

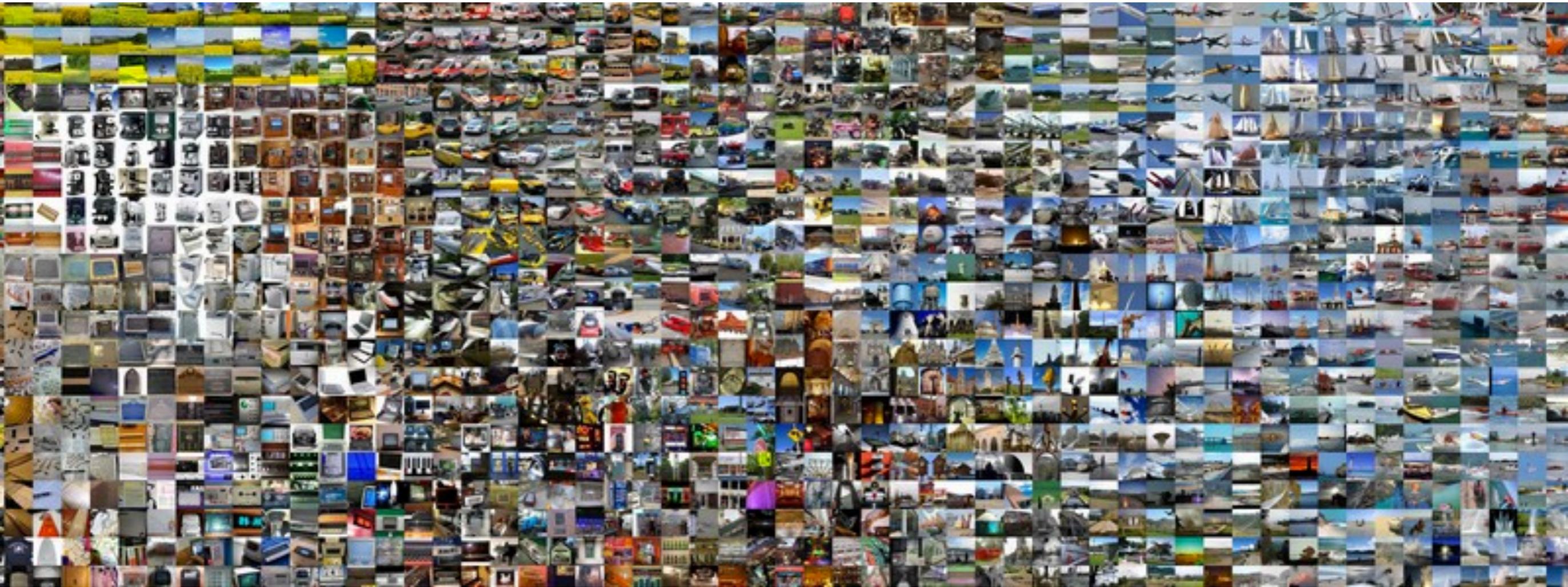
95-865:
Introduction to Neural Nets
and Deep Learning

George Chen

A few slides are by Phillip Isola

IMAGENET

Over 10 million images, 1000 object classes



2011: Traditional computer vision achieves accuracy ~74%

2012: Initial deep neural network approach accuracy ~84%

2015 onwards: Deep learning achieves accuracy 96%+

Russakovsky et al. ImageNet Large Scale Visual Recognition Challenge. IJCV 2015.

Deep Learning Takeover

Academia:

- Top computer vision conferences (CVPR, ICCV, ECCV) are now nearly all about deep learning
- Top machine learning conferences (ICML, NIPS) have *heavily* been taken over by deep learning

Heavily dominated by industry now!

Extremely useful in practice:

- Near human level image classification (including handwritten digit recognition)
- Near human level speech recognition
- Improvements in machine translation, text-to-speech
- Self-driving cars
- *Better* than humans at playing Go





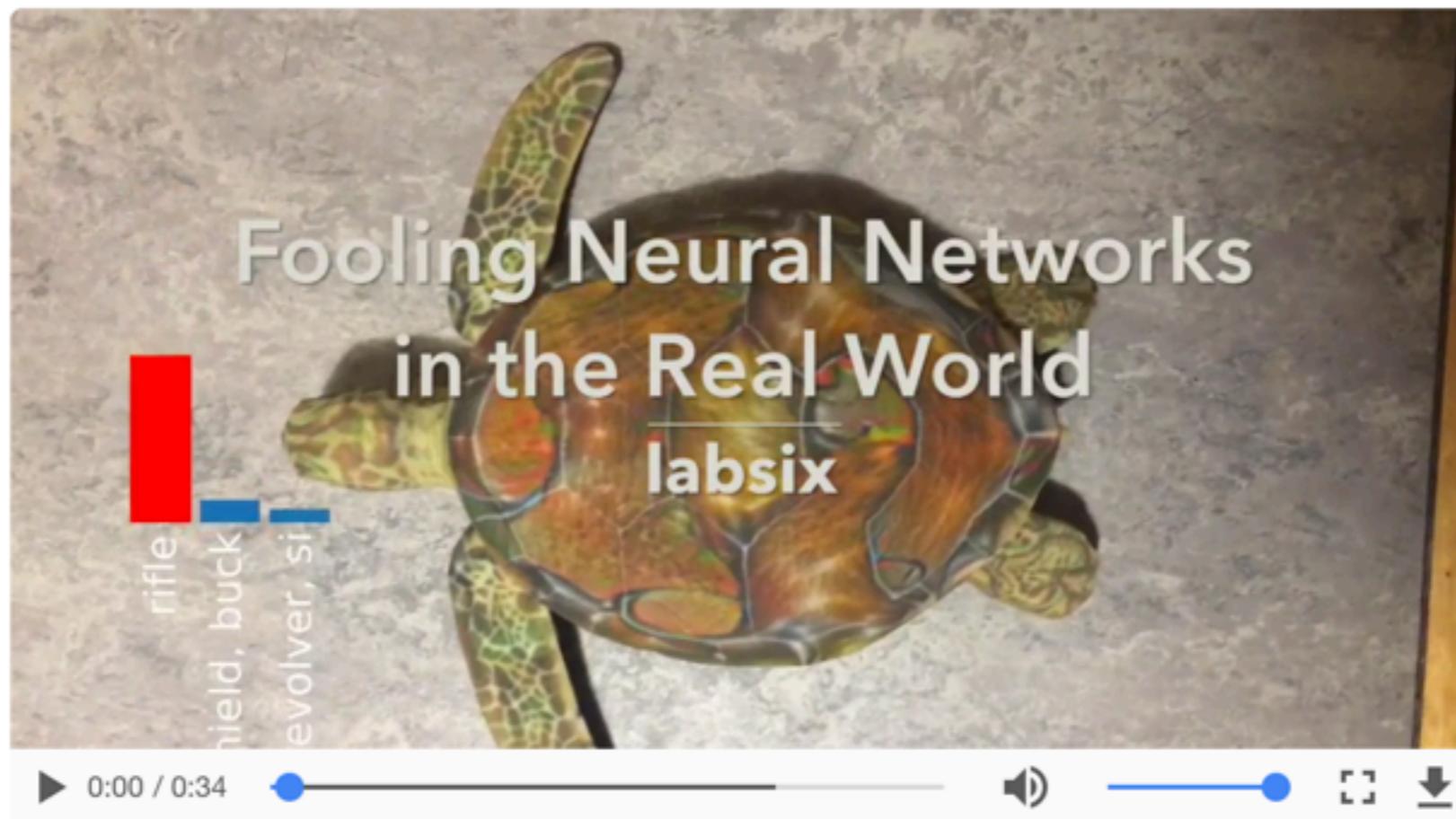
Google DeepMind's AlphaGo vs Lee Sedol, 2016

Is it all hype?

Fooling Neural Networks in the Physical World with 3D Adversarial Objects

31 Oct 2017 · 3 min read — shared on [Hacker News](#), [Lobsters](#), [Reddit](#), [Twitter](#)

We've developed an approach to generate *3D adversarial objects* that reliably fool neural networks in the real world, no matter how the objects are looked at.

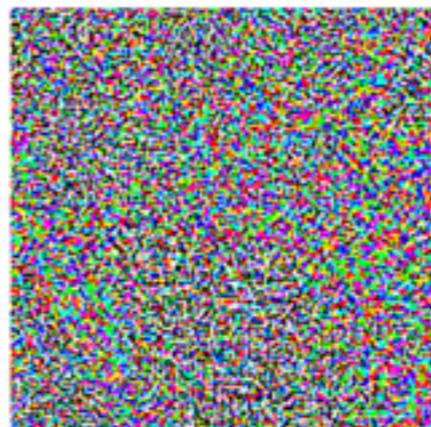


Neural network based classifiers reach near-human performance in many tasks, and they're used in high risk, real world systems. Yet, these same neural networks are particularly vulnerable to *adversarial examples*, carefully perturbed inputs that cause

Source: labsix



+ .007 ×



=



panda
~58% confidence

adversarial
noise

gibbon
~99% confidence

Source: Goodfellow, Shlens, and Szegedy. Explaining and Harnessing Adversarial Examples. ICLR 2015.



a cat is sitting on a toilet in a bathroom

Source: Gizmodo article "This Neural Network's Hilariously Bad Image Descriptions Are Still Advanced AI". September 16, 2015. (They're using the NeuralTalk image-to-caption software.)

Another AI Winter?

~1970's: First AI winter over symbolic AI

~1980's: Second AI winter over "expert systems"

Every time: Lots of hype, explosion in funding, then bubble bursts



Michael Jordan [Follow](#)

Michael I. Jordan is a Professor in the Department of Electrical Engineering and Computer Sciences and the Department of Statistics at UC Berkeley.

Apr 18 · 16 min read

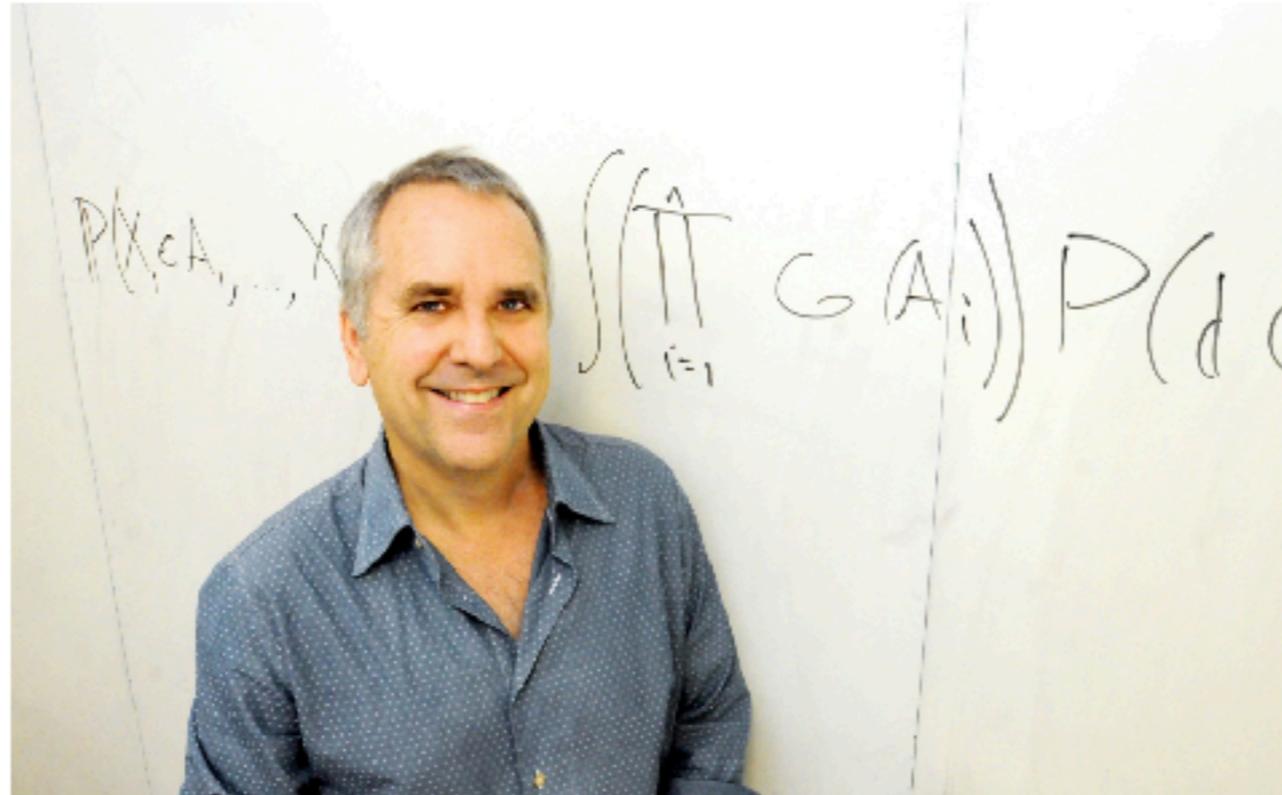


Photo credit: Peg Skorpinski

Artificial Intelligence—The Revolution Hasn't Happened Yet

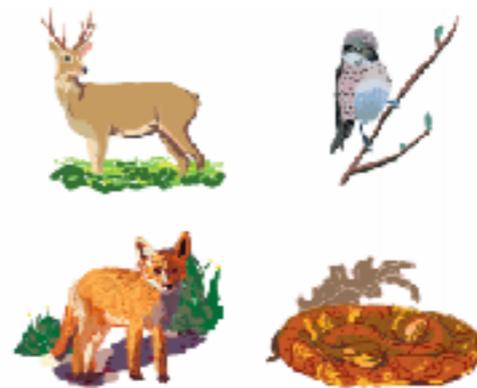
Artificial Intelligence (AI) is the mantra of the current era. The phrase is intoned by technologists, academicians, journalists and venture capitalists

<https://medium.com/@mijordan3/artificial-intelligence-the-revolution-hasnt-happened-yet-5e1d5812e1e7>

What is deep learning?



