

nlp4kor

<https://github.com/bage79/nlp4kor>

<https://facebook.com/nlp4kor>

왕초보를 위한 CNN



Luis Serrano

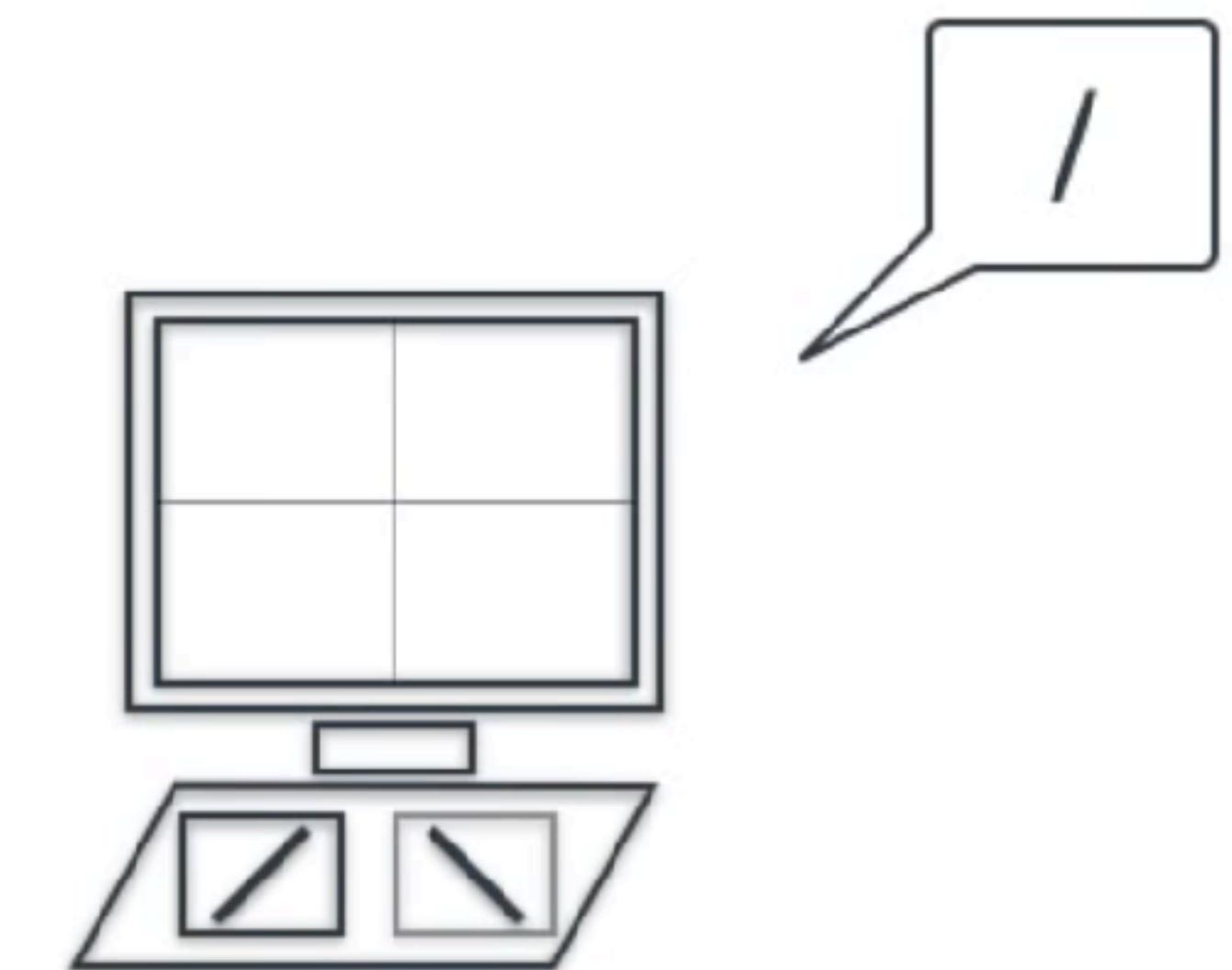
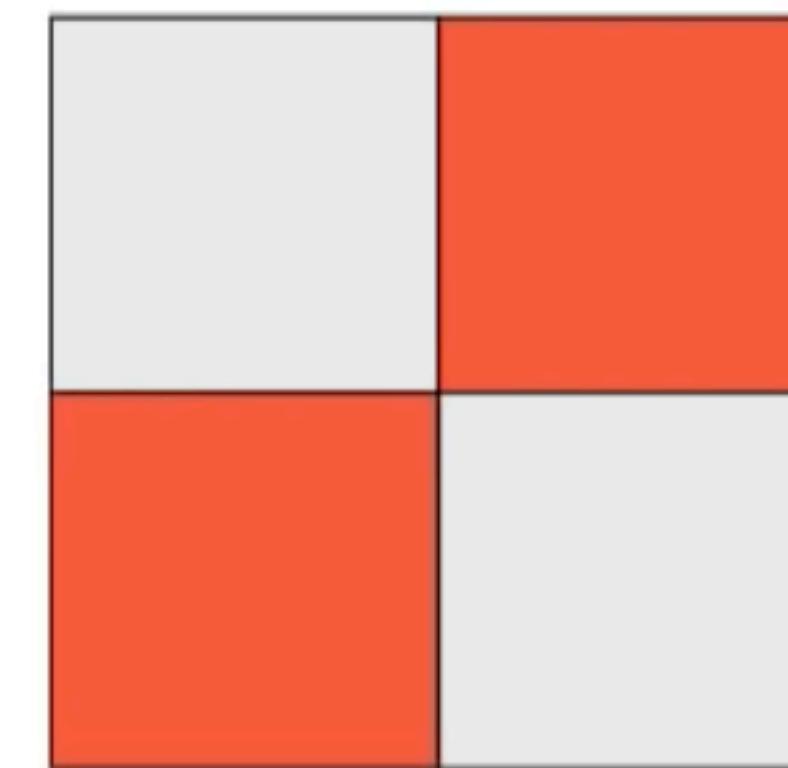
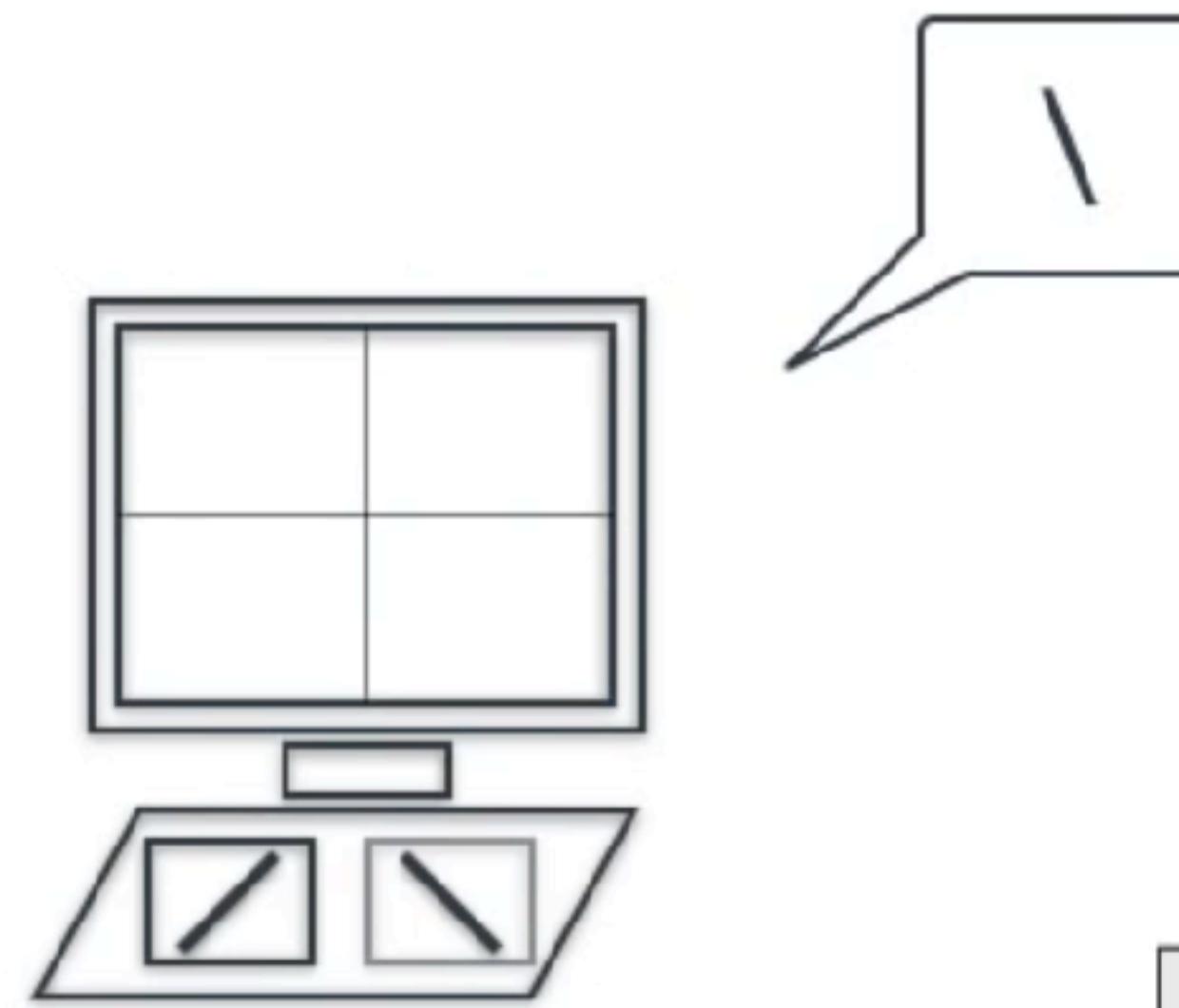
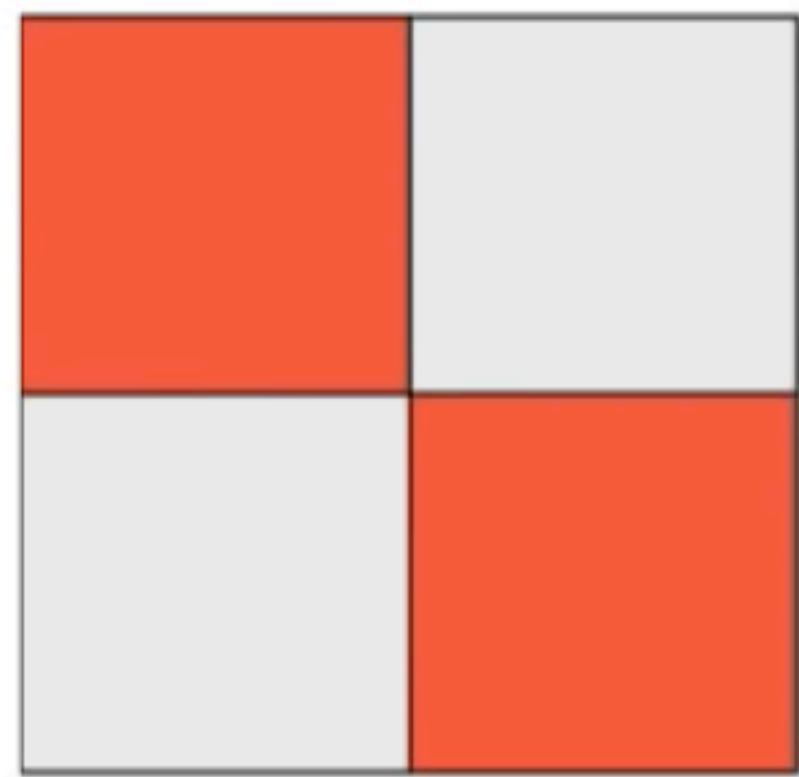
<https://youtu.be/2-OI7ZB0MmU>

