厦门大学信息学院 本科选修课
2021-2022 第二学期

模 式 识 别
Pattern Recognition

主 讲：王 程

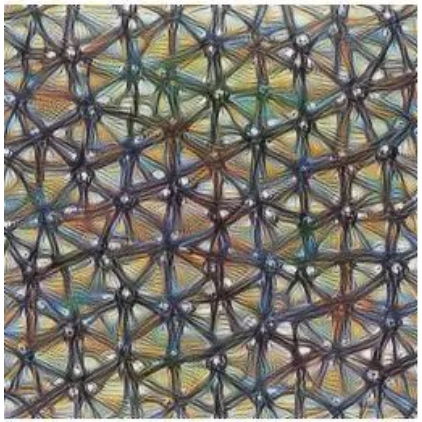
第八章

卷积神经网络（CNN）

陈龙彪
longbiaochen@xmu.edu.cn

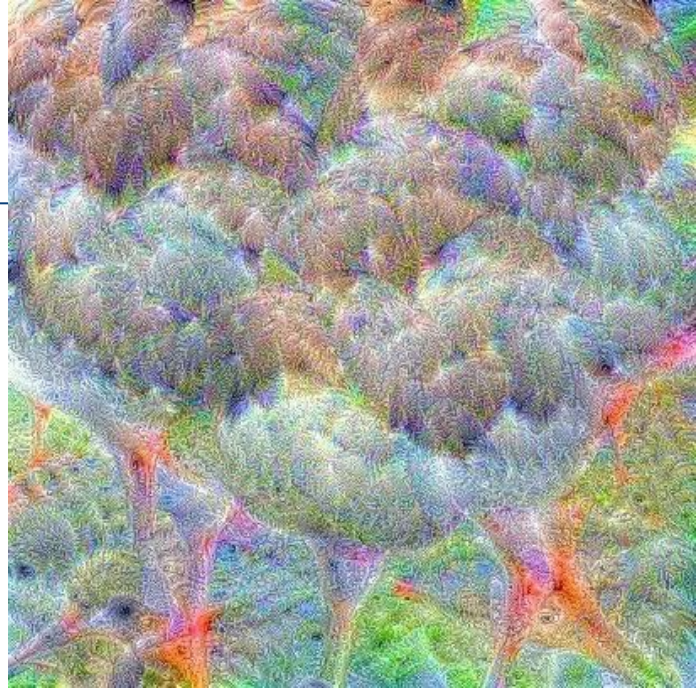
厦门大学信息学院

mixed3a, channel 31



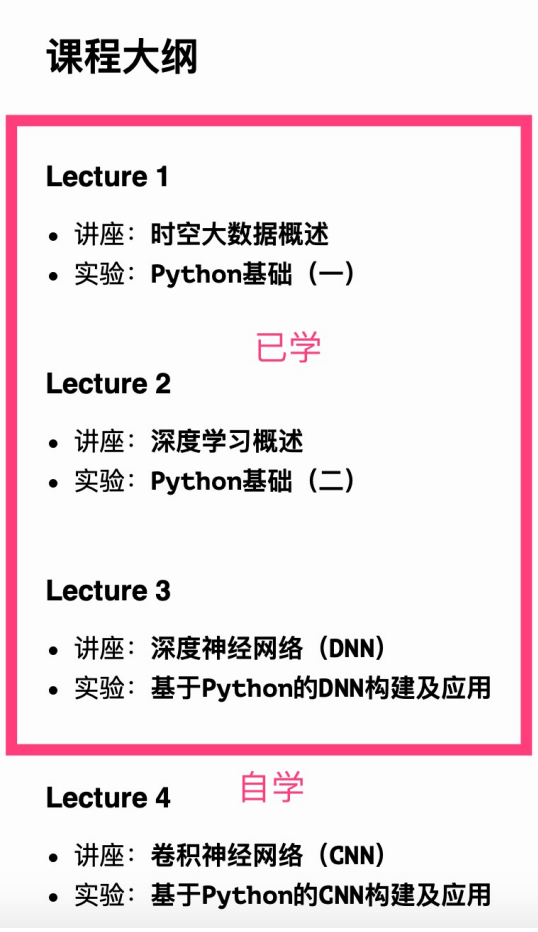
mixed4a, channel 11

mixed5a, channel 14



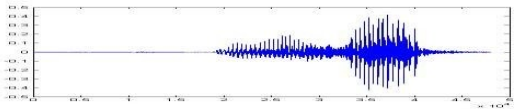
内容回顾

- **在线开放课程主页：**
  - https://longbiaochen.com/courses/steel/
- 翻转课堂学习模式
  - 课前：线上预习
  - 课堂：
    - 回顾总结
    - 知识点扩展
    - 应用展示：CNN小程序讨论
    - 启发与反思



深度学习≈寻找一个好函数

- 语音识别

$f($  $)$ = “How are you”

- 图片识别

$f($  $)$ = “Cat”

- 下围棋

$f($  $)$ = “5-5” (下一步棋的结果)

- 对话系统

f(“导航回家”) = “你说啥？”

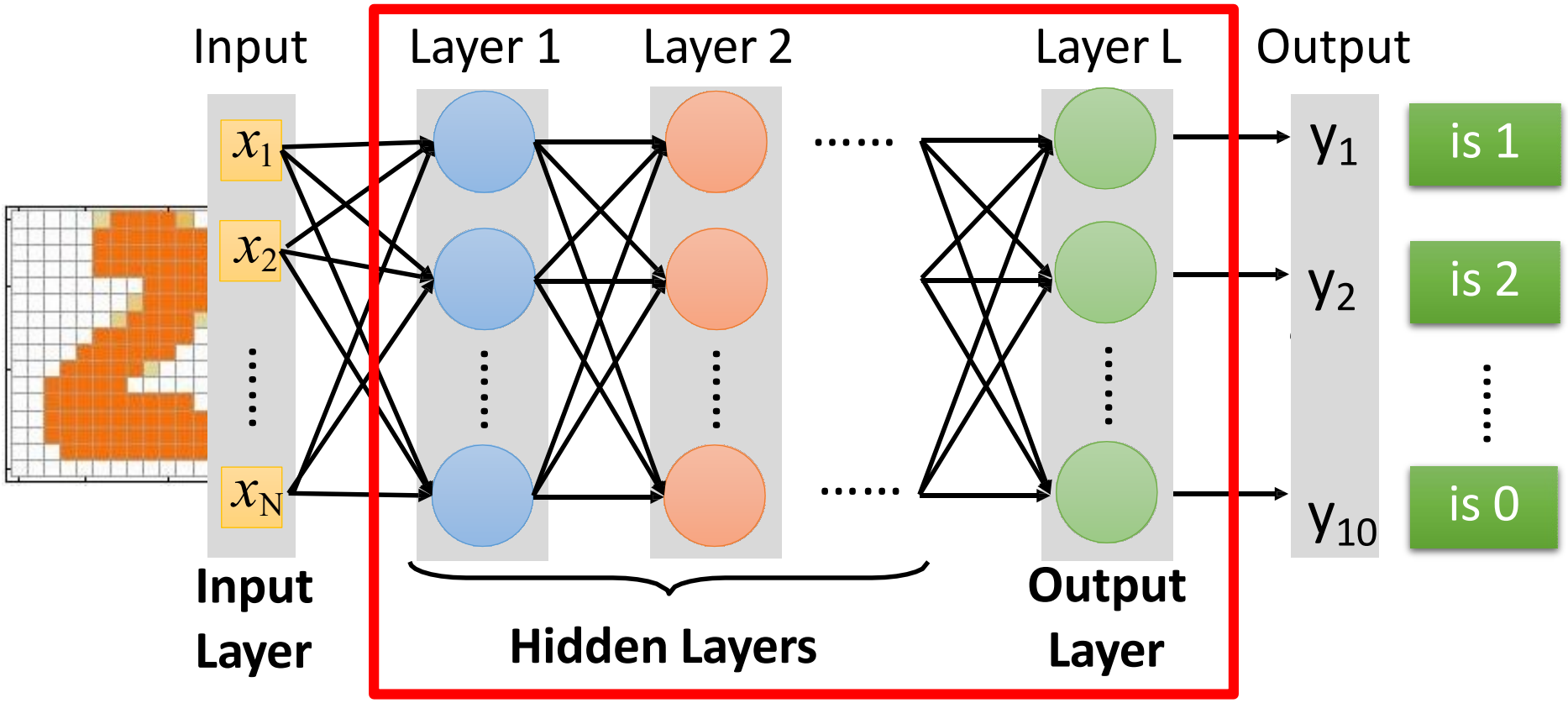
(你说的内容) (系统的回答)

深度学习框架

图像识别:

