

CSPB3202 Artificial Intelligence

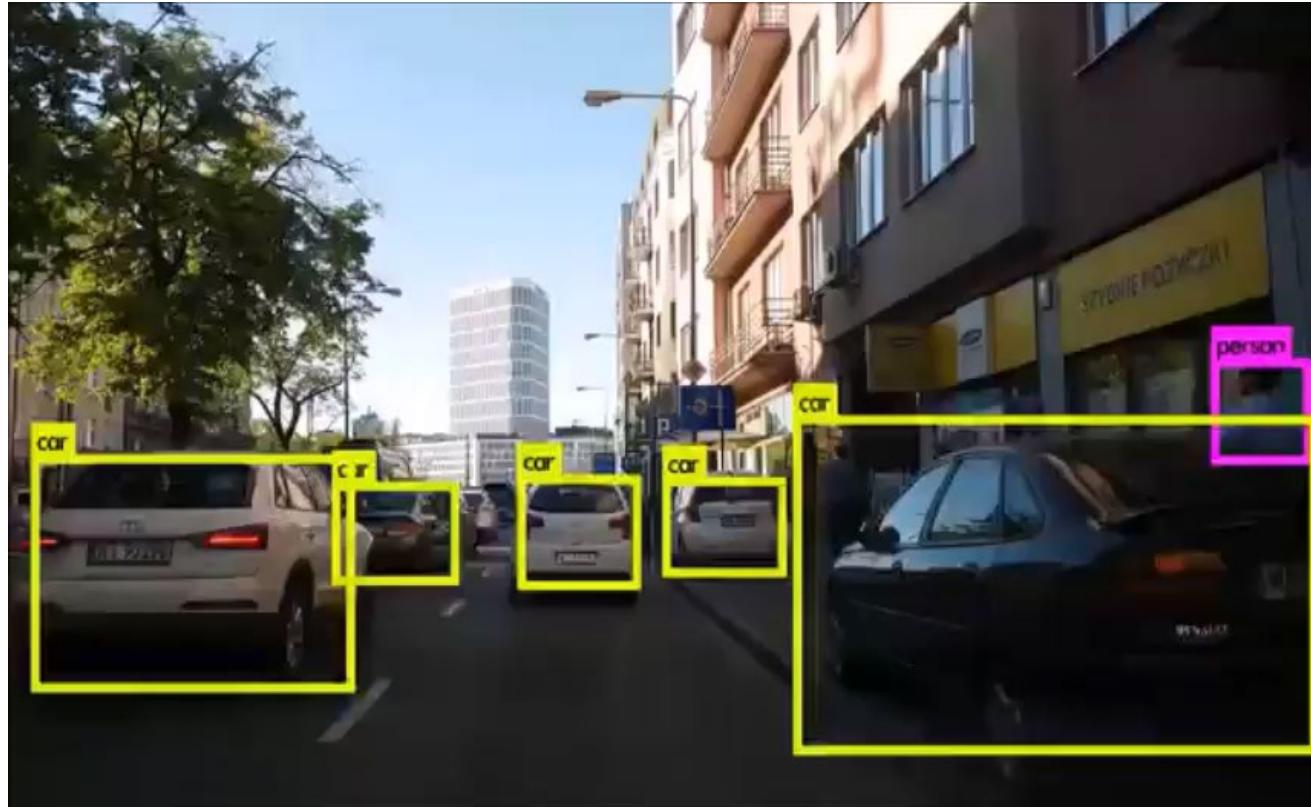
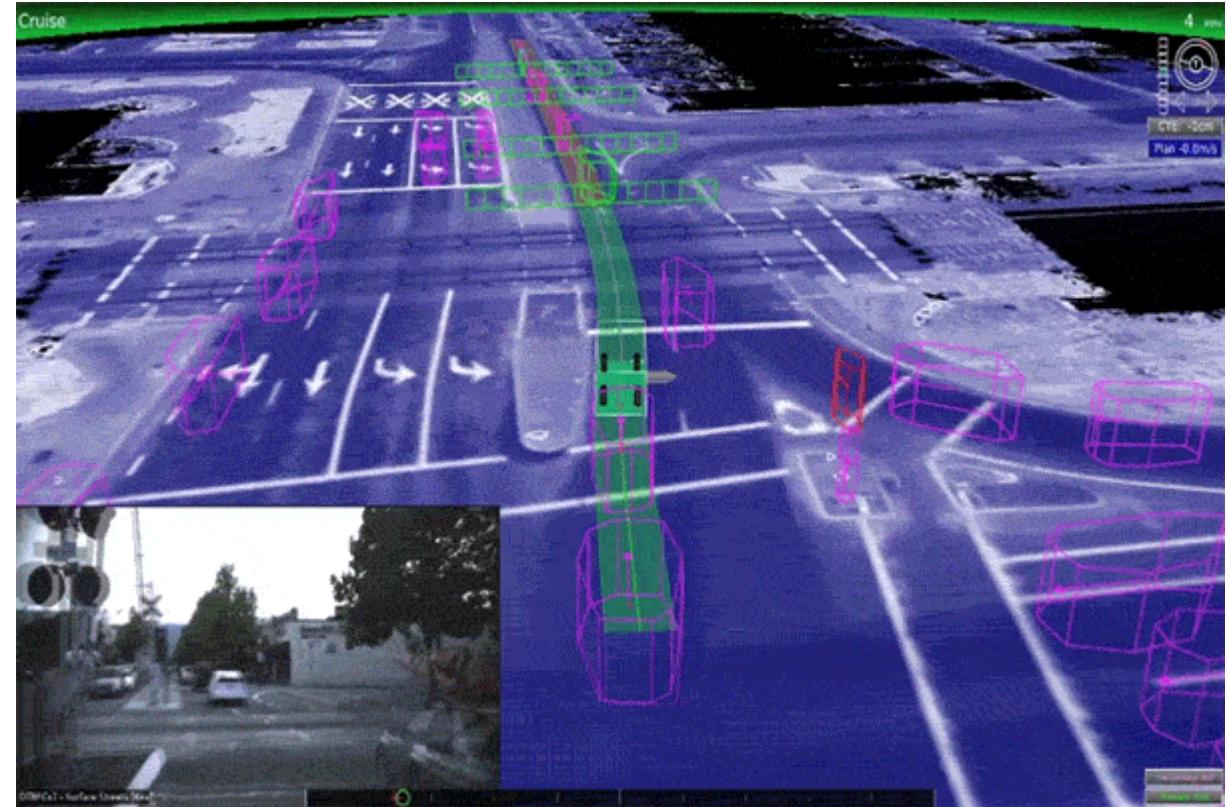
2020 Spring

# Convolutional Neural Networks (1)

Geena Kim



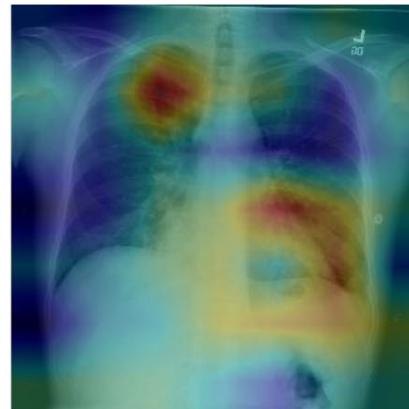
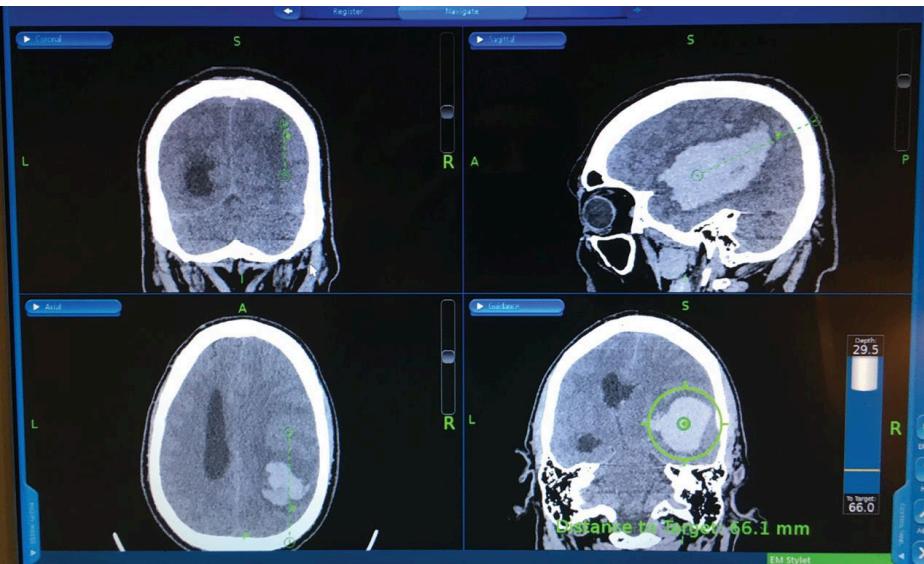
# Why Deep Learning?



<https://www.youtube.com/watch?v=EhcpGpFHCrw>

# Why Deep Learning?

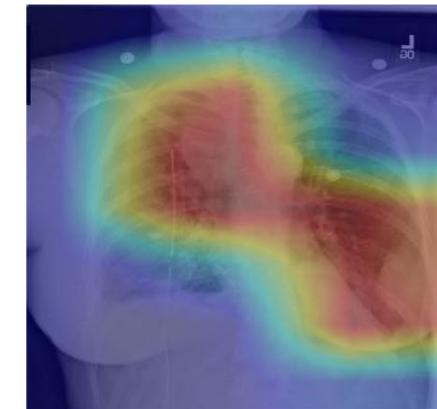
## Medical Imaging, AI in Medicine



(a) Patient with multifocal community acquired pneumonia. The model correctly detects the airspace disease in the left lower and right upper lobes to arrive at the pneumonia diagnosis.



(b) Patient with a left lung nodule. The model identifies the left lower lobe lung nodule and correctly classifies the pathology.



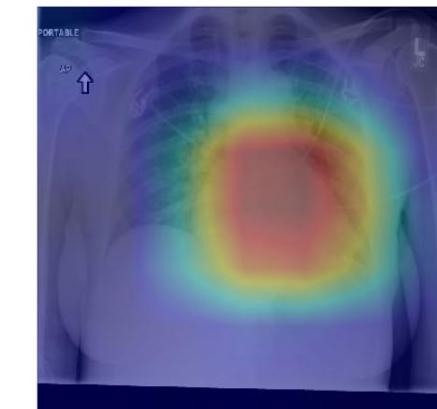
(c) Patient with primary lung malignancy and two large masses, one in the left lower lobe and one in the right upper lobe adjacent to the mediastinum. The model correctly identifies both masses in the X-ray.



(d) Patient with a right-sided pneumothorax and chest tube. The model detects the abnormal lung to correctly predict the presence of pneumothorax (collapsed lung).



(e) Patient with a large right pleural effusion (fluid in the pleural space). The model correctly labels the effusion and focuses on the right lower chest.



