Neural networks

Deep learning - motivation

NEURAL NETWORK

Topics: multilayer neural network

- Could have L hidden layers:
  - layer input activation for $k>0$ $(\mathbf{h}^{(0)}(\mathbf{x}) = \mathbf{x})$

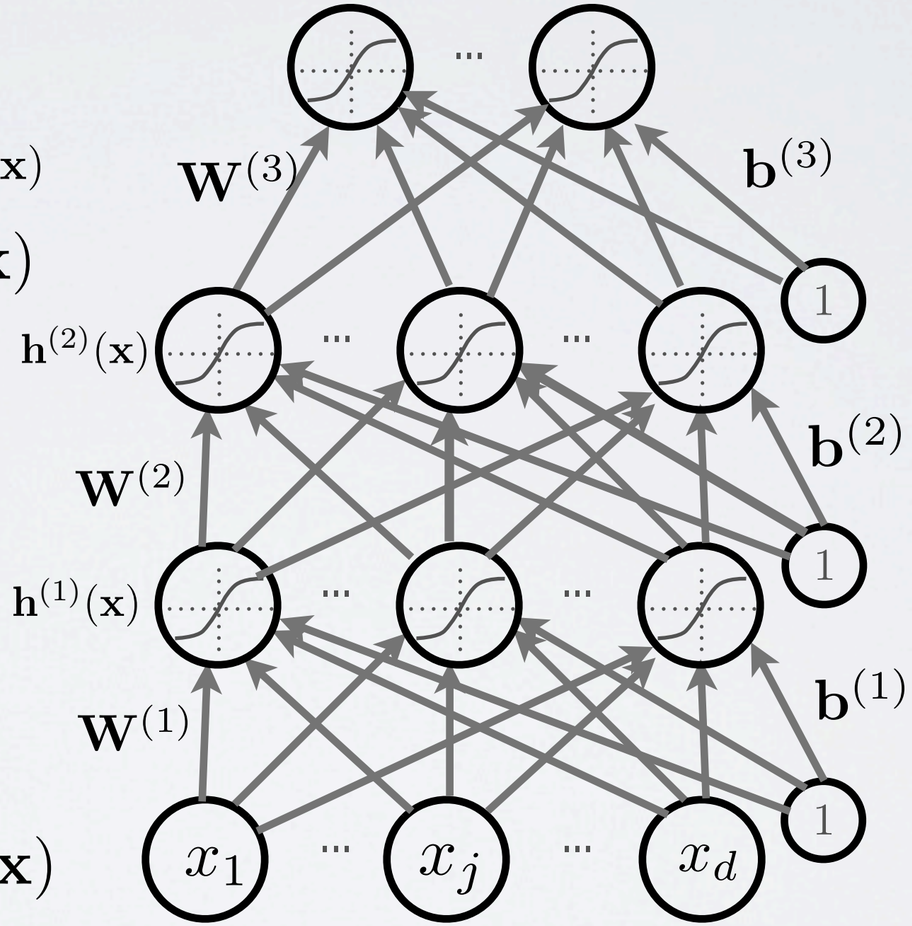
$$\mathbf{a}^{(k)}(\mathbf{x}) = \mathbf{b}^{(k)} + \mathbf{W}^{(k)}\mathbf{h}^{(k-1)}(\mathbf{x})$$

  - hidden layer activation (k from 1 to L):

$$\mathbf{h}^{(k)}(\mathbf{x}) = \mathbf{g}(\mathbf{a}^{(k)}(\mathbf{x}))$$

  - output layer activation ($k=L+1$):

$$\mathbf{h}^{(L+1)}(\mathbf{x}) = \mathbf{o}(\mathbf{a}^{(L+1)}(\mathbf{x})) = \mathbf{f}(\mathbf{x})$$



