CNN神经网络

Steven Tang

主要课程内容

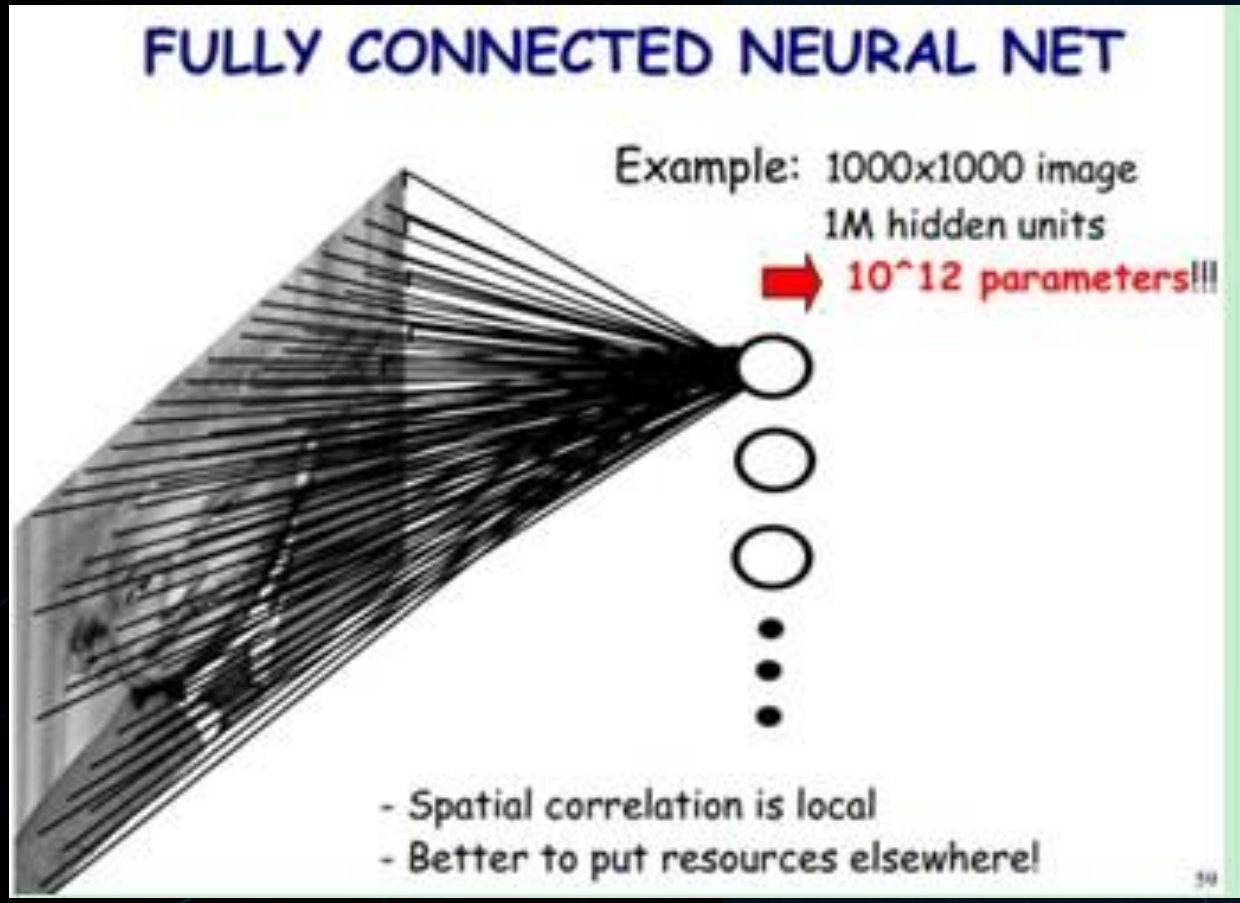
什么是特征提取？

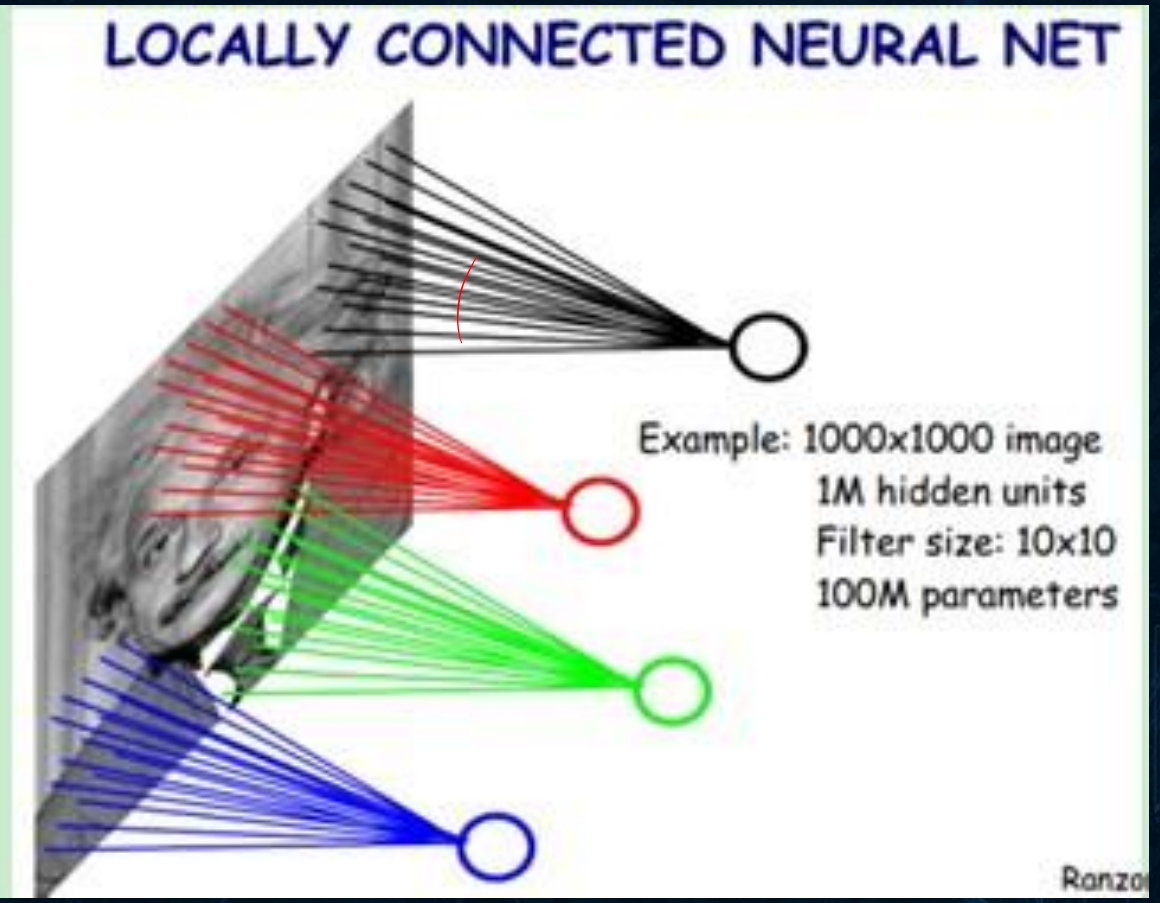
卷积的概念理解

卷积层，池化层，全连接层

简单的卷积神经网络框架

卷积操作回顾





卷积操作回顾

INPUT IMAGE

18	54	51	239	244	188
55	121	75	78	95	88
35	24	204	113	109	221
3	154	104	235	25	130
15	253	225	159	78	233
68	85	180	214	245	0

