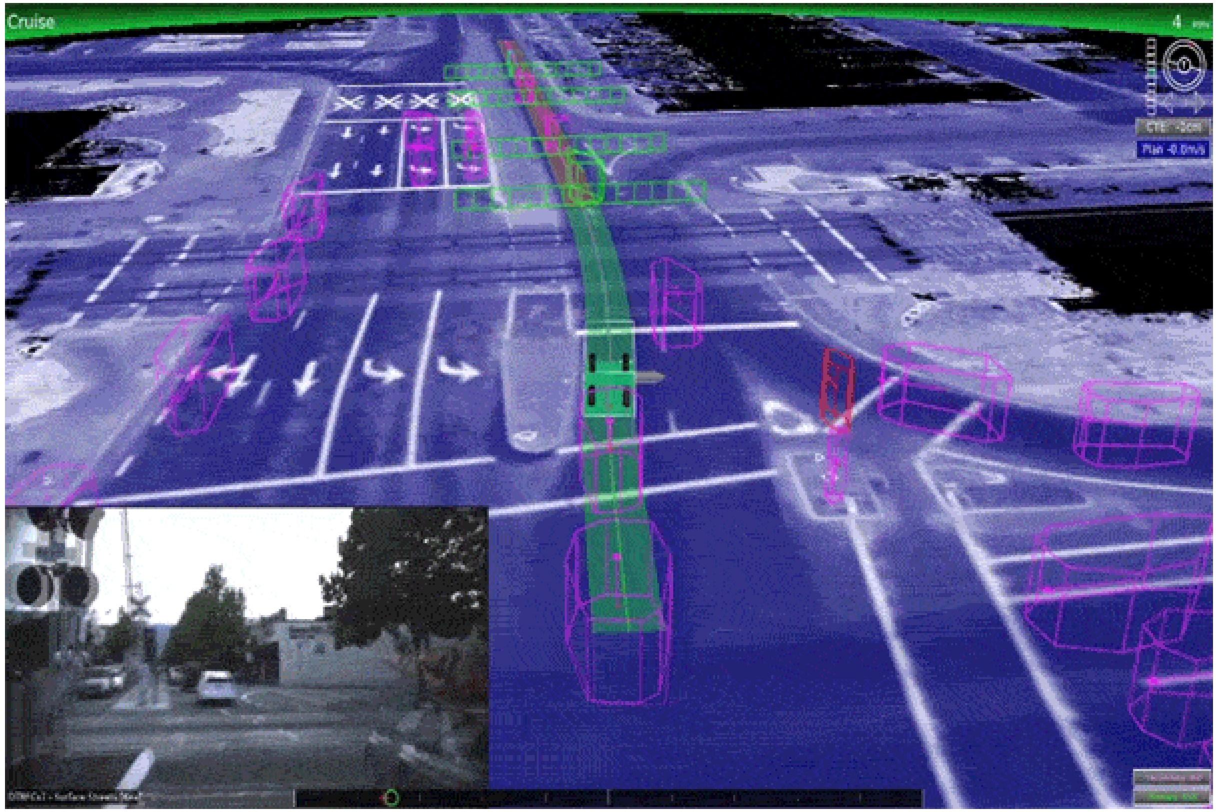
CSCI4622 Machine Learning

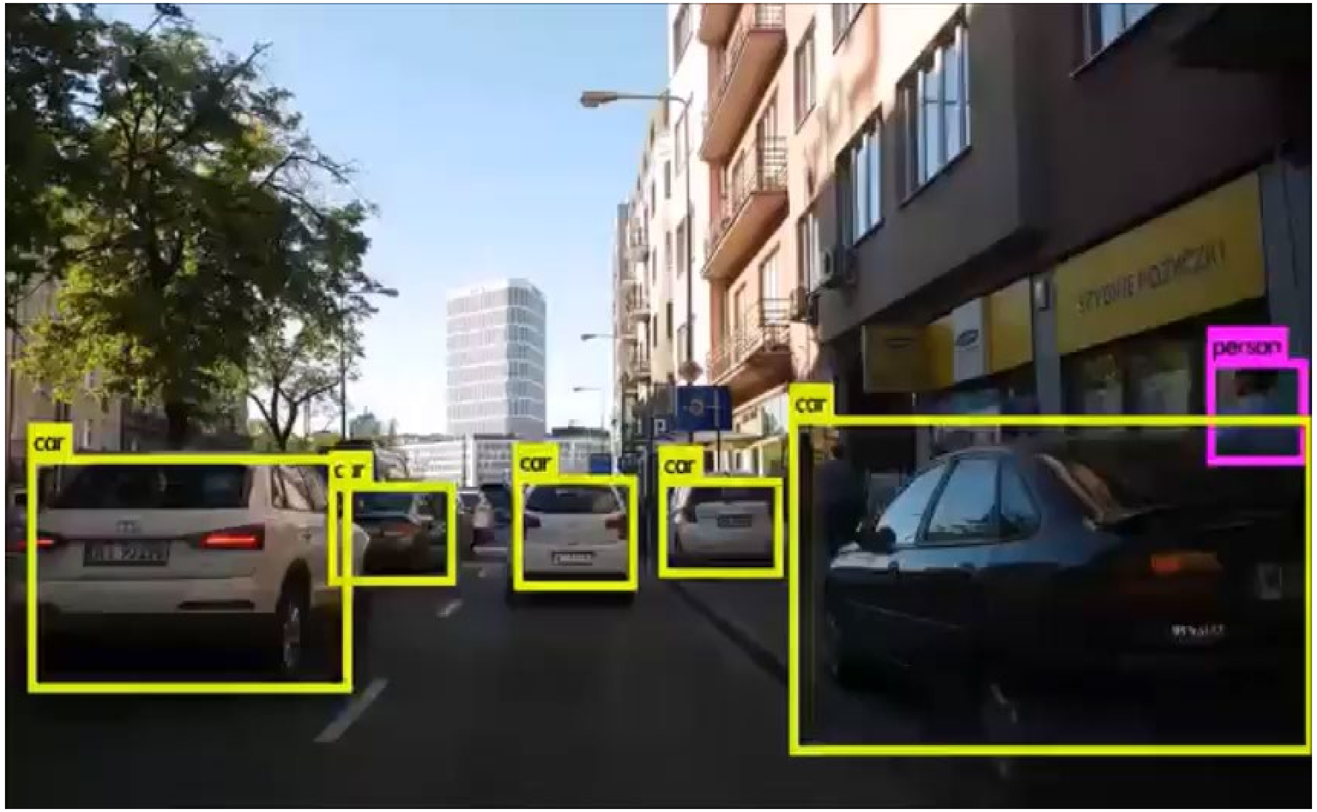
2020 Spring

Convolutional Neural Networks (1)

Geena Kim

Why Deep Learning?

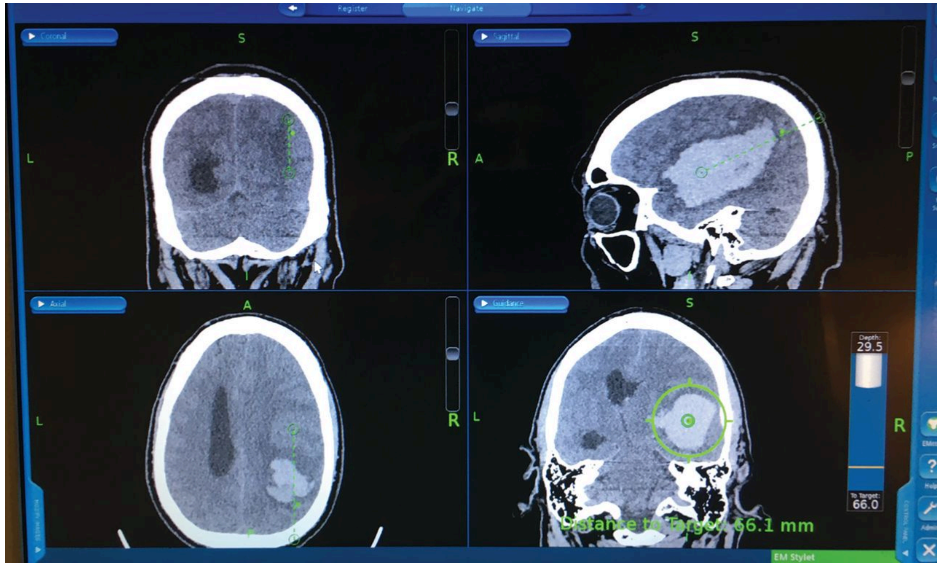


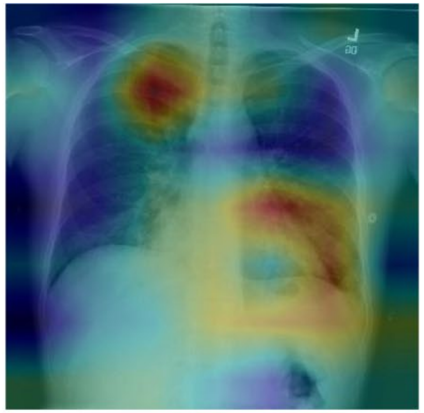


https://www.youtube.com/watch?v=EhcpGpFHCrw

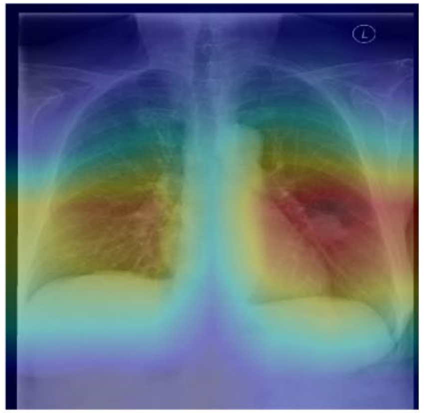
Why Deep Learning?

Medical Imaging, AI in Medicine

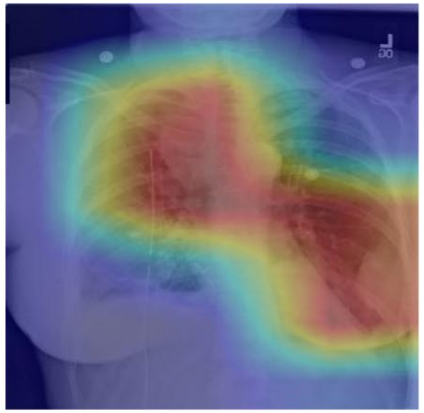




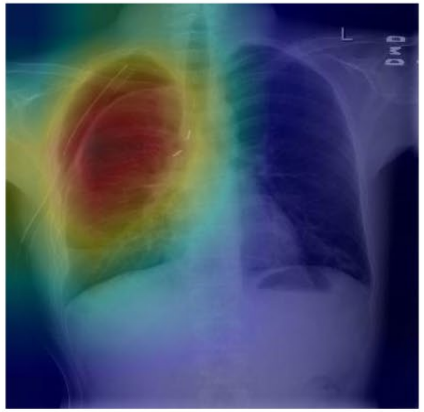
(a) Patient with multifocal community acquired pneumonia. The model correctly detects the airspace disease in the left lower and right upper lobes to arrive at the pneumonia diagnosis.



