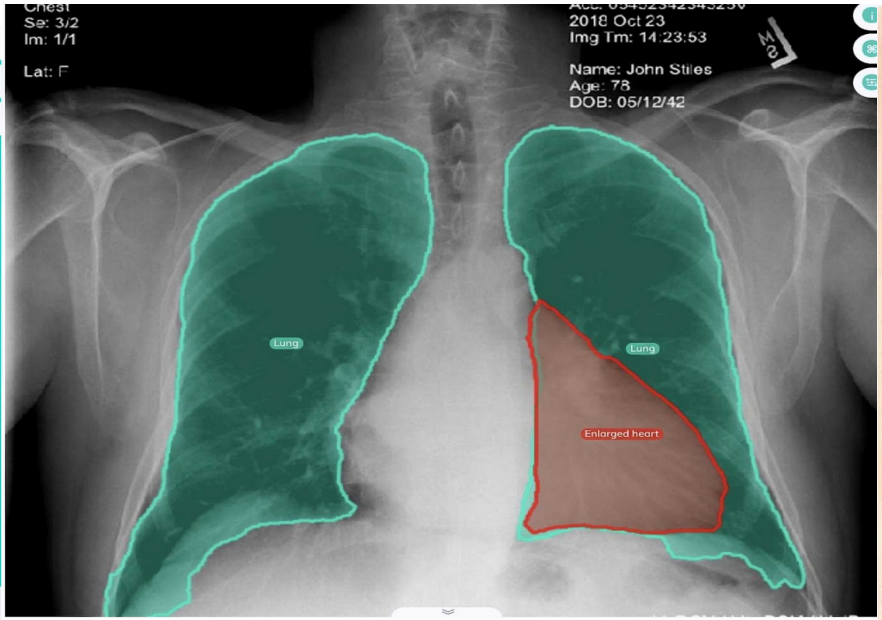
Introduction to Computer Vision

Solomon Ubani, Ph.D.

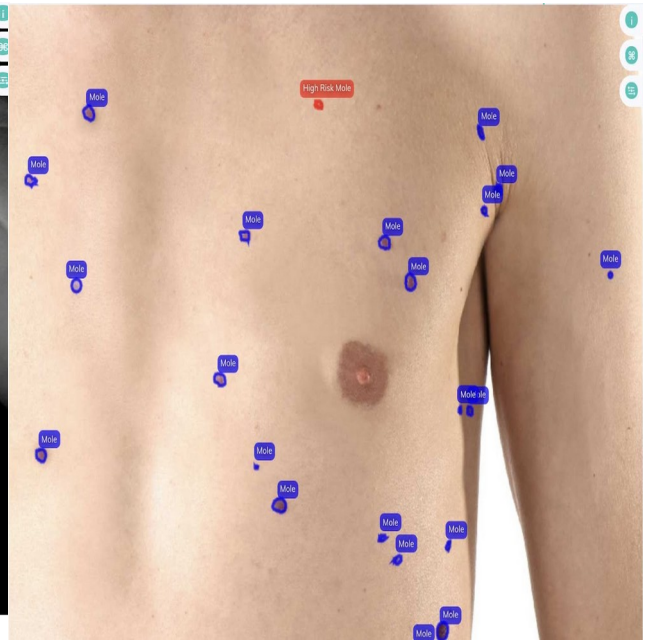
What is Computer Vision

- CV deals with how computers can gain high-level understanding from digital images or videos (Wikipedia).
- some applications of computer vision

Self driving cars



x-ray analysis



cancer detection

Why is computer vision difficult?

Viewpoint variation

Illumination

Scale

Why is computer vision difficult?

Intra-class variation

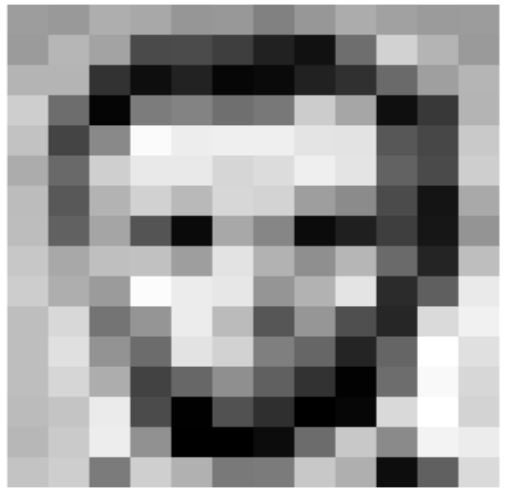
Motion (Source: S. Lazebnik)

Background clutter

Occlusion

Computer Vision (CV)

