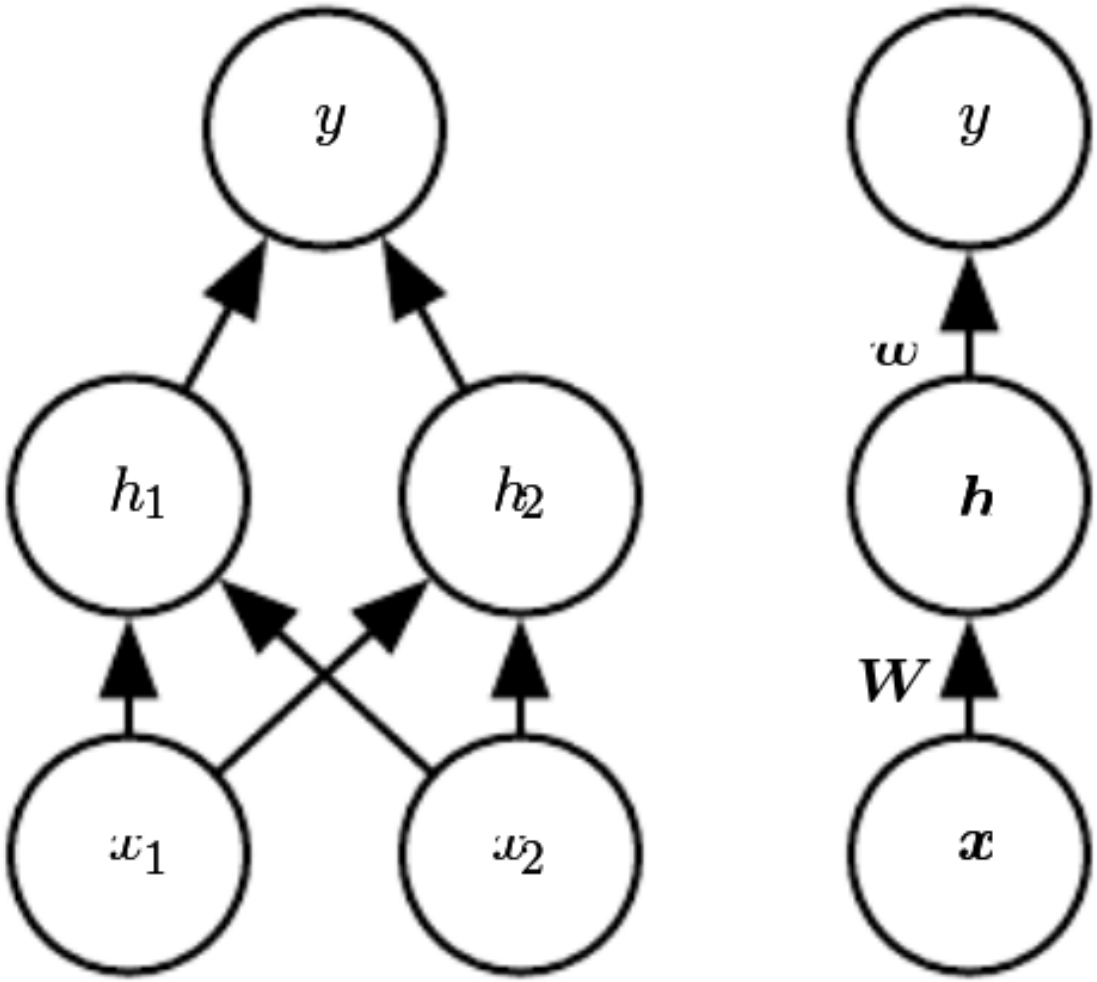
Quick introduction to CNNs

Goodfellow et al: Deep Learning 2016, Zimmermann: VIR lectures

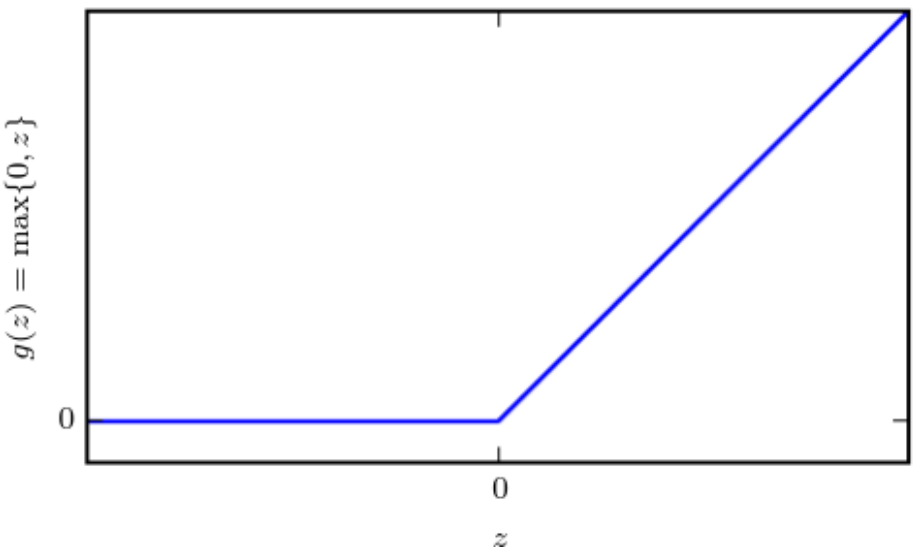
Jan Kybic

2020

