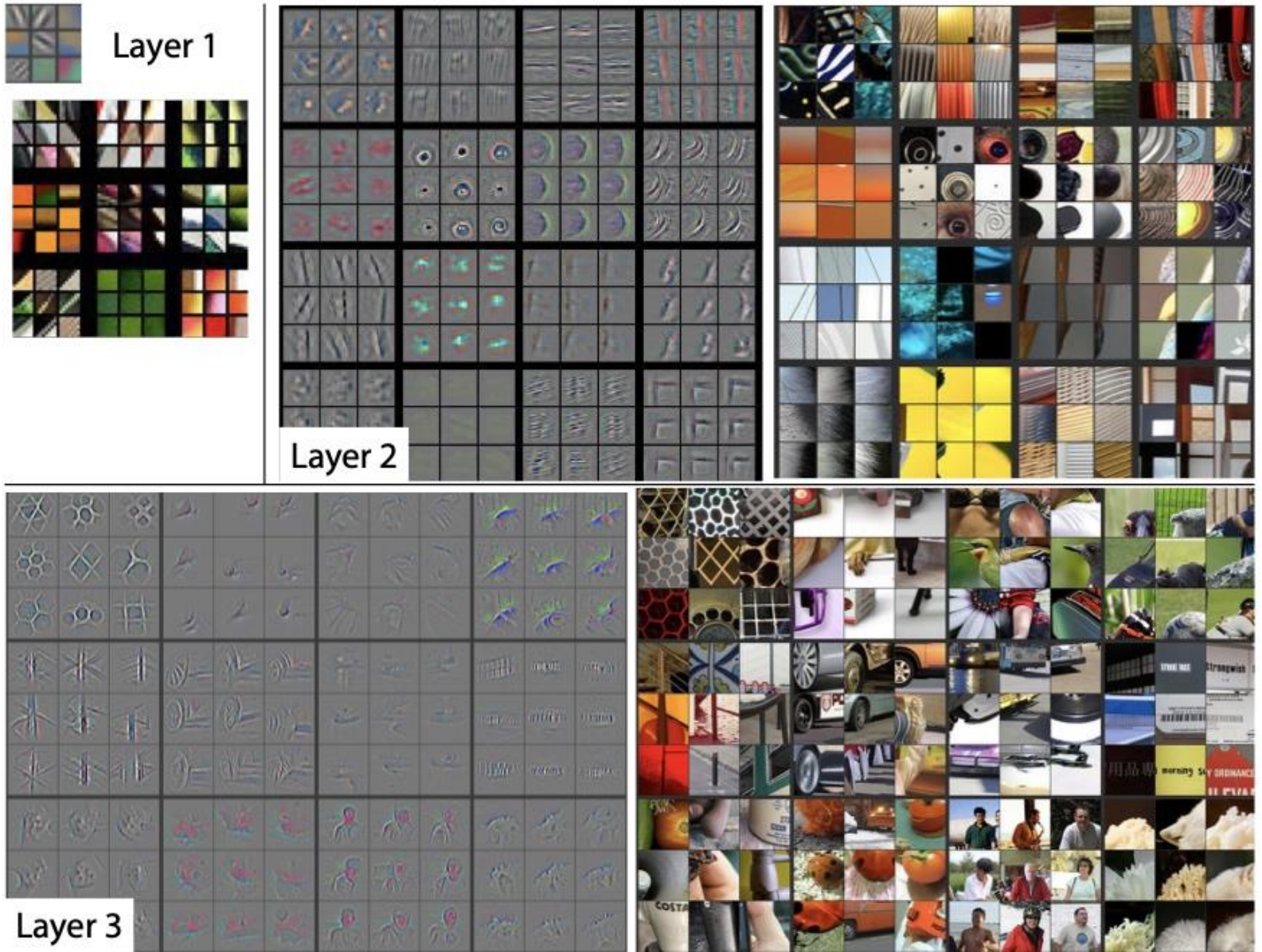
第一层卷积滤波器：局部图像模板
（通常学习边缘、颜色等纹理）



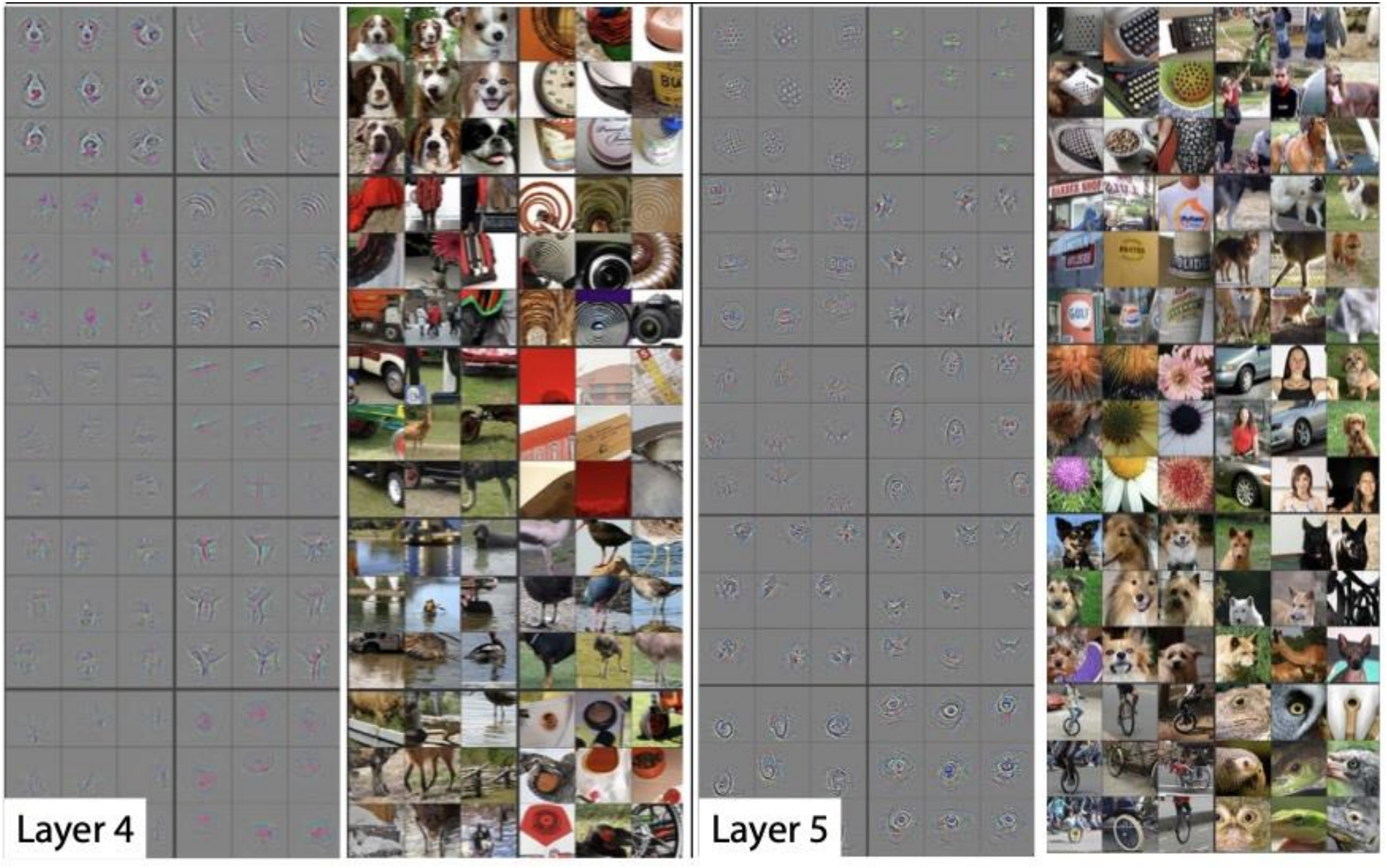
卷积神经网络

南京大学 NANJING UNIVERSITY

- 滤波器

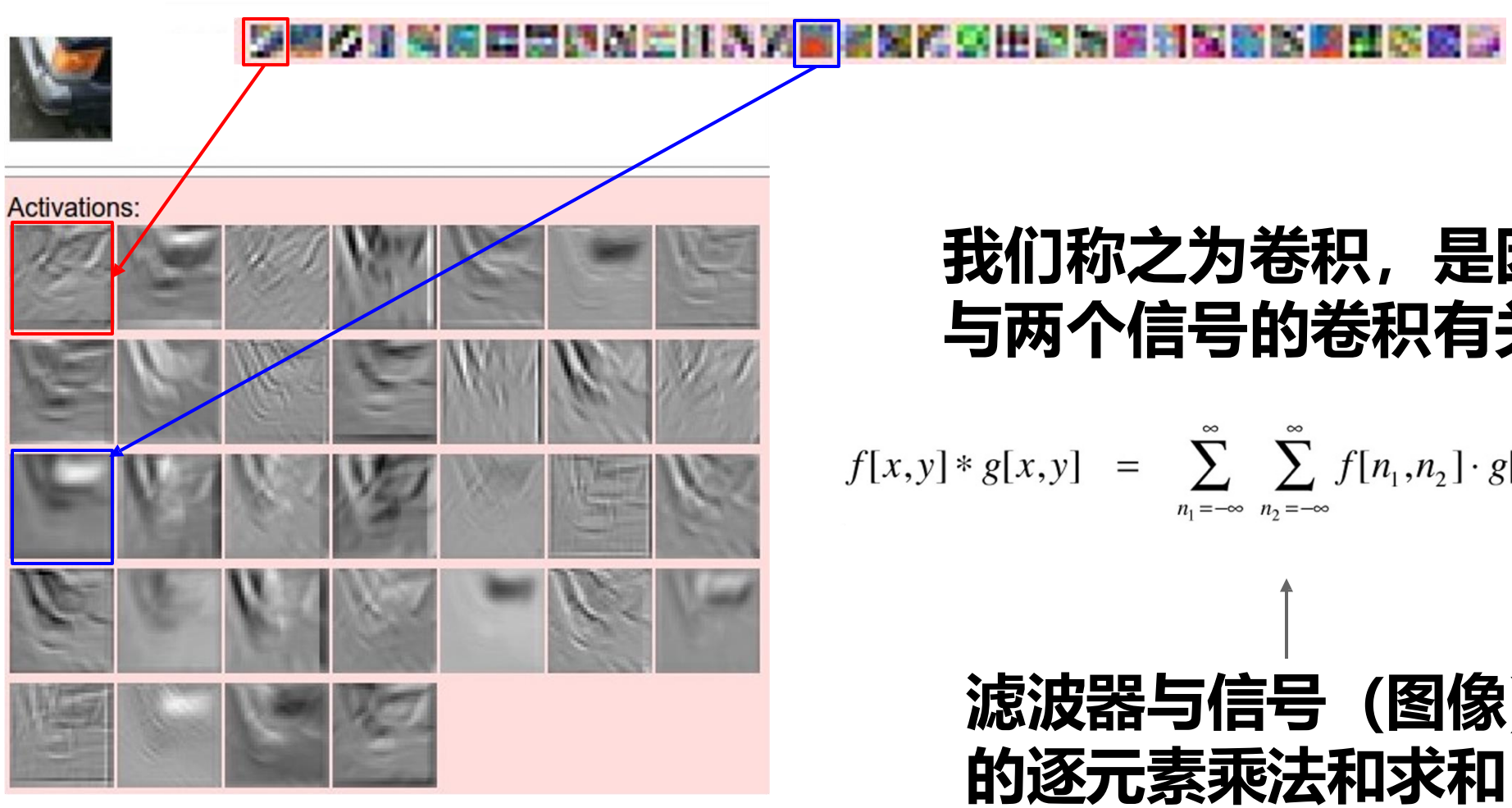


卷积神经网络



卷积神经网络

- 卷积神经网络　　一个滤波器-->一个激活图



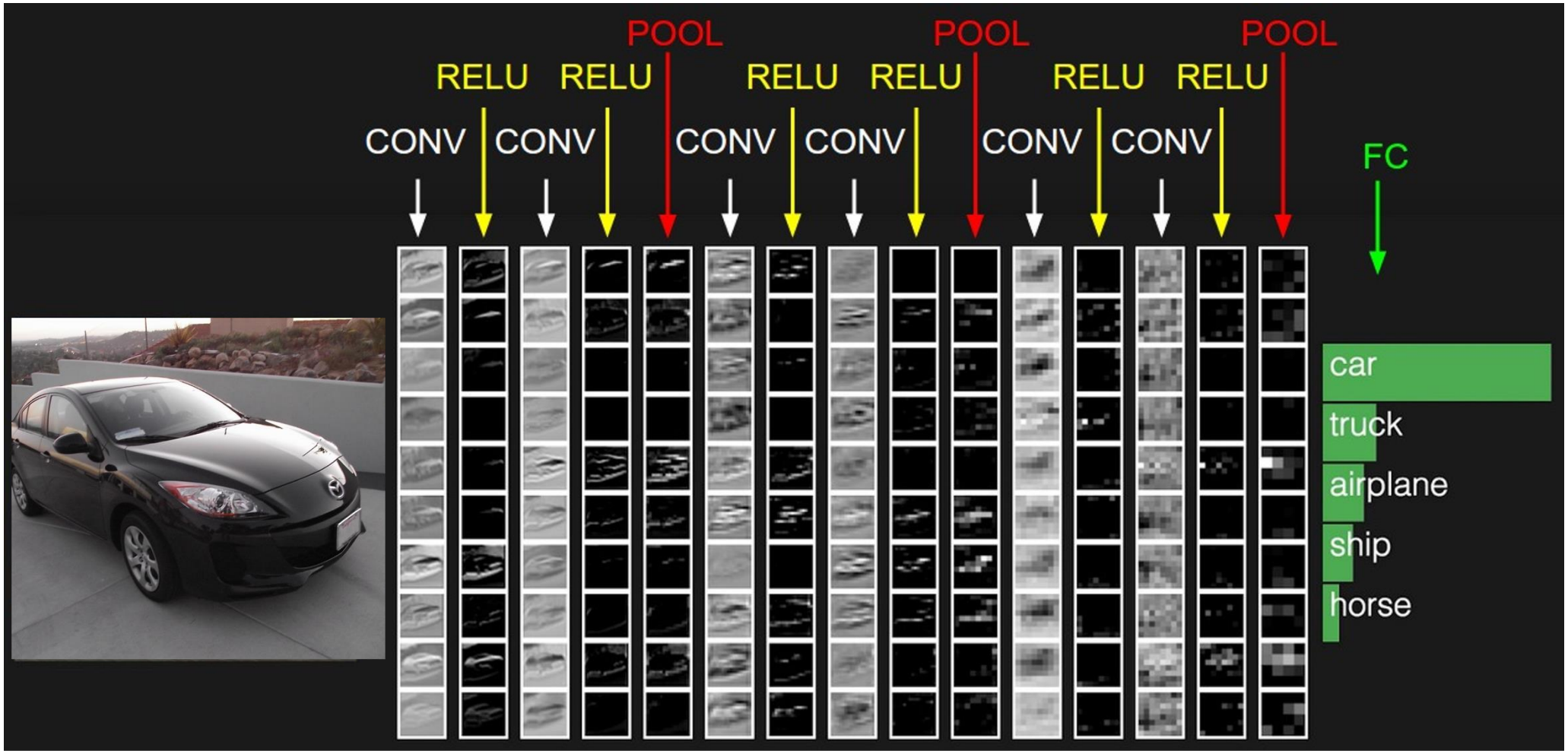
我们称之为卷积，是因为它与两个信号的卷积有关：

$$f[x,y] * g[x,y] = \sum_{n_1=-\infty}^{\infty} \sum_{n_2=-\infty}^{\infty} f[n_1,n_2] \cdot g[x-n_1, y-n_2]$$

