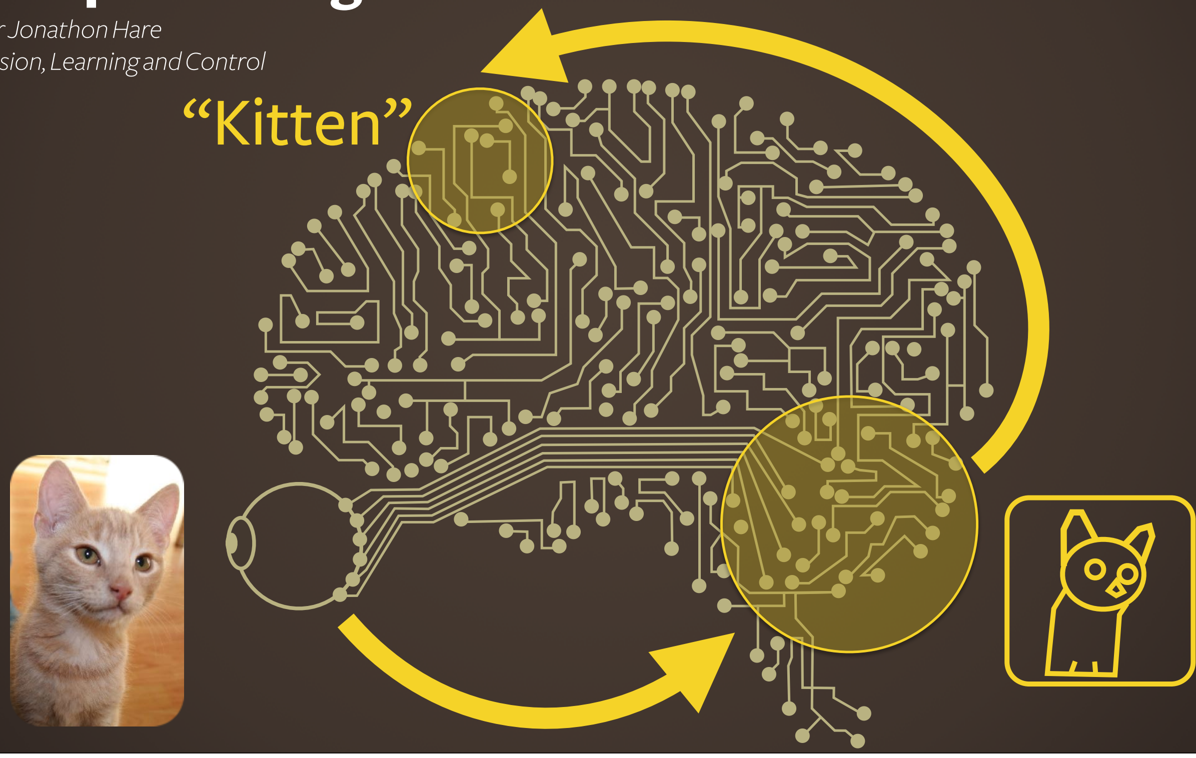
Electronics and Computer Science

UNIVERSITY OF Southampton

The biological inspiration and motivation of Deep Learning

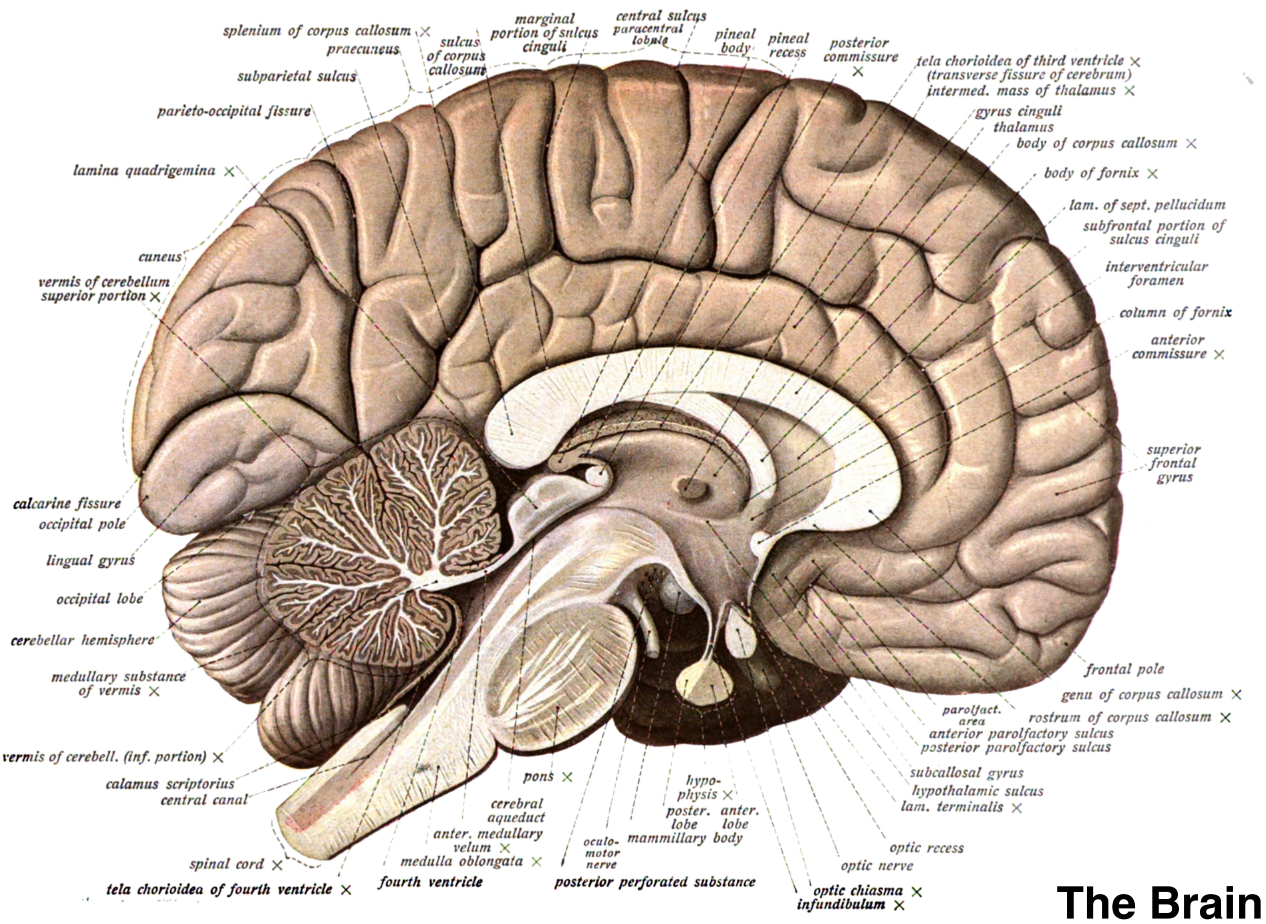
Dr Jonathon Hare
Vision, Learning and Control



- Biological systems, neuroscience and psychology have long served as inspiration and motivation for Artificial Intelligence

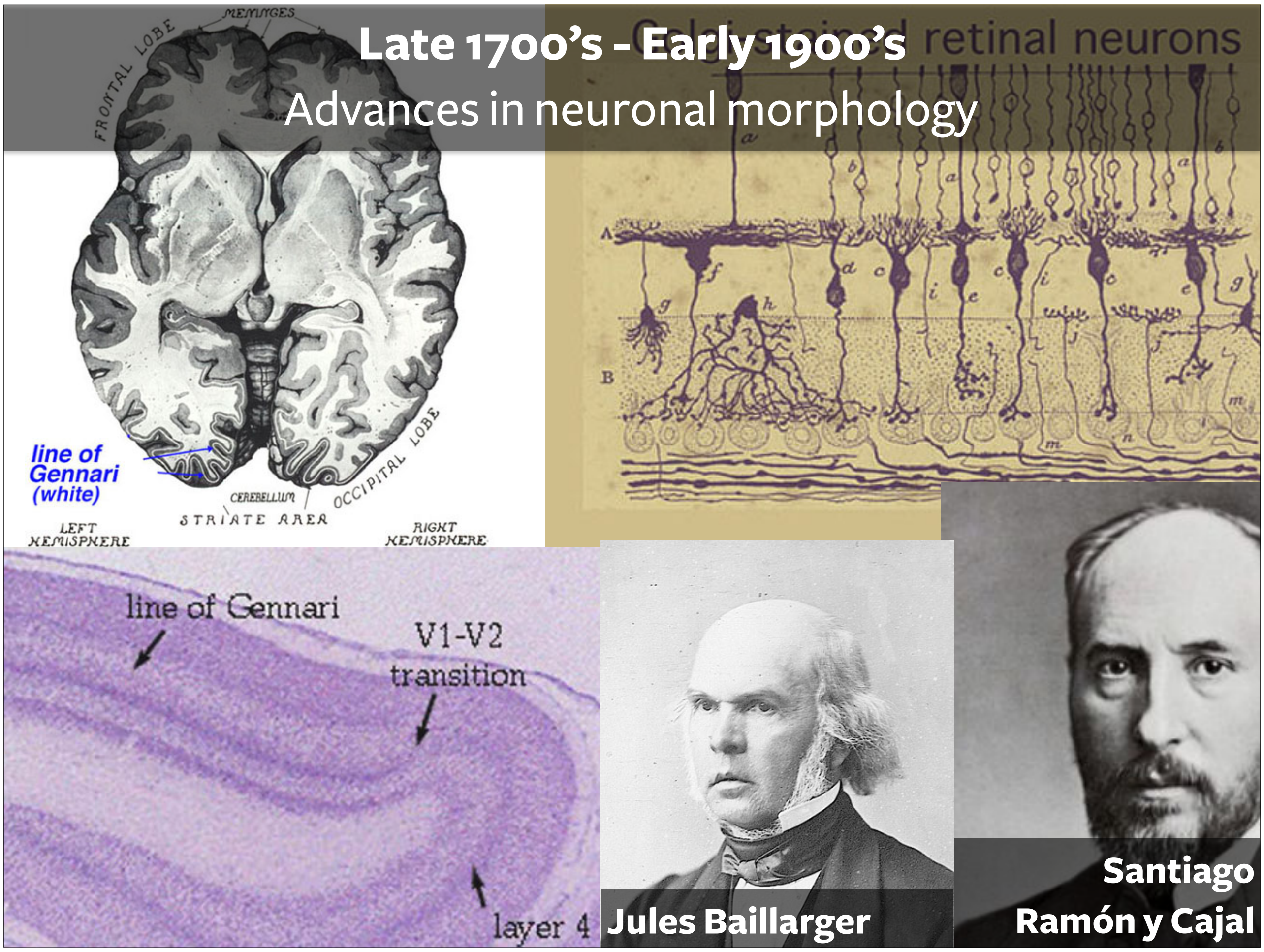
- There is no argument that current deep learning models and approaches are **not biologically plausible.**