- Images are represented as Pixels in Computer Vision tasks

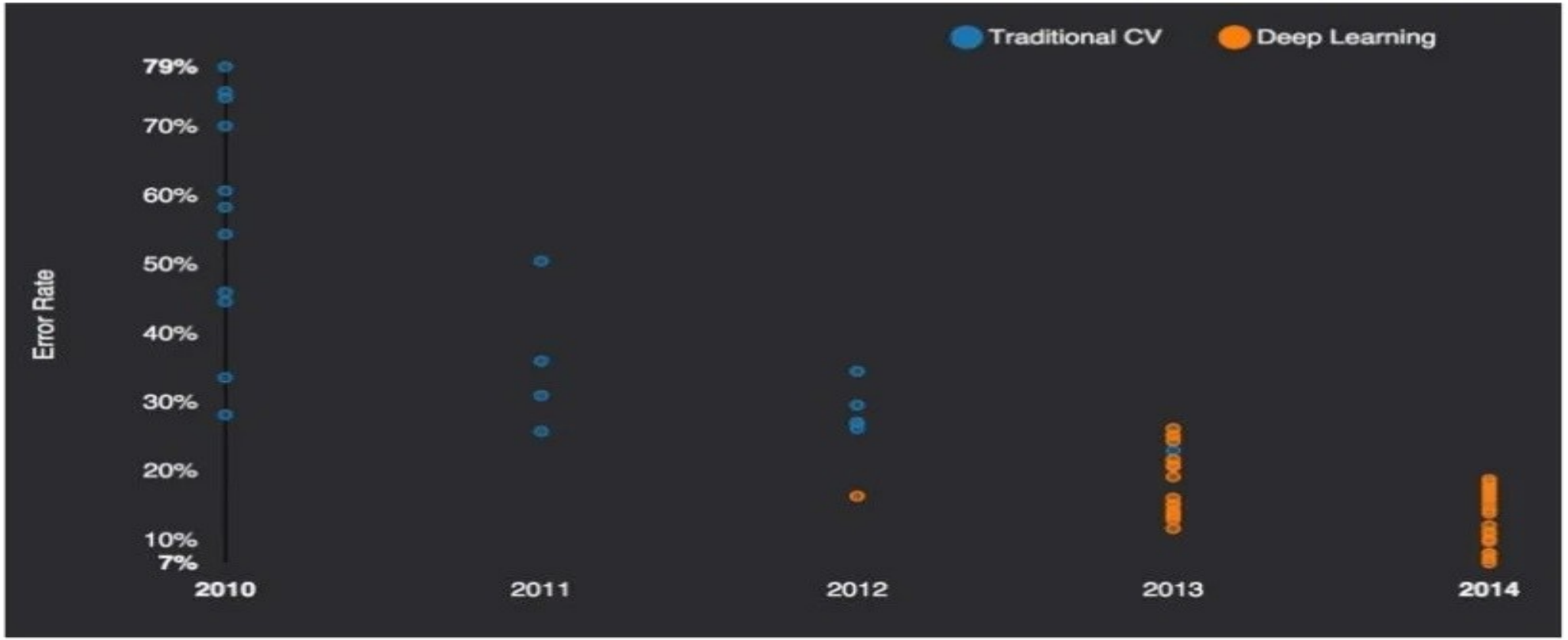


157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

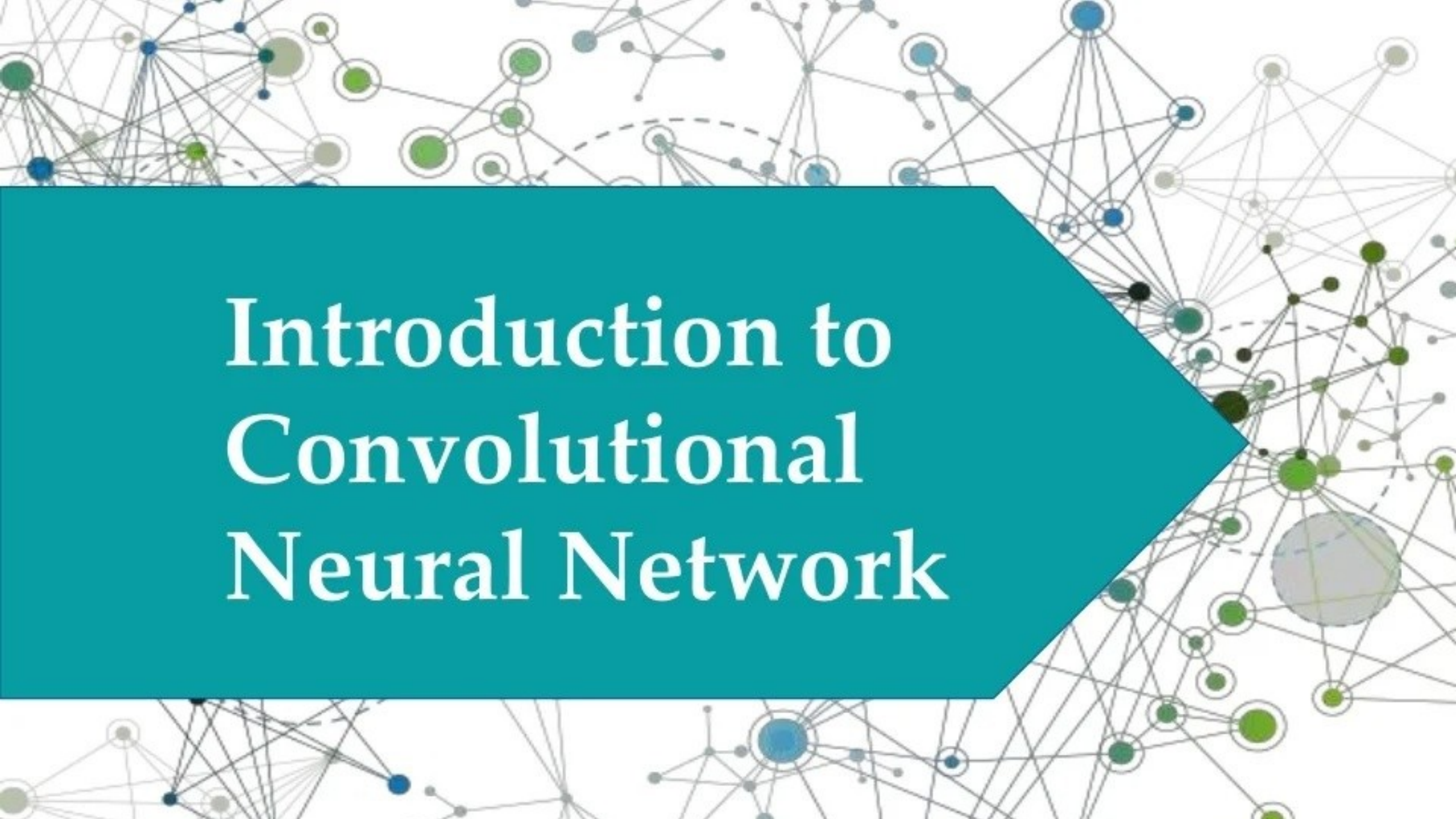
157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	75	20	169
189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	190
205	174	155	252	236	231	149	178	228	43	95	234
