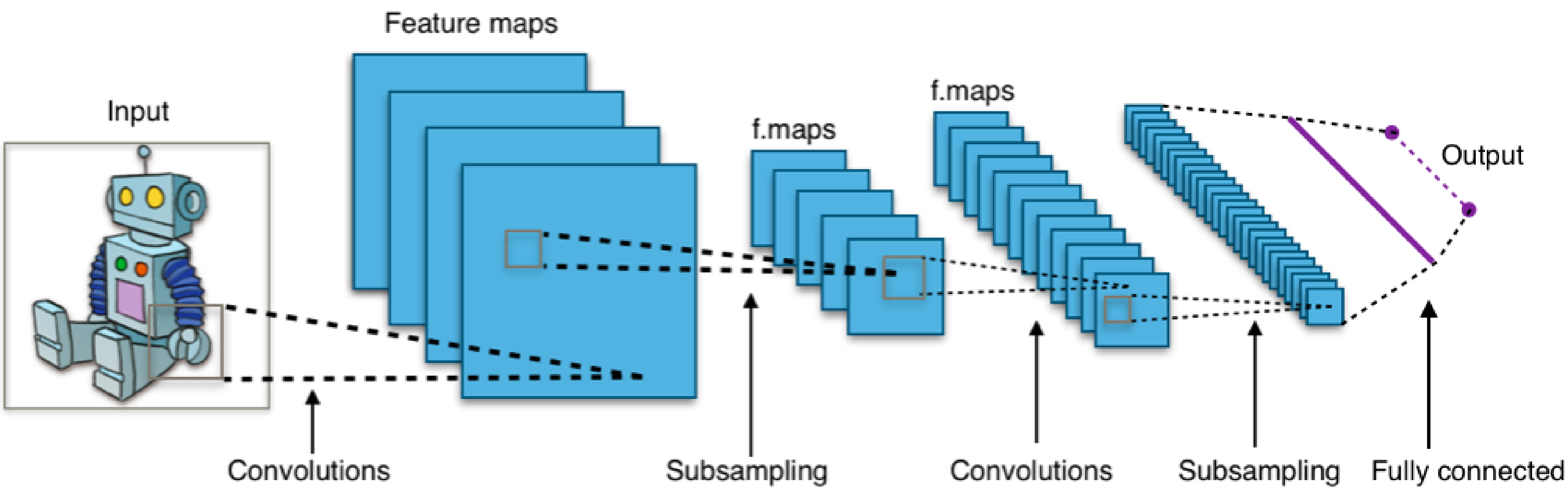
CS5670: Computer Vision

Noah Snavely

Lecture 24: Convolutional neural networks



Slides from Andrej Karpathy, Fei-Fei Li, Kavita Bala, and Sean Bell

Today

- Deep learning
- Field is in rapid motion
- Readings: No standard textbooks yet!
- Some good resources:
  - https://sites.google.com/site/deeplearningsummerschool/
  - http://www.deeplearningbook.org/
  - http://www.cs.toronto.edu/~hinton/absps/NatureDeepReview.pdf

Announcements

- Final project (P5), due Tuesday, 5/9, by 11:59pm, to be done in groups of two

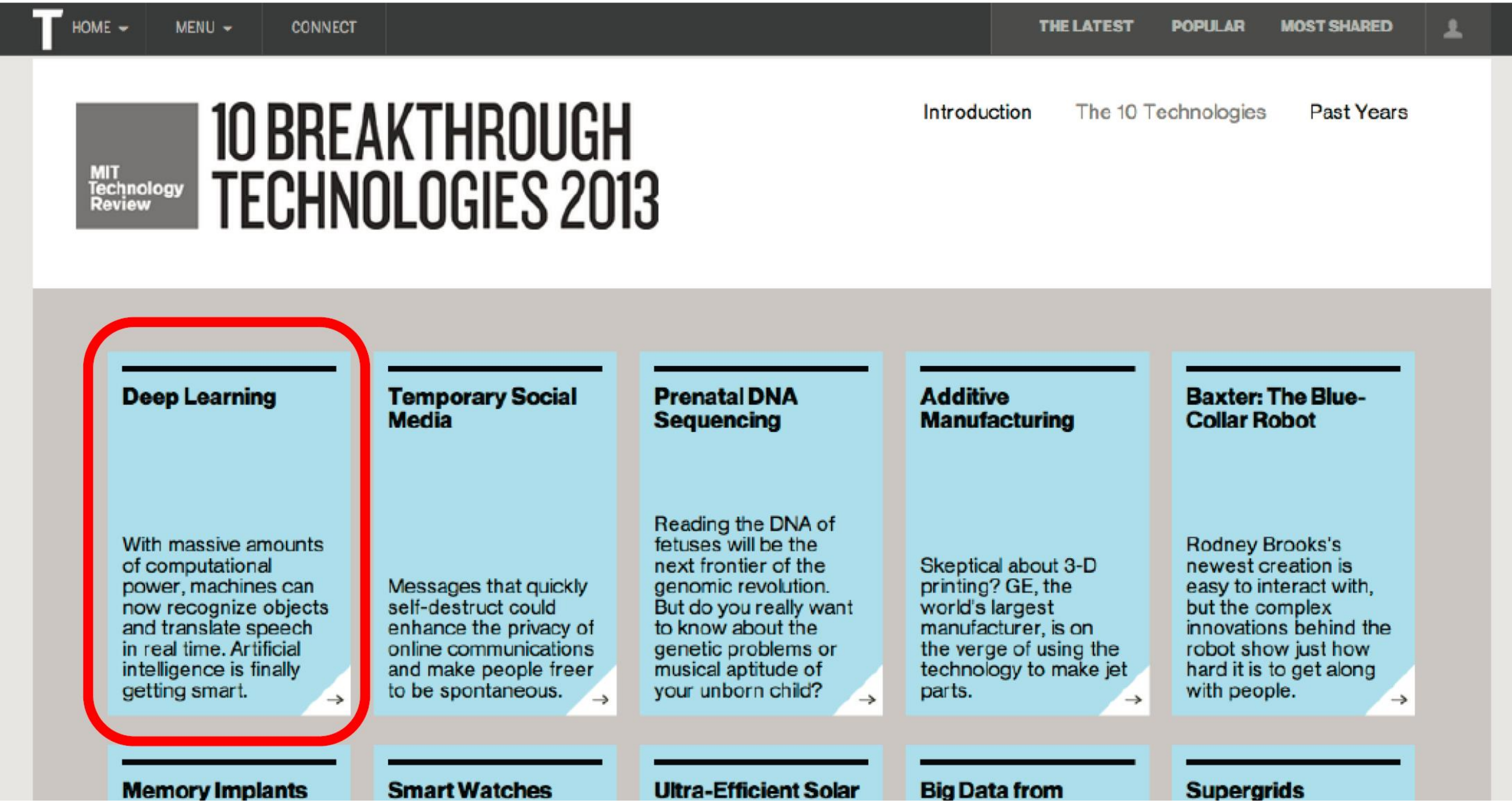
- Final exam will be handed out in class Tuesday, due Friday, 5/12, by 5pm

Aside: “CNN” vs “ConvNet”

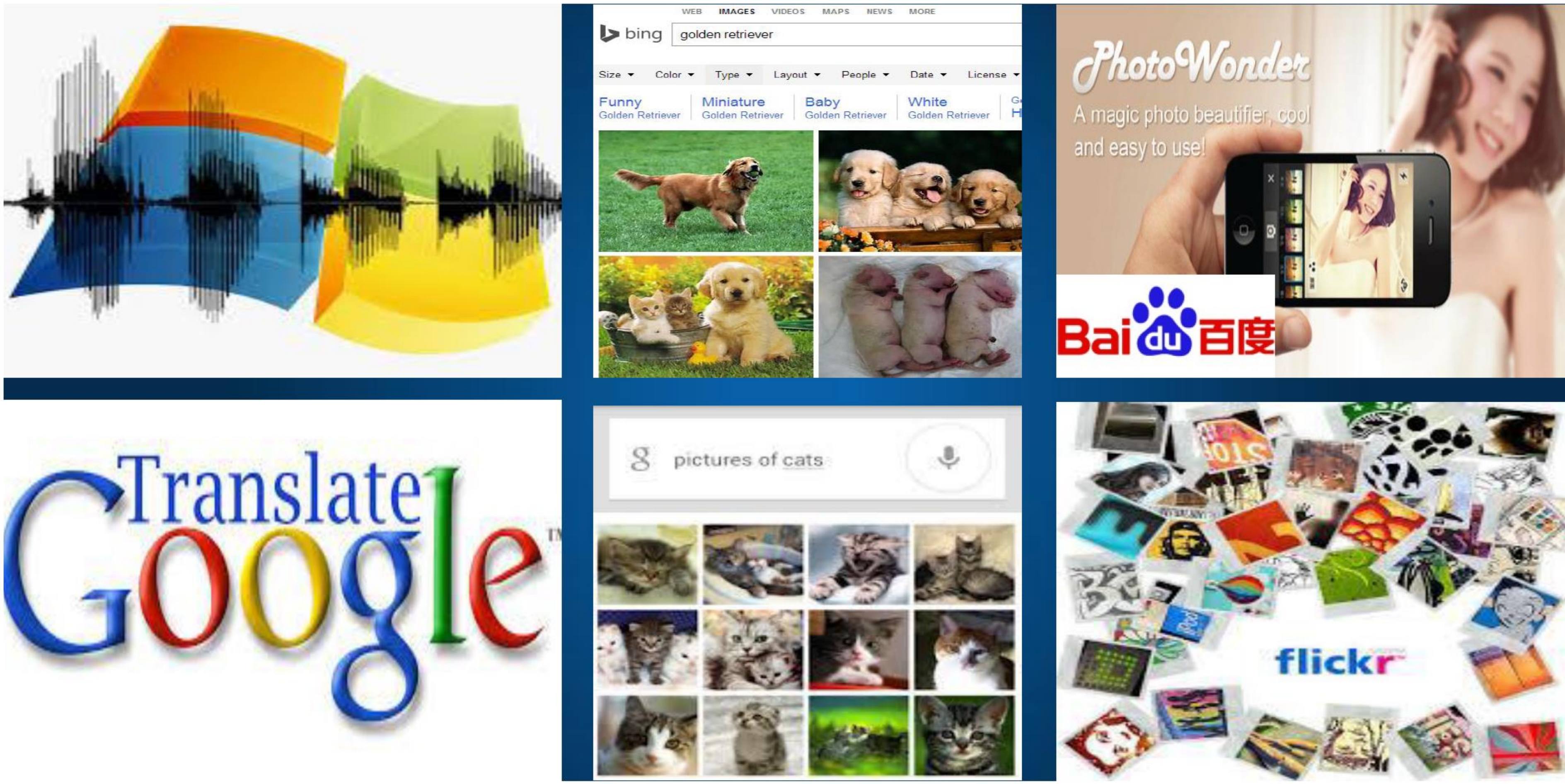
Note:

- There are many papers that use either phrase, but
- “ConvNet” is the preferred term, since “CNN” clashes with other things called CNN

Motivation

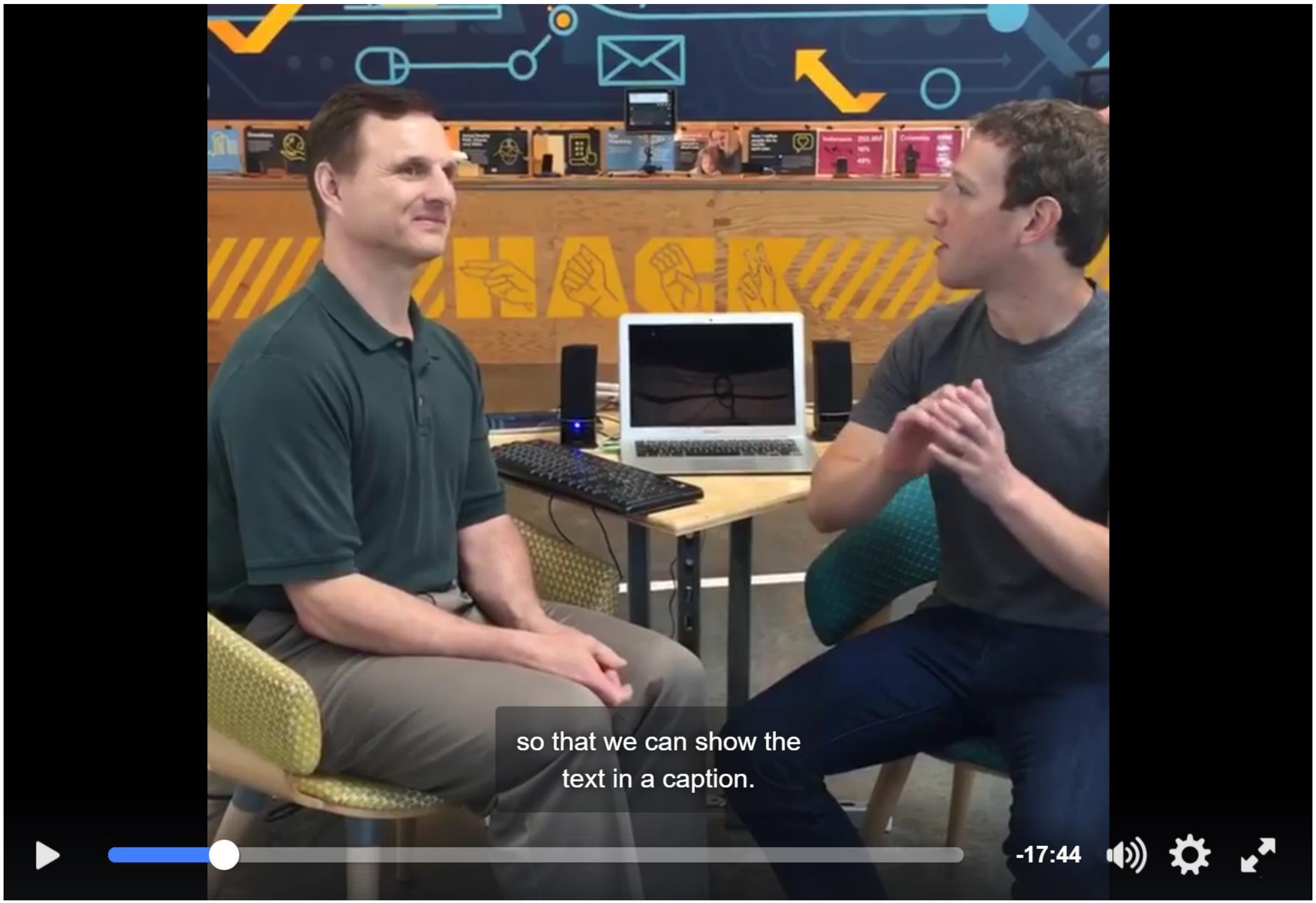

HOME
MENU
CONNECT
THE LATEST
POPULAR
MOST SHARED
MIT Technology Review
10 BREAKTHROUGH TECHNOLOGIES 2013
Introduction
The 10 Technologies
Past Years
Deep Learning
With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.
Temporary Social Media
Messages that quickly self-destruct could enhance the privacy of online communications and make people freer to be spontaneous.
Prenatal DNA Sequencing
Reading the DNA of fetuses will be the next frontier of the genomic revolution. But do you really want to know about the genetic problems or musical aptitude of your unborn child?
Additive Manufacturing
Skeptical about 3-D printing? GE, the world's largest manufacturer, is on the verge of using the technology to make jet parts.
Baxter: The Blue-Collar Robot
Rodney Brooks's newest creation is easy to interact with, but the complex innovations behind the robot show just how hard it is to get along with people.
Memory Implants
Smart Watches
Ultra-Efficient Solar
Big Data from
Supergrids

Products



[Slide credit: Deva Ramanan]

Helping the Blind



https://www.facebook.com/zuck/videos/10102801434799001/

(Unrelated) Dog vs Food

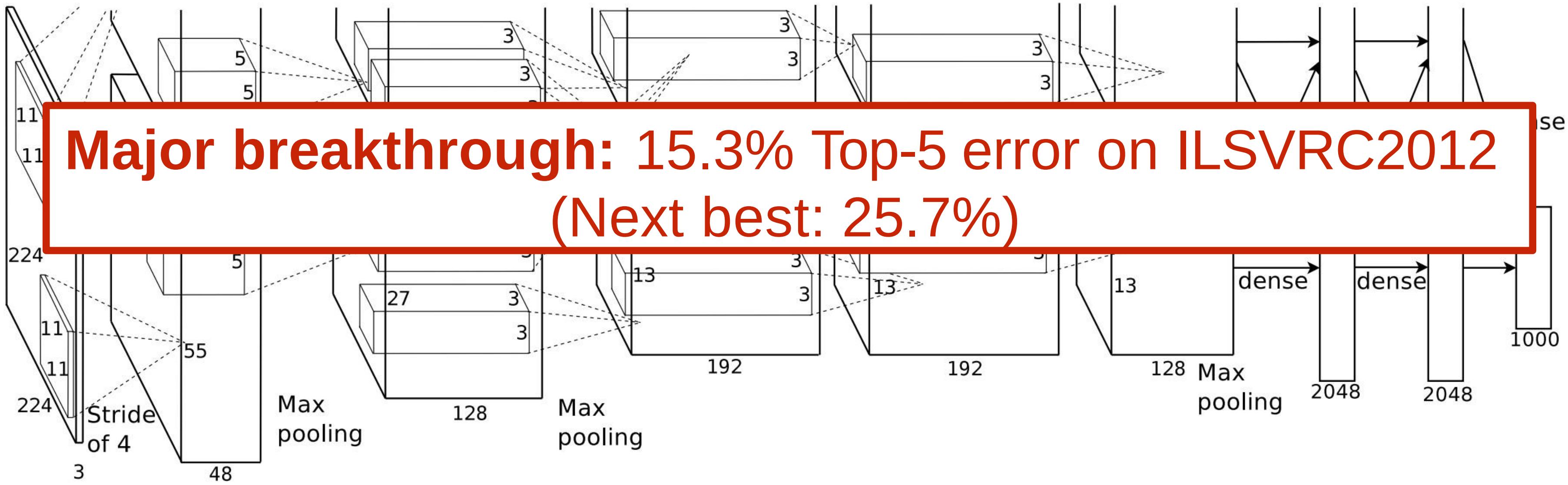
[Karen Zack, @teenybiscuit]

(Unrelated) Dog vs Food

[Karen Zack, @teenybiscuit]

CNNs in 2012: “SuperVision” (aka “AlexNet”)

