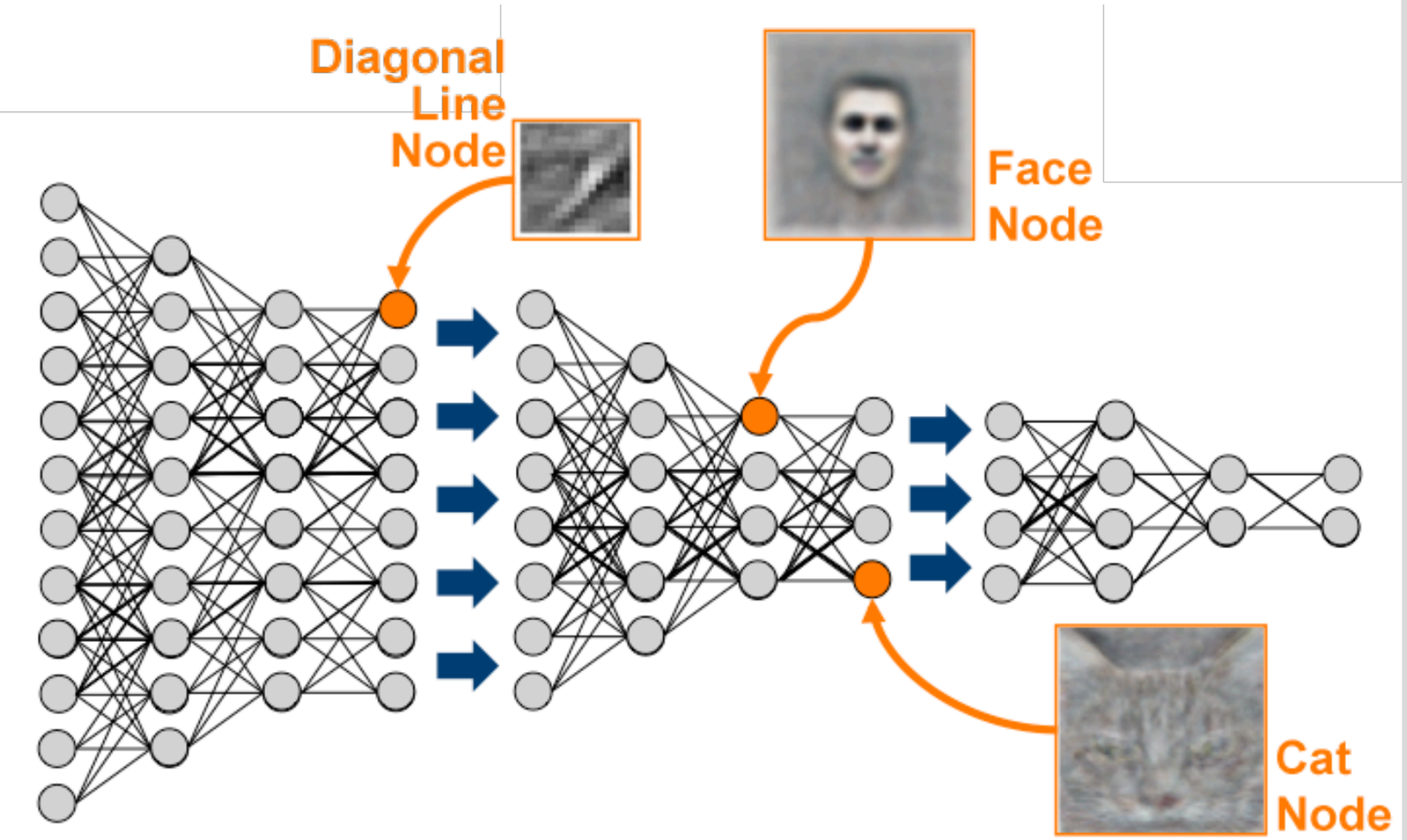
Learning to Deep Learn using Python, Keras, Theano, TensorFlow and a GPU

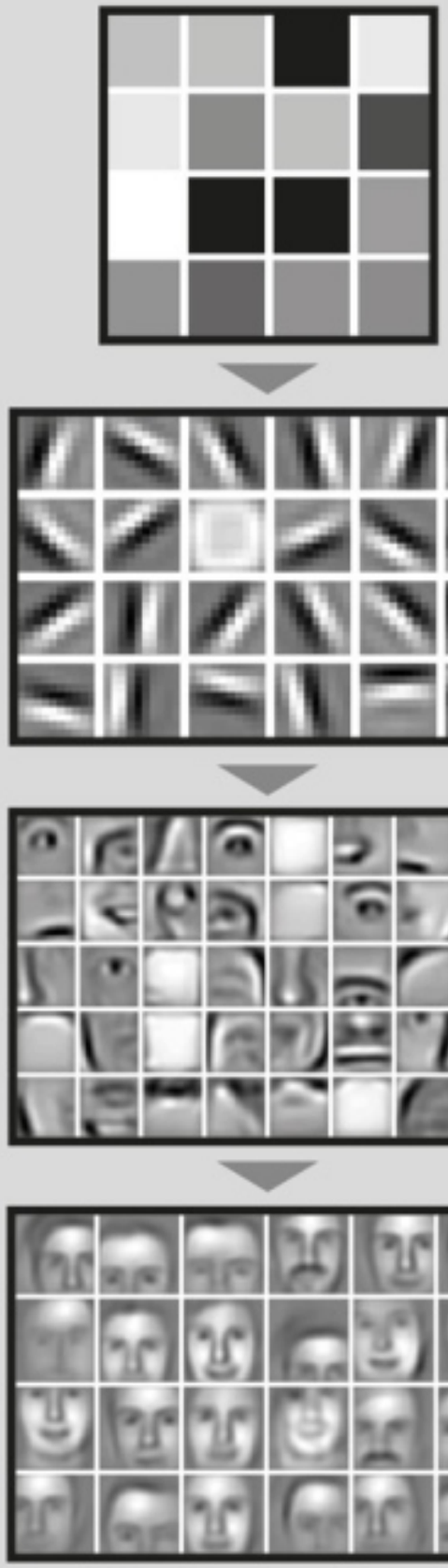
Jonathon Hare

Deep learning: learning layers of features



FACIAL RECOGNITION

Deep-learning neural networks use layers of increasingly complex rules to categorize complicated shapes such as faces.