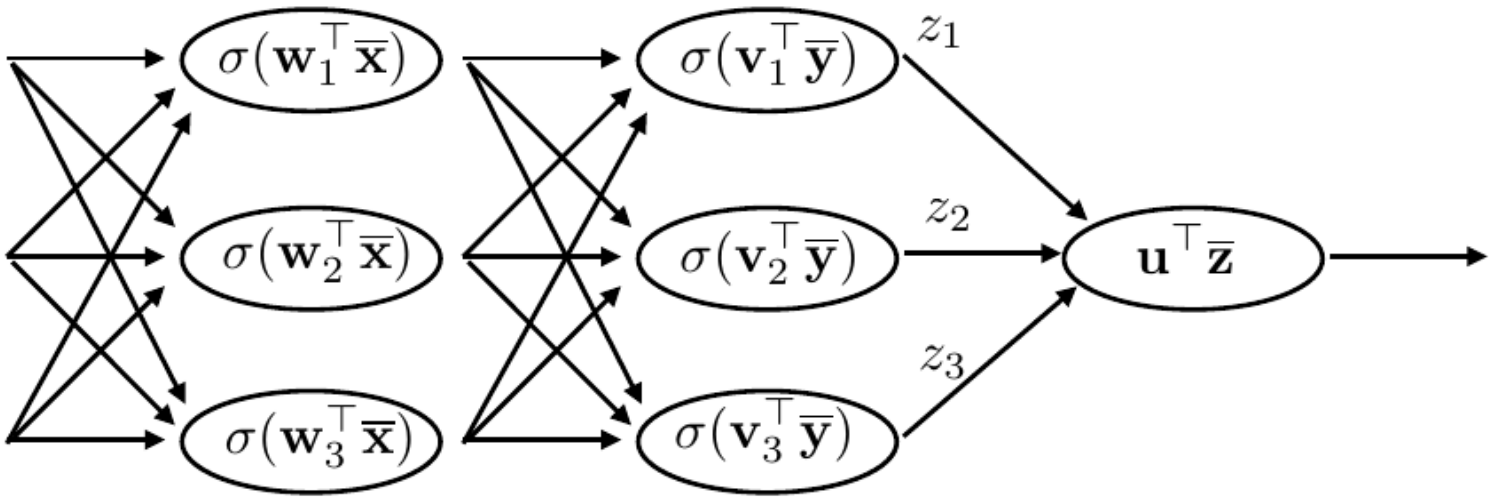
Neural network



ReLU Nonlinearity



Fully connected network

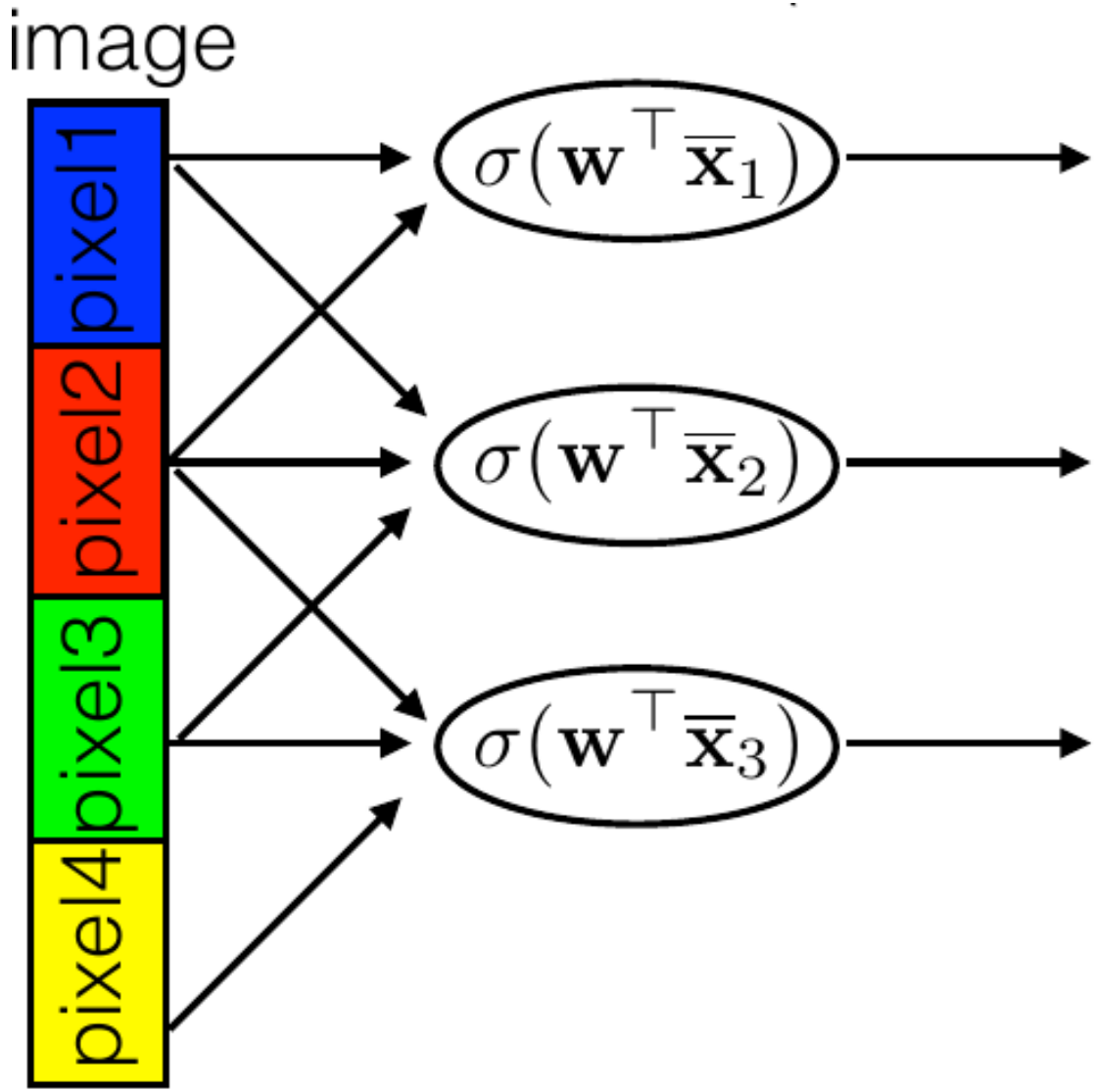


- loss function
- backpropagation

Softmax

