The logo of The University of Texas at Dallas is a circular seal. It features a large, stylized 'UTD' in the center. The words 'THE UNIVERSITY OF TEXAS AT DALLAS' are written around the top inner edge of the circle, and 'EST. 1969' is at the bottom. Two small stars are positioned on either side of the 'EST. 1969' text.

Convolutional Neural Networks I: Activation Function and FC Layer

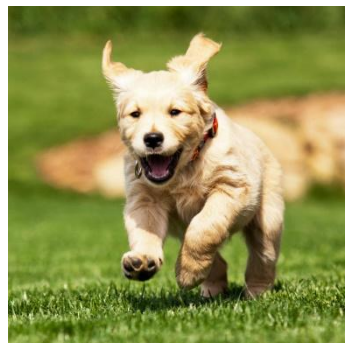
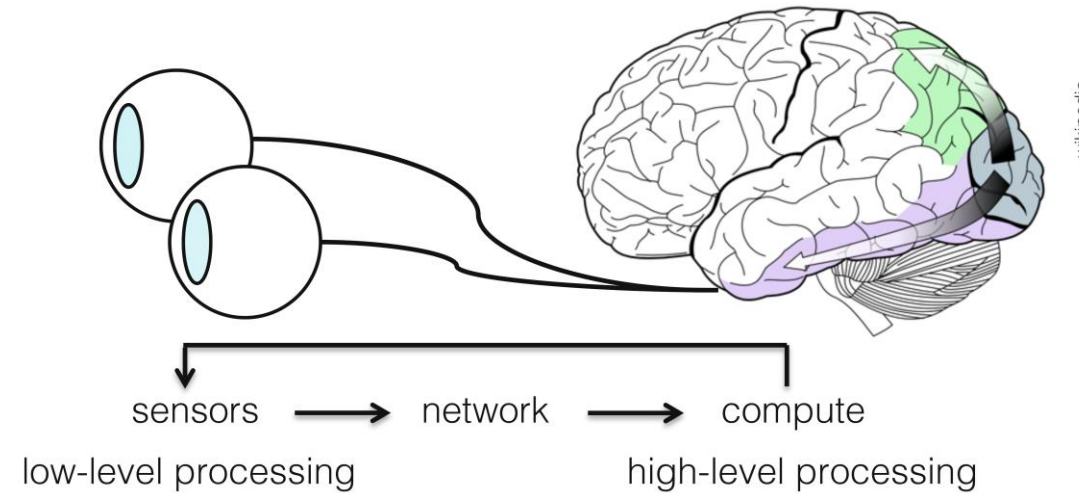
CS 4391 Introduction Computer Vision

