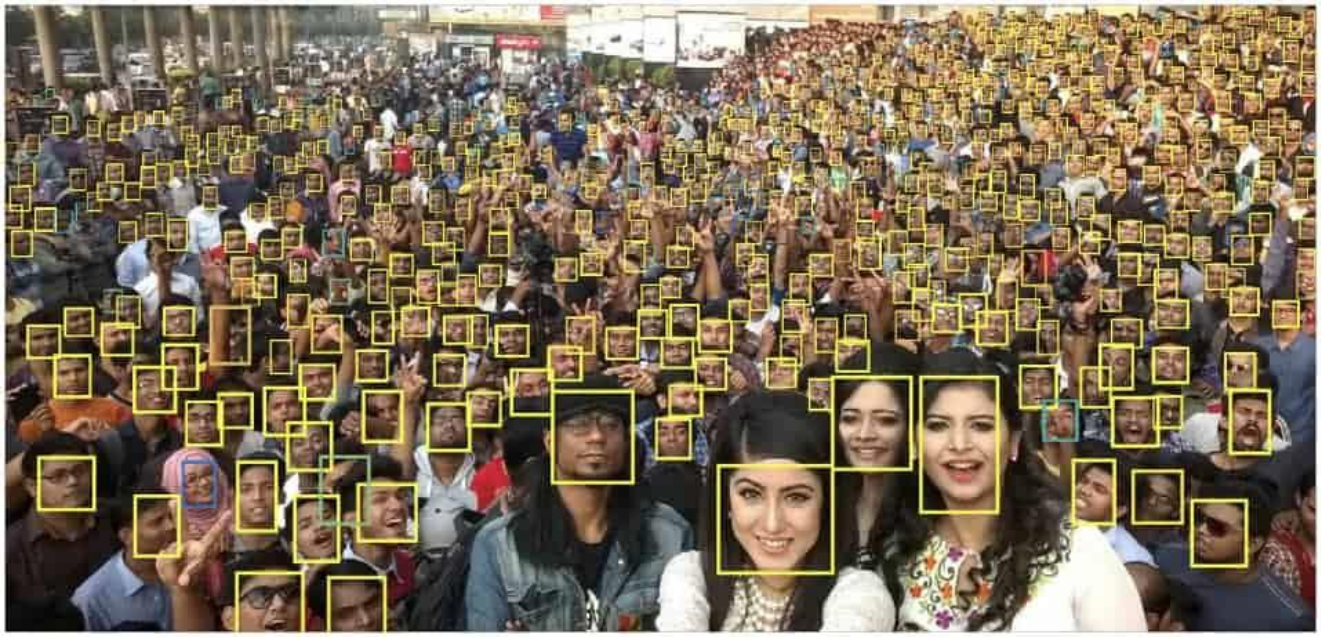
Robotics
JACOBS UNIVERSITY

Introduction to Convolutional Neural Networks



Prof. Francesco Maurelli

Arturo Gomez Chavez

Jacobs University Bremen gGmbH
Robotics, Computer Science and
Electrical Engineering
a.gomezchavez@jacobs-university.de

Before you start…

- These slides are meant to highlight main points of discussion about the topic, and are a compilation of the next sources. Please read them for full understanding:
  - General and intuitive introduction of CNNs:
    - https://medium.freecodecamp.org/an-intuitive-guide-to-convolutional-neural-networks-260c2de0a050
  - A more math based explanation:
    - https://towardsdatascience.com/applied-deep-learning-part-4-convolutional-neural-networks-584bc134c1e2
  - Very intuitive video tutorial of CNNs:
    - https://www.youtube.com/watch?v=JiN9p5vWHDY
  - A very condensed summary of recent developments in computer vision in the last years:
    - http://www.themtank.org/a-year-in-computer-vision

Motivation

- Need of computer vision algorithms that are not only **ACCURATE** but **FAST** and with **EASY DEPLOYMENT**.

