Deep Learning and Applications (to computer vision)

MLDM

2018-2019

Damien Muselet

Damien.muselet@univ-st-etienne.fr

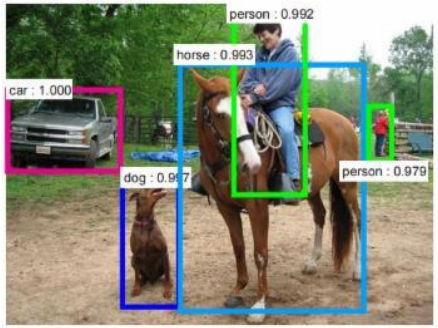
About this course…

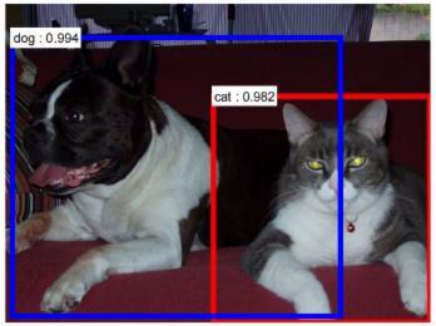
Lectures : 5 x 2 hours

Presentation of the project
Reminders about CNN
Practical session : Image classification with TensorFlow
Object detection
Practical session : Autoencoder (and RPN ?)
Domain adaptation, Generative Adversarial Networks, Recurrent Networks (LSTM), …

One project

Object detection in images





The project

Detection of the 24 patches of a color checker using deep learning.

The project

Detection of the 24 patches of a color checker using deep learning.

The project

Detection of the 24 patches of a color checker using deep learning.






High variability :

- Scale
- Light
- Sensors
- Small rotations

The project

The data :

- Available on this webpage :
  http://cvil.eecs.yorku.ca/projects/public_html/illuminant/illuminant.html

- 1853 images from 9 different cameras

The annotations

CAMERA	RAW	JPEG	PNG	MASK	GROUNDTRUTH
Canon EOS-1Ds Mark III	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Canon EOS 600D	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Fujifilm X-M1	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Nikon D5200	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Olympus E-PL6	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Panasonic Lumix DMC-GX1	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Samsung NX2000	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT
Sony SLT-A57	ZIP1 ZIP2 ZIP3 ZIP4 ZIP5	ZIP1 ZIP2	ZIP1 ZIP2 ZIP3 ZIP4	ZIP	MAT

And another extra camera with smaller number of images is also provided here:

CAMERA	RAW	JPEG	PNG	MASK	GROUNDTRUTH
Nikon D40	ZIP	ZIP	ZIP	ZIP	MAT

The images you have to use

The project

The provided PNG images have to be pre-processed !

The color components are linearly related to the energy quantities captured by the sensors (as in RAW images). You have to apply an inverse gamma correction (with γ=2.2) to obtain color components expressed in a classical display color space, such as sRGB.

R, G, B in [0;4095] → R, G, B in [0;1] → $R^{(1/2.2)}$, $G^{(1/2.2)}$, $B^{(1/2.2)}$

The project

Tests and results :

Among the 9 subsets of images, use 8 to learn (Train and Validation) and 1 to test.
Return 9 results **and** the average over these 9 runs.

The quality criteria is the mean Average Precision (mAP) for detection with IoU ≥0.5.

Reminder :

		Groundtruth	
		+	-
Algorithm decision	+	TP	FP
	-	FN	TN

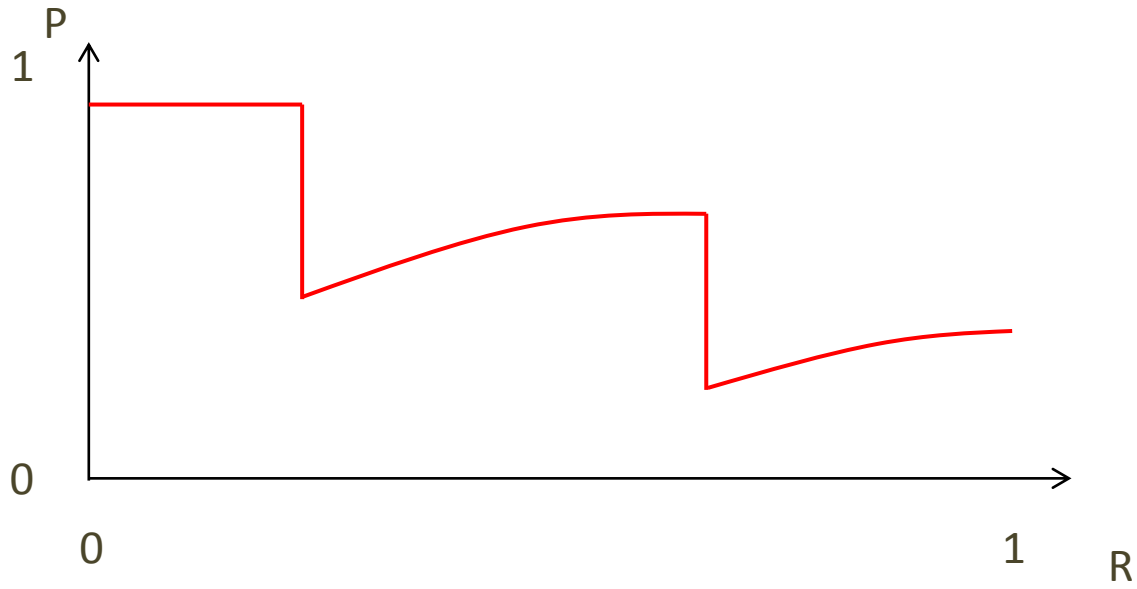
$$\text{Precision} = \frac{TP}{TP+FP}$$

$$\text{Recall} = \frac{TP}{TP+FN}$$

Output of your algorithm : a confidence to have a patch at this position.
For one threshold on this confidence → one Precision and one Recall.
For a set of thresholds → a curve of Precision vs Recall.

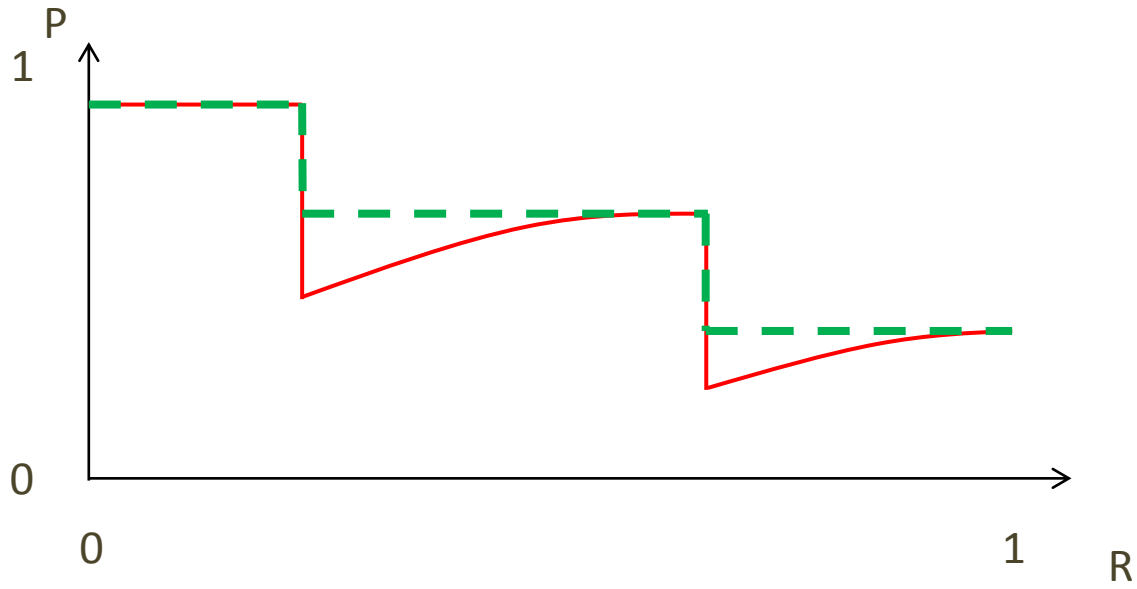
The project

Tests and results :



The project

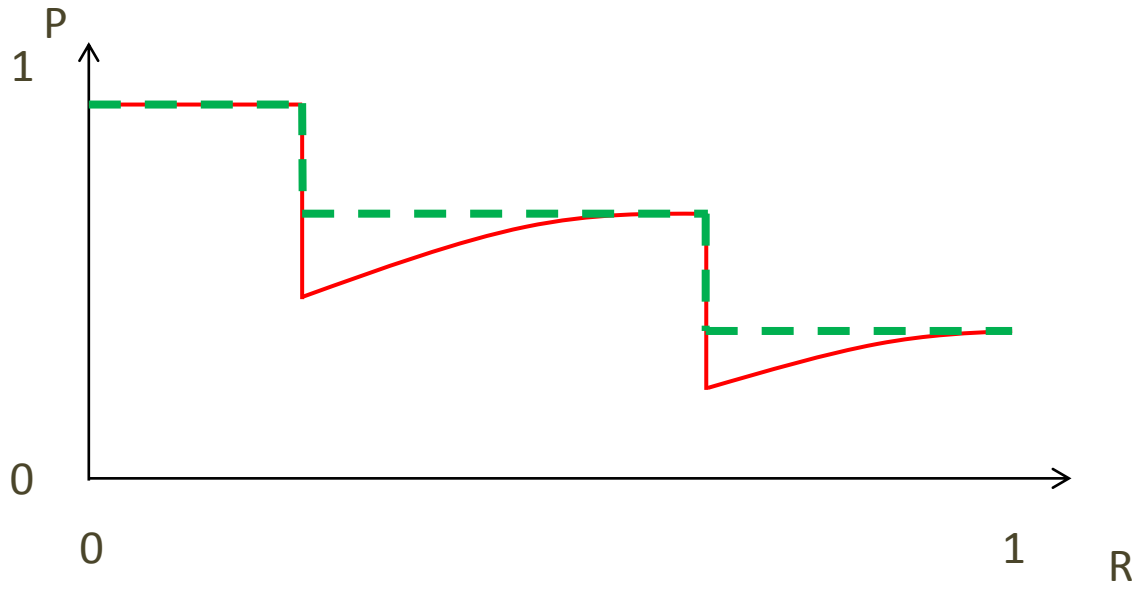
Tests and results :



$$P_m(R) = \max_{r \geq R} P(r)$$

The project

Tests and results :