Layer 1: The computer identifies pixels of light and dark.

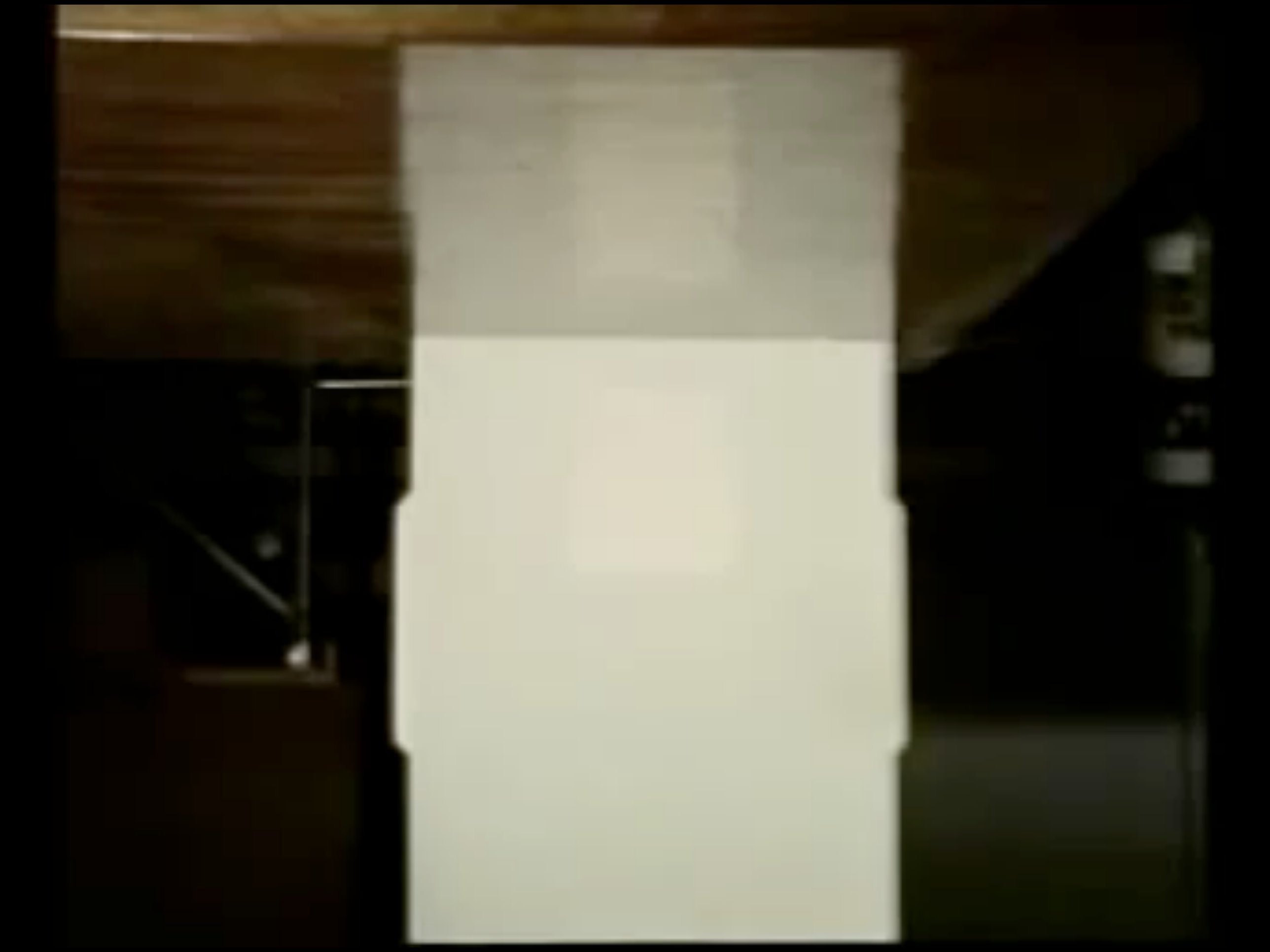
Layer 2: The computer learns to identify edges and simple shapes.

Layer 3: The computer learns to identify more complex shapes and objects.

Layer 4: The computer learns which shapes and objects can be used to define a human face.

Motivation

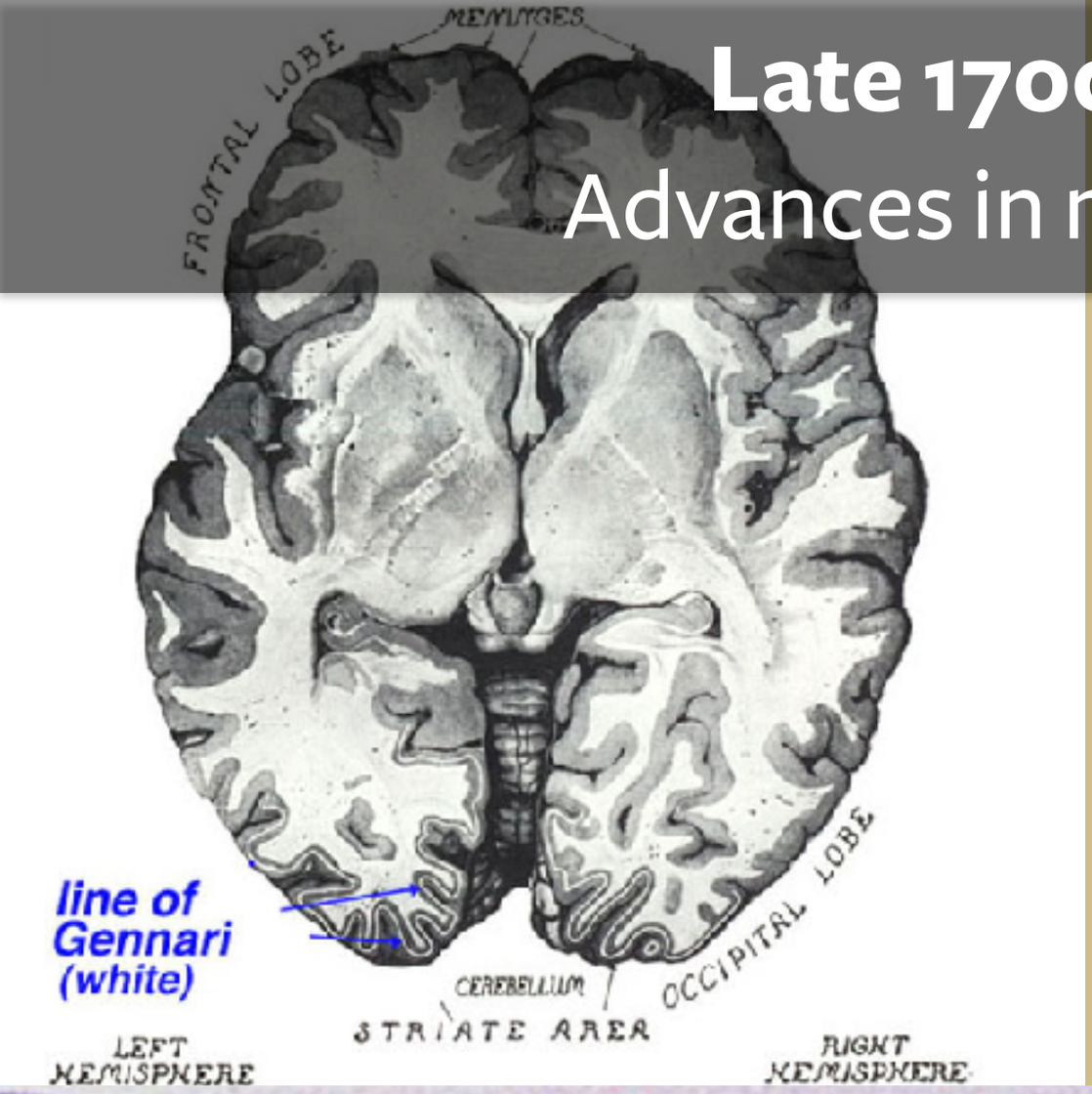
Is vision innate or acquired?