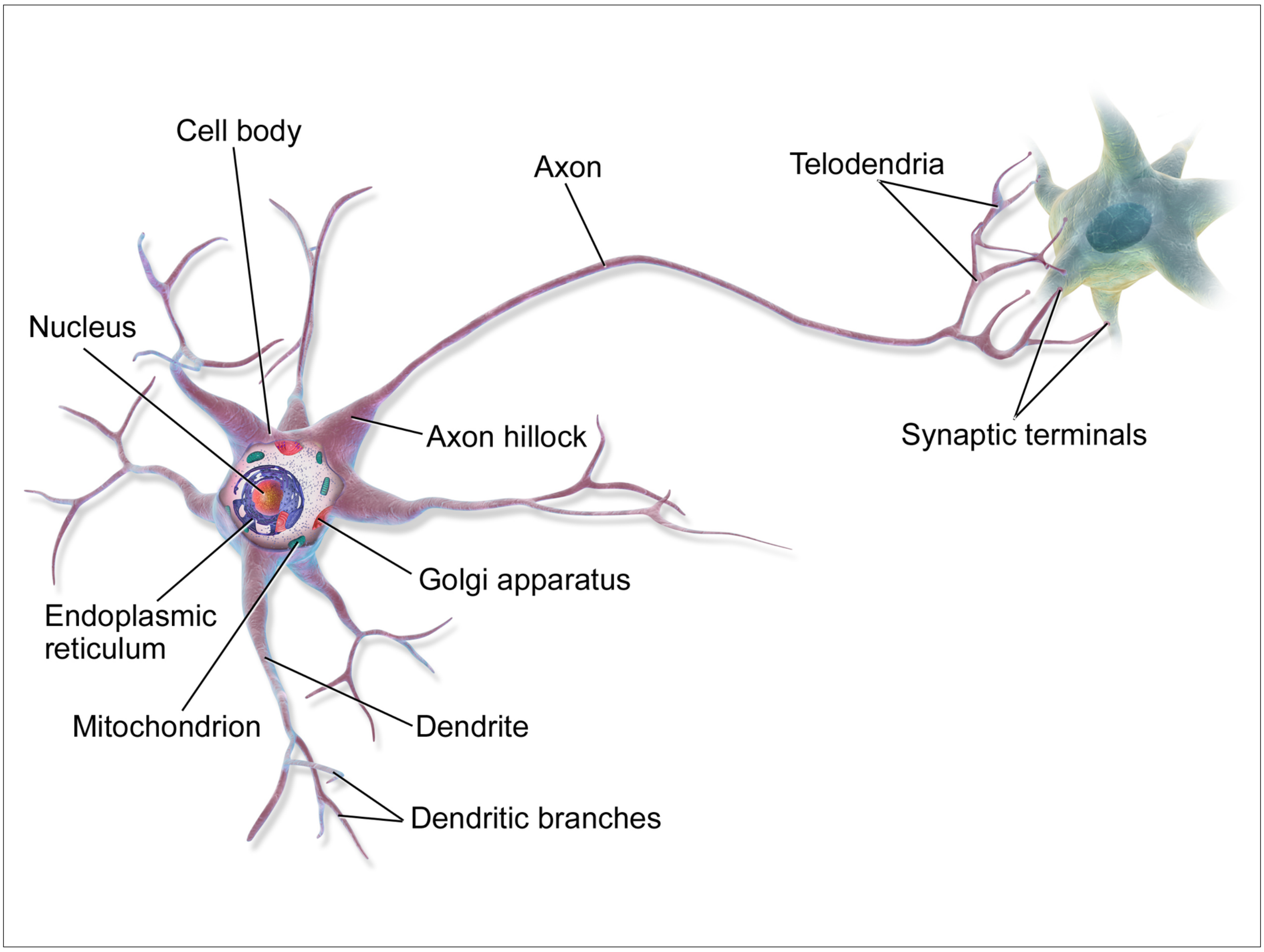
- But, throughout the development of neural networks and deep learning, biology has played a role…



