

AI in Computer Vision & Facial Recognition

CSE 352
Prof. Anita Wasilewska

Team 9
Haoran Yang
Wangshu Luo
Xiaohan Chen
Zijun He



Sources

<https://www.youtube.com/watch?v=40riCqvRoMs&t=936s>

<https://research.googleblog.com/2016/10/how-robots-can-acquire-new-skills-from.html>

http://cs231n.stanford.edu/slides/2017/cs231n_2017_lecture1.pdf

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

<http://www3.cs.stonybrook.edu/~cse352/L12NN.pdf>

<https://www.newscientist.com/article/2114900-concerns-as-face-recognition-tech-used-to-identify-criminals/>

<https://www.oreilly.com/ideas/the-ethics-of-face-recognition>

<https://www.cnet.com/news/iphone-xs-face-id-supposedly-got-hacked-we-have-questions/>

<https://www.appannie.com/en/apps/ios/app/b612-selfie-from-the-heart/>

<https://blog.emojipedia.org/apples-new-animoji/>

<http://wvtf.org/post/peeling-onion>

<http://stanfordflipside.com/2017/05/going-against-the-grain-this-dog-is-for-sure-going-to-hell/>

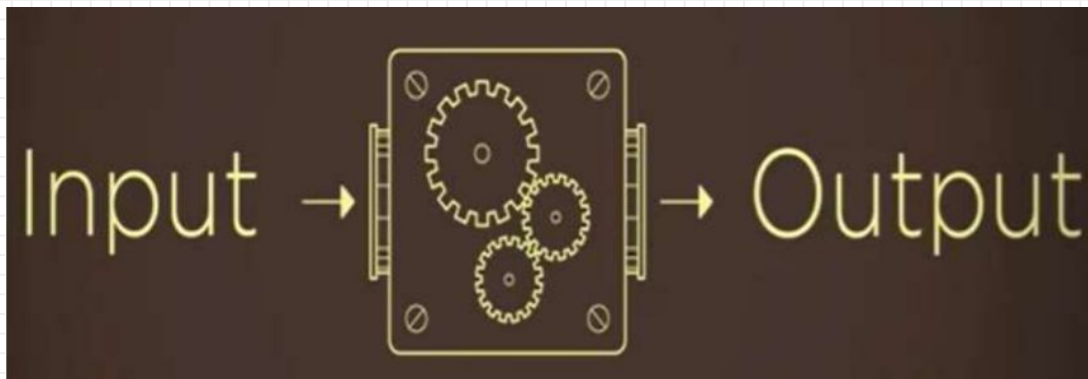
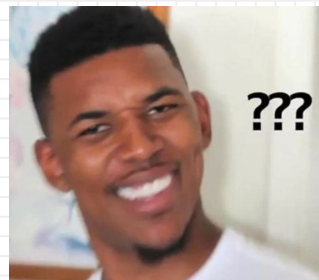


Overview

- Object Recognition
- Algorithm in Facial Recognition
- Applications in Facial Recognition
- Concerns in Facial Recognition



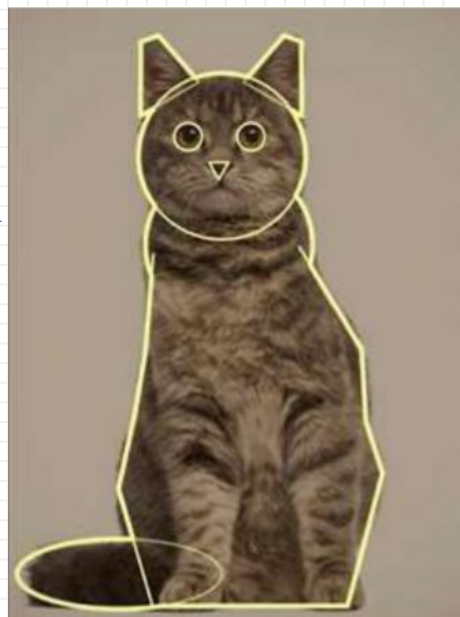
How does machine recognize an object?



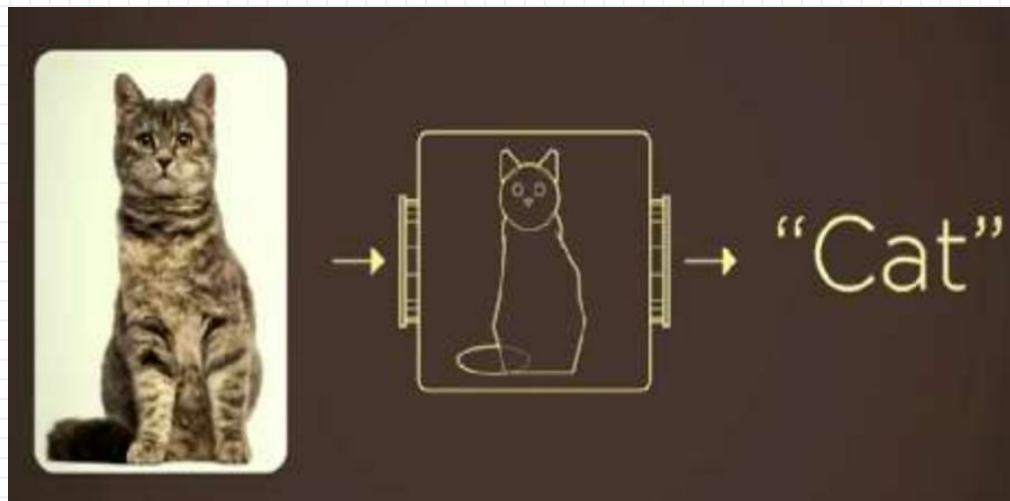
We give computer the algorithms and tell it what to do



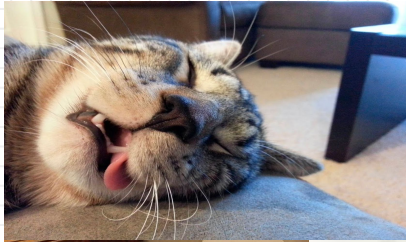
Design Model



- Round Face
- Pointy Ears
- Chubby Body
- Long Tail



What about these cats?



