Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neuro 130/230. Harvard College/GSAS 78454

Visual Object Recognition
Computational Models and Neurophysiological Mechanisms
Neurobiology 230. Harvard College/GSAS 78454

Class 1 [09/01/2021]. Introduction to Vision

Note: no class on 09/06/2021

Class 2 [09/13/2021]. Natural image statistics and the retina

Class 3 [09/20/2021]. The Phenomenology of Vision

Class 4 [09/27/2021]. Learning from Lesions

Class 5 [10/04/2021]. Primary Visual Cortex

Note: no class on 10/11/2021

Class 6 [10/18/2021]. Adventures into *terra incognita*

Class 7 [10/25/2021]. From the Highest Echelons of Visual Processing to Cognition

Class 8 [11/01/2021]. First Steps into in silico vision

Class 9 [11/08/2021]. Teaching Computers how to see

Class 10 [11/15/2021]. Computer Vision

Class 11 [11/22/2021]. Connecting Vision to the rest of Cognition

Class 12 [11/29/2021]. Visual Consciousness

FINAL EXAM, PAPER DUE 12/14/2021. No extensions.

An image is worth a million words

Who is there?

What is there?

Where is this?

What are they doing?

What is their relationship?

What will happen next?

Let there be light
The light switch theory

Photosynthesis: ~ 3,500 million years ago

Trilobites, circa 500 million years ago

Parker, A. (2004). *In the blink of an eye: how vision sparked the big bang of evolution*.

Why visual recognition?
Selective advantage of visual processing

- Navigation
- Assessing danger
- Identifying food
- Social interactions
- Detecting far away signals
  (cf. tactile & auditory senses)
- High speeds
  (cf. olfactory signals)
- Detecting patterns such as constellations
- Reading/Symbols

Four fundamental properties of visual recognition

1. Selectivity

2. Invariance

3. Speed

4. Large capacity

Fundamental properties of visual recognition
Selectivity

Selectivity: discriminating among many (similar) objects

ò ó   c ç   ϕ ψ   Б В   ٢ ١   ژ ز   ≈ =

📪 📫   ☺ ☹   ✶ ★   b d   1 I   o O   ■ ◆

Fundamental properties of visual recognition
Invariance

Invariance: recognizing an object in spite of changes in scale, position, illumination, contrast, viewpoint, cue, clutter, background, etc.

Fundamental properties of visual recognition
Speed

~10 frames/sec

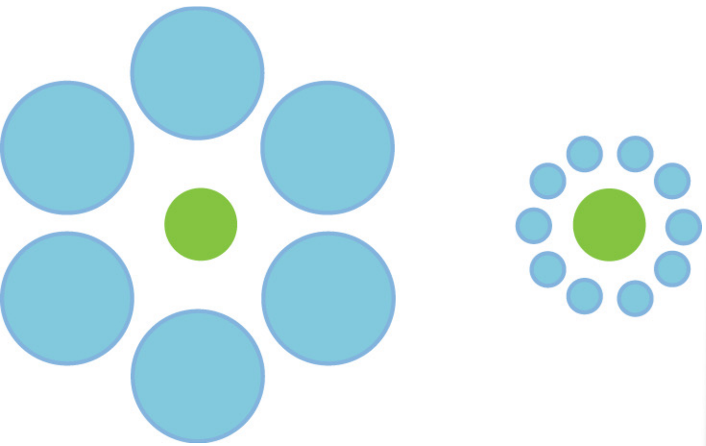
Potter & Levy 1969. Recognition memory for a rapid sequence of pictures;
Thorpe et al 1996. Speed of processing in the human visual system.

Fundamental properties of visual recognition

Capacity

Standing L (1973) Learning 10,000 pictures;
Shepard RN (1987) Toward a universal law of generalization for psychological science;
Biederman I (1987) Recognition-by-components: A theory of human image understanding.

Vision is a construct



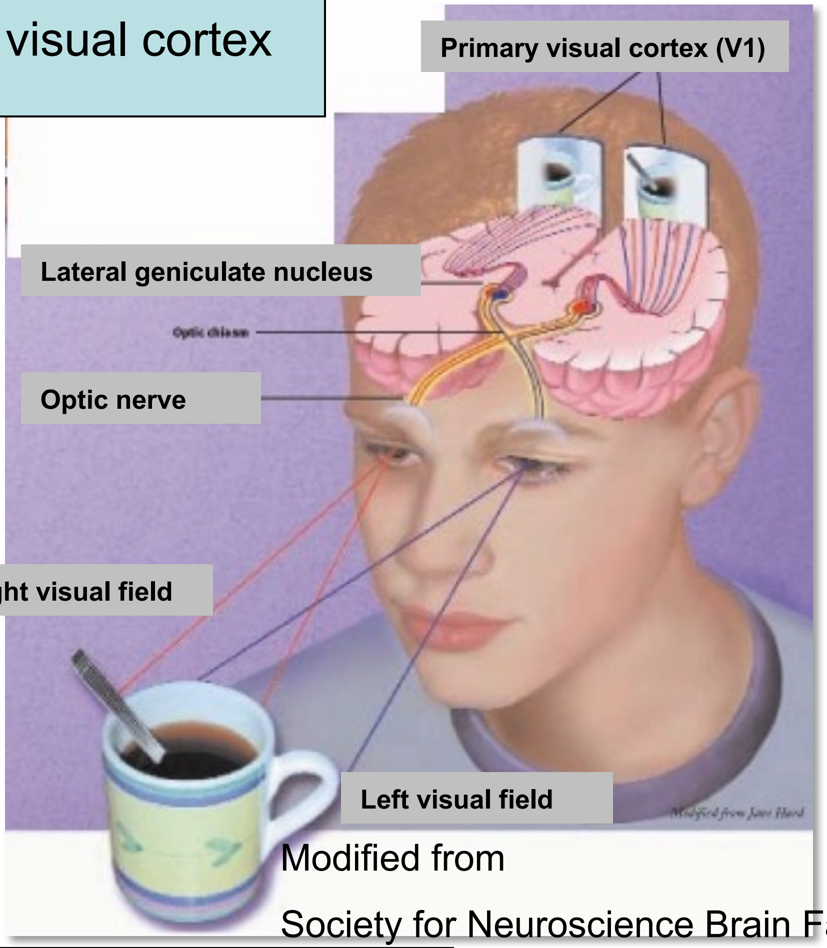
In the ~~eye~~ brain of the beholder

What color is this dress?

