greatlearning
Power Ahead

Introduction to Convolutional Neural Networks

Topics you have covered in week 1 videos

- Image data structures and types
- Convolution of Filters/kernels on Images
- Convolutional neural networks vs Fully connected networks
- Convolutional layers
- Average and Max Pooling layers
- Forward prop and Backprop in CNNs
- Implementation of CNNs in Keras

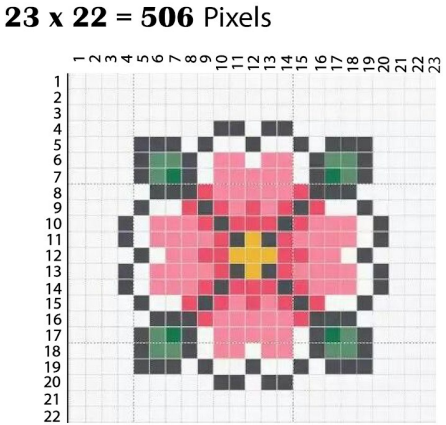
greatlearning
Power Ahead

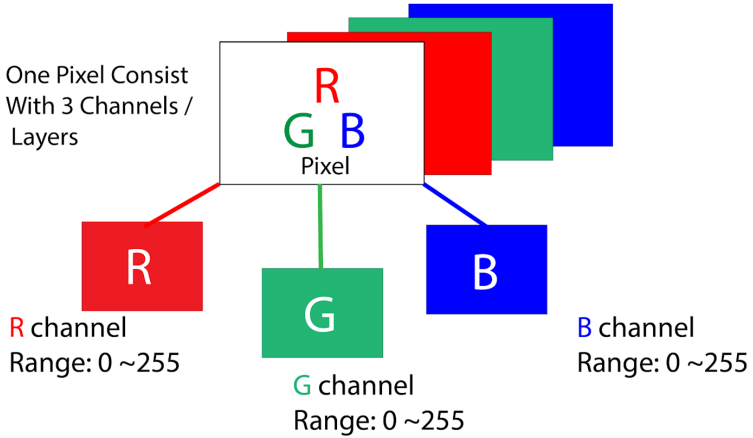
Session agenda

- Basics of Image handling in python
- Understanding filtering
- Understanding CNNs
- Case study
- Questions

Image basics

Digital Images

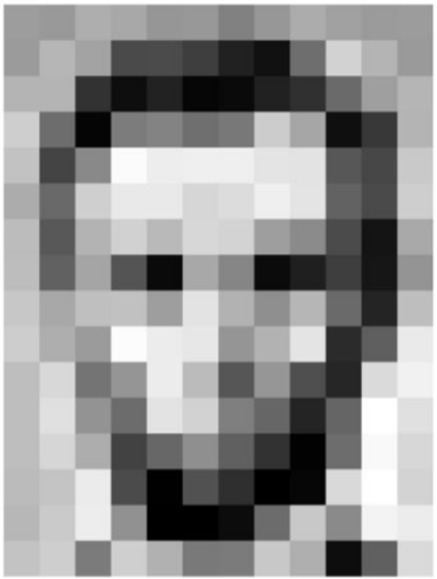




Digital Images are 2D or 3D matrices depending upon if they are greyscale or colour images. An entire image is discretized into small units called pixels. Each pixel is identified by it's row and column number. Each pixel in a grayscale or black and white image has a single value (Range 0~255) denoting the intensity of the image at that pixel location.
Each pixel in a colour image has respectively 3 values(Range 0~255) denoting the intensities of Red, Blue and Green channels.
The size of the image is generally denoted by the number of rows and columns.