滤波器与信号（图像）的逐元素乘法和求和

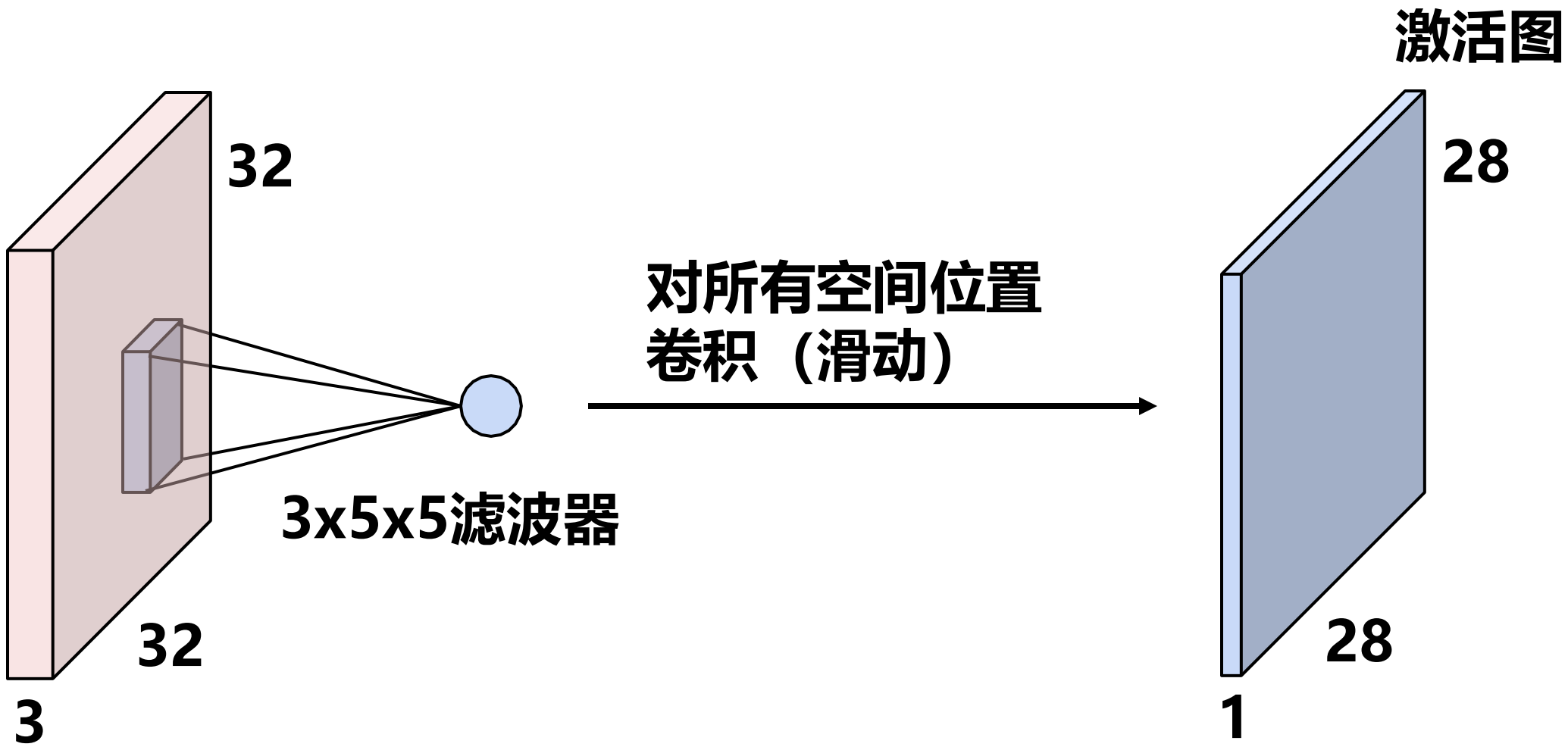
南京大学 NANJING UNIVERSITY

卷积神经网络



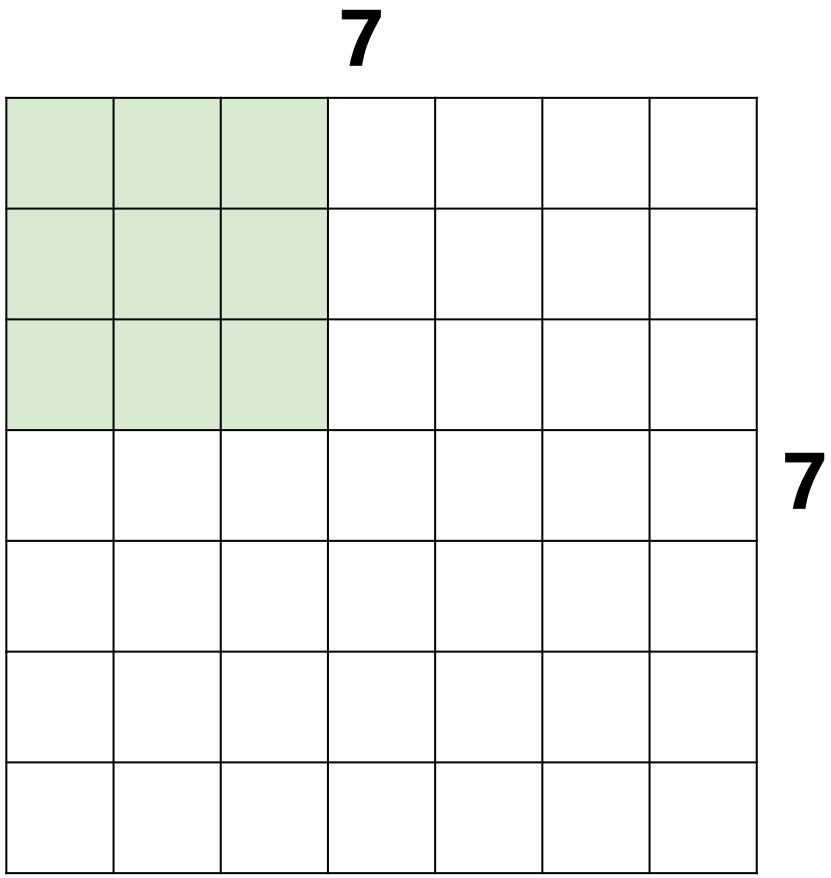
卷积神经网络

- 卷积输出的空间尺寸



南京大学 NANJING UNIVERSITY

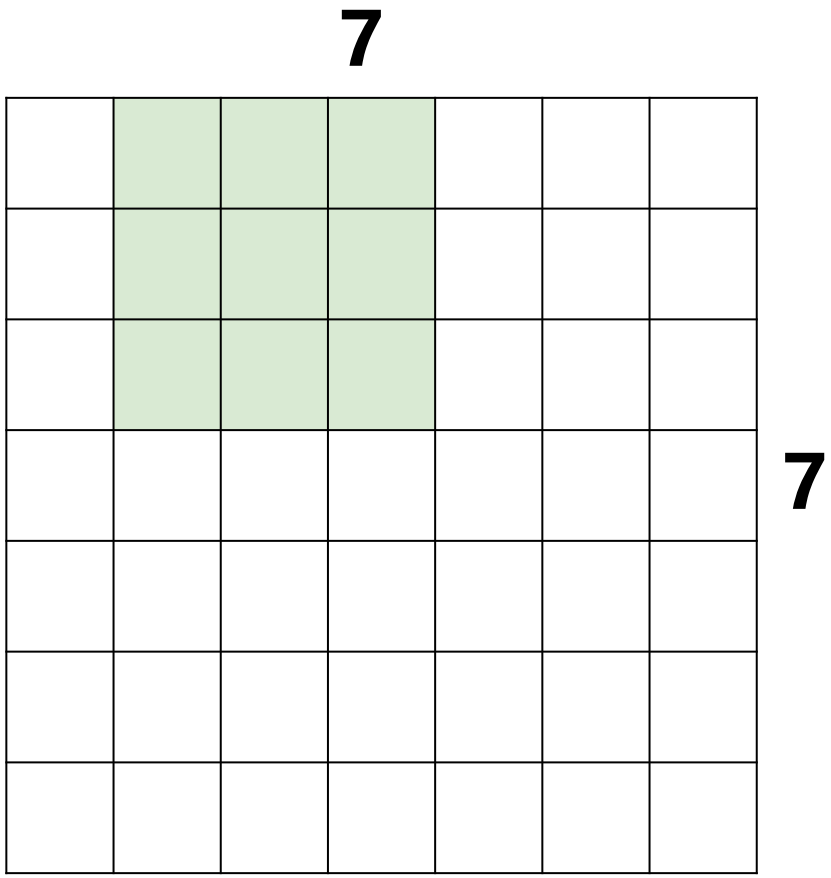
- 卷积输出的空间尺寸



7x7输入
3x3滤波器

南京大学 NANJING UNIVERSITY

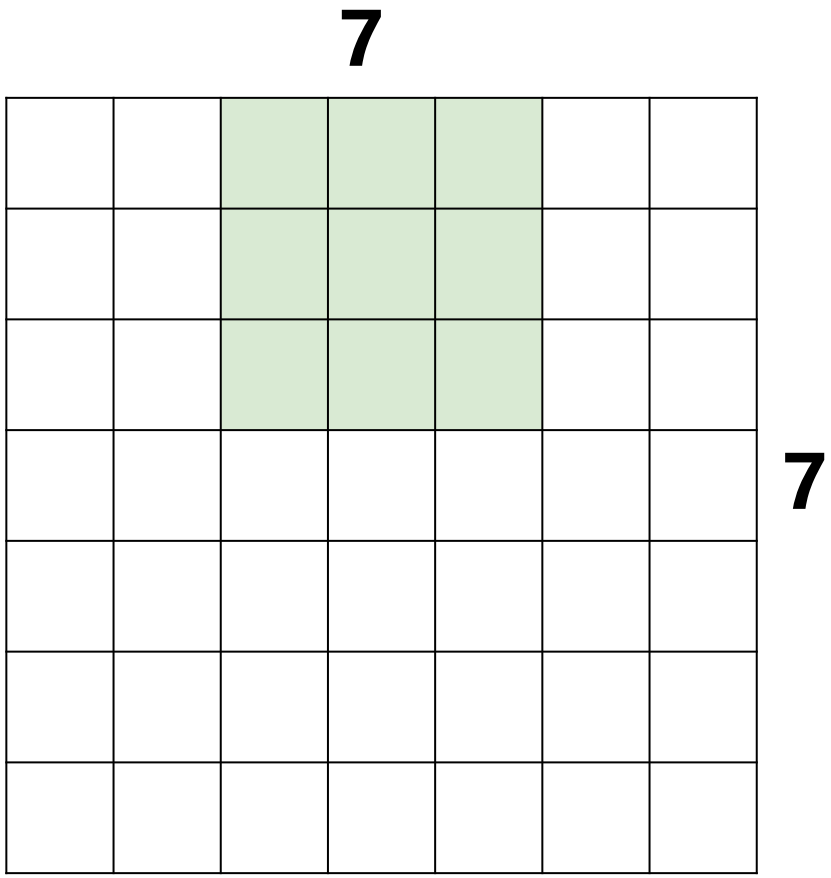
- 卷积输出的空间尺寸



7x7输入
3x3滤波器

南京大学 NANJING UNIVERSITY

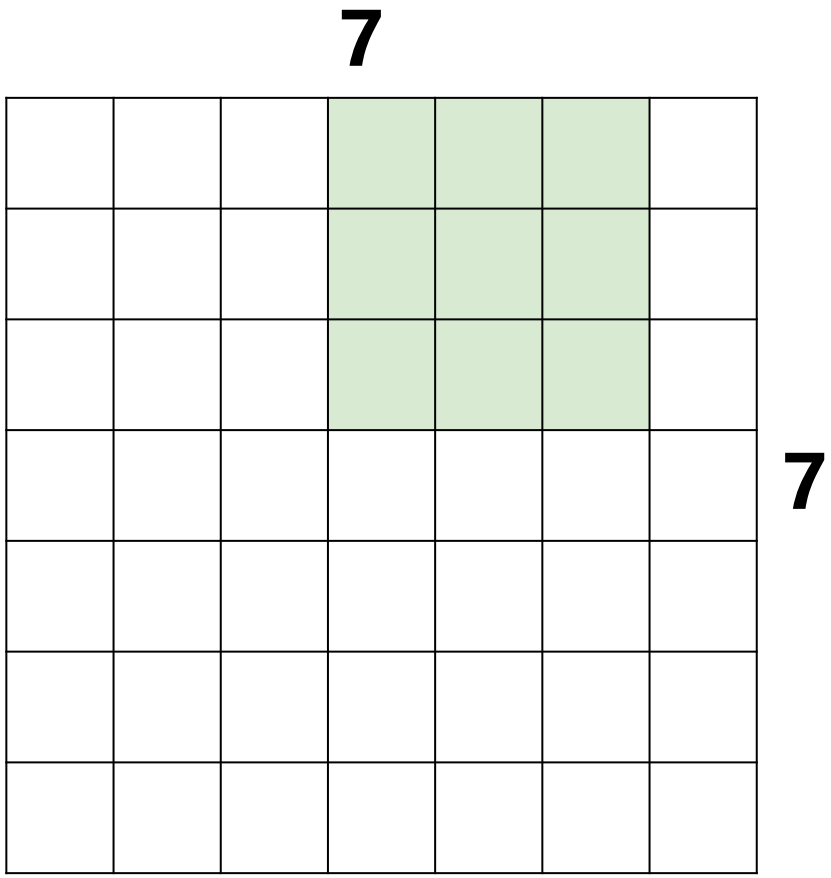
- 卷积输出的空间尺寸



7x7输入
3x3滤波器

南京大学 NANJING UNIVERSITY

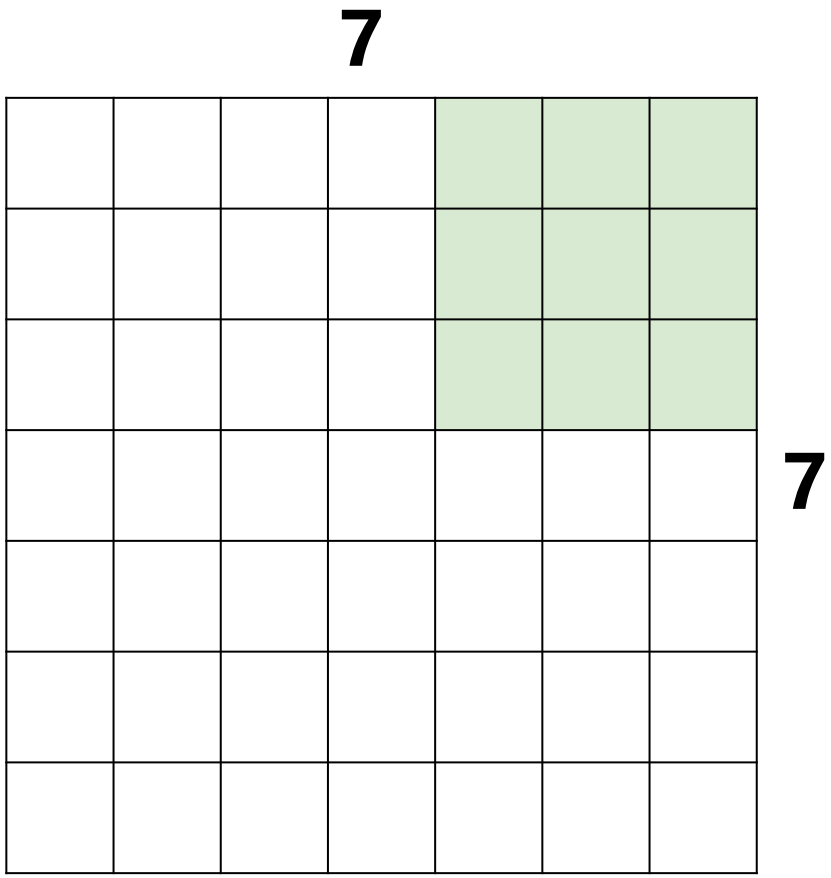
- 卷积输出的空间尺寸



7x7输入
3x3滤波器

南京大学 NANJING UNIVERSITY

卷积输出的空间尺寸

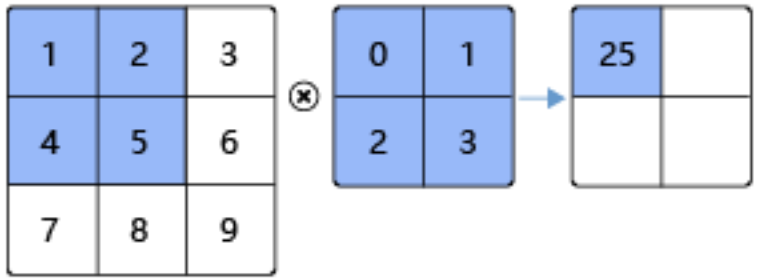


7x7输入
3x3滤波器

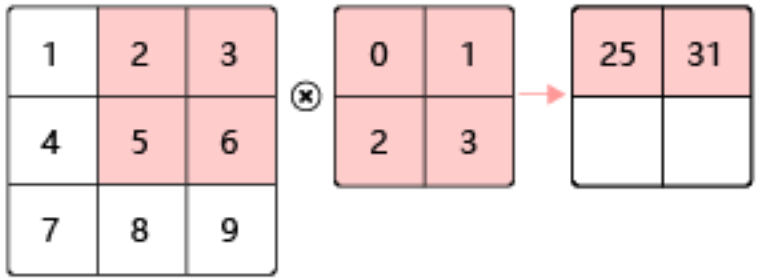
5x5输出

南京大学 NANJING UNIVERSITY

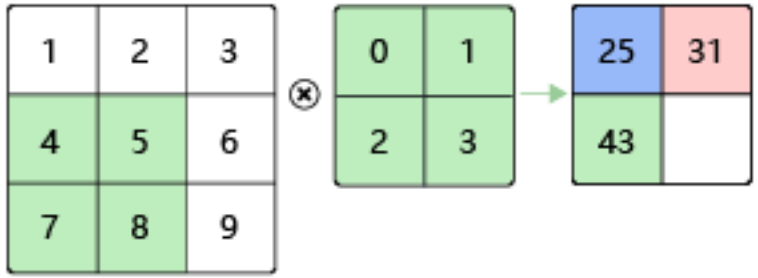
卷积层



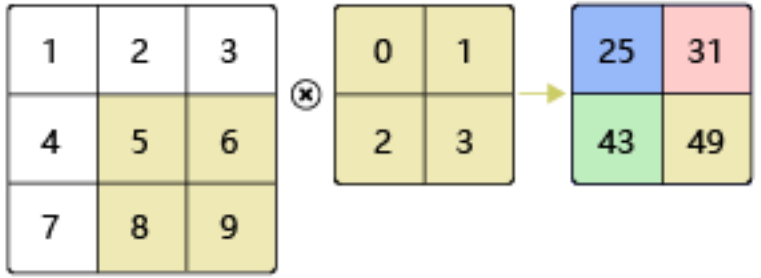
(a) 0x1+1x2+2x4+3x5=25



(b) 0x2+1x3+2x5+3x6=31



(c) 0x4+1x5+2x7+3x8=43

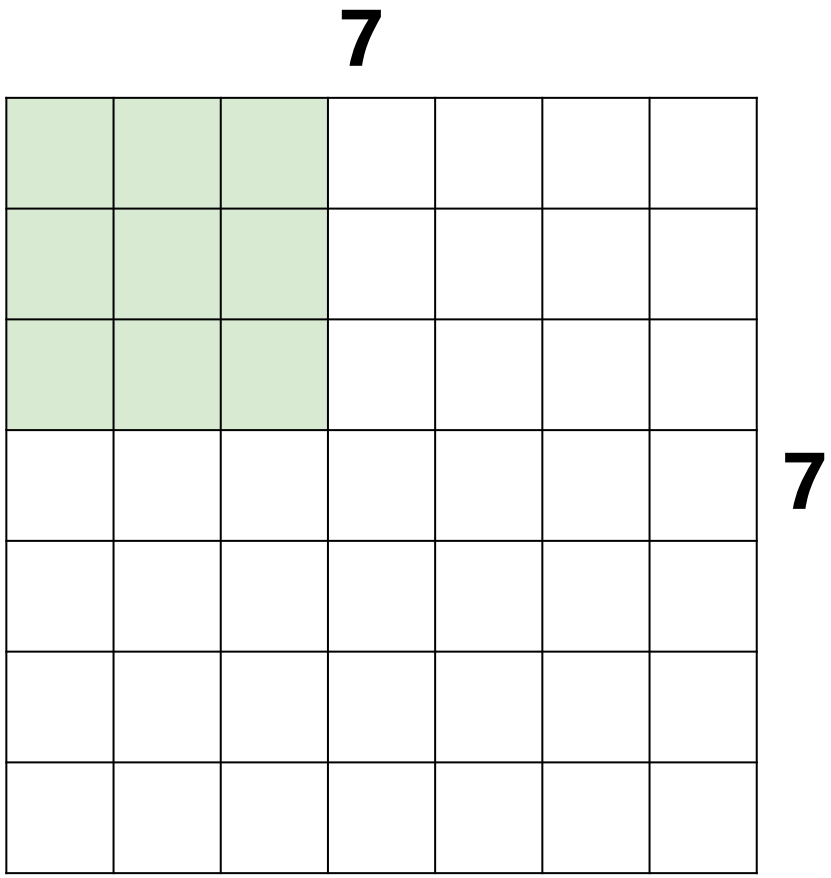


(d) 0x5+1x6+2x8+3x9=49

卷积神经网络

南京大学 NANJING UNIVERSITY

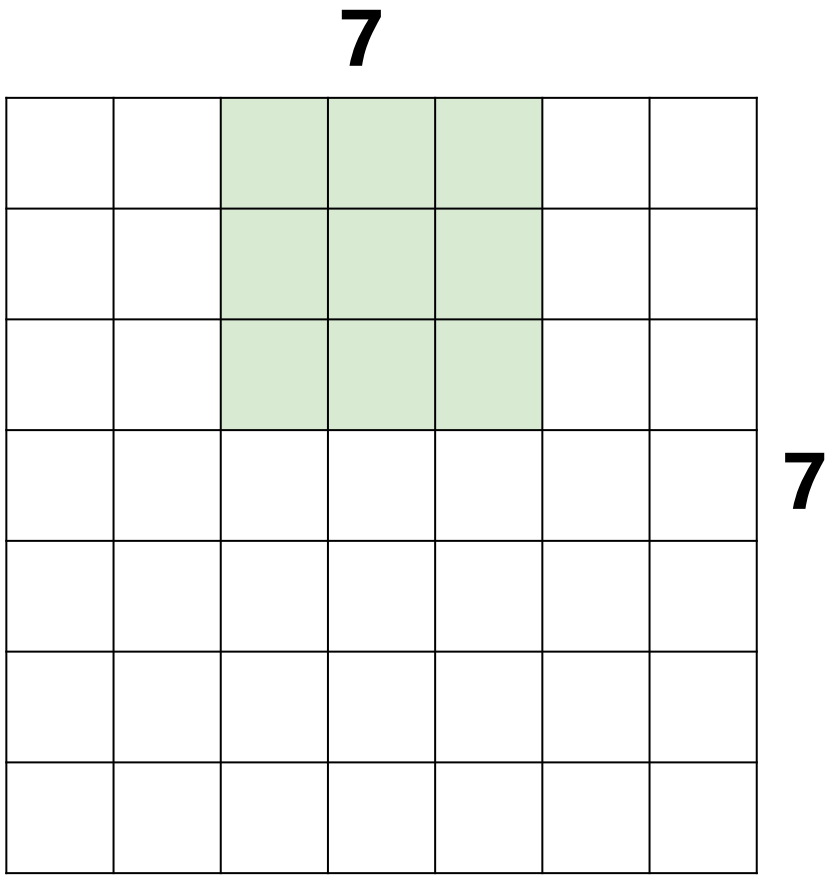
- 卷积输出的空间尺寸



7x7输入
3x3滤波器
2步长

南京大学 NANJING UNIVERSITY

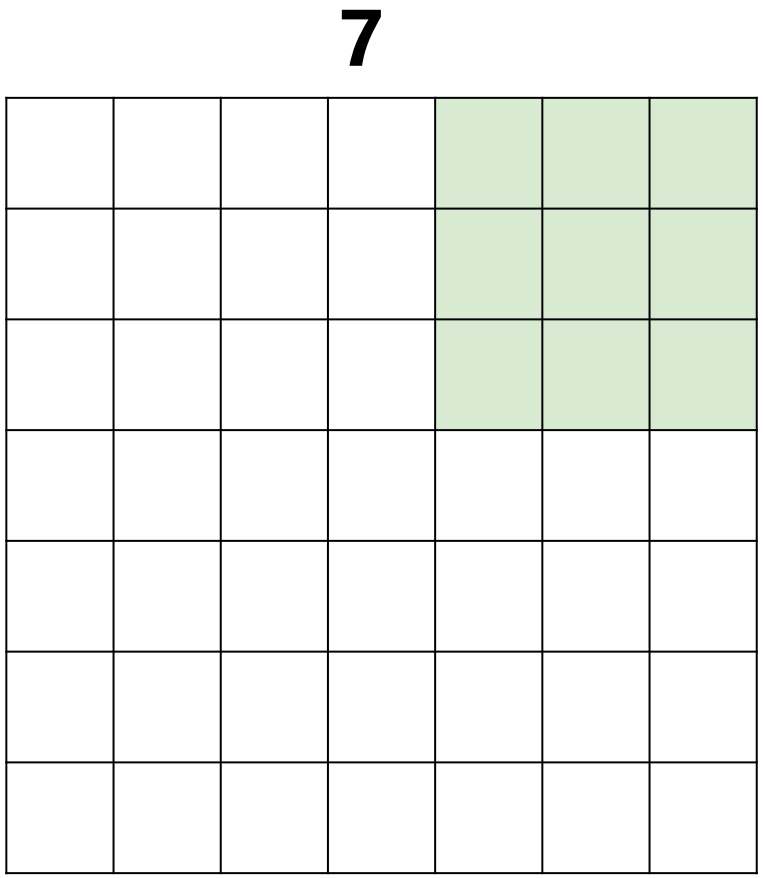
卷积输出的空间尺寸



7x7输入
3x3滤波器
2步长

卷积神经网络

- 卷积输出的空间尺寸

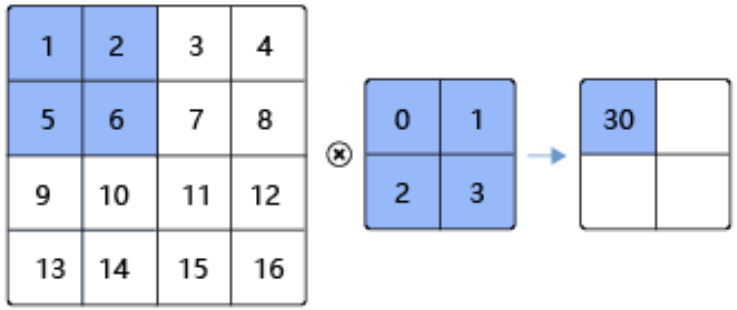


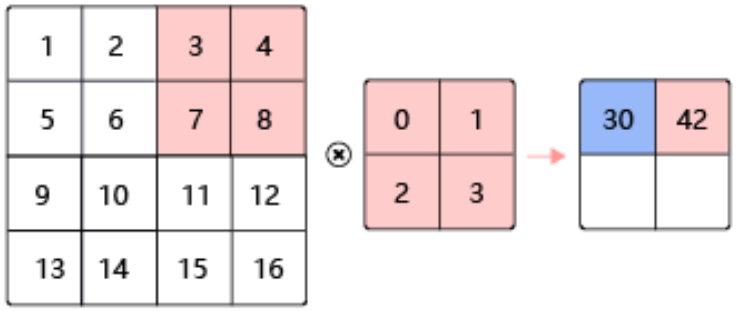
7x7输入
3x3滤波器
2步长