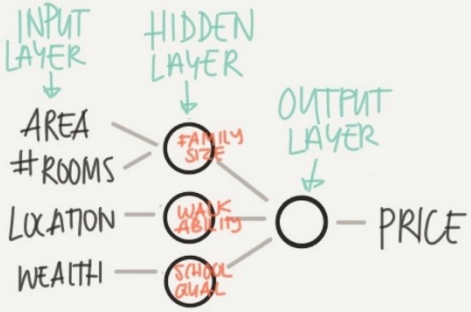
$f($  $)=$ “cat”

深度神经网络结构



你需要决定网络结构，让你的函数有好的性能

INTRO TO DEEP LEARNING



SUPERVISED LEARNING

INPUT: X	OUTPUT: y	NN TYPE
HOME FEATURES	PRICE	STANDARD NN
AD + USER INFO	WILL CLICK ON AD (0/1)	
IMAGE	OBJECT (1...1000)	CONV. NN (CNN)
AUDIO	TEXT TRANSCRIPT	RECURRENT NN (RNN)
ENGLISH	CHINESE	
IMAGE/RADAR	POS OF OTHER CARS	CUSTOM/HYBRID

NNs CAN DEAL WITH BOTH STRUCTURED & UNSTRUCTURED DATA

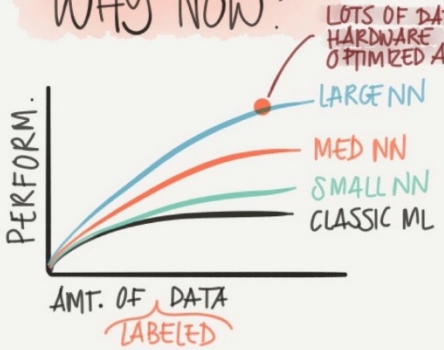
STRUCTURED

"THE QUICK BROWN FOX"

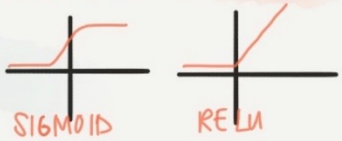
UNSTRUCTURED

HUMANS ARE GOOD AT THIS

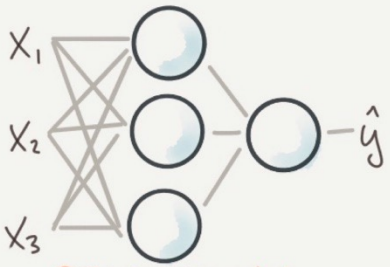
WHY NOW?



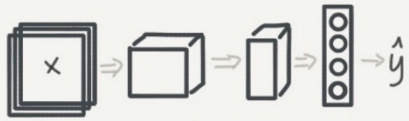
ONE OF THE BIG BREAKTHROUGHS HAS BEEN MOVING FROM SIGMOID TO RELU FOR FASTER GRADIENT DESCENT



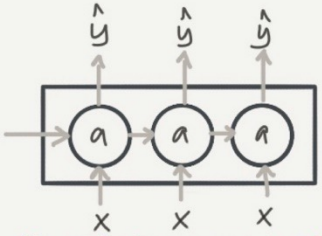
NETWORK ARCHITECTURES



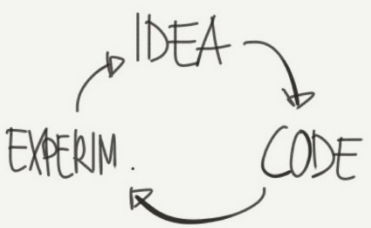
STANDARD NN



CONVOLUTIONAL NN



RECURRENT NN

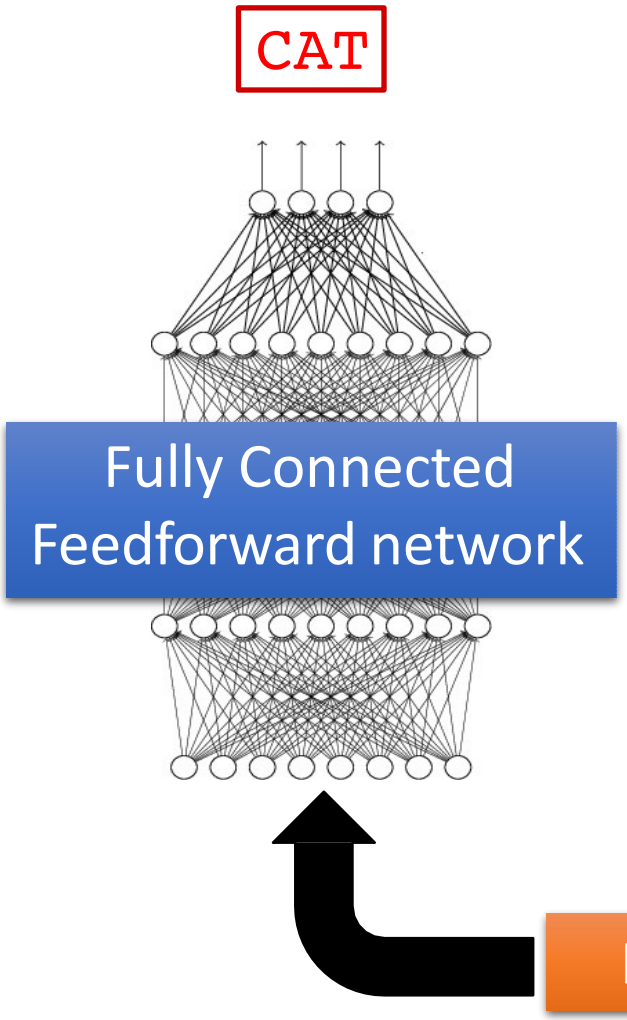


FASTER COMPUTATION IS IMPORTANT TO SPEED UP THE ITERATIVE PROCESS

你问我答

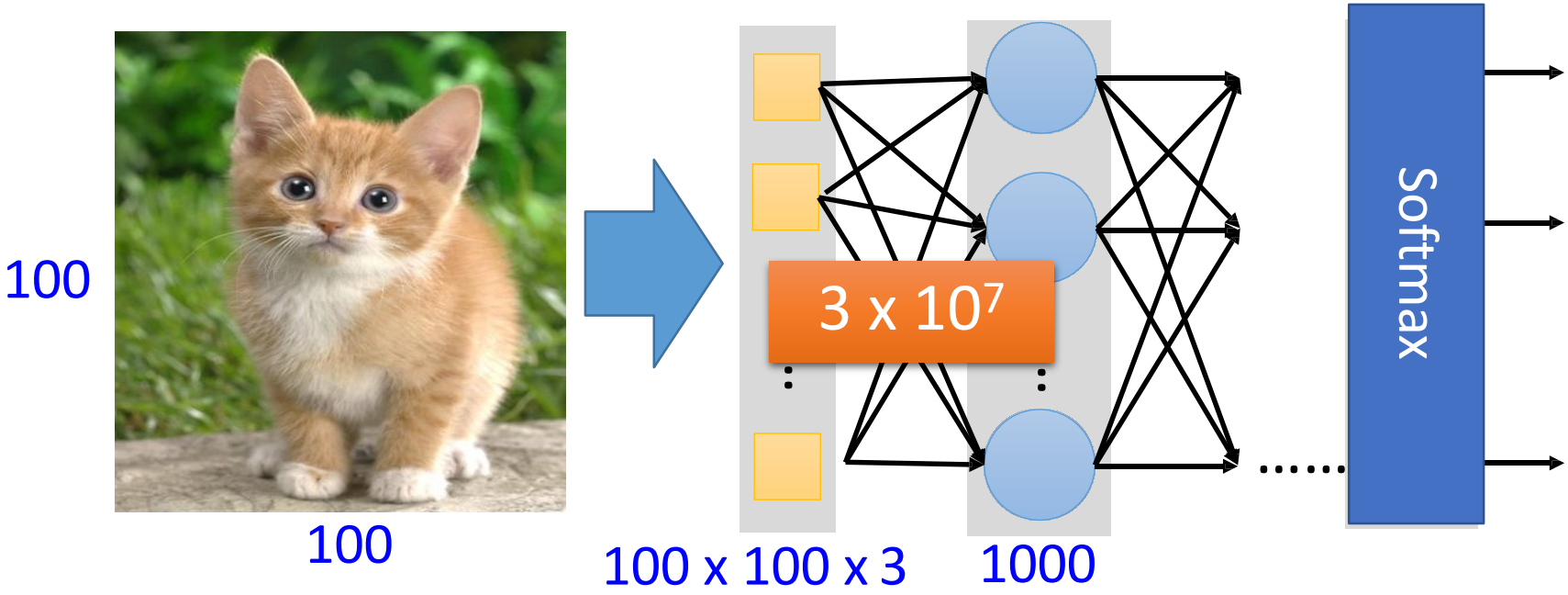
重点难点知识点扩展

CNN网络结构


CAT
Fully Connected
Feedforward network
Convolution
Max Pooling
Convolution
Max Pooling
Flatten
可以重复
多次

为什么CNN 可以用于图像识别?