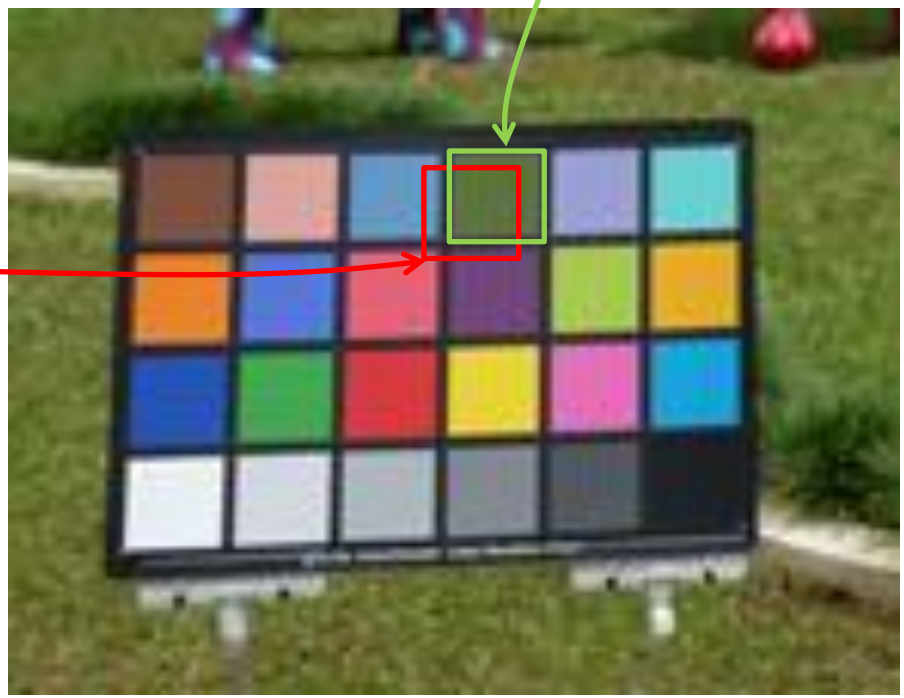
$$P_m(R) = \max_{r \geq R} P(r)$$

$$AP = \frac{1}{11} \times \sum_{R \text{ in } \{0,0.1,0.2,\ldots,1\}} P_m(R)$$

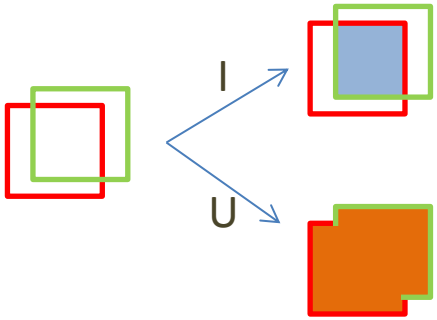
mAP is the mean AP over the classes.

The project

Tests and results :



IoU : Intersection over Union



$$IoU = \frac{\text{I}}{\text{U}}$$

For this project, we consider that a detection is correct if IoU ≥ 0.5

The project

- Groups of 2 or 3 students
- Read papers about object detection (faster R-CNN, Yolo, SSD, …)
- Use a pre-trained network and fine-tune it on your data
- Report the results (quantitative and qualitative)
- Analyze the results
- Propose improvements
- Write a nice report with CVPR template (8 pages + references)
  - ➔ Deadline : **Friday, 11th january 8am**
  - ➔ Explain what have been done, results, show images, work division between you
- **Oral presentation on 16th january** (20 minutes for presentation + demo)
- **Send the codes in an archive before the presentation.**

What you can do :

- Data augmentation (from the data you already have or synthetic)
- Acquire your own images (color checker available if needed)
- Add occlusions
- Rotated boxes
- Test on videos

We will meet **at each course** and discuss your progress, and whenever you want :
damien.muselet@univ-st-etienne.fr

About this course…

Lectures : 5 x 2 hours

Presentation of the project

Reminders about CNN

Practical session : Image classification with TensorFlow

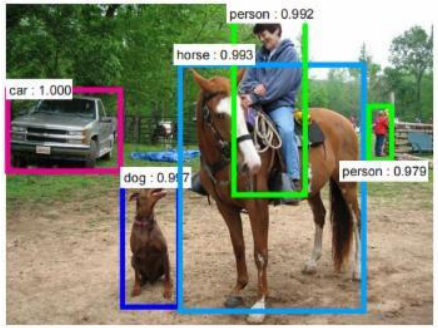
Object detection

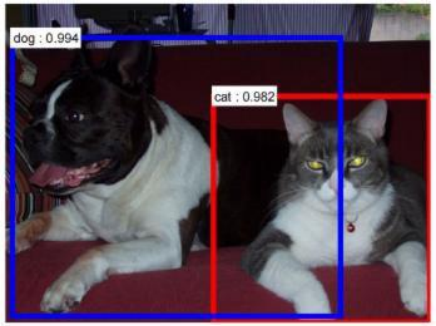
Practical session : Autoencoder (and RPN ?)

