

# Convolutional Neural Networks

# Overview

CNN Architectures and Parts

Why do they work?

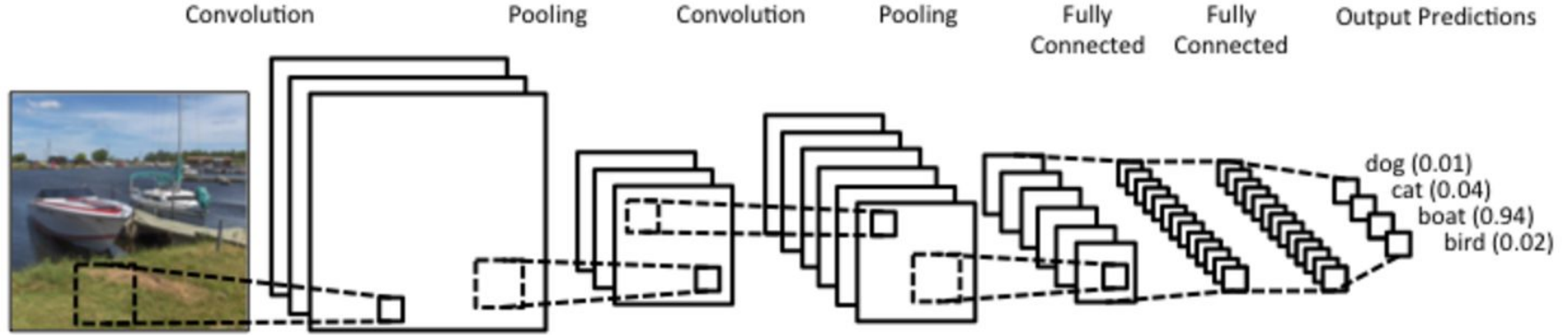
What are the benefits?