“AlexNet” — Won the ILSVRC2012 Challenge

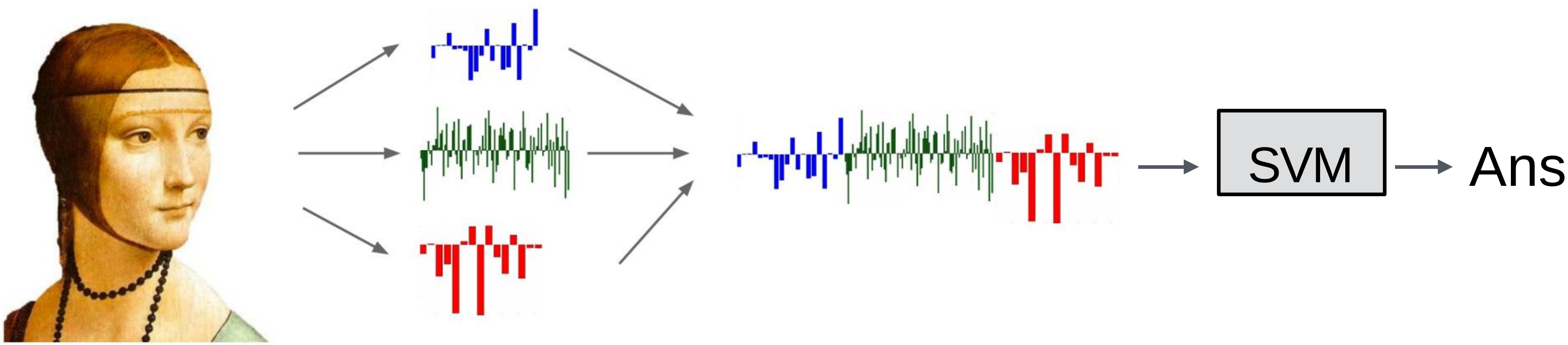


Major breakthrough: 15.3% Top-5 error on ILSVRC2012 (Next best: 25.7%)

Figure 2: An illustration of the architecture of our CNN, explicitly showing the delineation of responsibilities between the two GPUs. One GPU runs the layer-parts at the top of the figure while the other runs the layer-parts at the bottom. The GPUs communicate only at certain layers. The network’s input is 150,528-dimensional, and the number of neurons in the network’s remaining layers is given by 253,440–186,624–64,896–64,896–43,264–4096–4096–1000.

[Krizhevsky, Sutskever, Hinton. NIPS 2012]

Recap: Before Deep Learning



Input Pixels

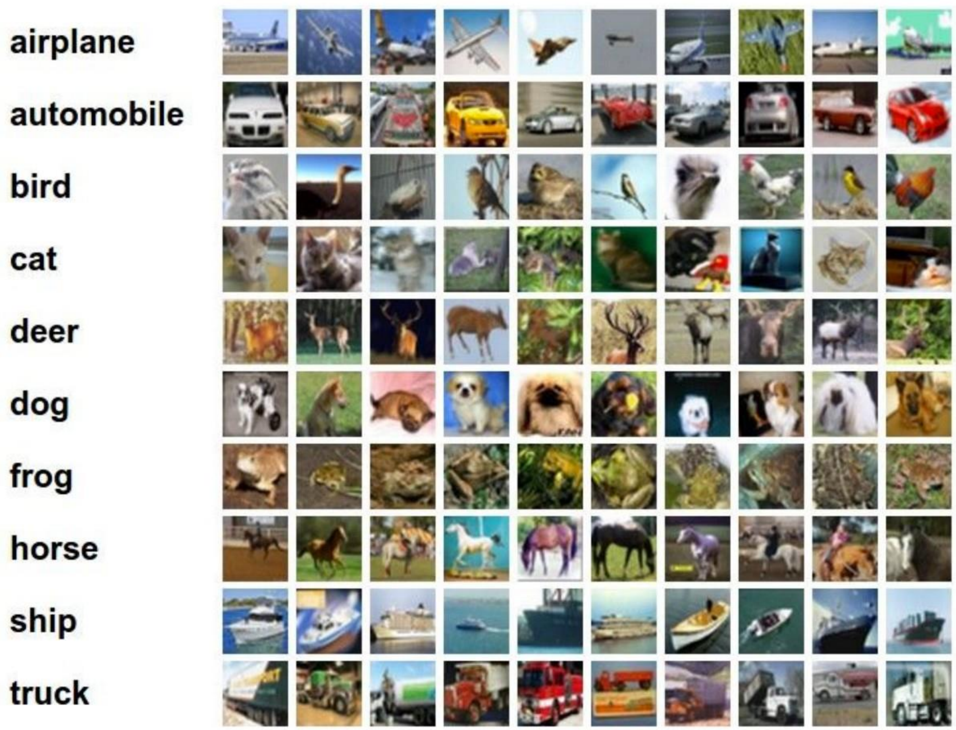
Extract Features

*Concatenate into a vector **x***

Linear Classifier

Figure: Karpathy 2016

Why use features?
Why not pixels?



$$f(x_i, W, b) = Wx_i + b$$

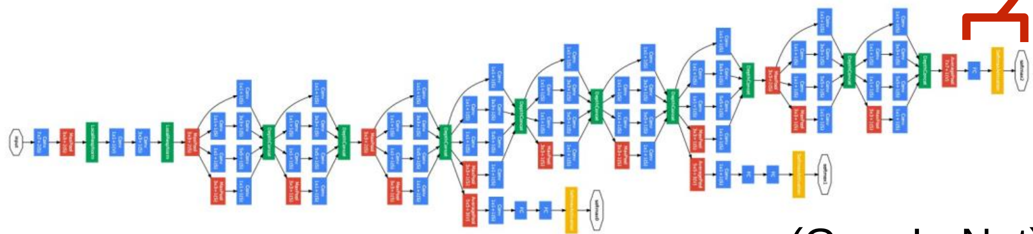
Q: What would be a very hard set of classes for a linear classifier to distinguish?

(assuming x = pixels)

Slide from Karpathy 2016

The last layer of (most) CNNs are linear classifiers

This piece is just a linear classifier



→ Ans

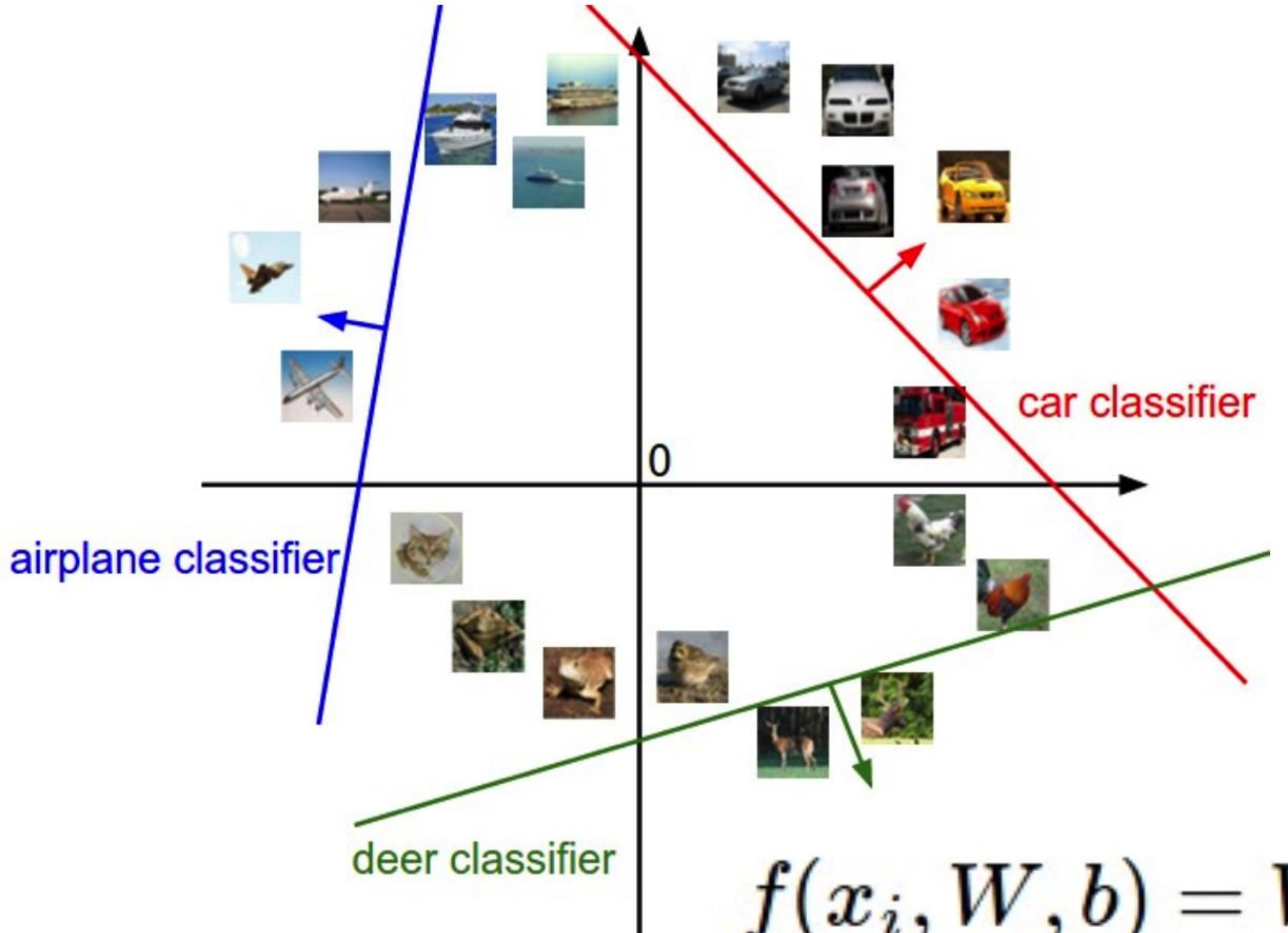
(GoogLeNet)