The Brain

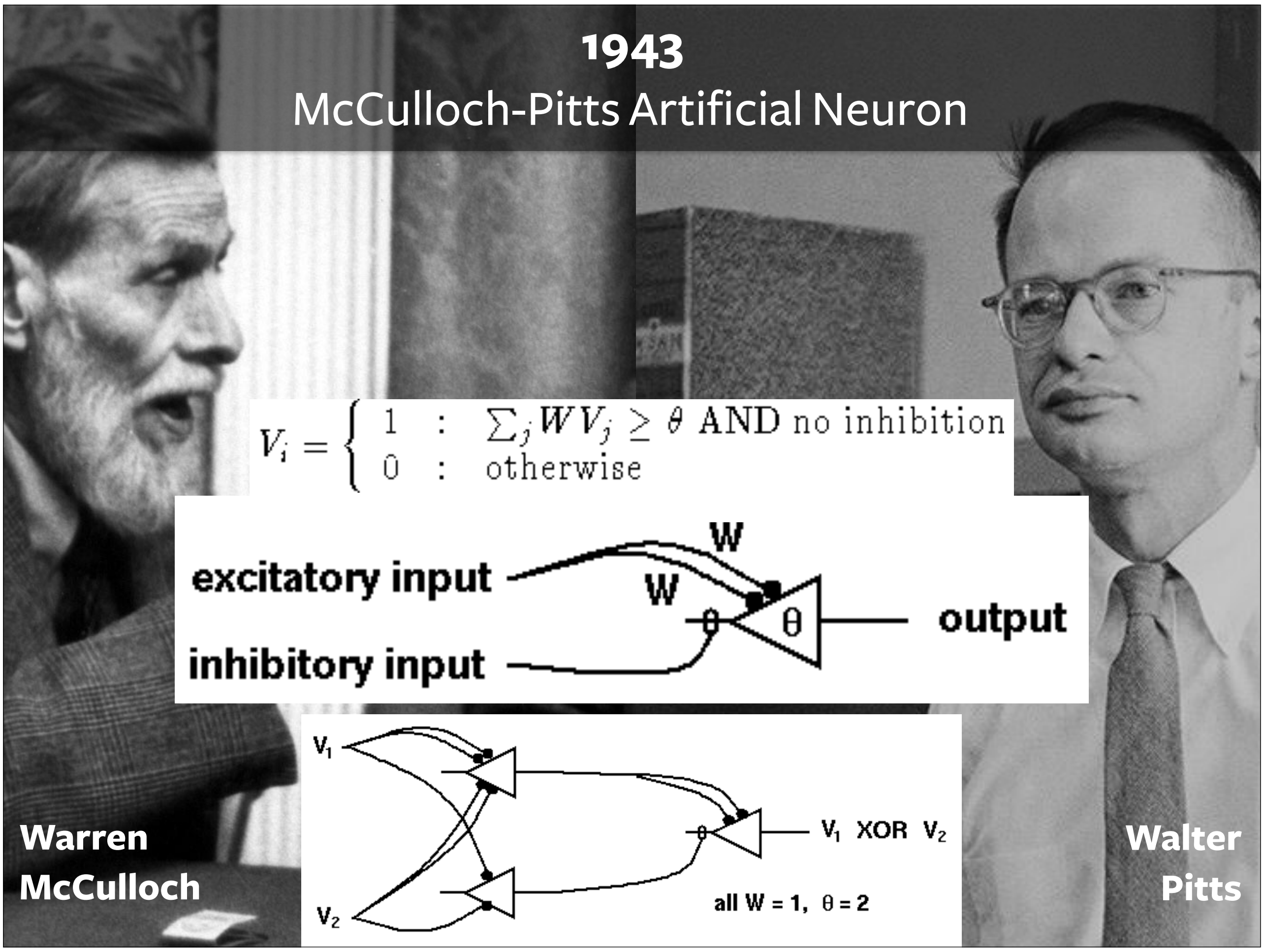
Biological Neurons

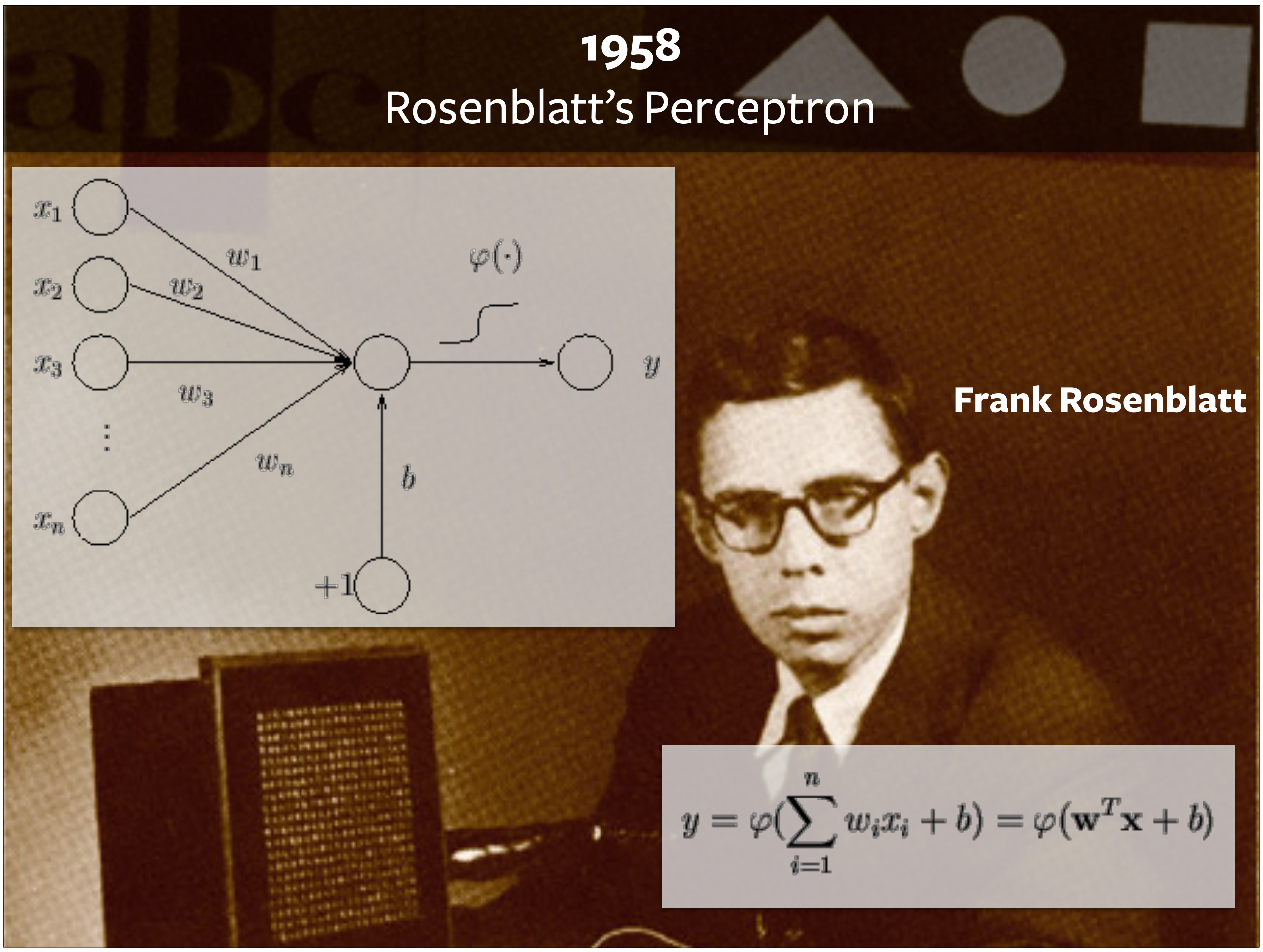

Late 1700's - Early 1900's
Advances in neuronal morphology
retinal neurons
MENINGES
FRONTAL LOBE
OCCIPITAL LOBE
line of Gennari (white)
CEREBELLUM
STRIATE AREA
LEFT HEMISPHERE
RIGHT HEMISPHERE
line of Gennari
V1-V2 transition
layer 4
Jules Baillarger
Santiago Ramón y Cajal


Cell body
Axon
Telodendria
Nucleus
Axon hillock
Synaptic terminals
Golgi apparatus
Endoplasmic reticulum
Mitochondrion
Dendrite
Dendritic branches

- Biological Neurons “spike”
    - If sufficient spikes are received by the dendrites within a time period the neuron fires its “action potential”
        - Some inputs are inhibitory - they reduce the potential of the neuron to fire
    - Once it has fired there is a short refractory period which inhibits firing again
- There are also a number of other interesting dynamical properties
    - e.g. short-term synaptic depression

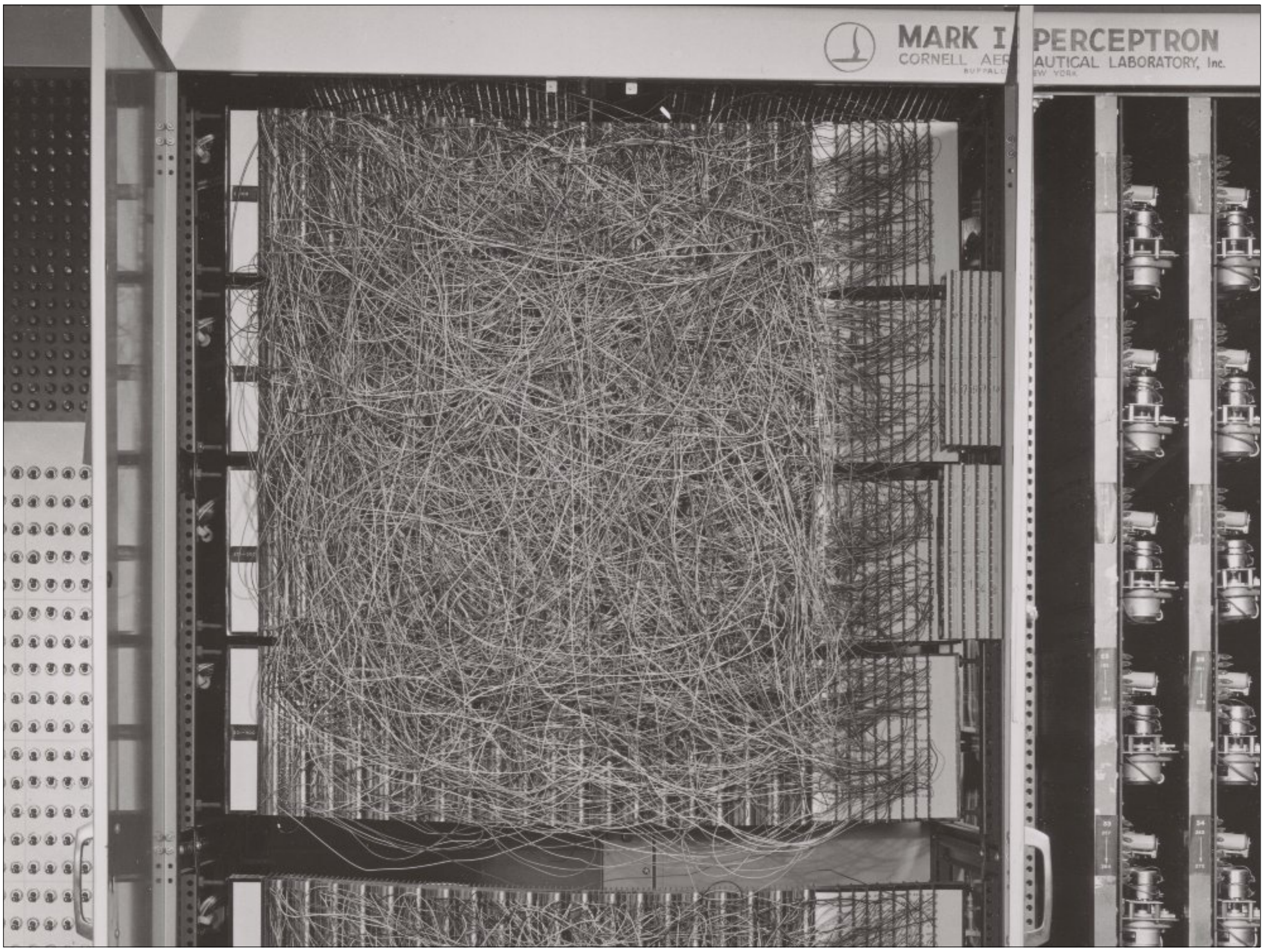
A Formalism for Approaching the Operation of a Single Neuron