Domain adaptation, Generative Adversarial Networks, Recurrent Networks (LSTM), …

One project

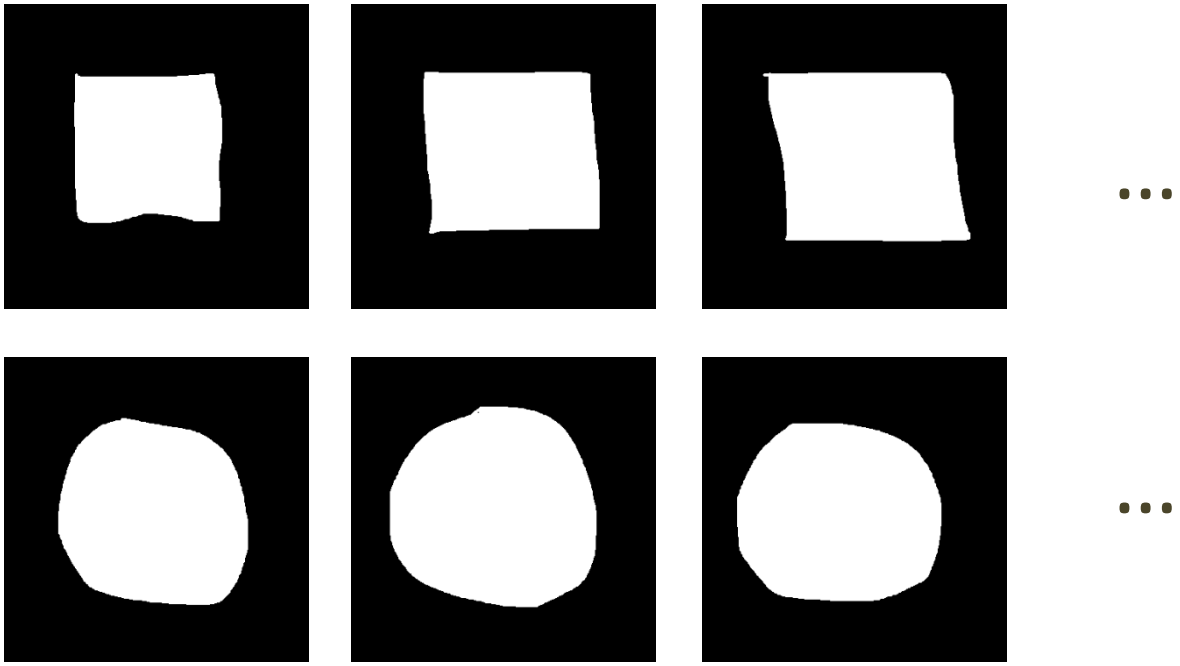
Object detection in images





Reminders about CNN

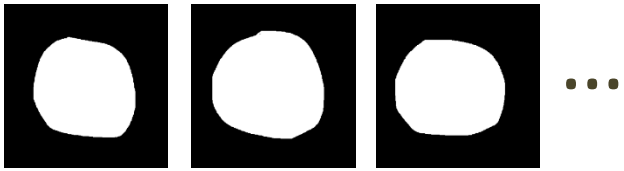
Squares versus circles…



Reminders about CNN

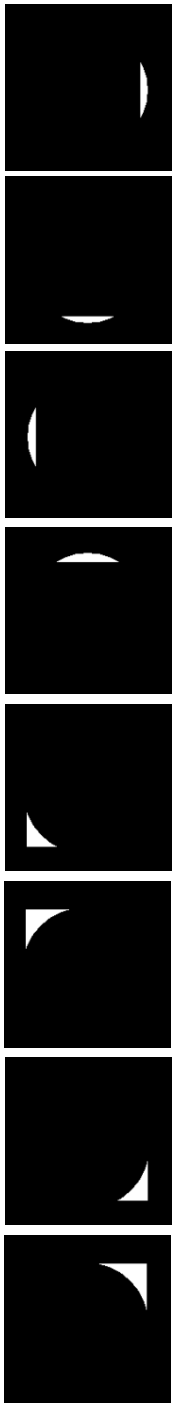
Squares versus circles…

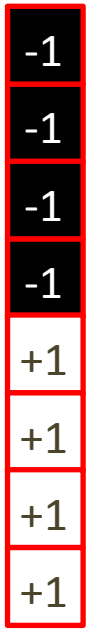
Class (+1)  …

Class (-1)  …

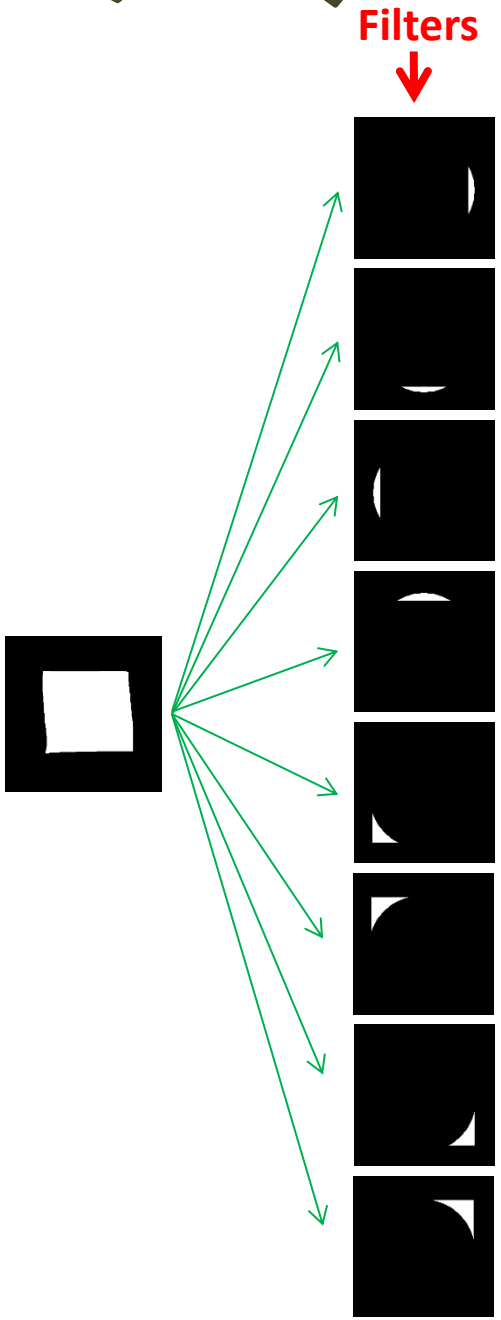
➔ Hand-crafted filters:

Each filter is L_1 normalized to 1

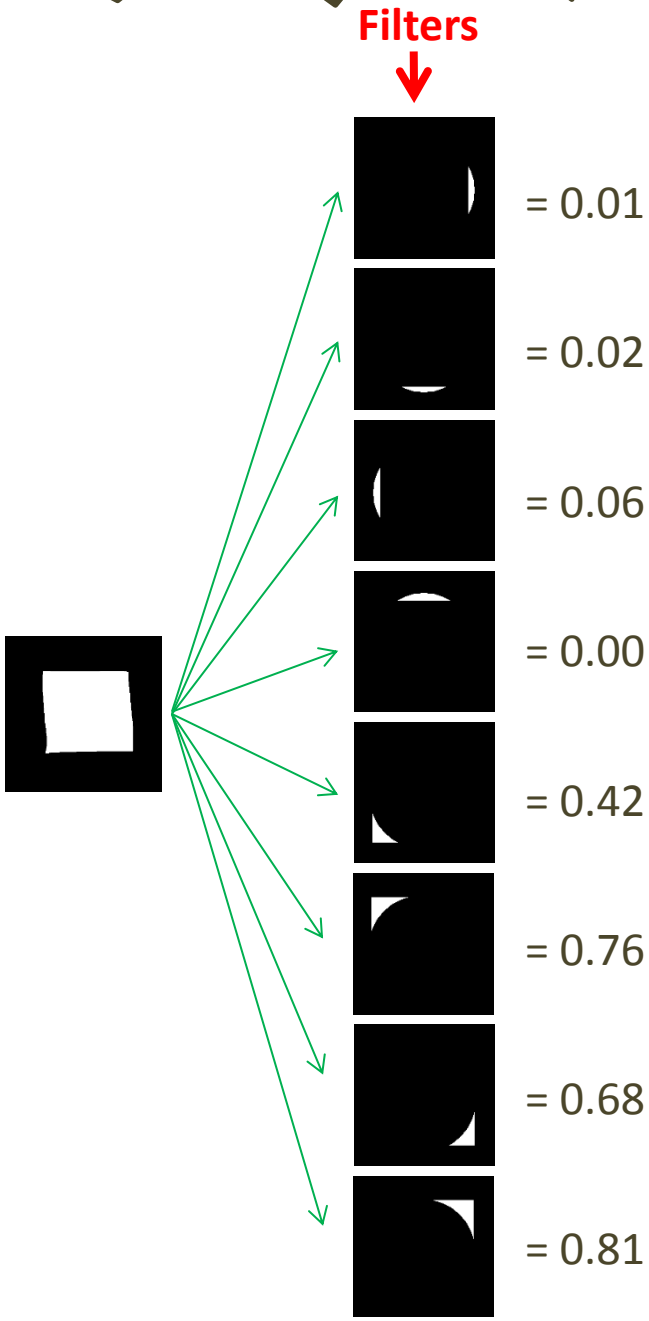




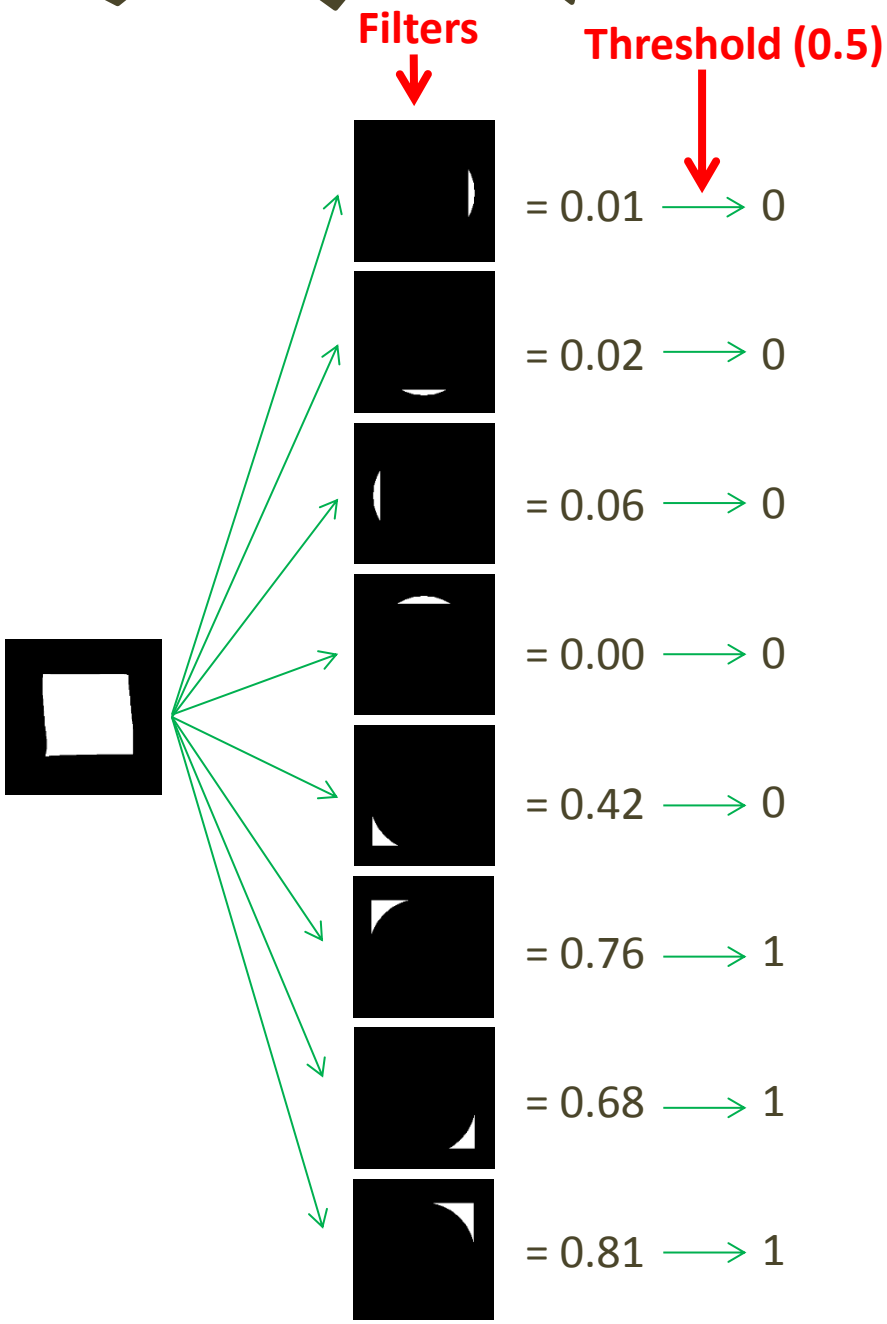
Reminders about CNN



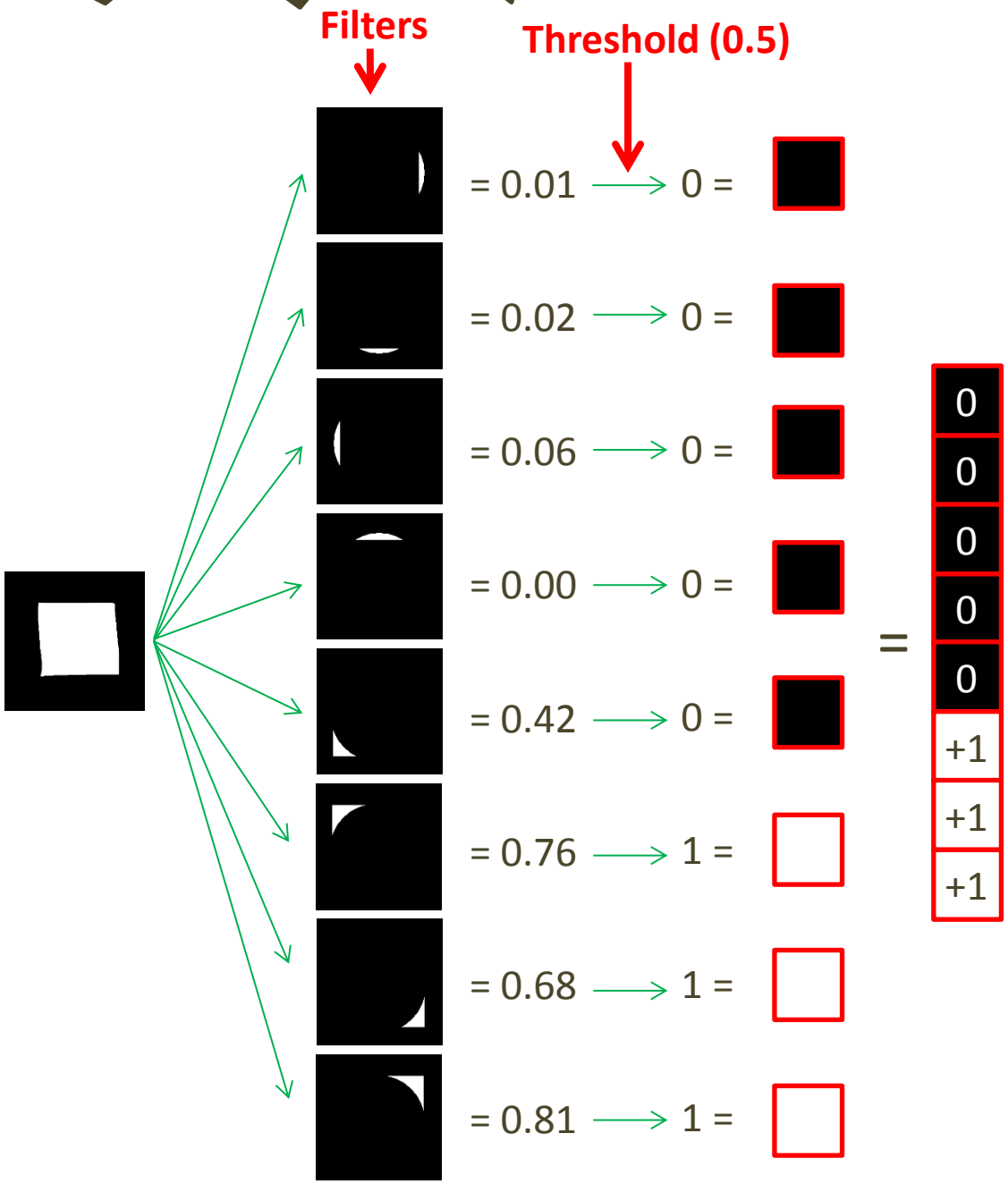
Reminders about CNN



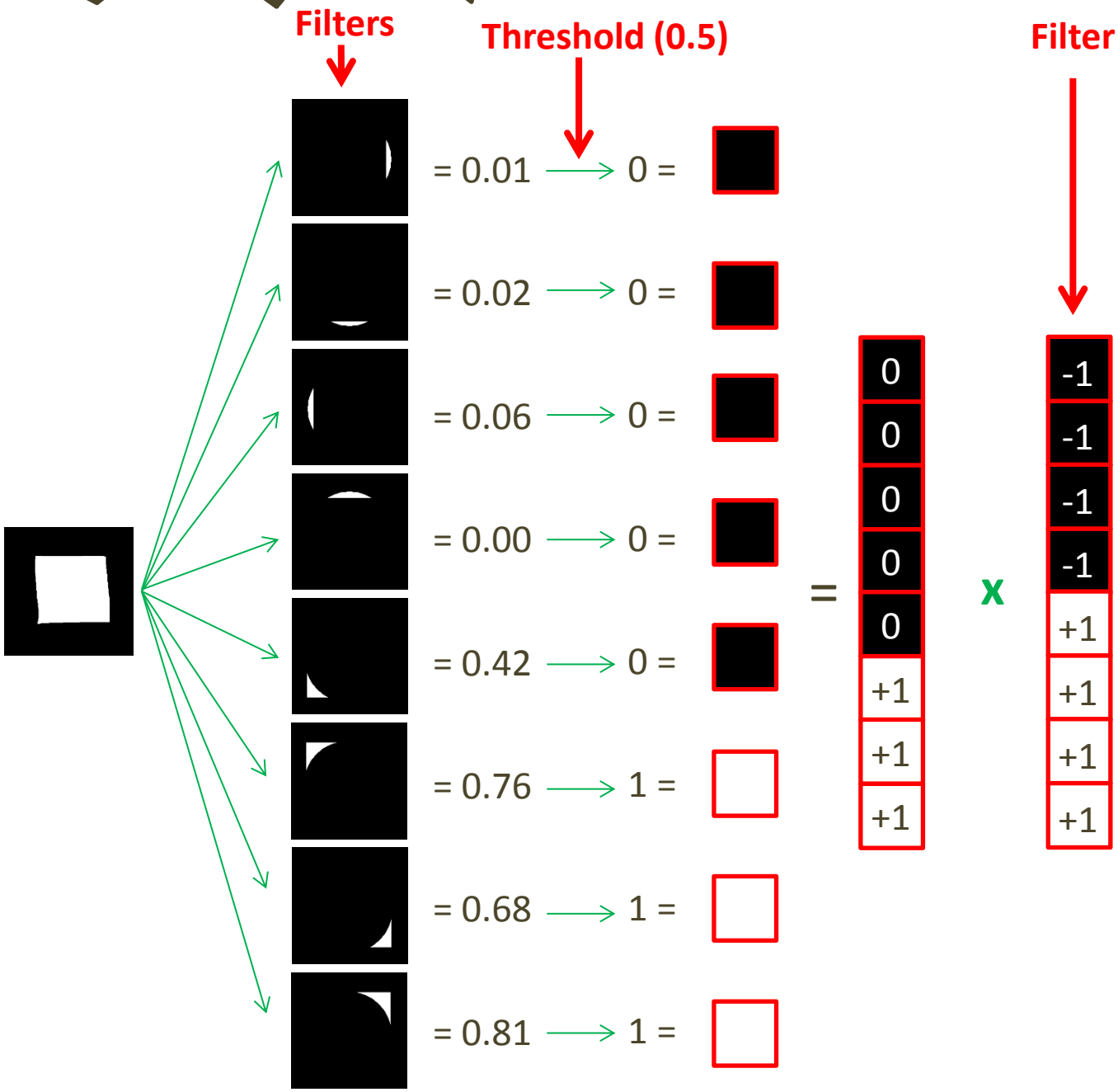
Reminders about CNN



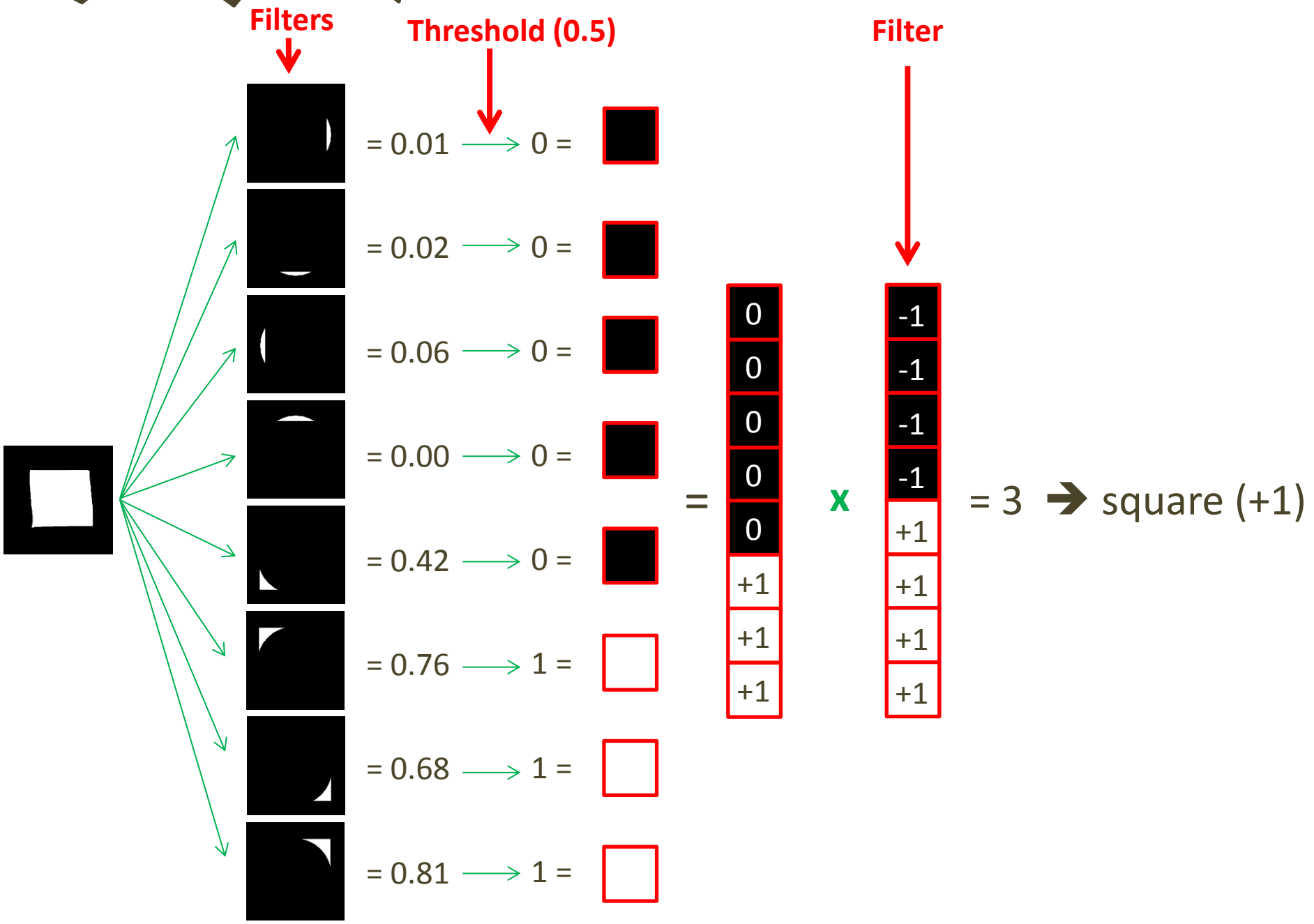