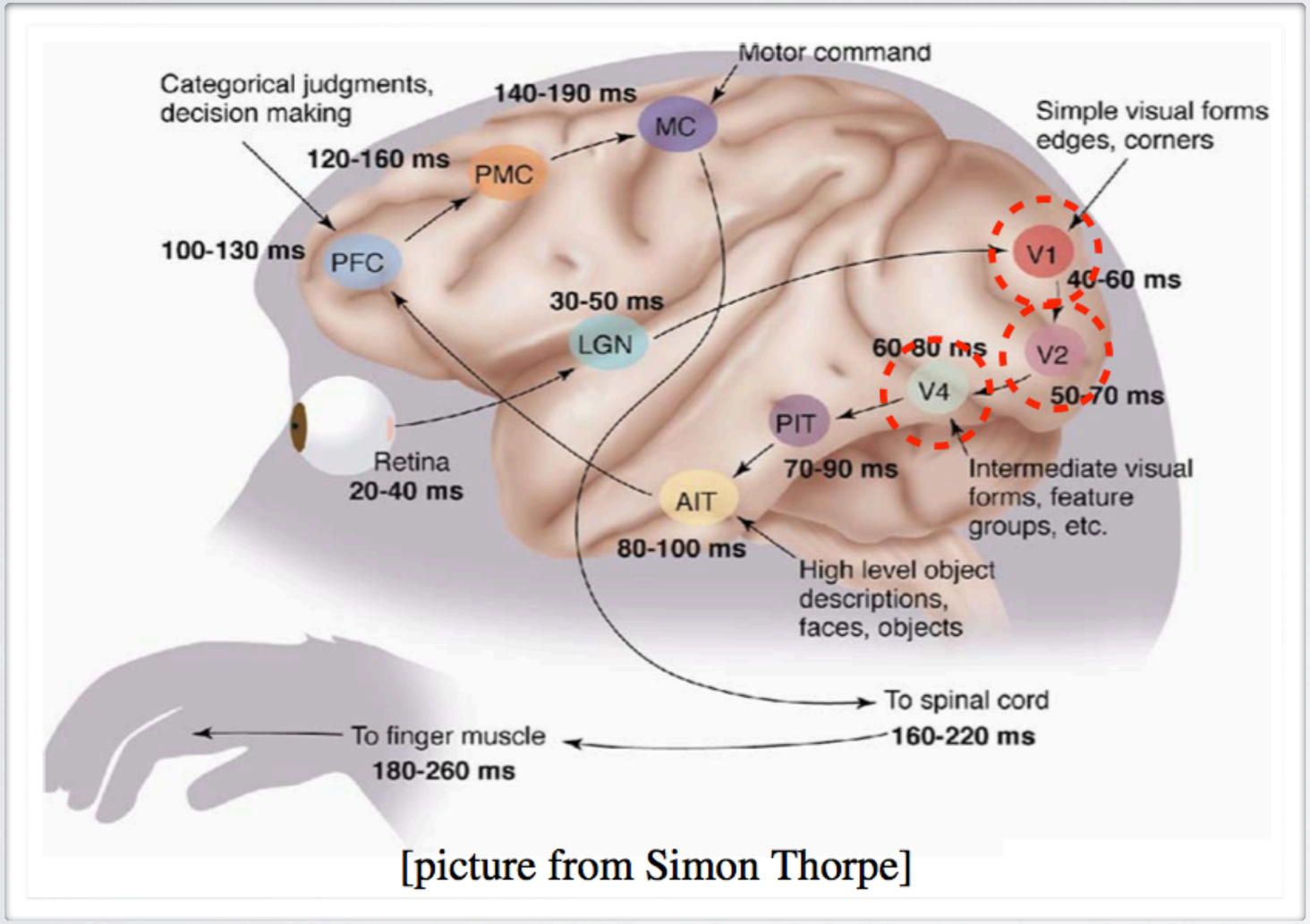
DEEP LEARNING

Topics: deep learning, distributed representation

- Deep learning is research on learning models with multilayer representations
  - multilayer (feed-forward) neural network
  - multilayer graphical model (deep belief network, deep Boltzmann machine)
- Each layer corresponds to a "distributed representation"
  - units in layer are not mutually exclusive
    - each unit is a separate feature of the input
    - two units can be "active" at the same time
  - they do not correspond to a partitioning (clustering) of the inputs
    - in clustering, an input can only belong to a single cluster