Visual recognition is instantiated by the most powerful computational device on Earth

Visual system circuitry

Class 5: Primary visual cortex

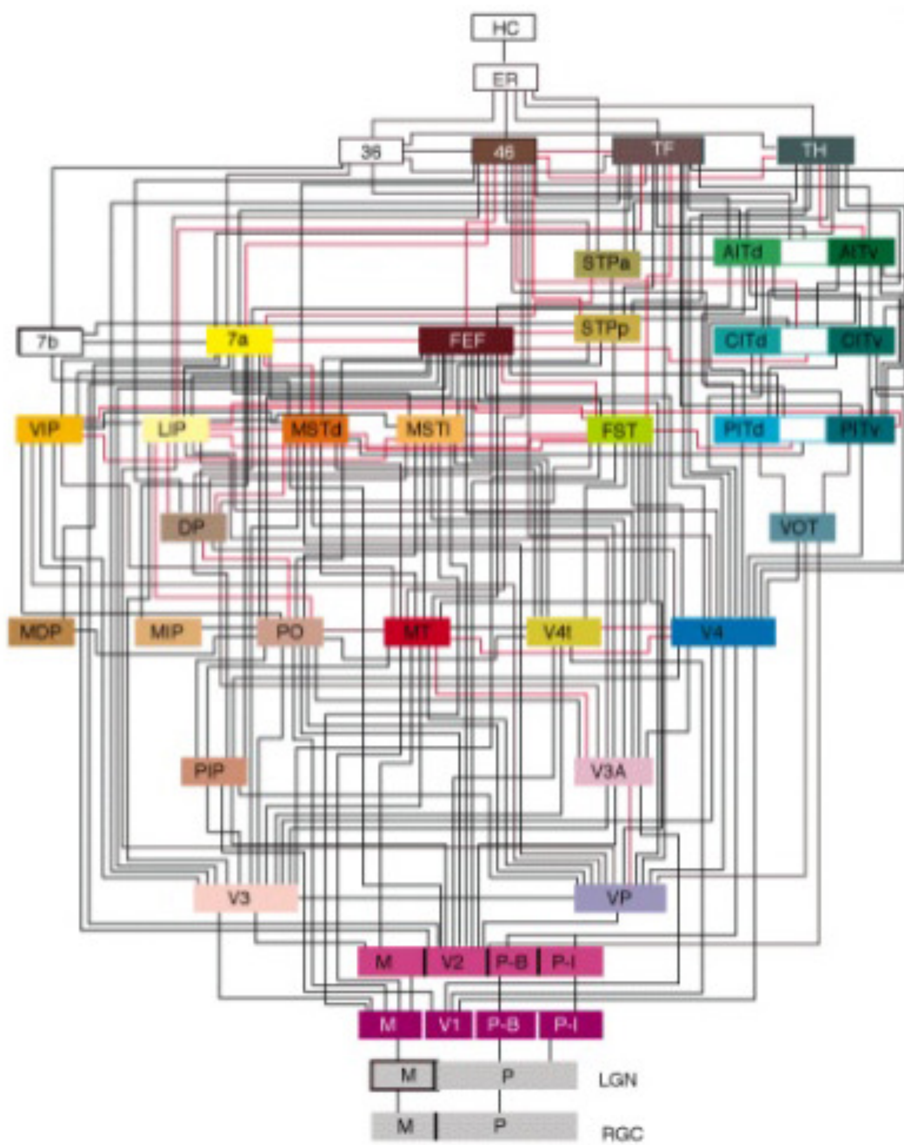


Modified from

Society for Neuroscience Brain Facts

Class 2: Natural image statistics and the retina

Visual system circuitry



Number of neurons in the human brain: ~10^{11}

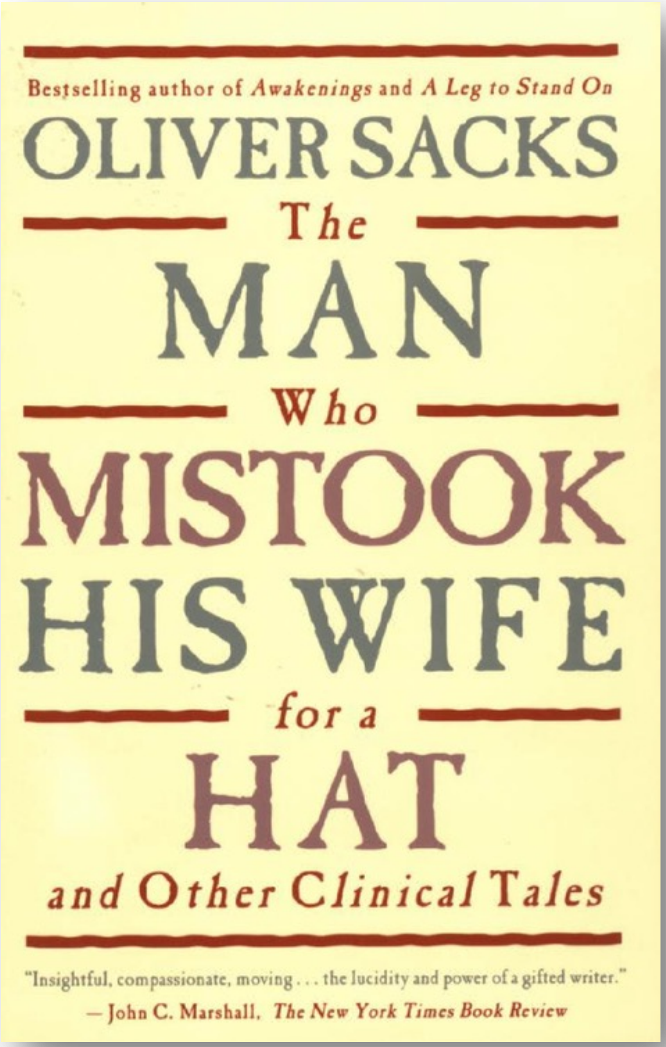
Number of synapses in the human brain: ~10^{15}