Reminders about CNN



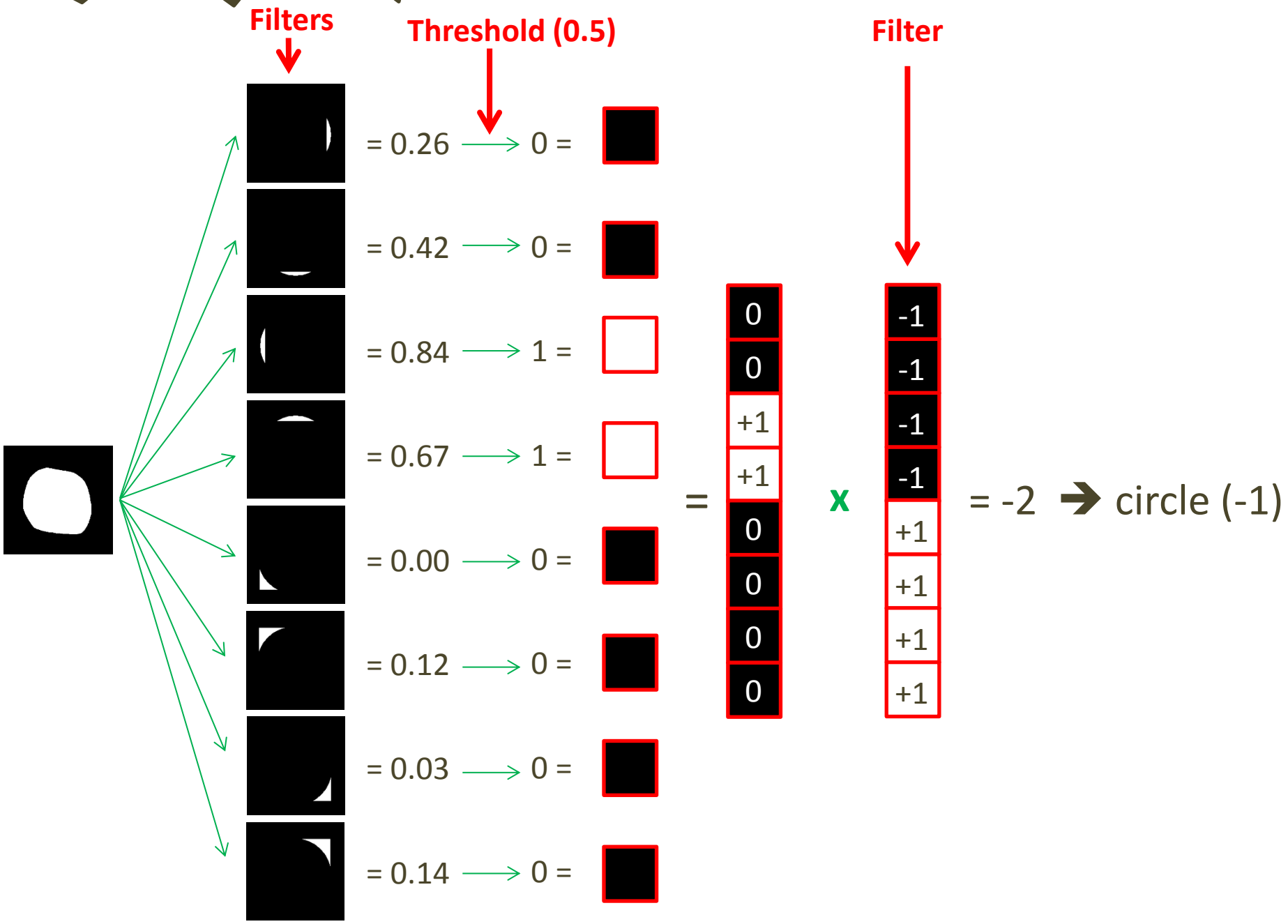
Reminders about CNN



Reminders about CNN



Reminders about CNN



Reminders about CNN

Filters

"Convolution"

-1	0	+1
-1	0	+1
-1	0	+1

*

*

-1	-1	-1
0	0	0
+1	+1	+1

Reminders about CNN

Filters

"Convolution"