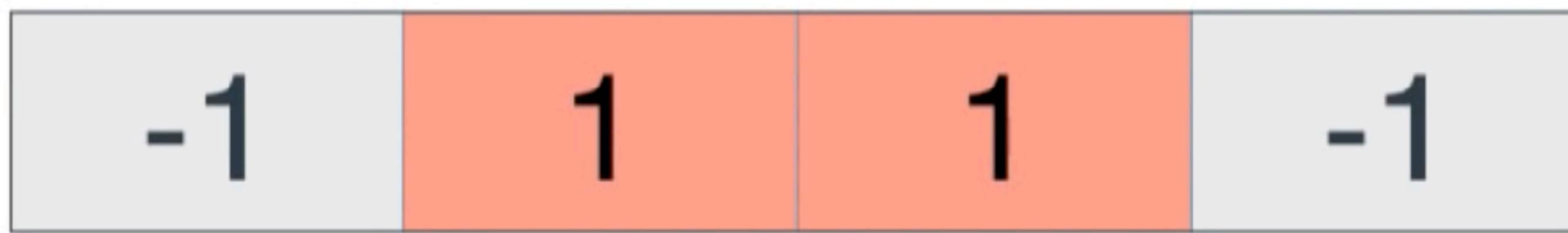
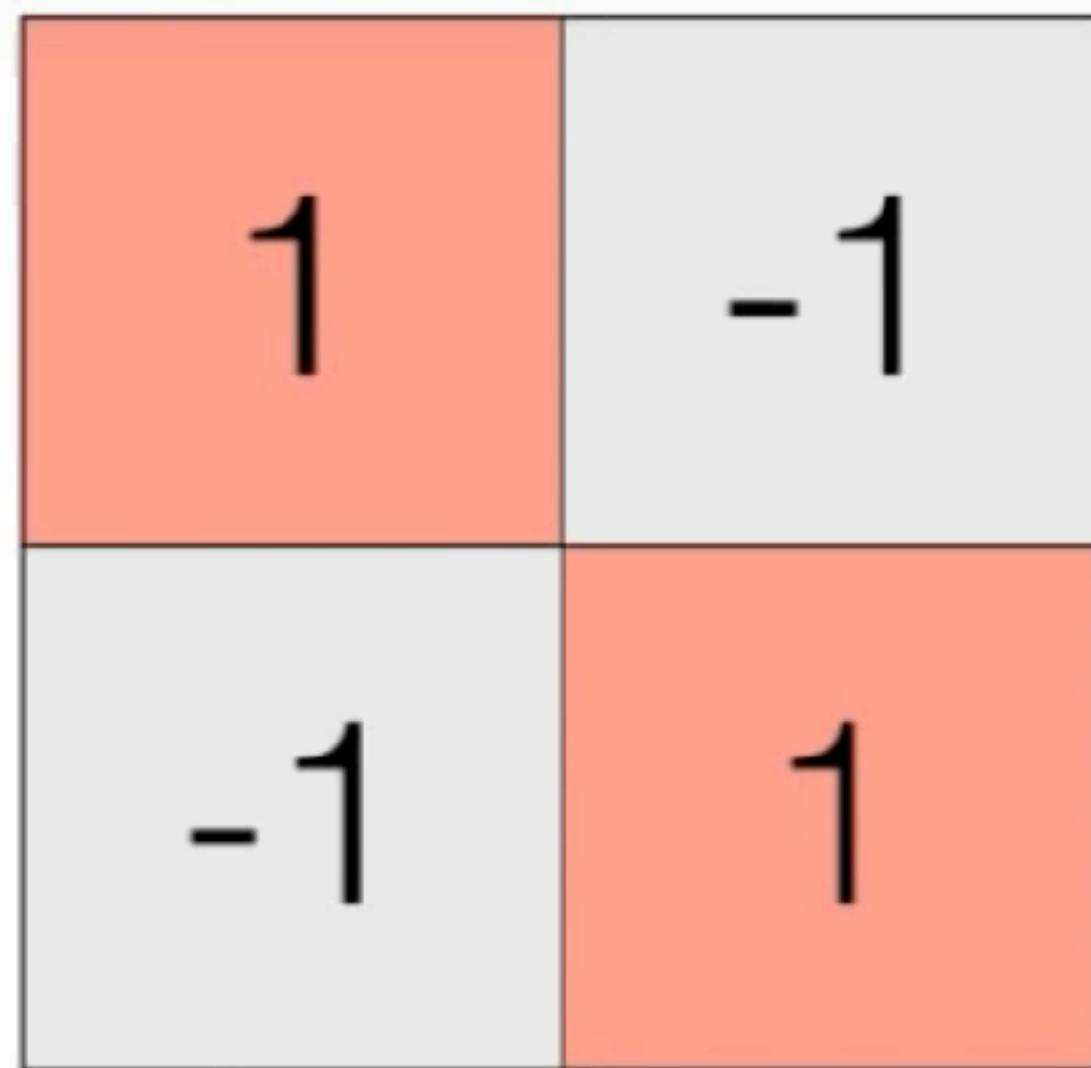


