Object Classification using Convolutional Neural Network(CNN) for Advanced Driver Assistance Systems (ADAS).

Sai Sravan Manne
EDM13B018

Guided by:
Dr. Binsu J Kailath

Discussion Points:

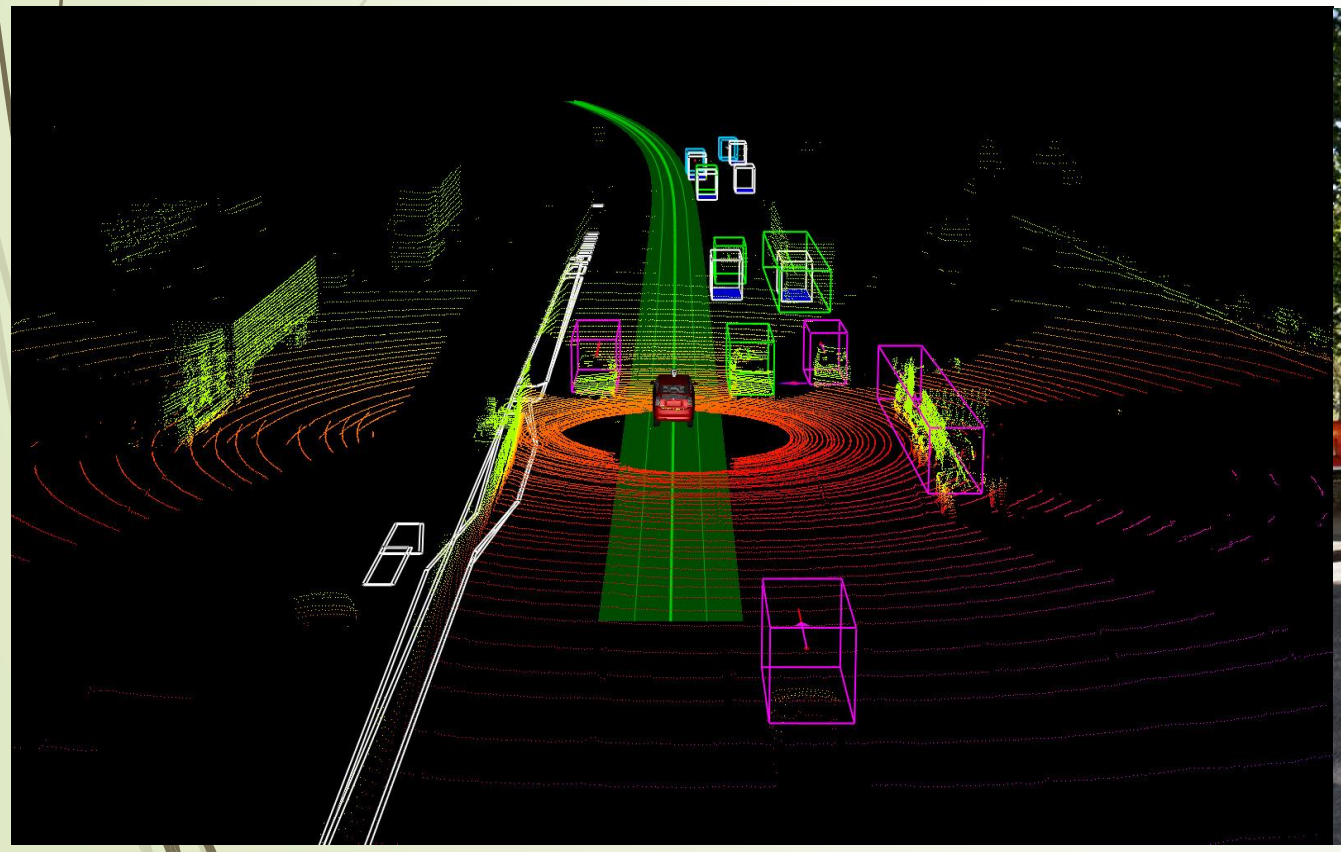
1. What is ADAS?
2. Which one is better, camera or LIDAR (Light Detection and Ranging) ?
3. Why CNN?
4. Basic architecture of CNN?
5. Work to be done.