-1	0	+1
-1	0	+1
-1	0	+1

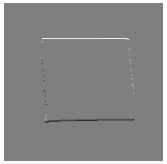
=

*

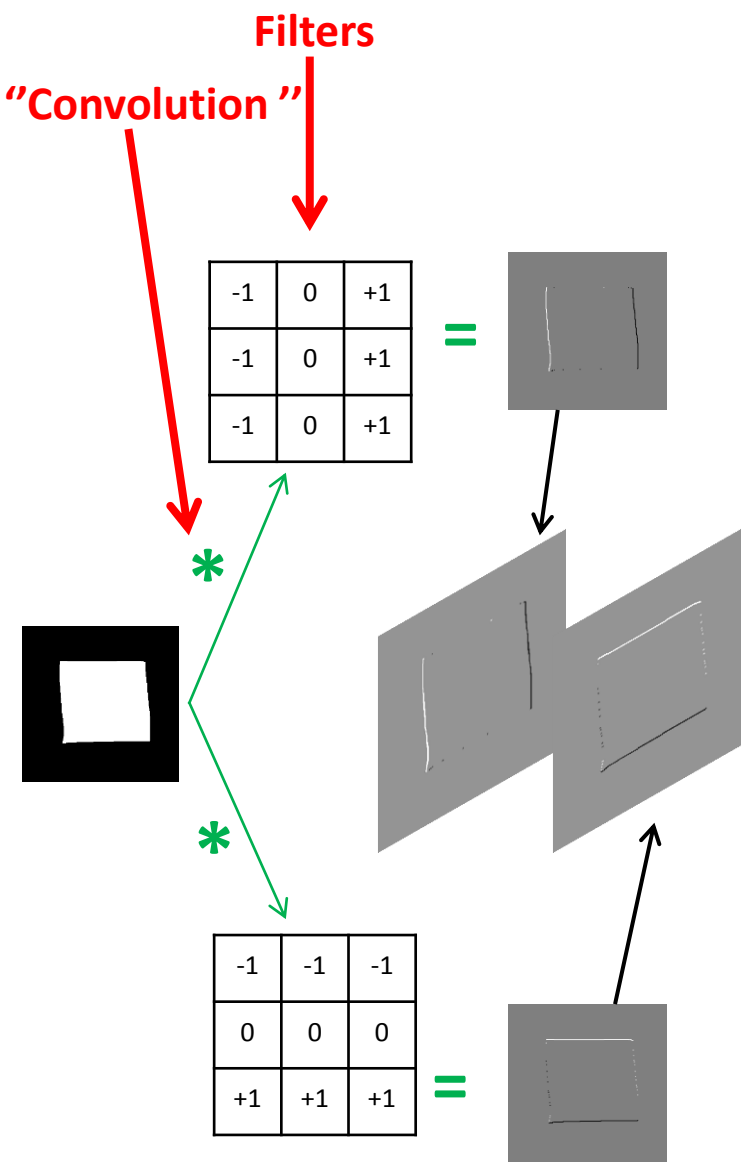
*

-1	-1	-1
0	0	0
+1	+1	+1

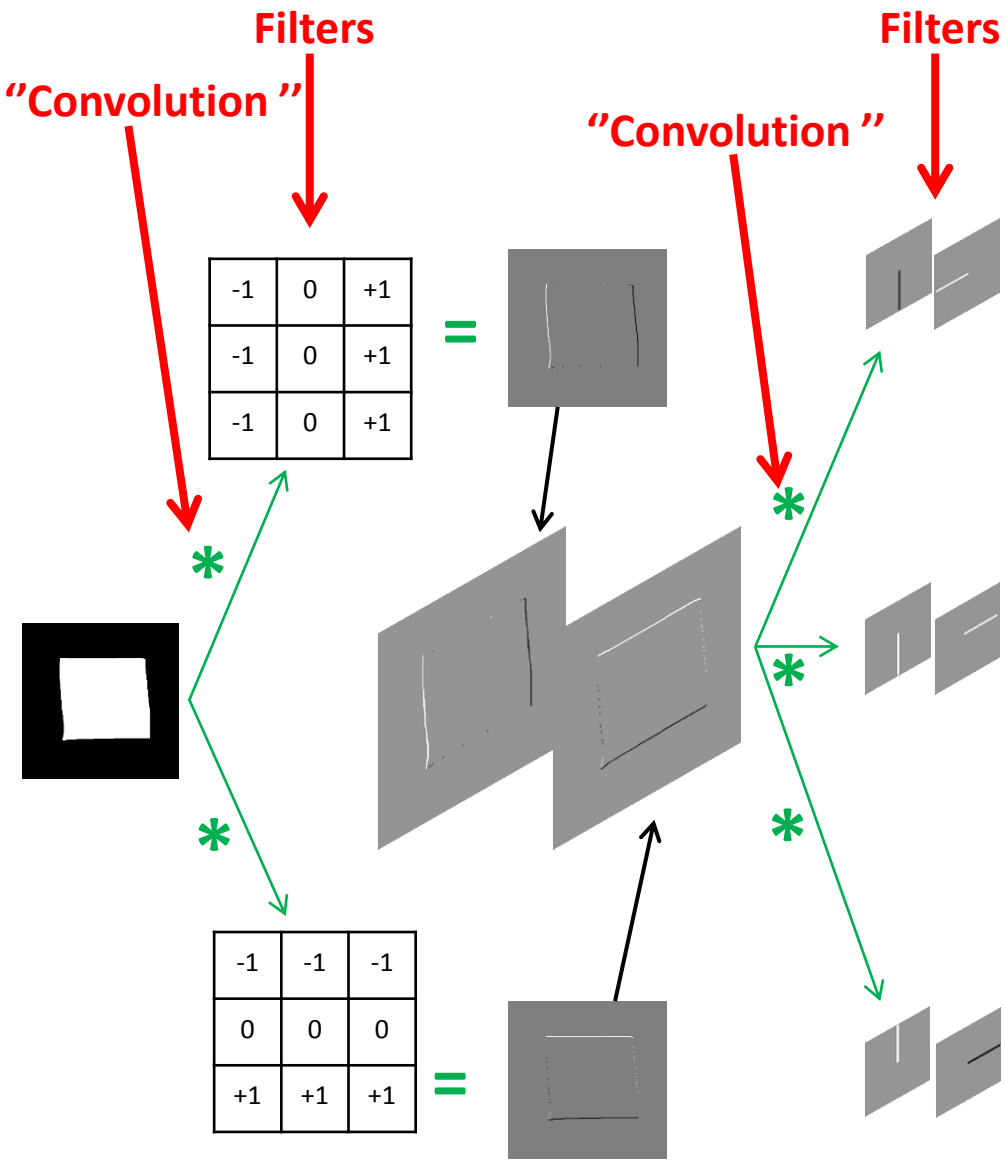
=



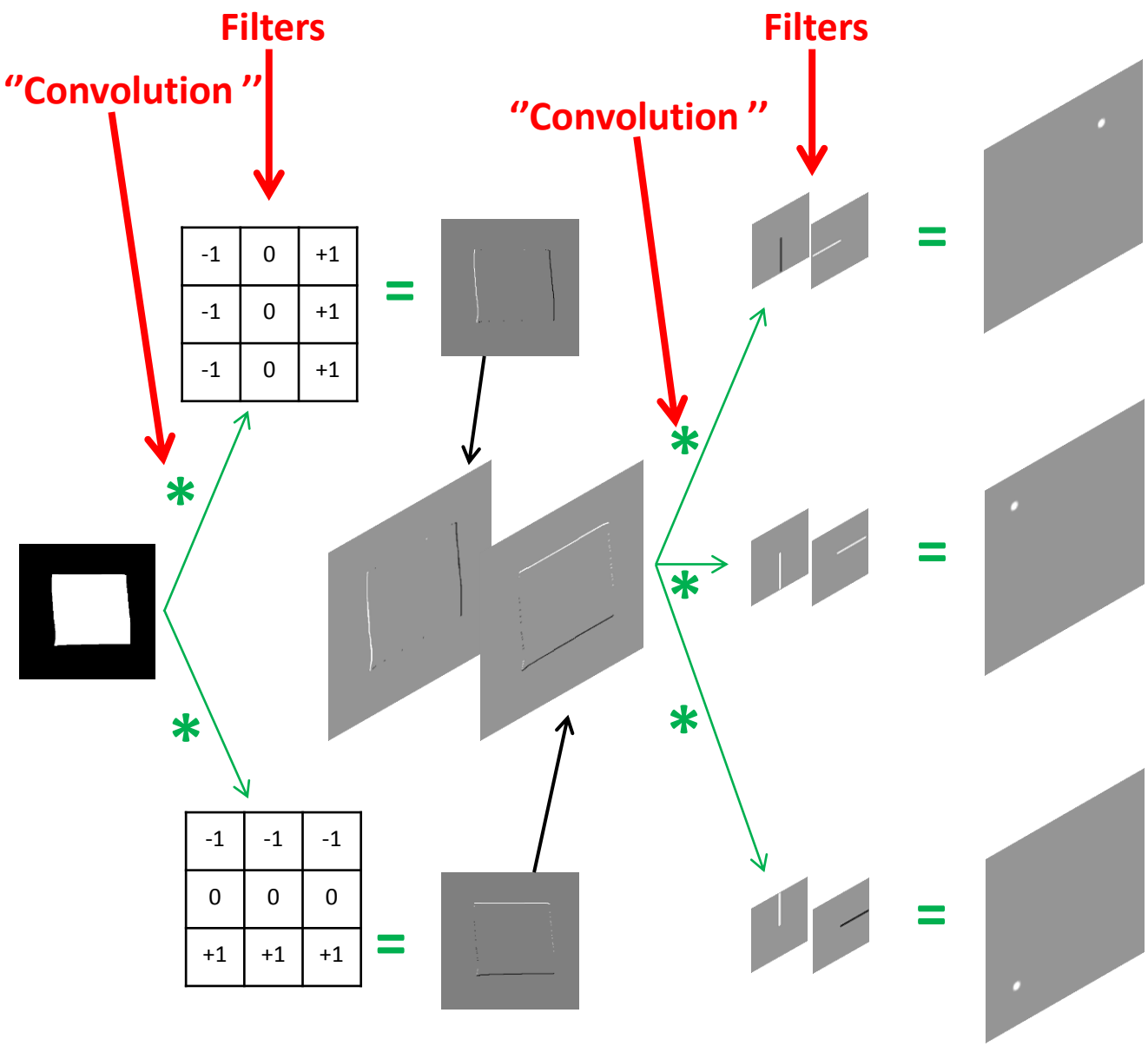
Reminders about CNN



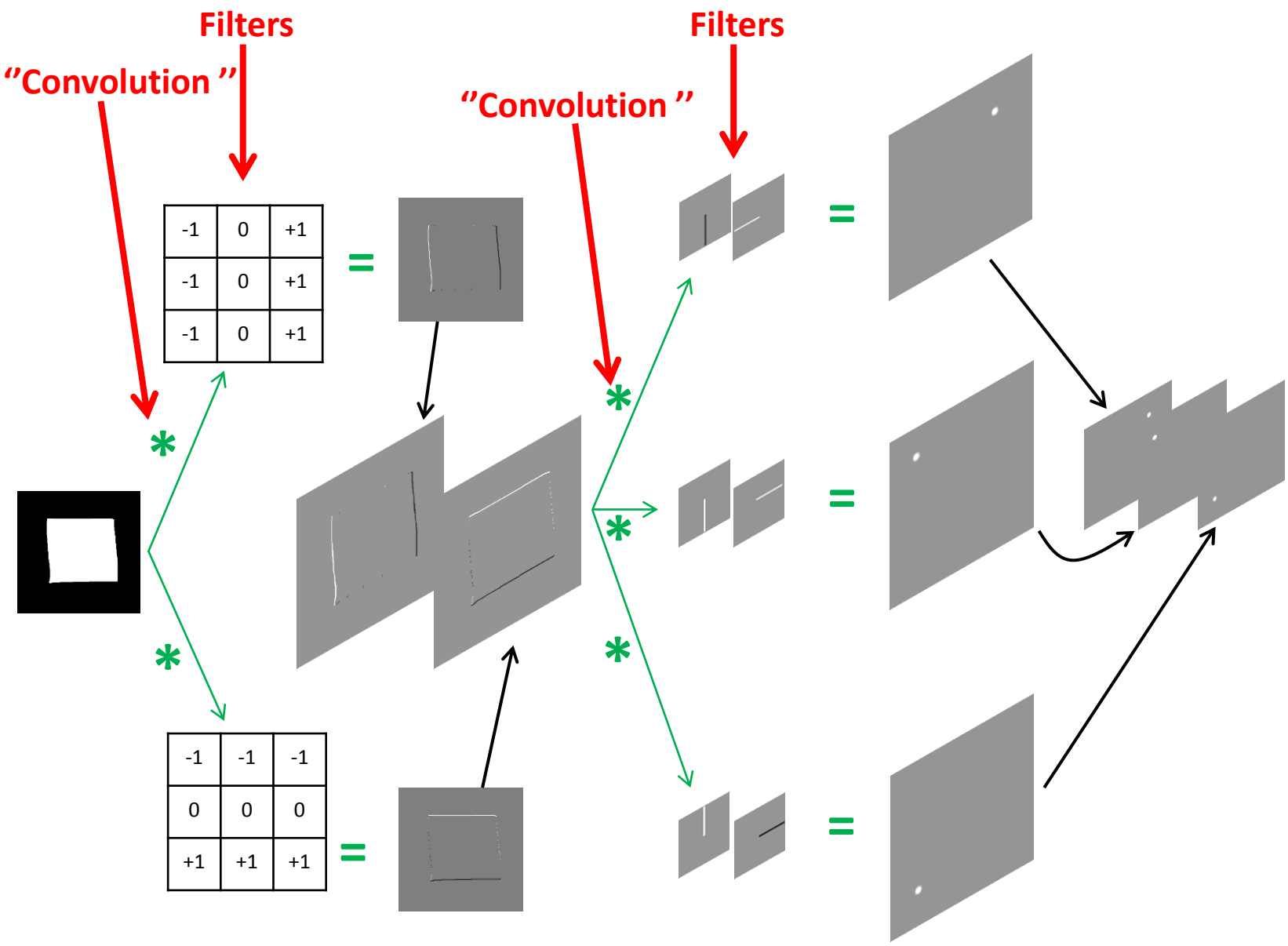
Reminders about CNN



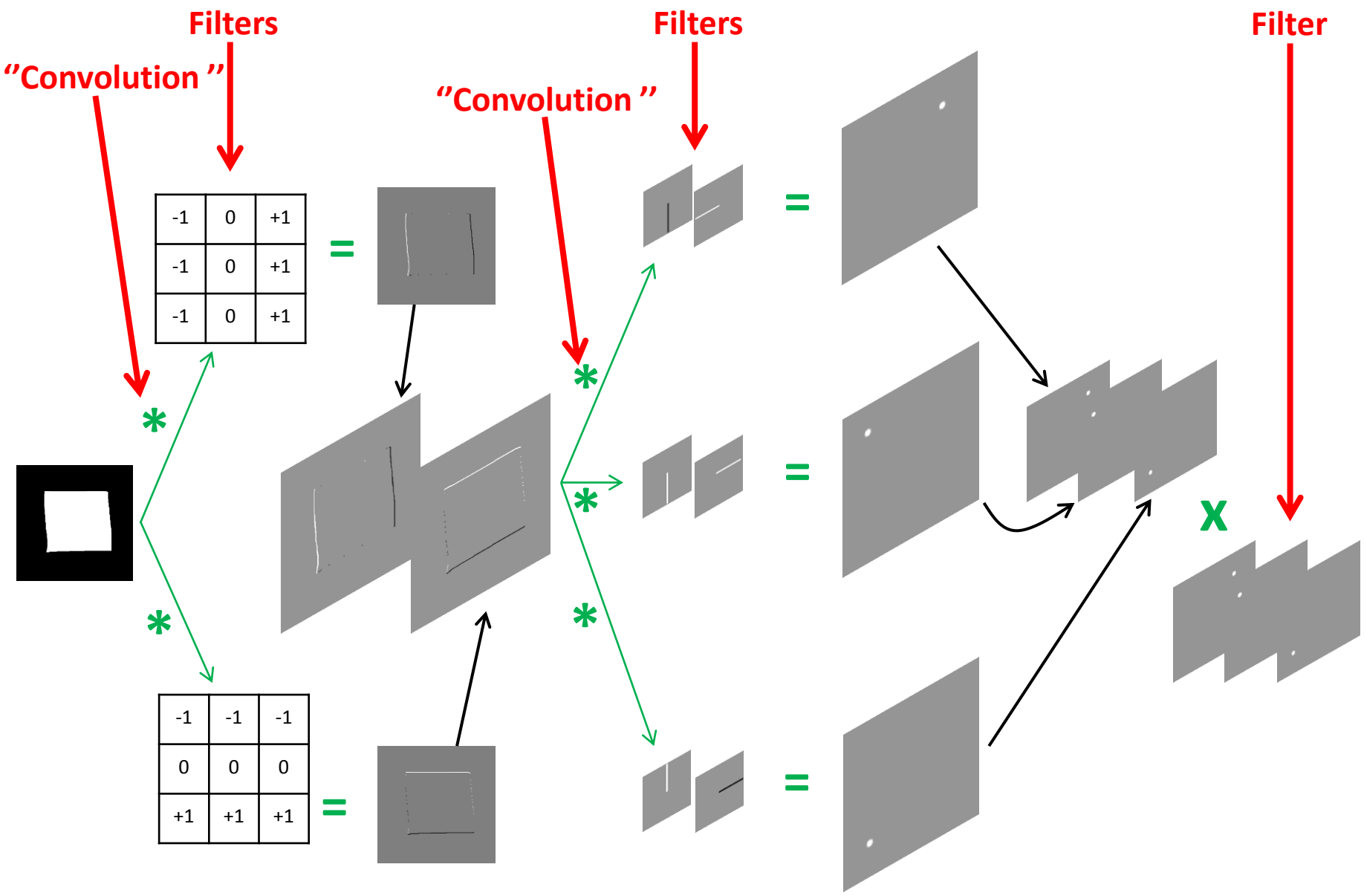
Reminders about CNN



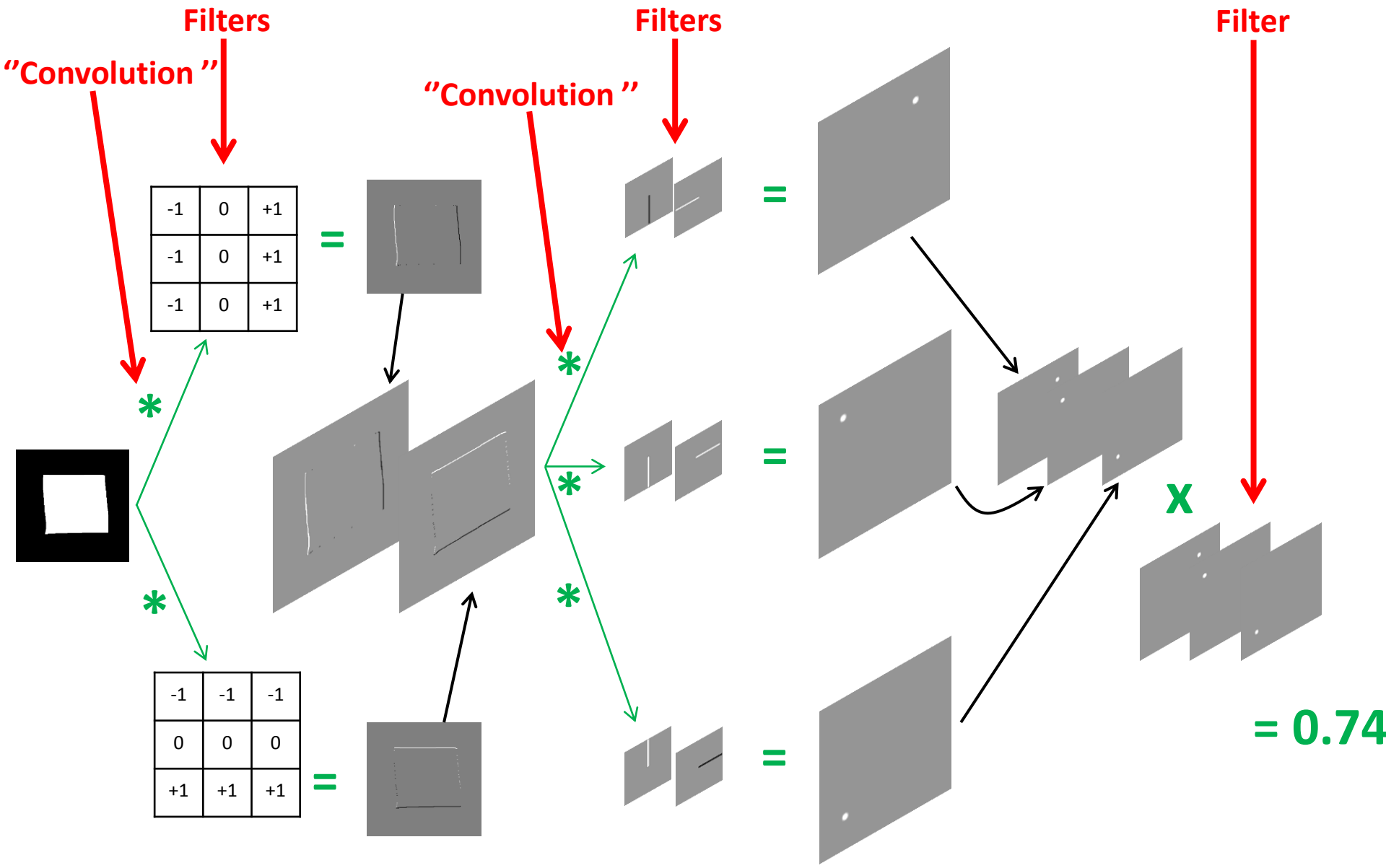
Reminders about CNN



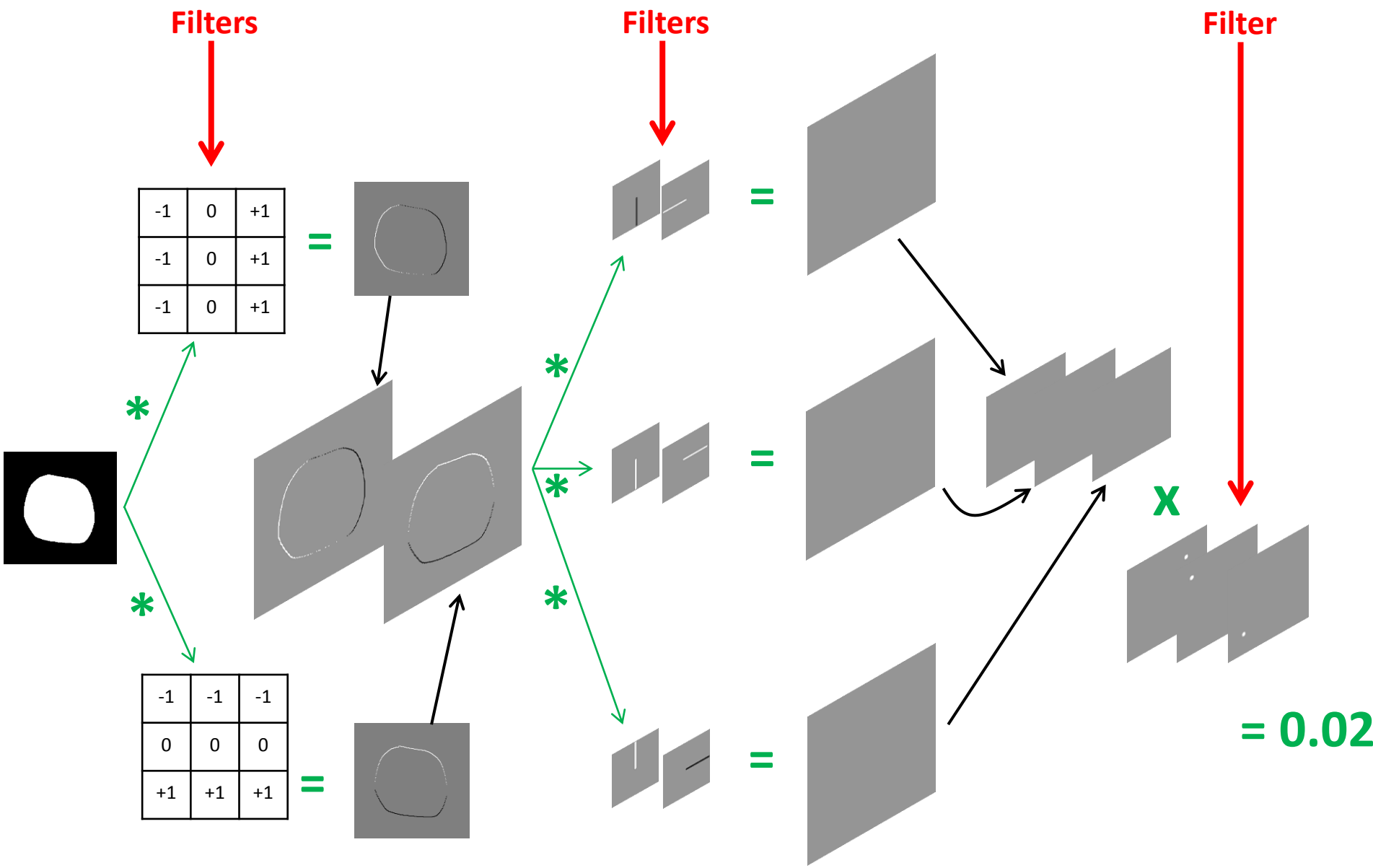
Reminders about CNN



Reminders about CNN



Reminders about CNN



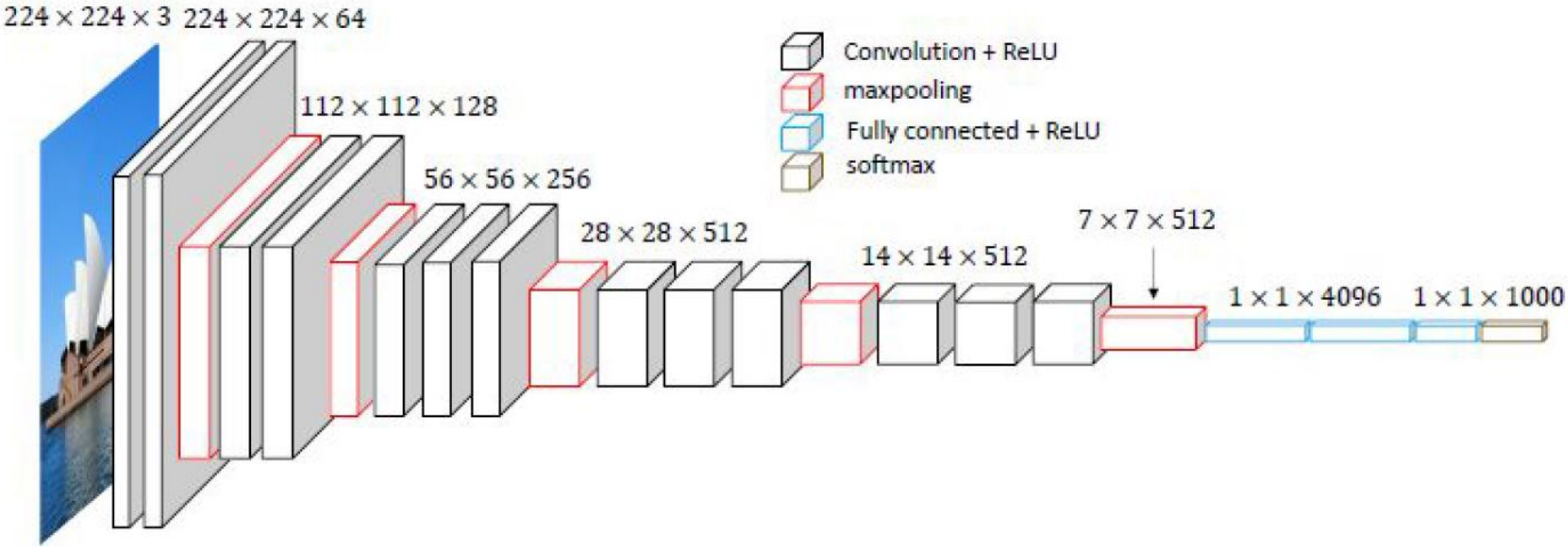
Reminders about CNN

Handcrafted features (eg. SIFT+ BOW) → Classification (eg. SVM) → *''Bike''*

CNN → *''Bike''*