Classification units



PIT/AIT



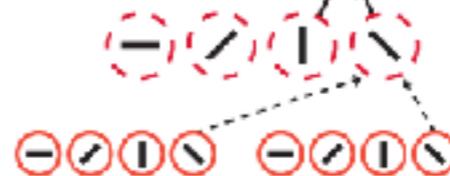
V4/PIT



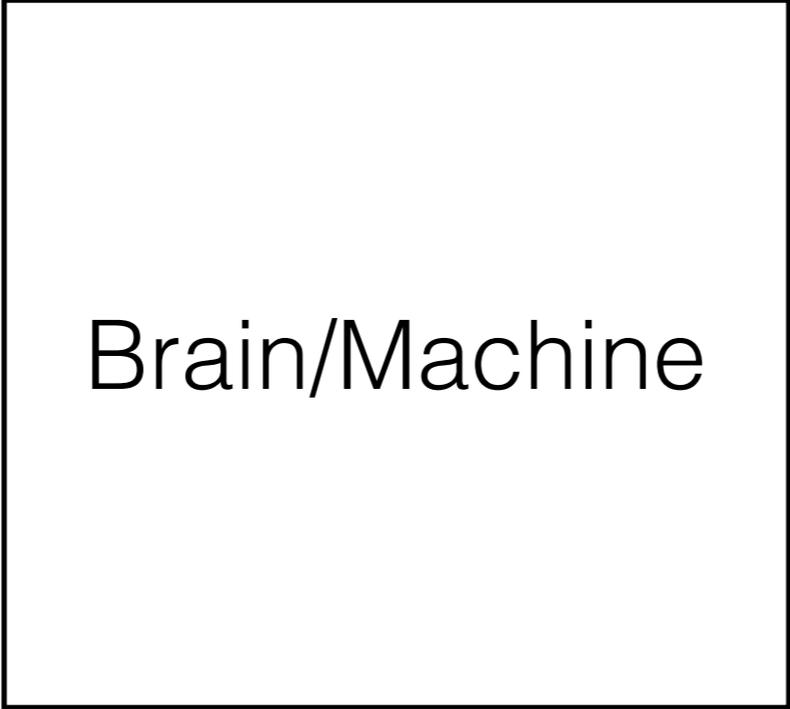
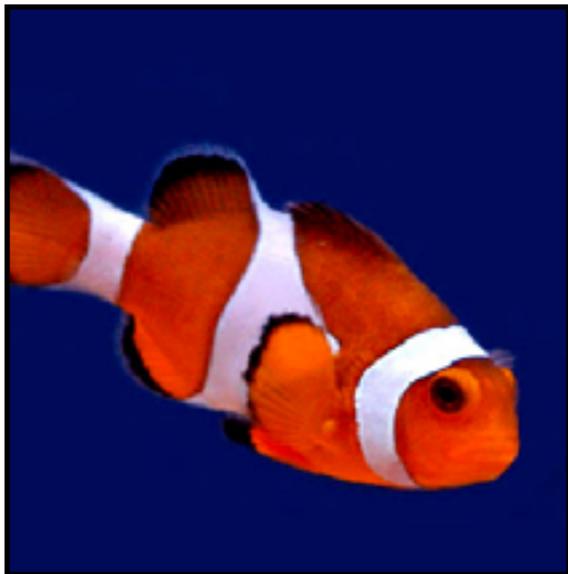
V2/V4



V1/V2

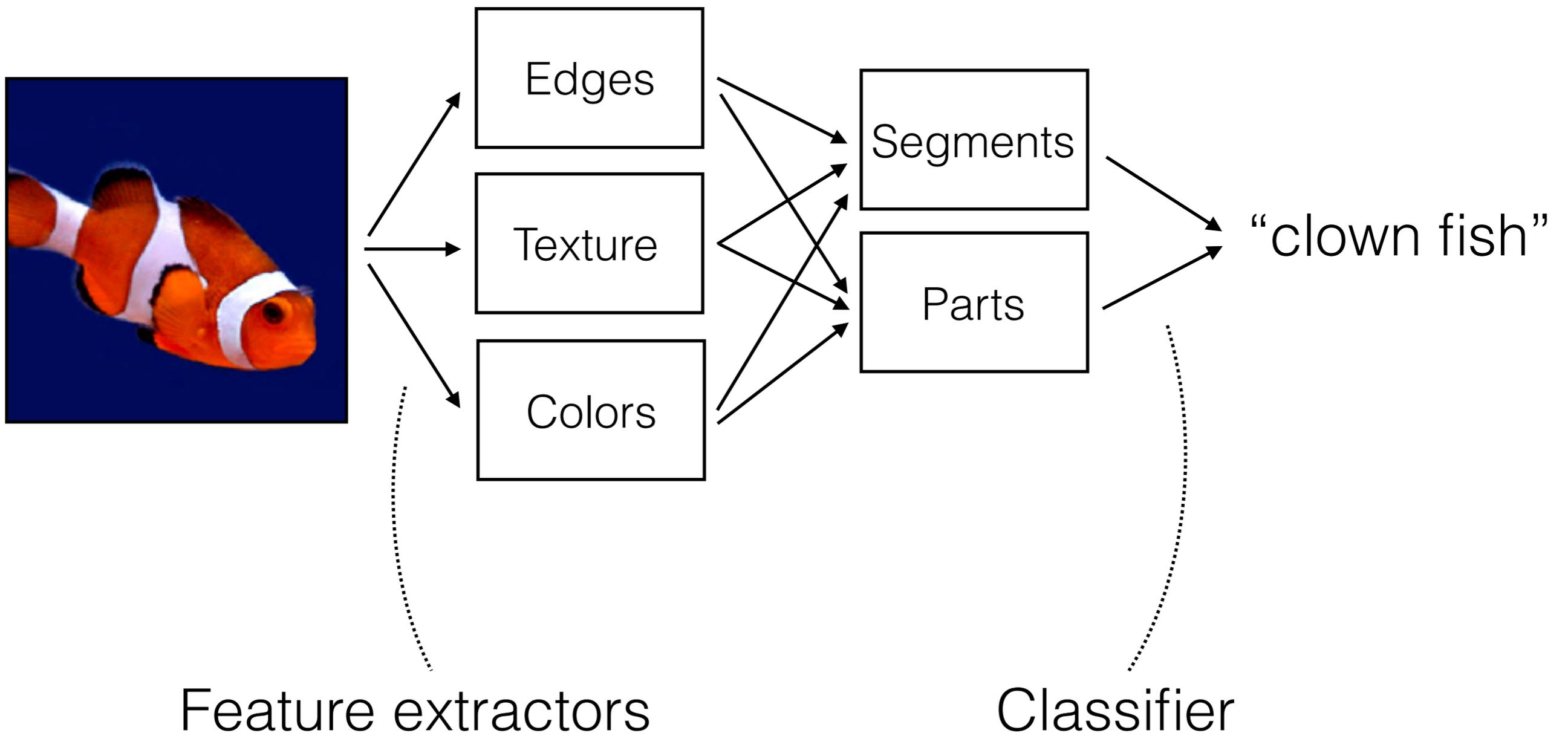


Basic Idea



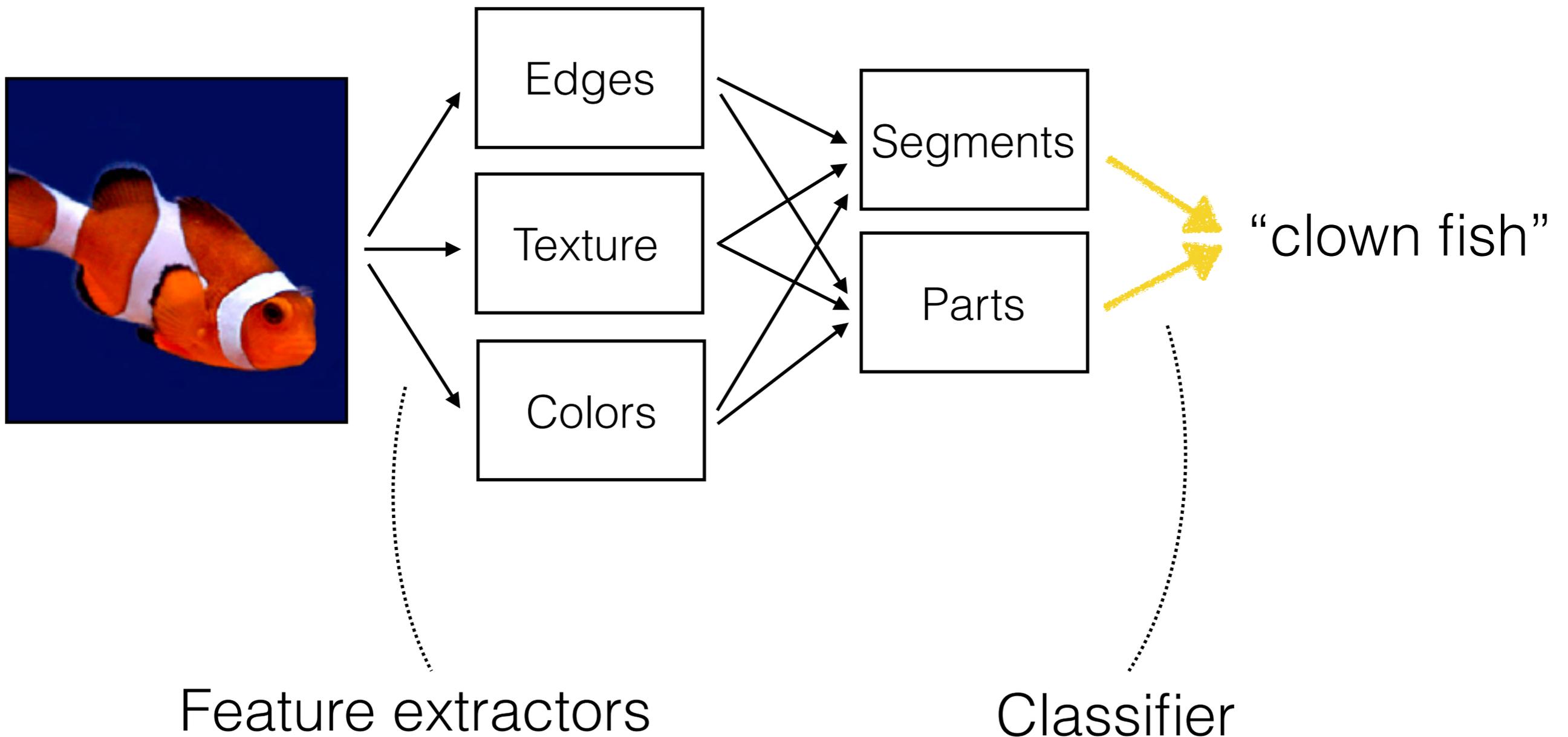
“clown fish”