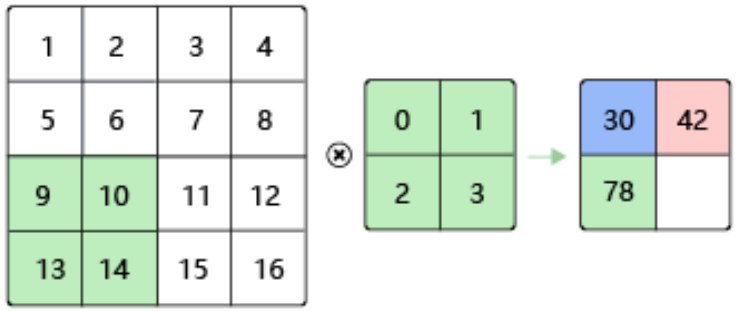
3x3输出

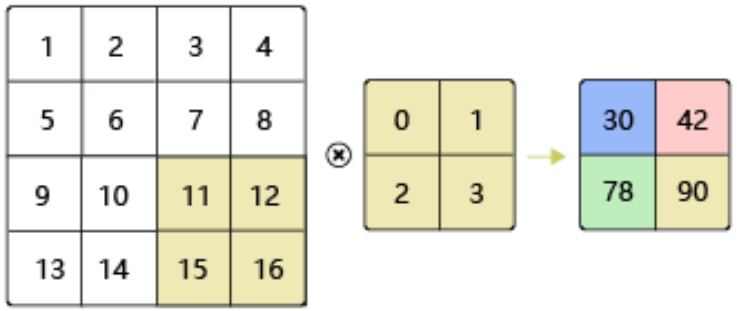
卷积神经网络

卷积层

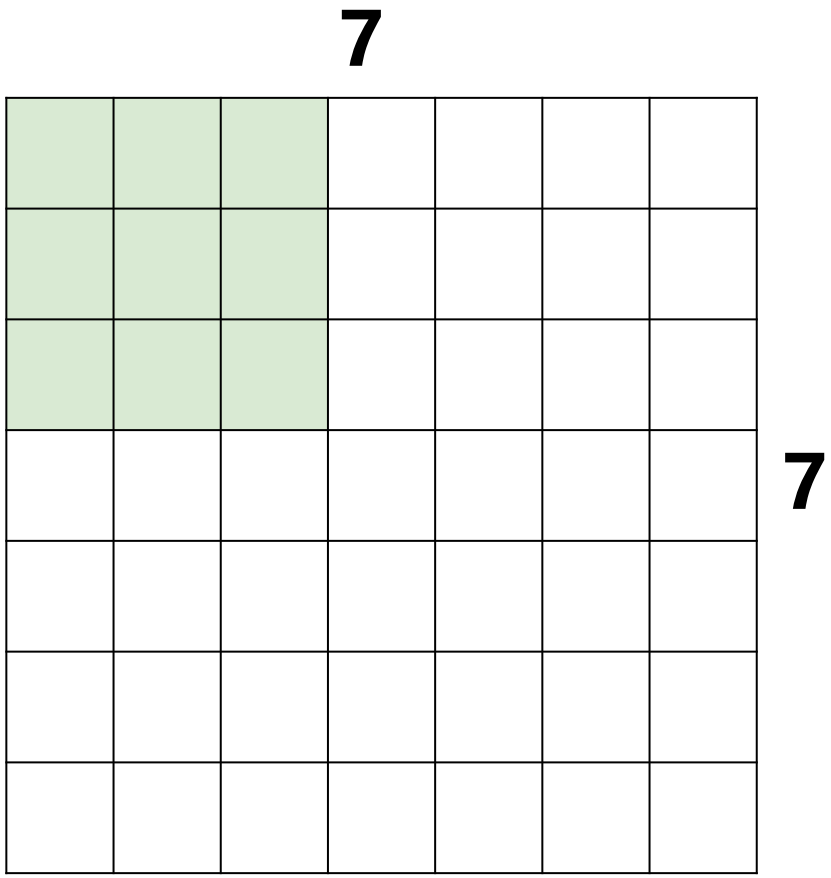

1 2 3 4
5 6 7 8
9 10 11 12
13 14 15 16
0 1
2 3
30


1 2 3 4
5 6 7 8
9 10 11 12
13 14 15 16
0 1
2 3
30 42


1 2 3 4
5 6 7 8
9 10 11 12
13 14 15 16
0 1
2 3
30 42
78


1 2 3 4
5 6 7 8
9 10 11 12
13 14 15 16
0 1
2 3
30 42
78 90

- 卷积输出的空间尺寸



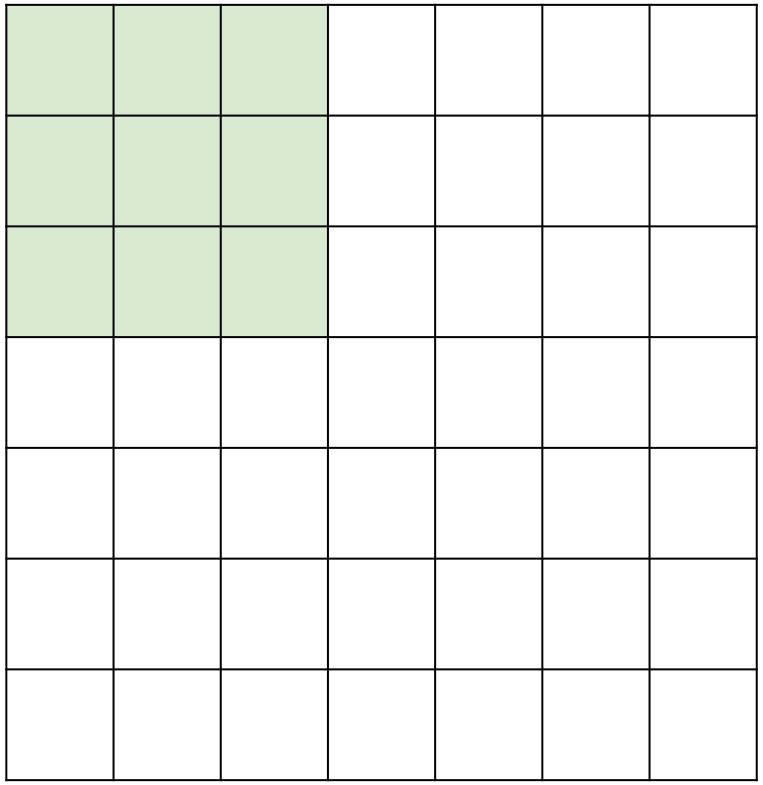
7x7输入
3x3滤波器
3步长?

不可以!
尺寸不匹配

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积输出的空间尺寸



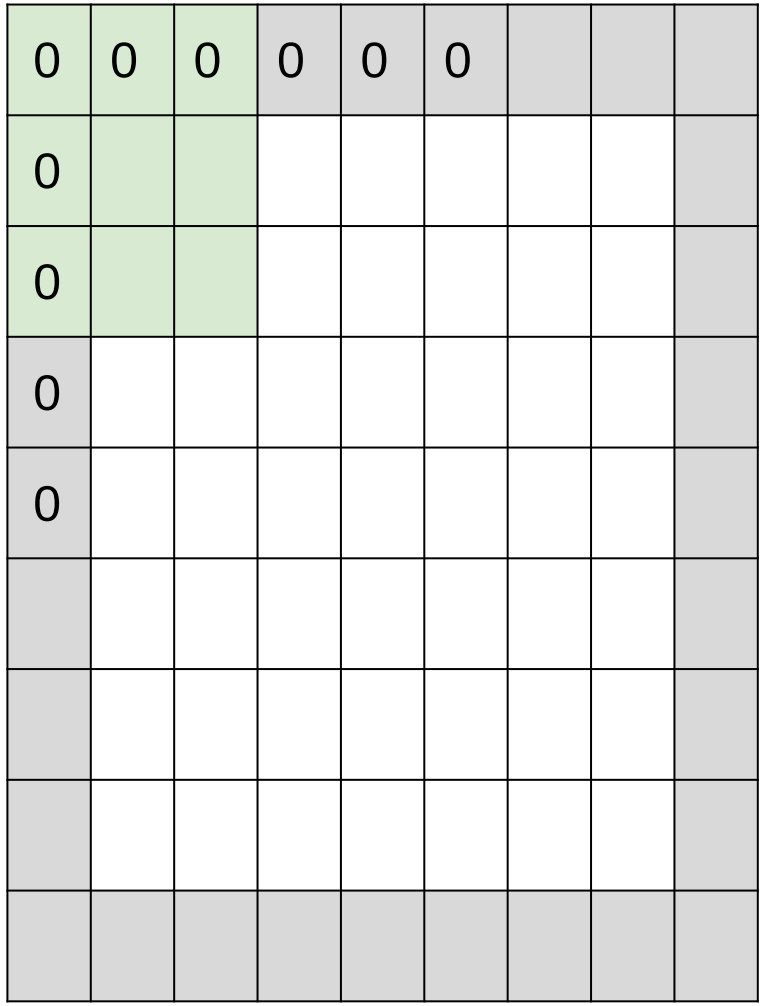
NxN 输入
FxF 滤波器
S 步长
输出大小：(N-F)/S+1

例如：N = 7, F = 3:
步长 1 => (7 - 3)/1 + 1 = 5
步长 2 => (7 - 3)/2 + 1 = 3
步长 3 => (7 - 3)/3 + 1 = 2.33

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积输出的空间尺寸



7x7 输入 NxN
3x3 滤波器 FxF
1 步长 S
1 Padding P
输出尺寸?

南京大学 NANJING UNIVERSITY

卷积输出的空间尺寸

0	0	0	0	0	0
0	13	14	15	16	0
0	9	10	11	12	0
0	5	6	7	8	0
0	1	2	3	4	0
0	0	0	0	0	0

(a)padding=1

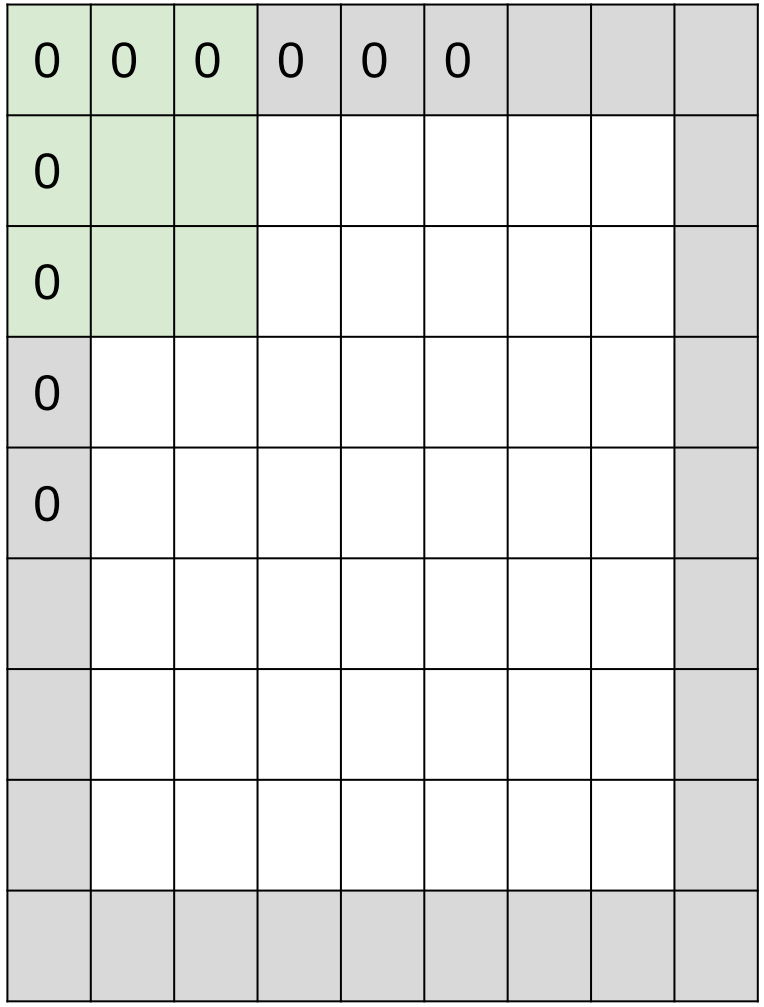
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0
0	0	13	14	15	16	0	0
0	0	9	10	11	12	0	0
0	0	5	6	7	8	0	0
0	0	1	2	3	4	0	0
0	0	0	0	0	0	0	0
0	0	0	0	0	0	0	0

(b)padding=2

卷积神经网络

南京大学 NANJING UNIVERSITY

卷积输出的空间尺寸



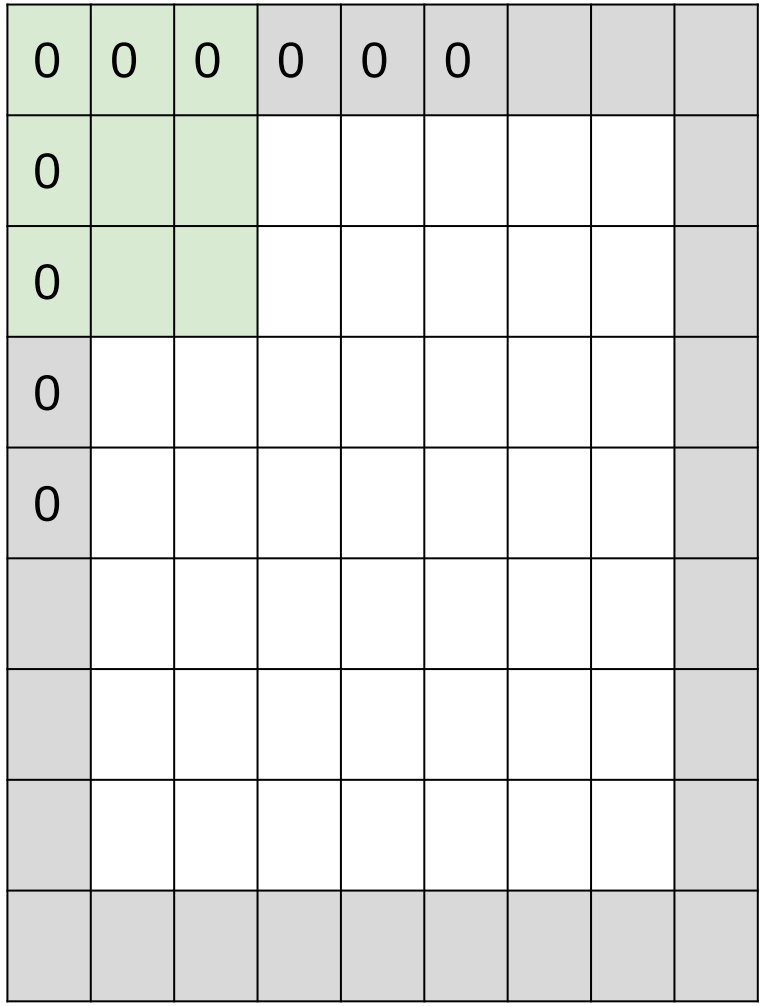
7x7 输入 NxN
3x3 滤波器 FxF
1 步长 S
1 Padding P
7x7 输出