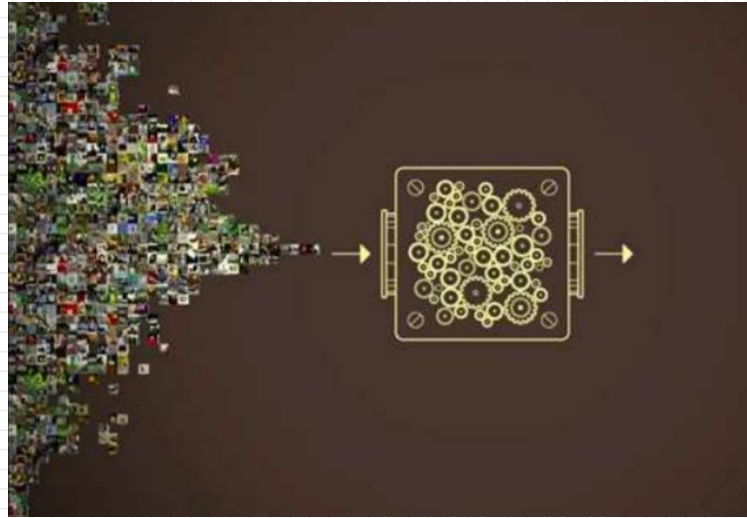
?



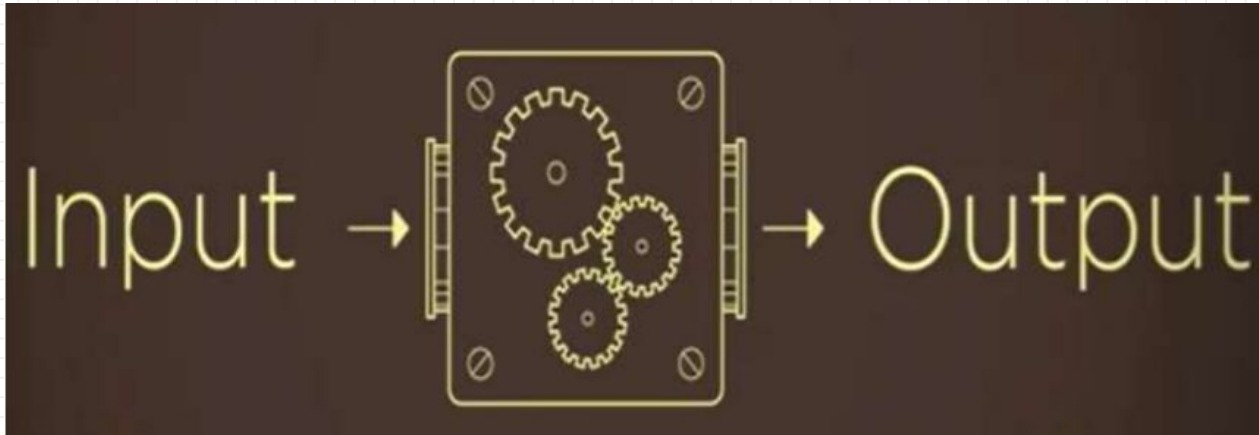
?

A child can learn about the world without being told everything.

Instead of focusing on how to write better algorithms, we use large amount of data to train the computer.



Deep Learning



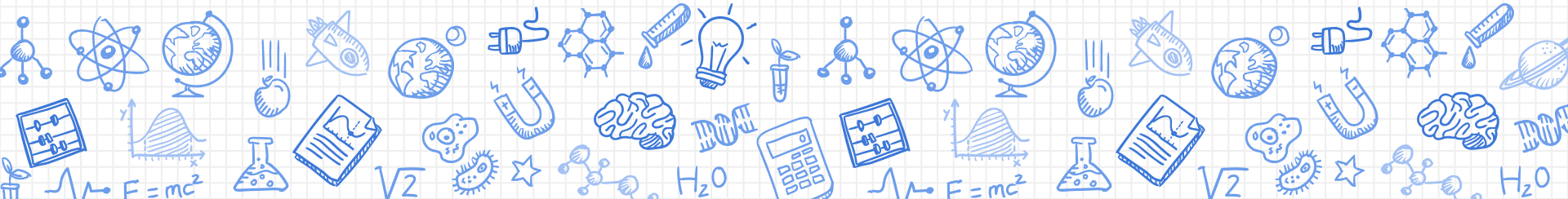
No longer tell computer what to do, let it learn by itself

Unsupervised Learning



A Quick Game!!! 🎈

Let's learn some Chinese



Now u r an algorithm-free machine

What does 踢 mean in Chinese?

- A. Eat
- B. Fart
- C. Kick
- D. Ball

What does 踏 mean in Chinese?

- A. Step On
- B. Knee Down
- C. Push up
- D. Cook

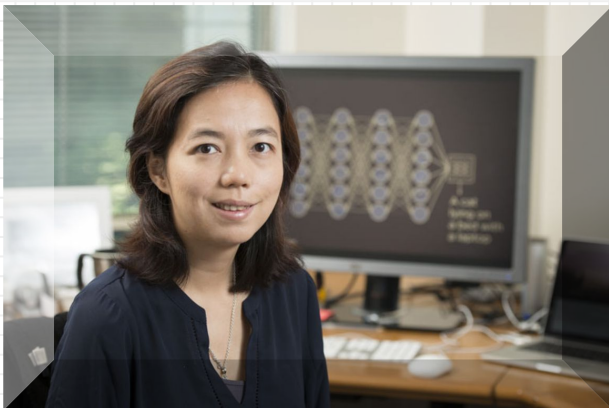
What does 蹂 mean in Chinese?

- A. Sneak
- B. Trample
- C. Pour
- D. Pick

What does 蹾 mean in Chinese?

- A. Wash
- B. Whisper
- C. Jump
- D. Cry





IMAGENET

14,197,122 images, 21841 synsets indexed

[Explore](#) [Download](#) [Challenges](#) [Publications](#) [CoolStuff](#) [About](#)

Not logged in. [Login](#) | [Signup](#)

ImageNet is an image database organized according to the **WordNet** hierarchy (currently only the nouns), in which each node of the hierarchy is depicted by hundreds and thousands of images. Currently we have an average of over five hundred images per node. We hope ImageNet will become a useful resource for researchers, educators, students and all of you who share our passion for pictures.

[Click here](#) to learn more about ImageNet, [Click here](#) to join the ImageNet mailing list.