- 处理图像时，全连接网络的输入层会非常大



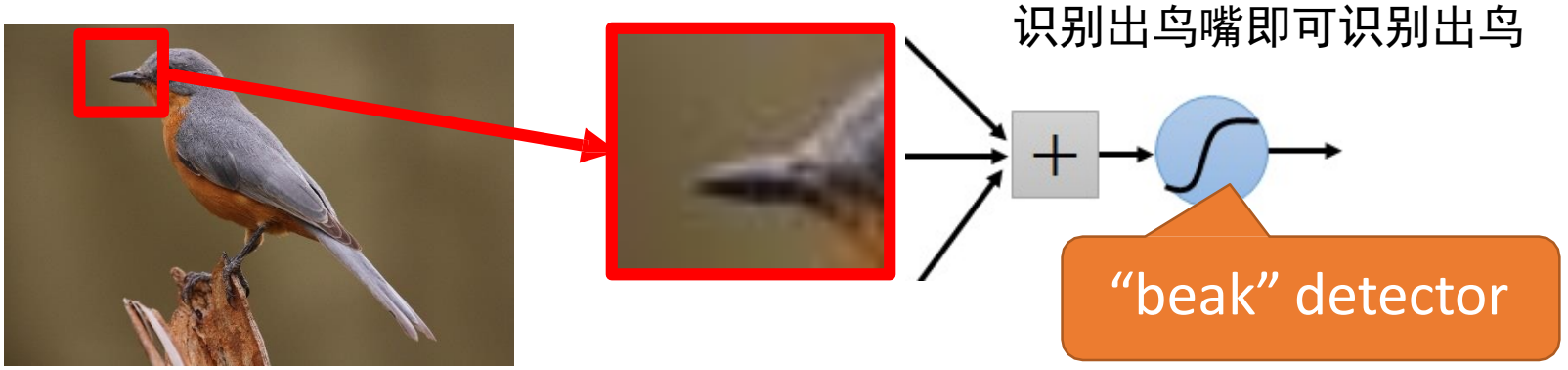
可以通过考虑图像识别的性质来简化全连接网络吗?

为什么CNN 用于图像识别?

1. 一些模式比整张图片小得多

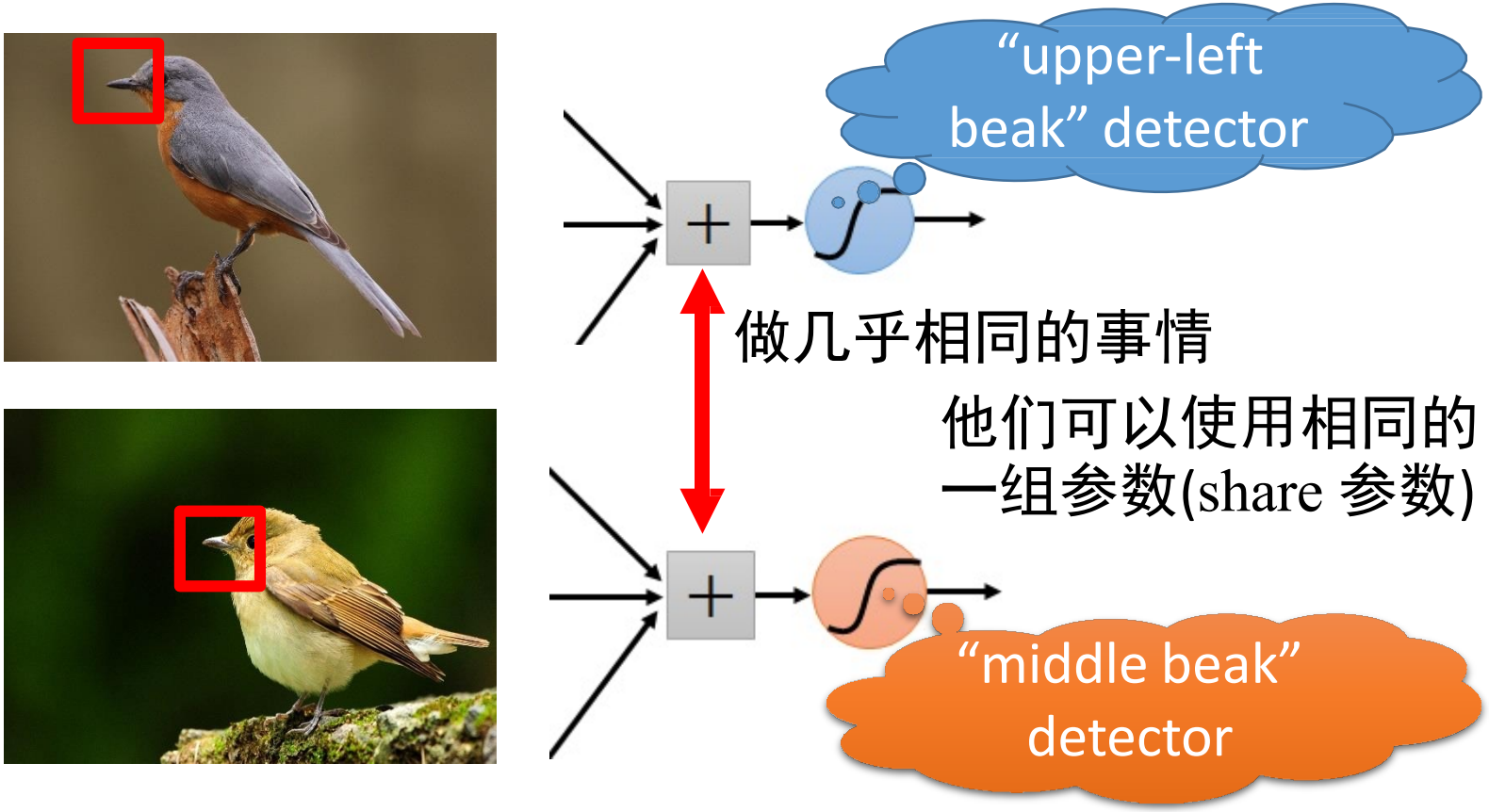
一个神经元不需要看到整个图像去发现模式

通过较少的参数连接到小区域



为什么CNN 用于图像识别?

2. 同样的模式可能出现在图像的不同区域



为什么CNN 用于图像识别?