190	216	116	149	236	187	86	150	79	38	218	241
190	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	103	143	96	50	2	109	249	215
187	196	235	75	1	81	47	0	6	217	255	211
183	202	237	145	0	0	12	108	200	138	243	236
195	206	123	207	177	121	123	200	175	13	96	218

A brief History

The Big Bang aka “One net to rule them all”

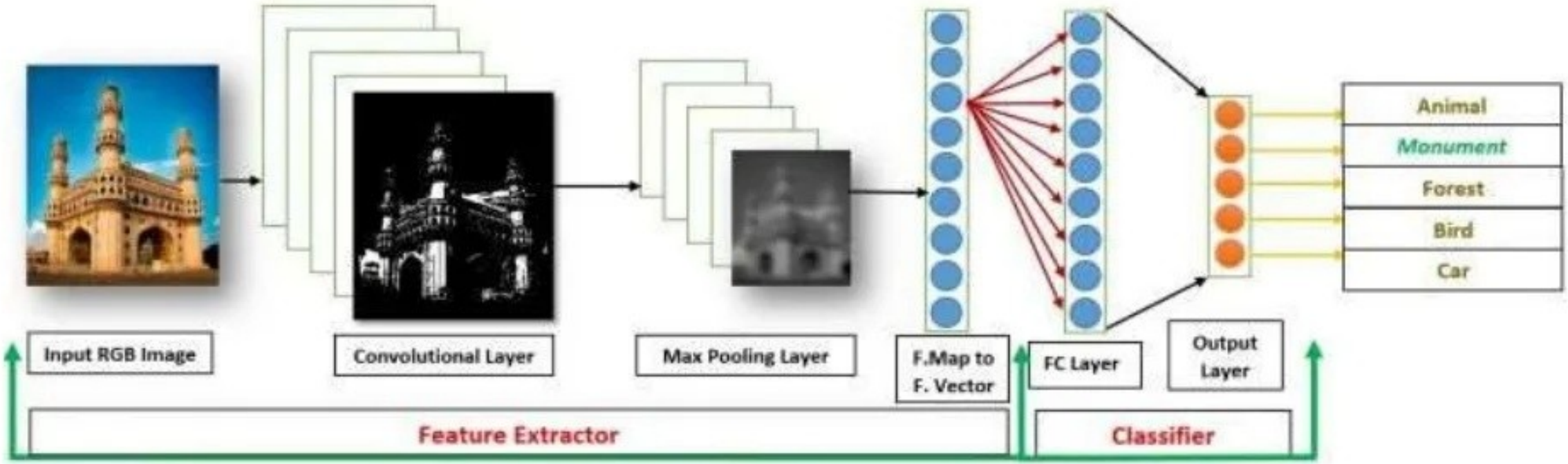


ImageNet: The “computer vision World Cup”