- Algorithms that can be processed on mobile systems
  - **(EDGE COMPUTING)**

Motivation

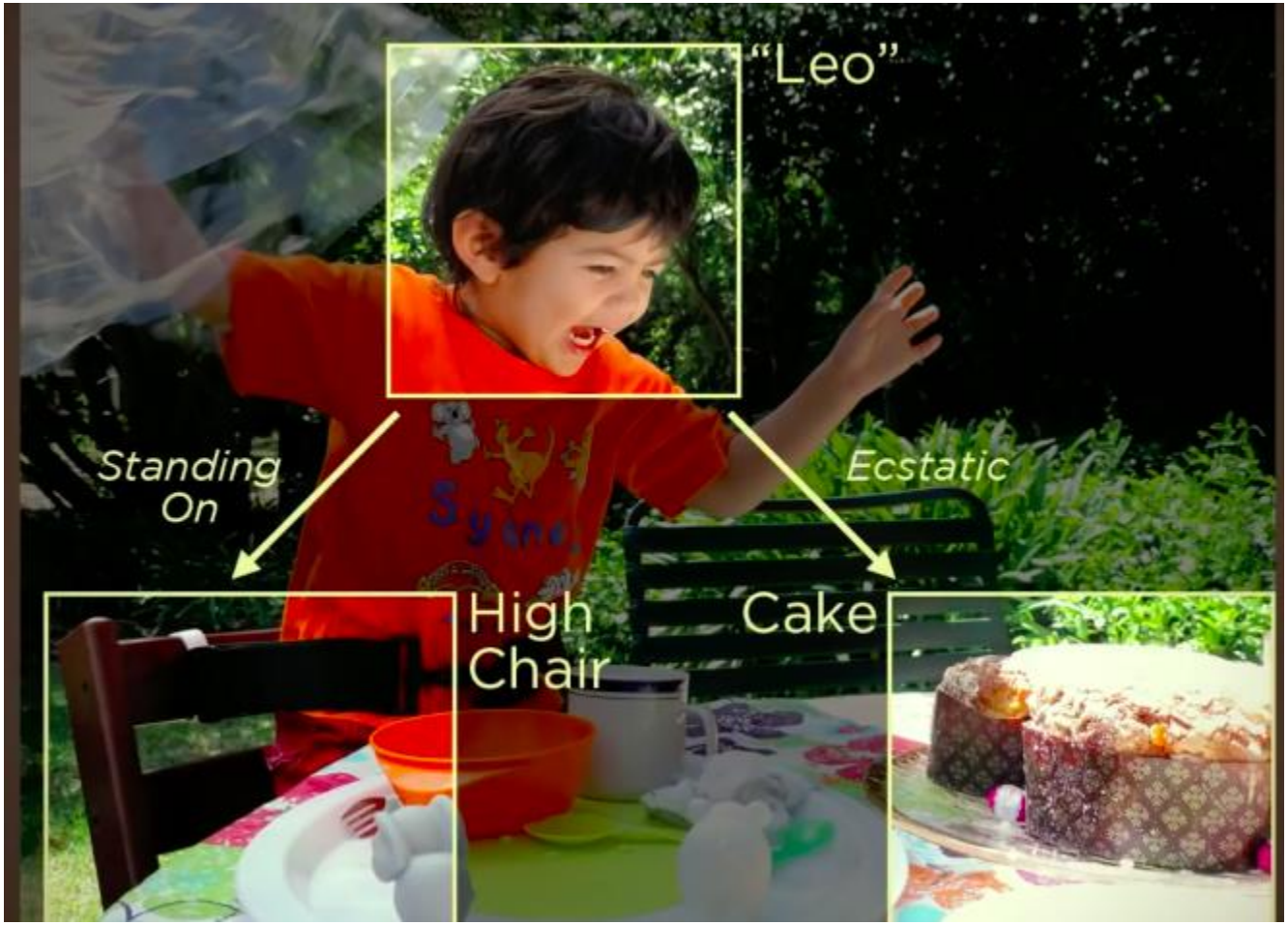
- Not so far from fiction anymore.
- DeepGlint, presented in CVPR 2016.
- https://www.youtube.com/watch?v=xhp47v5OBXQ

Introduction – The Brain

- What do you see in the picture ???

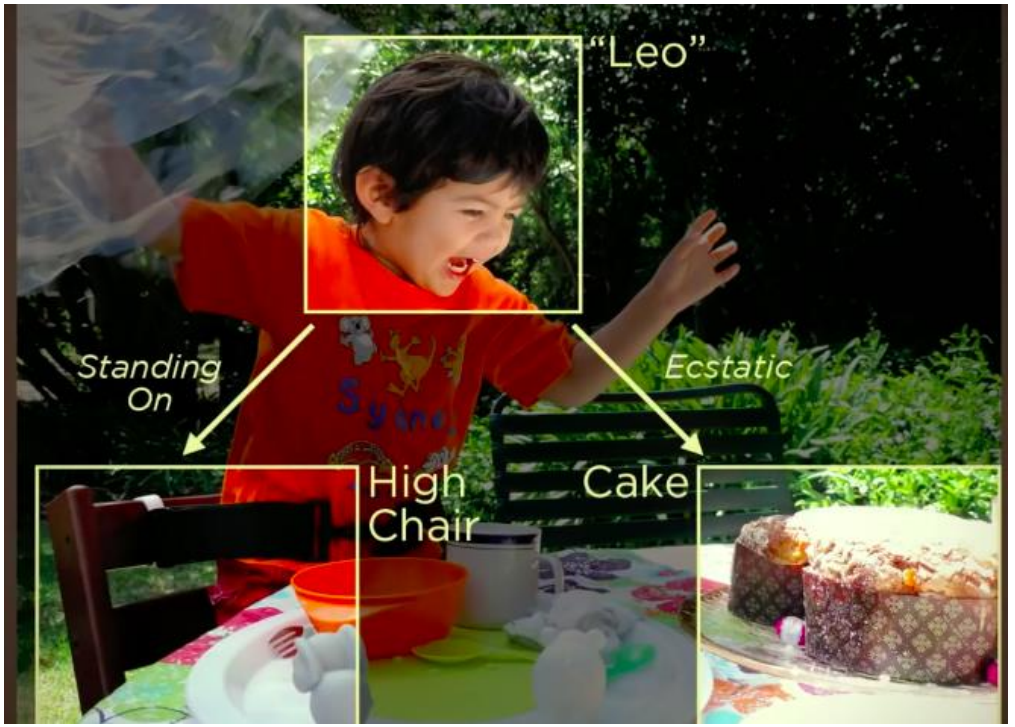
Introduction – The Brain

- Happy kid VS kid attacking a cake.



