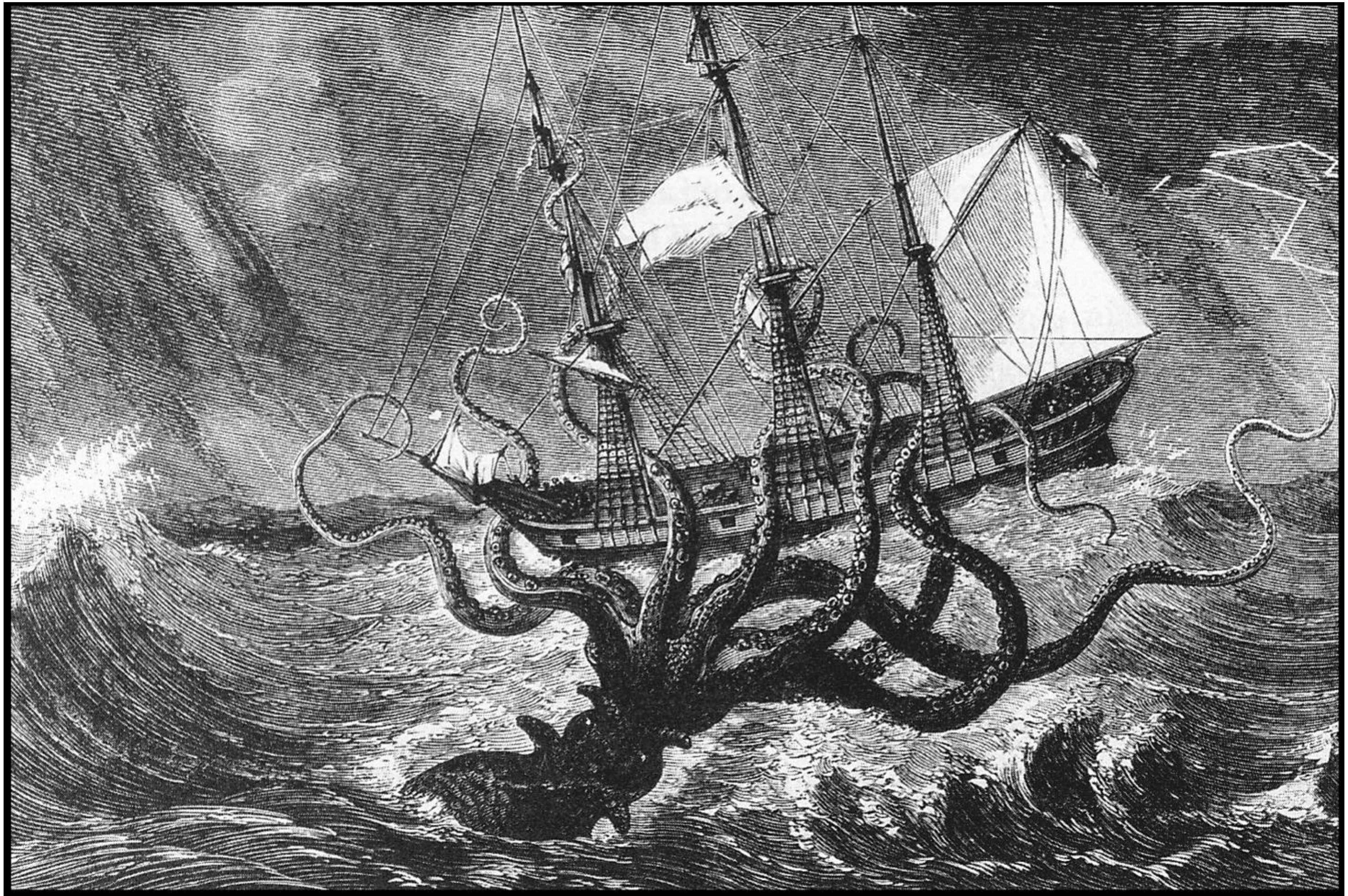
Deep Learning

Deep Learning

Course Website: https://sites.google.com/view/iiits-deep-learning

Course Email Id: iiits.dl@gmail.com

What is “deep learning”?



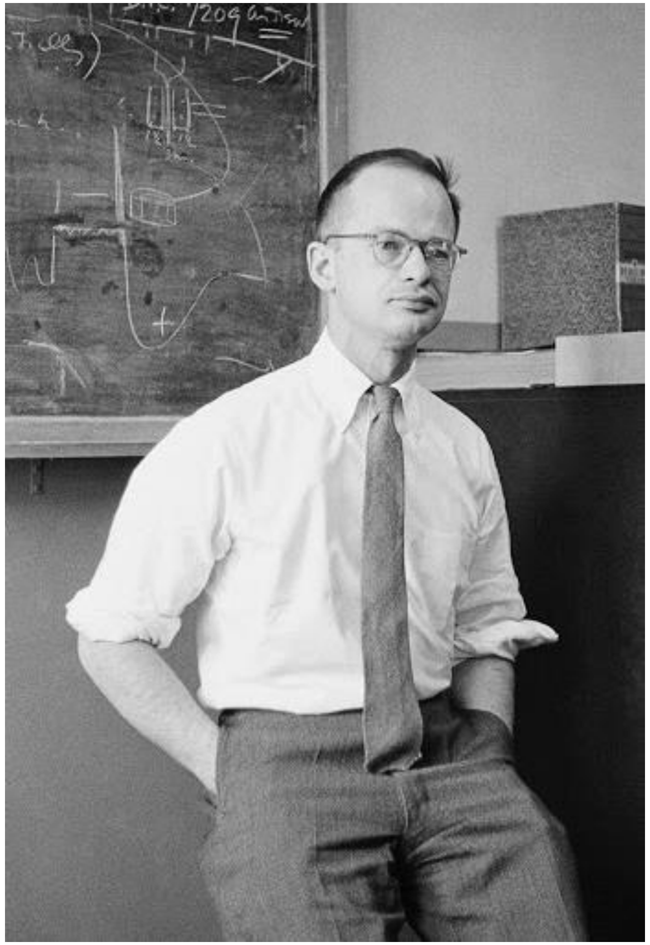
Deep Learning

- What is “learning”?
  - Improving performance through experience
  - Getting a computer to do well on a task without manually building in competence
- What is “deep”?
  - Learning using multi-layer neural networks

- What is the relationship between deep learning, ML, and AI?

An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
  - Fascinating reading: The Man Who Tried to Redeem the World with Logic, Nautilus, 2/5/2015



Walter Pitts (1923-1969)

An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt's perceptron

Frank Rosenblatt (1928-1971)

NEW NAVY DEVICE LEARNS BY DOING

Psychologist Shows Embryo of Computer Designed to Read and Grow Wiser

WASHINGTON, July 7 (UPI) —The Navy revealed the embryo of an electronic computer today that it expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence.

The embryo—the Weather Bureau's $2,000,000 "704" computer—learned to differentiate between right and left after fifty attempts in the Navy's demonstration for newsmen.

The service said it would use this principle to build the first of its Perceptron thinking machines that will be able to read and write. It is expected to be finished in about a year at a cost of $100,000.

Dr. Frank Rosenblatt, designer of the Perceptron, conducted the demonstration. He said the machine would be the first device to think as the human brain. As do human beings, Perceptron will make mistakes at first, but will grow wiser as it gains experience, he said.

Dr. Rosenblatt, a research psychologist at the Cornell Aeronautical Laboratory, Buffalo, said Perceptrons might be fired to the planets as mechanical space explorers.

Without Human Controls

The Navy said the perceptron would be the first non-living mechanism "capable of receiving, recognizing and identifying its surroundings without any human training or control."

The "brain" is designed to remember images and information it has perceived itself. Ordinary computers remember only what is fed into them on punch cards or magnetic tape.

Later Perceptrons will be able to recognize people and call out their names and instantly translate speech in one language to speech or writing in another language, it was predicted.

Mr. Rosenblatt said in principle it would be possible to build brains that could reproduce themselves on an assembly line and which would be conscious of their existence.

1958 New York Times...

In today's demonstration, the "704" was fed two cards, one with squares marked on the left side and the other with squares on the right side.

Learns by Doing

In the first fifty trials, the machine made no distinction between them. It then started registering a "Q" for the left squares and "O" for the right squares.

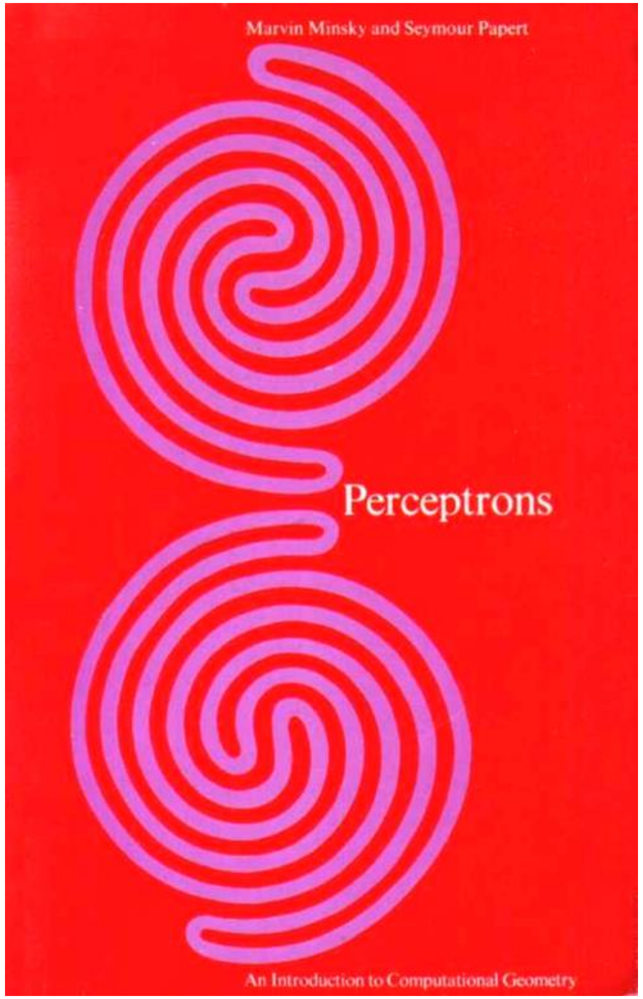
Dr. Rosenblatt said he could explain why the machine learned only in highly technical terms. But he said the computer had undergone a "self-induced change in the wiring diagram."