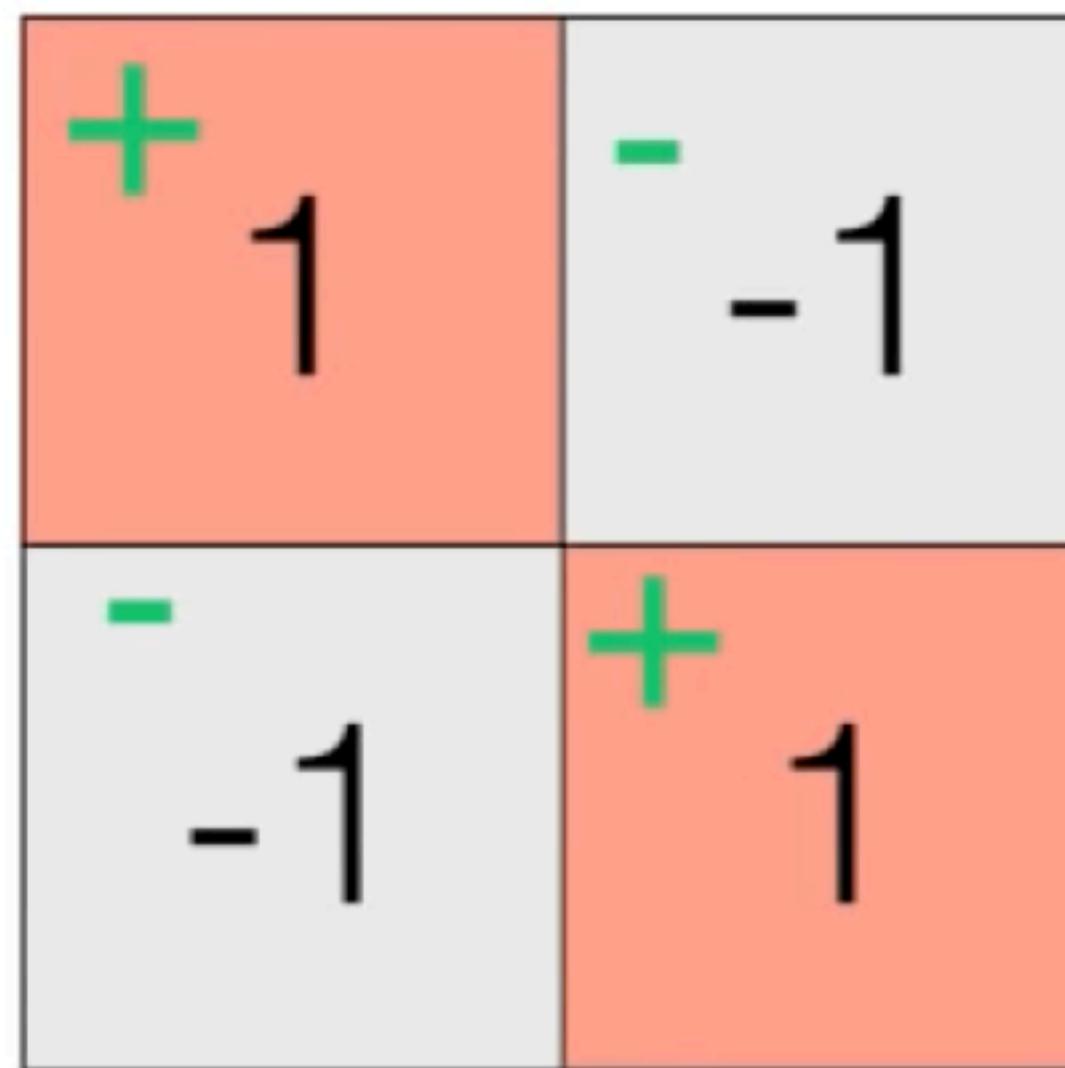
Fei-Fei Li

http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture5.pdf

Image recognition software







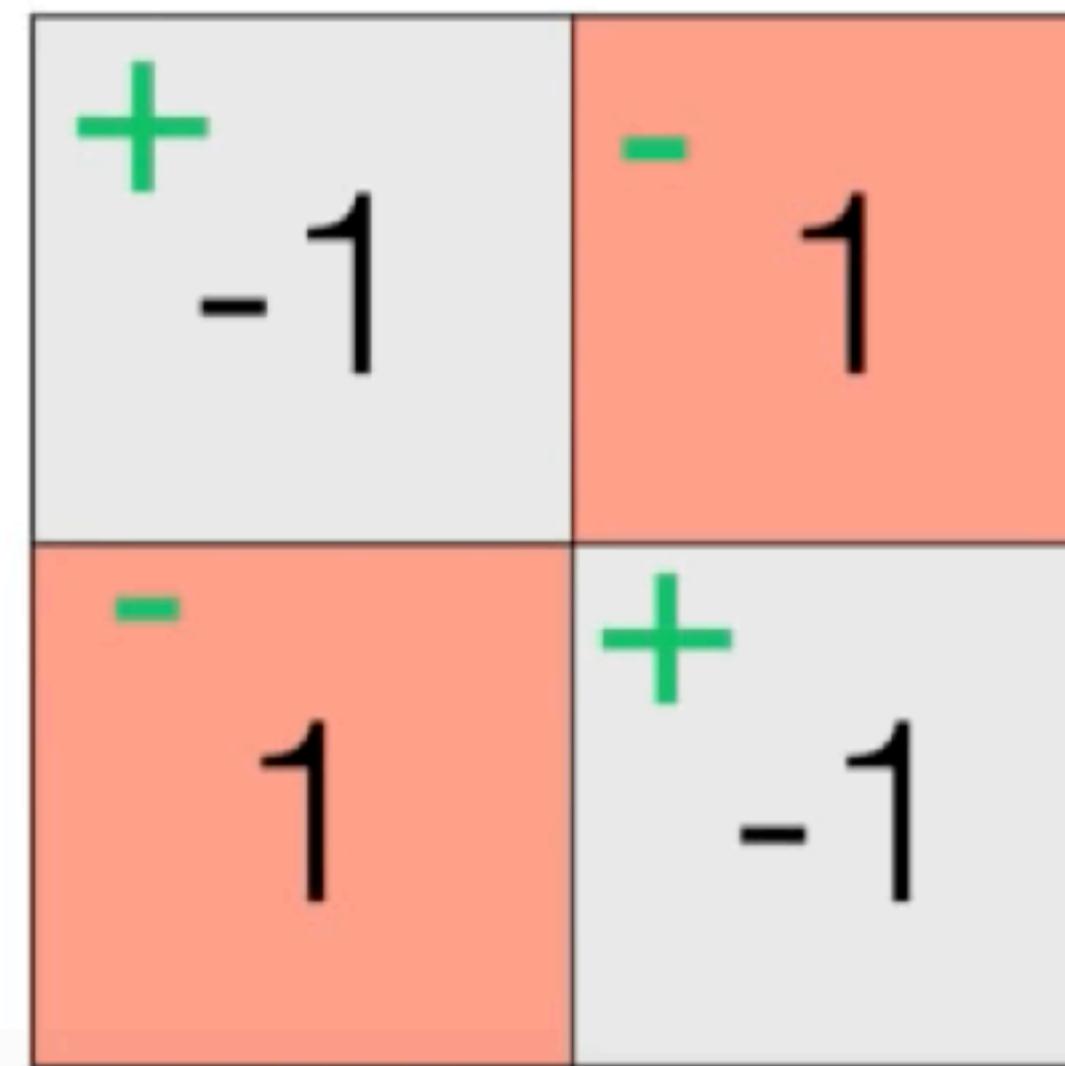
1	-1
-1	1

\\



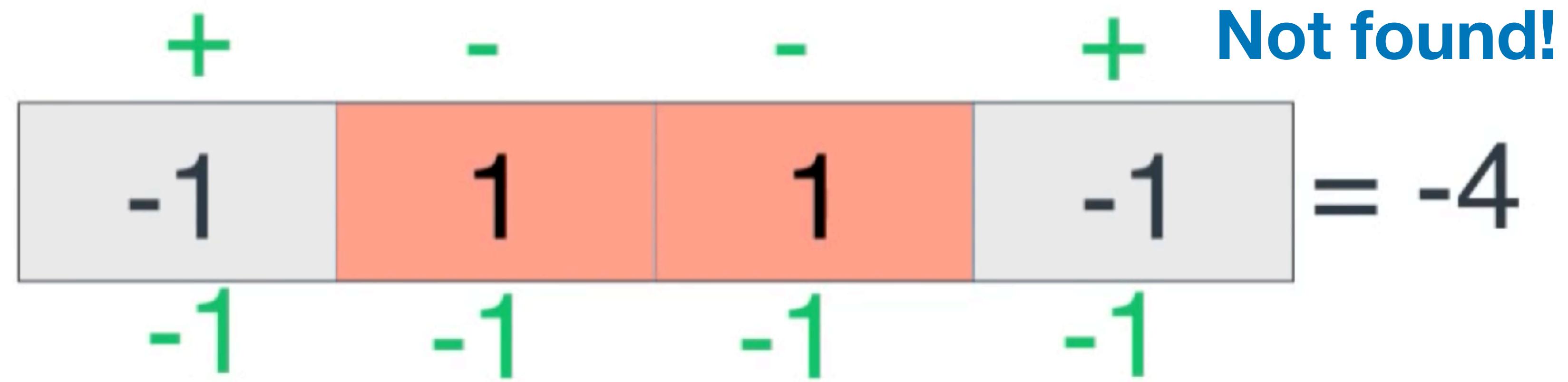
1	-1	-1	1
+1	+1	+1	+1

Found!

$$= 4$$


-1	1
1	-1

/

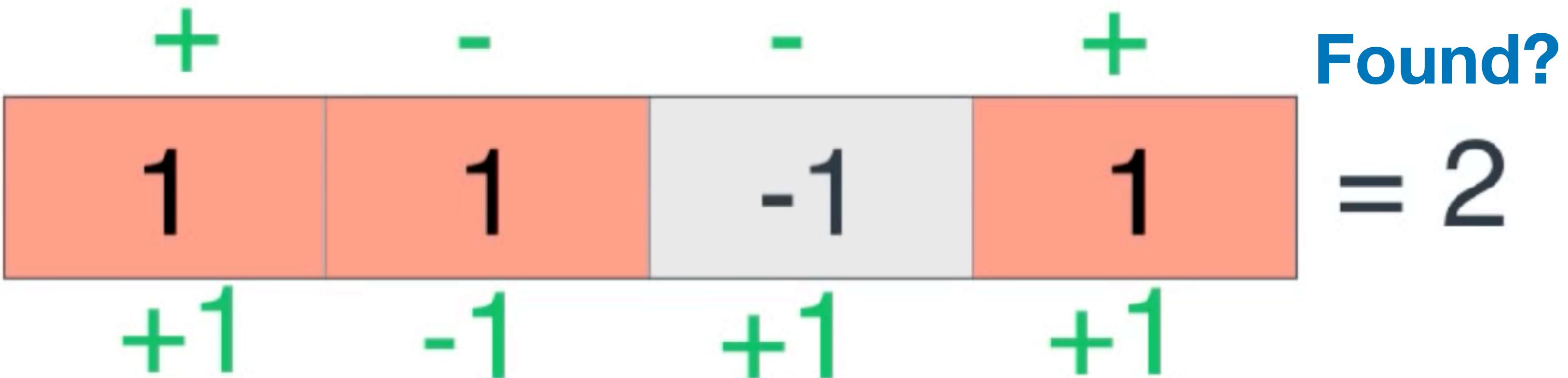


-1	1	1	-1
-1	-1	-1	-1

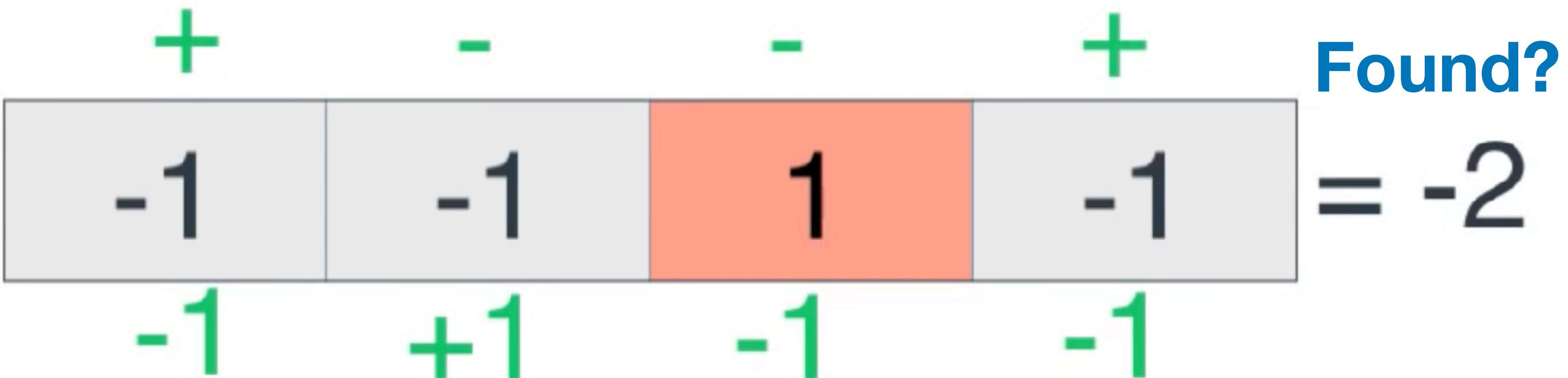
Not found!