DEEP LEARNING

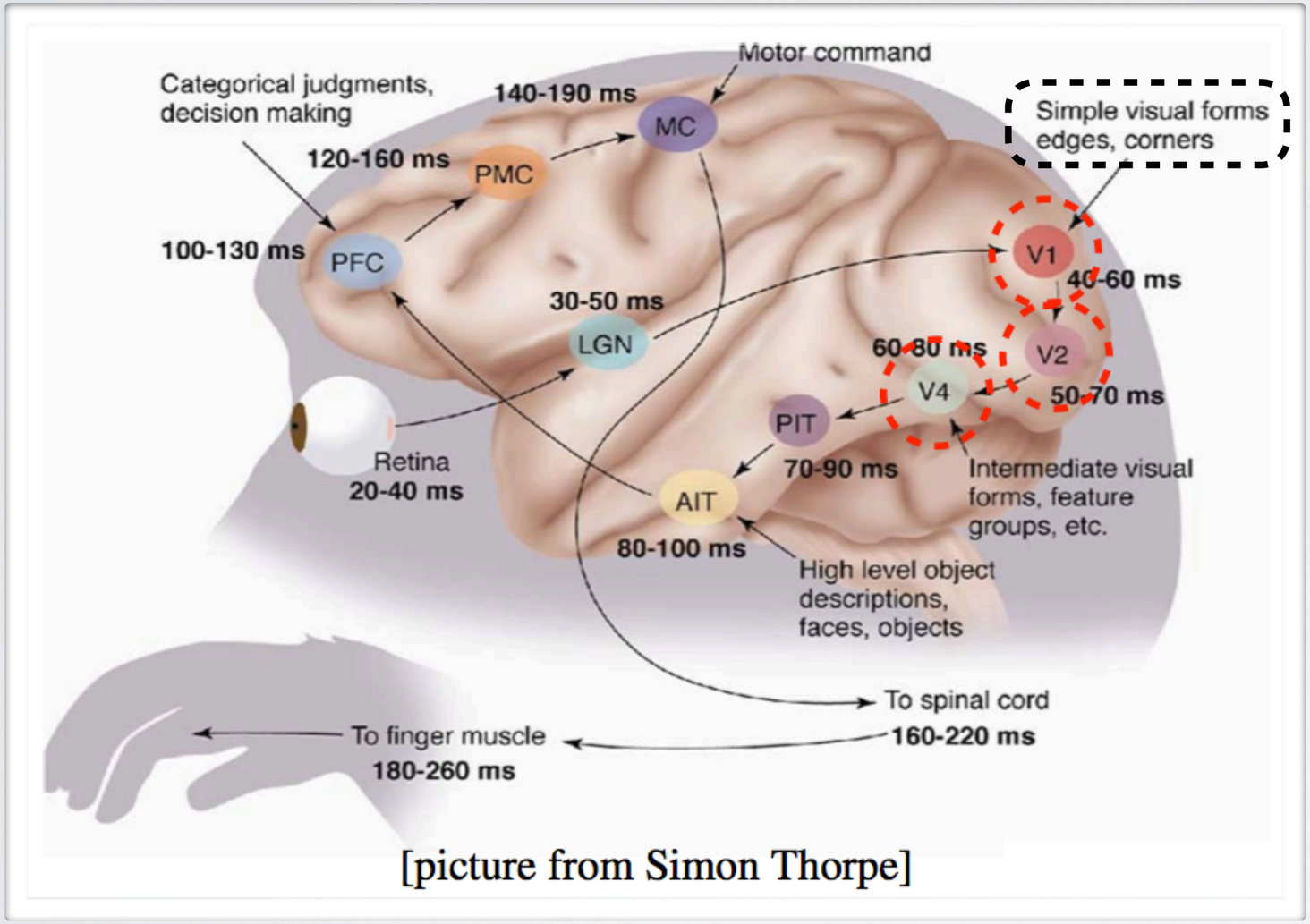
Topics: inspiration from visual cortex



[picture from Simon Thorpe]

DEEP LEARNING

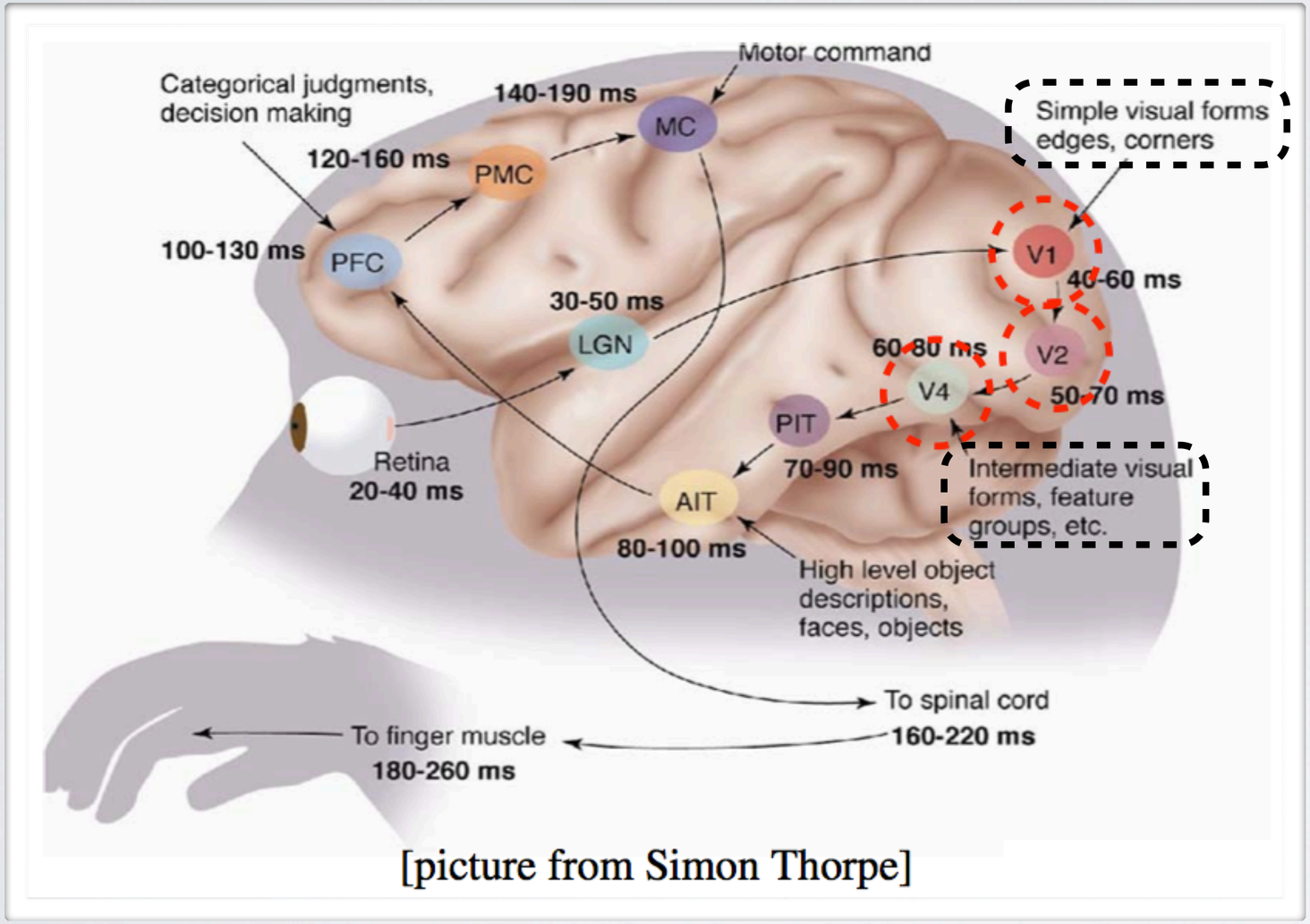
Topics: inspiration from visual cortex



[picture from Simon Thorpe]