(f) Patient with congestive heart failure and cardiomegaly (enlarged heart). The model correctly identifies the enlarged cardiac silhouette.  
**P. Rajpurkar et al., arXiv:1711.05225v3**

# Why Deep Learning?



News Startups Mobile Gadgets Enterprise Social Europe

Trending Amazon Tesla Microsoft

eBay

shopping

Search

eCommerce

eCommerce

Popular Posts

***Shopping,  
e-commerce***

## eBay launches visual search tools that let you shop using photos from your phone or web

Posted Oct 26, 2017 by **Sarah Perez** (@sarahintampa)



Next Story



### Crunchbase

**eBay**

FOUNDED  
1995

#### OVERVIEW

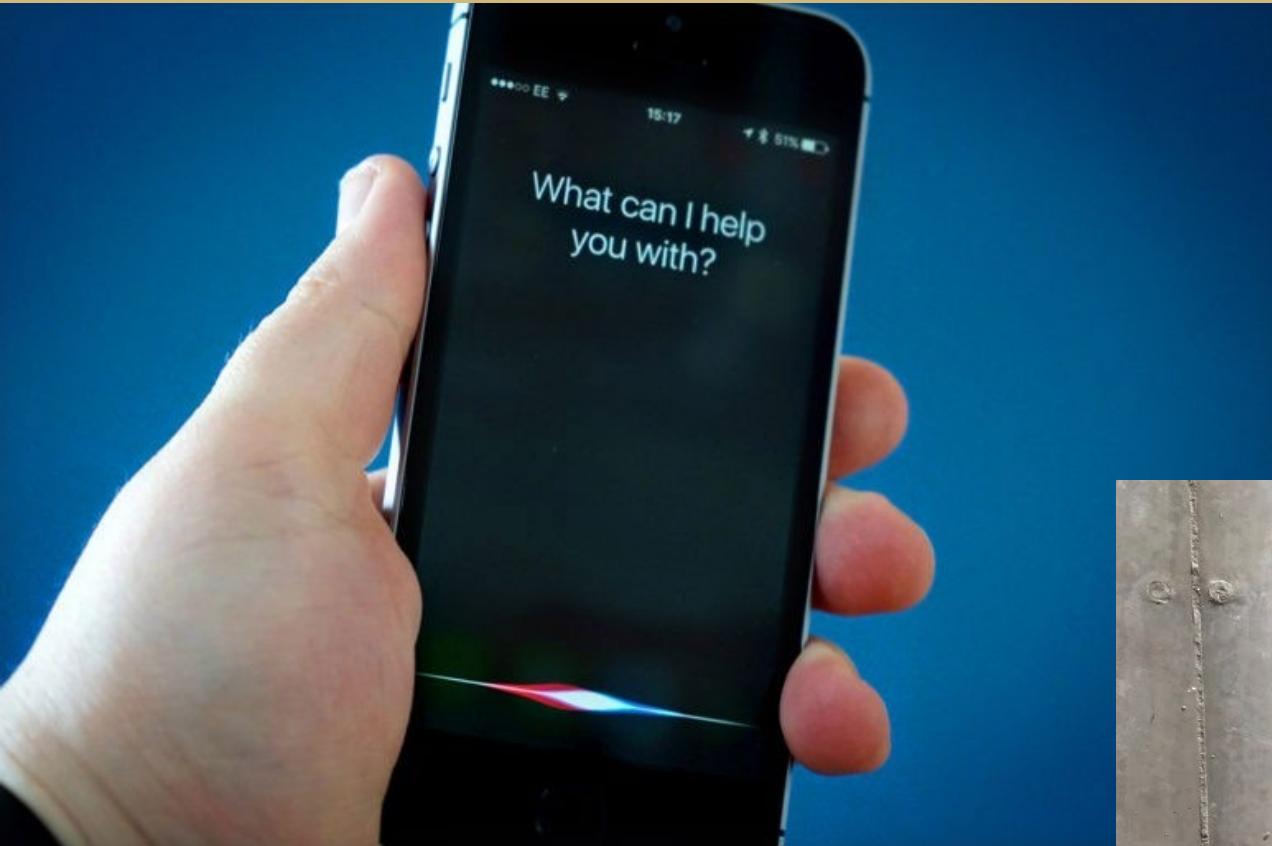
eBay is an online marketplace. The platform connects millions of buyers with sellers globally utilizing PayPal to ensure secure transactions. eBay products can be sold either via a silent auction in which users are able to input the maximum price they are willing to pay and for which the site will automatically increase bids as necessary up to that maximum, or via the Buy It

# Success of deep learning

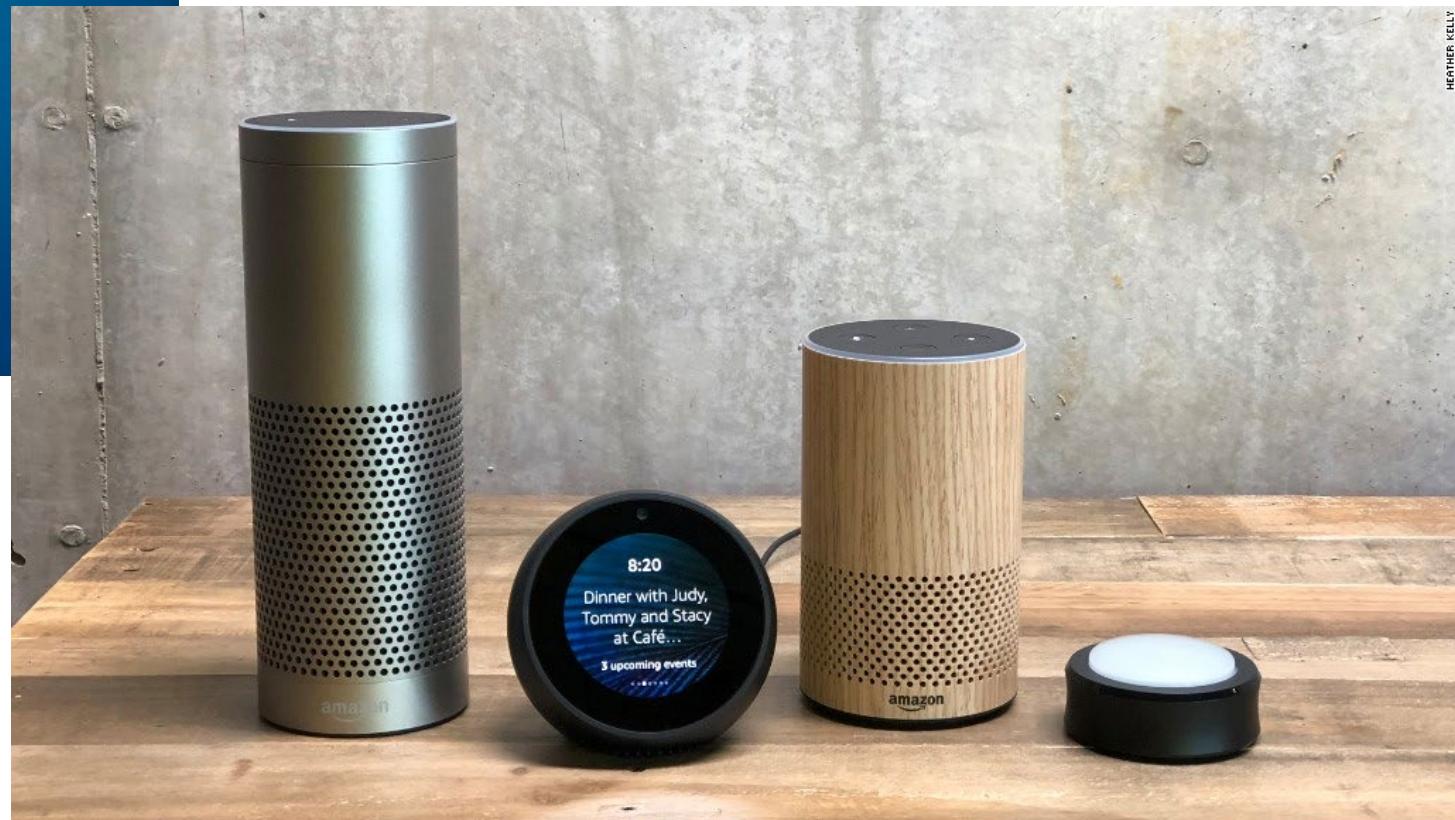


A few sample faces — all completely fake — created by [ThisPersonDoesNotExist.com](https://ThisPersonDoesNotExist.com)

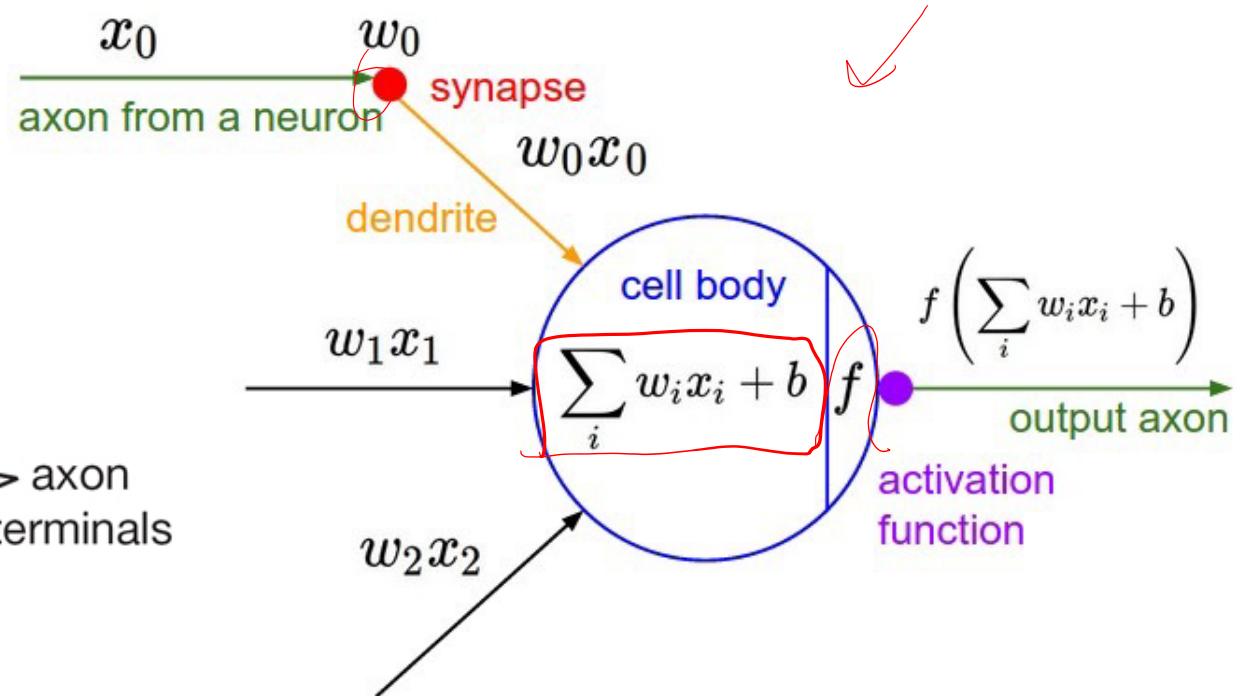
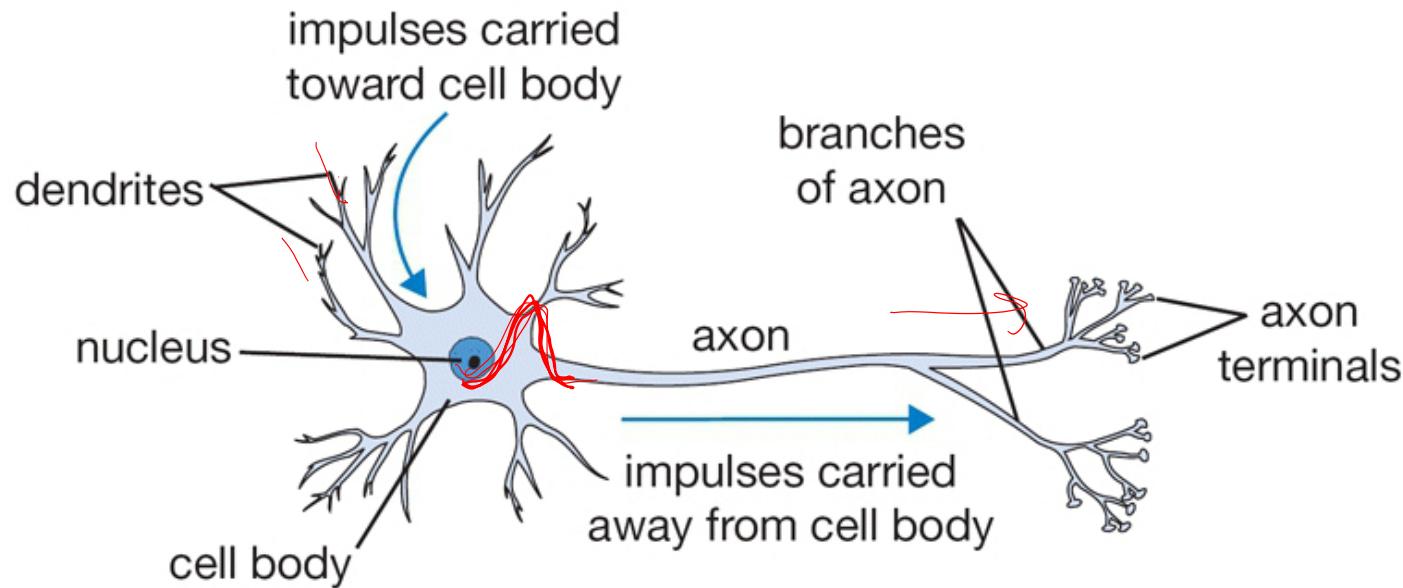
# Why Deep Learning?