Applications

CAFFE and Tensorflow

# CNN Architecture and Parts

# CNN Architecture



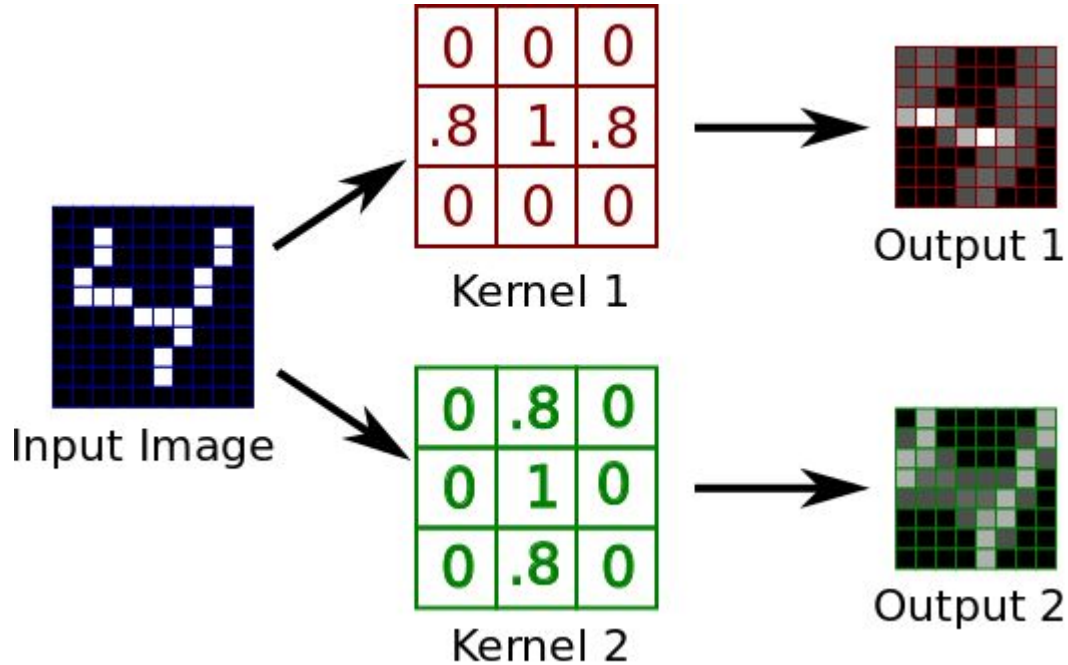
Typical architecture -

- Multiple convolutional & max pooling layers

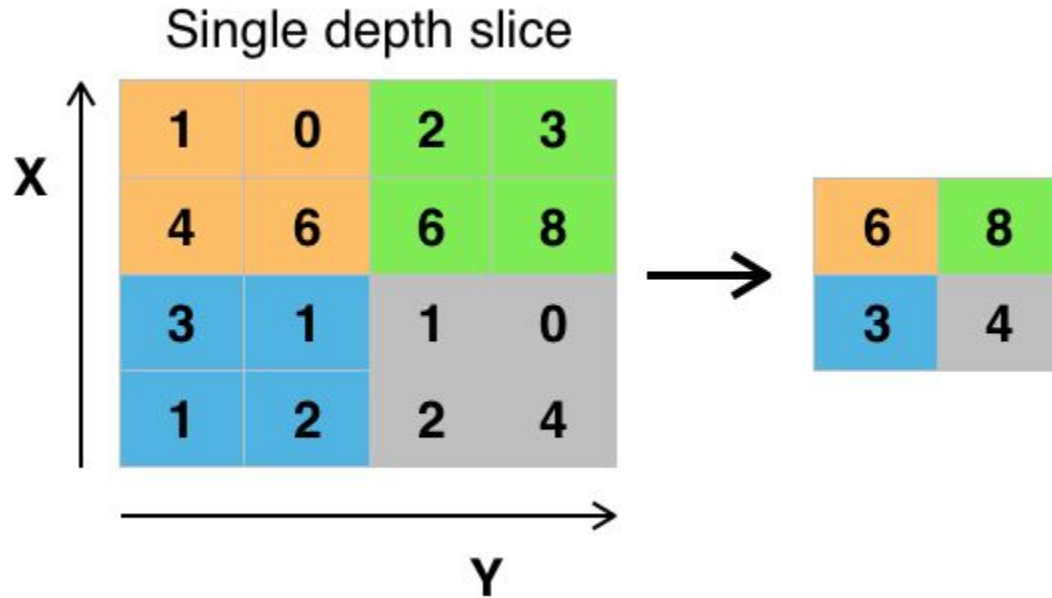
- Fully connected layers

- Softmax prediction