Advanced Driver Assistance Systems (ADAS)

Some examples:

- Automatic parking
- Automotive navigation system
- Blind spot monitor
- Collision avoidance system
- Driver drowsiness detection
- Driver Monitoring System
- Forward Collision Warning
- Parking sensor
- Wrong-way driving warning

Which sensor is better,
Camera or LIDAR (Light Detection and Ranging) ?



Self Driving Vehicle

AUTOMOTIVE

Uber's "Self-Driving" Test Cars to Be Overseen by Driver and Engineer

The multimillion-dollar Pittsburgh pilot program will be open to passengers, and comes complete with a special driver and engineer in each vehicle

World's first self-driving taxis debut in Singapore

By ANNABELLE LIANG and DEE-ANN DURBIN Aug. 25, 2016 2:38 PM EDT

Tesla Motors

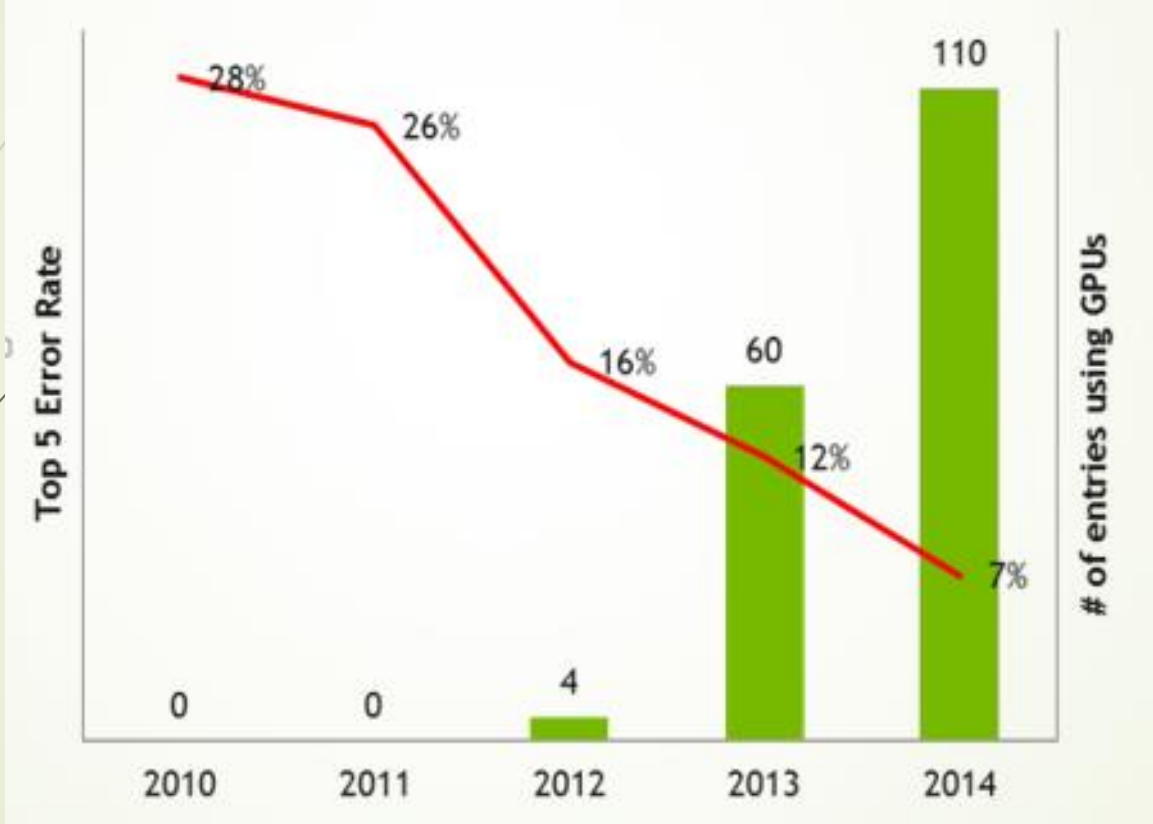
Energy storage company

From $70,000

Why Convolutional Neural Network(CNN) ?