Input Pixels

Perform everything with a big neural network, trained end-to-end

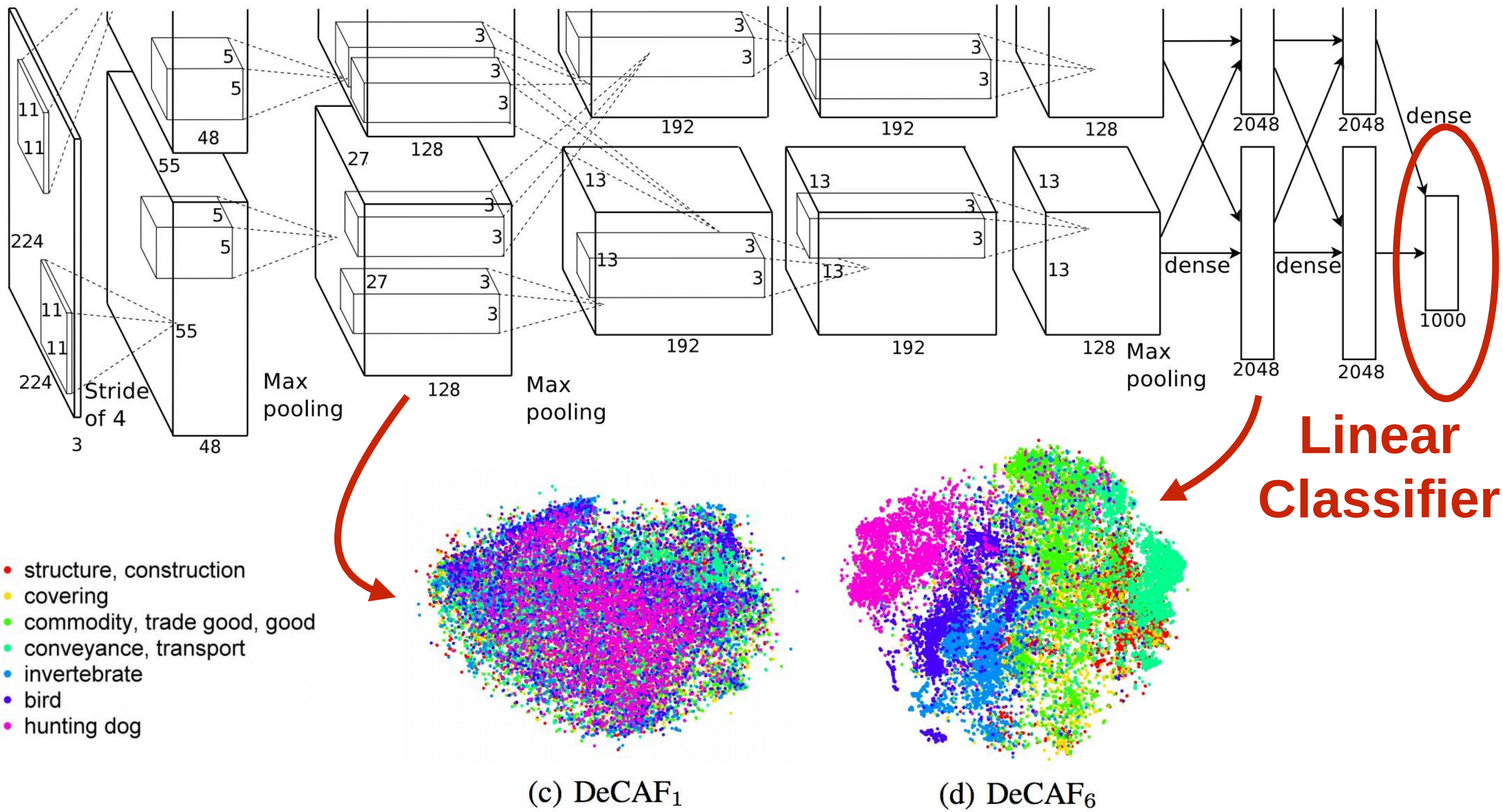
Key: perform enough processing so that by the time you get to the end of the network, the classes are linearly separable

Linearly separable classes



$$f(x_i, W, b) = Wx_i + b$$

Example: Visualizing AlexNet in 2D with t-SNE



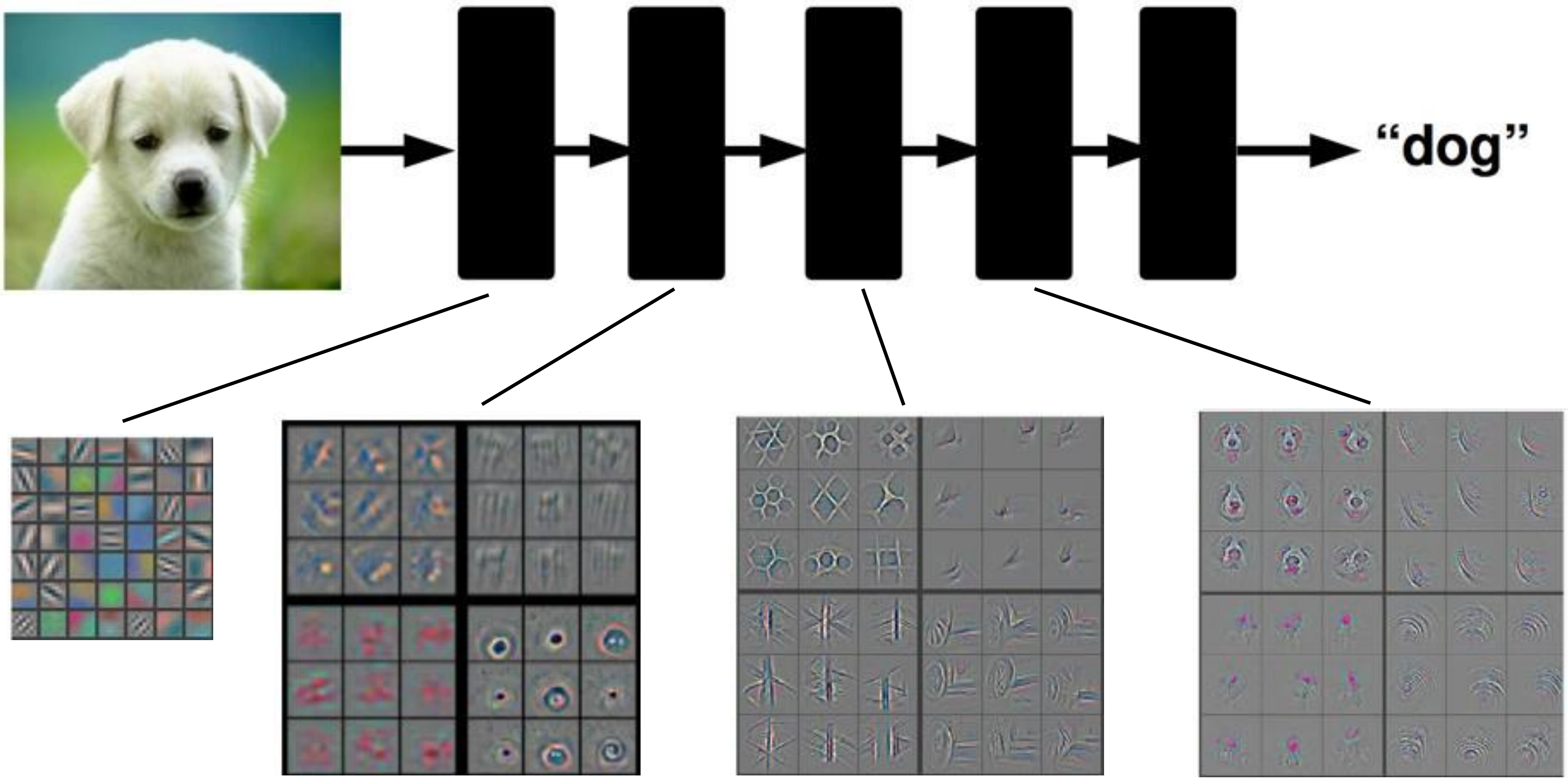
(2D visualization using t-SNE)

[Donahue, "DeCAF: DeCAF: A Deep Convolutional …", arXiv 2013]

Roadmap for today

- Neural networks
- Convolutional neural networks
- Optimization matters!
  - Backpropagation algorithm

Feature hierarchy with ConvNets
End-to-end models



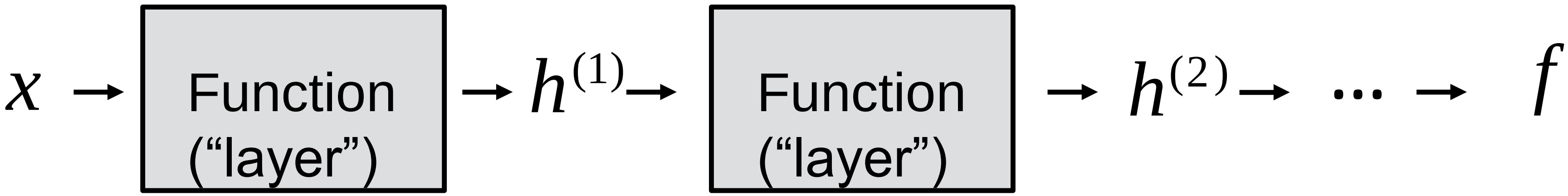
[Zeiler and Fergus, "Visualizing and Understanding Convolutional Networks", 2013]

