Convolutional Neural Networks (CNN)

ECE57000: Artificial Intelligence

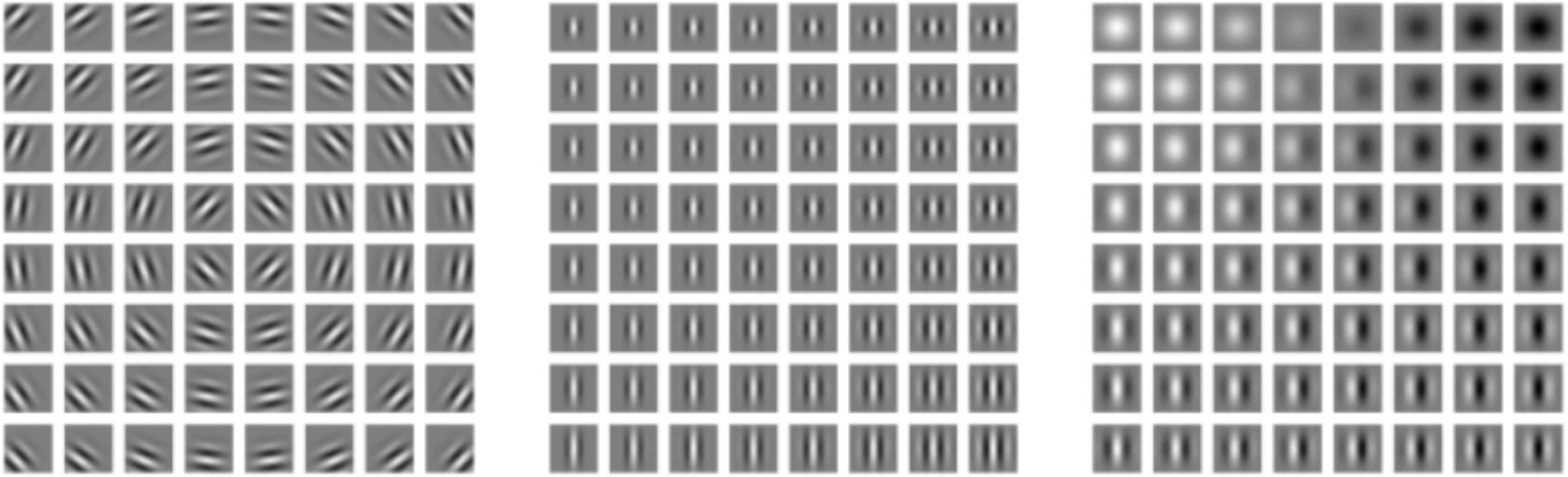
David I. Inouye

2020

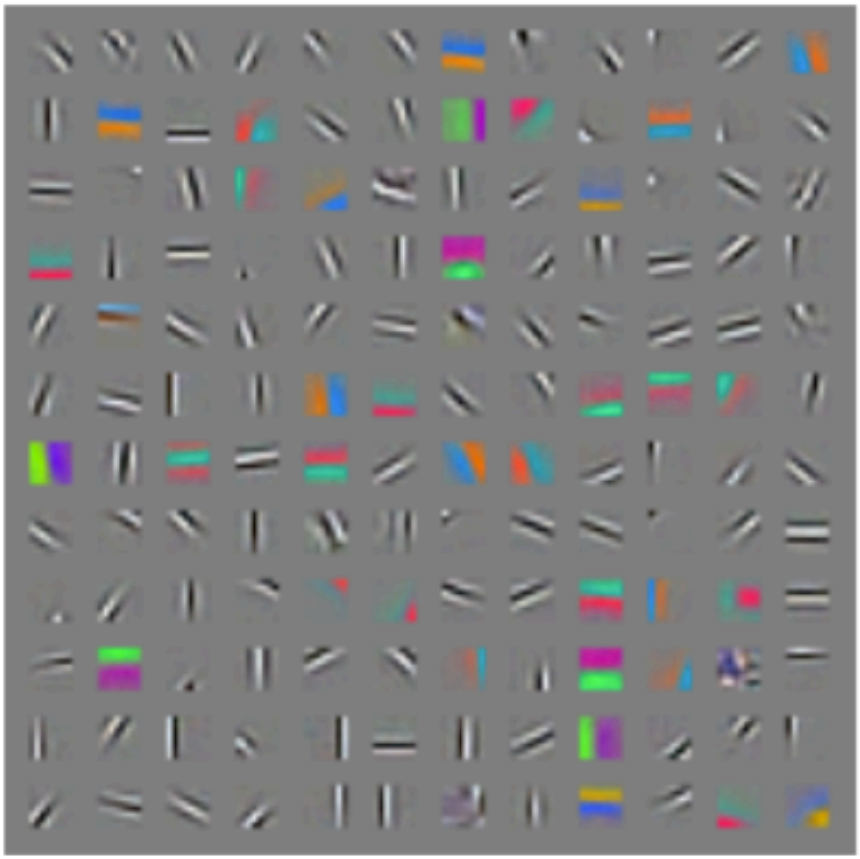
Why convolutional networks?

