Neuro4ML

Dan Goodman & Marcus Ghosh

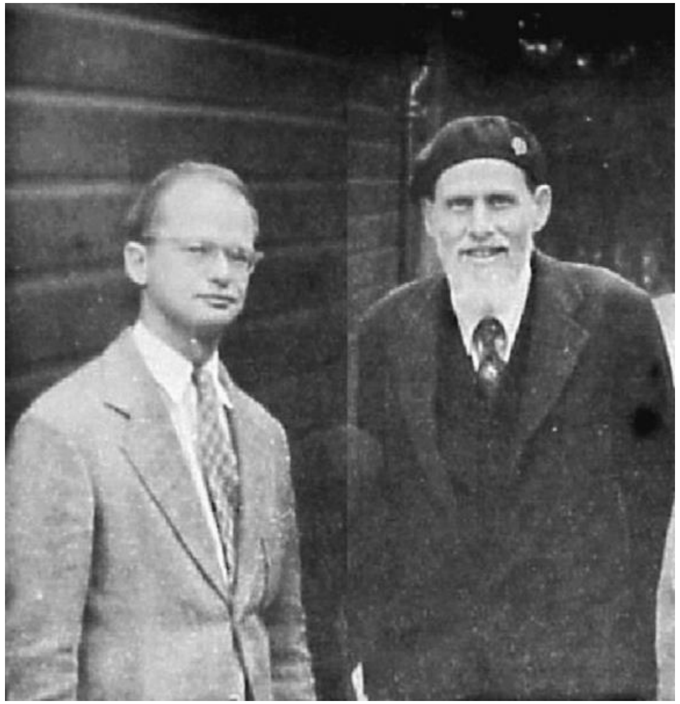
Neuroscience ❤ ML
History

Imperial College London

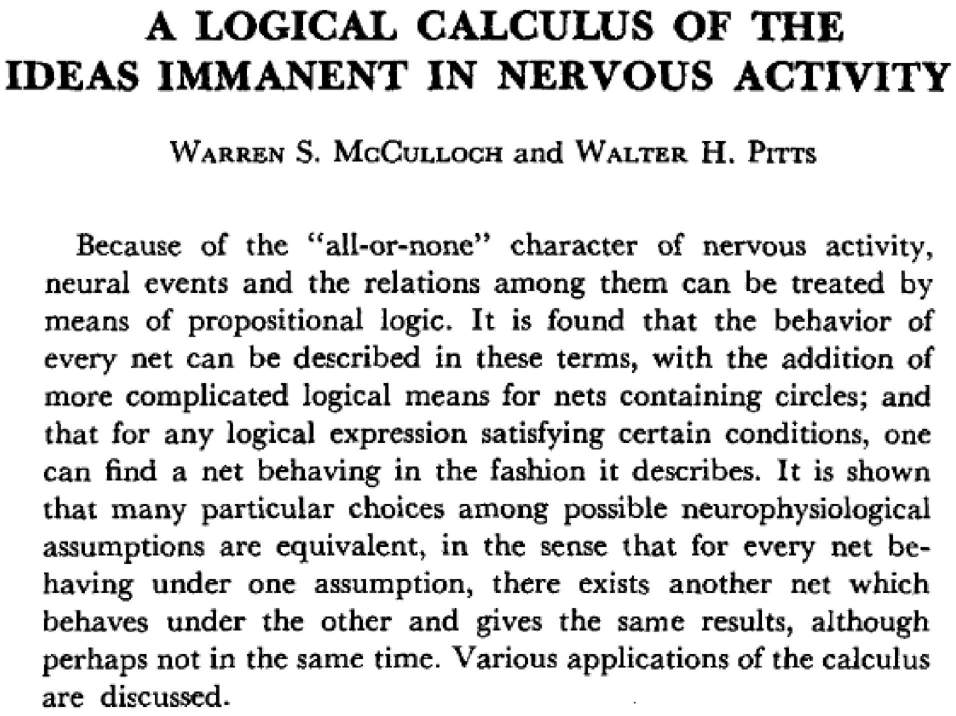
https://neural-reckoning.org

The neural network

McCulloch and Pitts (1943)