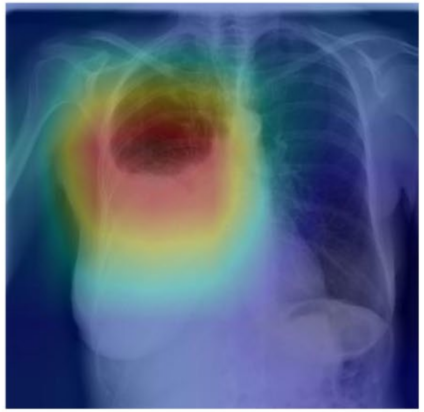
(b) Patient with a left lung nodule. The model identifies the left lower lobe lung nodule and correctly classifies the pathology.



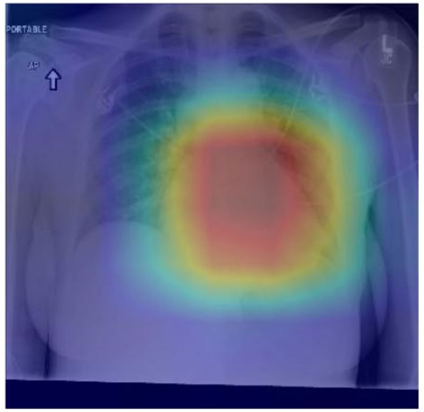
(c) Patient with primary lung malignancy and two large masses, one in the left lower lobe and one in the right upper lobe adjacent to the mediastinum. The model correctly identifies both masses in the X-ray.



(d) Patient with a right-sided pneumothroax 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 el., 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 –

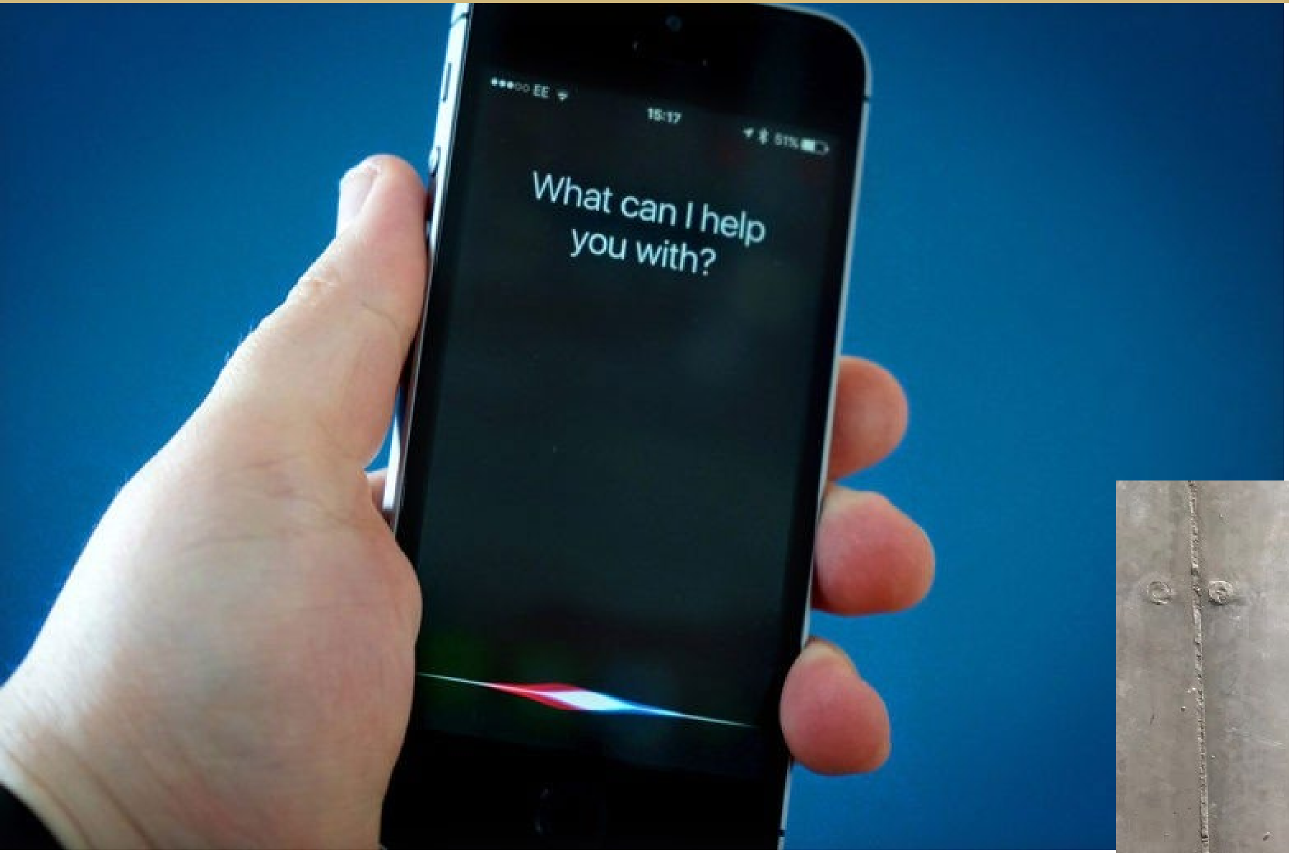
FOUNDED
1995

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

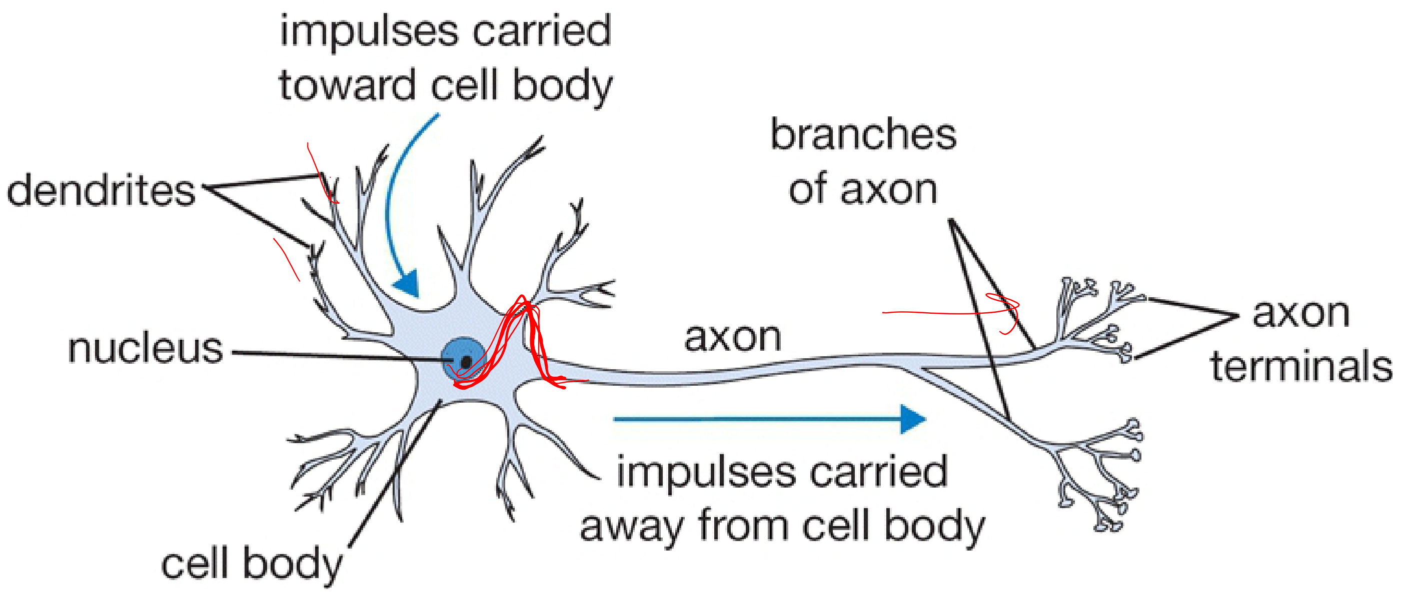
A few sample faces — all completely fake — created by ThisPersonDoesNotExist.com

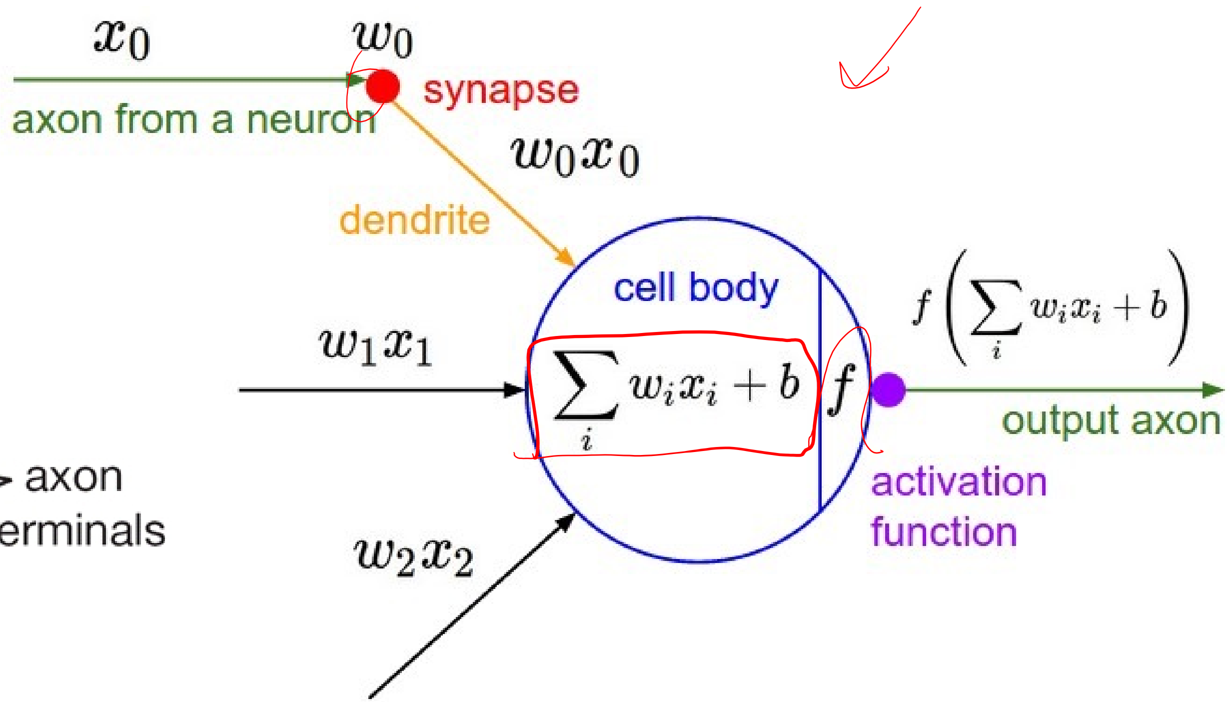
Why Deep Learning?