Learning Feature Hierarchy

- Learn hierarchy
- All the way from pixels → classifier
- One layer extracts features from output of previous layer

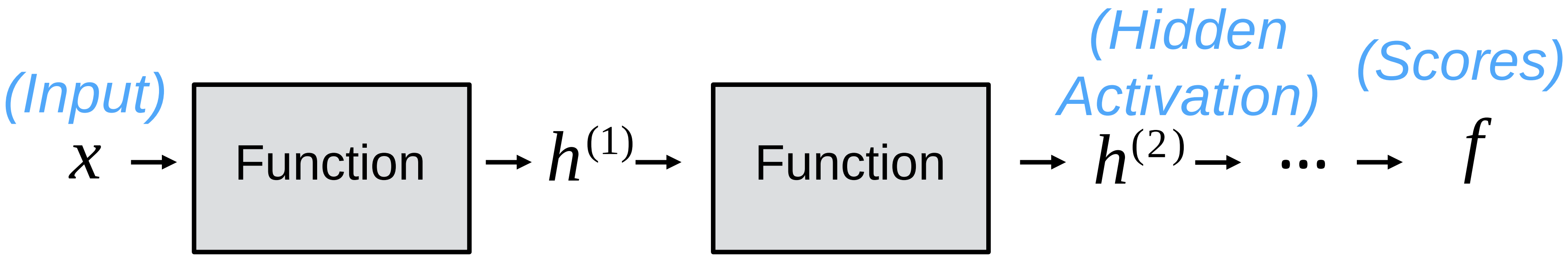
Slide: R. Fergus

Neural Networks: The Big Picture

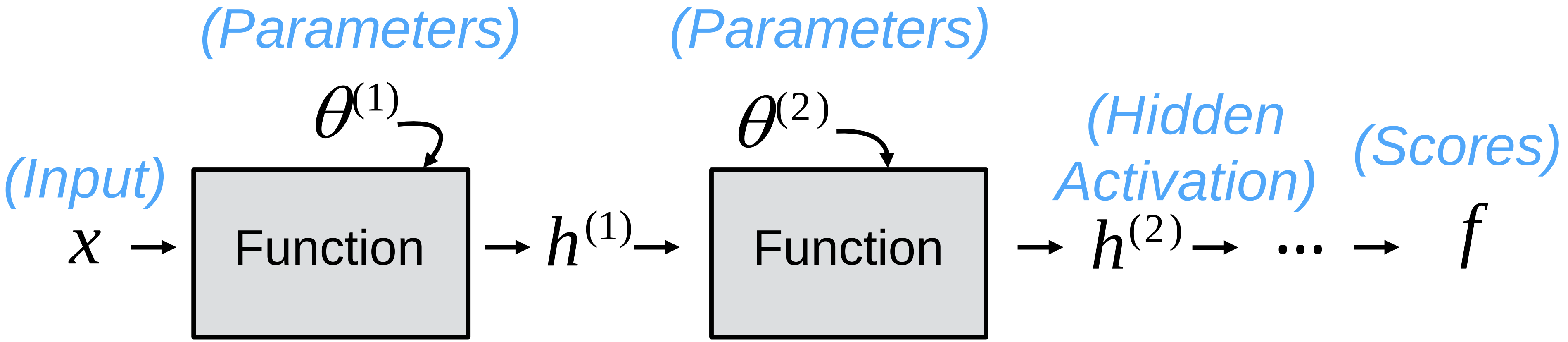


Key idea: **composition of simpler functions** called "layers" (e.g., multiple linear layers (not just one))

