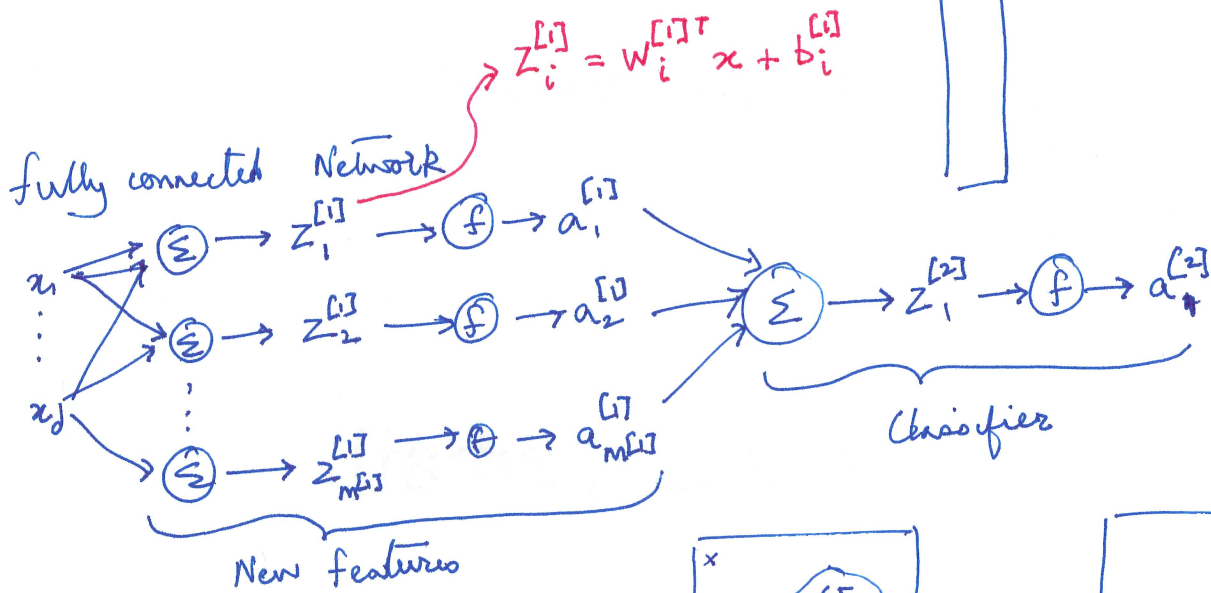
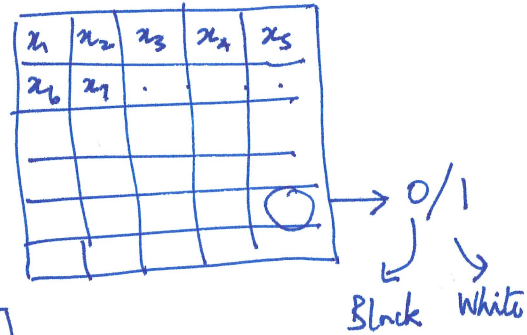


Image Classification: Images  $\rightarrow$  classify

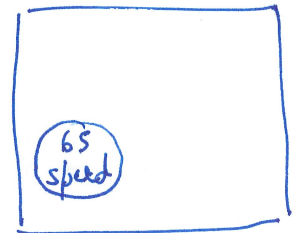
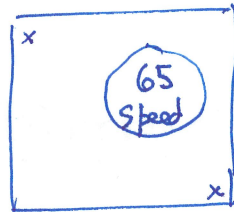
AlexNet (2012): Convolution Neural Network (CNN)