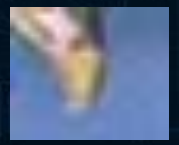
WEIGHT

1	0	1
0	1	0
1	0	1

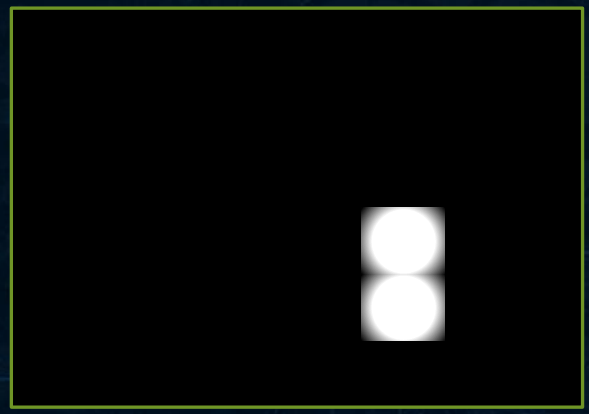
429

卷积操作的关键（通道）

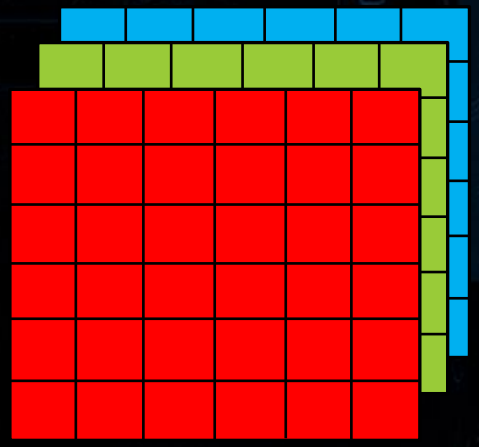
*



=

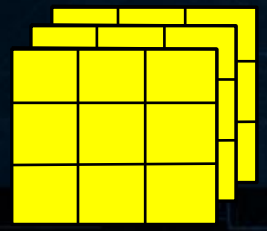


输入



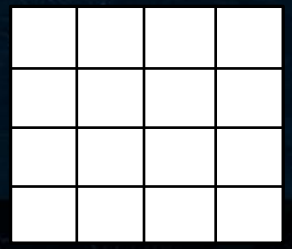
*

卷积核



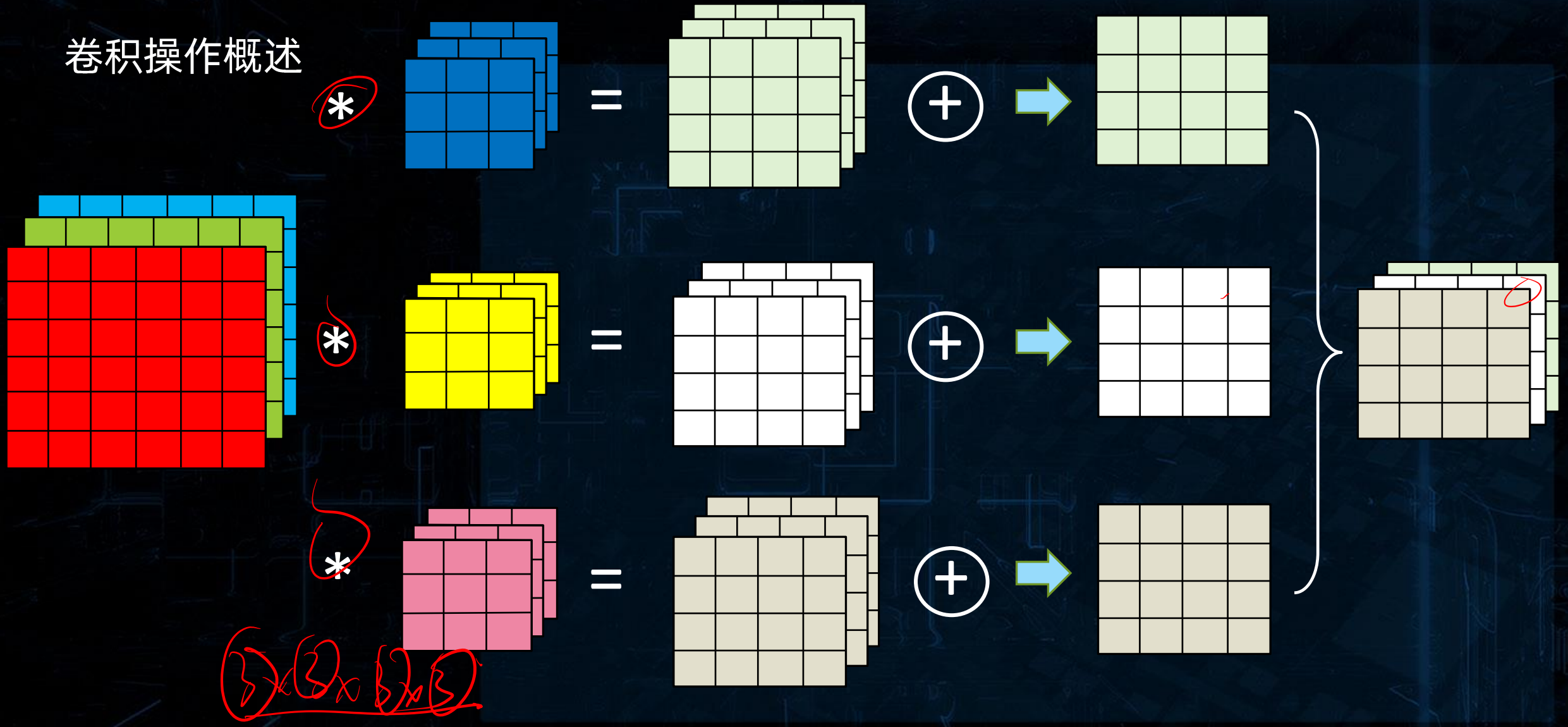
=

输出

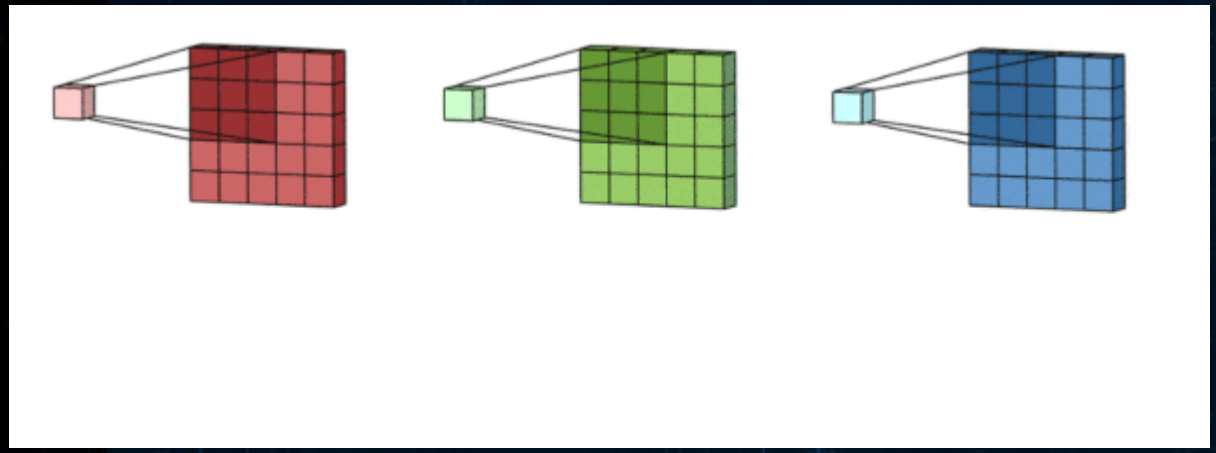


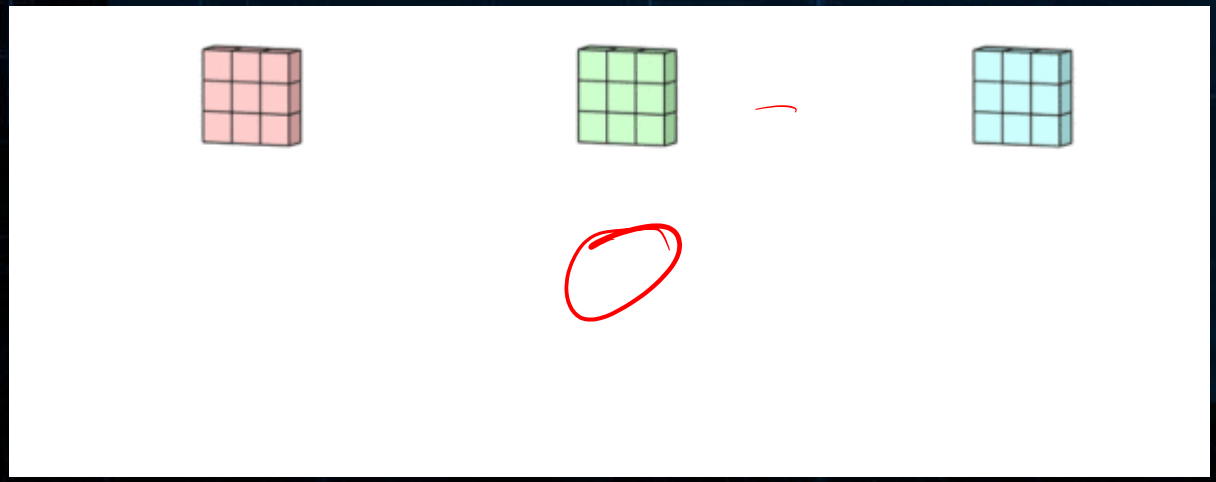
渡一教育 Duyi Education

卷积操作概述