W. Pitts and W. McCulloch, 1949
https://doi.org/10.1016/j.biosystems.2006.08.010

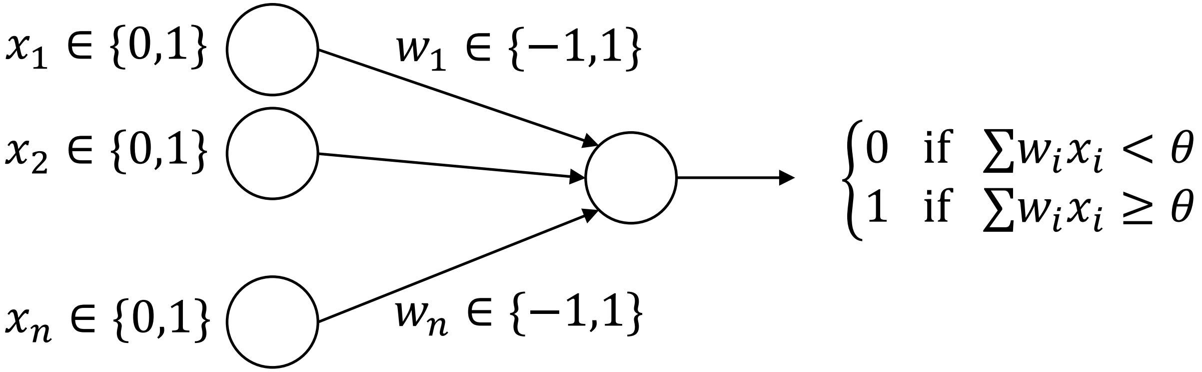


A LOGICAL CALCULUS OF THE IDEAS IMMANENT IN NERVOUS ACTIVITY

WARREN S. MCCULLOCH and WALTER H. PITTS

Because of the "all-or-none" character of nervous activity, neural events and the relations among them can be treated by means of propositional logic. It is found that the behavior of every net can be described in these terms, with the addition of more complicated logical means for nets containing circles; and that for any logical expression satisfying certain conditions, one can find a net behaving in the fashion it describes. It is shown that many particular choices among possible neurophysiological assumptions are equivalent, in the sense that for every net behaving under one assumption, there exists another net which behaves under the other and gives the same results, although perhaps not in the same time. Various applications of the calculus are discussed.

https://doi.org/10.1007/BF02478259



The digital computer

McCulloch and Pitts (1943)

Von Neumann (1945)

CPU RAM

2.6 The three specific parts CA, CC together C and M correspond to the associative neurons in the human nervous system. It remains to discuss the equivalents of the sensory or afferent and the motor or efferent neurons. These are the input and the output organs of the device, and we shall now consider them briefly.

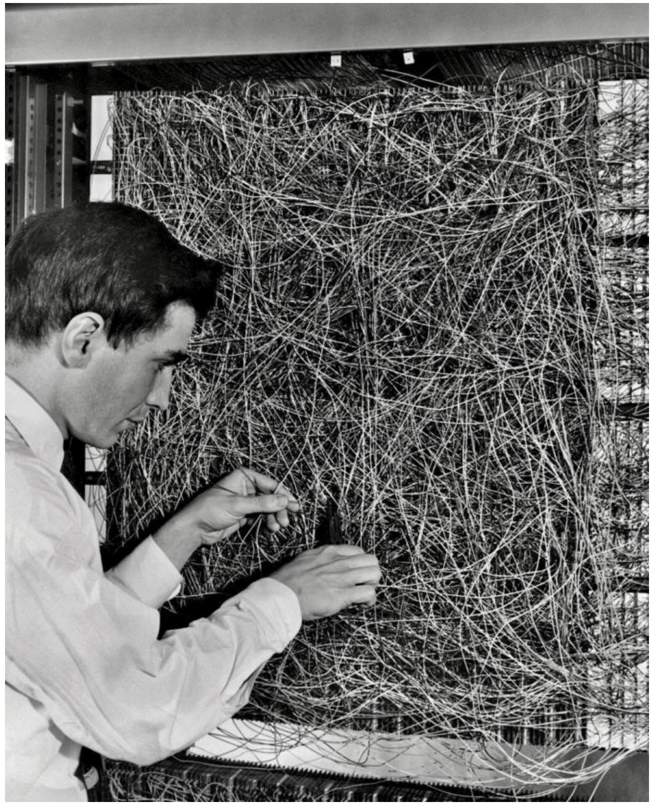
First draft of a report on the EDVAC
https://library.si.edu/digital-library/book/firstdraftofrepo00vonn