Images:   $\rightarrow$  RGB. Today: B/W images

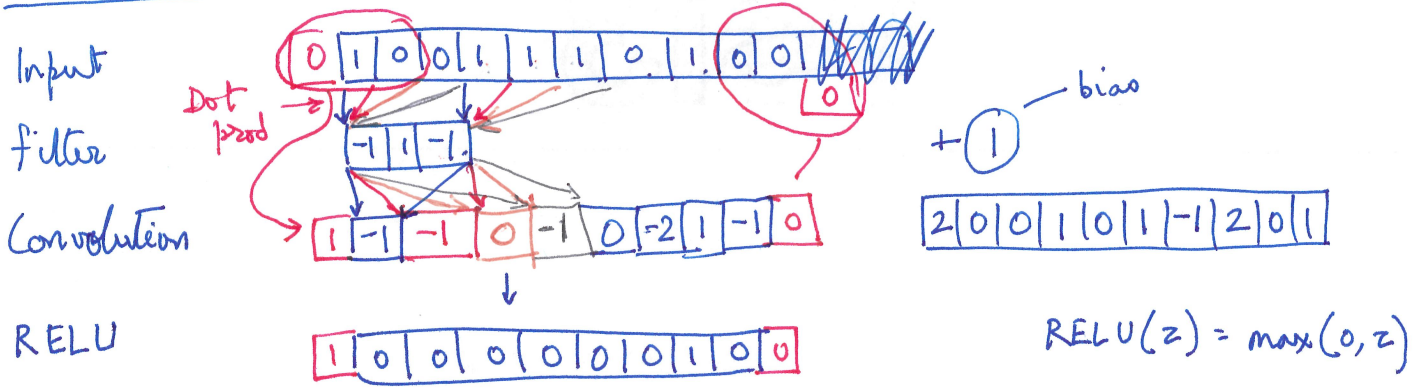


Properties of Images:

- Spatial locality
- Translation Invariance



Convolution in 1-D:



Padding: Adding 0's around the image and then apply convolution to this "expanded image".

Filter:  $f \in \mathbb{R}^f \rightarrow$  any vector of real numbers. + bias

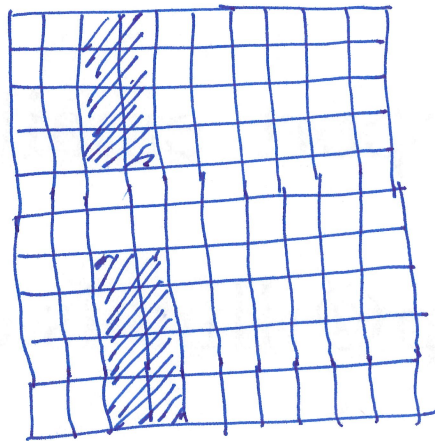
convolution:  $f^T \cdot x_{[i, i+1, \dots, i+f]} + b$ .

Example: Input  $\in \mathbb{R}^{10}$ , Output  $\in \mathbb{R}^{10}$ , filter size 3.

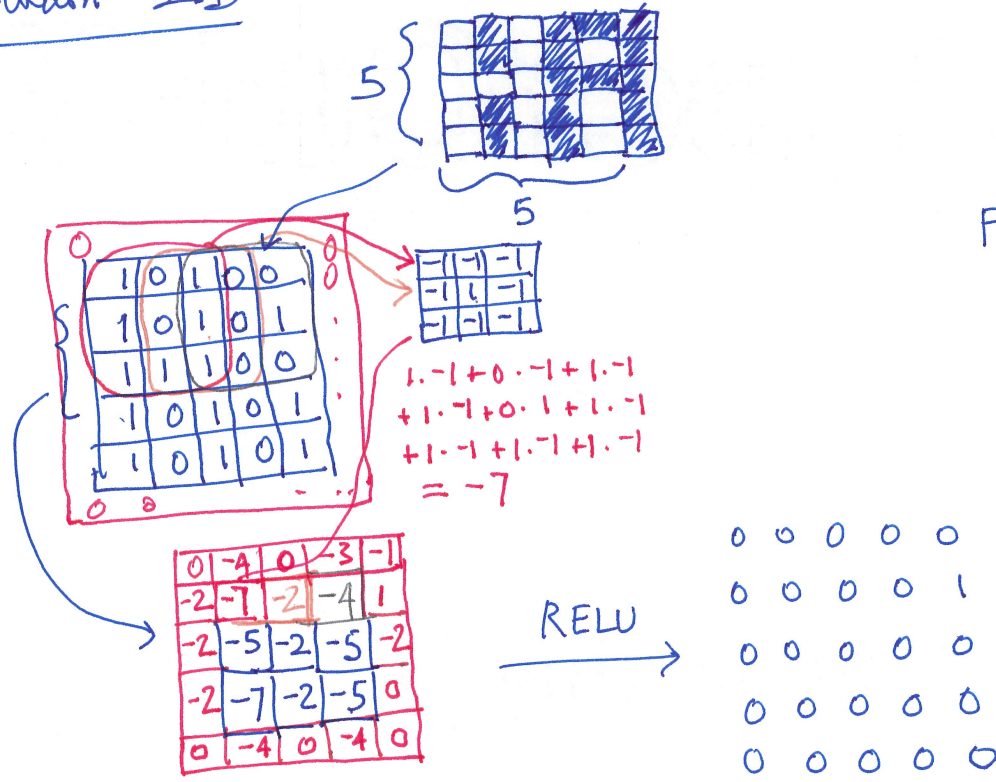
# parameters for convolution: 4

# parameter for fully connected:  $10 \times 11 = 110$

Convolution 2-D.