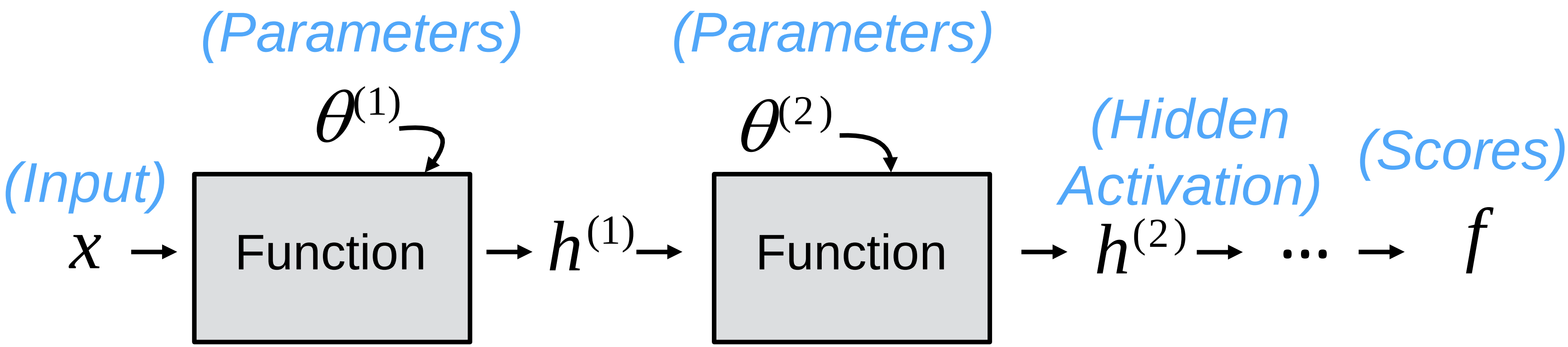
Neural Networks: The Big Picture



Neural Networks: The Big Picture

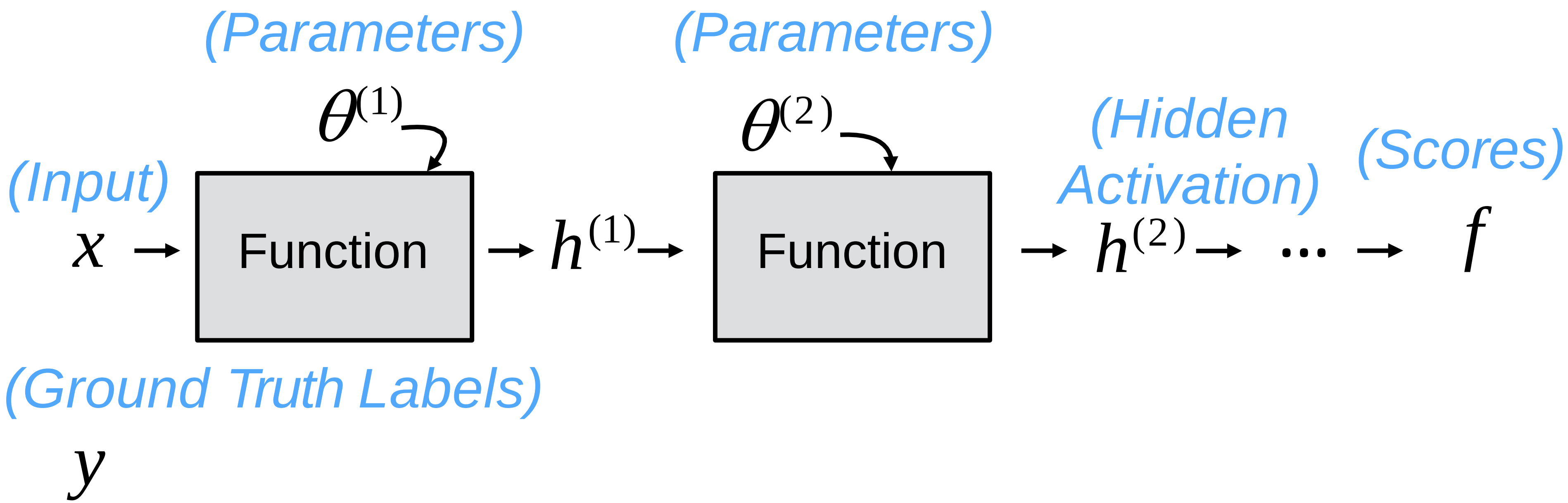


Neural Networks: The Big Picture



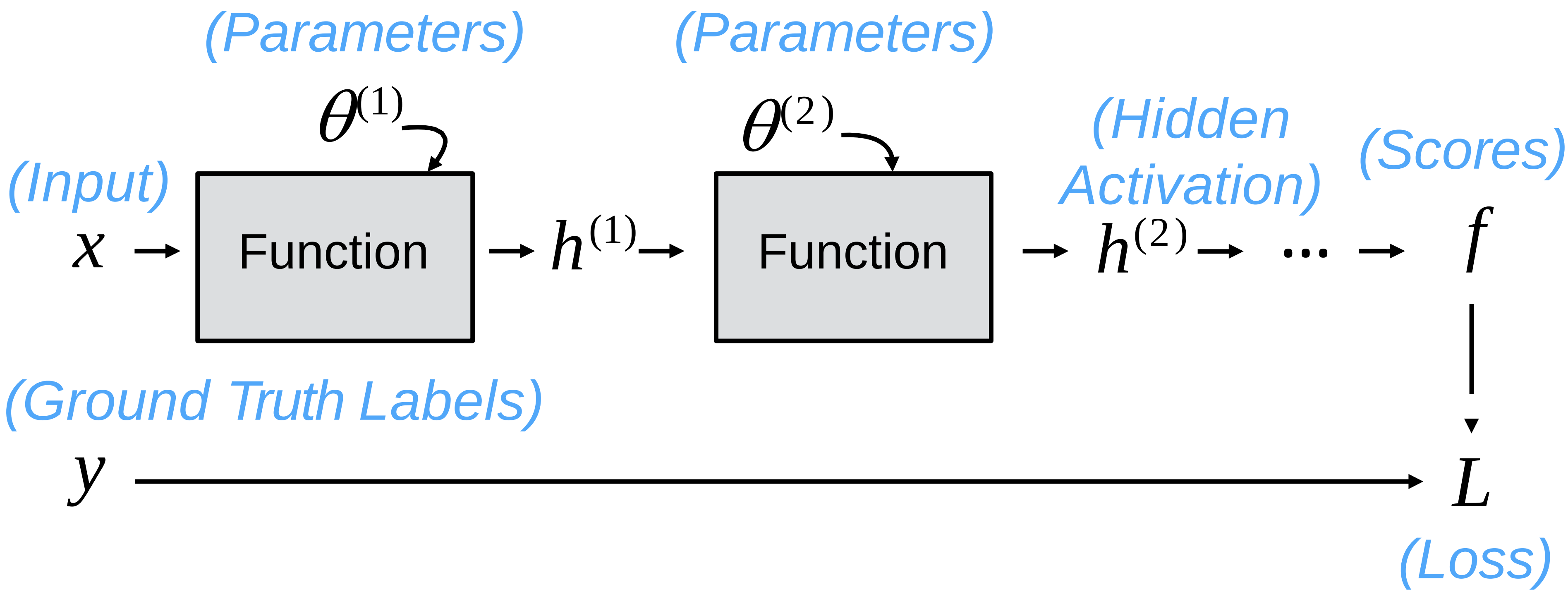
Here, θ represents whatever parameters that layer is using (e.g. for a "linear layer" $\theta^{(1)} = \{ W^{(1)}, b^{(1)} \}$).

Neural Networks: The Big Picture



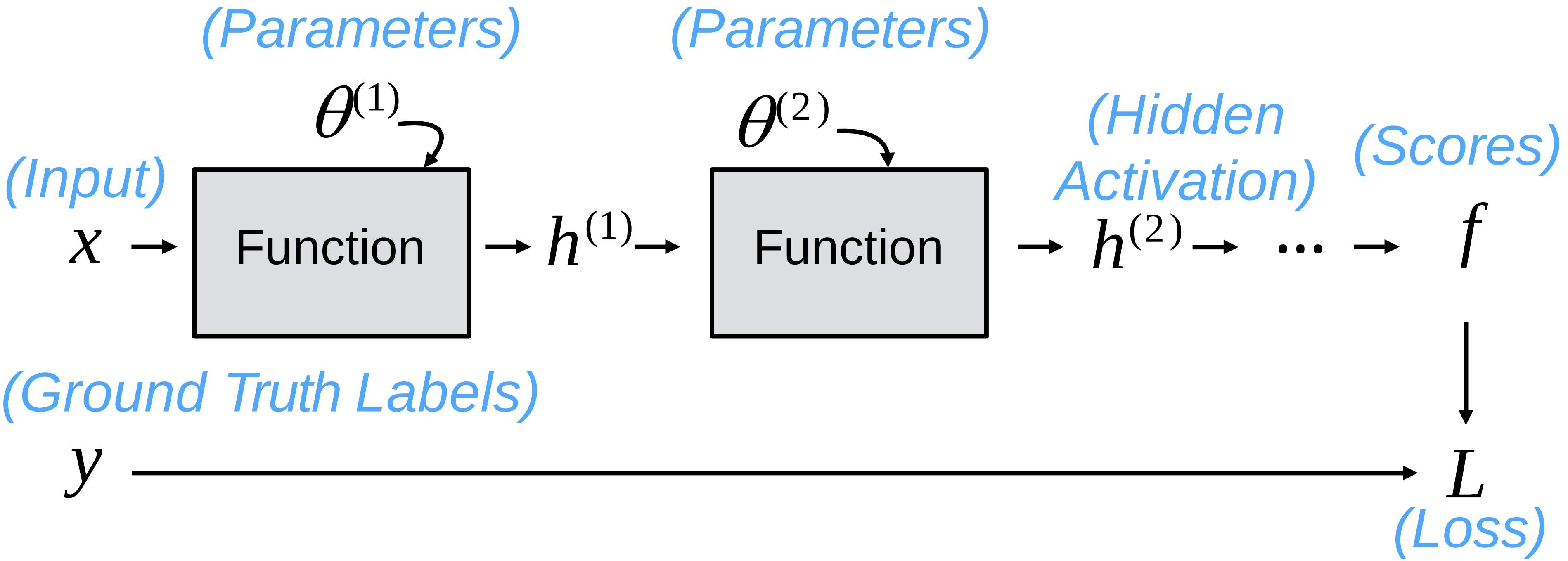
Here, θ represents whatever parameters that layer is using (e.g. for a "linear layer" $\theta^{(1)} = \{ W^{(1)}, b^{(1)} \}$).

Neural Networks: The Big Picture



Here, θ represents whatever parameters that layer is using (e.g. for a "linear layer" $\theta^{(1)} = \{ W^{(1)}, b^{(1)} \}$).

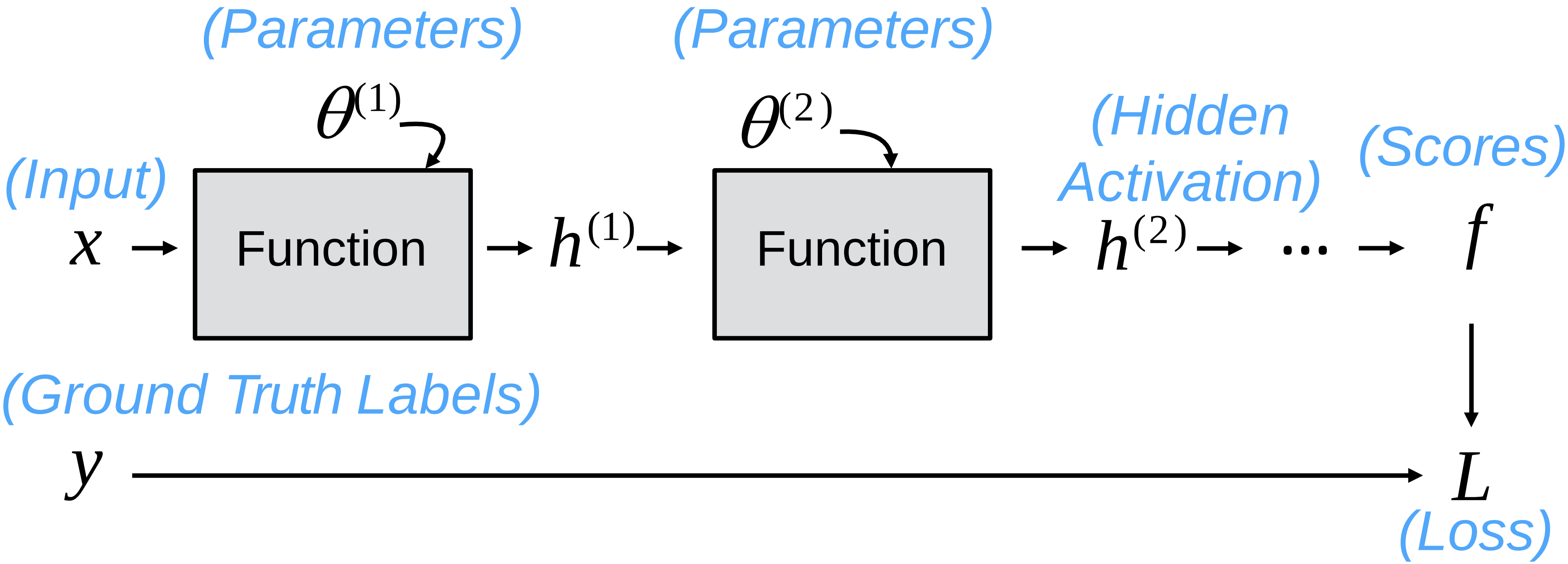
Neural Networks: The Big Picture



Here, θ represents whatever parameters that layer is using (e.g. for a "linear layer" $\theta^{(1)} = \{ W^{(1)}, b^{(1)} \}$).

Recall: the loss "L" measures how far the predictions "f" are from the labels "y". The most common loss is Softmax.