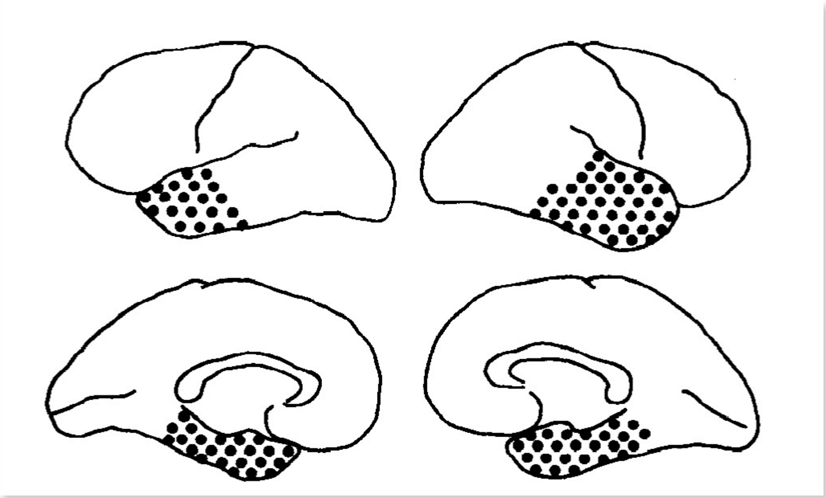
Earth population: ~ $7x10^9$

Connections (?): ~ $7x10^{12}$

Felleman and Van Essen. *Cerebral Cortex* 1991

Figuring out how the brain works from lesion studies





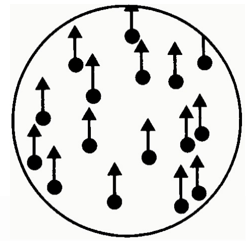
Distribution of lesion sites in cases of face agnosia

Classes 4: Lesions and neurological studies

Damasio et al. *Face agnosia and the neural substrates of memory.* Annual Review of Neuroscience (1990). **13**:89-109

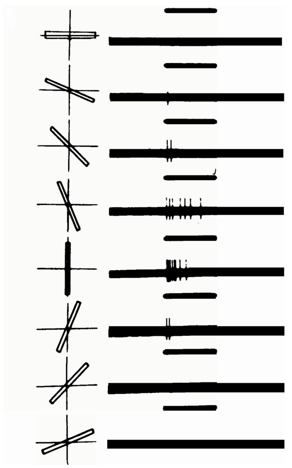
Functional anatomy of the primate visual system

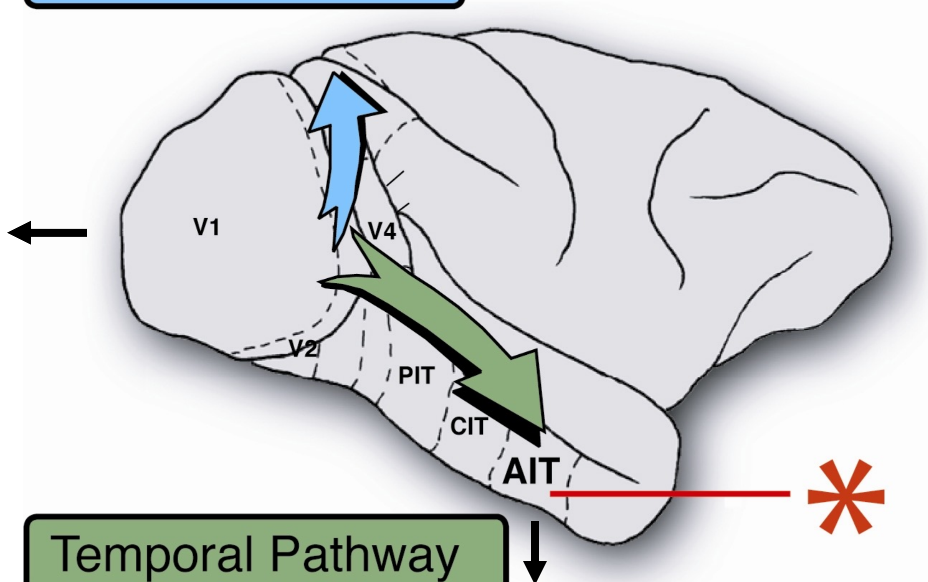
Class 6:
Adventures into *terra incognita*: beyond primary visual cortex

