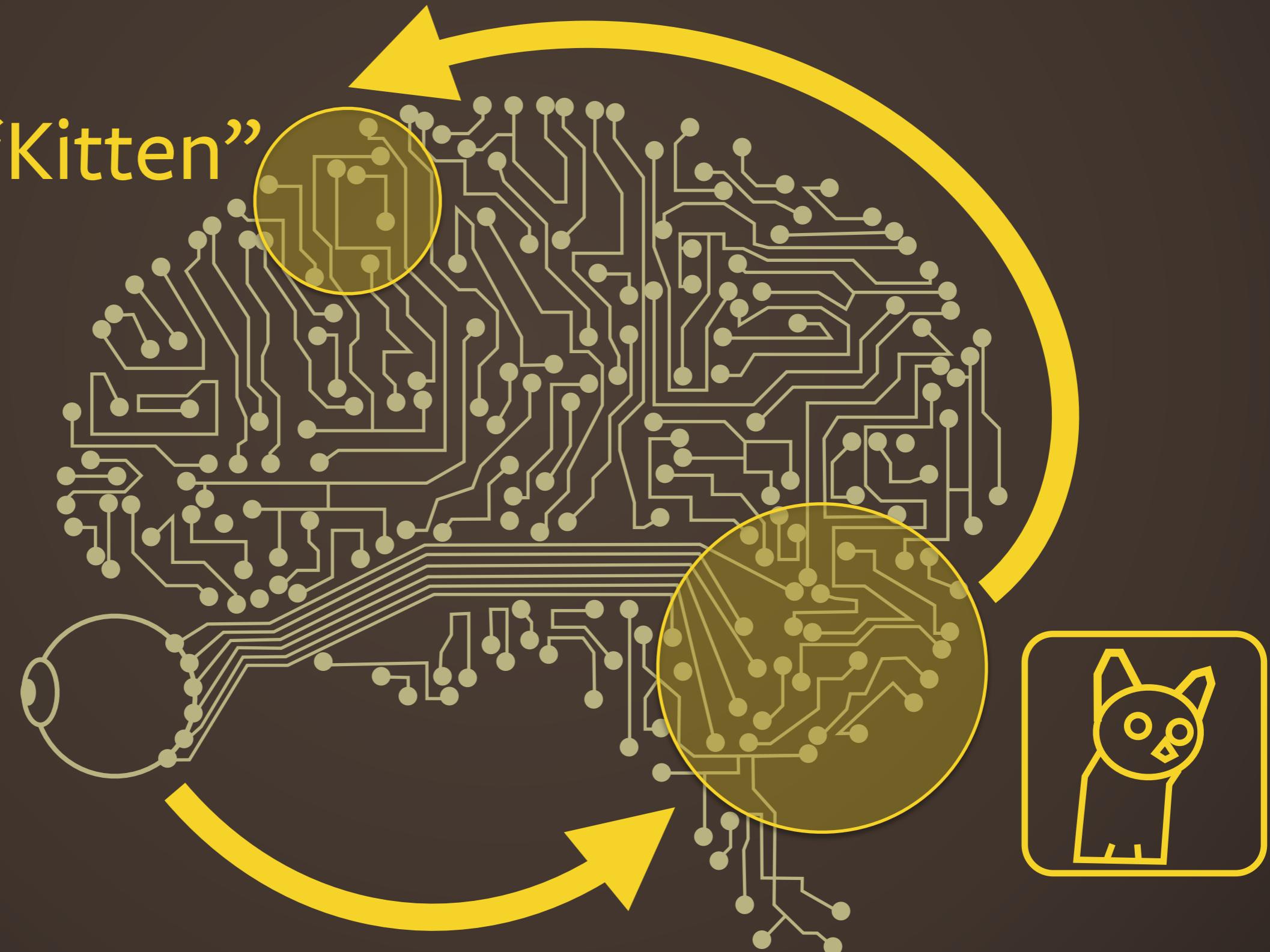


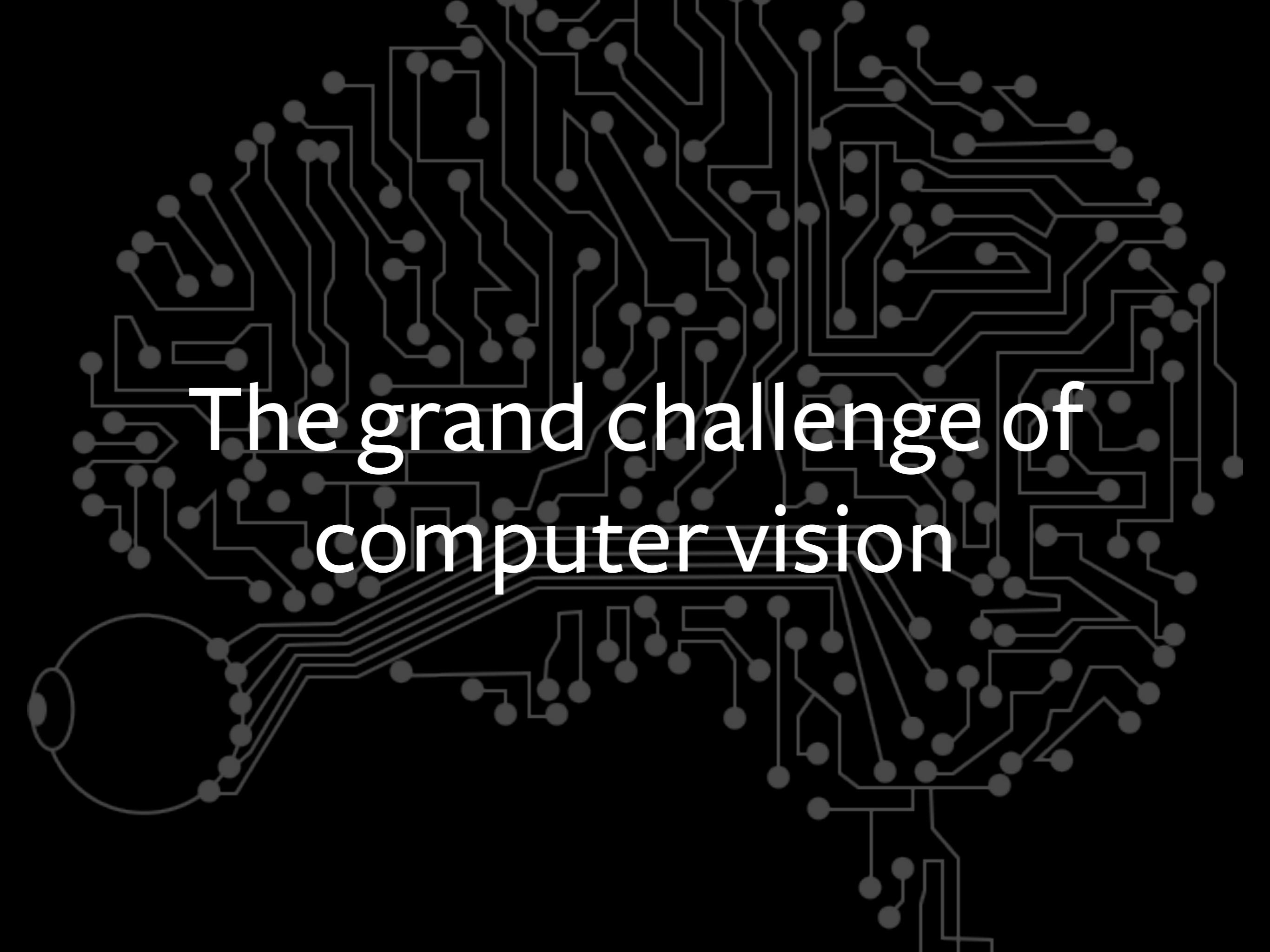
Computer Vision - A Retrospective: *Teaching machines to see*

Dr Jonathon Hare



“Kitten”





The grand challenge of
computer vision

Computer Vision research has always
been inspired by the way humans
“see” and perceive the world





Setting the Scene: A potted history of our understanding of:

- Biological Vision
 - Computation
- Machine Learning
- Computer Vision





Circa 300 BC to AD 200

*emission/extramission
theory*
(championed by Euclid &
Ptolemy)

versus

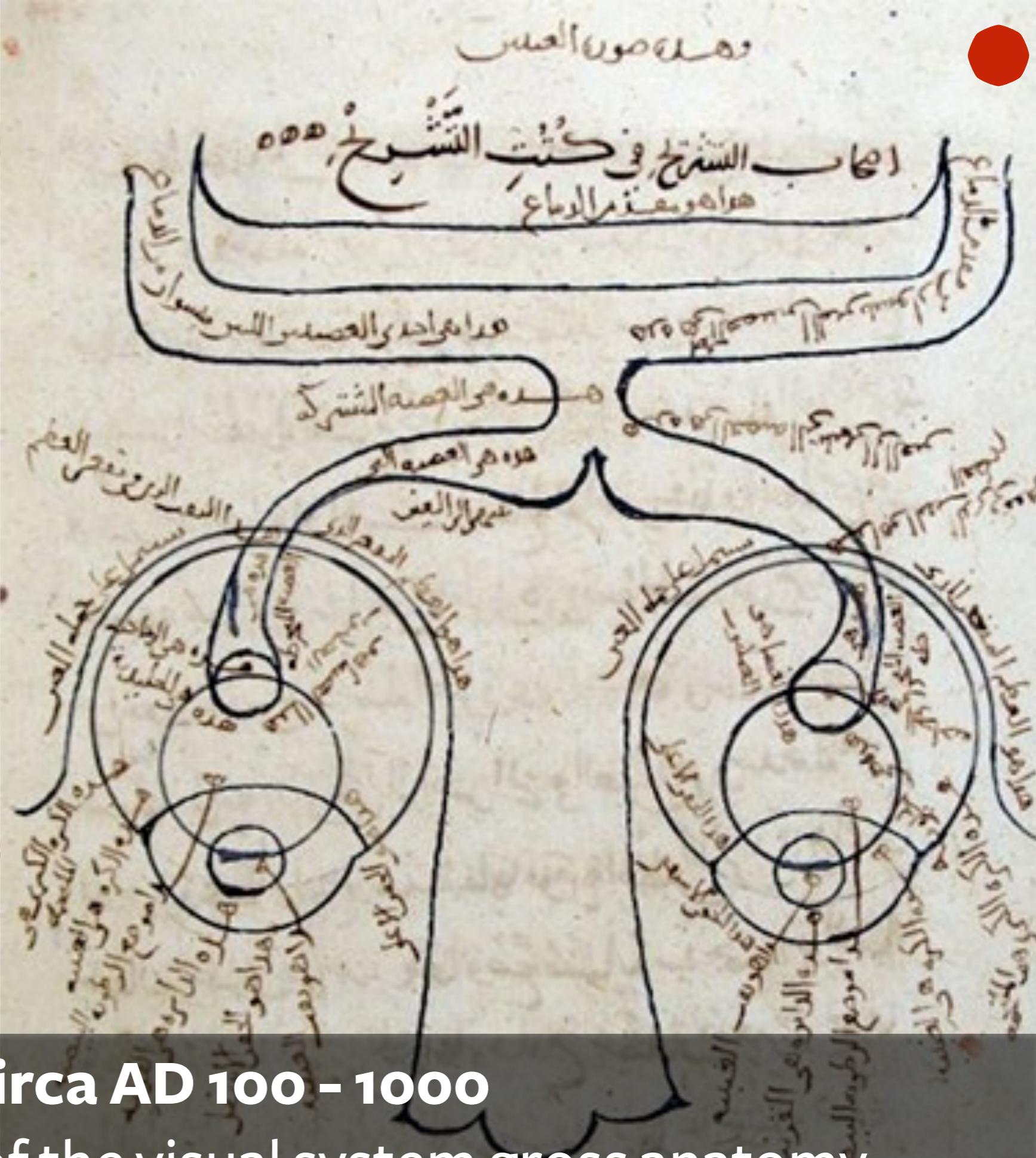
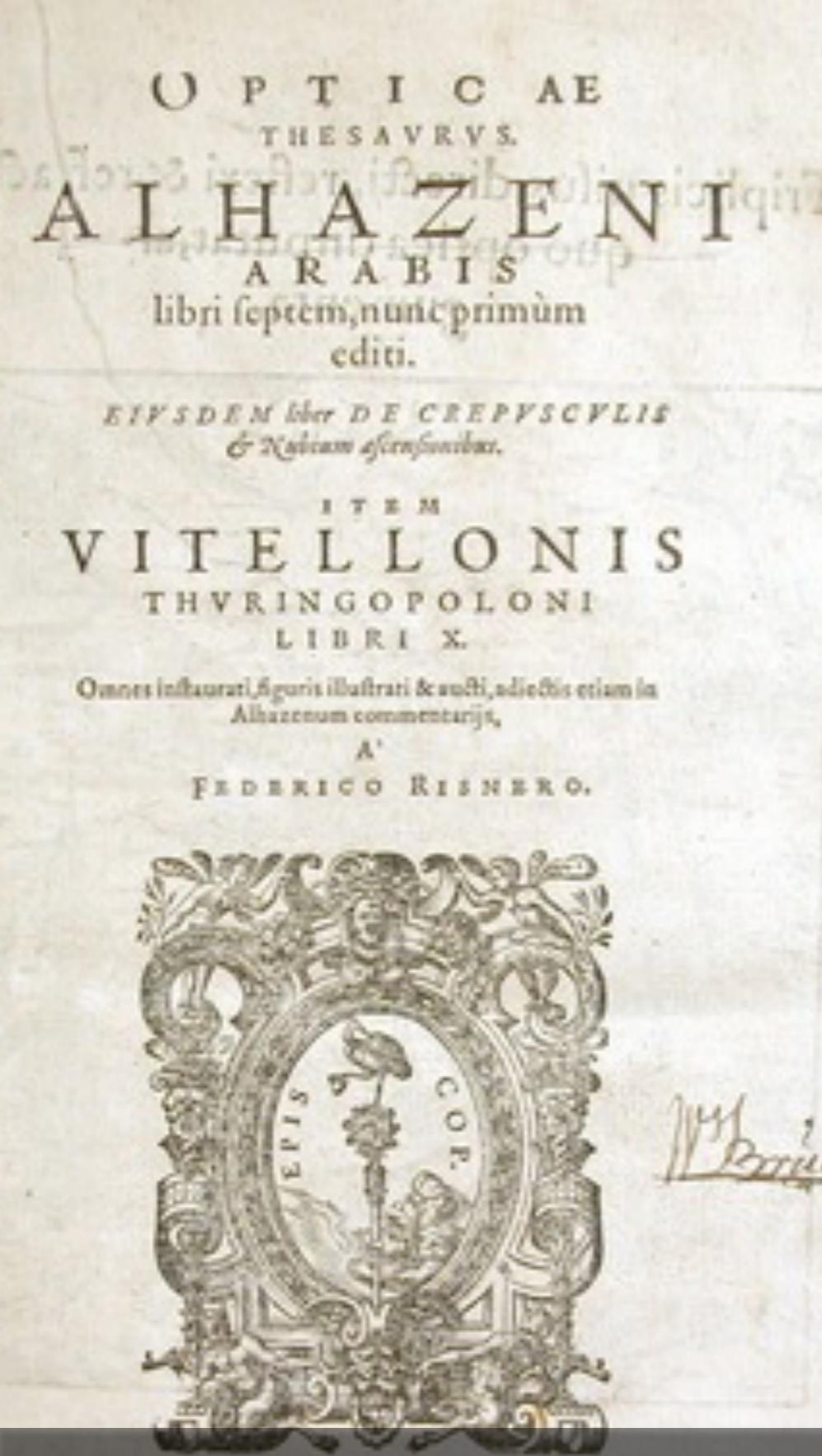
intromission theory
(championed by
Aristotle)