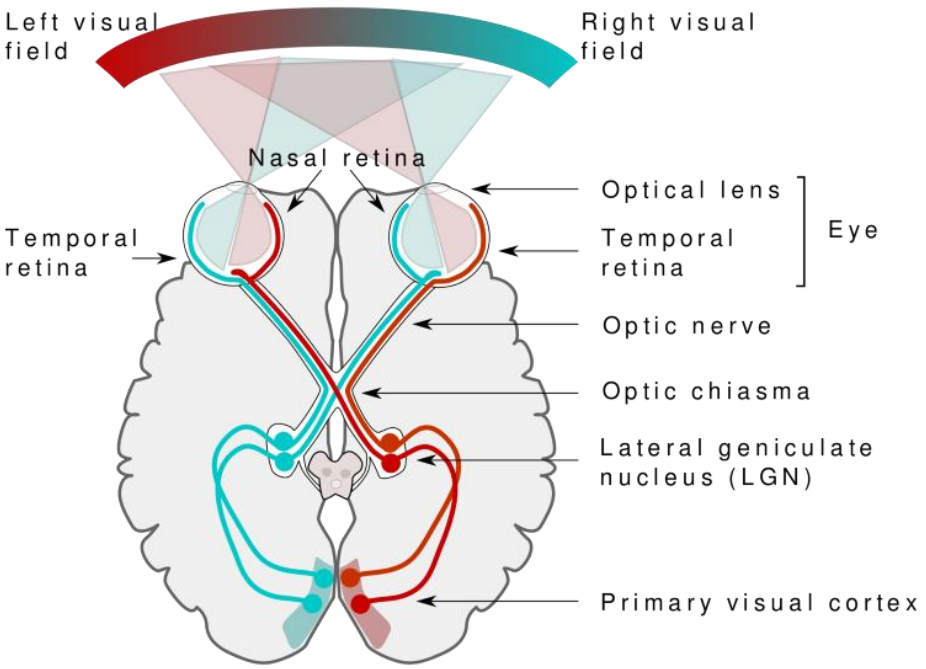
Introduction – The Brain

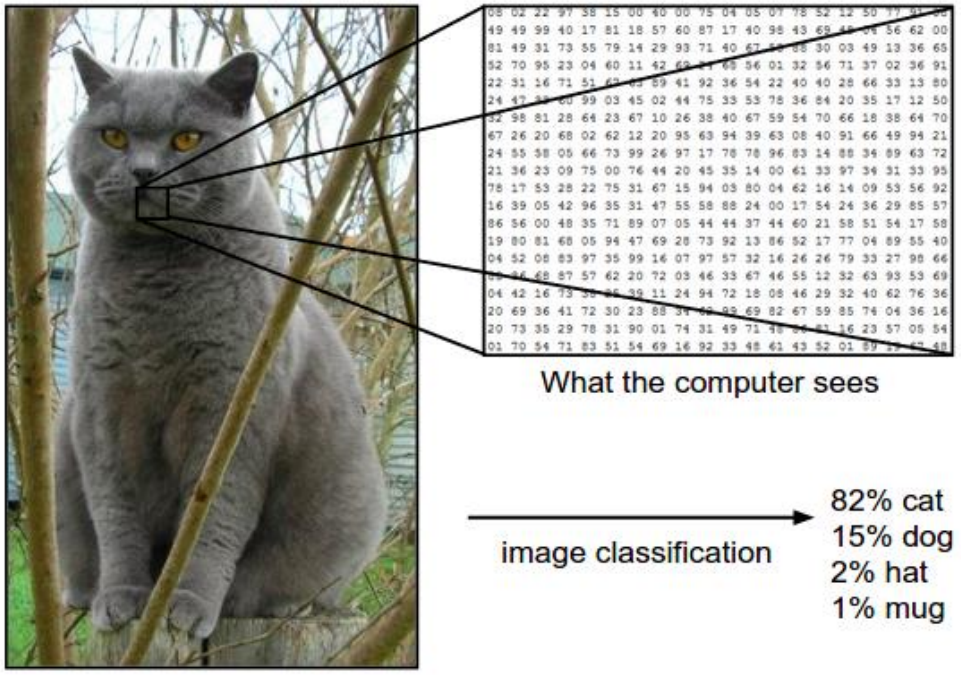
- In a daily basis humans:
  - Label
  - Make predictions
  - Recognize patterns
  - **Create models**
- We are trained since birth



Introduction – The Brain

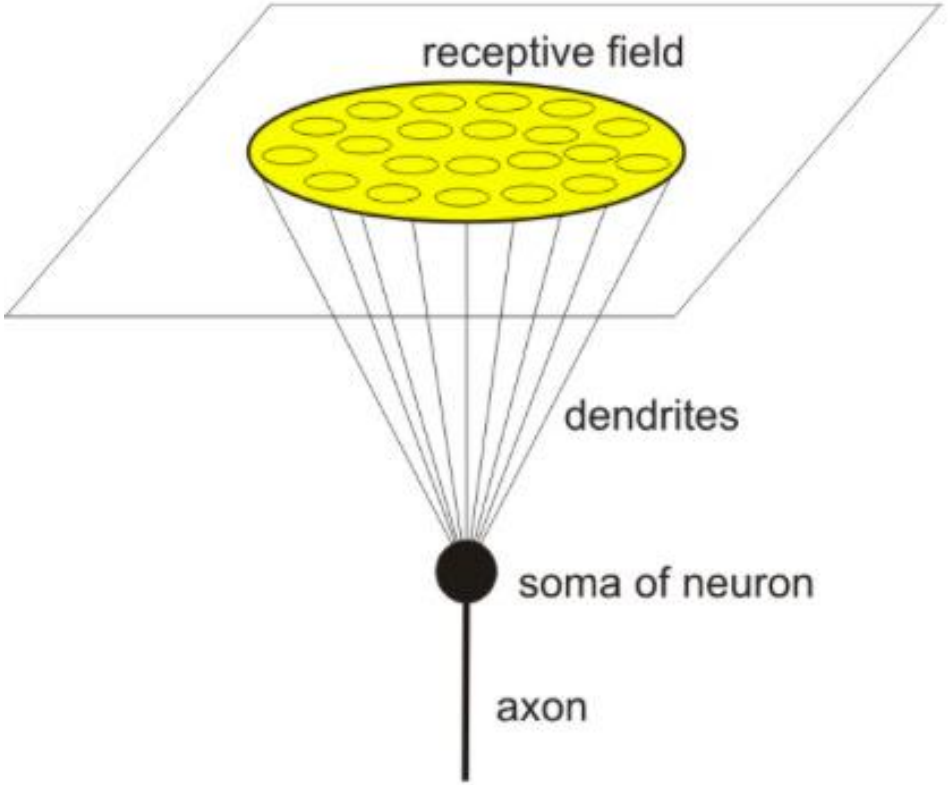
- 500 million years of evolution → visual pathway
- Millions of pictures → Adapt what the computer sees to mimic the brain