*Voice Recognition,  
Smart Devices*



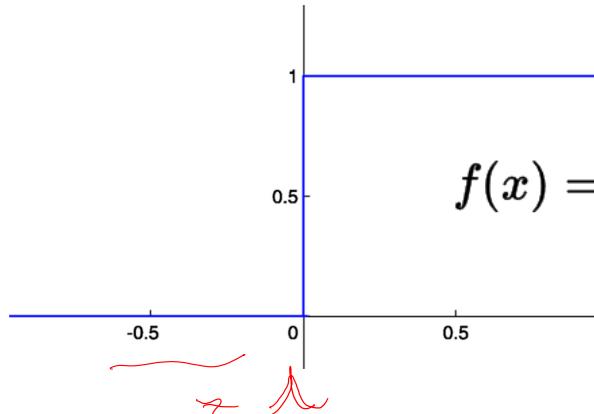
# What is Artificial Neuron (Perceptron)



# Activation functions

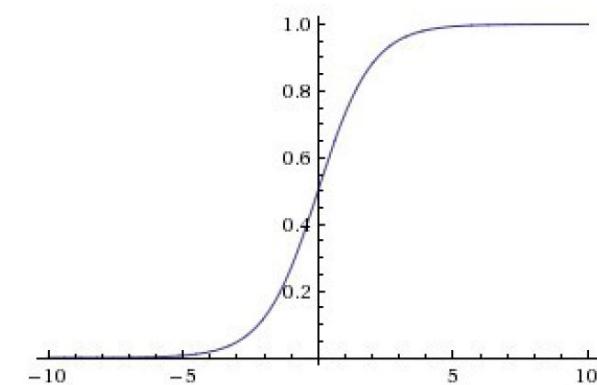
Output layers

- Binary Threshold (Step function)



$$f(x) = \begin{cases} 1 & \text{if } w \cdot x + b > 0 \\ 0 & \text{otherwise} \end{cases}$$

- Sigmoid



$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

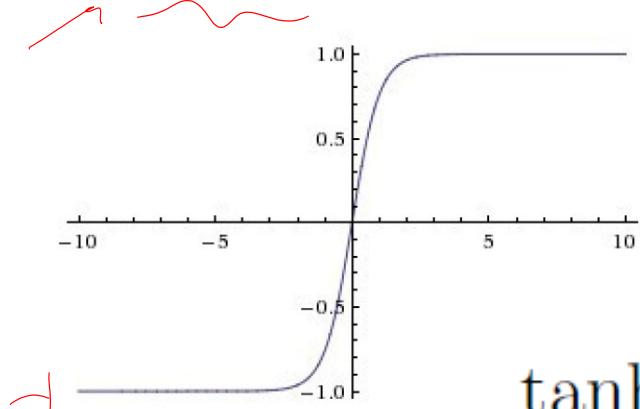
$$P \in \mathbb{R}[0, 1]$$

$$z = \mathbf{W} \cdot \mathbf{X} + b$$

- Softmax

$$P(y=j | \mathbf{x}) = \frac{e^{\mathbf{x}^\top \mathbf{w}_j}}{\sum_{k=1}^K e^{\mathbf{x}^\top \mathbf{w}_k}}$$

- Tanh

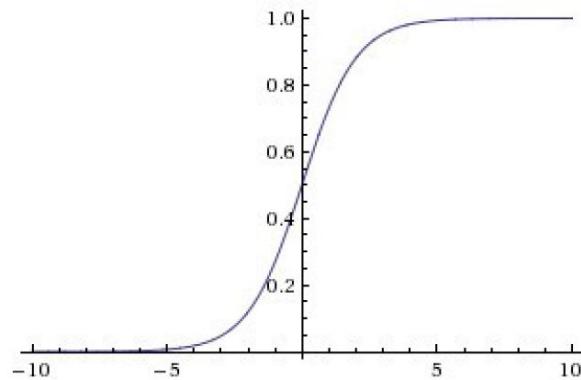


$$\tanh(z) = \frac{e^z - e^{-z}}{e^z + e^{-z}}$$

# Activation functions

- Sigmoid

~ binary

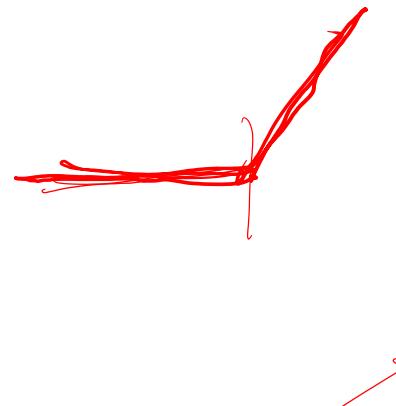
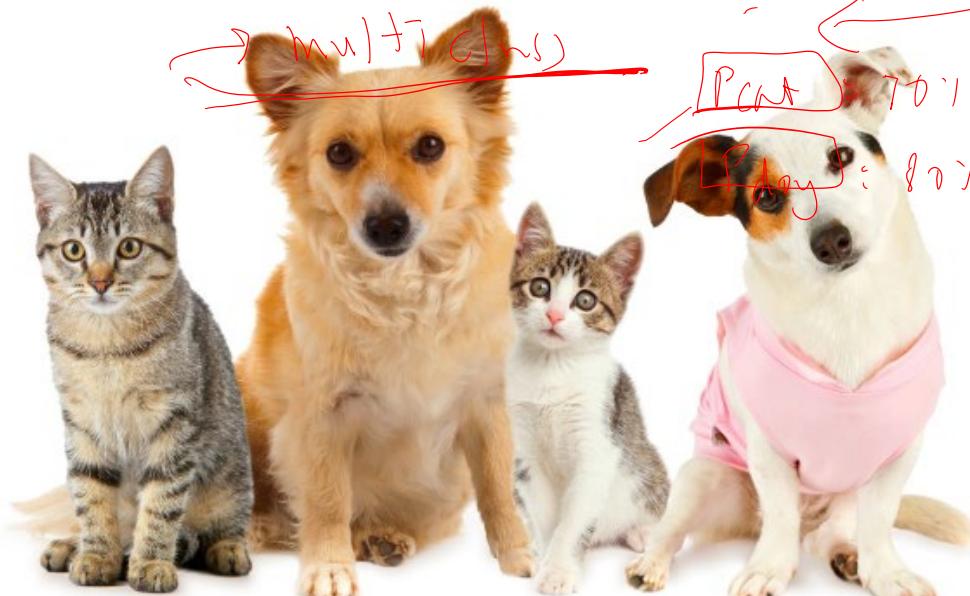


$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

$$P = \sigma(z)$$

$$P \in \mathbb{R}[0, 1]$$

$$z = \mathbf{W} \cdot \mathbf{X} + b$$



- Softmax

~ multiclass

$$P(y = j | \mathbf{x}) = \frac{e^{\mathbf{x}^\top \mathbf{w}_j}}{\sum_{k=1}^K e^{\mathbf{x}^\top \mathbf{w}_k}}$$

2.  $P(Y=1|\mathbf{x}) + P(Y=2|\mathbf{x}) = 1$

Cat 0.9

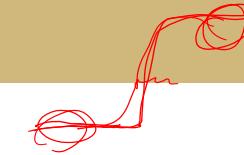
Dog 0.05

Bird 0.05

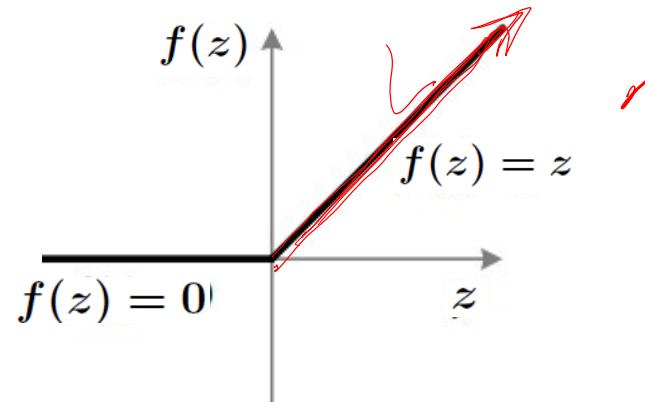
OR



# Activation functions



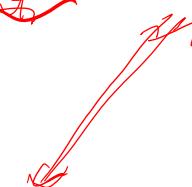
- Rectified Linear (ReLU)