Voice Recognition,
Smart Devices

What is Artificial Neuron (Perceptron)

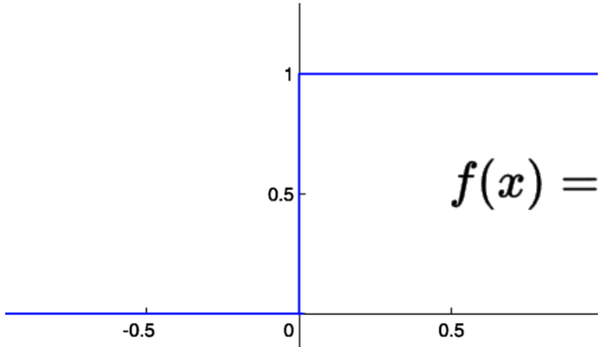




Activation functions

Output layers

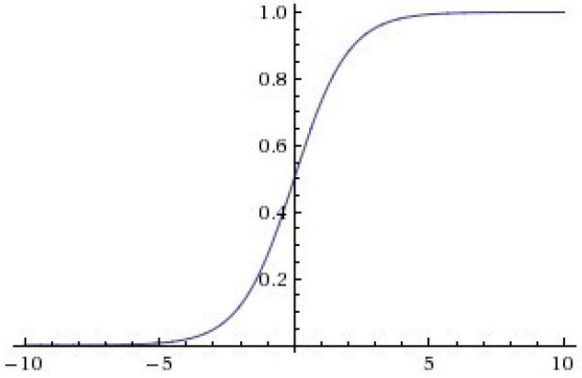
- Binary Threshold (Step function)



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

z

- Sigmoid



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

$$P = \sigma(z)$$

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

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

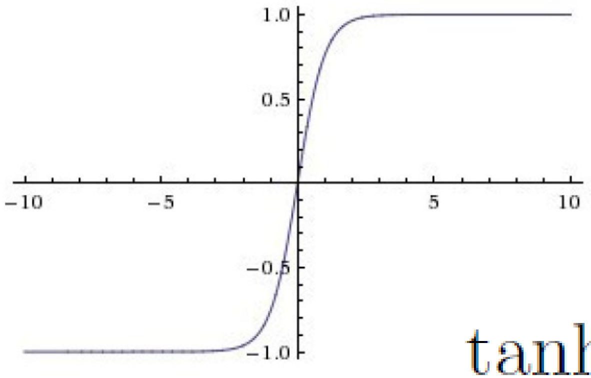
- Softmax

0 - 1

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

- Tanh

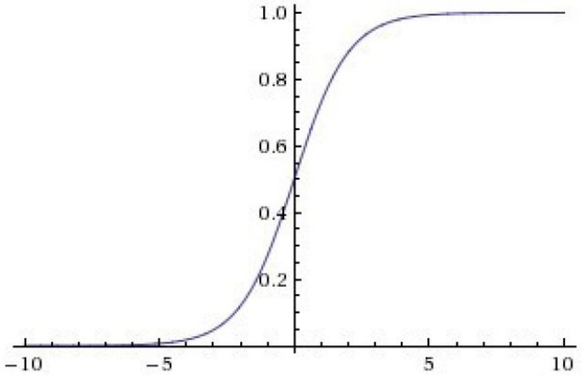
Angle



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

Activation functions

- Sigmoid ~ binary

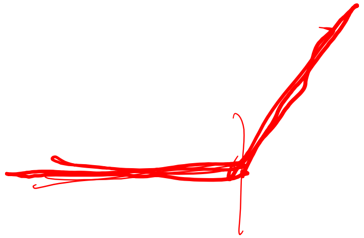


$$\sigma(z) = \frac{1}{1+e^{-z}}$$

$$P = \sigma(z)$$

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

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



- Softmax ~ multiclass

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

1

2

$P(y=1|x)+P(y=2|x)$
$=1$

multiclass

P cat : 70%

P dog : 80%

cat 0.9

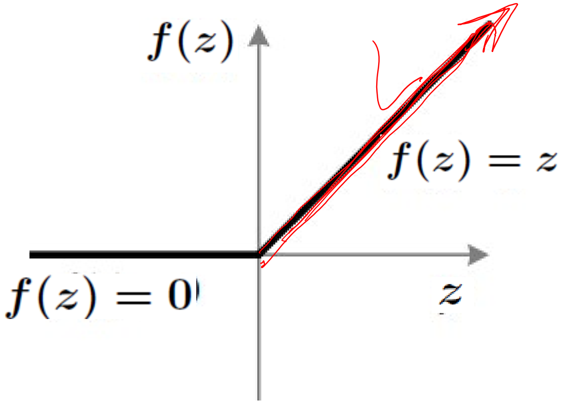
dog 0.05

bird 0.05

OR

Activation functions

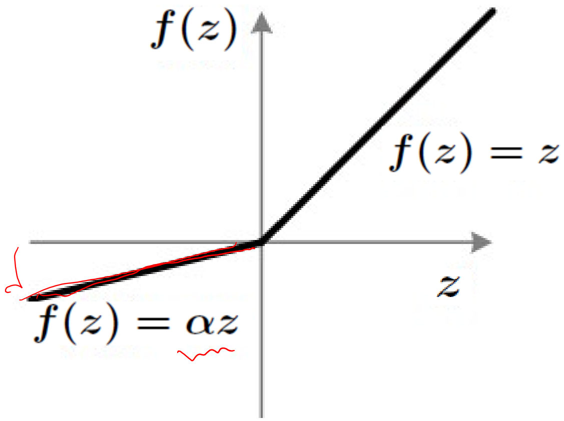
- Rectified Linear (ReLU)



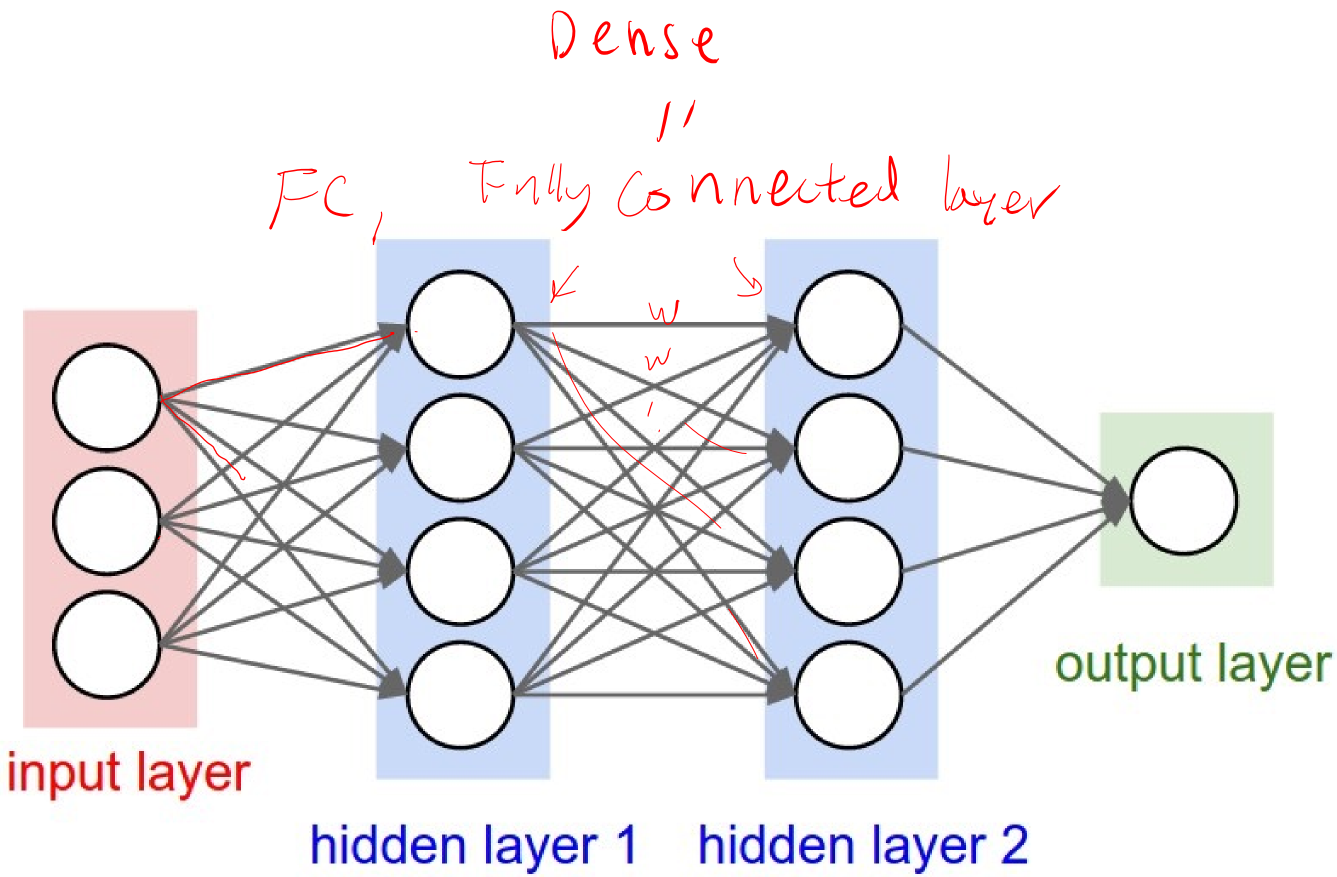
- Better convergence
- Does not saturate
- Less computation

regression

- Parametric Rectified Linear (PReLU) or Leaky ReLU



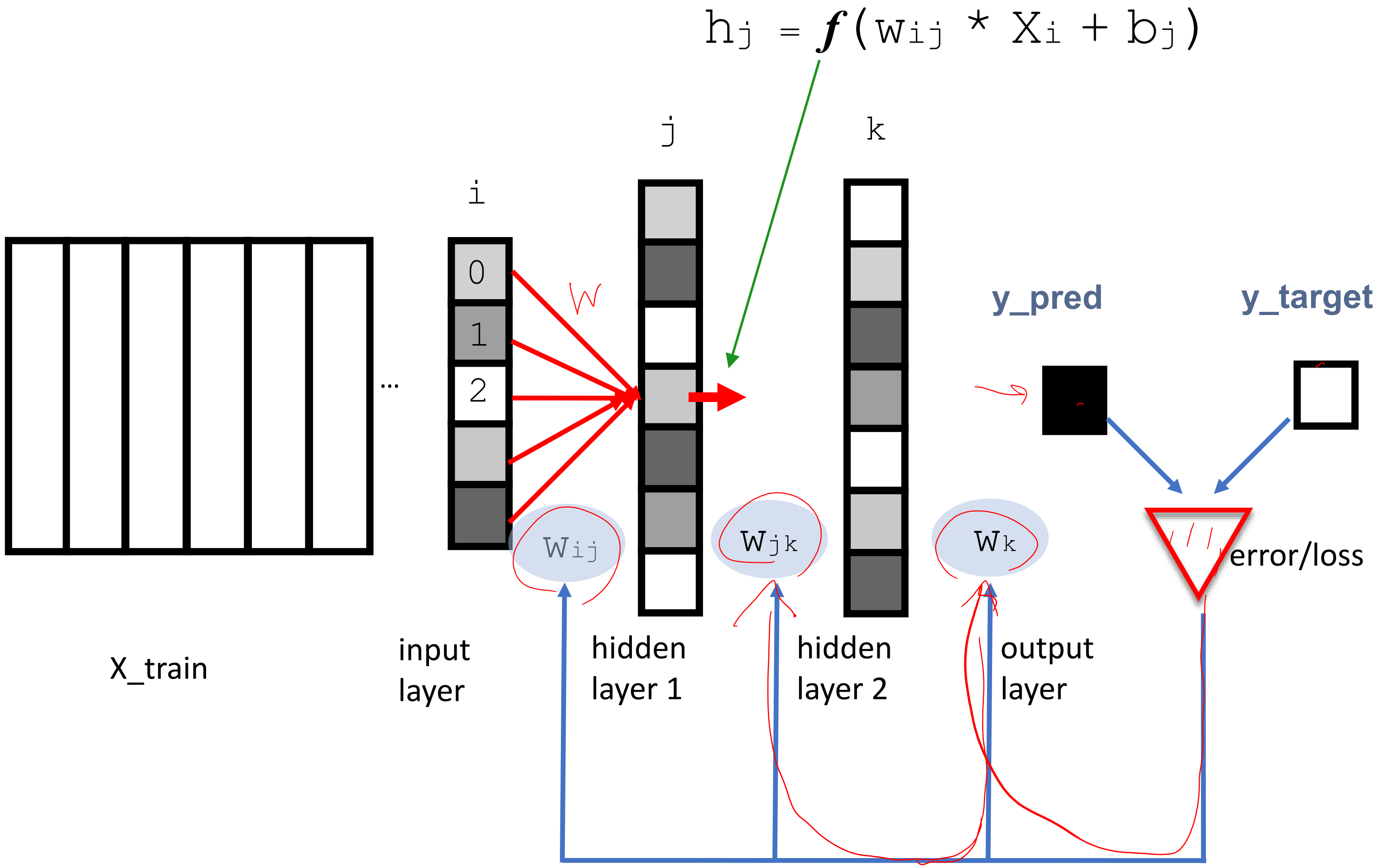
Multi-layer Perceptron (Neural network)