DEEP LEARNING

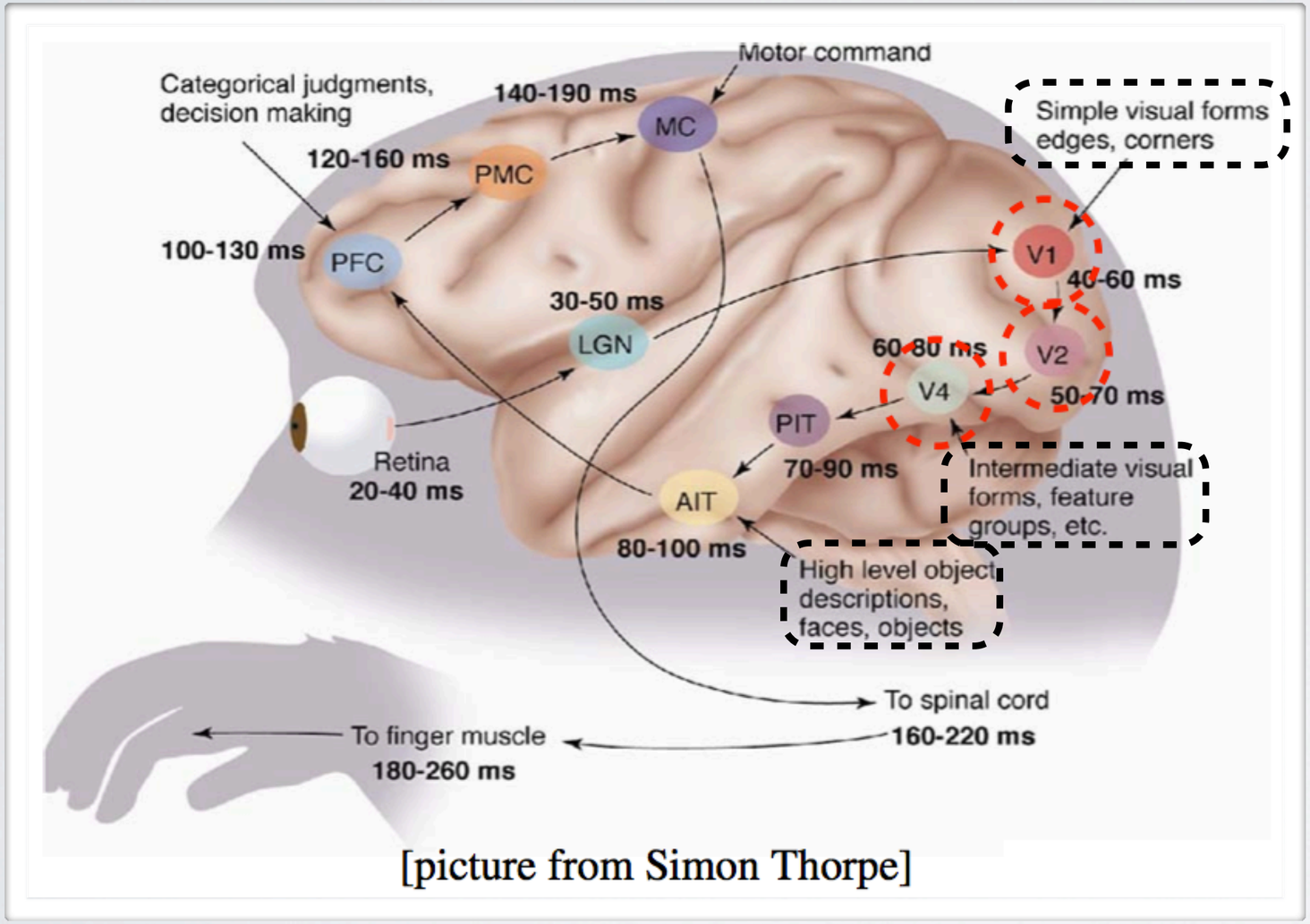
Topics: inspiration from visual cortex



[picture from Simon Thorpe]

