BONUS LECTURES

Make the most out of it!!!

LECTURE 12

AGENDA FOR TODAY

- Overview of Deep Learning
- Convolutional Neural Networks
- Recurrent Neural Networks
- LSTMs
- Case study

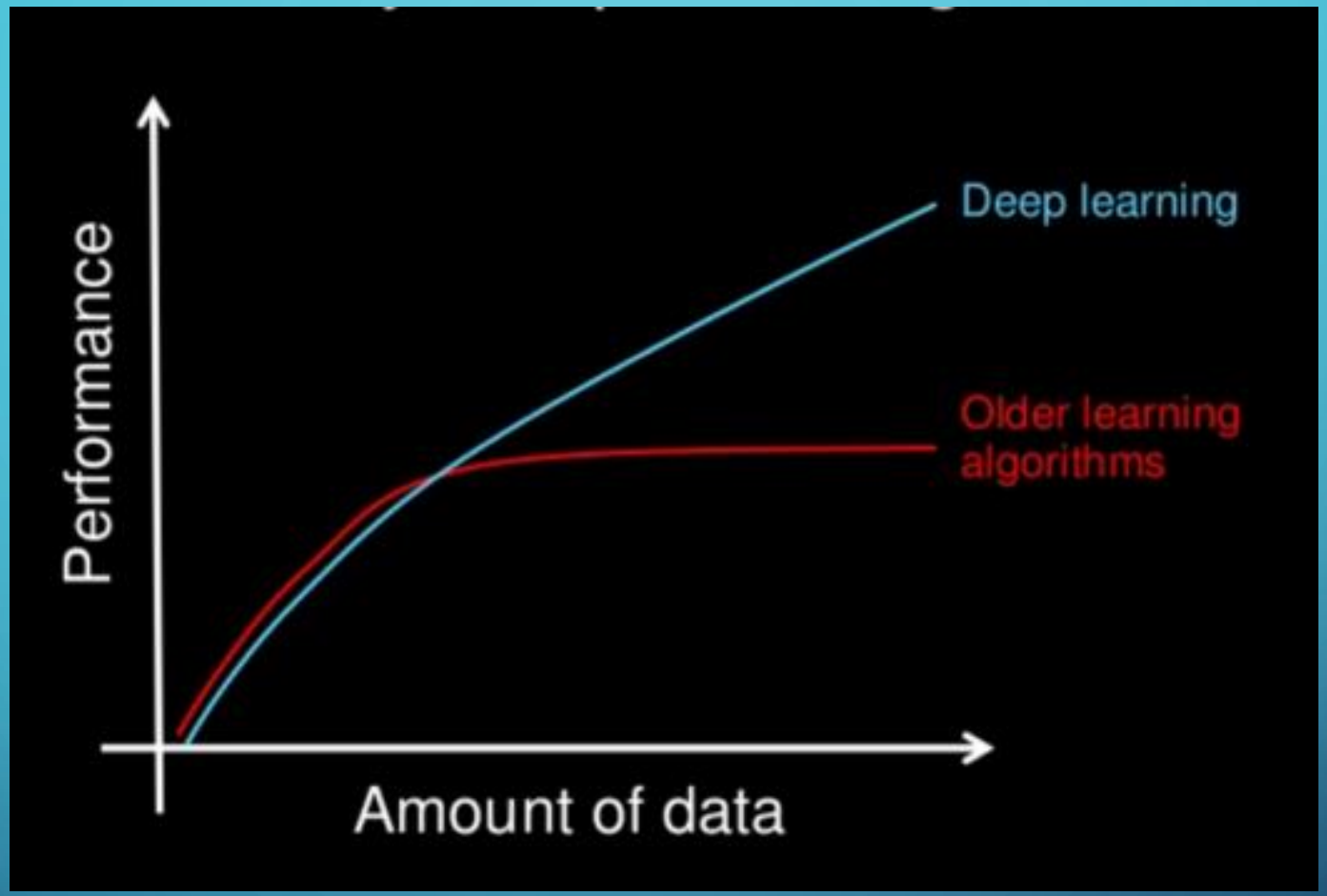
DEEP LEARNING

- Simple Definition: Deep Learning is Large Neural Networks

WHY DEEP LEARNING IS TAKING OFF??

Scale drives deep learning progress

- Data
- Computation
- Algorithms



One reason that deep learning has taken off like crazy is because it is fantastic at supervised learning

- Jeff Dean is a Wizard and Google Senior Fellow in the Systems and Infrastructure Group at Google
- In a 2016 talk titled "Deep Learning for Building Intelligent Computer Systems" he made a comment in the similar vein, that deep learning is really all about large neural networks.

"When you hear the term deep learning, just think of a large deep neural net. Deep refers to the number of layers typically and so this kind of the popular term that's been adopted in the press. I think of them as deep neural networks generally."

- Yoshua Bengio, another leader in deep learning:

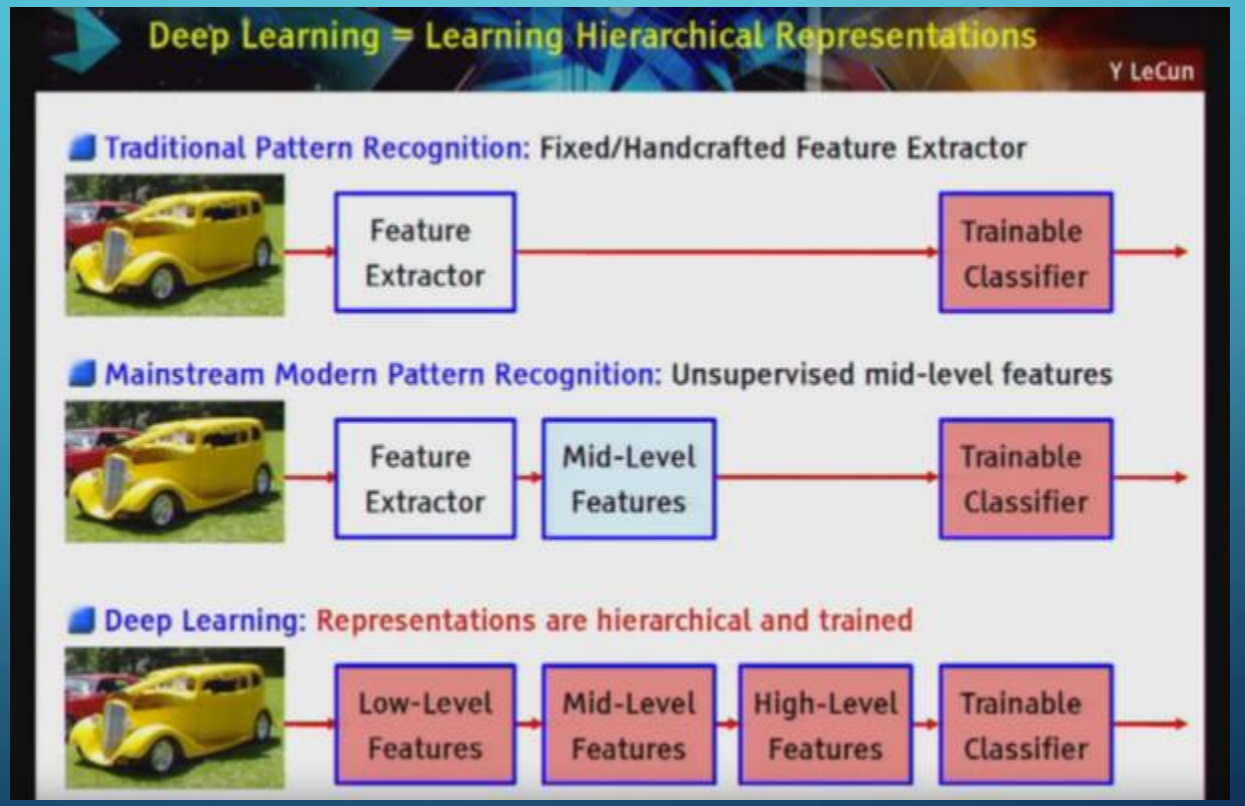
"Deep learning algorithms seek to exploit the unknown structure in the input distribution in order to discover good representations, often at multiple levels, with higher-level learned features defined in terms of lower-level features"

- Peter Norvig, Director of Research at Google:

"a kind of learning where the representation you form have several levels of abstraction, rather than a direct input to output"

- Yann LeCun, the Director of Facebook research:

"deep learning [is] … a pipeline of modules all of which are trainable. … deep because [has] multiple stages in the process of recognizing an object and all of those stages are part of the training"

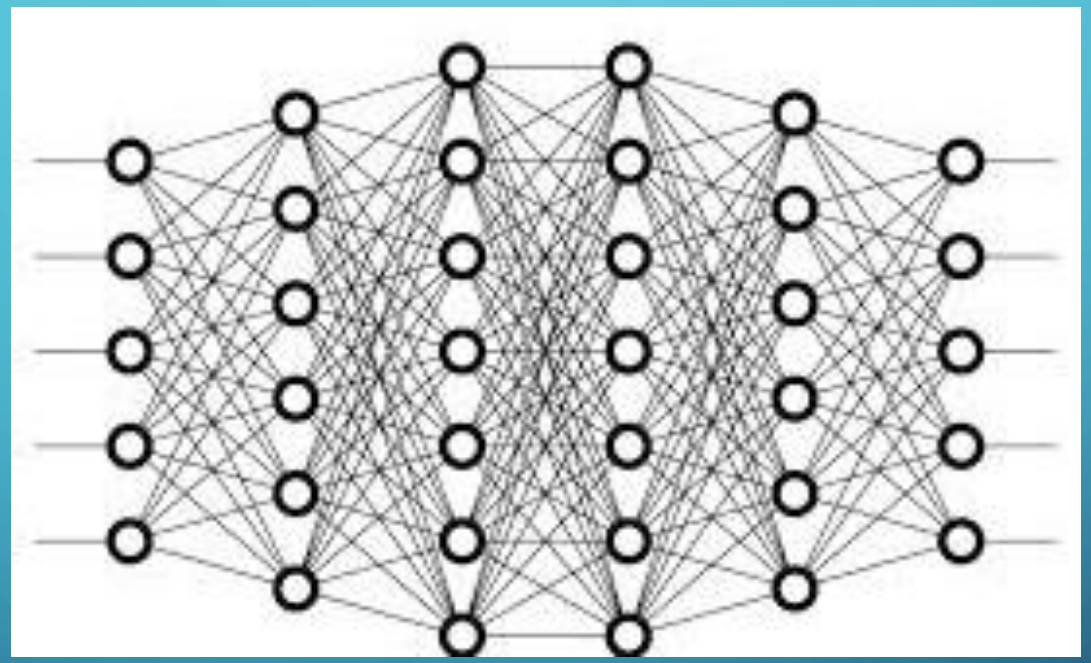


THE MOST POPULAR DL ALGORITHMS

- Multilayer Perceptron Network
- Convolutional Neural Network
- Long Short-Term Memory Recurrent Neural Networks

MULTILAYER PERCEPTRON NETWORK

- Just another term for Vanilla Neural Network which we saw earlier
- Usually if we go beyond 4-5 hidden layers then we enter the territory of Deep Learning



CONVOLUTIONAL NEURAL NET

- Primary application is in the field of Computer Vision

COMPUTER VISION PROBLEMS

Image Classification

64x64

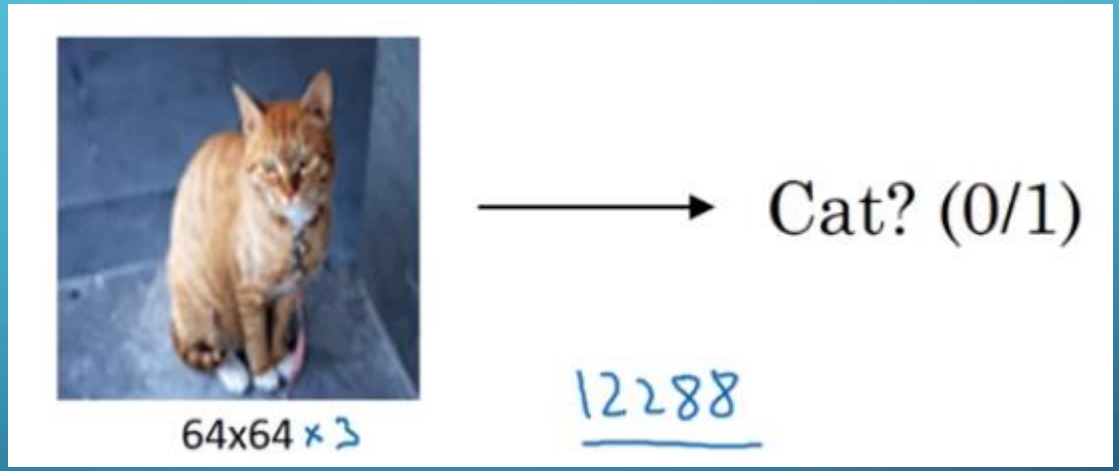
→ Cat? (0/1)

Object detection

Neural Style Transfer

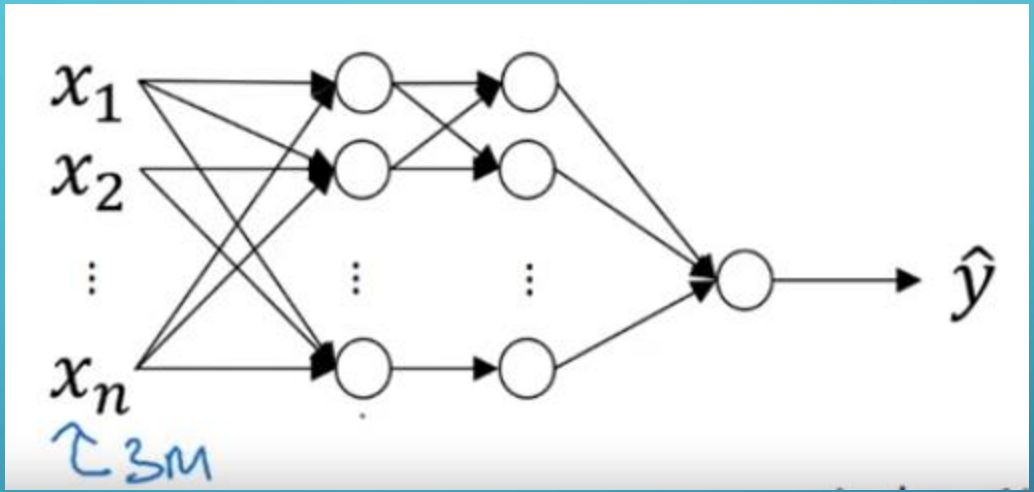
WHY CNN??