Design Parameters

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

How Neural Network Training Works



Weight update rule (Gradient descent)

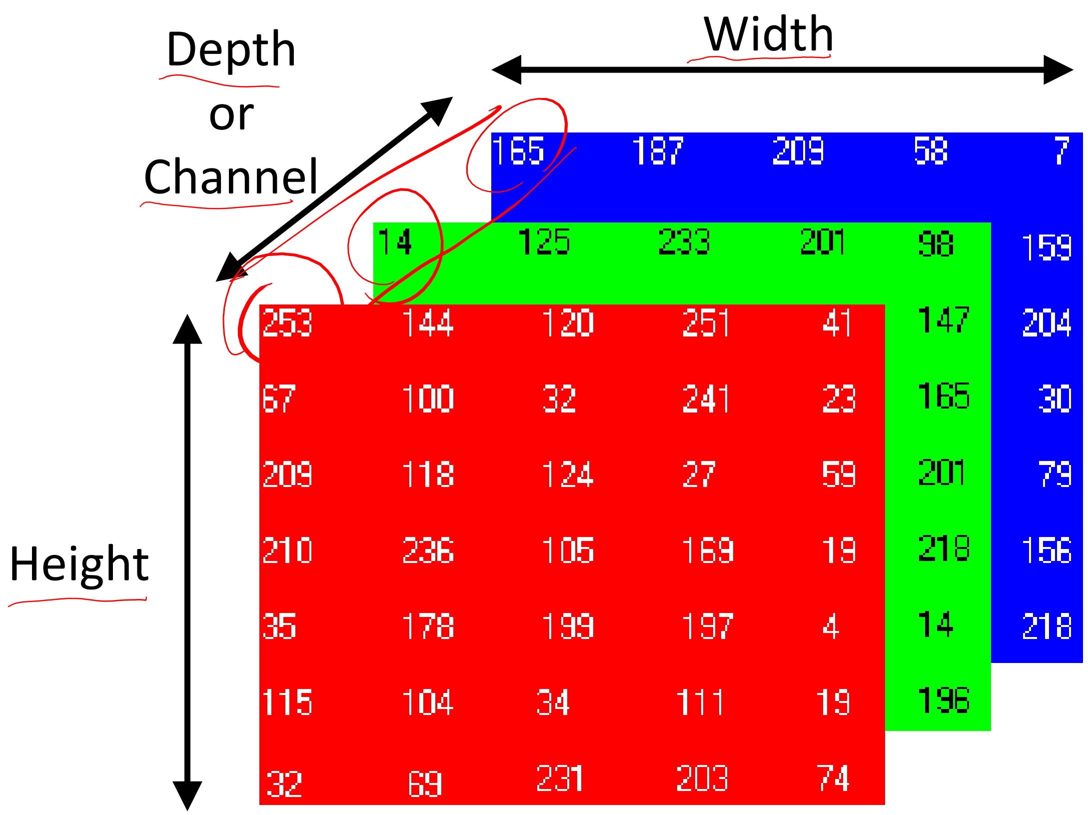
$$W_{ij} \leftarrow W_{ij} - \alpha \frac{\partial \mathcal{L}}{\partial W_{ij}}$$

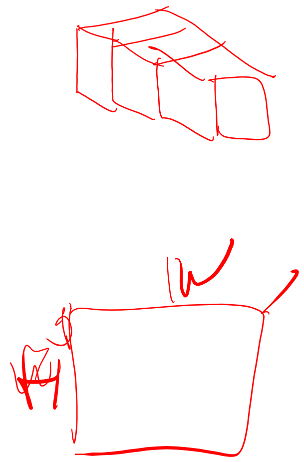
Chain rule

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

Dealing with images

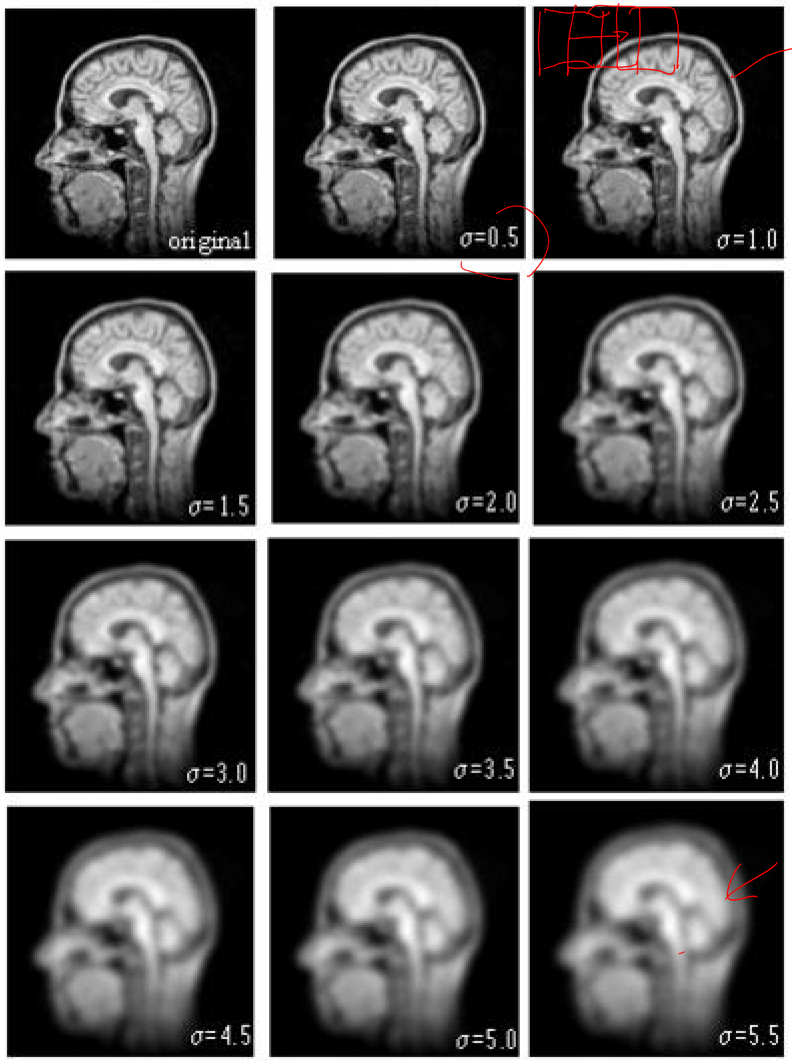
An image is a multi-dimension array





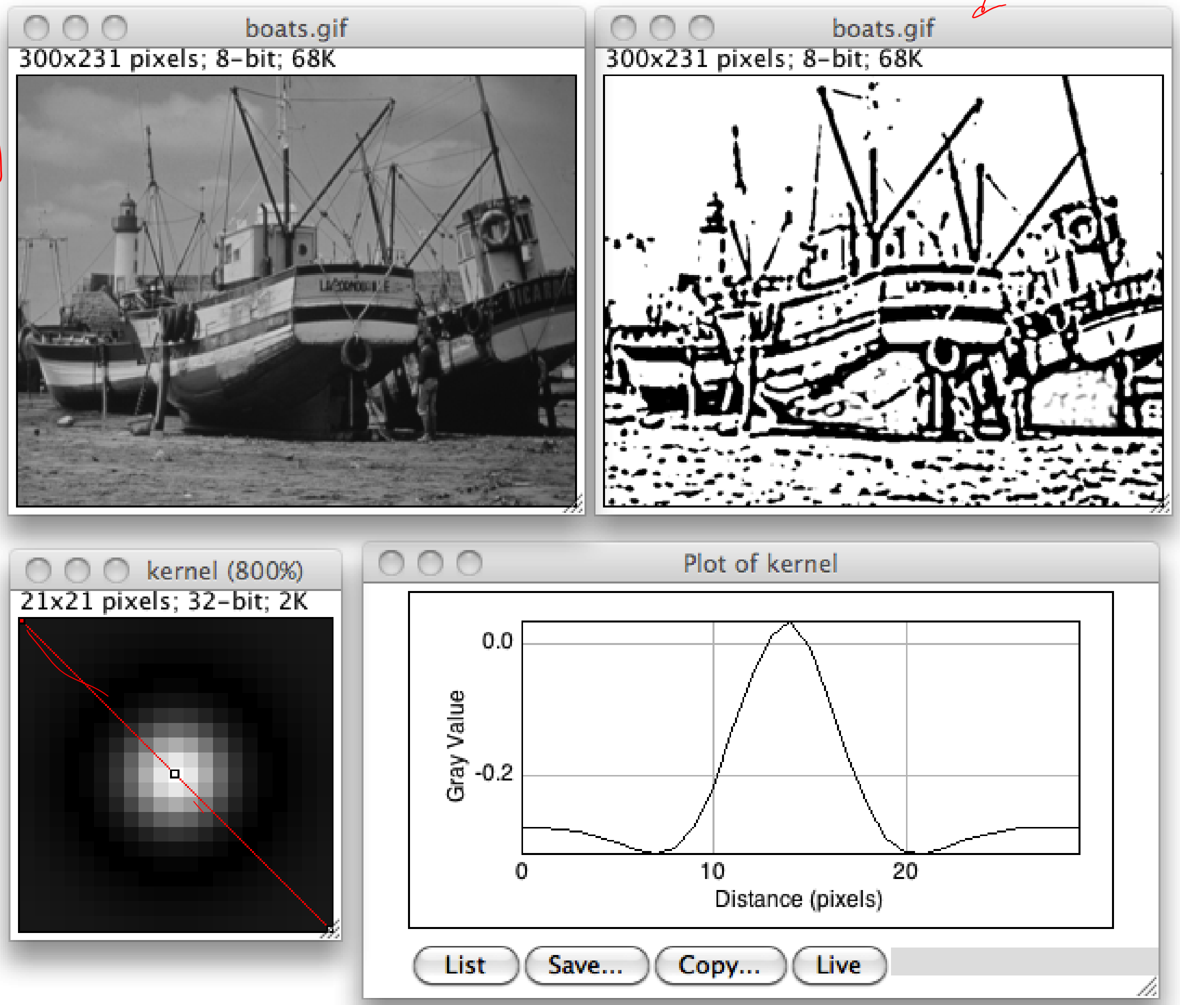
Dealing with images- filters

Gaussian Blurring



feature map

Mexican hat filtering



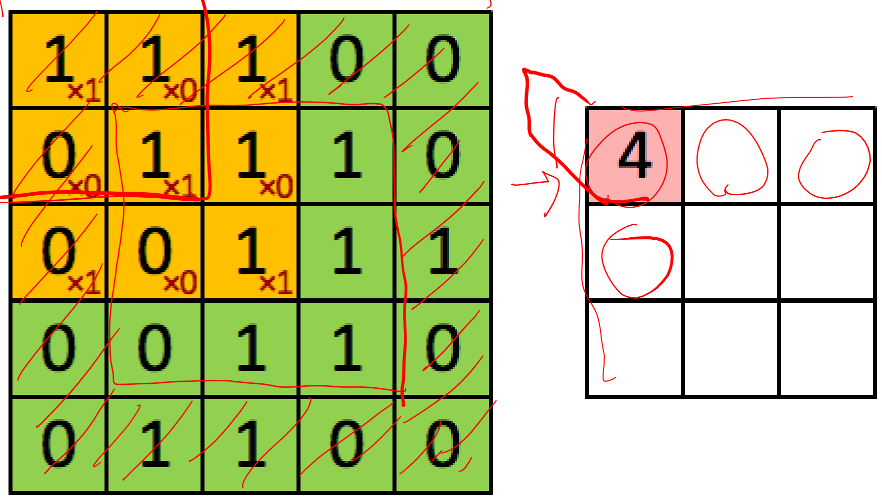
What is Convolution (2D) in an Image?