# Parts - Convolutional Layer

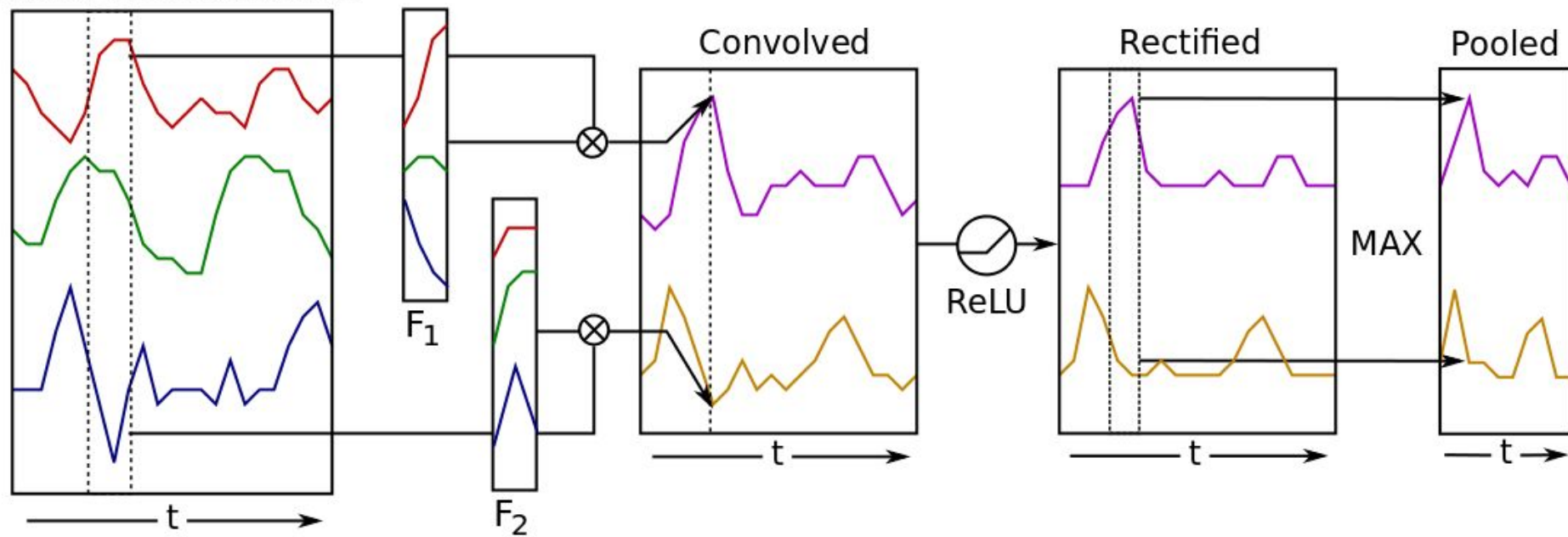


# Parts - Max Pooling

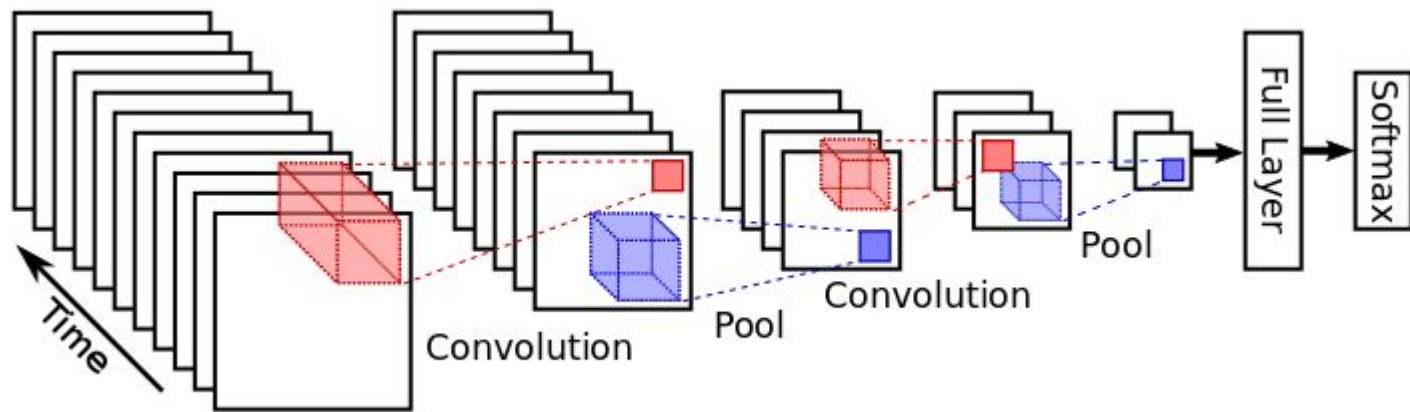


# 1D CNN

Sensor Measurements



# 3D CNN

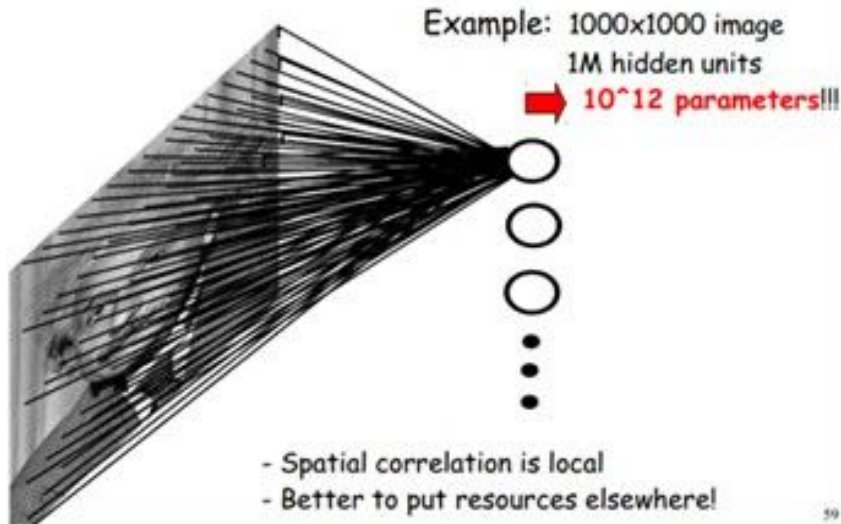




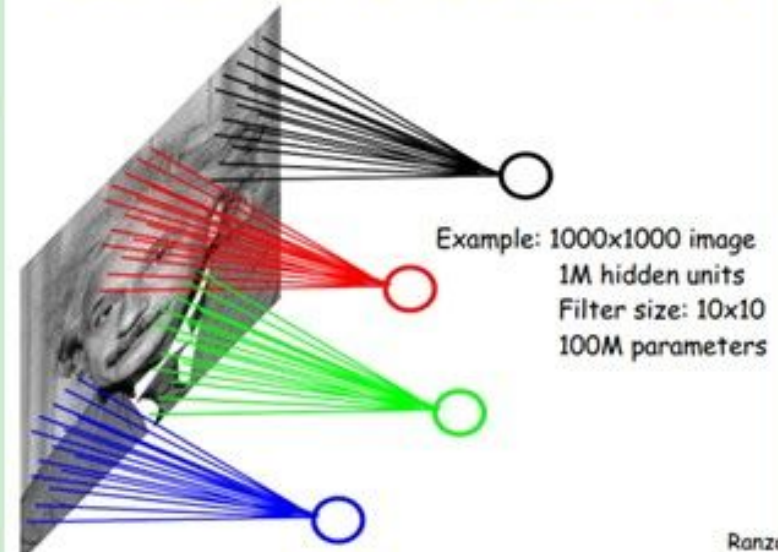
Why do they work?