$$\text{softmax}(\boldsymbol{z})_i = \frac{\exp(z_i)}{\sum_j \exp(z_j)}.$$

Convolutional NN



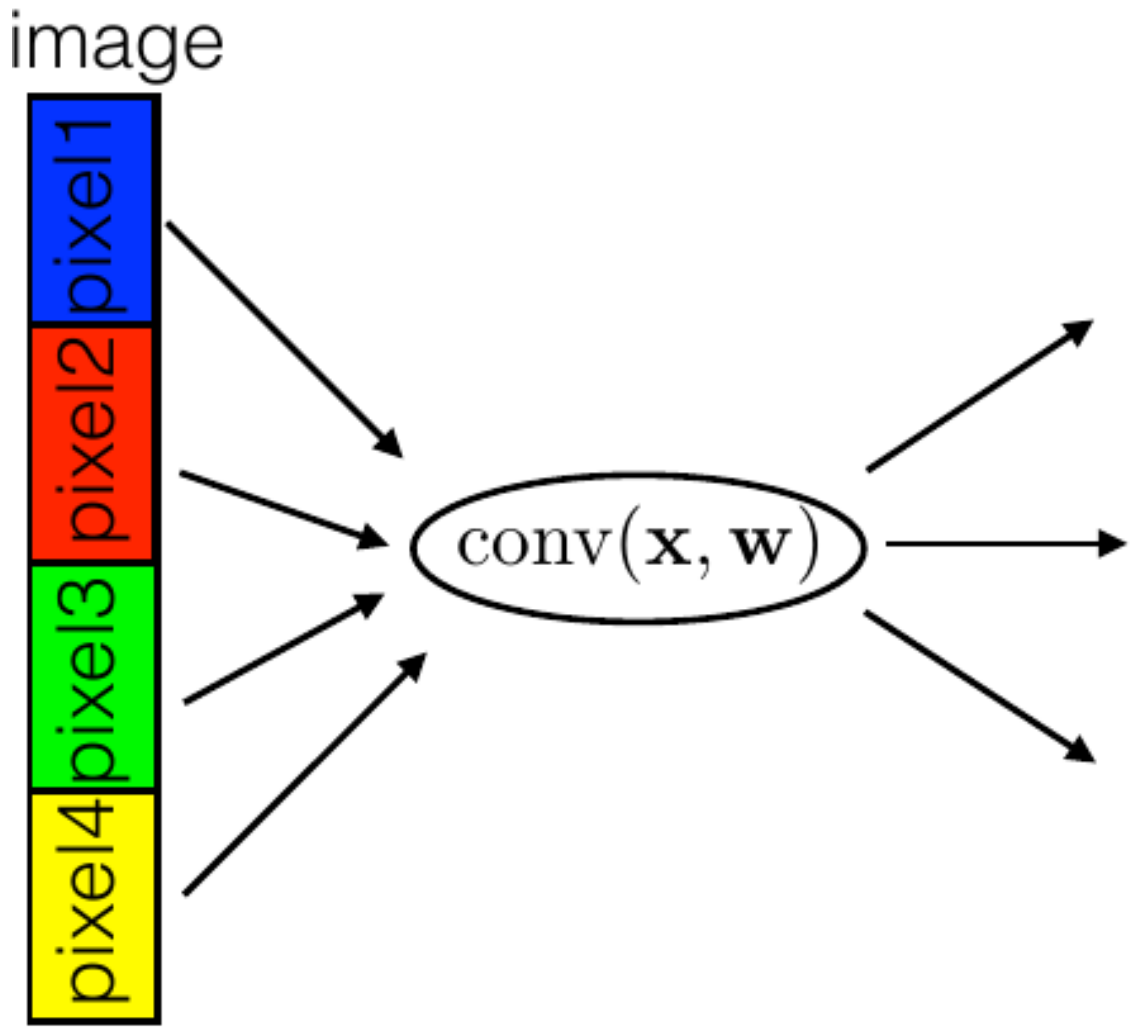
- translation invariance, parameter sharing, convolution