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

Image

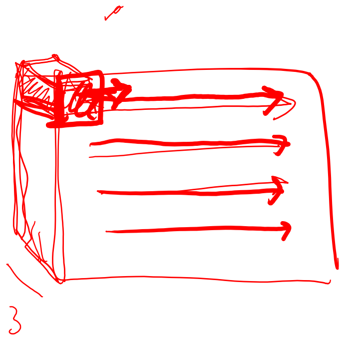
1	0	1
0	1	0
1	0	1

filter



Image

Convolved Feature

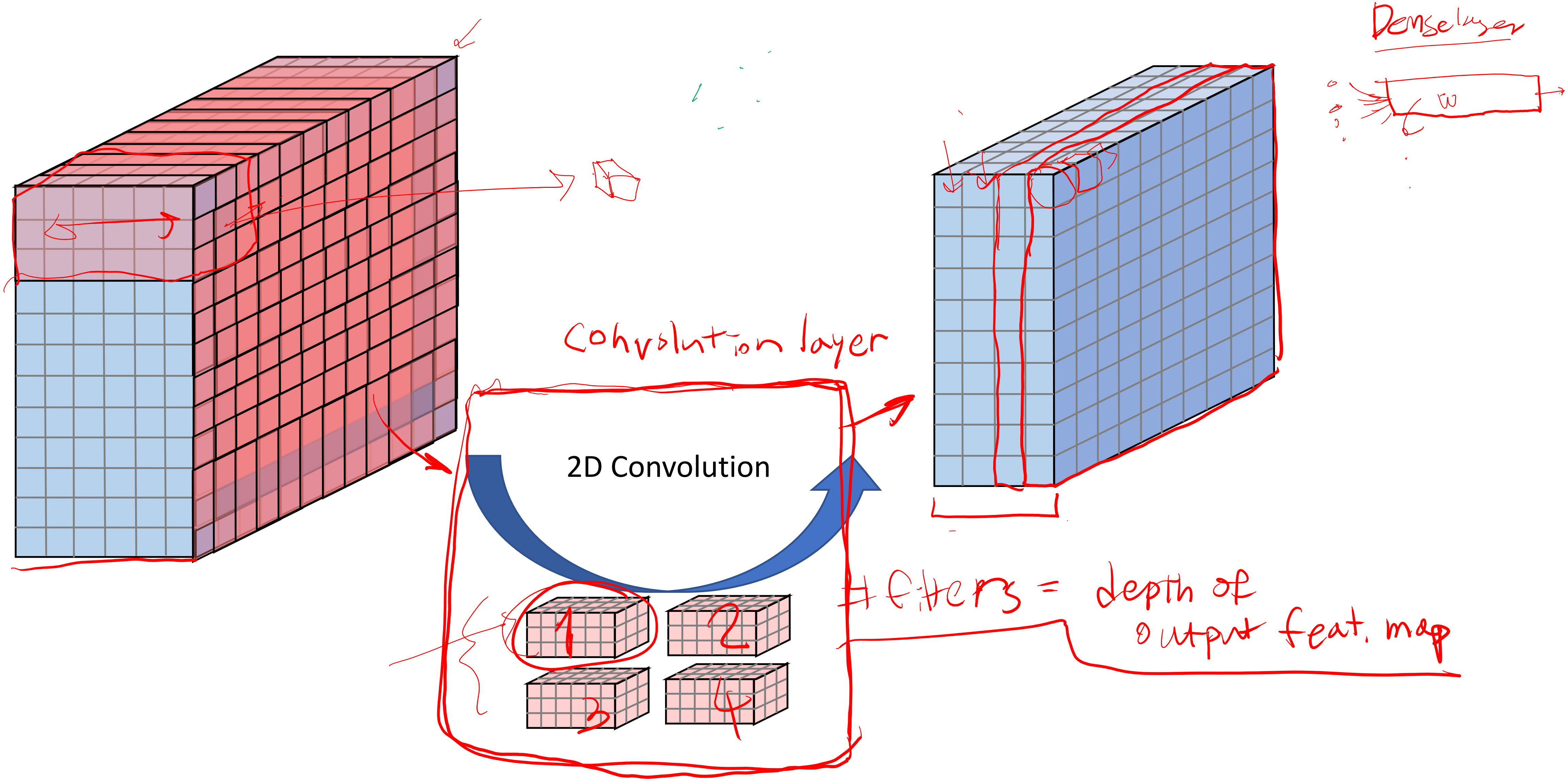


$$\left(\begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} * \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}\right)[2,2]$$

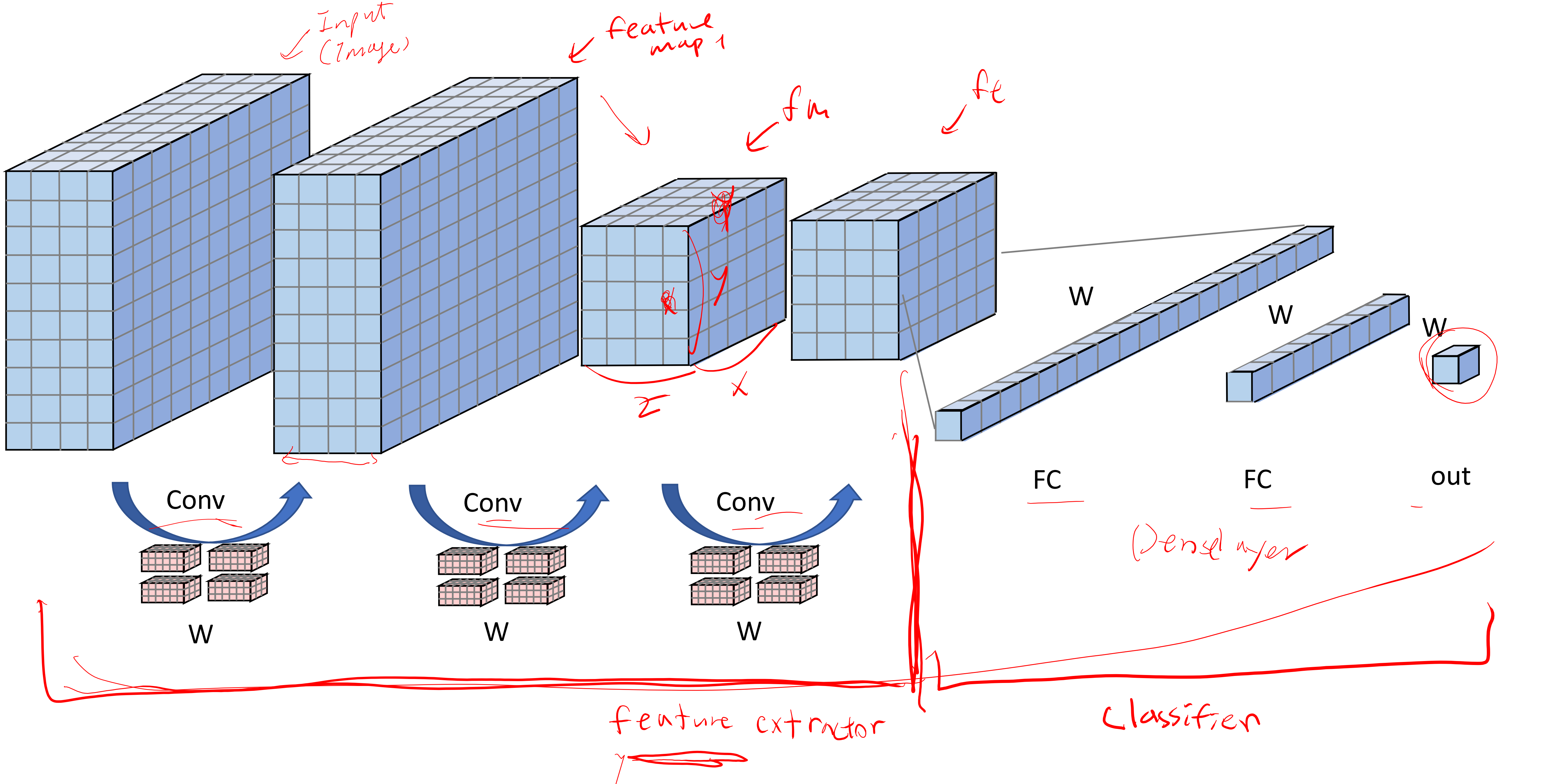
$$= (i \cdot 1) + (h \cdot 2) + (g \cdot 3) + (f \cdot 4) + (e \cdot 5) + (d \cdot 6) + (c \cdot 7) + (b \cdot 8) + (a \cdot 9)$$

Convolutional Layer

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

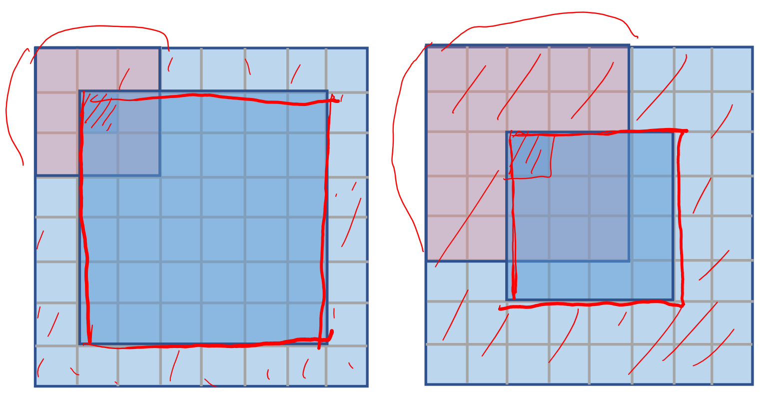


Convolutional Neural Network

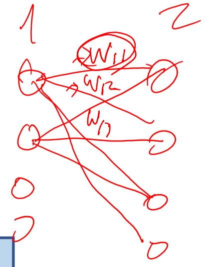


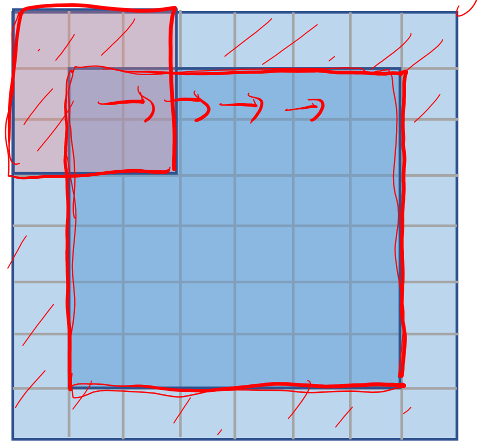
Convolutional Layer- hyper parameters

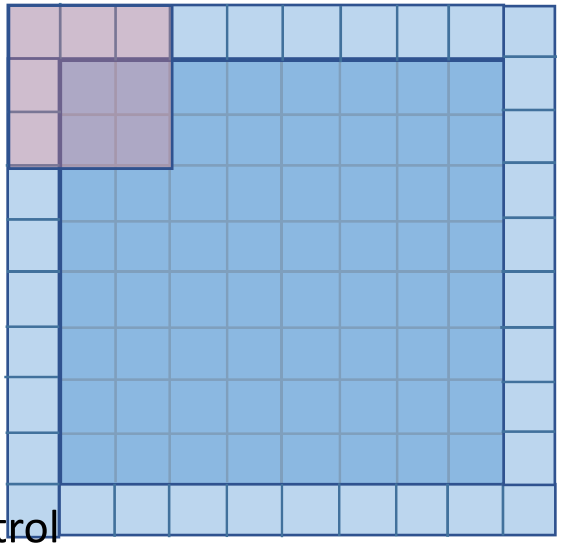
- Filter size (3x3, 5x5, 1x1....)