# CNN vs Fully Connected

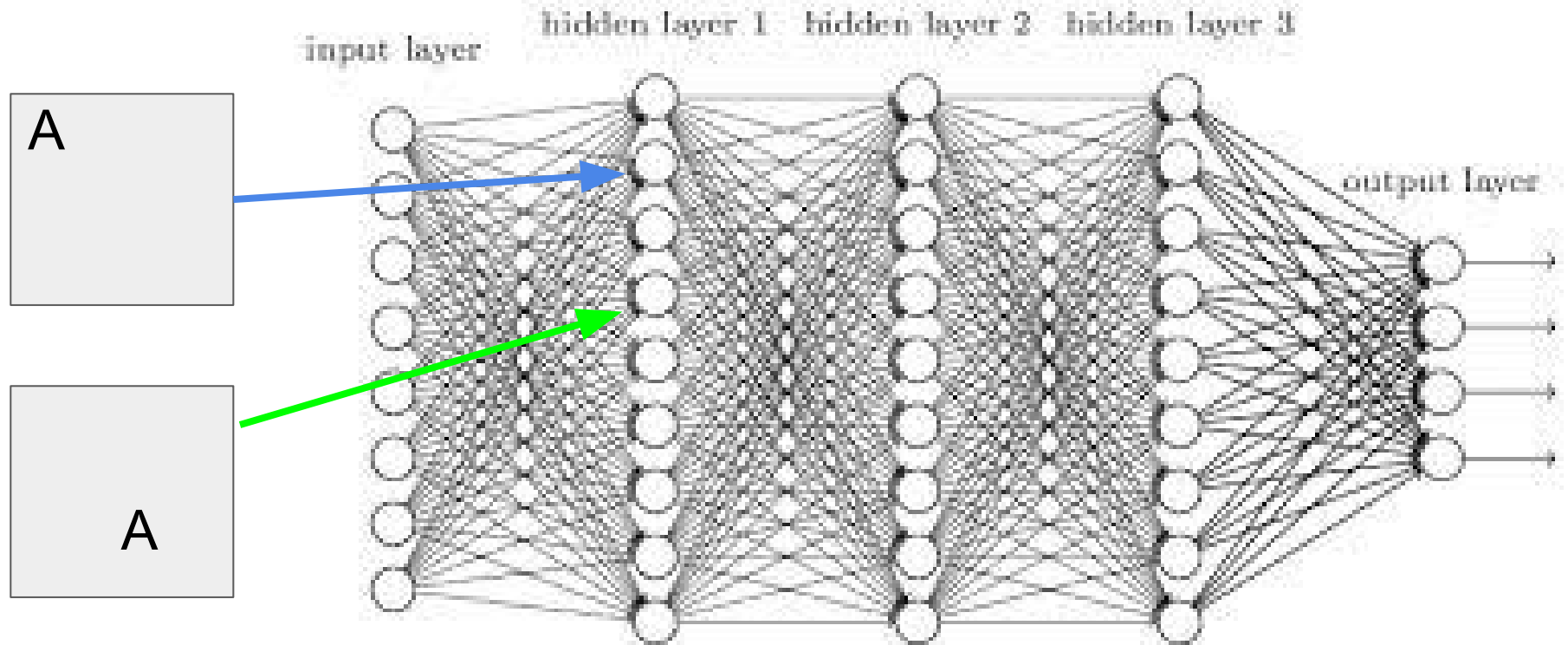
## FULLY CONNECTED NEURAL NET



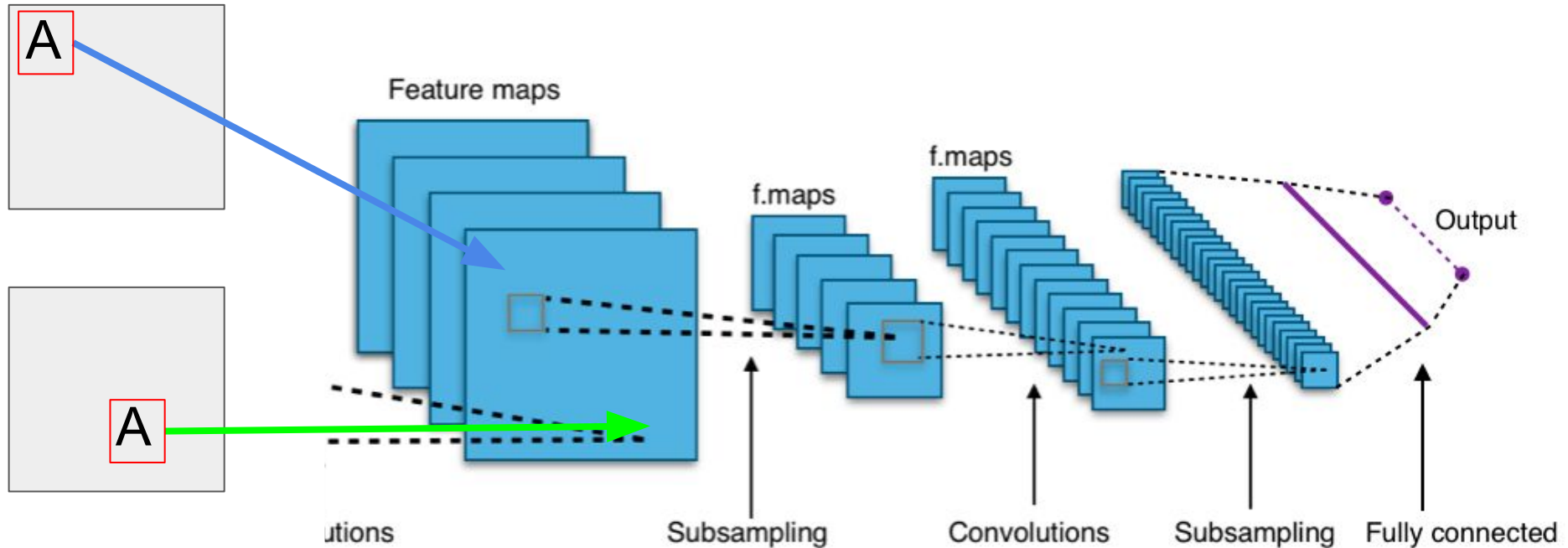
## LOCALLY CONNECTED NEURAL NET



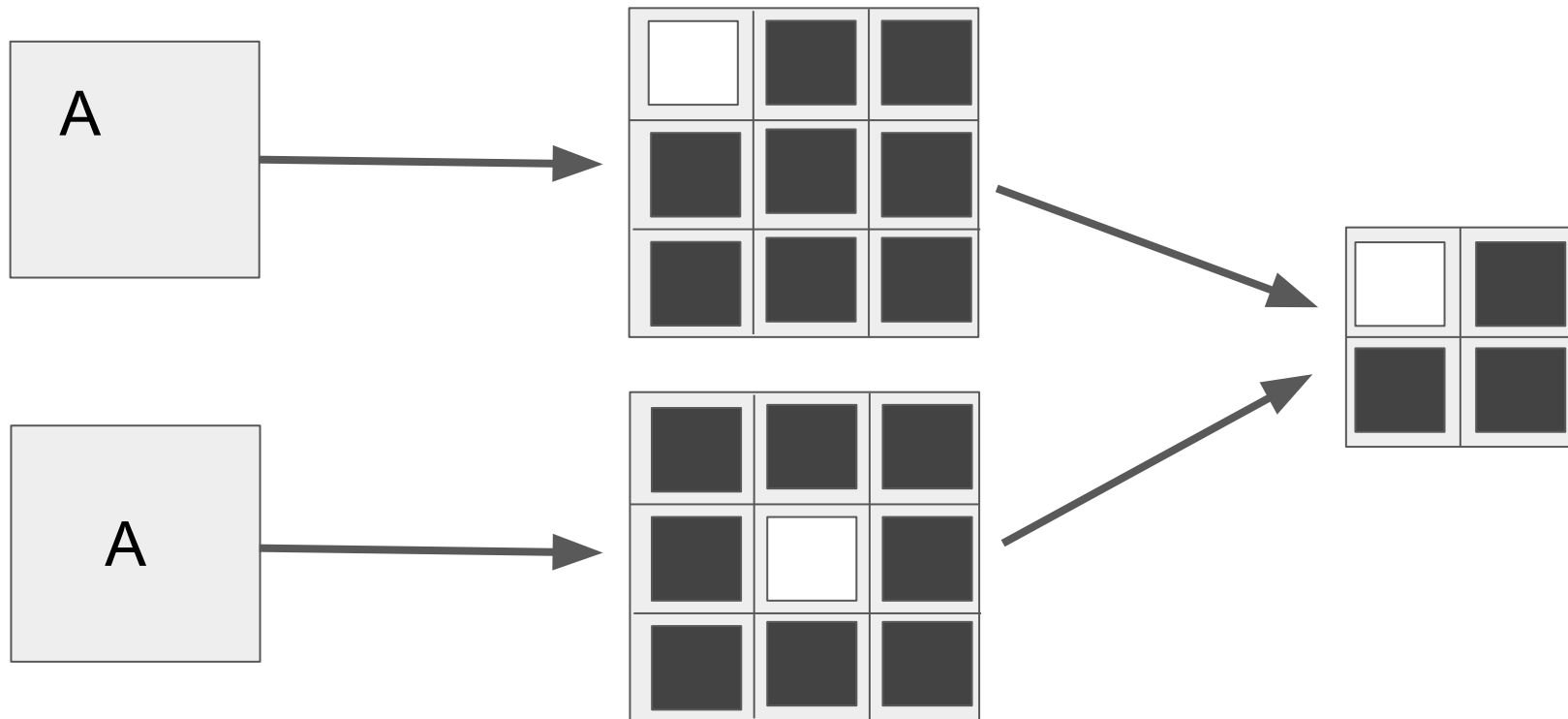
# CNN vs. Fully Connected



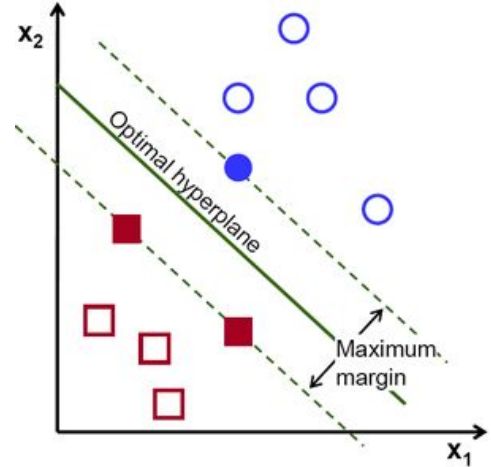
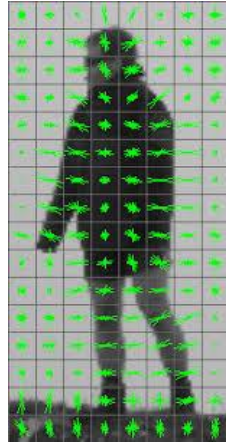