输出：(N+2P-F)/S+1

卷积神经网络

南京大学 NANJING UNIVERSITY

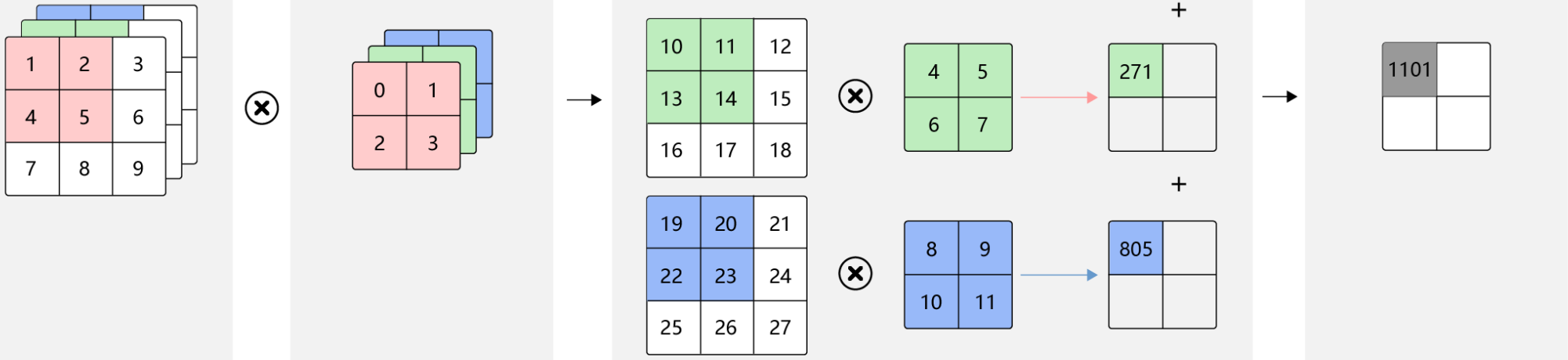
卷积输出的空间尺寸



为了保持输出的空间尺寸大小：
CONV层步幅为1
滤波器大小为FxF
零填充为（F-1）/2

F = 3 => zero pad with 1
F = 5 => zero pad with 2
F = 7 => zero pad with 3

多输入通道卷积



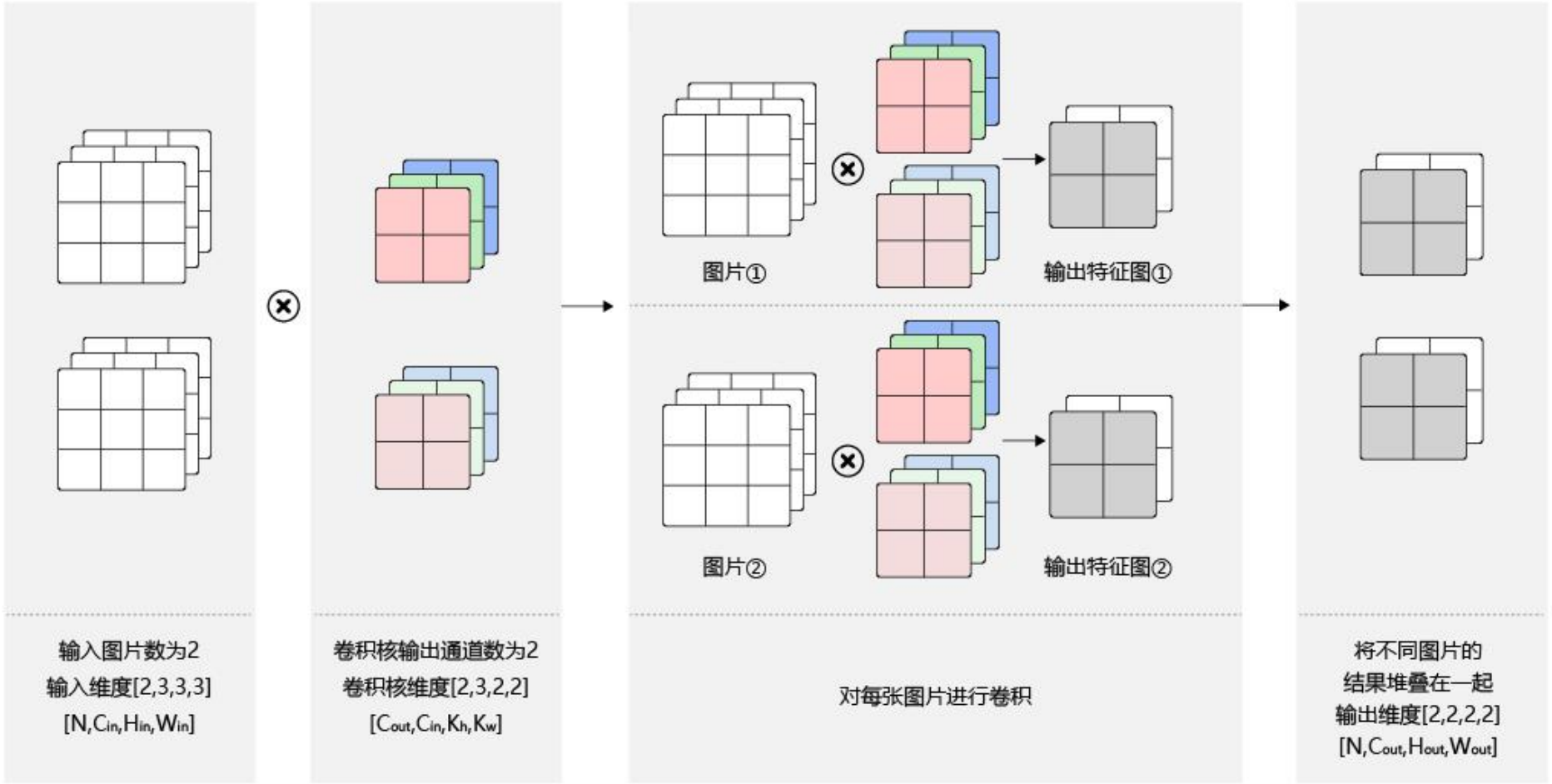
输入通道数为3
输入维度[3,3,3]
$[C_{in},H_{in},W_{in}]$

卷积核输入通道数为3
卷积核维度[3,2,2]
$[C_{in},K_h,K_w]$

对每个输入通道分别计算，
并将每个通道的输出结果求和

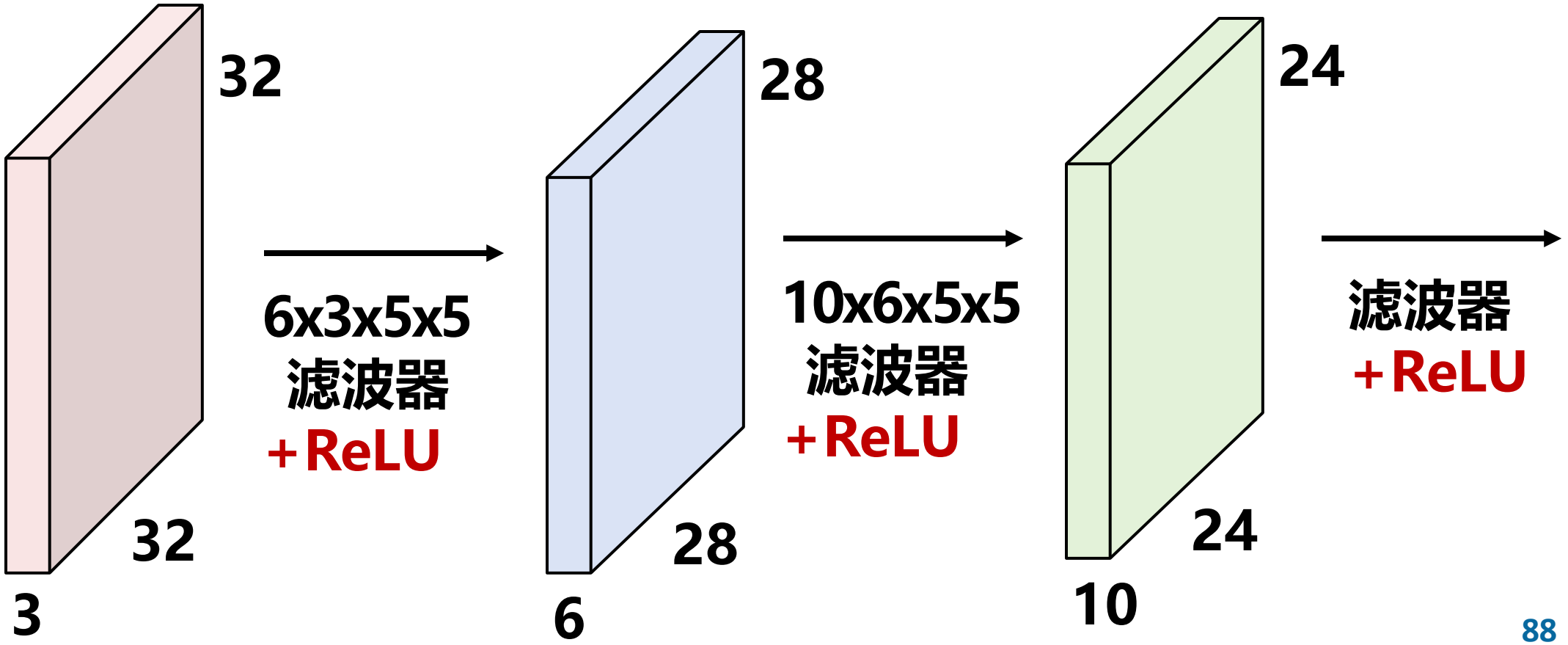
最终输出结果
输出维度[2,2]
$[H_{out},W_{out}]$

多输出通道卷积


图片①
输出特征图①
图片②
输出特征图②
输入图片数为2
输入维度[2,3,3,3]
[N,Cin,Hin,Win]
卷积核输出通道数为2
卷积核维度[2,3,2,2]
[Cout,Cin,Kh,Kw]
对每张图片进行卷积
将不同图片的
结果堆叠在一起
输出维度[2,2,2,2]
[N,Cout,Hout,Wout]

卷积神经网络

- 激活图尺寸：32x32→28x28→24x24



卷积神经网络

- 卷积输出的空间尺寸
- 假设：
  - 输入：3x32x32
  - 滤波器：10x5x5
  - 步长：1
  - Padding：2
- 输出多少?

卷积神经网络

- 卷积输出的空间尺寸
- 假设：
  - 输入：3x32x32
  - 滤波器：10x5x5
  - 步长：1
  - Padding：2
- 输出：32x32x10，即 (32+2*2-5)/1+1 = 32

卷积神经网络

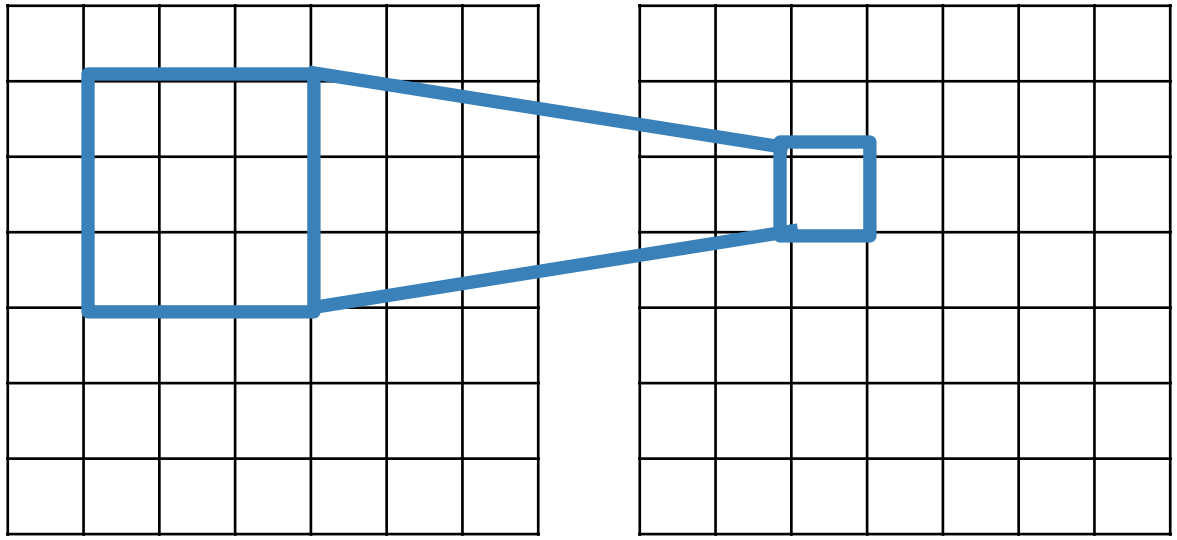
- 卷积输出的空间尺寸
- 假设：
  - 输入：3x32x32
  - 滤波器：10x5x5
  - 步长：1
  - Padding：2
- 参数量?

卷积神经网络

- 卷积输出的空间尺寸
- 假设：
  - 输入：3x32x32
  - 滤波器：10x5x5
  - 步长：1
  - Padding：2
- 参数量：(5x5x3+1)x10