Reminders about CNN

One concrete example



Questions :

How many layers ?

How many filters in the first layer ?

Size of each filter ?

How many weights in the first Conv layer ? In the first FC layer ?

Receptive field of a 1x1 window in the feature map before the first FC layer ?

What are MaxPooling, SoftMax, Convolution, ReLU ?

Reminders about CNN

Convolution

$$I * f\ (x, y) = \sum_{i=-d}^{+d} \sum_{j=-d}^{+d} f(i,j)\, I(x-i, y-j)$$

5	3	1	4	9	5	2	6	7	8
1	5	2	4	6	2	8	6	2	5
4	5	2	5	4	6	9	8	5	2
1	4	5	2	6	3	2	5	8	7
1	0	2	5	0	1	4	0	1	5
2	1	0	5	3	6	9	8	5	2
1	4	1	2	3	0	2	5	4	0
4	1	2	8	7	9	8	7	9	2

$*$

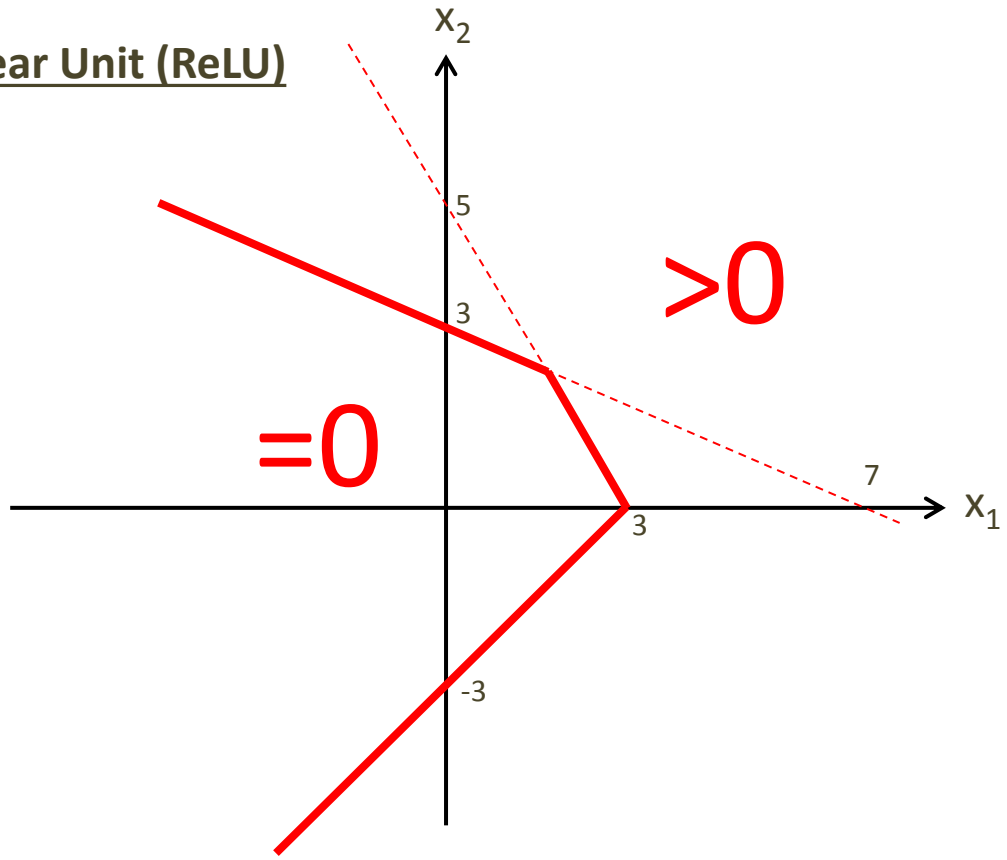
2	1	1
3	1	5
2	4	0

$=$

	43								

Reminders about CNN