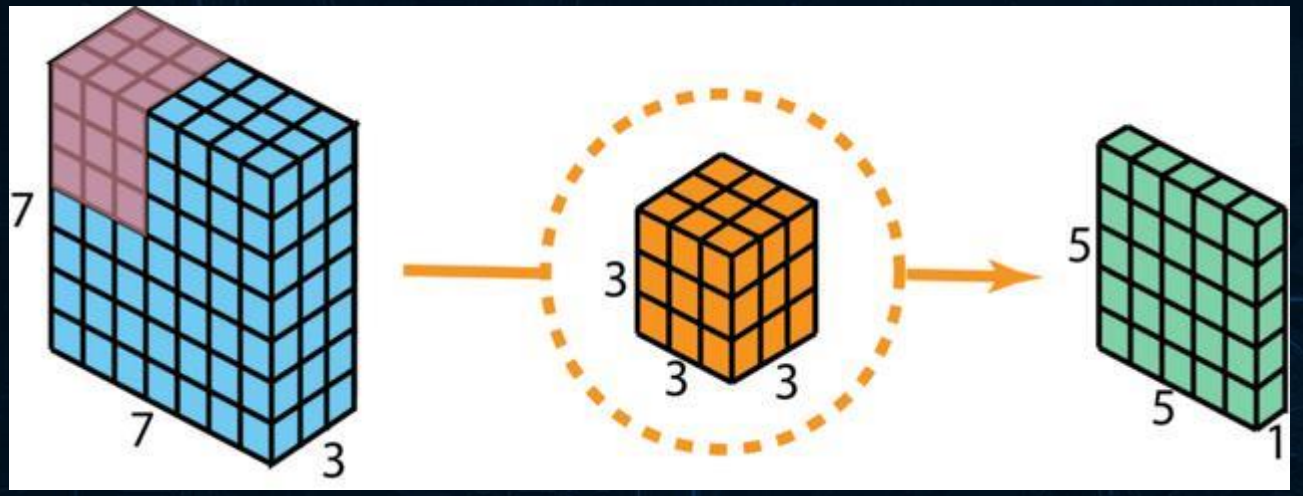


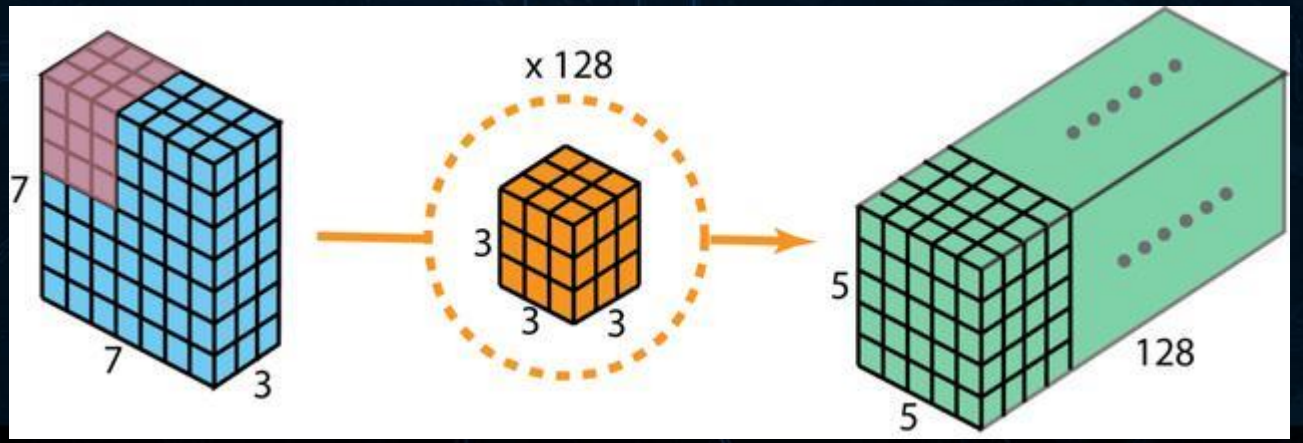
多通道卷积

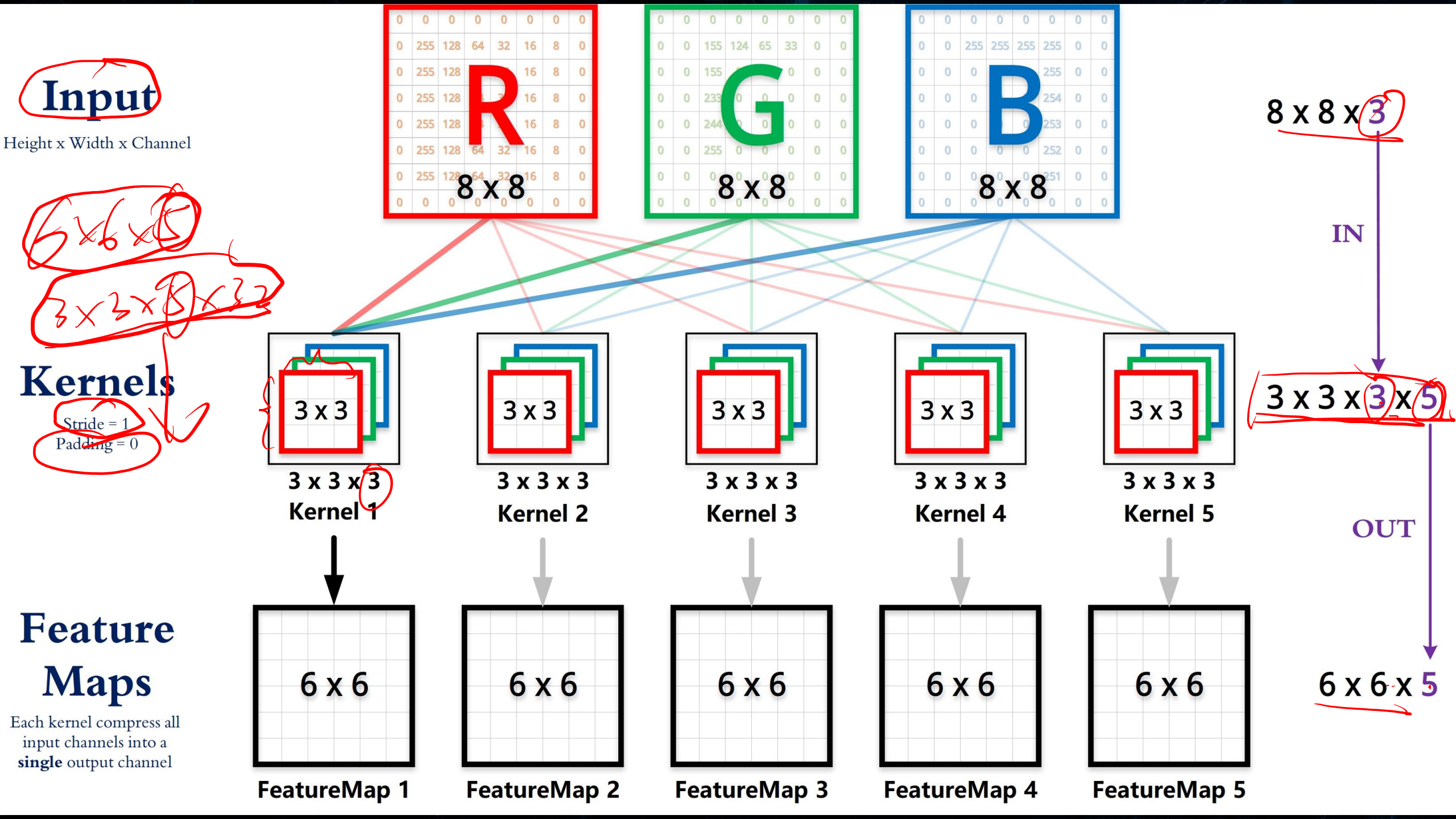




卷积核的各个维度（长，宽，通道数，卷积核个数）

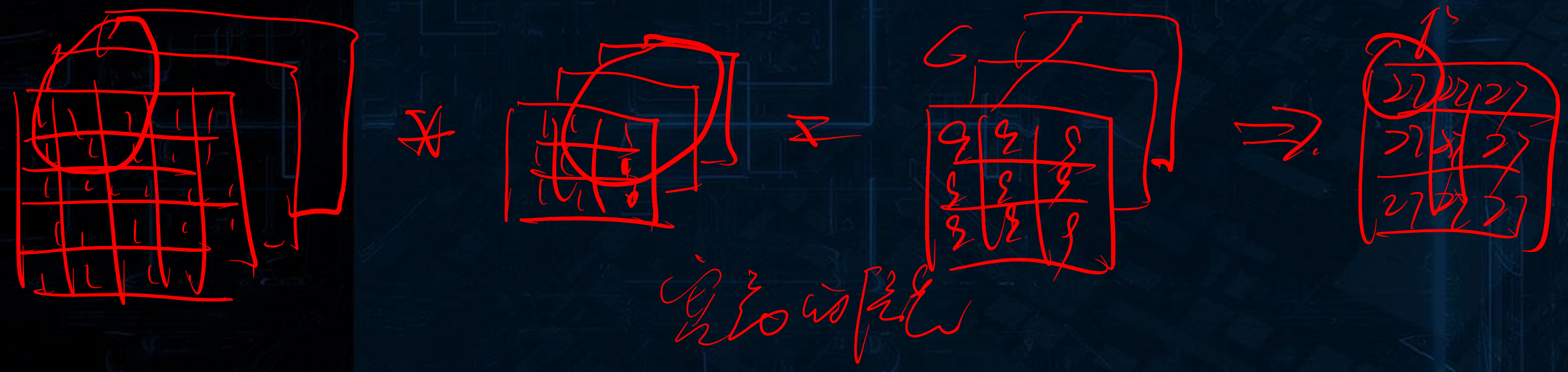




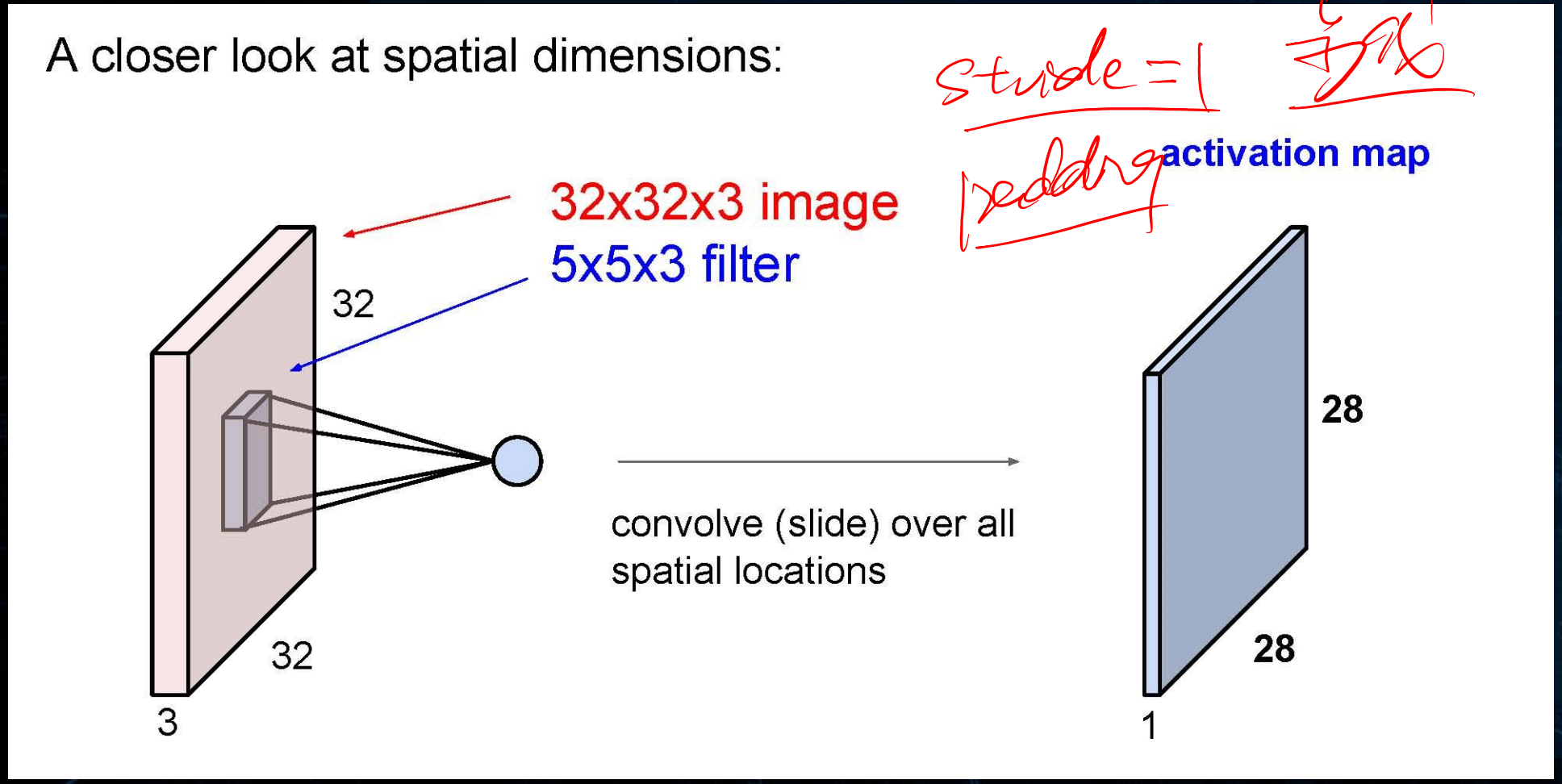

Input
Height x Width x Channel
R
G
B
8 x 8
8 x 8 x 3
IN
Kernels
Stride = 1
Padding = 0
3 x 3
3 x 3 x 3
Kernel 1
Kernel 2
Kernel 3
Kernel 4
Kernel 5
3 x 3 x 3 x 5
OUT
Feature Maps