$$= -4$$

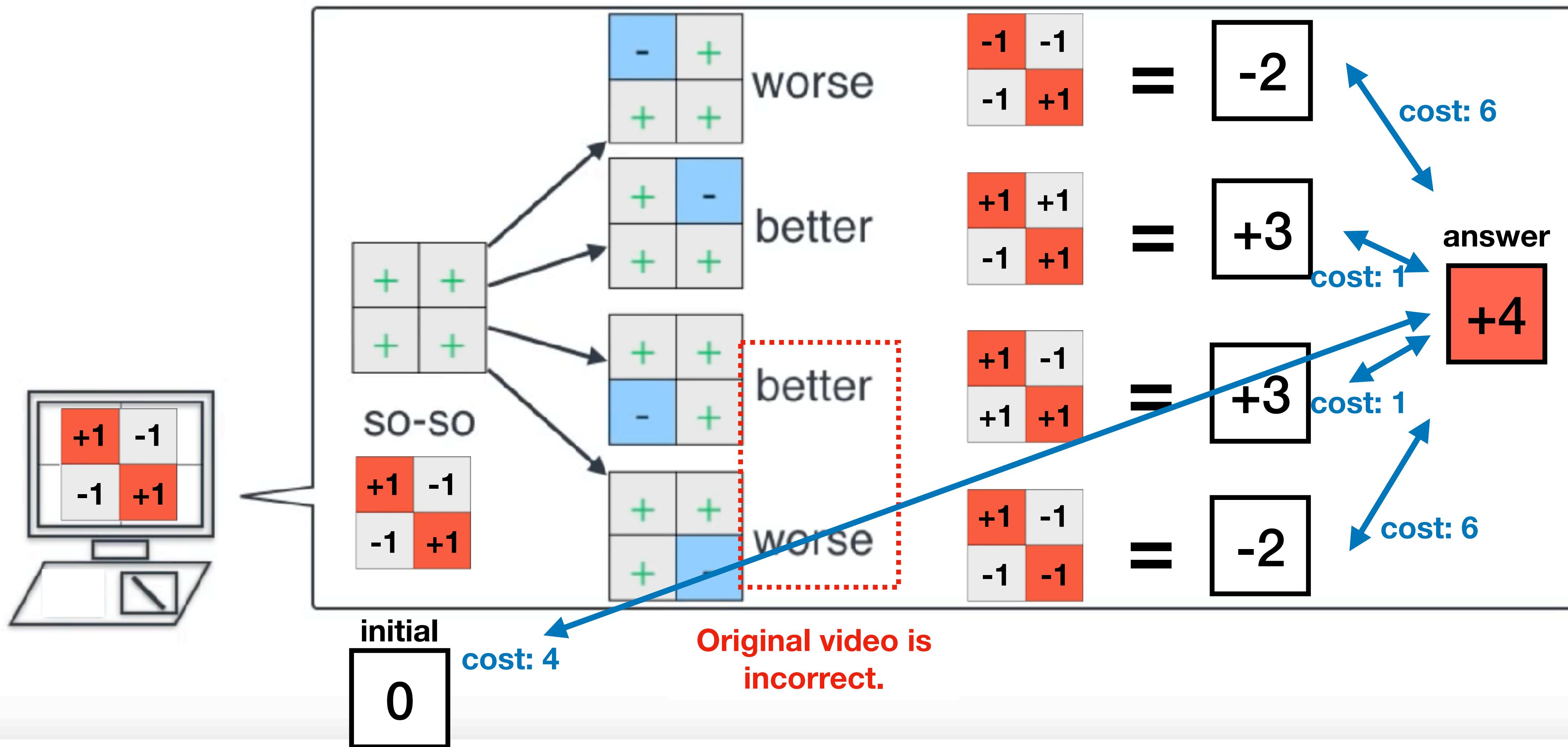
+	1
-	1
-	-1
-1	1



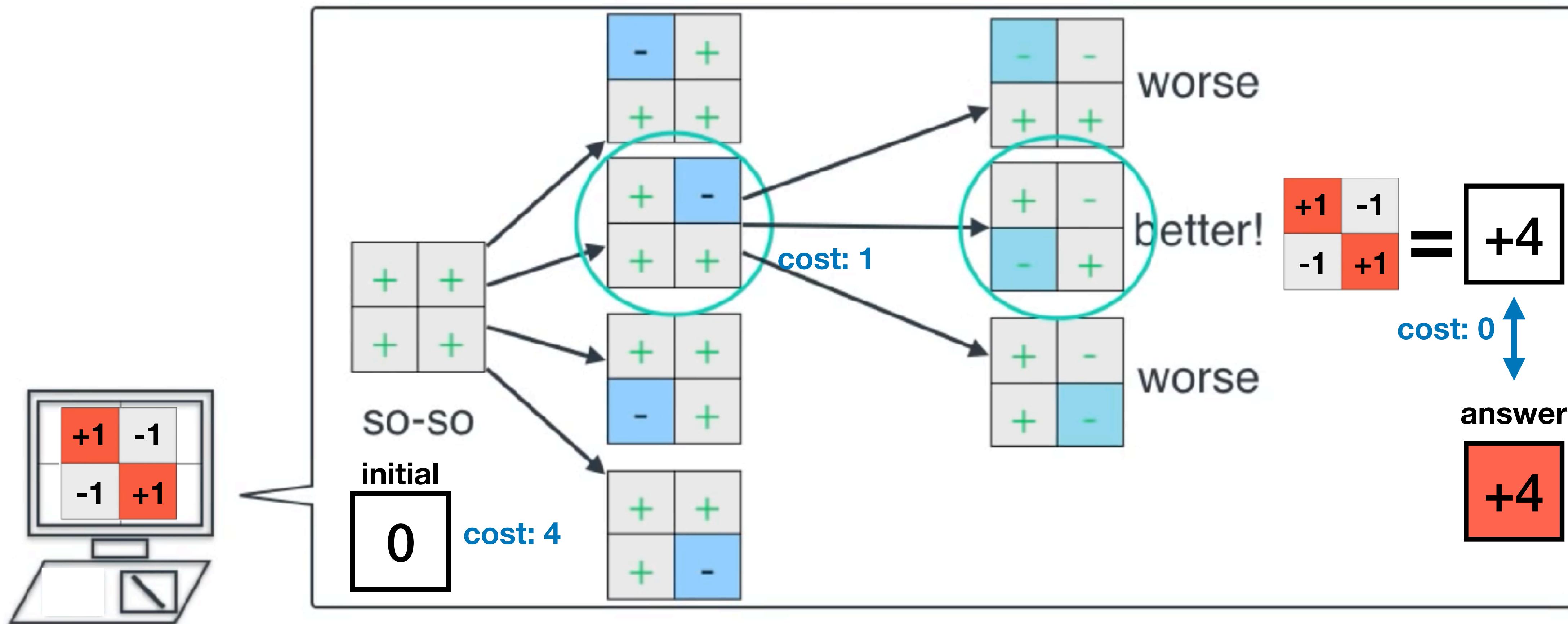
+	-1
-	-1
-	1
1	-1



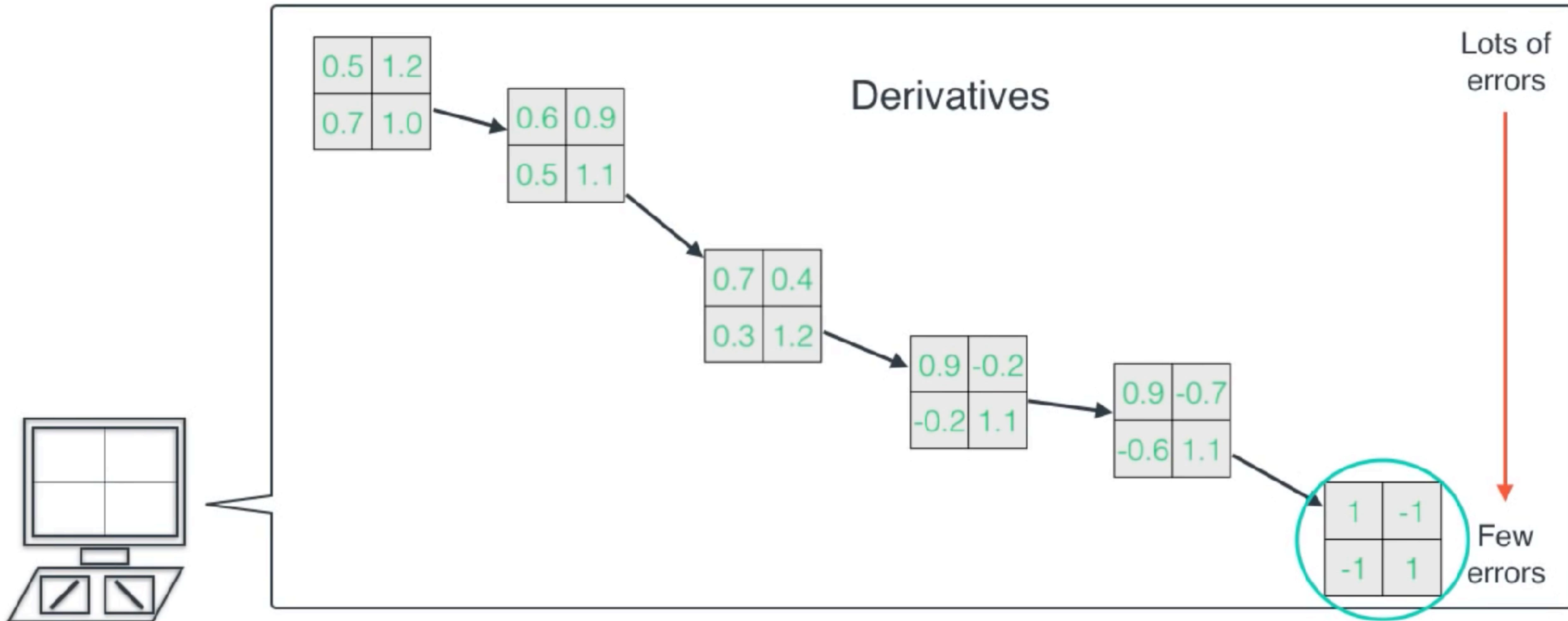
Artificial Intelligence



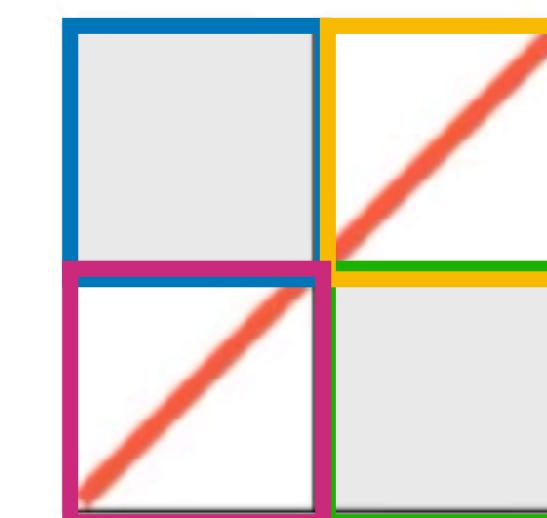
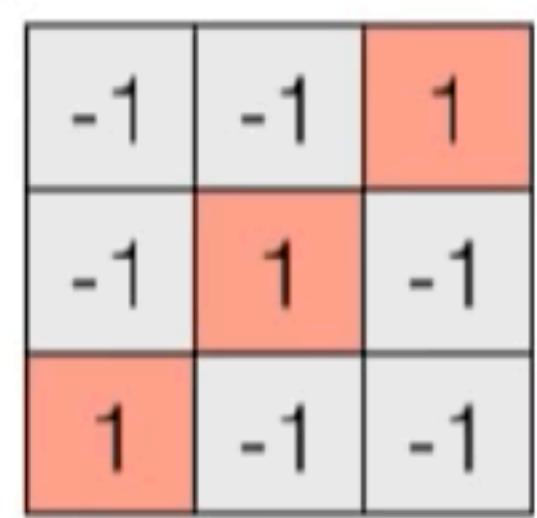
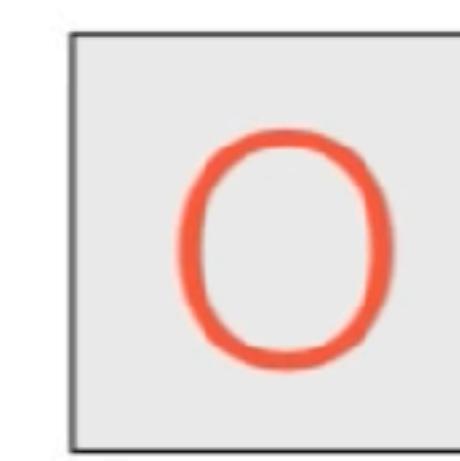