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(m,n) K(i-m, j-n).$$

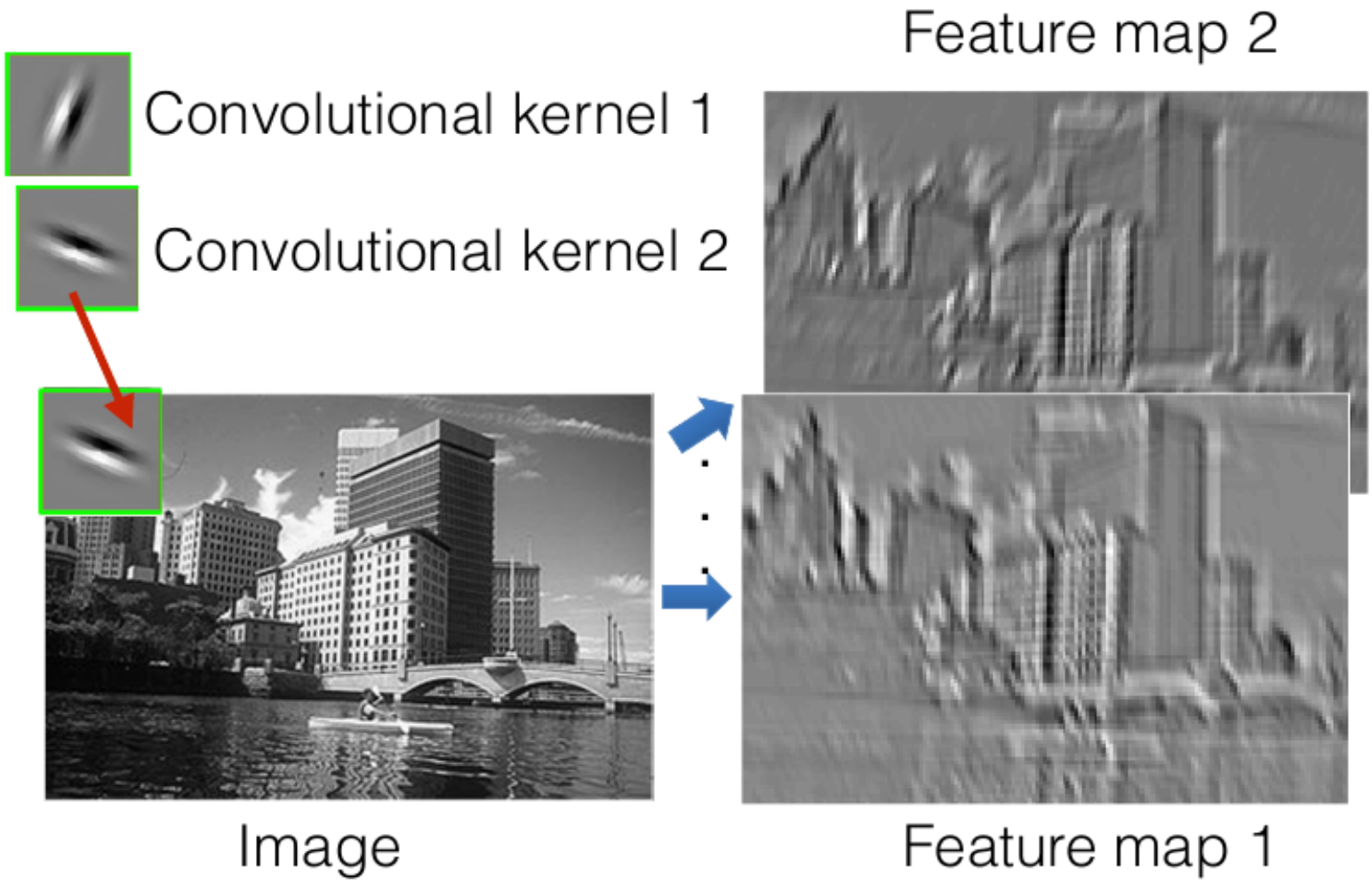
Convolution

$$\begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{bmatrix} = \text{conv}\left(\begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}, \begin{bmatrix} w_{11} & w_{12} \\ w_{21} & w_{22} \end{bmatrix} \right)$$

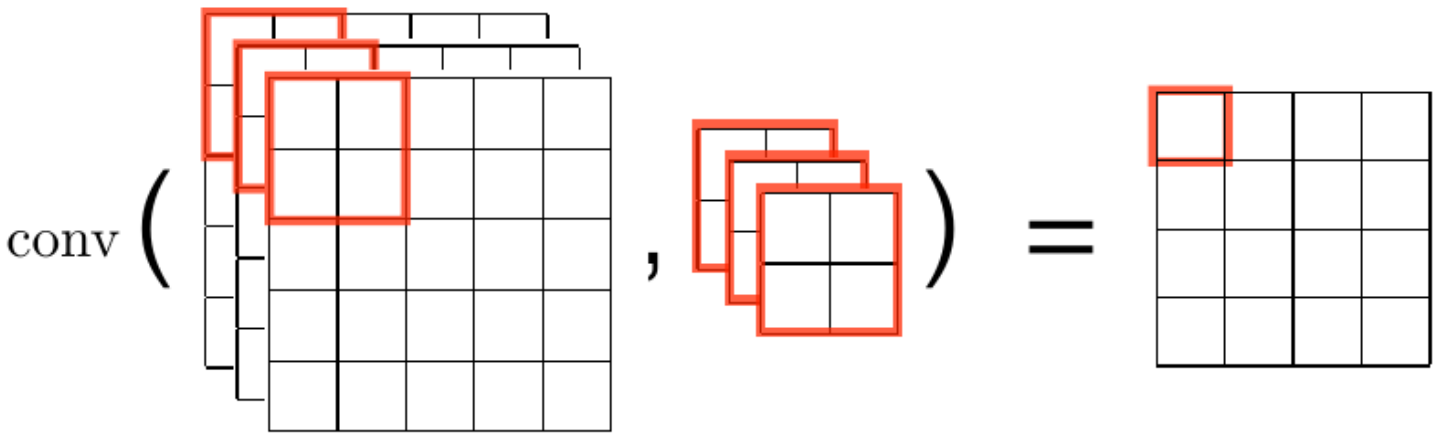
Convolutional NN (2)