The neural network

McCulloch and Pitts (1943)

Von Neumann (1945)

Rosenblatt (1958)



Psychological Review
Vol. 65, No. 6, 1958

THE PERCEPTRON: A PROBABILISTIC MODEL FOR INFORMATION STORAGE AND ORGANIZATION IN THE BRAIN[1]

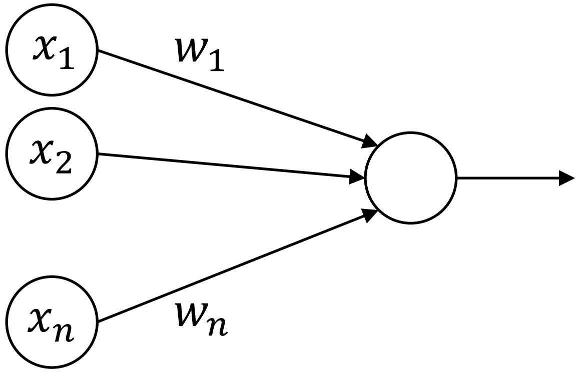
F. ROSENBLATT

Cornell Aeronautical Laboratory

If we are eventually to understand the capability of higher organisms for perceptual recognition, generalization, recall, and thinking, we must first have answers to three fundamental questions:

1. How is information about the physical world sensed, or detected, by the biological system?
2. In what form is information stored, or remembered?
3. How does information contained in storage, or in memory, influence recognition and behavior?

and the stored pattern. According to this hypothesis, if one understood the code or "wiring diagram" of the nervous system, one should, in principle, be able to discover exactly what an organism remembers by reconstructing the original sensory patterns from the "memory traces" which they have left, much as we might develop a photographic negative, or translate the pattern of electrical charges in the "memory" of a digital computer. This hypothesis is appealing in its simplicity and ready intelligibility, and a large family of theoretical brain



$$\begin{cases} 0 & \text{if } \sum w_i x_i < \theta \\ 1 & \text{if } \sum w_i x_i \geq \theta \end{cases}$$

See [Goodman et al. 2013 "Decoding neural responses to temporal cues for sound localization"]

The neural network

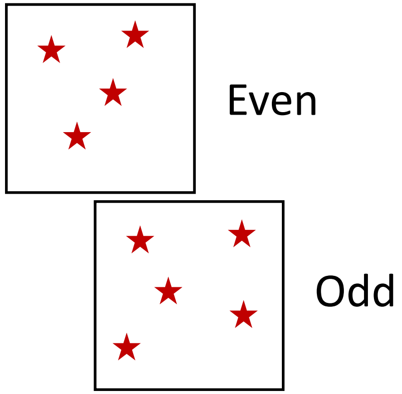
McCulloch and Pitts (1943)

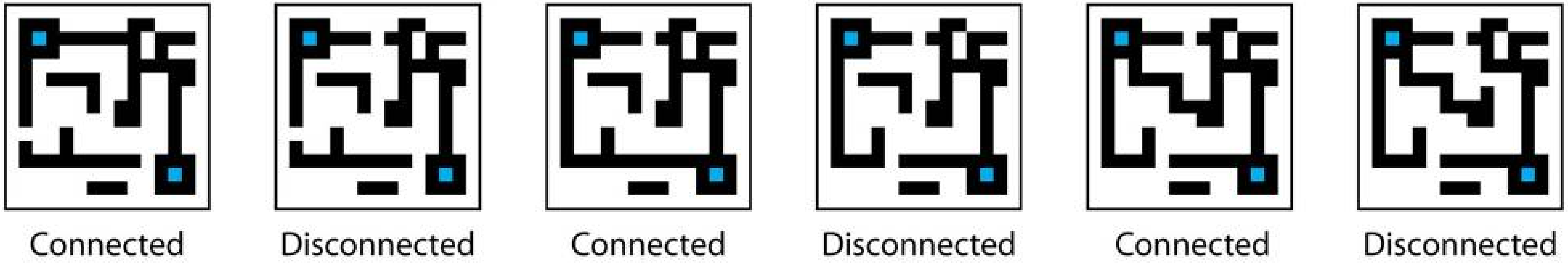
Von Neumann (1945)