Each kernel compress all input channels into a **single** output channel
6 x 6
FeatureMap 1
FeatureMap 2
FeatureMap 3
FeatureMap 4
FeatureMap 5
6 x 6 x 5

思考题

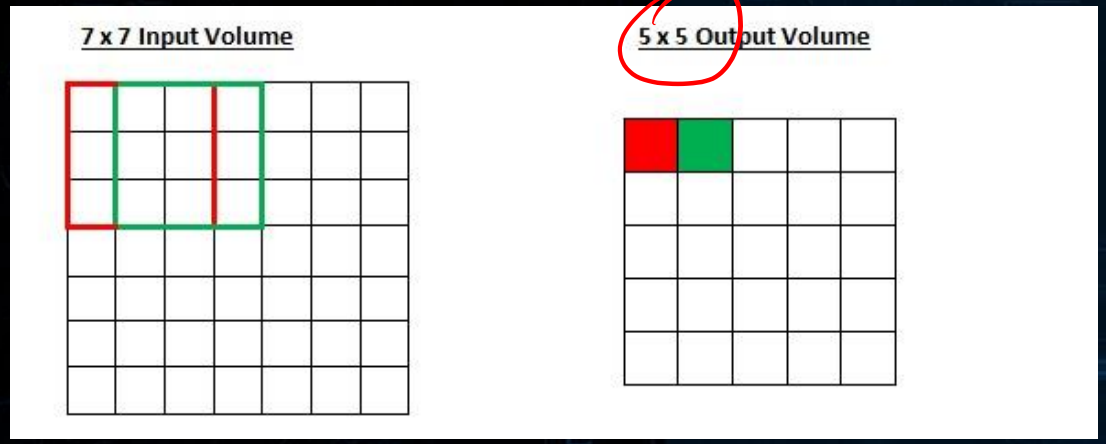
一个5*5*3的图像假设每个像素的值为1，被一个3*3*3的卷积核卷积，且该卷积核每个值也均为一，得到的输出图的尺寸和值分别为多少？

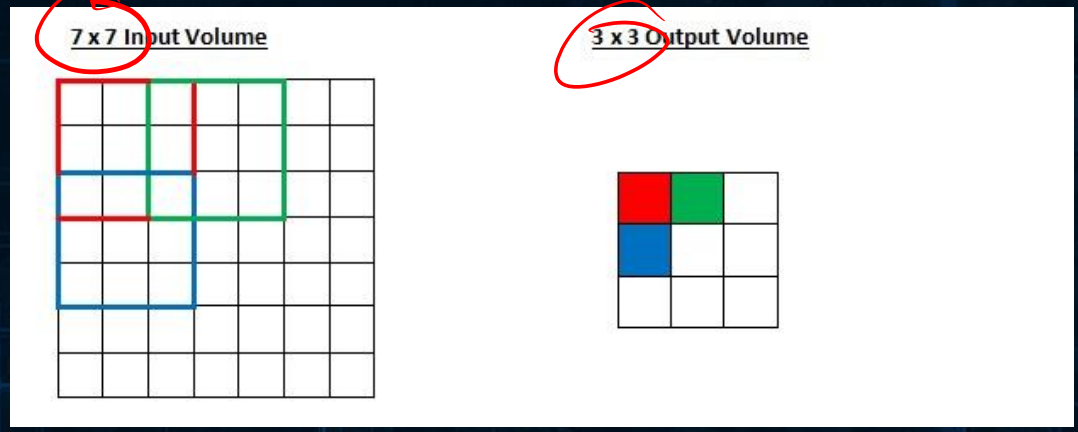


卷积层（卷积核和特征图）


A closer look at spatial dimensions:
32x32x3 image
5x5x3 filter
activation map
32
32
3
convolve (slide) over all
spatial locations
28
28
1