Euclid

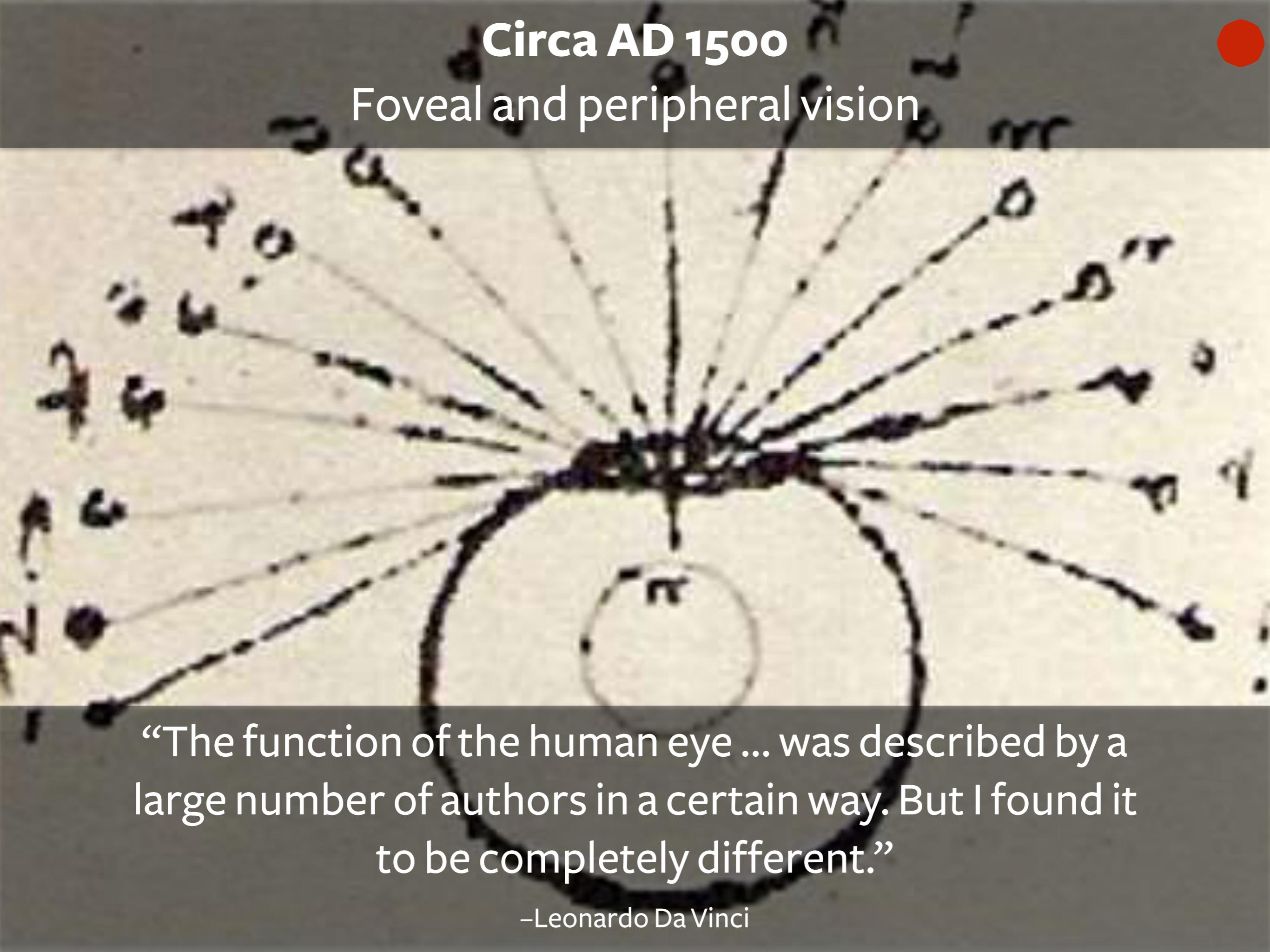


Aristotle



Circa AD 1500

Foveal and peripheral vision

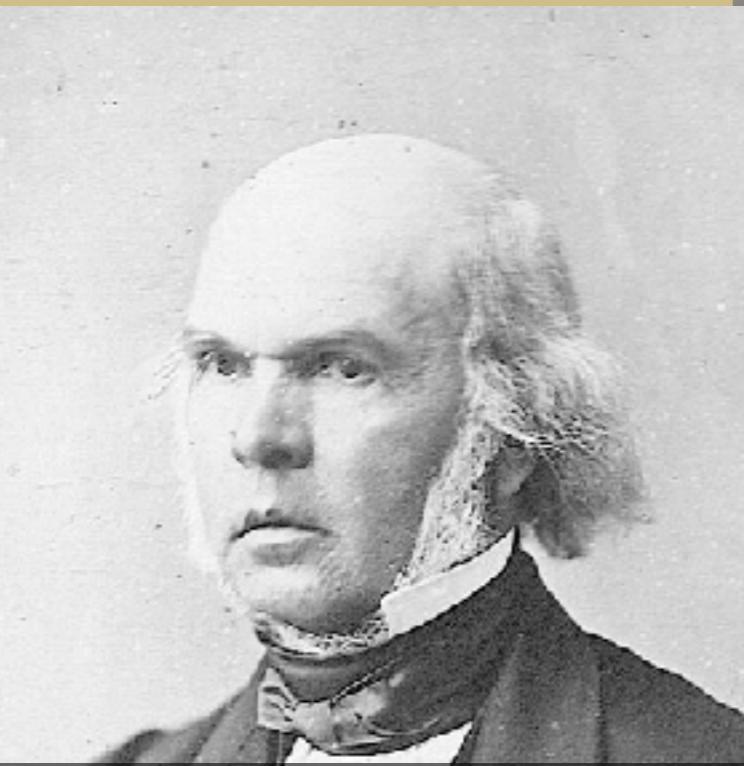
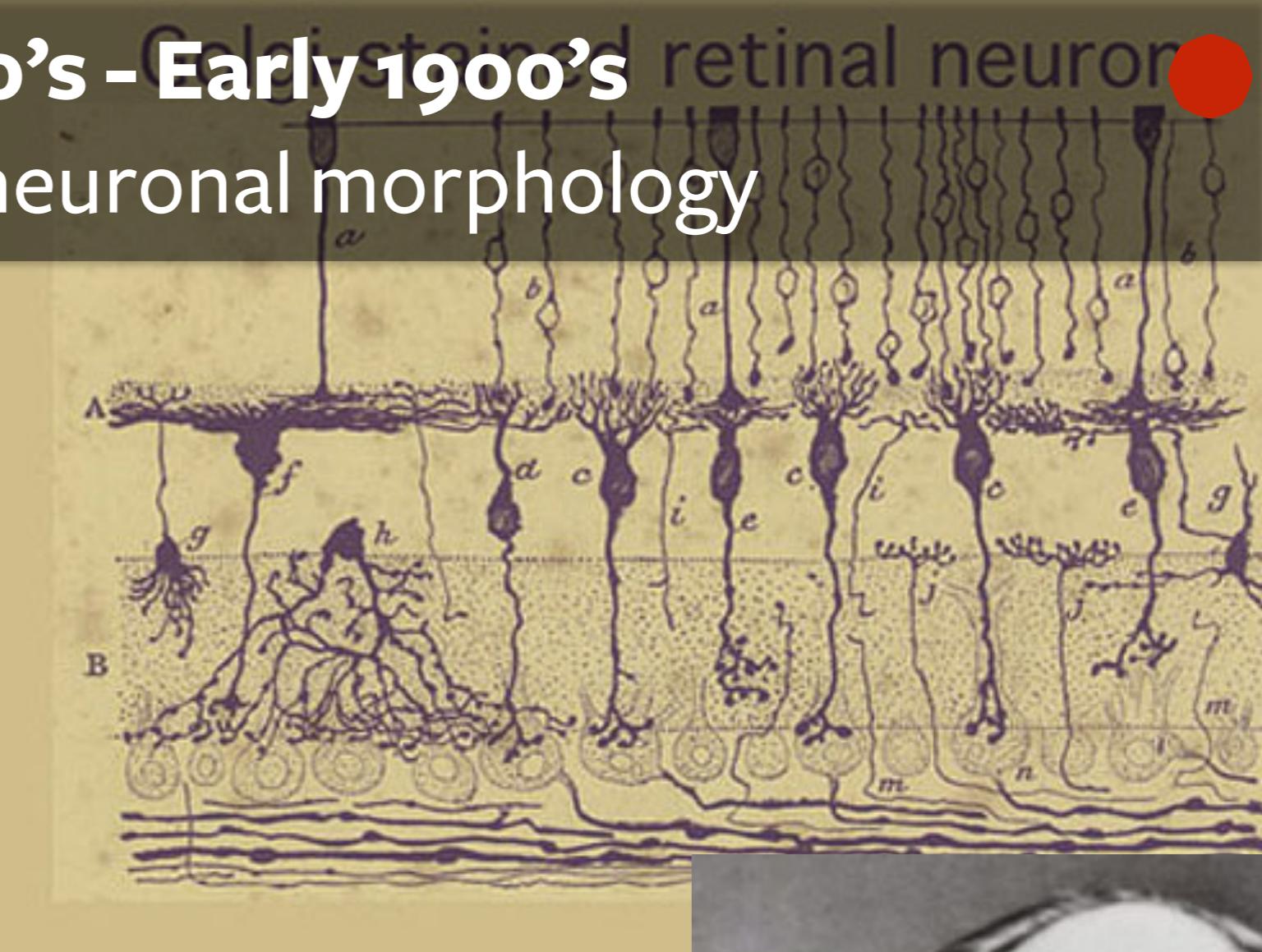
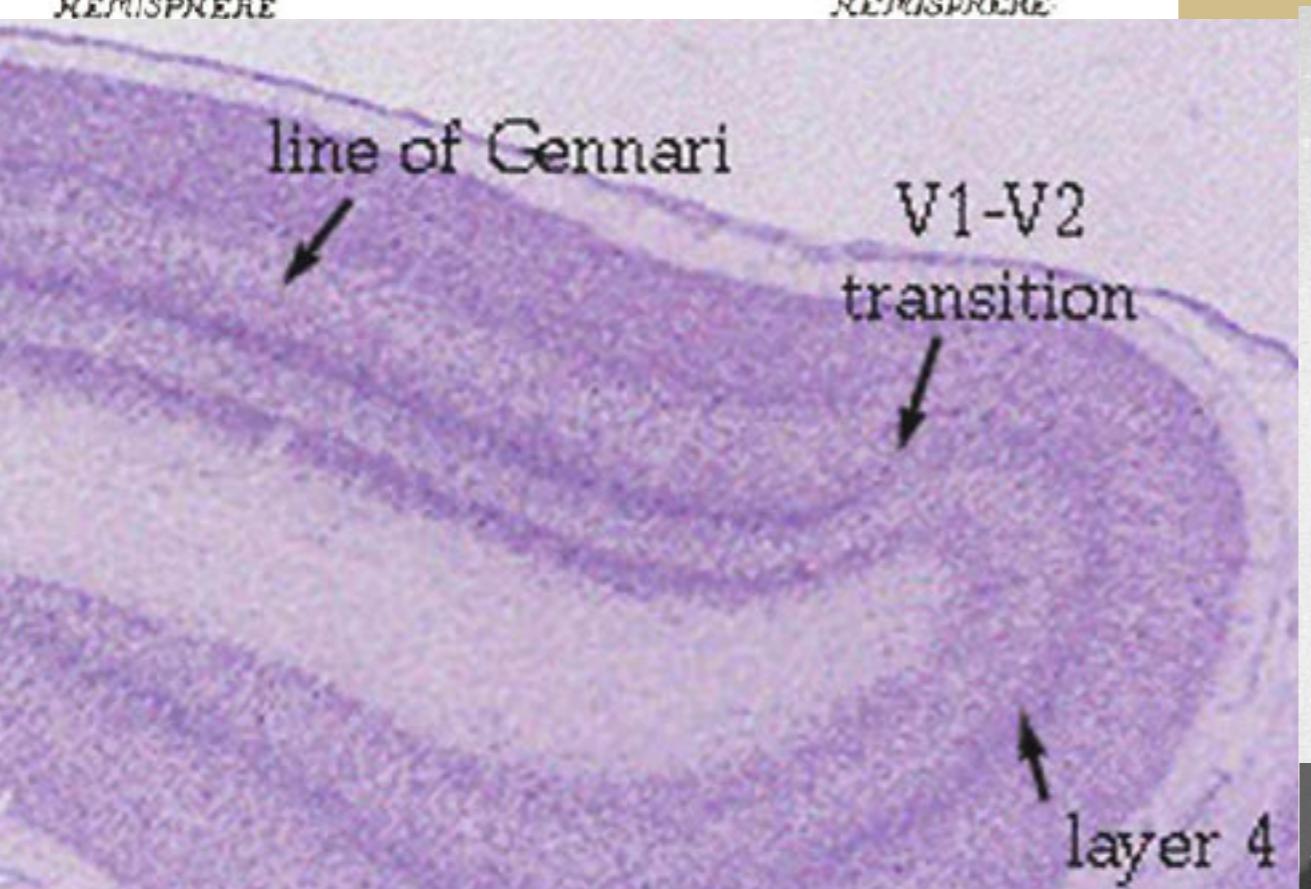
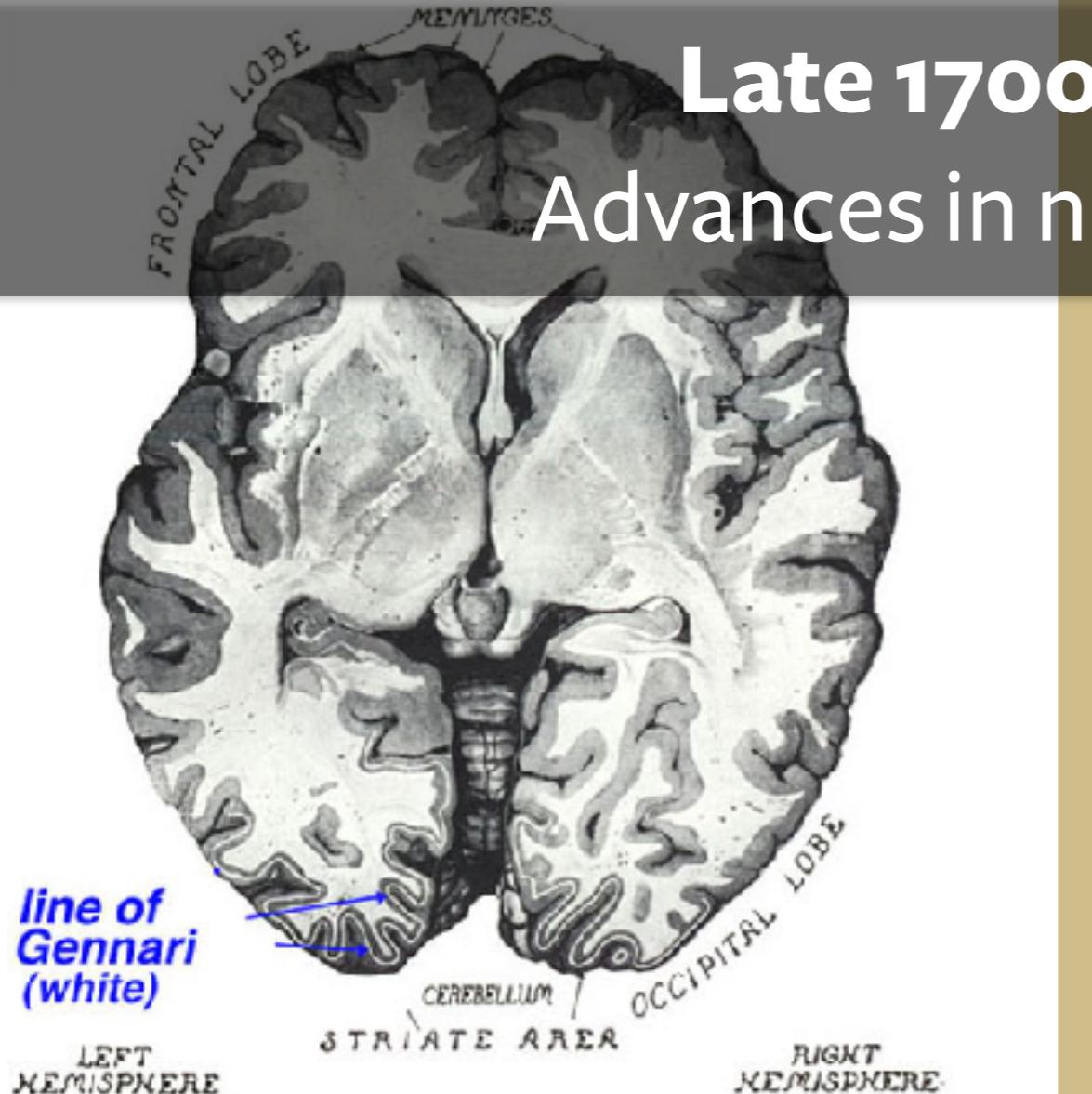