The first Perceptron will have about 1,000 electronic "association cells" receiving electrical impulses from an eye-like scanning device with 400 photo-cells. The human brain has 10,000,000,000 responsive cells, including 100,000,000 connections with the eyes.

An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt's perceptron
- 1969: Minsky and Papert Perceptrons book
  - Fascinating reading: M. Olazaran, A Sociological Study of the Official History of the Perceptrons Controversy, *Social Studies of Science*, 1996

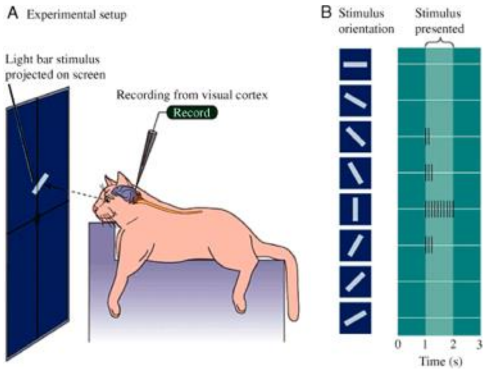
Perceptrons: an introduction to computational geometry



An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt's perceptron
- 1969: Minsky and Papert Perceptrons book
- 1980: Fukushima's Neocognitron
  - Video (short version)
  - Inspired by the findings of Hubel & Wiesel
    about the hierarchical organization
    of the visual cortex in cats and monkeys (1959-1977)

Kunihiko Fukushima



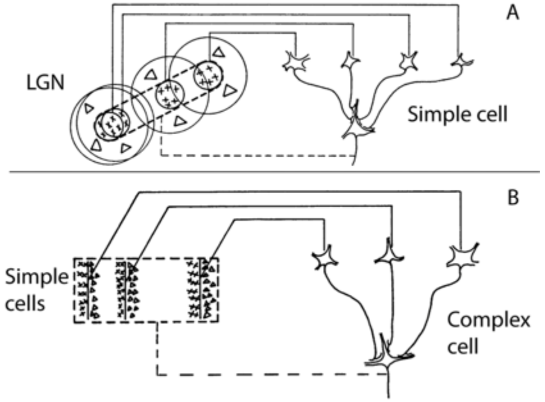
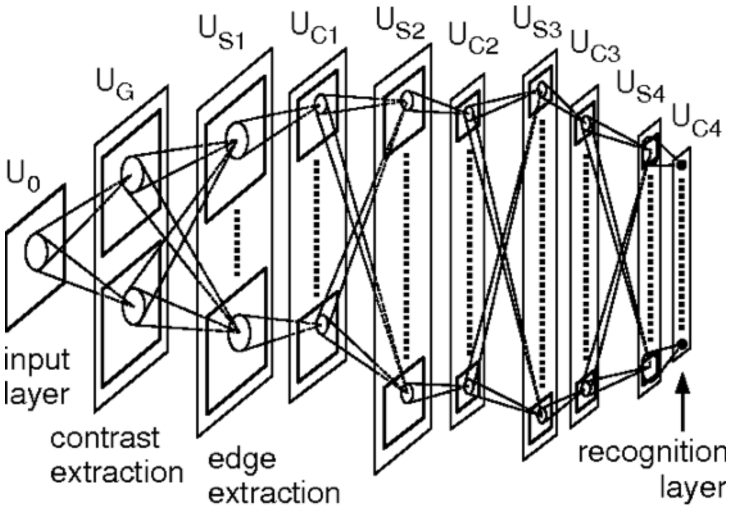


Image source



An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt’s perceptron
- 1969: Minsky and Papert Perceptrons book
- 1980: Fukushima’s Neocognitron
- 1986: Back-propagation
  - Origins in control theory and optimization: Kelley (1960), Dreyfus (1962), Bryson & Ho (1969), Linnainmaa (1970)
  - Application to neural networks: Werbos (1974)
  - Popularized by Rumelhart, Hinton & Williams (1986)

An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt’s perceptron
- 1969: Minsky and Papert Perceptrons book
- 1980: Fukushima’s Neocognitron
- 1986: Back-propagation
- 1989 – 1998: Convolutional neural networks
  - LeNet to LeNet-5

Yann LeCun

PROC. OF THE IEEE, NOVEMBER 1998 7

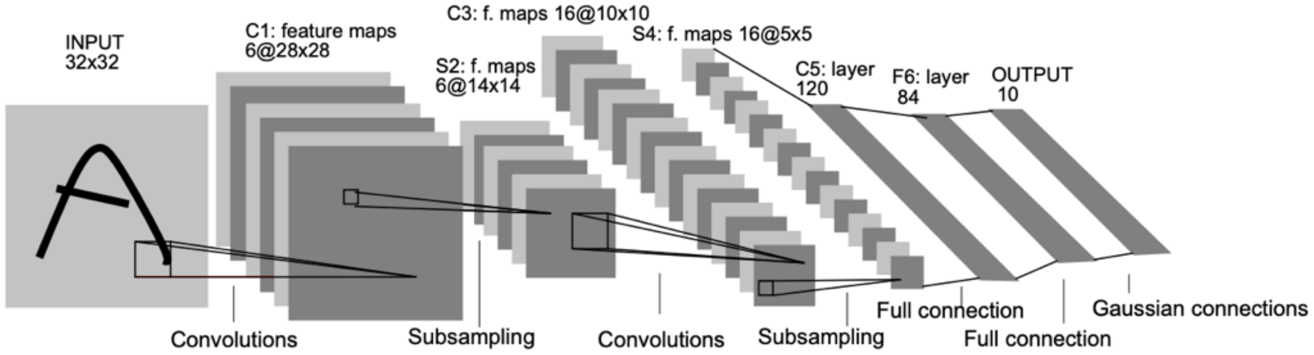
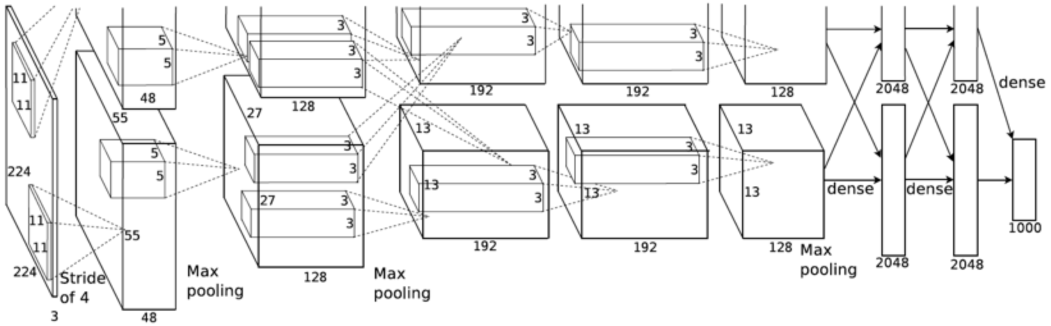
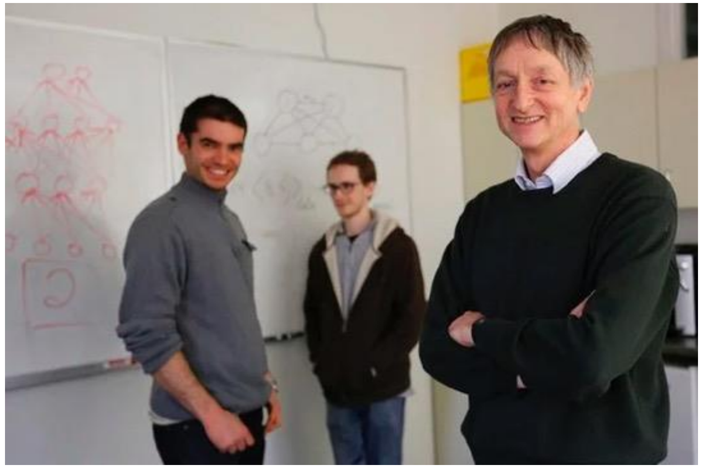


Fig. 2. Architecture of LeNet-5, a Convolutional Neural Network, here for digits recognition. Each plane is a feature map, i.e. a set of units whose weights are constrained to be identical.