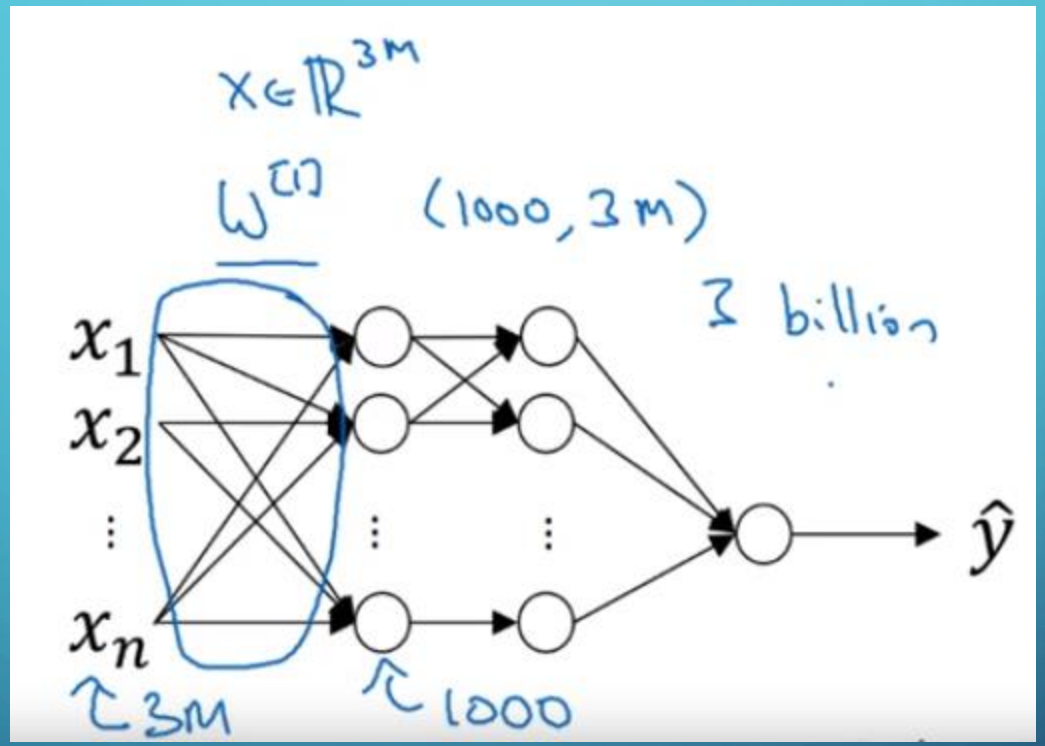
- Deep Vanilla Neural Networks are not efficient for large images



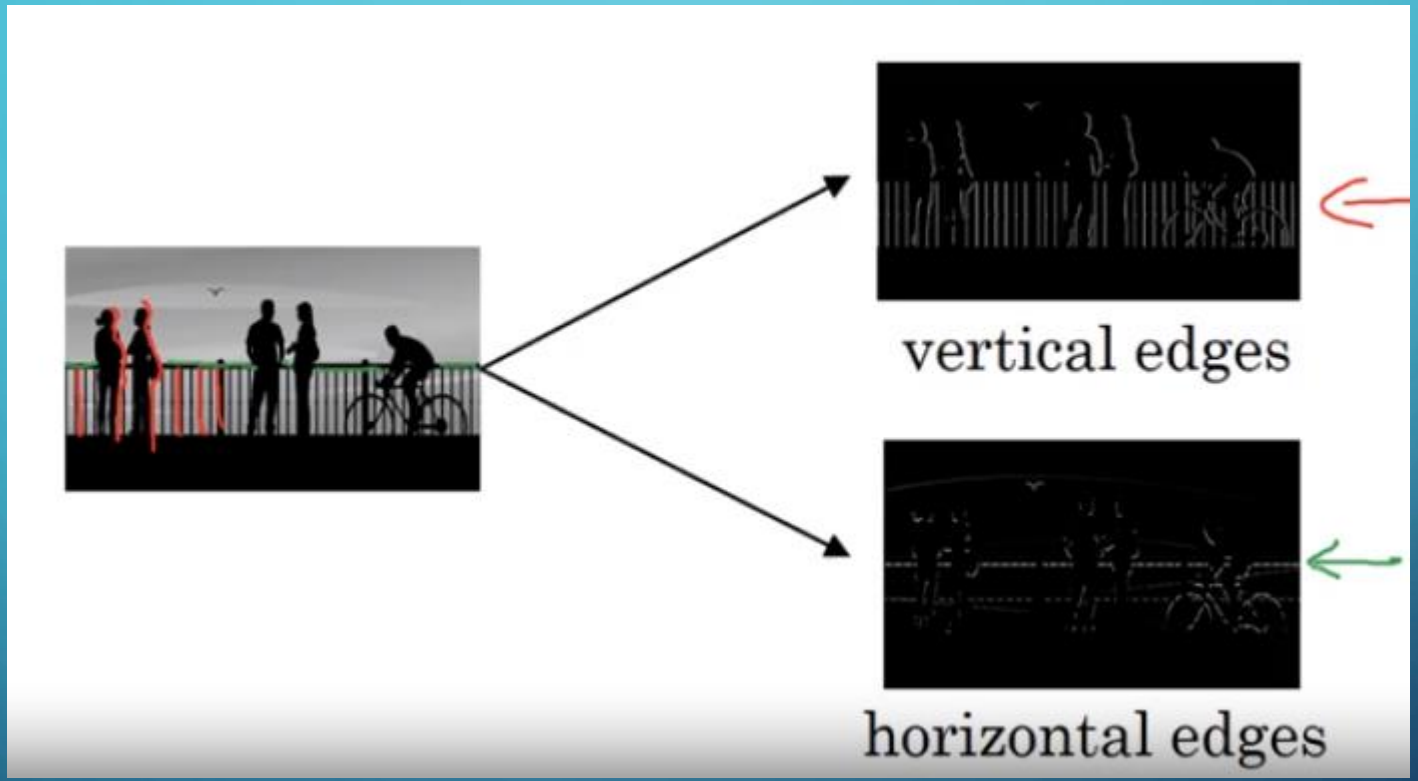
$1000 \times 1000 \times 3$

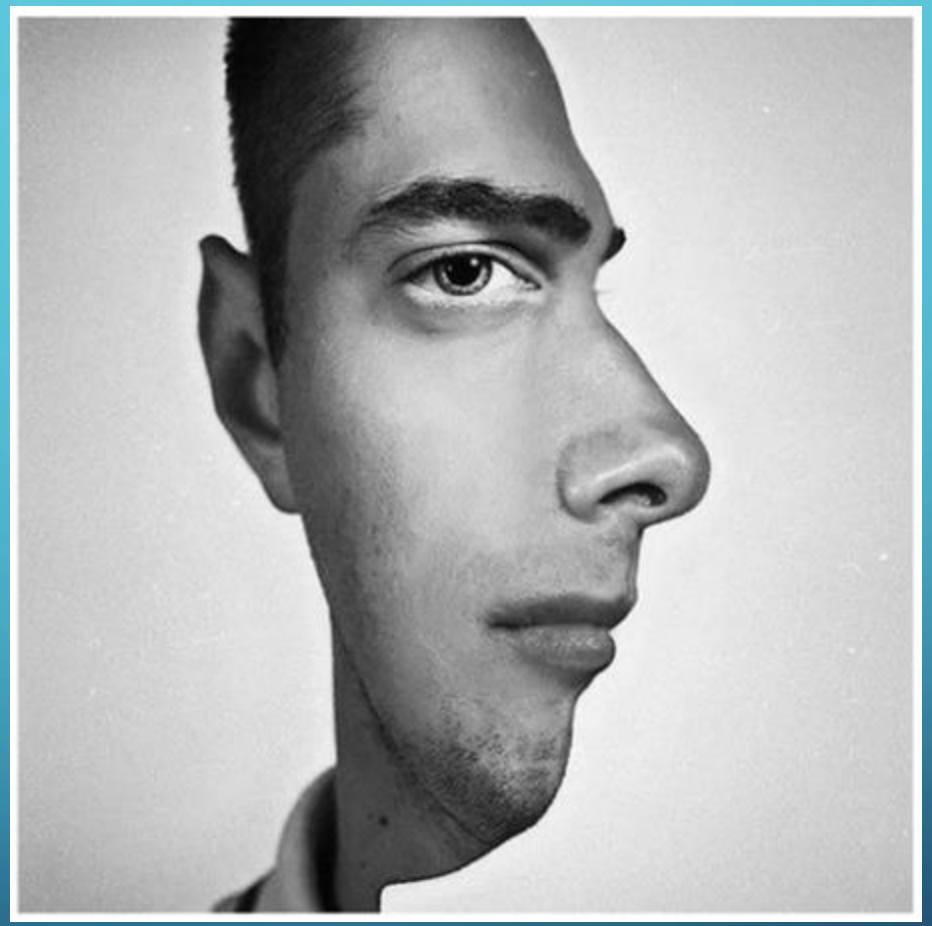
$= 3$ million


x_1
x_2
x_n
$\hat{y}$
3M

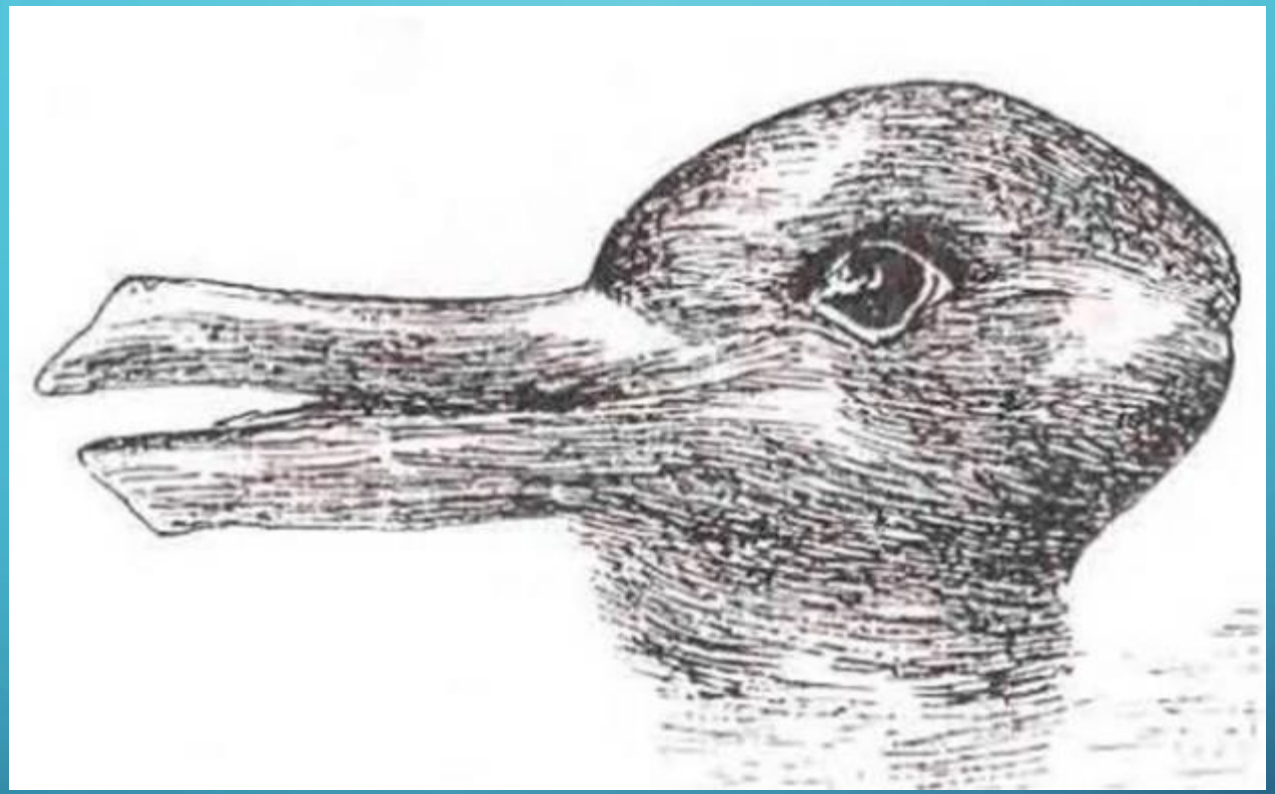

$x \in \mathbb{R}^{3M}$
$W^{[1]}$
(1000, 3M)
3 billion
x_1
x_2
x_n
$\hat{y}$
3M
1000

BRAIN ONLY LOOKS FOR FEATURES!!