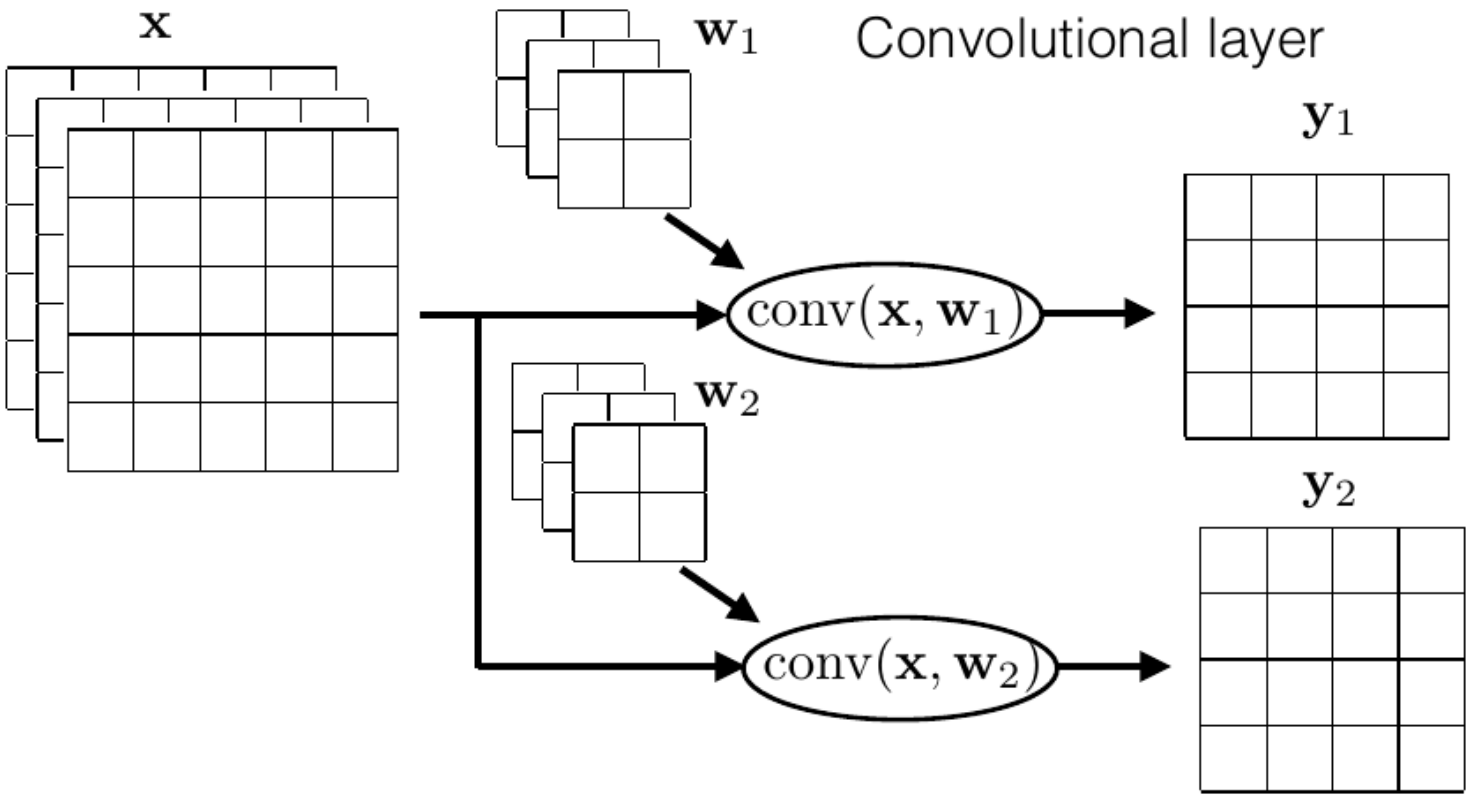
Feature map


Feature map 2
Convolutional kernel 1
Convolutional kernel 2
Image
Feature map 1

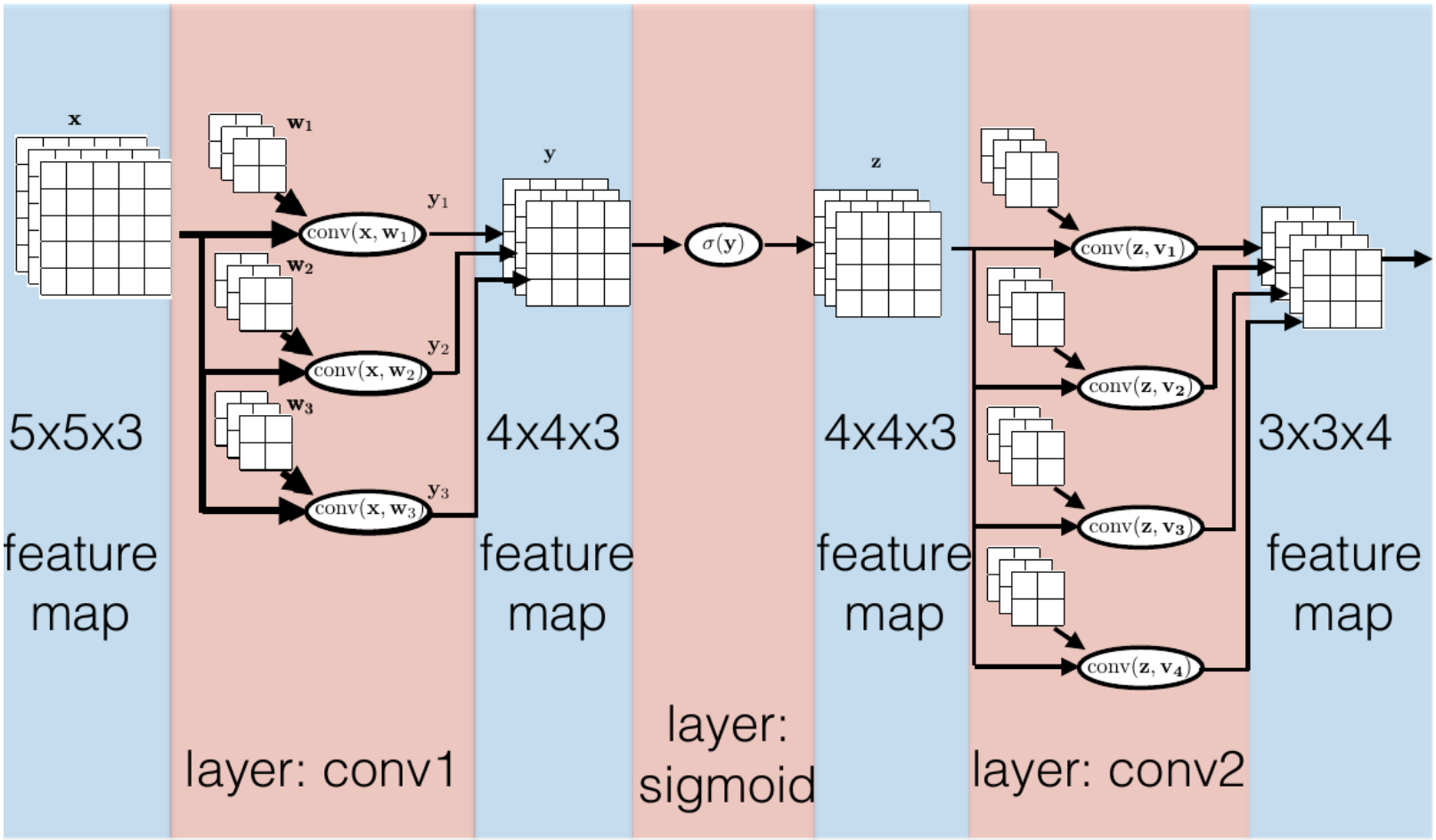