- Better convergence
- Does not saturate
- Less computation

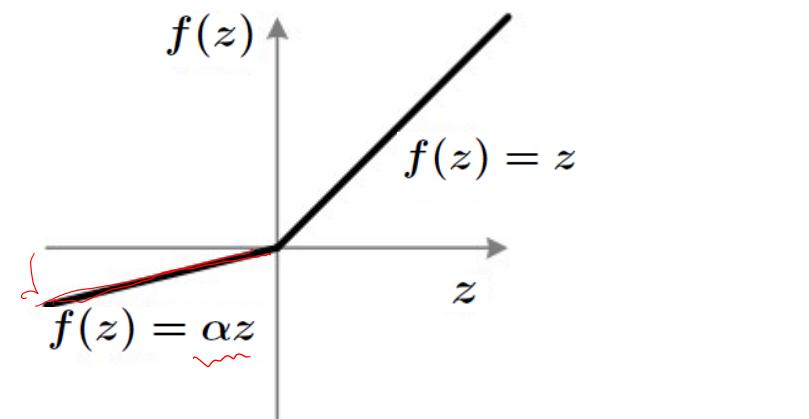
• regression

• fast

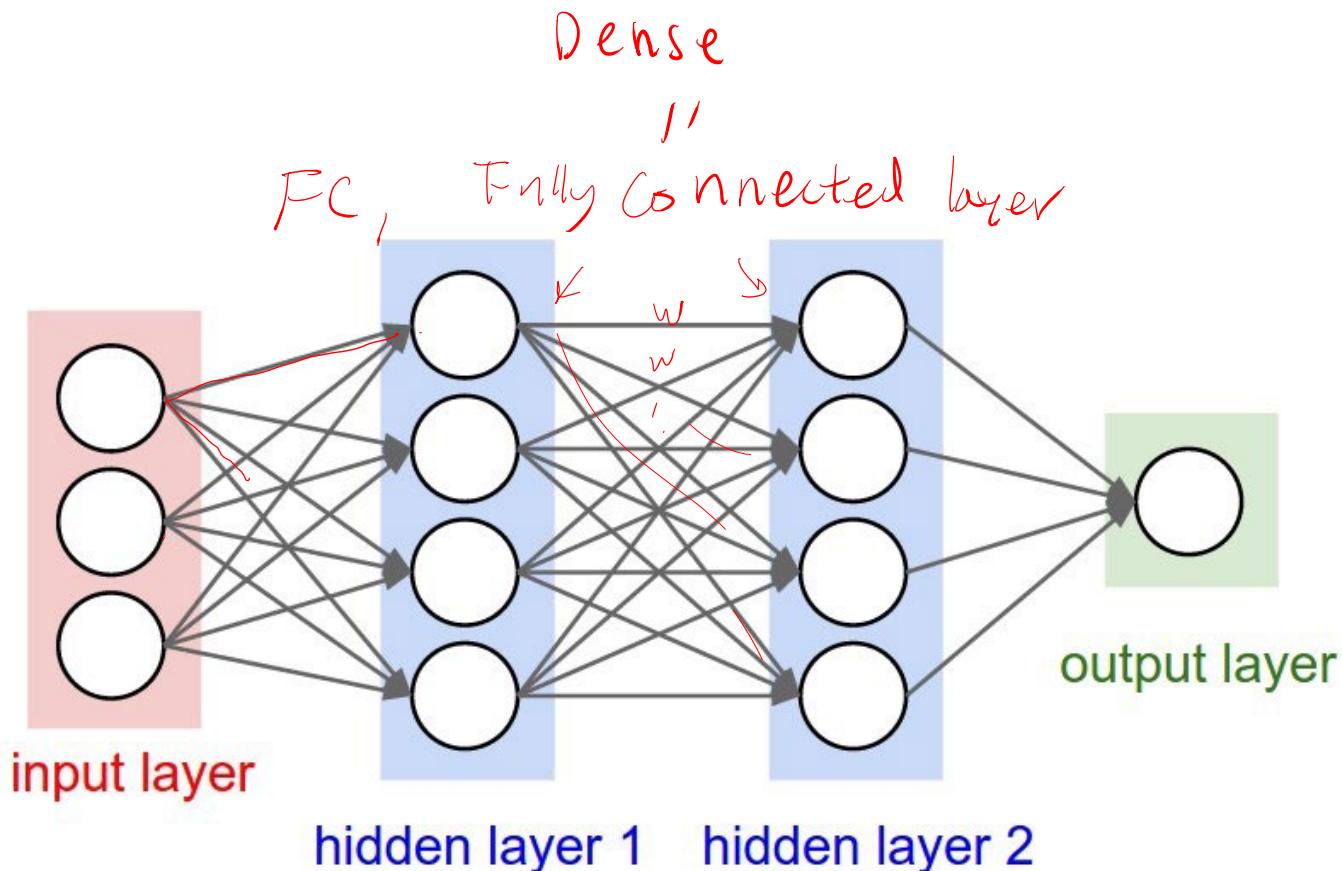


- Parametric Rectified Linear (PReLU) or Leaky ReLU

*m*



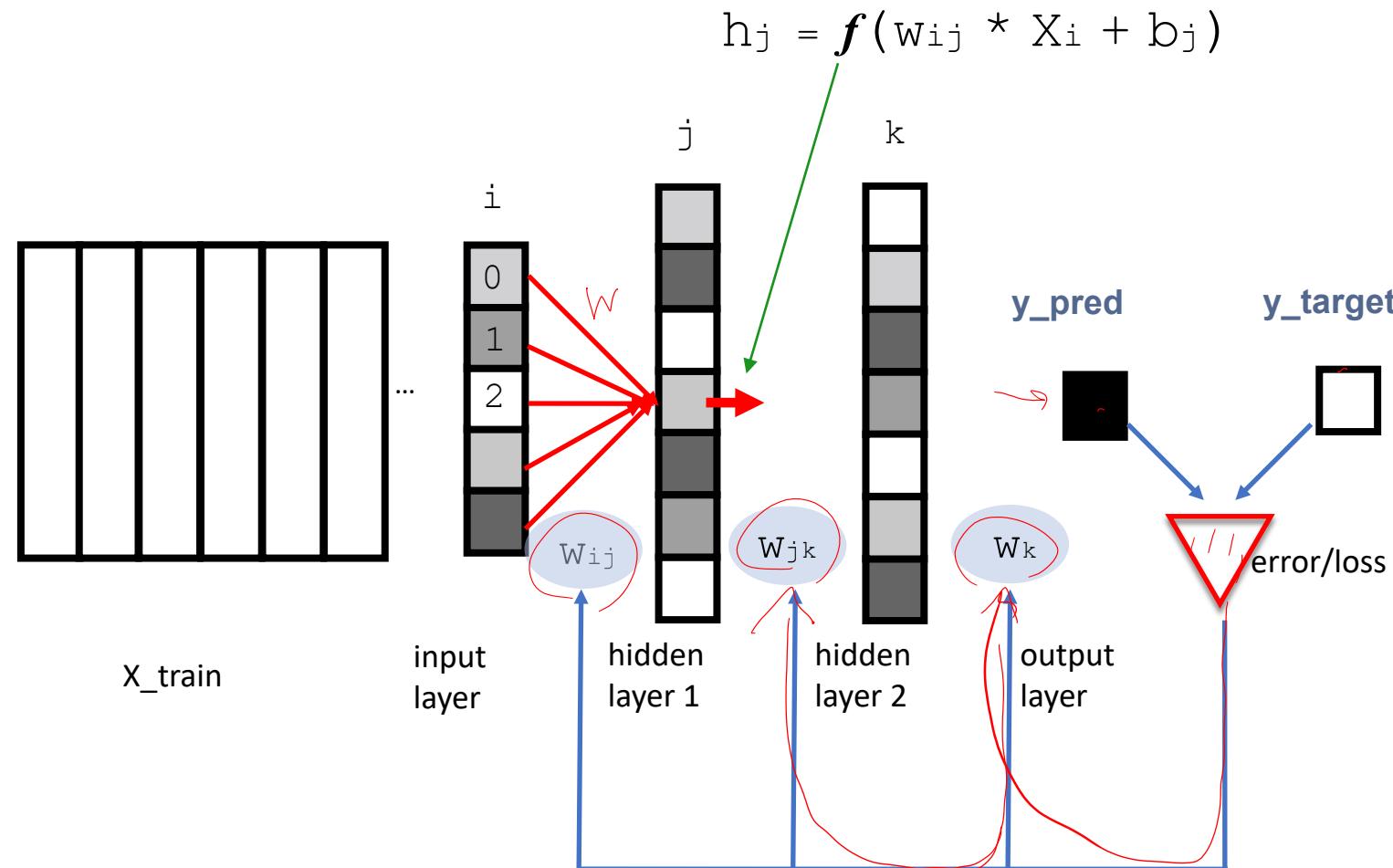
# Multi-layer Perceptron (Neural network)



## Design Parameters

- Architecture
- Number of layers
- Number of neurons in a layer
- Activation functions

# How Neural Network Training Works



**Weight update rule  
(Gradient descent)**

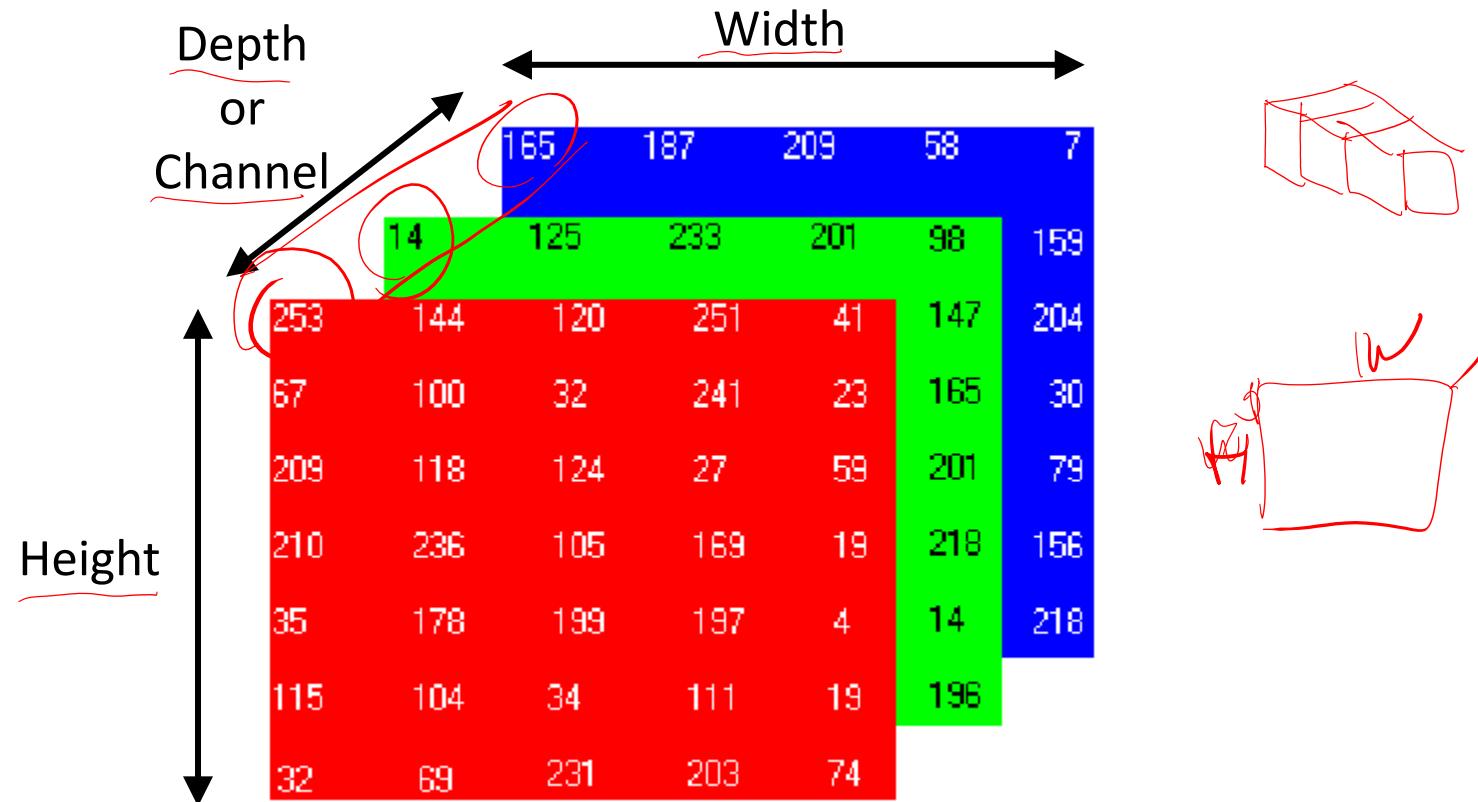
$$W_{ij} \leftarrow W_{ij} - \alpha \frac{\partial \mathcal{L}}{\partial W_{ij}}$$

**Chain rule**

$$\frac{\partial f(g(x))}{\partial x} = \frac{\partial f}{\partial g} \frac{\partial g}{\partial x}$$