vertical edges
horizontal edges



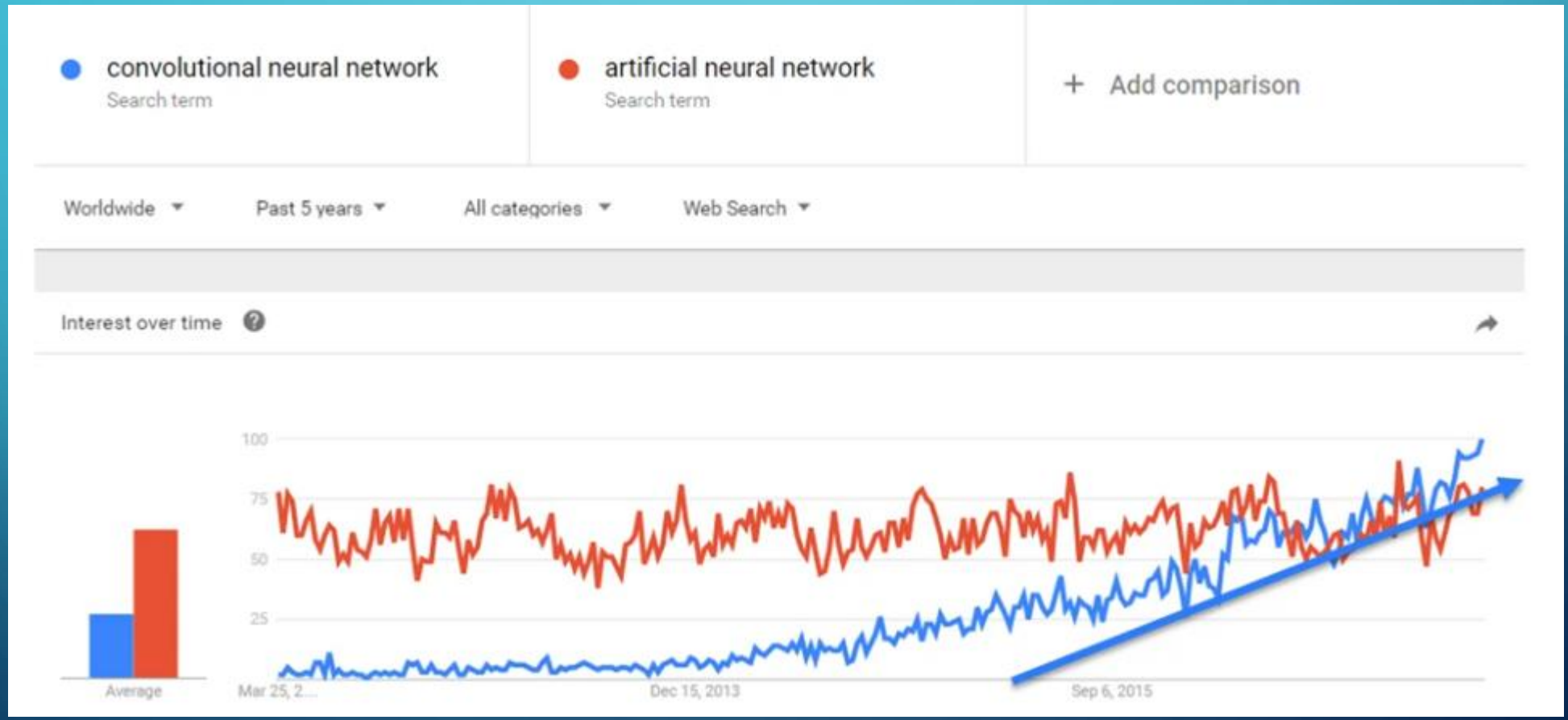
Old vs Young Lady



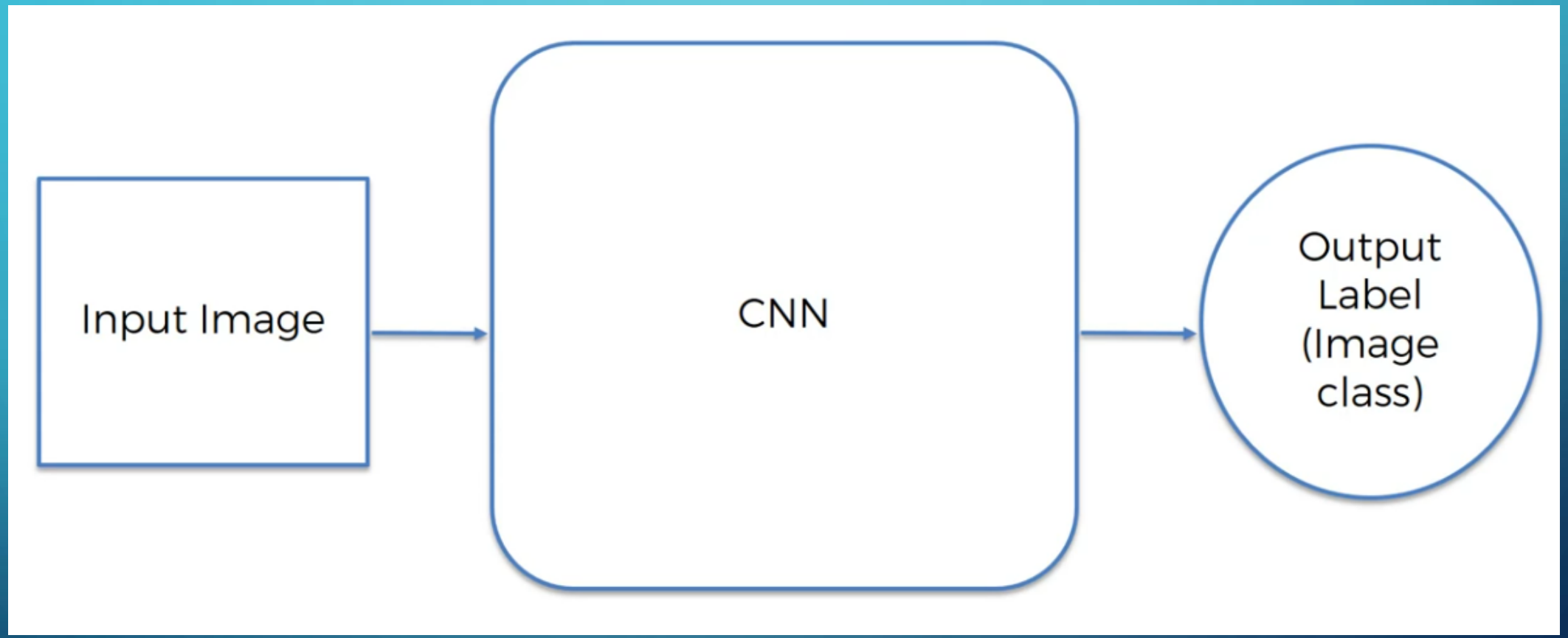
Duck or Rabbit

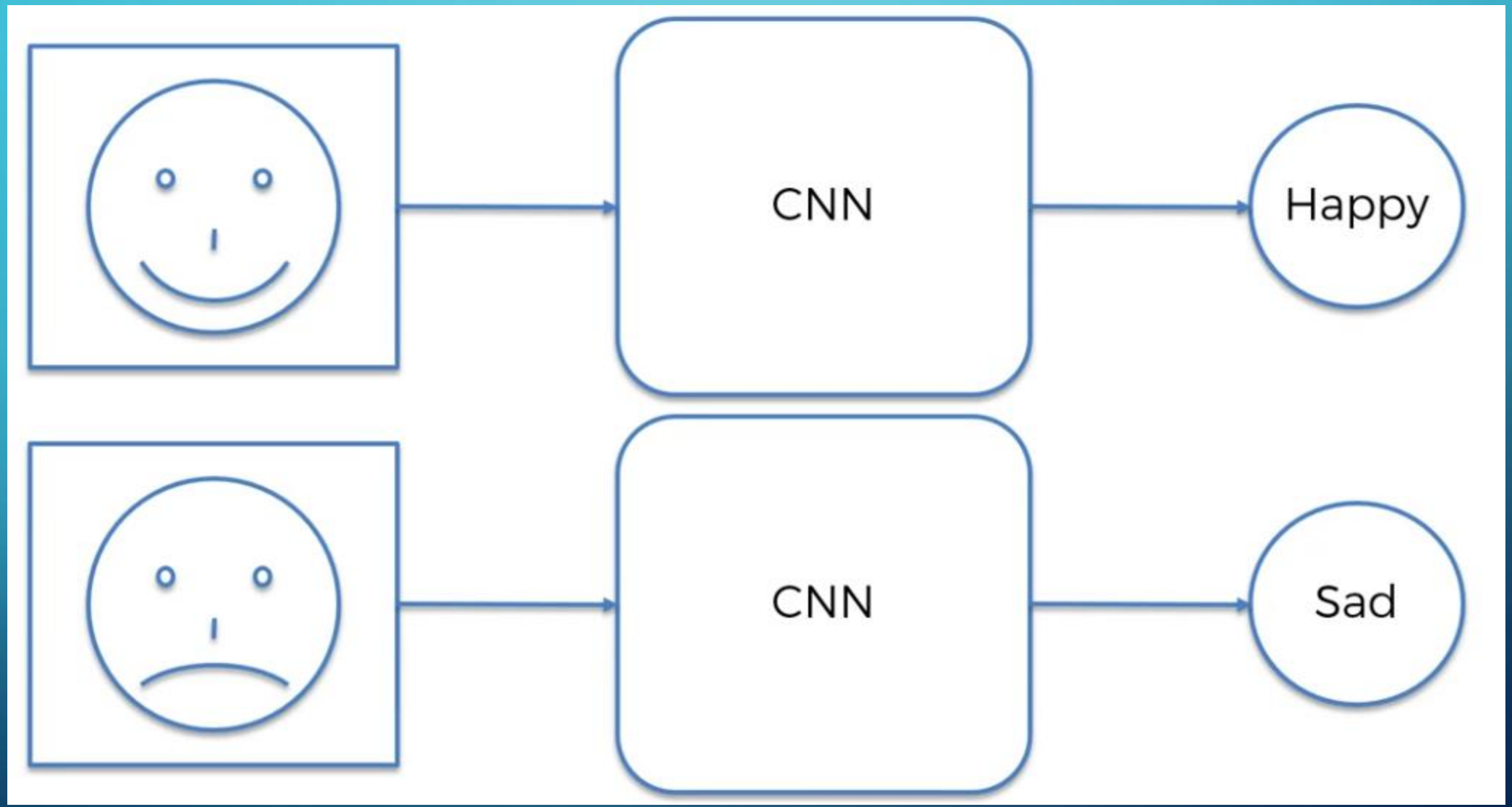
BRAIN IS CONFUSED!!

POPULARITY OF CNN

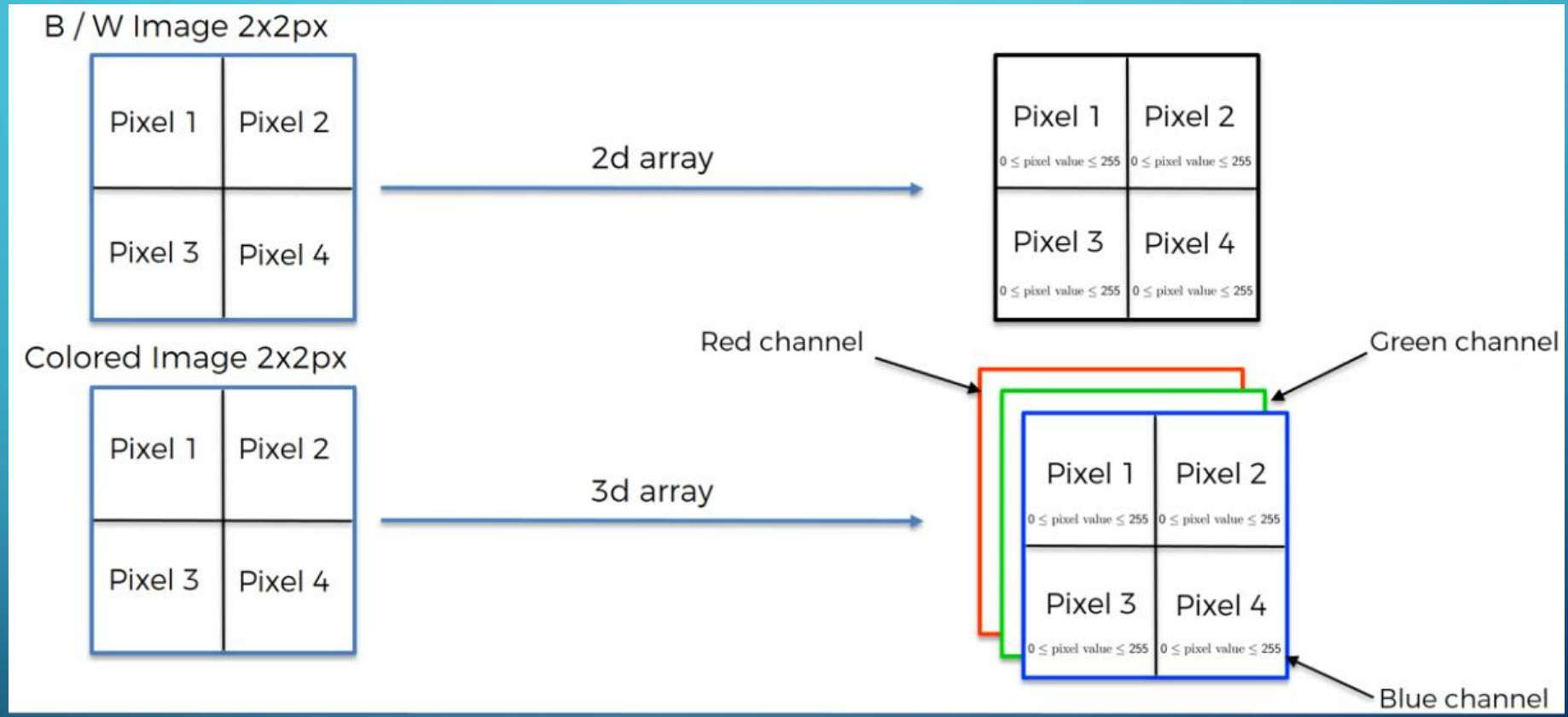


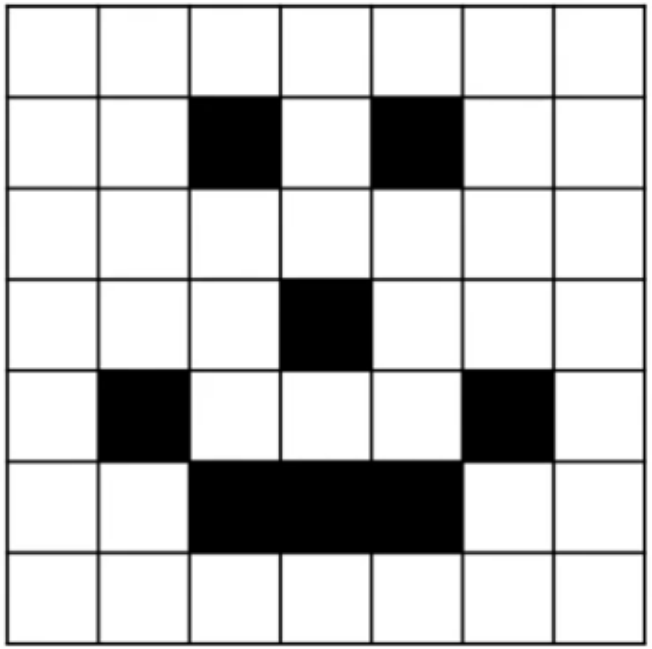
PIONEERS


Input Image
CNN
Output Label (Image class)


CNN
Happy
CNN
Sad

TYPES OF IMAGES





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

MAJOR CONCEPTS OF CNN:

- Convolution
- Max Pooling
- Flattenning
- Full Connection

CONVOLUTION

Vertical edge detection

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

6×6

Vertical edge detection

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

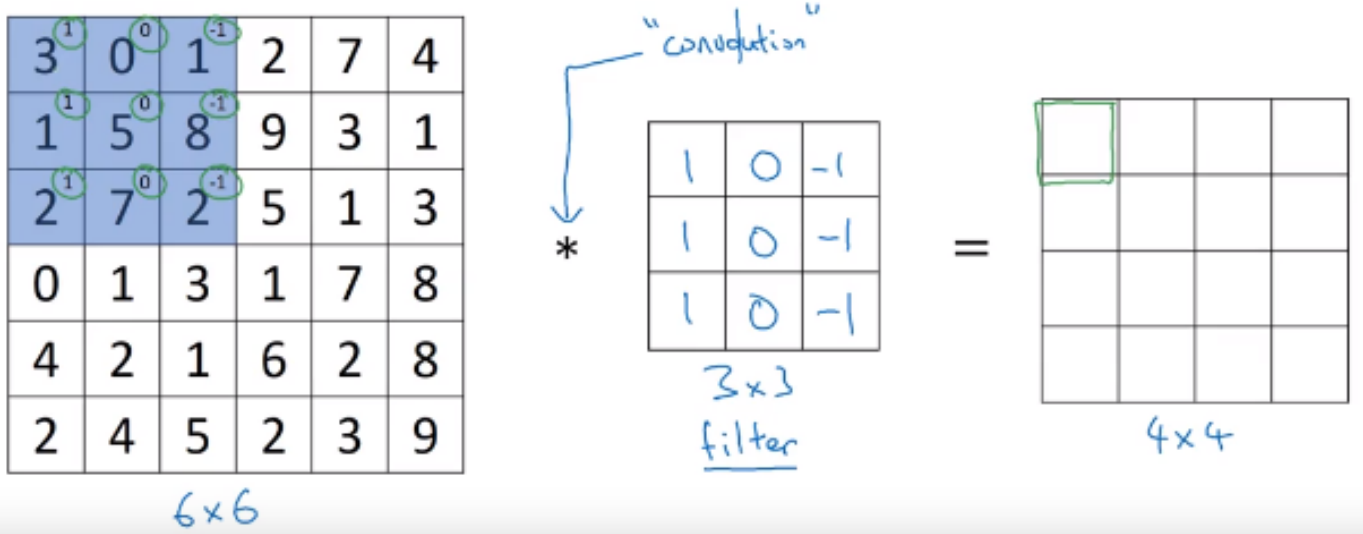
6×6

*

1	0	-1
1	0	-1
1	0	-1

3×3
filter

Vertical edge detection



Vertical edge detection

$3\times1 + 1\times1 + 2\times1 + 0\times0 + 5\times0 + 7\times0 + 1\times-1 + 8\times-1 + 2\times-1 = -5$

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

6×6

"convolution"

*

1	0	-1
1	0	-1
1	0	-1

3×3 filter

=

-5			

4×4

Vertical edge detection

$3\times1+1\times1+2\times1+0\times0+5\times0+7\times0+1\times-1+8\times-1+2\times-1=-5$

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

6×6

"convolution"

*

1	0	-1
1	0	-1
1	0	-1

3×3 filter

python: conv-forward.