Statistics of high level categories

High level category	# synset (subcategories)	Avg # images per synset	Total # images
amphibian	94	591	56K
animal	3822	732	2799K
appliance	51	1164	59K
bird	856	949	812K
covering	946	819	774K
device	2385	675	1610K
fabric	262	690	181K
fish	566	494	280K
flower	462	735	339K
food	1495	670	1001K
fruit	309	607	188K
fungus	303	453	137K
furniture	187	1043	195K
geological formation	151	838	127K
invertebrate	728	573	417K
mammal	1138	821	934K
musical instrument	157	891	140K
plant	1666	600	999K

Since 2010, the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) is a competition where research teams evaluate their algorithms on the given data set, and compete to achieve higher accuracy on several visual recognition tasks



48940 Workers
167 Countries

The Image Classification Challenge:

1,000 object classes
1,431,167 images



Output:

Scale
T-shirt
Steel drum
Drumstick
Mud turtle



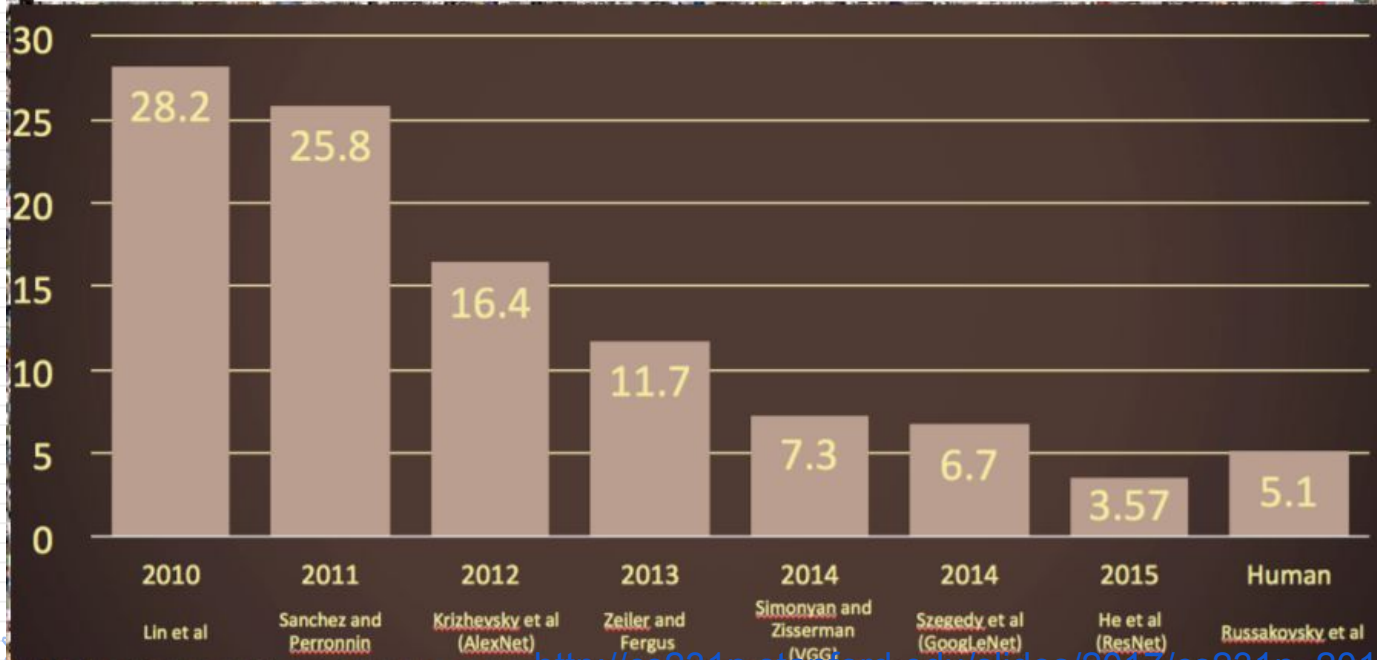
Output:

Scale
T-shirt
Giant panda
Drumstick
Mud turtle

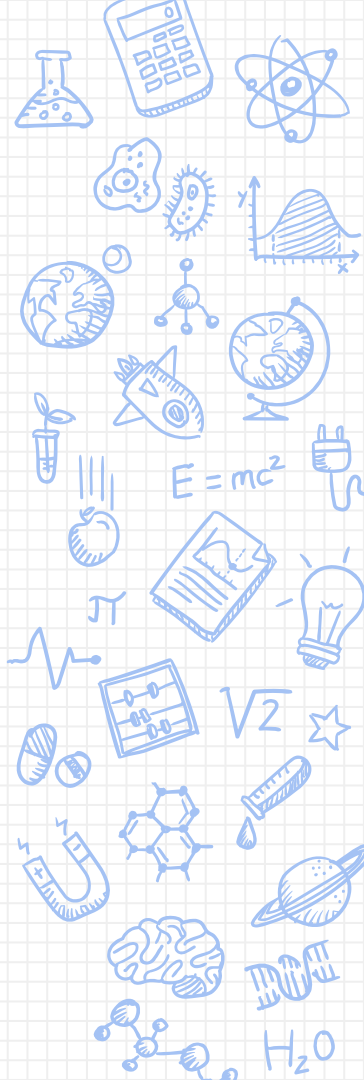


The Image Classification Challenge:

1,000 object classes
1,431,167 images



Facial Recognition





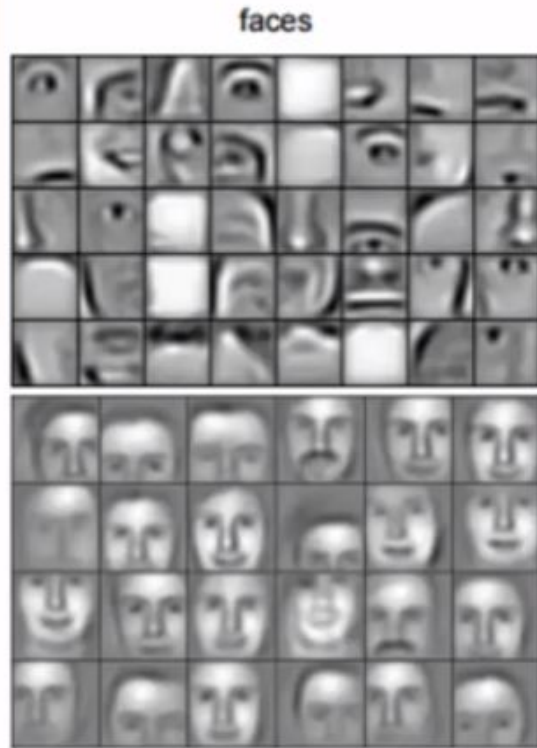
Algorithm in Facial Recognition

Convolutional Neural Network (CNN)

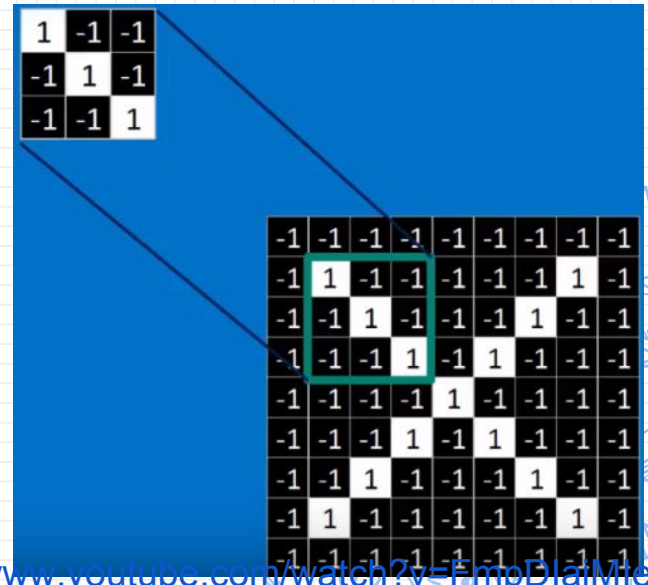
- Convolution
- Normalization
- Pooling

Convolution

Features



Filtering



Filtering

The diagram illustrates the process of filtering a 9x9 grid. It shows a 3x3 kernel being applied to a 9x9 grid to produce a 7x7 filtered grid.