“The function of the human eye ... was described by a large number of authors in a certain way. But I found it to be completely different.”

—Leonardo Da Vinci

Late 1700's - Early 1900's
Colored stained retinal neurons

Advances in neuronal morphology



Jules Baillarger

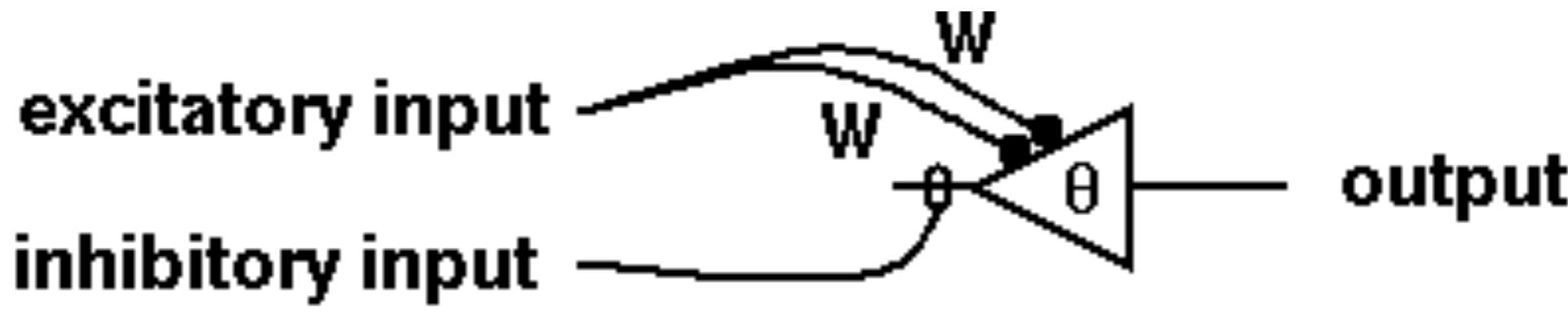


Santiago
Ramón y Cajal

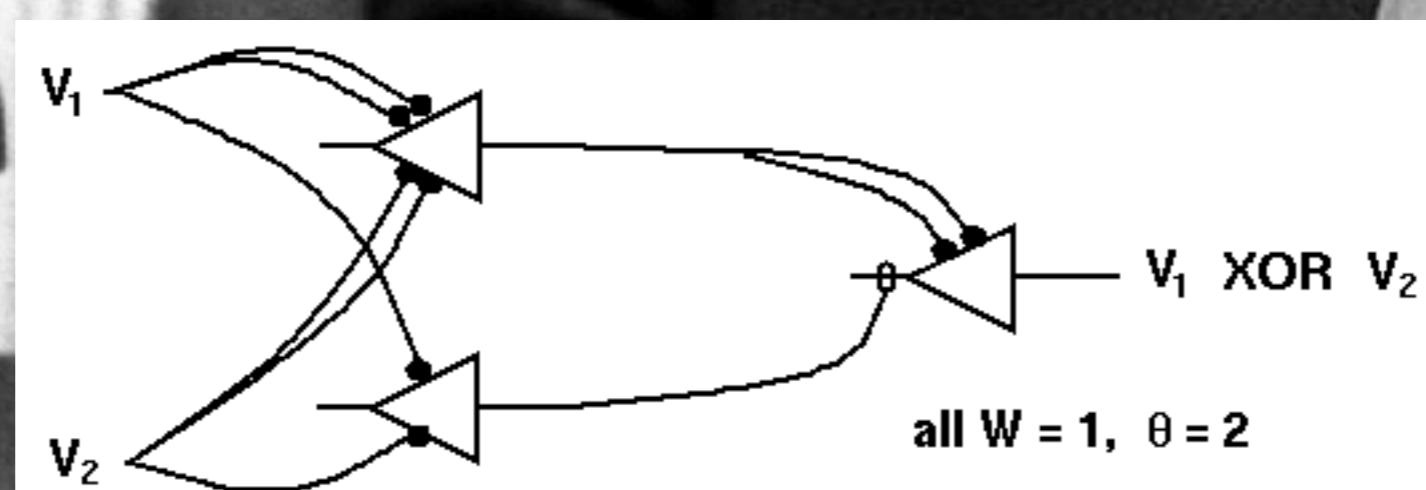
1943

McCulloch-Pitts Artificial Neuron

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



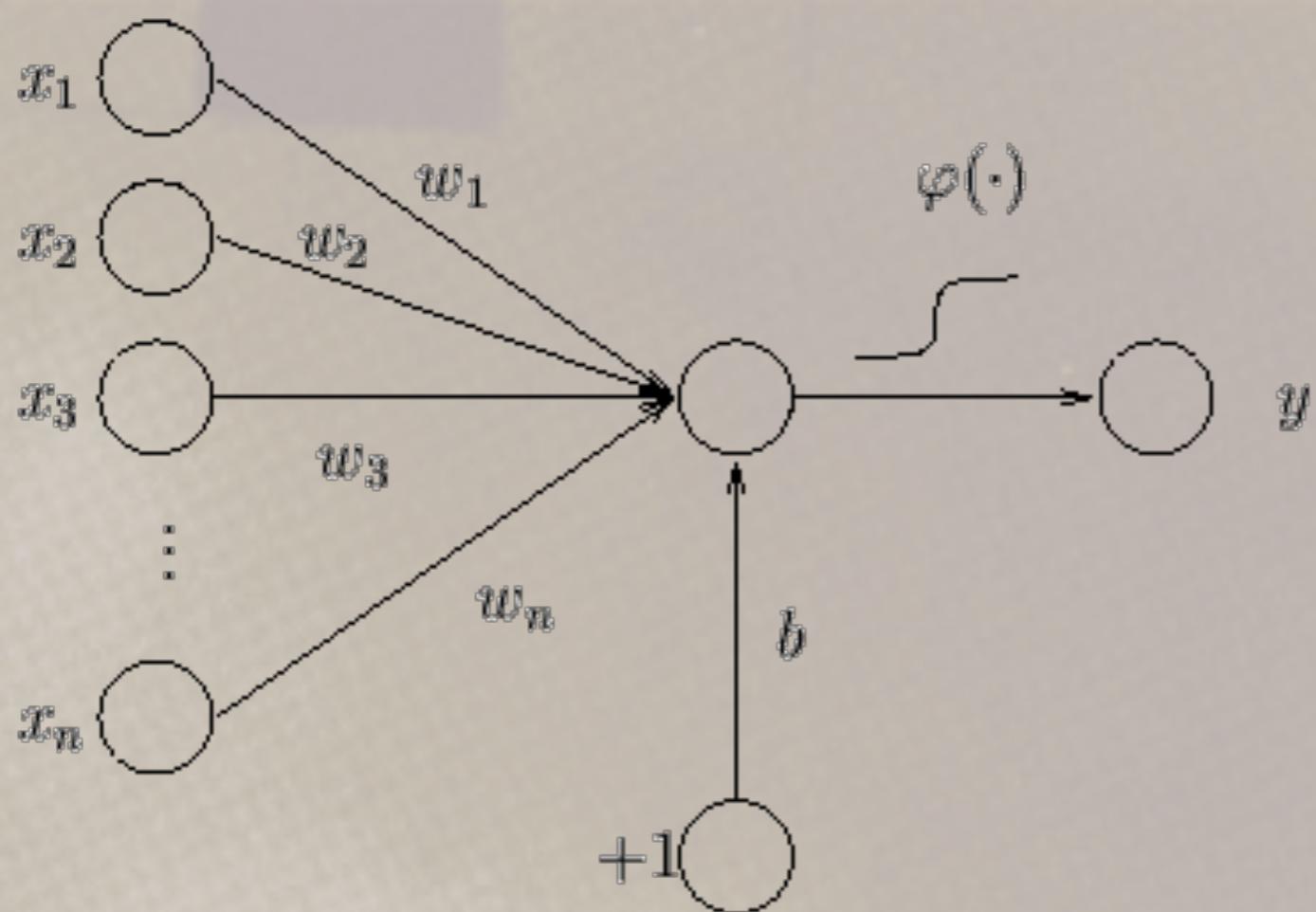
Warren
McCulloch



Walter
Pitts

1958

Rosenblatt's Perceptron



Frank Rosenblatt

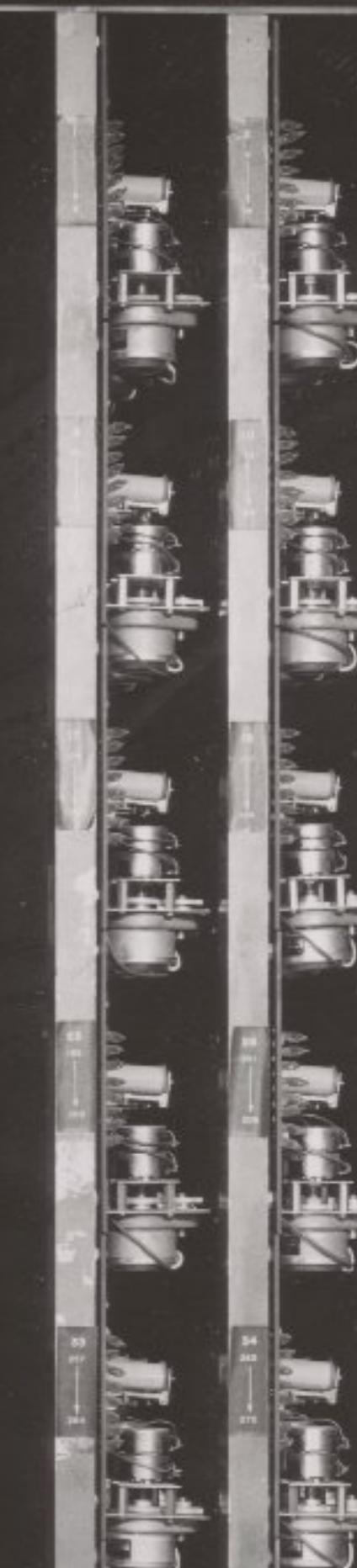
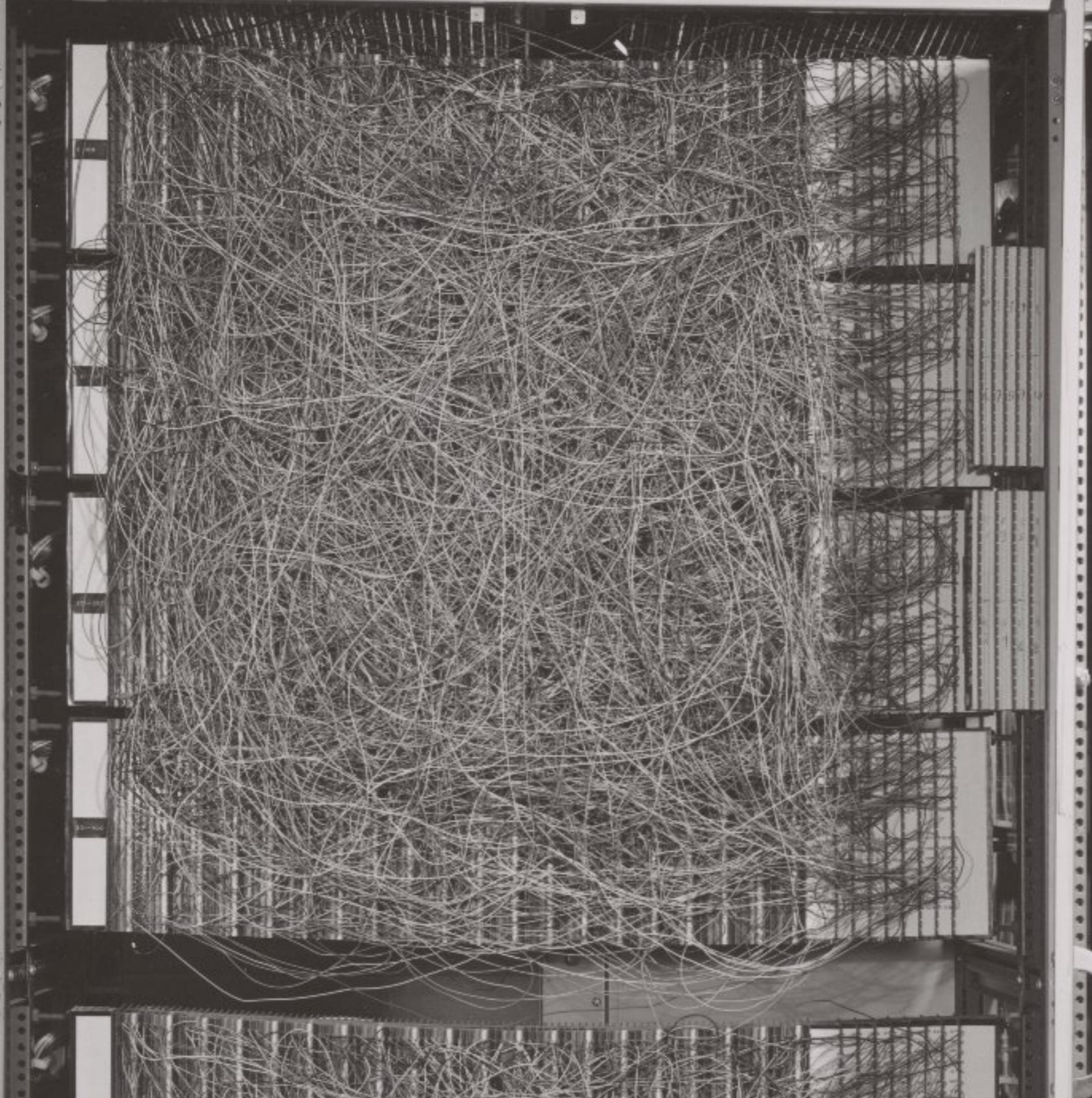


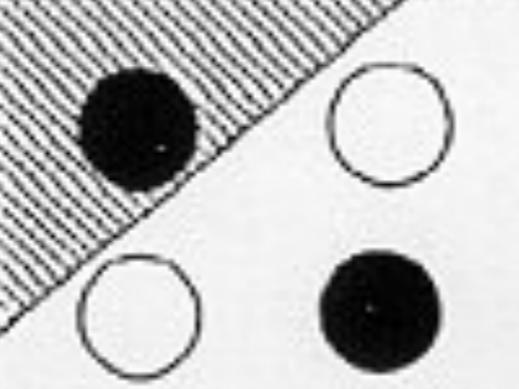