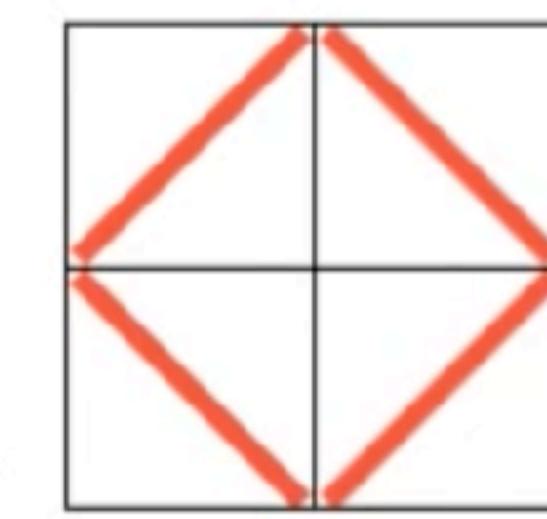
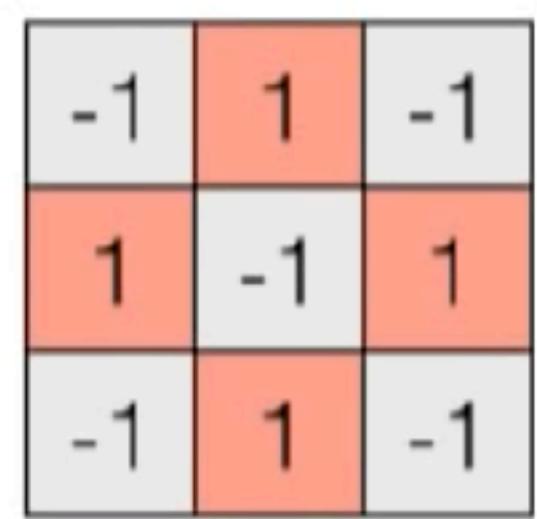
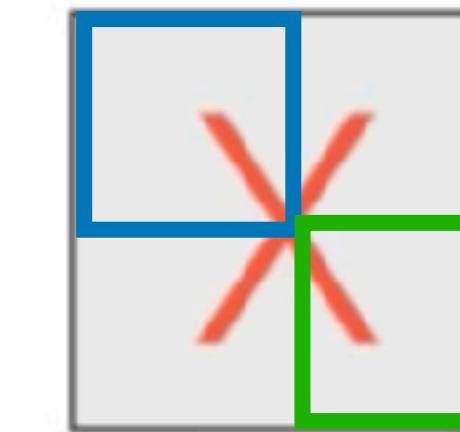
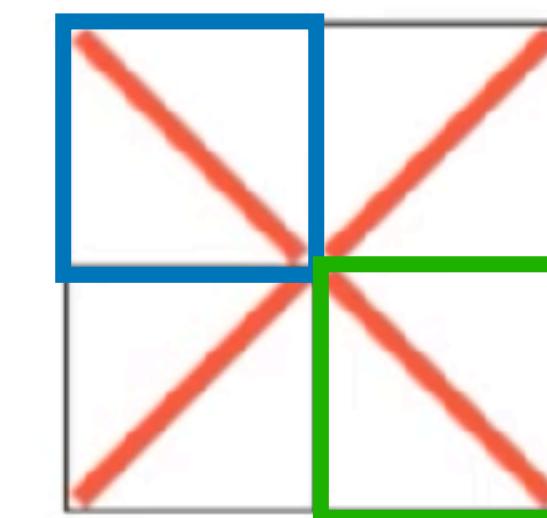
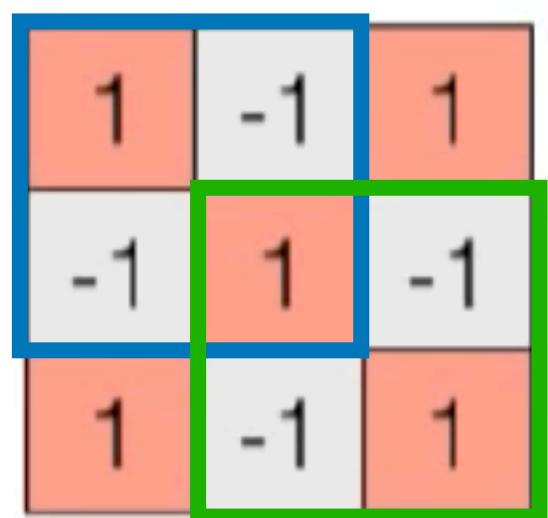
Artificial Intelligence



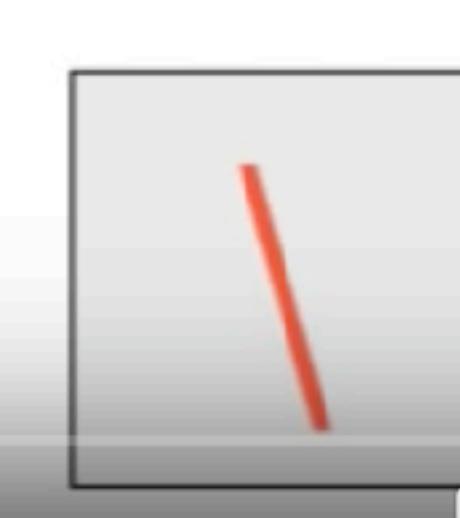
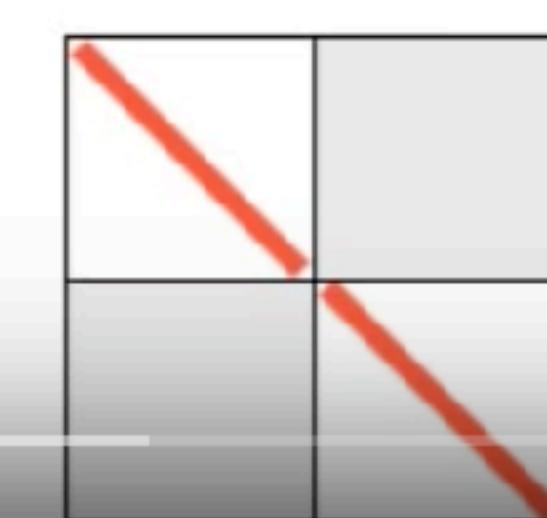
Gradient Descent