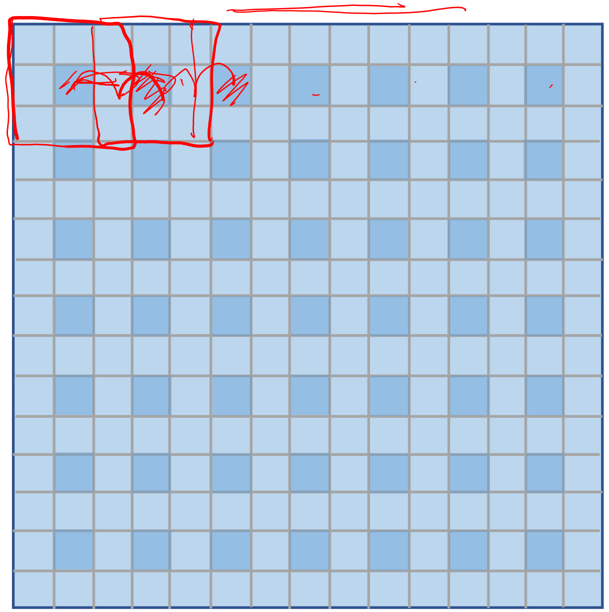
- Padding

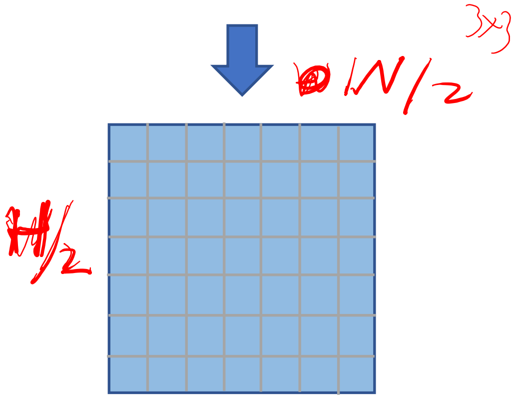




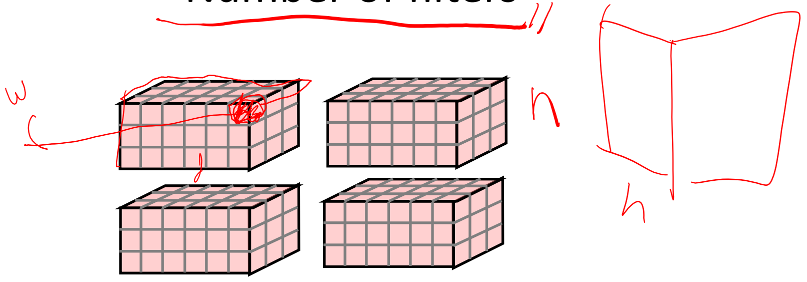


- Stride





- Number of filters



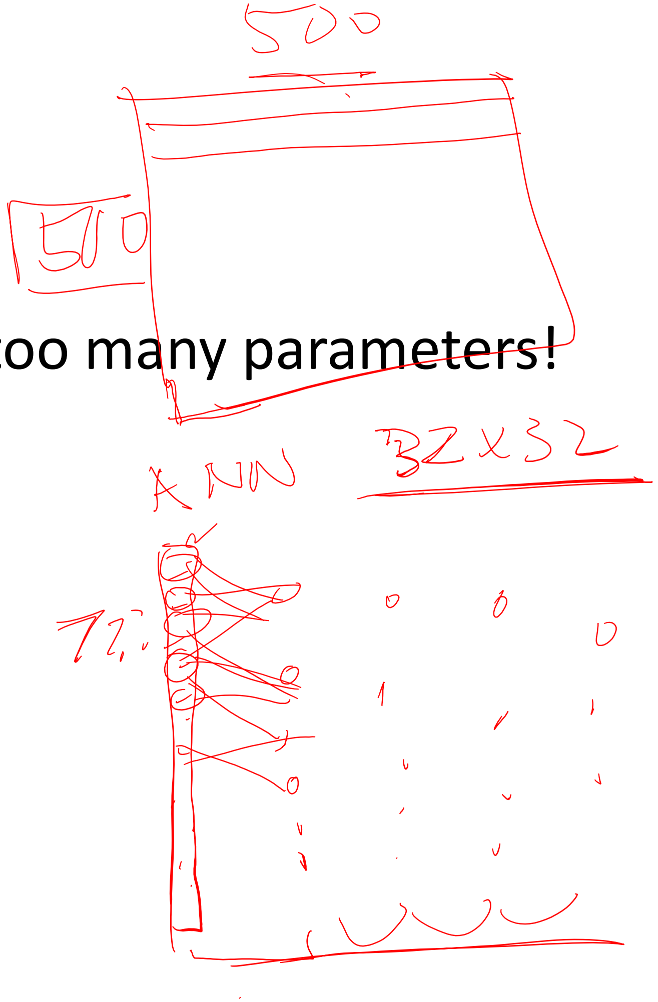
** Parameters and Hyper parameters are different!

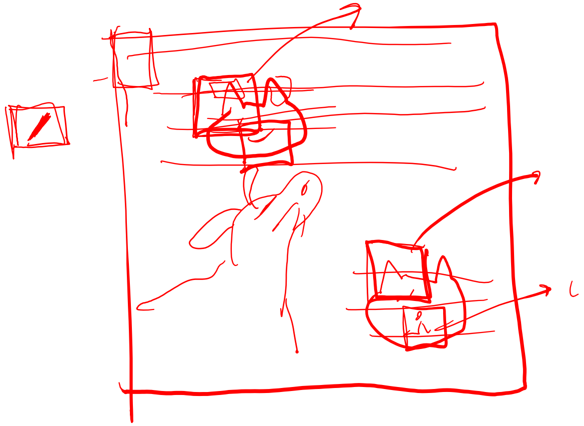
Parameters = Weights to be optimized.

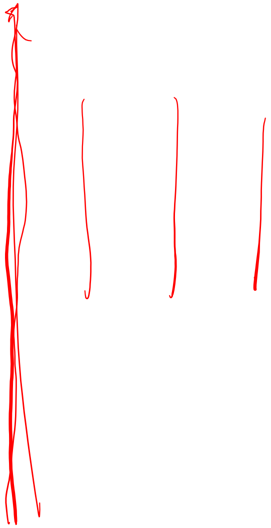
Hyperparameters = design parameters you can control

Why CNN? What do the convolution filters do?

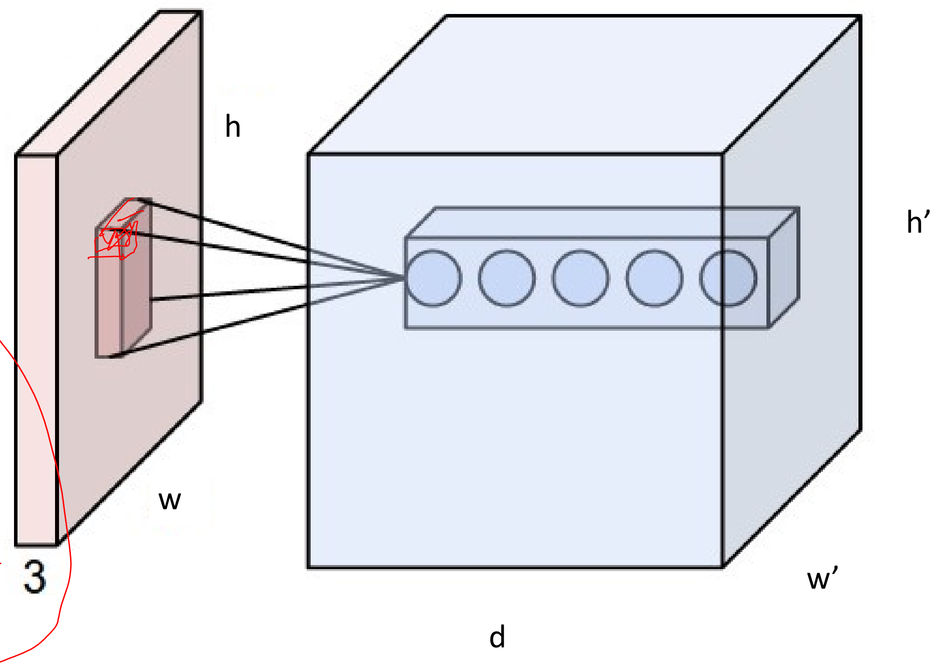
- Images have big pixels!
- Fully-connected neural network would have too many parameters!
- Translational invariance in images



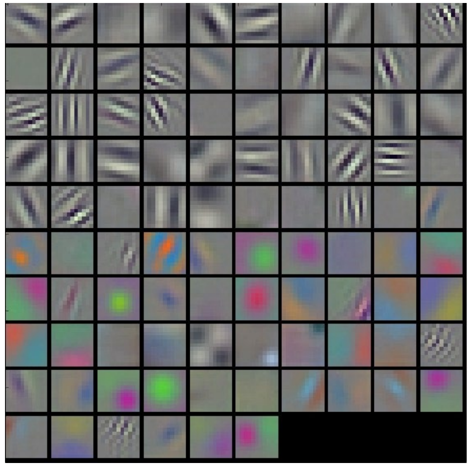




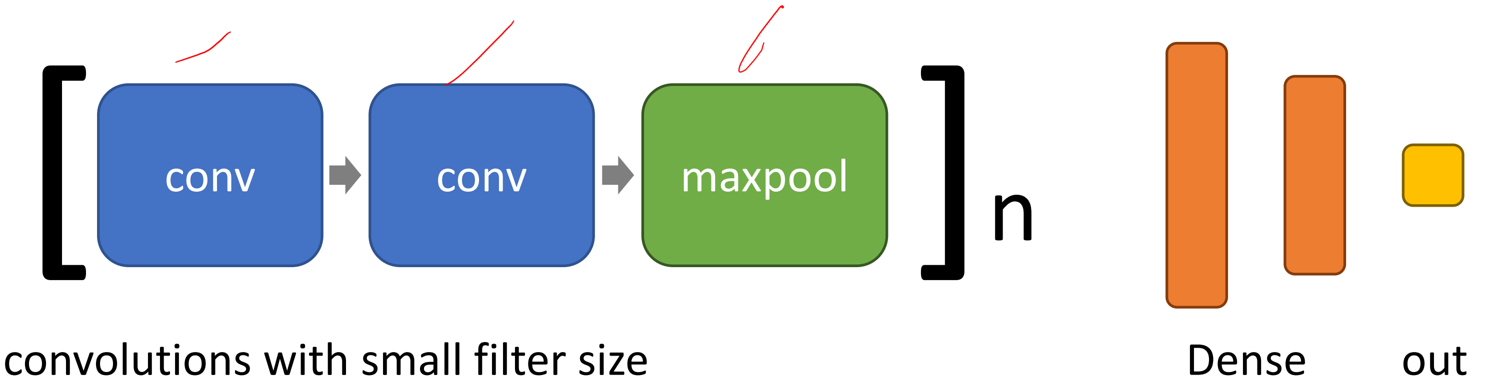
Why CNN? What do the convolution filters do?