Object Recognition



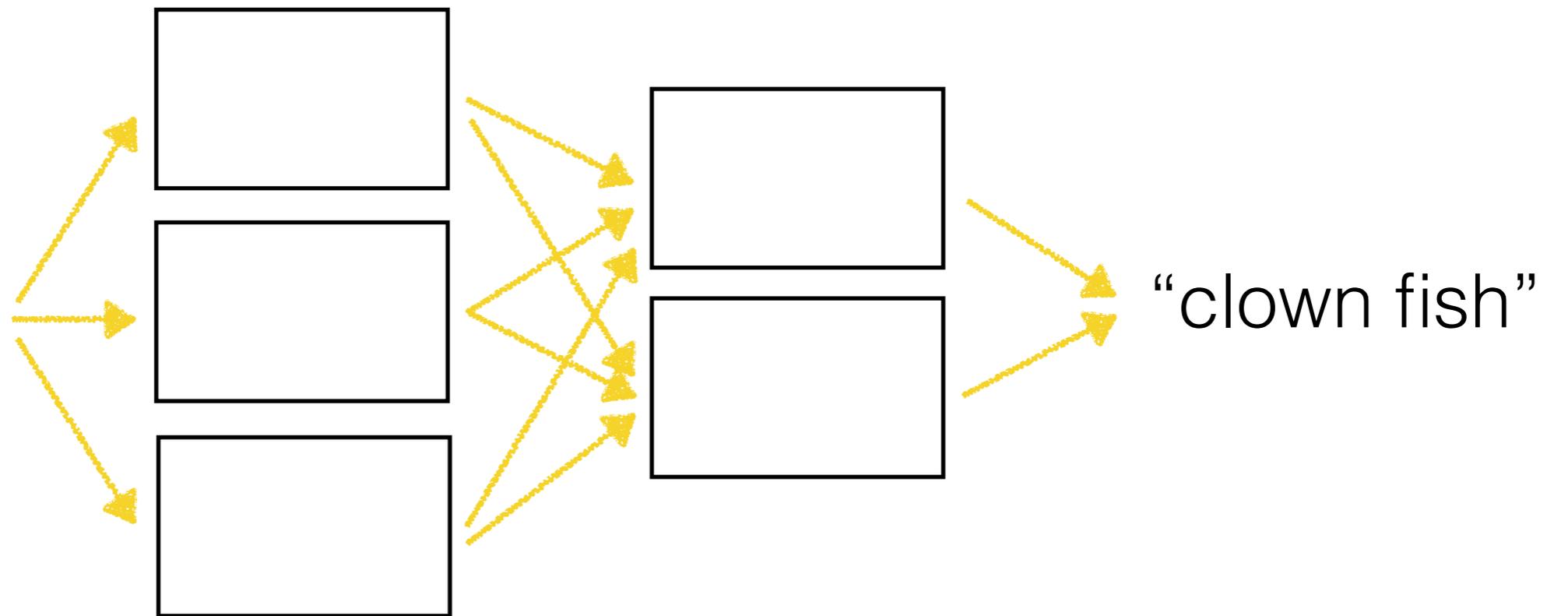
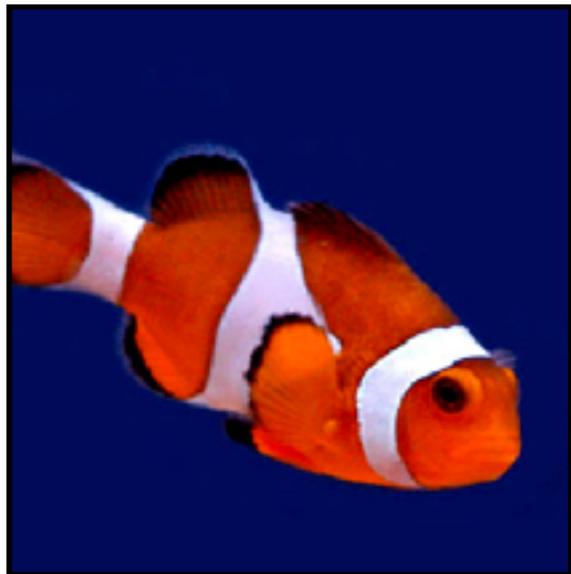
Object Recognition