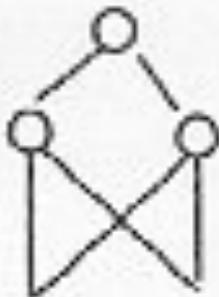
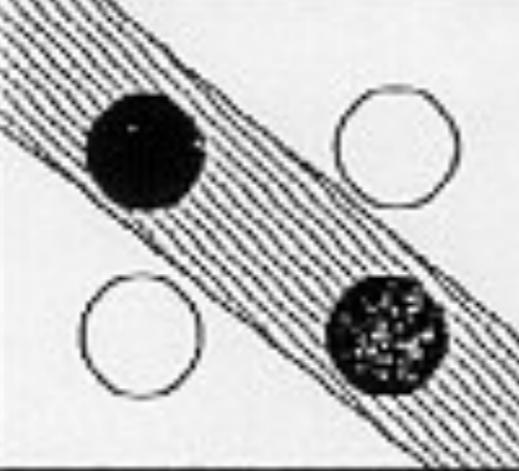
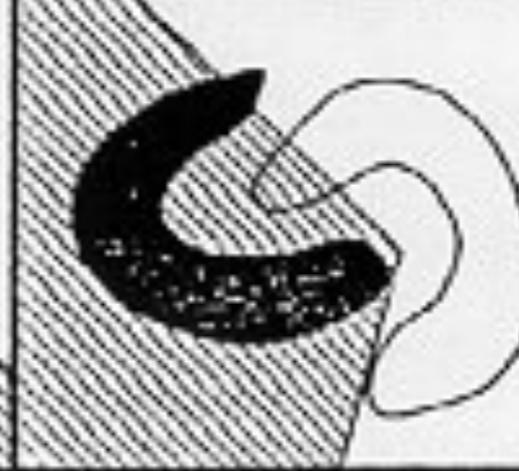
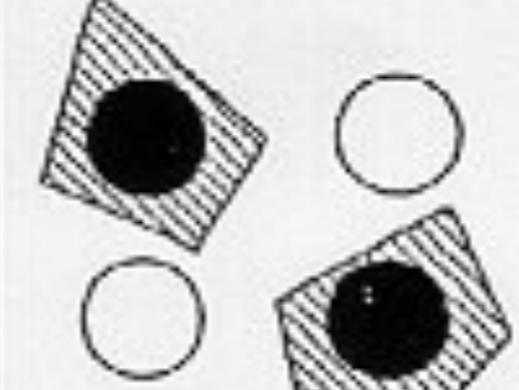
$$y = \varphi\left(\sum_{i=1}^n w_i x_i + b\right) = \varphi(\mathbf{w}^T \mathbf{x} + b)$$



MARK I
CORNELL AERONAUTICAL LABORATORY, Inc.
BUFFALO, NEW YORK

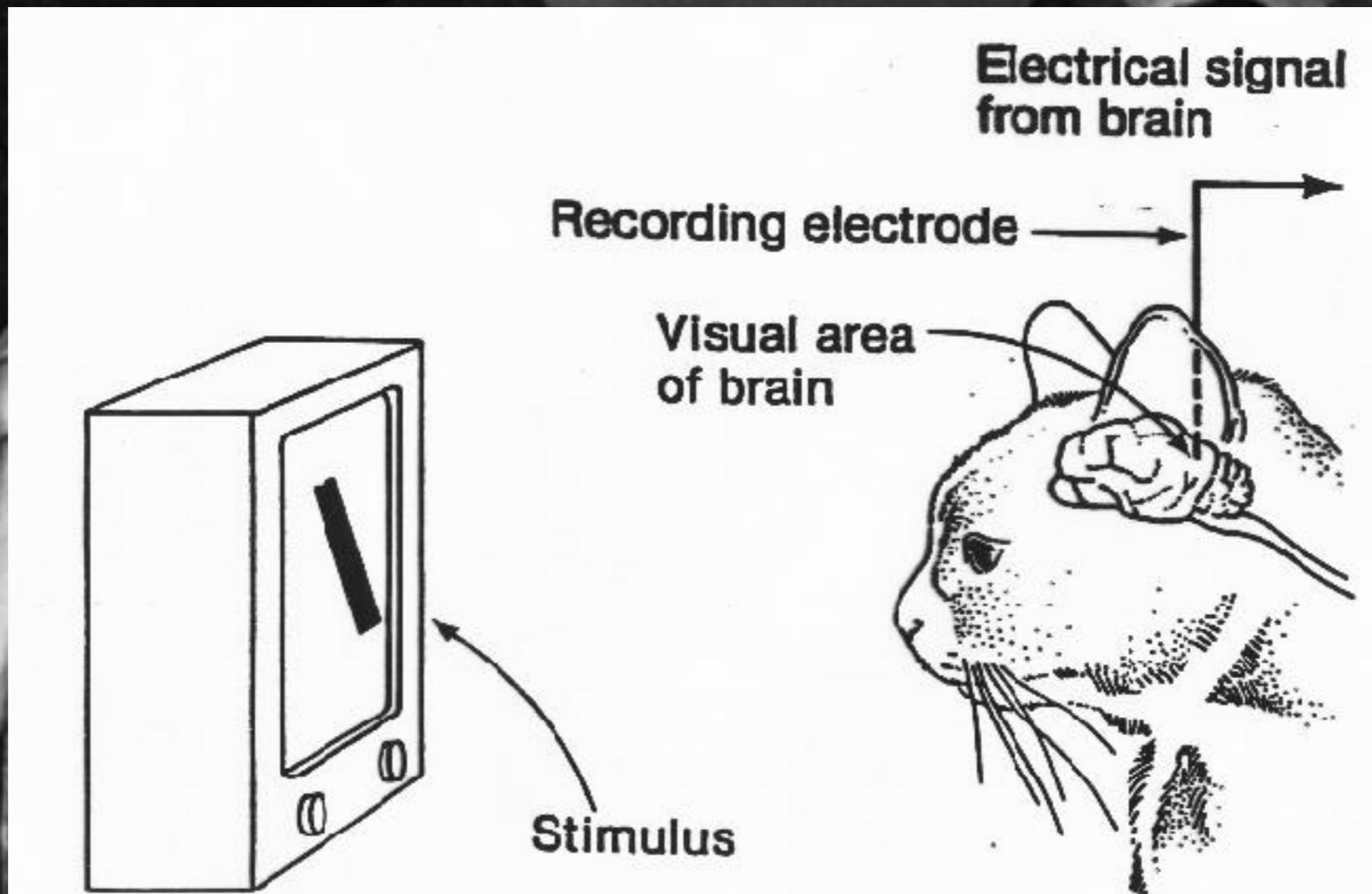
PERCEPTION
AUTICAL LABORATORY, Inc.
EW YORK



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)			

1959

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



David
Hubel

Torsten
Wiesel



1966

Computer vision "summer project"

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
PROJECT MAC

Artificial Intelligence Group
Vision Memo. No. 100.