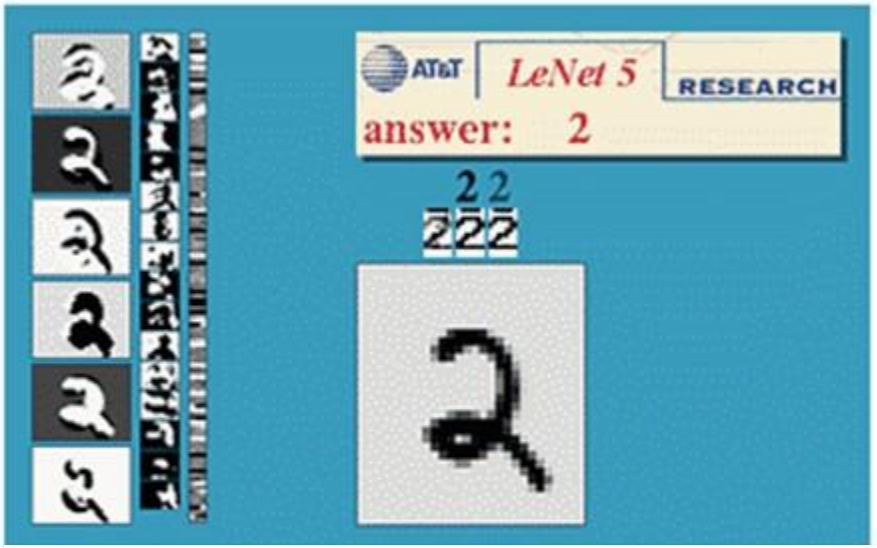
Introduction – The Brain

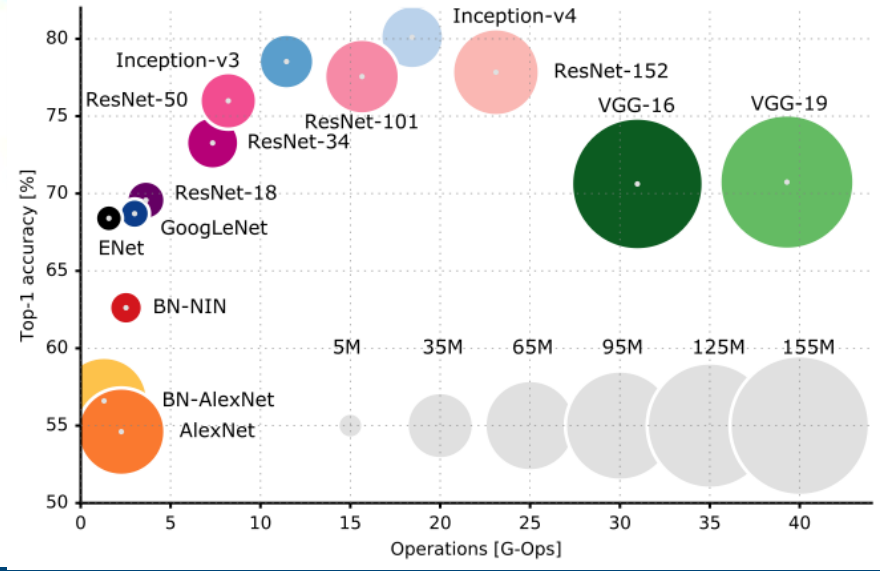
- A receptive field is a single sensory neuron on the retina.
- Different events trigger each receptive field.
- For example, when they eye perceives a line or and edge, or even a familiar shape.
- Others trigger with movement
- 1960s D.H Hubel and T.N Wiesel research on mammals.
- **STACK HIERARCHICALY MILLIONS OF THIS**



History

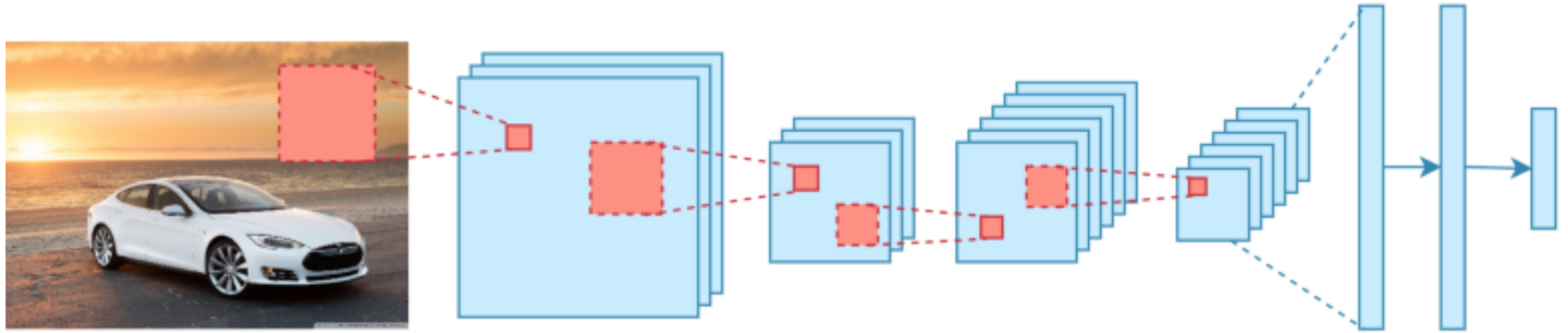
- 1980 – Fukushima – Neocognitron.
- 1998 – Yoshua Bengio, Yan Le Cun, et.al. – Convolutional Neural Network LeNet-5
- 2012 – AlexNet – Alex Krizhervsky
- And then …





CNN - Architecture

- Input image
- Convolutional layer (Features computation)
- Pooling layer (Dimensionality reduction)
- Fully Connected Layer (Classification)
- Softmax Layer (Classification)



