Image Classification: Images $\rightarrow$ classify

AlexNet (2012): Convolution Neural Network (CNN)

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

Fully connected Network

New features

Properties of Images:
- Spatial locality
- Translation invariance
Convolution in 1-D:

Input: 1 0 0 1 1 0 1 0 1 0
Filter: -1 -1 -1
Convolution: 1 -1 -1 0 1 0 -2 1 -1 0

ReLU: \[ \text{ReLU}(x) = \max(0, x) \]

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: \( f^T \times [i, i+1, \ldots, i+f] + b \).

Example: Input \( \mathbb{R}^{10} \), Output \( \mathbb{R}^{10} \), Filter size 3.
- # parameters for convolution: 4
- # parameters 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

Input of appropriate size.

Stride: Position you skip before evaluating pre-activation function.

Maxpool → Convolution

Image → Convolution → Max Pool → Convolution → Max Pool

Convolution → MaxPool

Logistic/Softmax