GEORG-AUGUST-UNIVERSITÄT
GÖTTINGEN

Deep Learning

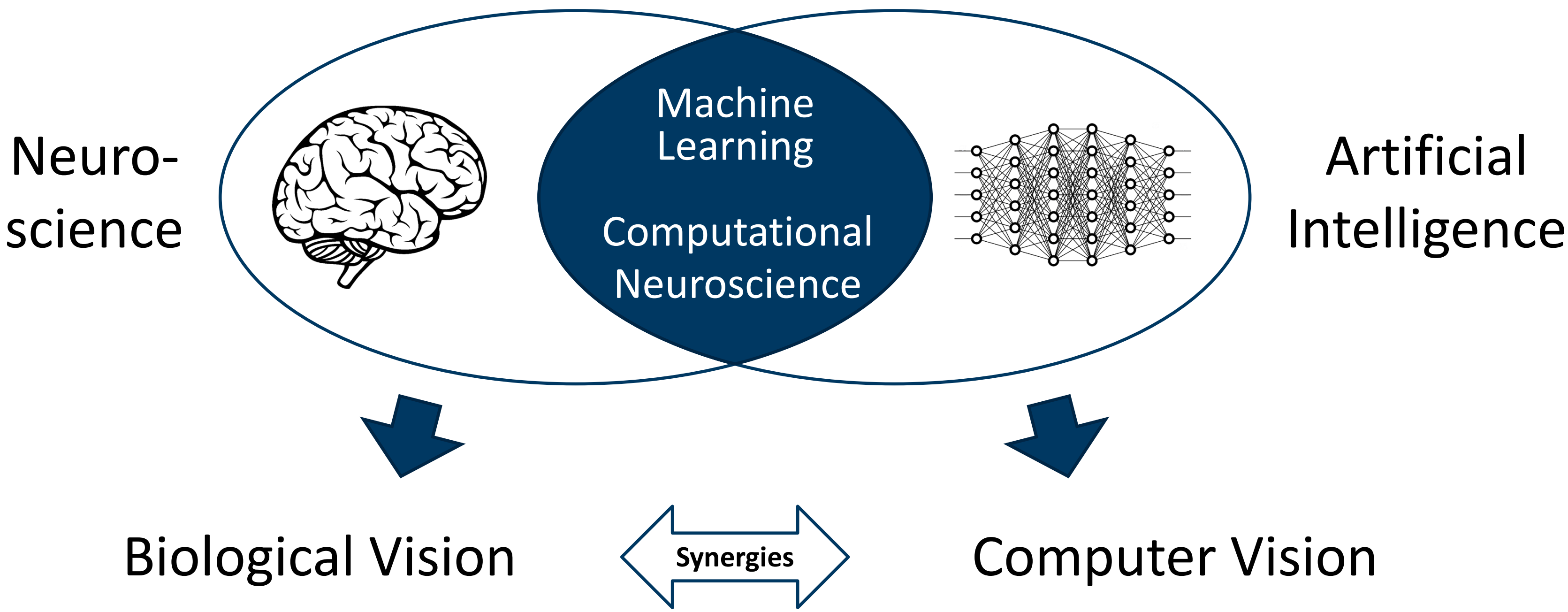
Lecture 1: Introduction and Overview

Alexander Ecker
Institut für Informatik, Uni Göttingen

alxecker https://alexanderecker.wordpress.com

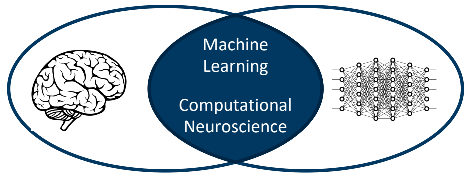
Neural Data Science Group

Algorithms of vision in brains and machines



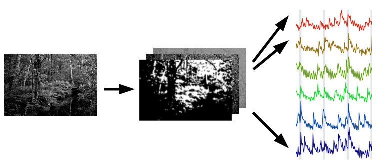
Neural Data Science Group

Neuro-
science



Artificial
Intelligence

Biological Vision

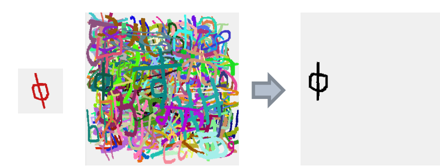


→ Large-scale system identification with CNNs
→ Structure & function

Synergies

→ Neural Style Transfer

Computer Vision



→ Image clutter
→ Few-shot learning
→ Object-centric vision

The group (started Oct. 1st)

Alexander Ecker

Group leader

Marissa Weis

PhD student

Vitus Benson

Undergraduate student

...