1. Extremely complicated to design.
2. Requires a GPU(Graphics Processing Unit) for implementation.
3. Requires huge datasets to train.

IMAGENET Large Scale Visual Recognition Challenge (ILSVRC)



The image classification challenge

#1000 object classes

#1,431,167 images

2012- Alex Net

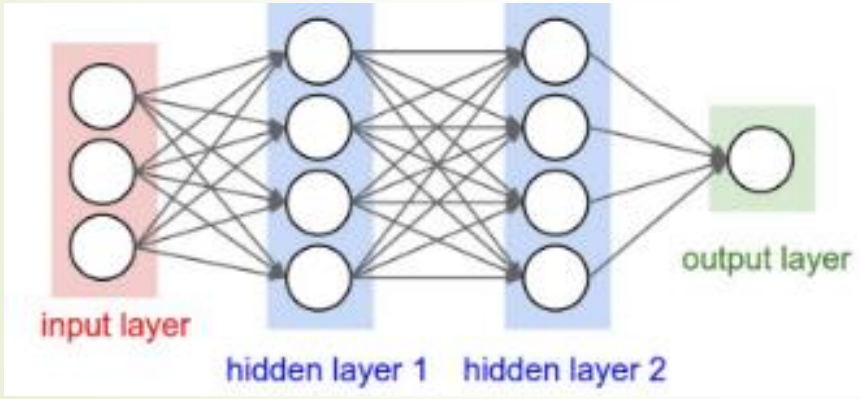
2013- ZF Net

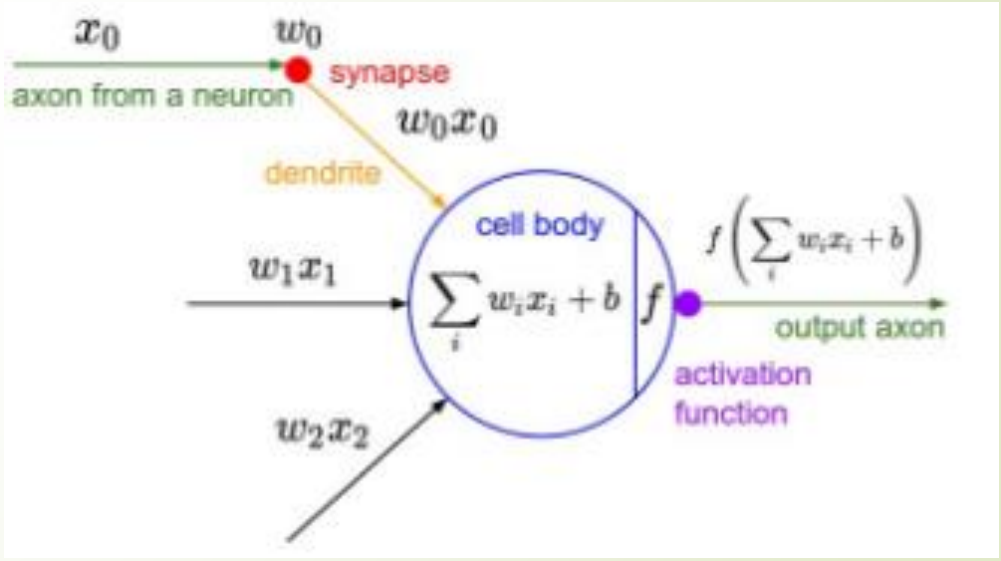
2014- GoogLeNet

2015-ResNet

CNN Architecture:

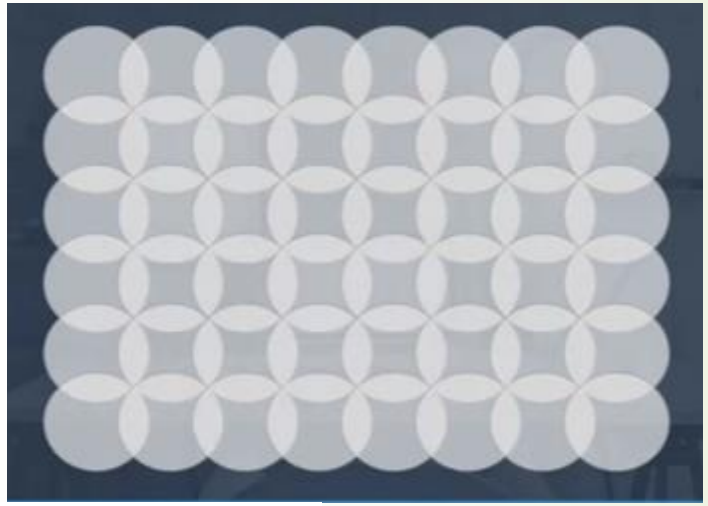
Input image size: it dictates the complexity of the net.
For example: 200x200x3, would lead to neurons that have 200*200*3 = 120,000 weights.