An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt's perceptron
- 1969: Minsky and Papert Perceptrons book
- 1980: Fukushima's Neocognitron
- 1986: Back-propagation
- 1989 – 1998: Convolutional neural networks
- 2012: AlexNet





A Krizhevsky,
I Sutskever,
GE Hinton

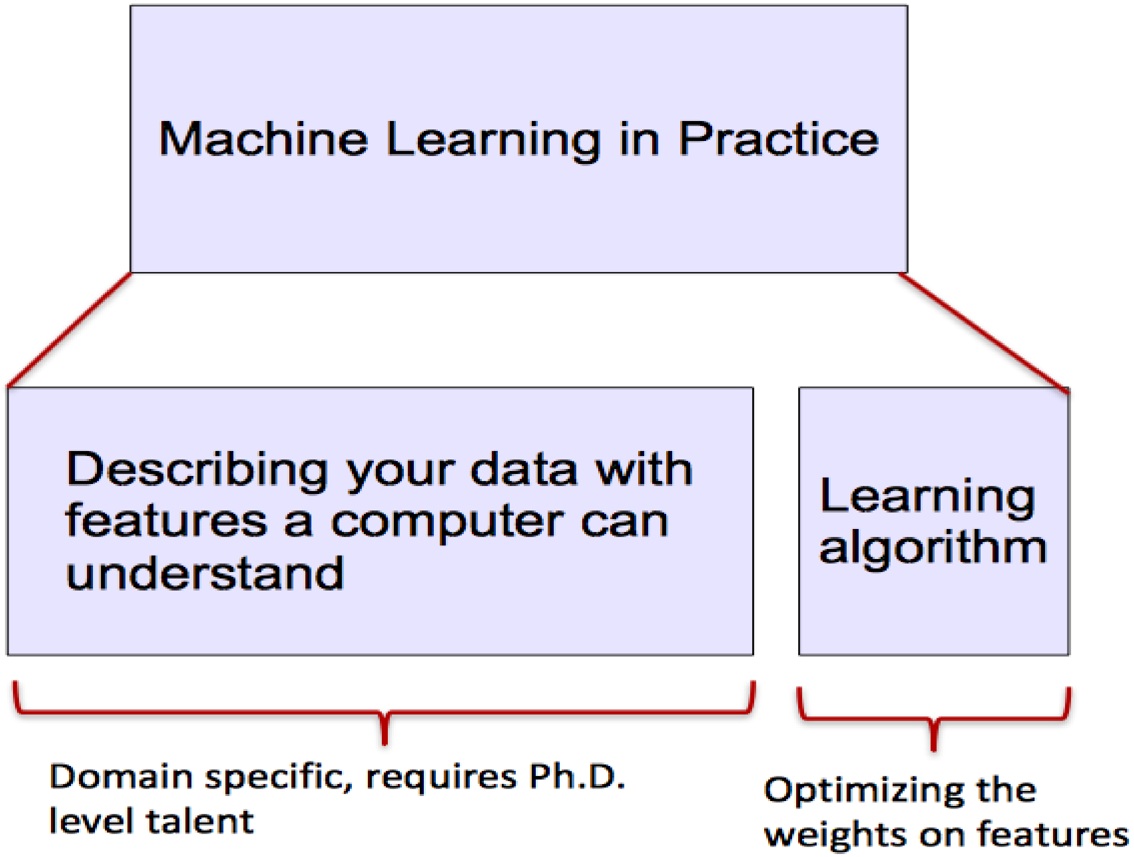
Photo source

An incomplete timeline of deep learning

- 1943: McCulloch and Pitts neurons
- 1958: Rosenblatt's perceptron
- 1969: Minsky and Papert Perceptrons book
- 1980: Fukushima's Neocognitron
- 1986: Back-propagation
- 1989 – 1998: Convolutional neural networks
- 2012: AlexNet
- 2018: ACM Turing Award
  to Hinton, LeCun, and Bengio

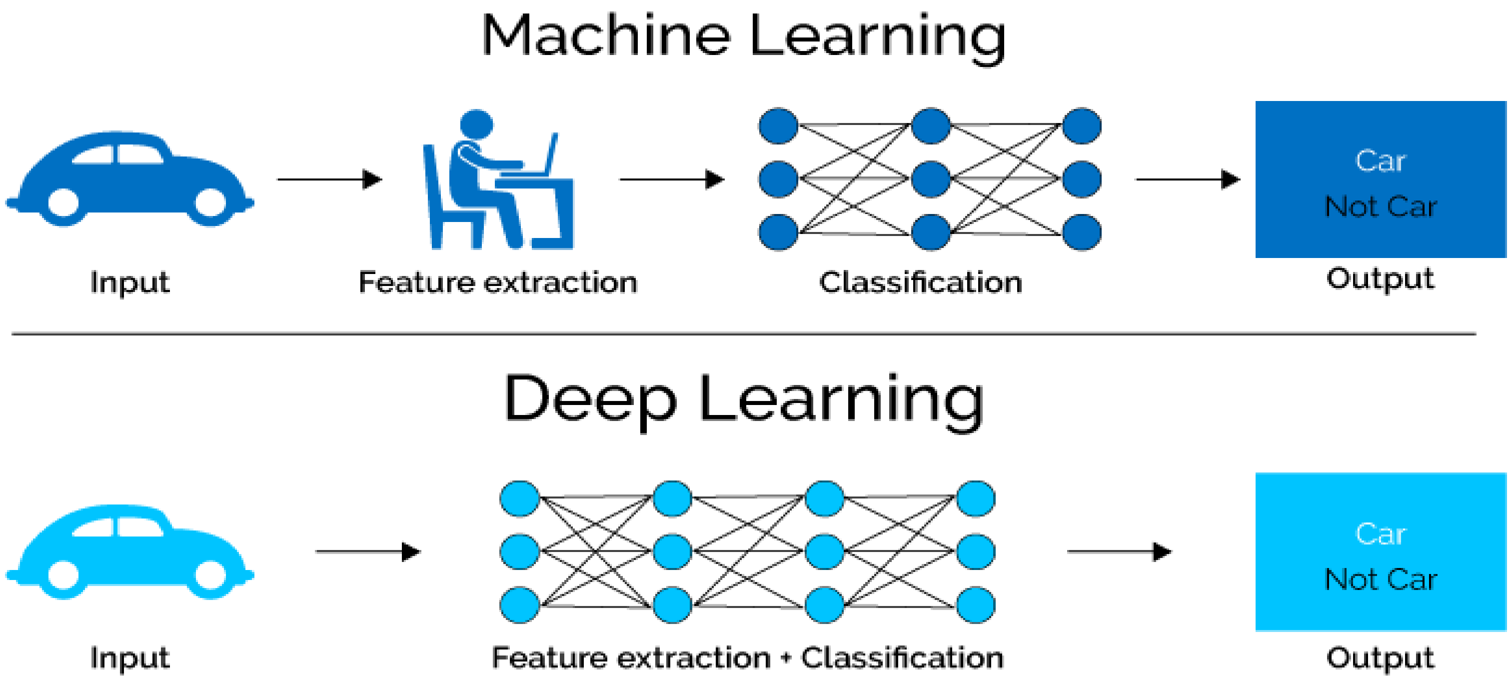
ML vs. Deep Learning

- **Most machine learning methods work well because of human-designed representations and input features**
- **ML becomes just optimizing weights to best make a final prediction**



What is Deep Learning (DL) ?

- **A machine learning subfield of learning representations of data. Exceptional effective at learning patterns.**
- **Deep learning algorithms attempt to learn (multiple levels of) representation by using a hierarchy of multiple layers**
- **If you provide the system tons of information, it begins to understand it and respond in useful ways.**

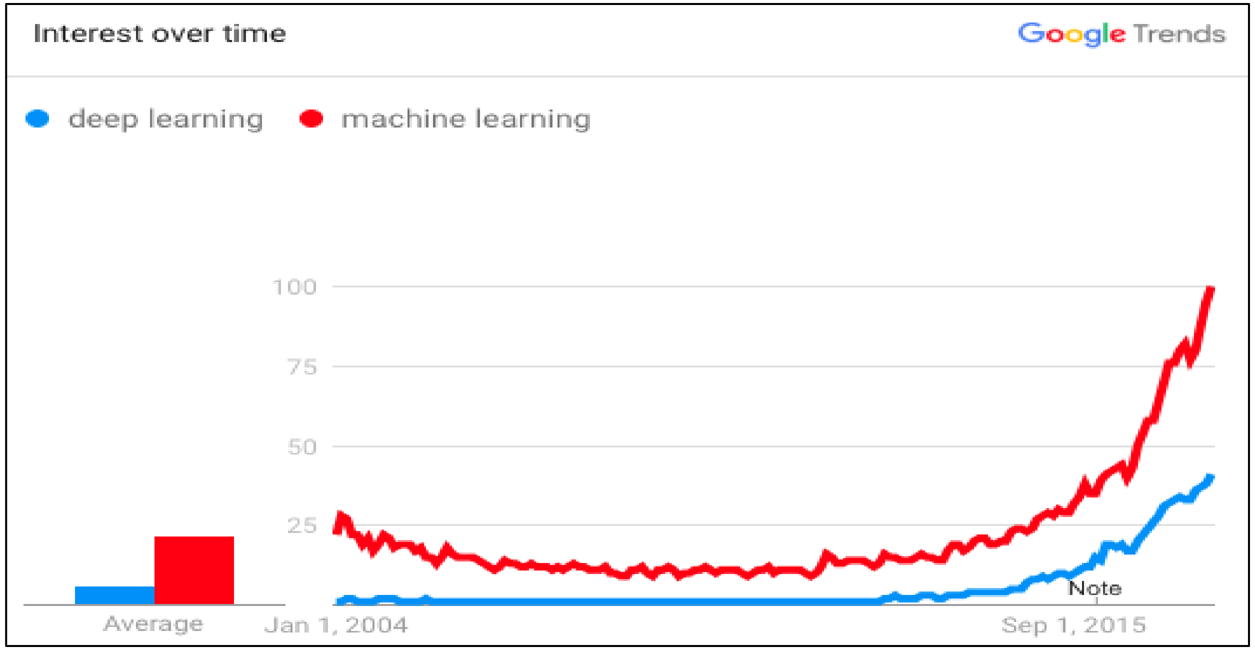


https://www.xenonstack.com/blog/static/public/uploads/media/machine-learning-vs-deep-learning.png

Why is DL useful?