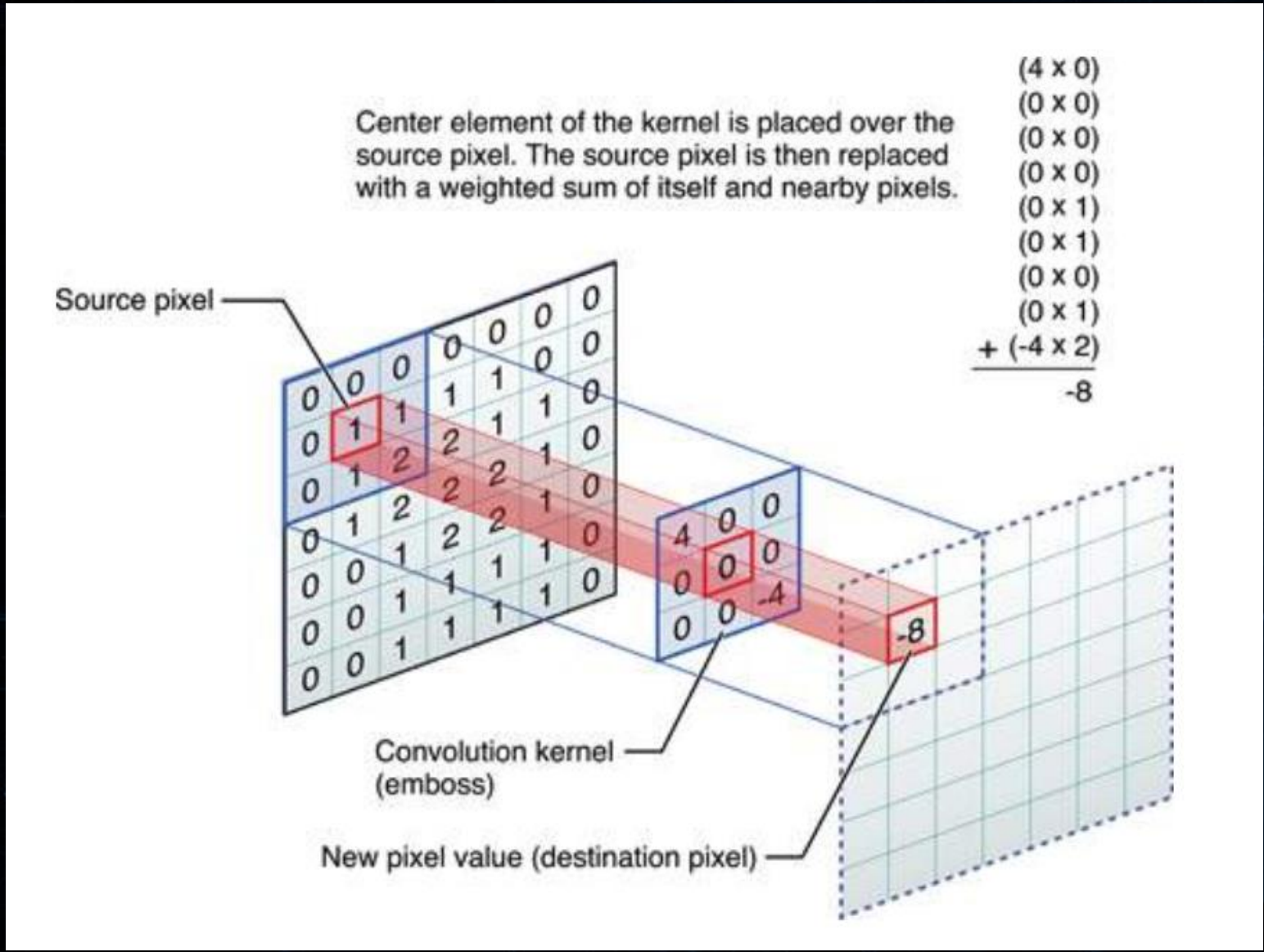
Padding：Valid，以及stride



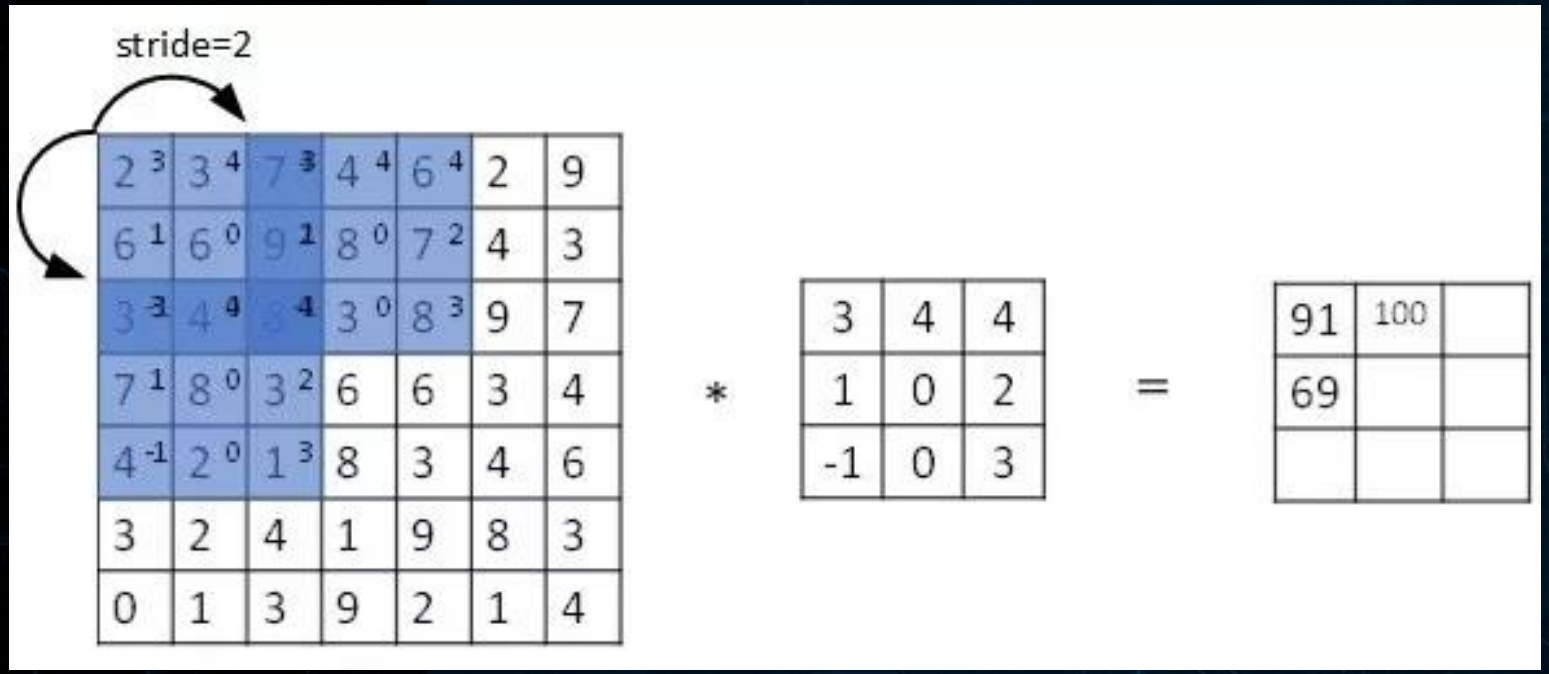


Stride：移动滤波器平移时的像素的数量，当stride=1，输出的尺寸和输入的尺寸大体相同，stride=2时候，输出尺寸相当于输入的一半。

卷积的内积操作

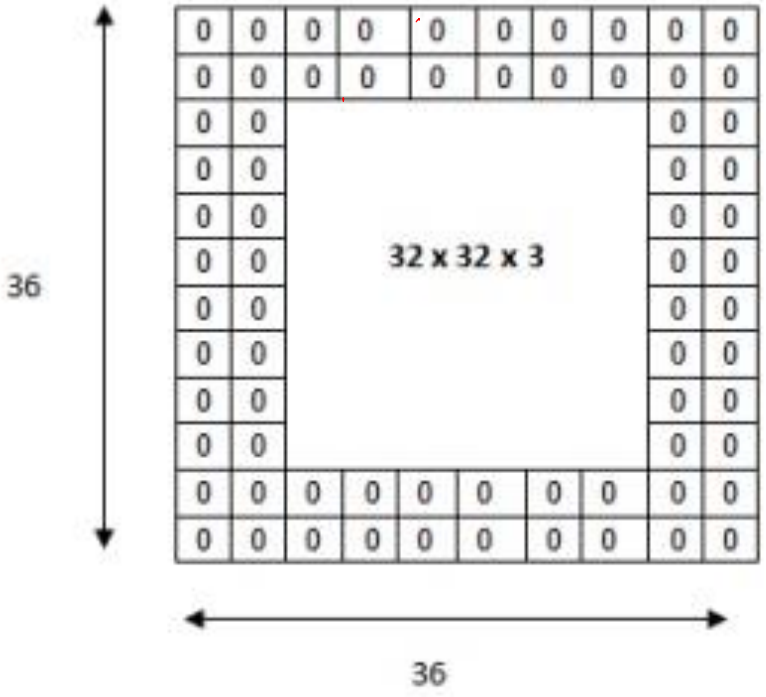

Center element of the kernel is placed over the source pixel. The source pixel is then replaced with a weighted sum of itself and nearby pixels.
(4 x 0)
(0 x 0)
(0 x 0)
(0 x 0)
(0 x 1)
(0 x 1)
(0 x 0)
(0 x 1)
+ (-4 x 2)
-8
Source pixel
Convolution kernel (emboss)
New pixel value (destination pixel)

stride

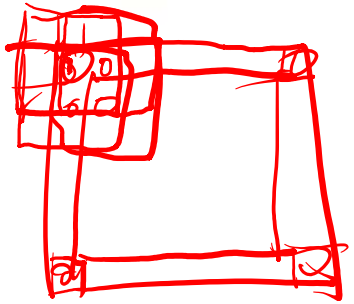


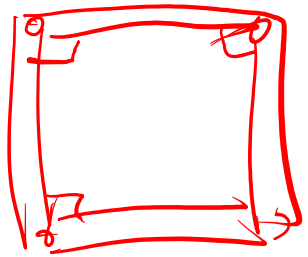
Stride：移动滤波器平移时的像素的数量，当stride=1，输出的尺寸和输入的尺寸大体相同，stride=2时候，输出尺寸相当于输入的一半。

Padding: Same



The input volume is 32 x 32 x 3. If we imagine two borders of zeros around the volume, this gives us a 36 x 36 x 3 volume. Then, when we apply our conv layer with our three 5 x 5 x 3 filters and a stride of 1, then we will also get a 32 x 32 x3 output volume.

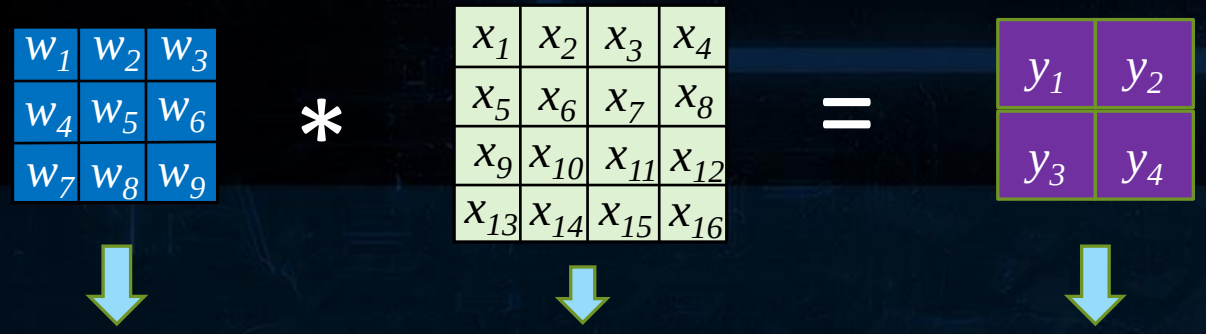




当你把 5 x 5 x 3 的过滤器用在 32 x 32 x 3 的输入上时，会发生什么？输出的大小会是 28 x 28 x 3。

如何维持输出的维度仍然为： 32 x 32 x 3 ？

将卷积看成矩阵乘法



w_1	w_2	w_3	0	w_4	w_5	w_6	0	w_7	w_8	w_9	0	0	0	0	0
0	w_1	w_2	w_3	0	w_4	w_5	w_6	0	w_7	w_8	w_9	0	0	0	0
0	0	0	0	w_1	w_2	w_3	0	w_4	w_5	w_6	0	w_7	w_8	w_9	0
0	0	0	0	0	w_1	w_2	w_3	0	w_4	w_5	w_6	0	w_7	w_8	w_9