3. 对图像二次采样不会改变对物体的认知

- 只要不是太过分（电子包浆）

二次采样

我们可以对图像进行二次采样以使图像变小

→ 更少的参数来处理图像

CNN网络结构

属性 1

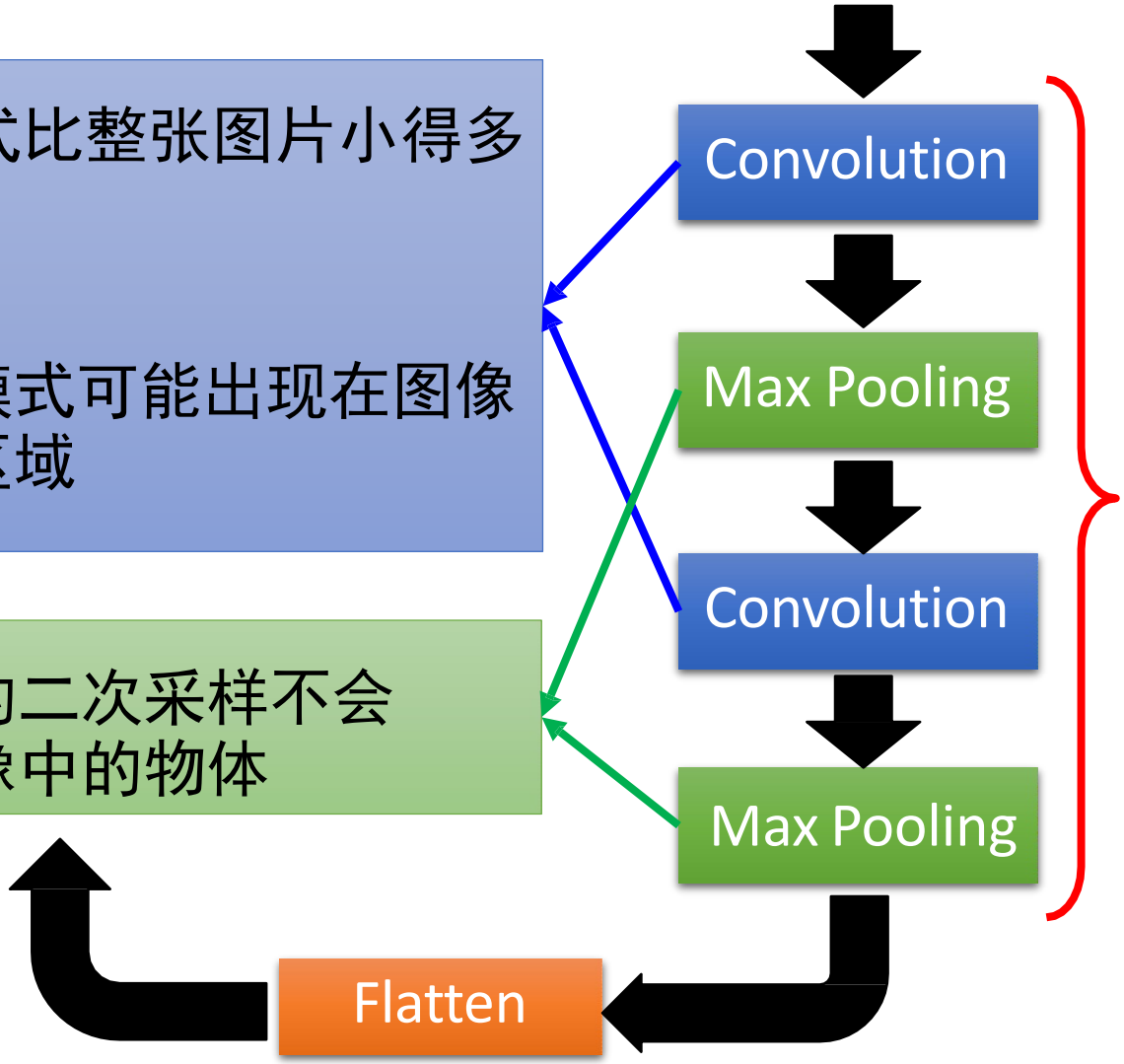
- 一些模式比整张图片小得多

属性 2

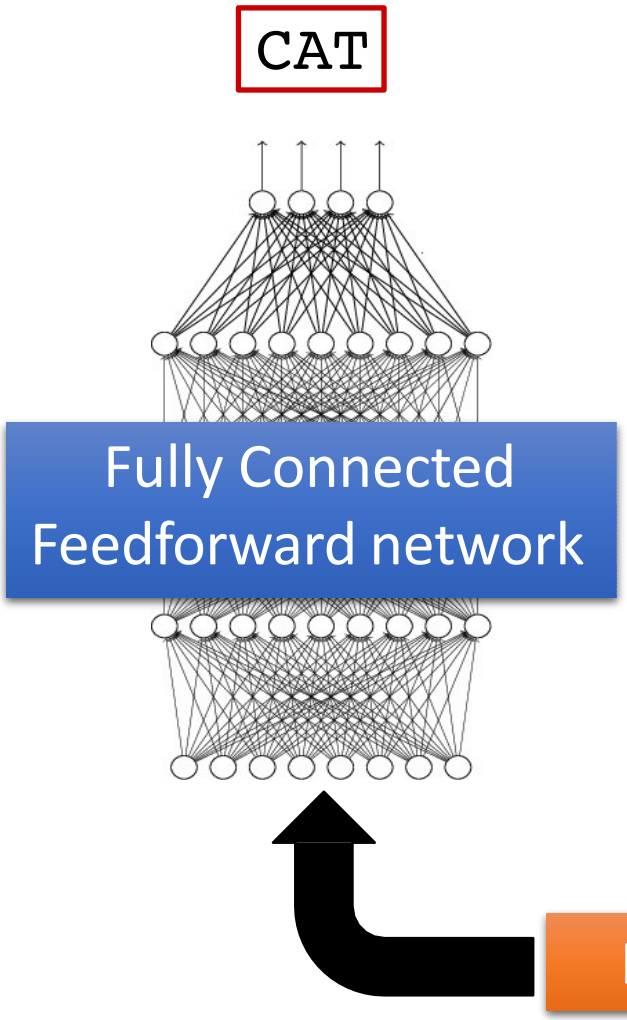
- 同样的模式可能出现在图像的不同区域

属性 3

- 对图像的二次采样不会改变图像中的物体



CNN网络结构


CAT
Fully Connected
Feedforward network
Convolution
Max Pooling
Convolution
Max Pooling
Flatten
可以重复
多次

CNN – Convolution

- Kernel: 卷积核

注意：这些kernel都是要学习的参数！

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

6 x 6 image

1	-1	-1
-1	1	-1
-1	-1	1

Kernel 1 矩阵

-1	1	-1
-1	1	-1
-1	1	-1

Kernel 2 矩阵

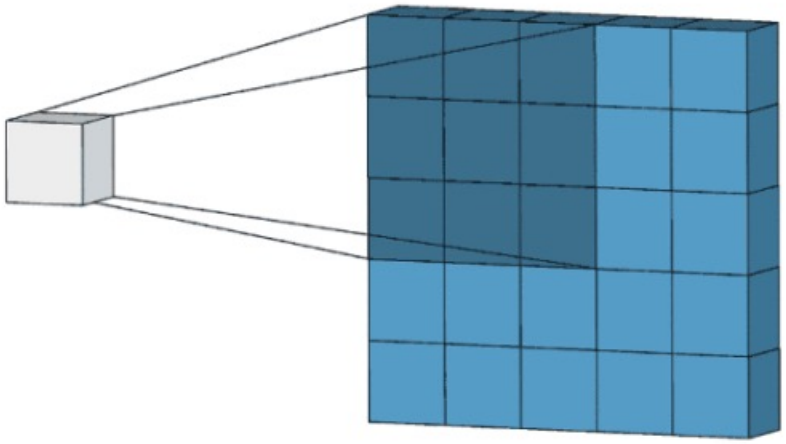
⋮

每个 kernel 探测一个小的模式(3 x 3)

属性一：一些模式比整张图片小得多

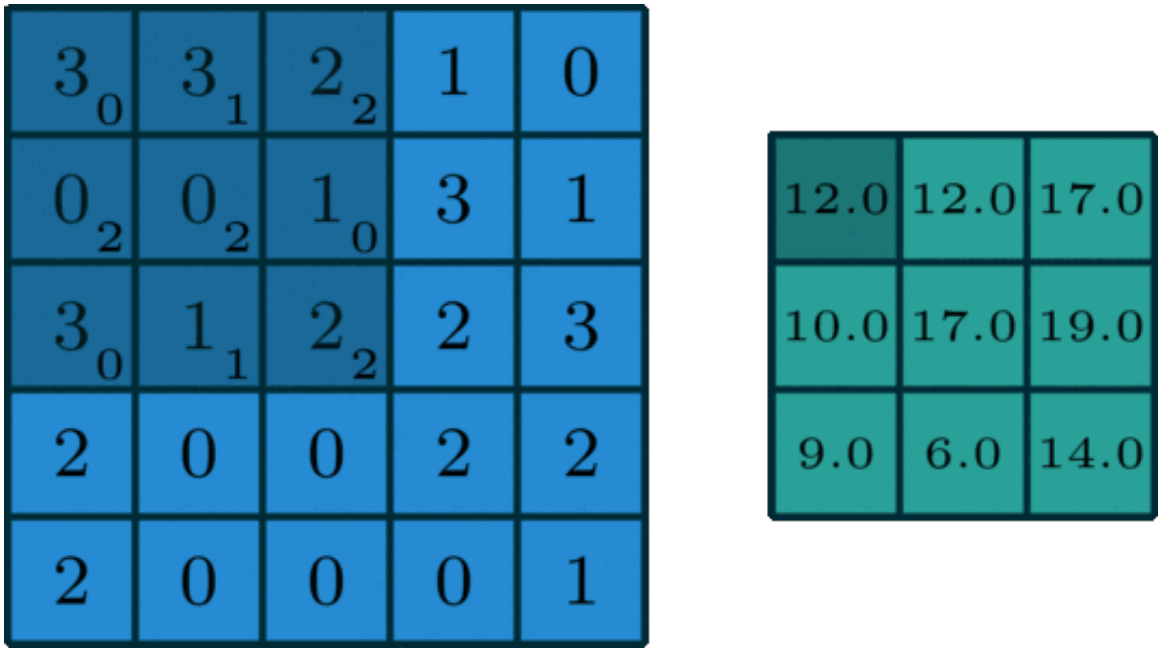
CNN – Convolution

- **卷积是一个相当简单的操作：**先从一个小小的权重矩阵，也就是卷积核（kernel）开始，让它逐步在二维输入数据上“扫描”。卷积核“滑动”的同时，计算权重矩阵和扫描所得的数据矩阵的乘积，然后把结果汇总成一个输出像素