# CNN vs Fully Connected



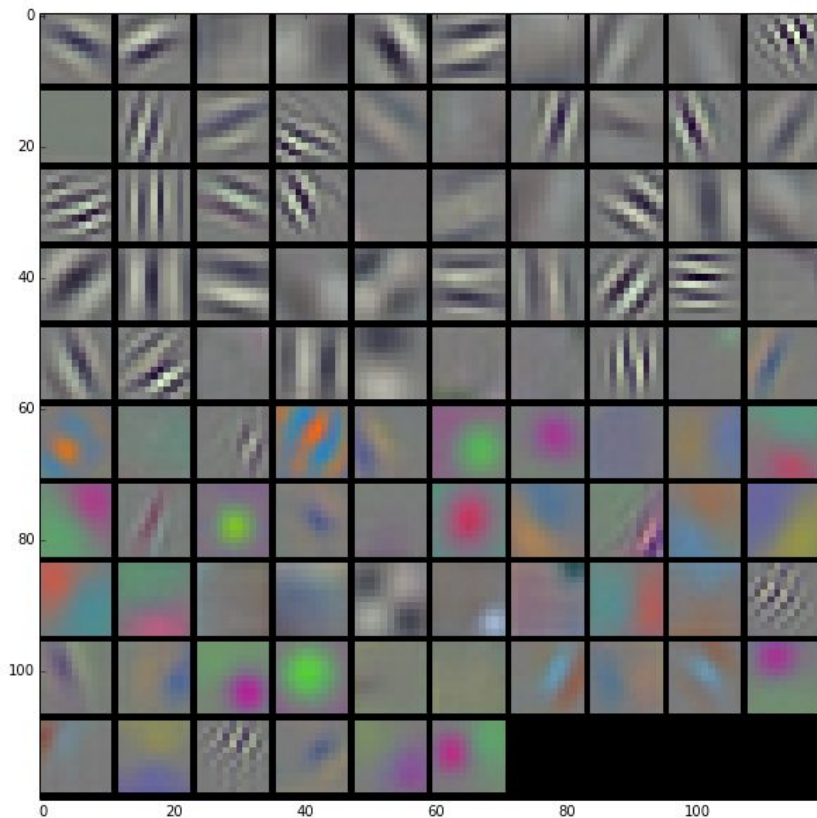
# Pooling



# Hand Engineered Features

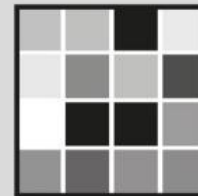


# Automatic Feature Extraction

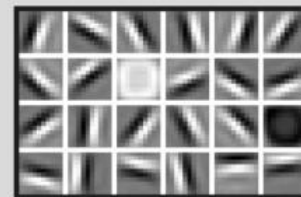


## FACIAL RECOGNITION

Deep-learning neural networks use layers of increasingly complex rules to categorize complicated shapes such as faces.