=

-5	-4	0	8
-10	-2	2	3
0	-2	-4	-7
-3	-2	-3	-16

4×4

Vertical edge detection

10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0
10	10	10	0	0	0

6x6

*

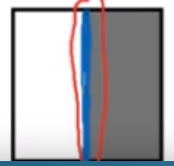
1	0	-1
1	0	-1
1	0	-1

3x3

=

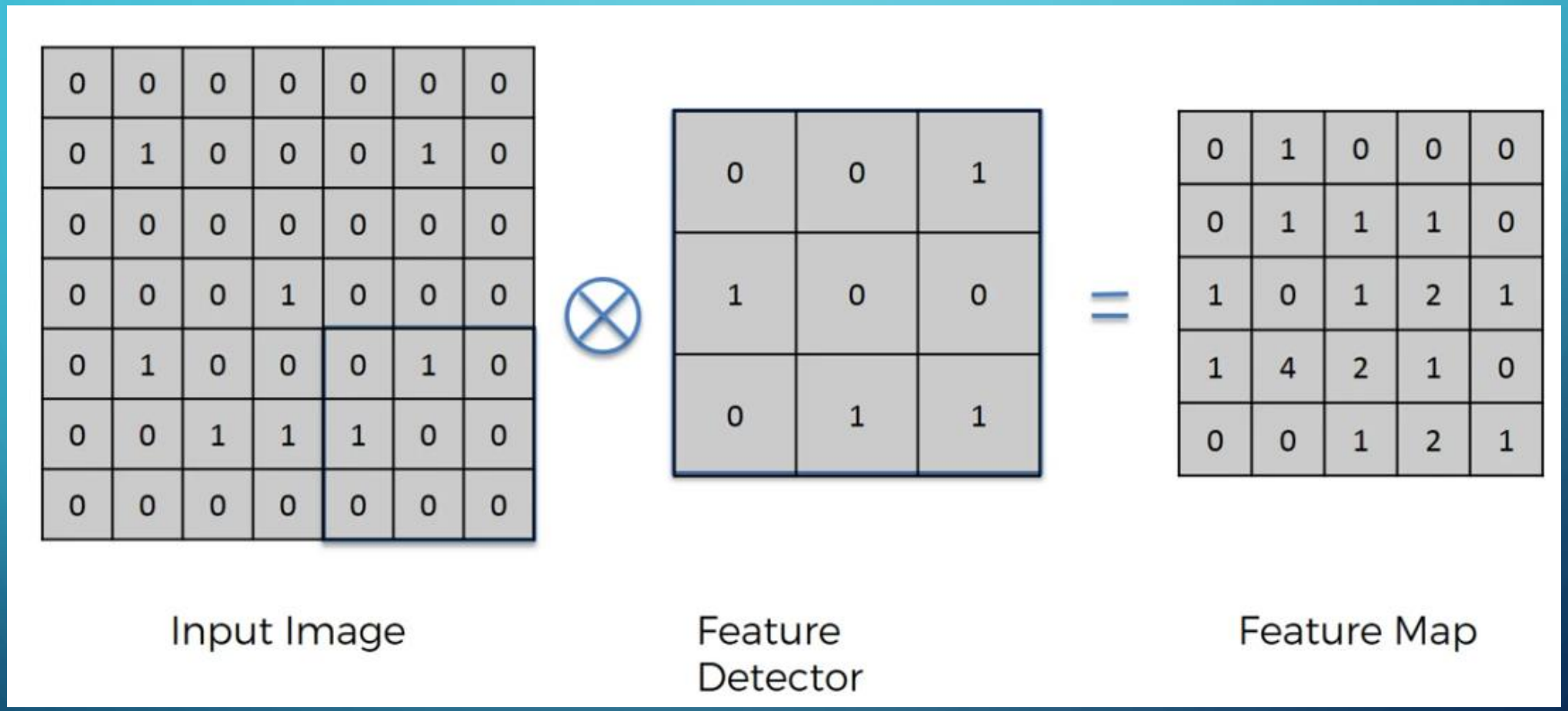
0	30	30	0
0	30	30	0
0	30	30	0
0	30	30	0

4x4

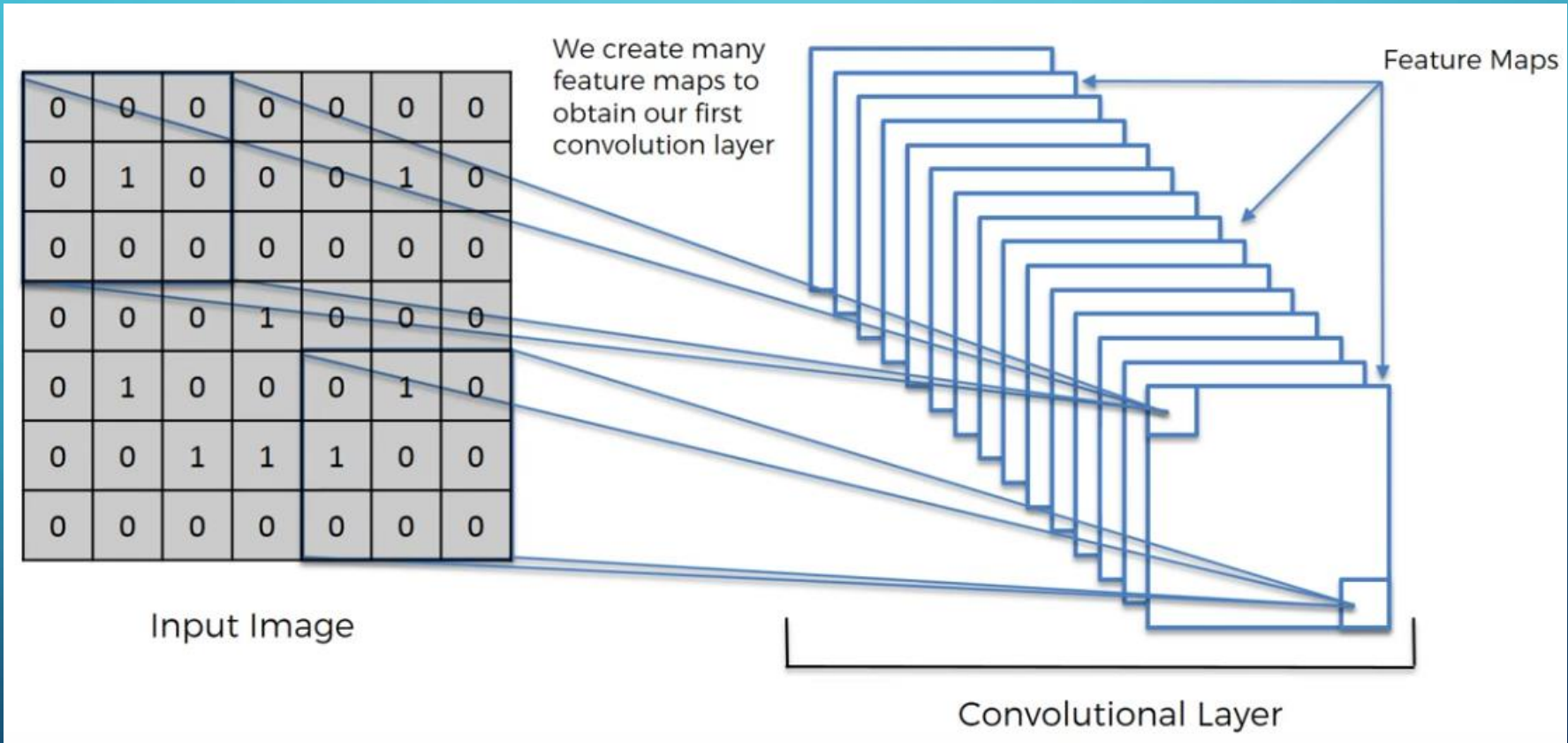


*

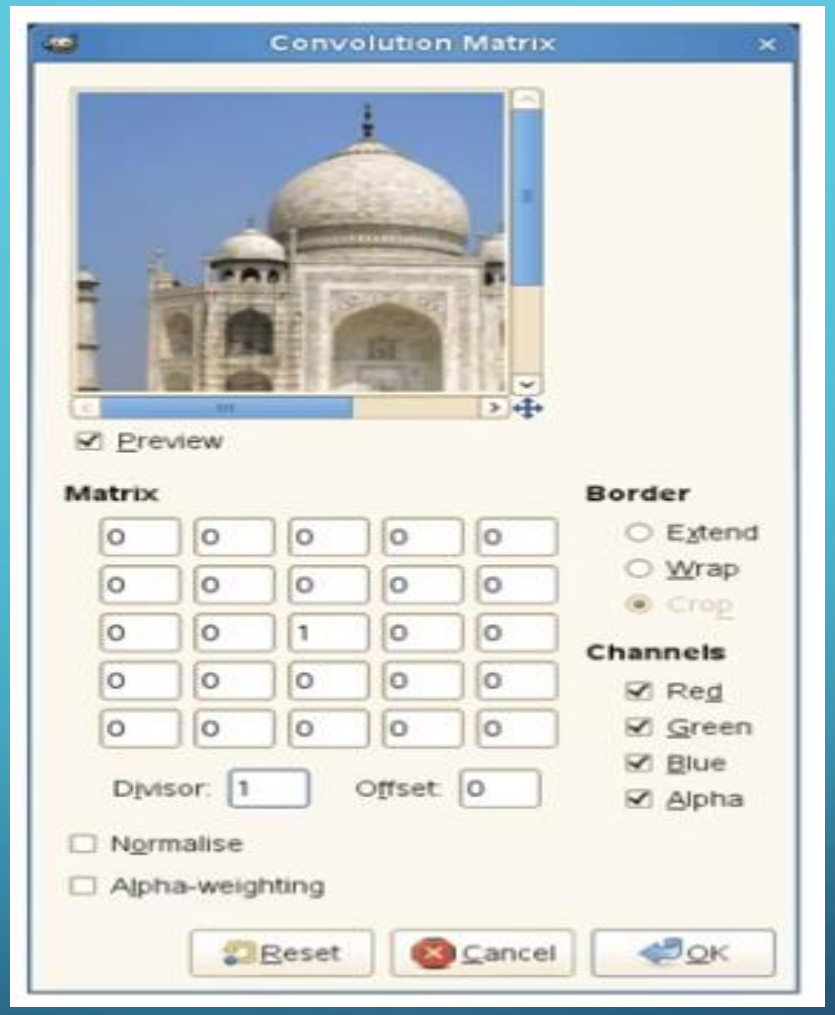
- We learnt the concept of convolution and filters
- Lets see some other terms


0 0 0 0 0 0 0
0 1 0 0 0 1 0
0 0 0 0 0 0 0
0 0 0 1 0 0 0
0 1 0 0 0 1 0
0 0 1 1 1 0 0
0 0 0 0 0 0 0
0 0 1
1 0 0
0 1 1
0 1 0 0 0
0 1 1 1 0
1 0 1 2 1
1 4 2 1 0
0 0 1 2 1
Input Image
Feature Detector
Feature Map

MULTIPLE FILTERS/FEATURE MAPS



INTUITION


Convolution Matrix
Preview
Matrix
0 0 0 0 0
0 0 0 0 0
0 0 1 0 0
0 0 0 0 0
0 0 0 0 0
Divisor: 1
Offset: 0
Normalise
Alpha-weighting
Border
Extend
Wrap
Crop
Channels
Red
Green
Blue
Alpha
Reset
Cancel
OK

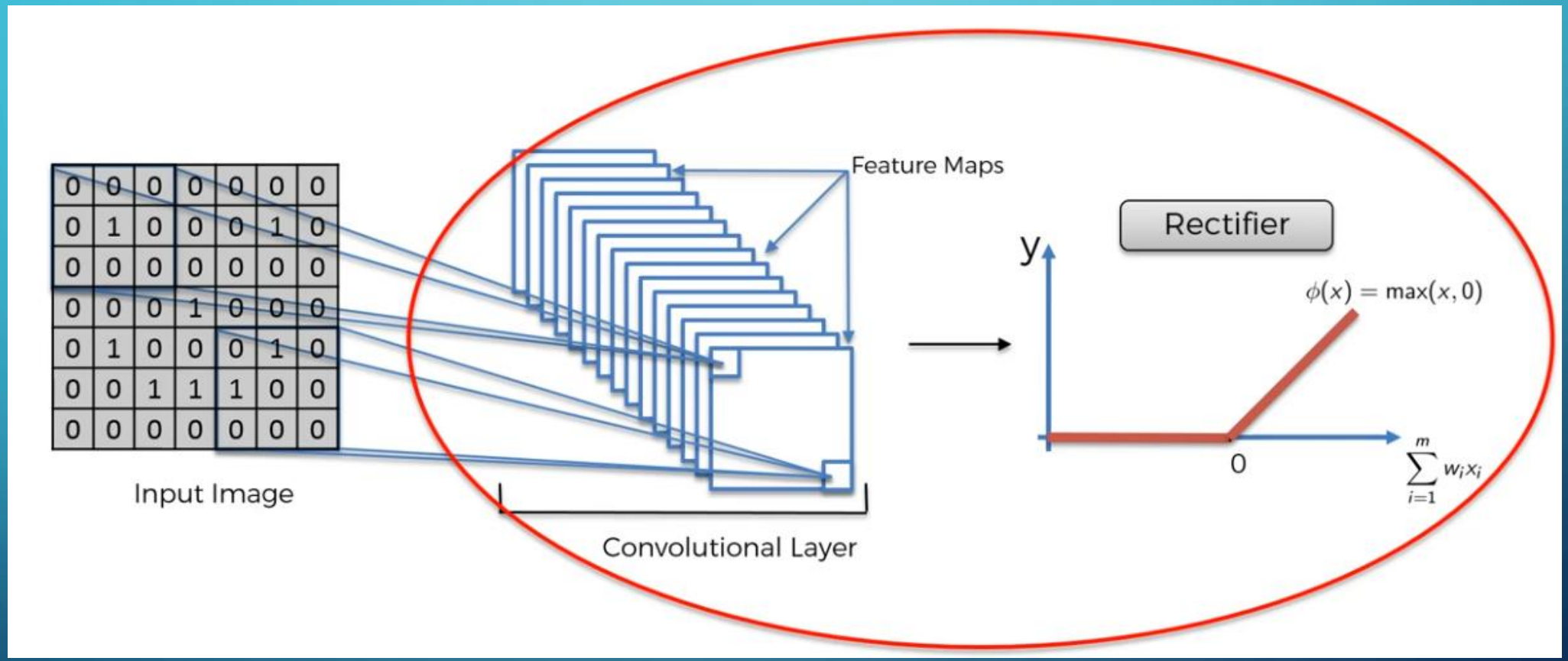
Sharpen:

0	0	0	0	0
0	0	-1	0	0
0	-1	5	-1	0
0	0	-1	0	0
0	0	0	0	0

CNN DECIDES THE FILTER VALUES

- Size of the filter gives the number of parameters required to be learnt by that filter