Multiple input channels



Multiple output channels

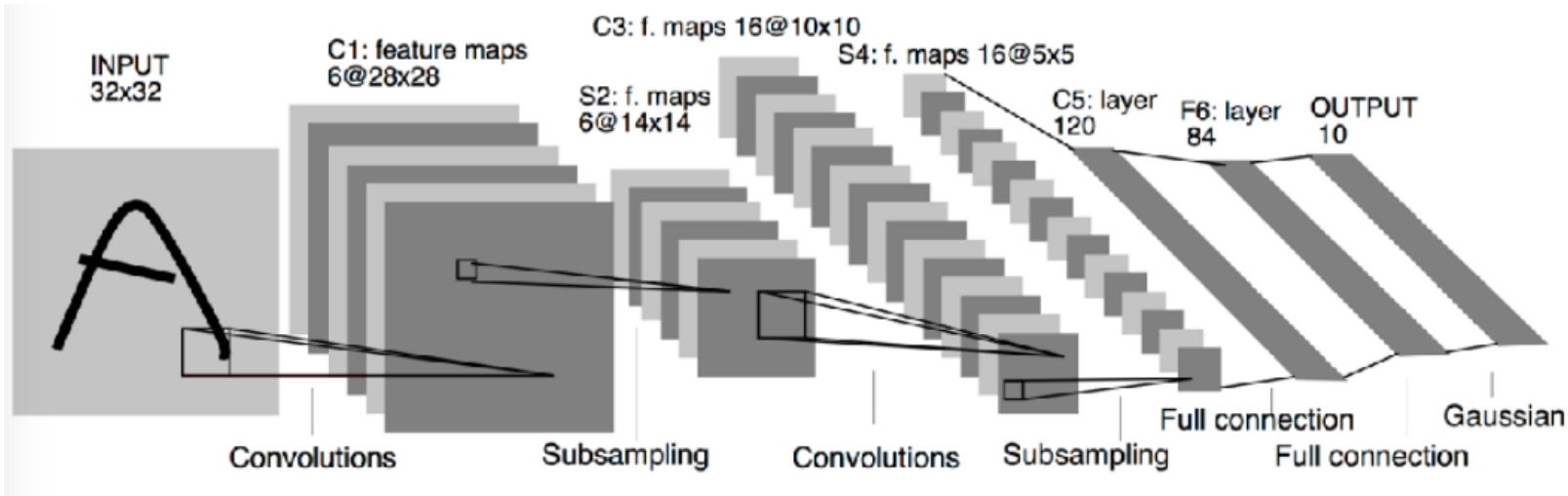


Multilayer CNN



Letter recognition CNN

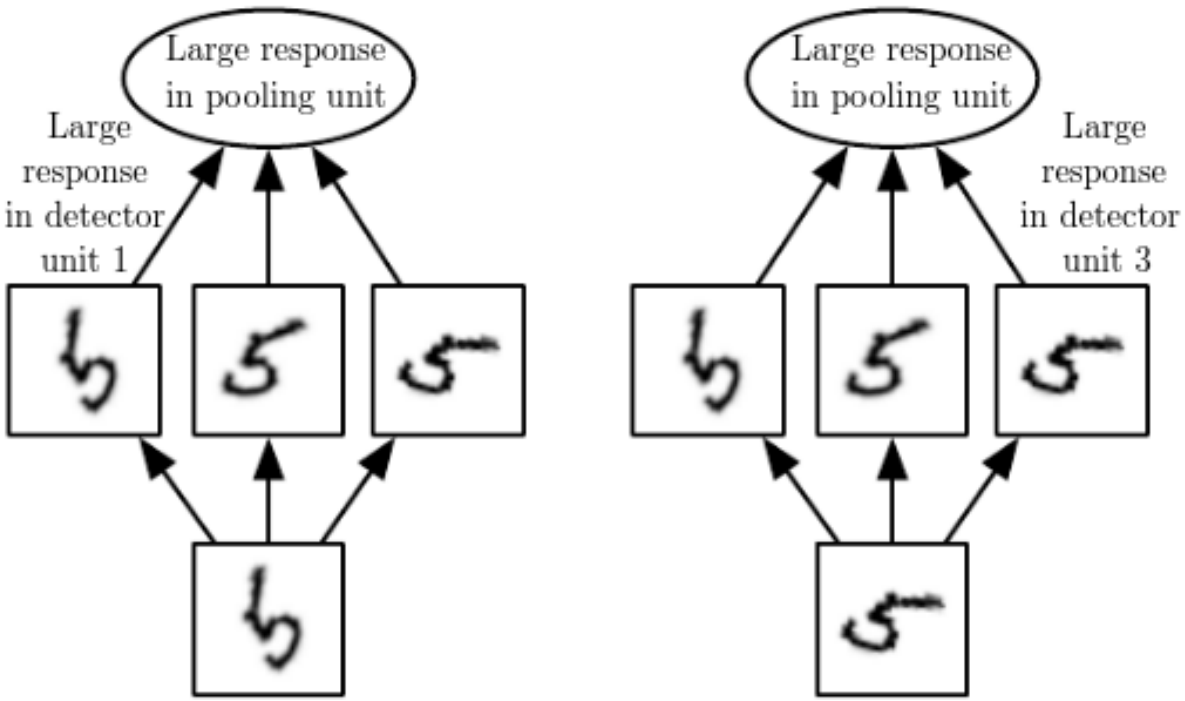