CNN - Convolution

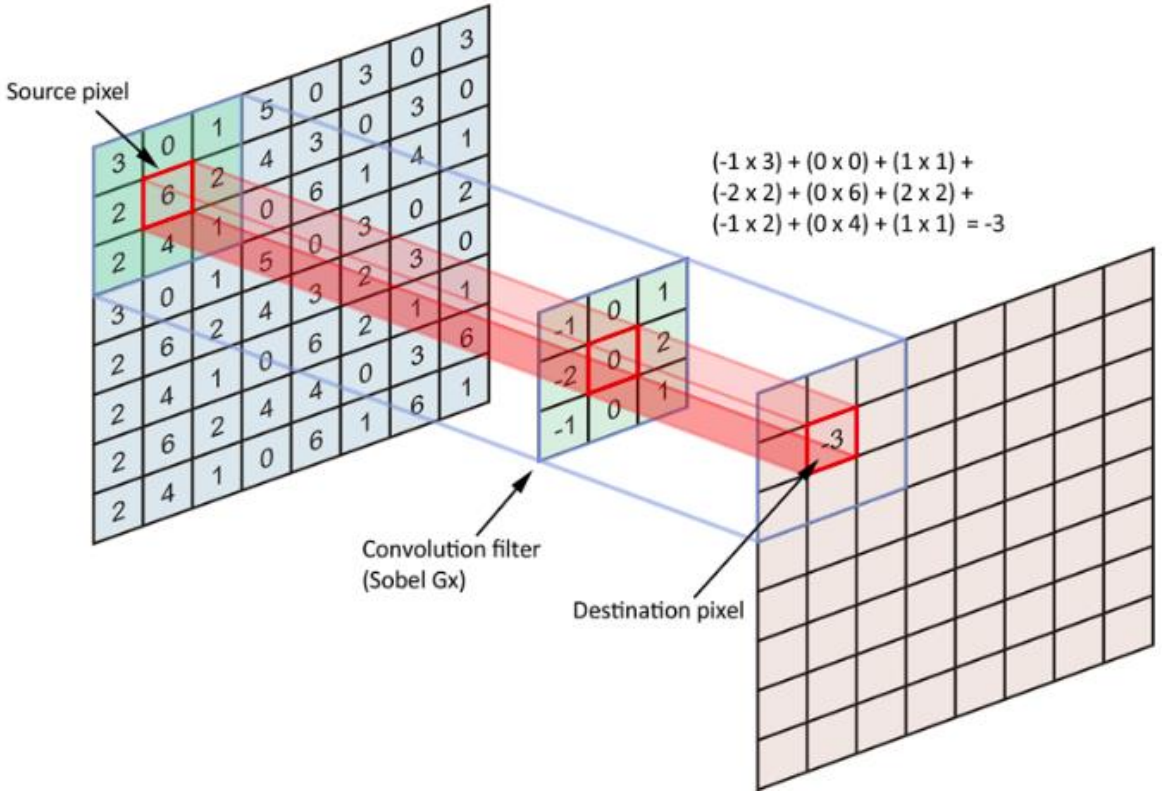
- Use of a filter or kernel to do a feature map.
- Features such as SIFT, SURF, ORB, HOG apply the same principle.