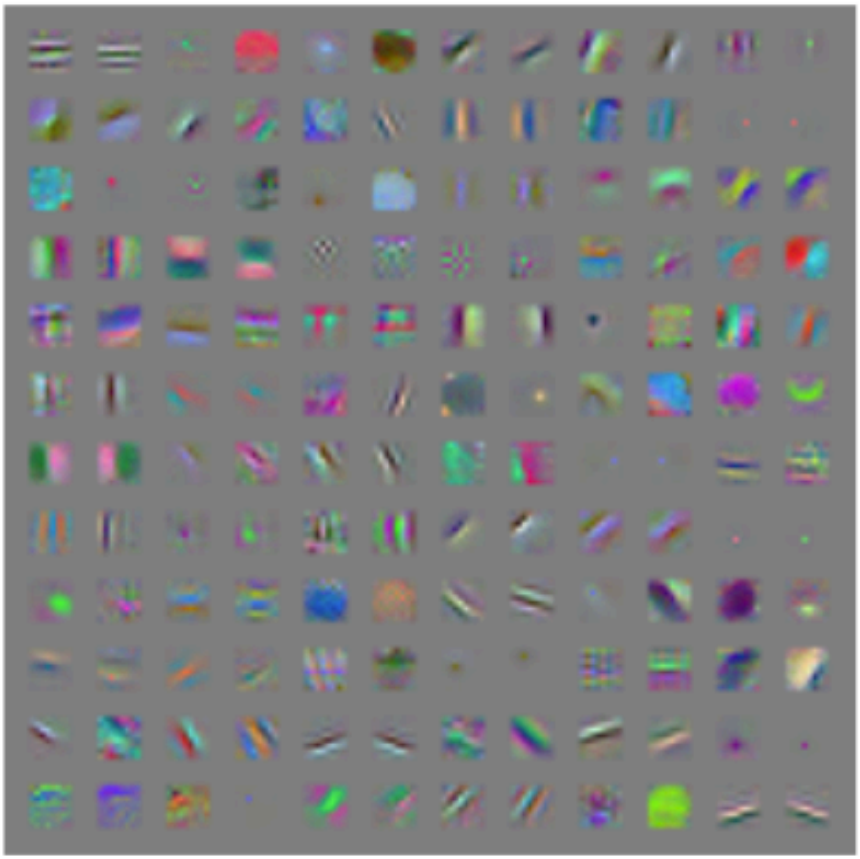
- Neuroscientific inspiration
- Computational reasons
  - Sparse computation (compared to full deep networks)
  - Shared parameters (only a small number of shared parameters)
  - Translation invariance

Motivation for convolution networks:
Gabor functions derived from neuroscience experiments are simple convolutional filters [DL, ch. 9]

