# Convolutional Neural Network

Fully connected layers: each hidden unit looks at an entire image

Convolutional layers: each column or set of hidden units look at a small region of the image

Convolution: $[2, -1, 1] * [1, 1, 2]$

$= 2 \times [1, 1, 2, 0, 0] + (-1) \times [0, 1, 1, 2, 0] + 1 \times [0, 0, 1, 1, 2] = [2, 1, 4, -1, 2]$

$a * b = b * a \;;\; a * (\lambda_1 b + \lambda_2 c) = \lambda_1 a * b + \lambda_2 a * c$

2D Convolution:

$$\begin{matrix} 1 & 3 & 1 \\ 0 & -1 & 1 \\ 2 & 2 & -1 \end{matrix} * \begin{matrix} 1 & 2 \\ 0 & -1 \end{matrix} :$$

kernel/filter

$$1 \times \begin{matrix} 1 & 3 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & 2 & -1 & 0 \\ 0 & 0 & 0 & 0 \end{matrix} + 2 \times \begin{matrix} 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 2 & 2 & -1 \\ 0 & 0 & 0 & 0 \end{matrix}$$

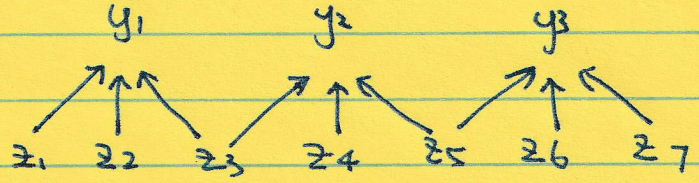
$$+ 0 \times \begin{matrix} 0 & 0 & 0 & 0 \\ 1 & 3 & 1 & 0 \\ 0 & -1 & 1 & 0 \\ 2 & 2 & -1 & 0 \end{matrix} + (-1) \times \begin{matrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & -1 & 1 \\ 0 & 2 & 2 & -1 \end{matrix}$$

$$= \begin{matrix} 1 & 5 & 7 & 2 \\ 0 & -2 & -4 & 1 \\ 2 & 6 & 4 & -3 \\ 0 & -2 & -2 & 1 \end{matrix}$$

the convolutional layer has a set of filters. Its output is a set of feature maps, each one obtained by convolving the image with a filter/kernel

Pooling layer: reduce the size of the representation and build invariance to small transformations.

Max pooling:




Backpropagation

Conv is defined as: $y_{i,t} = \sum_{j=1}^{J} \sum_{\tau=-R}^{R} w_{i,j,\tau} \cdot x_{j,t+\tau}$

in which there are J input feature maps, I output feature maps, and convolutional kernels have radius R

$$\bar{x}_{j,t} = \sum_{\tau} \bar{y}_{i,t-\tau} \cdot \frac{\partial y_{i,t-\tau}}{\partial x_{j,t}}$$

$$\bar{x}_{j,t} = \sum_{\tau} \bar{y}_{i,t-\tau} \cdot w_{i,j,\tau} \Rightarrow \bar{x}_j = \bar{y}_j * w_{i,j}$$