Image and text source: dzone

Digital Images

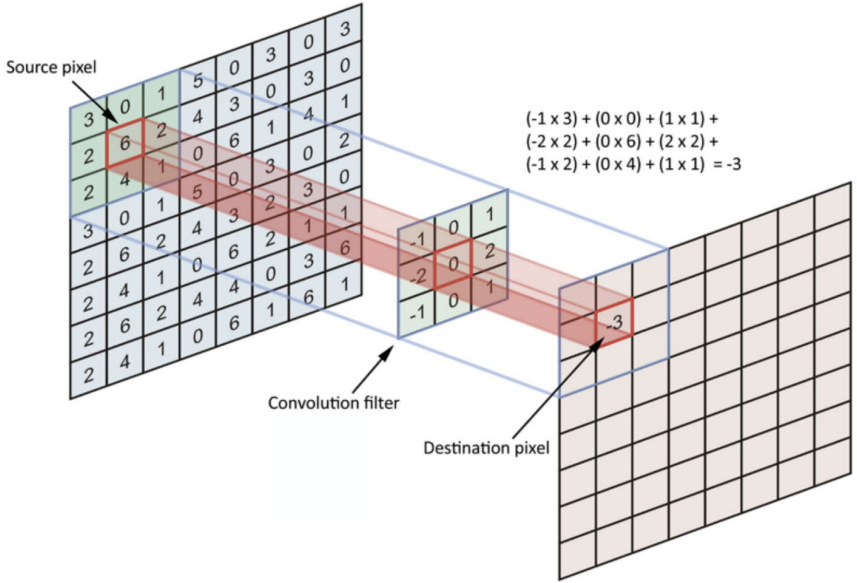


157	153	174	168	150	152	129	151	172	161	155	156
155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	159	181
206	109	5	124	131	111	120	204	166	15	56	180
194	68	137	251	237	239	239	228	227	87	71	201
172	105	207	233	233	214	220	239	228	98	74	206
188	88	179	209	185	215	211	158	139	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

Image and text source: dzone

Image filtering



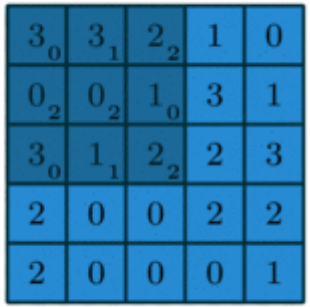
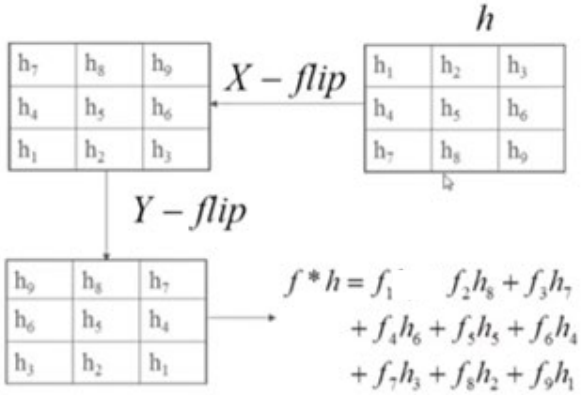


Image source: towardsdatascience

Convolution vs Correlation



Convolution:

$$G = h * F \qquad G[i,j] = \sum_{u=-k}^{k} \sum_{v=-k}^{k} h[u,v]F[i-u, j-v]$$

Correlation:

$$G = h \otimes F \qquad G[i,j] = \sum_{u=-k}^{k} \sum_{v=-k}^{k} h[u,v]F[i+u, j+v]$$

Convolution is basically flipping the kernel via X-axis and Y-axis and then performing a correlation with the resultant kernel