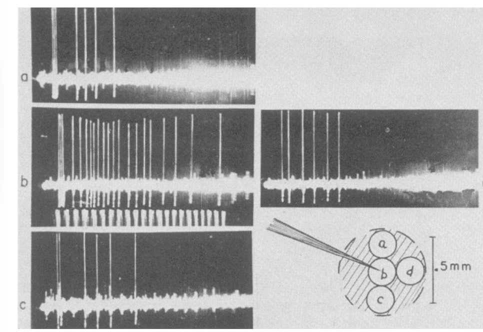


Newsome *et al* (1989)
Nature **341**:52-54

Parietal Pathway





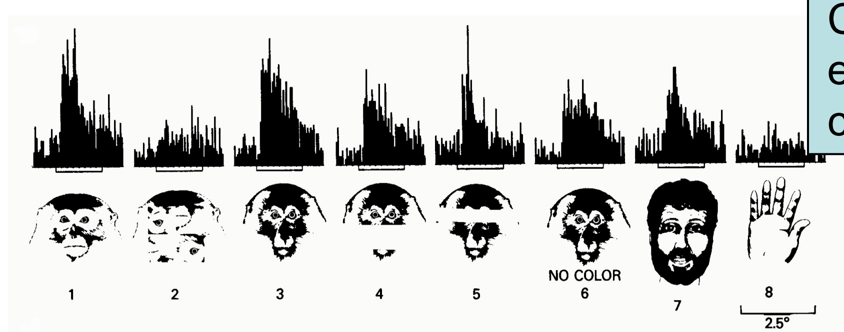


Hubel and Wiesel (1959)
J. Physiol. **148**: 574-591

Kuffler, S. (1953)
J. Neurophys. **16**: 37-68

Temporal Pathway

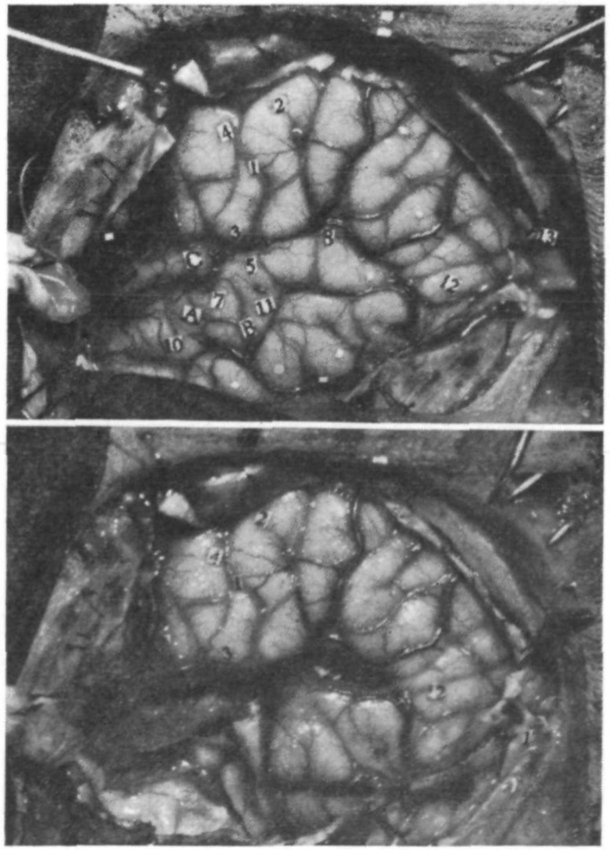
Class 5:
Primary
visual cortex



Class 7: From the highest echelons of vision to cognition

Desimone *et al* (1984)
*J. Neurosci. 4:*2051-2062

Electrical stimulation in the human brain

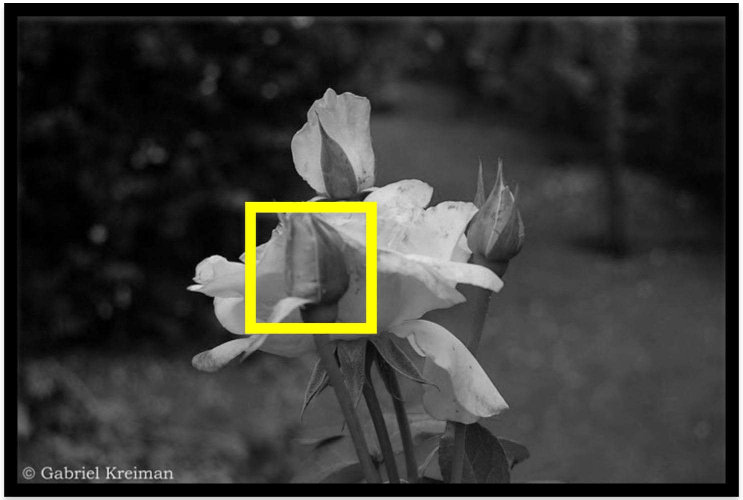


Before the removal was carried out, stimulation at points 5 and 7 produced the following experiential responses.

5. Patient did not reply.
5. Repeated. "Something."
5. Patient did not reply.
5. Repeated. "Something."
5. Repeated again. "People's voices talking." When asked, he said he could not tell what they were saying. They seemed to be far away.
5. Stimulation without warning. He said, "Now I hear them." Then he added, "A little like in a dream."
7. "Like footsteps walking—on the radio."
7. Repeated. "Like company in the room."
7. Repeated. He explained "it was like being in a dance hall, like standing in the doorway—in a gymnasium—like at the Kenwood Highschool." He added, "If I wanted to go there it would be similar to what I heard just now."
7. Repeated. Patient said, "Yes, yes, yes." After withdrawal of the stimulus, he said it was "like a lady was talking to a child. It seemed like it was in a room, but it seemed as though it was by the ocean—at the seashore."
7. Repeated. "I tried to think." When asked whether he saw something or heard something, he said, "I saw and heard. It seemed familiar, as though I had been there."
5. Repeated (20 minutes after last stimulation at 5). "People's voices." When asked, he said, "Relatives, my mother." When asked if it was over, he said, "I do not know." When asked if he also realized he was in the operating room, he said "Yes." He explained it seemed like a dream.
5. Repeated. Patient said, "I am trying." After withdrawal of the electrode he said, "It seemed as if my niece and nephew were visiting at my home. It happened like that many times. They were getting ready to go home, putting their things on—their coats and hats." When asked where, he said, "In the dining room—the front room—they were moving about. There were three of them and my mother was talking to them. She was rushed—in a hurry. I could not see them clearly or hear them clearly."