Le Cun et al, 1998



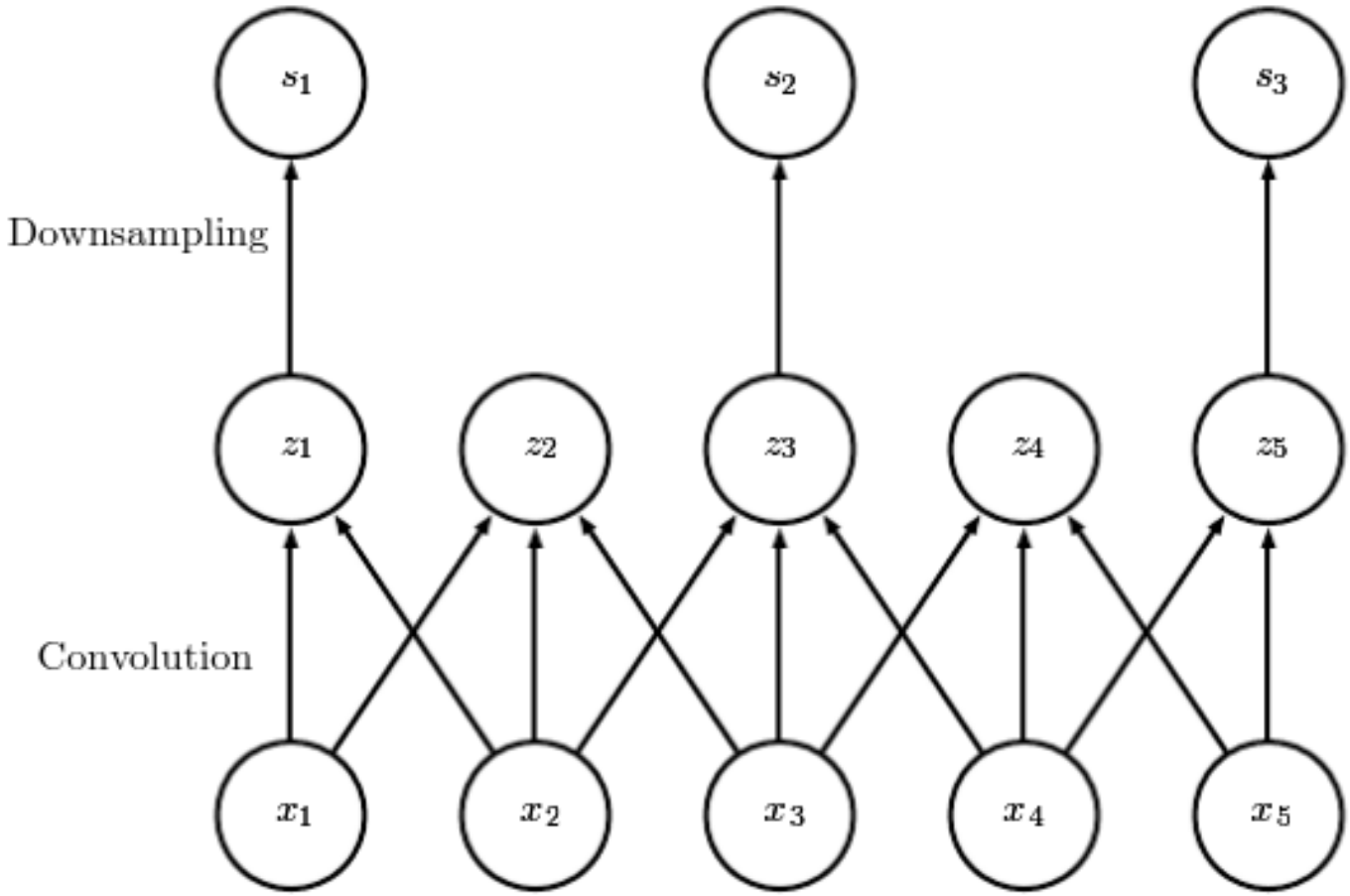
Subsampling/pooling

- reduce size (e.g. $2\times$)
- maximum over a region (max-pooling)
- (weighted) averaging