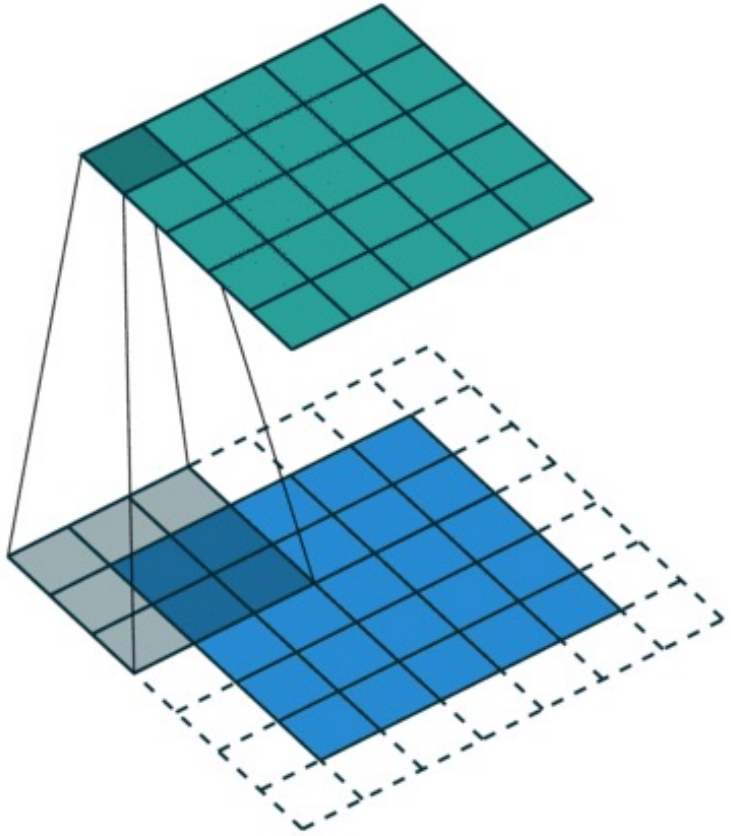
CNN – Convolution

- 输出的特征是原输入特征的加权和
- 权重是卷积核自带的值
- 像素位置所处的地方大致相同



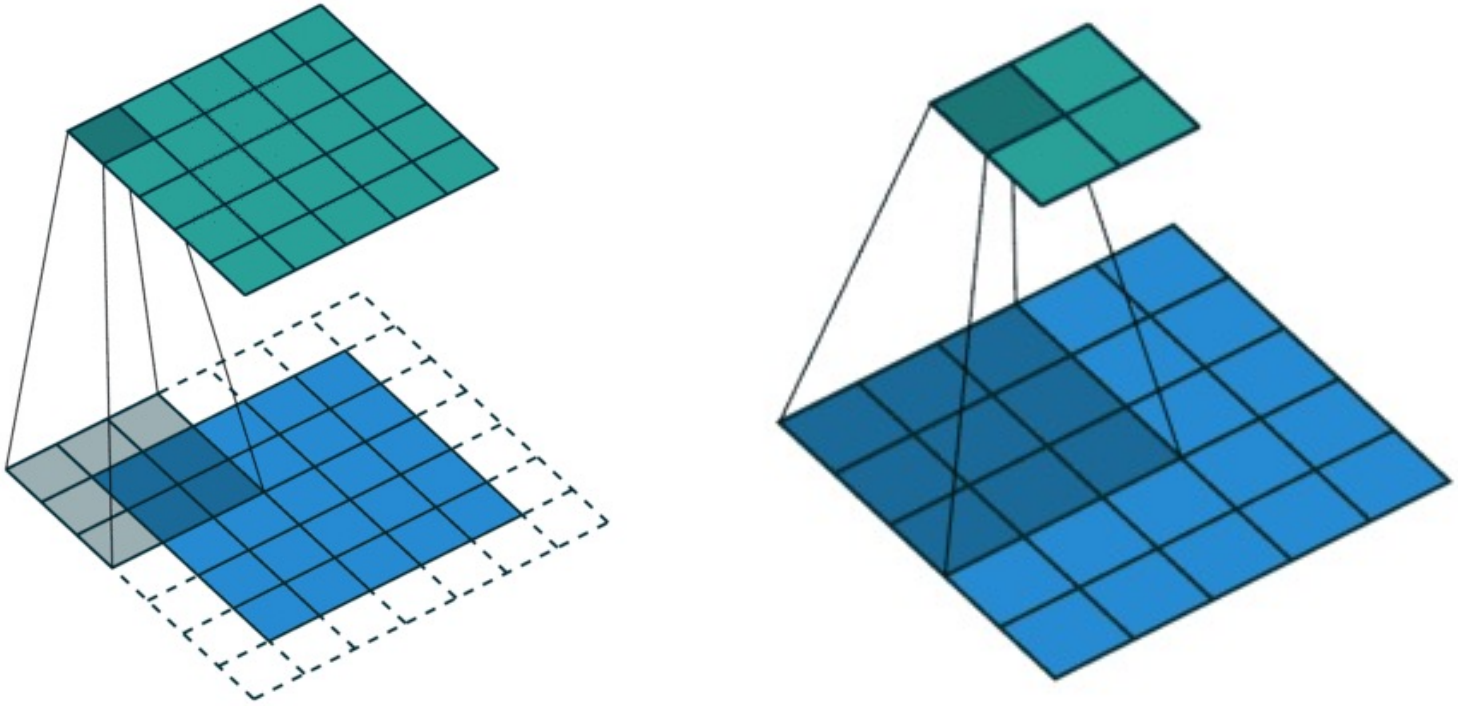
CNN – Convolution

- **Padding: 填充**
  - 用额外的“假”像素填充边缘（值一般为0）



CNN – Convolution

- **Stride: 步长**
  - 每次滑动的行数和列数称为Stride



CNN – Convolution

步长=1

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

6 x 6 image

1	-1	-1
-1	1	-1
-1	-1	1

Kernel

-3	-1	-3	-1
-3	-1	0	-3
-3	-3	0	1
-3	-2	-2	-1

属性二：同样的模式可能出现在图像的不同区域

CNN – Convolution

对于每一个 kernel 重复卷积操作

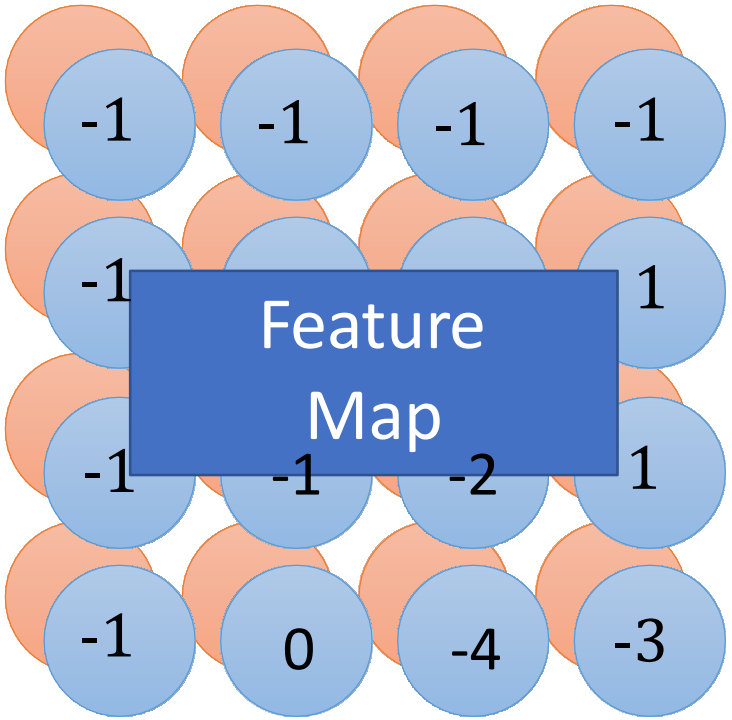
-1	1	-1
-1	1	-1
-1	1	-1

Kernel 2

步长=1