Kernel:

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

Original 9x9 Grid:

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

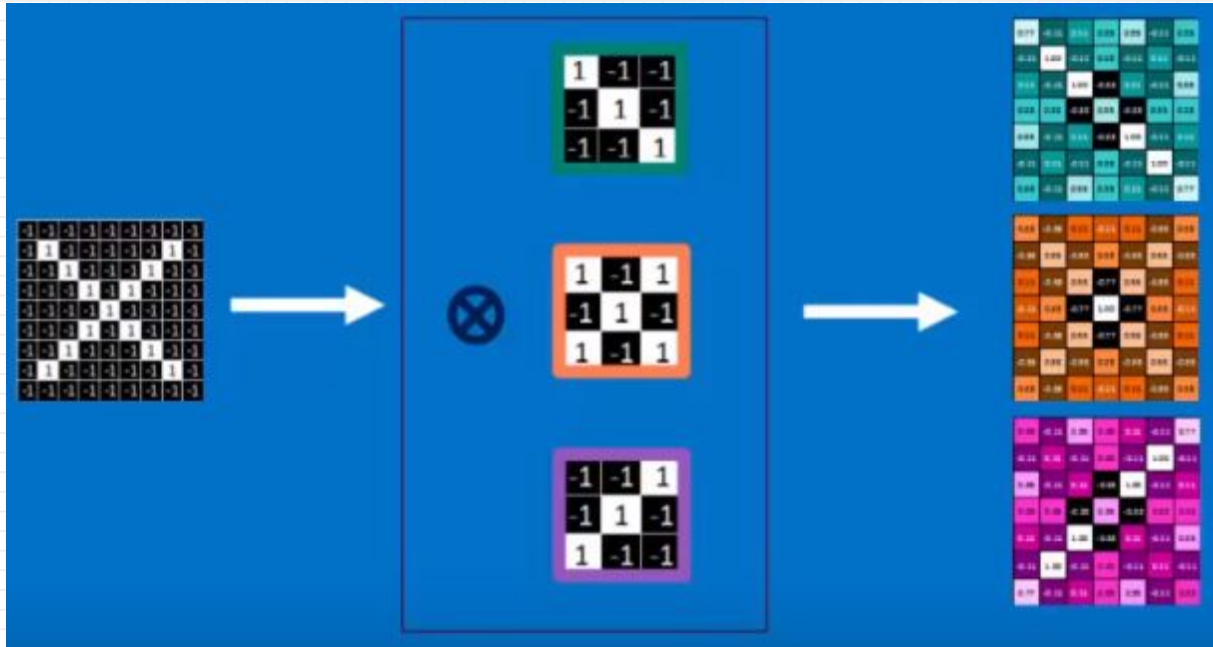
Filtered 7x7 Grid:

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

Calculation:

$$\frac{1 + 1 - 1 + 1 + 1 + 1 - 1 + 1 + 1}{9} = .55$$

Filtering

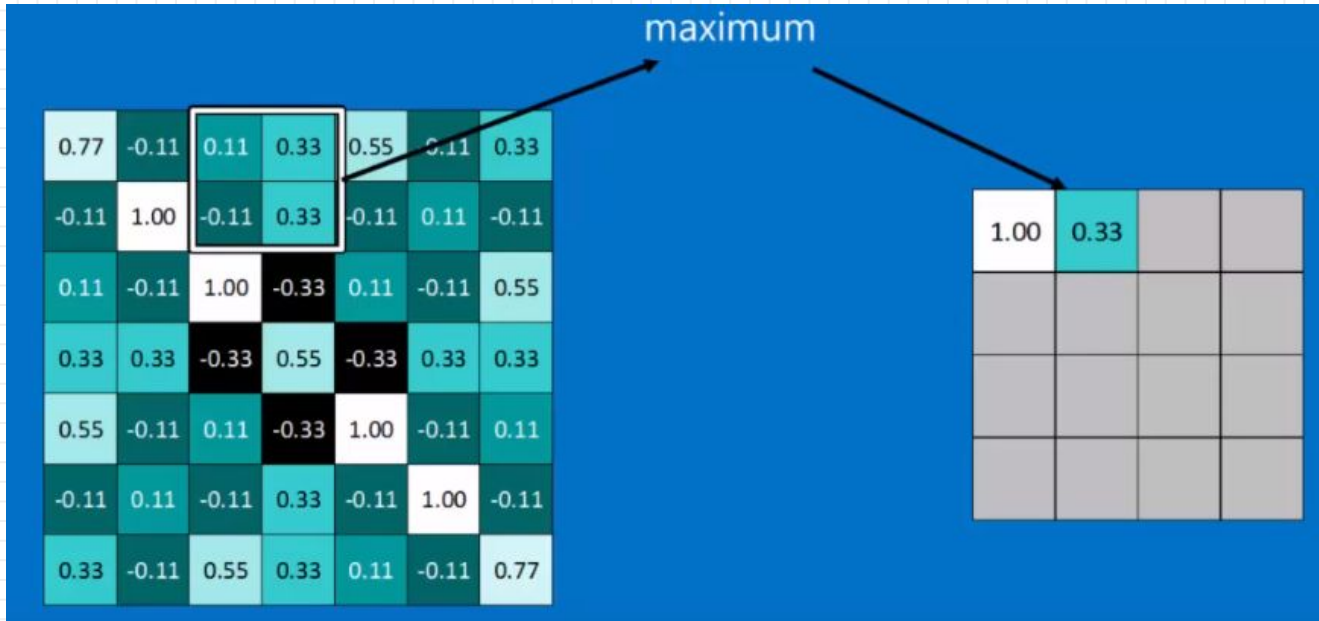


Normalization

Rectified Linear Units(ReLUs)