... and growing ...

You?

Context of this course

Starting next summer term: two **machine learning lectures** every year targeting M.Sc. and advanced B.Sc. students

1. Summer term: „**Machine Learning**“ *(starting next semester)*
2. Winter term: „**Deep Learning**“ *(this course)*

Additionally planned, summer or winter term:

3. Practical course (Fachpraktikum)
   „**Deep Learning for image processing and image recognition**“

This term we will cover the basics of machine learning in this course

Course topics

Overview and introduction (today)

Review of math basics

- linear algebra
- Probability
- information theory

Review of machine learning basics

- Empirical risk minimization
- Overfitting
- No free lunch theorem

Perceptron and multi-layer perceptron

Training deep neural networks

- Stochastic gradient descent
- Regularization
- Batch normalization

Convolutional neural networks (CNN)

Deep generative image models

- (Variational) Autoencoder
- Autoregressive models (PixelCNN)
- Generative adversarial networks (GAN)

Image-to-image models

- Domain translation networks
- Style transfer
- Superresolution

Recurrent neural networks (RNN)

- Gated recurrent unit (GRU)
- Long-short-term memory module (LSTM)
- Transformer networks

Deep word and text embeddings

- Word2Vec
- ELMo + BERT

Deep Reinforcement Learning

Recommended literature

Primary book:

- **Goodfellow, Bengio, Courville: Deep Learning**
  https://www.deeplearningbook.org

Good general machine learning book:

- Christopher Bishop: Pattern Recognition and Machine Learning
  https://www.microsoft.com/en-us/research/people/cmbishop/prml-book/

Good resource for Deep Reinforcement Learning (last week only):

- OpenAI: Spinning Up in Deep RL
  https://openai.com/blog/spinning-up-in-deep-rl/

Recommended prerequisites

Generally:

- Linear algebra
- Basic probability theory
- *(Basic machine learning concepts → Not required this term)*

B.Sc. Applied Data Science / Applied Computer Science:

- B.Inf.1131 – Data Science I: Algorithmen und Prozesse
- B.Inf.1234 – Machine Learning *(offered next term)*

→ Recommended for 5th semester

M.Sc. Applied Computer Science

- M.Inf.1151 – Data Science und Big Data Analytics
- M.Inf.XXXX – Machine Learning *(offered next term)*

→ Recommended for 3rd semester

General organization

Communication

- Announcements via **Stud.IP → Sign up** and monitor announcements

Lectures

- Mon & Thur, 2–4pm, MN 09 (this room)
- Starting Mon, Nov 25
- No lectures from Dec 2–13 (travel)

Homeworks

- Weekly homework assignments
- Groups of 2–3 students
- Weekly tutorial where students present their solution to tutors

Grade

- Final exam *(tentative date: Feb 10–21)*
- Requirement for taking exam: $N - 1$ homework solutions presented

Homeworks

Tutorials

- Room: Computer Pool Provisorium Informatik
- Mon 12–14
- Thur 16–18

Sign up for tutorials via Stud.IP until Nov. 10

- We will use a lottery (total: 2x52 spots; same procedure as Info I+II+III)
- Sign up for both tutorials if you can make it; you will be assigned to one

Tech stack

- Python 3 + PyTorch
- Jupyter notebooks

Special offer for MSc & PhD students

You are

- Familiar with basic ML concepts
- Fluent in Python
- (ideally some experience with PyTorch)