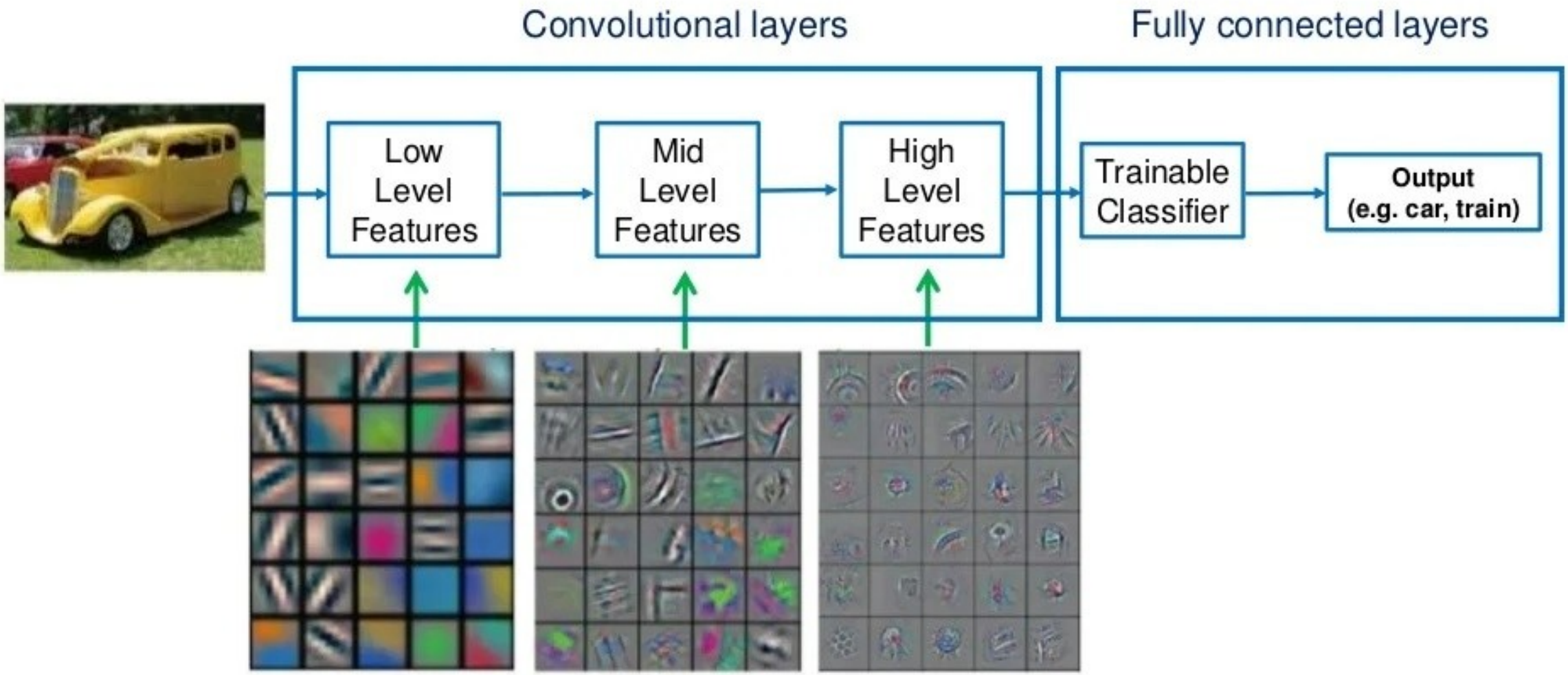

Introduction to
Convolutional
Neural Network

Introduction

Convolutional Neural Networks (**CNNs**) **learns multi-level features and classifier in a joint fashion** and performs much **better than traditional approaches** for various image classification and segmentation problems.