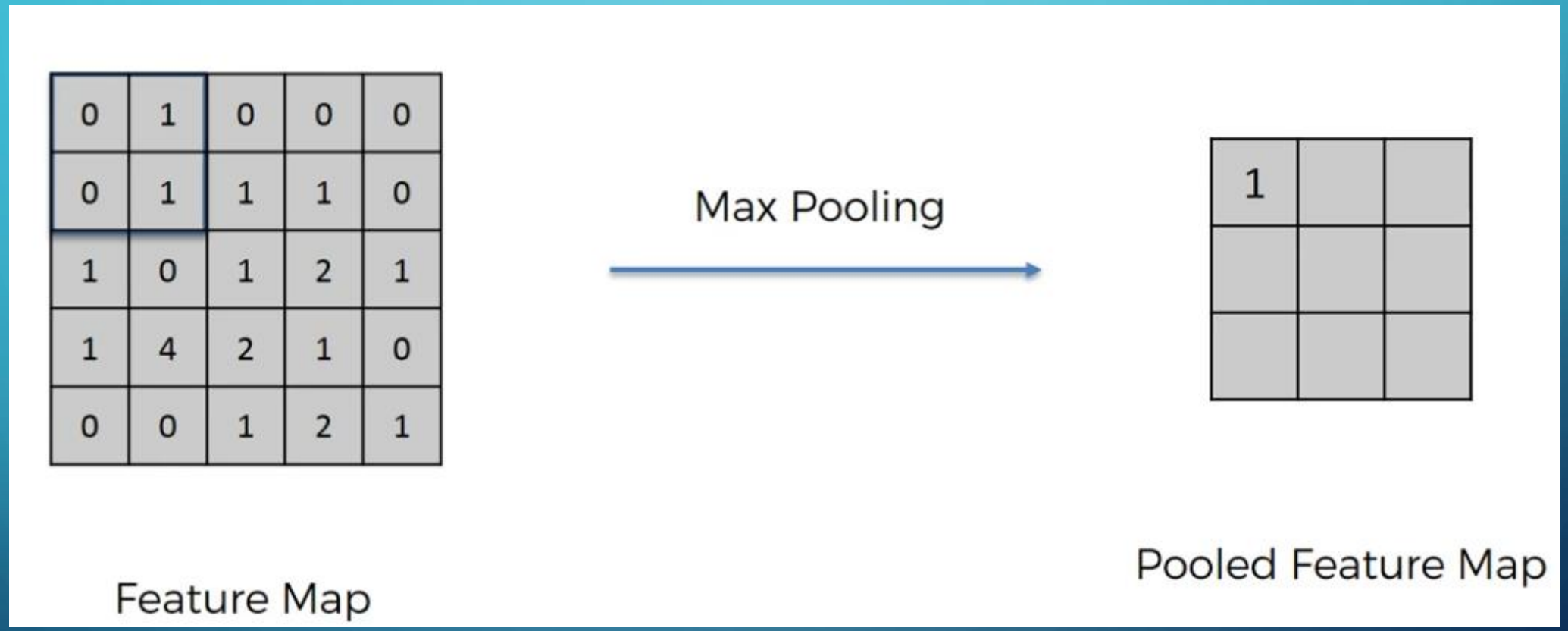
LETS APPLY THE ACTIVATION: RELU

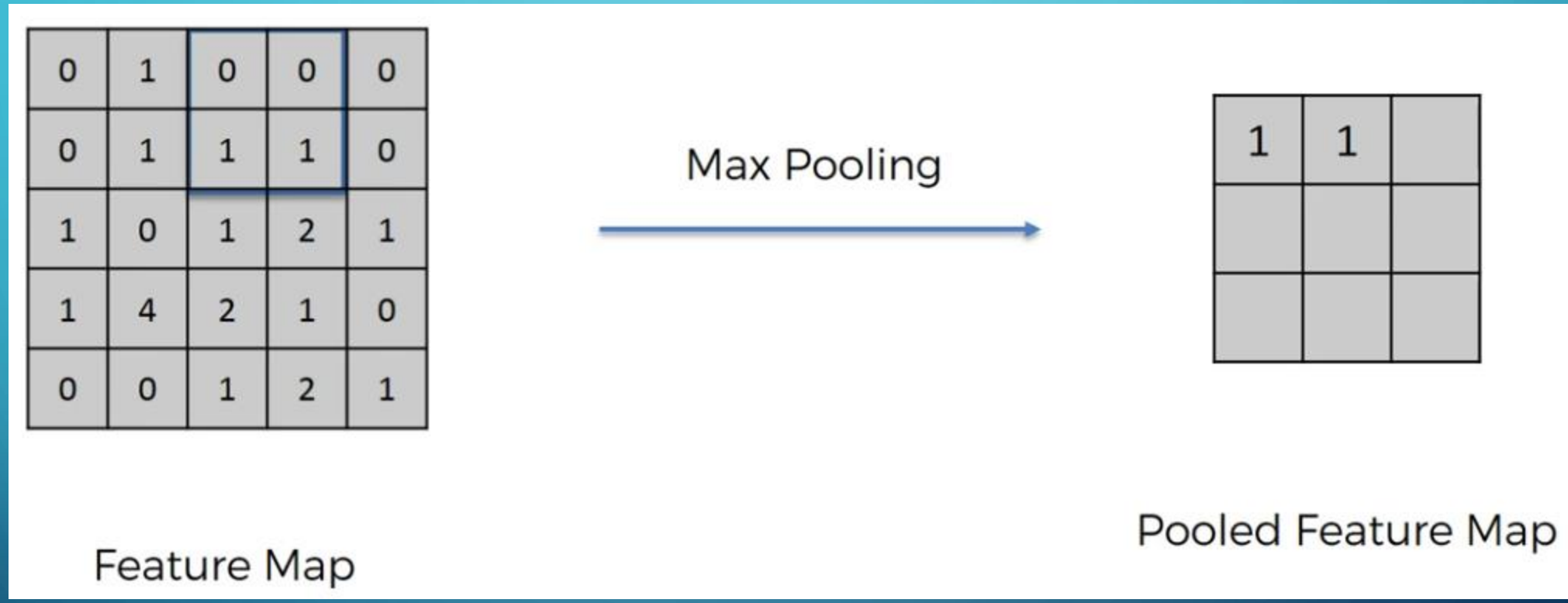

0 0 0 0 0 0 0
0 1 0 0 0 1 0
0 0 0 0 0 0 0
0 0 0 1 0 0 0
0 1 0 0 0 1 0
0 0 1 1 1 0 0
0 0 0 0 0 0 0
Input Image
Feature Maps
Convolutional Layer
Rectifier
y
$\phi(x) = \max(x, 0)$
0
$\sum_{i=1}^{m} w_i x_i$

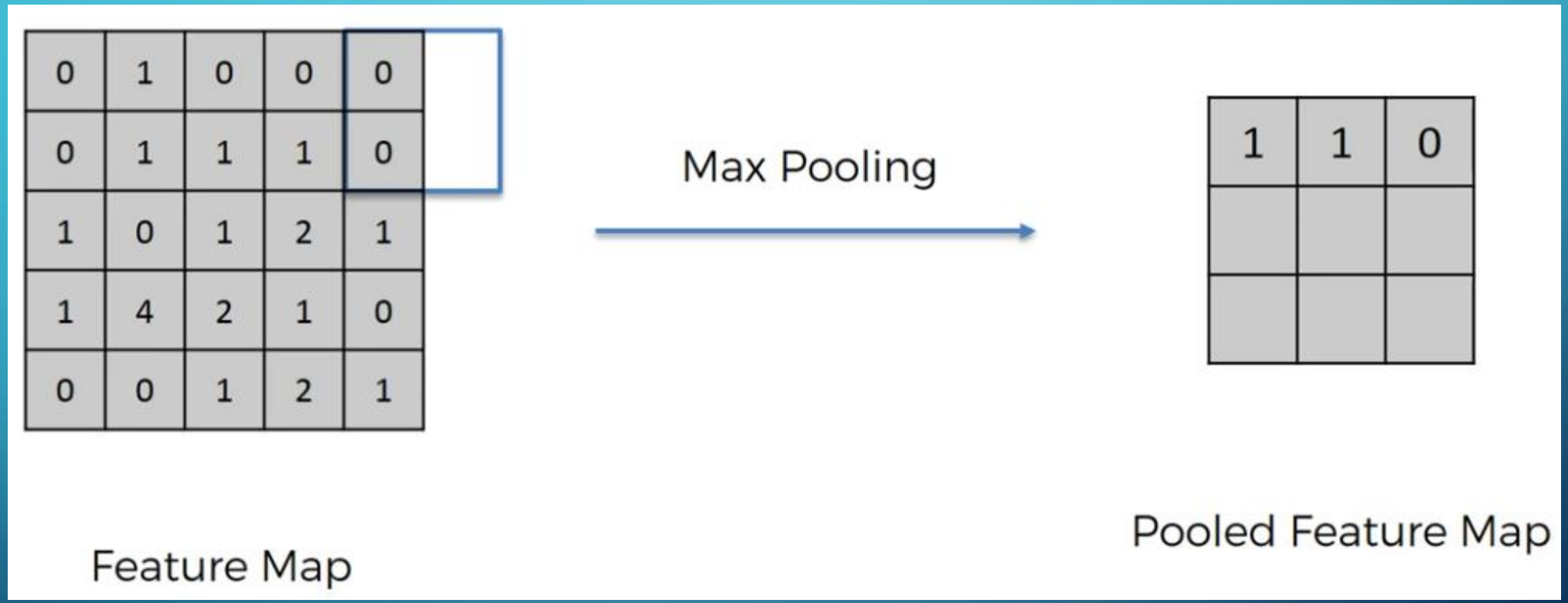
- Basic idea is to introduce non-linearity
- No extra parameters required

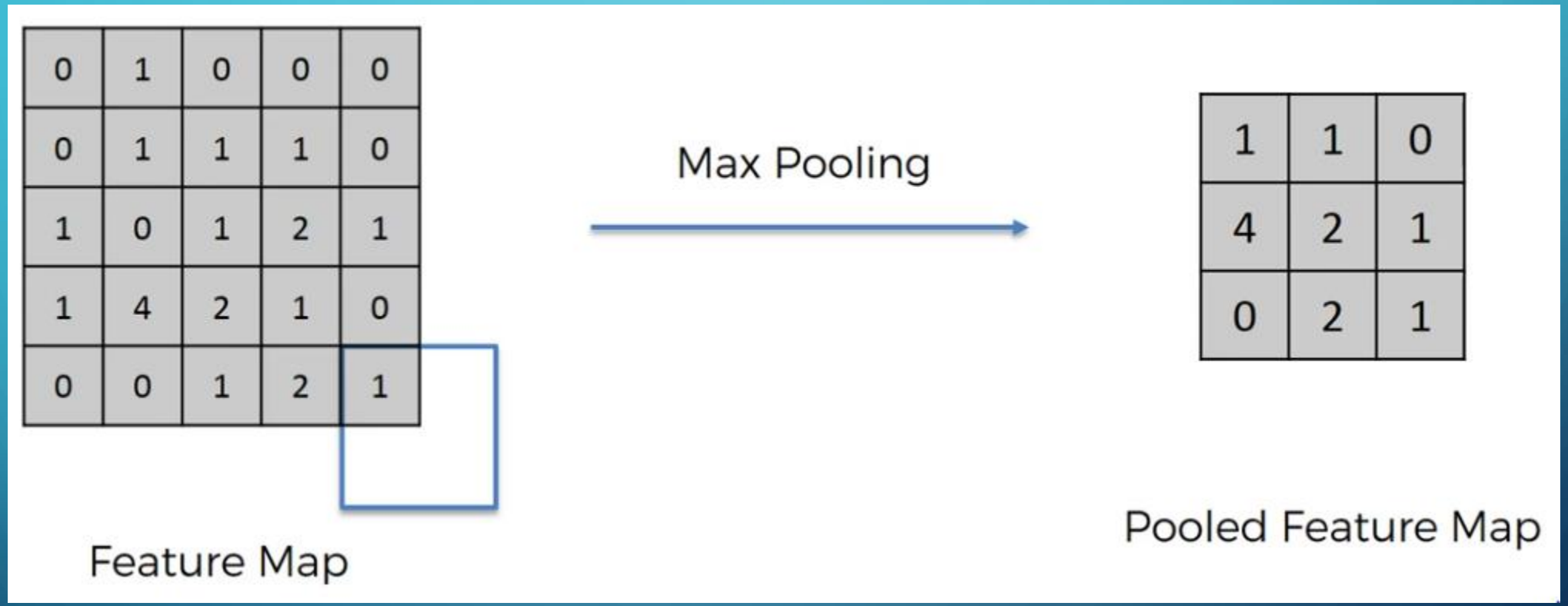
MAX POOLING

- Helps to introduce spatial invariance


0 1 0 0 0
0 1 1 1 0
1 0 1 2 1
1 4 2 1 0
0 0 1 2 1
Max Pooling
1
Feature Map
Pooled Feature Map


0 1 0 0 0
0 1 1 1 0
1 0 1 2 1
1 4 2 1 0
0 0 1 2 1
Max Pooling
1 1
Feature Map
Pooled Feature Map


0 1 0 0 0
0 1 1 1 0
1 0 1 2 1
1 4 2 1 0
0 0 1 2 1
Max Pooling
1 1 0
Feature Map
Pooled Feature Map


0 1 0 0 0
0 1 1 1 0
1 0 1 2 1
1 4 2 1 0
0 0 1 2 1
Feature Map
Max Pooling
1 1 0
4 2 1
0 2 1
Pooled Feature Map

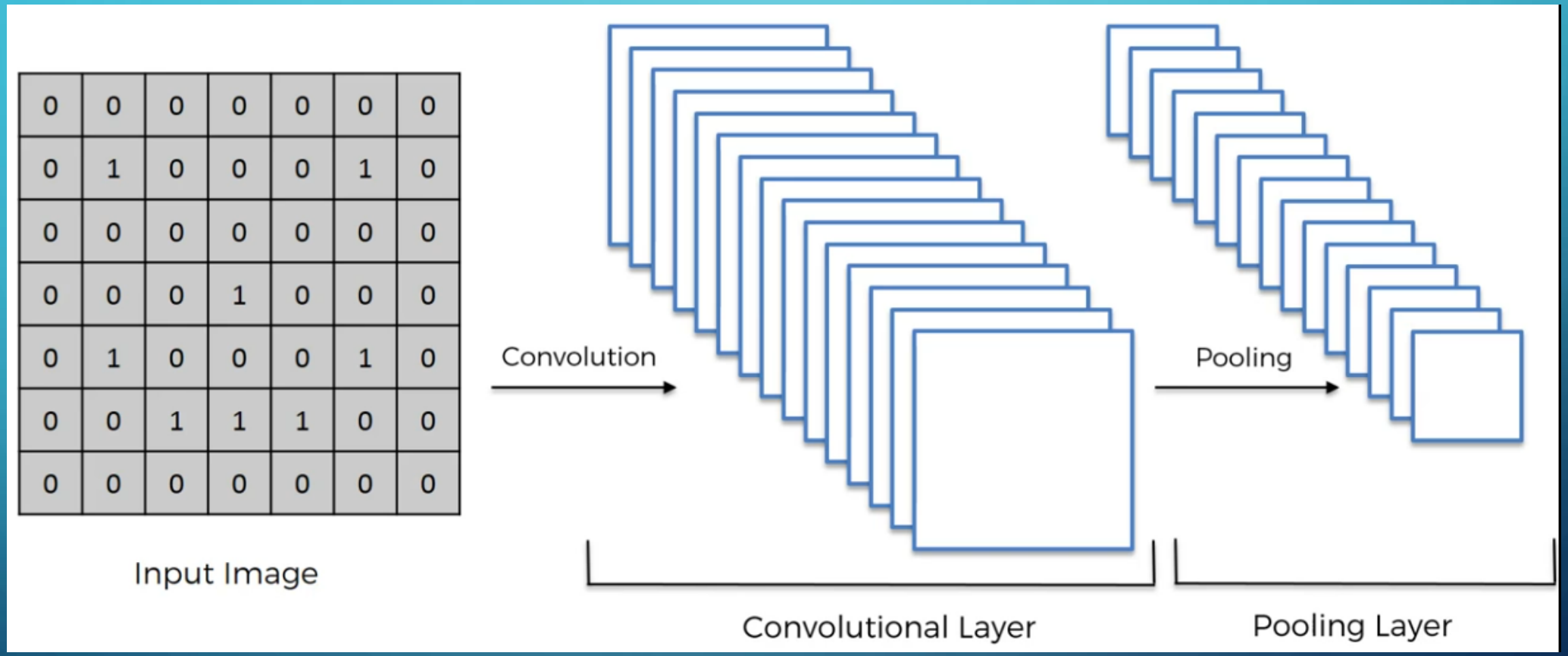
Additional Reading:

Evaluation of Pooling Operations in Convolutional Architectures for Object Recognition

By Dominik Scherer et al. (2010)

Link:

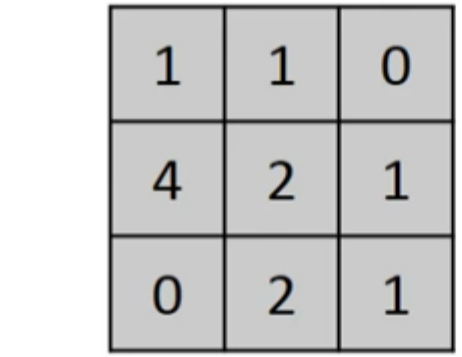
http://ais.uni-bonn.de/papers/icann2010_maxpool.pdf


0 0 0 0 0 0 0
0 1 0 0 0 1 0
0 0 0 0 0 0 0
0 0 0 1 0 0 0
0 1 0 0 0 1 0
0 0 1 1 1 0 0
0 0 0 0 0 0 0
Input Image
Convolution
Pooling
Convolutional Layer
Pooling Layer

DEMO

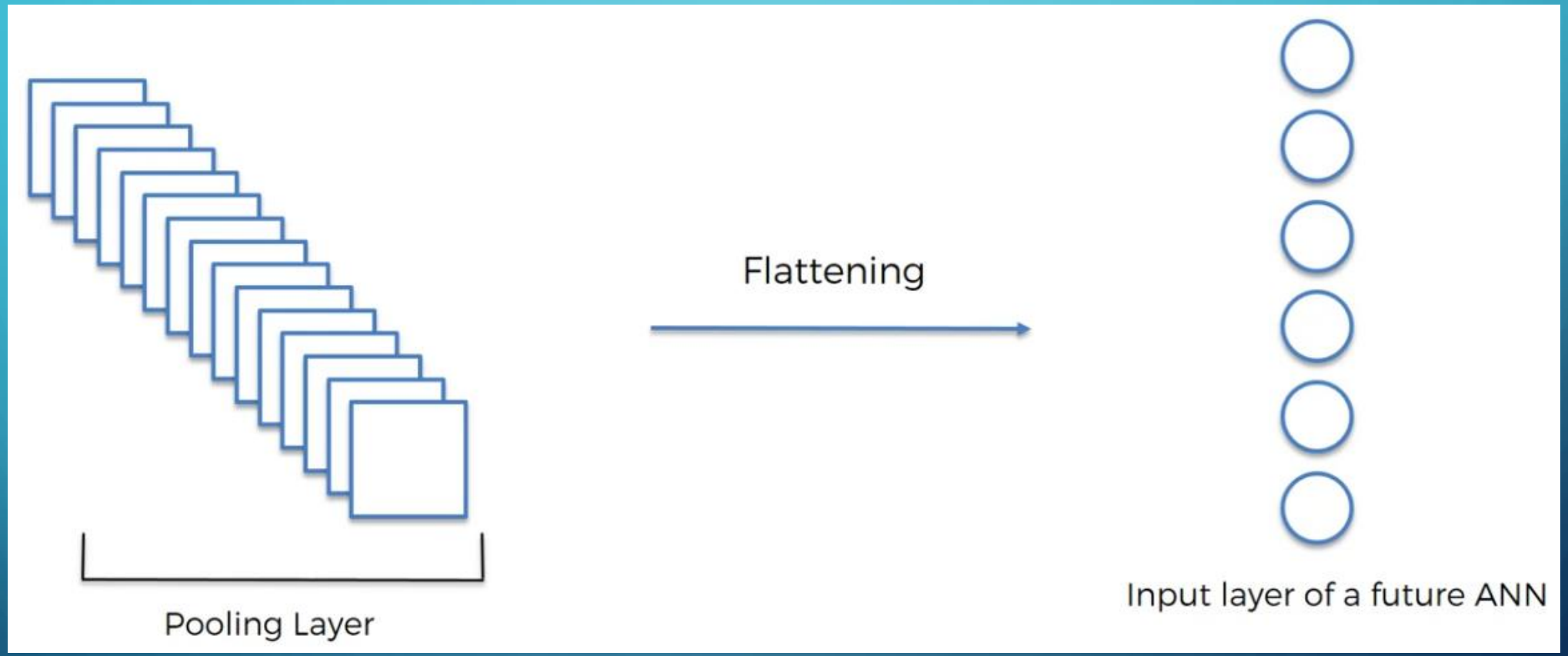
- Open http://scs.ryerson.ca/~aharley/vis/conv/flat.html