Then you can

- **Design one homework exercise instead of solving all of them**

How it works

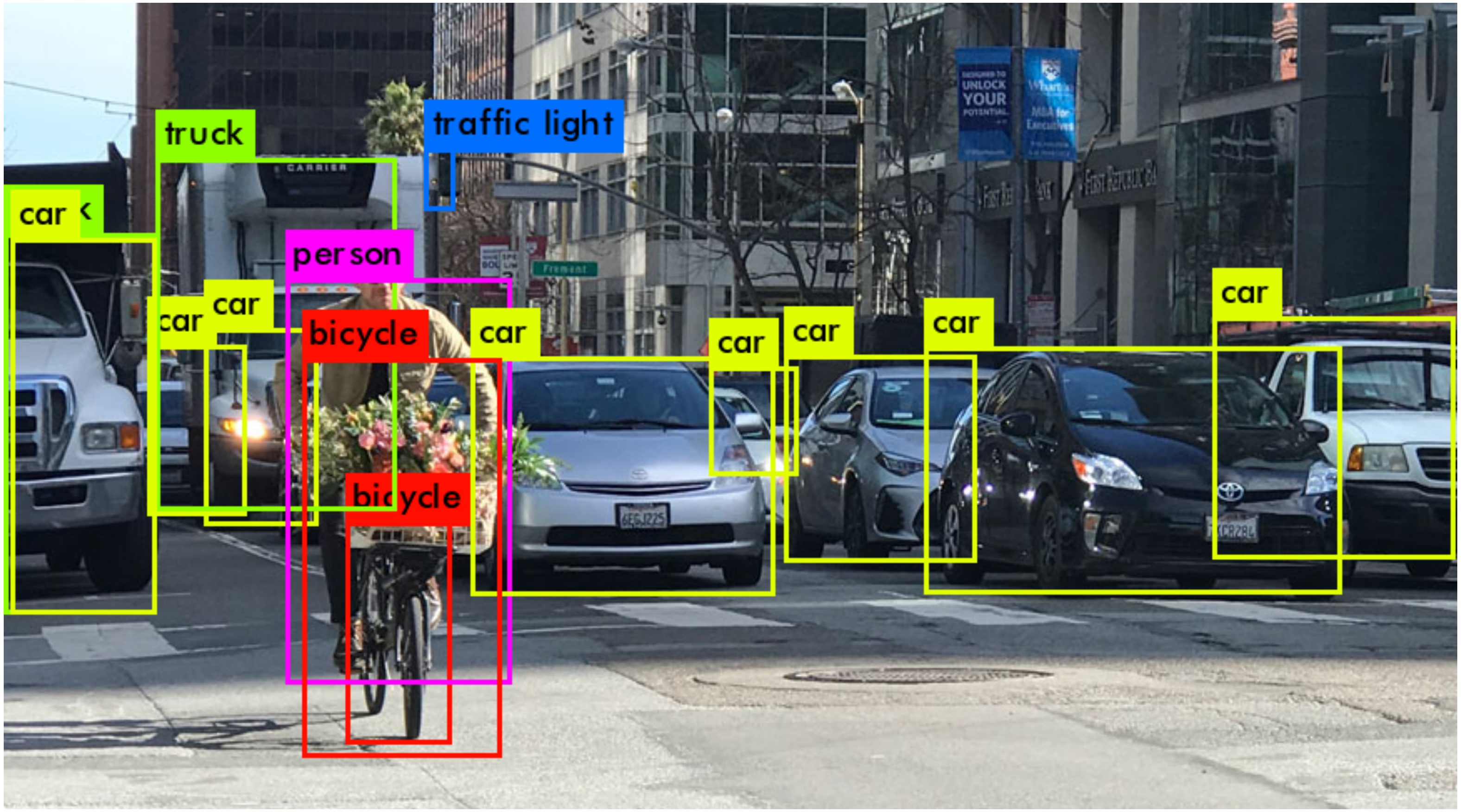
- Teams of 2–4 students
- Design one homework exercise (with my help, starting next week)
- Lead tutorials for this exercise