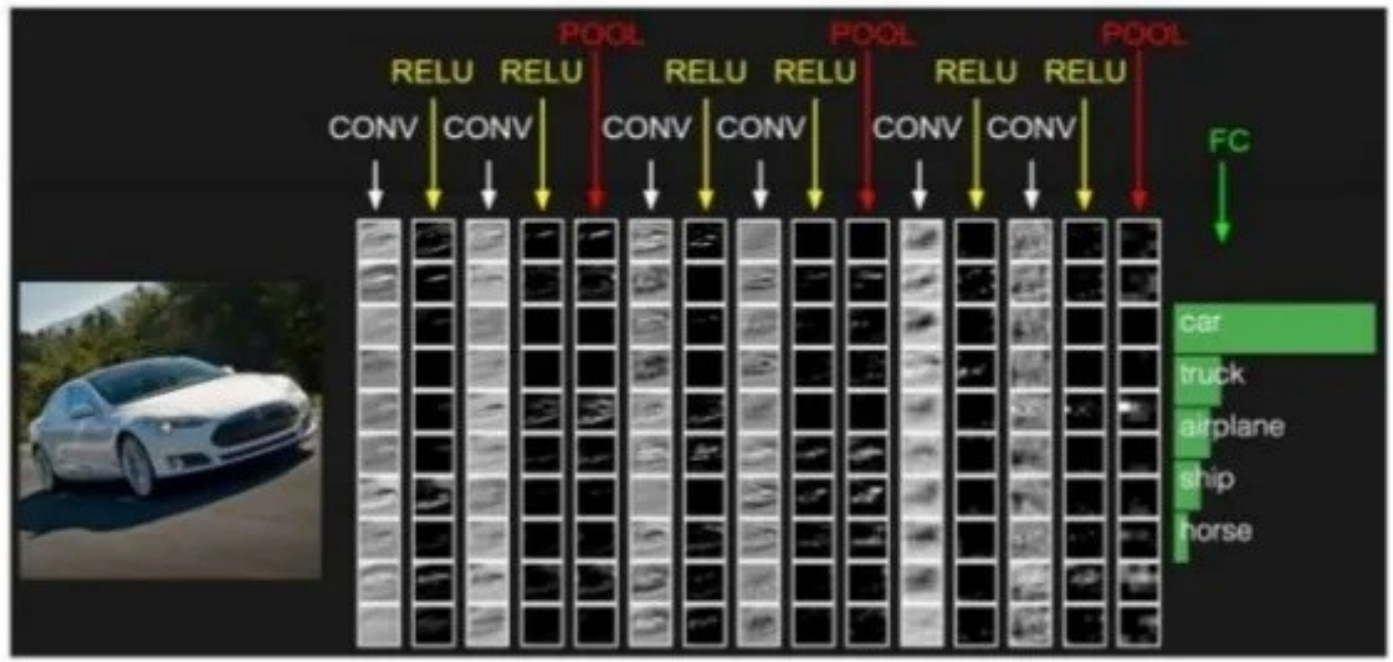
CNN – What do they learn?



CNN - Components

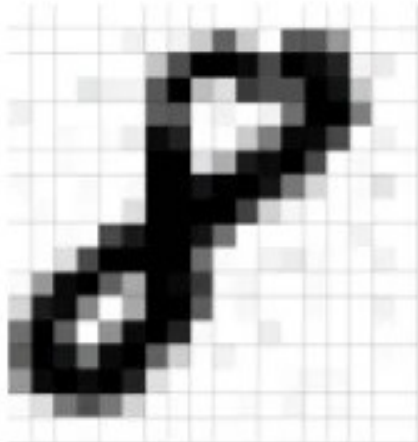
There are four main components in the CNN:

1. Convolution
2. Non-Linearity
3. Pooling or Sub Sampling
4. Classification (Fully Connected Layer)



Input

- An Image is a matrix of pixel values.
- If we consider a gray scale image, the value of each pixel in the matrix will range from 0 to 255.
- If we consider an RGB image, each pixel will have the combined values of R, G and B.



What We See

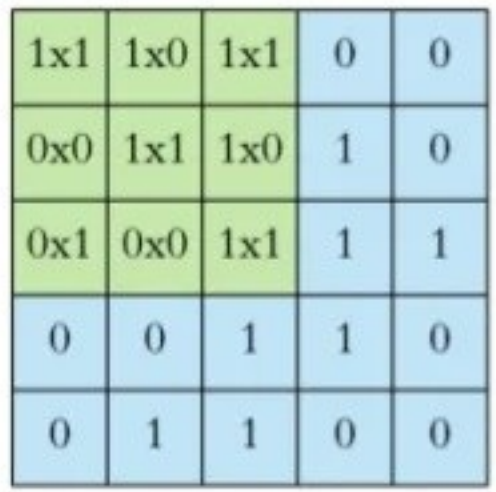
What Computers See

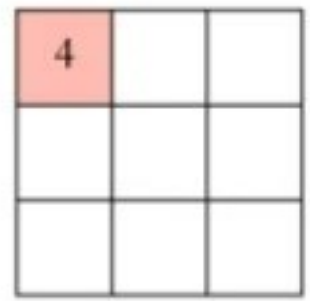
Convolution

The primary purpose of Convolution in case of a CNN is to extract features from the input image.