Learned



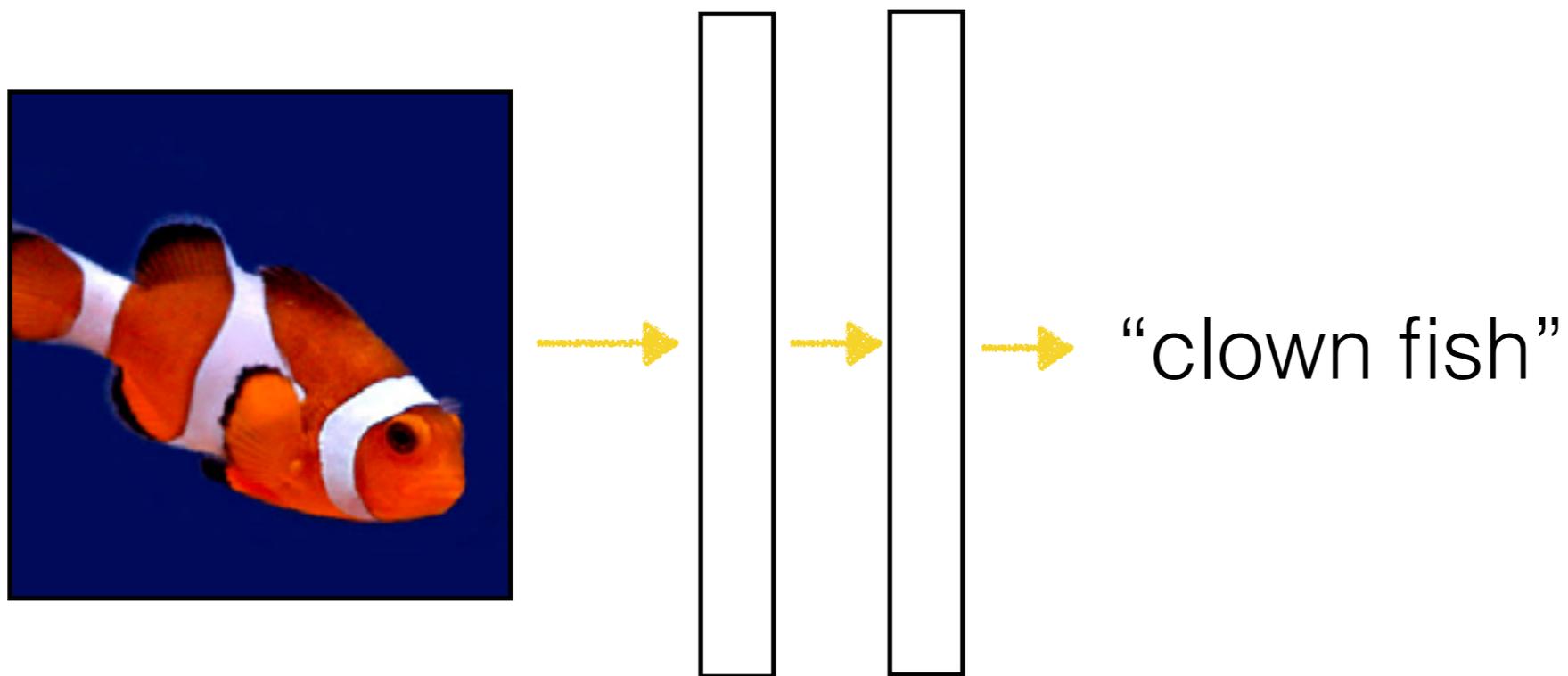
Neural Network

Learned



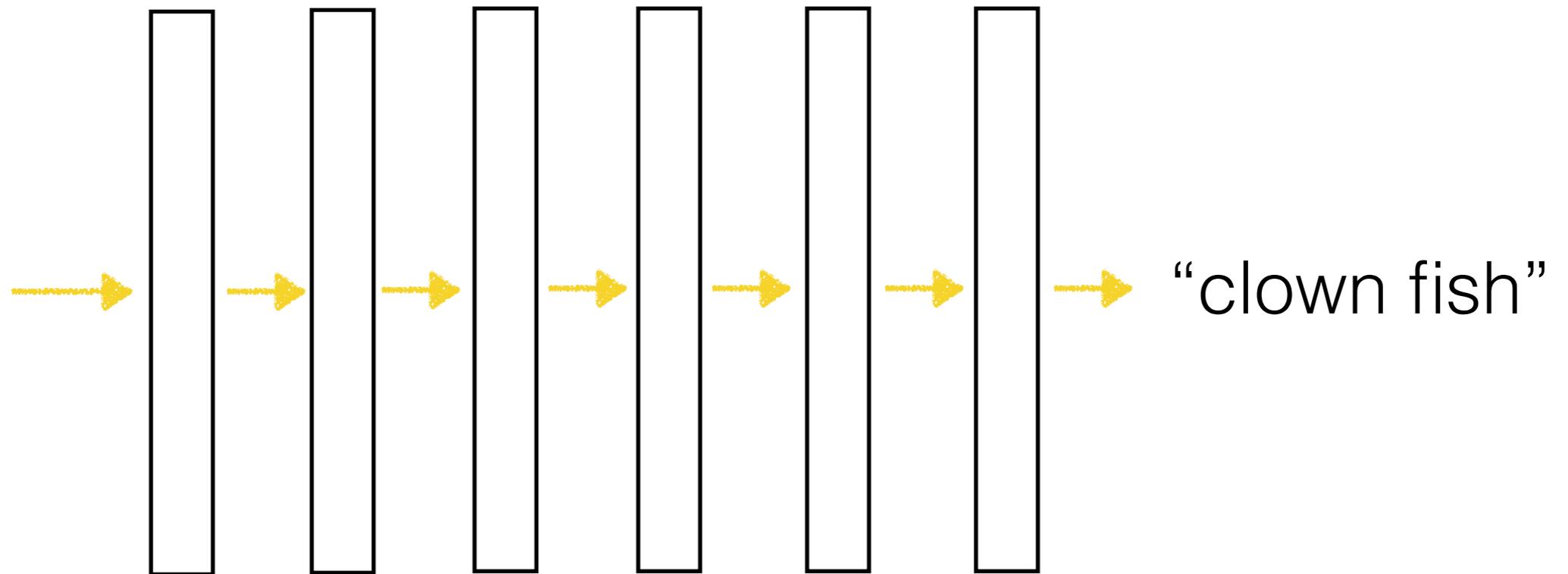
Neural Network

Learned

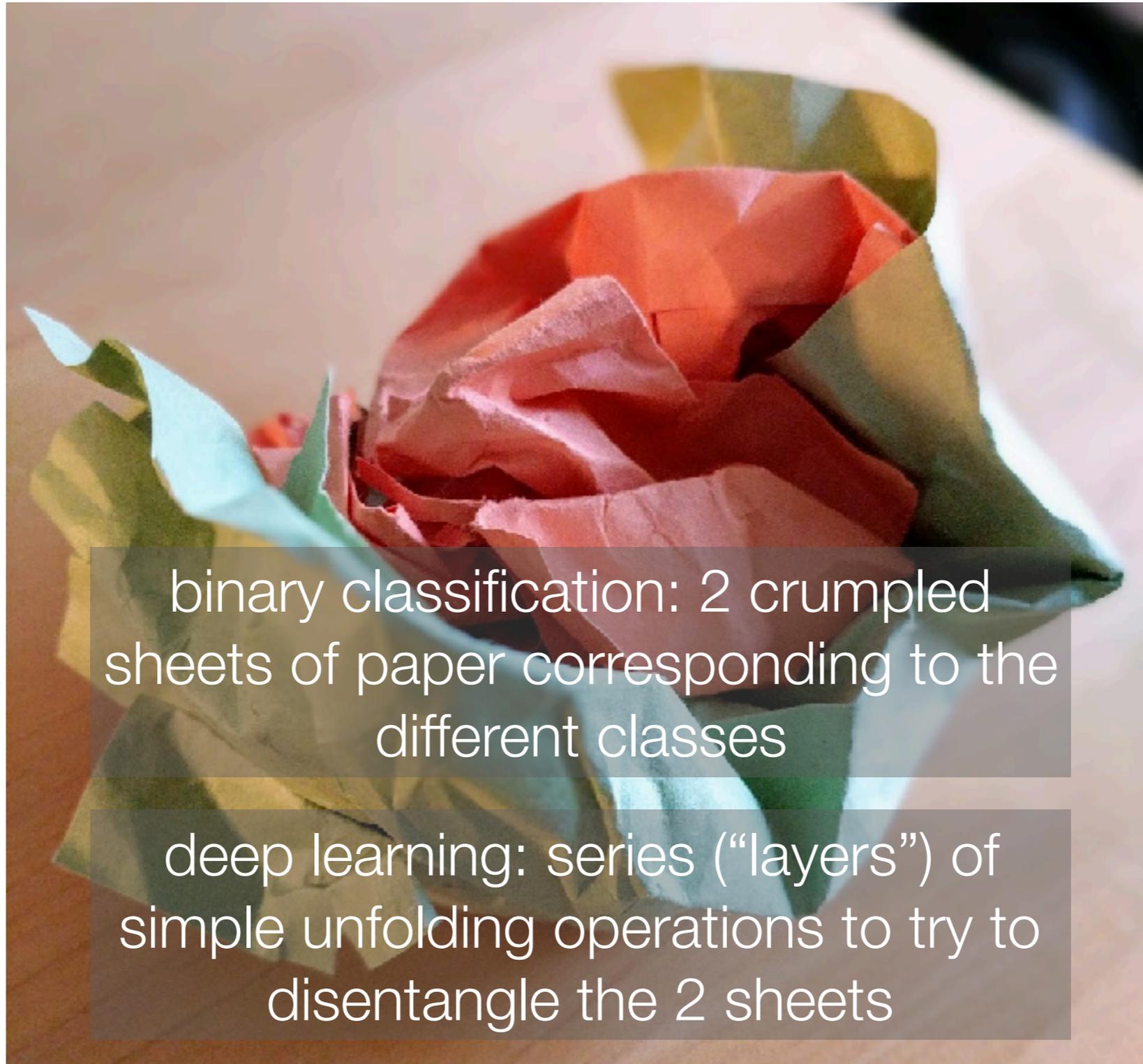


Deep Neural Network

Learned



Crumpled Paper Analogy



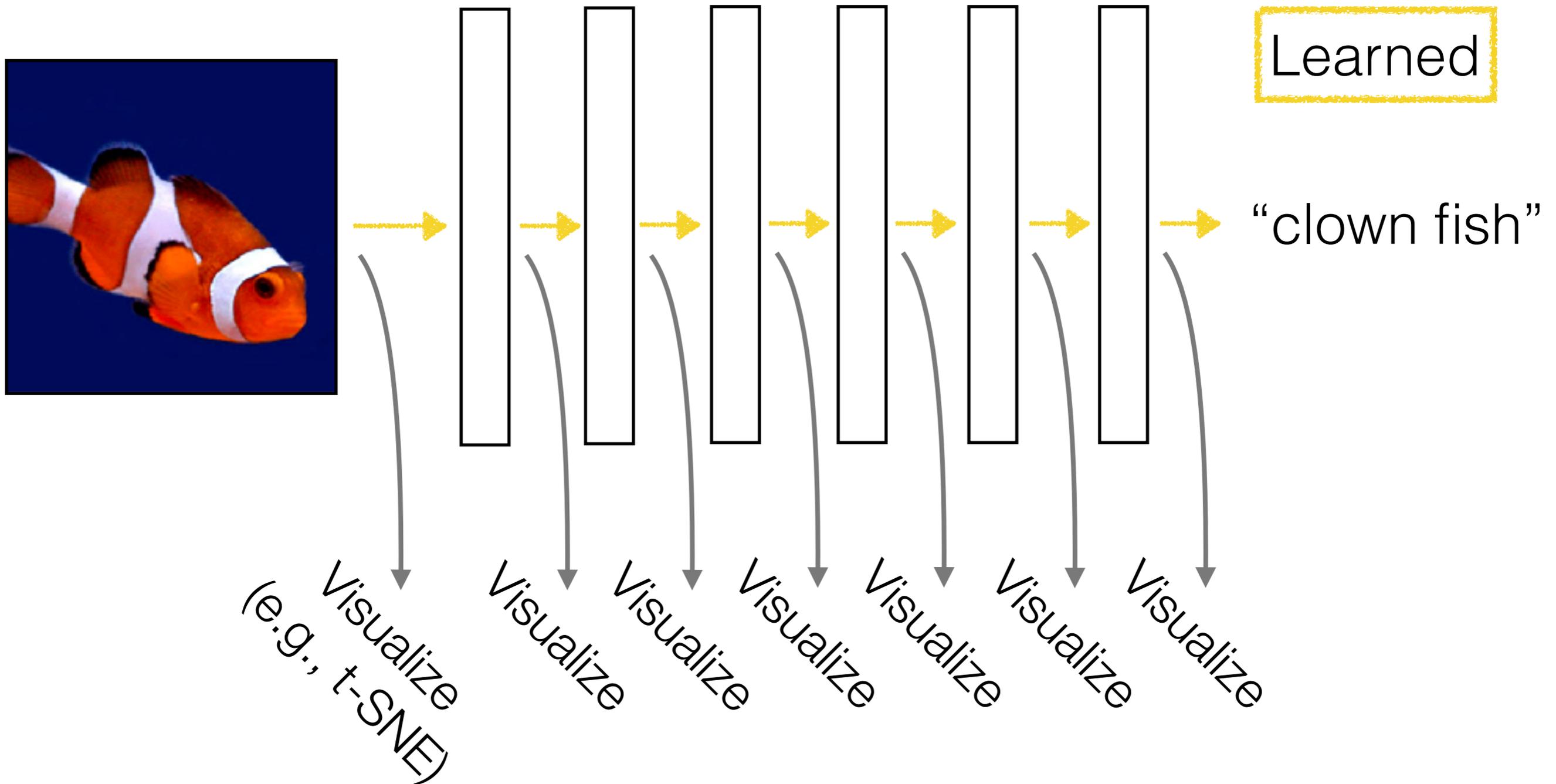
binary classification: 2 crumpled sheets of paper corresponding to the different classes

deep learning: series (“layers”) of simple unfolding operations to try to disentangle the 2 sheets

Analogy: Francois Chollet, photo: George Chen

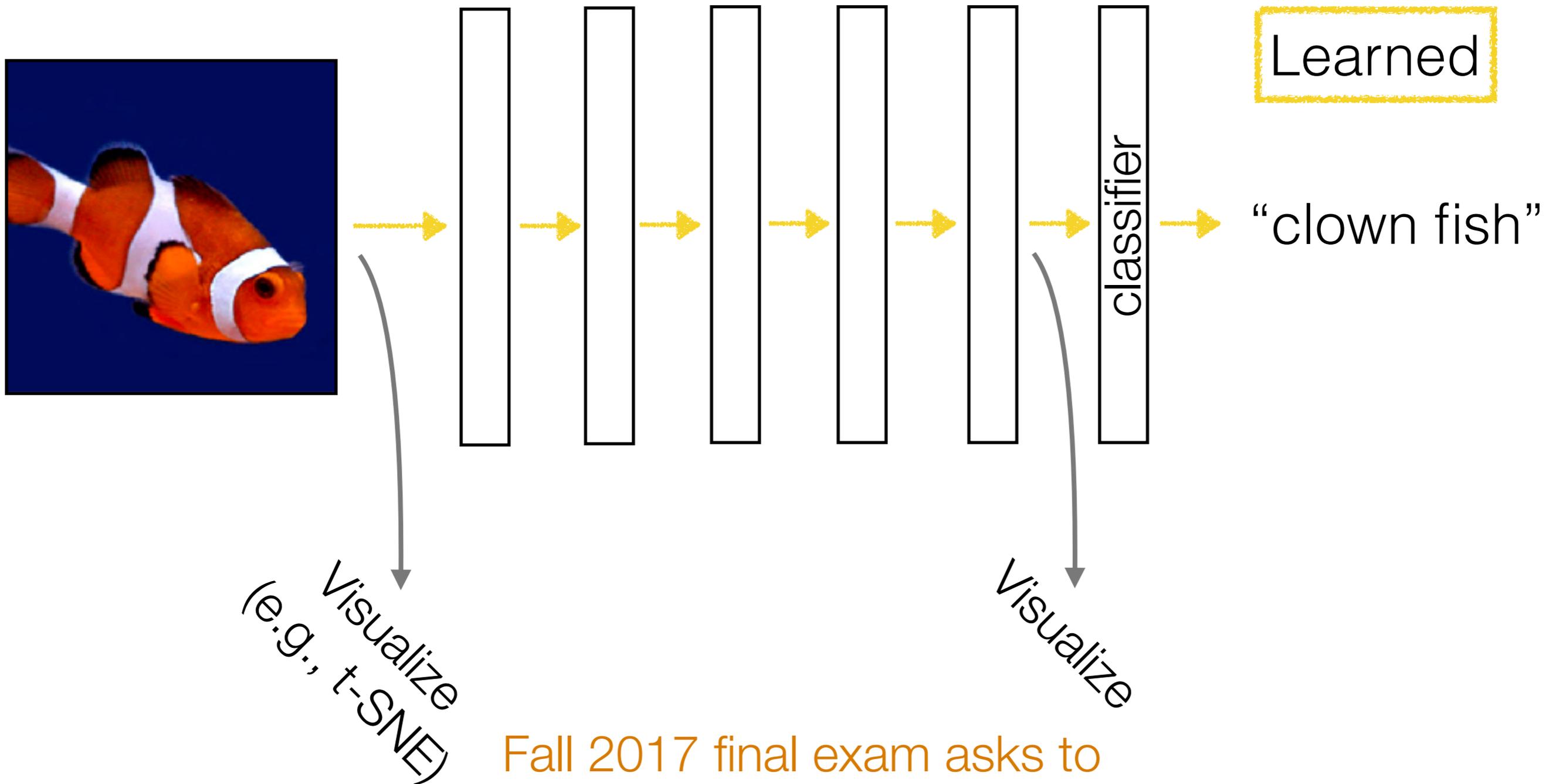
Representation Learning

Each layer's output is *another way we could represent the input data*



Representation Learning

Each layer's output is *another way we could represent the input data*



Fall 2017 final exam asks to compare these 2 visualizations

Why Does Deep Learning Work?