Neural Networks: The Big Picture

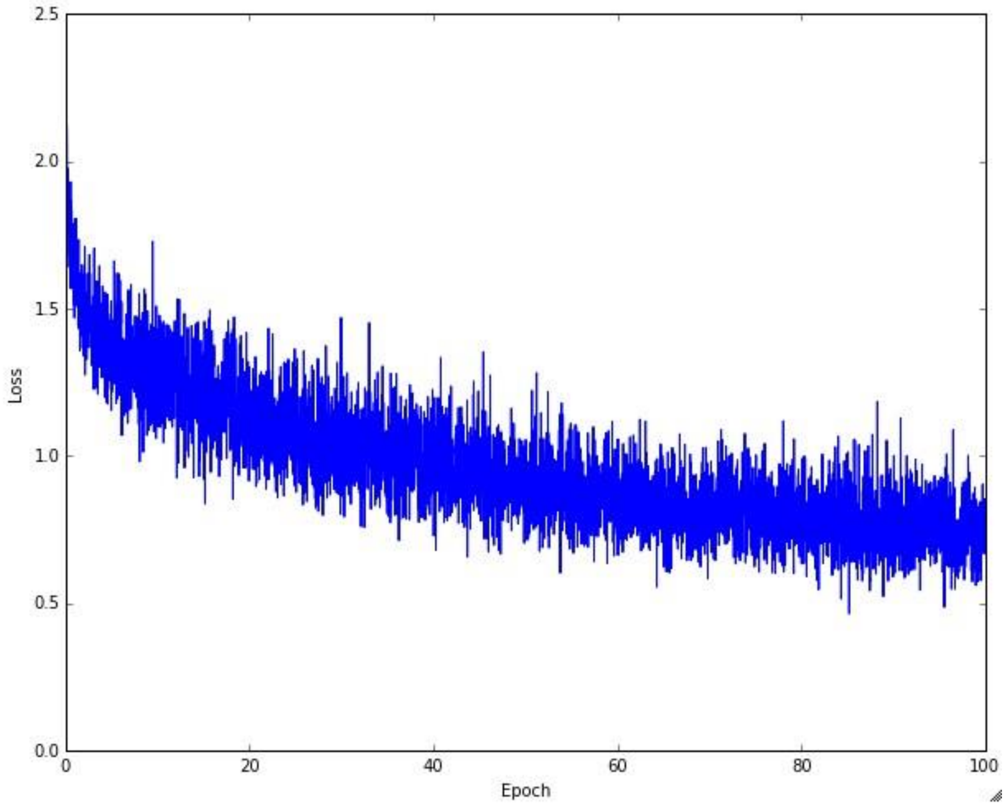


Here, θ represents whatever parameters that layer is using (e.g. for a "linear layer" $\theta^{(1)} = \{ W^{(1)}, b^{(1)} \}$).

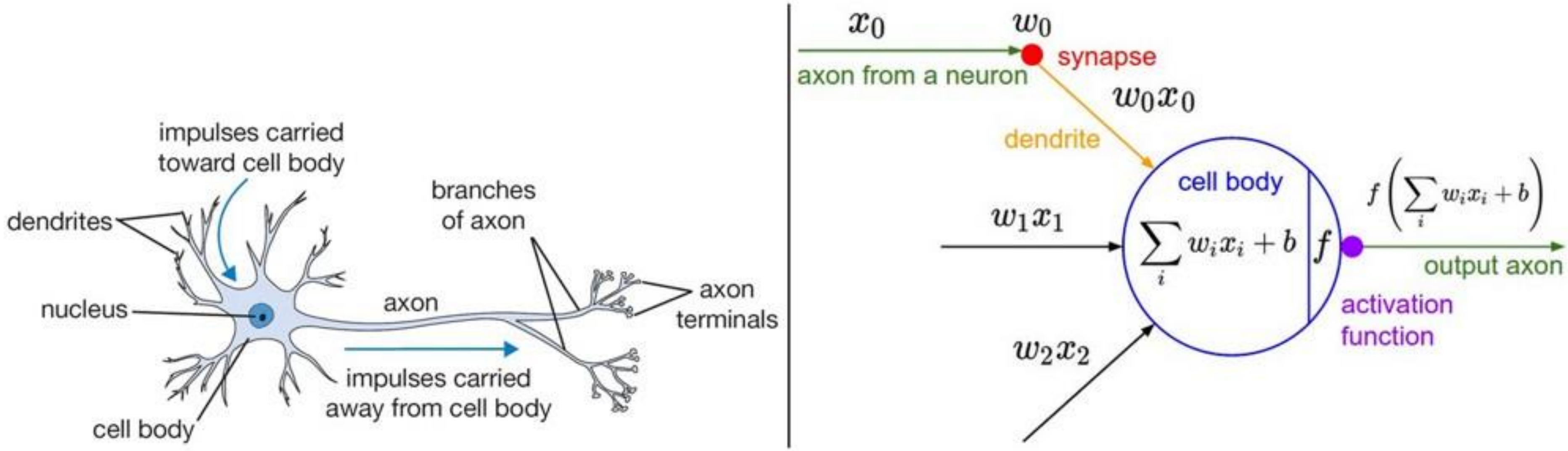
Key problem: Adjust the weights in all layers to minimize the training loss. We do this with *backpropagation.*

Training Deep Neural Networks

- Must run many iterations of batch gradient descent
- With lots of other tweaks



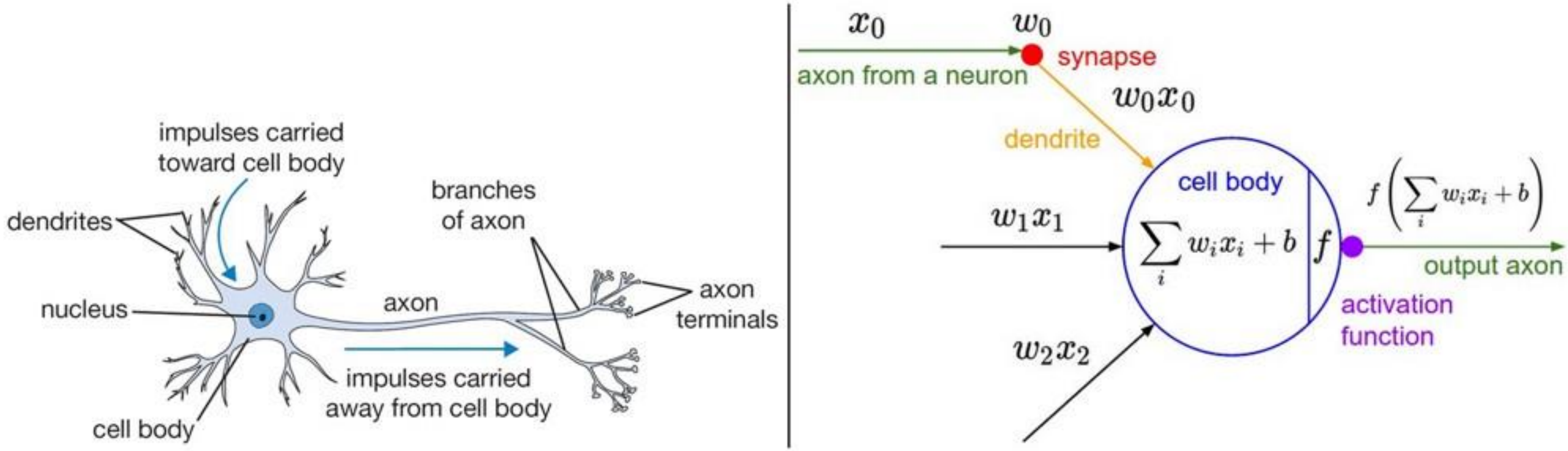
Aside: Inspiration from Biology



A cartoon drawing of a biological neuron (left) and its mathematical model (right).

Figure: Andrej Karpathy

Aside: Inspiration from Biology

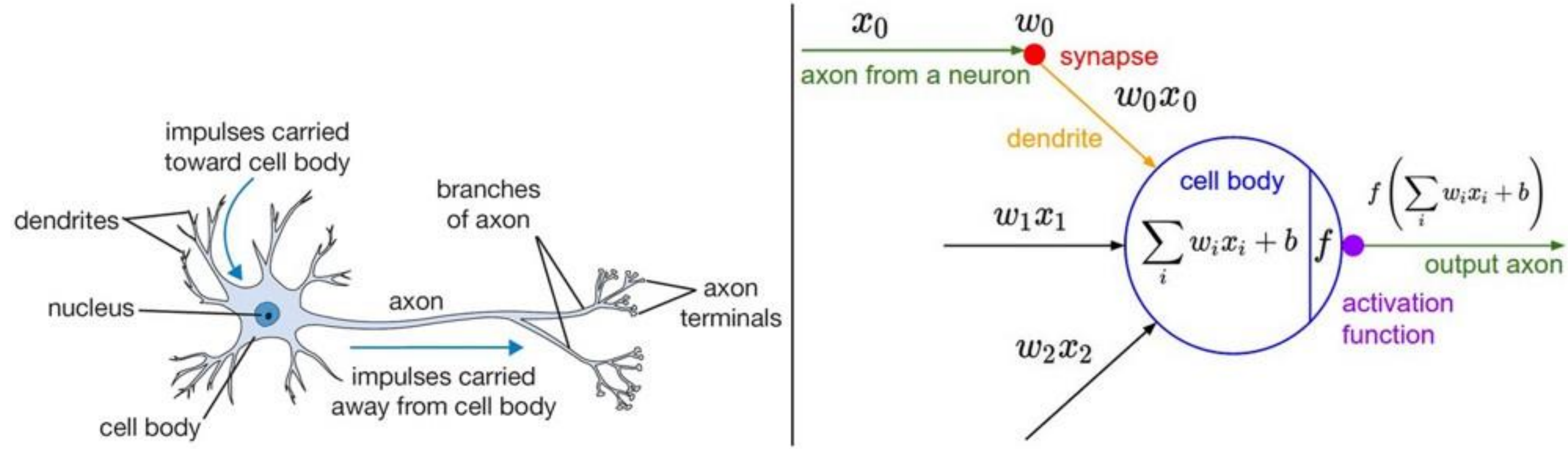


A cartoon drawing of a biological neuron (left) and its mathematical model (right).