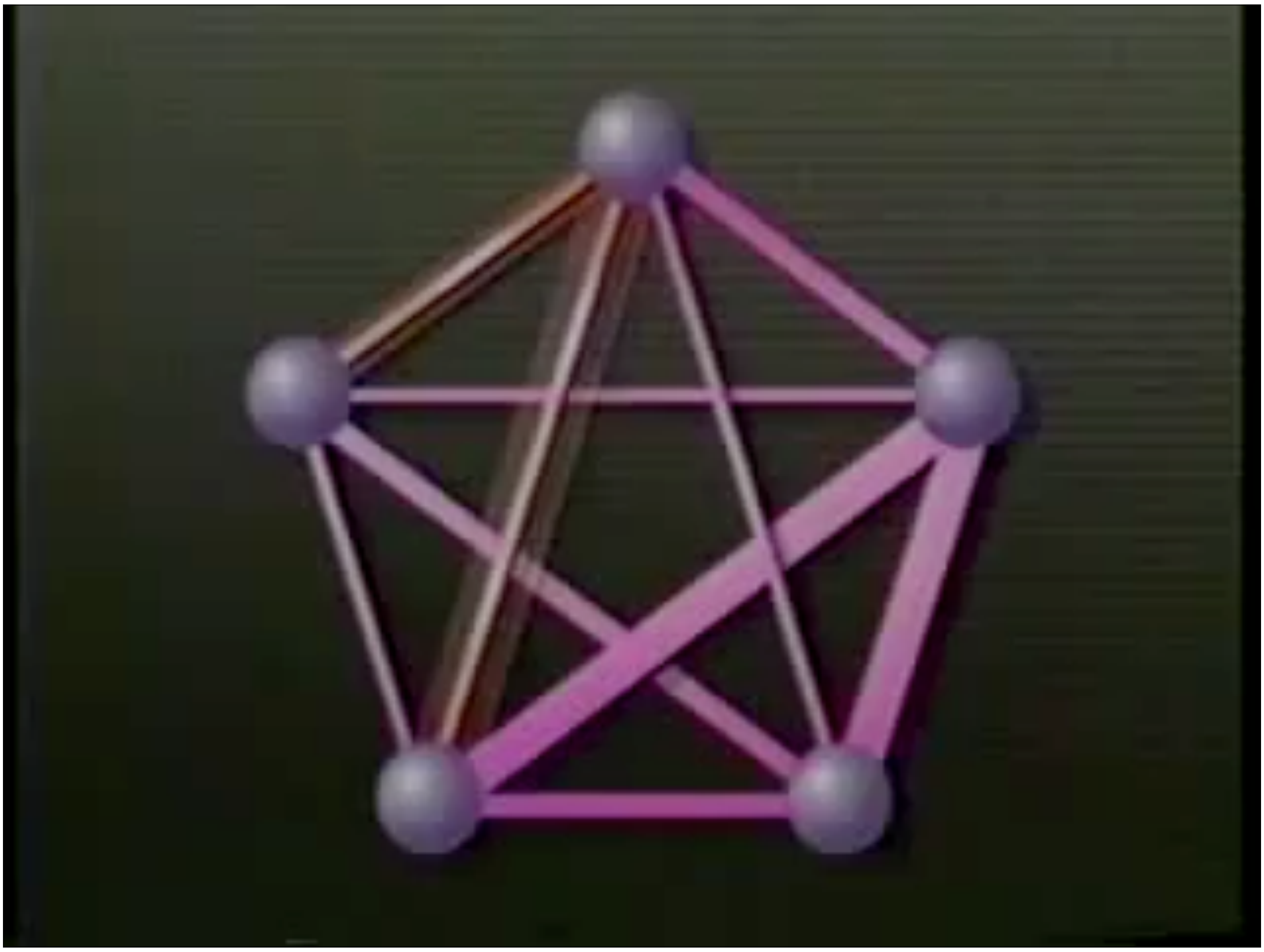

1943
McCulloch-Pitts Artificial Neuron
$V_i = \begin{cases} 1 & : \sum_j W V_j \geq \theta \text{ AND no inhibition} \\ 0 & : \text{otherwise} \end{cases}$
excitatory input
inhibitory input
W
W
θ
θ
output
V1
V2
V1 XOR V2
all W = 1, θ = 2
Warren McCulloch
Walter Pitts


1958
Rosenblatt's Perceptron
x_1
x_2
x_3
x_n
w_1
w_2
w_3
w_n
$\varphi(\cdot)$
y
b
$+1$
Frank Rosenblatt
$y = \varphi(\sum_{i=1}^{n} w_i x_i + b) = \varphi(\mathbf{w}^T\mathbf{x} + b)$

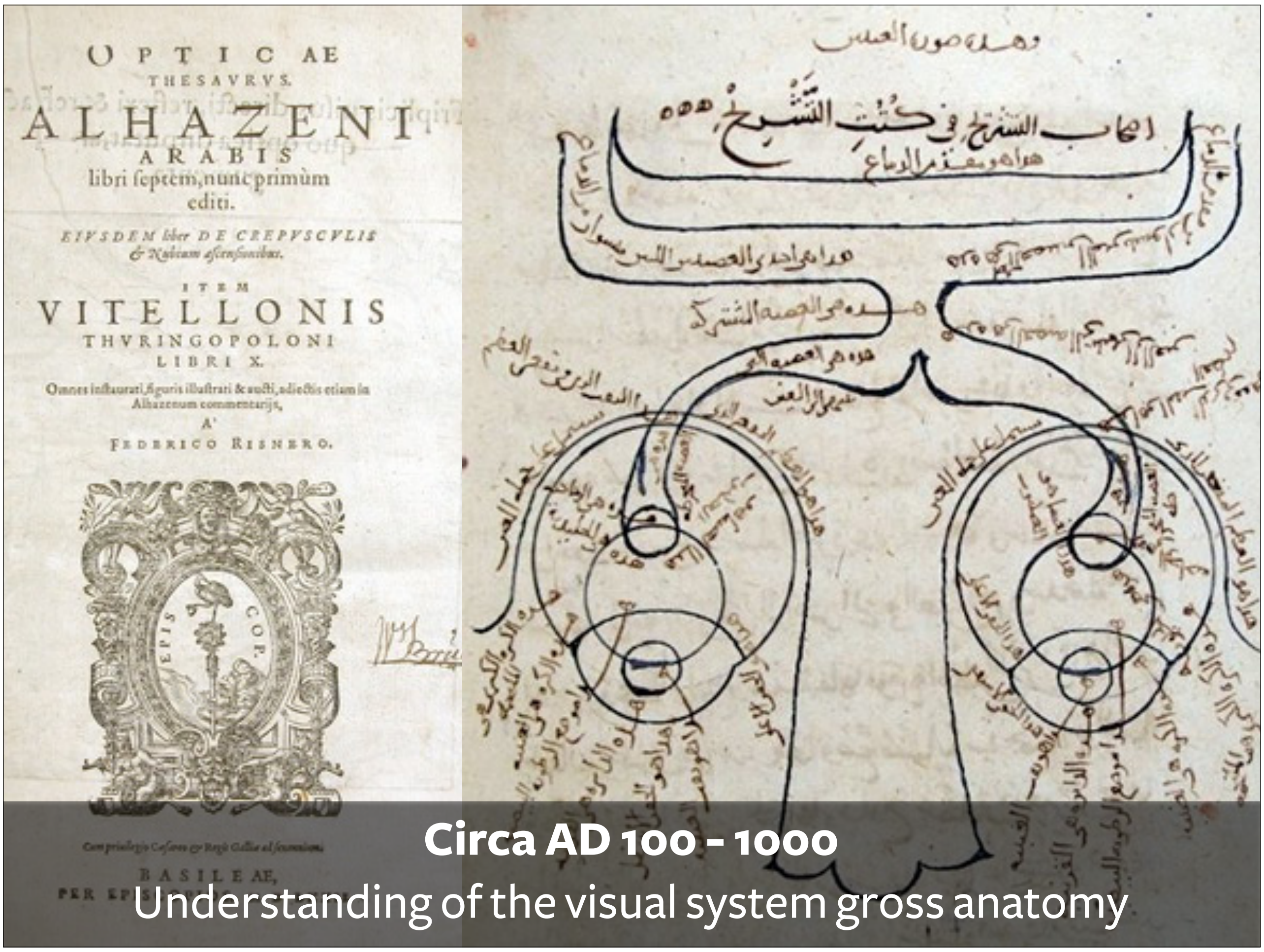
$$y = \varphi(\sum_{i=1}^{n} w_i x_i + b) = \varphi(\mathbf{w}^T \mathbf{x} + b)$$