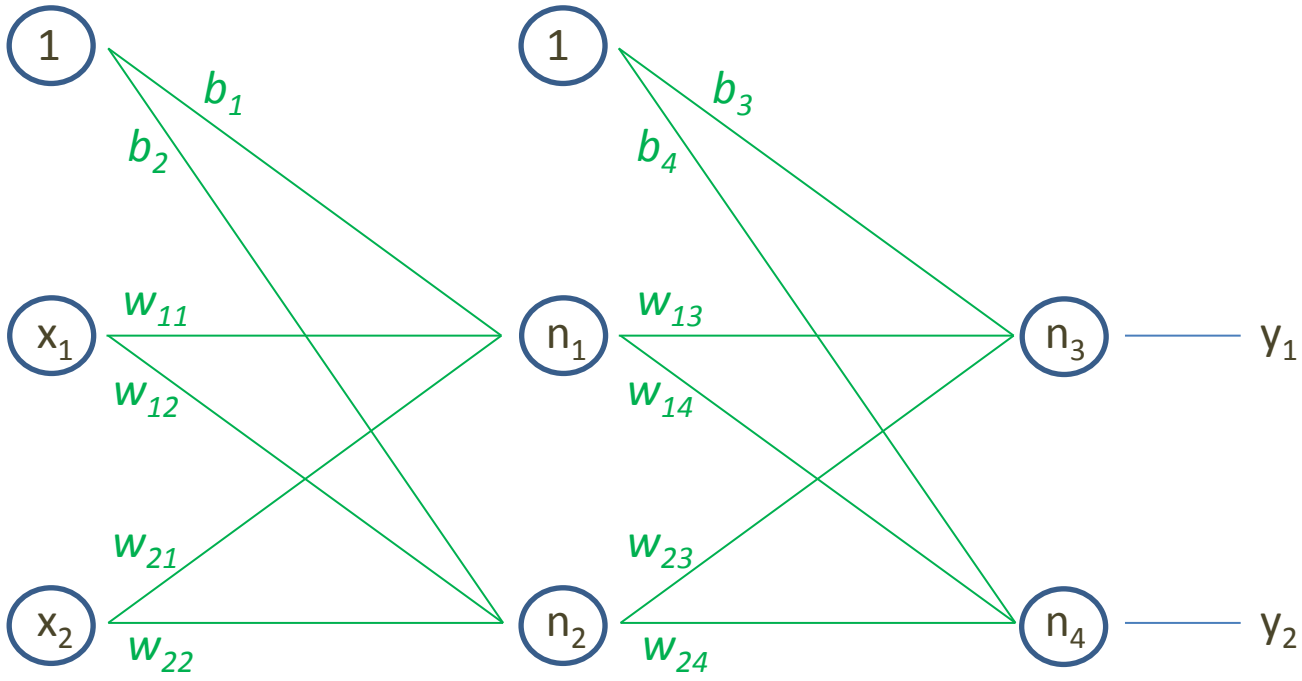
Rectified Linear Unit (ReLU)



Exercise: Propose a NN with ReLU activations that solves this problem.

Reminders about CNN

Back-propagation : Stochastic Gradient Descent and Chain Rule



b_1	0.2
b_2	-0.1
w_{11}	-0.3
w_{12}	0.2
w_{21}	0.4
w_{22}	0.2
b_3	-0.1
b_4	0.6
w_{13}	0.5
w_{14}	-0.2
w_{23}	0.7
w_{24}	0.1

Square Loss and sigmoids as activation functions.

$\begin{pmatrix} x_1=0.1 \\ x_2=0.3 \end{pmatrix}$ ➔ expected output $\begin{pmatrix} y_1 = 0.5 \\ y_2 = 0.8 \end{pmatrix}$

Exercise: *Real output ? Update w_{21} with η=1 (SGD).*

Reminders about CNN

Adam optimizer, learning rate decay, momentum, cross entropy, drop out, … in the next practical session.