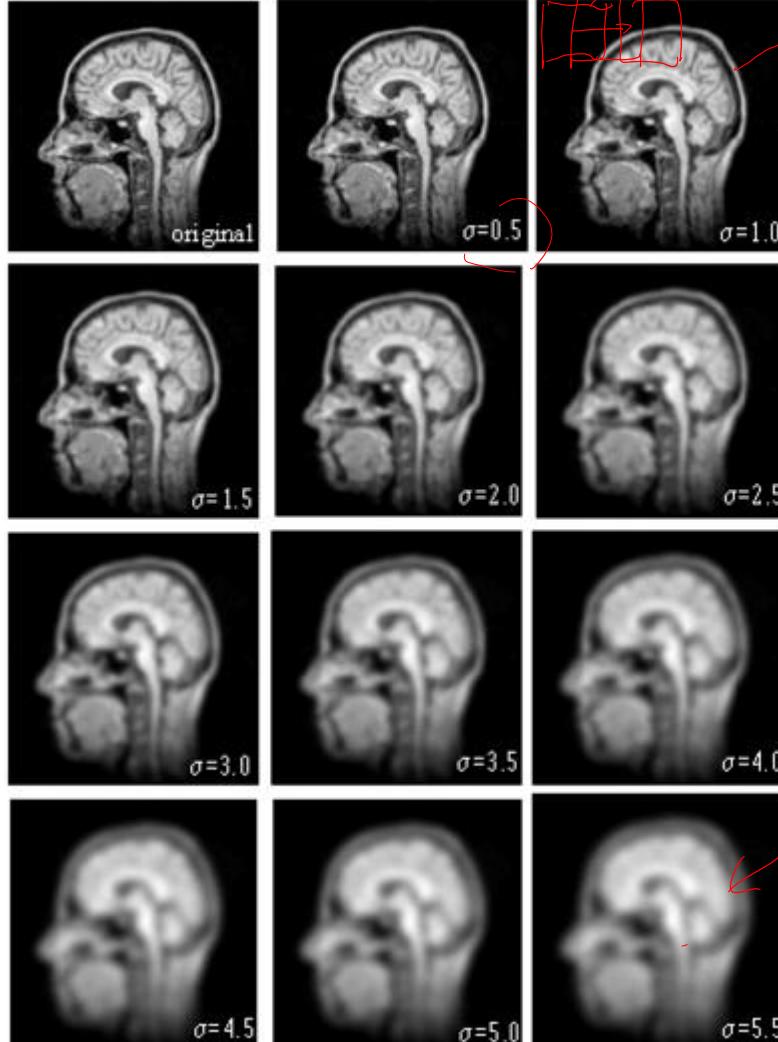
# Dealing with images

An image is a multi-dimension array



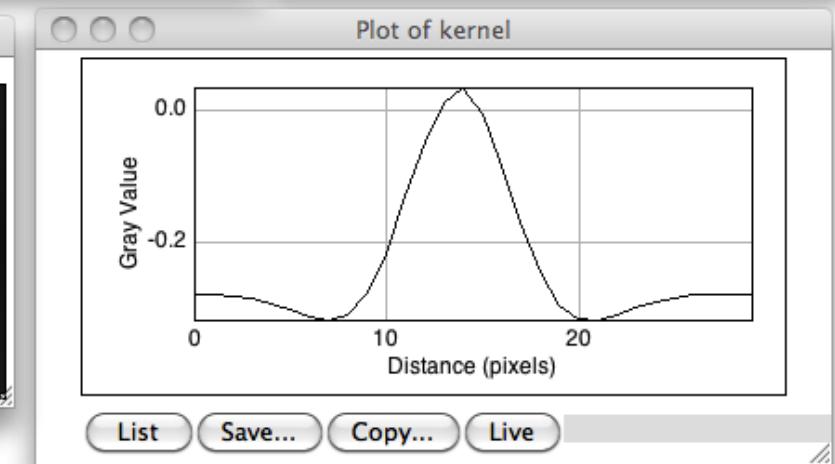
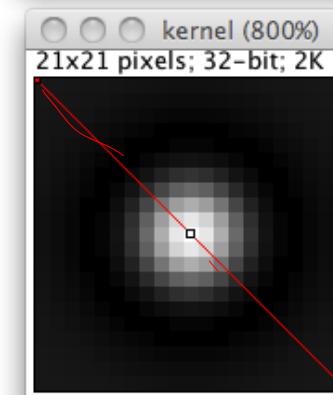
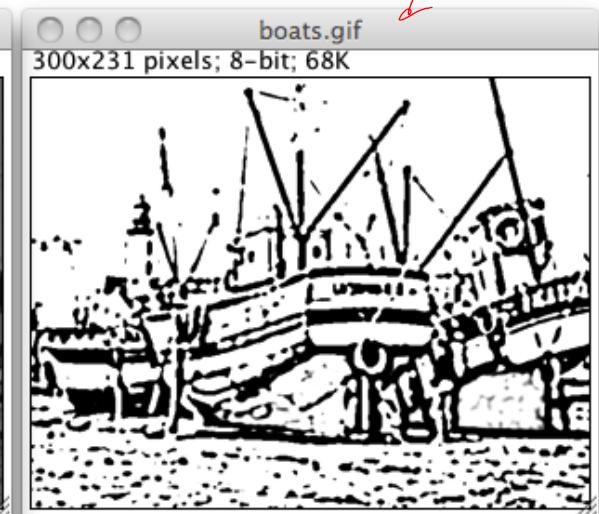
# Dealing with images- filters

Gaussian Blurring



feature map

Mexican hat filtering



# What is Convolution (2D) in an Image?

Diagram illustrating the receptive field of a central unit in a 5x5 input image. The receptive field is highlighted in red and covers a 3x3 area centered at the highlighted unit (1). A 3x3 filter is shown to the right.

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

filter

1	0	1
0	1	0
1	0	1

Image

Diagram illustrating the convolution process. A 3x3 filter is applied to a 5x5 input image. The result is a 3x3 convolved feature map. Red annotations show the receptive field of the central unit in the output map, which is highlighted in pink.

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

1	0	1
0	1	0
1	0	1

Image

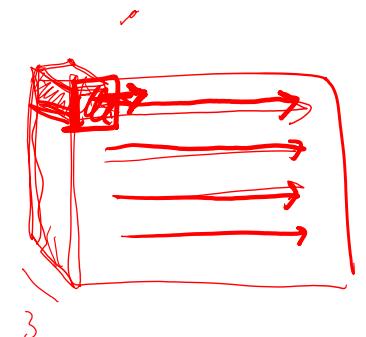
Diagram illustrating the resulting 3x3 convolved feature map. The central unit is highlighted in pink and has a value of 4.

4	0	0
0		
0		

Convolved  
Feature

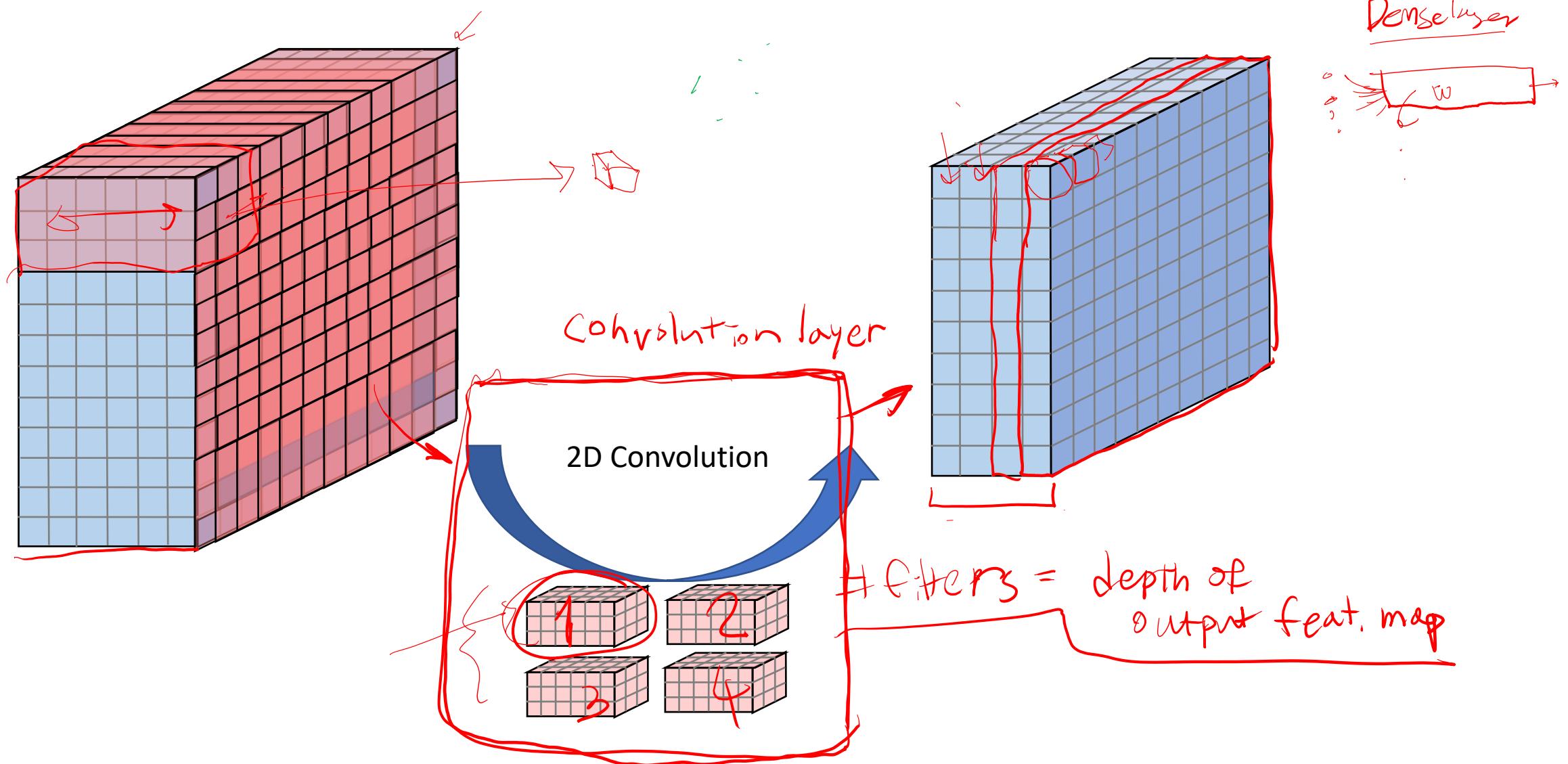
$$\left( \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} * \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \right) [2, 2]$$

$$= (i \cdot 1) + (h \cdot 2) + (g \cdot 3) + (f \cdot 4) + (e \cdot 5) + (d \cdot 6) + (c \cdot 7) + (b \cdot 8) + (a \cdot 9)$$

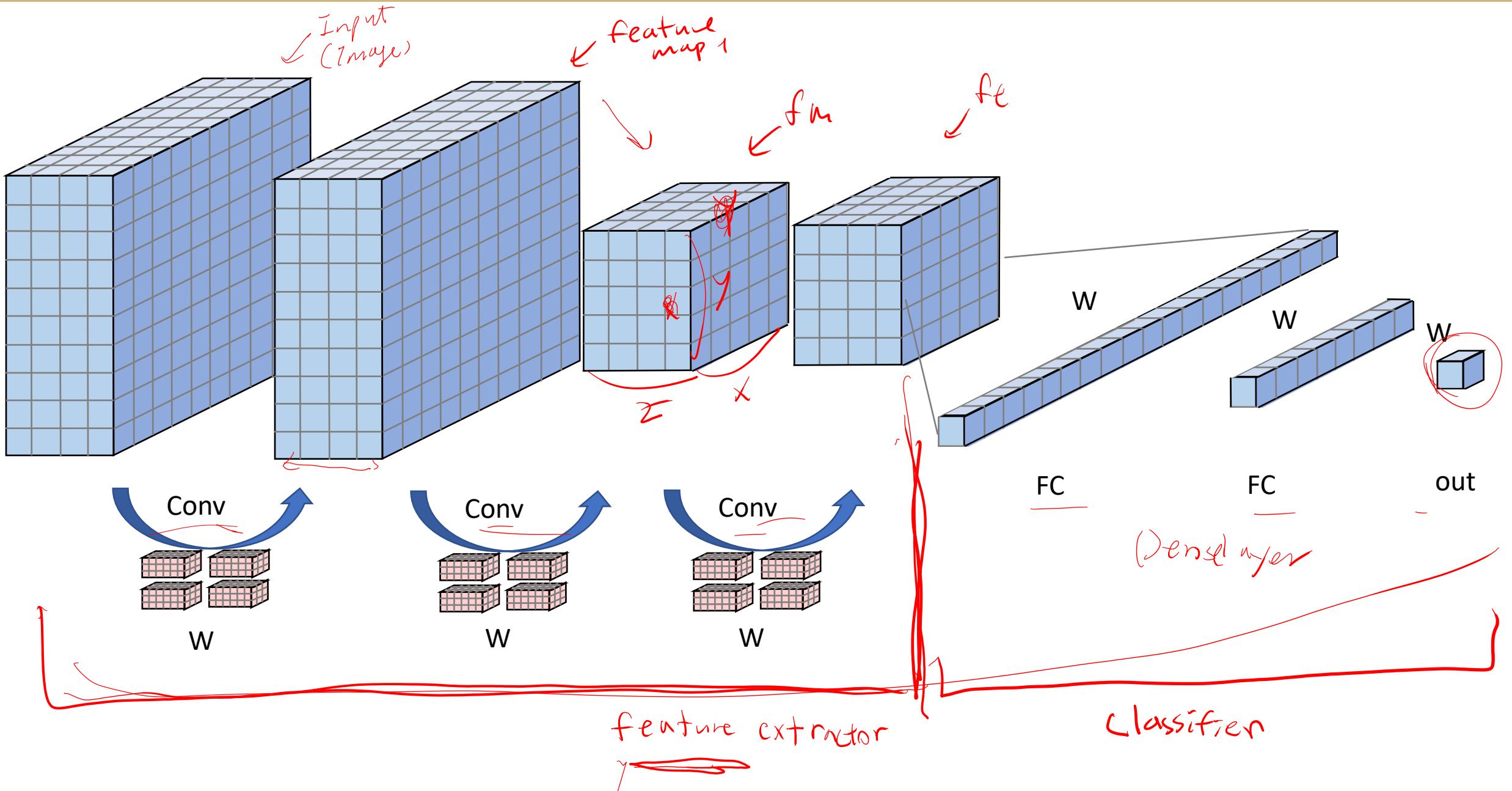


# Convolutional Layer

Q1. What is the feature map dimension, after a convolution layer with N filters?