Rosenblatt (1958)

Minsky (1960s)

A	B	A XOR B
0	0	0
0	1	1
1	0	1
1	1	0





PLoS Comput Biol. 2022 Jun; 18(6): e1010227. PMCID: PMC9258846
Published online 2022 Jun 21. doi: 10.1371/journal.pcbi.1010227 PMID: 35727818

Towards a more general understanding of the algorithmic utility of recurrent connections

Brett W. Larsen, Conceptualization, Formal analysis, Investigation, Methodology, Software, Visualization, Writing – original draft [1,2,3] and Shaul Druckmann, Conceptualization, Formal analysis, Investigation, Methodology, Software, Supervision, Writing – original draft [2,3,*]

Backpropagation

McCulloch and Pitts (1943)

Von Neumann (1945)

Rosenblatt (1958)

Minsky (1960s)

Rumelhart, Hinton, Williams (1986)

“ The learning procedure, in its current form, is not a plausible model of learning in brains. However, applying the procedure to various tasks shows that interesting internal representations can be constructed by gradient descent in weight-space, and this suggests that it is worth looking for more biologically plausible ways of doing gradient descent in neural networks. ”

Lillicrap et al. 2016, 2020

Backprop / gradient descent in brief

Want to minimise scalar loss $\mathcal{L}$

Parameters vector θ

Compute the gradient

Gradient $(\nabla\mathcal{L})_i = \frac{\partial\mathcal{L}}{\partial\theta_i}$

Vector chain rule:

$$x \in \mathbb{R}^n \xrightarrow{f} y \in \mathbb{R}^m \xrightarrow{g} z \in \mathbb{R}^p$$

Jacobian matrix $J_{ij}^f = \frac{\partial f_i}{\partial x_j}$

Chain rule is $J^{g\circ f} = J^g J^f$

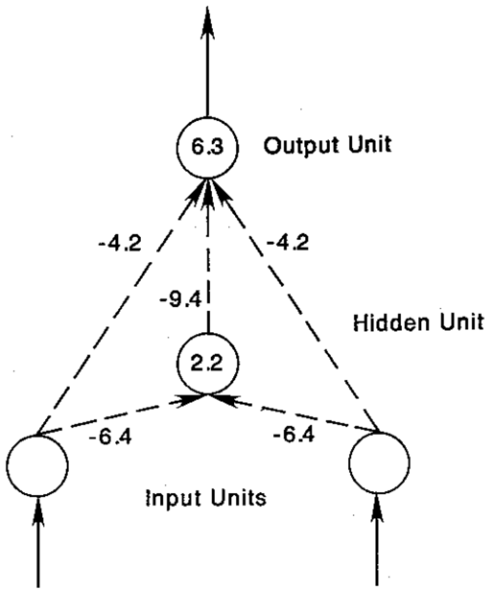
Efficiently compute gradient just by matrix and vector multiplications.

Gradient descent

Learning rate α

Update parameters: $\theta \to \theta - \alpha\nabla\mathcal{L}$

Better algorithms exist!