- increases receptive field
- approximate invariance to shifts (for detection)

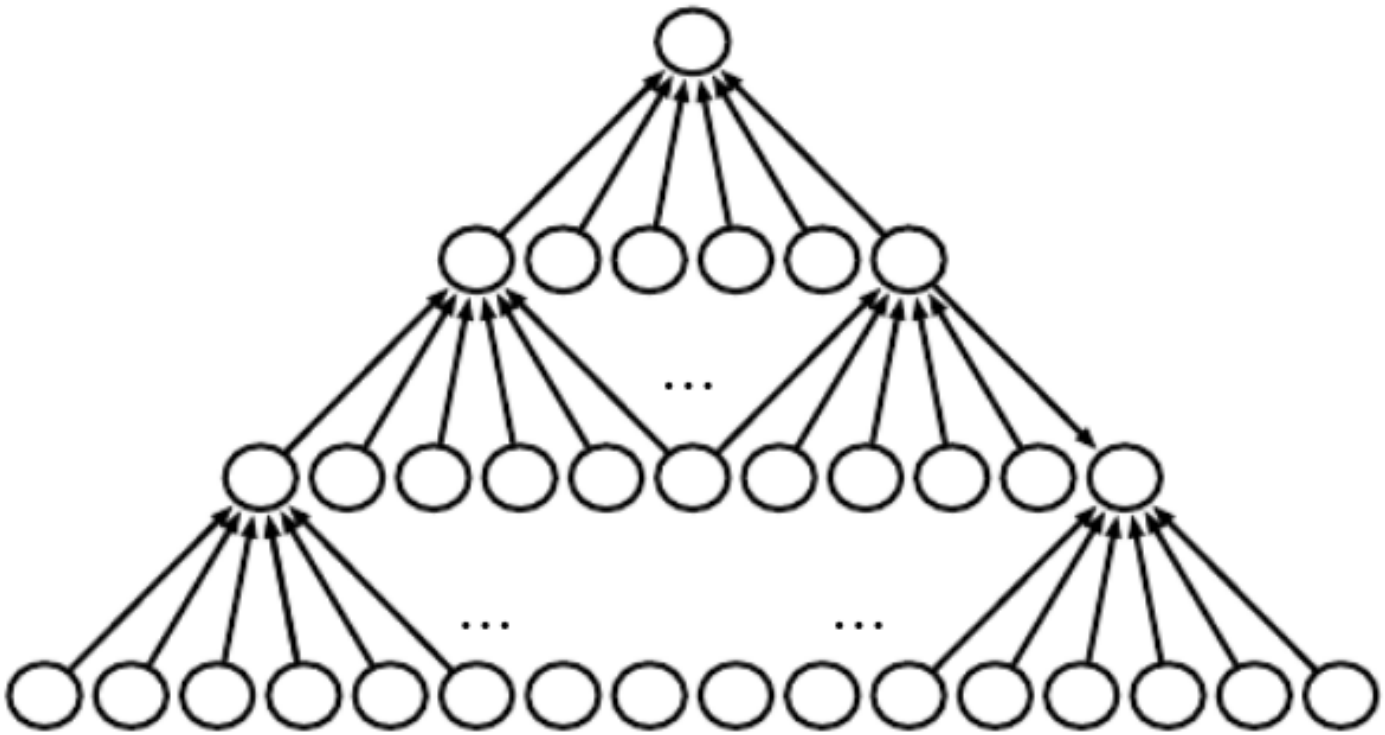
Learning invariances



Strided convolution



No padding



Zero padding