Penfield & Perot. *The brain's record of auditory and visual experience. A final summary and discussion. Brain* (1963) **86**:595-696

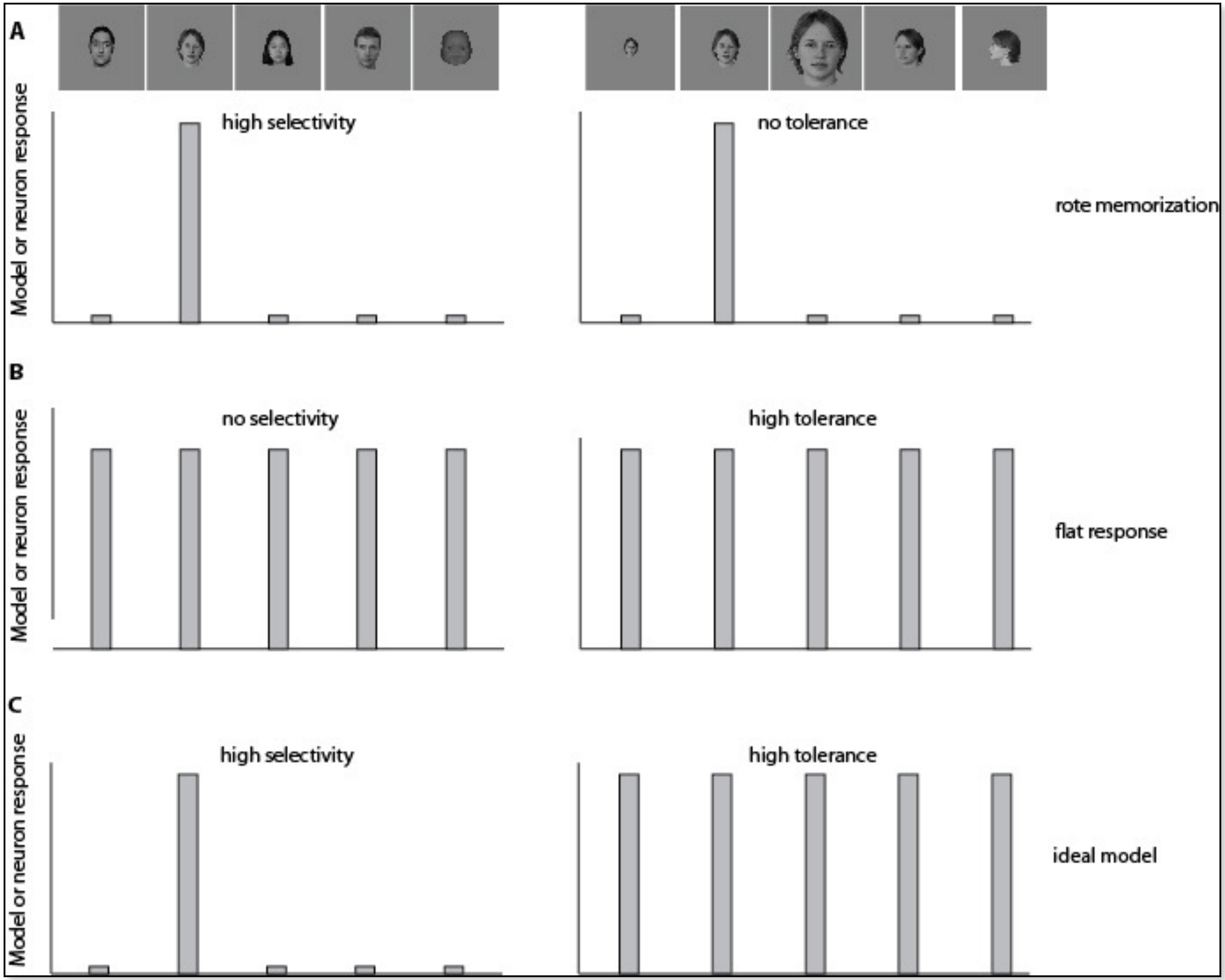
A flower, as seen by a computer



23	16	13	12	13	13	12	12	12	14	16	19	21	22	25	24	20	90	127	101
31	22	13	13	12	12	11	11	13	16	18	18	23	22	21	19	39	83	96	78
34	24	16	14	13	12	21	14	13	17	15	22	15	29	42	82	147	118	63	36
30	20	15	13	14	12	26	34	10	11	79	139	88	91	119	174	172	137	96	78
20	14	12	12	14	14	21	77	35	16	136	148	110	109	127	137	168	157	144	175
13	10	10	12	15	16	14	81	86	52	155	123	91	114	149	120	154	139	138	186
9	9	9	11	14	17	18	54	110	111	143	99	105	104	148	128	103	148	162	172
9	8	9	11	14	18	20	26	97	99	99	91	116	116	141	139	77	88	117	156
9	9	12	12	15	18	15	29	107	99	88	86	121	124	115	123	79	78	98	92
9	10	11	13	15	16	30	97	121	112	98	68	102	125	115	101	100	60	105	109
9	9	11	14	17	13	96	127	145	115	95	60	90	114	118	98	107	72	60	111
9	10	12	13	16	17	117	128	122	114	89	65	94	108	118	116	117	93	59	67
10	10	10	7	9	78	152	127	118	114	77	72	95	109	116	120	128	96	68	50
7	1	10	54	114	166	145	121	125	113	65	70	97	107	110	107	103	93	67	54
33	92	129	151	157	158	146	130	125	104	66	77	100	105	111	108	94	85	62	58
145	144	135	142	151	152	149	137	131	98	69	82	102	111	102	93	89	84	59	54
125	125	140	156	144	150	145	133	128	98	74	87	110	110	106	93	86	80	56	48
147	147	161	143	143	144	138	129	121	94	69	86	107	106	102	91	82	77	50	43
182	181	164	140	143	140	132	128	121	97	71	82	100	109	97	91	93	80	44	40
188	174	143	147	146	144	137	127	119	97	78	83	100	105	104	92	86	81	46	38

Classes 8-11: Can computers see the way we do? Computer vision

Why is vision difficult?