Actually the ideas behind deep learning are old (~1980's)

- Big data



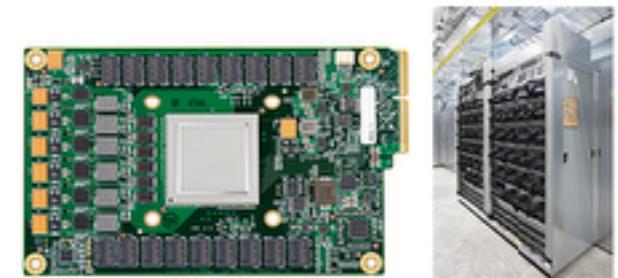
- Better hardware



CPU's
& Moore's law



GPU's



TPU's

- Better algorithms

Structure Present in Data Matters

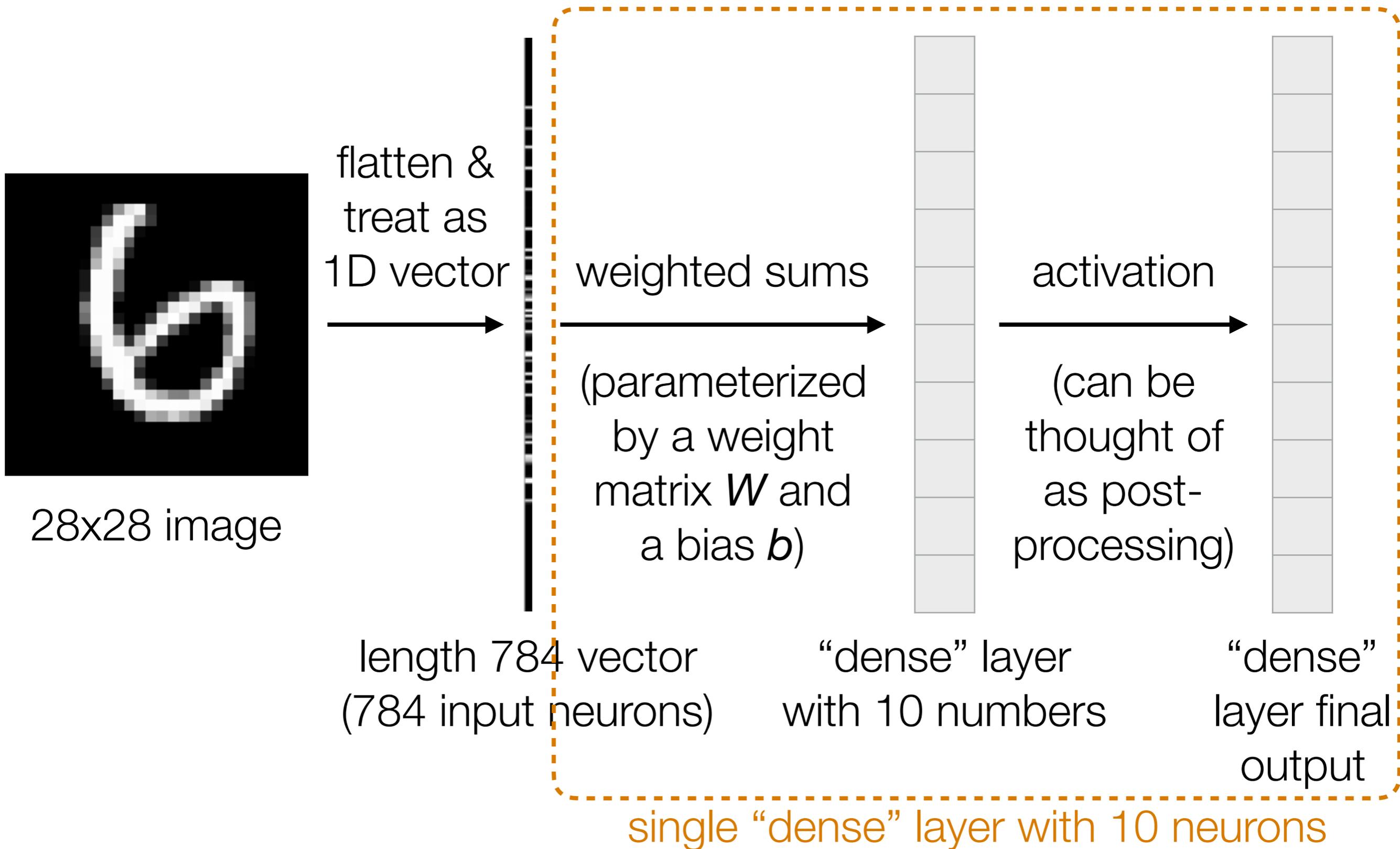
Neural nets aren't doing black magic

- **Image analysis:** convolutional neural networks (convnets) neatly incorporates basic image processing structure
- **Time series analysis:** recurrent neural networks (RNNs) incorporates ability to remember and forget things over time
 - Note: text is a time series
 - Note: video is a time series