Features from kernels

Edge detection

$$* \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix} =$$

Kernel

Sharpen

$$* \begin{bmatrix} 0 & -1 & 0 \\ -1 & 5 & -1 \\ 0 & -1 & 0 \end{bmatrix} =$$

Image source: towardsdatascience

Features from kernels

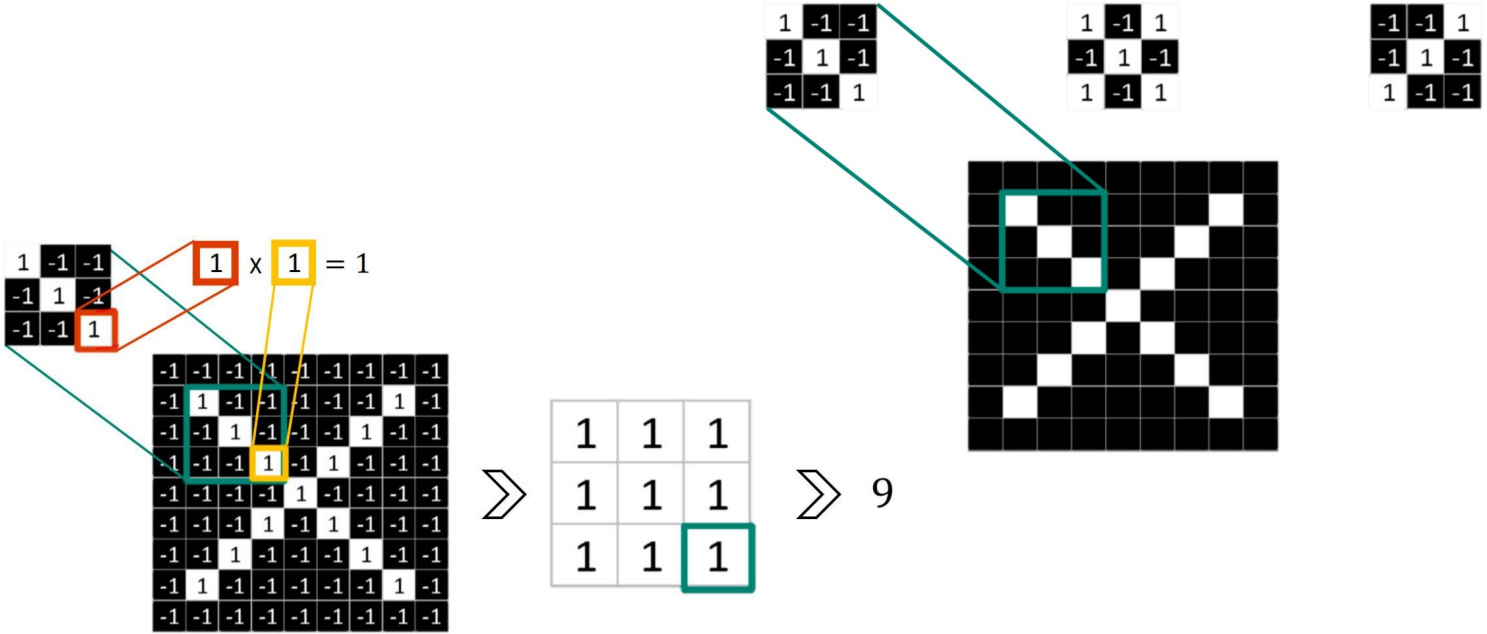
Original

Sharpen

Edge Detect

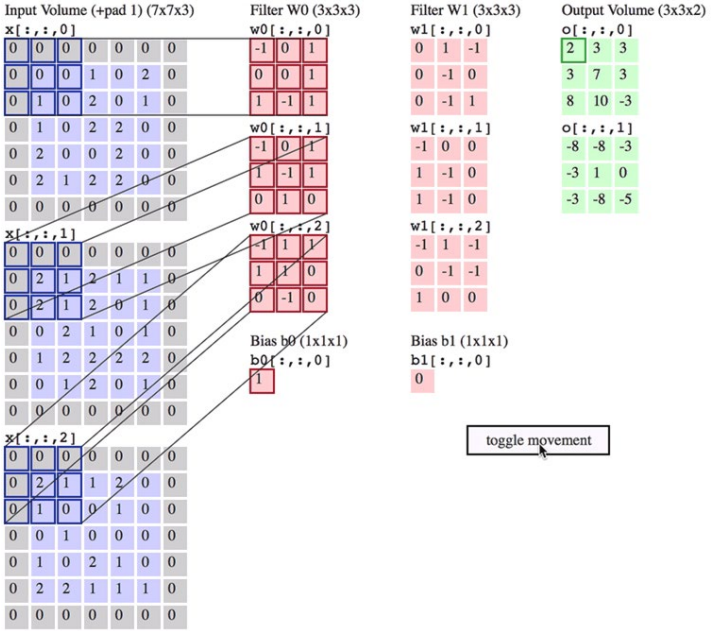
Stronger Edge Detect

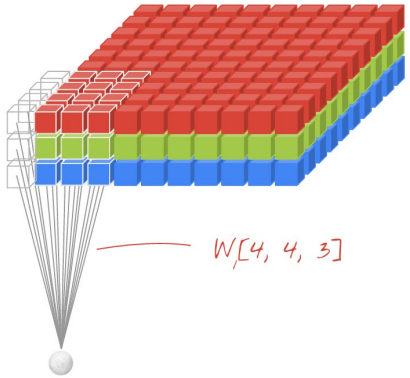
Features from kernels