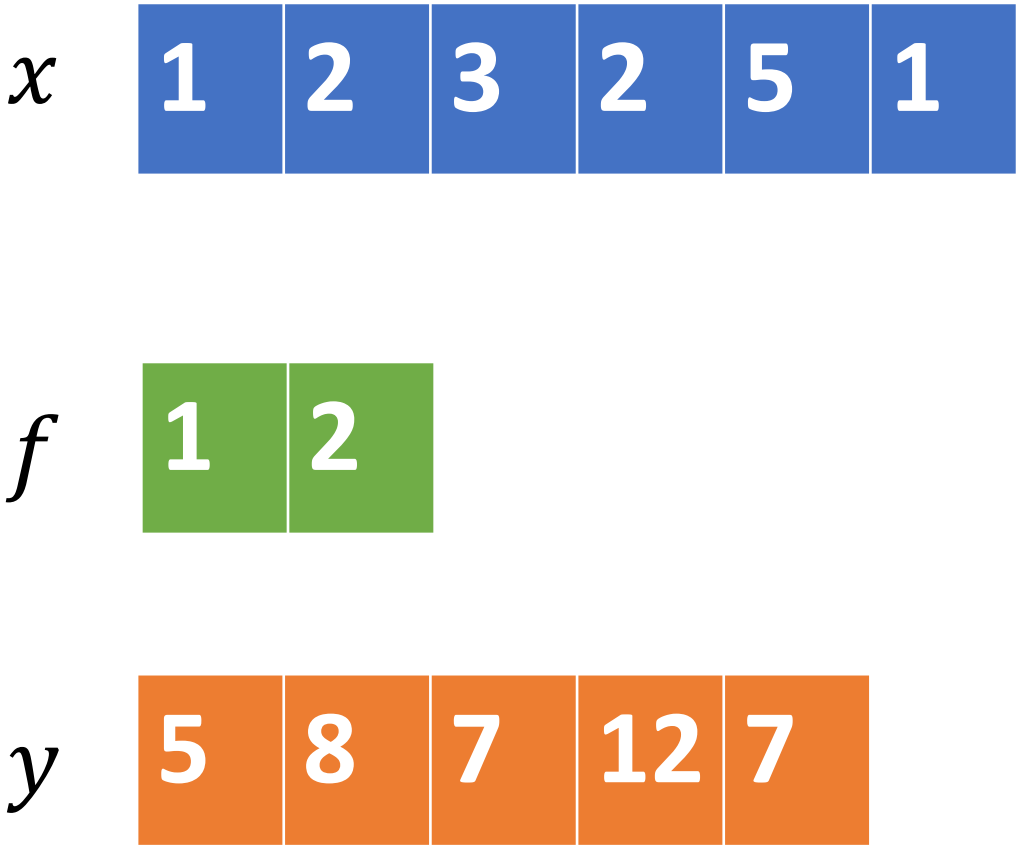


Convolutional networks automatically learn filters similar to Gabor functions [DL, ch. 9]