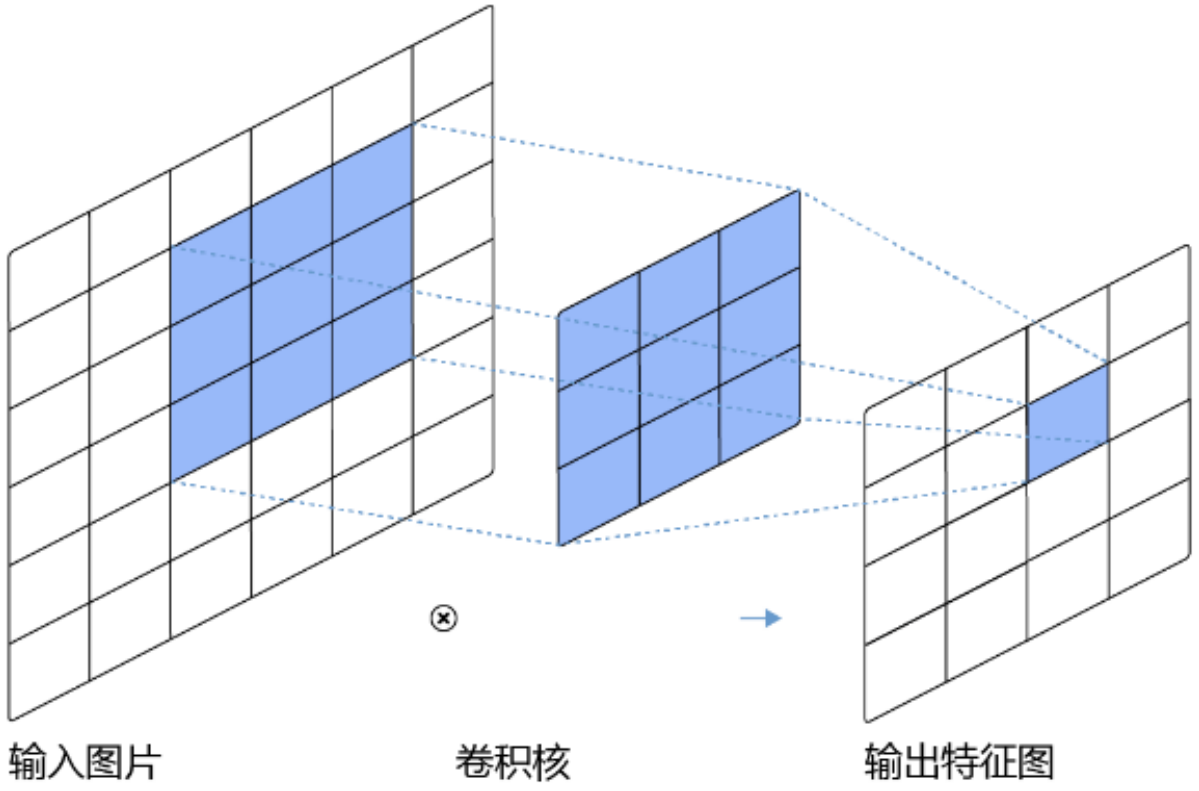
卷积神经网络

- 卷积层的感受野
- 对于核大小为K的卷积，输出中的每个元素都取决于输入中的K x K感受野



卷积层的感受野

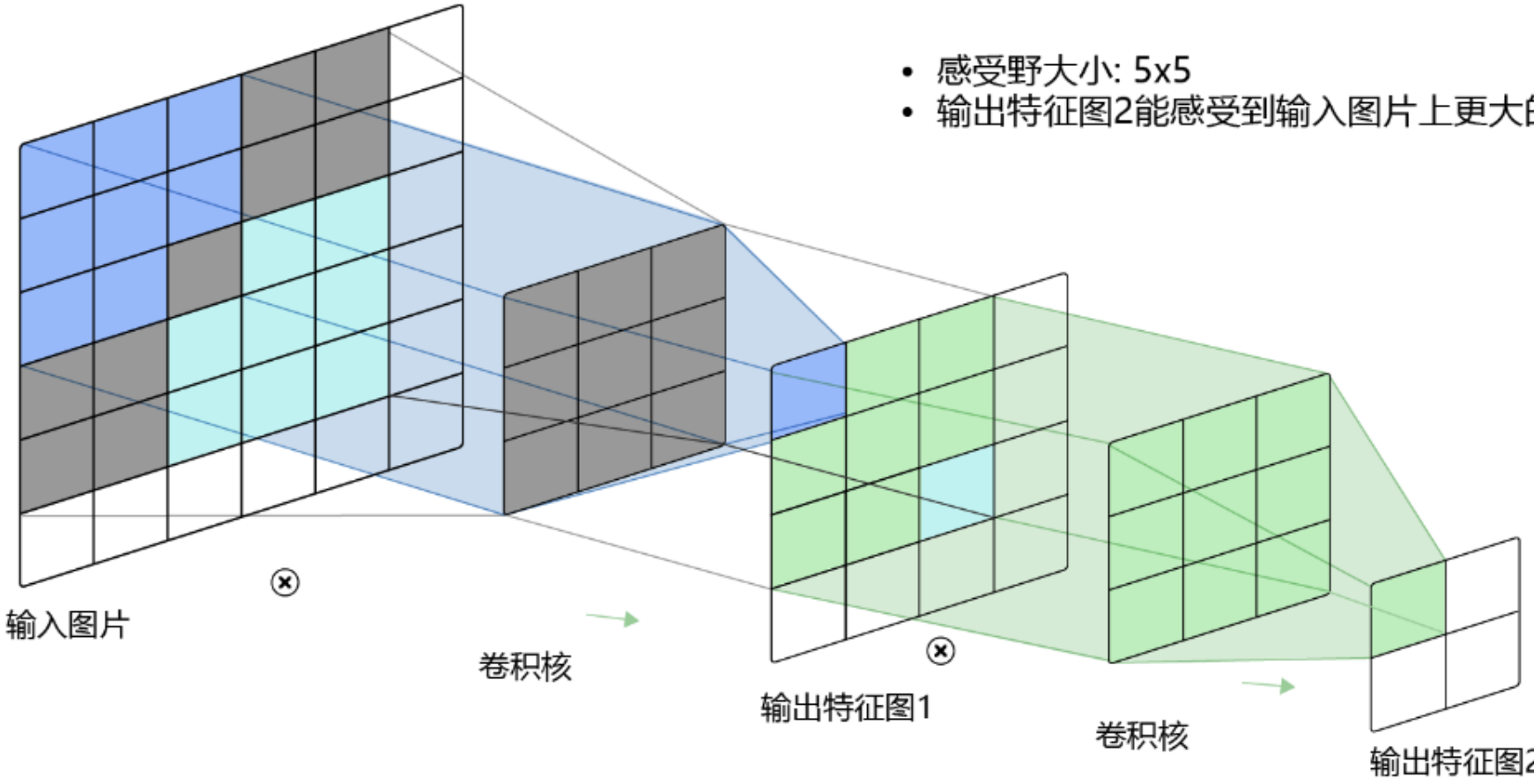
单层卷积



- 感受野大小: 3x3
- 输出特征图上的像素点所能感受到的输入数据的范围

卷积层的感受野

两层卷积



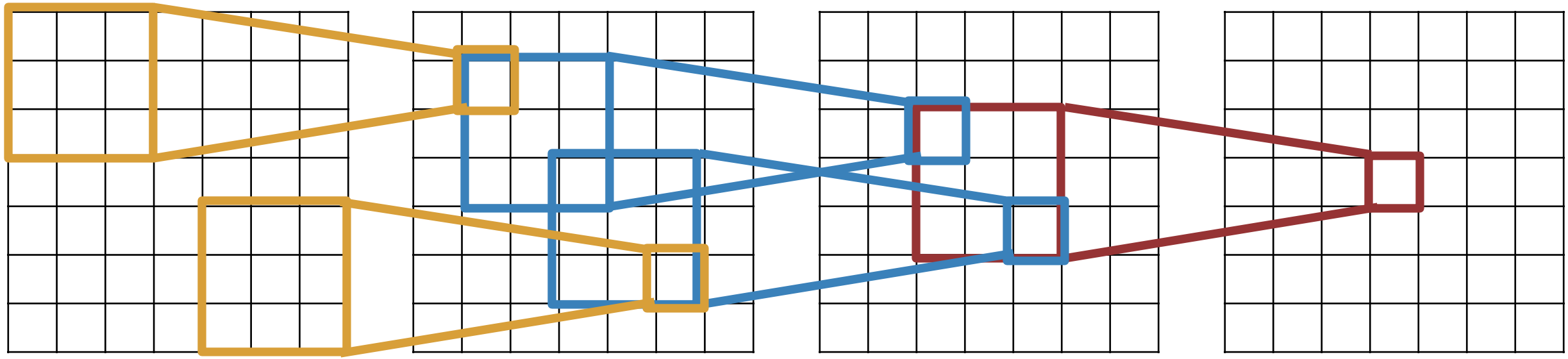
- 感受野大小: 5x5
- 输出特征图2能感受到输入图片上更大的范围

卷积神经网络

- 卷积层的感受野
- 每次连续的卷积都会使感受野大小增加K-1。

  对于L层，感受野大小为1+L*（K-1）

注意区分：
在前一层上的感受野
在输入图像上的感受野



卷积神经网络

- 卷积层的感受野
- 问题：对于大图像，我们需要使用许多卷积层，才能“看到”整个图像
- 解决方案：下采样
  - 大步长
  - 池化层

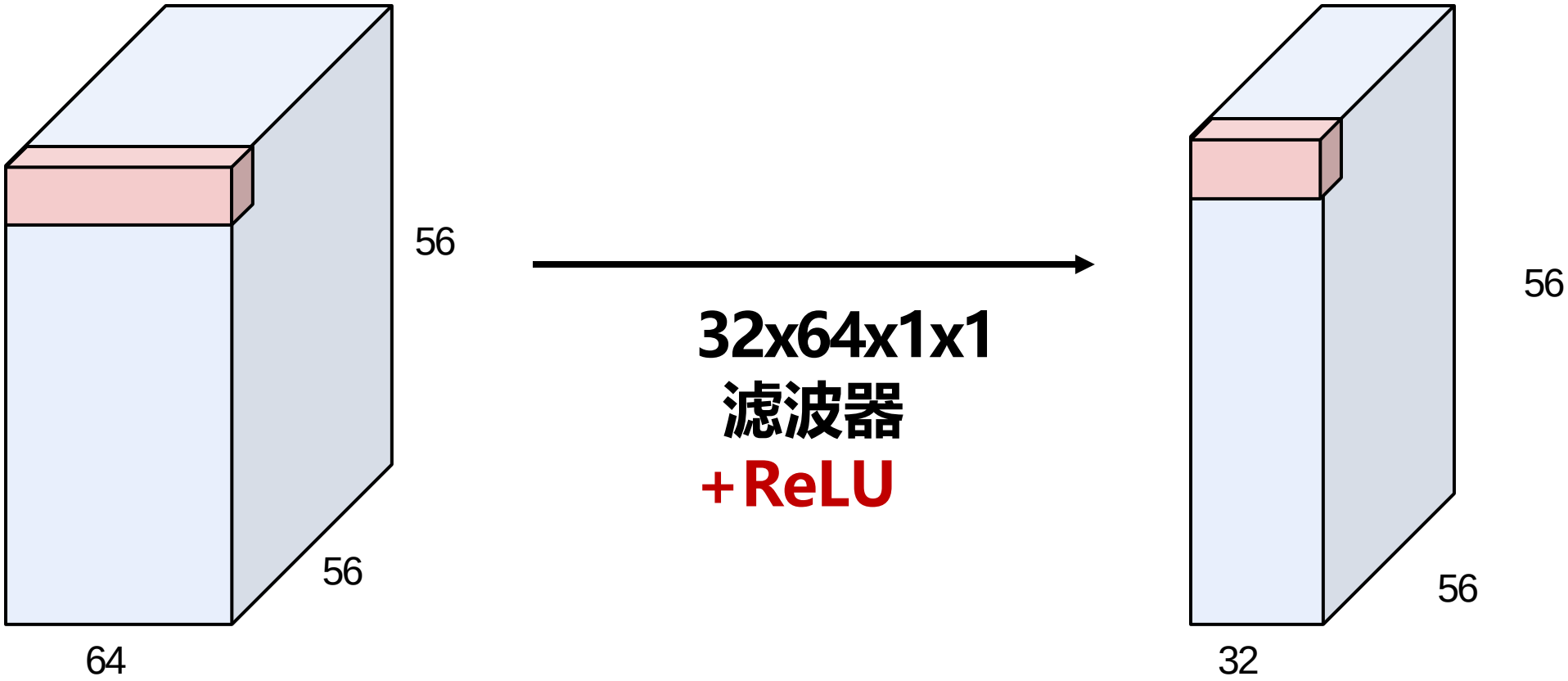
卷积神经网络

- 卷积层总结
- 输入： $W_1 \times H_1 \times C$
- 卷积超参：
  - 滤波器数量 K
  - 滤波器大小 F
  - 步长 S
  - Padding P
- 输出：$W_2 \times H_2 \times C$
  - $W_2 = (W_1 - F + 2P)/S + 1$
  - $H_2 = (H_1 - F + 2P)/S + 1$
- 参数量：
  - F^2CK+K，其中K是bias数量

卷积神经网络

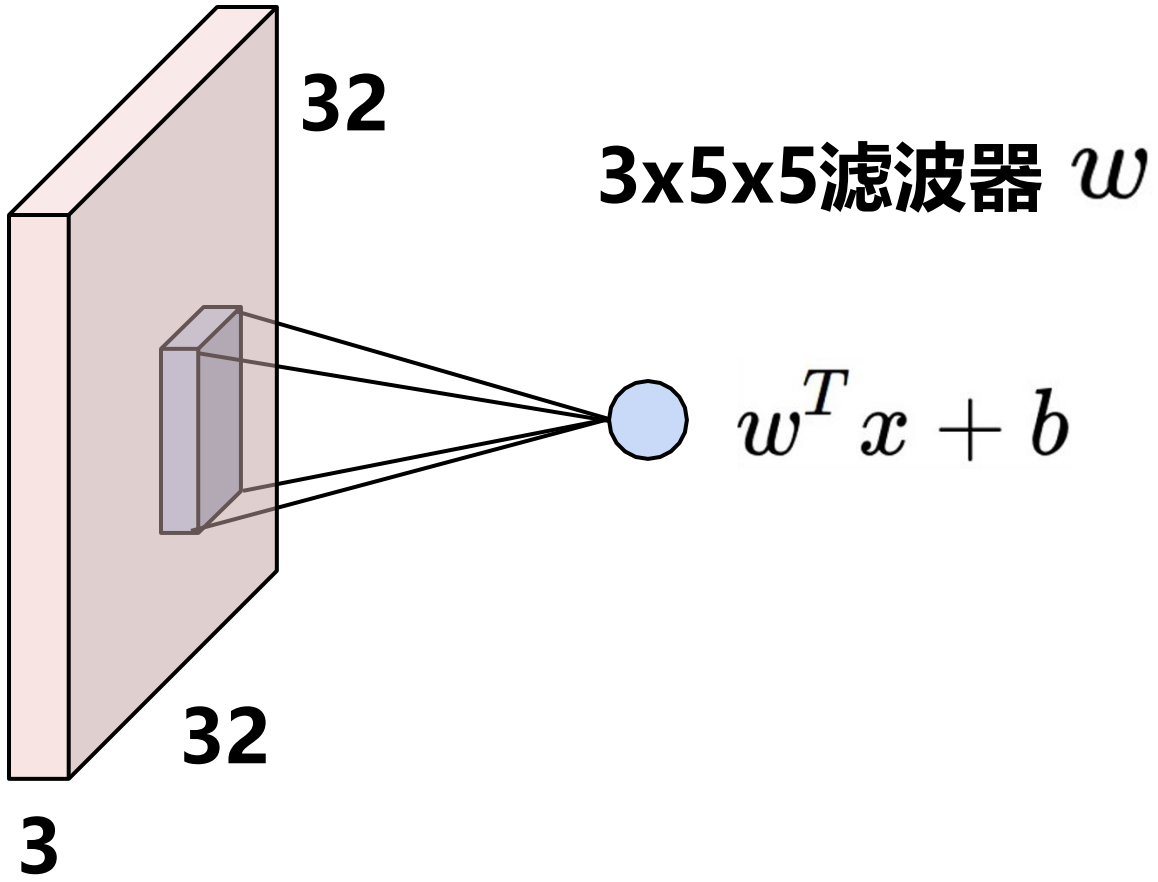
- 卷积层总结
- 输入： W_1 x H_1 x C
- 卷积超参：
  - 滤波器数量 K
  - 滤波器大小 F
  - 步长 S
  - Padding P
- K = (2的指数, e.g. 32, 64, 128, 512)
- F = 3, S = 1, P = 1
- F = 5, S = 1, P = 2
- F = 5, S = 2, P = ？ (匹配尺寸)
- F = 1, S = 1, P = 0