Layer 1: The computer identifies pixels of light and dark.



Layer 2: The computer learns to identify edges and simple shapes.



Layer 3: The computer learns to identify more complex shapes and objects.

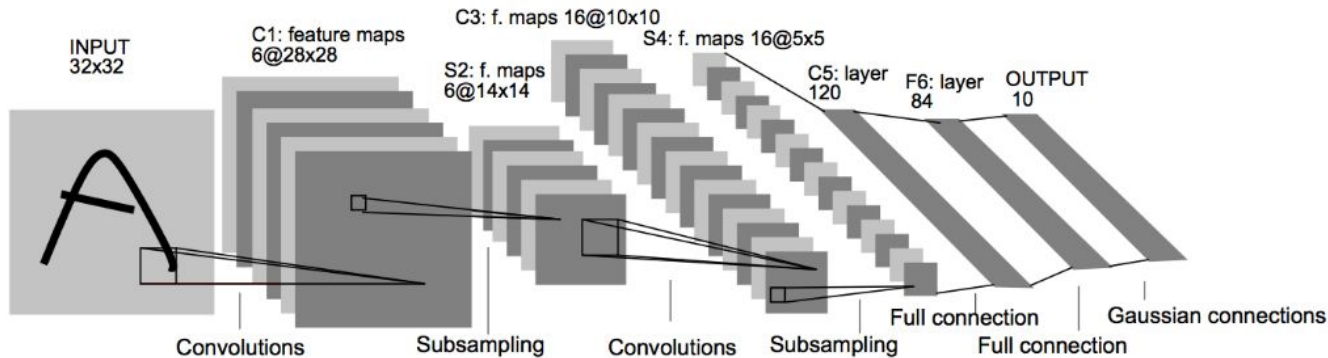


Layer 4: The computer learns which shapes and objects can be used to define a human face.