Towards a theory of object recognition

Computational models can

- Integrate existing data

- Explain apparently disparate observations

- Quantify and formalize knowledge

- Suggest experimentally-testable predictions

- Provide a useful engineering tool

The summer vision project

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

PROJECT MAC

Artificial Intelligence Group July 7, 1966

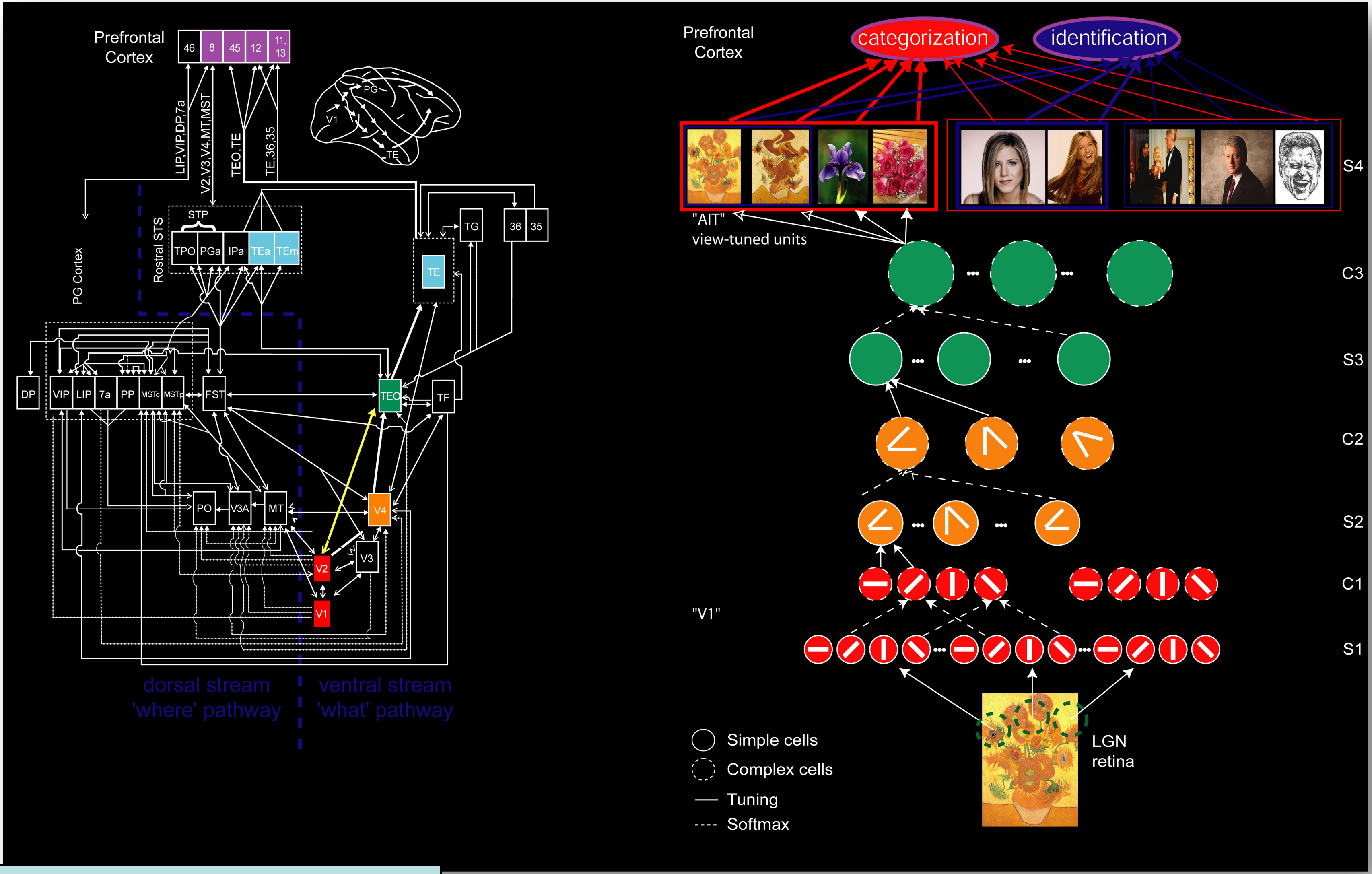
Vision Memo. No. 100.

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 participate in the construction of a system complex enough to be a real landmark in the development of "pattern recognition".

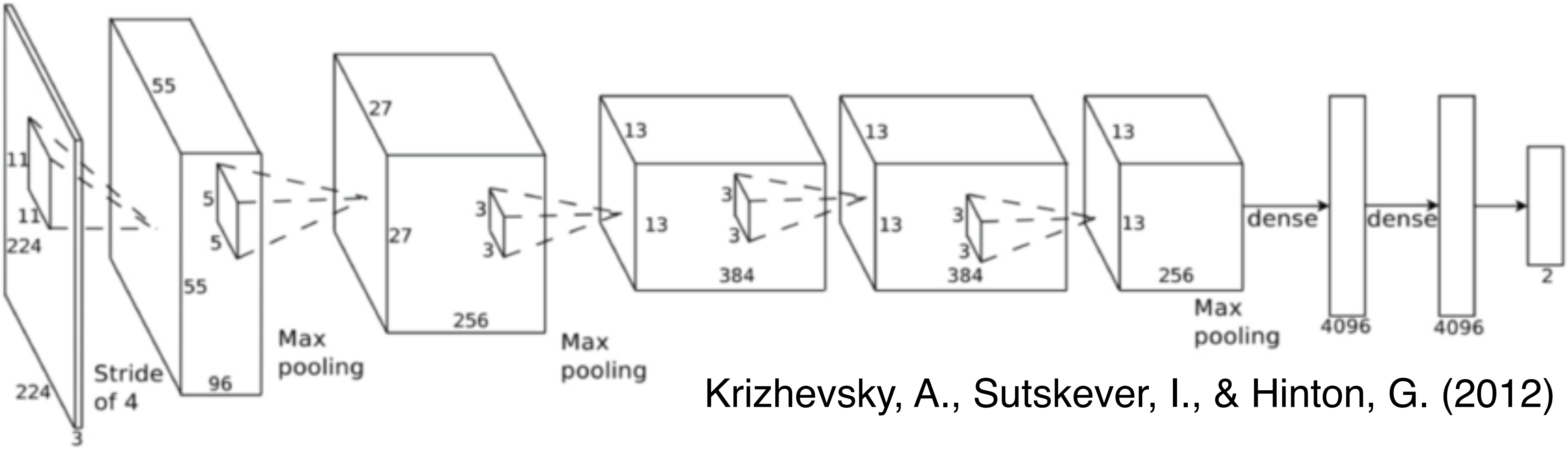
A feed-forward hierarchical model of ventral cortex