1x1	1x0	1x1	0	0
0x0	1x1	1x0	1	0
0x1	0x0	1x1	1	1
0	0	1	1	0
0	1	1	0	0

4		

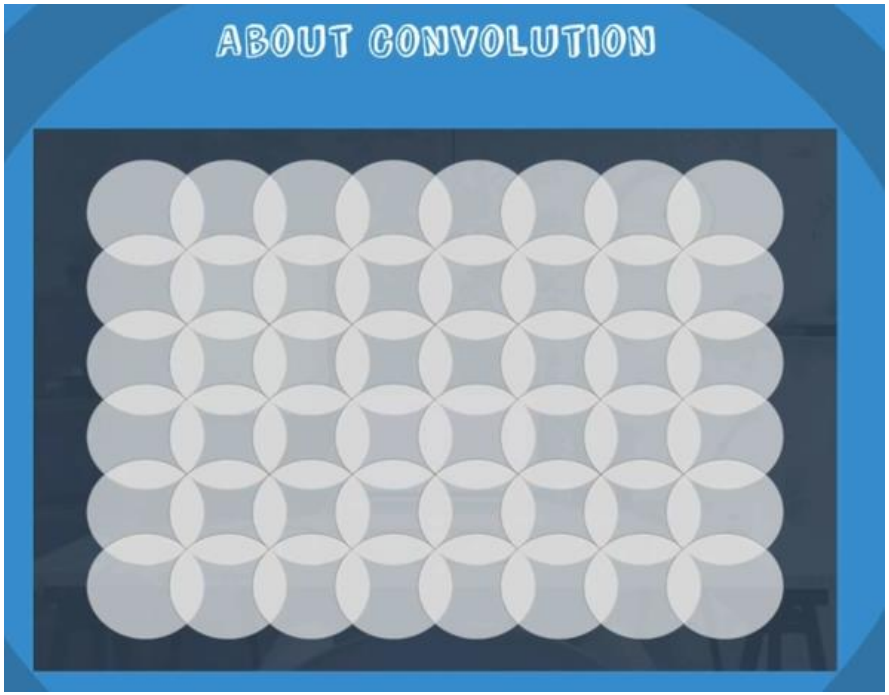
CNN - Convolution

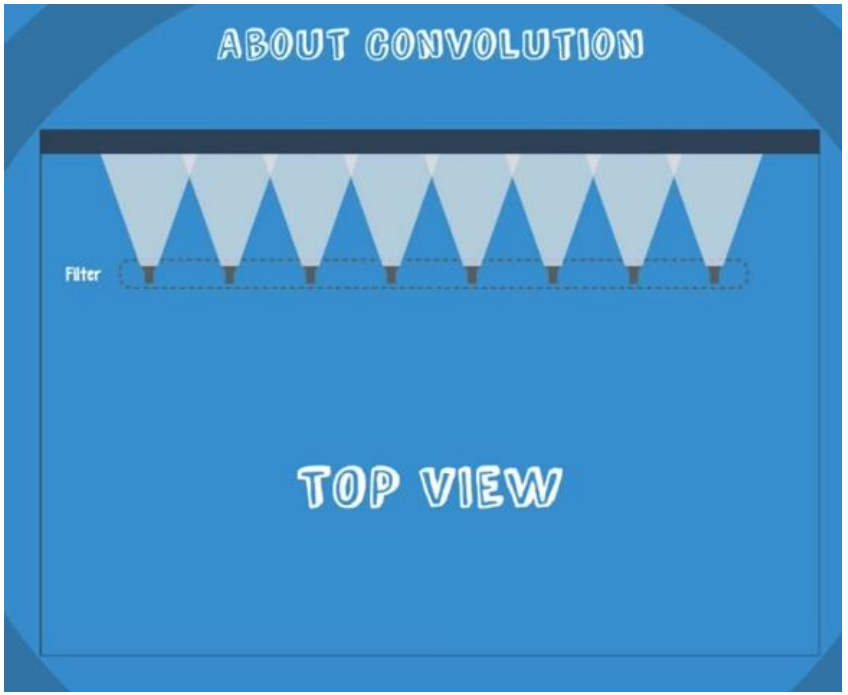
- Use of a filter or kernel to do a feature map.
- Features such as SIFT, SURF, ORB, HOG apply the same principle.