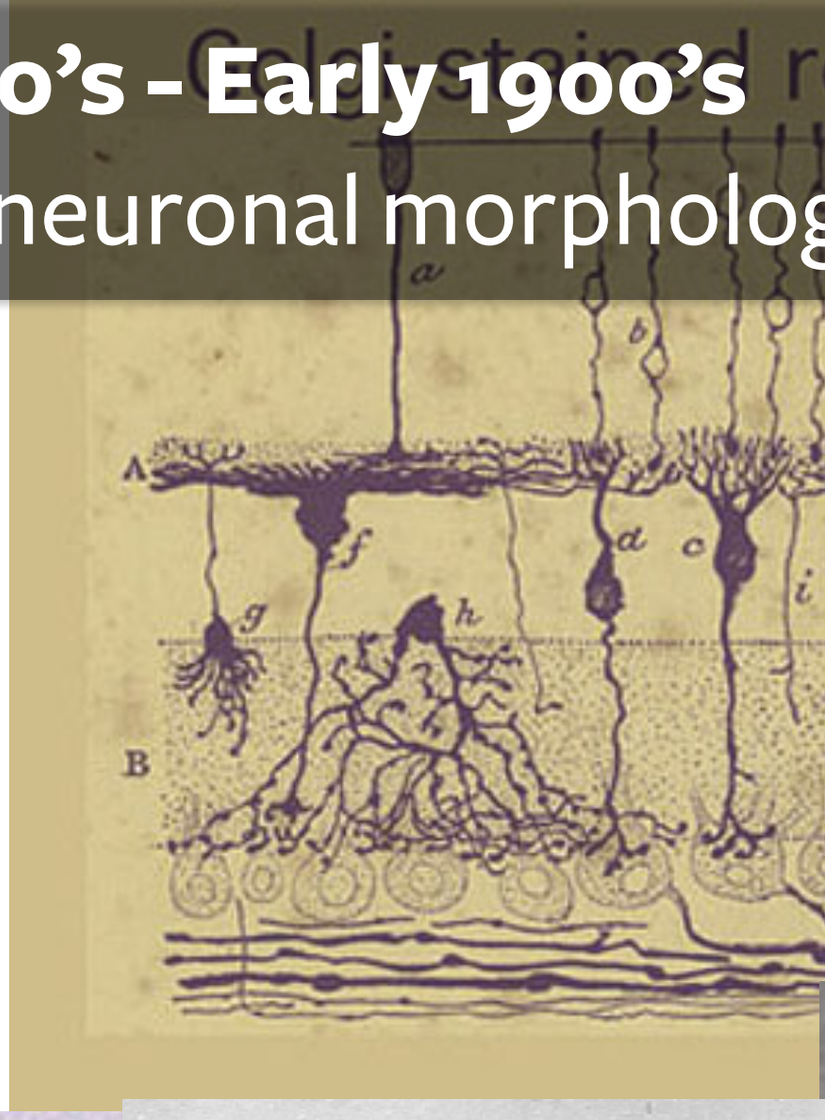


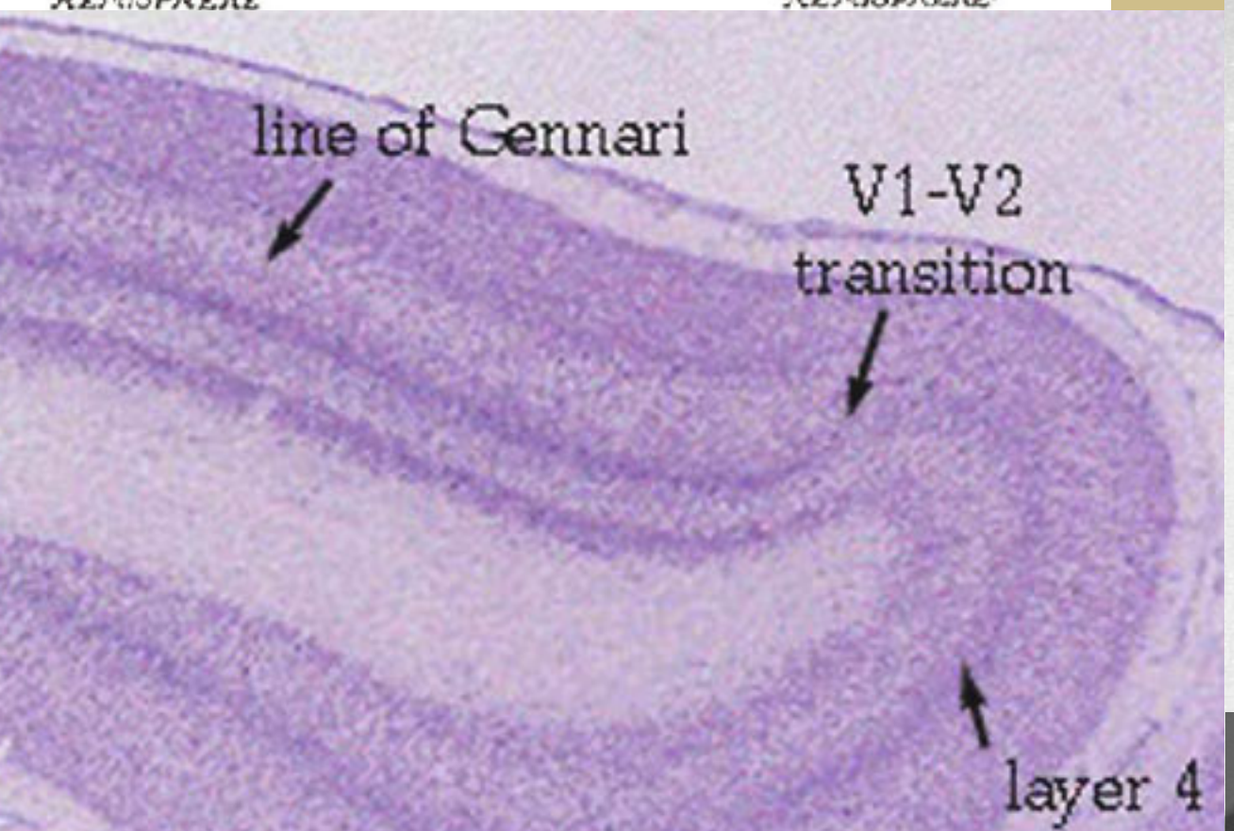
Historical Context

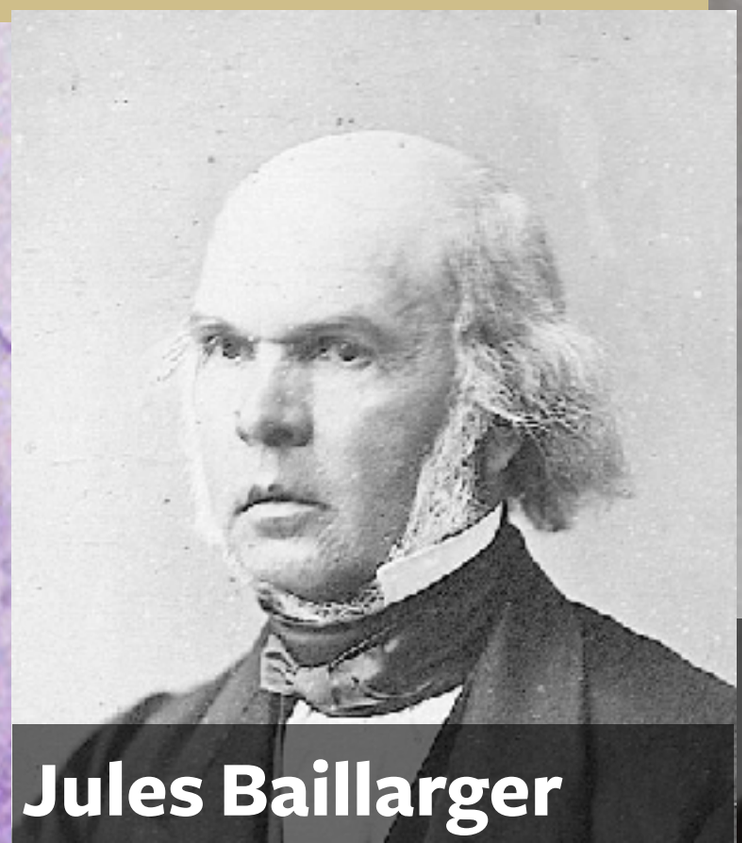
Late 1700's - Early 1900's

Advances in neuronal morphology







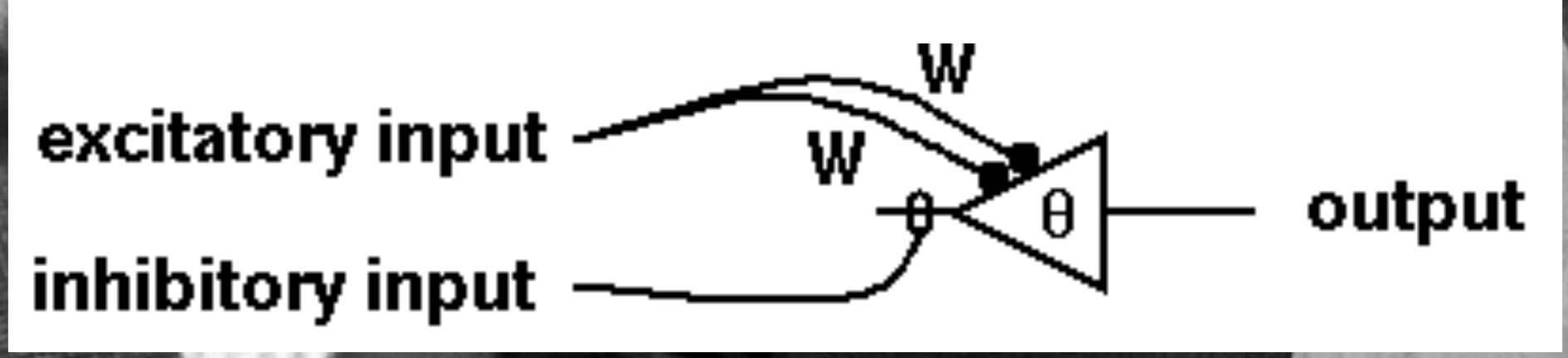