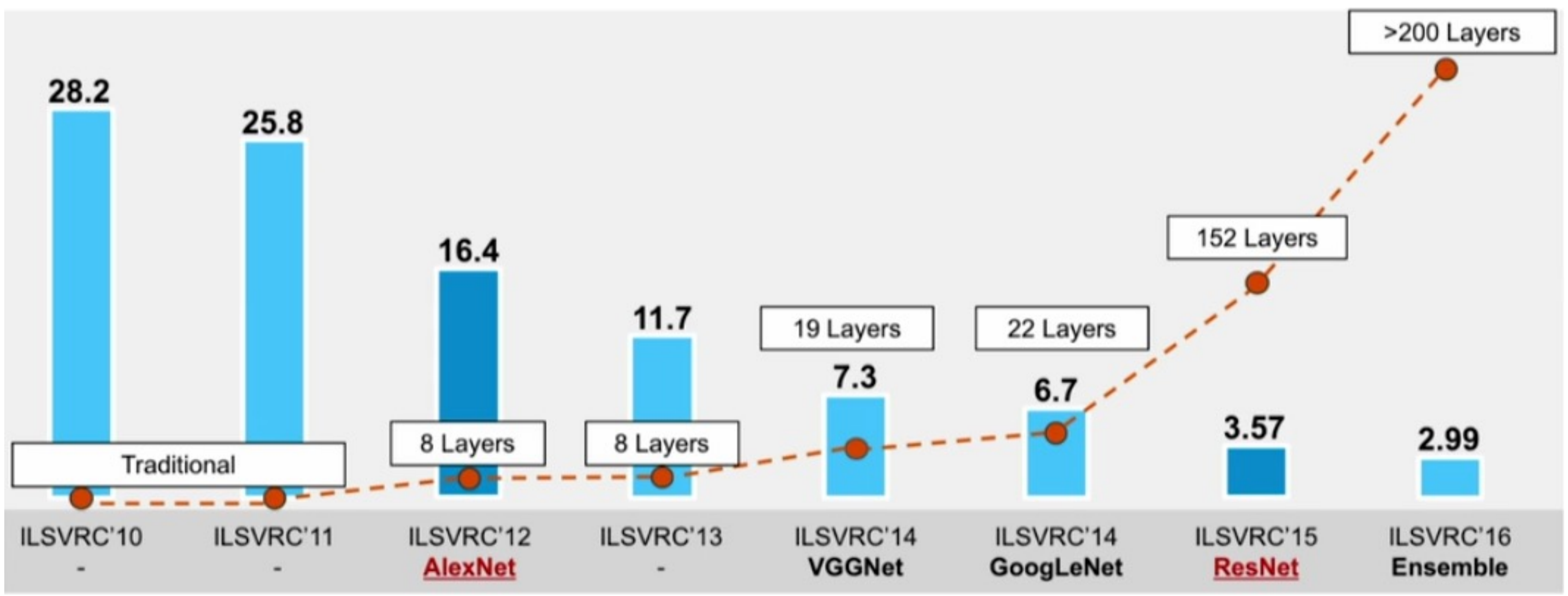
Classes 9-10: Computer vision

Fukushima. *Biological Cybernetics* 1980;
Serre, Kreiman, Cadieu, Knoblich, Poggio, Progress in Brain Research 2007

Rapid progress in image classification tasks



Krizhevsky, A., Sutskever, I., & Hinton, G. (2012)

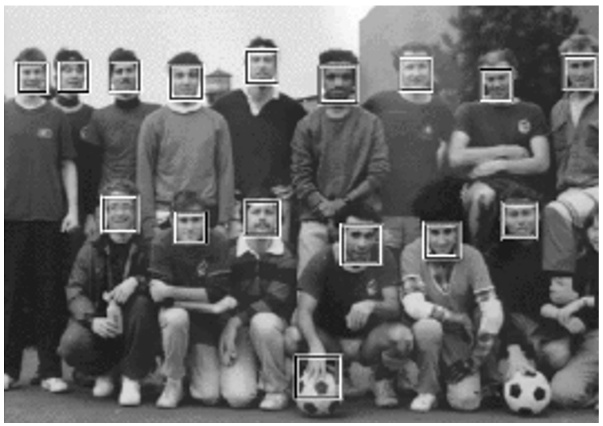


Detection, segmentation, recognition

Face detection

Segmentation

Recognition



Classes 10-11: Can computers see the way we do?
Computer vision

Fig. 7: The 101 object categories and the background clutter category. Each category contains between 45 to 400 images. Two randomly chosen samples are shown for each category. The categories were selected prior to the experimentation, collected by operators not associated with the experiment, and no category was excluded from the experiments. The last row shows examples from the background dataset. This dataset is

Why visual recognition?
Applications

- Face recognition
- Pedestrian recognition
- Self-driving cars
- Robot navigation
- Clinical applications
- Security
- Intelligent image understanding

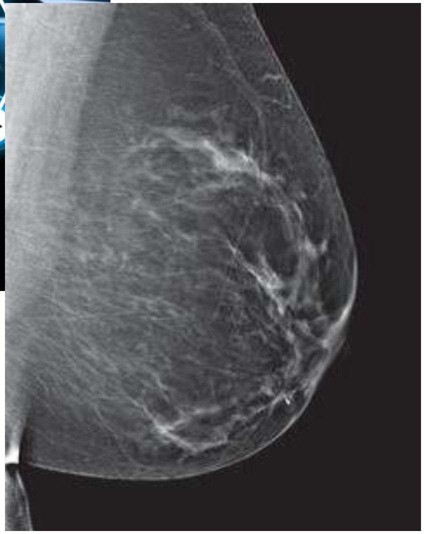


Image captioning

Caption Bot

I think it's a group of people standing next to a man in a suit and tie.

How did I do?

A Turing test for vision

Can machines (be taught to) see the world the way we do?