- Manually designed features are often over-specified, incomplete and take a long time to design and validate
- Learned Features are easy to adapt, fast to learn
- Deep learning provides a very flexible, (almost?) universal, learnable framework for representing world, visual and linguistic information.
- Can learn both unsupervised and supervised
- Effective end-to-end joint system learning
- Utilize large amounts of training data

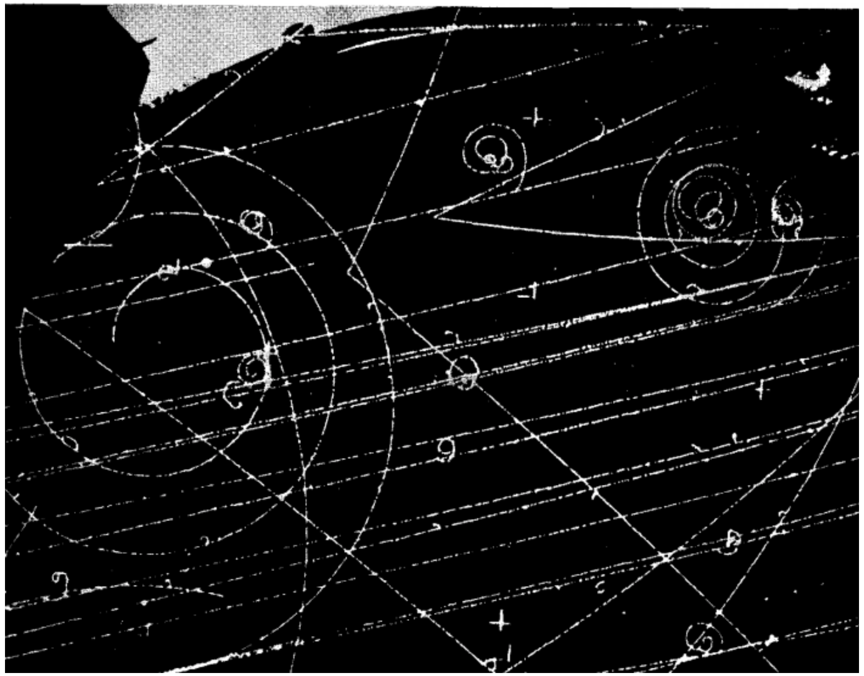
- In ~2010 DL started outperforming other ML techniques
- first in speech and vision, then NLP



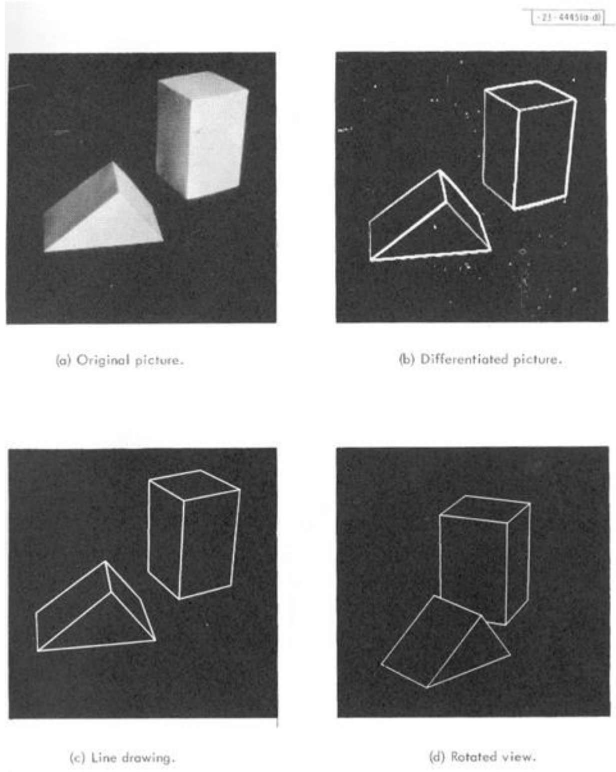
Successes of deep learning

- Vision
- Speech and Language
- Games
- Robotics

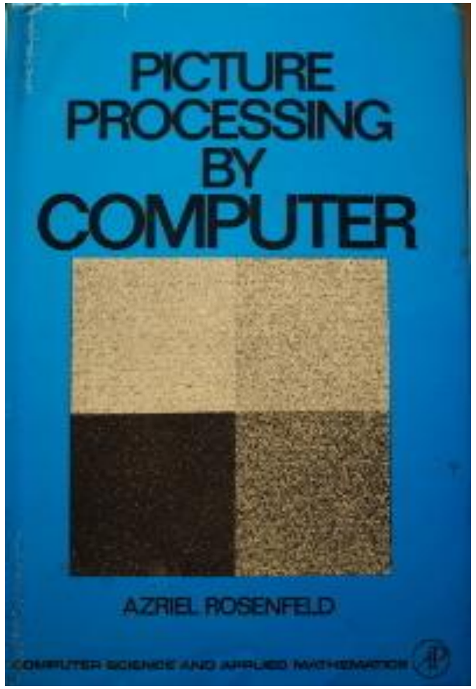
Vision: Origins



[Hough, 1959]



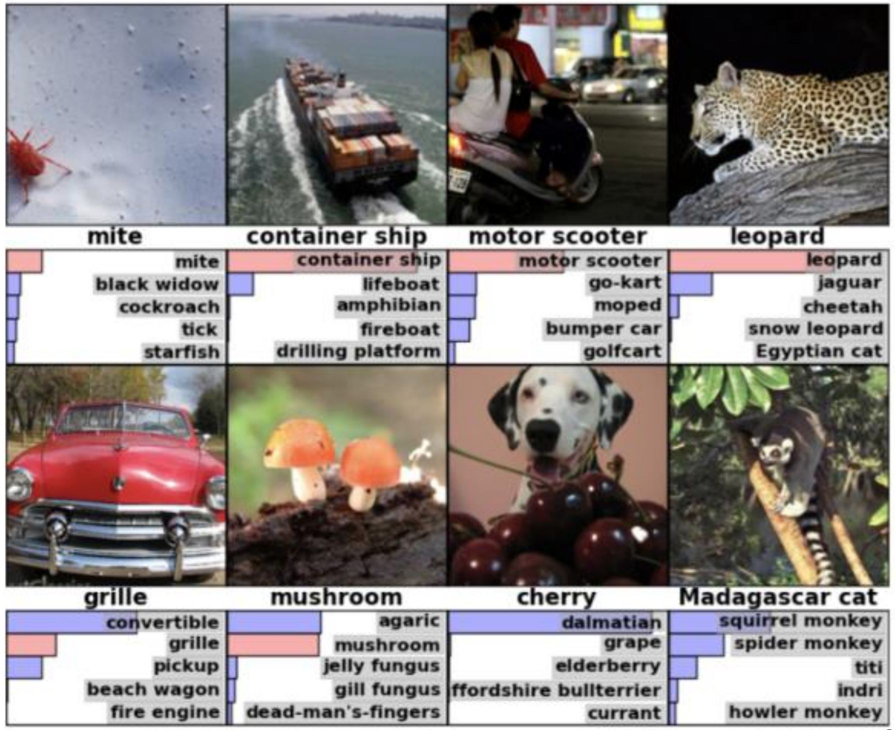
[Roberts, 1963]

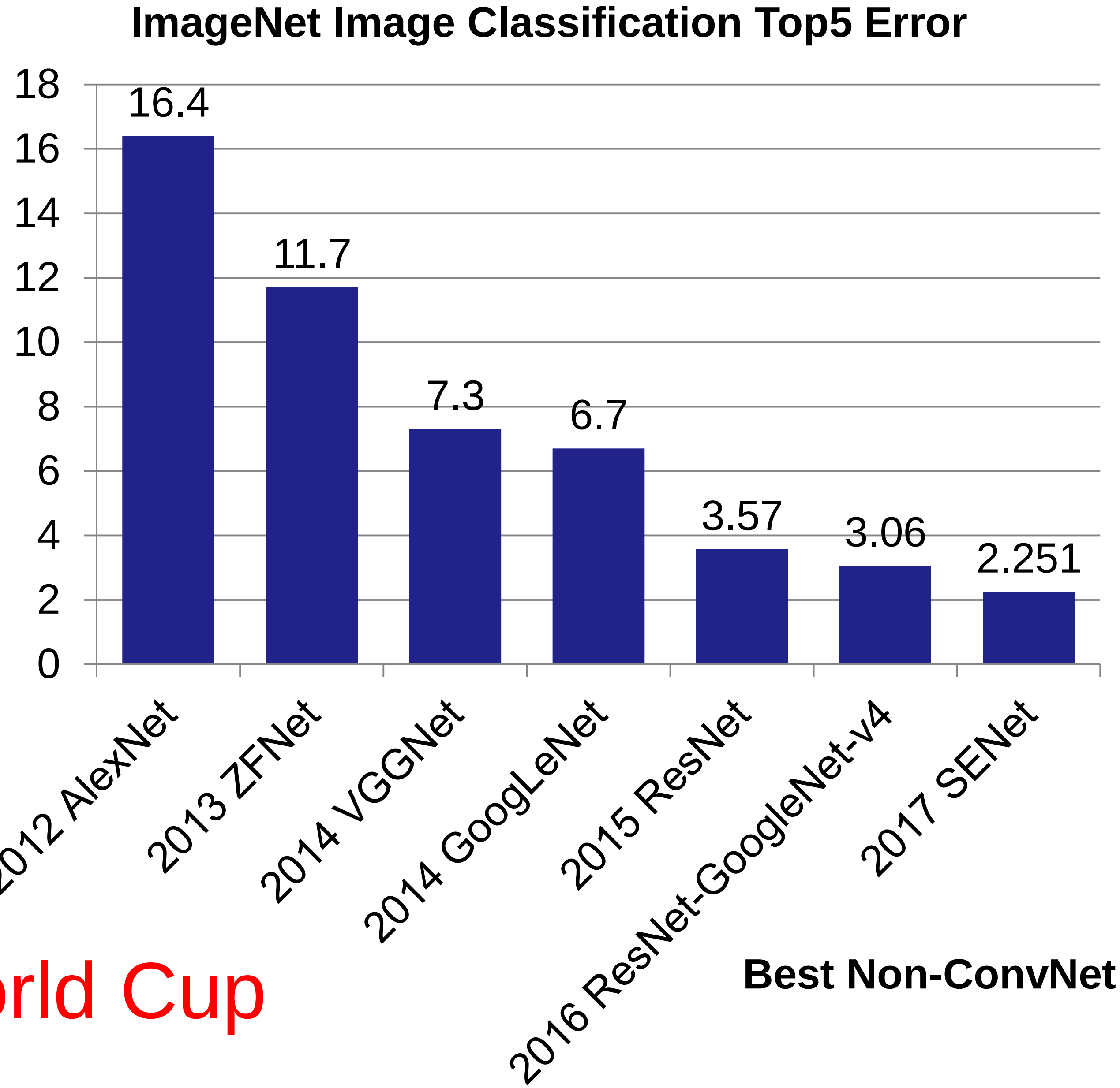


Rosenfeld, 1969

State of the Art in Vision: ImageNet Challenge

ILSVRC





ImageNet: The CV World Cup

Best Non-ConvNet in 2012: 26.2%

Vision: Detection, segmentation

K. He, G. Gkioxari, P. Dollar, and R. Girshick, [Mask R-CNN](),
ICCV 2017 (Best Paper Award)

Vision: Image generation

- Faces: 1024x1024 resolution, CelebA-HQ dataset

T. Karras, T. Aila, S. Laine, and J. Lehtinen, Progressive Growing of GANs for Improved Quality, Stability, and Variation, ICLR 2018