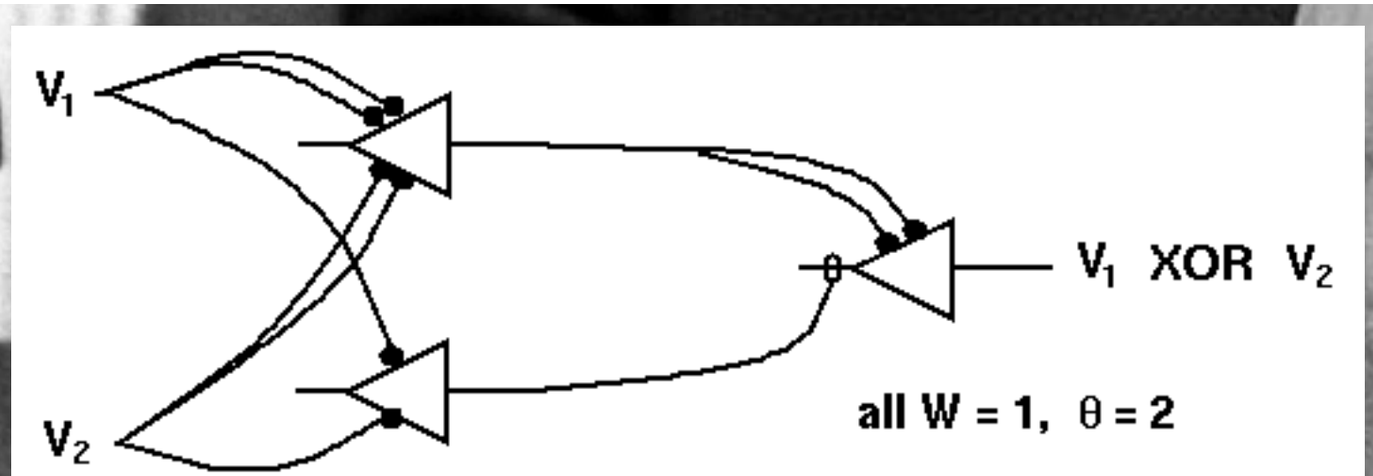
Jules Baillarger

Santiago Ramón y Cajal

1943
McCulloch-Pitts Artificial Neuron

$$V_i = \begin{cases} 1 & : \quad \sum_j W V_j \geq \theta \text{ AND no inhibition} \\ 0 & : \quad \text{otherwise} \end{cases}$$