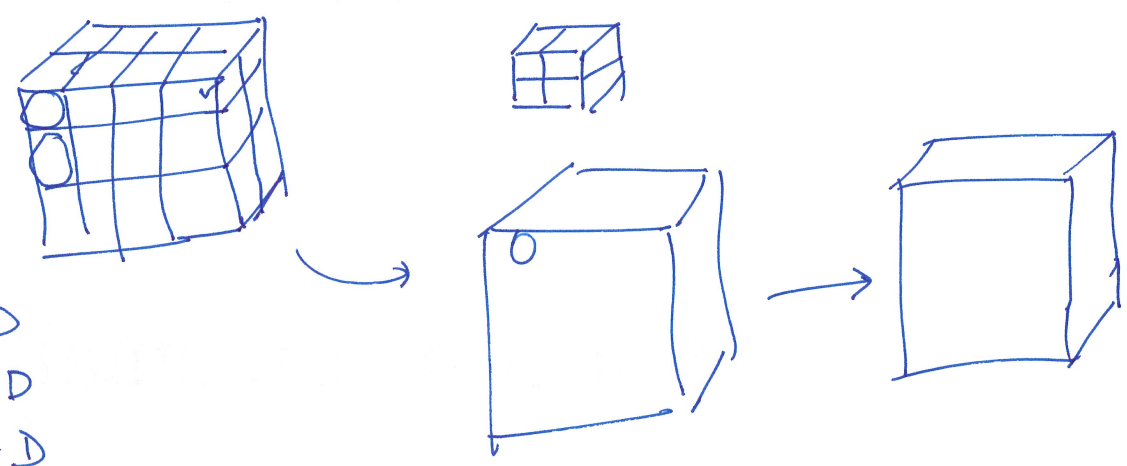


# Convolution 2-D

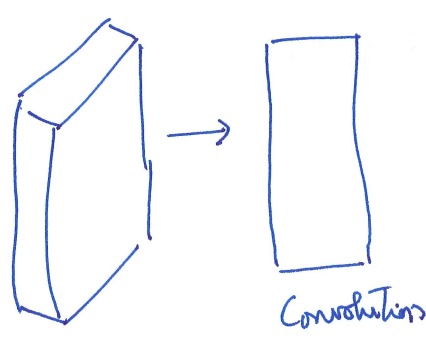


Filter :  $F \in \mathbb{R}^{f_1 \times f_2}$   
 Bias :  $b$

# Convolution 3-D

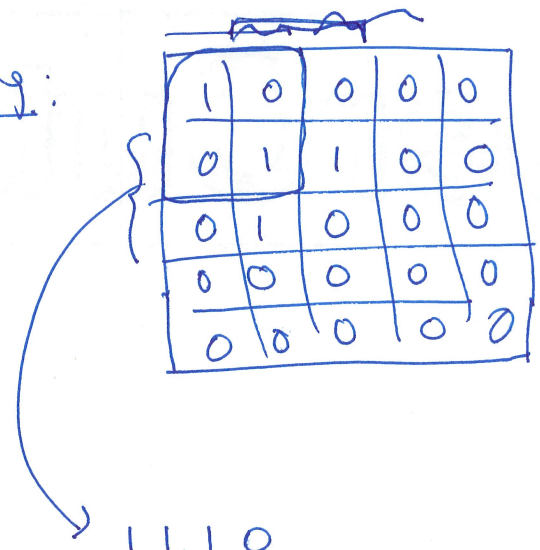


Vector : 1-D  
 Matrix : 2-D  
 Tensor : 3-D



Max Pooling:

2x2



max in sub-block of input of appropriate size.

1 1 1 0  
 1 1 1 0  
 1 1 0 0  
 0 0 0 0

Stride: # position you skip before evaluating pre-activation fn.



# parameters: 0

Stride: 1

Image → ~~Conv~~ Convolution → Max Pool → Convolution → Max Pool