*

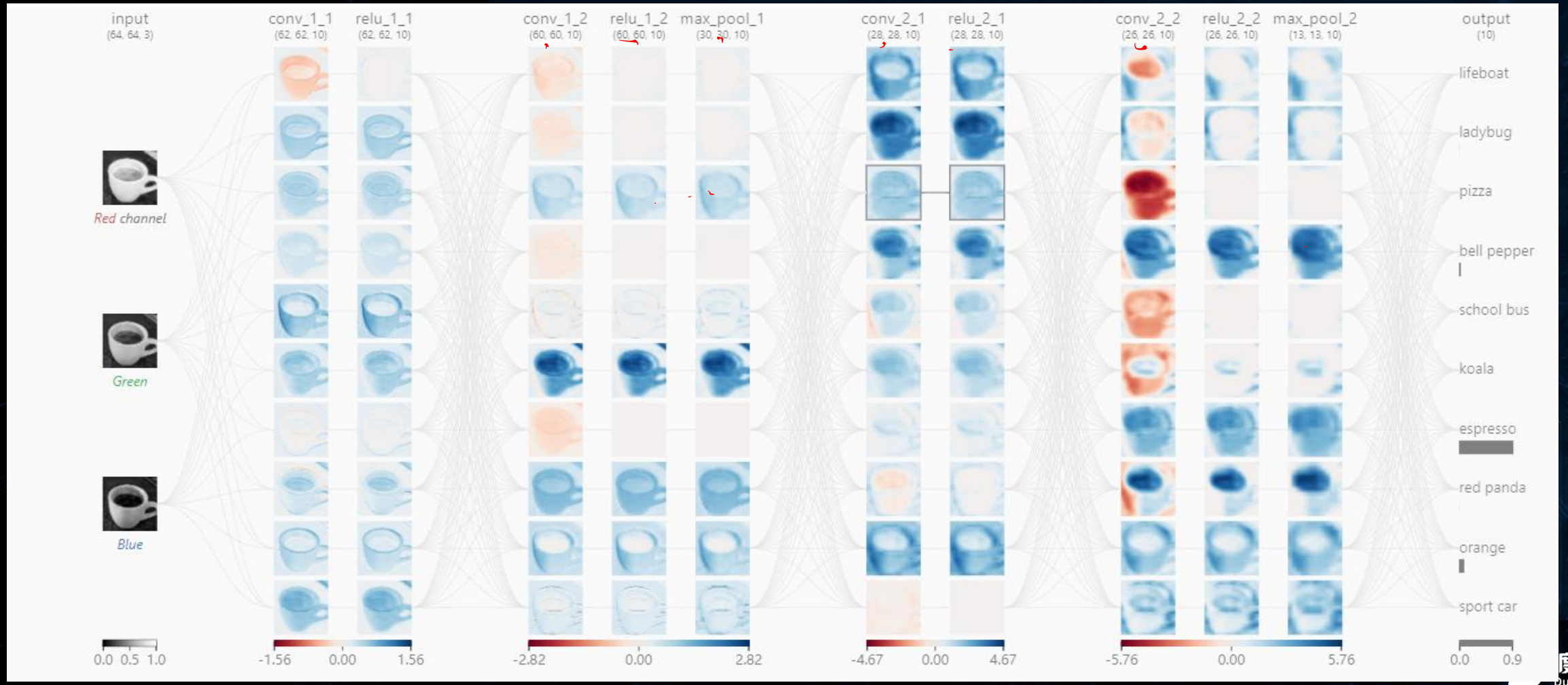
x_1
x_2
x_3
x_4
x_5
x_6
x_7
x_8
x_9
x_{10}
x_{11}
x_{12}
x_{13}
x_{14}
x_{15}
x_{16}

=

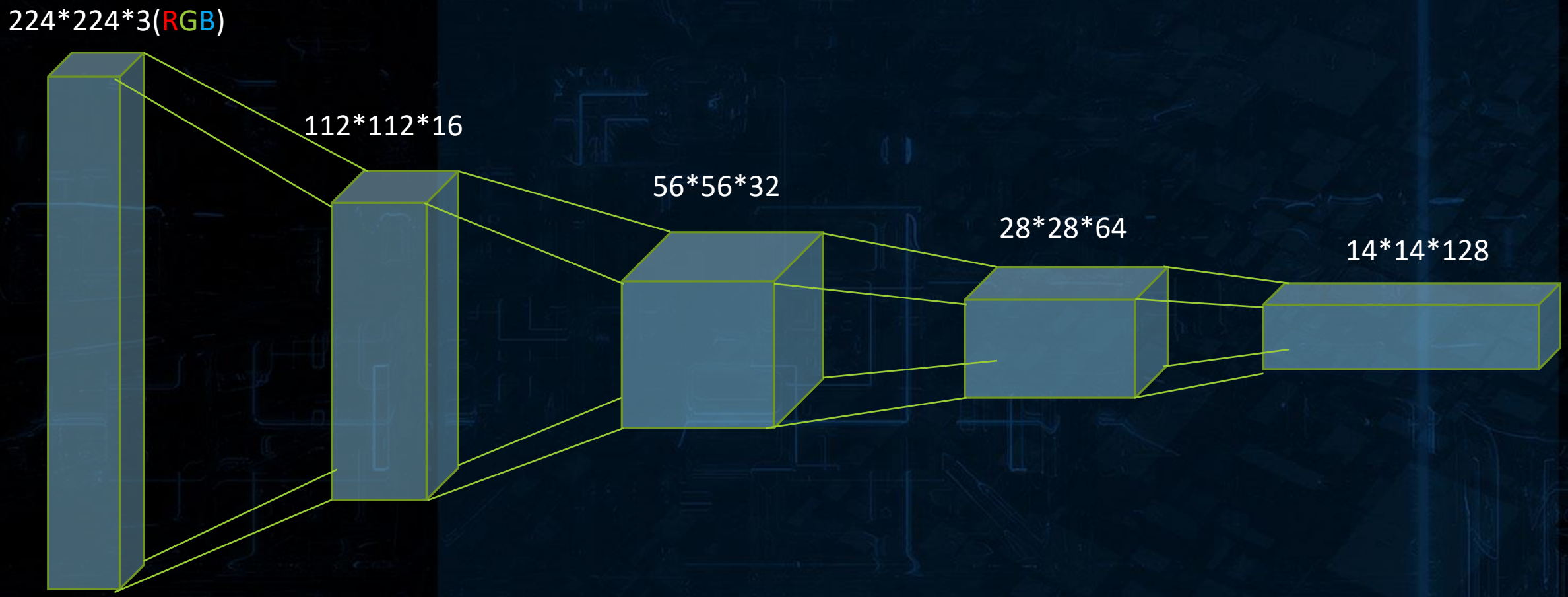
y_1
y_2
y_3
y_4

x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}

卷积神经网络框架：卷积->池化->卷积->池化->全连接->softmax



卷积网络概述



卷积层（卷积核和特征图）

每一个过滤器的输出被堆叠在一起，形成卷积图像的纵深维度。假设我们有一个 32*32*3 的输入。我们使用 5*5*3，带有 valid padding 的 10 个过滤器。输出的维度将会是 28*28*10

卷积神经网络输出shape练习

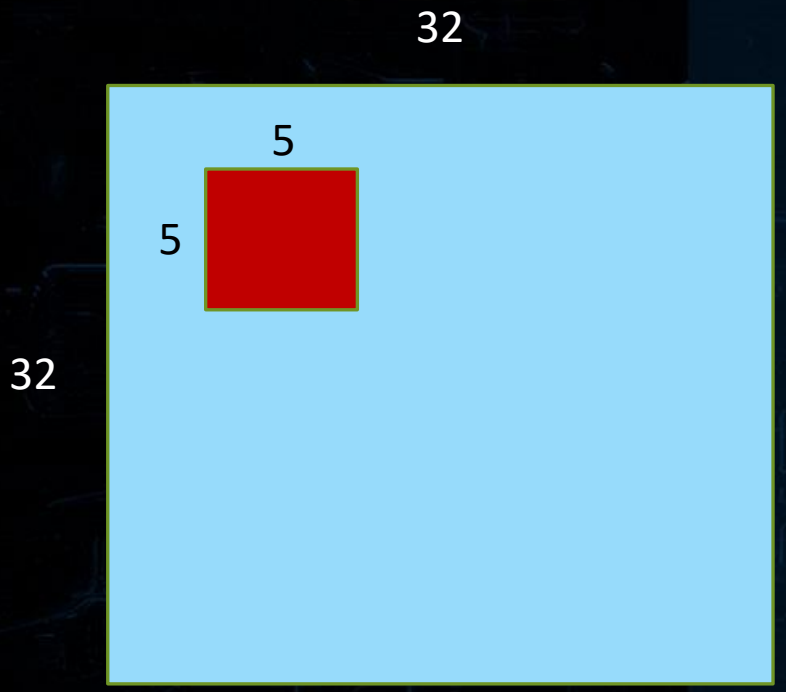
H=高度， W=宽度， D=深度
Input shape=32*32*3(H*W*D)
20个卷积核，卷积核的尺寸为8*8*3
高和宽的stride=2
Padding size=1
求输出的shape:应该为 H*W*D