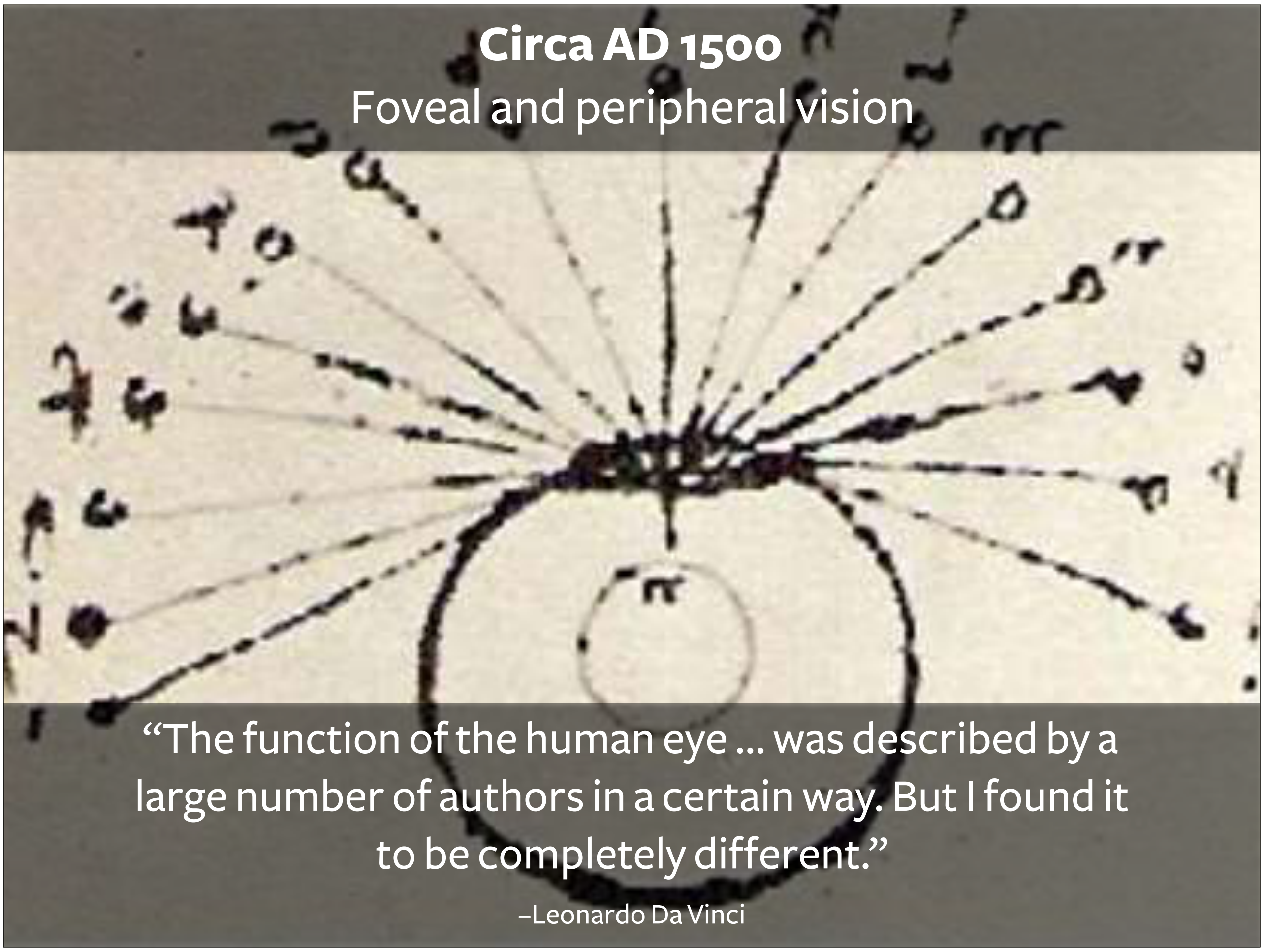
- Rosenblatt’s model captures many of the key points of a biological neuron:
  - Output is a function of the sum of inputs
  - Negative weights account for inhibitory connections
  - The “activation function” can ensure that output is only produced once a threshold is exceeded (although we use many variants of these days)



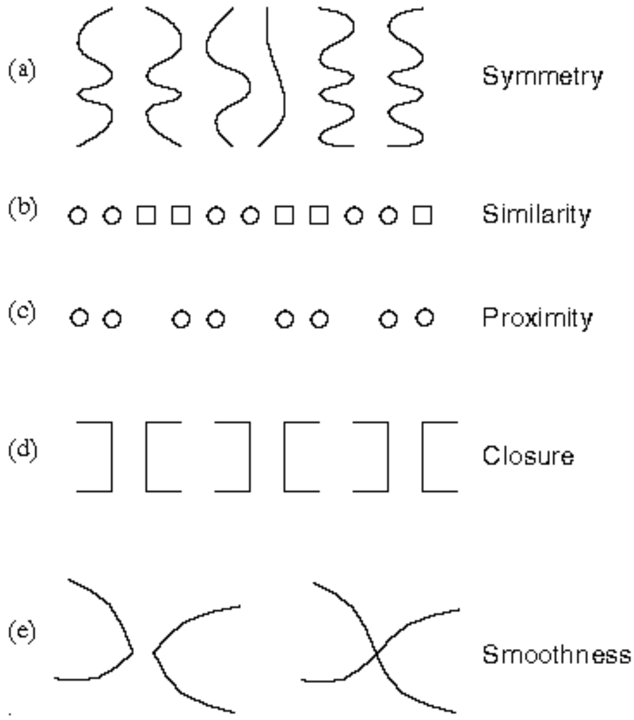


Back to a real brain: The visual system

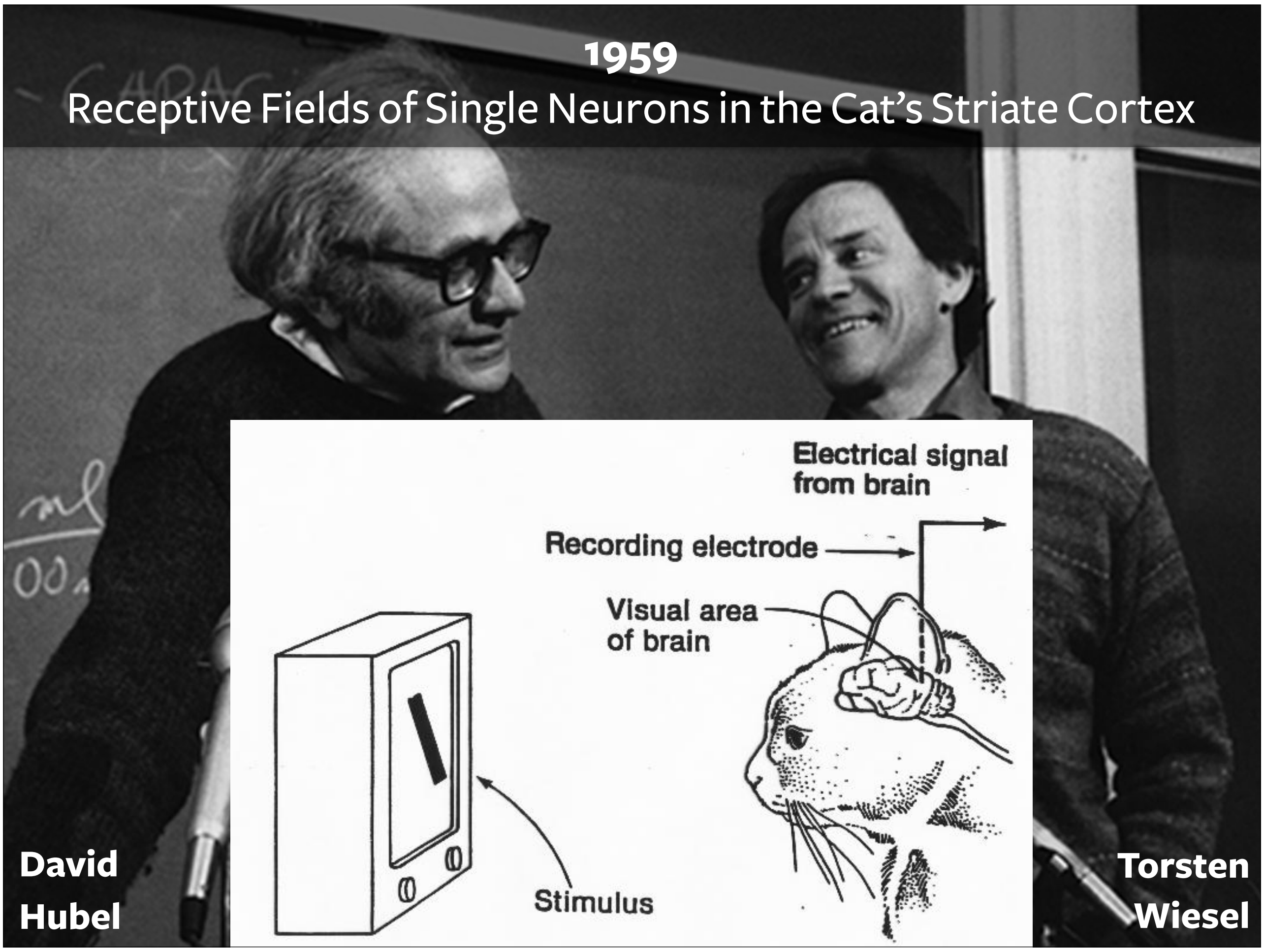

OPTICAE
THESAVRVS.
ALHAZENI
ARABIS
libri septem, nunc primùm
editi.
EIVSDEM liber DE CREPVSCVLIS
& Nubium ascensionibus.
ITEM
VITELLONIS
THVRINGOPOLONI
LIBRI X.
Omnes instaurati, figuris illustrati & aucti, adiectis etiam in
Alhazenum commentarijs,
A'
FEDERICO RISNERO.
Circa AD 100 - 1000
Understanding of the visual system gross anatomy