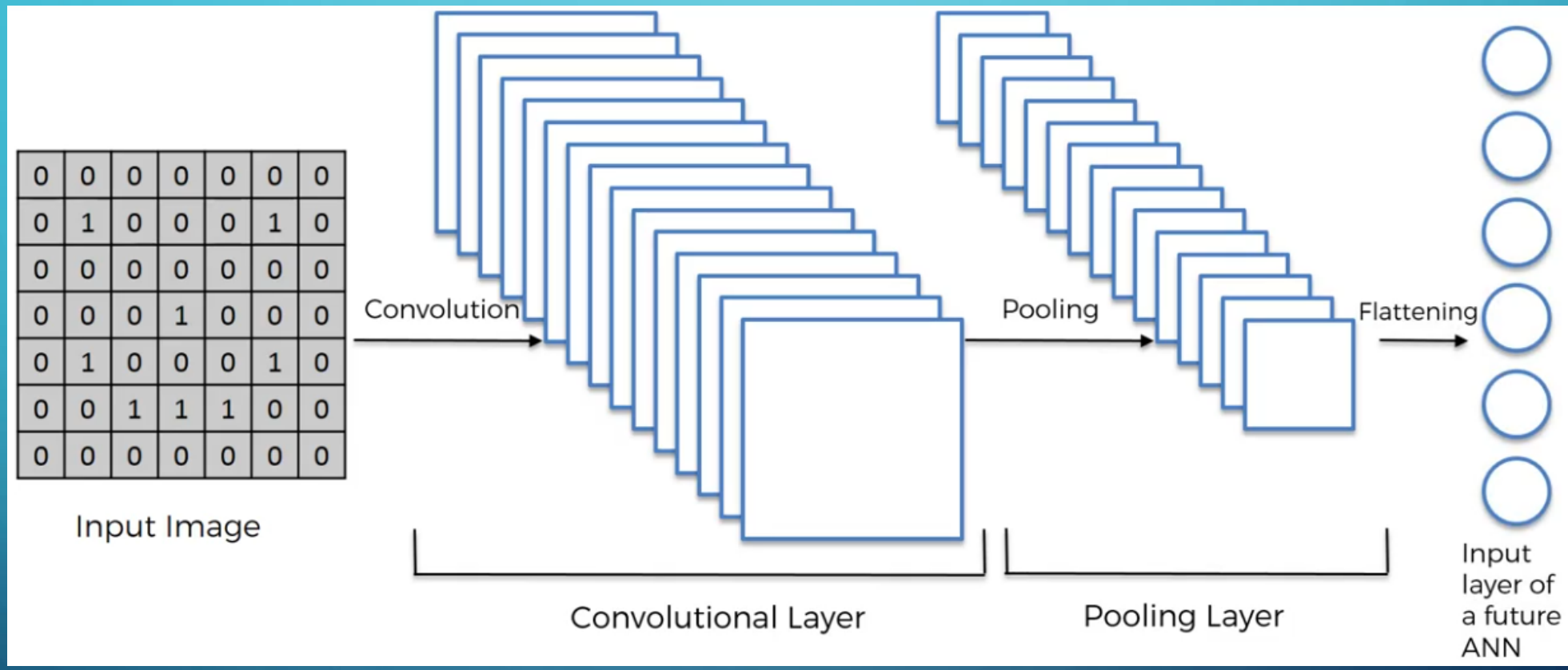
FLATTENING



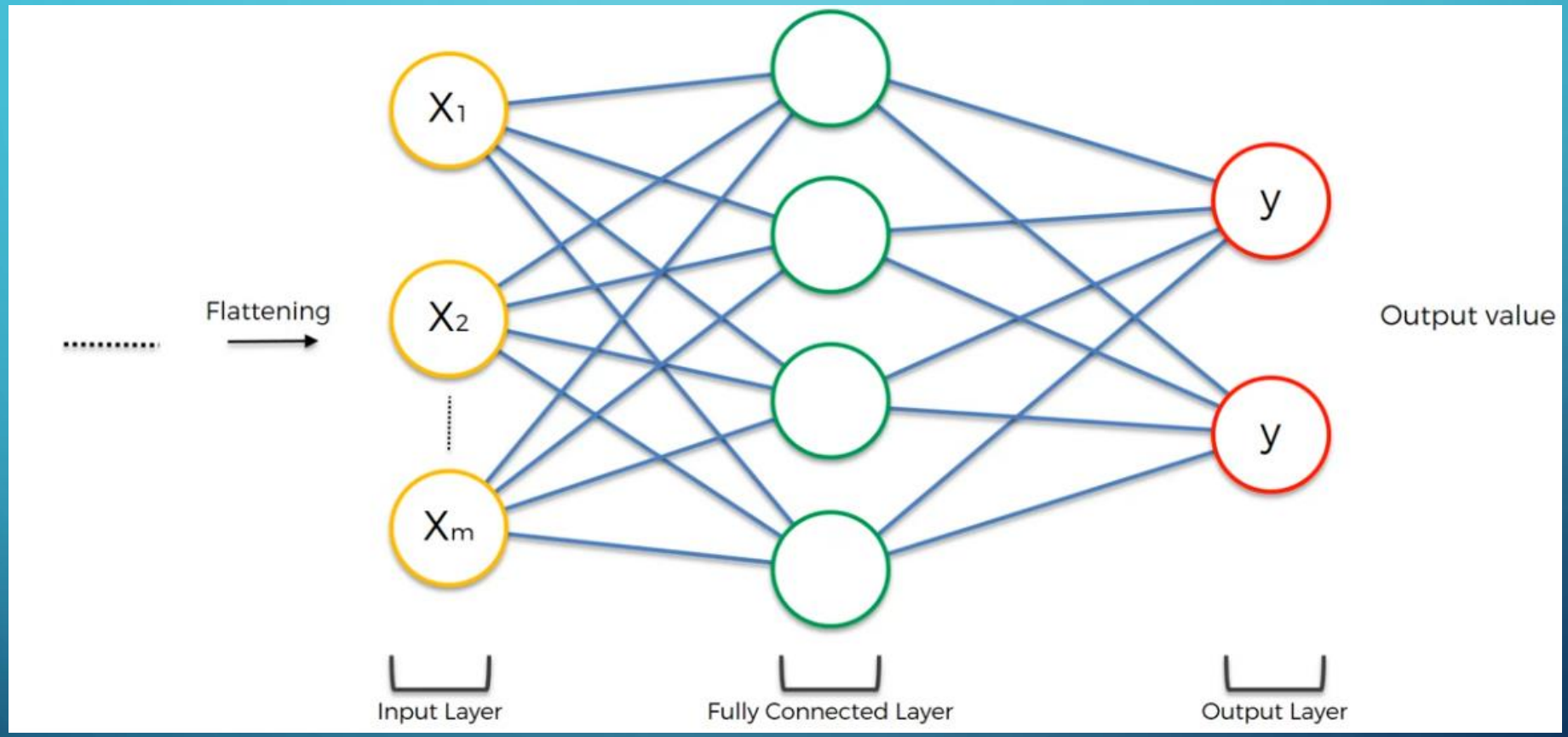
Pooled Feature Map

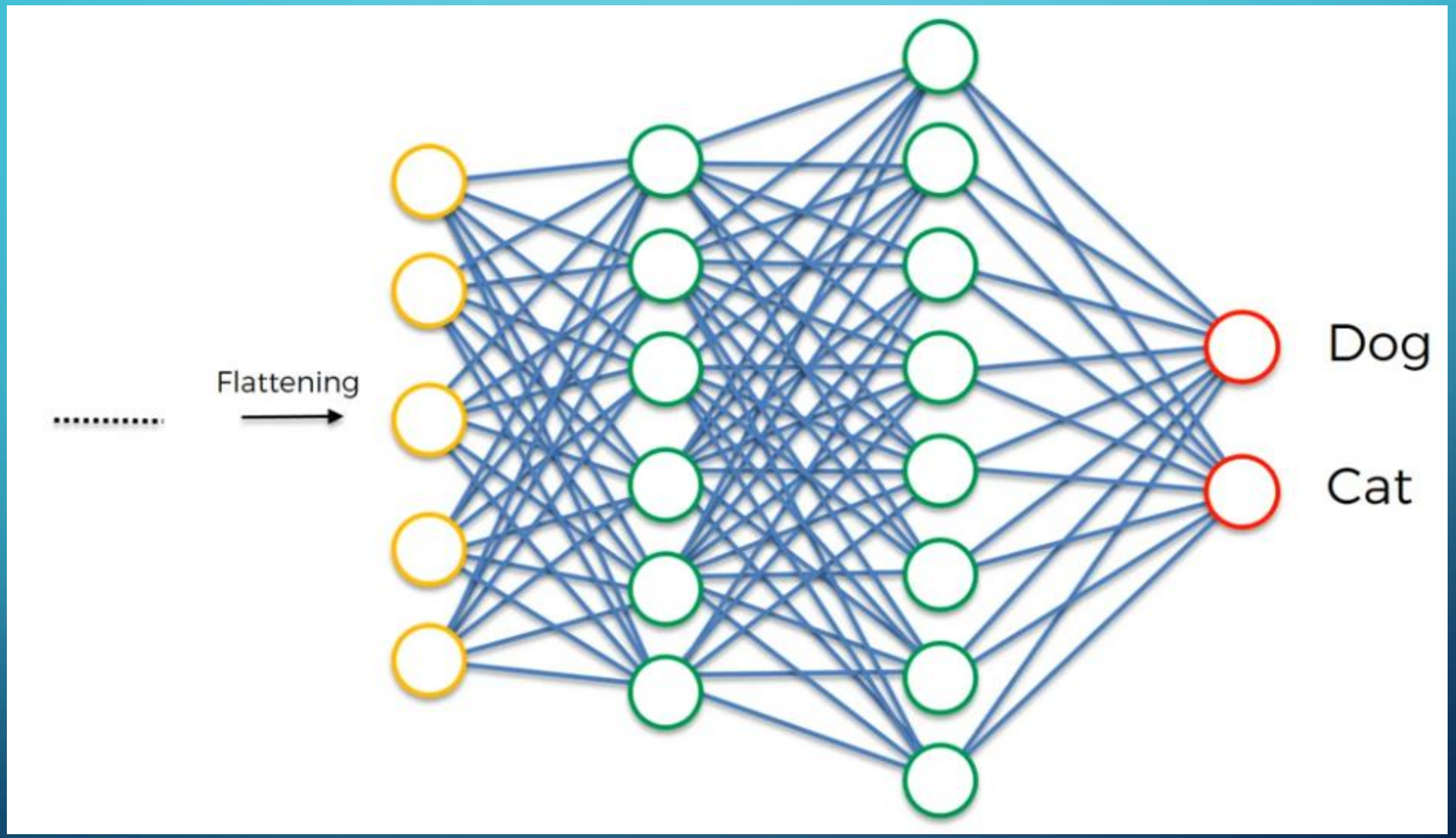
Flattening


Flattening
Pooling Layer
Input layer of a future ANN

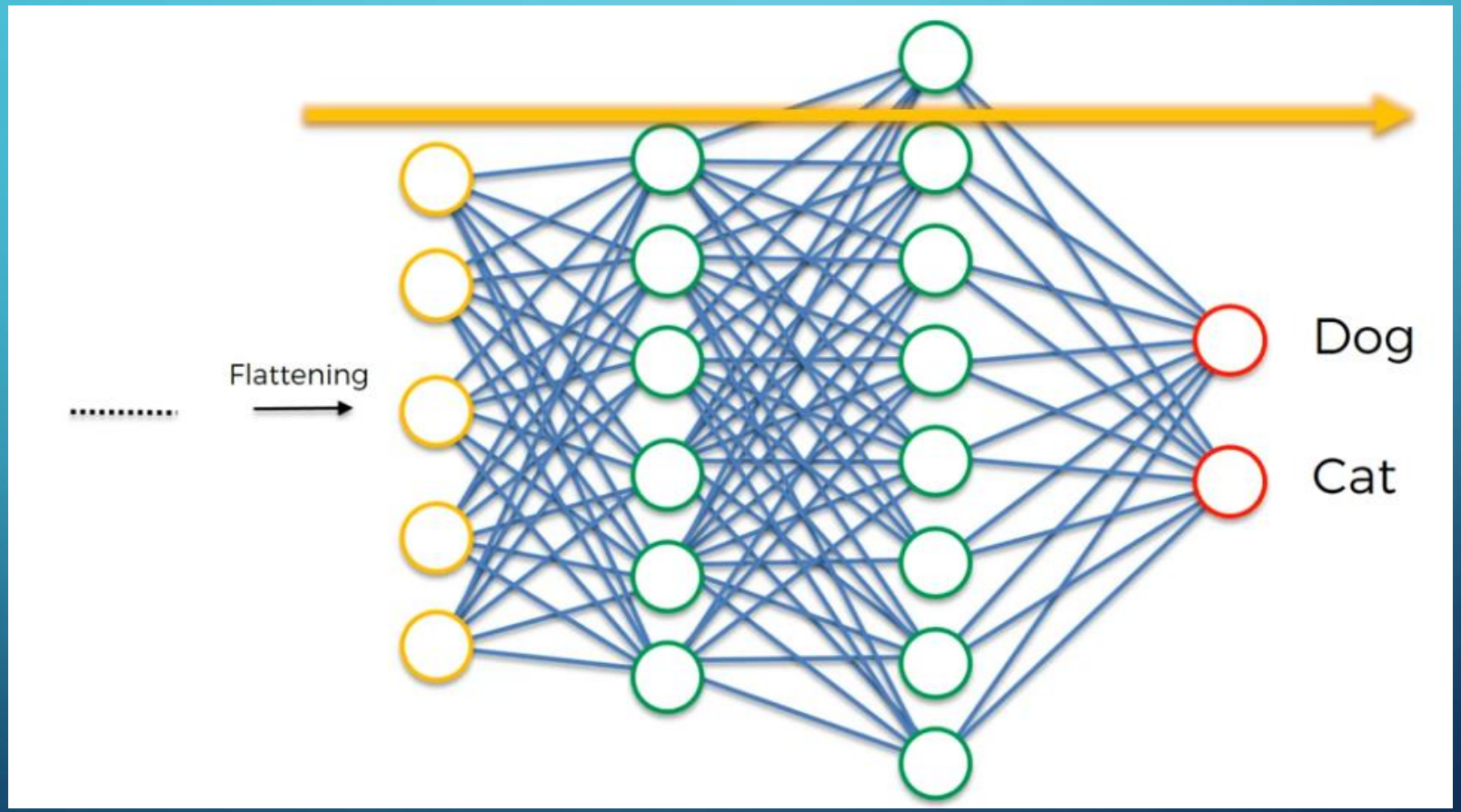

0 0 0 0 0 0 0
0 1 0 0 0 1 0
0 0 0 0 0 0 0
0 0 0 1 0 0 0
0 1 0 0 0 1 0
0 0 1 1 1 0 0
0 0 0 0 0 0 0
Input Image
Convolution
Pooling
Flattening
Convolutional Layer
Pooling Layer
Input layer of a future ANN

FULL CONNECTION

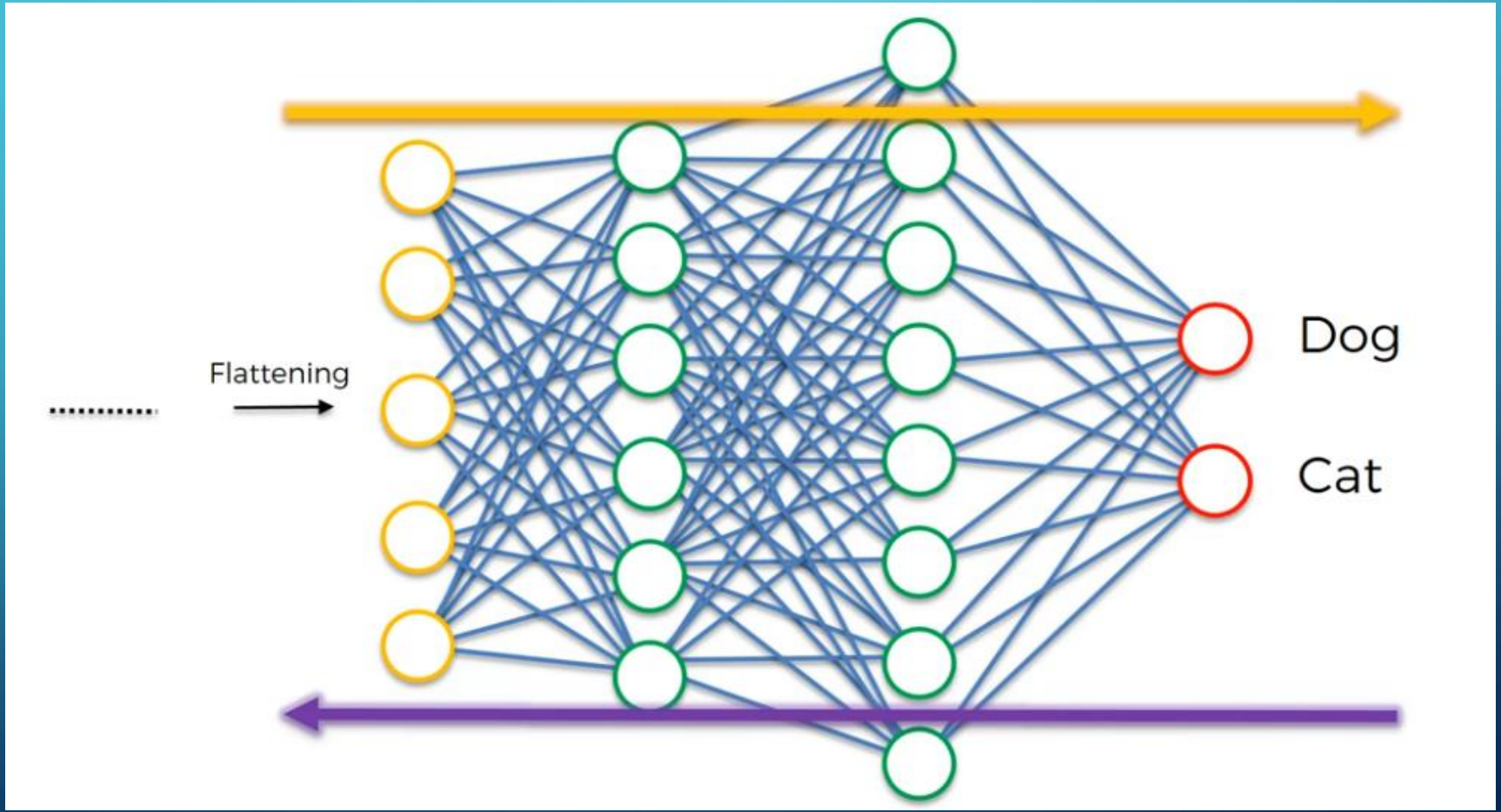

Flattening
X1
X2
Xm
y
y
Output value
Input Layer
Fully Connected Layer
Output Layer