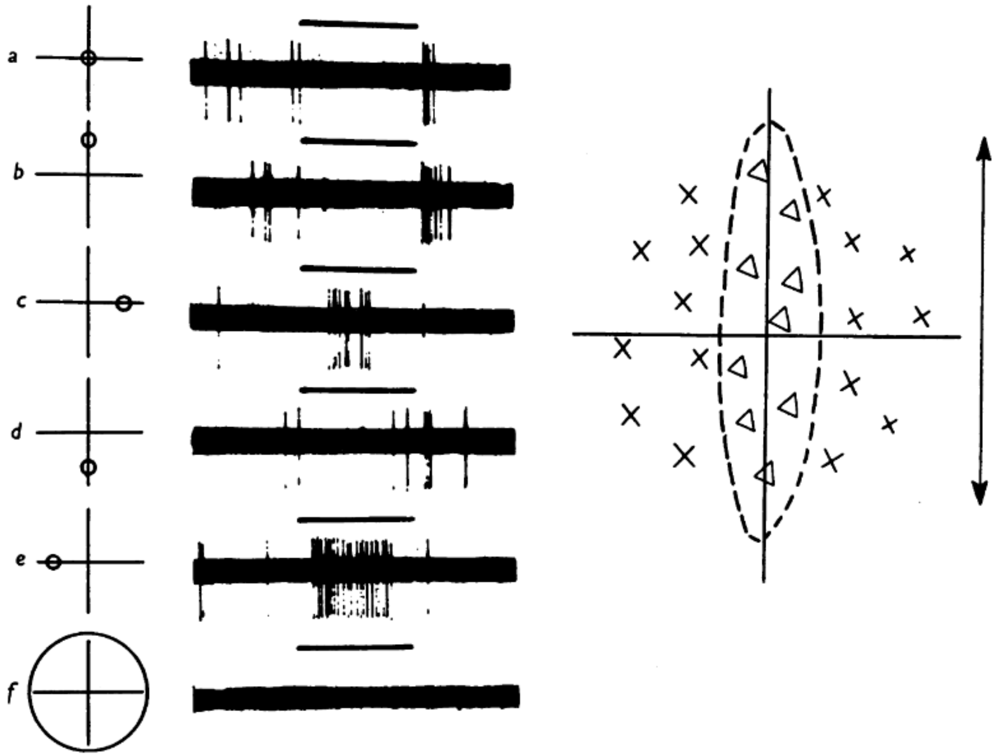


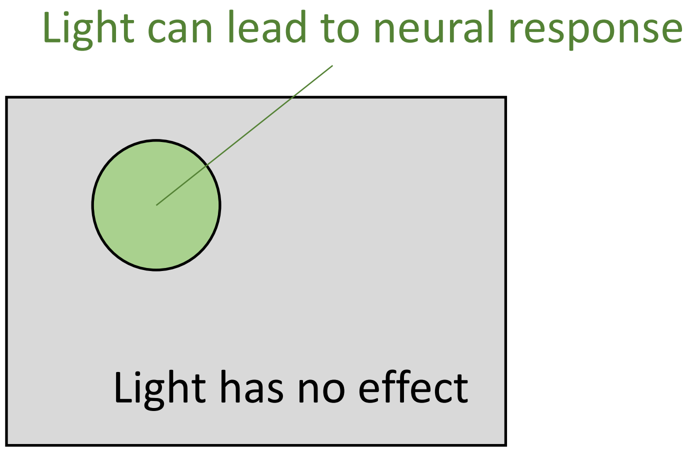
Discovered XOR network

The visual system and convolutional neural networks

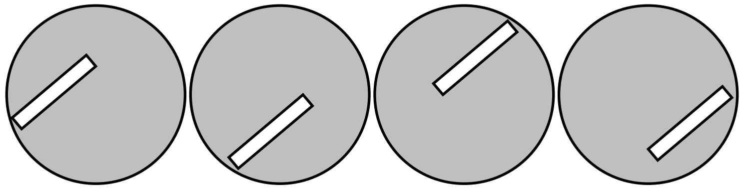
Hubel and Wiesel (1959, 1962)

Simple cell

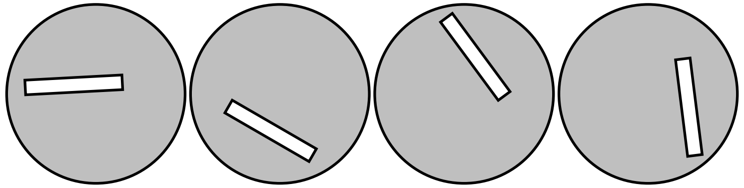




Complex cell



Same response: active



Same response: inactive

The visual system and convolutional neural networks

Hubel and Wiesel (1959, 1962)