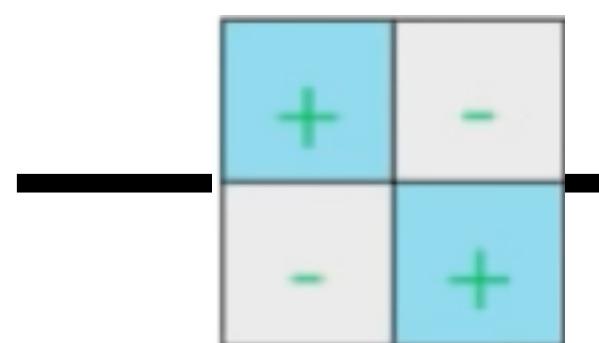
Convolutional Neural Network



Fully Connected
Layer

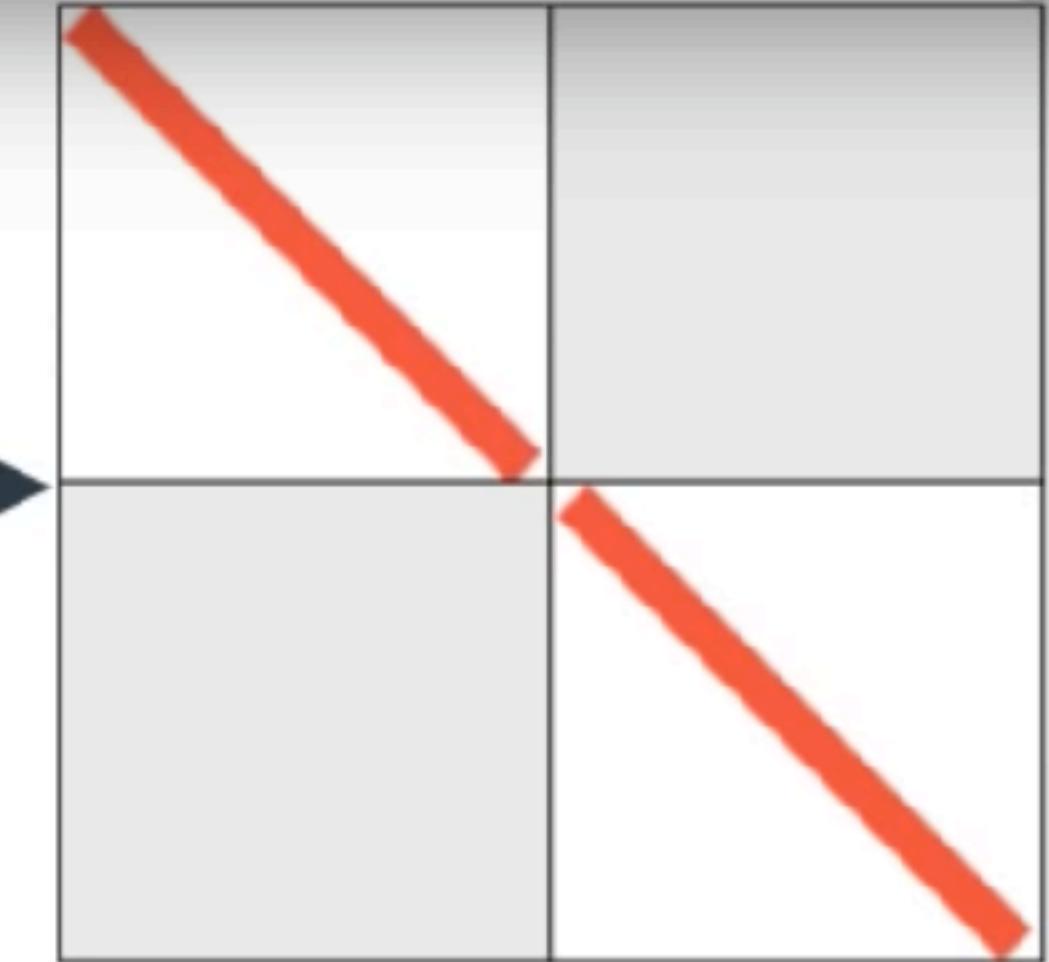


1	-1	-1
-1	1	-1
-1	-1	1



4	-2
-2	4

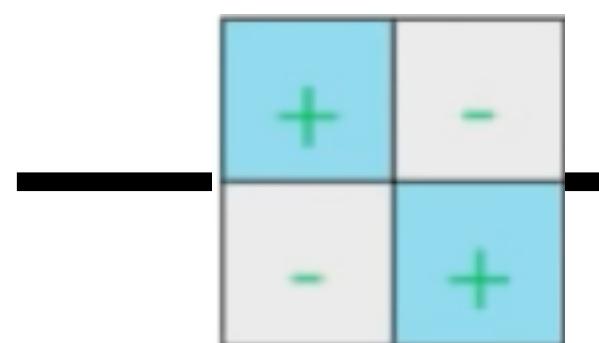
found!
~~4~~



2x2 pooling

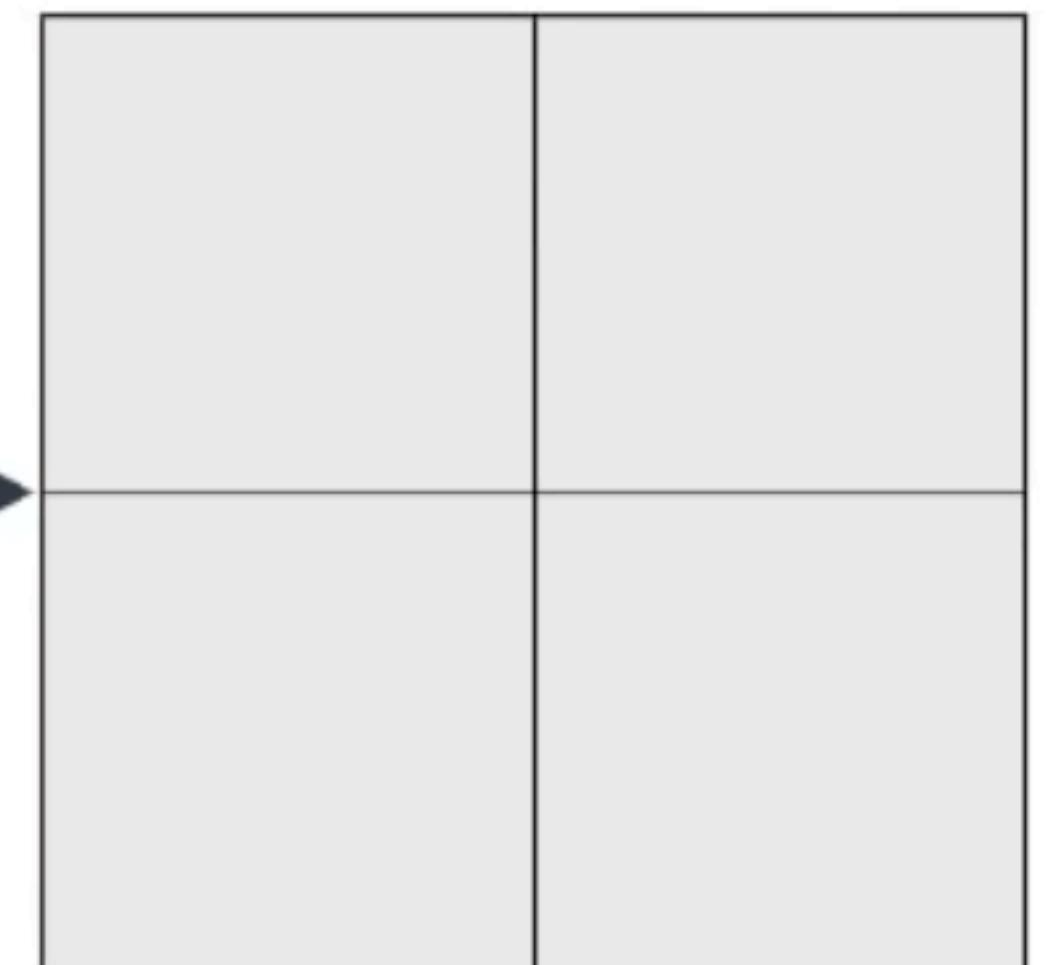
1x1 pooling

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



2	-4
-4	2

2
not found!