Interested? Send me an email: ecker@cs.uni-goettingen.de

The AI revolution

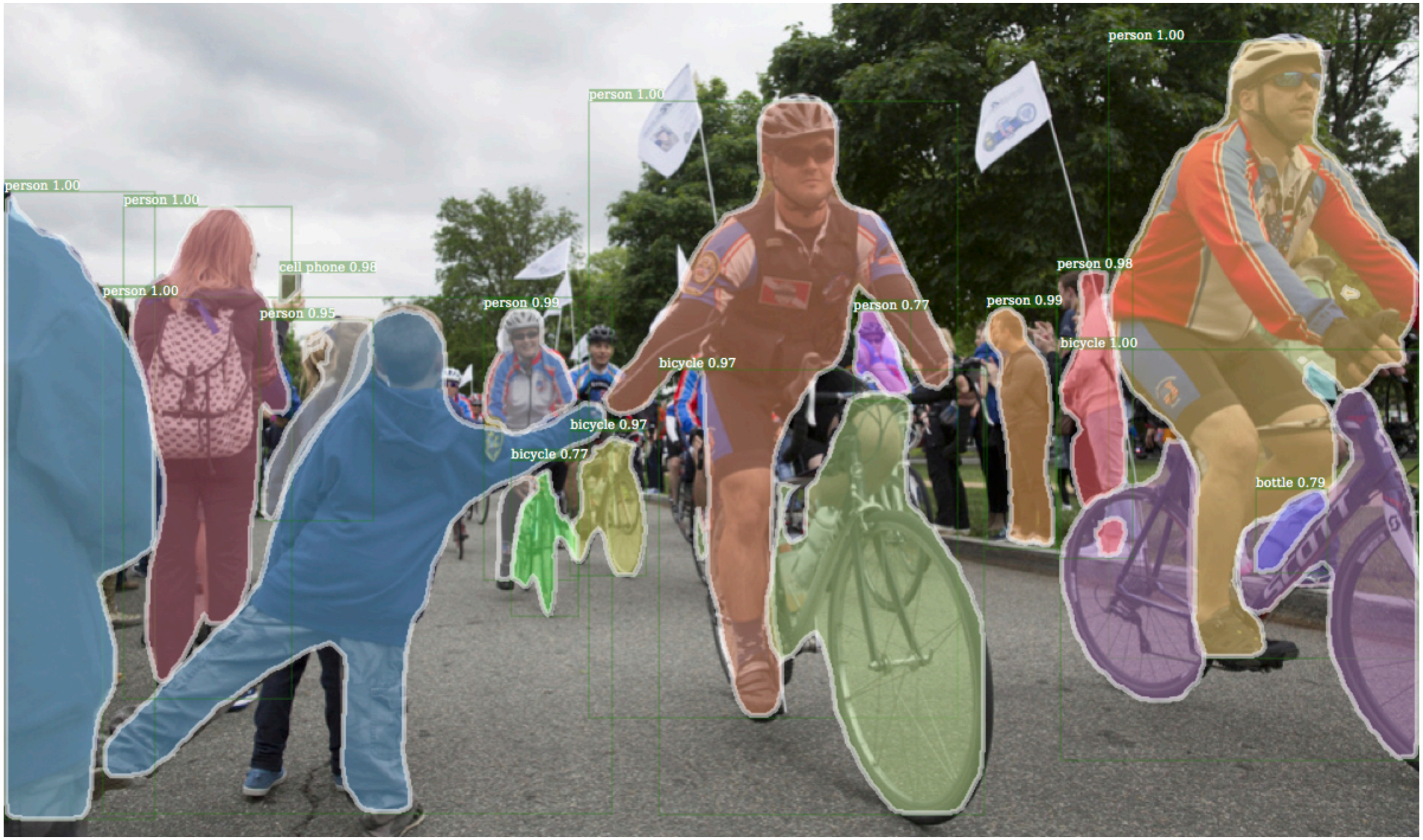
2012 – PRESENT