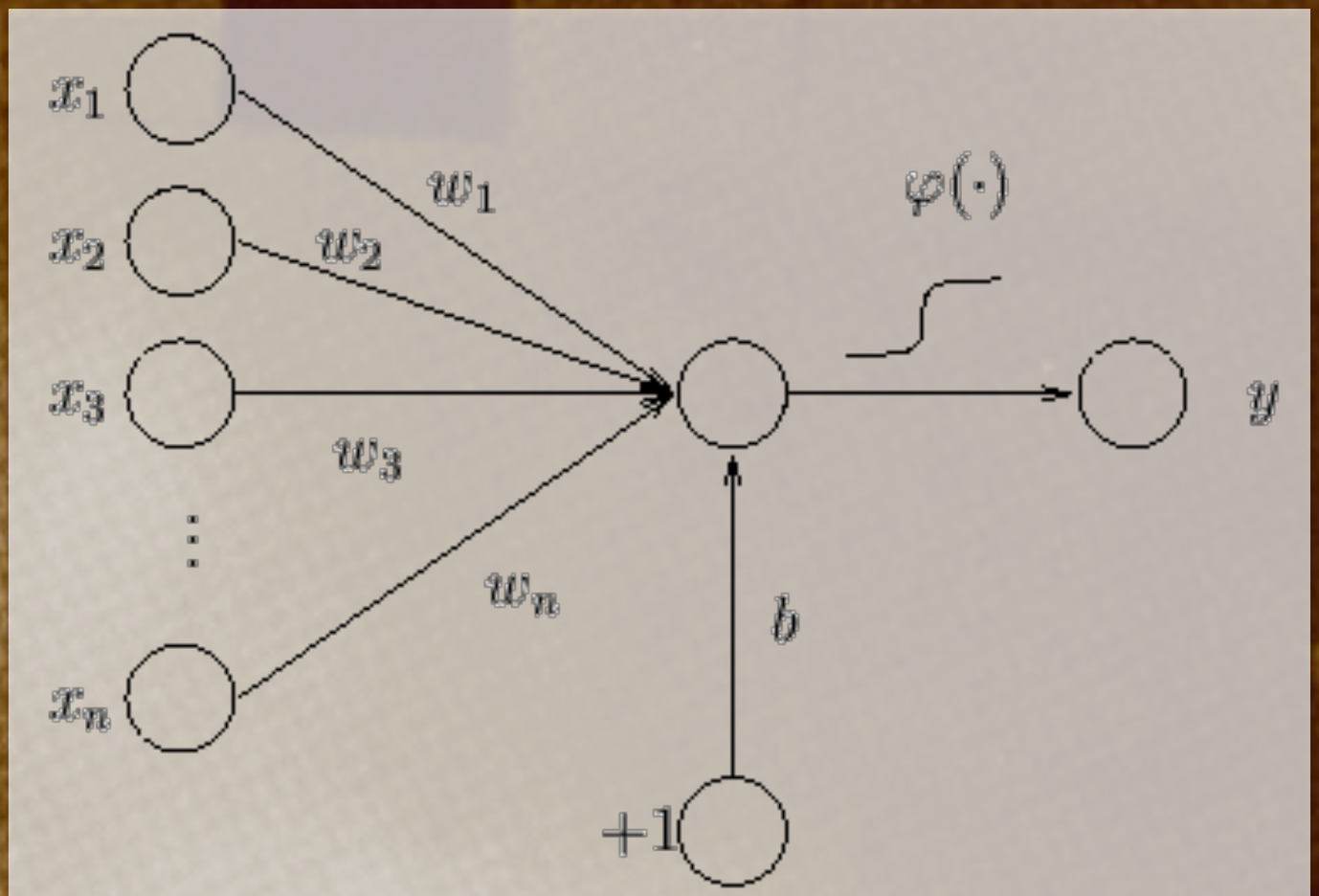




Warren McCulloch

Walter Pitts

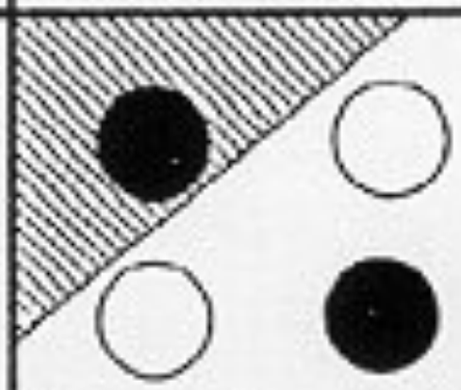
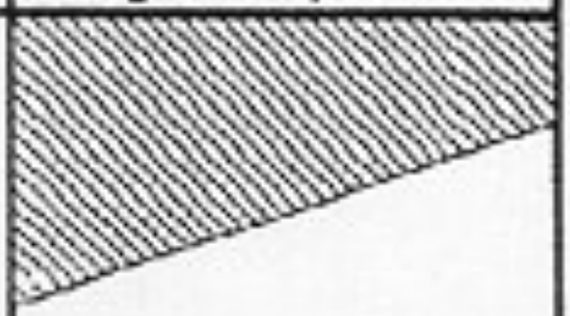
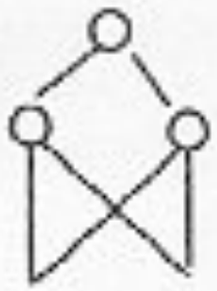
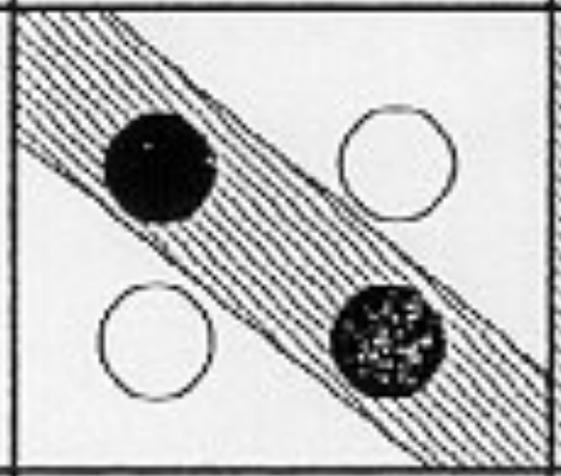
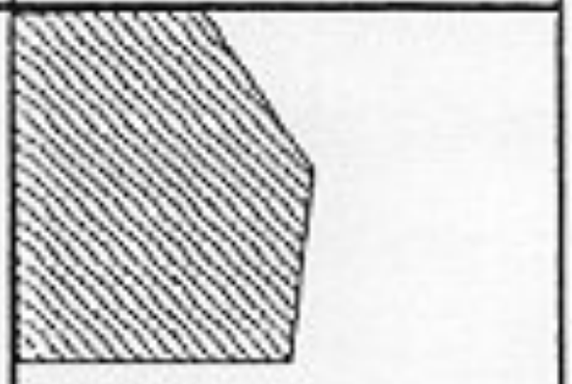
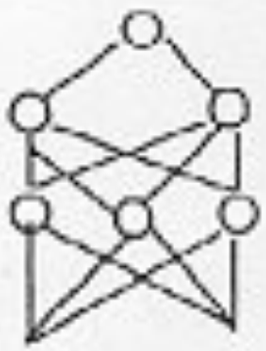
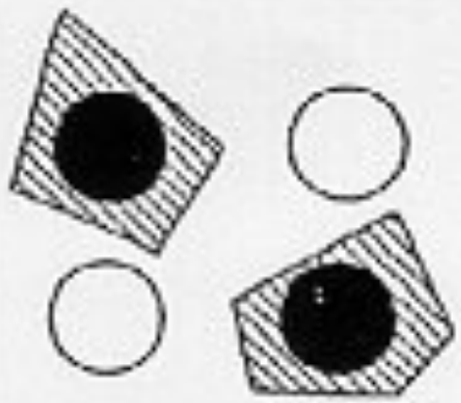
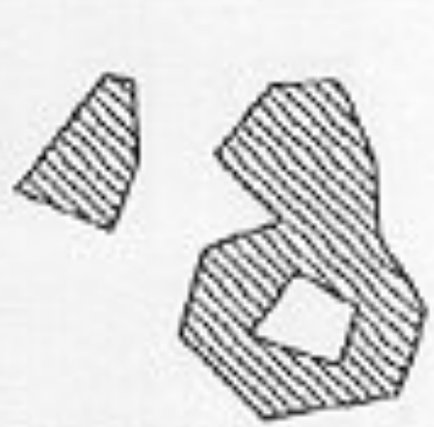
1958
Rosenblatt's Perceptron