Pooling



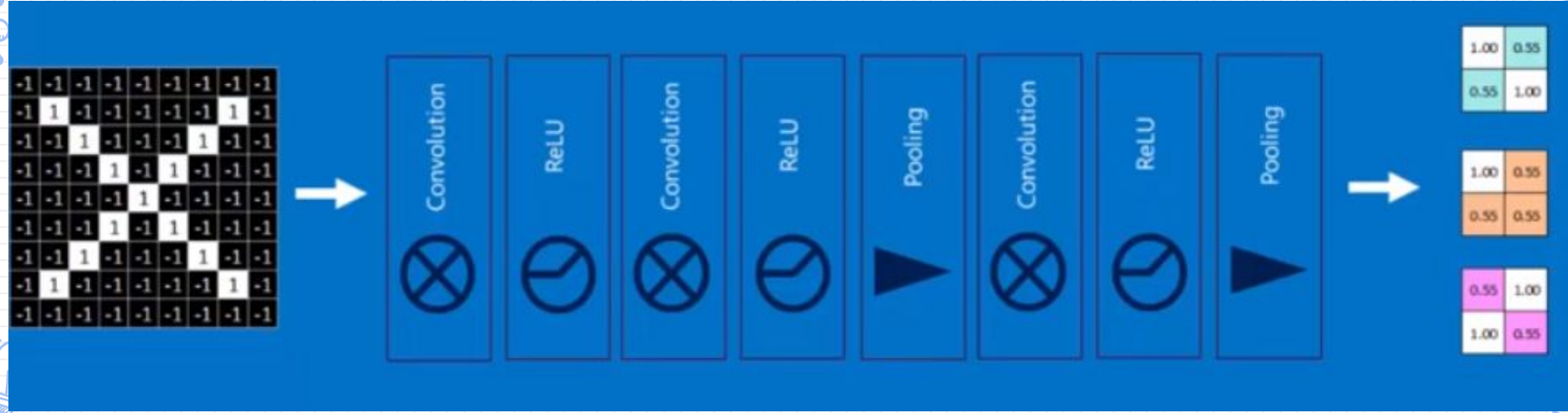
Pooling

0.77	-0.11	0.11	0.33	0.55	-0.11	0.33
-0.11	1.00	-0.11	0.33	-0.11	0.11	-0.11
0.11	-0.11	1.00	-0.33	0.11	-0.11	0.55
0.33	0.33	-0.33	0.55	-0.33	0.33	0.33
0.55	-0.11	0.11	-0.33	1.00	-0.11	0.11
-0.11	0.11	-0.11	0.33	-0.11	1.00	-0.11
0.33	-0.11	0.55	0.33	0.11	-0.11	0.77