Instructor Yu Xiang

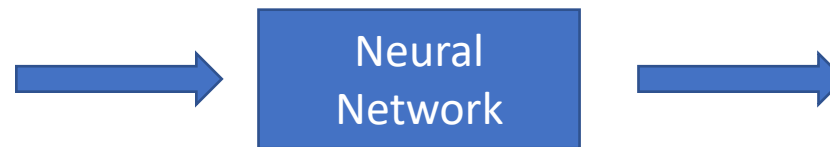
The University of Texas at Dallas

Some slides of this lecture are courtesy Stanford CS231n

Visual Perception vs. Computational Perception



Image

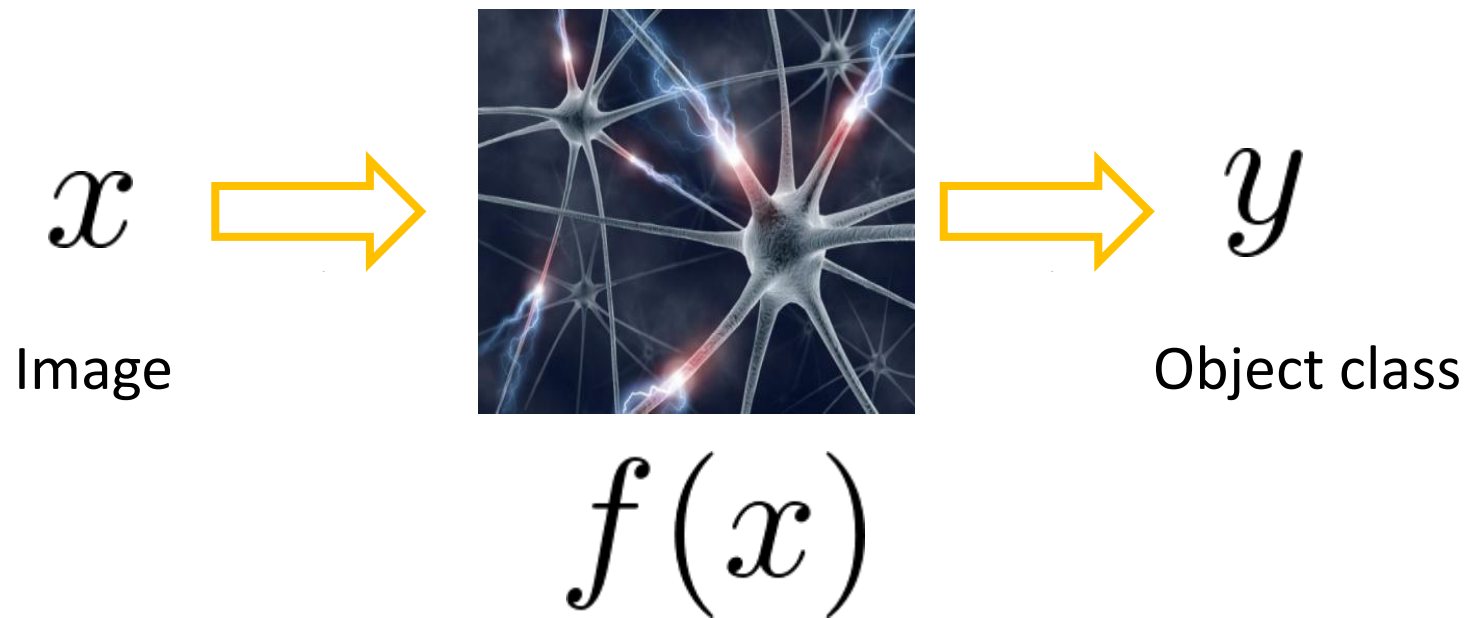


High-level information

- Depth
- Motion
- Object classes
- Object poses
- Etc.

Mathematic Models

- Try to model the human brain with computational models, e.g., neural networks

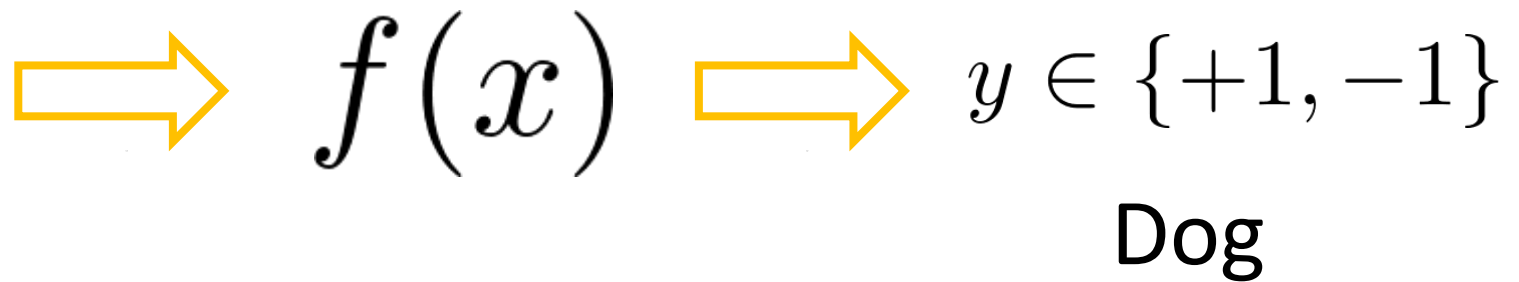


Mathematic Models

- What is the form of the function $f(x)$?
 - No idea!
 - Concatenate simple functions (neurons)