Convolutional neural networks

Convolutional layer





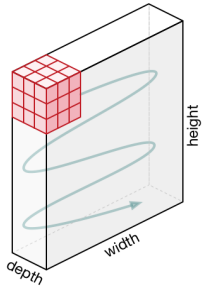


Image source: cs231 notes

Convolutional layer

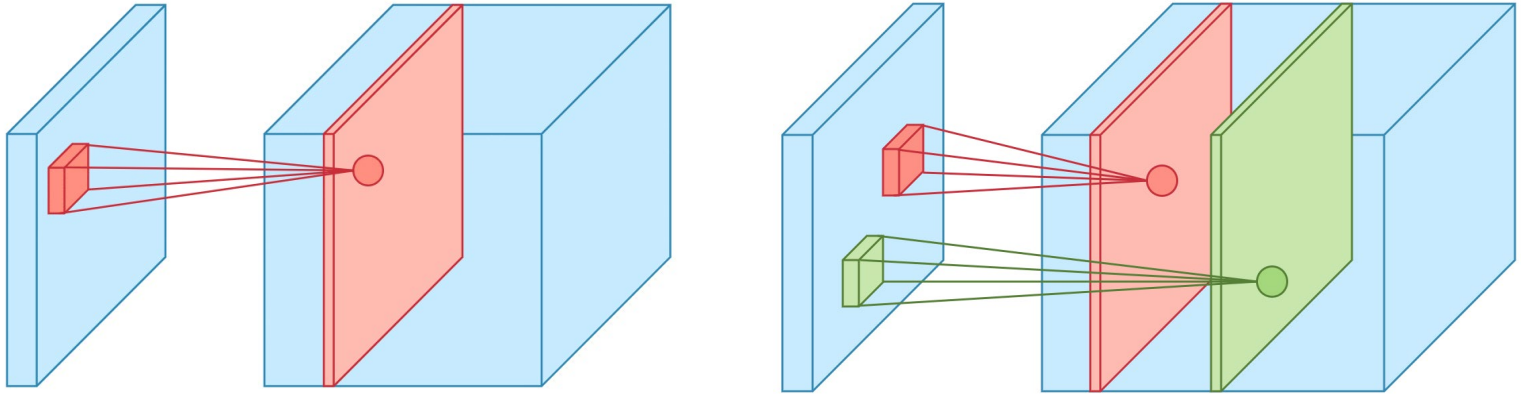


Image source: towardsdatascience

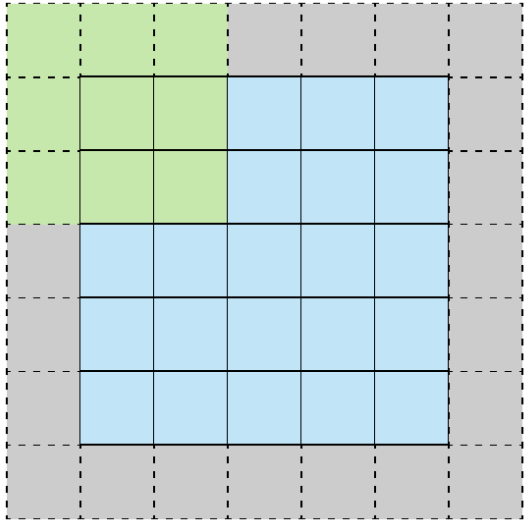
Convolutional layers

RGB Image	Filter 1 (4 filters)	Convolution layer 1	Filter 2	Convolution layer 2	Filter 3	Convolution layer 3
32x32x3	2x2x4 (s=2)	16x16x4	2x2x8 (s=2)	8x8x8	8x32	2x2x32 (s=2)