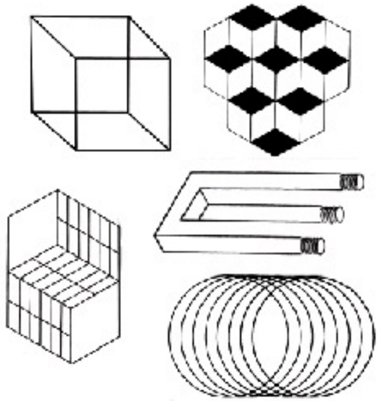
Alan Turing, 1950. Computing Machinery and Intelligence. “Can machines think?”

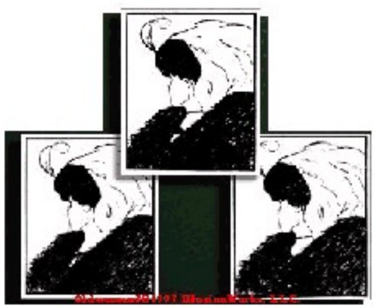
Ultraintelligence

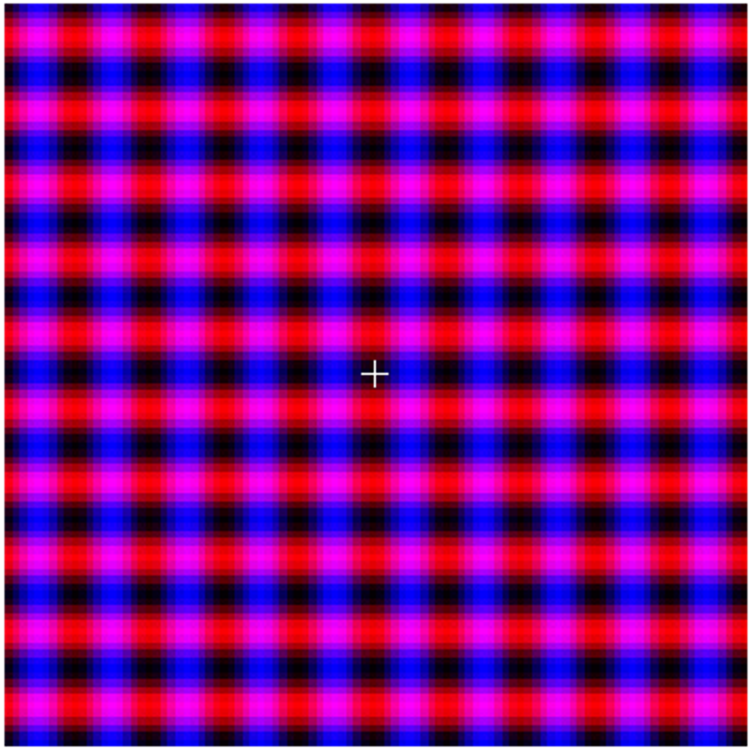
Let an ultraintelligent machine be defined as a machine that can far surpass all the intellectual activities of any man however clever. Since the design of machines is one of these intellectual activities, an ultraintelligent machine could design even better machines; there would then unquestionably be an "intelligence explosion," and the intelligence of man would be left far behind. **Thus the first ultraintelligent machine is the last invention that man need ever make . . .**

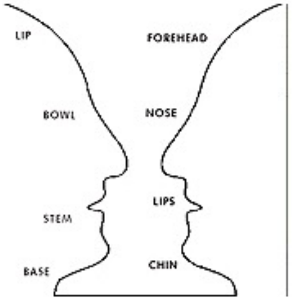
I.J.Good "Speculations Concerning the First Ultra-intelligent Machine", 1965

Bistable percepts and subjective perception









Blake R, Logothetis N (2002) Visual competition. Nature Reviews Neuroscience 3: 13-21.

Crick F, Koch C (1990) Towards a neurobiological theory of consciousness.

Class 12: Visual consciousness

Further reading

- **Kreiman, G. 2021. Biological and Computer Vision. Chapter 1. Cambridge University Press.**

- **Marr, D. (1982). Vision. San Francisco, Freeman publishers.**

- **Alan Turing, 1950. Computing Machinery and Intelligence. "Can machines think?" Mind 49: 433-460.**

- Felleman, D.J. and D.C. Van Essen. Cerebral Cortex, 1991. **1**: p. 1-47.
- Hubel, D. and T. Wiesel,. Journal of Physiology (London), 1959. **148**: p. 574-591.
- Newsome, W., K. Britten, and J. Movshon. Nature, 1989. **341**: p. 52-54.
- Kuffler, S.. Journal of Neurophysiology, 1953. 16: p. 37-68.
- Desimone, R., et al. Journal of Neuroscience, 1984. **4**(8): p. 2051-2062.
- Damasio, A., D. Tranel, and H. Damasio. Annual Review of Neuroscience, 1990. **13**: p. 89-109.
- Penfield, W. and P. Perot. Brain, 1963. **86**(4): p. 595-696.
- Fukushima, K. Biological Cybernetics, 1980. 36(4): p. 193-202.
- Serre, T., et al. Progress In Brain Research, 2007. 165C: p. 33-56.
- Potter M, Levy E (1969) Journal of Experimental Psychology 81: 10-15.
- Thorpe S, Fize D, Marlot C (1996) Nature 381: 520-522.
- Standing, L. (1973). Quarterly Journal of Experimental Psychology **25**(2): 207-222.
- Shepard, R. N. (1987). Science **237**: 1317-1323.
- Biederman, I. (1987). Psychological Review **24**(2): 115-147.
- Blake R, Logothetis N (2002) Visual competition. Nature Reviews Neuroscience 3: 13-21.
- Sacks, O. (1998). The man who mistook his wife for a hat. New York, Touchstone Books.
- Crick, F., & Koch, C. (1990). Seminars in the Neurosciences., *2*, 263-275.
- Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). Paper presented at NIPS, Montreal.
- Good, I. (1966). Speculations Concerning the First Ultraintelligent Machine. Advances in Computers, *6*, 31-88.
- Parker, A. (2004). *In the blink of an eye: how vision sparked the big bang of evolution*. New York: Basic Books.