CNN - Intuition

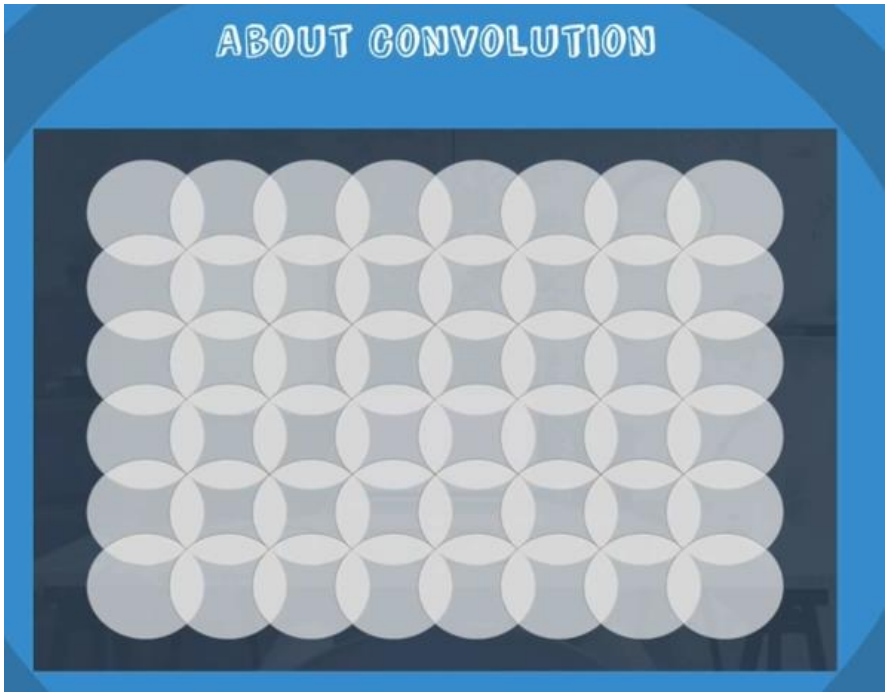
- Flashlight analogy

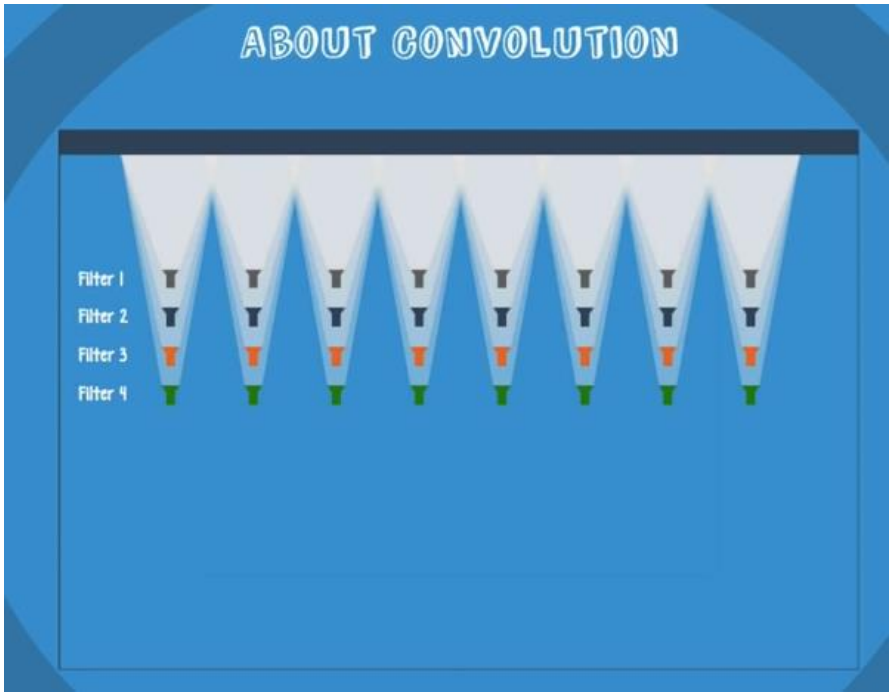




CNN - Intuition

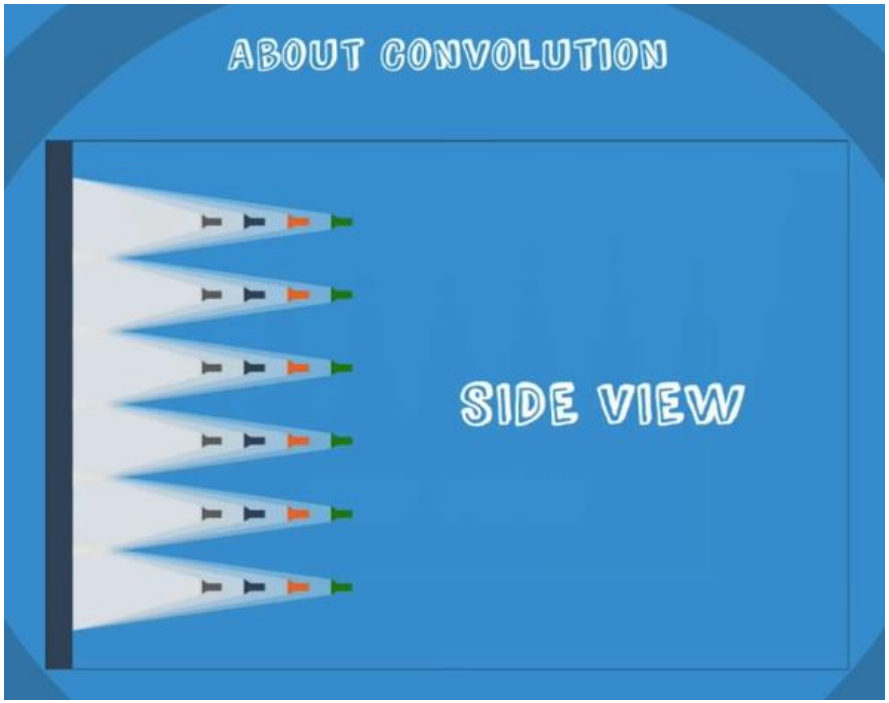
- Flashlight analogy – Each filter provides a feature

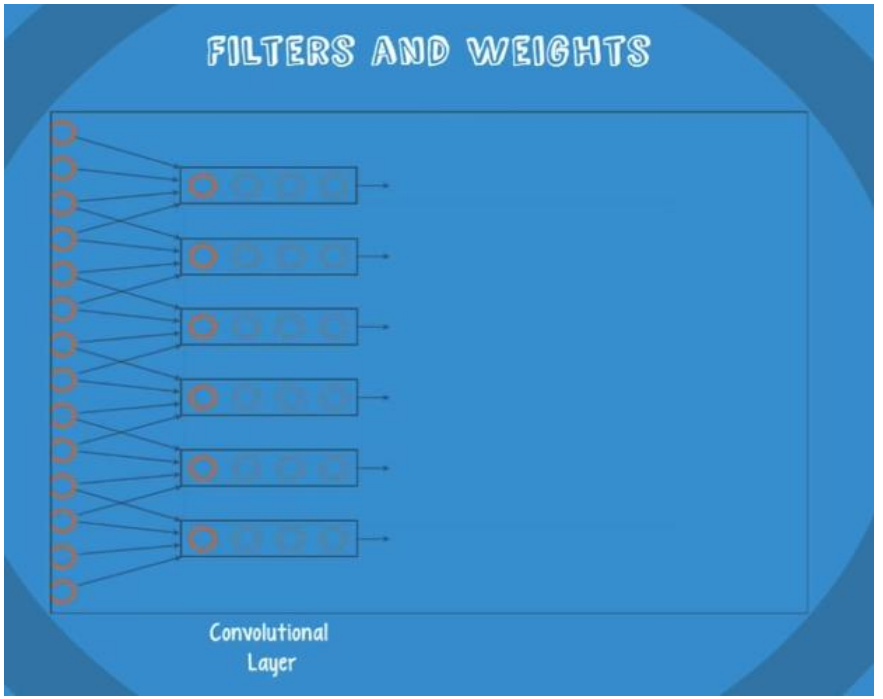




CNN - Intuition

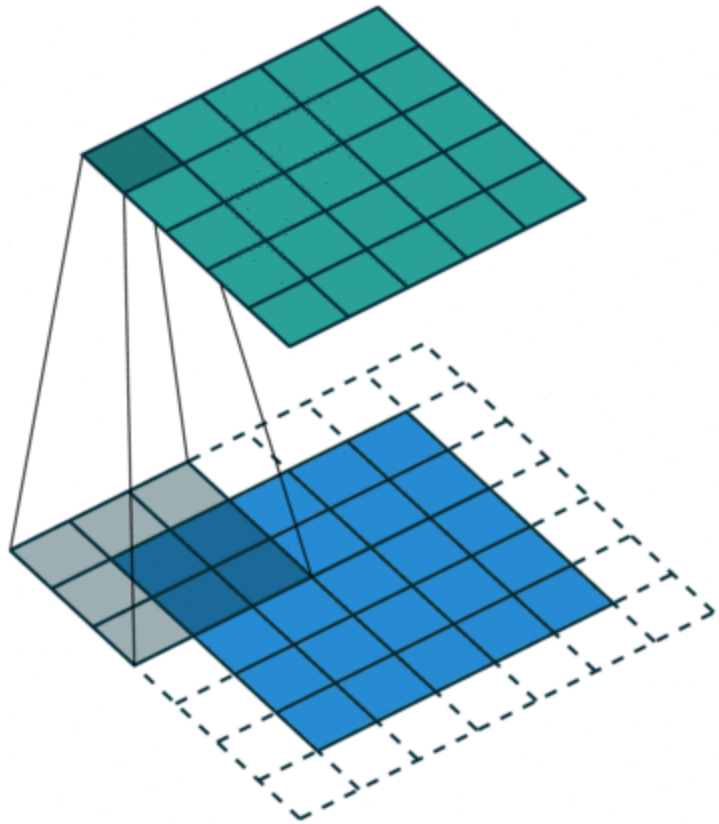
- Flashlight analogy – Each filter provides a feature