Follow-up work

Vision: Image generation

- BigGAN: Synthesize ImageNet images, conditioned on class label, up to 512 x 512 resolution

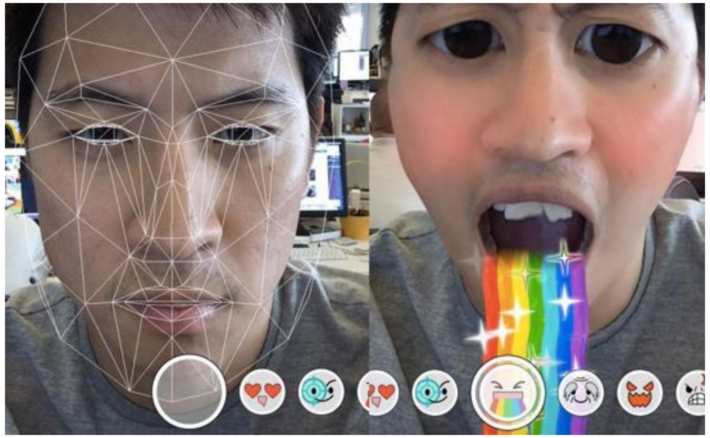
Difficult classes

A. Brock, J. Donahue, K. Simonyan, [Large scale GAN training for high fidelity natural image synthesis](), ICLR 2019

Vision

[Facebook accessibility tools for the visually impaired]

[AI beats human pathologists at detecting cancer]



[Technology behind Snapchat lenses]

Vision working too well? Face recognition

How China Uses High-Tech Surveillance to Subdue Minorities – New York Times, 5/22/2019

The Secretive Company That Might End Privacy As We Know It – New York Times, 1/18/2020

Wrongfully Accused by an Algorithm – New York Times, 6/24/2020

Vision working too well? DeepFakes

Harrison Ford Is Young Han In Solo Deepfake Video

Thanks to deepfake technology, the maligned Solo: A Star Wars Story now stars Harrison Ford instead of Alden Ehrenreich as the young Han.

BY DAN ZINSKI
2 DAYS AGO

Just a random recent example…

https://screenrant.com/star-wars-han-solo-movie-harrison-ford-video-deepfake/

https://www.youtube.com/watch?v=bC3uH4Xw4Xo

https://en.wikipedia.org/wiki/Deepfake

Vision working too well? DeepFakes

THE NEW YORKER

DEPT. OF TECHNOLOGY NOVEMBER 12, 2018 ISSUE

IN THE AGE OF A.I., IS SEEING STILL BELIEVING?

Advances in digital imagery could deepen the fake-news crisis—or help us get out of it.

By Joshua Rothman

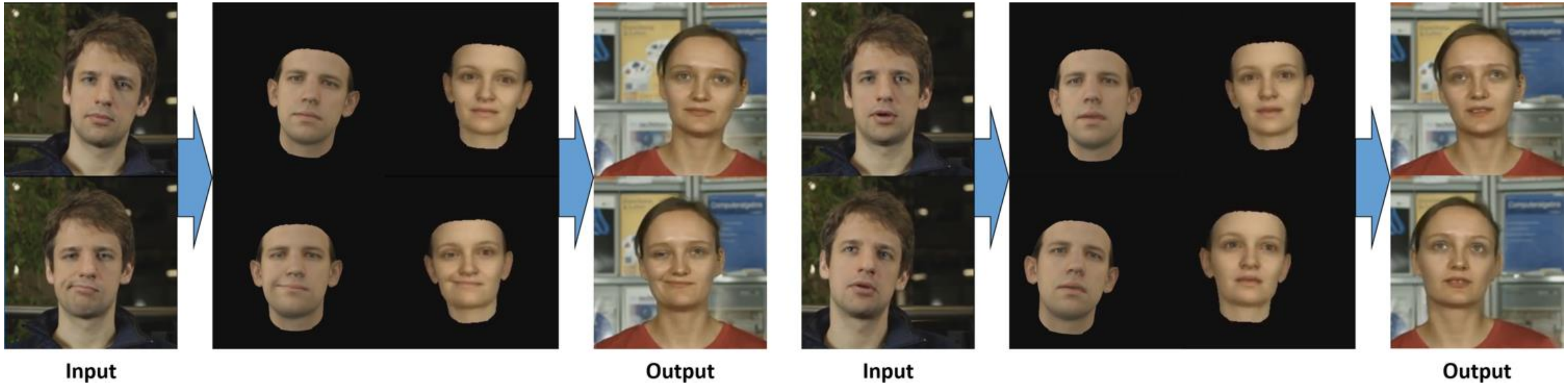
As synthetic media spreads, even real images will invite skepticism.

Illustration by Javier Jaén; photograph by Svetikd / Getty

https://www.newyorker.com/magazine/2018/11/12/in-the-age-of-ai-is-seeing-still-believing

Vision working too well? DeepFakes

- Example system: H. Kim et al., Deep video portraits, SIGGRAPH 2018

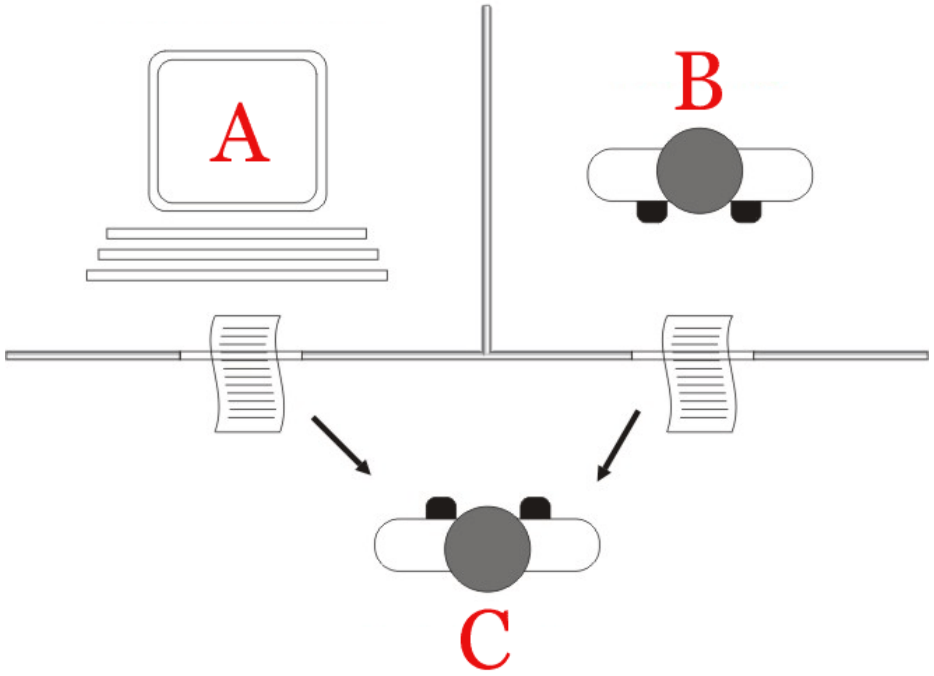


- Jordan Peele as Obama video

https://www.vox.com/2018/4/18/17252410/jordan-peele-obama-deepfake-buzzfeed

Natural language: Origins

- Turing test (1950)



The "standard interpretation" of the Turing test:

- Player C, the interrogator, is given the task of trying to determine which player – A or B – is a computer and which is a human.
- The interrogator is limited to using the responses to written questions to make the determination.

Natural language: Origins

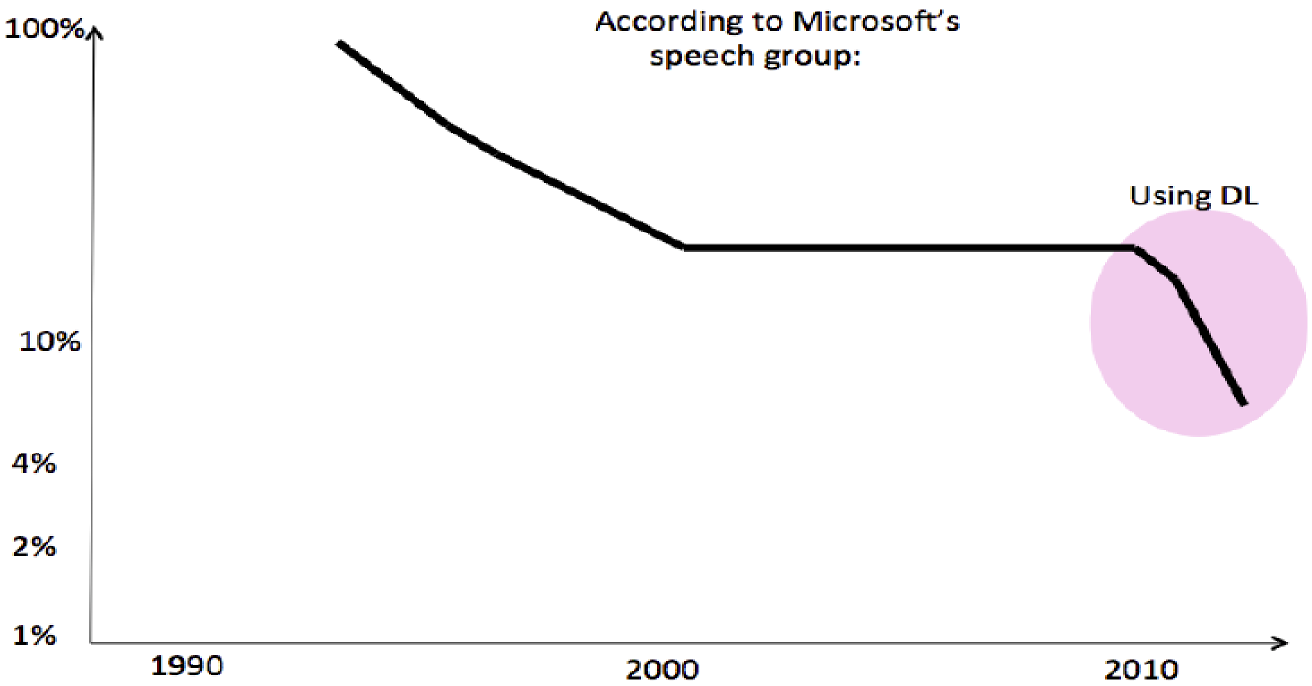
- Machine translation
  - 1954: Georgetown-IBM experiment
    - Completely automatic translation of more than sixty Russian sentences into English
    - Only six grammar rules, 250 vocabulary words, restricted to organic chemistry
    - Promised that machine translation would be solved in three to five years (press release)