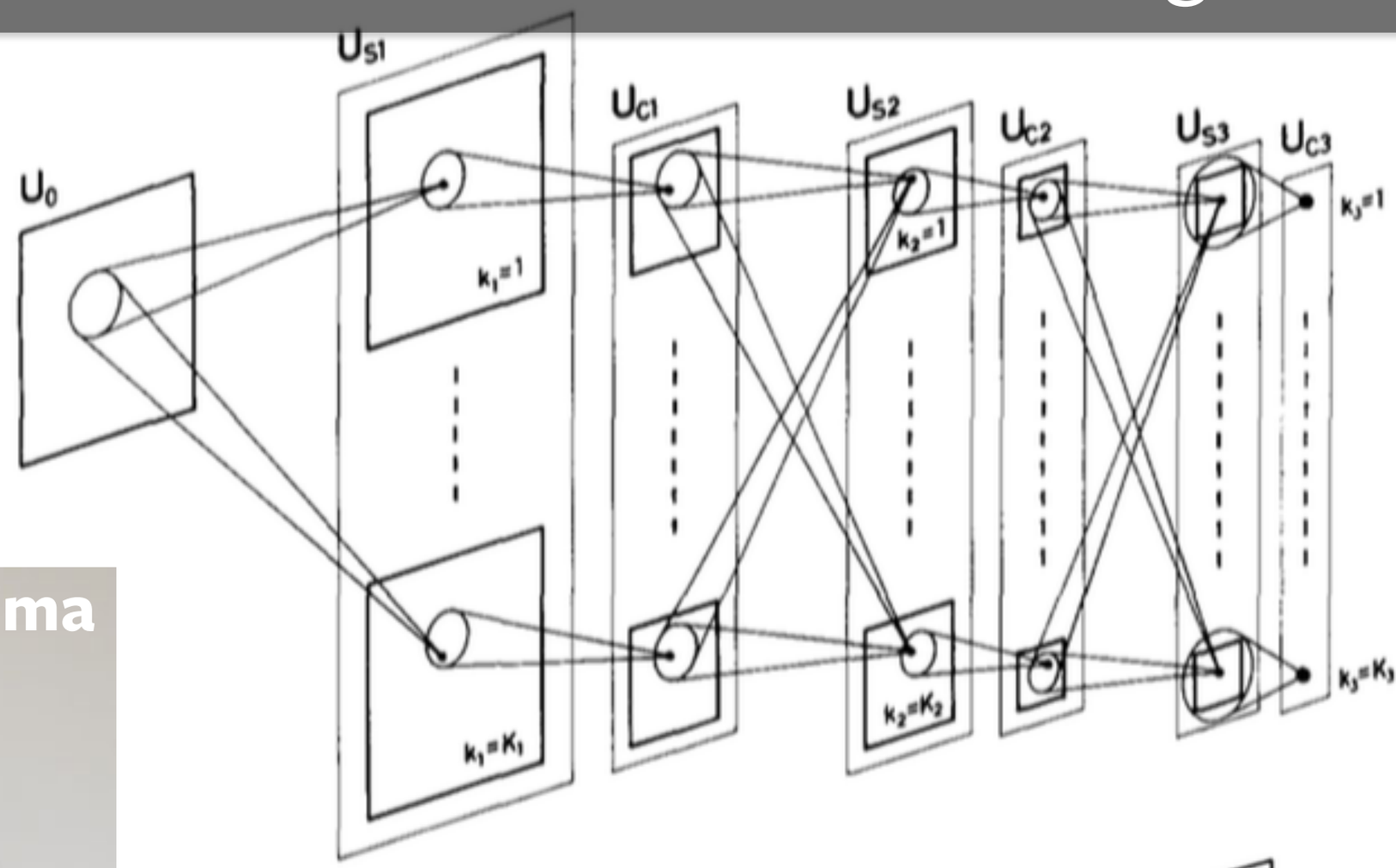
Frank Rosenblatt

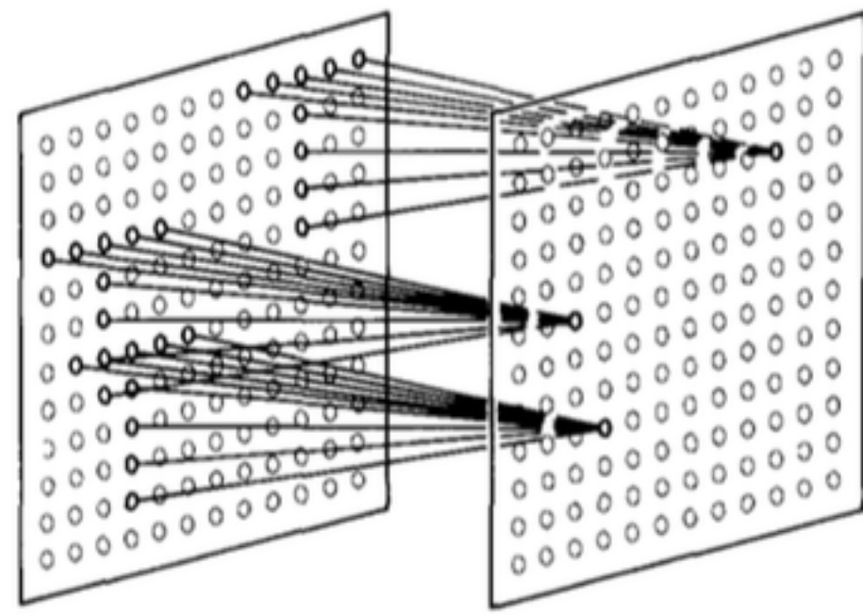
$$y = \varphi(\sum_{i=1}^{n} w_i x_i + b) = \varphi(\mathbf{w}^T \mathbf{x} + b)$$


MARK I
CORNELL AER
BUFFALO
PERCEPTRON
AUTICAL LABORATORY, Inc.
NEW YORK

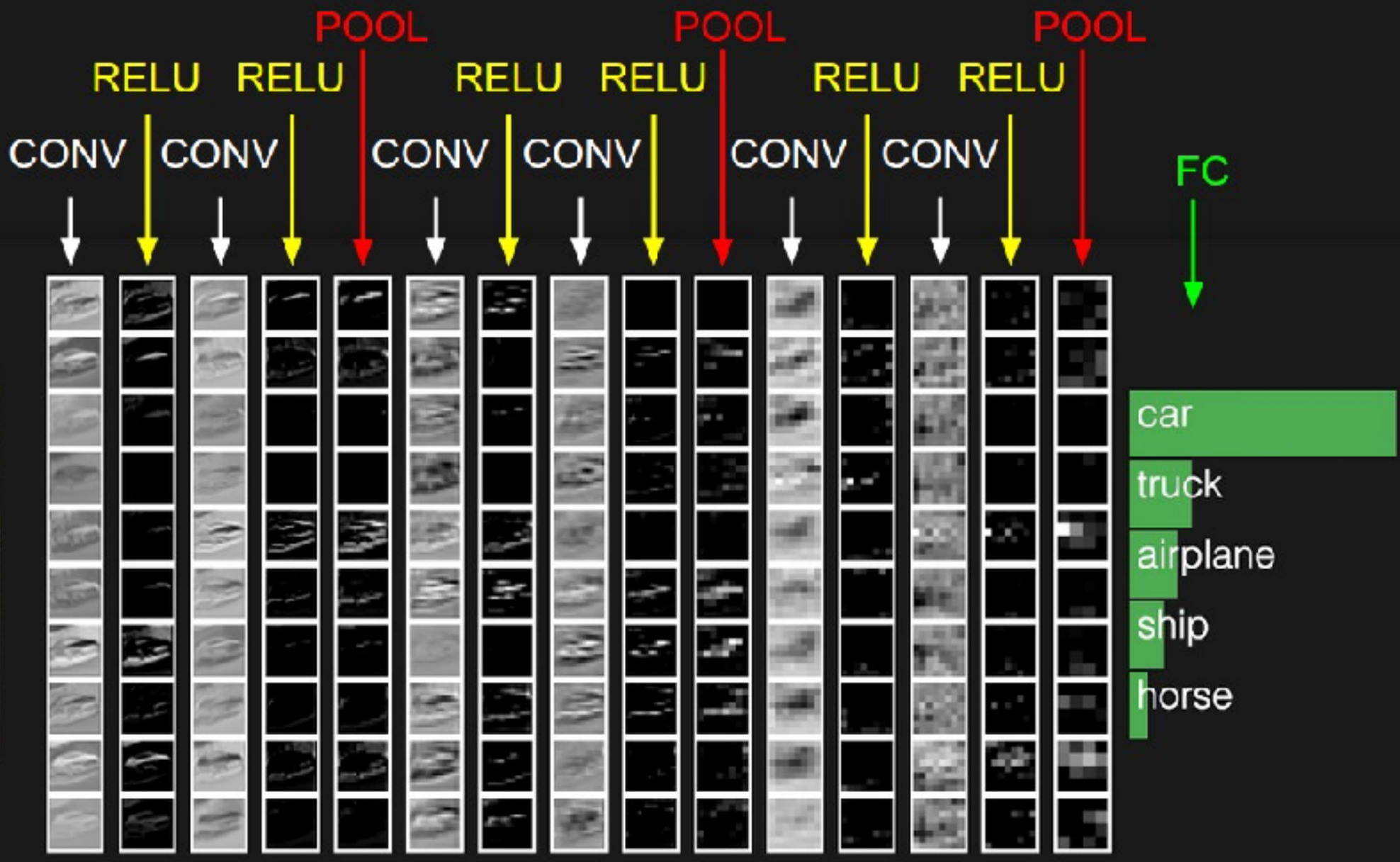
Structure	Description of decision regions	Exclusive-OR problem	Classes with meshed regions	General region shapes
 Single layer	Half plane bounded by hyperplane			
 Two layer	Arbitrary (complexity limited by number of hidden units)			
 Three layer	Arbitrary (complexity limited by number of hidden units)			