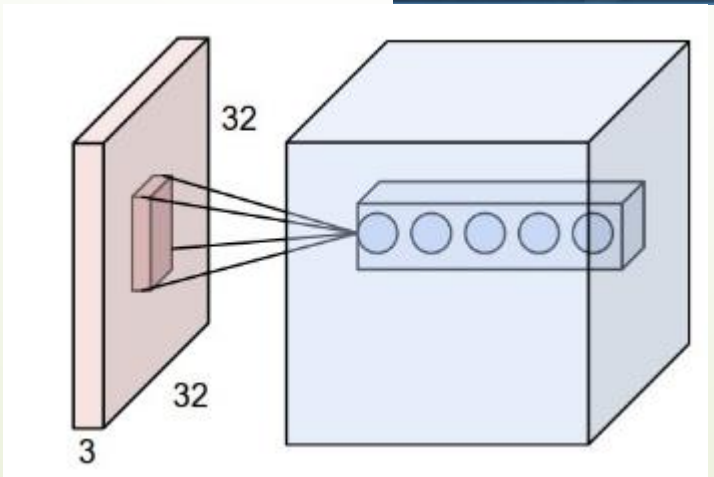




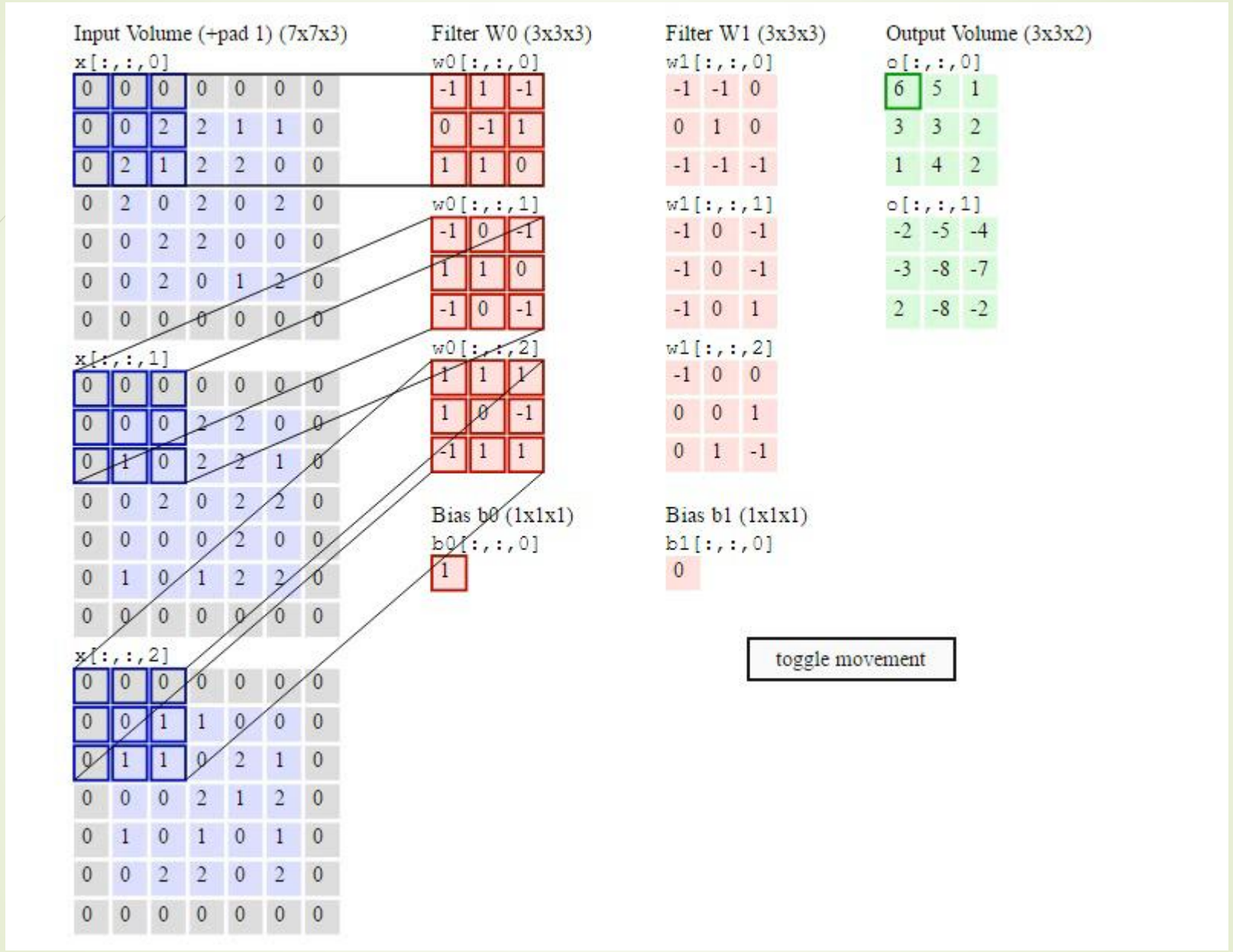
CNN Architecture:

Convolutional layer: the input image is divided into a set of overlapping blocks and each neuron is connected to one particular block



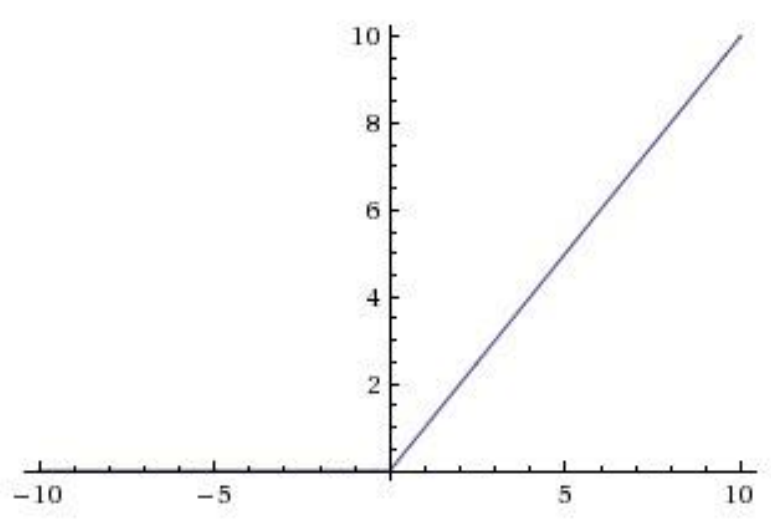


In a typical CNN there will be a minimum of 96 filters. All the neurons present in same filter will have same weight matrix.