Object detection



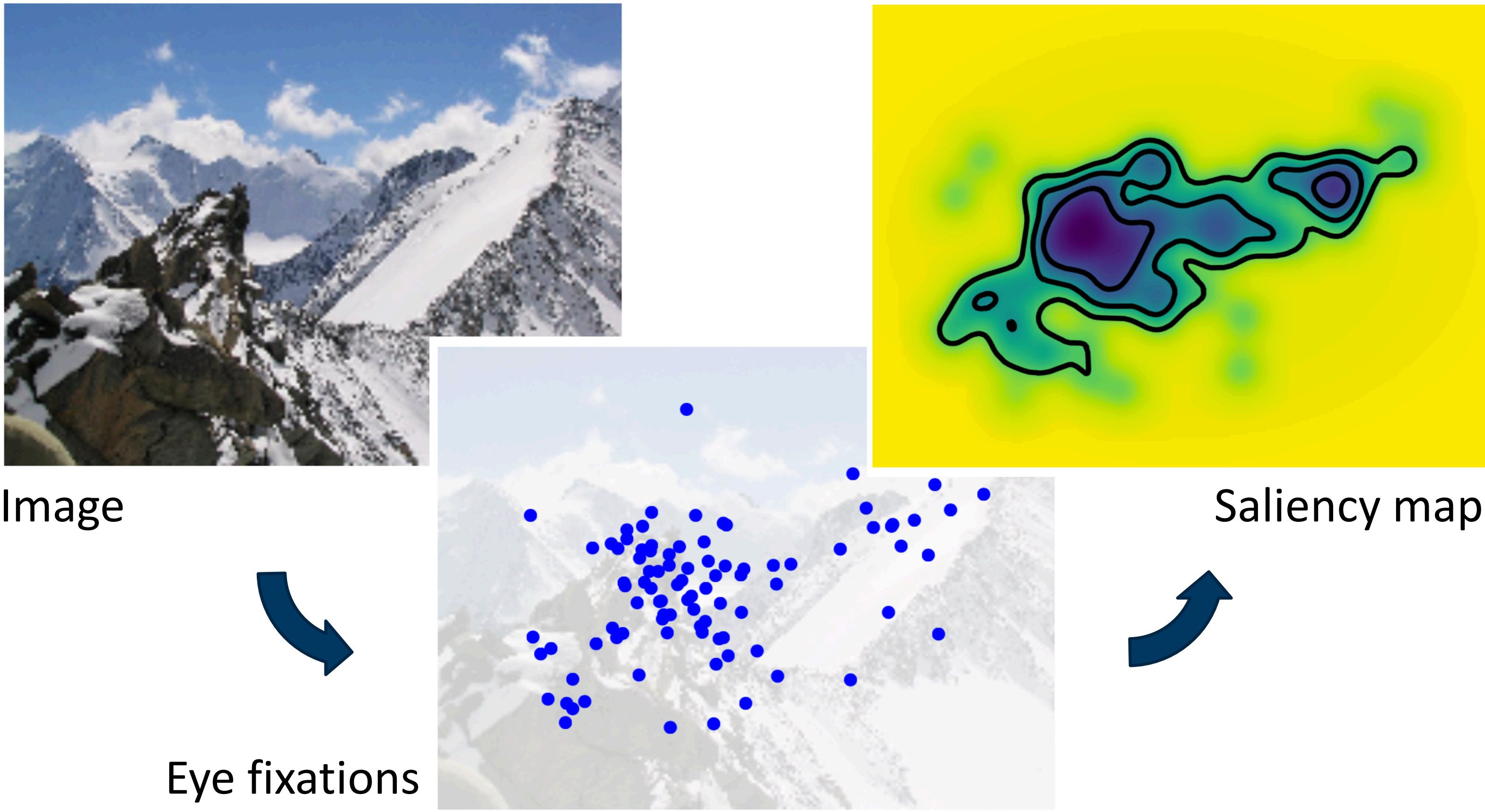
https://medium.com/@jonathan_hui/real-time-object-detection-with-yolo-yolov2-28b1b93e2088