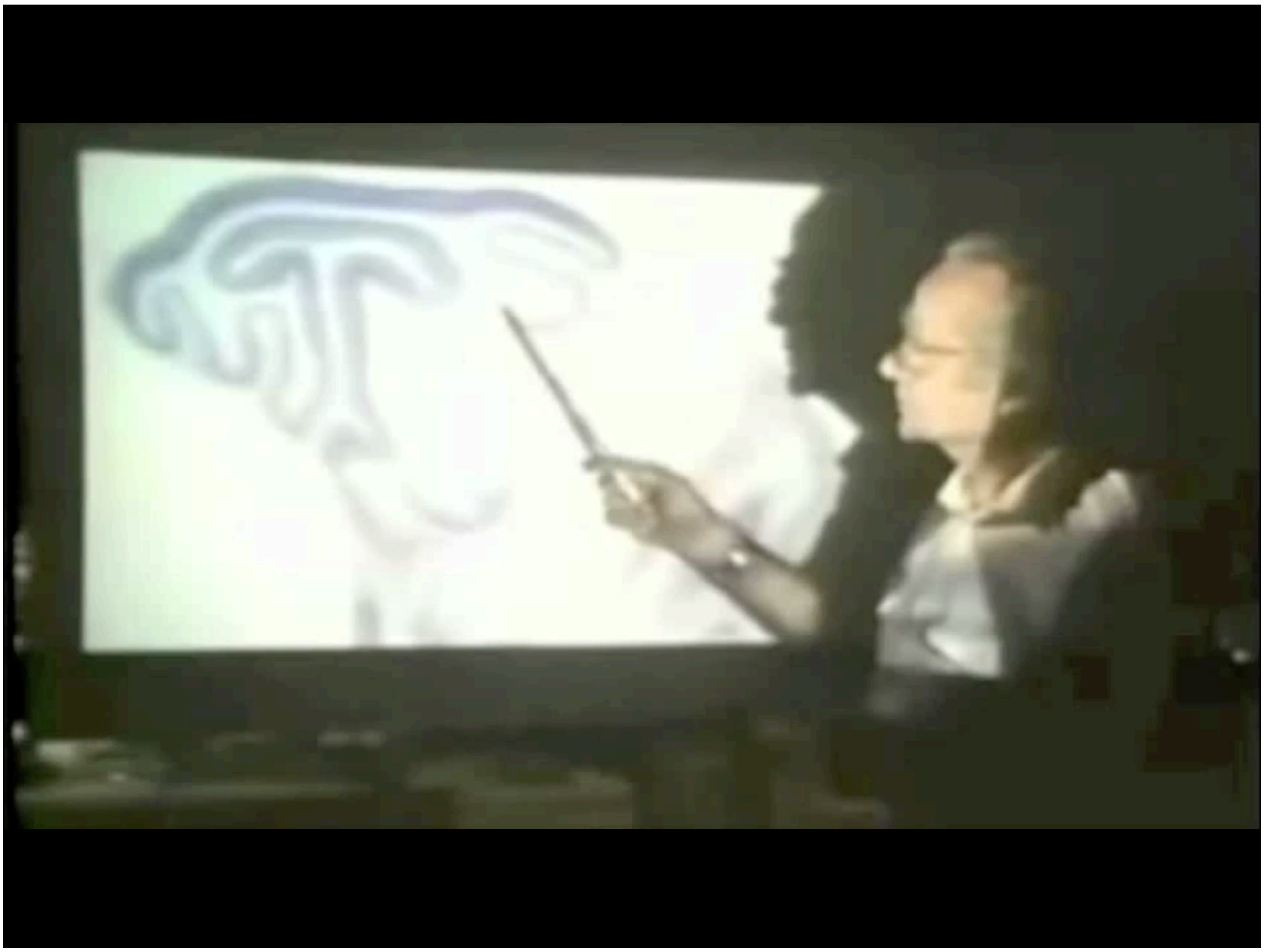

Circa AD 1500
Foveal and peripheral vision
“The function of the human eye ... was described by a large number of authors in a certain way. But I found it to be completely different.”
–Leonardo Da Vinci

1930's
Gestalt Laws of Perceptual Grouping



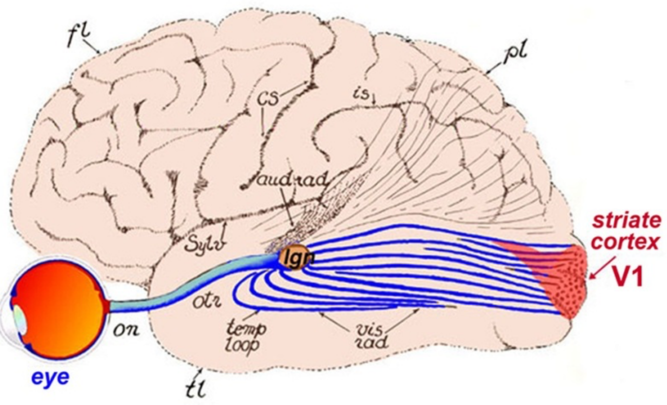
1959
Receptive Fields of Single Neurons in the Cat's Striate Cortex

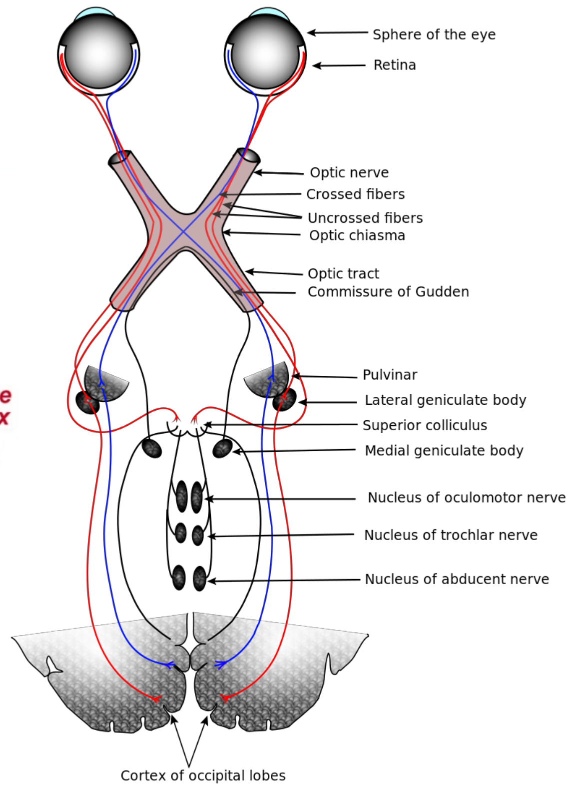




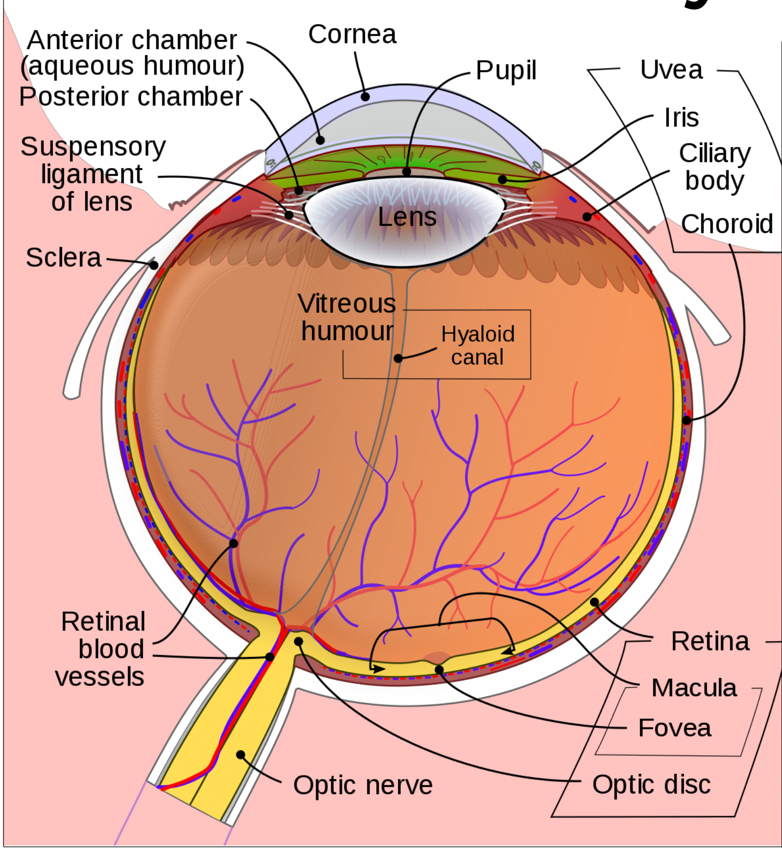
A modern understanding of the visual system