Sentences in Russian are punched into standard cards for feeding into the electronic data processing machine for translation into English

  - 1966: Automatic Language Processing Advisory Committee (ALPAC) report: machine translation is not living up to the hype

- Chatbots: ELIZA (1966)
  - Simulated a psychotherapist, could fool naïve users

State of the Art in Speech and Language



Deep Learning in Speech Recognition

Several big improvements in recent years in NLP

- ✓ **Machine Translation**
- ✓ **Sentiment Analysis**
- ✓ **Dialogue Agents**
- ✓ **Question Answering**
- ✓ **Text Classification …**

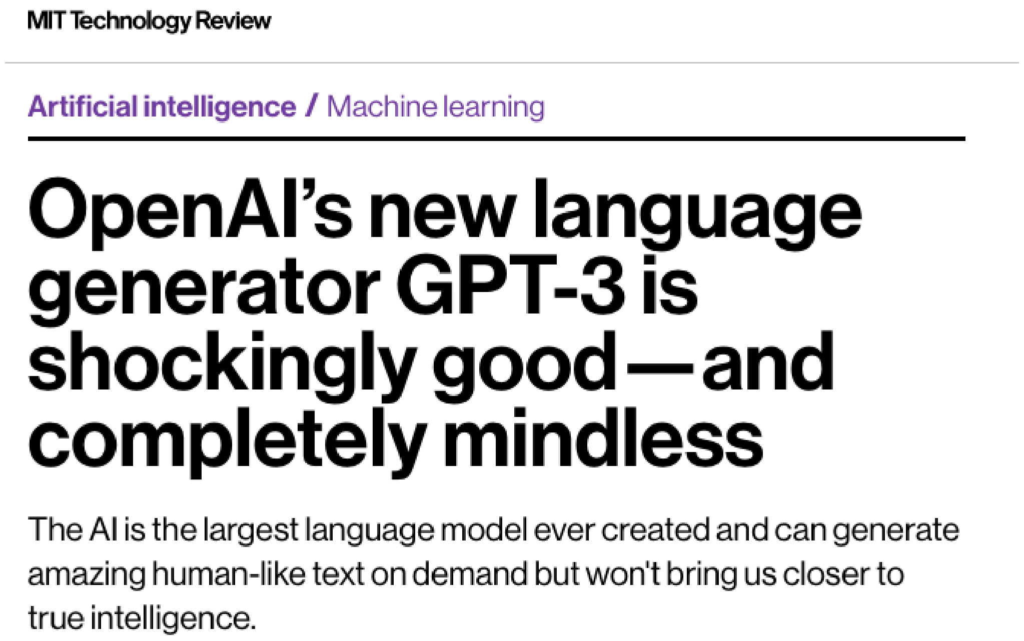
Leverage different levels of representation

- **words & characters**
- **syntax & semantics**

Successes in natural language

- Neural machine translation
  - The Great AI Awakening (Google Translator) – New York Times Magazine, 12/14/2016

- Language models: e.g., GPT-3 (Generative Pre-trained Transformer 3)



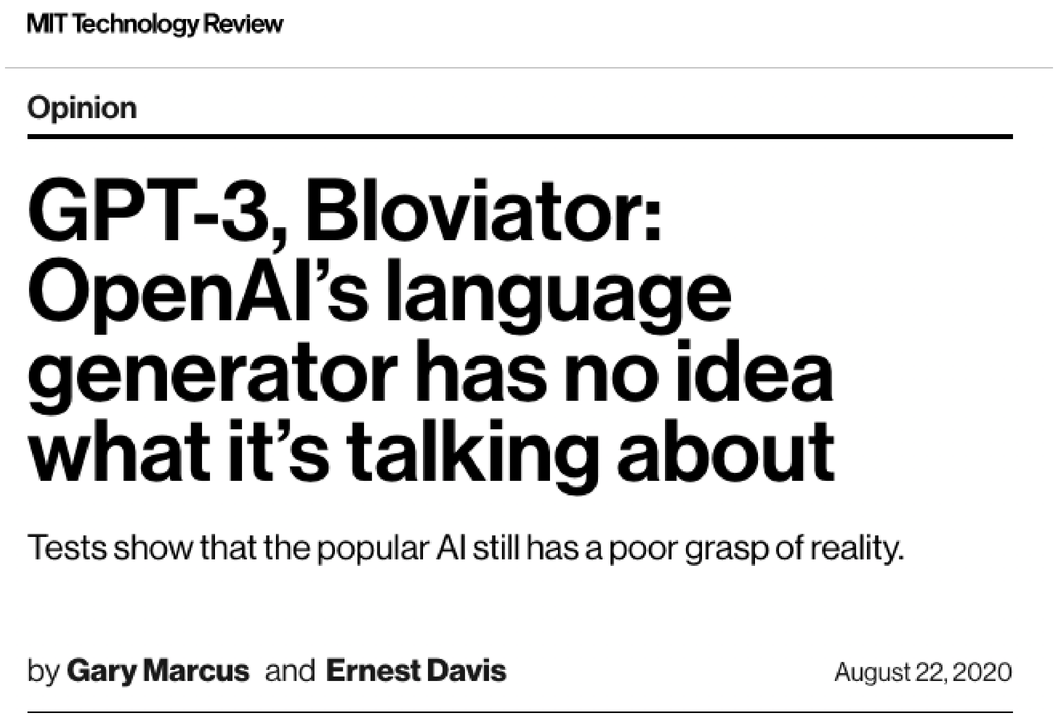
MIT Technology Review

Artificial intelligence / Machine learning

OpenAI's new language generator GPT-3 is shockingly good—and completely mindless

The AI is the largest language model ever created and can generate amazing human-like text on demand but won't bring us closer to true intelligence.

https://www.technologyreview.com/2020/07/20/1005454/openai-machine-learning-language-generator-gpt-3-nlp/



MIT Technology Review

Opinion

GPT-3, Bloviator: OpenAI's language generator has no idea what it's talking about

Tests show that the popular AI still has a poor grasp of reality.

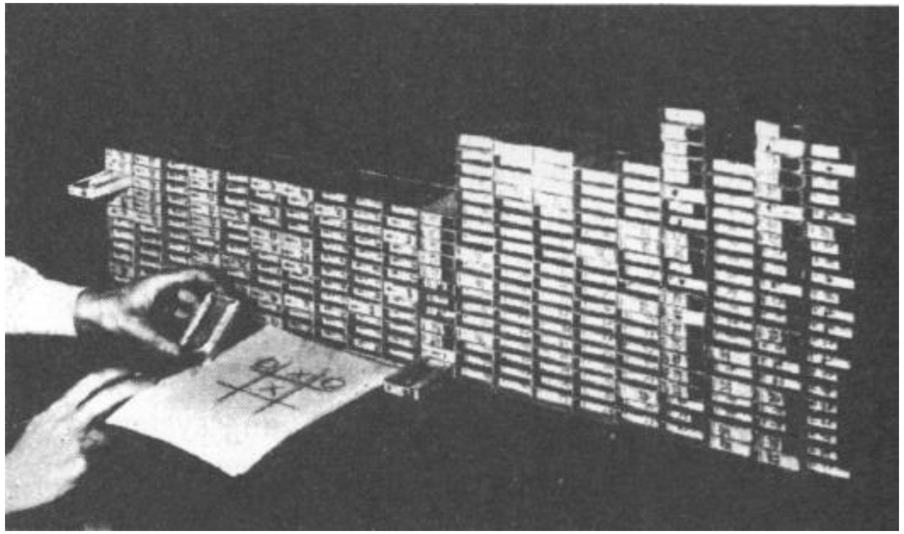
by **Gary Marcus** and **Ernest Davis** August 22, 2020

https://www.technologyreview.com/2020/08/22/1007539/gpt3-openai-language-generator-artificial-intelligence-ai-opinion/

Games: Origins

- 1952-1959: Arthur Samuel programmed a digital computer to learn to play checkers

- 1960: Donald Michie built a “machine” out of 304 matchboxes that could learn to play tic-tac-toe