x



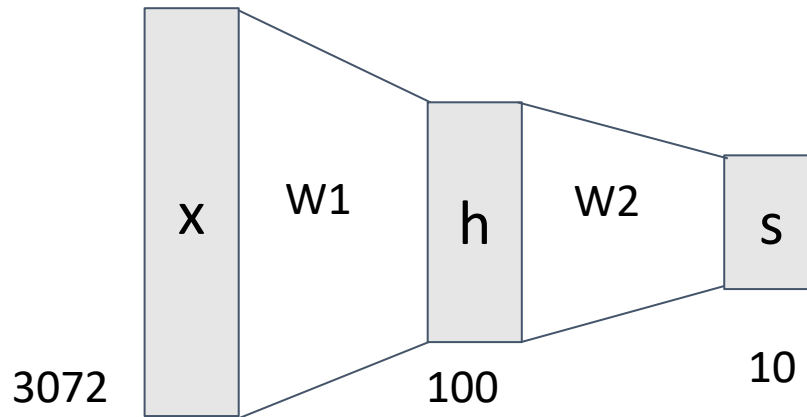
Neural Network: Concatenation of functions

Linear score function: $f = Wx$

2-layer Neural Network

$$f = f_2(f_1(x)) = W_2 \max(0, W_1 x)$$

Non-linearity

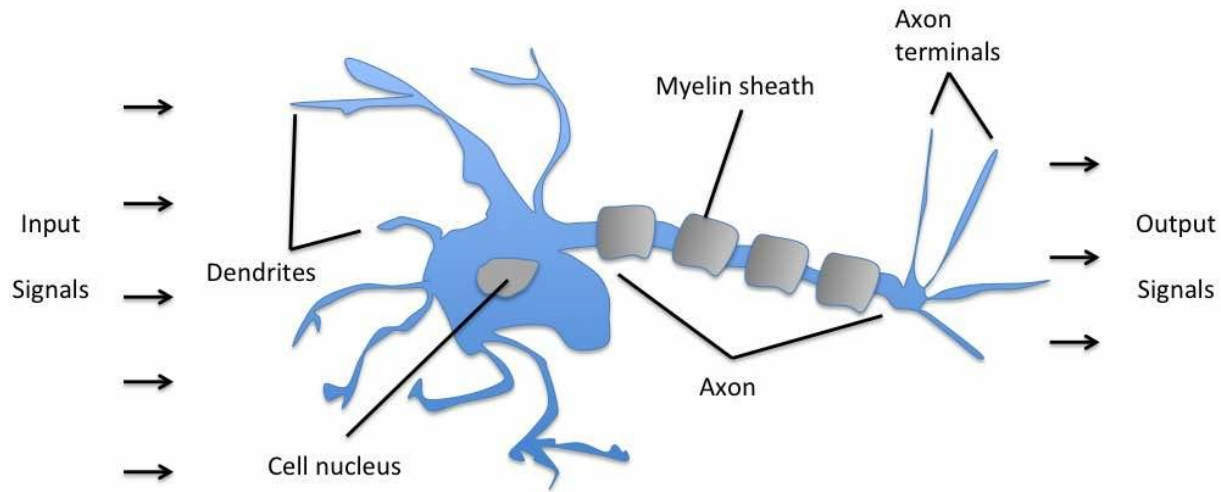


$$h = f_1(X)$$

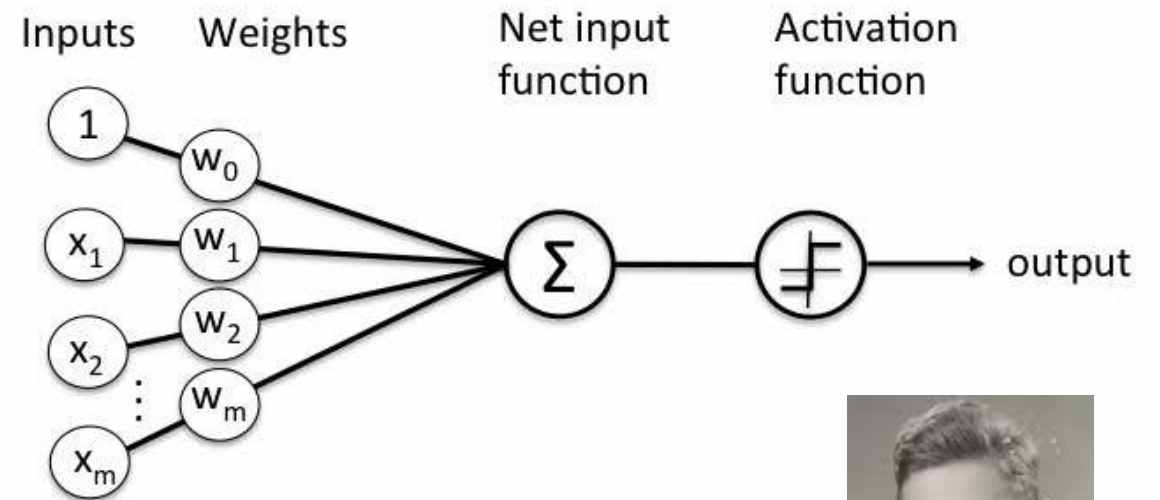
$$s = f_2(h)$$

Need to learn the weights!

Frank Rosenblatt's Perceptron



Schematic of a biological neuron.



$$\sigma(\mathbf{w}^T \mathbf{x} + b) = \begin{cases} 1 & \text{if } \mathbf{w}^T \mathbf{x} + b \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$



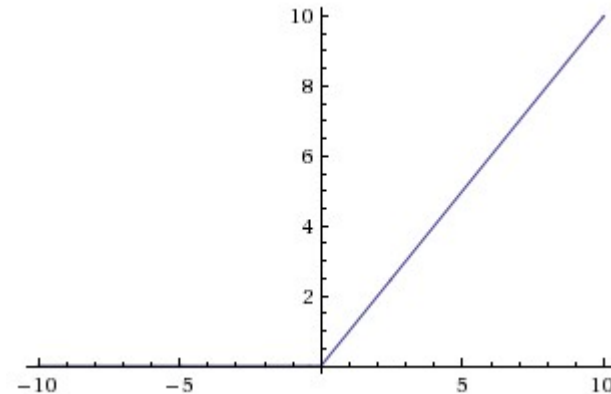
Frank Rosenblatt
(1928-1971)