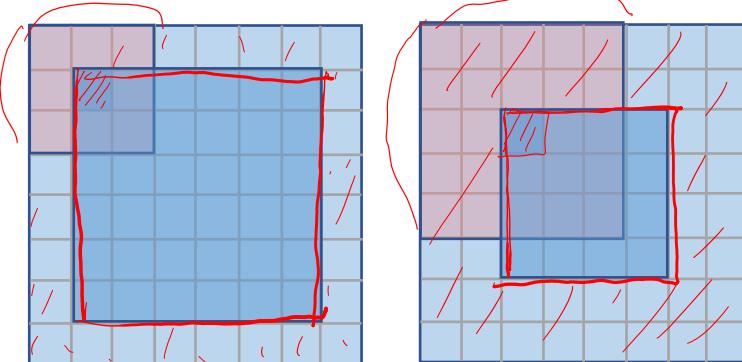


# Convolutional Neural Network

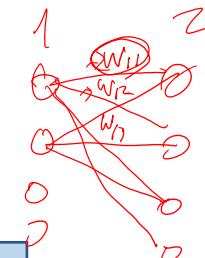
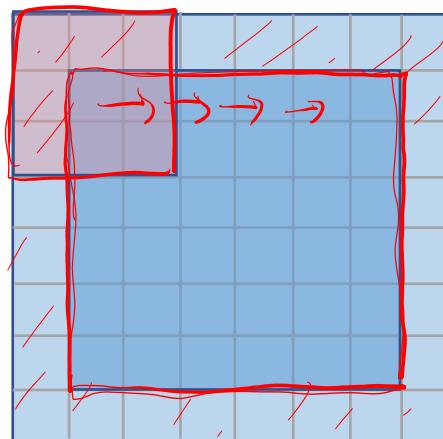


# Convolutional Layer- hyper parameters

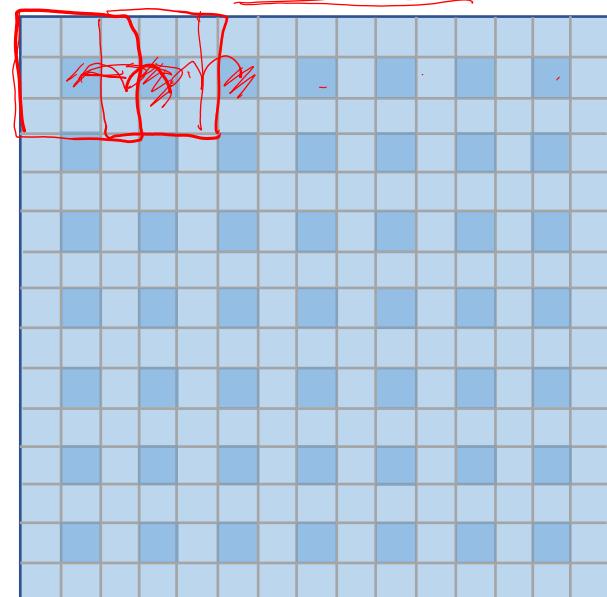
- Filter size ( ~~$3 \times 3$ ,  $5 \times 5$ ,  $1 \times 1$ ...~~) x2 28



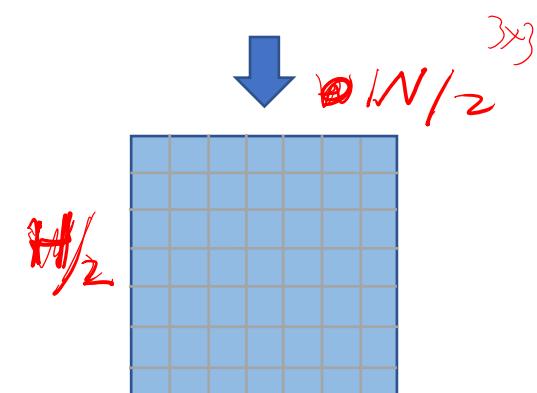
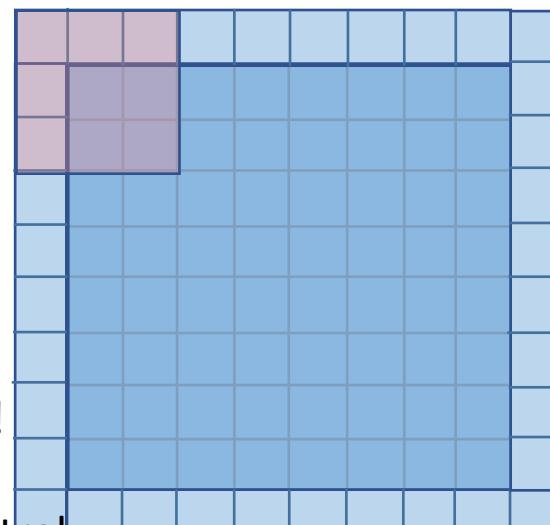
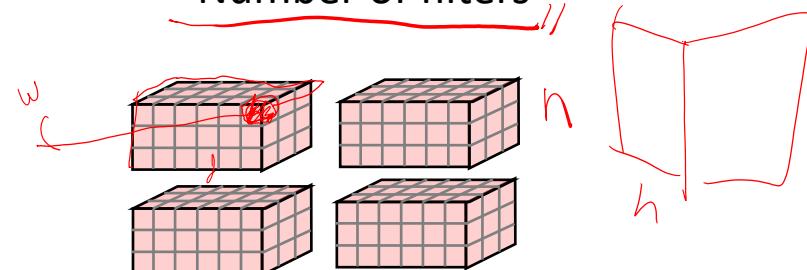
- Padding



- Stride = 2



- Number of filters 11



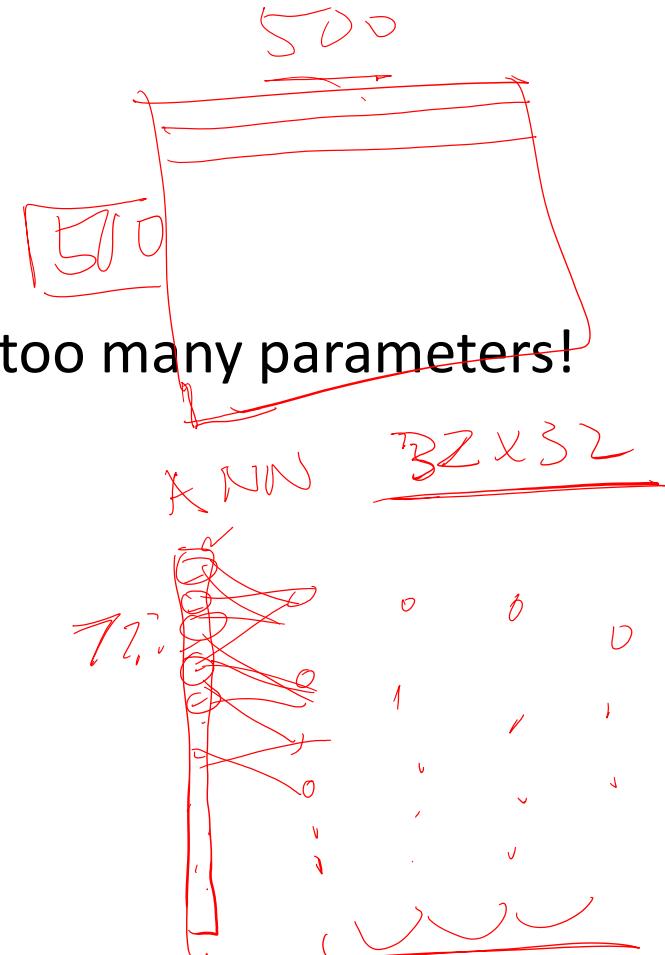
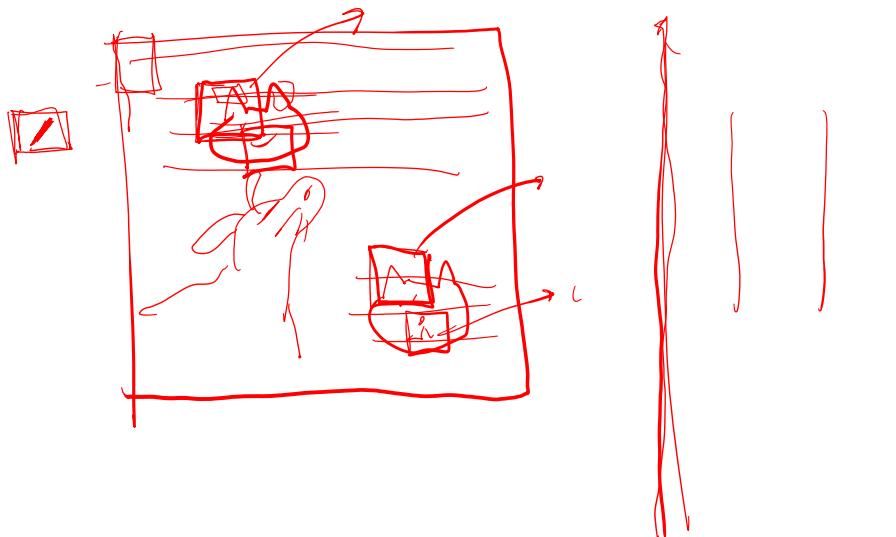
\*\* Parameters and Hyper parameters are different!

Parameters = Weights to be optimized.

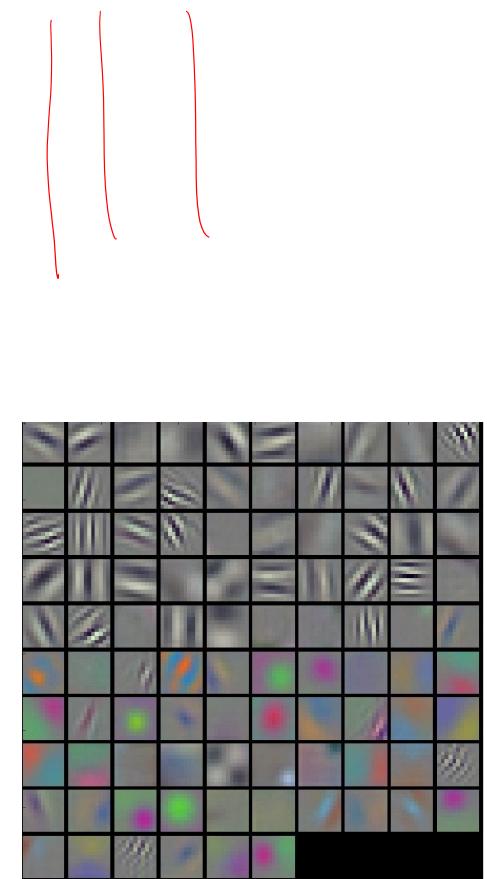
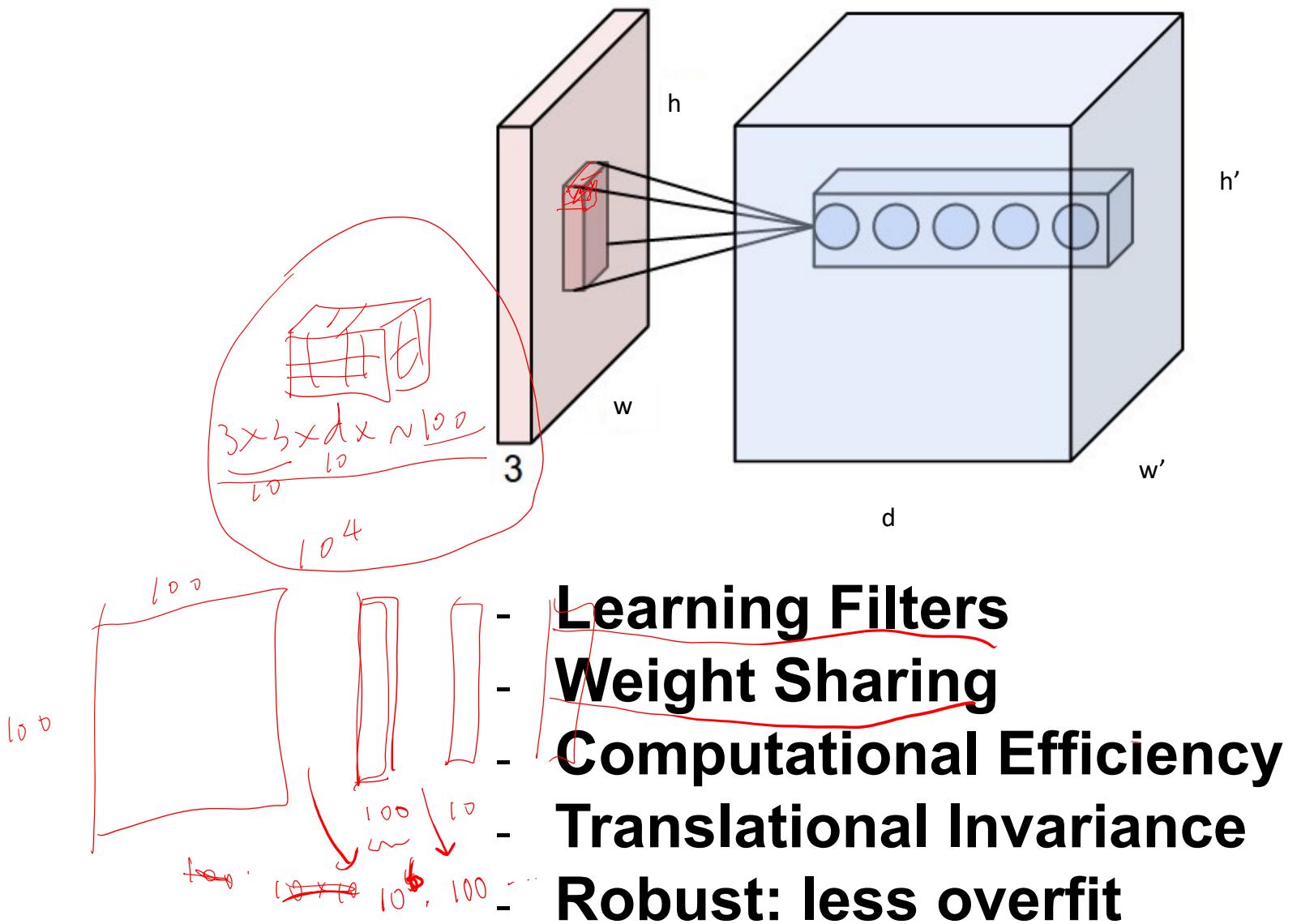
Hyperparameters = design parameters you can control