- **Learning Filters**
- **Weight Sharing**
- **Computational Efficiency**
- **Translational Invariance**
- **Robust: less overfit**



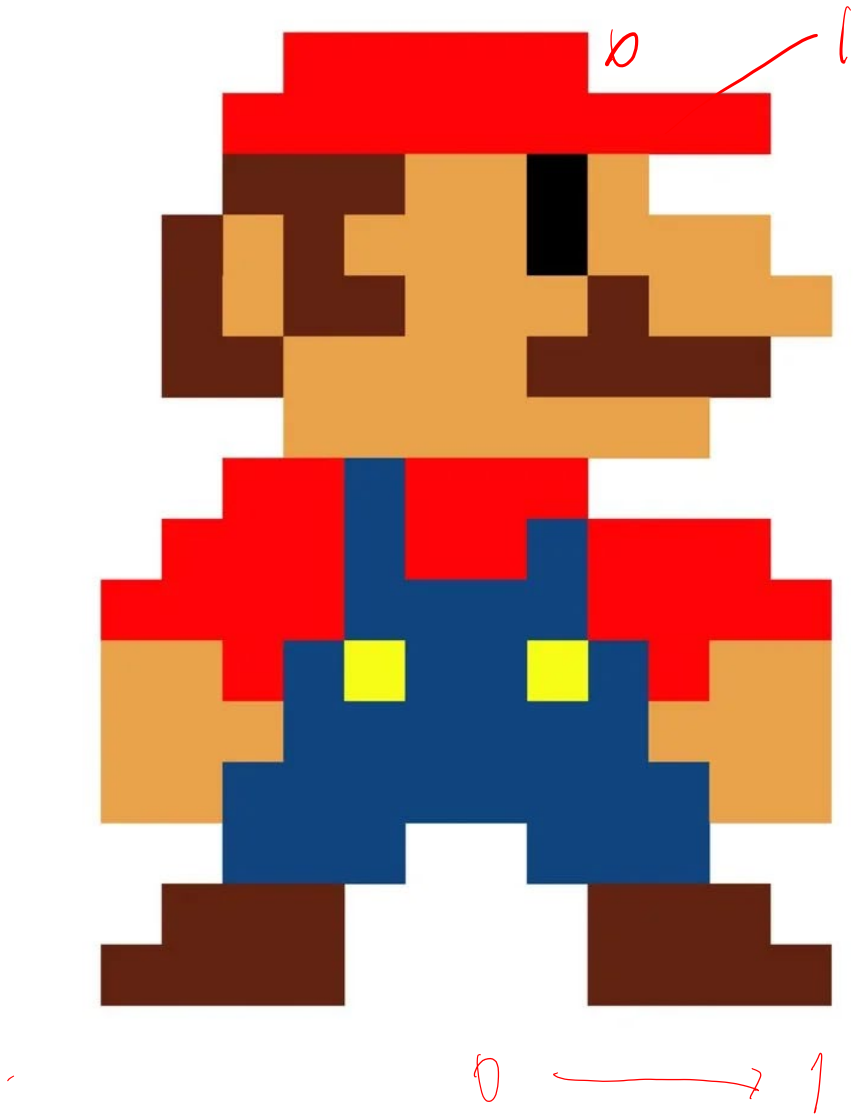
Typical CNN architecture

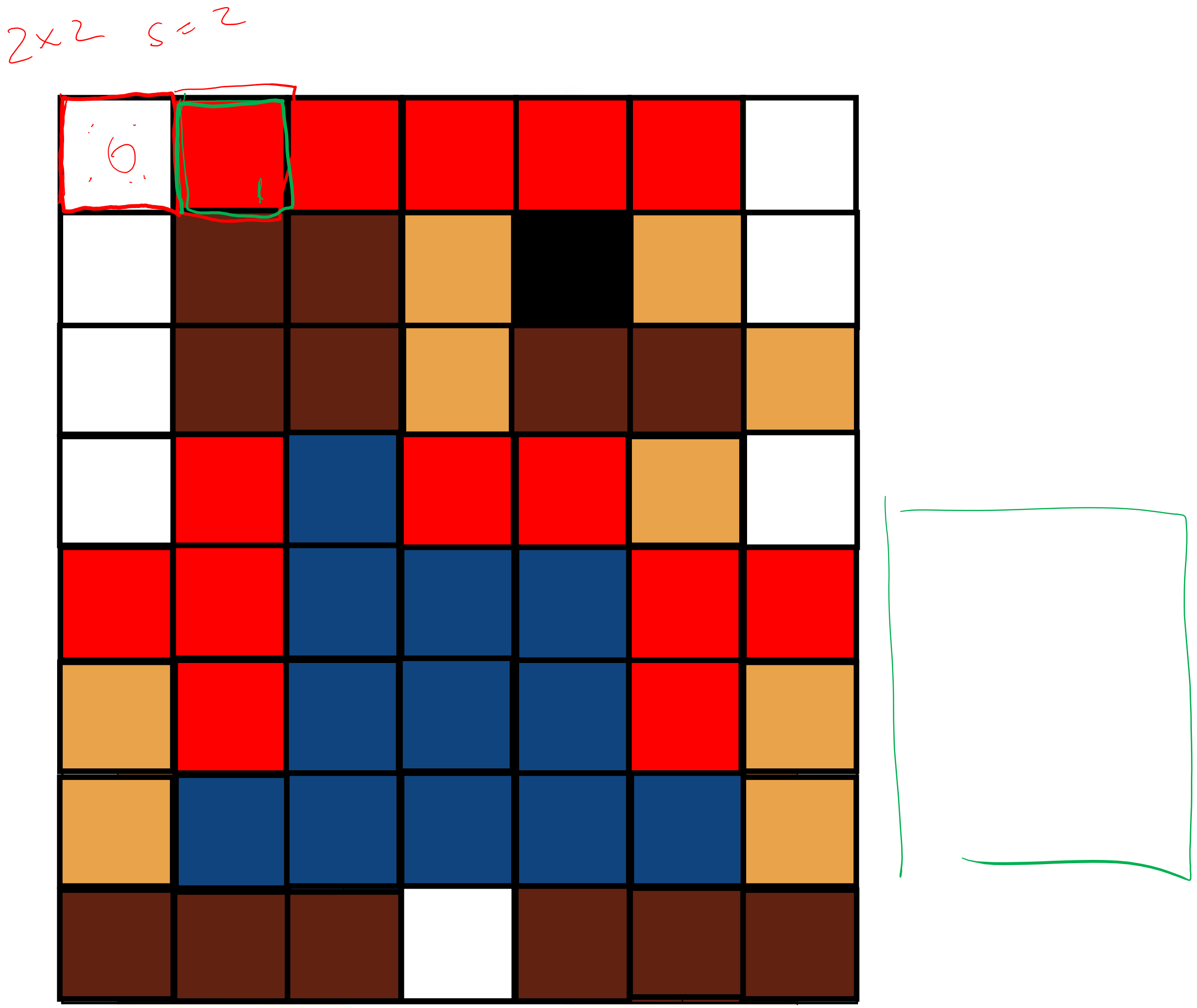


What is a Pooling layer?

- Pooling is like sub-sampling
- Pooling filter size usually is 2x2 (or 2n x 2n)
- Usually reduce the size to 1/N per each side (e.g. N=2 for 2x2)
- Max Pool
- Average Pool

Max pooling operation example

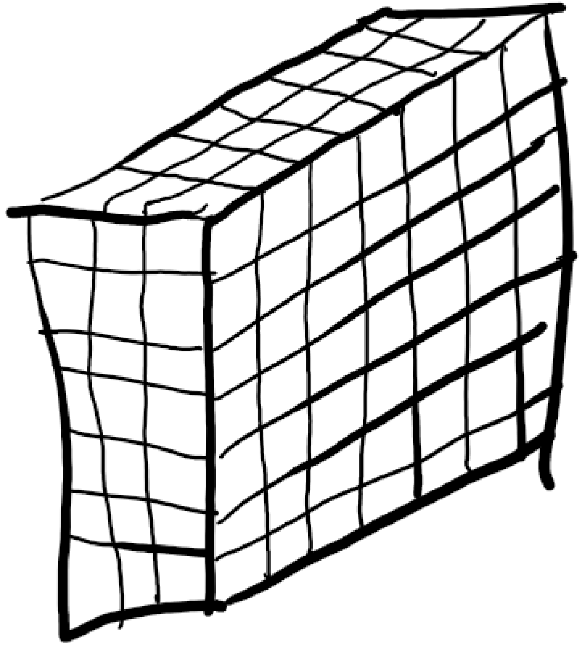




Another way to shrink the spatial dimension

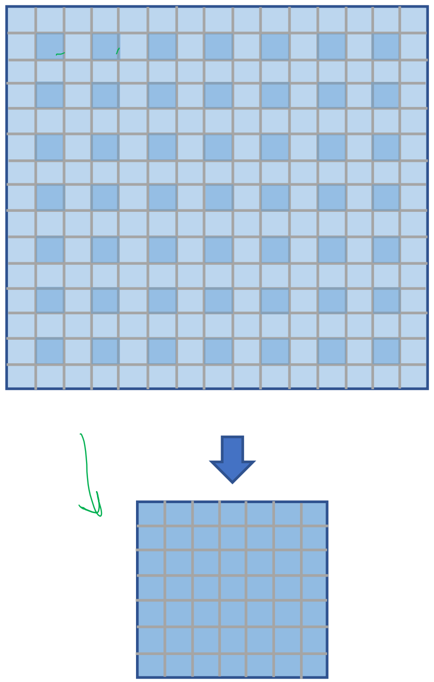
We can reduce the spatial dimension using convolution filters with a bigger stride