July 7, 1966

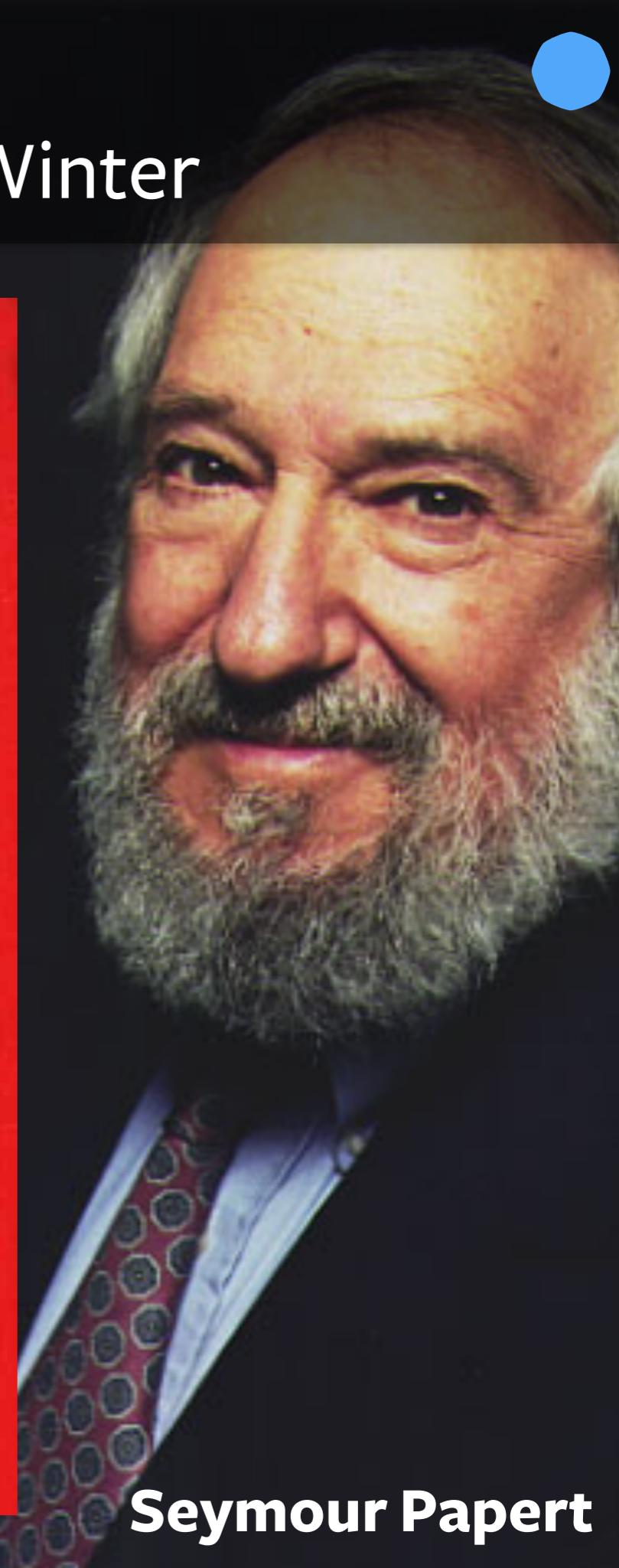
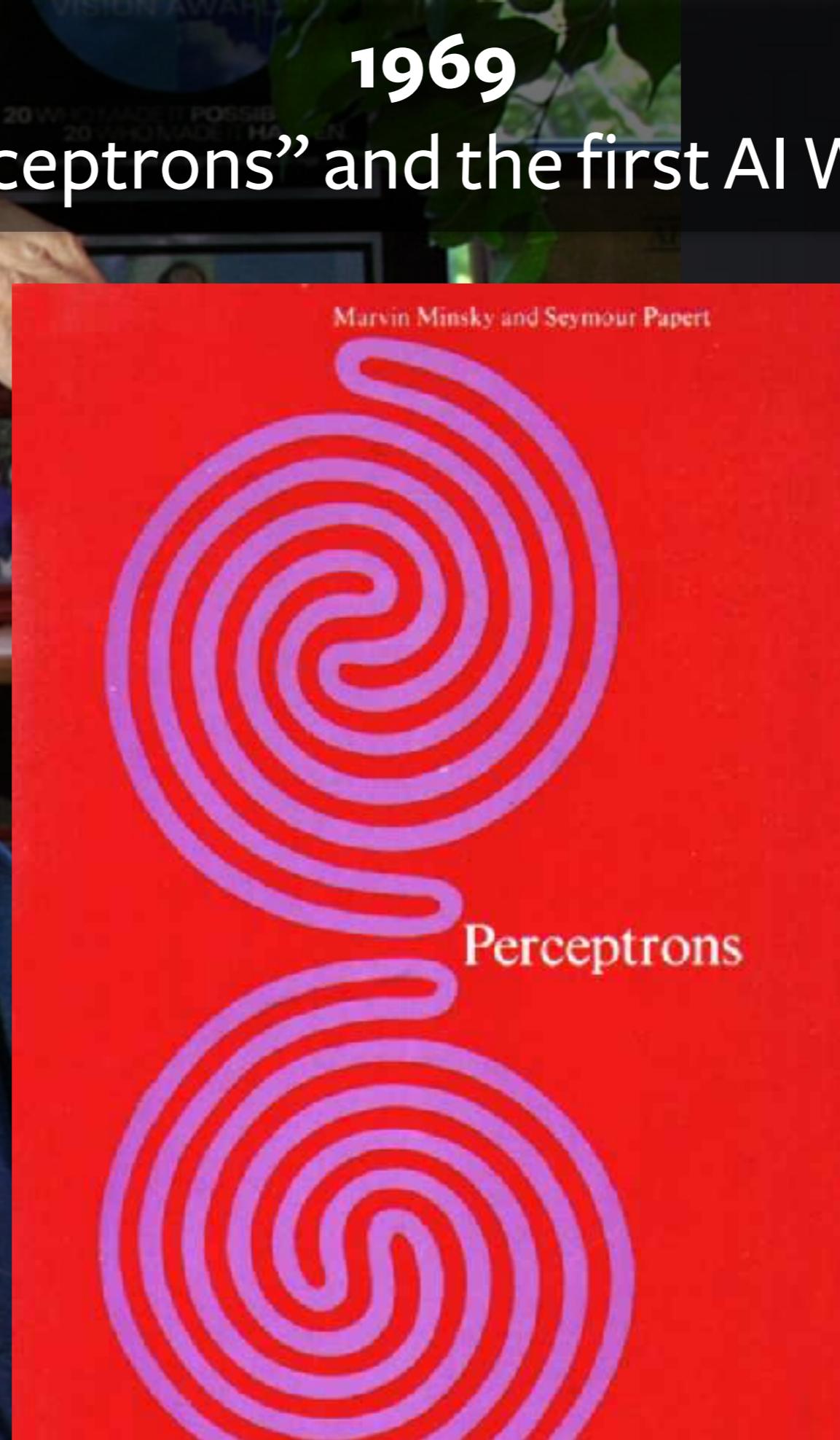
THE SUMMER VISION PROJECT

Seymour Papert

The summer vision project is an attempt to use our summer workers effectively in the construction of a significant part of a visual system. The particular task was chosen partly because it can be segmented into sub-problems which will allow individuals to work independently and yet

1969

“Perceptrons” and the first AI Winter

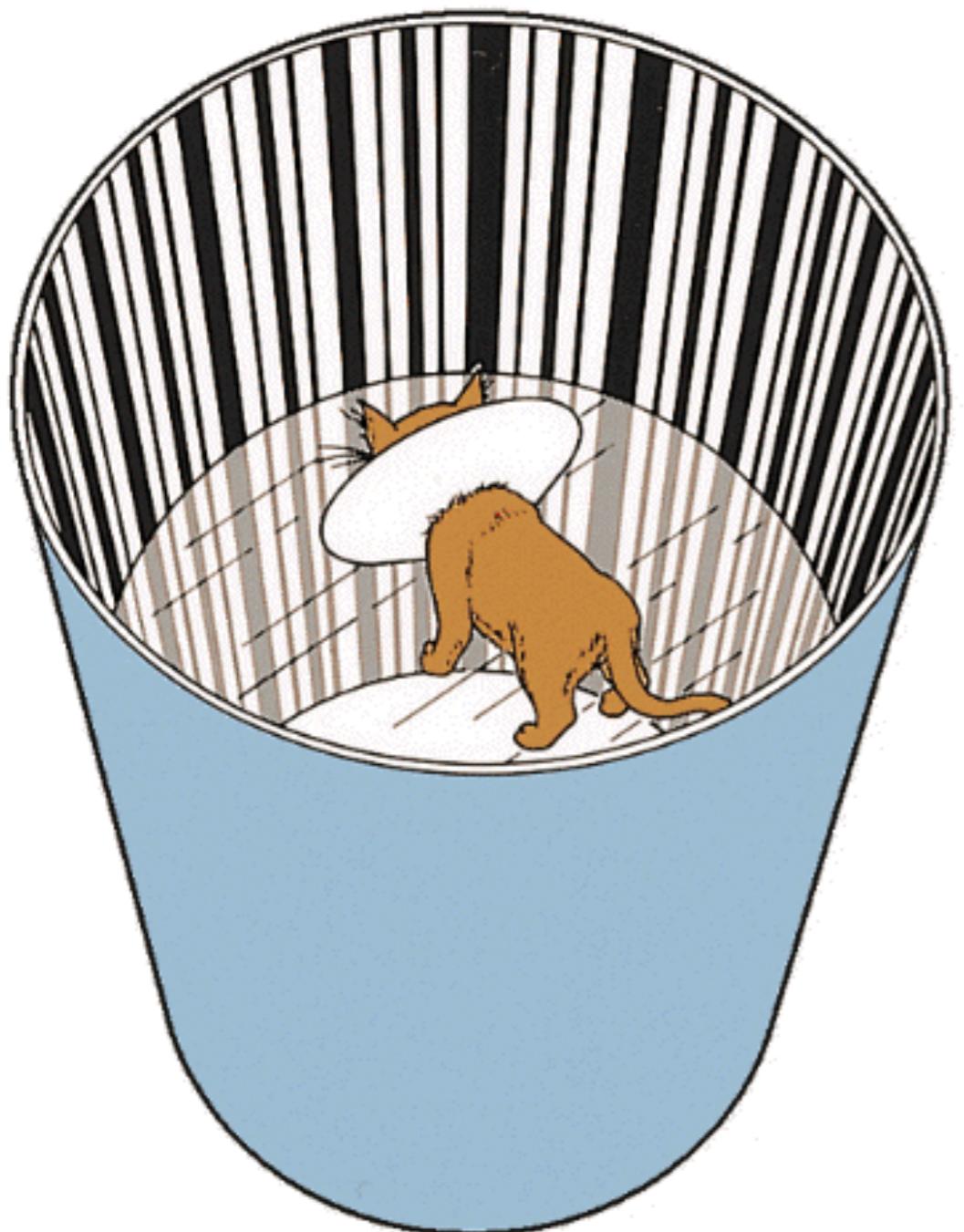


Marvin Minsky

Seymour Papert

1970

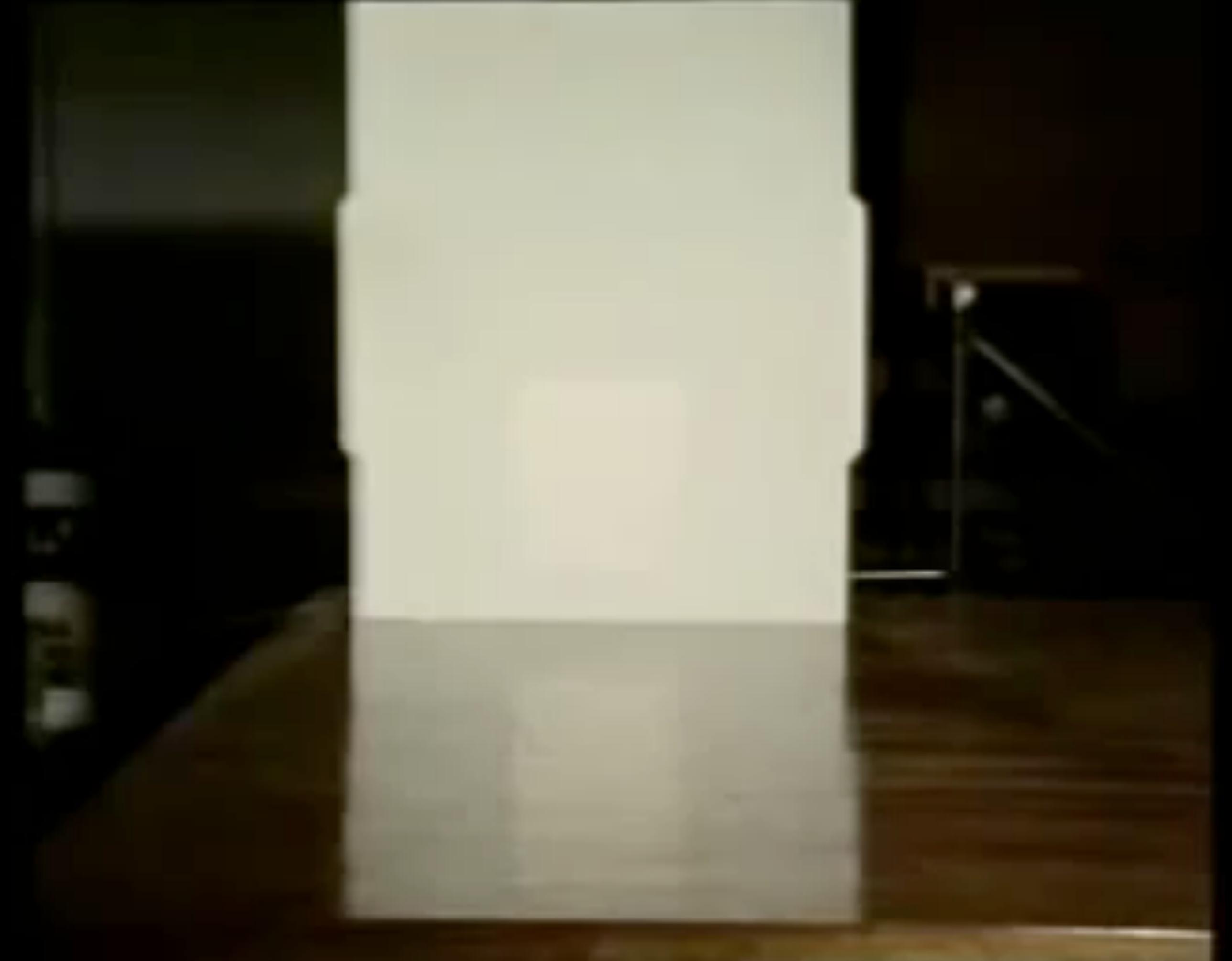
Is vision innate or acquired?