# Applications



# LeNet - Identify Numbers (MNIST Dataset!)

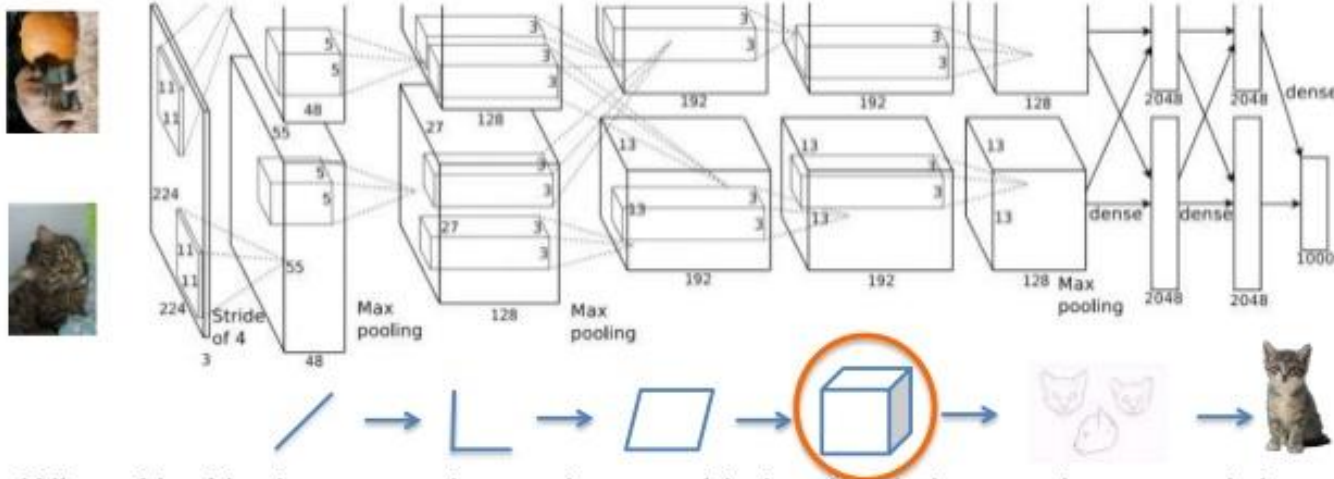


**CNN called LeNet by Yann LeCun (1998)**

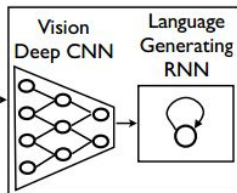
# AlexNet - ImageNet Classification

## AlexNet (Krizhevsky et al. 2012)

*The class with the highest likelihood is the one the DNN selects*













# Automatic Figure Captioning

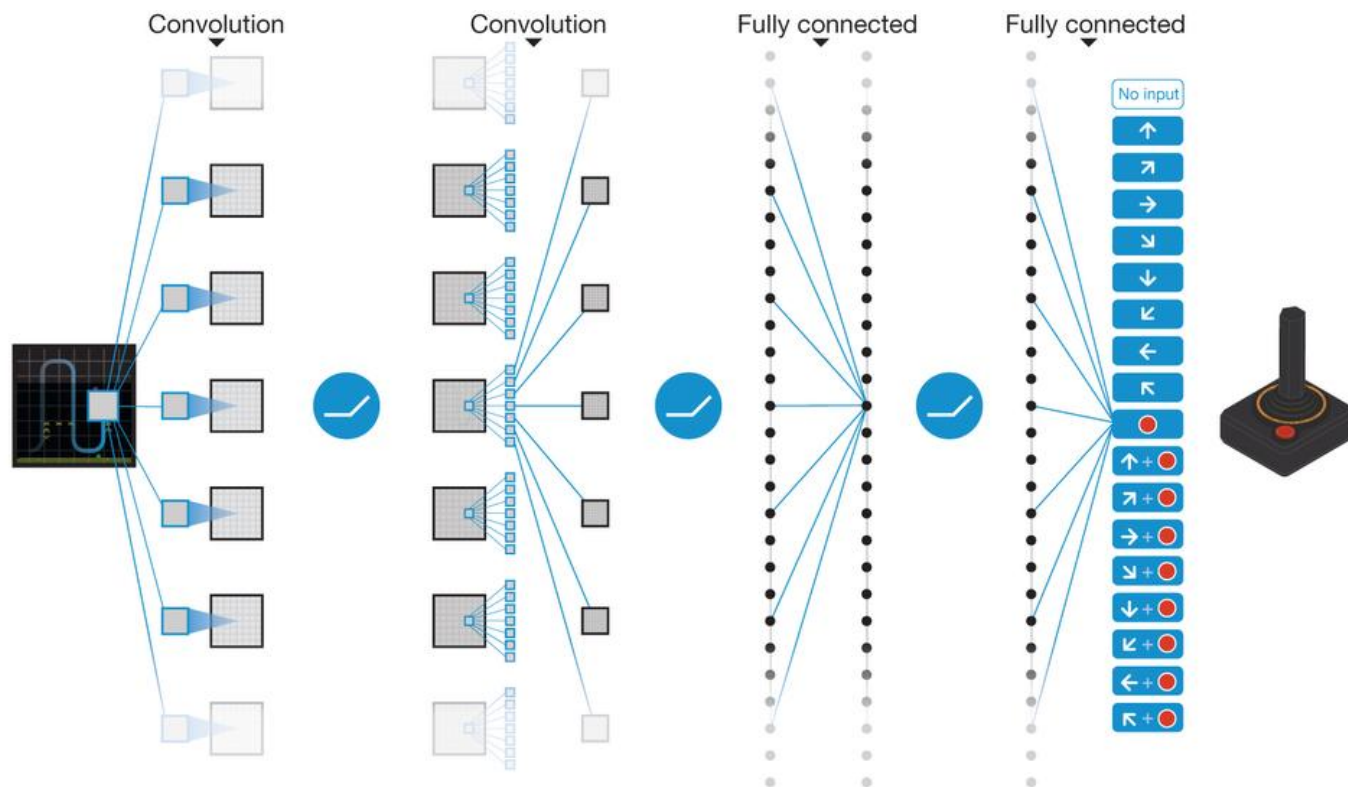


**A group of people shopping at an outdoor market.**

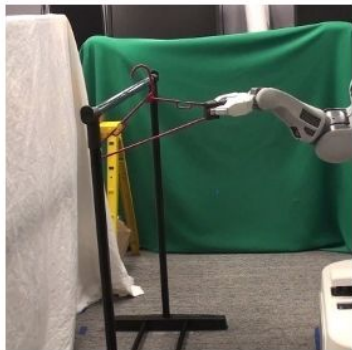
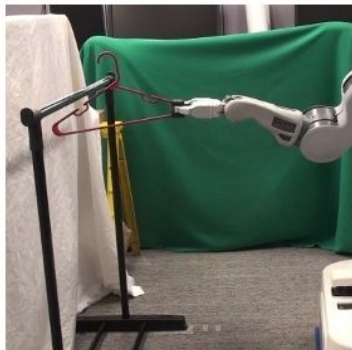
**There are many vegetables at the fruit stand.**

Describes without errors	Describes with minor errors	Somewhat related to the image	Unrelated to the image
 <p>A person riding a motorcycle on a dirt road.</p>	 <p>Two dogs play in the grass.</p>	 <p>A skateboarder does a trick on a ramp.</p>	 <p>A dog is jumping to catch a frisbee.</p>
 <p>A group of young people playing a game of frisbee.</p>	 <p>Two hockey players are fighting over the puck.</p>	 <p>A little girl in a pink hat is blowing bubbles.</p>	 <p>A refrigerator filled with lots of food and drinks.</p>
 <p>A herd of elephants walking across a dry grass field.</p>	 <p>A close up of a cat laying on a couch.</p>	 <p>A red motorcycle parked on the side of the road.</p>	 <p>A yellow school bus parked in a parking lot.</p>