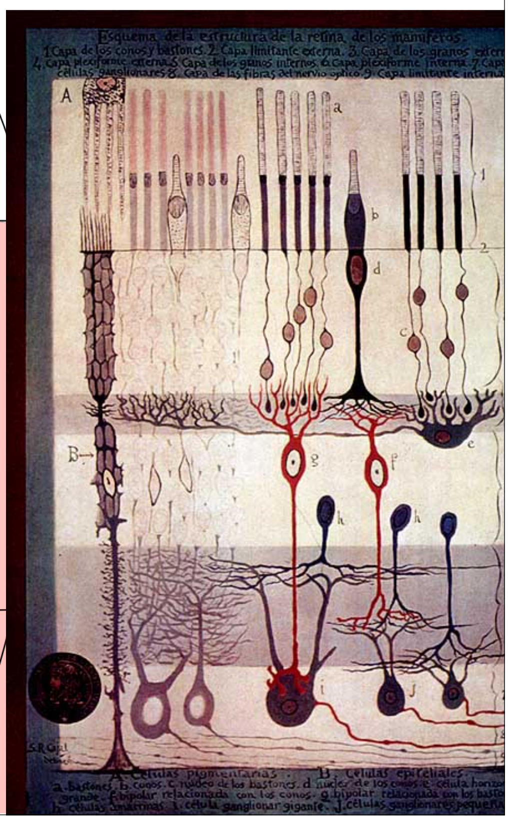
Gross Neuroanatomical Features



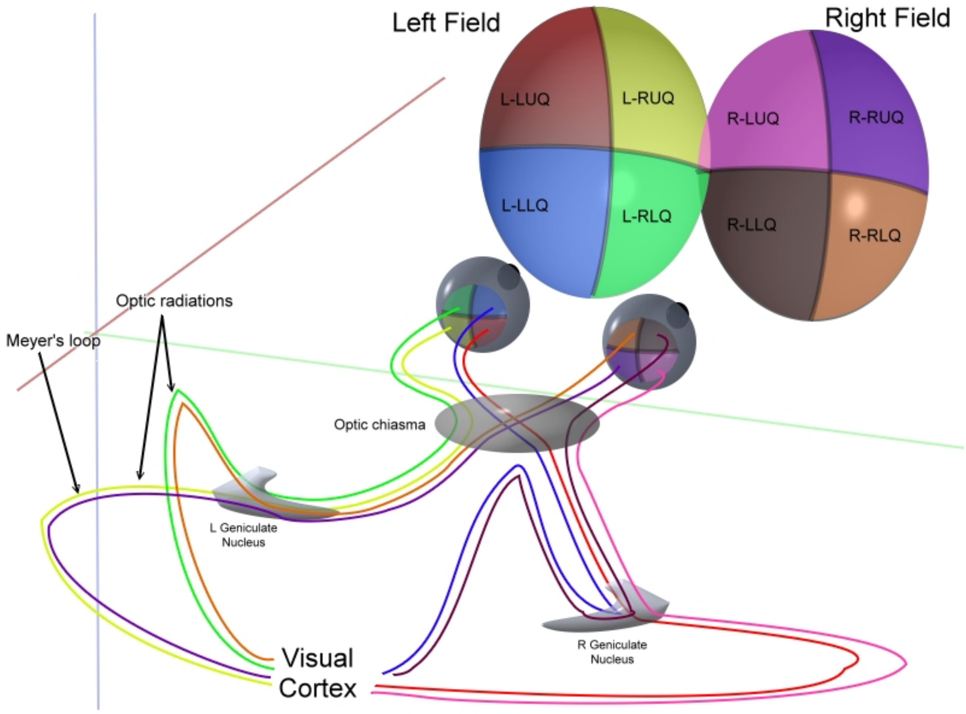


Neuroanatomy: the retina





Neuroanatomy: information flow along the optic nerve



Bandwidth of the optic nerve

- Estimated to be 8960kbps (https://www.newscientist.com/article/dn9633-calculating-the-speed-of-sight/)
  - ~1.1 megabytes per second
  - An uncompressed 640*480 8-bit RGB image is 640*480*3 bytes = 0.92 megabytes

Neuroanatomy: the striate cortex (aka primary visual cortex or V1)

- Receives information from the LGN, with retinotopic mapping preserved
- Characterised by a layered structure of cells organised into "hypercolumns"
  - Small "receptive fields"
  - Neurons adapted to firing on relatively simple features like edges of specific orientations

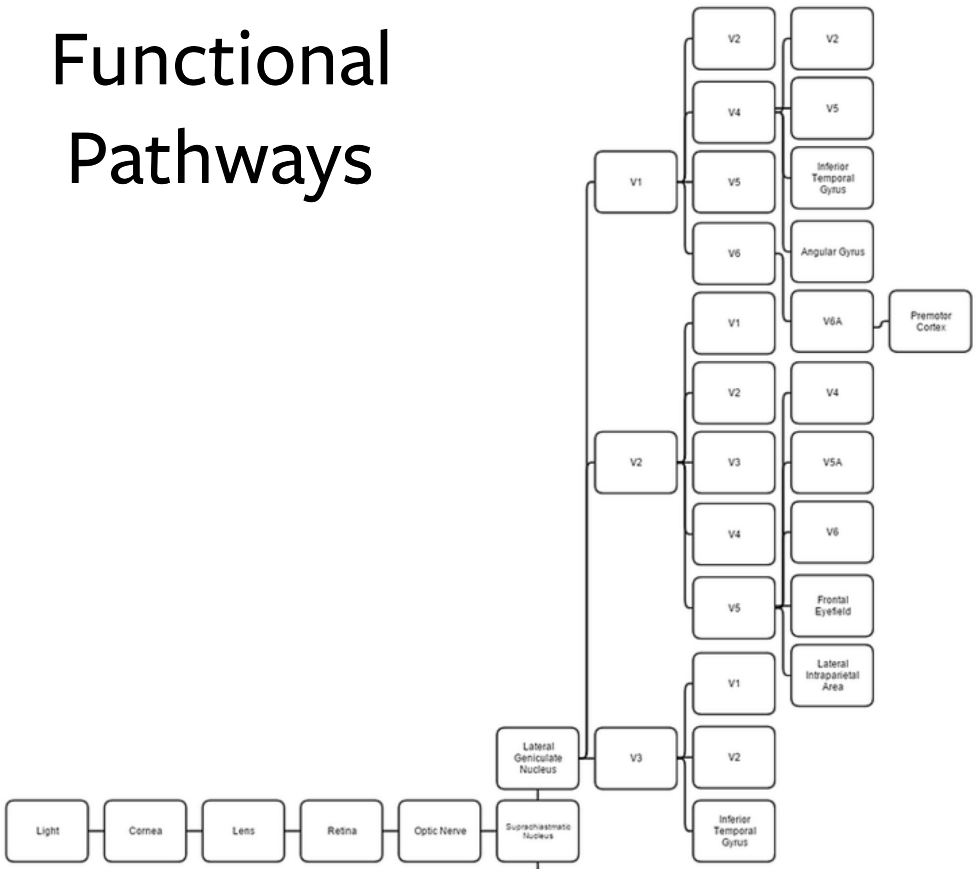
Neuroanatomy: the extrastriate cortex (aka V2-V5)

- V2: local receptive fields, forward connections to V3, V4 and V5, backward connections to V1
  - Cells tuned to moderately complex patterns
- V3: lots of controversy to what the extent of this bit is & what it does!
- V4: attentional modulation; tuned to moderately complex object features
- V5/Middle Temporal: cells sensitive to movement and direction
- V6/Dorsomedial: processing of ego-motion

Neuroanatomy: the Inferior Temporal Cortex (IT)

- Cells sensitive to specific types of high-level features
  - For example cells that fire when a face is present in the visual field