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

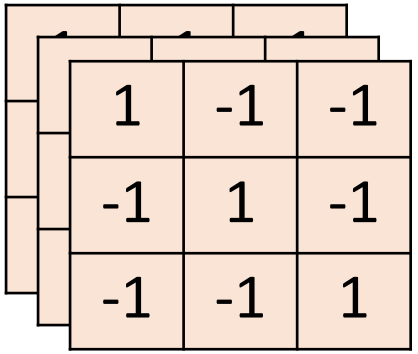
6 x 6 image



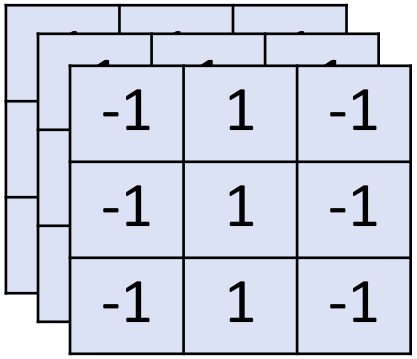
4 x 4 image * 2

CNN – Convolution

彩色图像

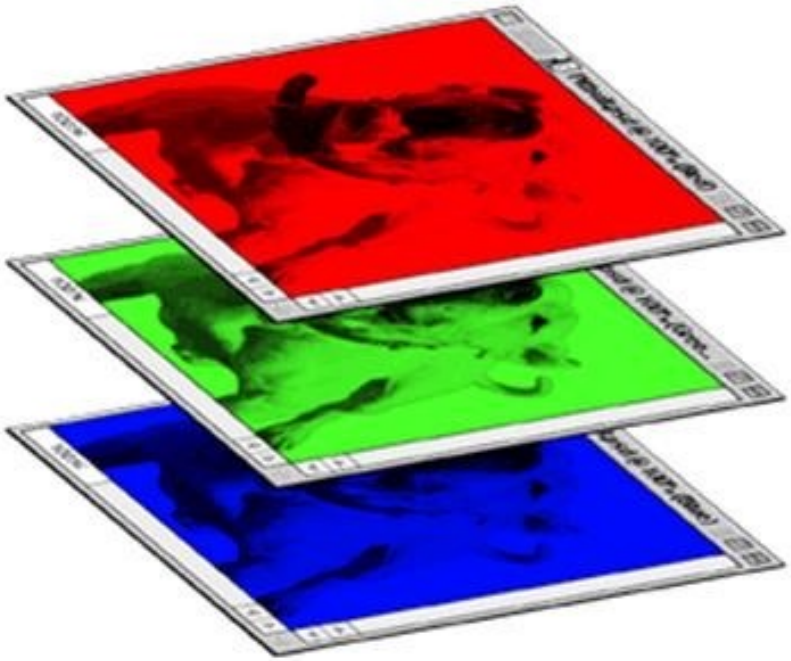


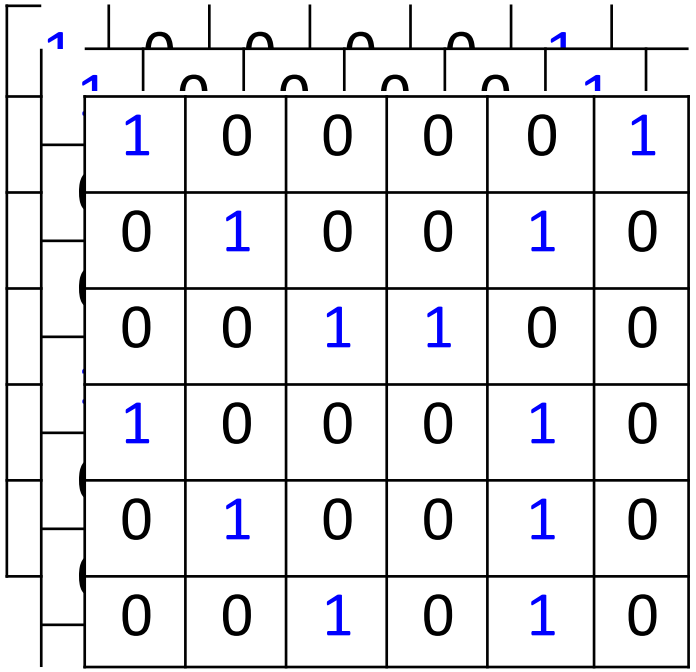
Kernel 1



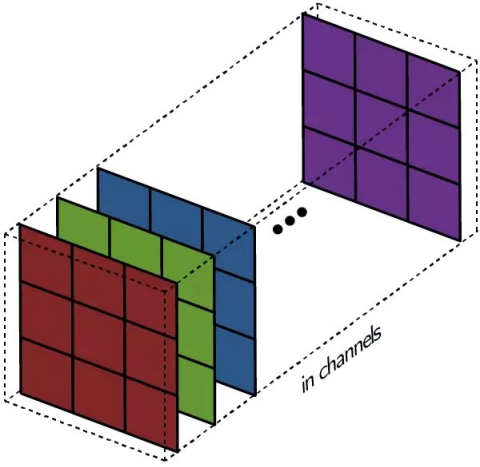
Kernel 2

三通道

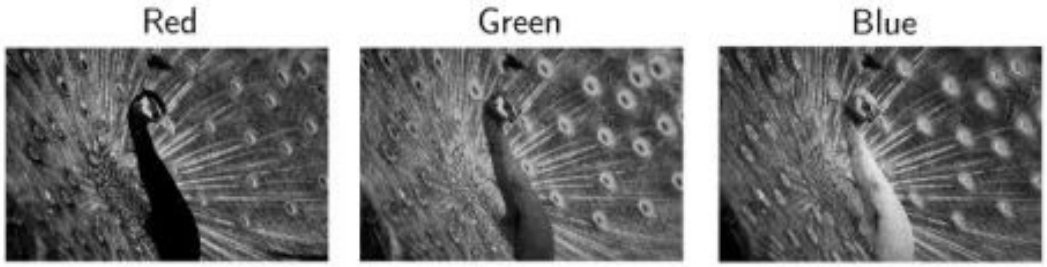


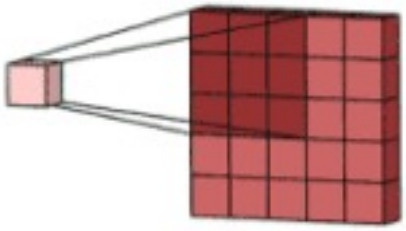


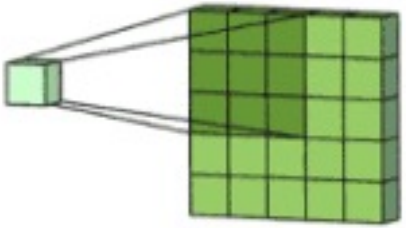
CNN – Convolution

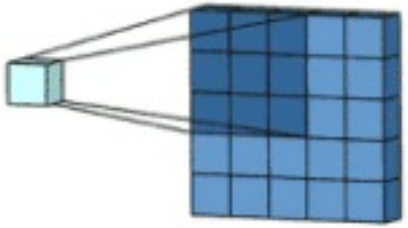


- Channel: 通道

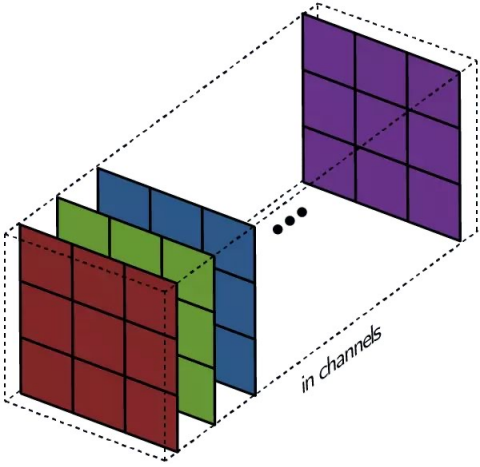




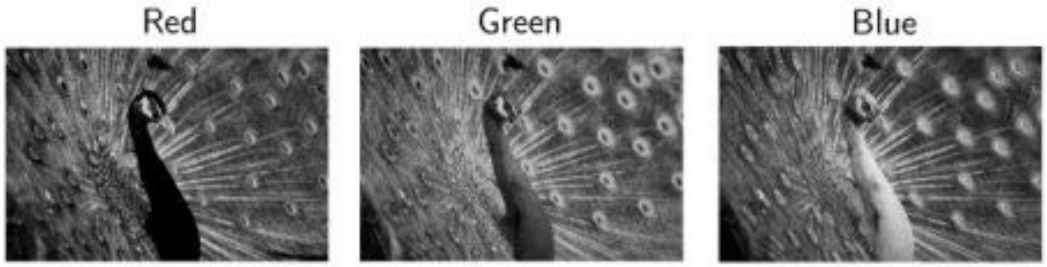