max pooling

1.00	0.33	0.55	0.33
0.33	1.00	0.33	0.55
0.55	0.33	1.00	0.11
0.33	0.55	0.11	0.77

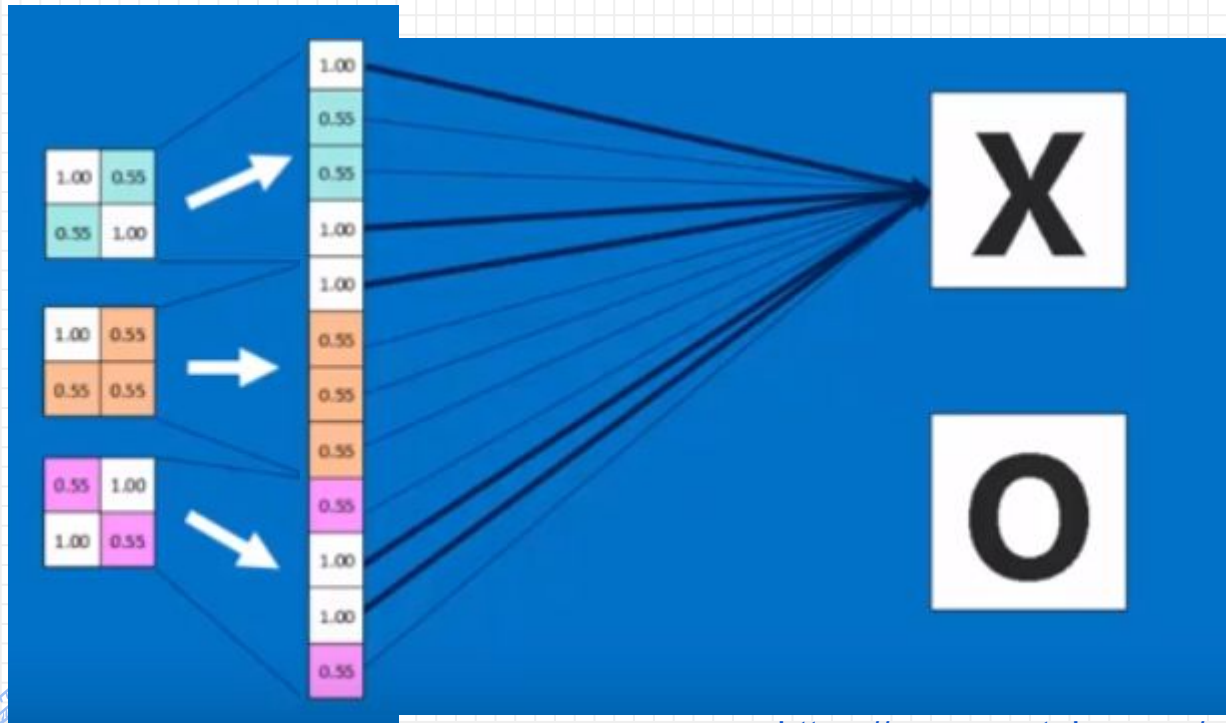
Stacking



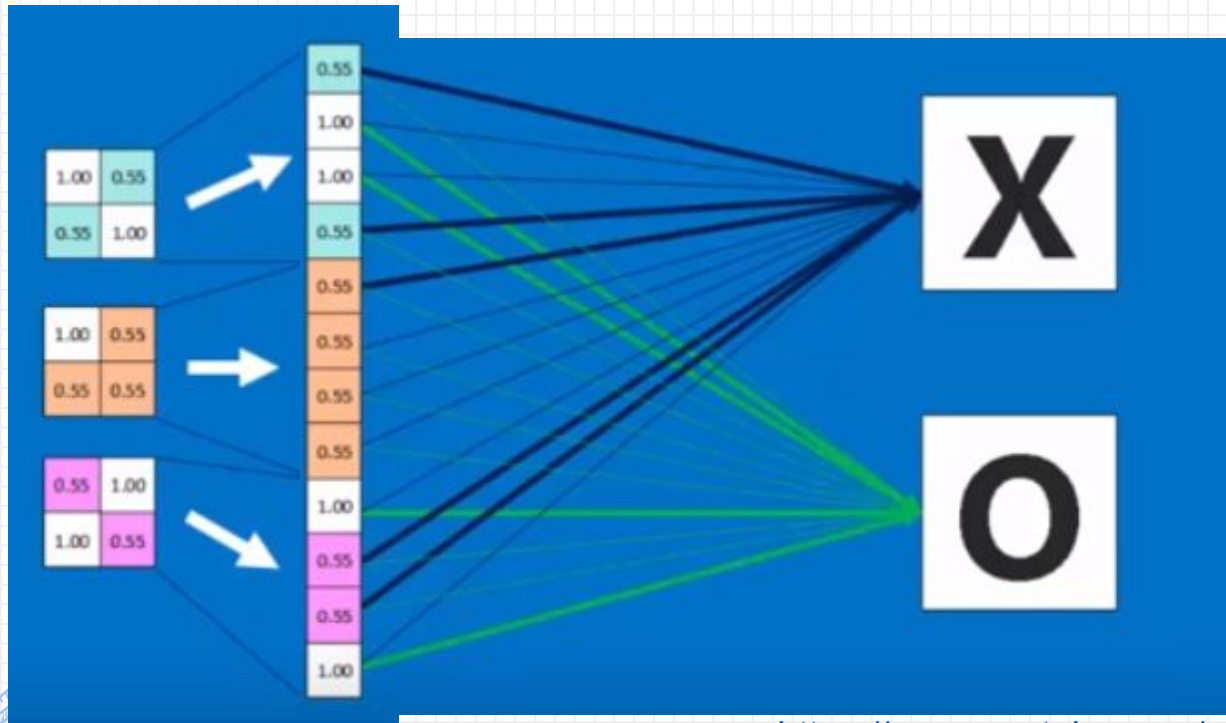
Fully Connected Layer



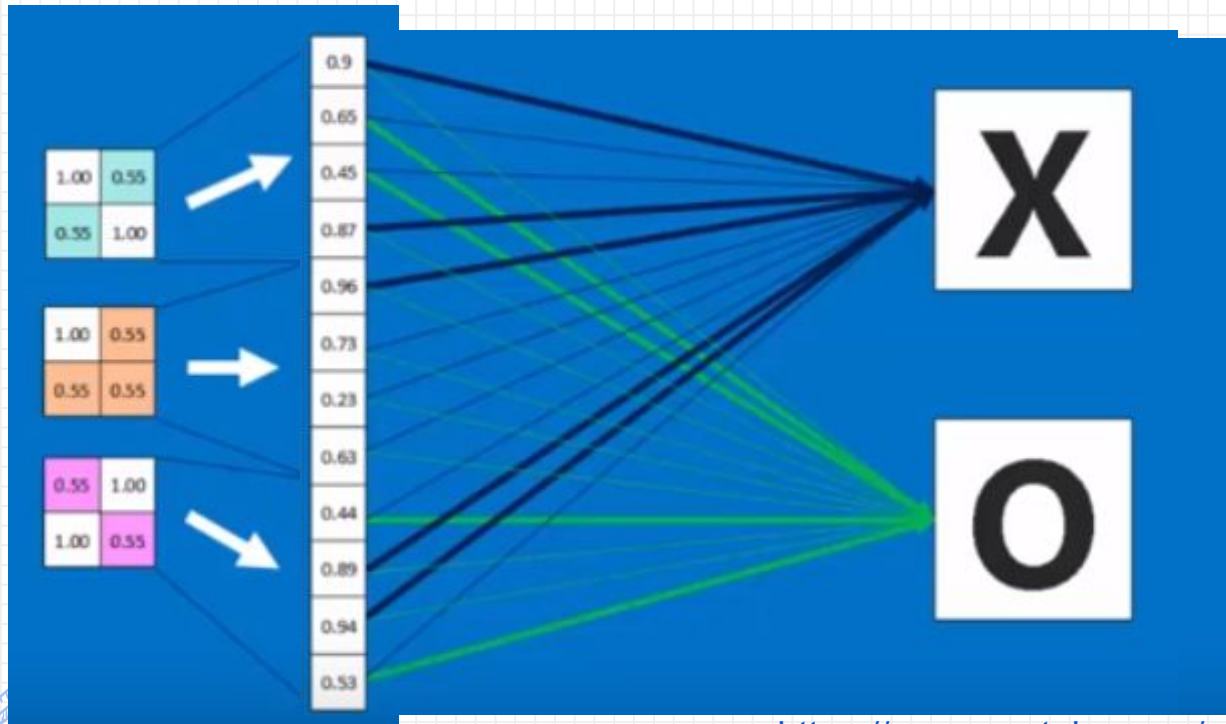
Fully Connected Layer



Fully Connected Layer

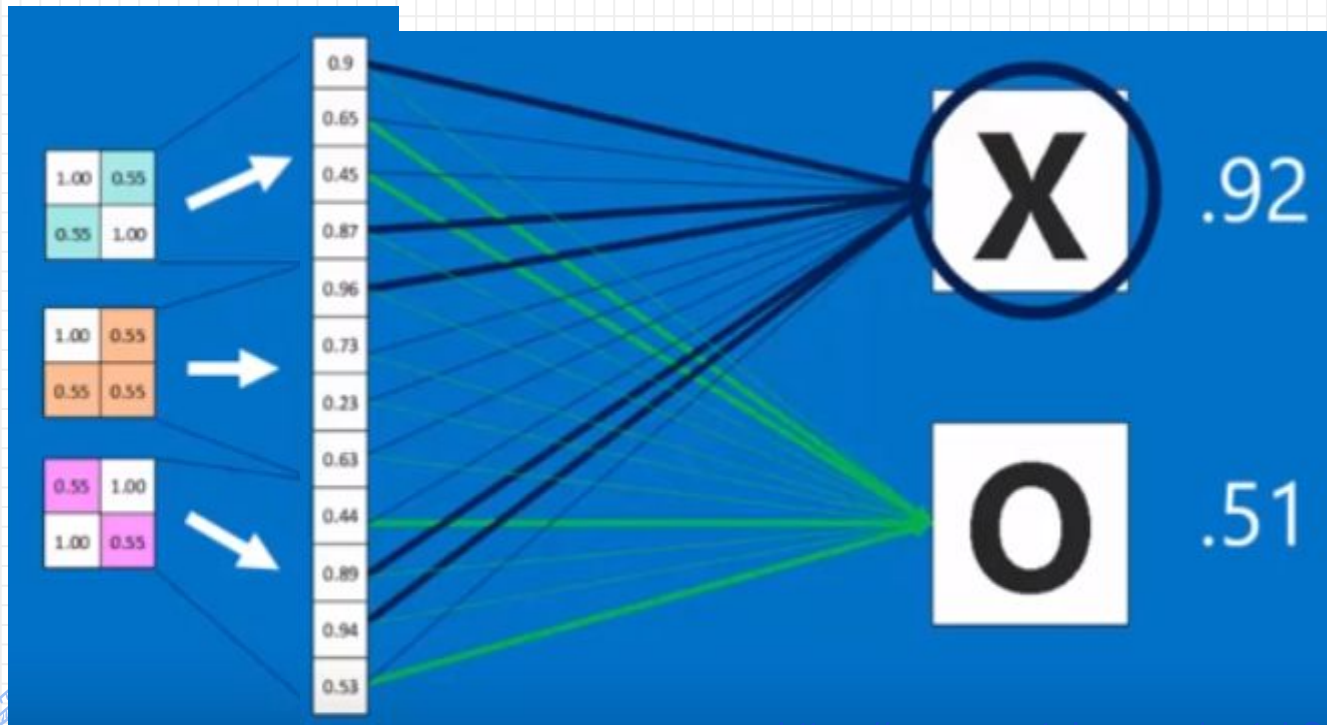


Fully Connected Layer

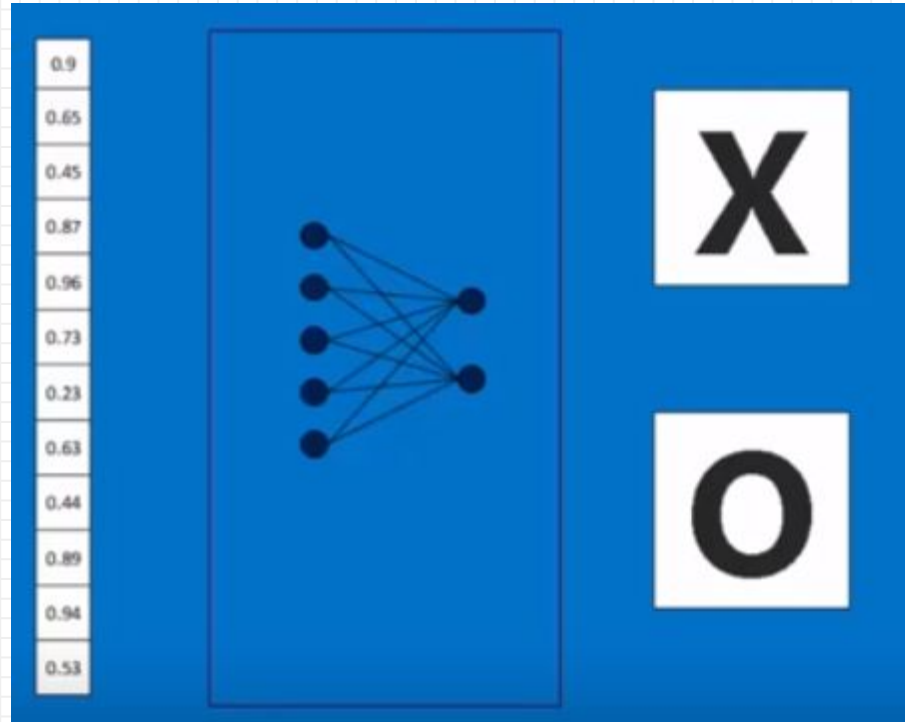


<https://www.youtube.com/watch?v=FmpDlaiMleA>

Fully Connected Layer

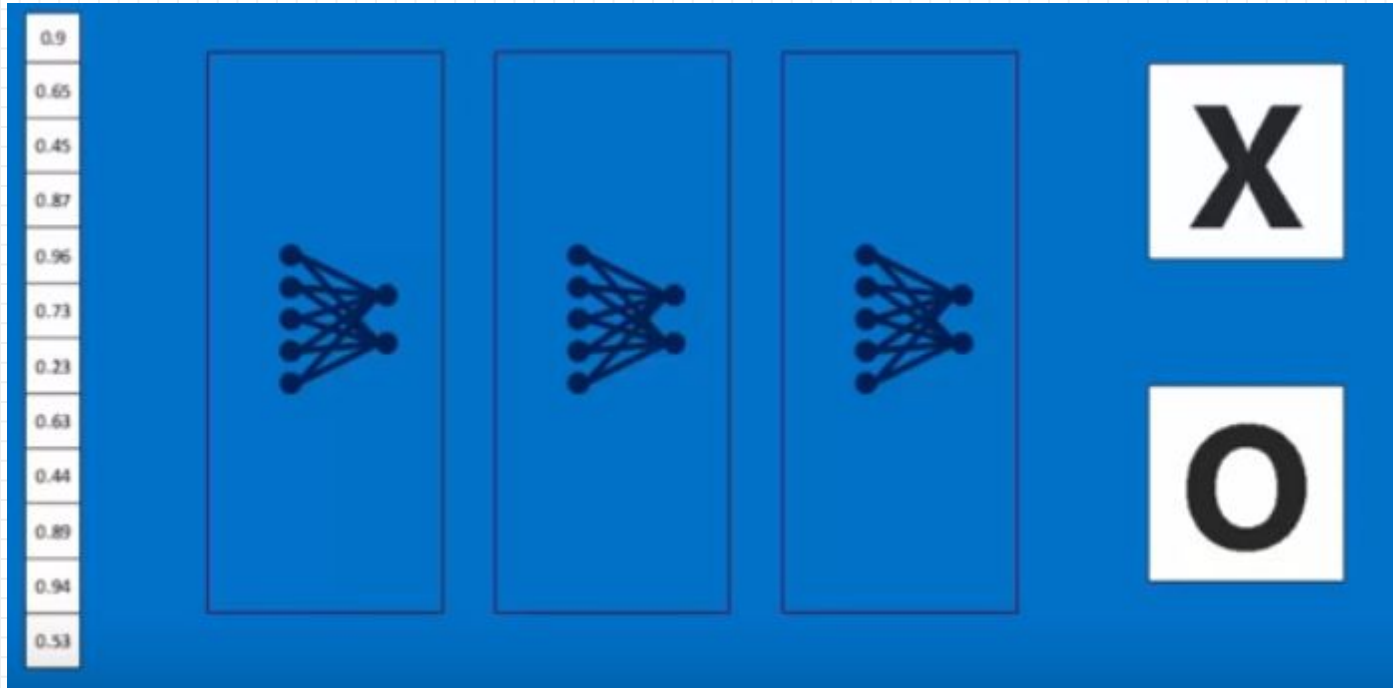


Fully Connected Layer



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