- 1x1卷积层



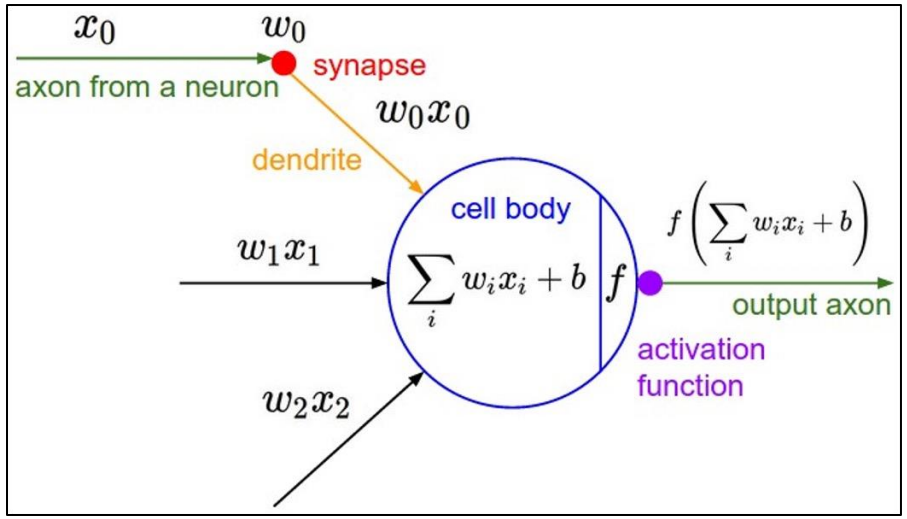
卷积神经网络

- 从大脑/神经元角度理解卷积层

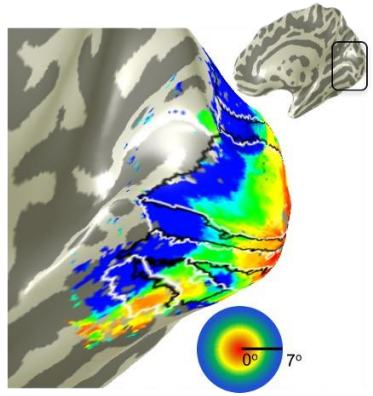


3x5x5滤波器 w

$w^T x + b$

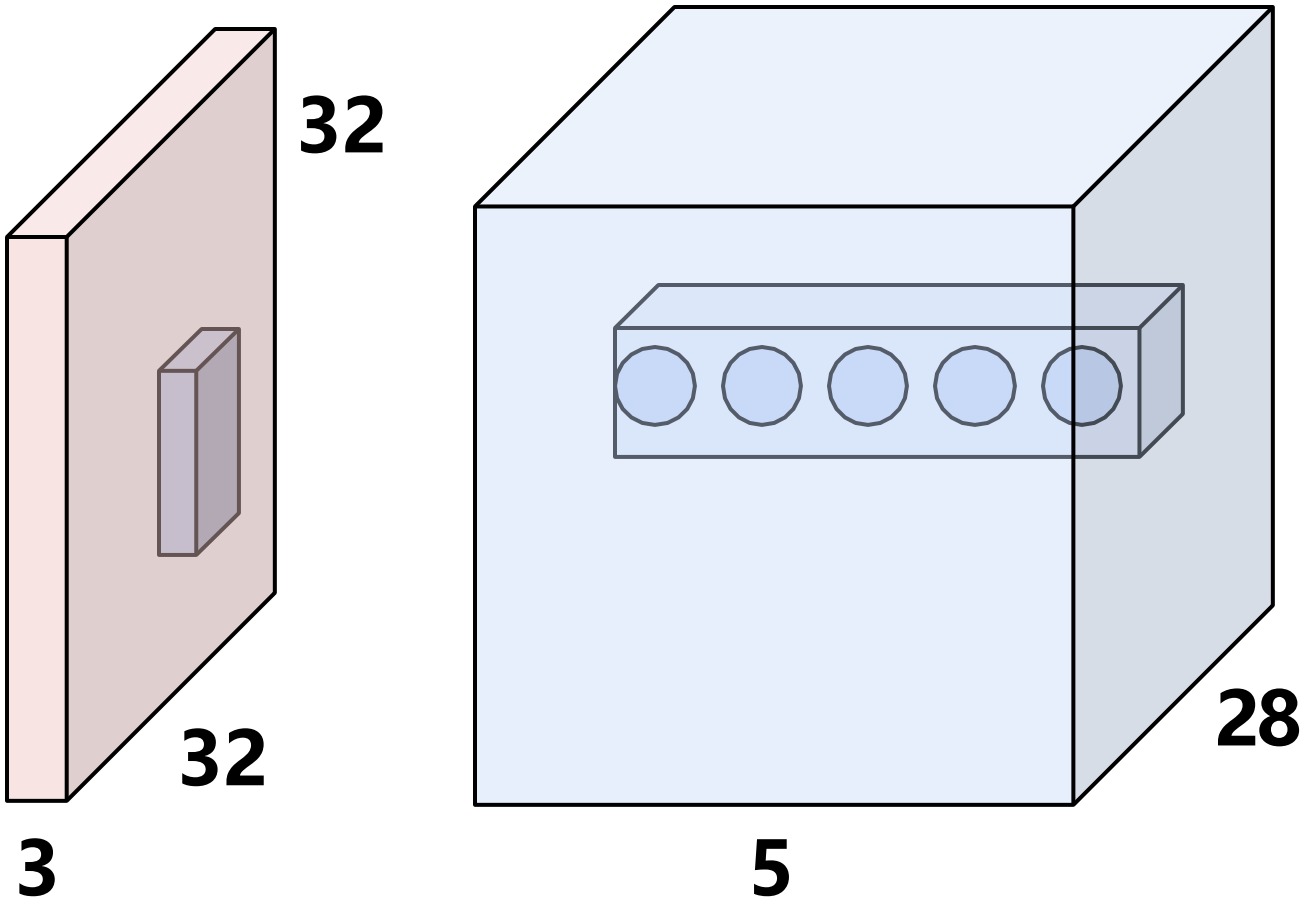


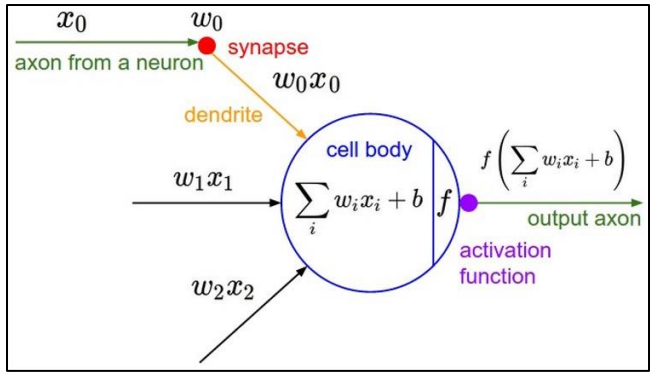
具有局部连接的神经元



卷积神经网络

- 从大脑/神经元角度理解卷积层

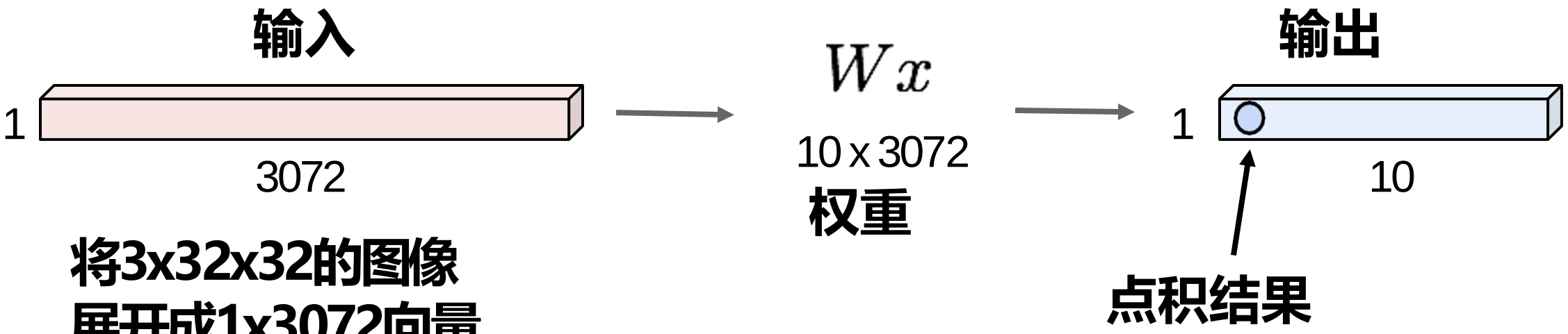




5个不同的神经元都在观察输入体积中的同一区域

卷积神经网络

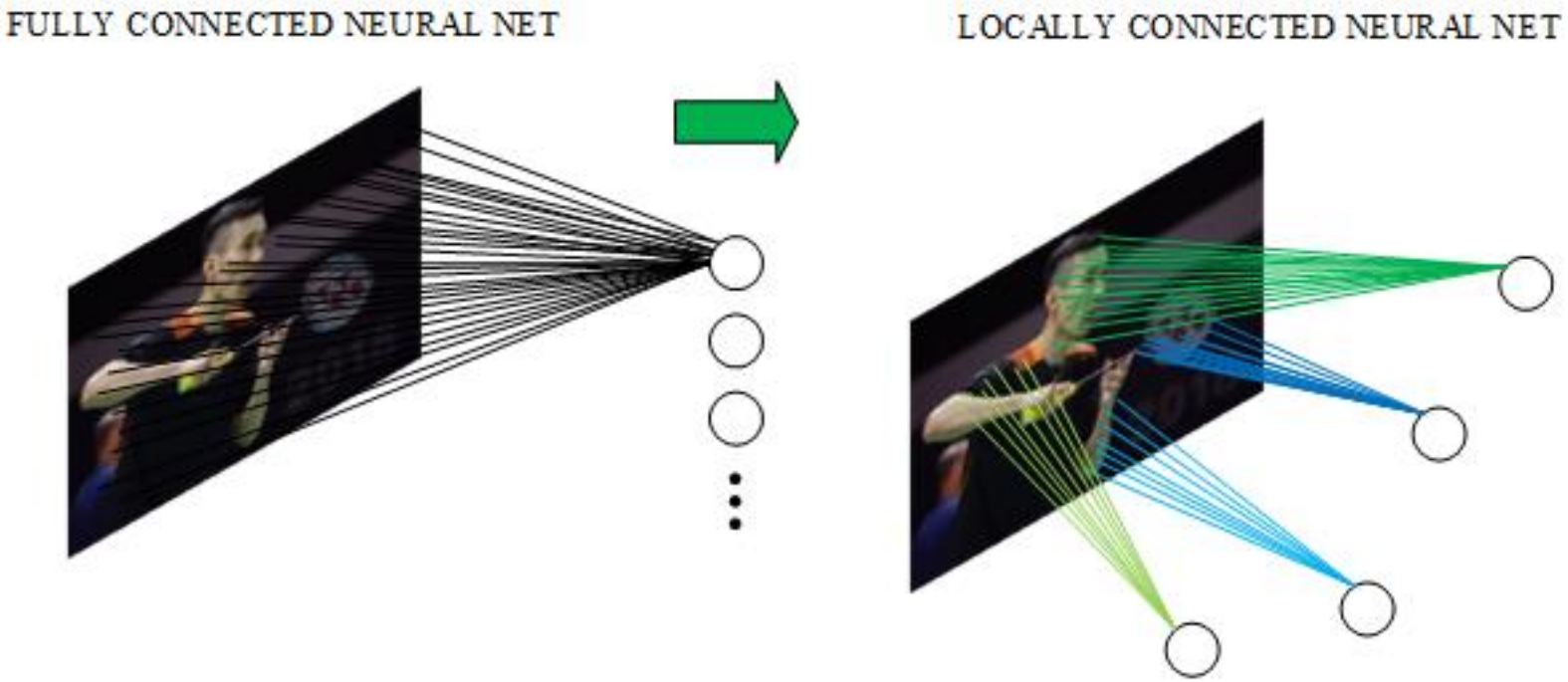
- 从大脑/神经元角度理解全连接层



卷积神经网络

南京大学 NANJING UNIVERSITY

卷积层 v.s. 全连接层

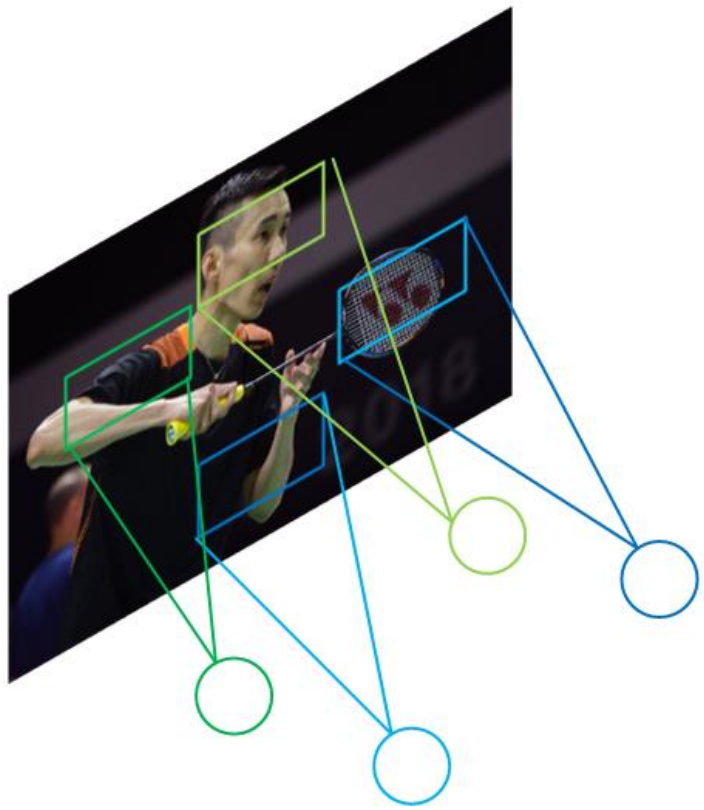


卷积神经网络

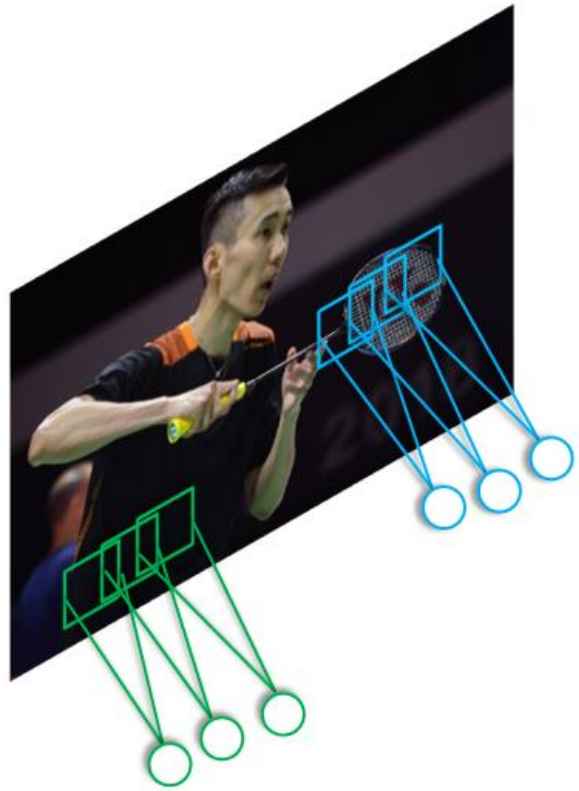
南京大学 NANJING UNIVERSITY

卷积层 v.s. 局部全连接层

LOCALLY CONNECTED NEURAL NET



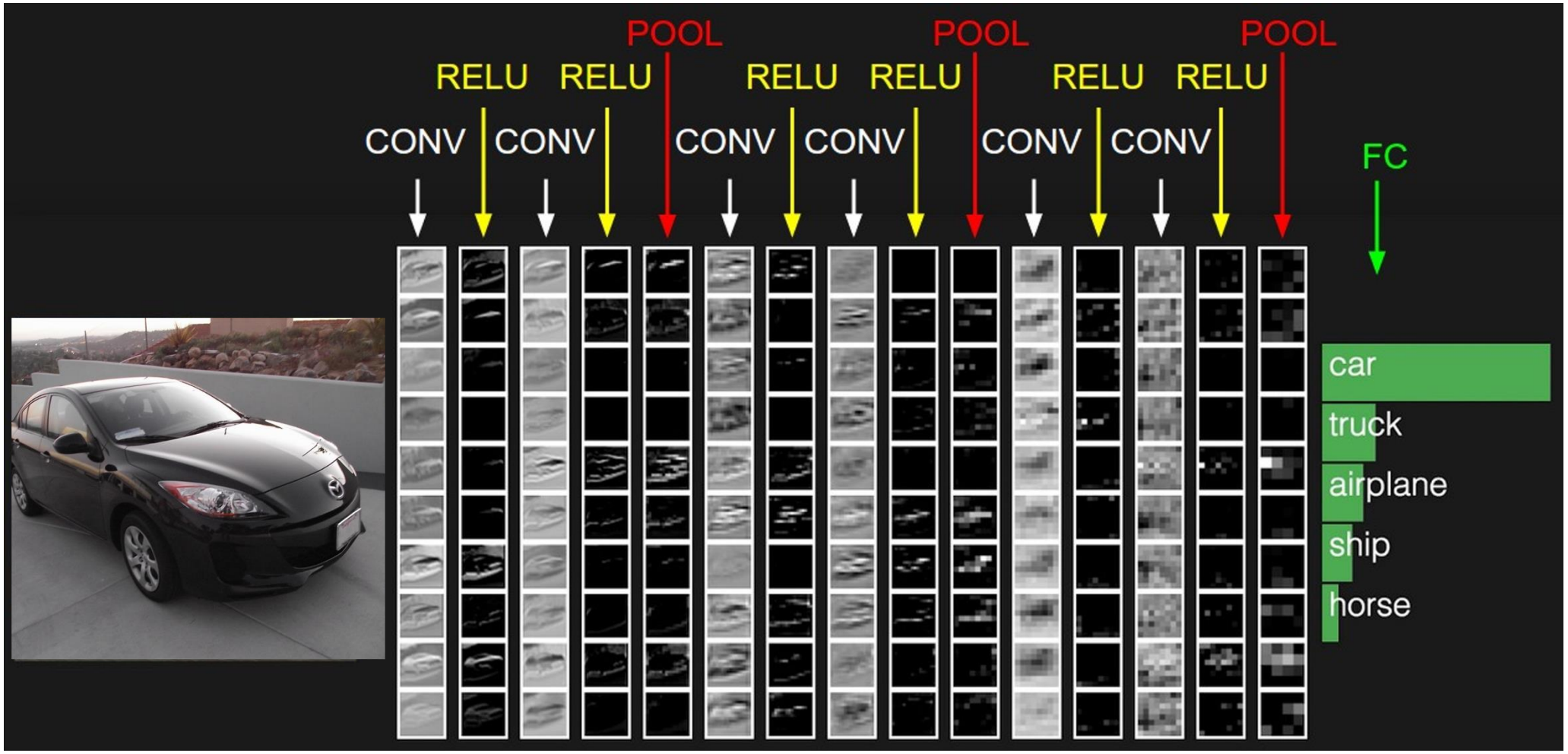
CONVOLUTIONAL NET



卷积神经网络

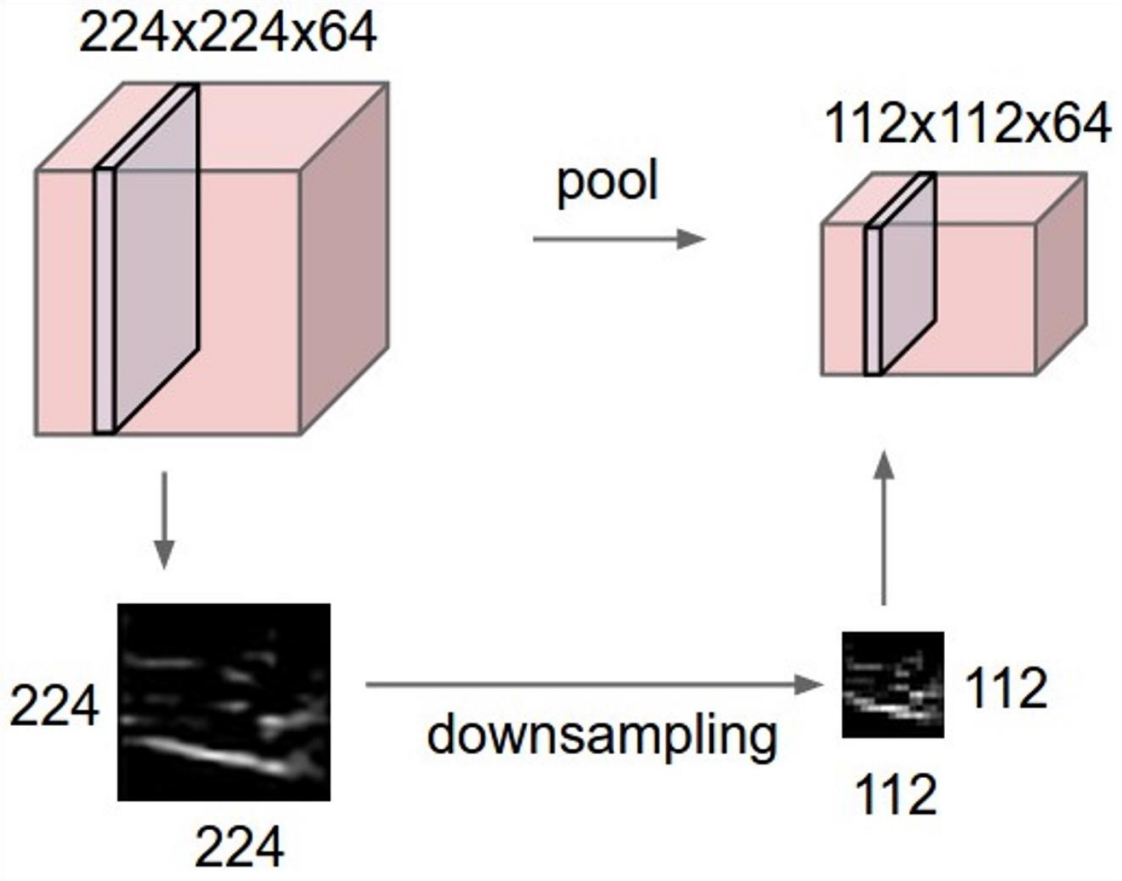
南京大学 NANJING UNIVERSITY

- 池化层

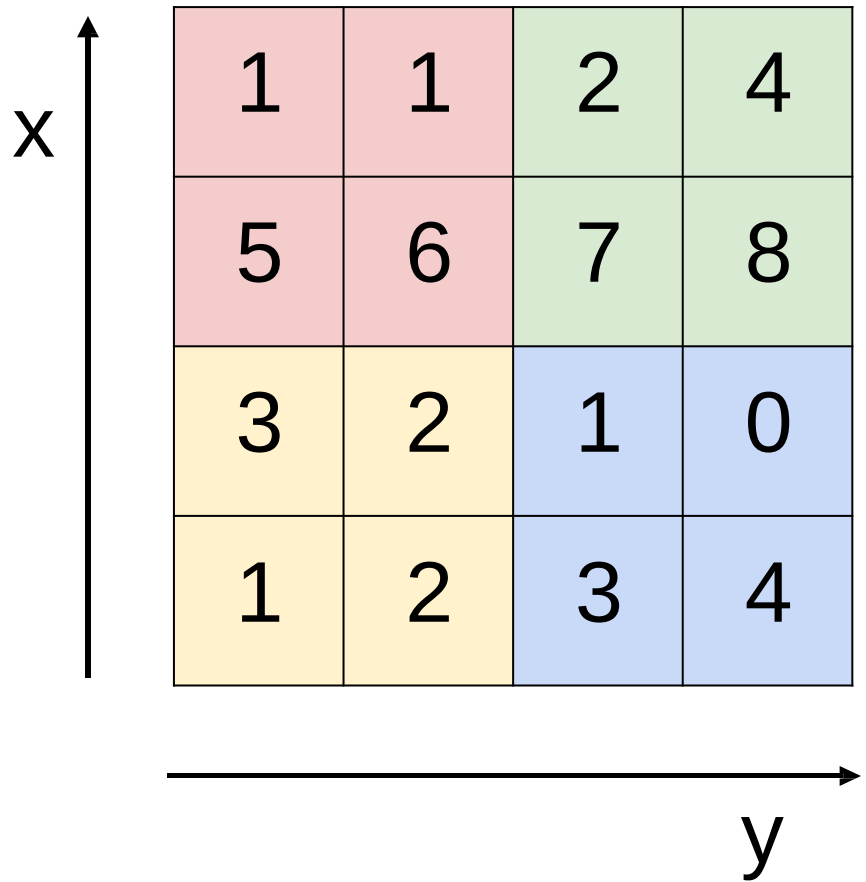


卷积神经网络

- 池化层
- 使特征更小，更易于处理
- 独立操作每个激活图



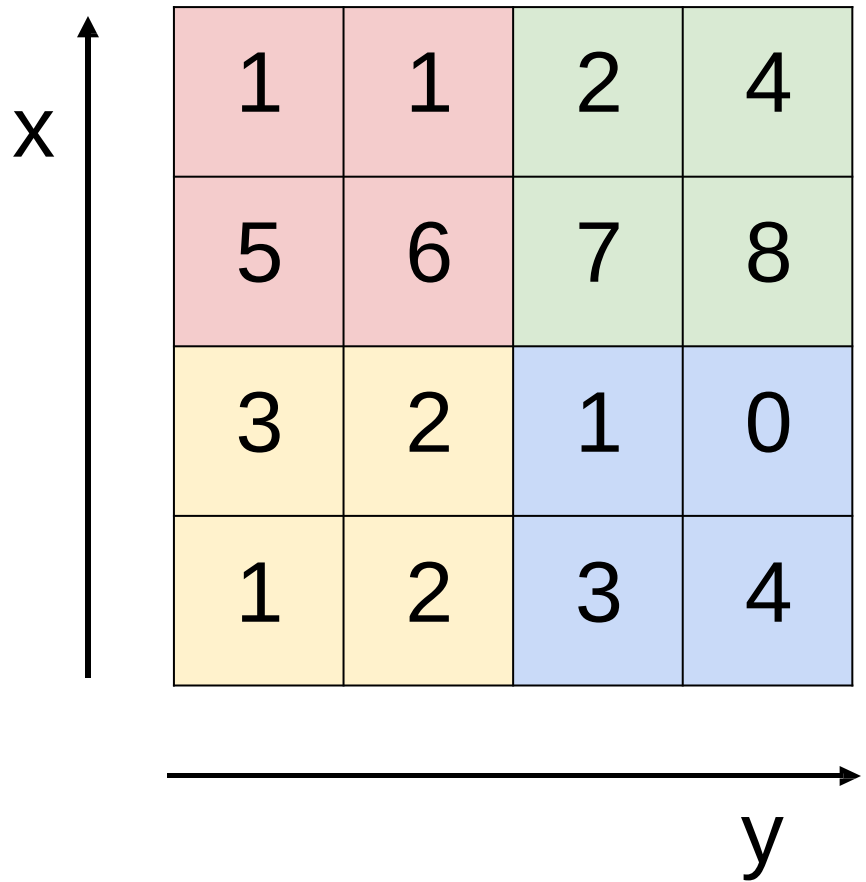
- 最大池化（Max Pooling）



Max Pooling
2x2 大小和 2 步长

6	8
3	4