DEEP LEARNING

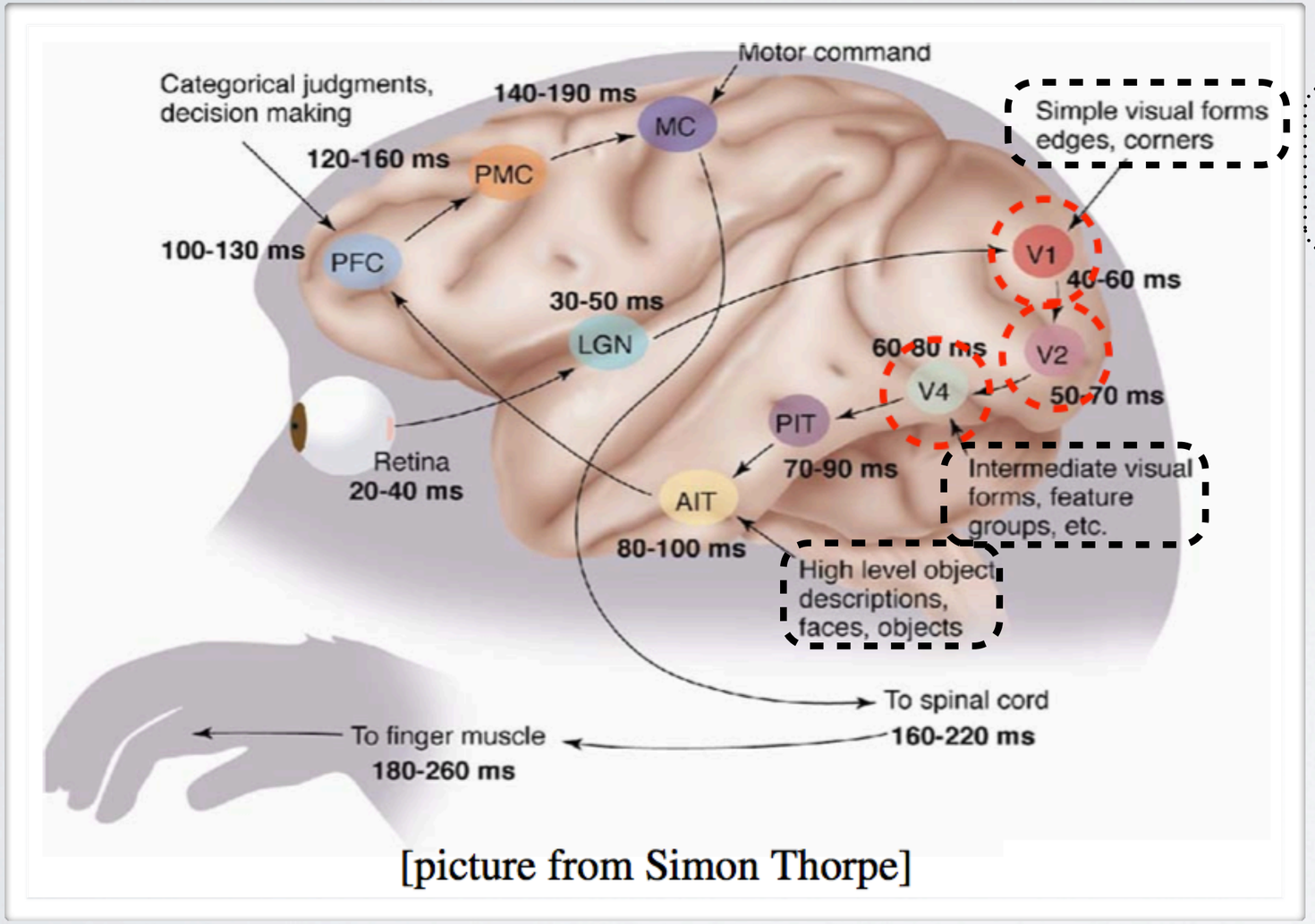
Topics: inspiration from visual cortex



[picture from Simon Thorpe]