Fully Connected Layer



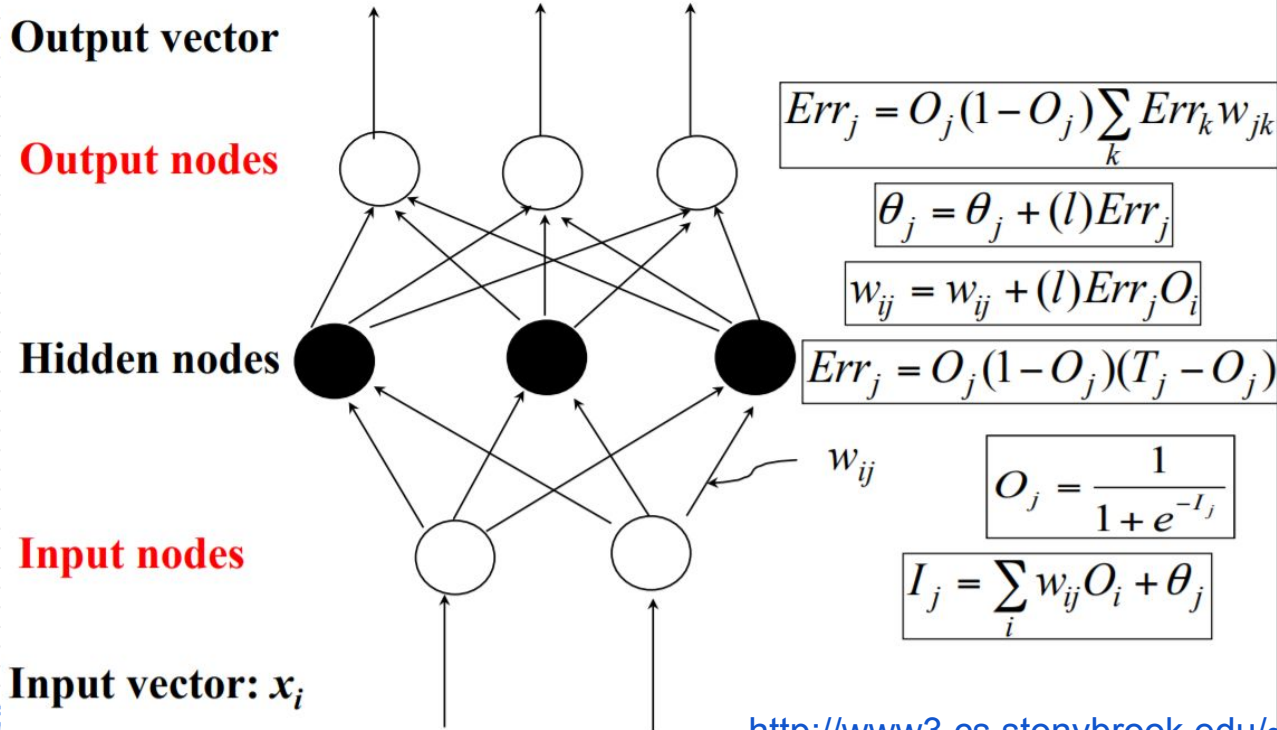
<https://www.youtube.com/watch?v=FmpDlaiMleA>

Q: Where do all features come from?

A: Backpropagation.



Backpropagation



Application

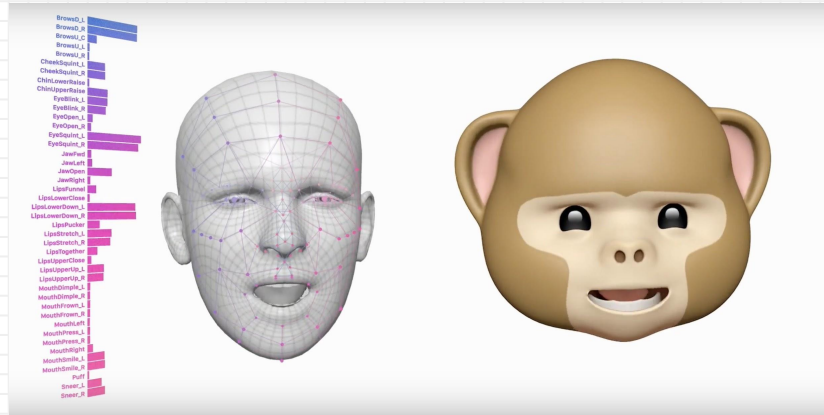
- B612 APP: express yourself daily through fun stickers and amazing AR filters.



Application

- Animoji

New feature in iOS 11. It is an animated versions of emoji.
It uses facial recognition to create 3D versions based on facial expression.



<https://blog.emojipedia.org/apples-new-animoji/>

Application

- Taking Attendance