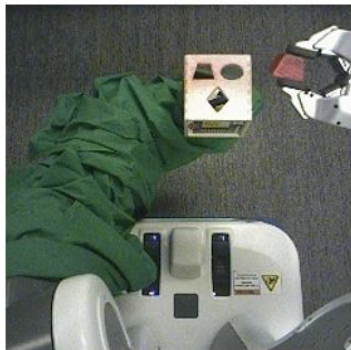
# Playing Atari



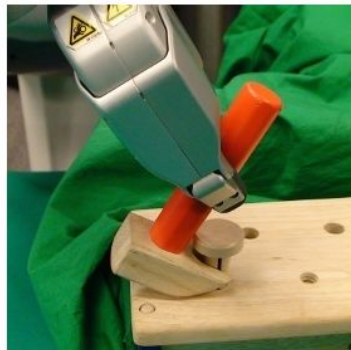
# Control Robots



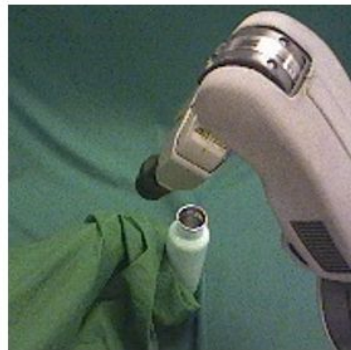
(a) hanger



(b) cube



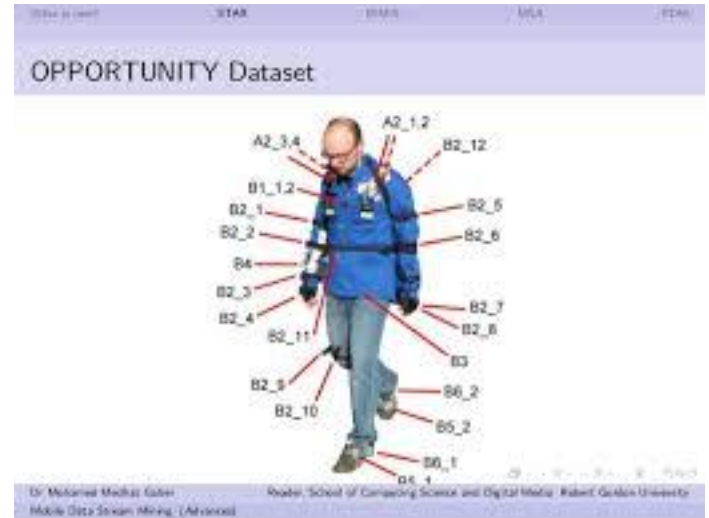
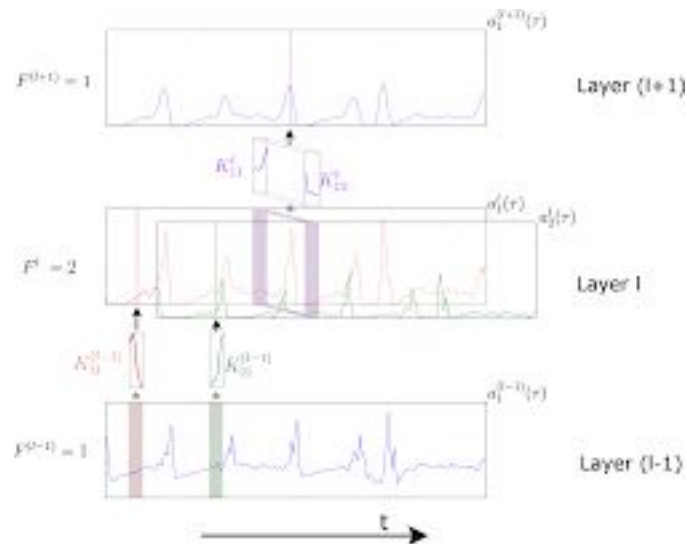
(c) hammer



(d) bottle

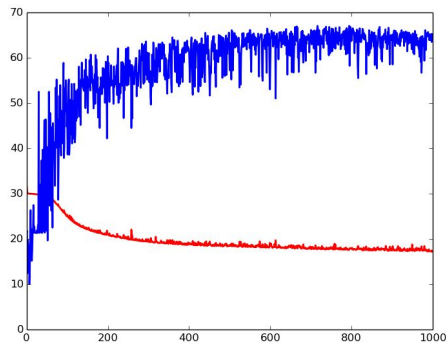


# Human Activity Recognition

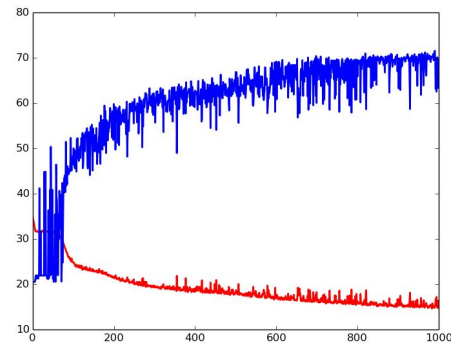


# Caffe and TensorFlow

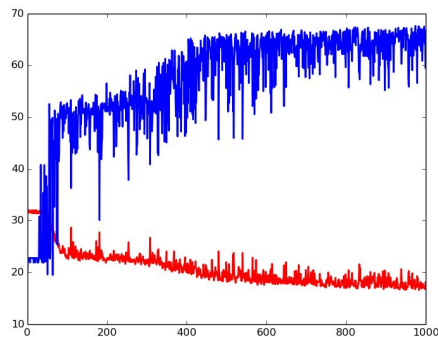
# HAR Demo



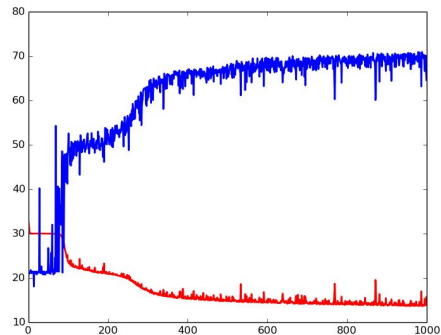
FFNN  
1 layer  
6,807 params



CNN  
1 layer  
4,387 params



FFNN  
2 layer  
15,907 params



CNN  
2 layer  
2,547 params