Colin Blakemore



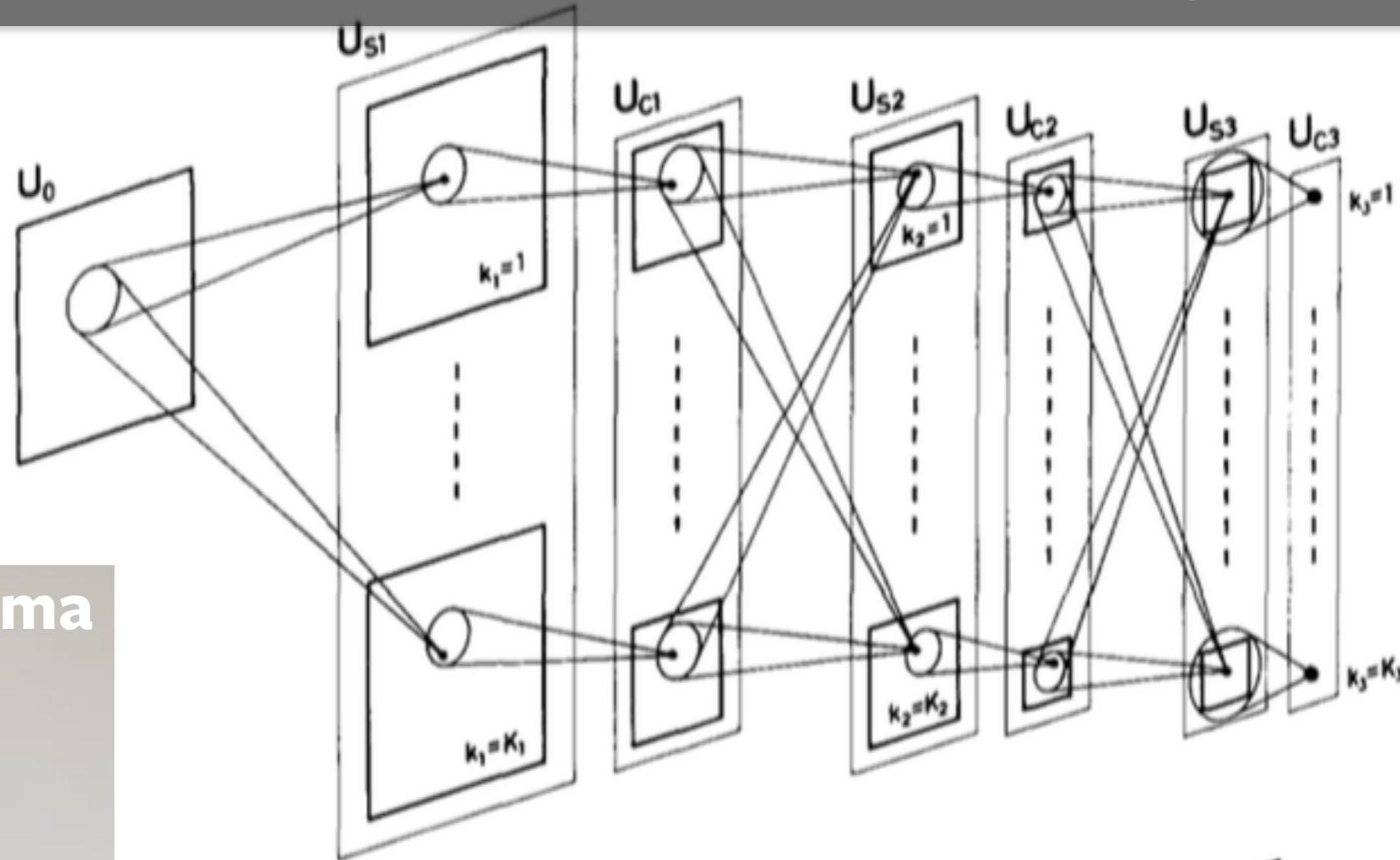
1960



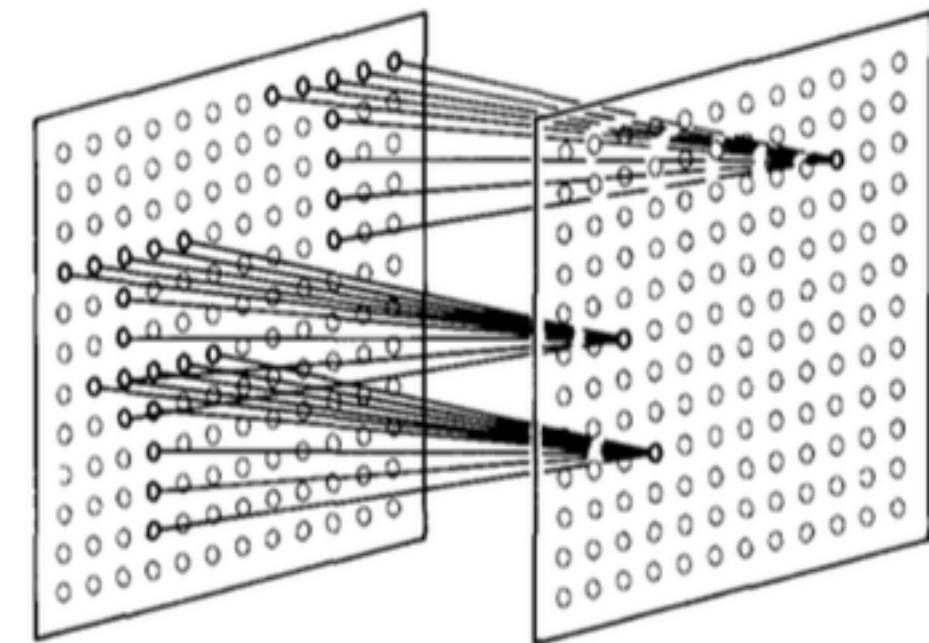
1979



Convolutional Neural Networks & Neocognitron



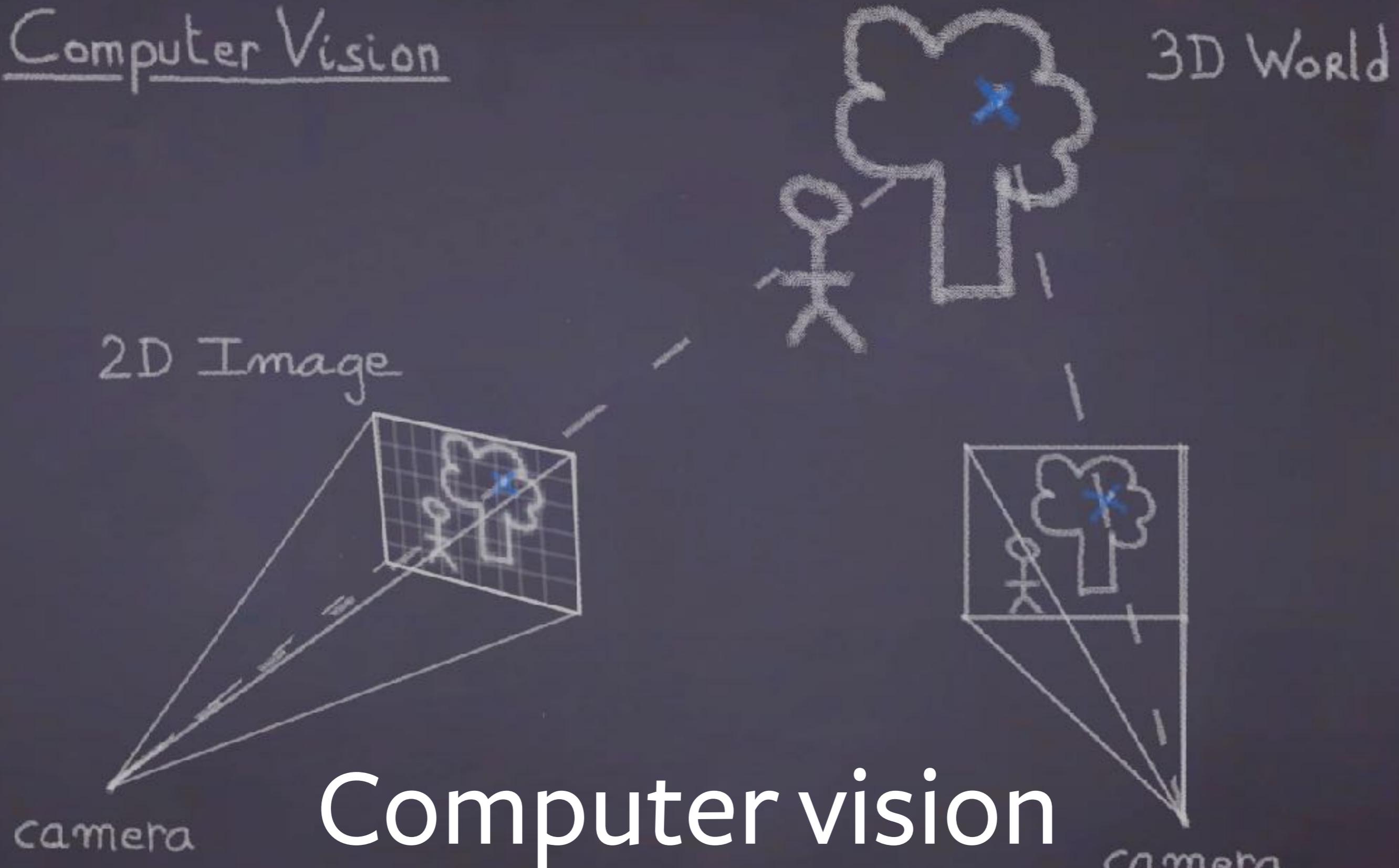
Kunihiko Fukushima



1980's Second AI Winter



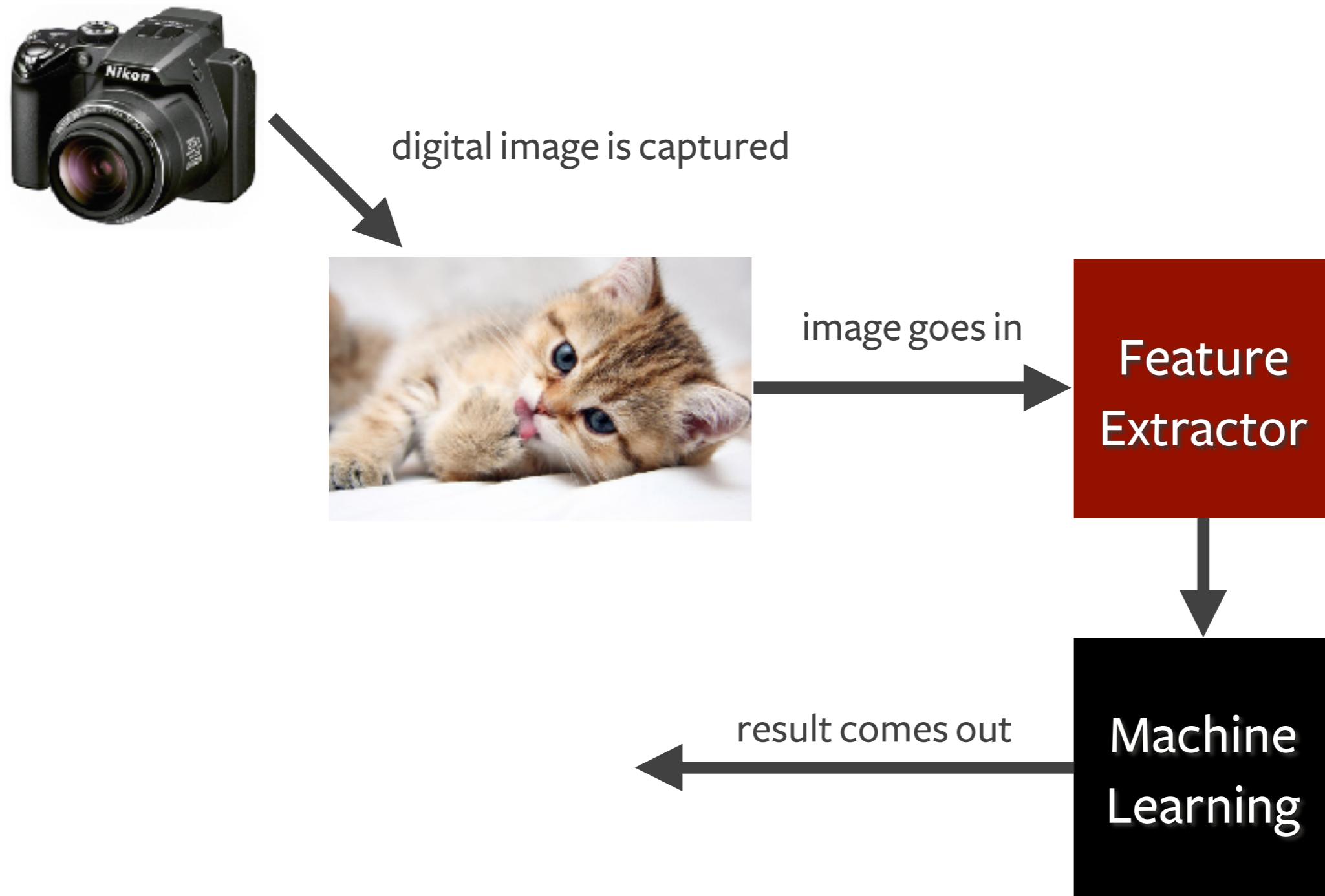
Computer Vision



Computer vision
from the late 80's

3D World

Classical approaches to computer vision take the following form:



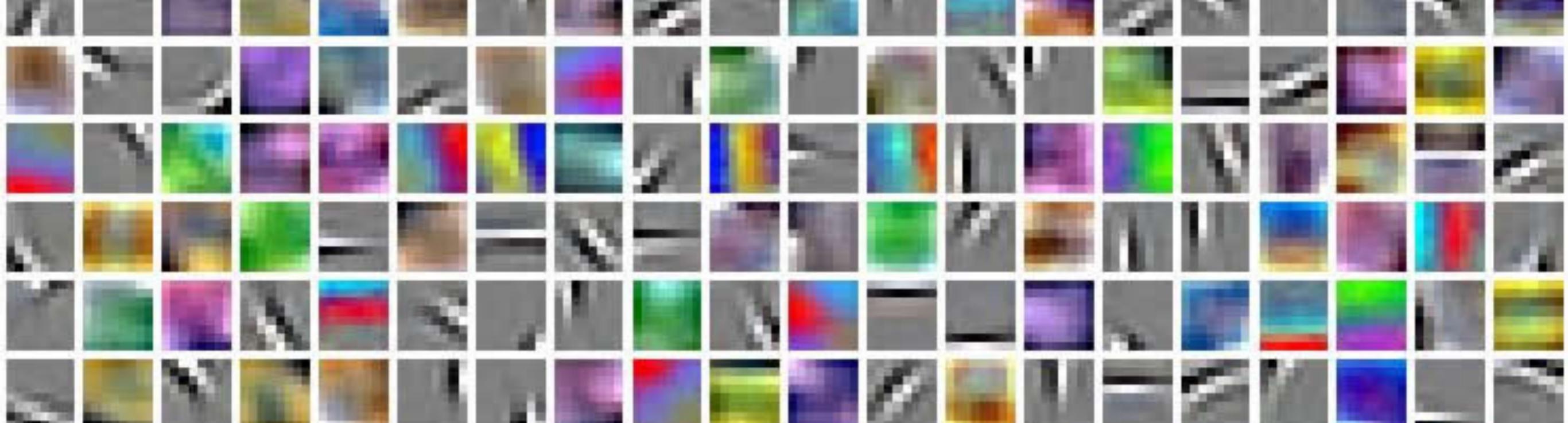
Focus on “Feature Engineering”