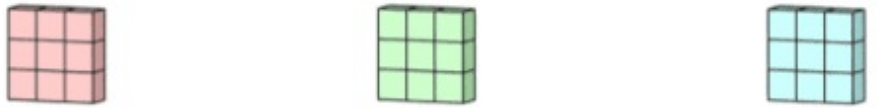


CNN – 彩色图像



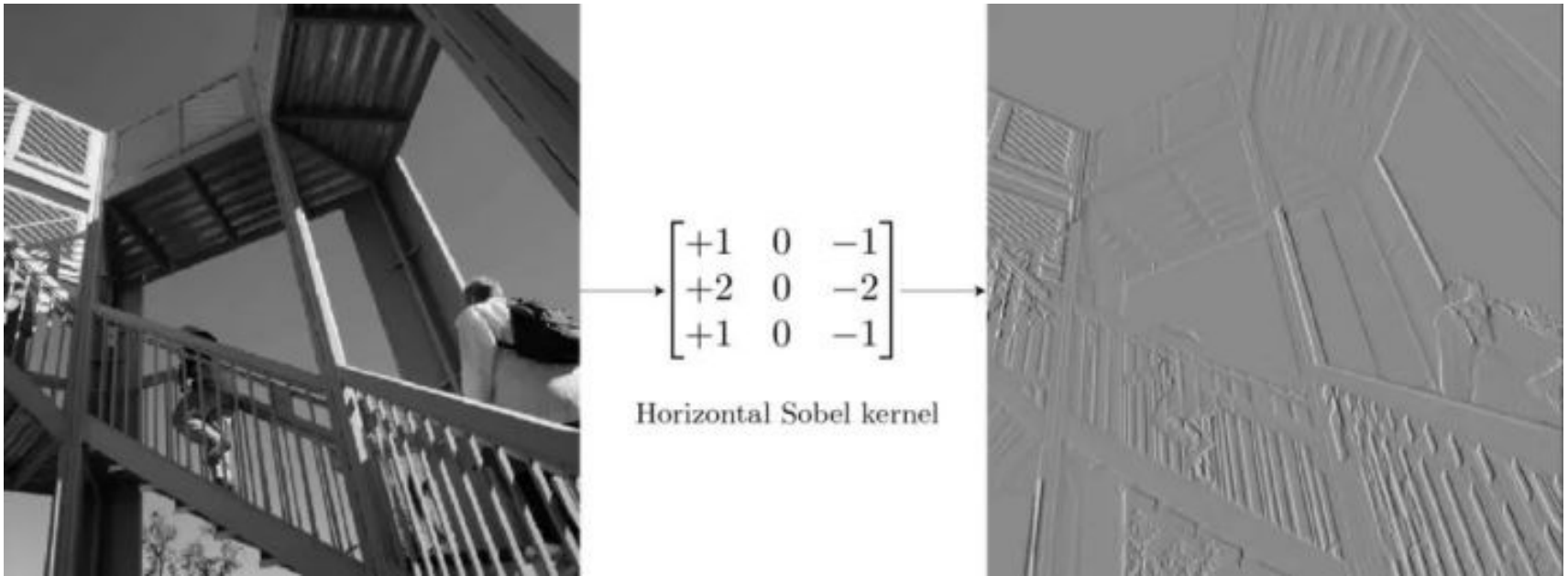
- Channel: 通道





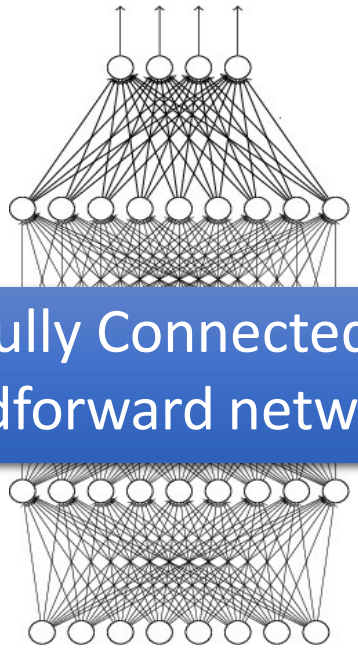
CNN – Convolution

- 卷积核的例子
  - 边缘检测核：Sobel edge detection



The whole CNN

cat dog ……



Fully Connected Feedforward network

Convolution

Max Pooling

Convolution

Max Pooling

Can repeat many times

Flatten

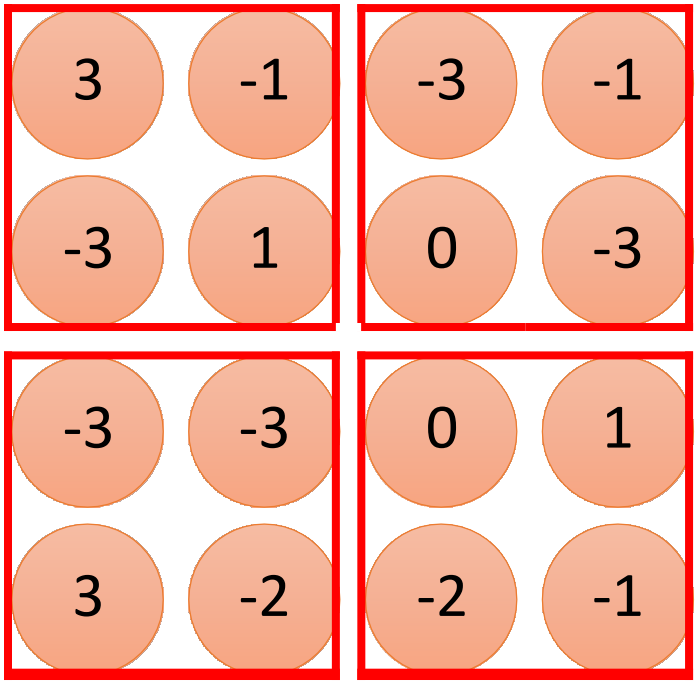
CNN – Max Pooling

1	-1	-1
-1	1	-1
-1	-1	1

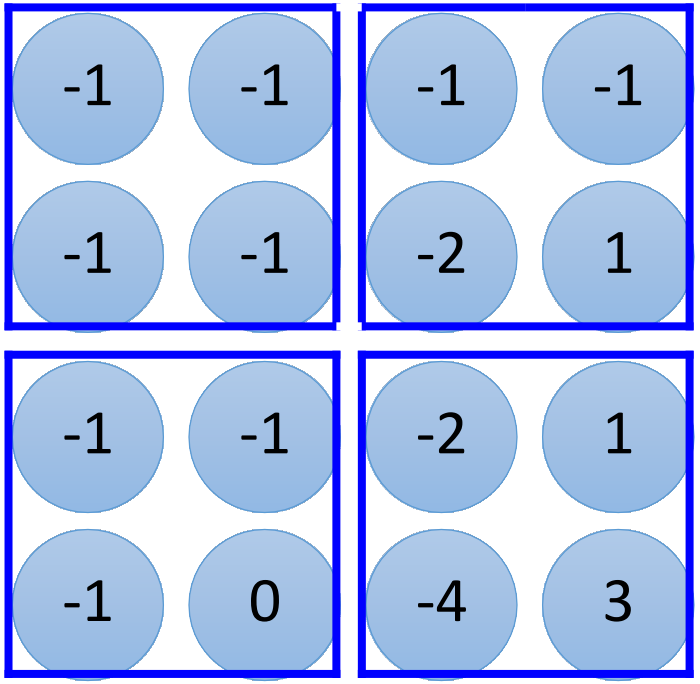
Filter 1

-1	1	-1
-1	1	-1
-1	1	-1

Filter 2

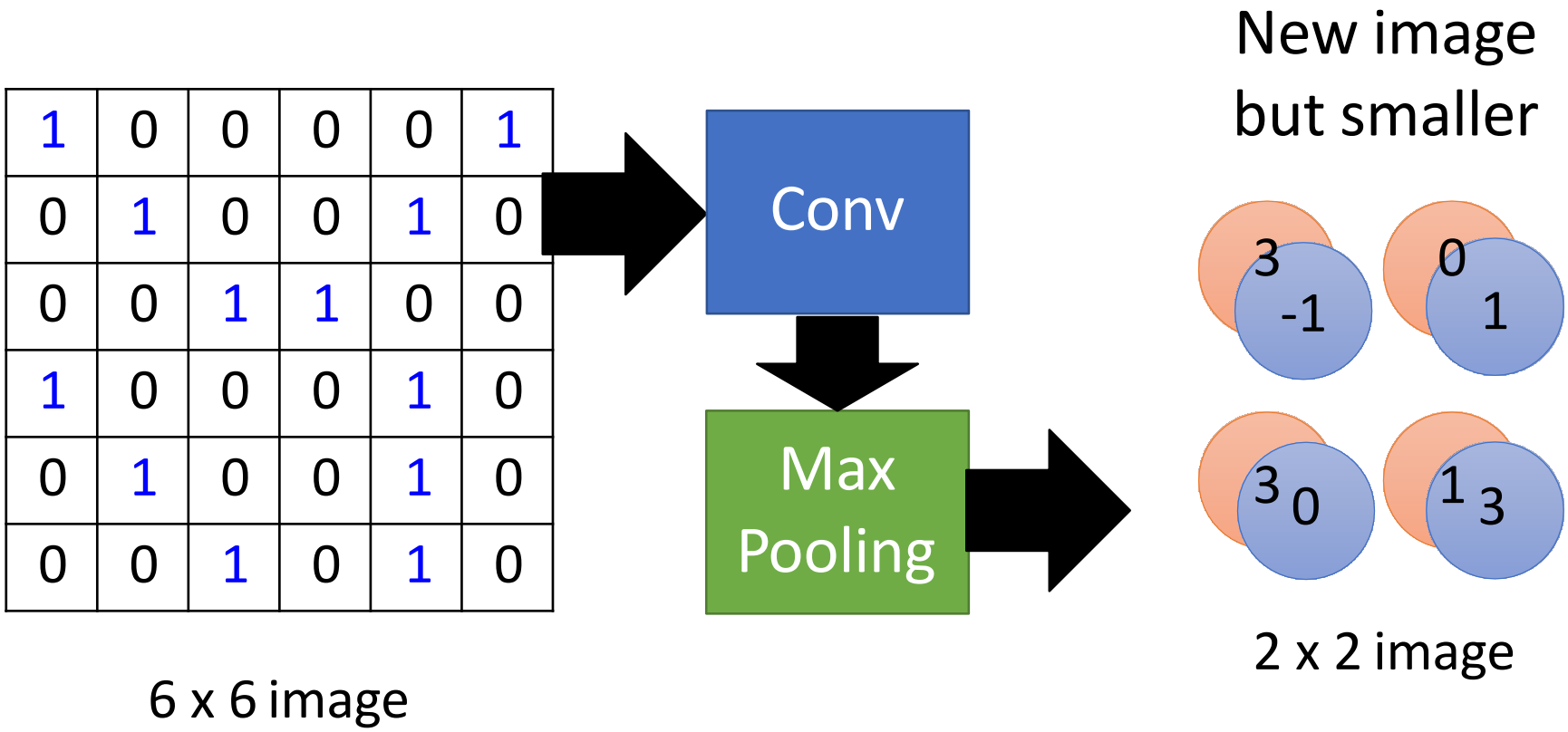


4 x 4 feature map