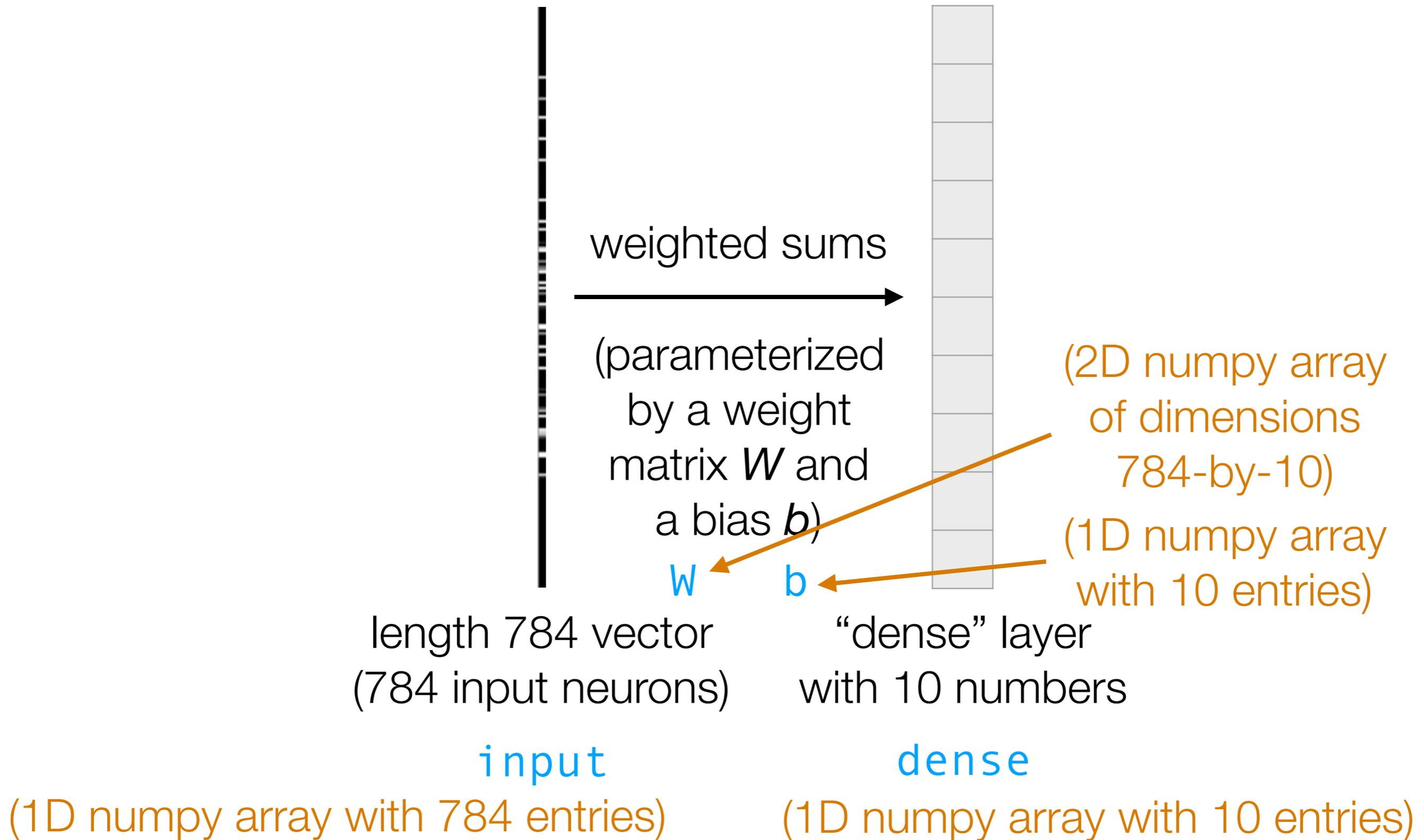
Handwritten Digit Recognition Example

Walkthrough of building a 1-layer and then a 2-layer neural net

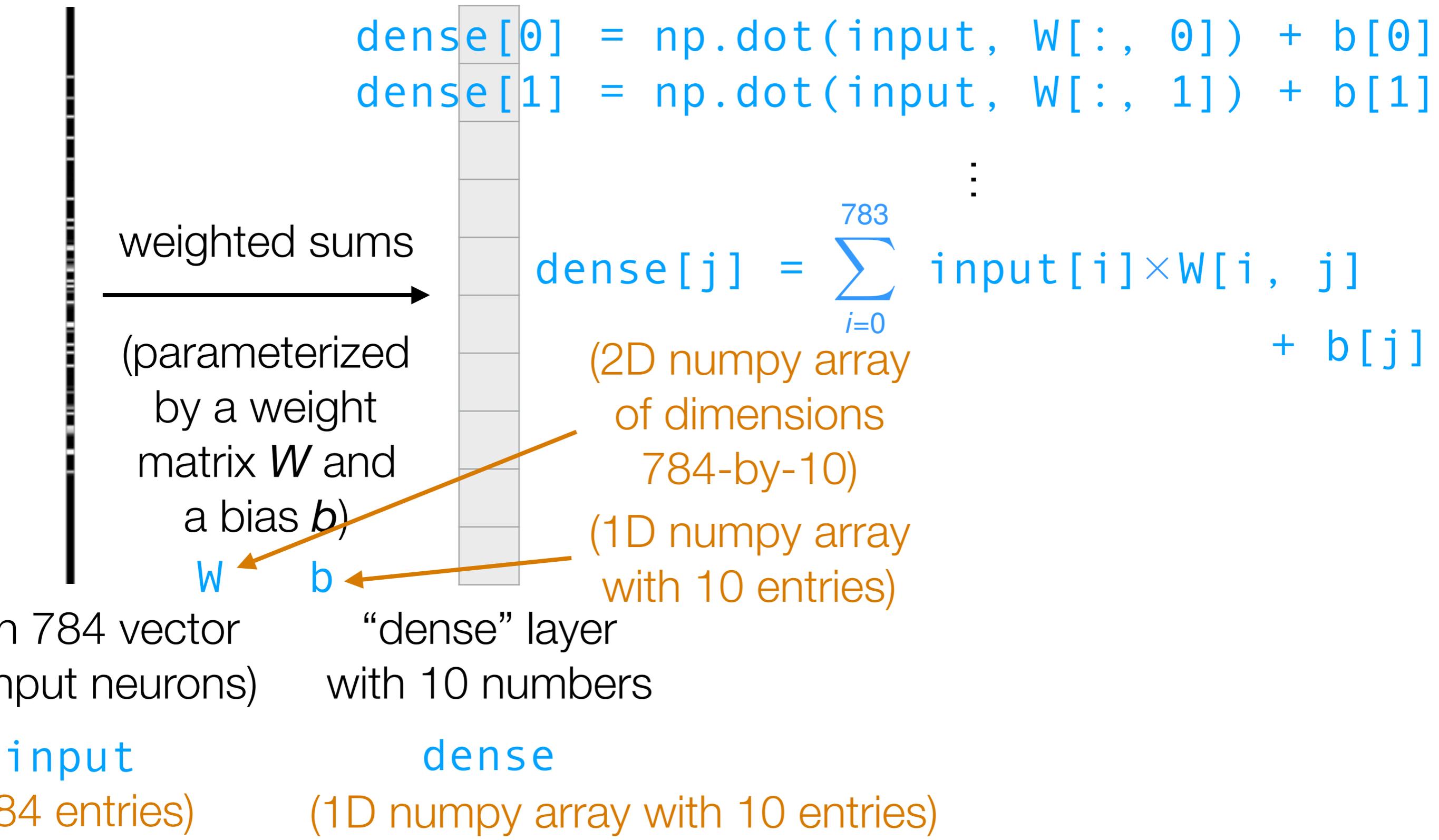
Handwritten Digit Recognition



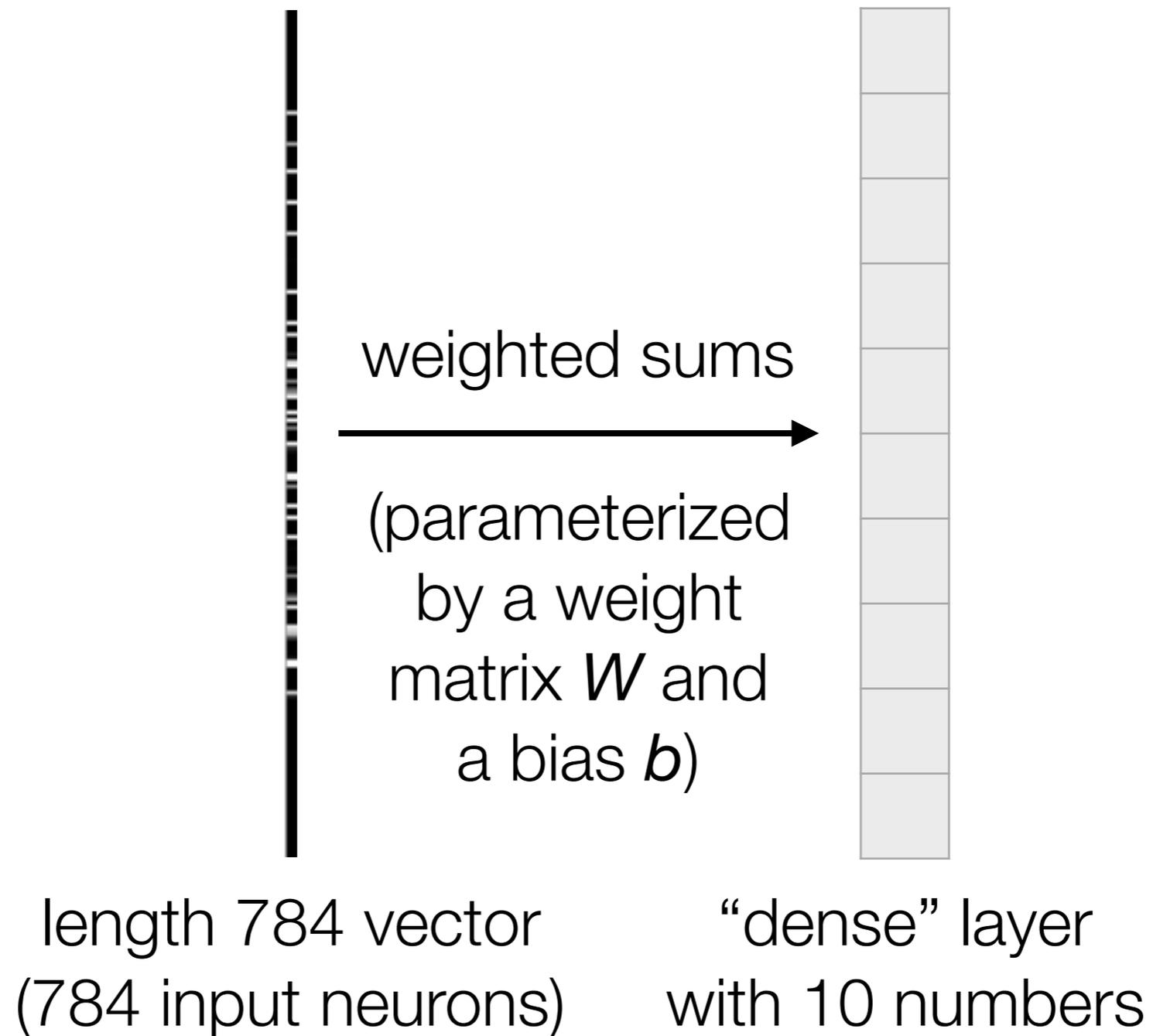
Handwritten Digit Recognition



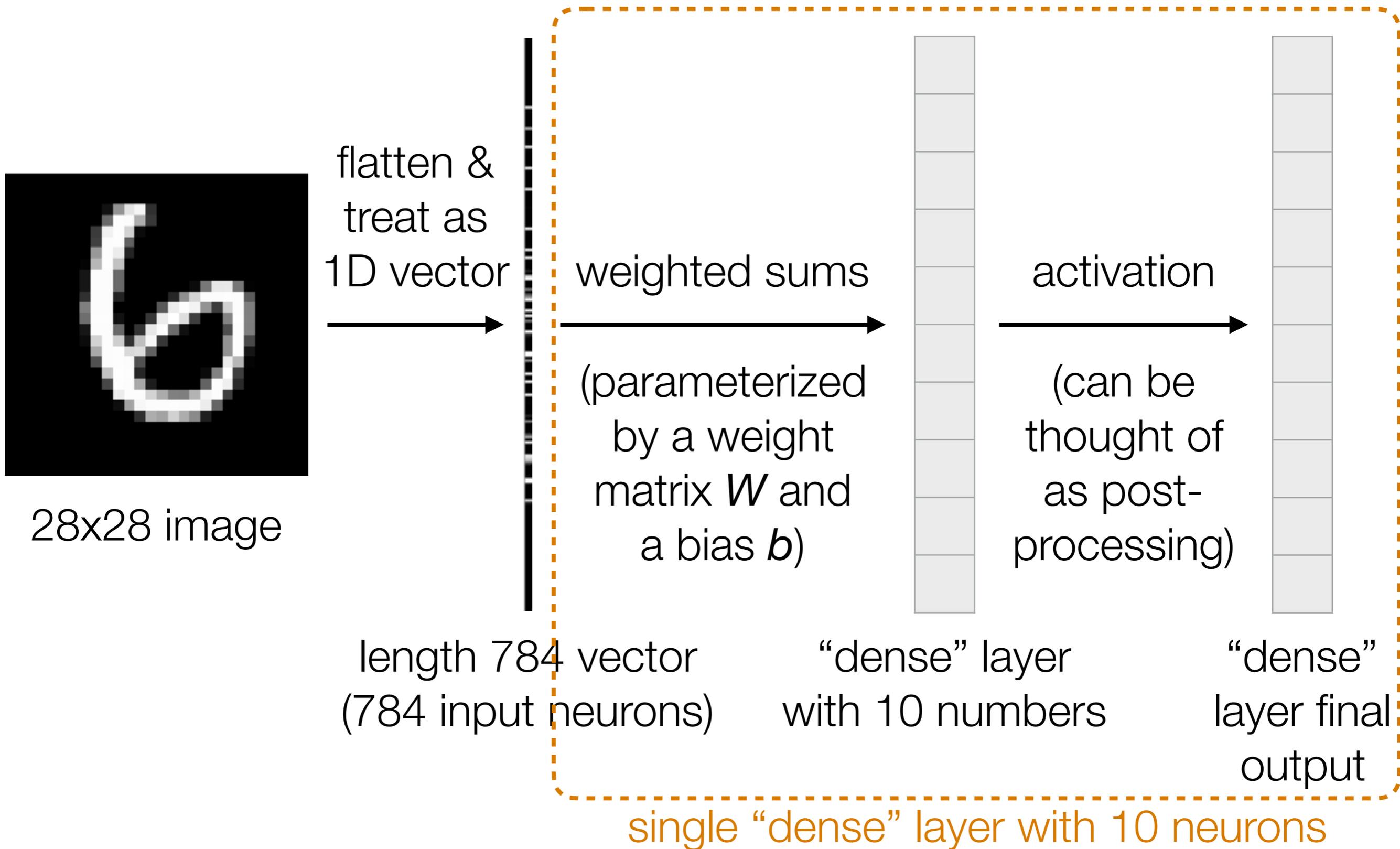
Handwritten Digit Recognition



Handwritten Digit Recognition



Handwritten Digit Recognition



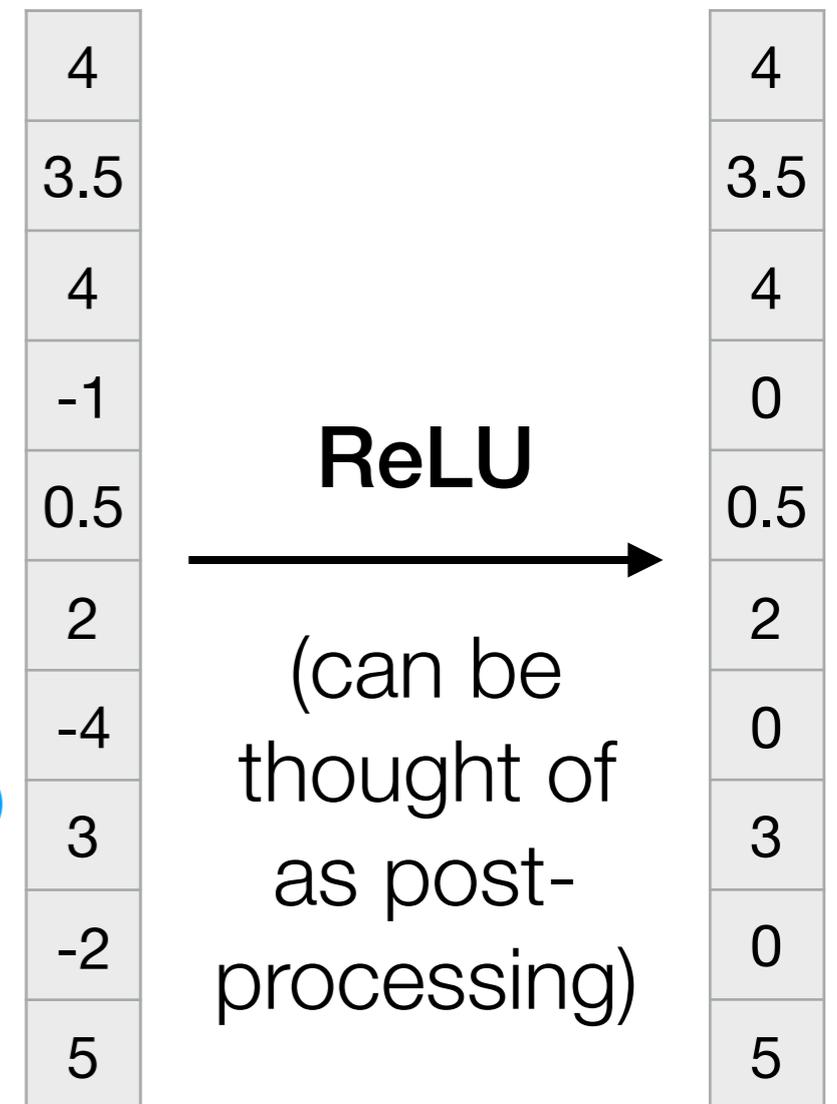
Handwritten Digit Recognition

Many different activation functions possible

Example: **Rectified linear unit (ReLU)**

zeros out entries that are negative

```
dense_final = np.maximum(0, dense)
```



“dense” layer
with 10 numbers

`dense`

“dense”
layer final
output

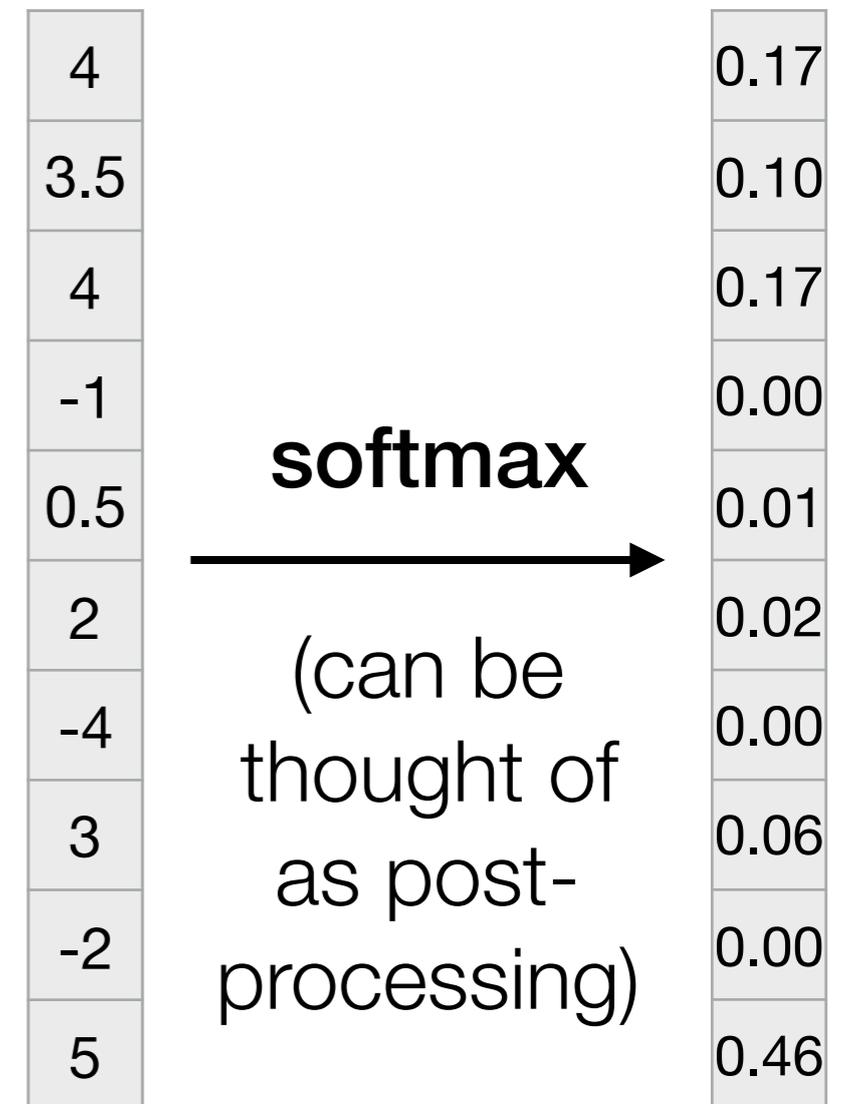
`dense_final`

Handwritten Digit Recognition

Many different activation functions possible

Example: **softmax** turns the entries in the dense layer (prior to activation) into a probability distribution (using the “softmax” transformation)

```
dense_exp = np.exp(dense)
dense_exp /= np.sum(dense_exp)
dense_final = dense_exp
```