1979
Convolutional Neural Networks & Neocognitron

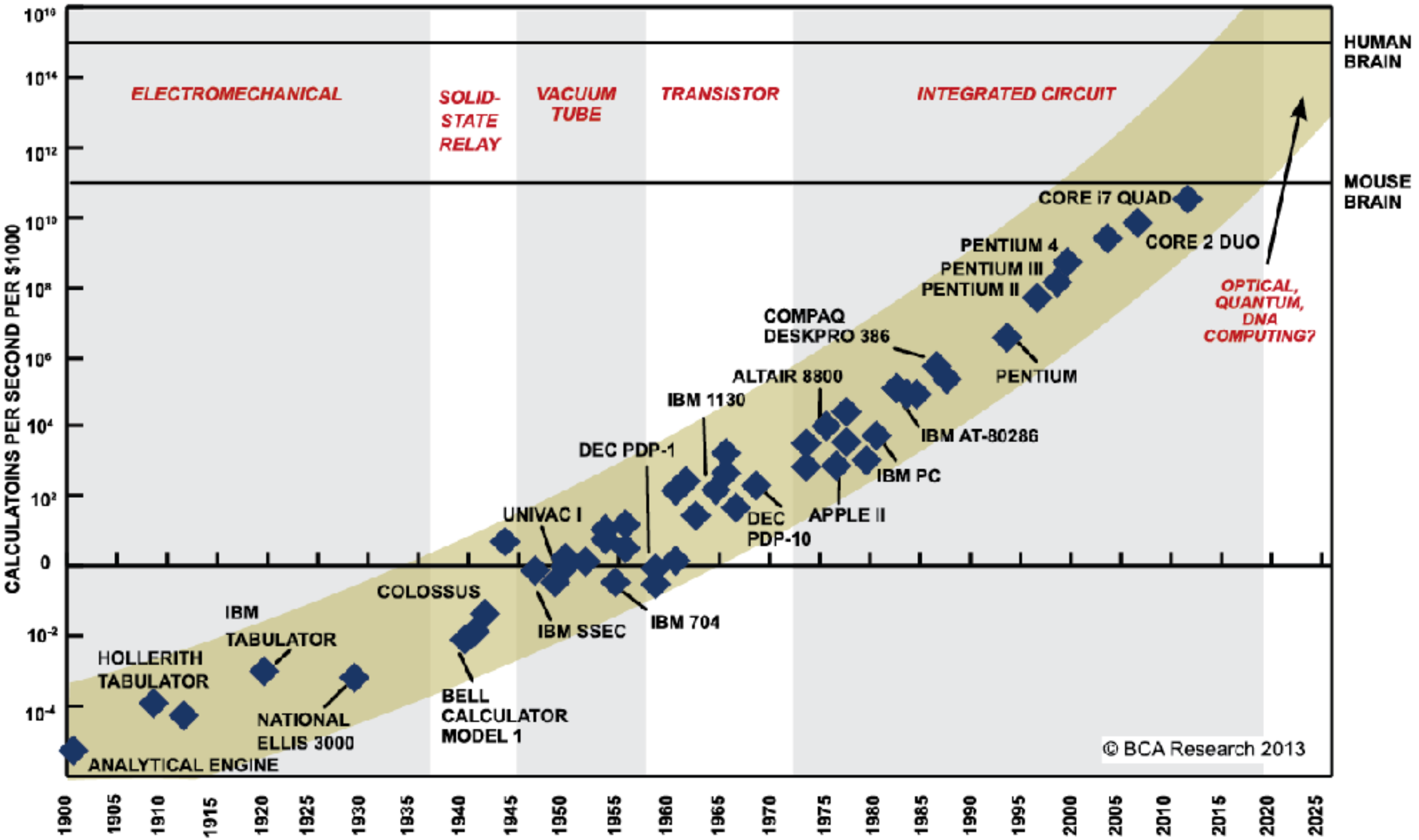




Kunihiko Fukushima


POOL
POOL
POOL
RELU
RELU
RELU
RELU
RELU
RELU
CONV
CONV
CONV
CONV
CONV
CONV
FC
car
truck
airplane
ship
horse

Moore's Law



SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, *THE VIKING PRESS*, 2006. DATAPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

The new Moore’s Law: Computer’s no longer get faster, just wider

Lets get hands-on

- We'll learn:
  - How to use the MNIST dataset in Keras.
  - How to develop and evaluate a baseline neural network model for the MNIST problem.
  - How to switch the backends used by Keras and run your code on the GPU.
  - How to implement and evaluate a simple Convolutional Neural Network for MNIST.
  - How to implement a close to state-of-the-art deep learning model for MNIST.
  - How to serialise and deserialise trained models.
  - How to load your own image created outside of the MNIST dataset, and pass it through the network.
  - How to visualise the filters learned by the network.
  - How to implement networks with branching and merging.