DEEP LEARNING

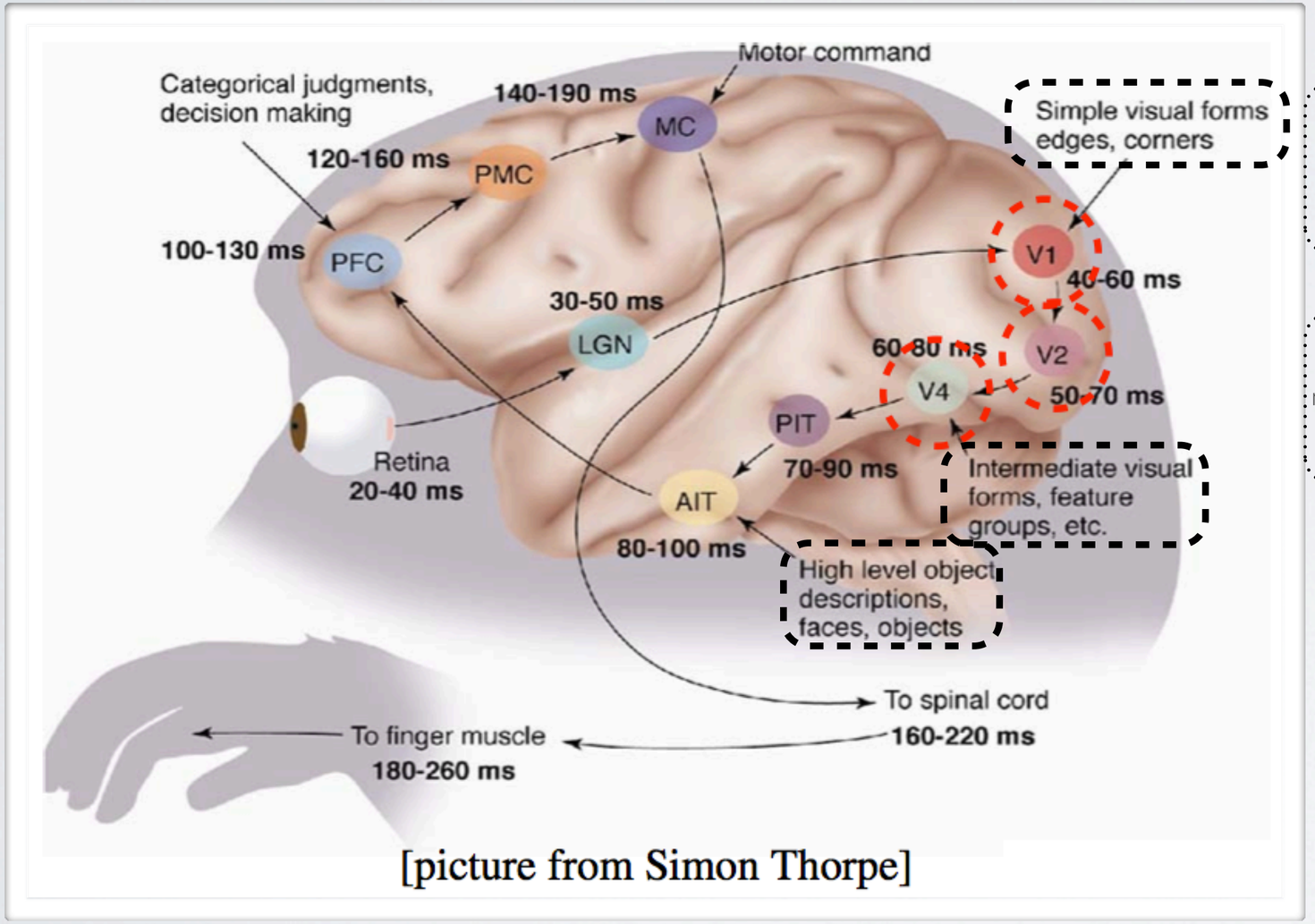
Topics: inspiration from visual cortex

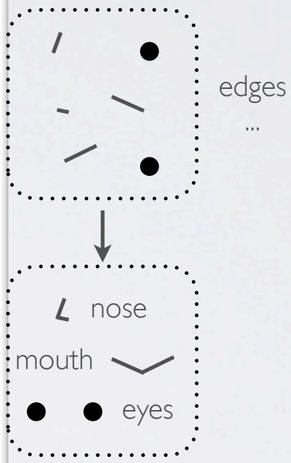


edges
...