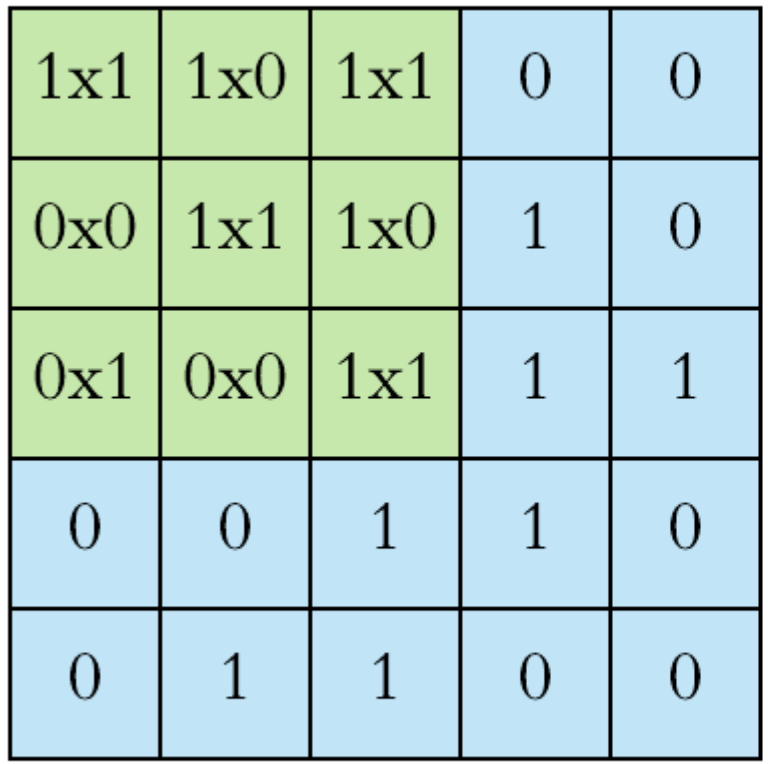
Filter / Kernel / Feature detector

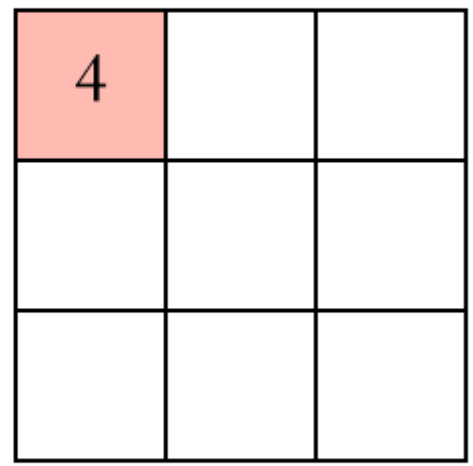
Image





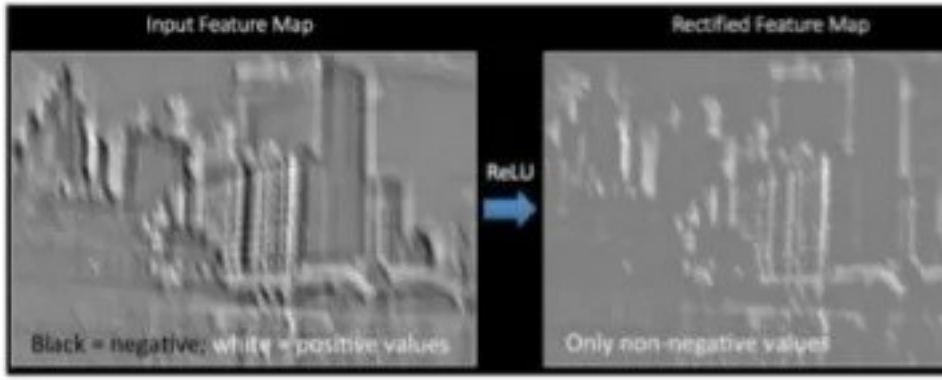
Convolved Feature /
Activation Map /
Feature Map