Input Volume (+pad 1) (7x7x3)
x[:,:,0]
0 0 0 0 0 0 0
0 0 2 2 1 1 0
0 2 1 2 2 0 0
0 2 0 2 0 2 0
0 0 2 2 0 0 0
0 0 2 0 1 2 0
0 0 0 0 0 0 0
x[:,:,1]
0 0 0 0 0 0 0
0 0 0 2 2 0 0
0 1 0 2 2 1 0
0 0 2 0 2 2 0
0 0 0 0 2 0 0
0 1 0 1 2 2 0
0 0 0 0 0 0 0
x[:,:,2]
0 0 0 0 0 0 0
0 0 1 1 0 0 0
0 1 1 0 2 1 0
0 0 0 2 1 2 0
0 1 0 1 0 1 0
0 0 2 2 0 2 0
0 0 0 0 0 0 0
Filter W0 (3x3x3)
w0[:,:,0]
-1 1 -1
0 -1 1
1 1 0
w0[:,:,1]
-1 0 -1
1 1 0
-1 0 -1
w0[:,:,2]
1 1 1
1 0 -1
-1 1 1
Bias b0 (1x1x1)
b0[:,:,0]
1
Filter W1 (3x3x3)
w1[:,:,0]
-1 -1 0
0 1 0
-1 -1 -1
w1[:,:,1]
-1 0 -1
-1 0 -1
-1 0 1
w1[:,:,2]
-1 0 0
0 0 1
0 1 -1
Bias b1 (1x1x1)
b1[:,:,0]
0
Output Volume (3x3x2)
o[:,:,0]
6 5 1
3 3 2
1 4 2
o[:,:,1]
-2 -5 -4
-3 -8 -7
2 -8 -2
toggle movement

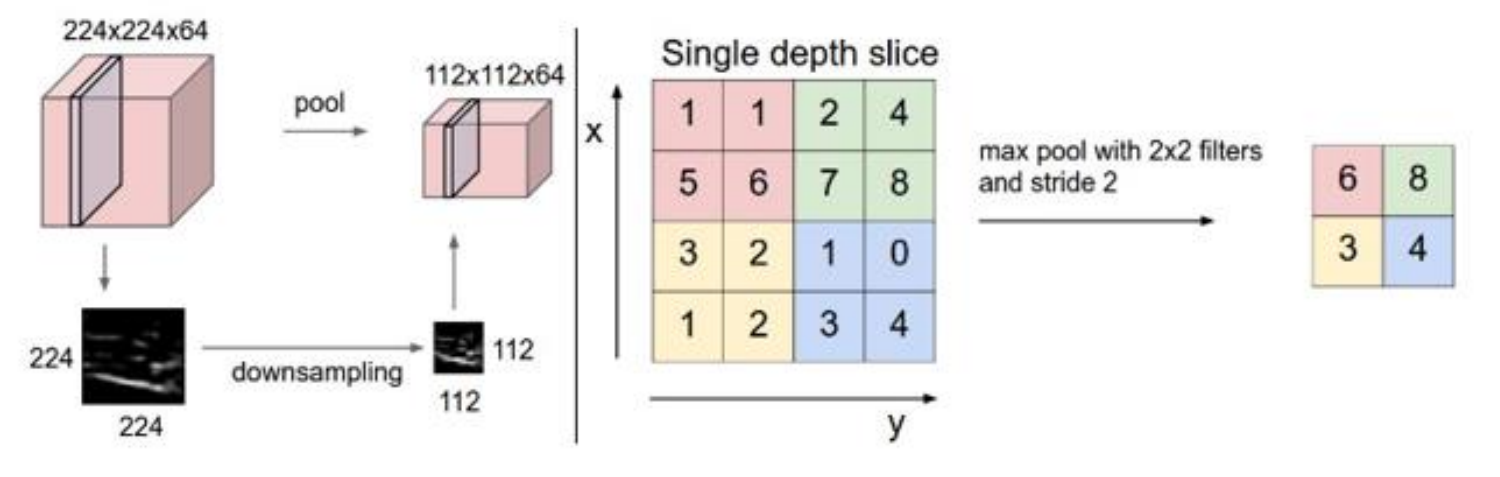
CNN Architecture:

Rectified Linear Units (ReLU) layer: This is a layer of neurons that applies the non-saturating activation function f(x) = max (0,x). Compared to other functions the usage of ReLU is preferable, because it results in the neural network training several times faster, without making a significant difference to generalisation accuracy



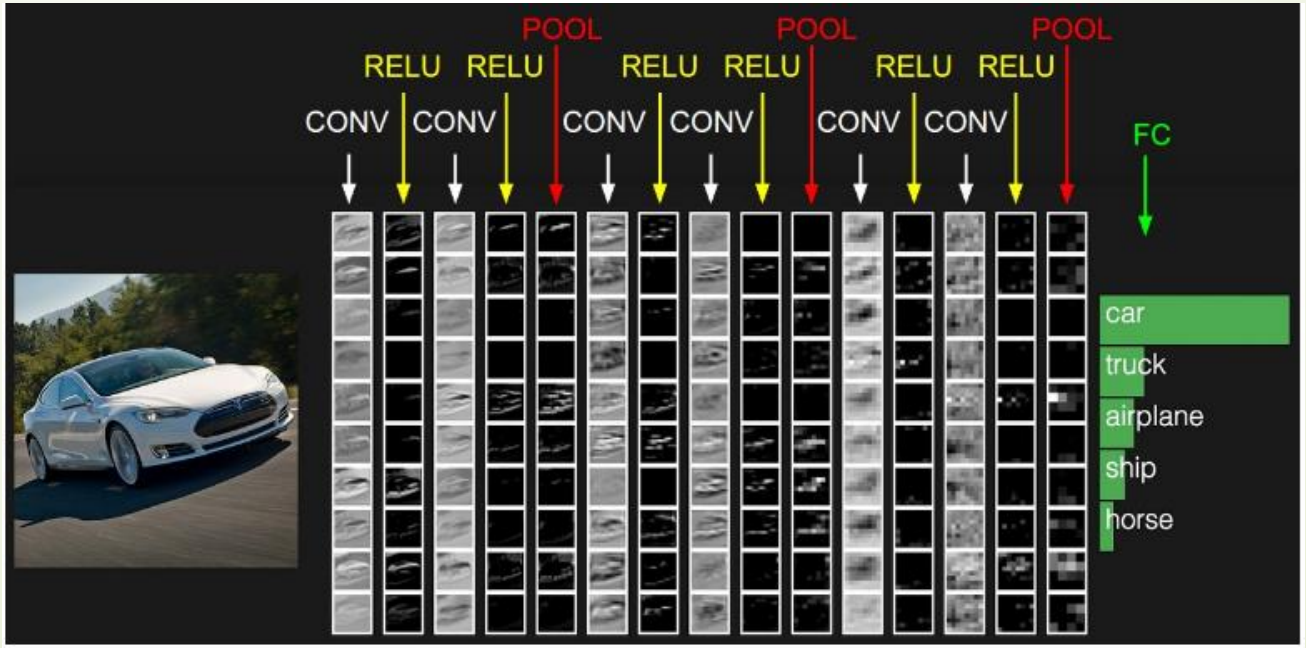
CNN Architecture:

Pooling layer:
Its function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network, and hence to also control overfitting.



CNN Architecture:

Fully connected layer:
Neurons in a fully connected layer have full connections to all activations in the previous layer, as seen in regular Neural Networks. Their activations can hence be computed with a matrix multiplication followed by a bias offset.



Work to be done

Further study will be made on the recently published CNN [4] i.e. GoogLeNet, Alexnet, ZF Net, ResNet(Microsoft Research Asia).

A new CNN will be developed (in Python using Caffe framework) integrating the latest advances with self-developed ideas with an aim to reduce the existing network complexity and also to improve the classification accuracy.

Thank You