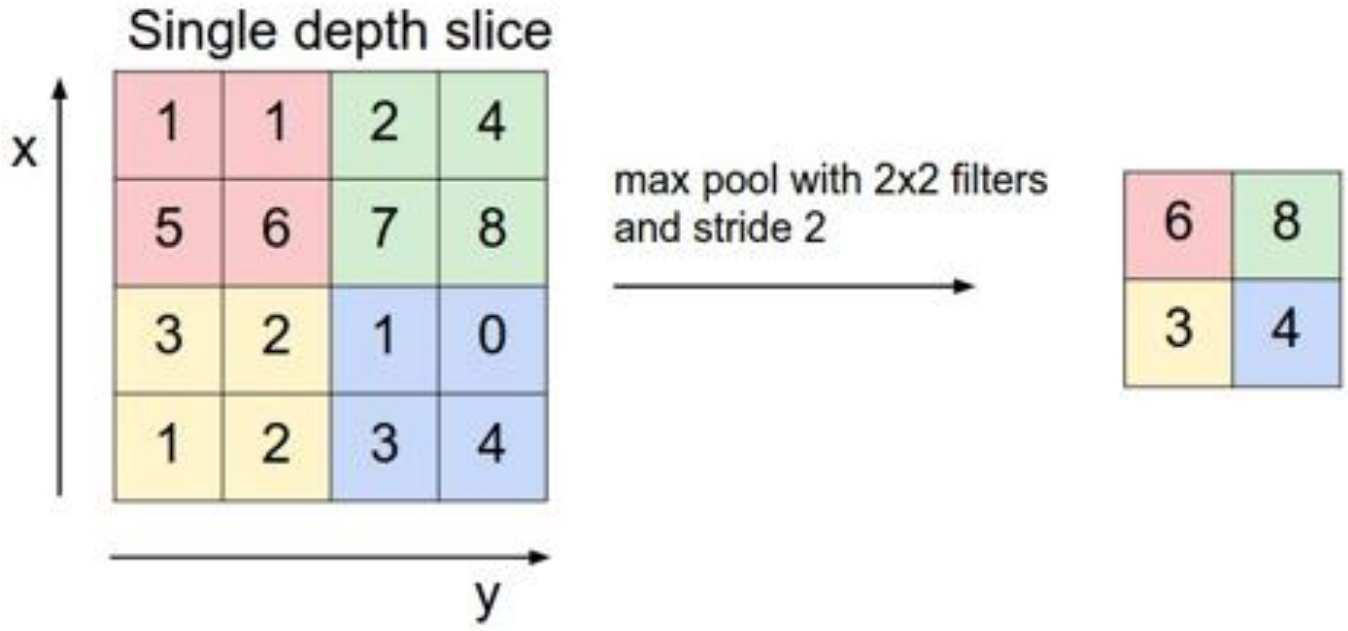
CNN - Parameters

- Kernel Size
- Filter count (No. Features)
- Stride (Displacement)
- Padding (Zeros at the edges)



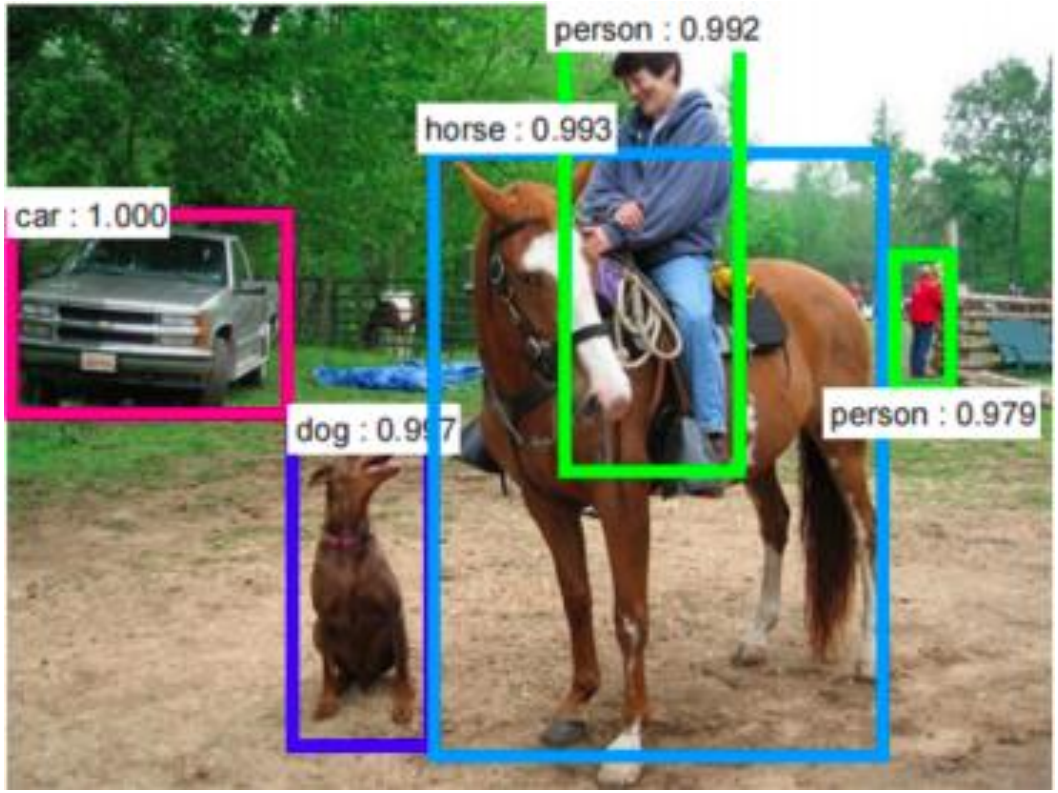
CNN - Pooling

- Dimensionality reduction



Success!!!

- Object detection



CNN - Visualization

- Very good toolbox for visualization of CNN

  https://www.youtube.com/watch?v=AgkfIQ4IGaM