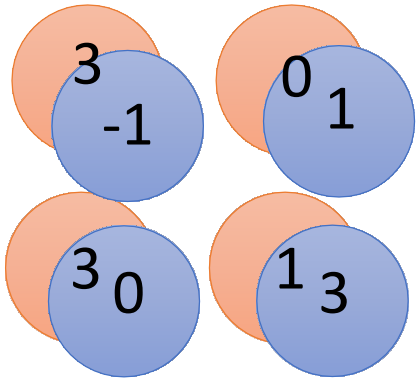


4 x 4 feature map

CNN网络结构


1 0 0 0 0 1
0 1 0 0 1 0
0 0 1 1 0 0
1 0 0 0 1 0
0 1 0 0 1 0
0 0 1 0 1 0
Conv
Max Pooling
New image but smaller
3 -1
0 1
3 0
1 3
6 x 6 image
2 x 2 image

CNN网络结构



新的图像

比原始图像更小

Convolution

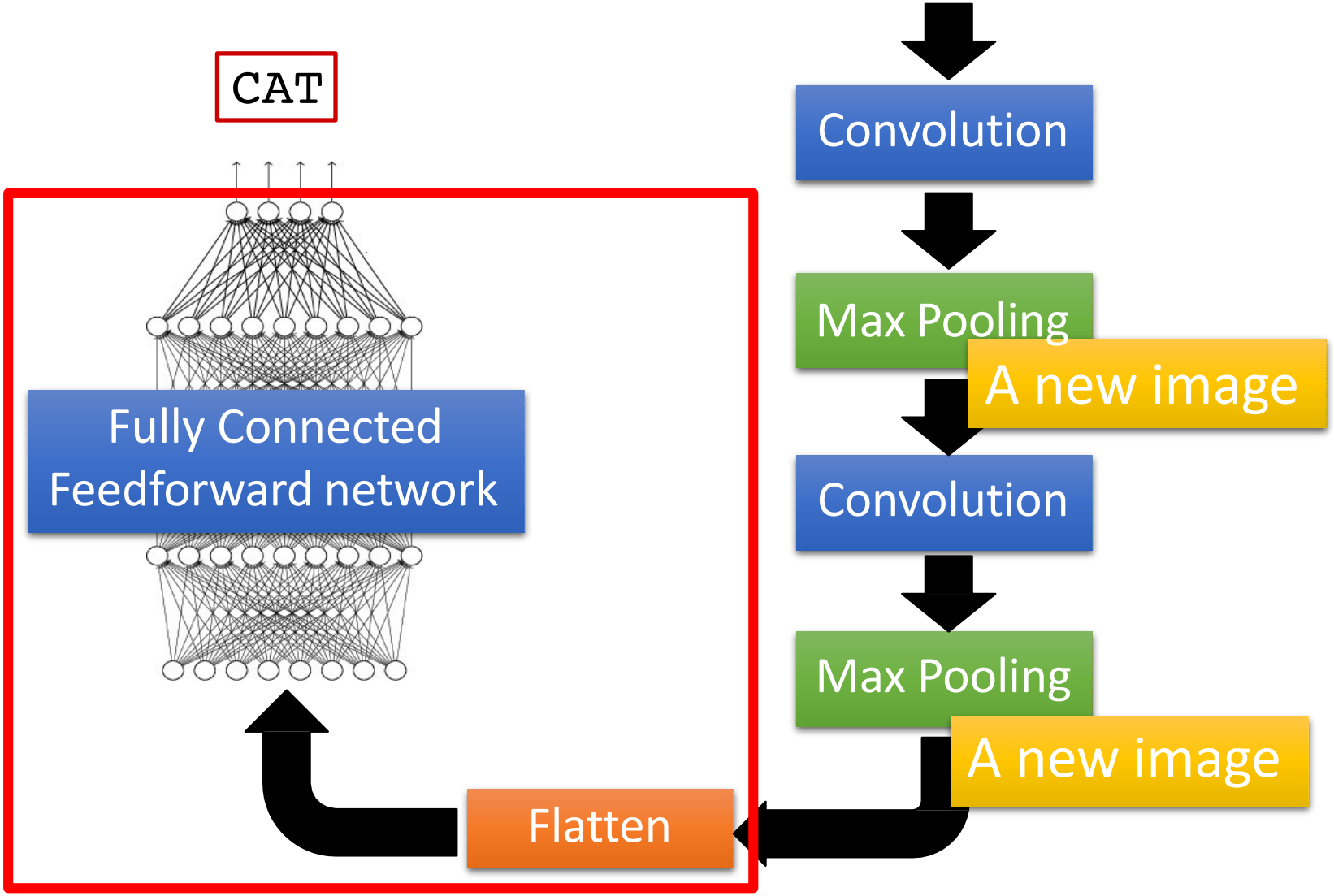
Max Pooling

Convolution

Max Pooling

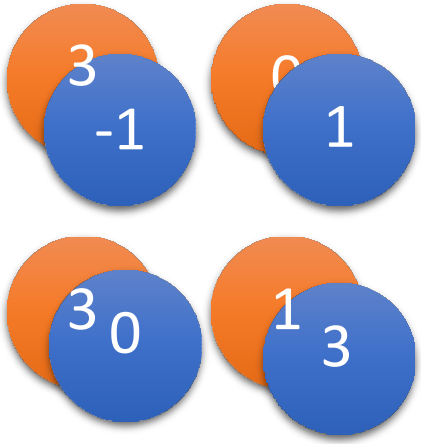
重复多次

CNN网络结构

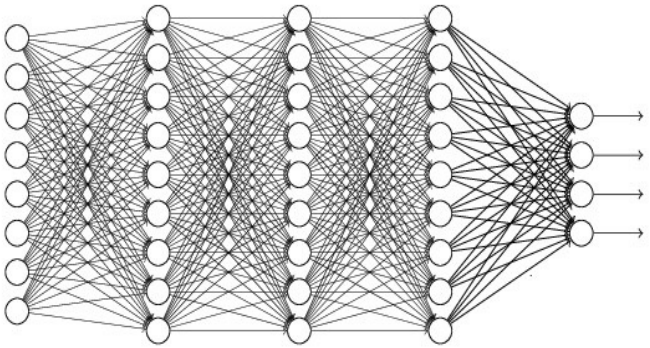

CAT
Convolution
Max Pooling
A new image
Convolution
Max Pooling
A new image
Flatten
Fully Connected
Feedforward network

Flatten

- 特征图 -> 特征向量



Flatten



Fully Connected Feedforward network

End