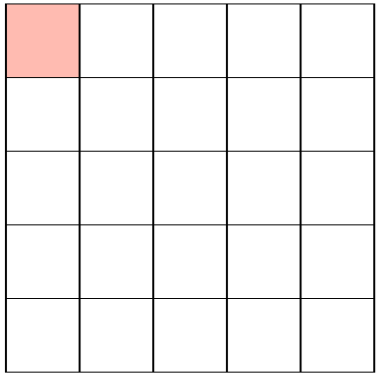
$$\text{output width} = \frac{W - F_w + 2P}{S_w} + 1$$

$$\text{output height} = \frac{H - F_h + 2P}{S_h} + 1$$

Padding



Stride 1 with Padding



Feature Map

Image source: towardsdatascience

Pooling Layers

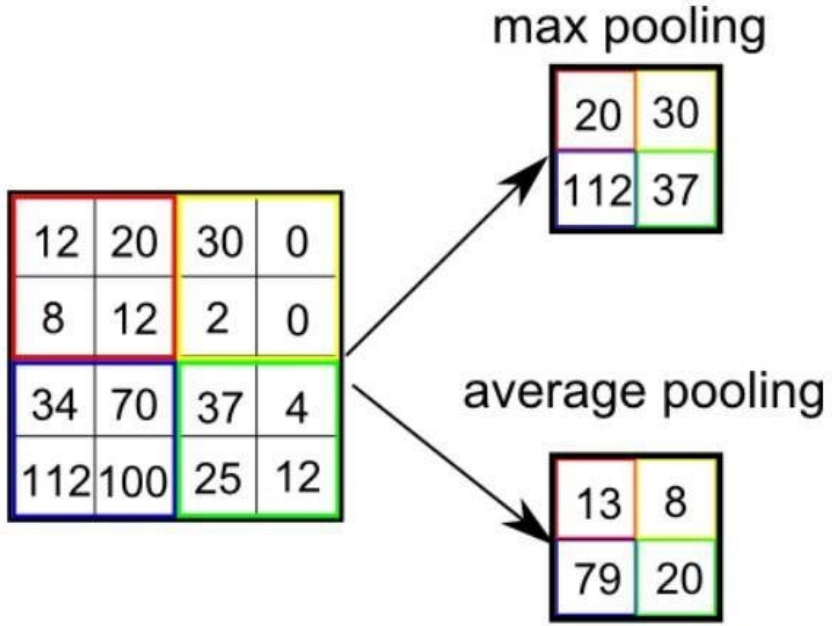


Image source: towardsdatascience

Flatten and fully connected layers

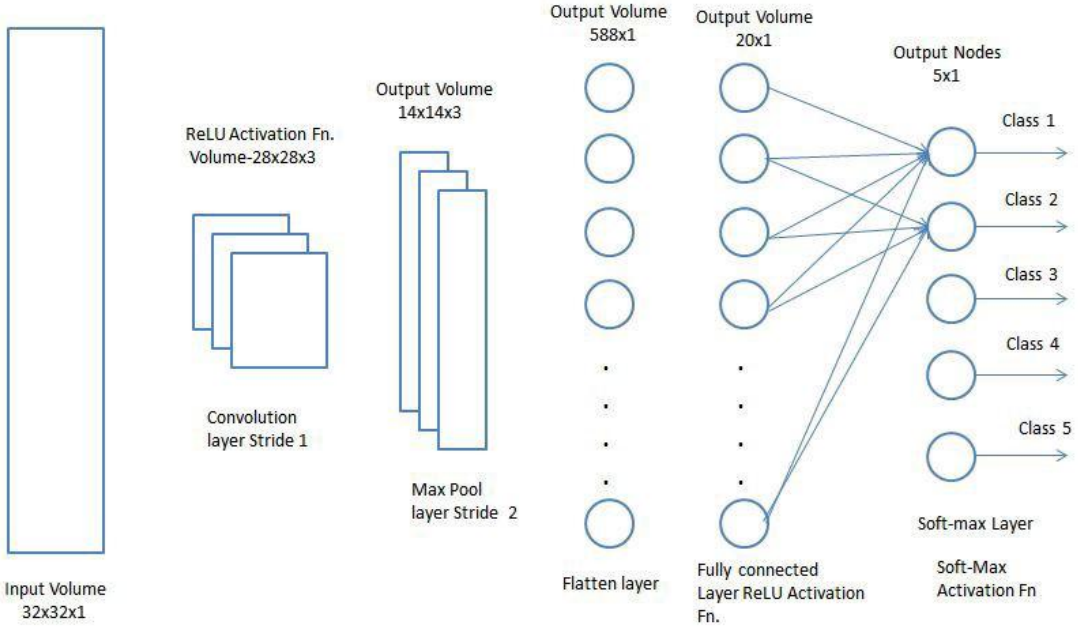