<http://wvtf.org/post/peeling-onion>

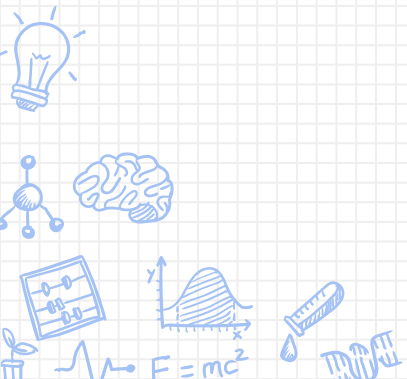
Application

- Finding Rover



<http://stanfordflipside.com/2017/05/going-against-the-grain-this-dog-is-for-sure-going-to-hell/>

Some Concerns About Facial Recognition



Concerns

What can your face say about you?

- ✗ Age
- ✗ Gender
- ✗ Mood



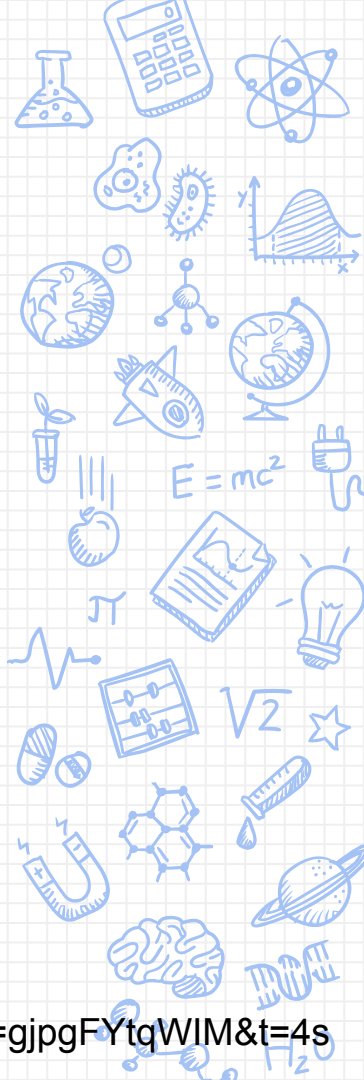
Concerns

Face recognition tech used to 'identify' criminals

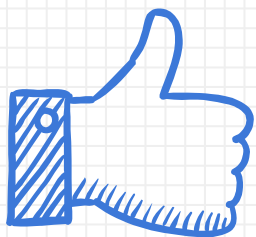
- ✗ Accuracy of up to 90%
- ✗ Different sources



Concerns



<https://www.youtube.com/watch?v=gjpgFYtgWIM&t=4s>



THANKS!

Any questions?