A few words about Convolutional Neural Networks

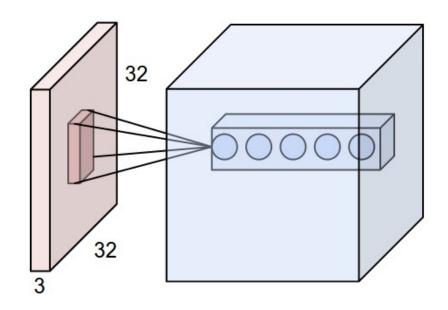
Motivation:

MLP can require a lot of parameters:

256x256 image x3 channels and 1000 nodes in first hidden layer MORE than 200 MILLION parameters!

The MLP hidden layer ignore all the spatial structure!

Convolutional layers associate each of their nodes with a weighted window («receptive field », « filter kernel »)

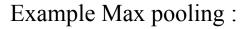


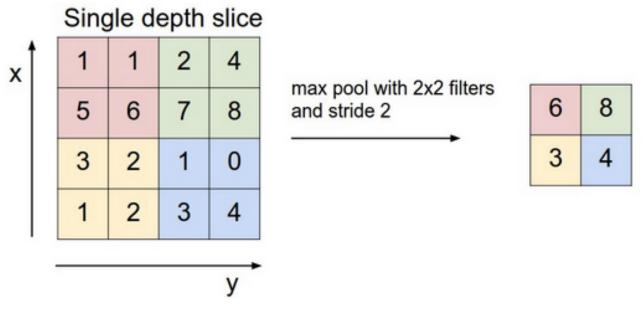
www Sources for figures & examples https://scikit-learn.org/stable/
https://skit-learn.org/stable/
https://skymind.ai/wiki/

CNN CONVOLUTIVE LAYER example

RGB channels	Filter 1	Filter 2	Output stack
Input Volume (+pad 1) (7x7x3)	Filter W0 (3x3x3) w0[:,:,0]	Filter W1 (3x3x3)	-
x[:,:,0] 0 0 0 0 0 0 0	-1 1 0	w1[:,:,0]	o[:,:,0] 2 -4 -3
0 2 0 1 2 2 0	0 T -1	1 1 -1	3 2 -4
	-1 -1 -1	1 0 1	6 13 4
0 2 2 2 1 2 0	w0[:,:,1]	w1[:,:,1]	o[:,:,1] 3 1 5
0 0 0 2 1 0 0			
0 0 0 2 0 0 0	1 0 1	0 1 0	-3 2 7
0 0 0 0 0 0 0	-1 -1 1	-1 -1 0	0 2 5
x[:,:,1]	w0[:,:,2]	w1[:,,2]	
0 0 0 0 0 0	0 1 1	0 0	
0 2 2 1 0 0 0	-1 0 0	0 -1 0	
0 1 1 0 2 0 0	-1 -1 -1	1 0 0	
0 1 0 1 1 0 0	7. /2 /	//	
0 2 2 1 2 0 0	Bjas b0 (1x1x1) b0[:,:,0]	Bias b1 (1x1x1) b1[:,:,0]	
0 2 1 2 1 2 0	1	0	
x[:,:,2] 0 0 0 0 0 0 0		toggle	e movement
0 0 0 1 0 0 0			
0 2 1 2 1 1 0			
0 1 0 2 2 1 0			
0 0 2 0 2 2 0			
0 0 0 2 1 0 0			
0 0 0 0 0 0 0			

Pooling / Downsampling with CNNs





- → Only the locations on the image that showed the strongest correlation to each feature (the maximum value) are preserved, and those maximum values combine to form a lower-dimensional space.
- \rightarrow Decreases the amount of storage and processing requirements but at the price of loss of information about lesser values

Example of alternating sequences of transformations involved in CNNs