Games: Origins

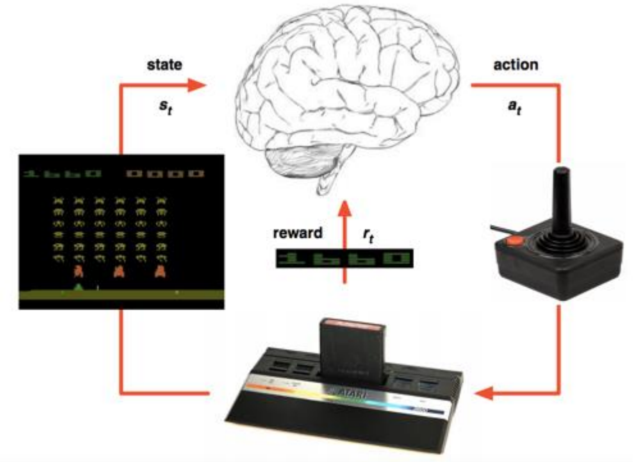
- *“In 1959 Arthur Samuel published a paper titled ‘Some Studies in Machine Learning Using the Game of Checkers’, the first time the phrase ‘Machine Learning’ was used”*

- *“Donald Michie’s description of reinforcement learning comes from 1961, and is the first use of the term reinforcement learning when applied to a machine process … There have been some developments in reinforcement learning since 1961, but only in details”*

Rodney Brooks essay, 8/28/2017

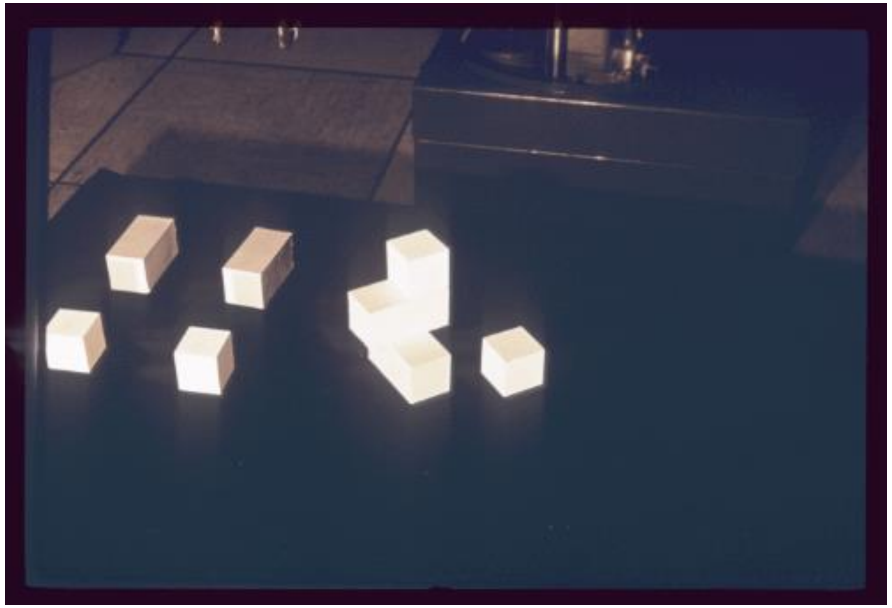
Successes in games

- 2013: DeepMind uses deep reinforcement learning to beat humans at some Atari games

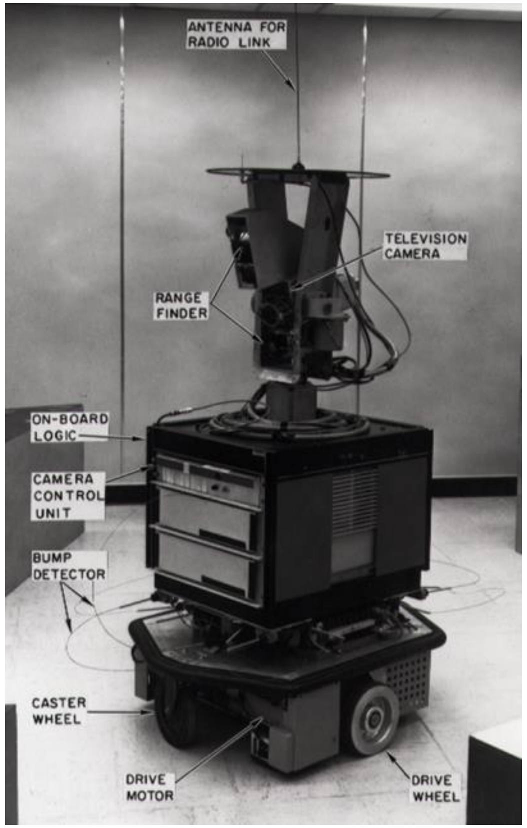


- 2016: DeepMind's AlphaGo system beats Go grandmaster Lee Sedol 4-1
- 2017: AlphaZero learns to play Go and chess from scratch
- 2019: DeepMind's StarCraft 2 AI is better than 99.8 percent of all human players

Robotics: Origins



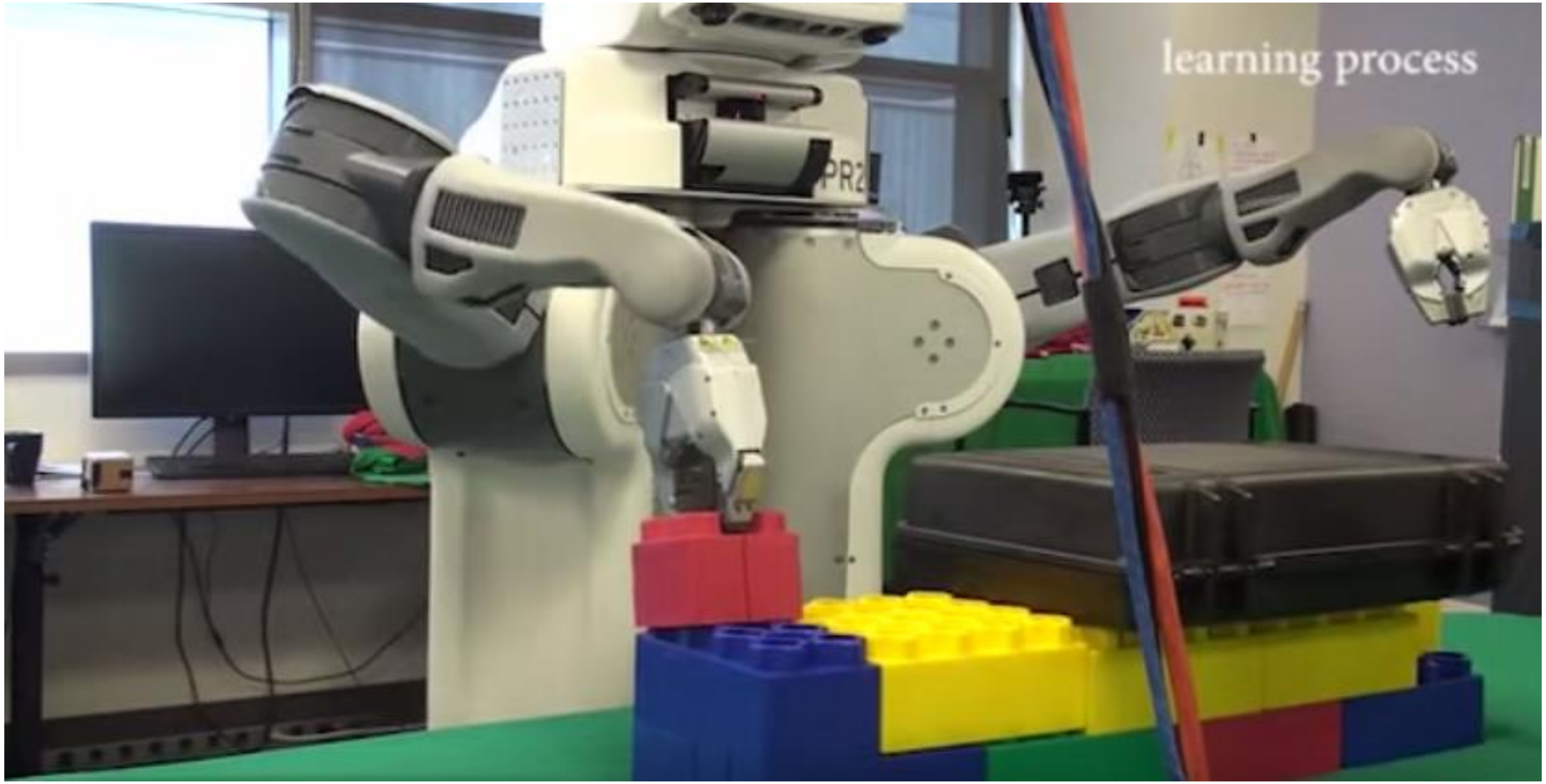
Blocks World
MIT, 1960s – 1970s
Copy demo (1970)