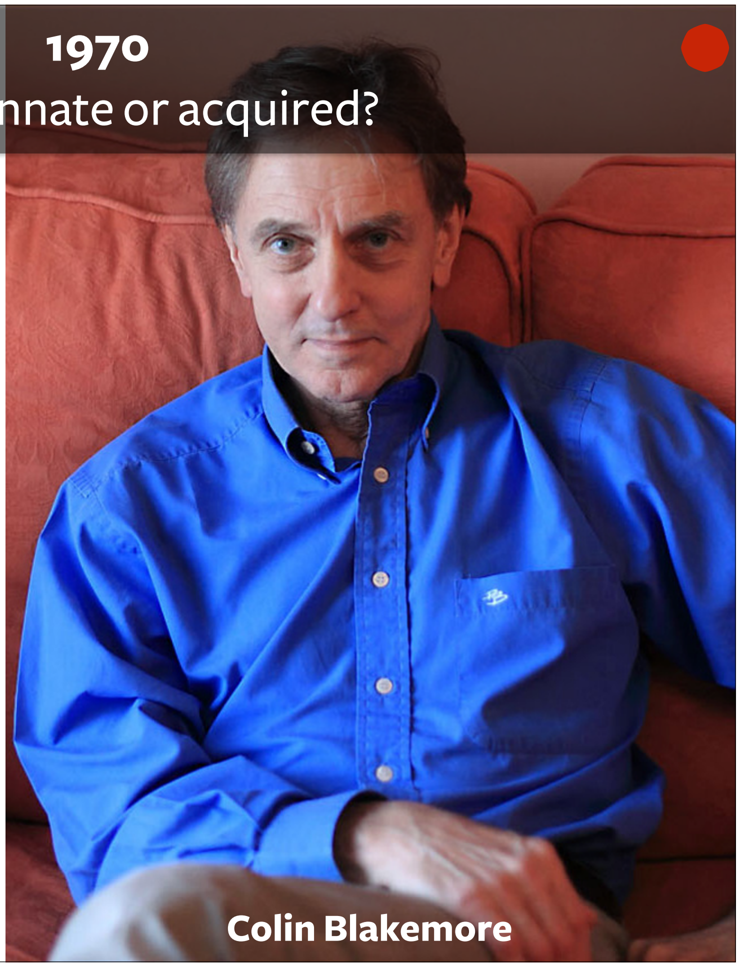
Functional Pathways

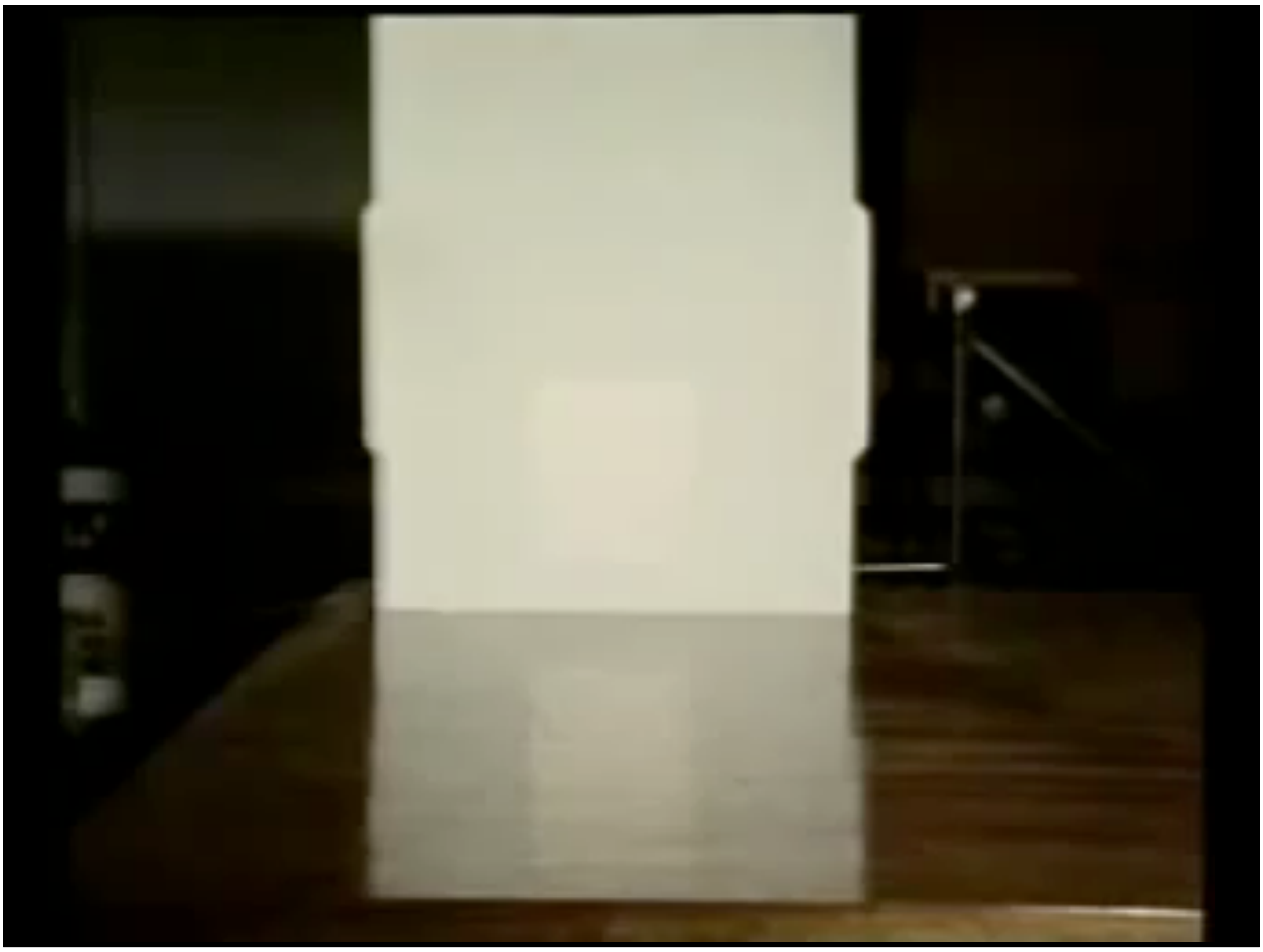


Nature or nurture... Do we “learn” to see?

1970

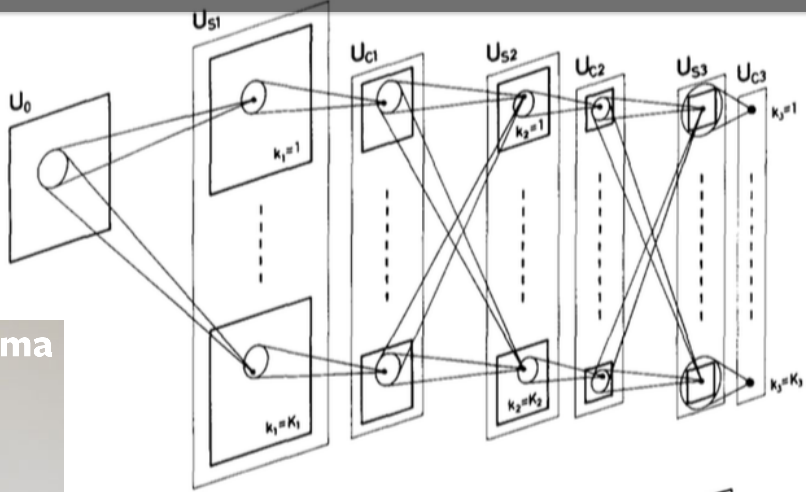
Is vision innate or acquired?

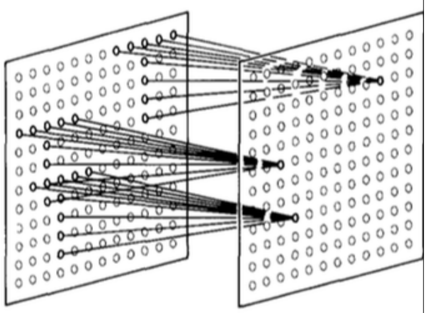




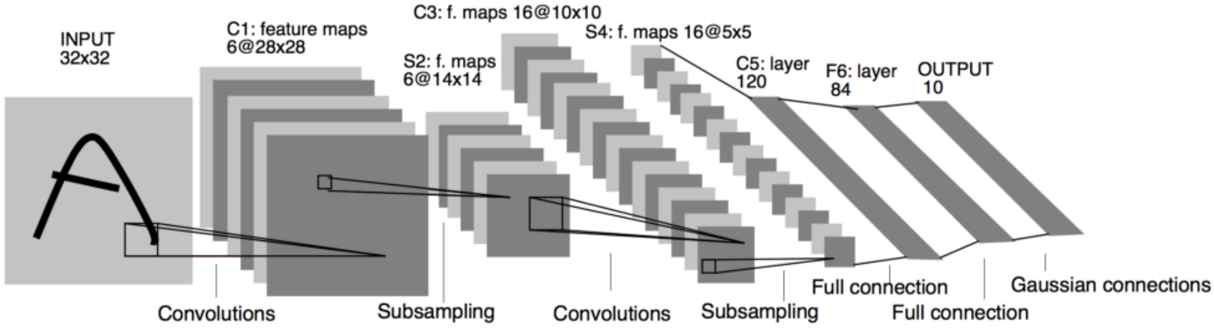
Back to computational models: feature learning, local receptive fields and preserved spatial mappings

1979
Neocognitron





1998
LeNet-5: Convolutional Neural Networks



A word of warning from Marvin Minsky about learning features

Parting words

- Deep learning architectures are definitely not biologically plausible, but they do take ideas from our understanding of the brain.
- We haven't talked about learning… is backprop biologically plausible?
    - Does all learning need to involve gradients?