Fukushima (1980) – Neocognitron

LeCun et al. (1989) – Convolutional Neural Network

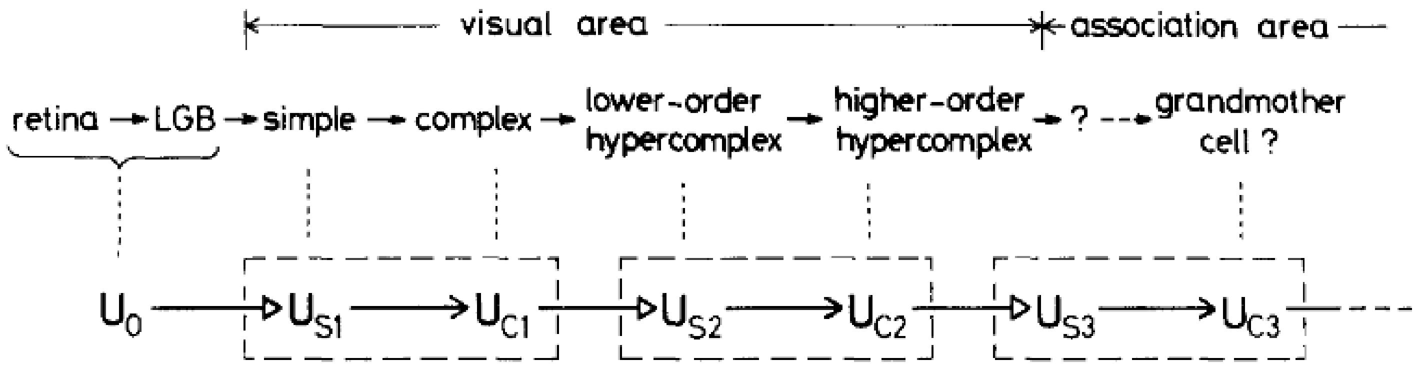
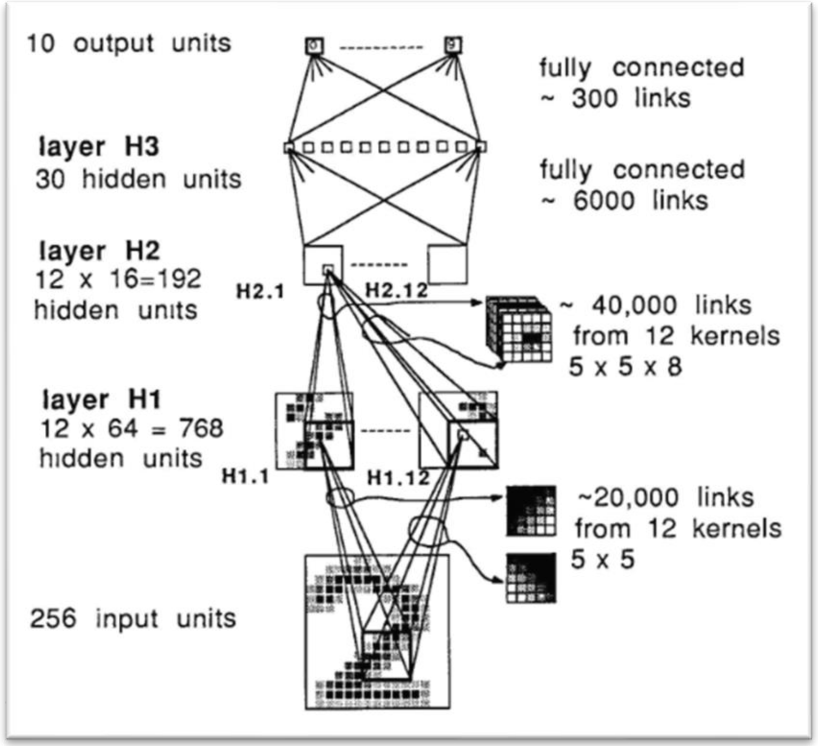


Fig. 1. Correspondence between the hierarchy model by Hubel and Wiesel, and the neural network of