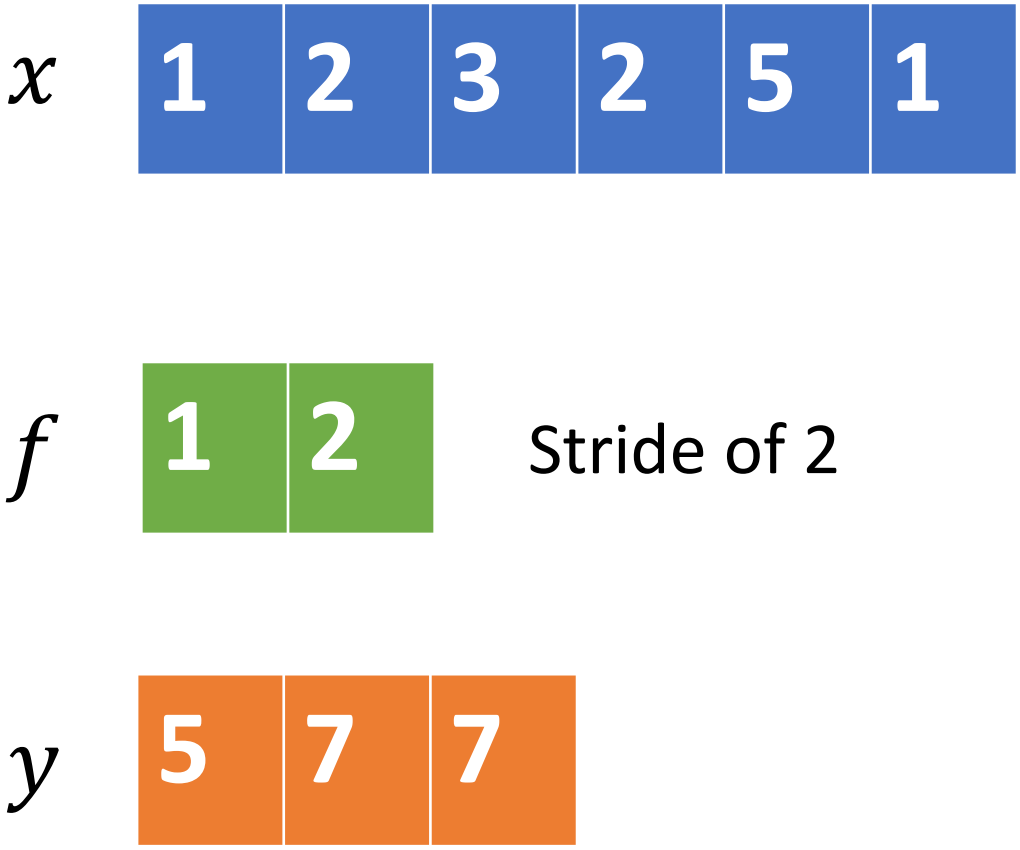




1D convolutions are similar but slightly different than signal processing / math convolutions



Padding or stride parameters alter the computation and output shape



1D convolutions are similar but slightly different than signal processing / math convolutions