Convolution Layer

**Pooling Layer
(subsampling)**



1

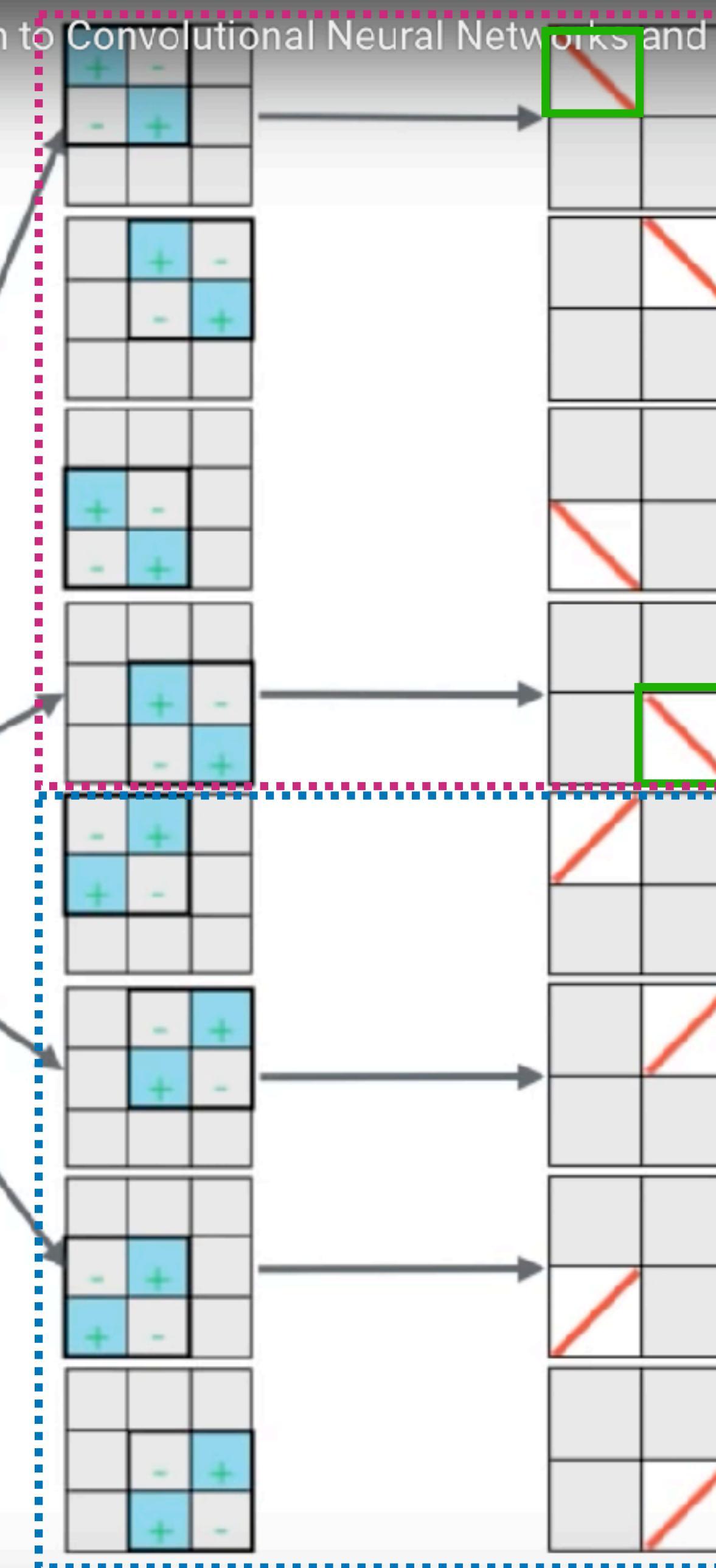
8

-8

4

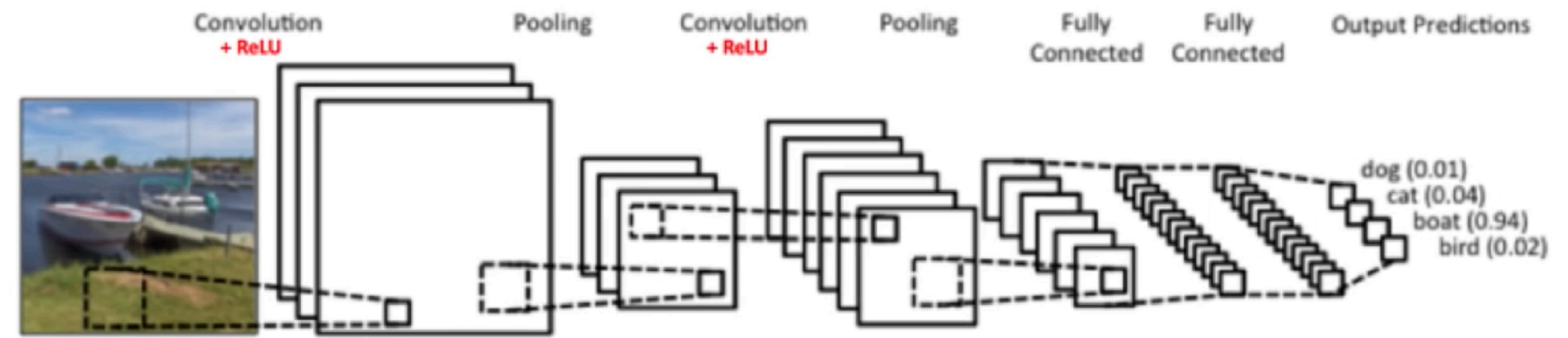
4

$$\begin{bmatrix} 1 & -1 & 1 \\ -1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix}$$

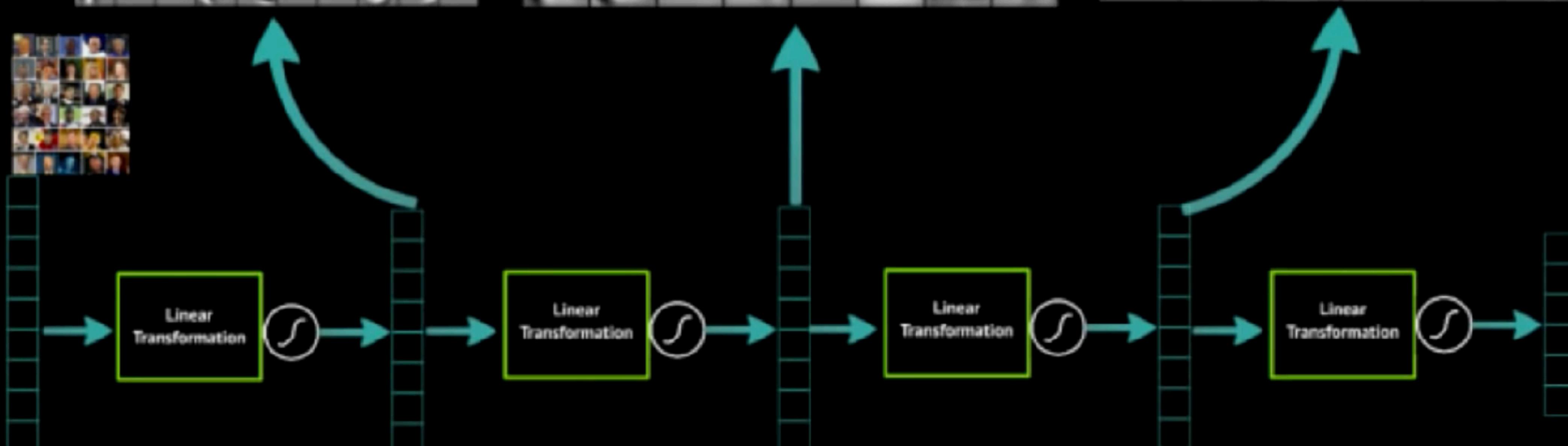
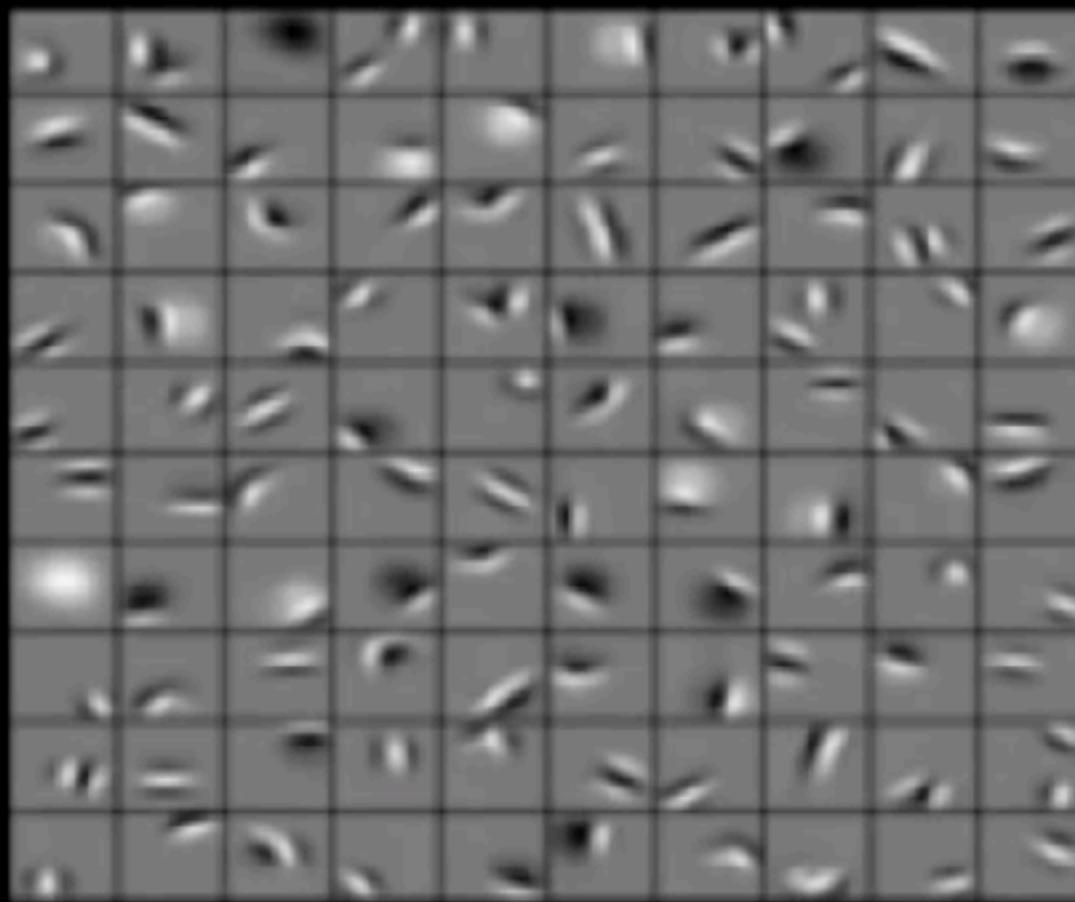