LeCun et al. 1989

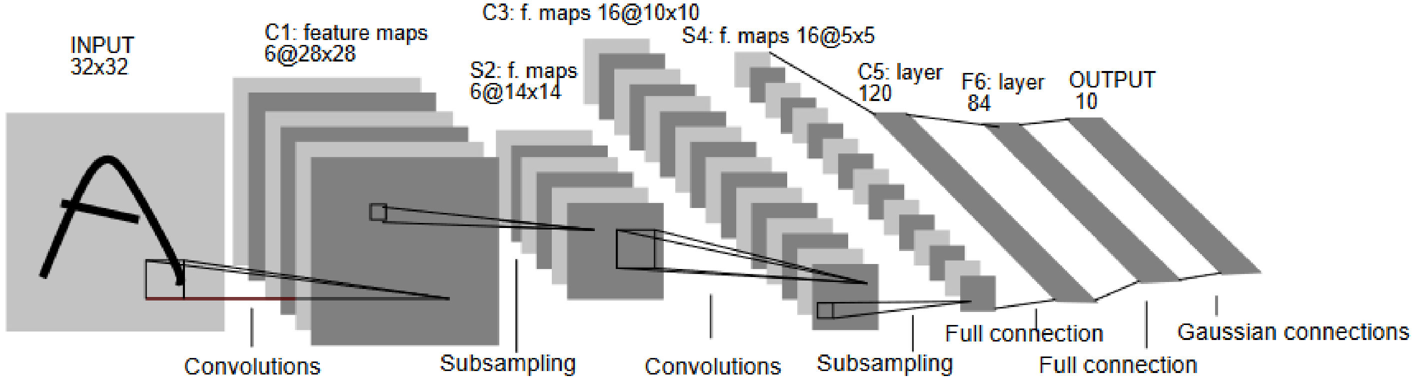
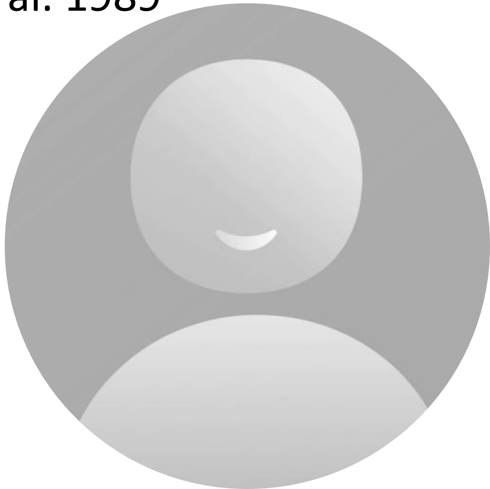


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

LeNet-5 from
LeCun et al. (1998)



The visual system and convolutional neural networks

Hubel and Wiesel (1959, 1962)

Fukushima (1980) – Neocognitron

LeCun et al. (1989) – Convolutional Neural Network

Back to neuroscience

Yamins et al. (2014) – Trained CNNs match neural data

Kell et al. (2018) – Same thing works for auditory cortex

Schrimpf et al. (2020) – Brain-Score leaderboard for best match to neural data

→ https://www.brain-score.org/

But…

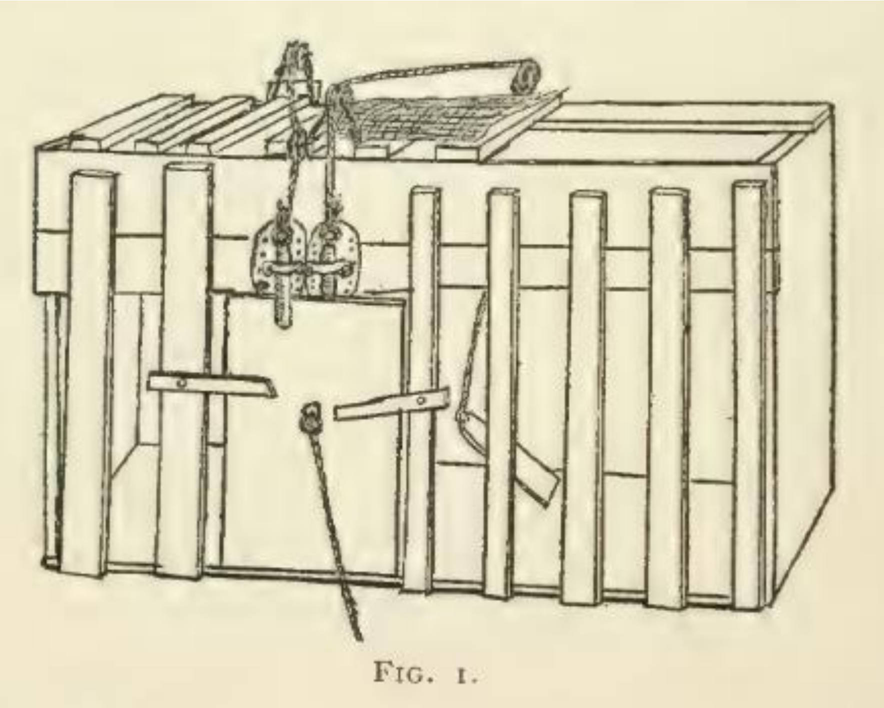
Jacob et al. (2021) – CNNs don't predict human behaviour out- of-distribution