Image source: towardsdatascience

CNNs for classification

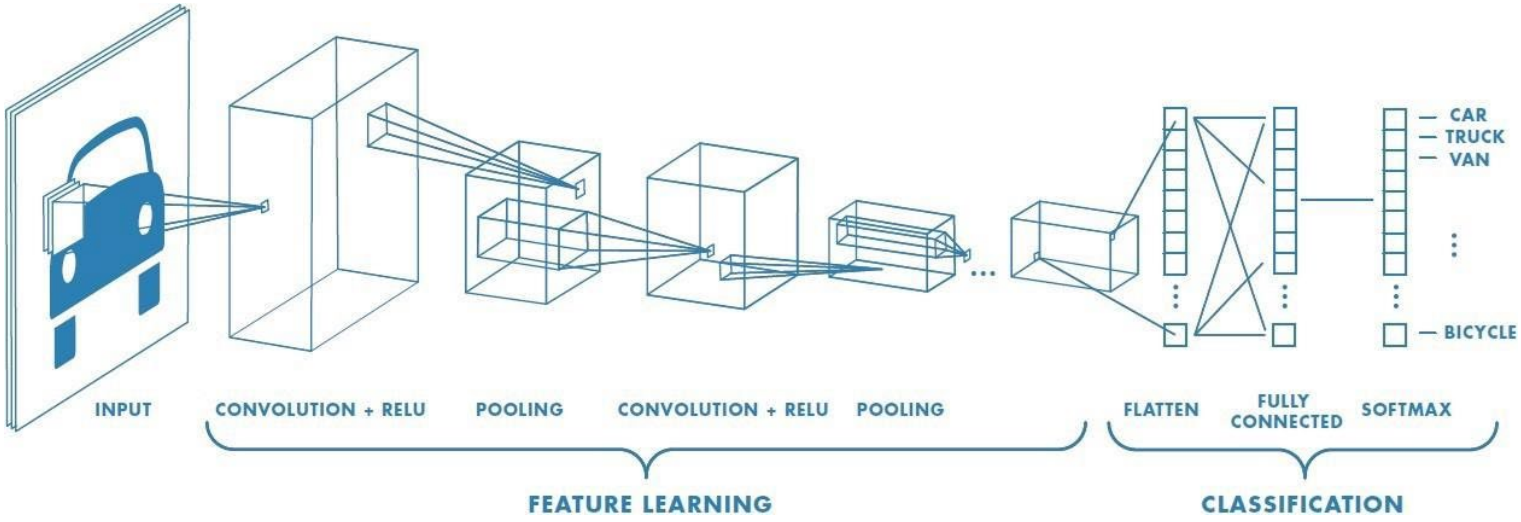


Image source: towardsdatascience

Summary

In summary - Deep CNNs

- Kernels are feature extractors
- CNNs are trained to learn the features
- Convolutional Layers
- Stride and Padding
- Max and Average Pooling Layers
- Flattening and Fully connected final layers
- CNN Classifier

Thank you! :)

Questions are always welcome