- 最大池化（Max Pooling）



Max Pooling
2x2 大小和 2 步长

6	8
3	4

没有可学习参数
引入空间不变性

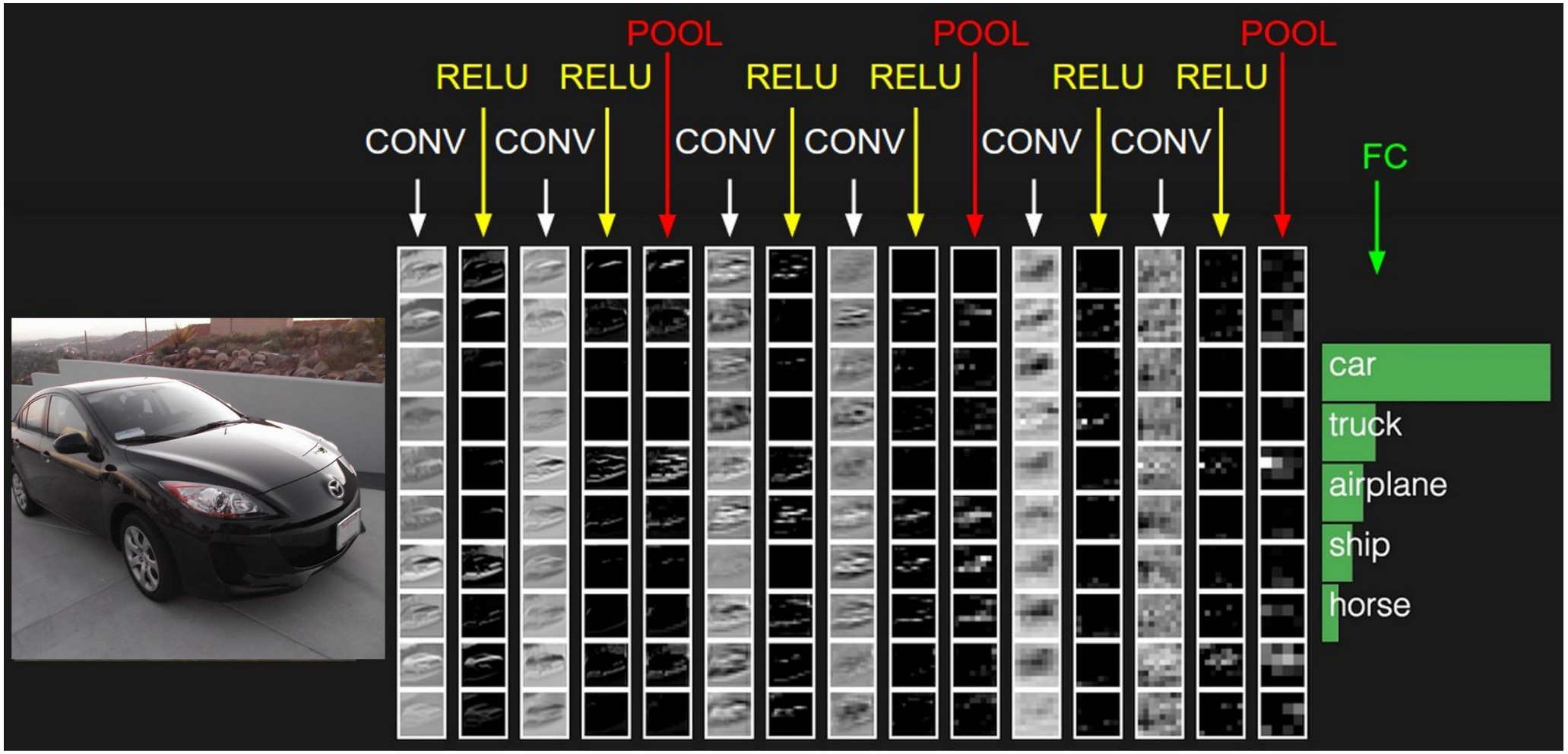
卷积神经网络

- 池化层总结
- 输入： $W_1 \times H_1 \times C$
- 卷积超参：
  - 滤波器大小 F
  - 步长 S
- 输出： $W_2 \times H_2 \times C$
  - $W_2 = (W_1 - F)/S + 1$
  - $H_2 = (H_1 - F)/S + 1$
- 参数量：
  - 0

卷积神经网络

南京大学 NANJING UNIVERSITY

全连接层



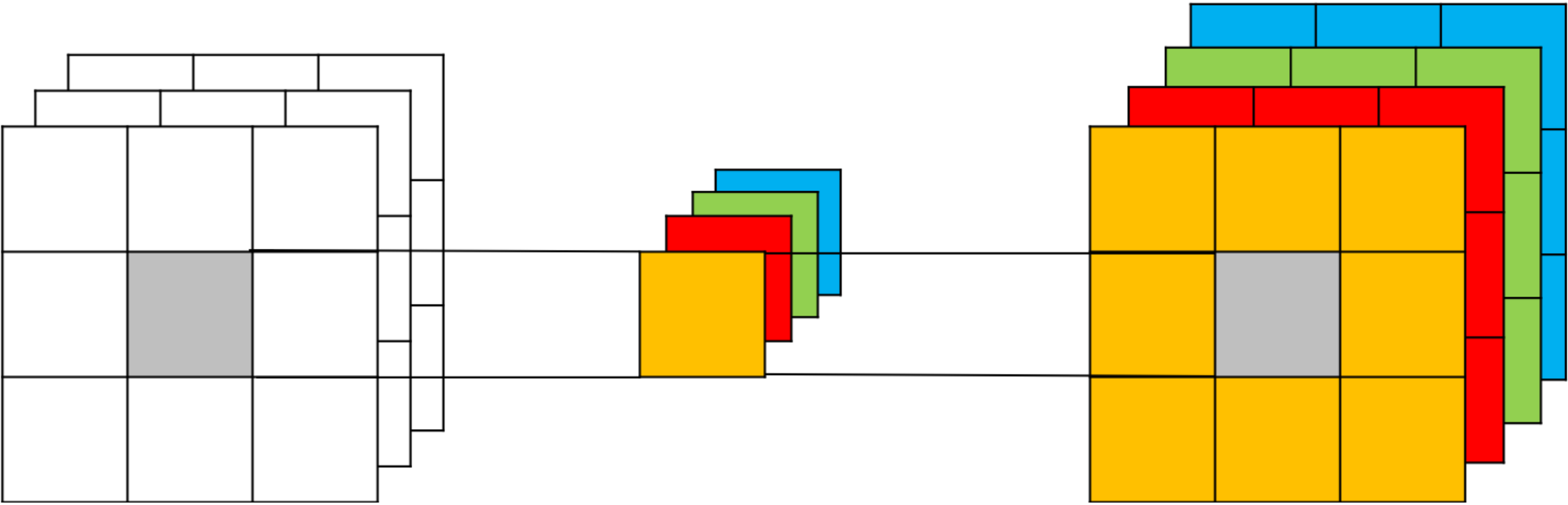
卷积神经网络

- 卷积层+池化层+全连层+激活函数
- 趋向于更小的滤波器和更深的层
- 趋向于舍弃池化层和全连接层
- 以往常见的卷积网络架构：
  - [(Conv-ReLU)*N-Pooling]*M-(FC-ReLU)*K-Softmax
  - N一般为~5，M较大，K一般在0~2
- 新型卷积网络（ResNet，GoogLeNet）提出了新的架构

卷积神经网络

南京大学 NANJING UNIVERSITY

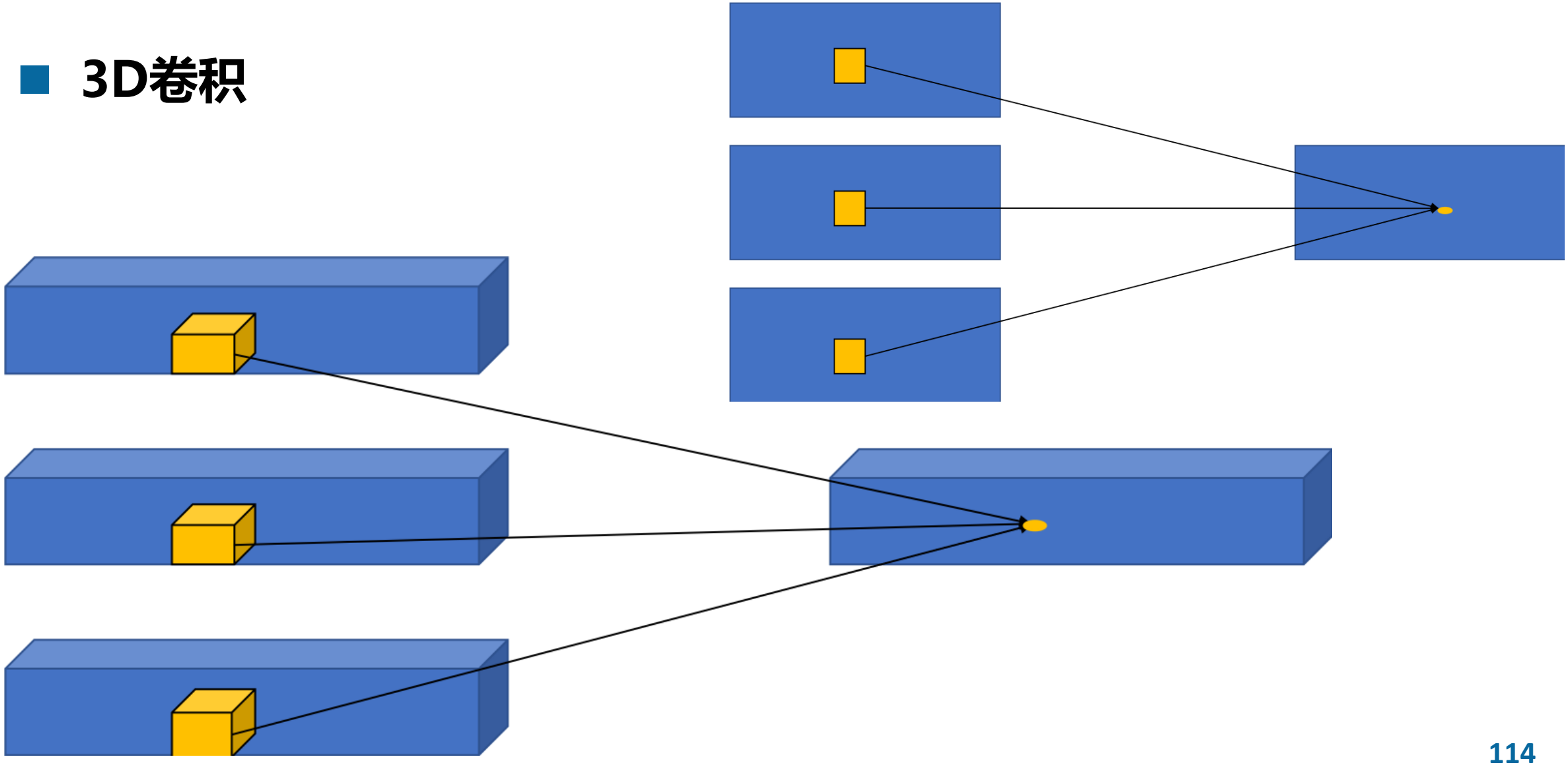
- 1x1卷积



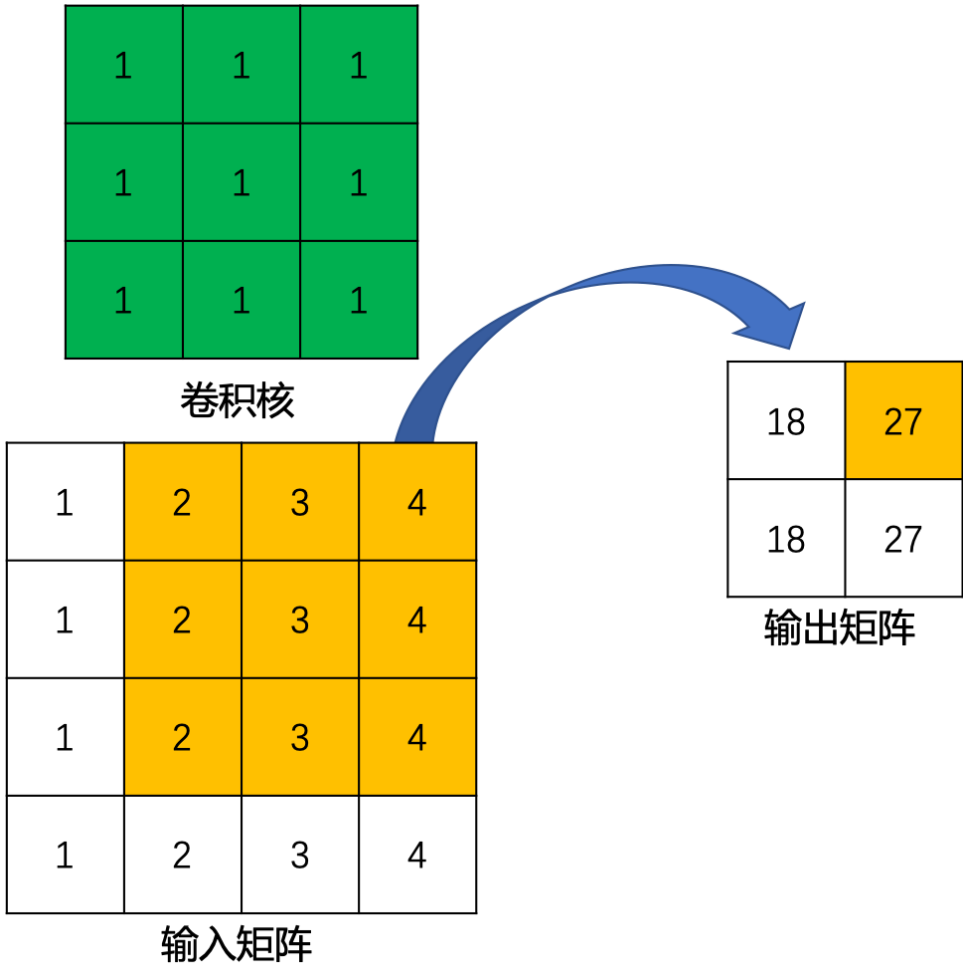
卷积神经网络

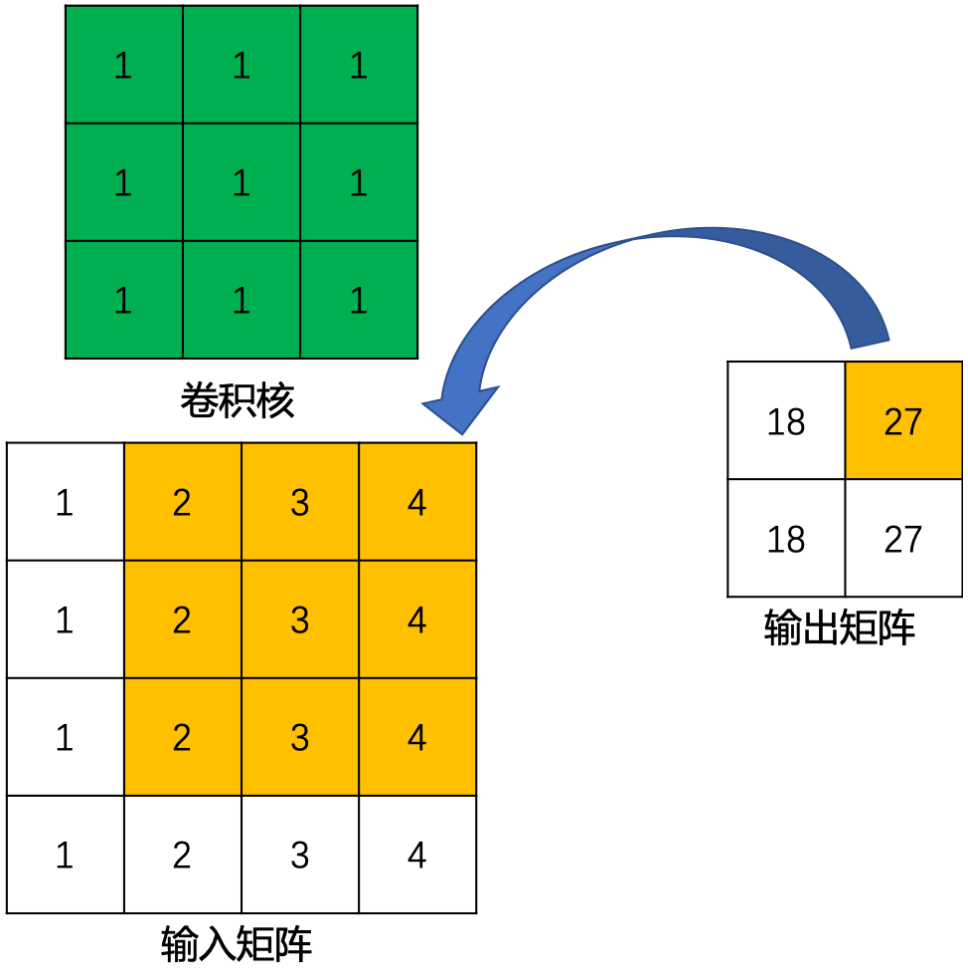
南京大学 NANJING UNIVERSITY

- 3D卷积



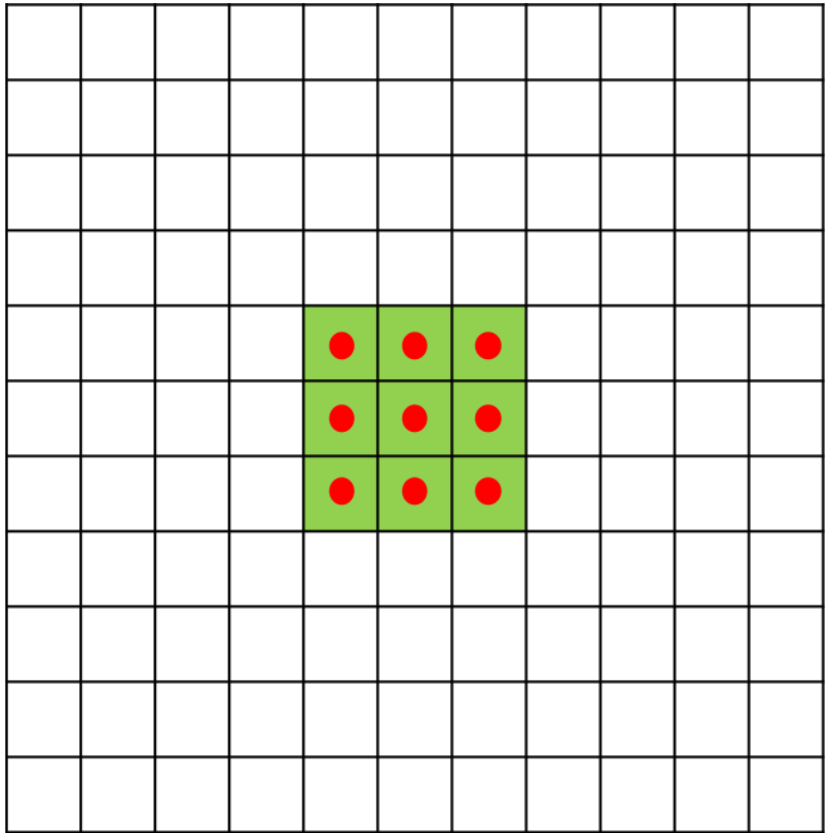
转置卷积

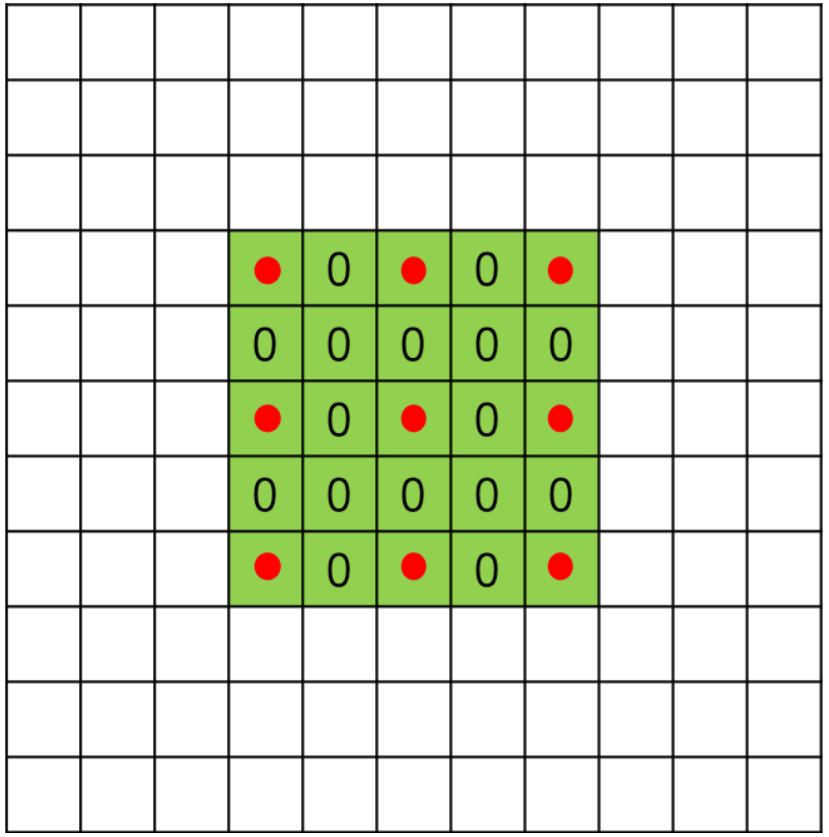

1 1 1
1 1 1
1 1 1
卷积核
18 27
18 27
输出矩阵
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4
输入矩阵