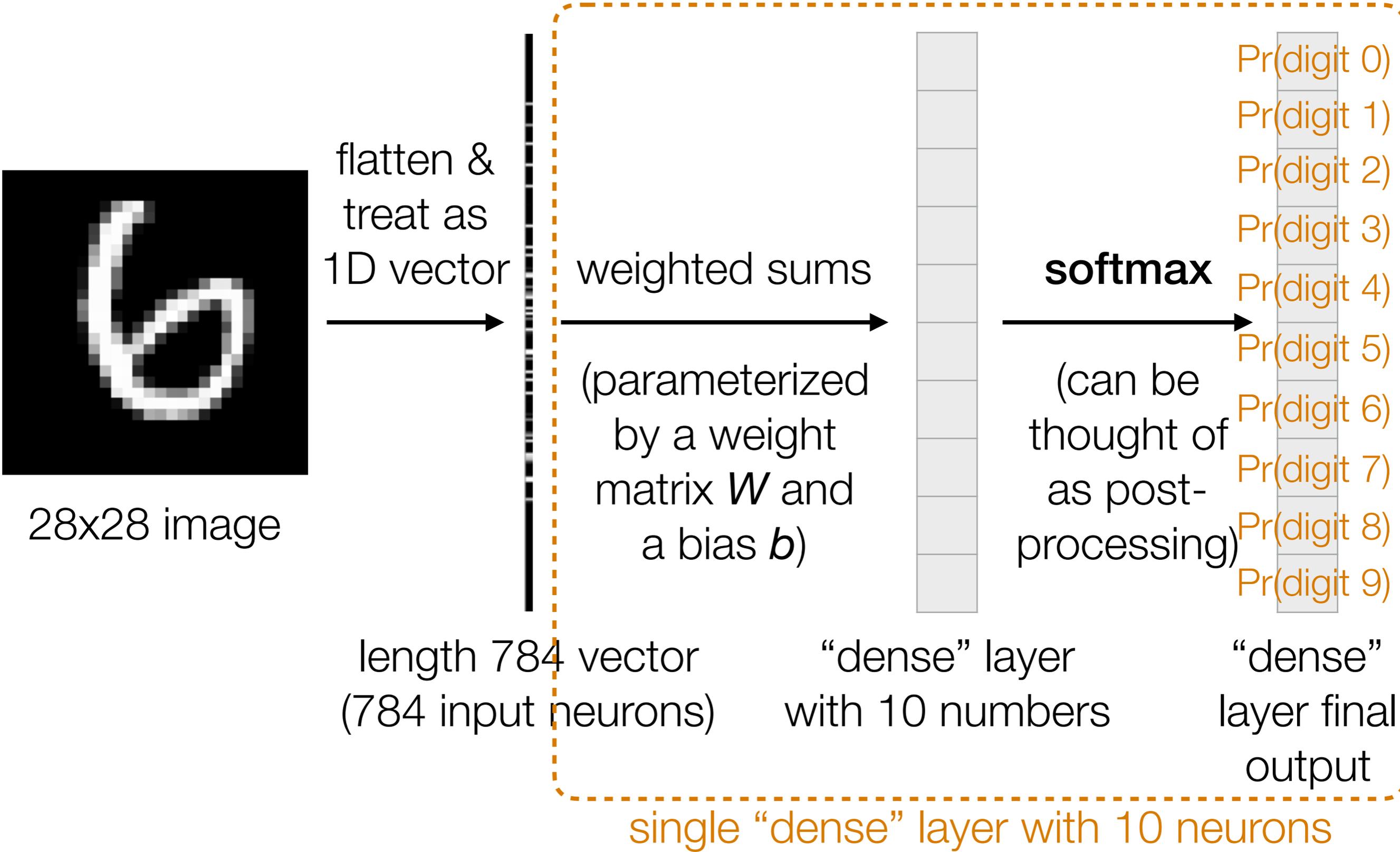
“dense” layer
with 10 numbers

`dense`

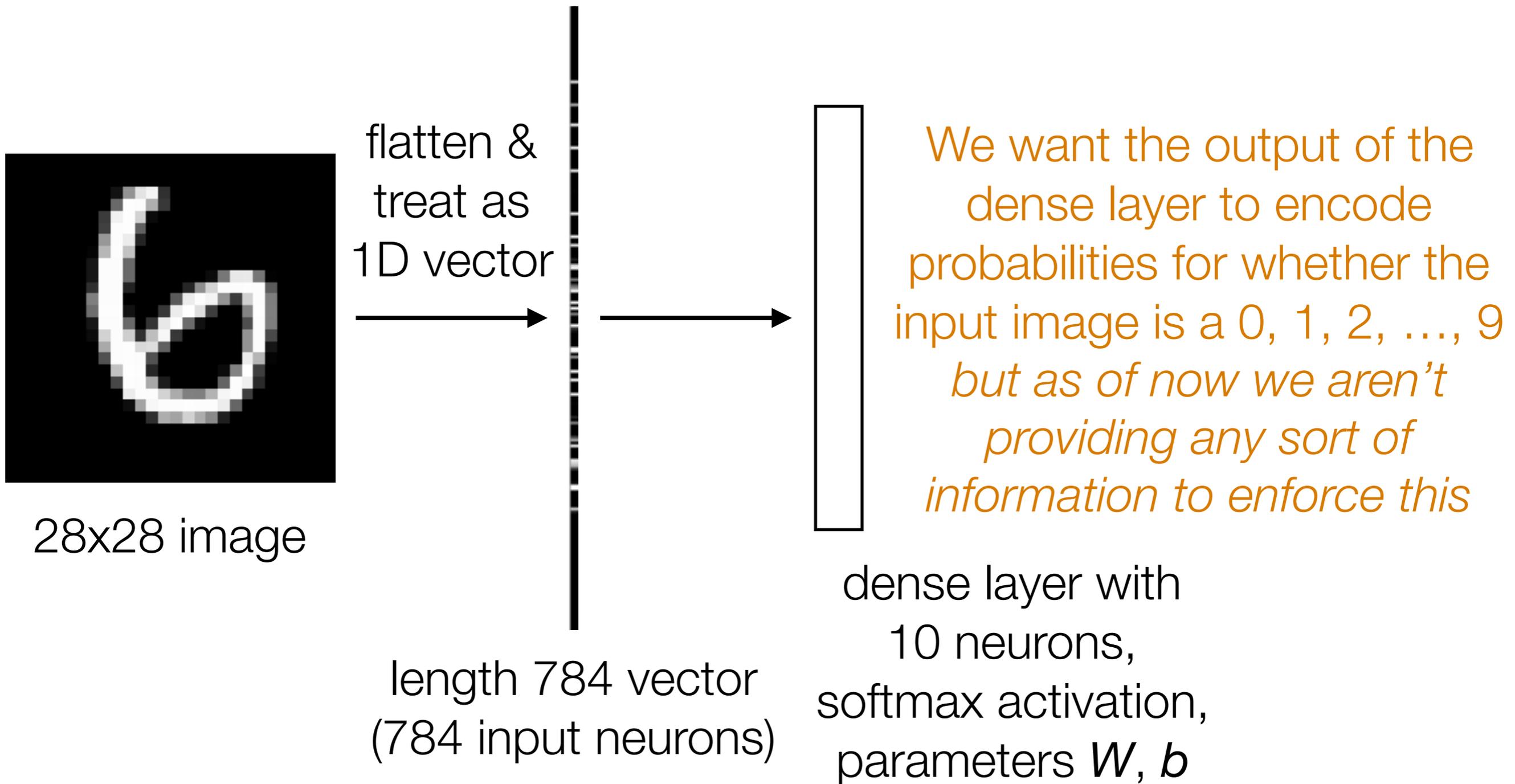
“dense”
layer final
output

`dense_final`

Handwritten Digit Recognition



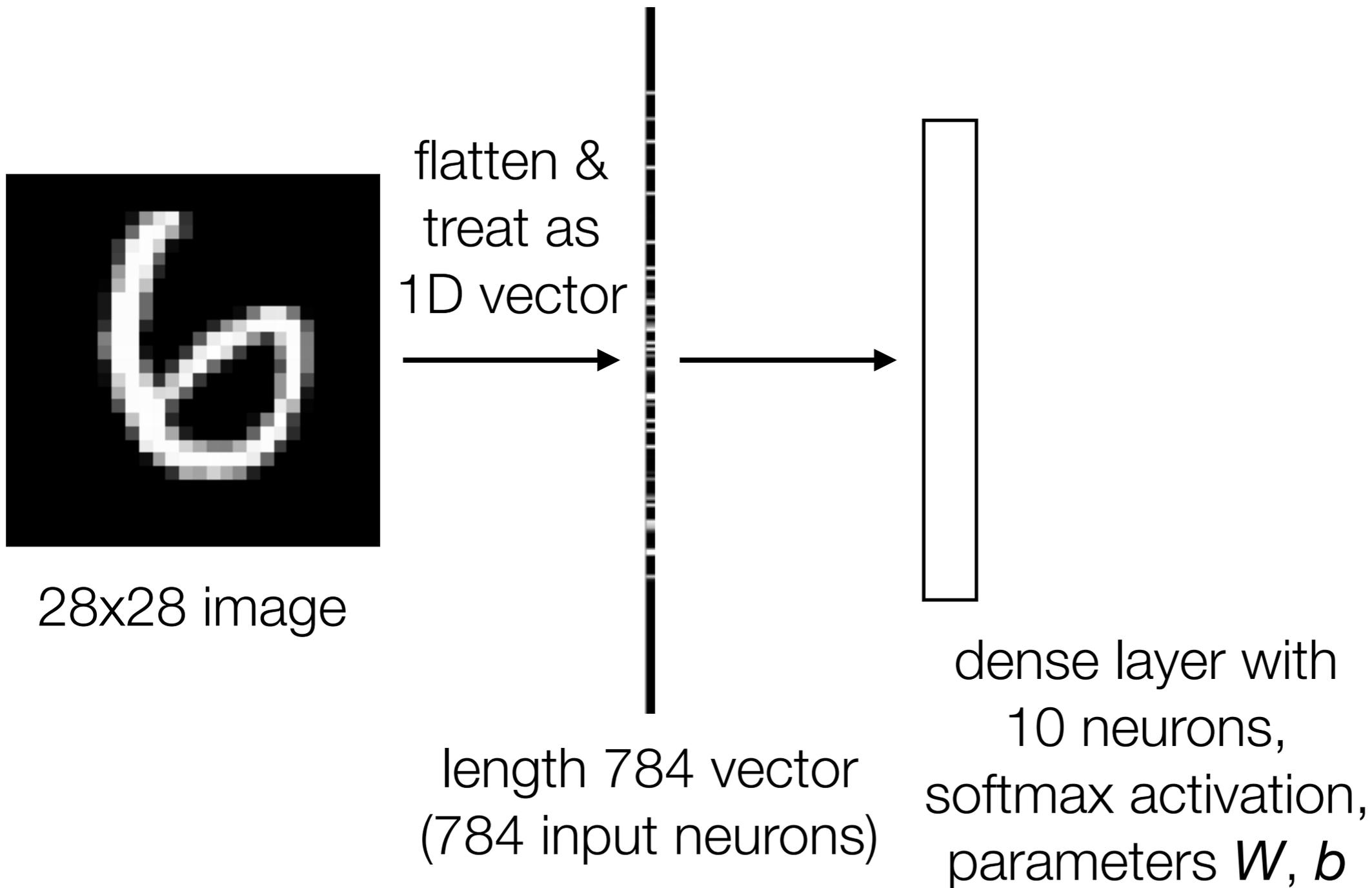
Handwritten Digit Recognition



Handwritten Digit Recognition

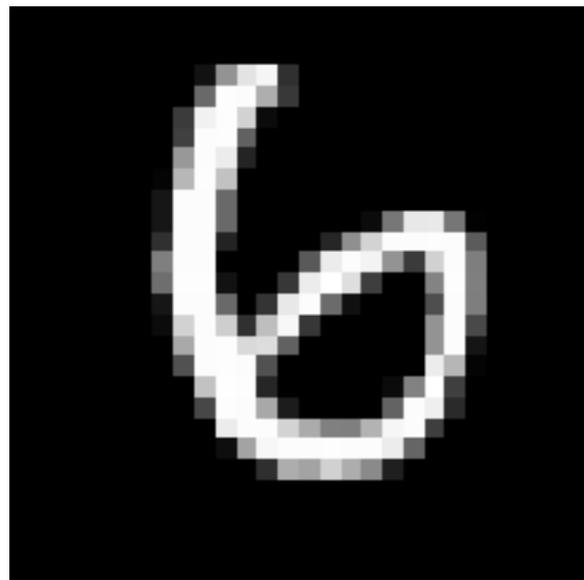
Demo part 1

Handwritten Digit Recognition



Handwritten Digit Recognition

Training label: 6



28x28 image

flatten &
treat as
1D vector



Loss/"error"



Error is averaged across training examples

Popular loss function for classification (> 2 classes):
categorical cross entropy

dense layer with 10 neurons, softmax activation, parameters W, b

$$\log \frac{1}{\text{Pr}(\text{digit } 6)}$$

Learning this neural net means learning W and b

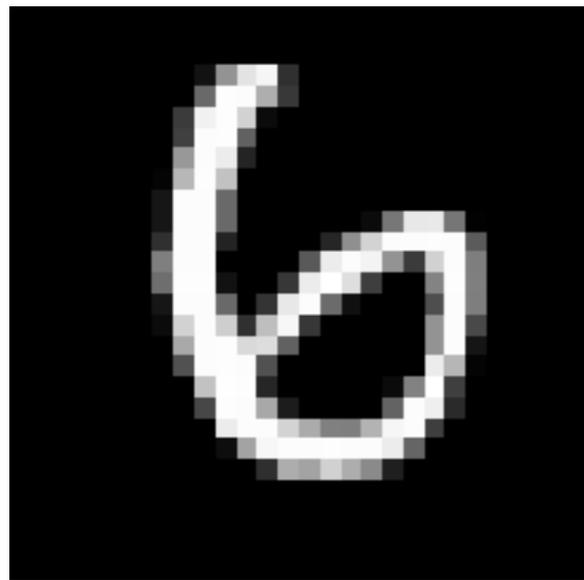
length 784 vector (784 input neurons)

Handwritten Digit Recognition

Demo part 2

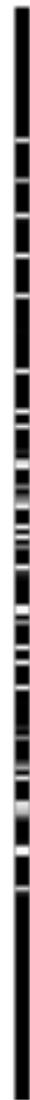
Handwritten Digit Recognition

Training label: 6



28x28 image

flatten &
treat as
1D vector



Loss/"error"



Error is averaged across training examples

Popular loss function for classification (> 2 classes):
categorical cross entropy

dense layer with 10 neurons, softmax activation, parameters W, b

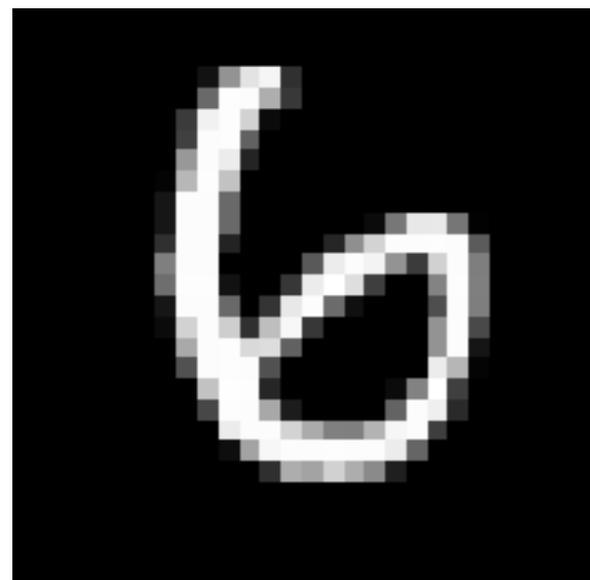
$$\log \frac{1}{\text{Pr}(\text{digit } 6)}$$

Learning this neural net means learning W and b

length 784 vector (784 input neurons)