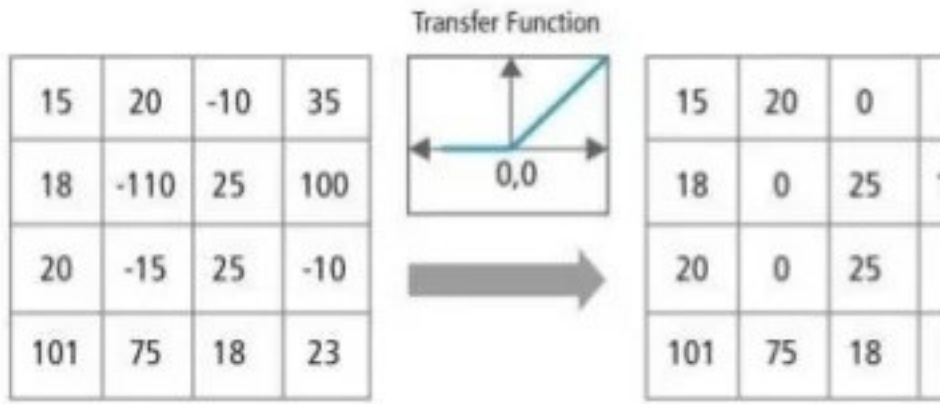

1x1 1x0 1x1 0 0
0x0 1x1 1x0 1 0
0x1 0x0 1x1 1 1
0 0 1 1 0
0 1 1 0 0


4

Non-Linearity (ReLU)

- Replaces all negative pixel values in the feature map by zero.
- The purpose of ReLU is to introduce non-linearity in CNN, since most of the real-world data would be non-linear.
- Other non-linear functions such as tanh (-1,1) or sigmoid (0,1) can also be used instead of ReLU (0,input).





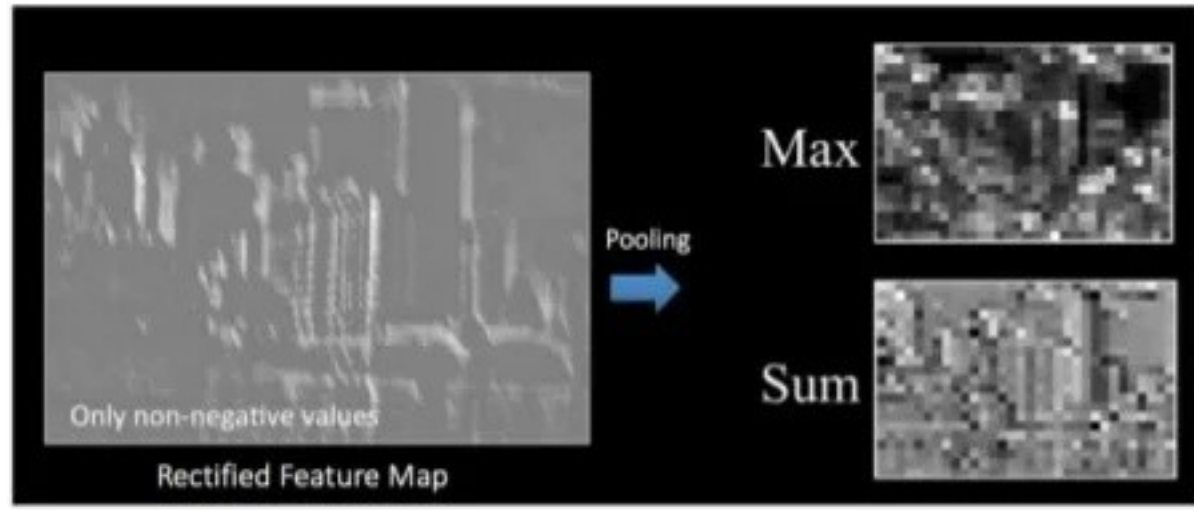
Pooling

Reduces the dimensionality of each feature map but retains the most importa information. Pooling can be of different types: Max, Average, Sum etc.