DEEP LEARNING

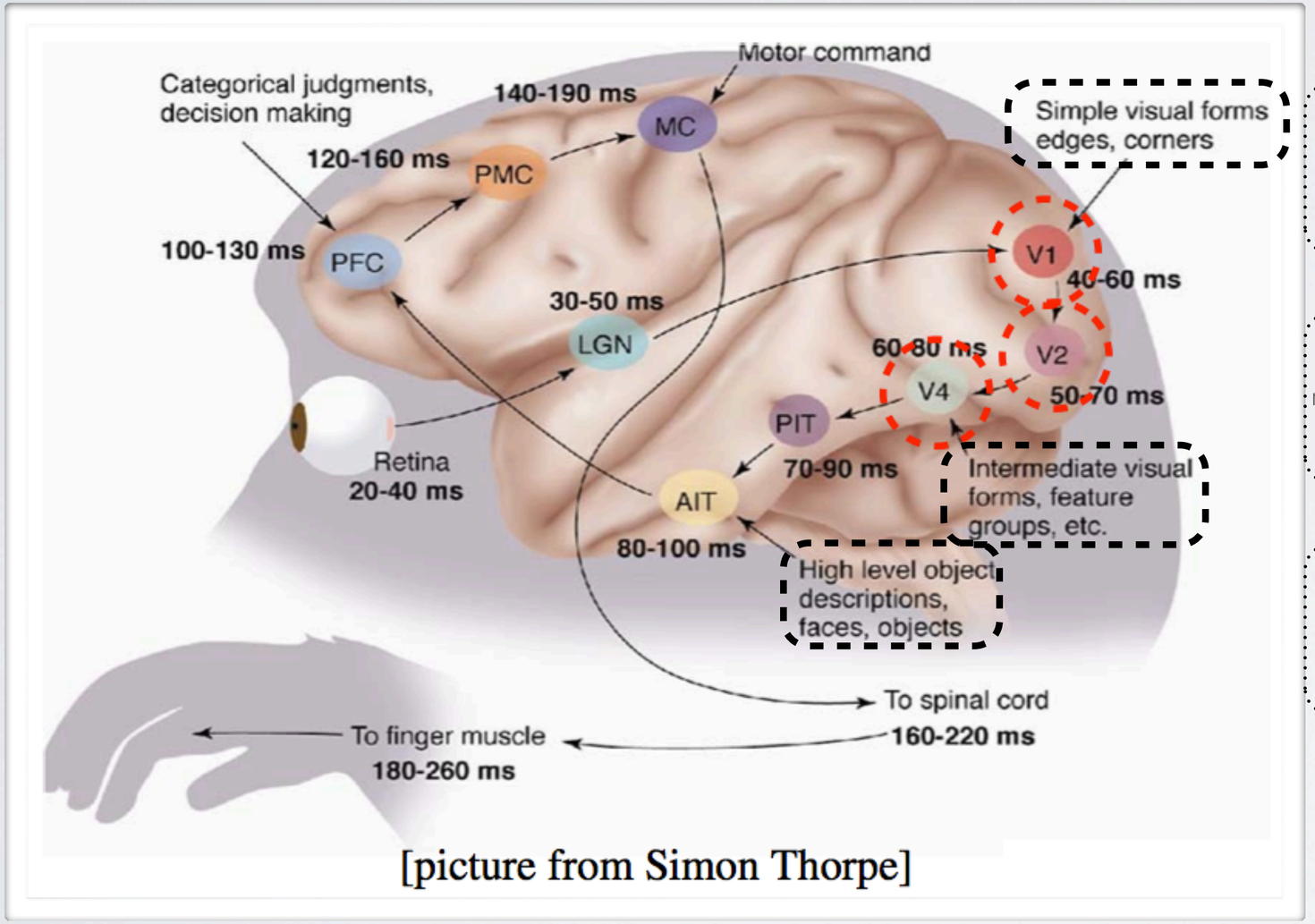
Topics: inspiration from visual cortex

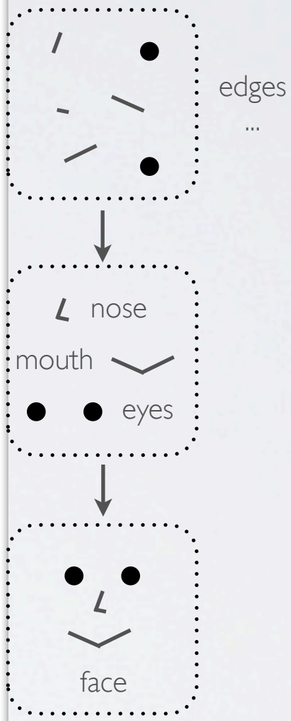




DEEP LEARNING

Topics: inspiration from visual cortex





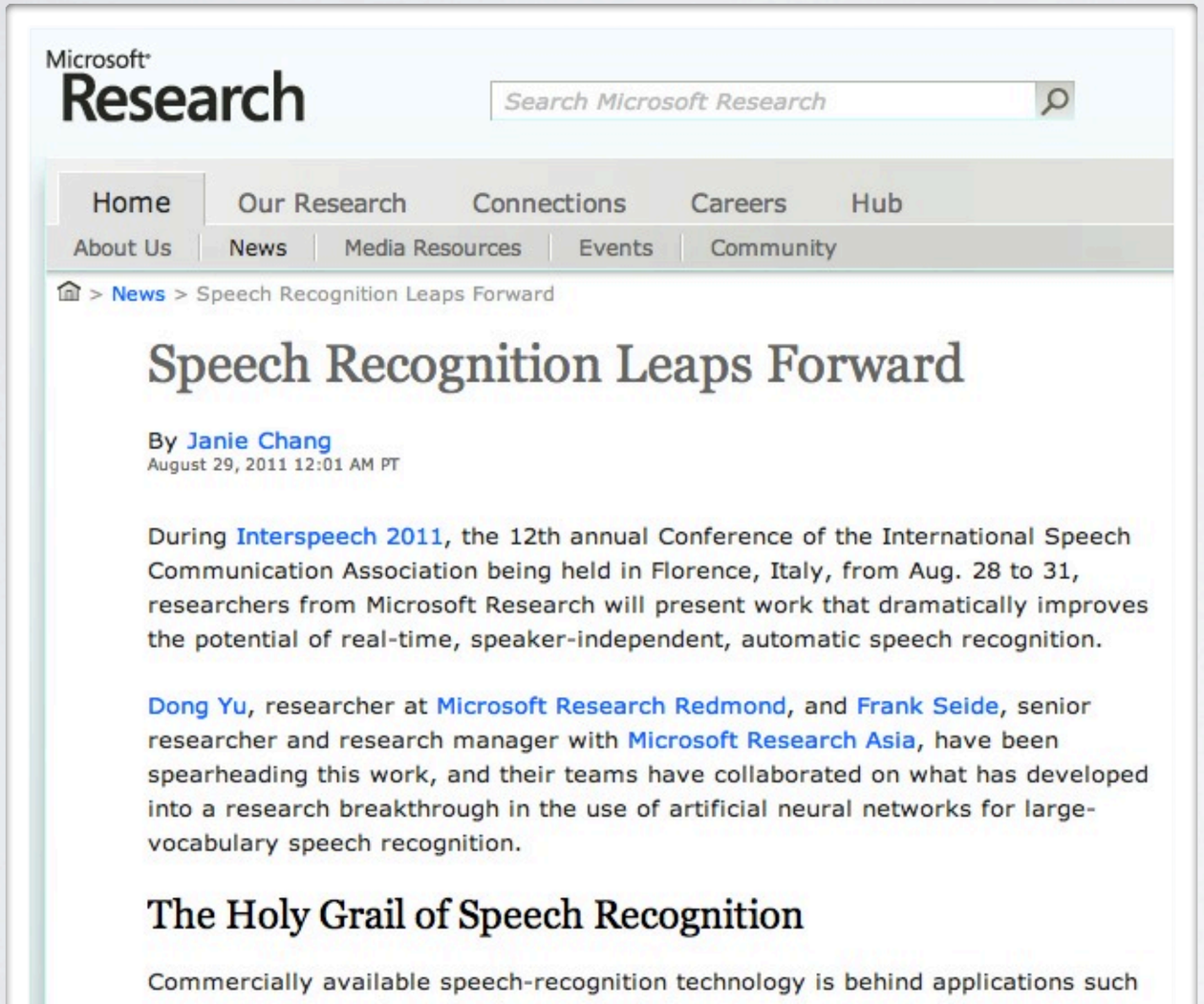
DEEP LEARNING

Topics: theoretical justification

- A deep architecture can represent certain functions (exponentially) more compactly
- Example: Boolean functions
  - a Boolean circuit is a sort of feed-forward network where hidden units are logic gates (i.e. AND, OR or NOT functions of their arguments)
  - any Boolean function can be represented by a "single hidden layer" Boolean circuit
    - however, it might require an exponential number of hidden units