Handwritten Digit Recognition

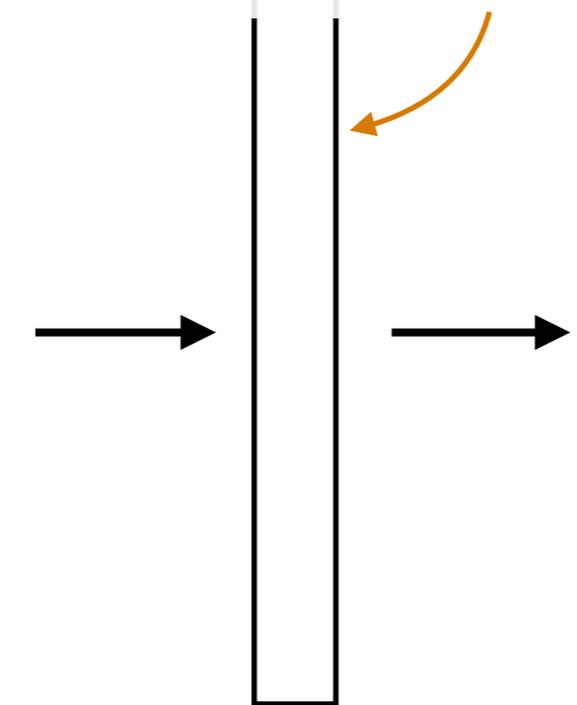
Training label: 6



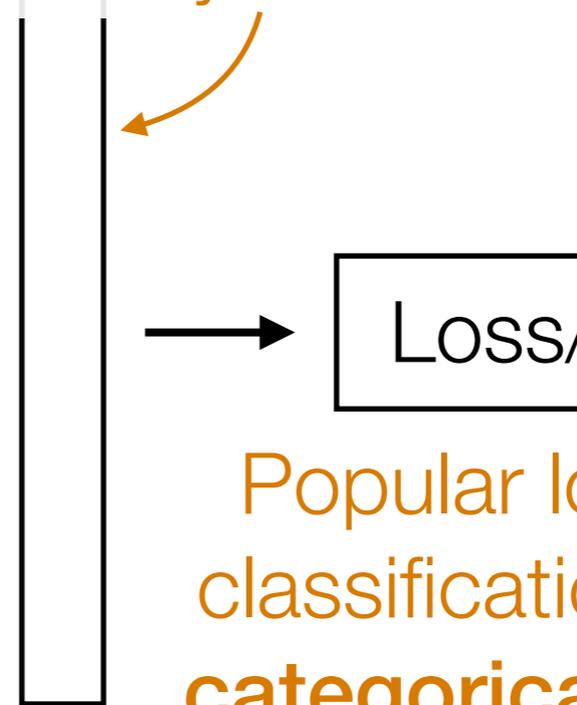
28x28 image

length 784 vector
(784 input neurons)

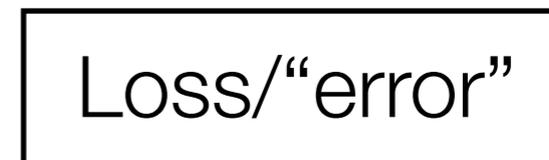
Learning this neural net means learning parameters of both dense layers!



dense layer with 512 neurons, ReLU activation



dense layer with 10 neurons, softmax activation



Popular loss function for classification (> 2 classes): **categorical cross entropy**

$$\log \frac{1}{\text{Pr}(\text{digit } 6)}$$

Error is averaged across training examples

error

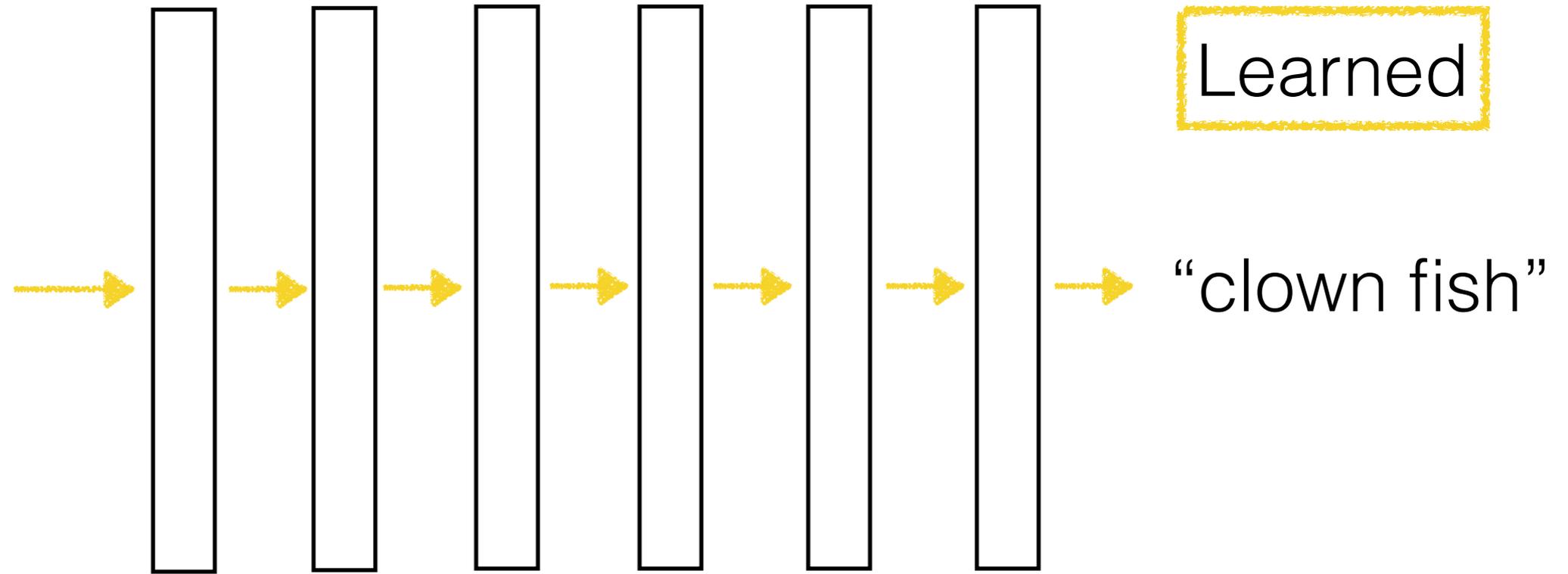
Handwritten Digit Recognition

Demo part 3

Architecting Neural Nets

- Increasing number of layers (depth) makes neural net more complex
 - Can approximate more functions
 - More parameters needed
 - More training data may be needed
- Designing neural net architectures is a bit of an art
 - How to select the number of neurons for intermediate layers?
 - Very common in practice: modify existing architectures that are known to work well (e.g., VGG-16 for computer vision/image processing)

Deep Learning



- Inspired by biological neural nets *but otherwise not the same at all* (biological neural nets do *not* work like deep nets)
- Learns a layered representation
 - Tries to get rid of manual feature engineering
 - Need to design constraints for what features are learned to account for structure in data (e.g., images, text, ...)