Flattening
Dog
Cat

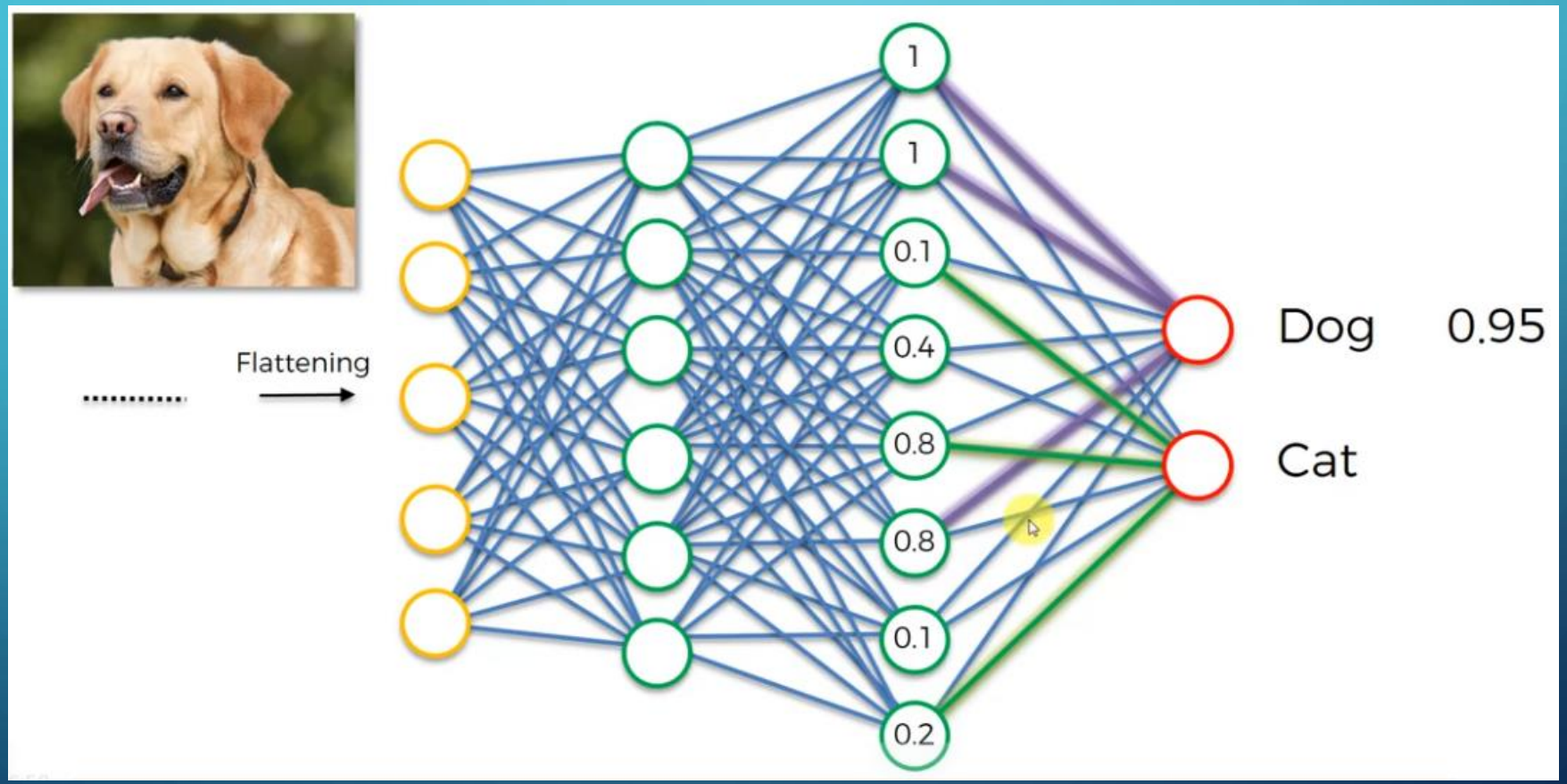
FORWARD PROPAGATION



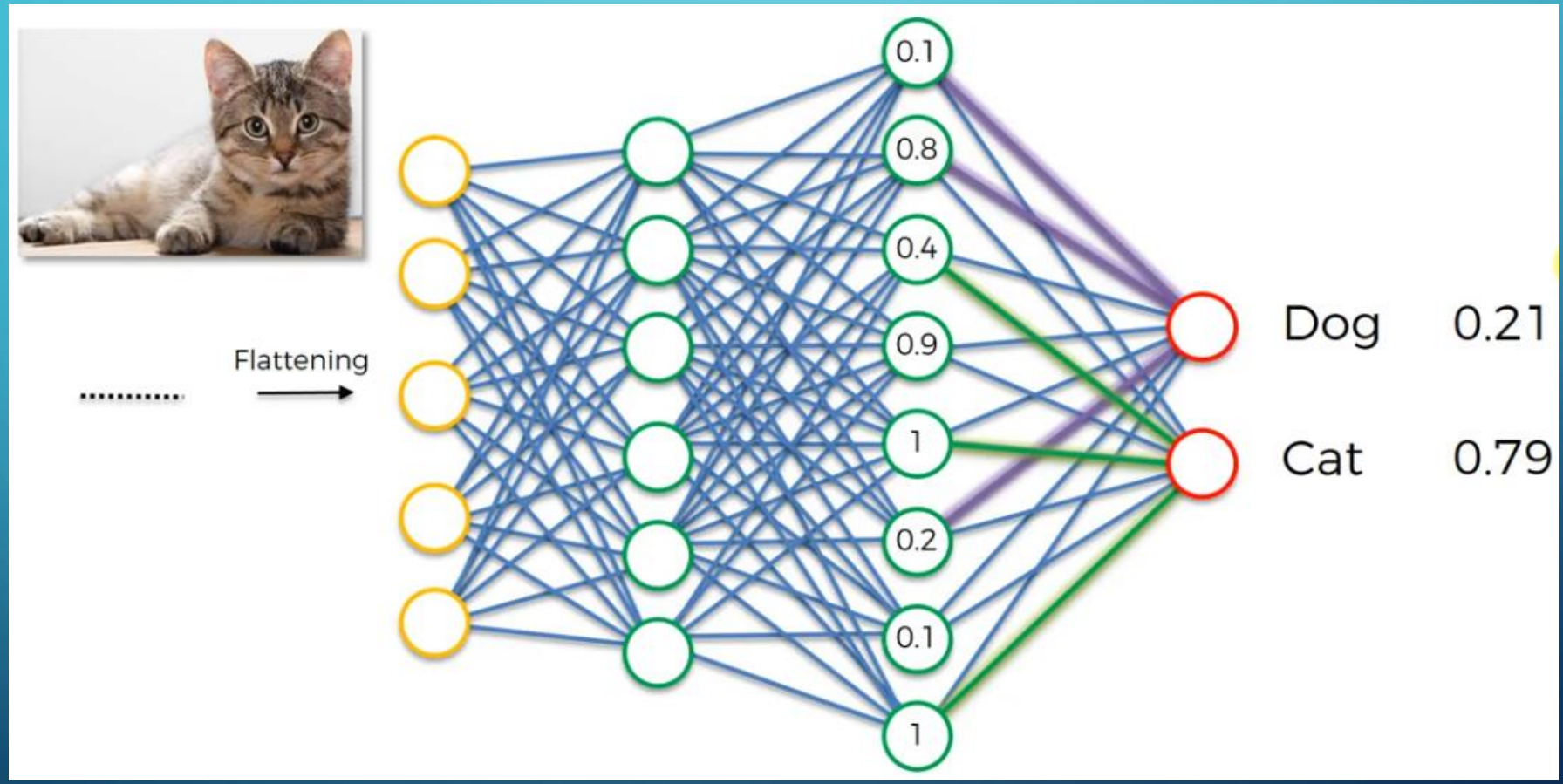
BACKWARD PROPAGATION



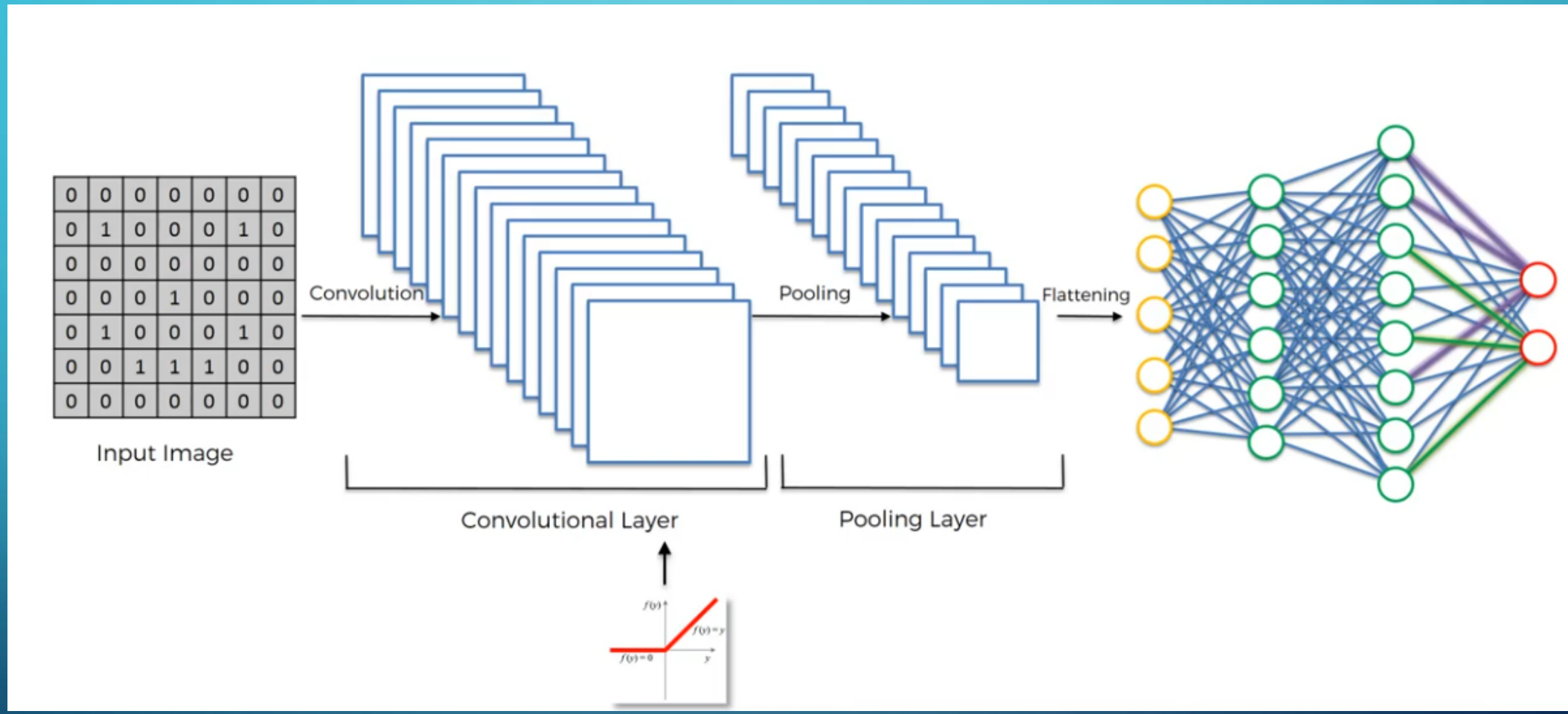
TESTING

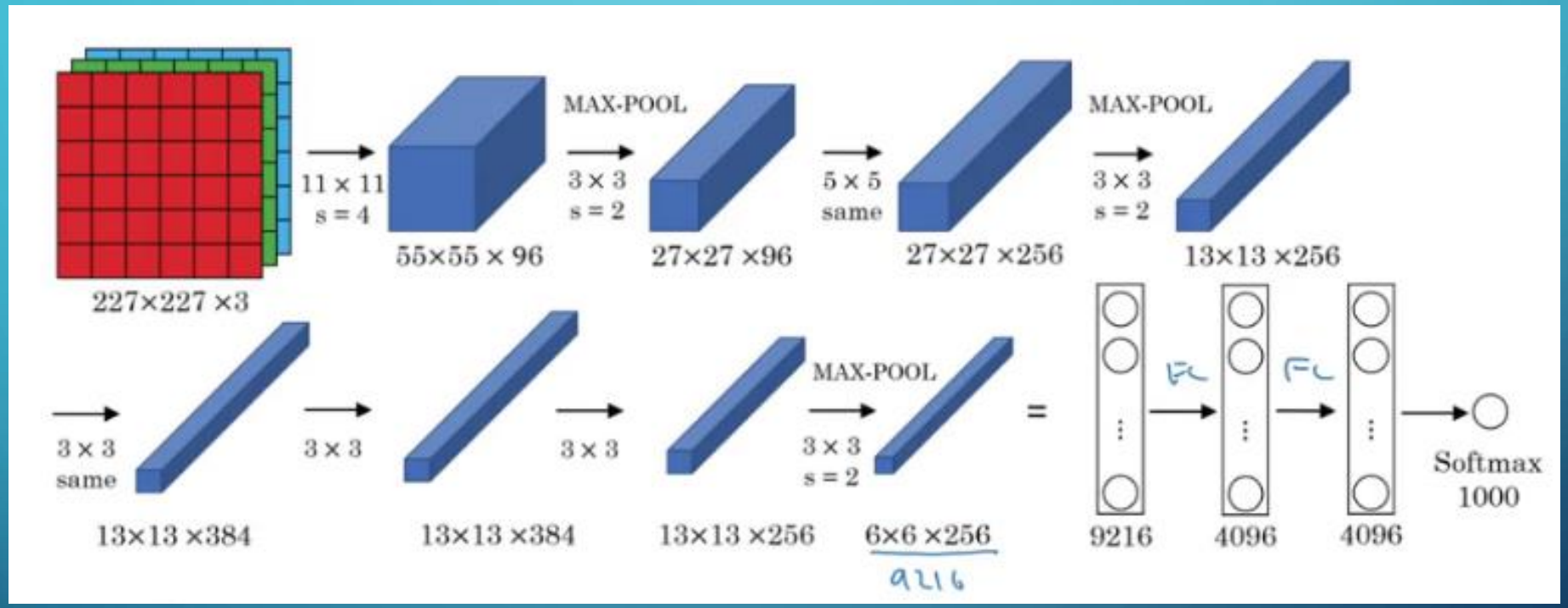


TESTING



SUMMARY

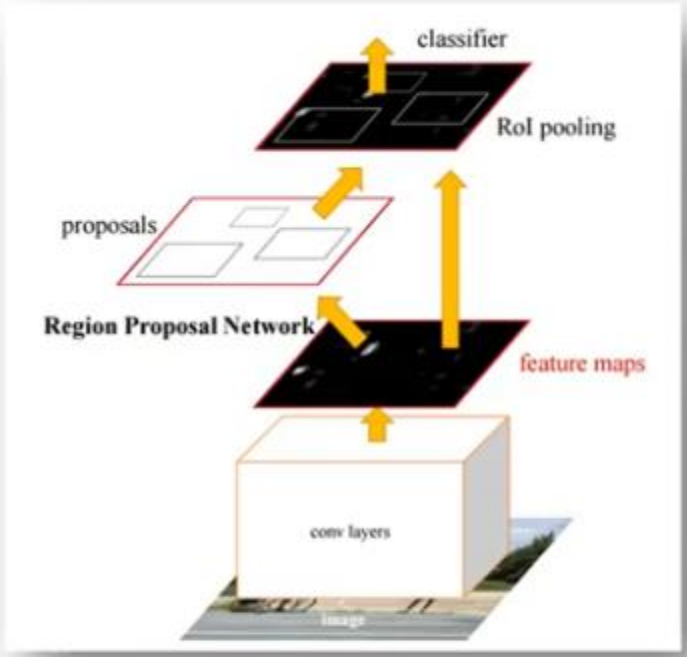

0 0 0 0 0 0 0
0 1 0 0 0 1 0
0 0 0 0 0 0 0
0 0 0 1 0 0 0
0 1 0 0 0 1 0
0 0 1 1 1 0 0
0 0 0 0 0 0 0
Input Image
Convolution
Pooling
Flattening
Convolutional Layer
Pooling Layer


227×227 ×3
11 × 11
s = 4
55×55 × 96
MAX-POOL
3 × 3
s = 2
27×27 ×96
5 × 5
same
27×27 ×256
MAX-POOL
3 × 3
s = 2
13×13 ×256
3 × 3
same
13×13 ×384
3 × 3
13×13 ×384
3 × 3
13×13 ×256
MAX-POOL
3 × 3
s = 2
6×6 ×256
9216
=
FC
FC
9216
4096
4096
Softmax
1000

Additional Reading:

The 9 Deep Learning Papers You Need To Know About (Understanding CNNs Part 3)

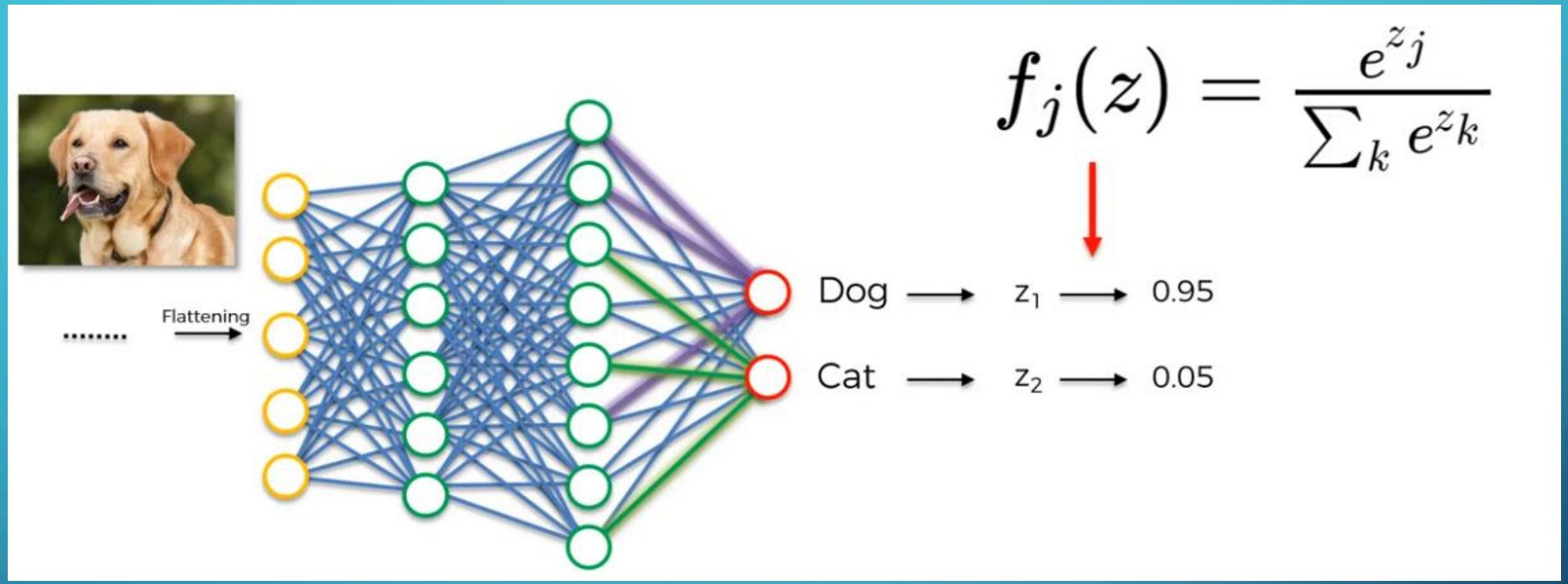
Adit Deshpande (2016)



Link:

https://adeshpande3.github.io/adeshpande3.github.io/The-9-Deep-Learning-Papers-You-Need-To-Know-About.html

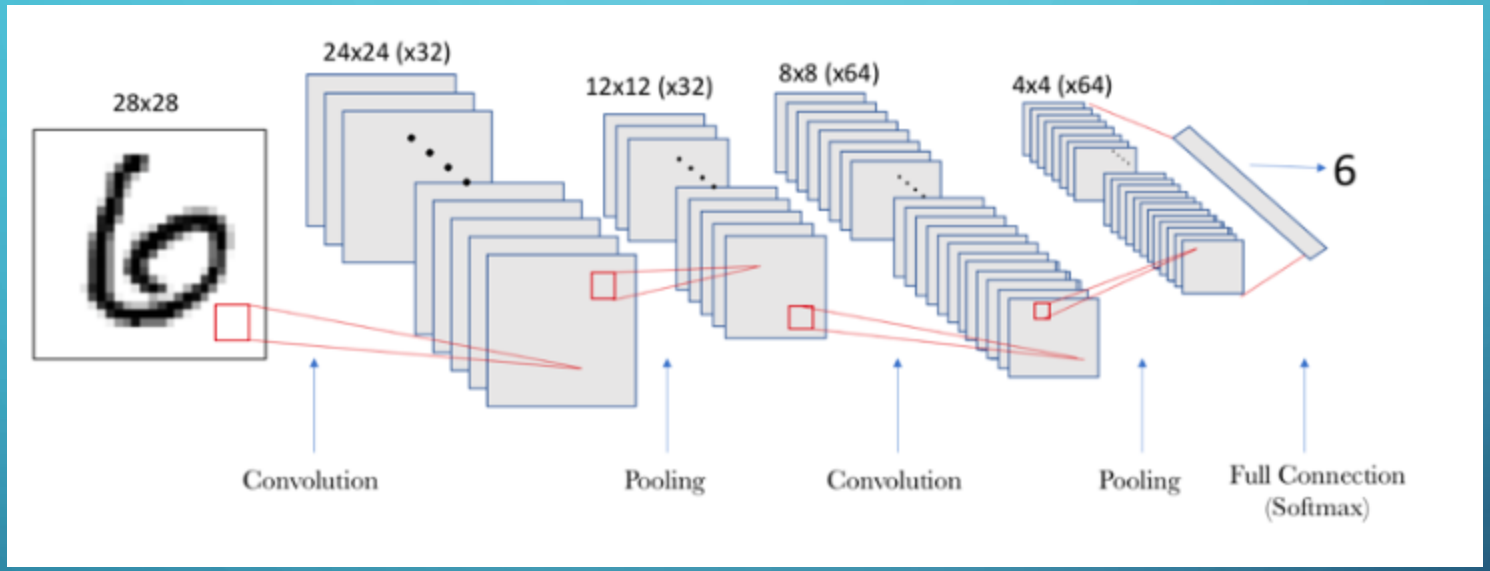
SOFTMAX



Here, f(z) is the softmax function

- Softmax function is the generalisation of the Logistic/Sigmoid function
- It squashes all the output categories between 0 and 1 such that their sum is equal to 1
- Hence, we can consider its output as the probablity
- We usually use softmax layer as the output layer in classification problems

LETS BUILD OUR OWN CNN!!


28x28
24x24 (x32)
12x12 (x32)
8x8 (x64)
4x4 (x64)
6
Convolution
Pooling
Convolution
Pooling
Full Connection
(Softmax)