# Why CNN? What do the convolution filters do?

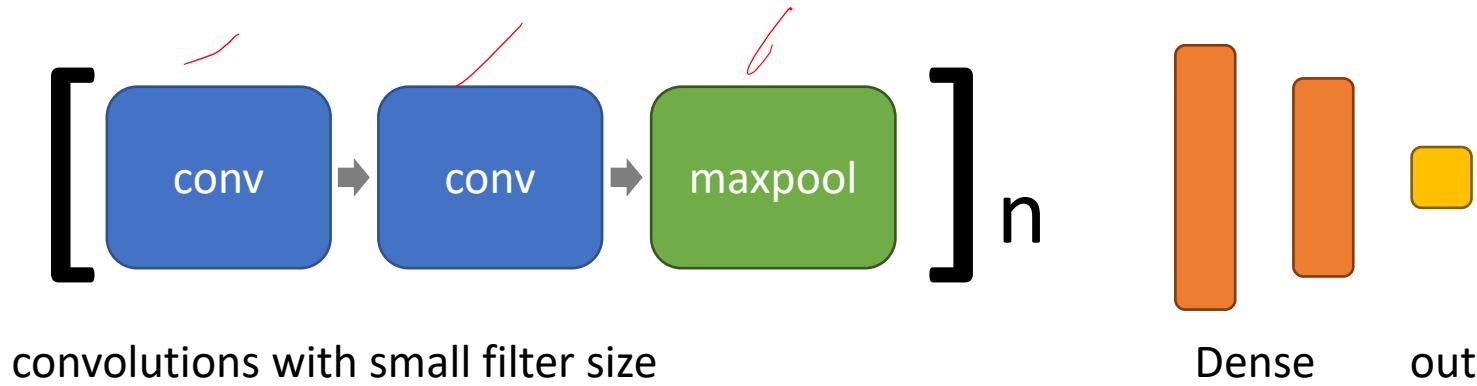
- Images have big pixels!
- Fully-connected neural network would have too many parameters!
- Translational invariance in images



# Why CNN? What do the convolution filters do?

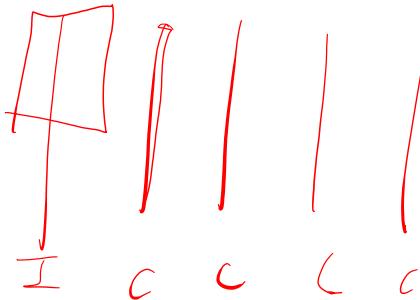


# Typical CNN architecture

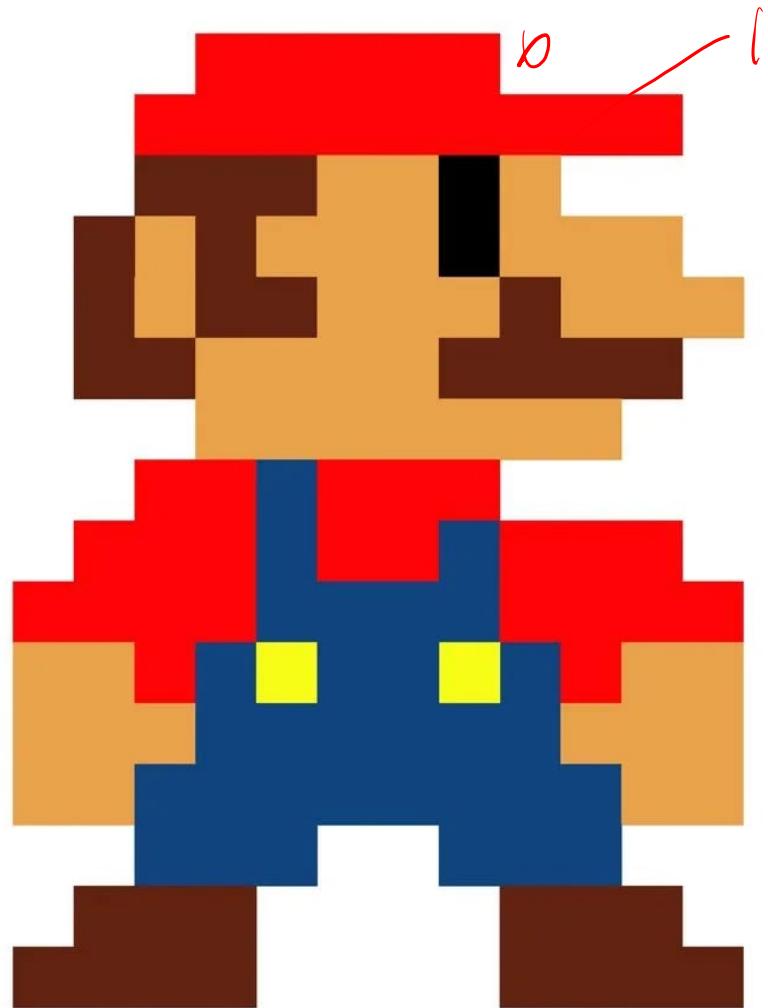


# What is a Pooling layer?

- Pooling is like sub-sampling
- Pooling filter size usually is  $2 \times 2$  (or  $2^n \times 2^n$ )
- Usually reduce the size to  $1/N$  per each side (e.g.  $N=2$  for  $2 \times 2$ )
- Max Pool
- Average Pool

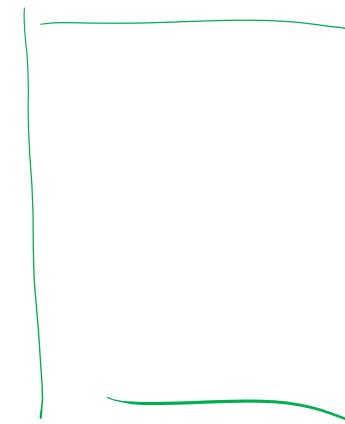
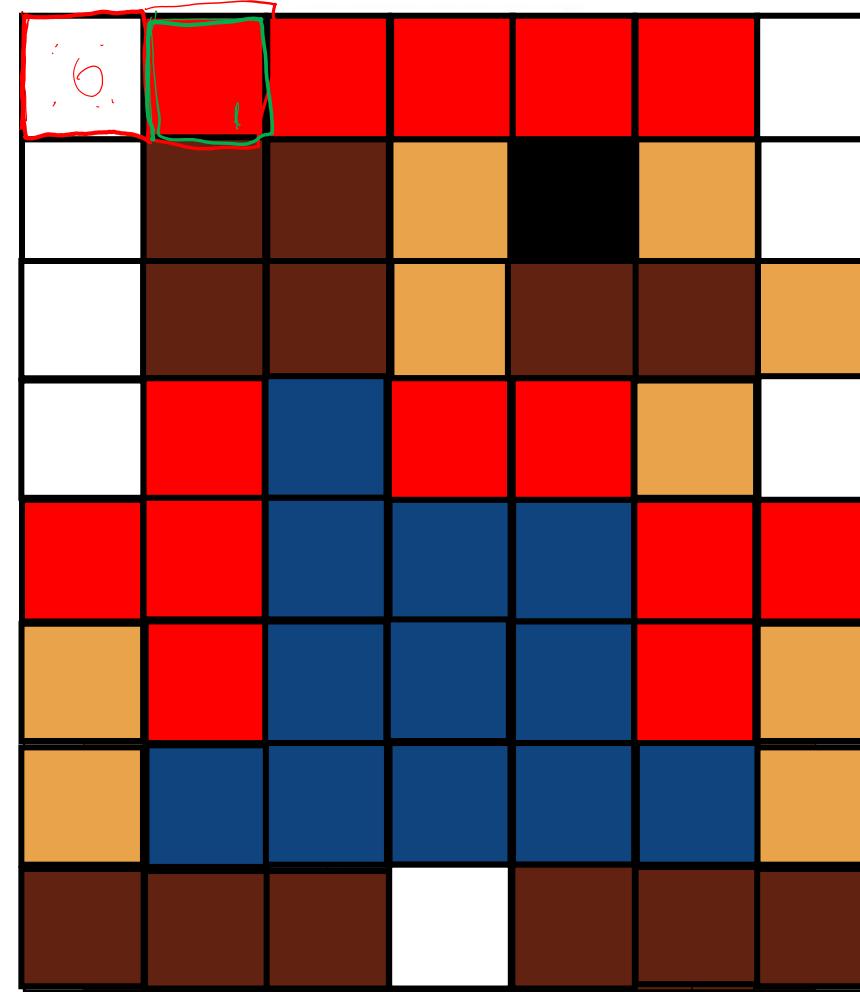


# Max pooling operation example



0 → 1

$2 \times 2 \quad s=2$

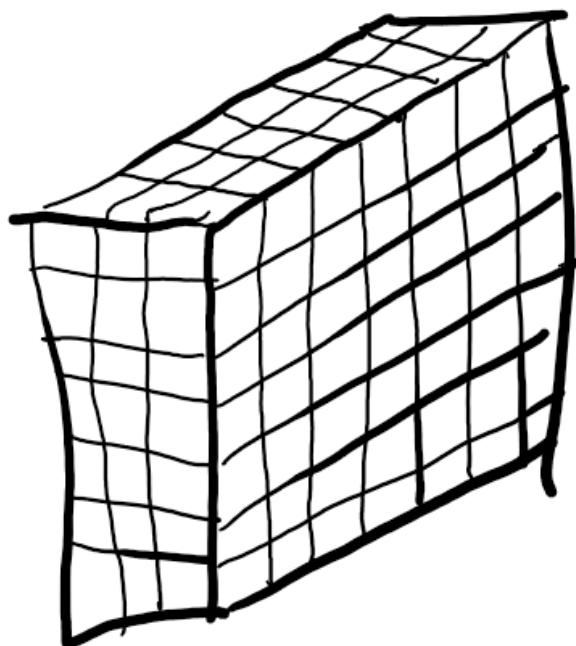


# Another way to shrink the spatial dimension

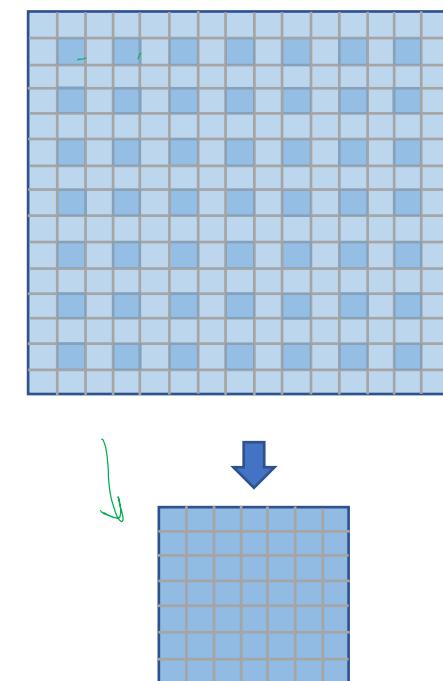
We can reduce the spatial dimension using convolution filters with a bigger stride