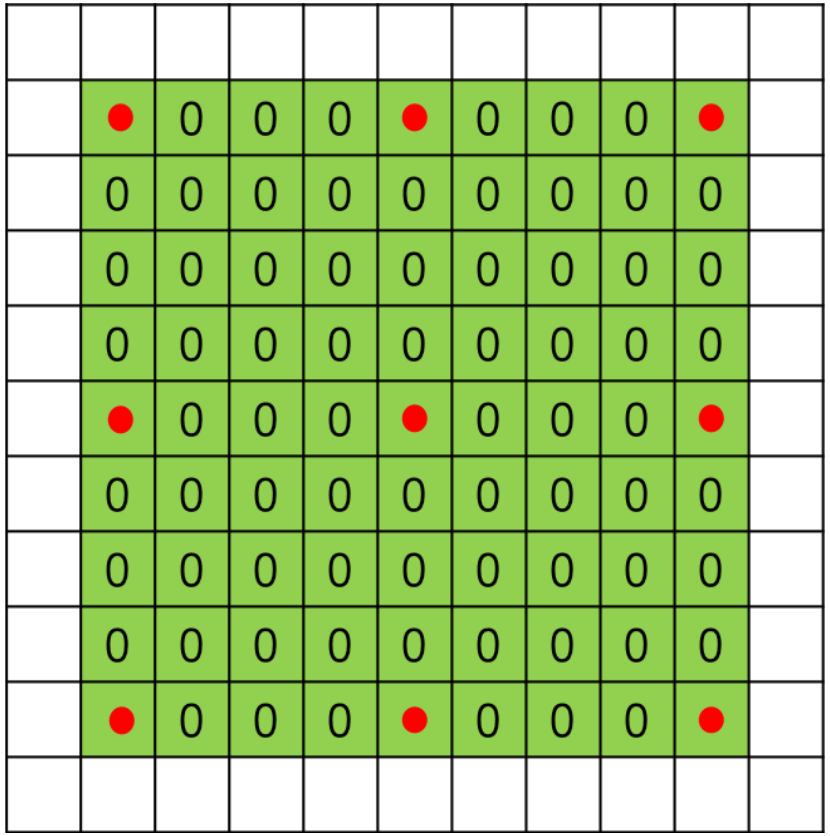

1 1 1
1 1 1
1 1 1
卷积核
18 27
18 27
输出矩阵
1 2 3 4
1 2 3 4
1 2 3 4
1 2 3 4
输入矩阵

南京大学 NANJING UNIVERSITY

空洞卷积

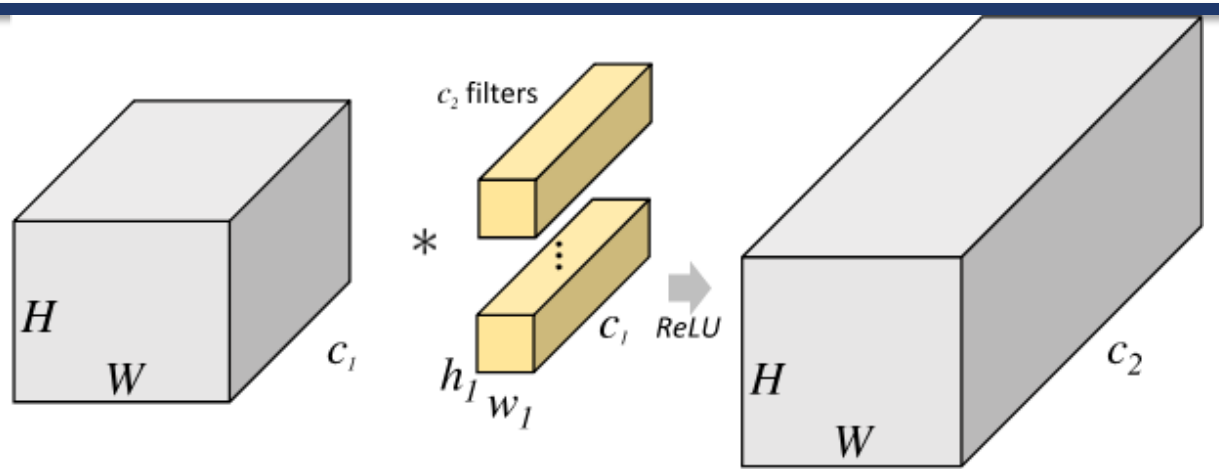


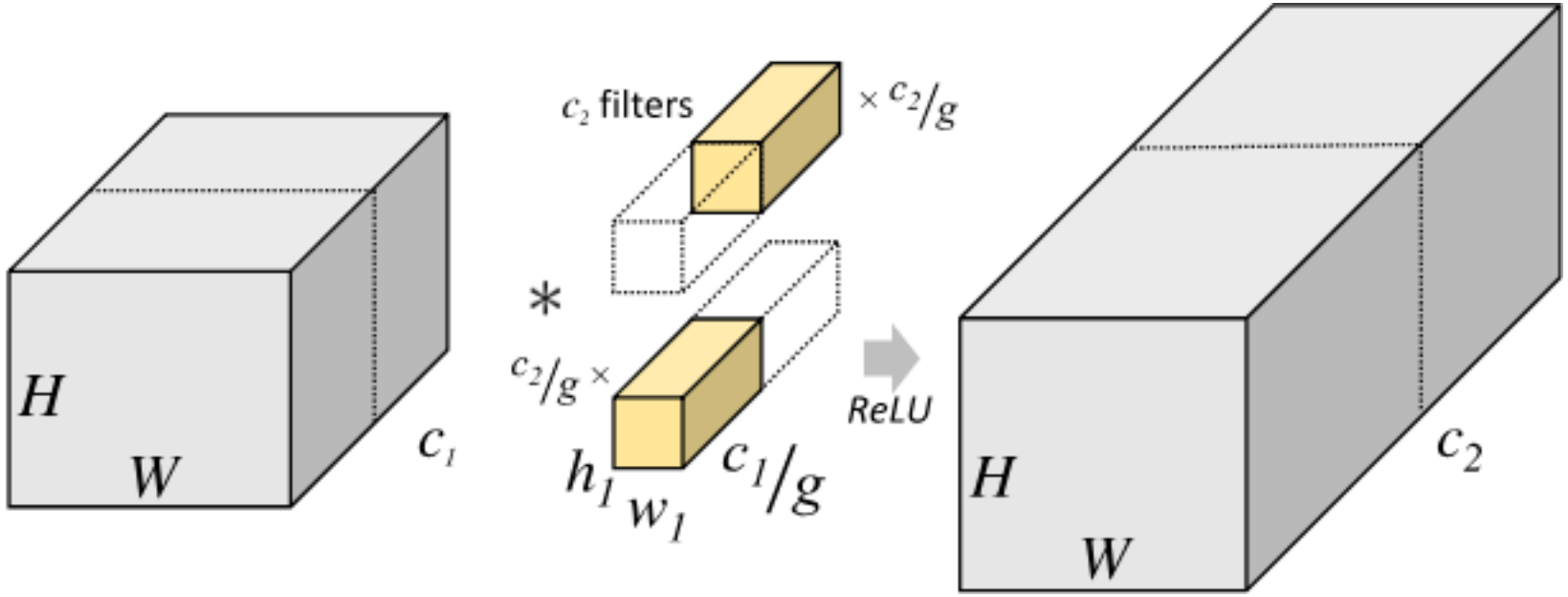




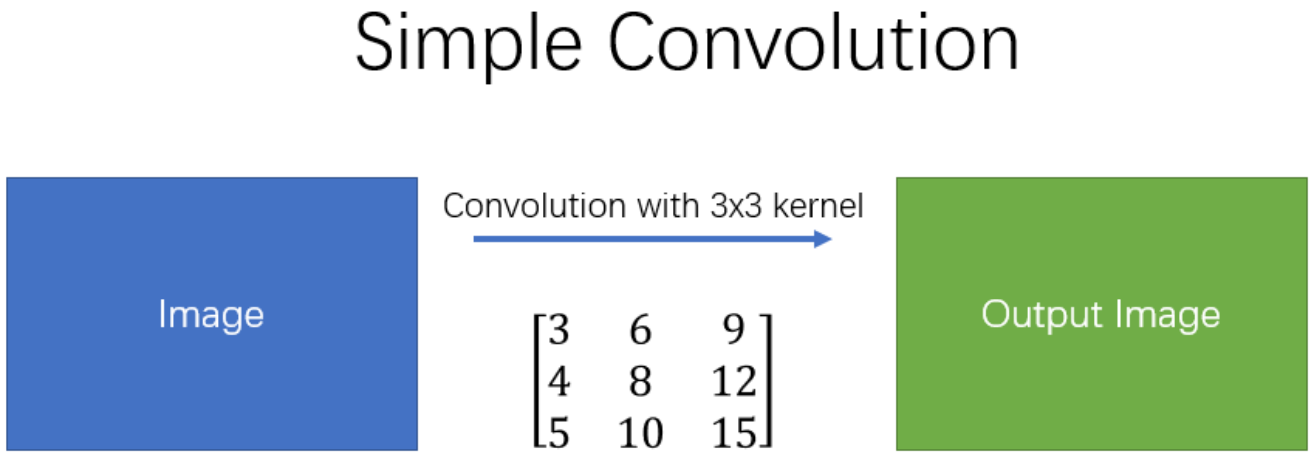
卷积神经网络

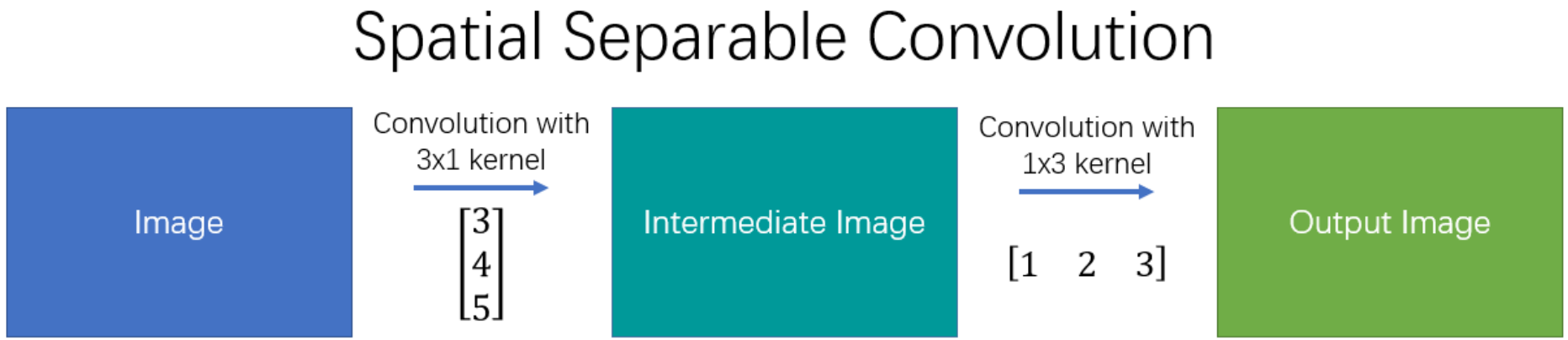
- 分组卷积


H
W
c_1
c_2 filters
h_1
w_1
c_1
ReLU
H
W
c_2


H
W
c_1
c_2 filters
$\times c_2/g$
$c_2/g \times$
h_1
w_1
c_1/g
ReLU
H
W
c_2

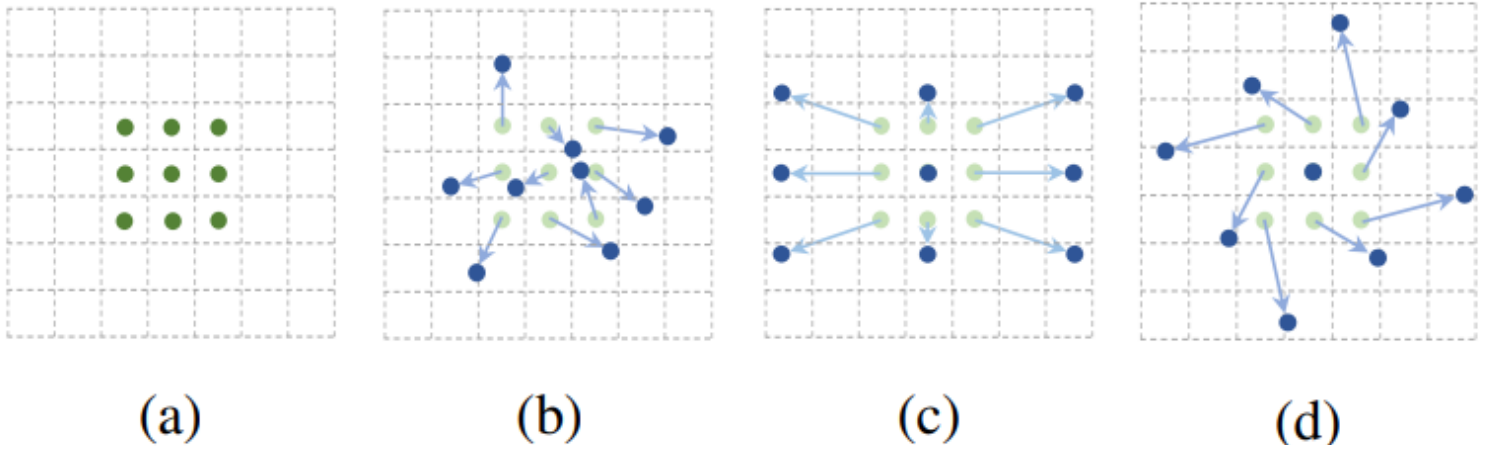
可分离卷积

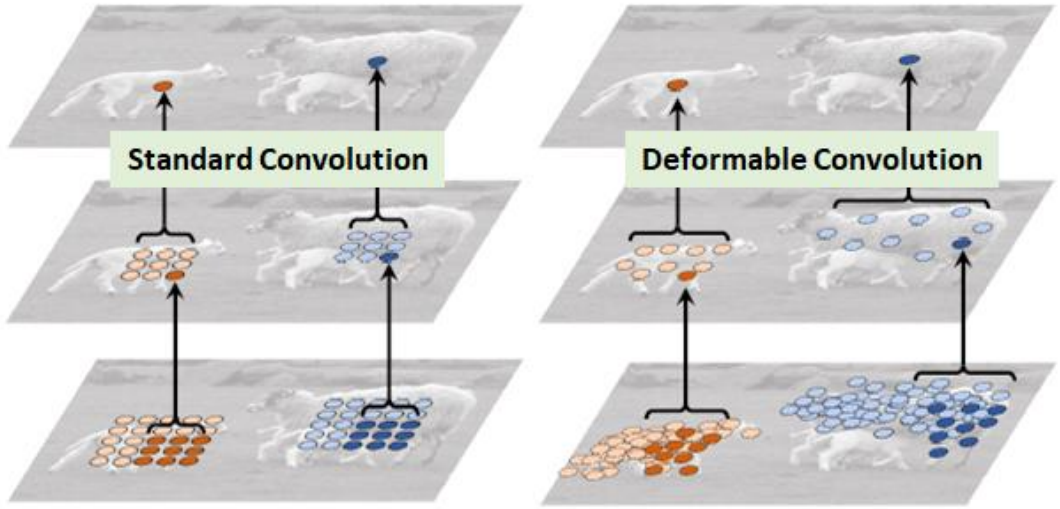

Simple Convolution
Image
Convolution with 3x3 kernel
3 6 9
4 8 12
5 10 15
Output Image

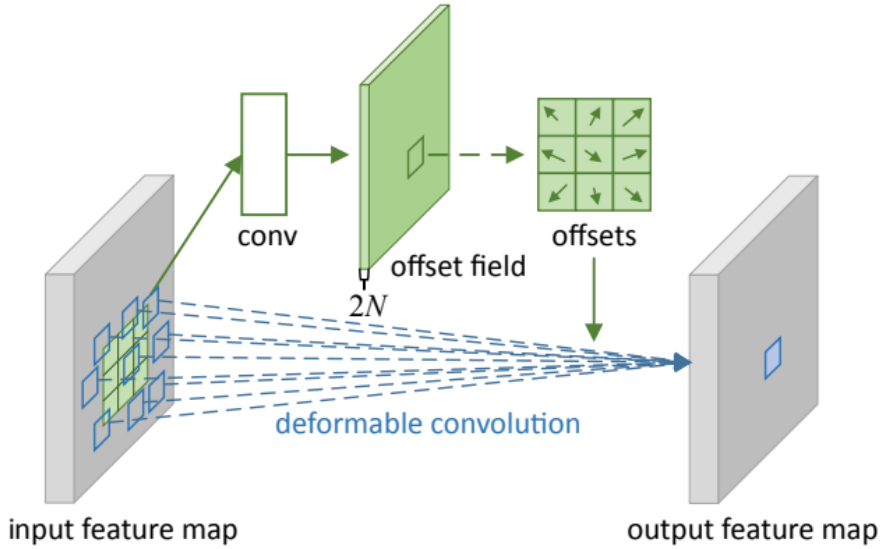

Spatial Separable Convolution
Image
Convolution with 3x1 kernel
3
4
5
Intermediate Image
Convolution with 1x3 kernel
1 2 3
Output Image

南京大学 NANJING UNIVERSITY

可变形卷积







卷积神经网络

南京大学 NANJING UNIVERSITY

- 下节课：卷积神经网络架构