Neural nets are **loosely inspired** by biology

Figure: Andrej Karpathy

Aside: Inspiration from Biology



A cartoon drawing of a biological neuron (left) and its mathematical model (right).

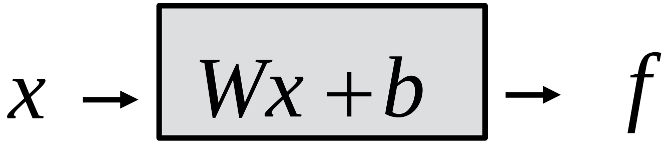
Neural nets are **loosely inspired** by biology

But they certainly are **not** a model of how the brain works, or even how neurons work

Figure: Andrej Karpathy

Simple Neural Net: 1 Layer

Let’s consider a simple 1-layer network:

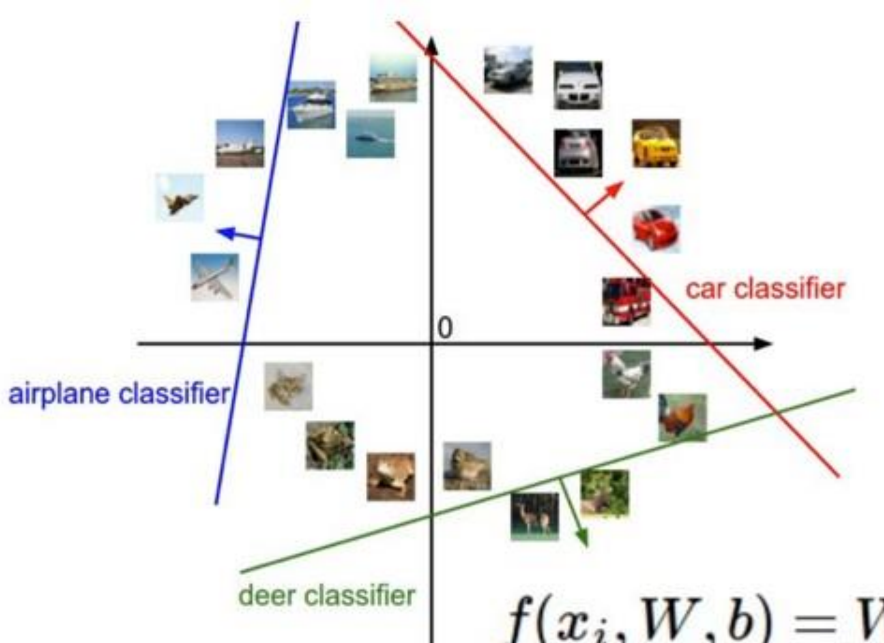


Simple Neural Net: 1 Layer

Let's consider a simple 1-layer network:

$$x \rightarrow \boxed{Wx + b} \rightarrow f$$

This is just what we've seen before ("linear classifier"):



$$f(x_i, W, b) = Wx_i + b$$

1 Layer Neural Net

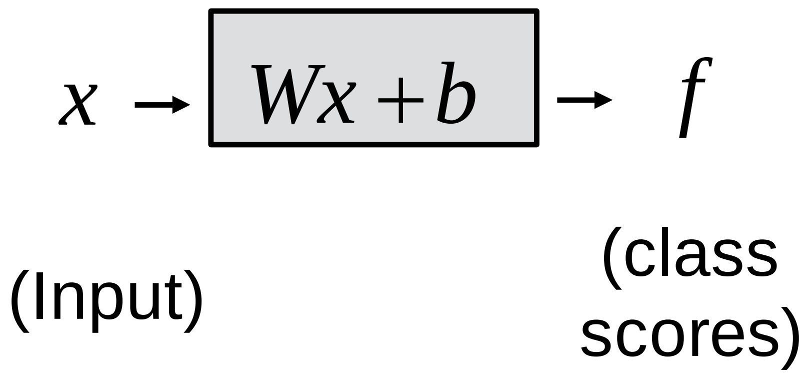
Block Diagram:

$$x \rightarrow \boxed{Wx + b} \rightarrow f$$

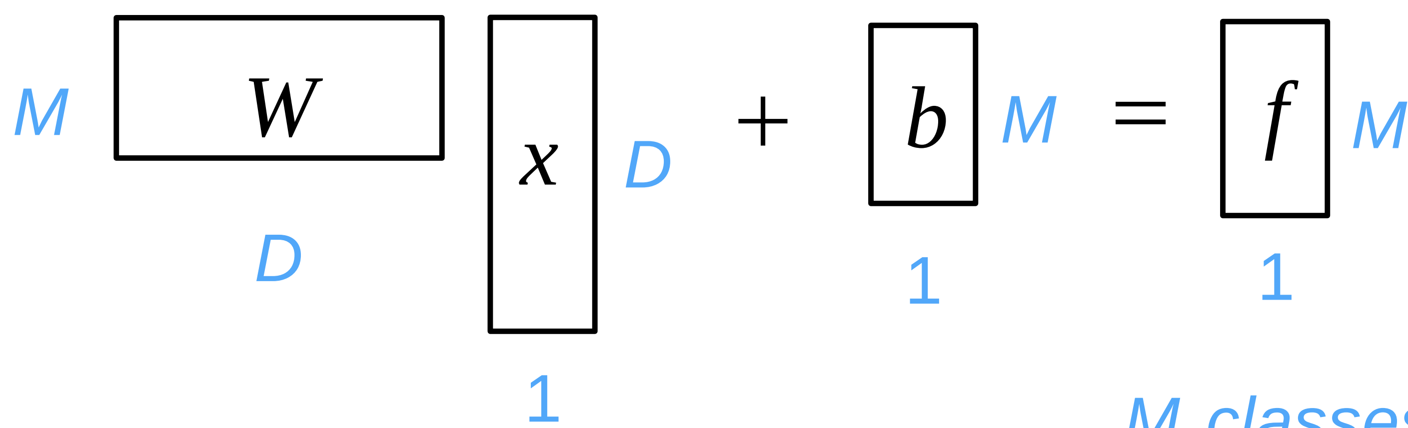
(Input) (class scores)

1 Layer Neural Net

Block Diagram:



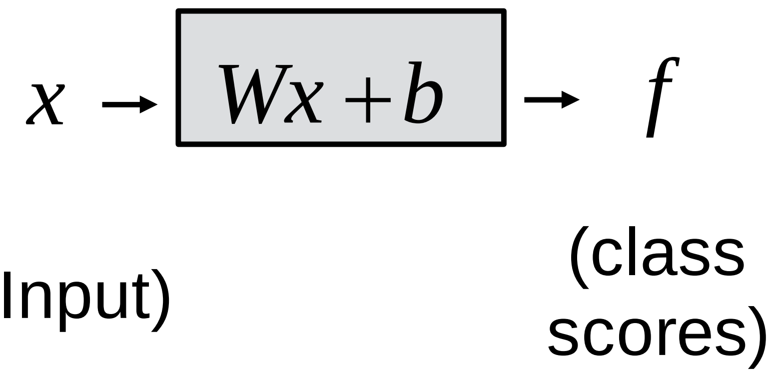
Expanded Block Diagram:



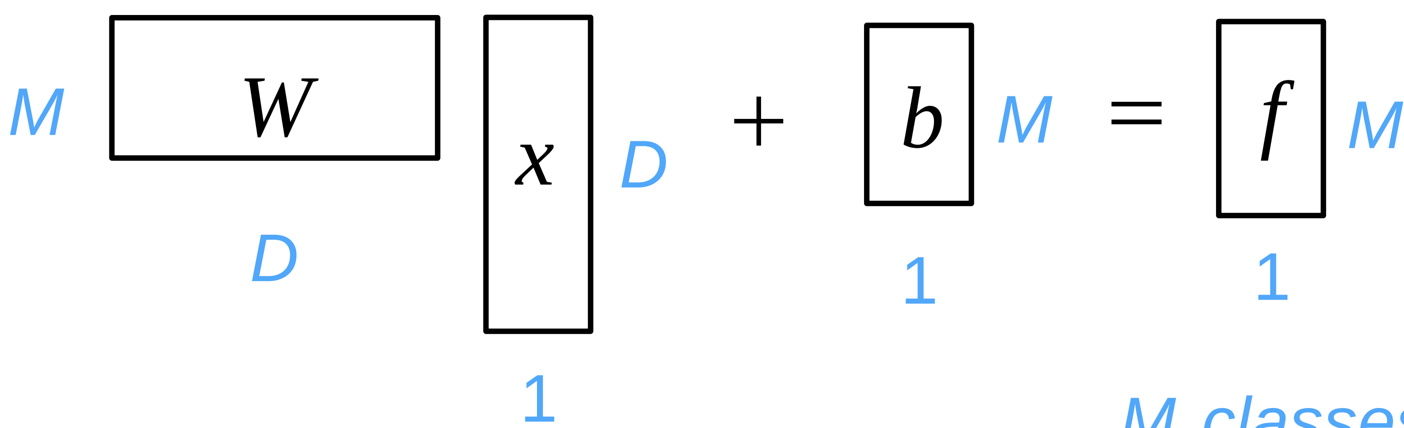
M classes
D features
1 example

1 Layer Neural Net

Block Diagram:



Expanded Block Diagram:



NumPy:

```
f = np.dot(W, x) + b
```

M classes
D features
1 example