Activation Functions

2-layer Neural Network

$$f = f_2(f_1(x)) = W_2 \max(0, W_1 x)$$

Rectified Linear Unit (ReLU)
 $\max(0, x)$

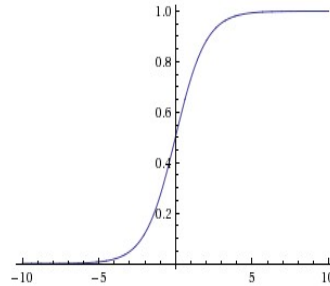


Introduce non-linearity to the network

Activation Functions

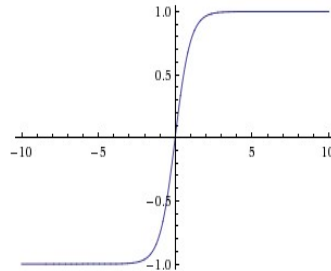
Sigmoid

$$\sigma(x) = 1/(1 + e^{-x})$$

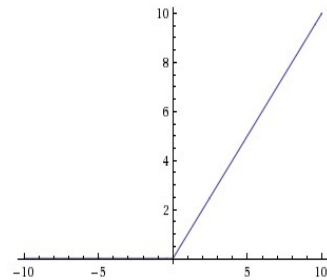


tanh $\tanh(x)$

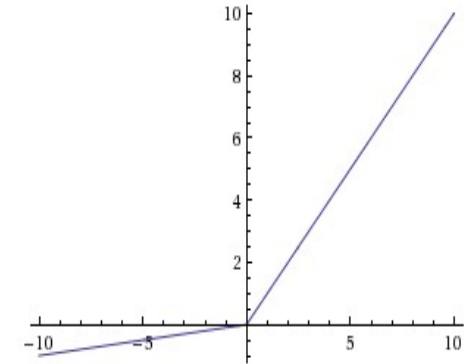
$$\frac{e^{2x} - 1}{e^{2x} + 1}$$



ReLU $\max(0, x)$

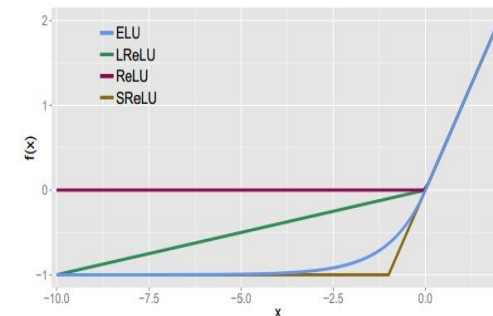


Leaky ReLU
 $\max(0.1x, x)$



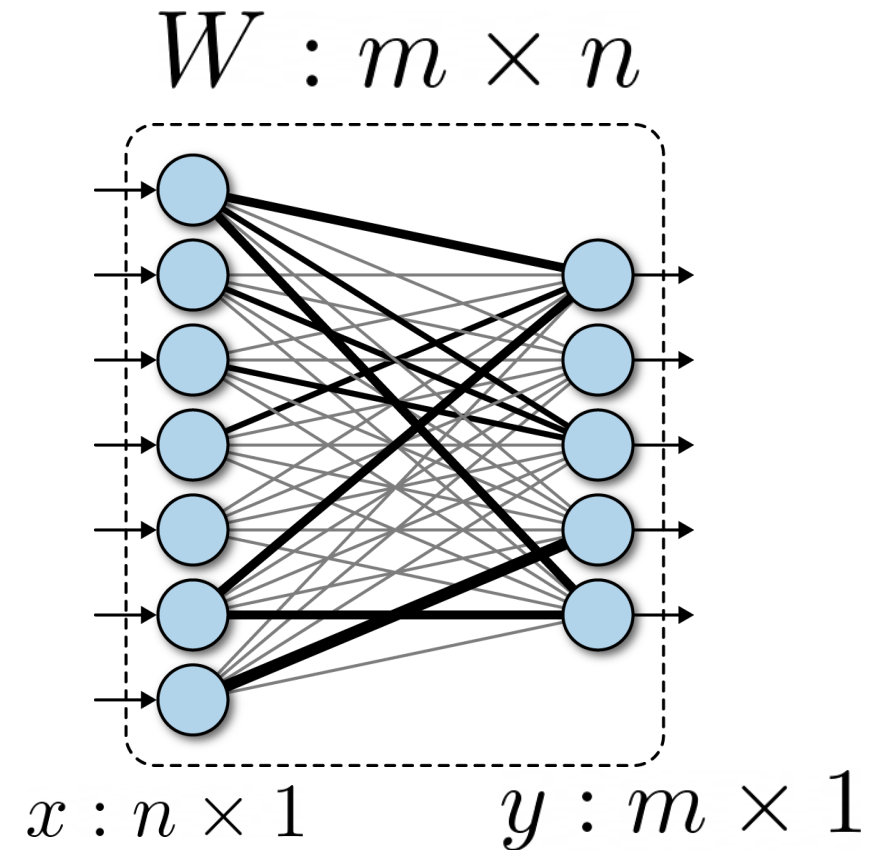
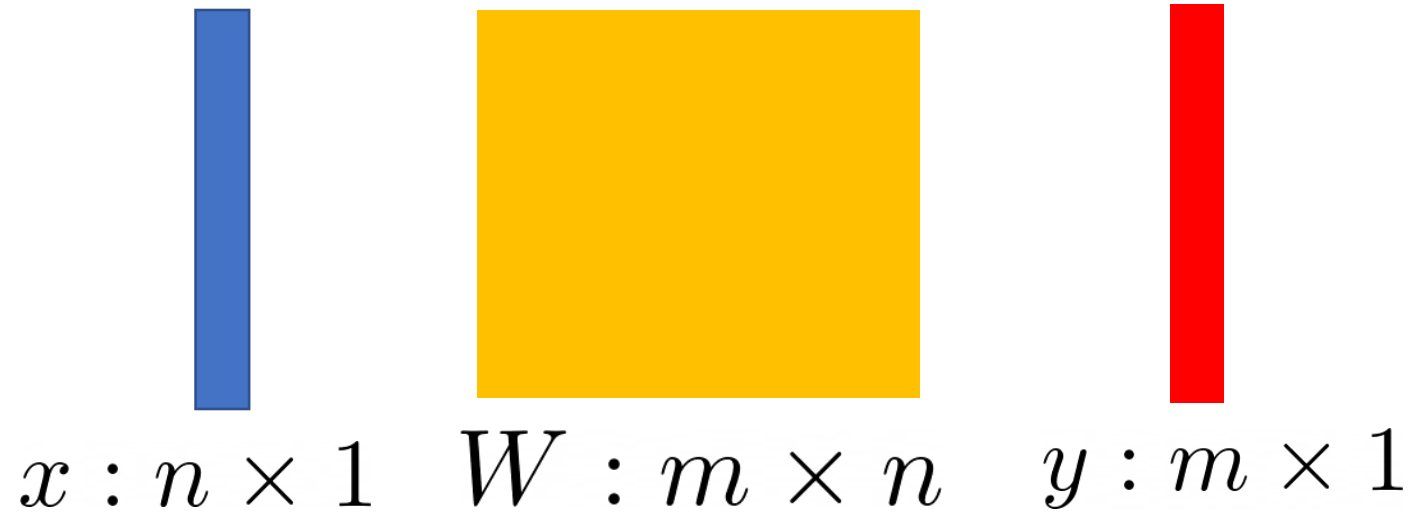
Maxout $\max(w_1^T x + b_1, w_2^T x + b_2)$

ELU Exponential Linear Unit
$$f(x) = \begin{cases} x & \text{if } x > 0 \\ \alpha (\exp(x) - 1) & \text{if } x \leq 0 \end{cases}$$



Fully Connected Layer

$$y = Wx$$



Fully Connected Layer

- What is the drawback of only using fully connected layers?

$$y = Wx$$

- Consider an image with 640 x 480
 - x is with dimension 307,200
 - The weight matrix of the fully connect layer is too large

Further Reading

- Stanford CS231n, lecture 5, Convolutional Neural Networks
<http://cs231n.stanford.edu/schedule.html>
- Deep learning with PyTorch
https://pytorch.org/tutorials/beginner/deep_learning_60min_blitz.html
- AlexNet (2012):
<https://papers.nips.cc/paper/2012/hash/c399862d3b9d6b76c8436e924a68c45b-Abstract.html>
- Vgg16 (2014): <https://arxiv.org/abs/1409.1556>
- GoogleNet (2014): <https://arxiv.org/abs/1409.4842>
- ResNet (2015): <https://arxiv.org/abs/1512.03385>