x | 1 | 2 | 3 | 2 | 5 | 1 |

f | 1 | 2 |

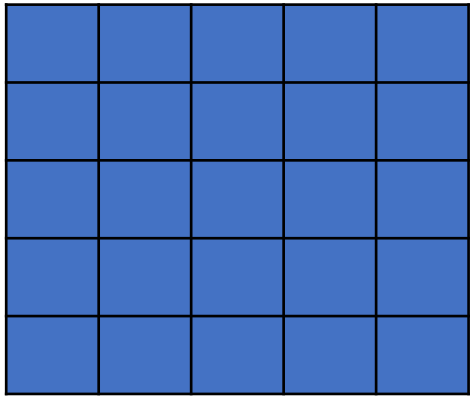
Zero padding of 1

y | 2 | 5 | 7 | 8 | 12 | 7 | 1 |

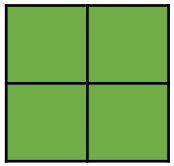
Switch to demo of 1D

2D convolutions are simple generalizations to matrices

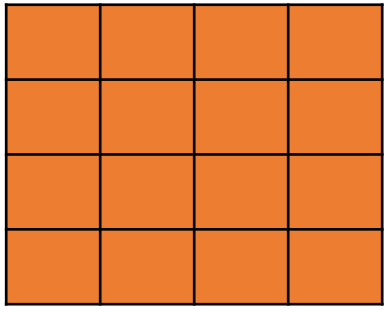
x



f



y



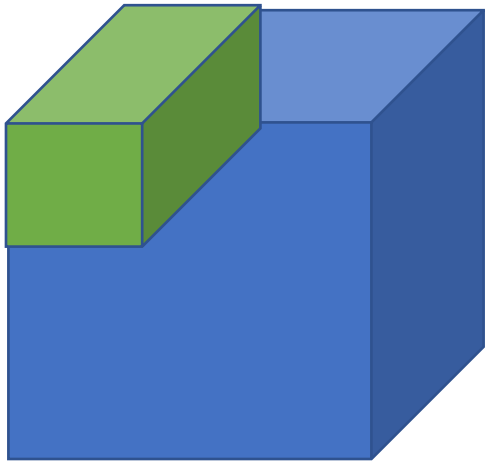
Stride of 2

y

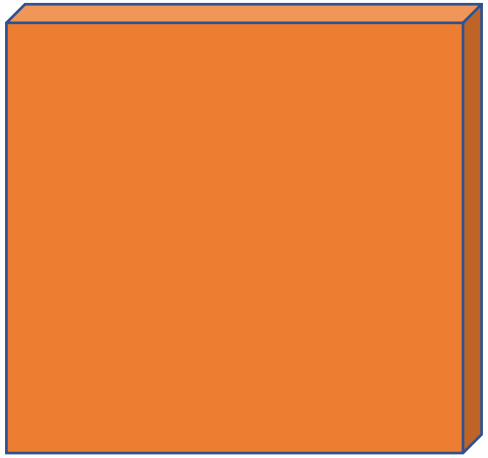
Switch to demo of 2D

3D convolutions are similar but usually channel dimension is assumed

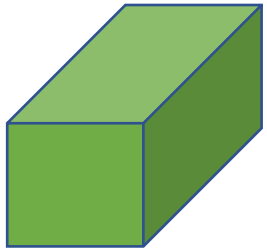
$$x \in \mathcal{R}^{c \times w \times h}$$



$$y \in \mathcal{R}^{1 \times w' \times h'}$$



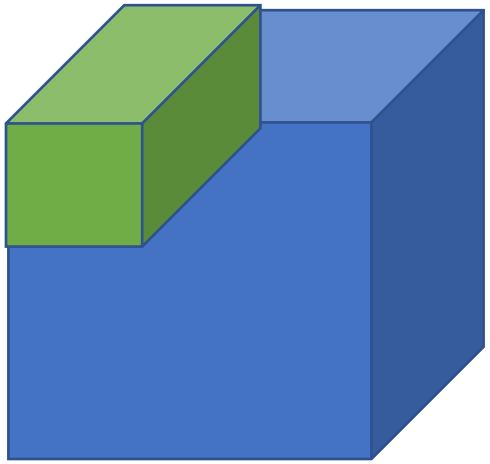
$$f \in \mathcal{R}^{c \times f_w \times f_h}$$



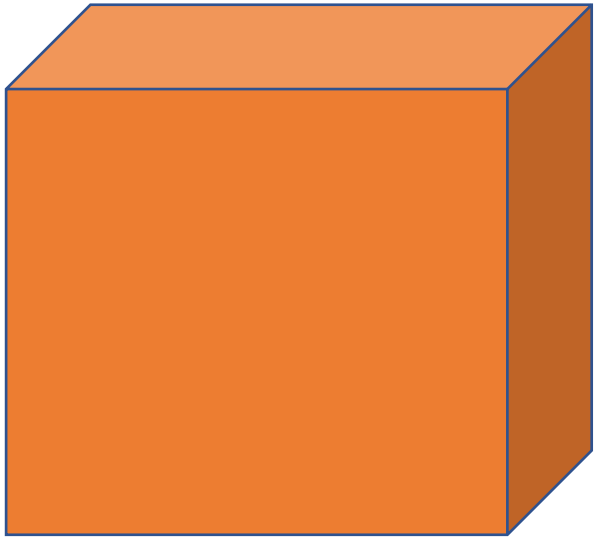
"$f_w \times f_h$ convolution" (channel dimension is assumed)

Multiple convolutions increase the output channel dimension

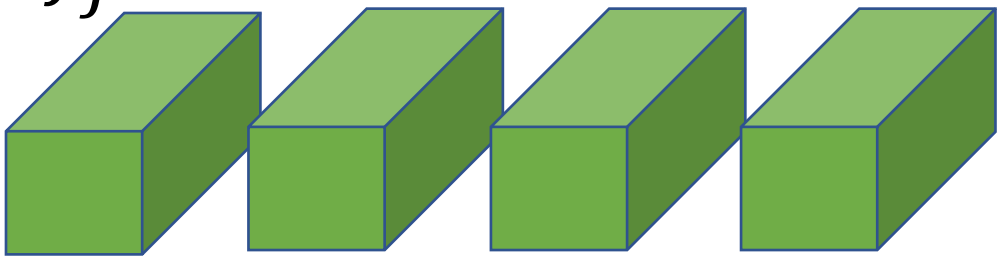
$$x \in \mathcal{R}^{c \times w \times h}$$



$$y \in \mathcal{R}^{4 \times w' \times h'}$$



$$f_j \in \mathcal{R}^{c \times f_w \times f_h}$$



Switch to demo of 3D

Standard Convolutional Layer Terminology [DL, ch. 9]