公式如下：

new_height = (input_height - filter_height + 2 * P)/S + 1

new_width = (input_width - filter_width + 2 * P)/S + 1

特征图尺寸练习



输入通道数为32

输出通道数为64

32 x 32 x 32

5 x 5

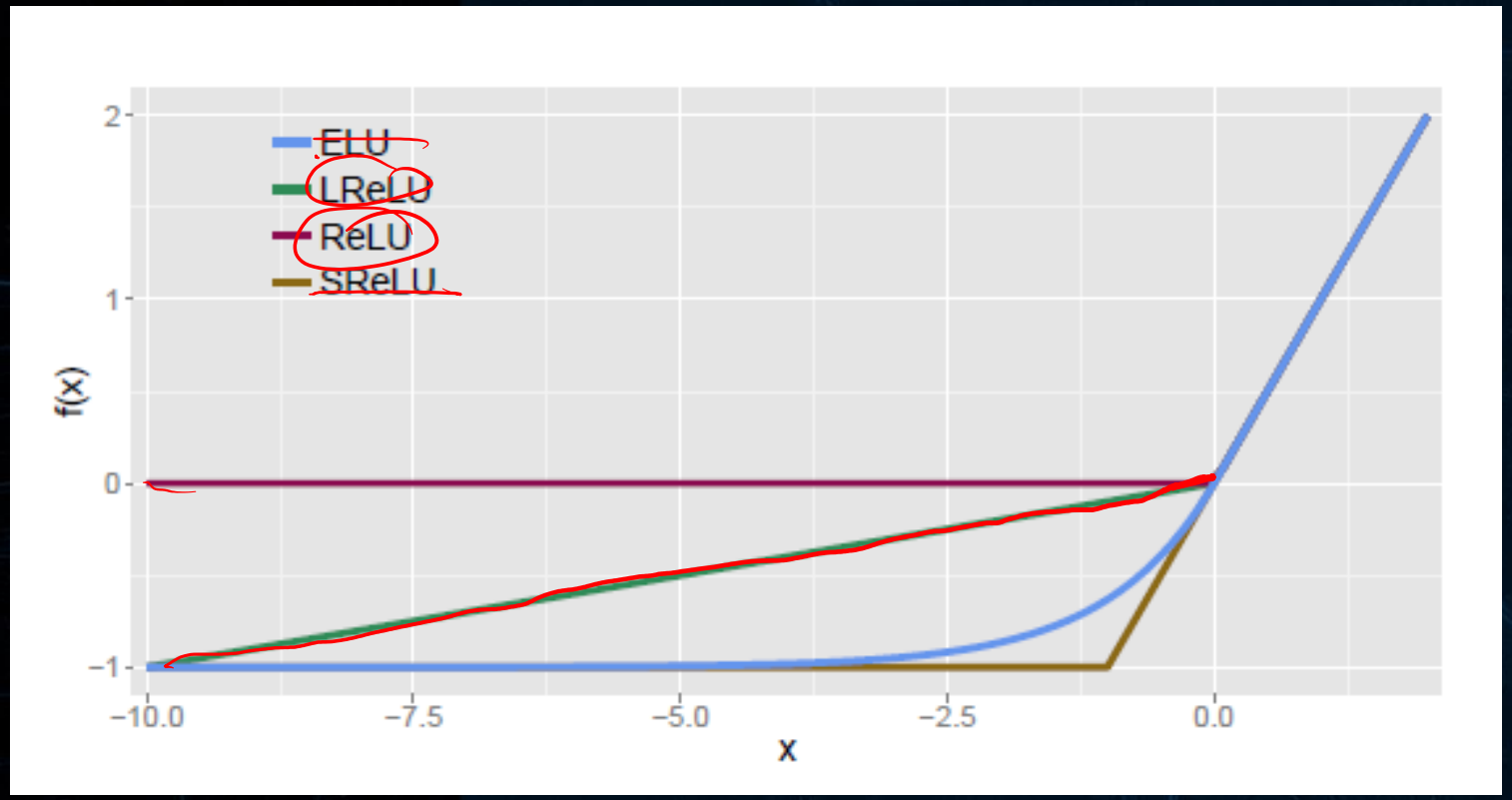
输出				
填充	步幅	宽	高	深度
Same	1	32	32	64
Valid	1	28	28	64
Valid	2	14	14	64

卷积神经网络-ReLU Layer

常用非线性激活函数

- Sigmoid
- Tanh
- ReLU
- Leaky ReLU
- ELU

卷积神经网络-ReLU Layer



The *exponential linear unit* (ELU) with $0 < \alpha$ is

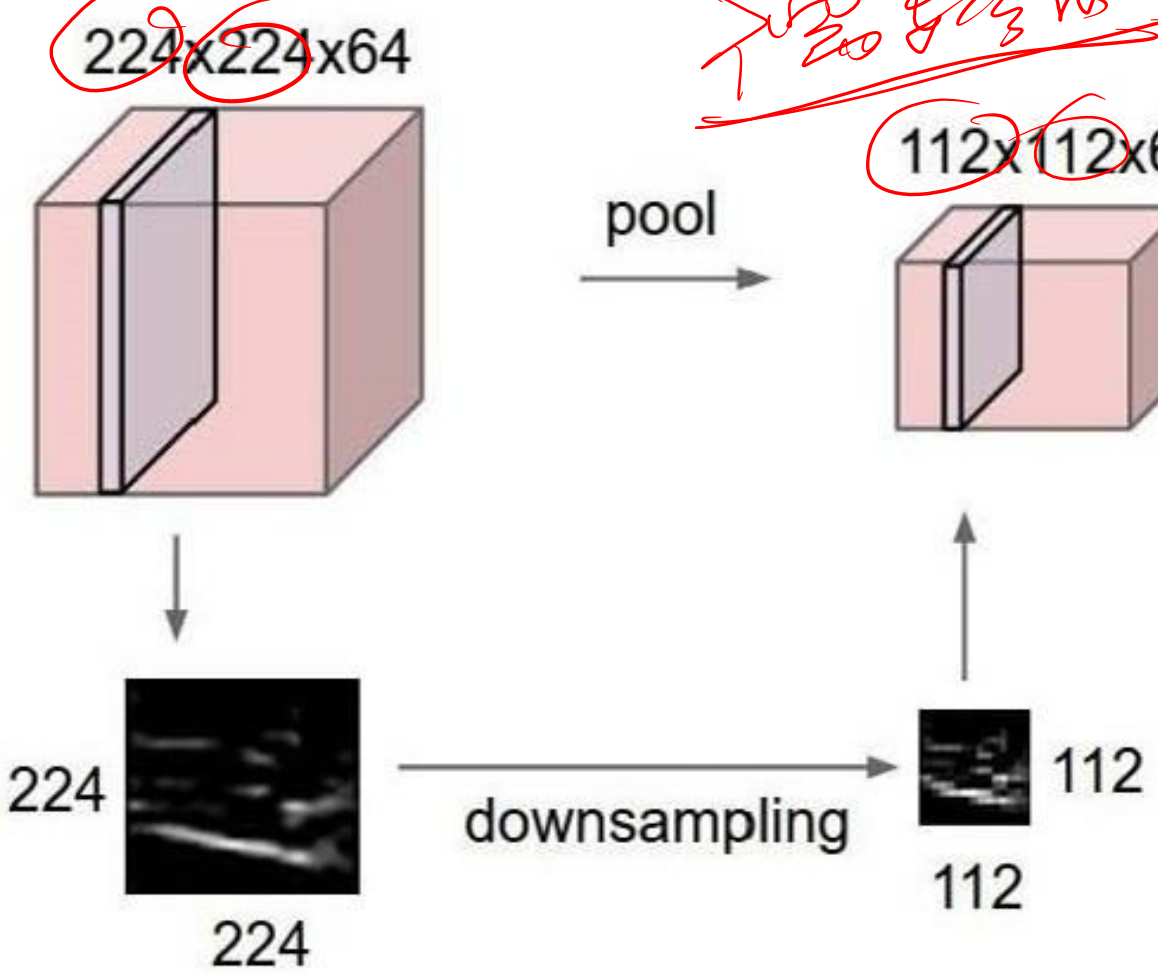
$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha\left(\exp(x) - 1\right) & \text{if } x \leq 0 \end{cases}, \quad f'(x) = \begin{cases} 1 & \text{if } x > 0 \\ f(x) + \alpha & \text{if } x \leq 0 \end{cases}.$$

卷积神经网络-ReLU Layer

激励层建议

- CNN尽量不要使用sigmoid，如果要使用，建议只在全连接层使用
- 首先使用RELU，因为迭代速度快，但是有可能效果不佳
- 如果使用RELU失效的情况下，考虑使用Leaky ReLu或者Maxout，此时一般情况都可以解决啦
- tanh激活函数在某些情况下有比较好的效果，但是应用场景比较少

池化层（pooling layer）



池化层（pooling layer)

- 对图像进行下采样并不会改变图像中特征的相对位置和目标的属性

bird

下采样（池化）

bird

我们能够通过下采样降低图像的分辨率

这样网络需要处理的参数就更少。

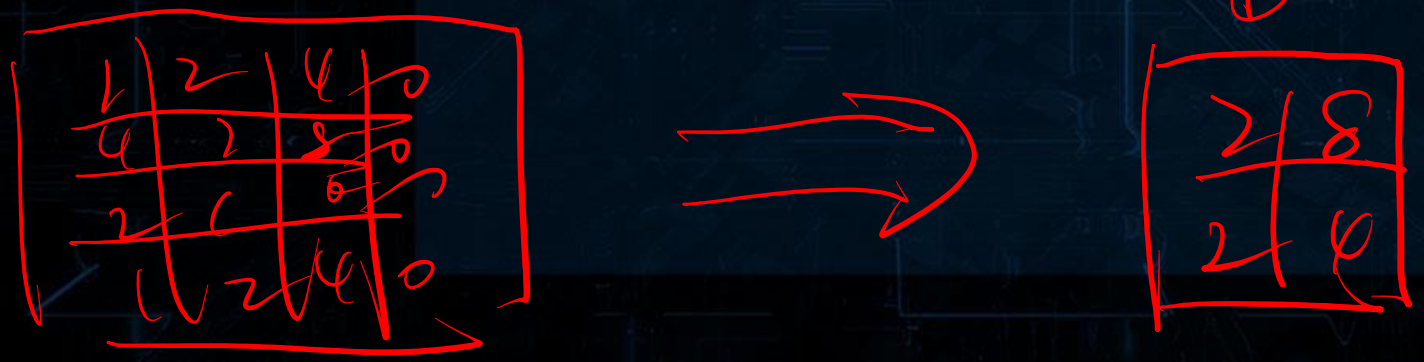
最大池化层（max pooling layer）

1	1	2	4
2	4	7	8
3	2	1	0
1	1	2	4

2*2区域下采样

步长为2

4	8
3	4



池化层（pooling layer)