  - it can be shown that there are Boolean functions which
    - require an exponential number of hidden units in the single layer case
    - require a polynomial number of hidden units if we can adapt the number of layers
  - See "Exploring Strategies for Training Deep Neural Networks" for a discussion

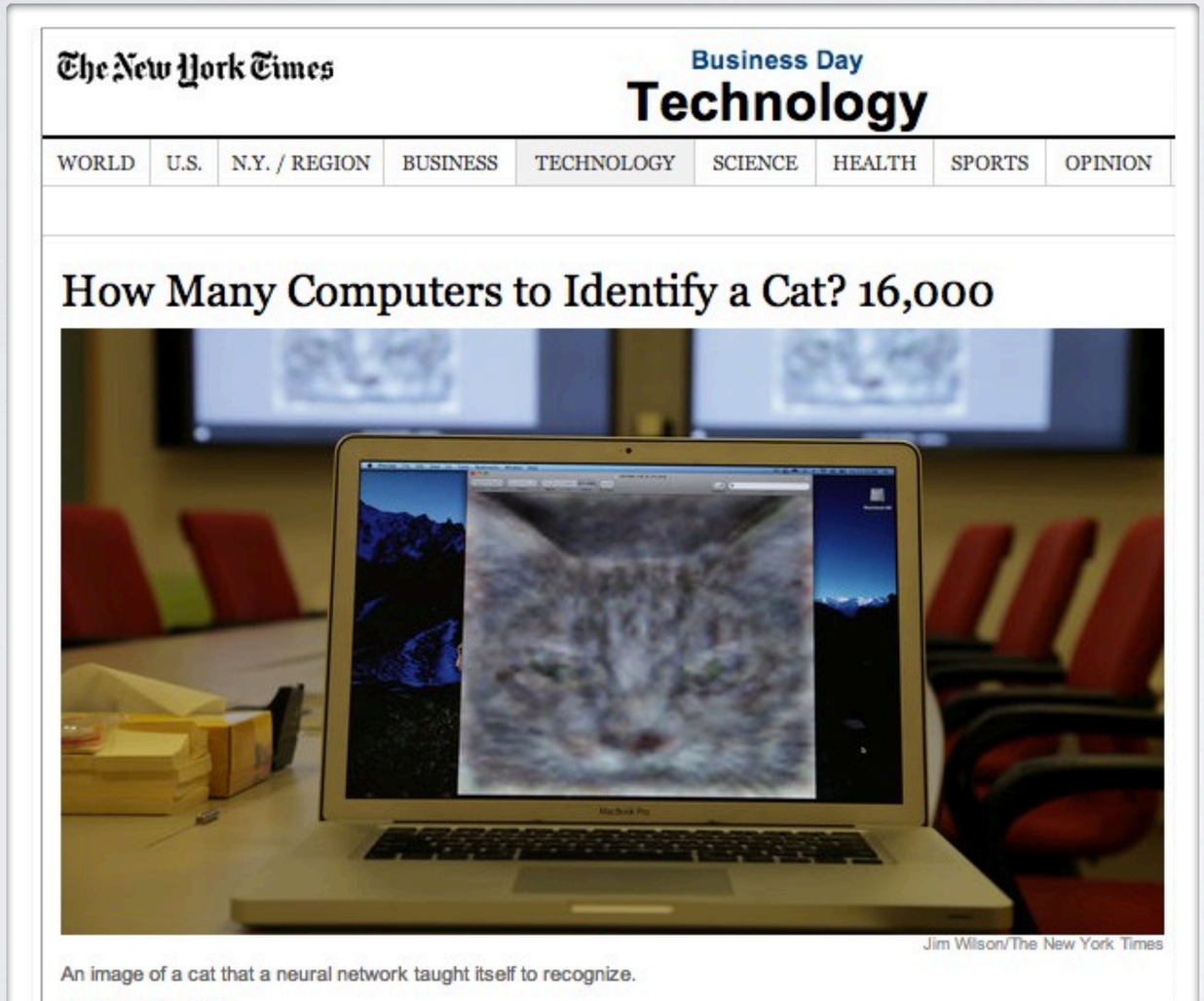
DEEP LEARNING

Topics: success stories (Microsoft Research)



DEEP LEARNING

Topics: success stories (Google)



The New York Times

Business Day
Technology

WORLD | U.S. | N.Y. / REGION | BUSINESS | TECHNOLOGY | SCIENCE | HEALTH | SPORTS | OPINION

How Many Computers to Identify a Cat? 16,000

Jim Wilson/The New York Times

An image of a cat that a neural network taught itself to recognize.