Weerts et al. (2022), Adolfi et al. (2023) – Same thing for auditory

Reinforcement learning

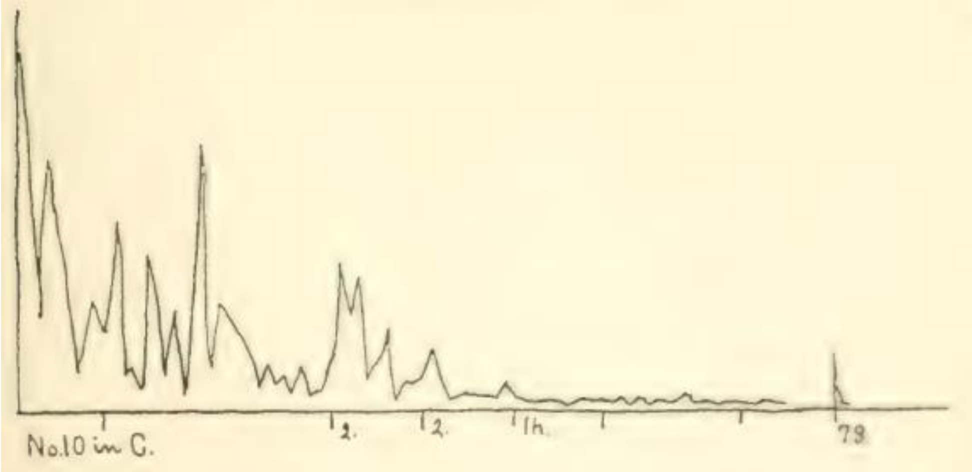
Thorndike (1898)

Pavlov (1927)



“ In the first place, most of the books do not give us a psychology, but rather a eulogy, of animals. They have all been about animal intelligence, never about animal stupidity… Dogs get lost hundreds of times and no one ever notices it or sends an account of it to a scientific magazine. But let one find his way from Brooklyn to Yonkers and the fact immediately becomes a circulating anecdote. ”

Time to escape cage



Trials

“our view abolishes and declares that the progress is not from little and simple to big and complicated, but from direct connections to indirect connections in which a stock of isolated elements plays a part, is from ‘ pure experience’ or undifferentiated feelings, to discrimination, on the one hand, to generalizations, abstractions, on the other ”

Reinforcement learning

Thorndike (1898)

Pavlov (1927)

Turing (1948)

Bellman (1950s)

Minsky (1950s-1960s)

Klopf (1980s)

Watkins (1989): Q-learning

Schultz (1980s-1990s): dopamine

Montague, Dayan, Sejnowski (1996): dopamine ↔ TD-learning

Hassabis:

- Postdoc with Dayan (2009)
- Founded DeepMind (2010)
- Hassabis et al. (2017)

Sutton and Barto “Reinforcement Learning: An Introduction”
http://incompleteideas.net/book/the-book.html

“ When a configuration is reached for which the action is undetermined, a random choice for the missing data is made and the appropriate entry is made in the description, tentatively, and is applied. When a pain stimulus occurs all tentative entries are cancelled, and when a pleasure stimulus occurs they are all made permanent. ”

Turing (1948)

Neuron
Review

Neuroscience-Inspired Artificial Intelligence

Demis Hassabis,[1,2,*] Dharshan Kumaran,[1,3] Christopher Summerfield,[1,4] and Matthew Botvinick[1,2]
[1]DeepMind, 5 New Street Square, London, UK
[2]Gatsby Computational Neuroscience Unit, 25 Howland Street, London, UK
[3]Institute of Cognitive Neuroscience, University College London, 17 Queen Square, London, UK
[4]Department of Experimental Psychology, University of Oxford, Oxford, UK
*Correspondence: dhcontact@google.com
http://dx.doi.org/10.1016/j.neuron.2017.06.011

What next?

Read more

Hassabis et al. (2017) “Neuroscience-Inspired Artificial Intelligence”

Zador et al. (2023) “Catalyzing next-generation Artificial Intelligence through NeuroAI”

Doerig et al. (2023) “The neuroconnectionist research programme”

Stay up to date

With Patrick Mineault’s “NeuroAI archive”

https://www.neuroai.science/

https://xcorr.net/