Exercise



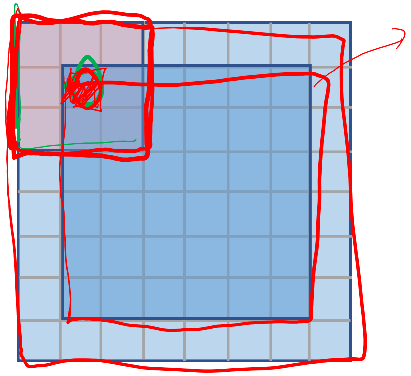
conv(3x3 , s=2)
w/o padding

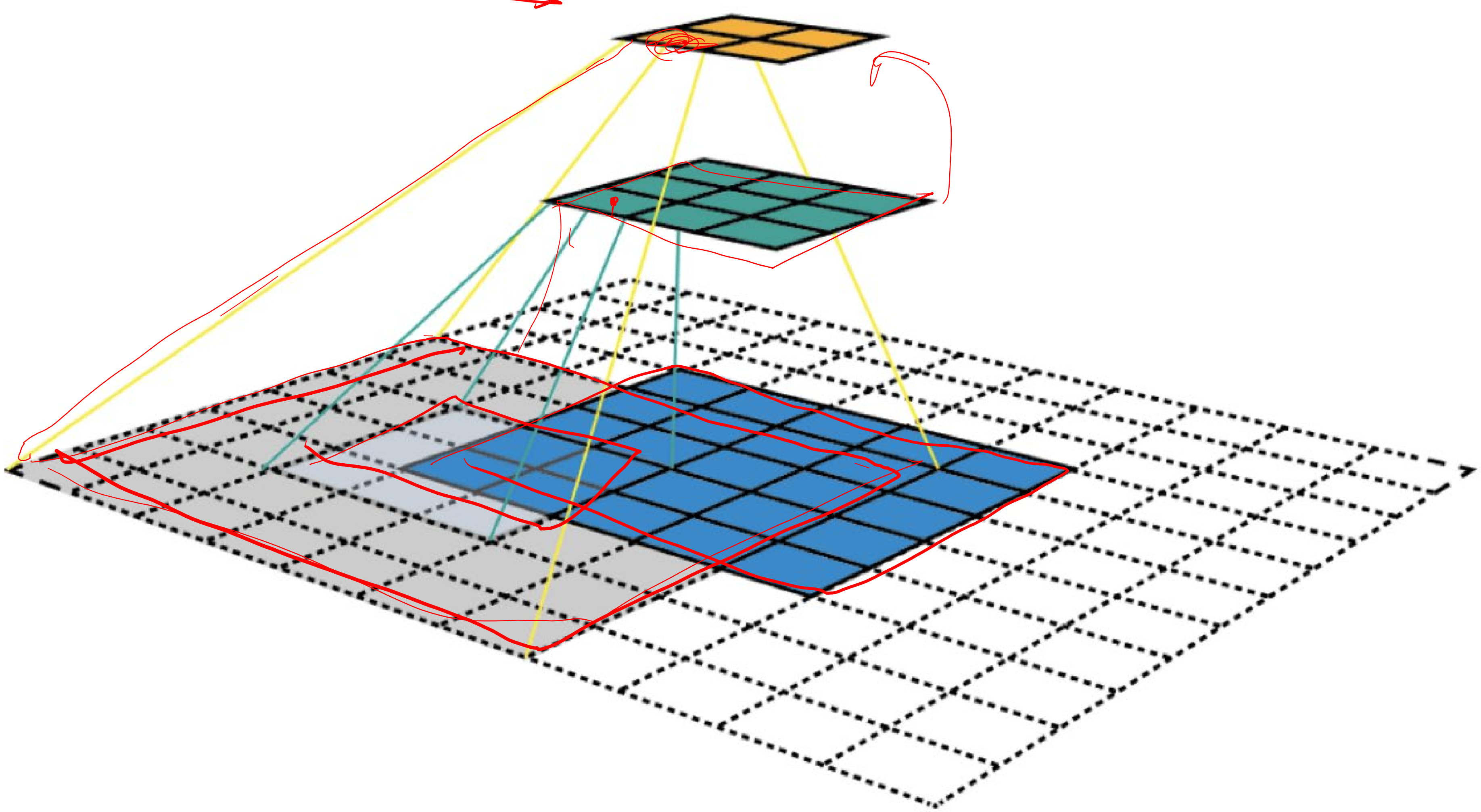
- Stride



What do the multiple convolution layers do?

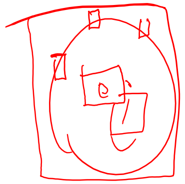
Field of view



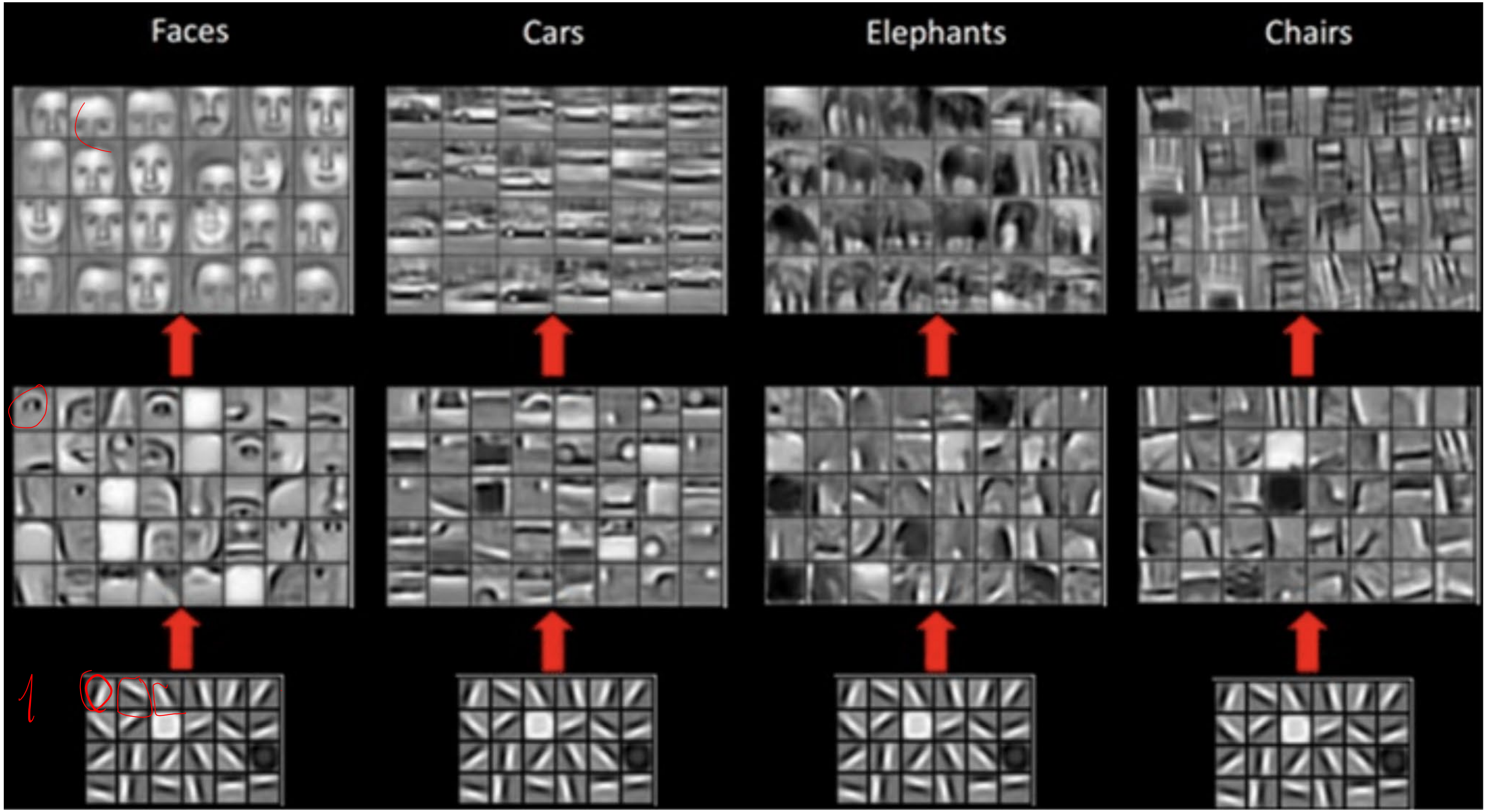


What do the multiple convolution layers do?

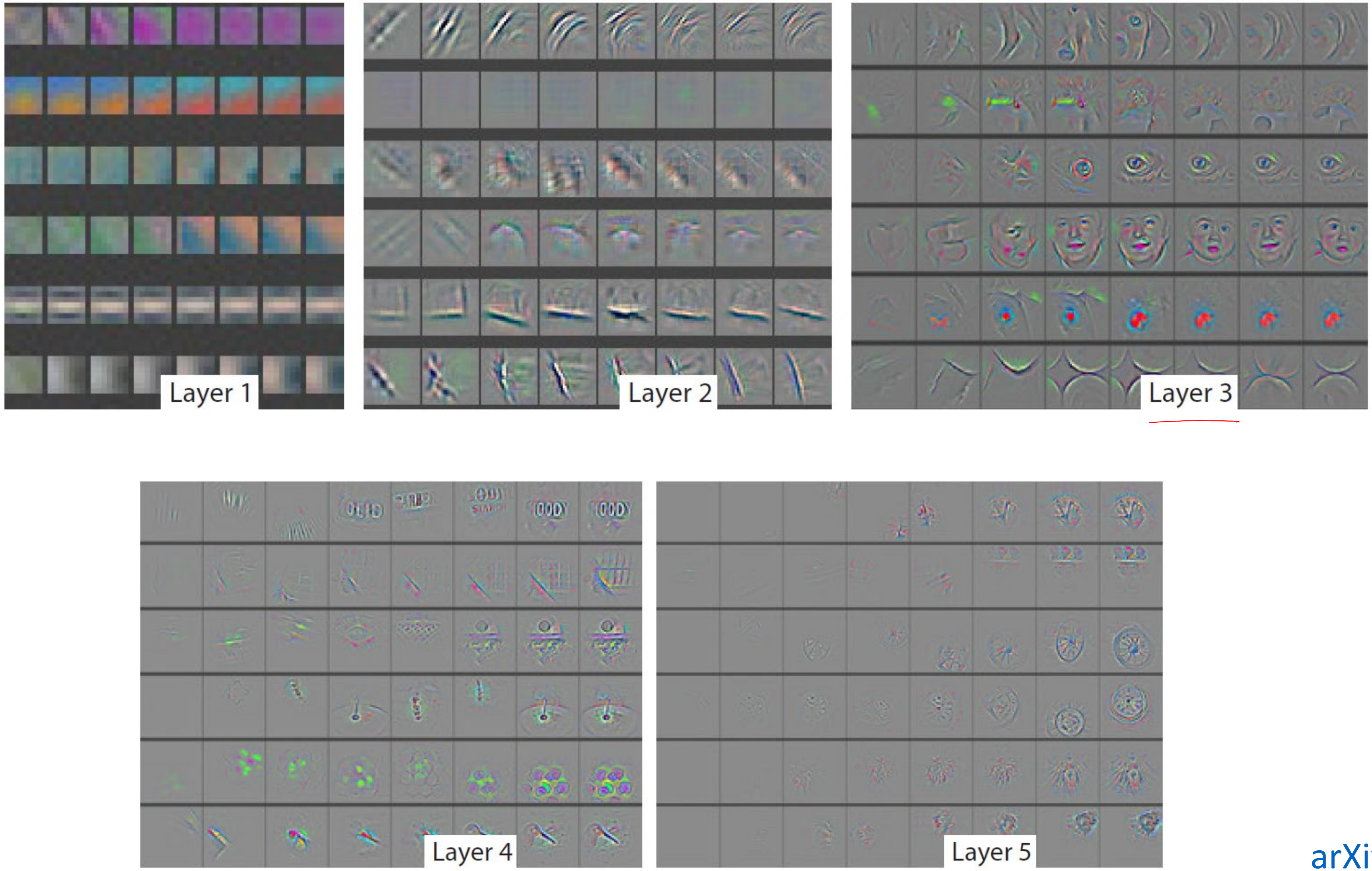
Geometrical shapes are made of pieces of small curves, lines, and color blobs



1



What do the multiple convolution layers do?



arXiv:1311.2901v3

Next time!

- CNN Architectures
- How to train CNNs