- Stride

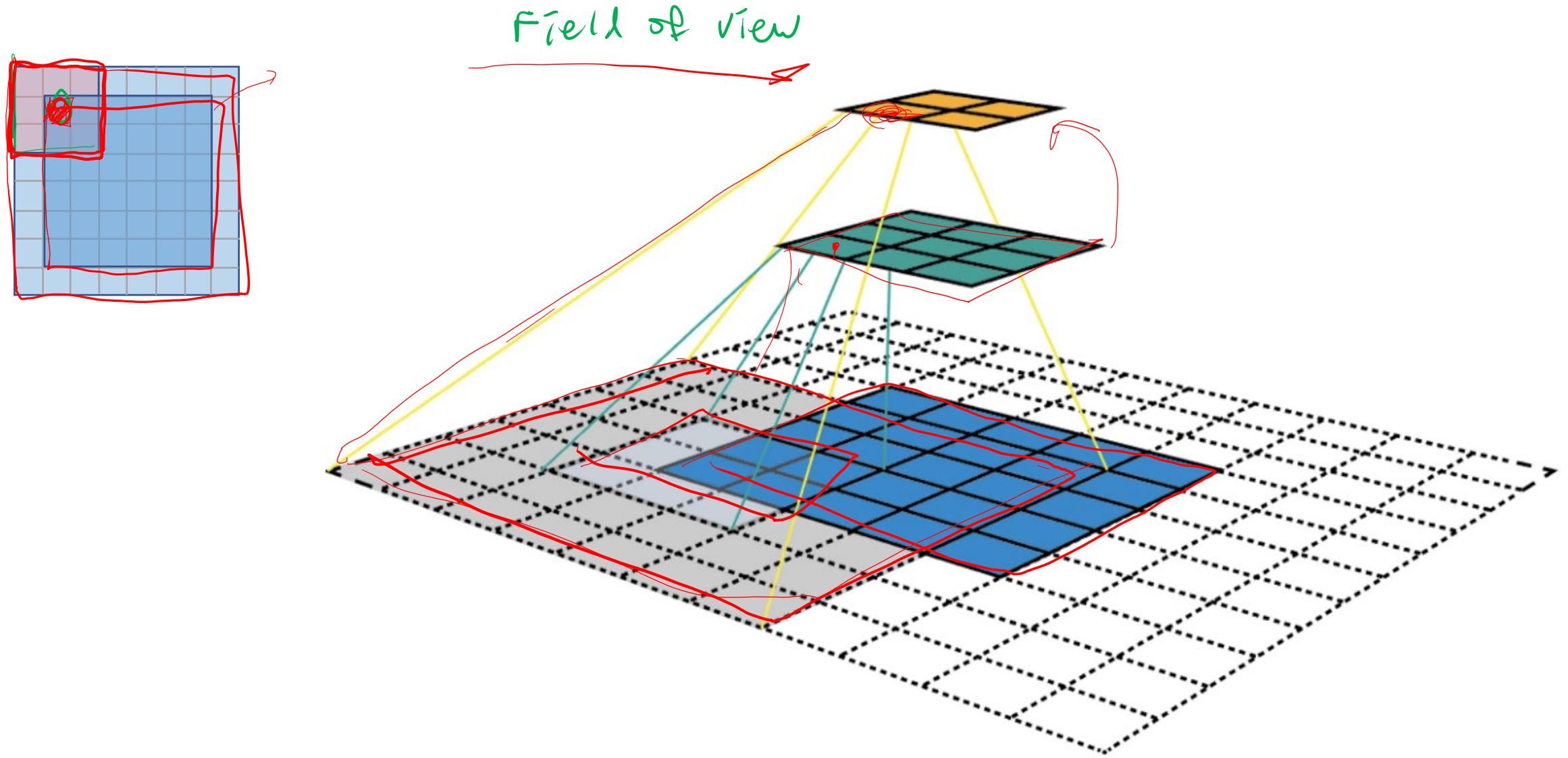
Exercise



$\text{conv}(3 \times 3, s=2)$   
w/o padding

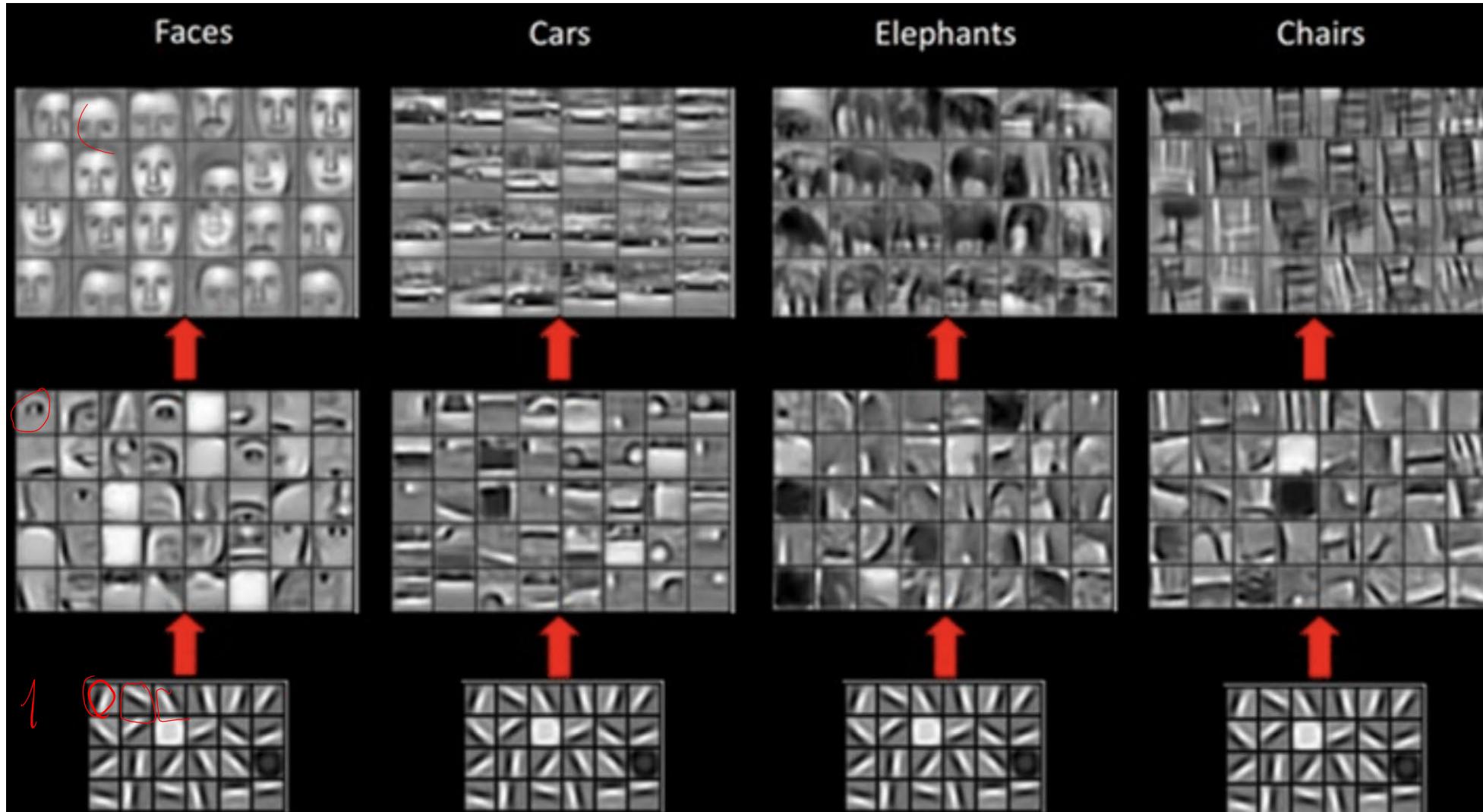


# What do the multiple convolution layers do?

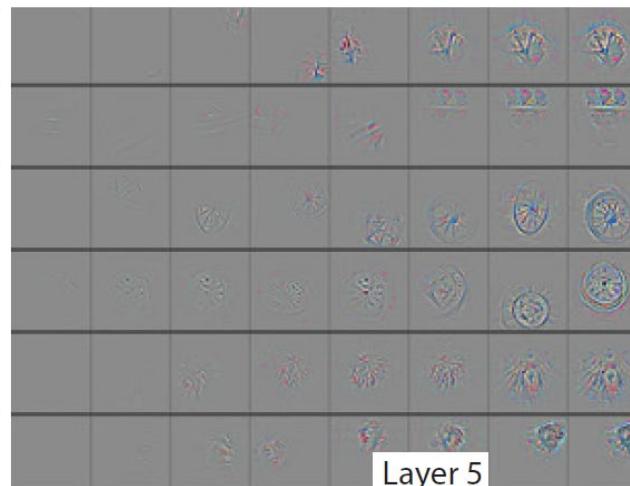
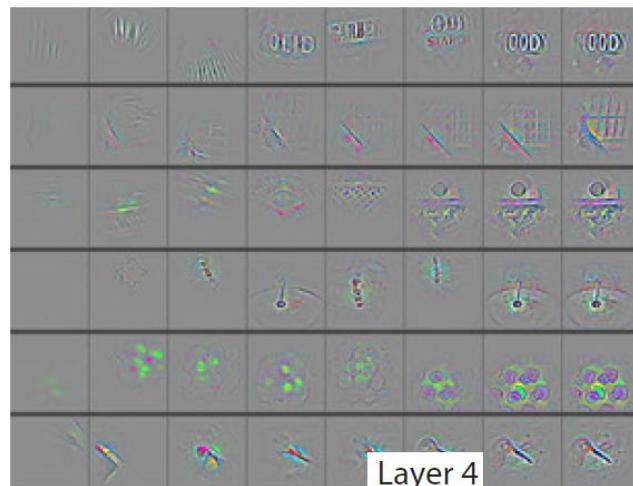
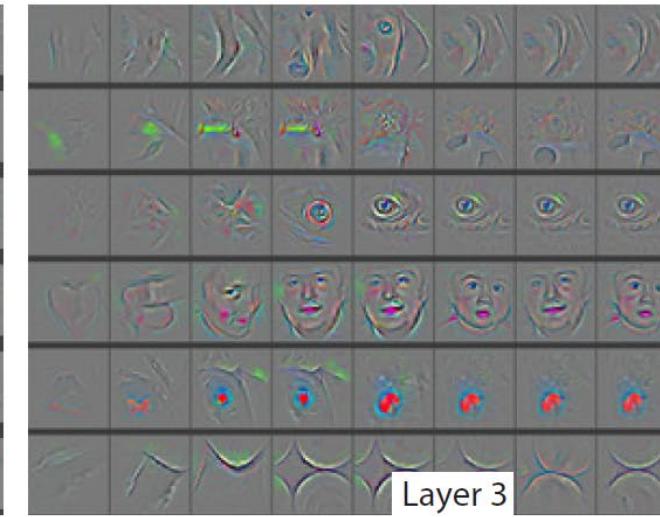
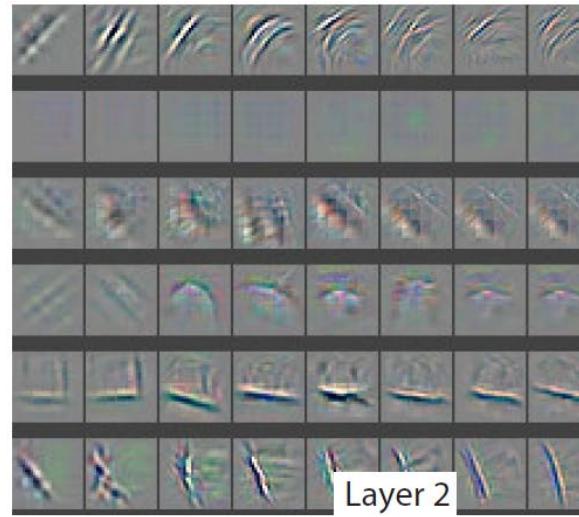
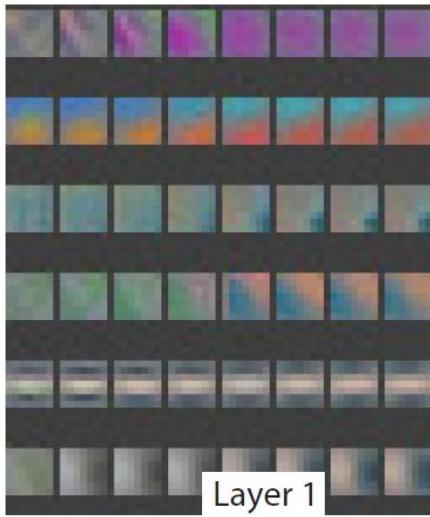


# What do the multiple convolution layers do?

Geometrical shapes are made of pieces of small curves, lines, and color blobs



# What do the multiple convolution layers do?



[arXiv:1311.2901v3](https://arxiv.org/abs/1311.2901v3)

# Next time!

- CNN Architectures
- How to train CNNs