1	1	2	4
2	4	7	8
3	2	1	0
1	1	2	4

2*2区域下采样

步长为1

4	7	8
4	7	7
3	2	4

平均池化层（average pooling layer）

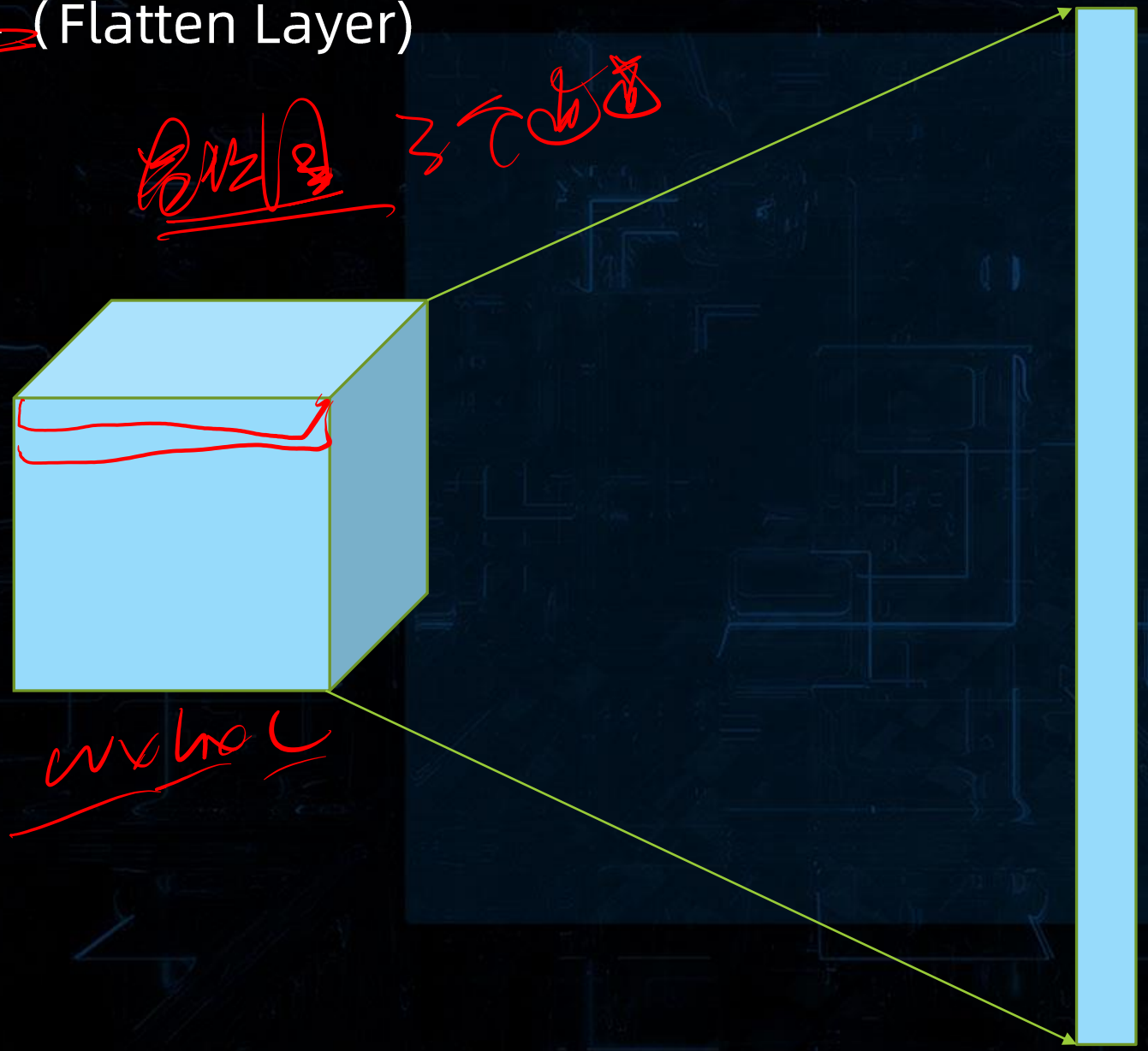
1	1	2	3
2	4	7	8
3	2	1	3
1	2	2	4

2*2区域下采样

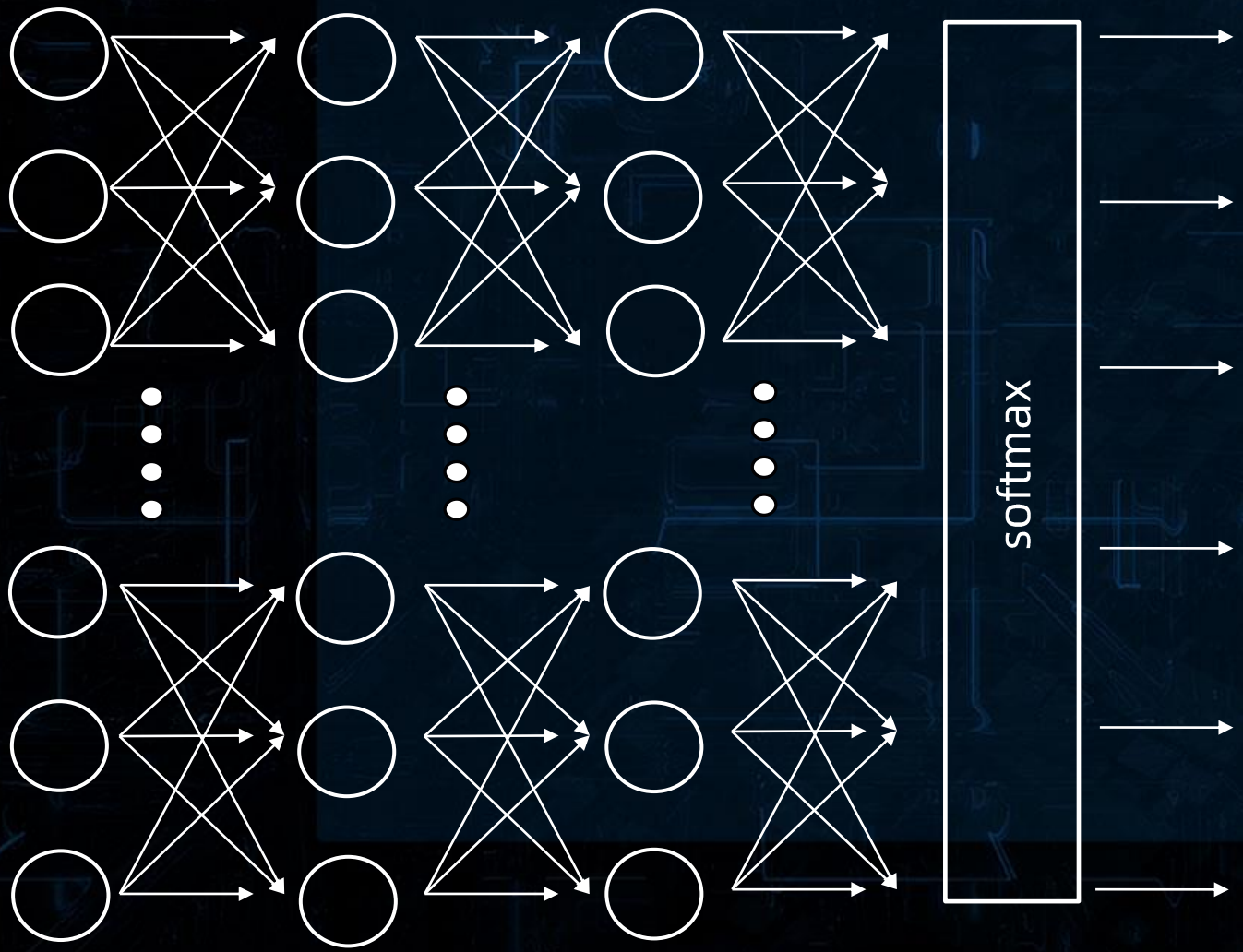
步长为2

2	5
2	2.5

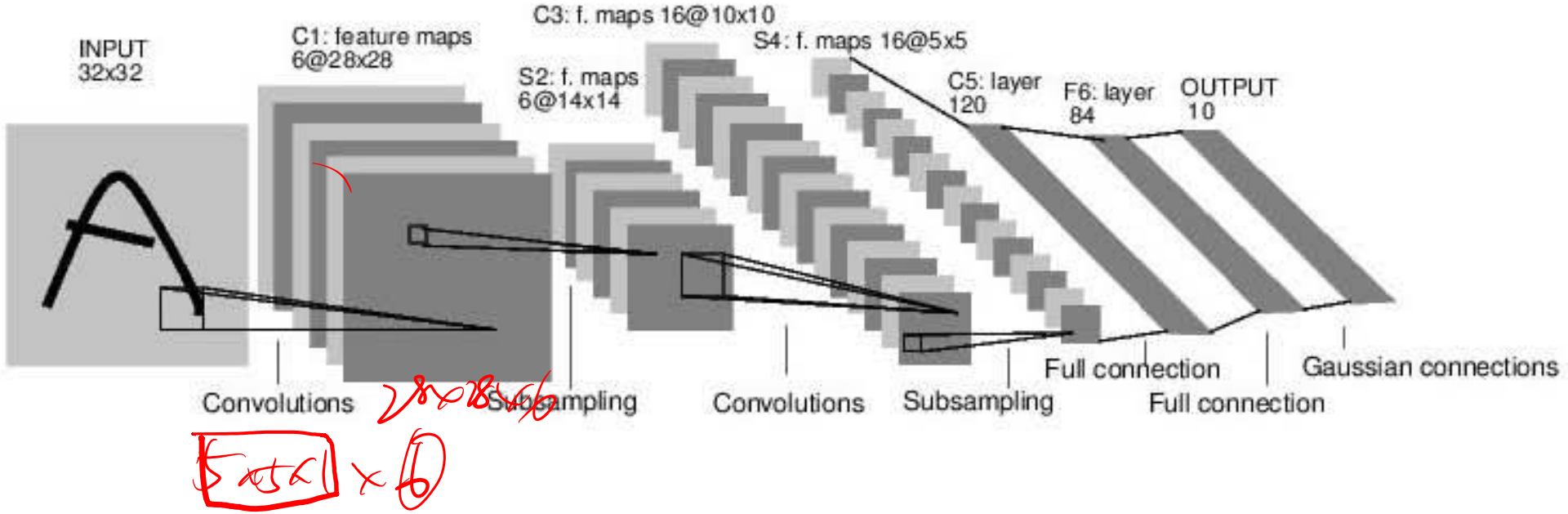
拉平层（Flatten Layer）



全连接层（FC layer）



Convolutional Neural Networks



[LeNet-5, LeCun 1980]

根据网络图计算参数

	特征图	特征图尺寸	参数
输入层	32*32*1	1024	
卷积层	28*28*6	4704	
池化层	14*14*6	1176	
卷积层	10*10*16	1600	
池化层	5*5*16	400	
全连接层	120*1	120	
全连接层	60*1	60	
Softmax层	10*1	10	