Low-level features:
“*Global features*”; edges; corners

High-level features:
“*Model-based features*”; objects;
feature combinations

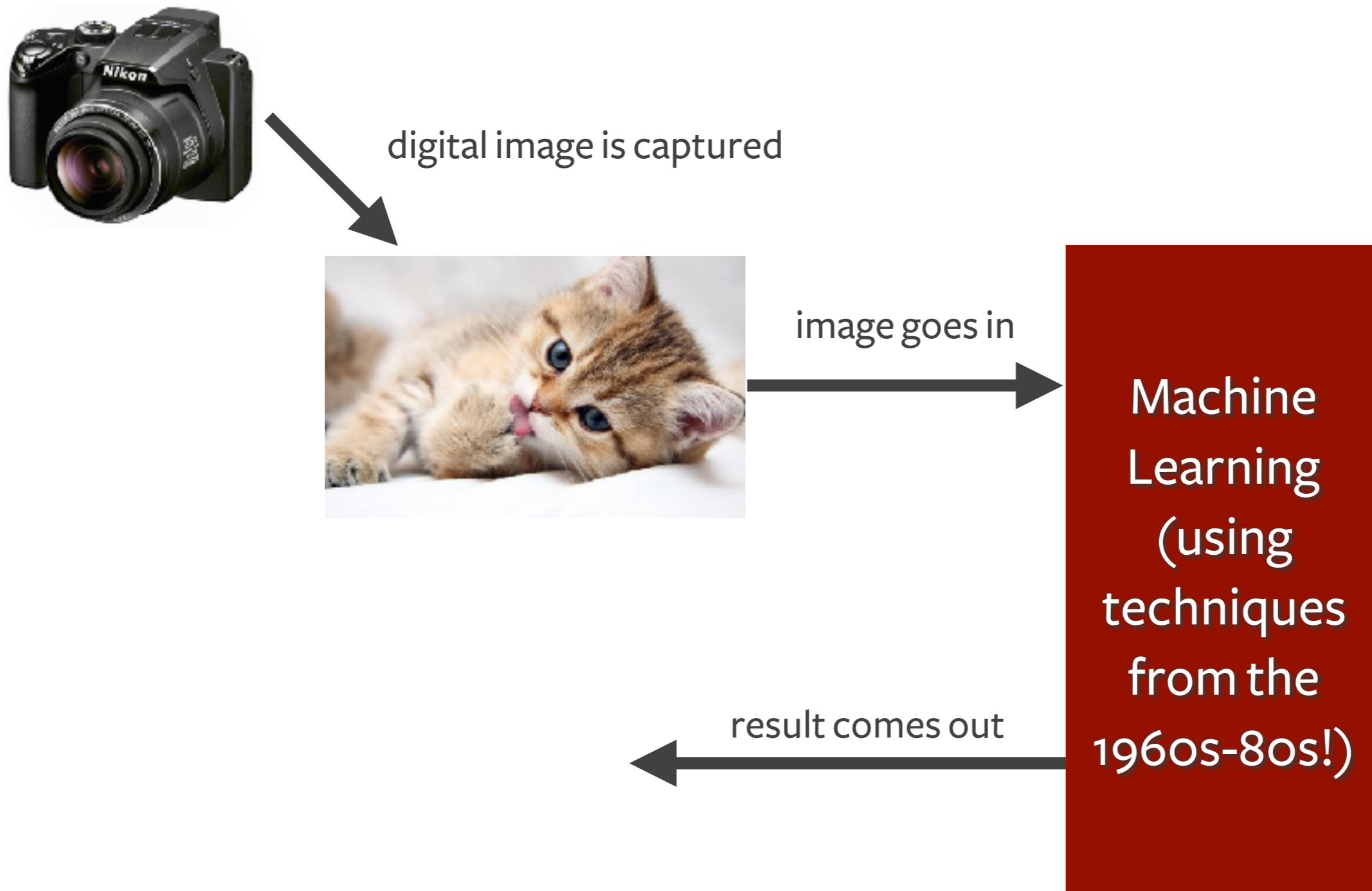




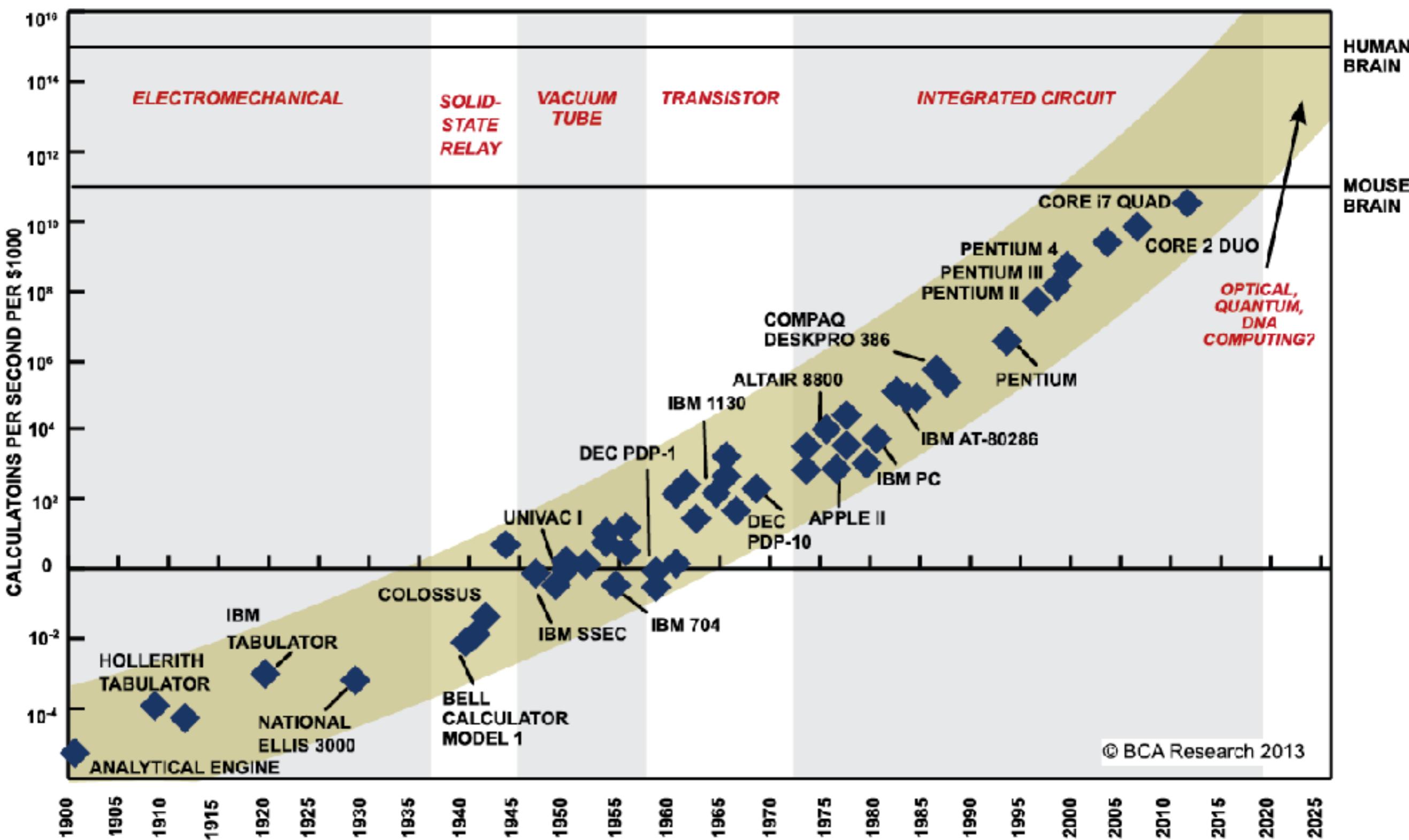
Vision since 2012: *Feature Learning*



Recent approaches to computer vision take the following form:



Moore's Law

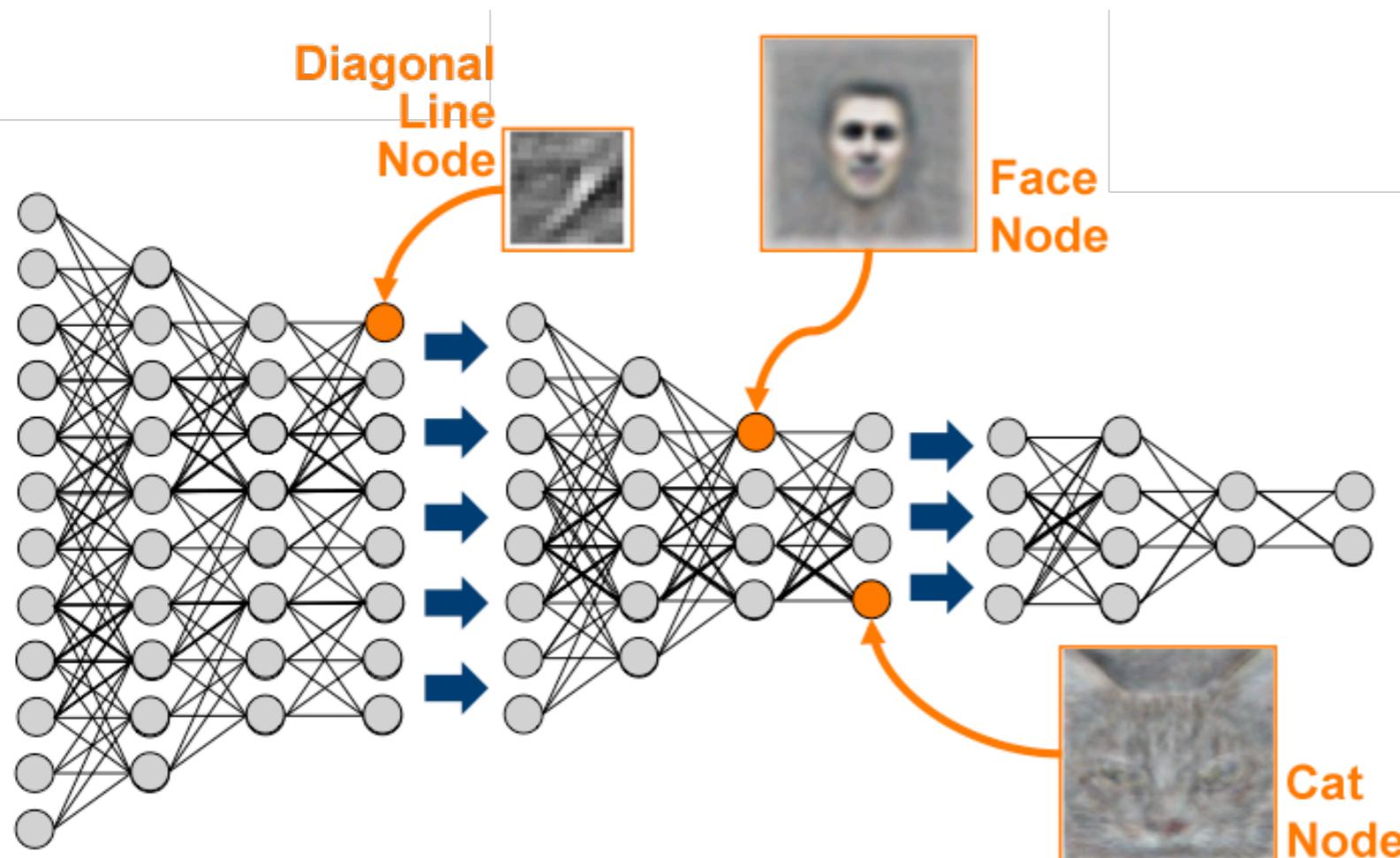


SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, THE VIKING PRESS, 2006. DATAPoints BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

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

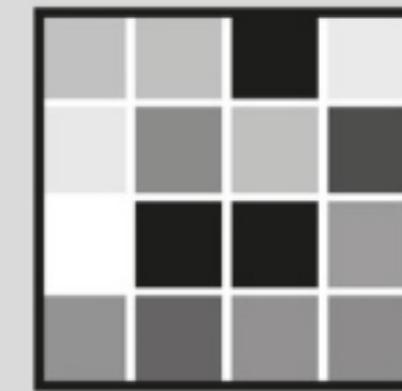


Deep learning: learning layers of features

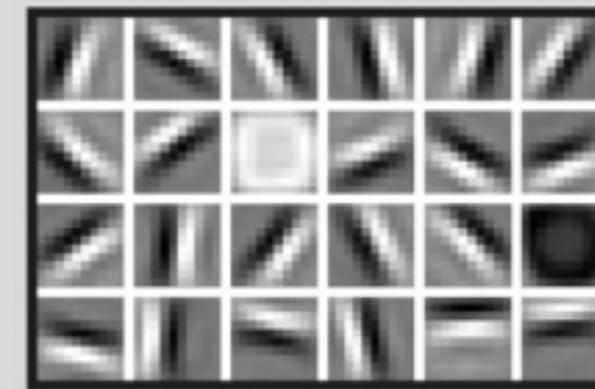


FACIAL RECOGNITION

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.

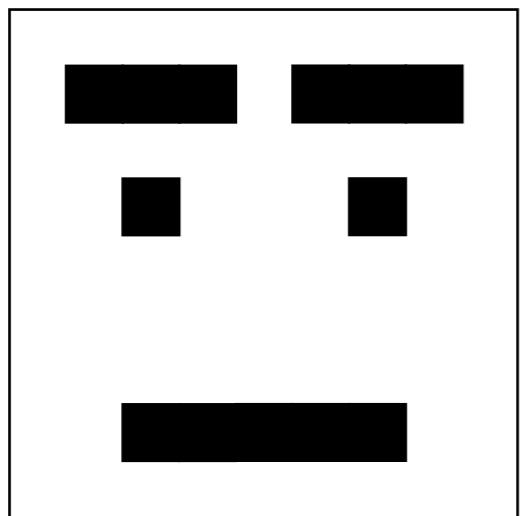


Layer 3: The computer learns to identify more complex shapes and objects.



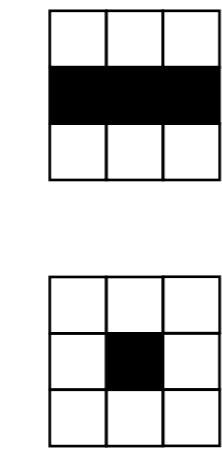
Layer 4: The computer learns which shapes and objects can be used to define a human face.