1	3	2	9
7	4	1	5
8	5	2	3
4	2	1	4

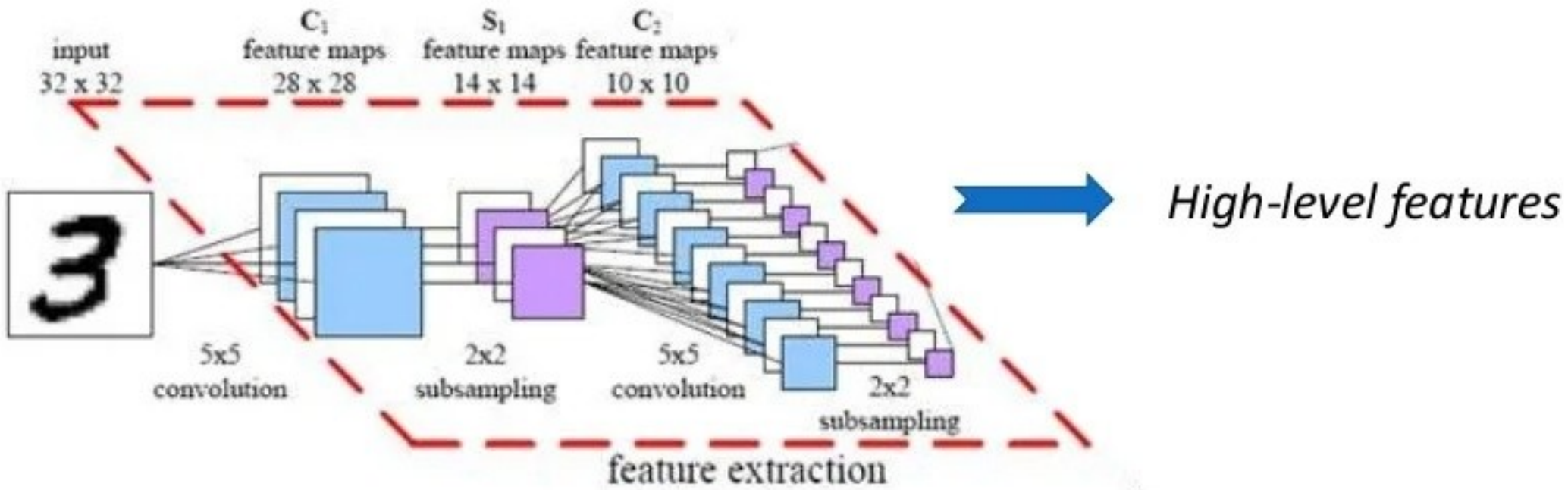
2×2 region

7	9
8	

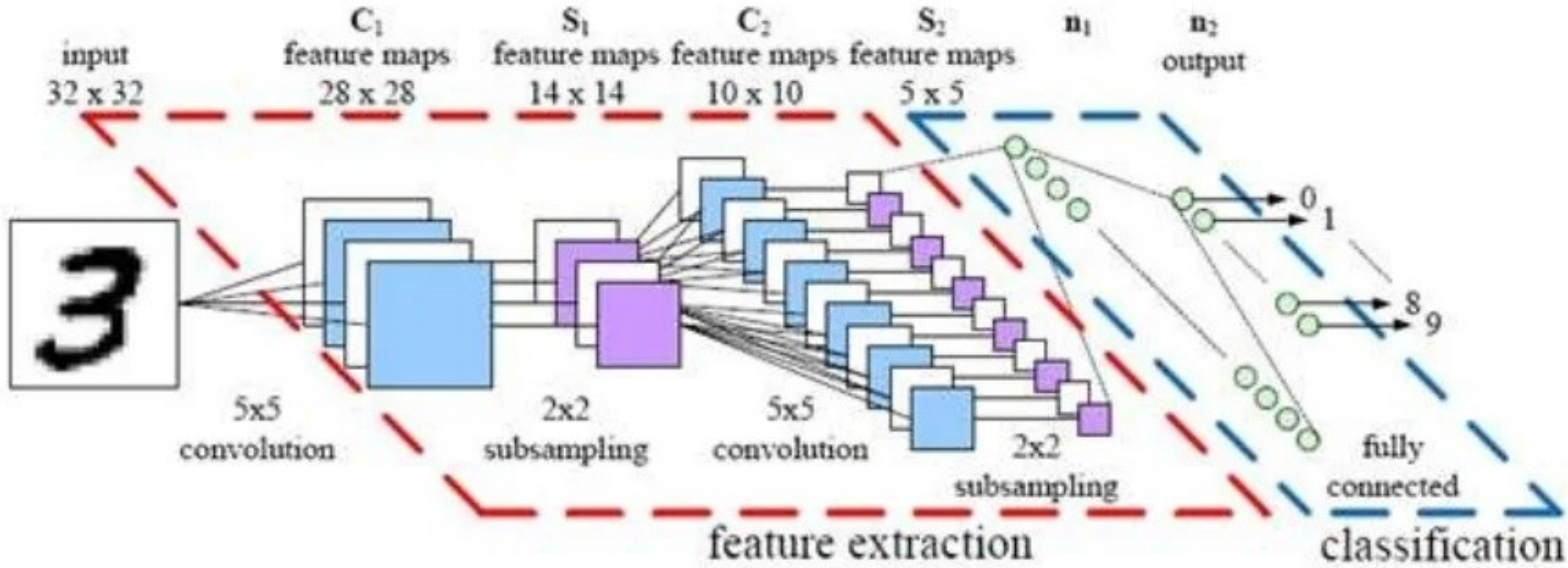


Story so far

- Together these layers extract the useful features from the images.
- The output from the convolutional and pooling layers represent high-level features of the input image.



Overall CNN Architecture



References

- https://towardsdatascience.com/everything-you-ever-wanted-to-know-about-computer-vision-heres-a-look-why-it-s-so-awesome-e8a58dfb641e