Pooling Layer

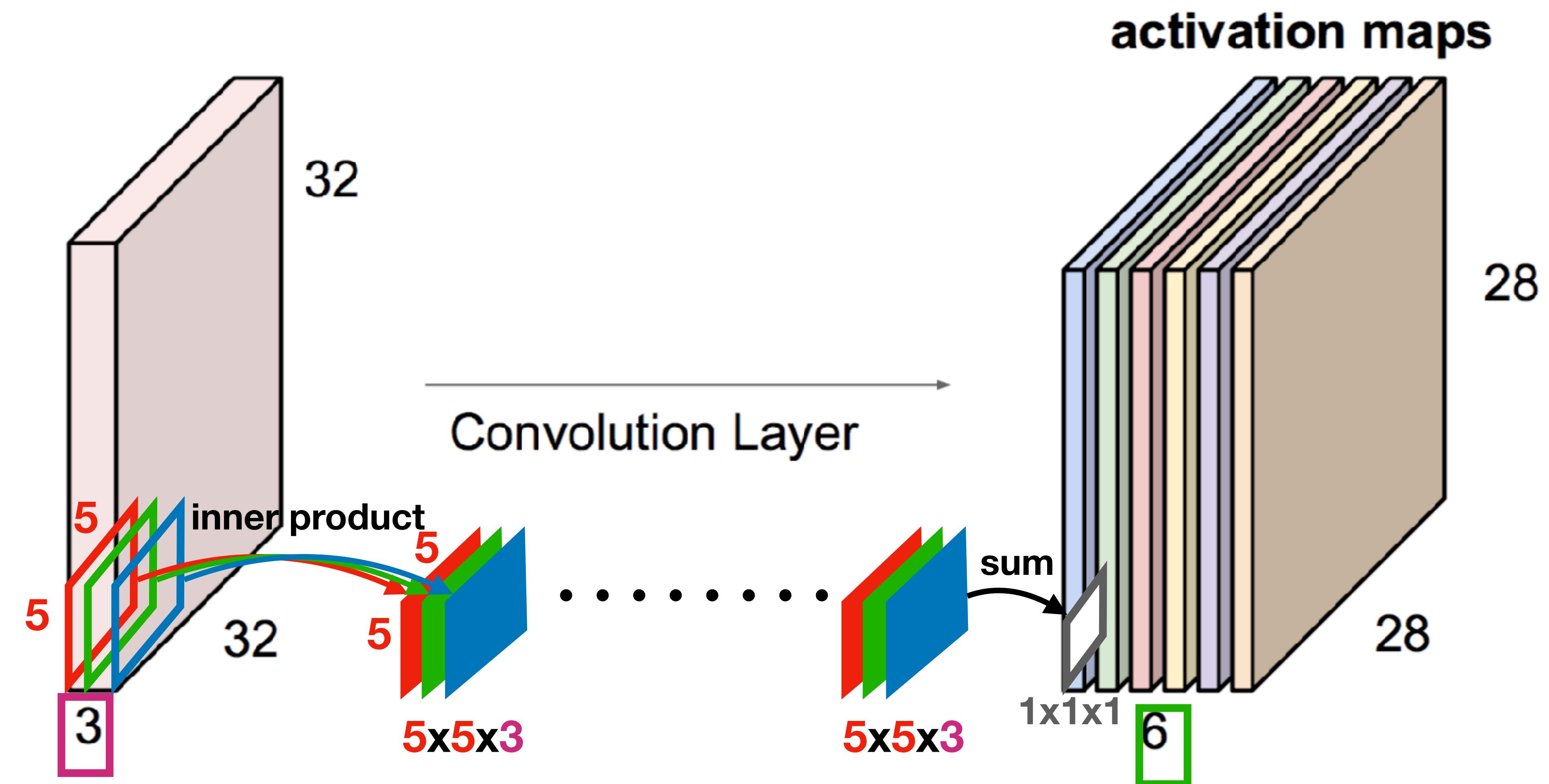
Fully Connected Layer



Deep Learning learns layers of features



For example, if we had **6** **5x5** filters, we'll get **6** separate activation maps:

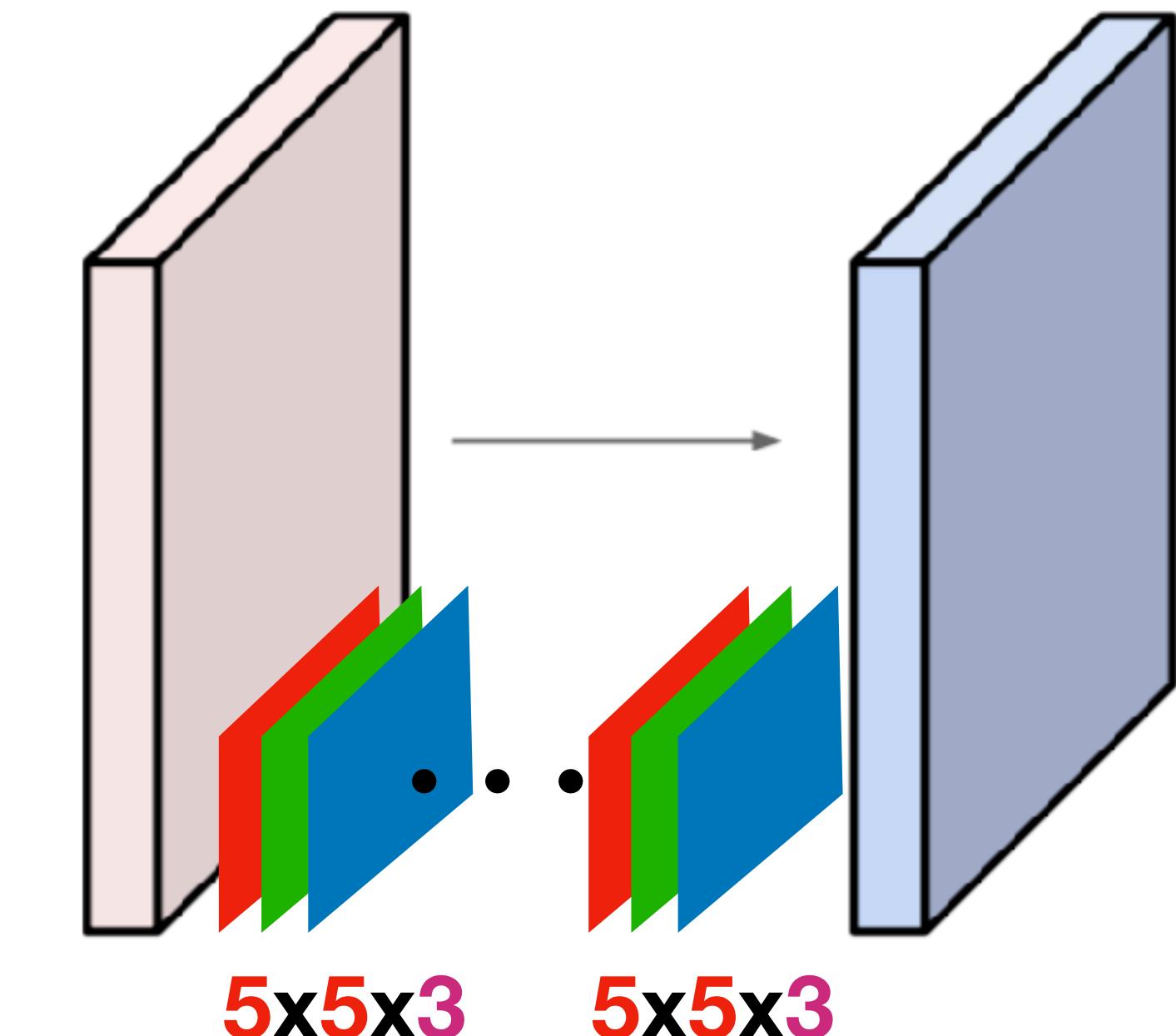


We stack these up to get a “new image” of size $28 \times 28 \times 6$!

Examples time:

Input volume: **32x32x3**

10 **5x5** filters with stride 1, pad 2



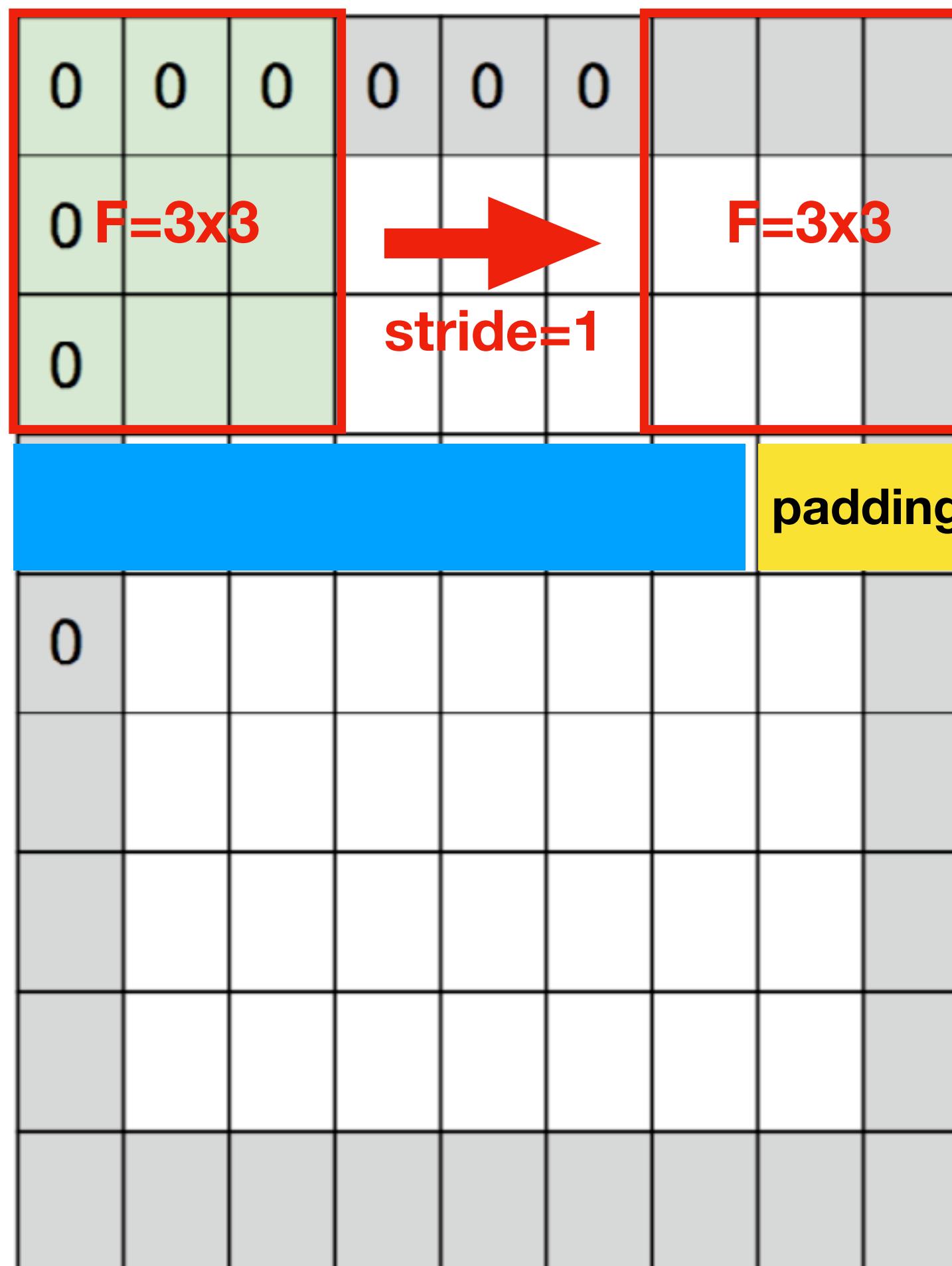
Number of parameters in this layer?

each filter has $5*5*3 + 1 = 76$ params

(+1 for bias)

$$\Rightarrow 76*10 = 760$$

In practice: Common to zero pad the border



e.g. input 7x7

3x3 filter, applied with **stride 1**

pad with 1 pixel border => what is the output?

7x7 output!

in general, common to see CONV layers with
stride 1, filters of size $F \times F$, and zero-padding with
 $(F-1)/2$. (will preserve size spatially)

e.g. $F = 3 \Rightarrow$ zero pad with 1

$F = 5 \Rightarrow$ zero pad with 2

$F = 7 \Rightarrow$ zero pad with 3

1x1 convolution