Shakey the Robot
SRI, 1966 – 1972
Video

Successes in embodied vision and robotics

- Sensorimotor learning



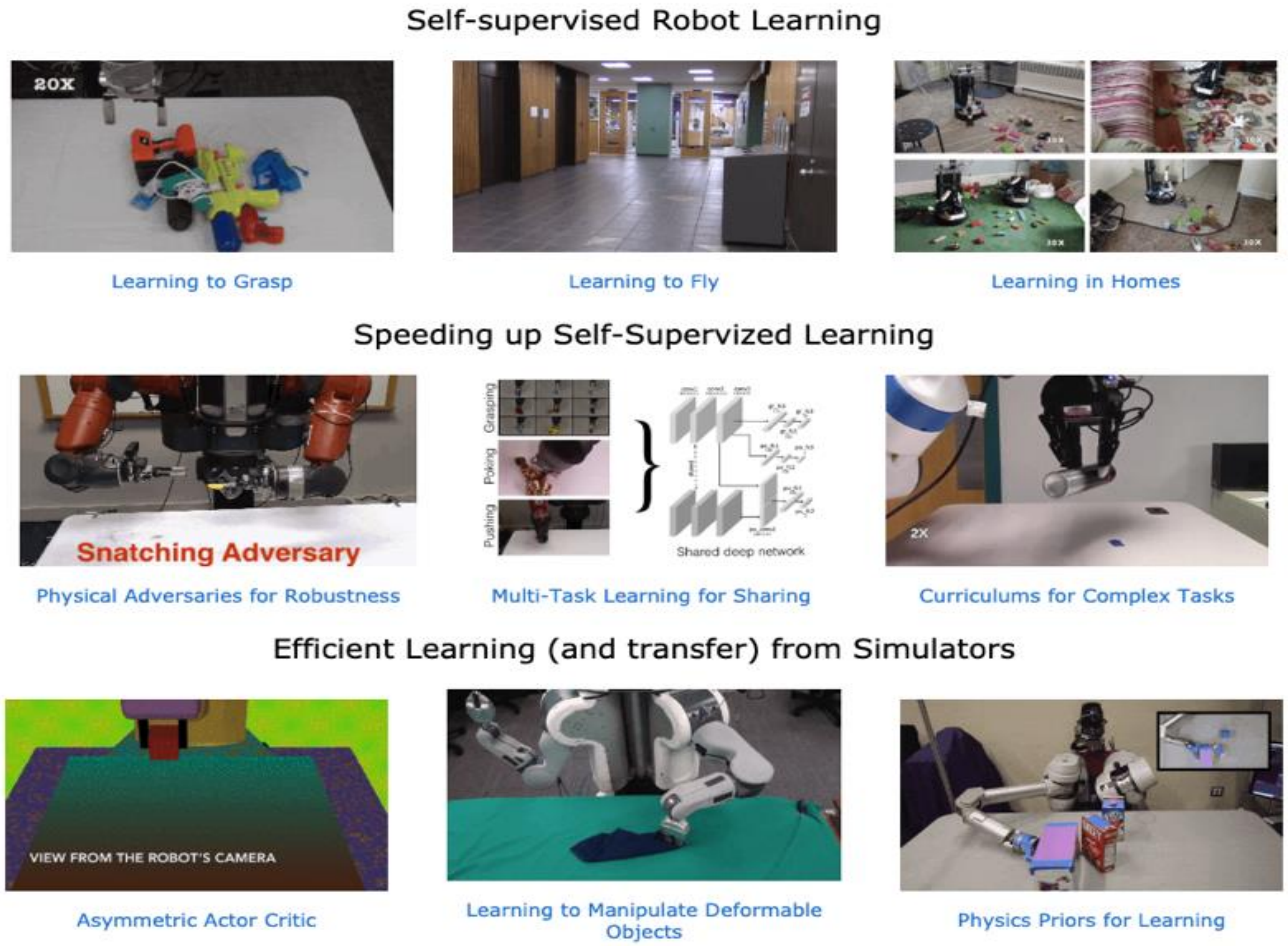
Overview video,

training video

S. Levine, C. Finn, T. Darrell, P. Abbeel, End-to-end training of deep visuomotor policies, JMLR 2016

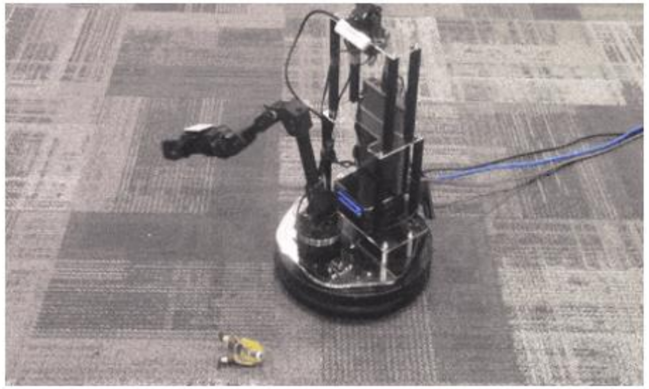
Embodied vision and robotics

A cross-section of topics:



Embodied platforms

- Simulation: AI2Thor, Habitat



- Real robots: PyRobot

- Robot on your smartphone: OpenBot

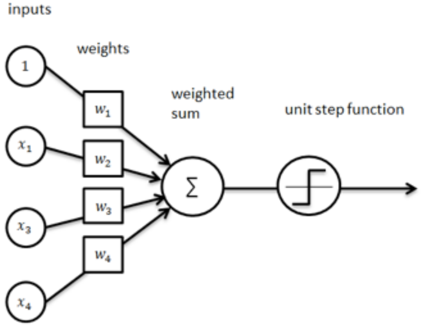
Self-driving cars

Image source

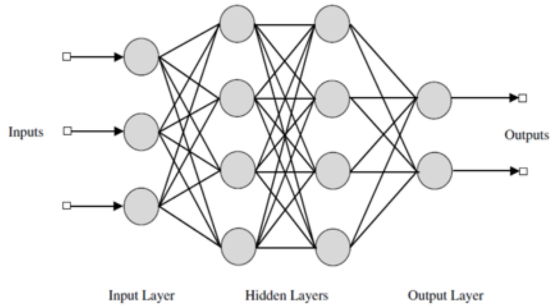
- Deep learning crucial for the global success of automotive autonomy – Automotive World, 6/26/2018

In this course

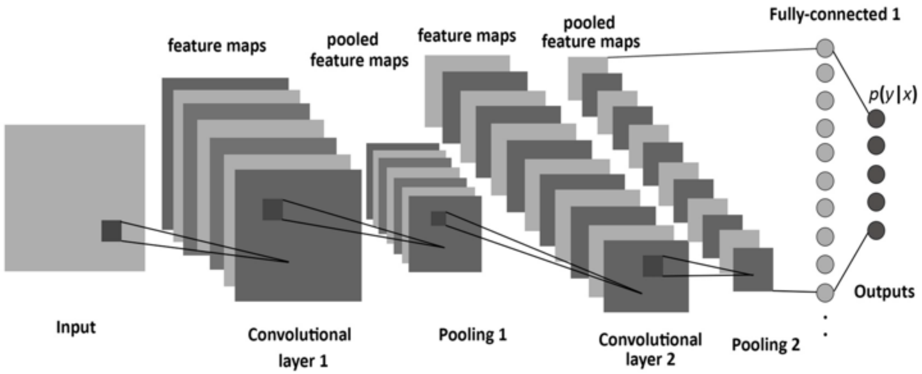
ML basics, linear classifiers



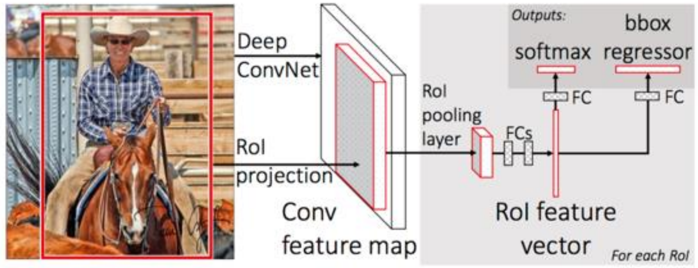
Multilayer neural networks, backpropagation



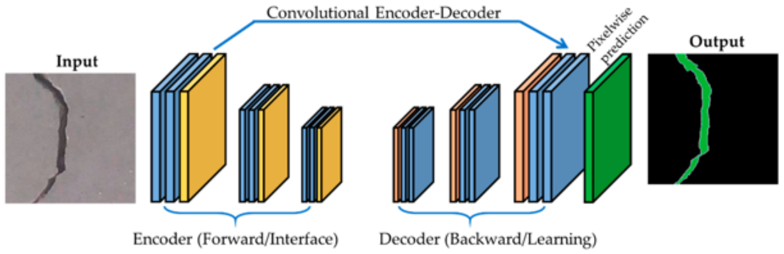
Convolutional networks for classification



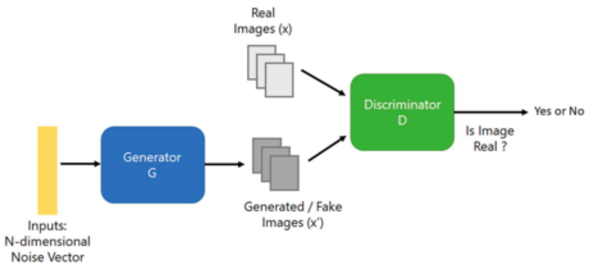
Networks for detection



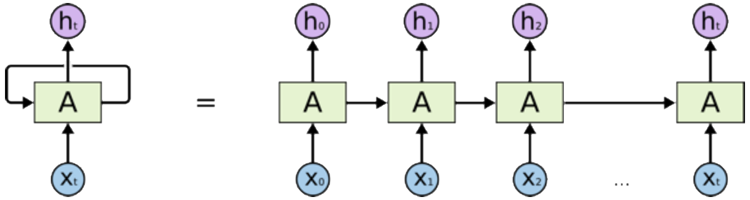
Networks for dense prediction



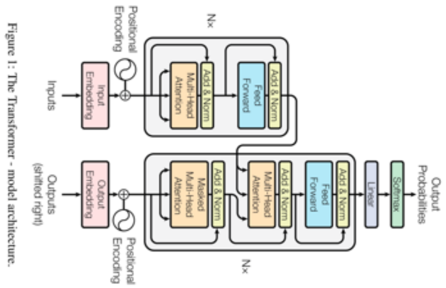
Generative models (GANs, VAEs)



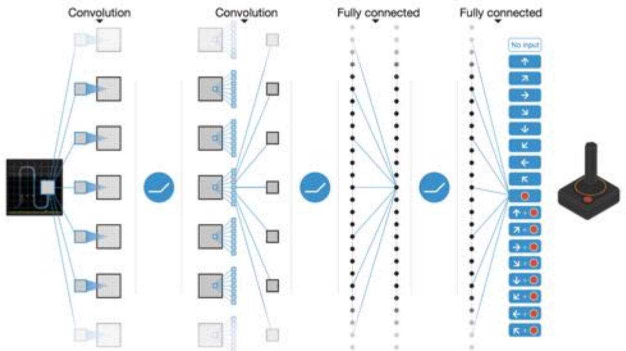
Recurrent models



Transformers



Deep reinforcement learning



Acknowledgement

- Deep Learning, Stanford University
- Introduction to Deep Learning, University of Illinois at Urbana-Champaign
- Introduction to Deep Learning, Carnegie Mellon University
- Convolutional Neural Networks for Visual Recognition, Stanford University
- Natural Language Processing with Deep Learning, Stanford University