Input
Image
(grey or
colour)



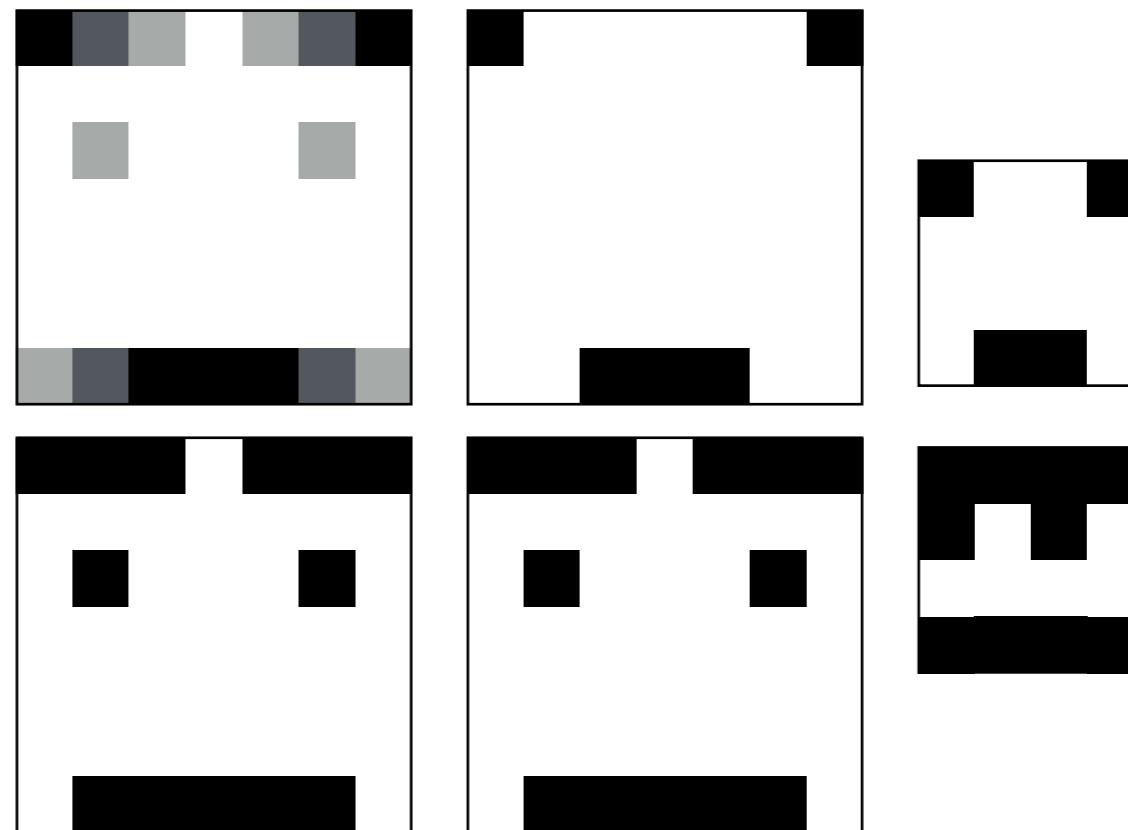
Convolution
Kernels
(multiband for
colour
images)

Response
Image
(multiband)



Rectified
Response
Image
(multiband)

Pooled
Response
Image
(multiband)

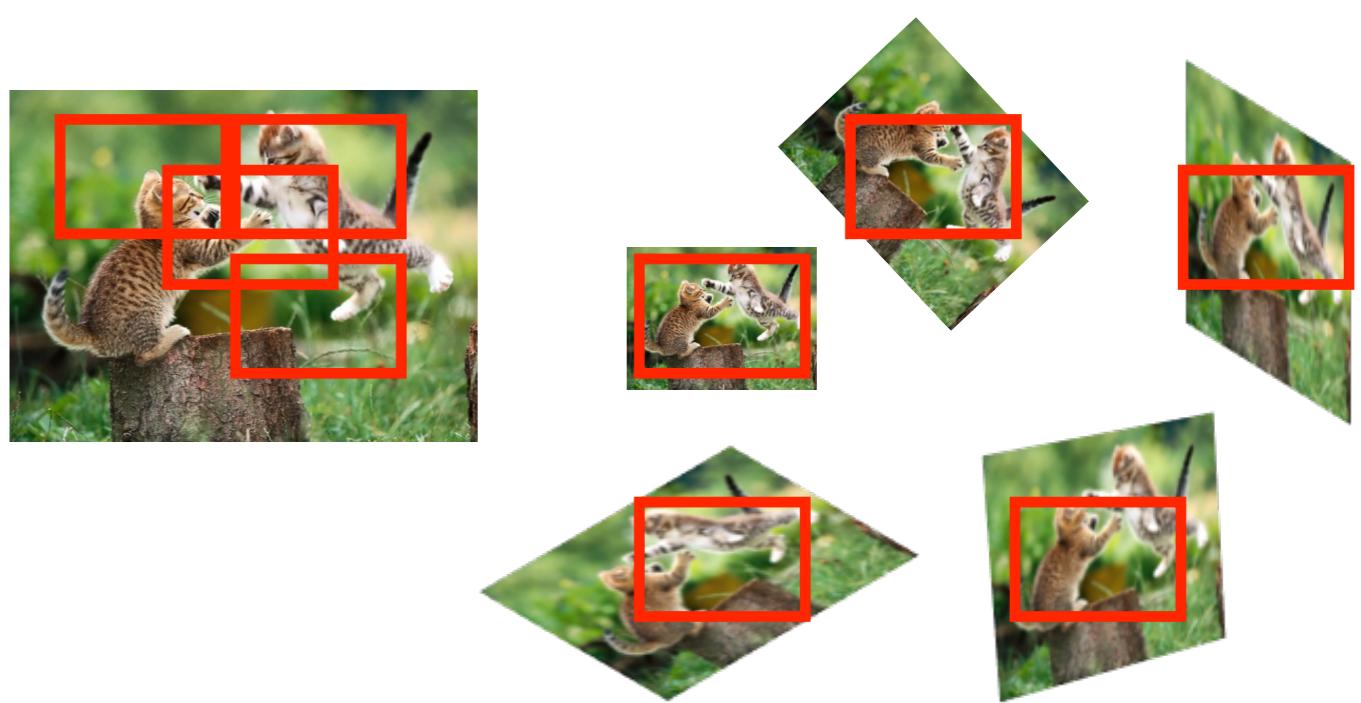




Demo

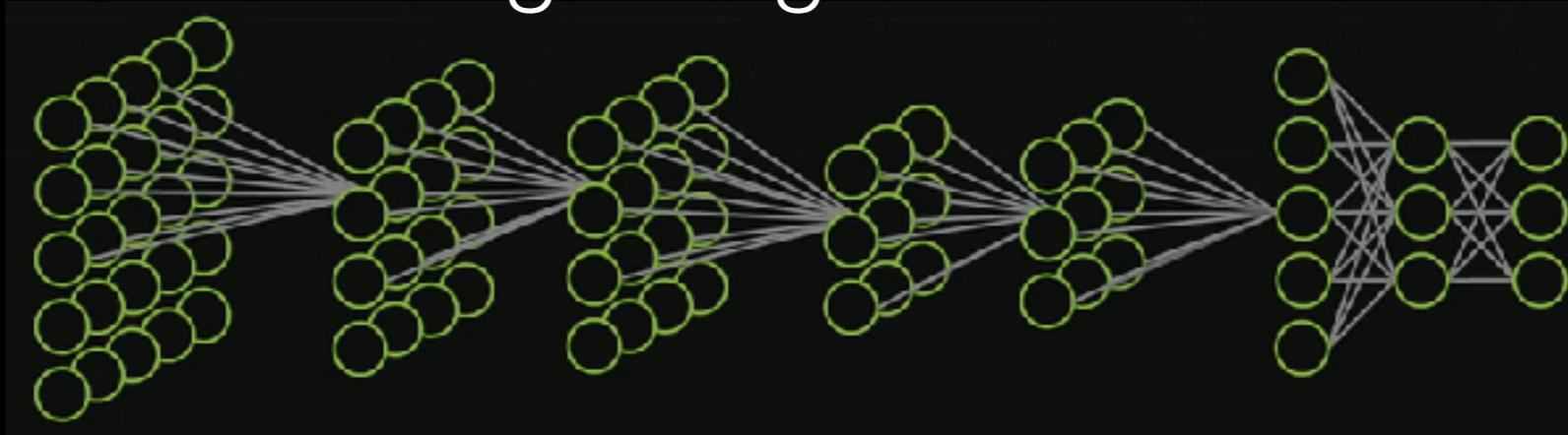
Tricks 'n Tips

- Lots of training data needed...
 - Use **data augmentation** with random transforms to create more from less
- Network overfits...
 - Use **dropout** when learning

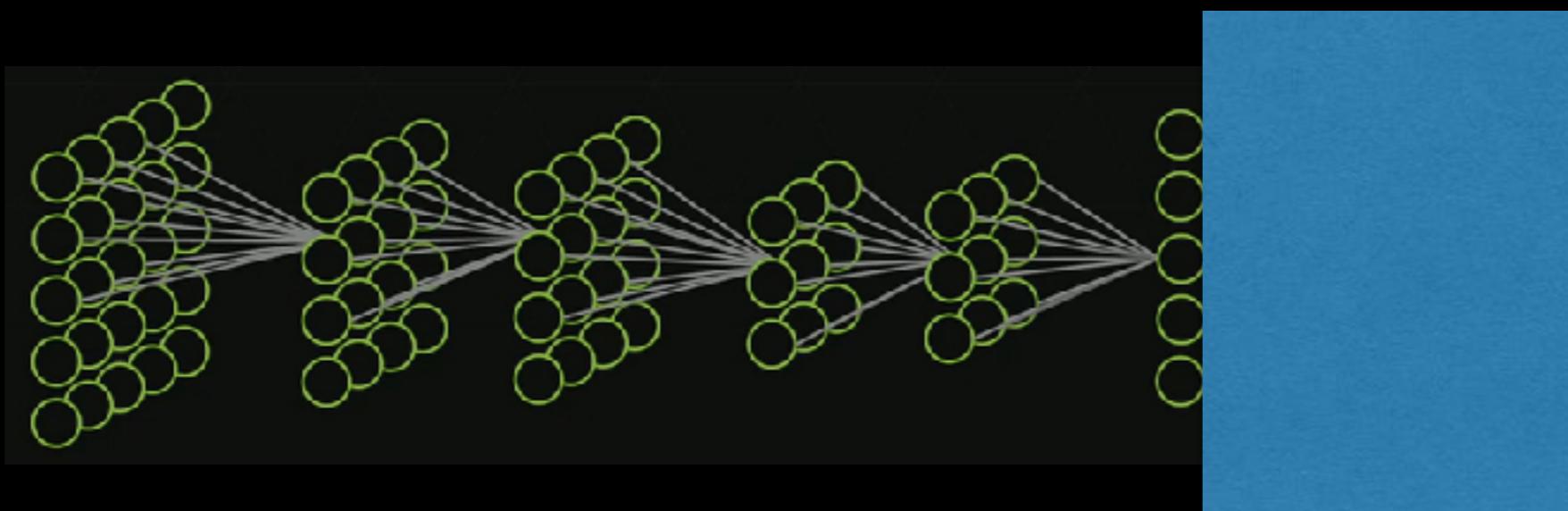


Transfer Learning

ConvNet trained on e.g. ImageNet



Take first bit of network and use as a feature extractor...



Train an MLP/ SVM/... on your
problem using the features
extracted from the net

Do computers dream of electric sheep?

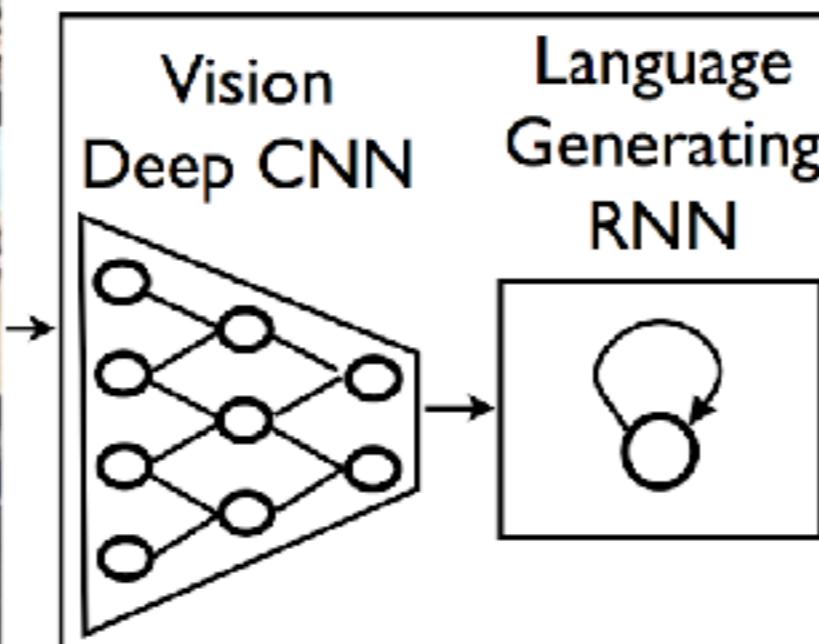
Inceptionism and Algorithmic Pareidolia







State-of-the-art computer vision: Recurrent networks for image captioning



A group of people
shopping at an
outdoor market.

There are many
vegetables at the
fruit stand.



“a man is climbing up a rock face”



“a motorcycle racer is driving a turn on a racetrack”



“a basketball player in a red uniform is trying to score a player in the air”



“a man in a red shirt is riding a bike on a snowy
hill”



“a surfer is jumping off a snowy hill”

Questions?

Mark & I hope you have enjoyed learning about
Computer Vision this semester

We'll meet after Christmas for a revision lecture
(probably in the Tuesday slot - I'll be in touch to confirm)

*If you've enjoyed this module, you might also like
Advanced Computer Vision (Mark), Advanced
Machine Learning (Adam) & Data Mining (me)*