Object segmentation



https://github.com/facebookresearch/Detectron

Predicting where people look



Kümmerer, Theis, Bethge, 2015, ICLR Workshop

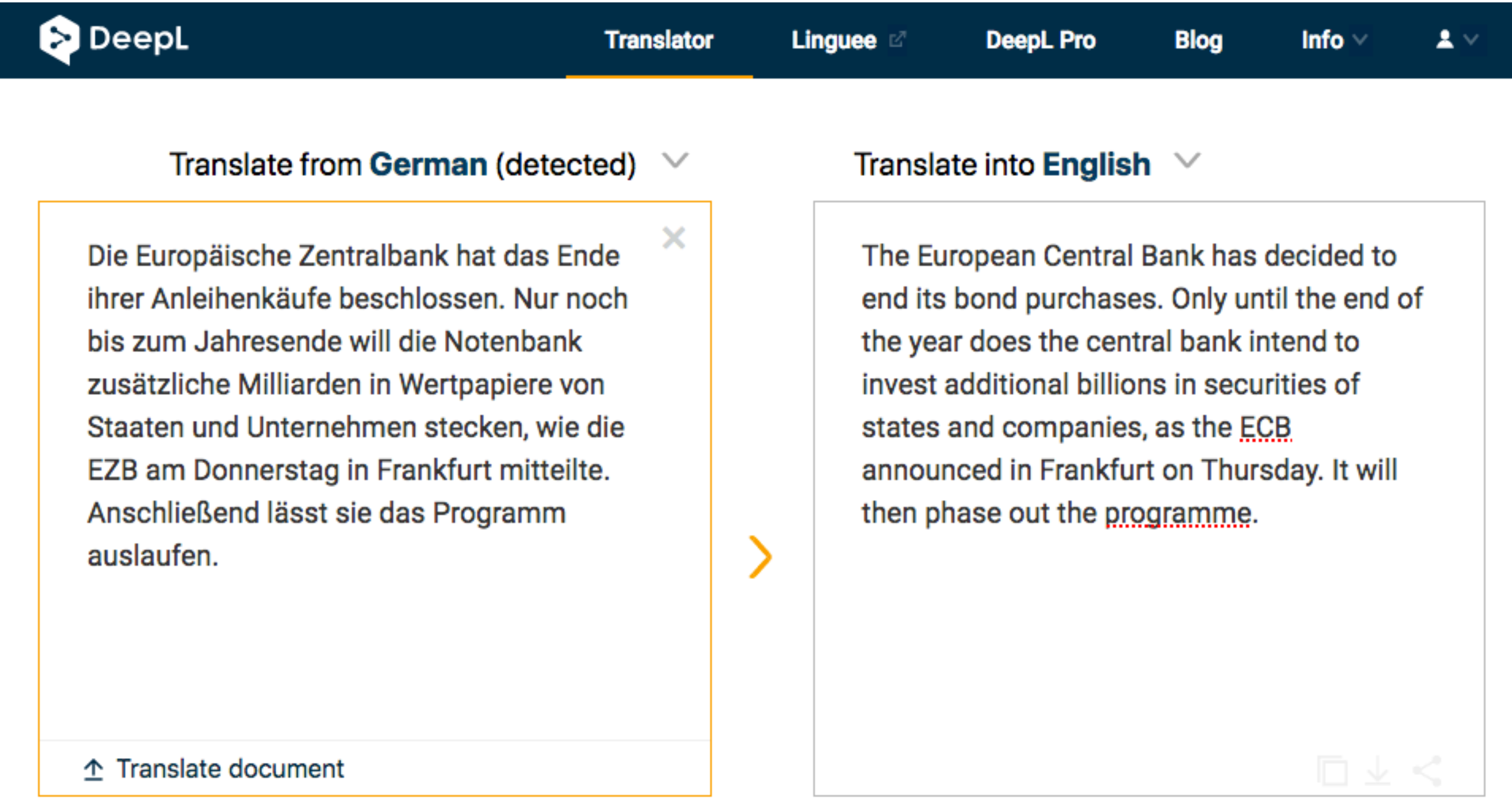
Smart cropping: Twitter

Before: center crop

Now: neural network to predict saliency

Lucas Theis @ Twitter
Kümmerer, Theis, Bethge, 2015, ICLR Workshop

Machine translation



https://deepl.com

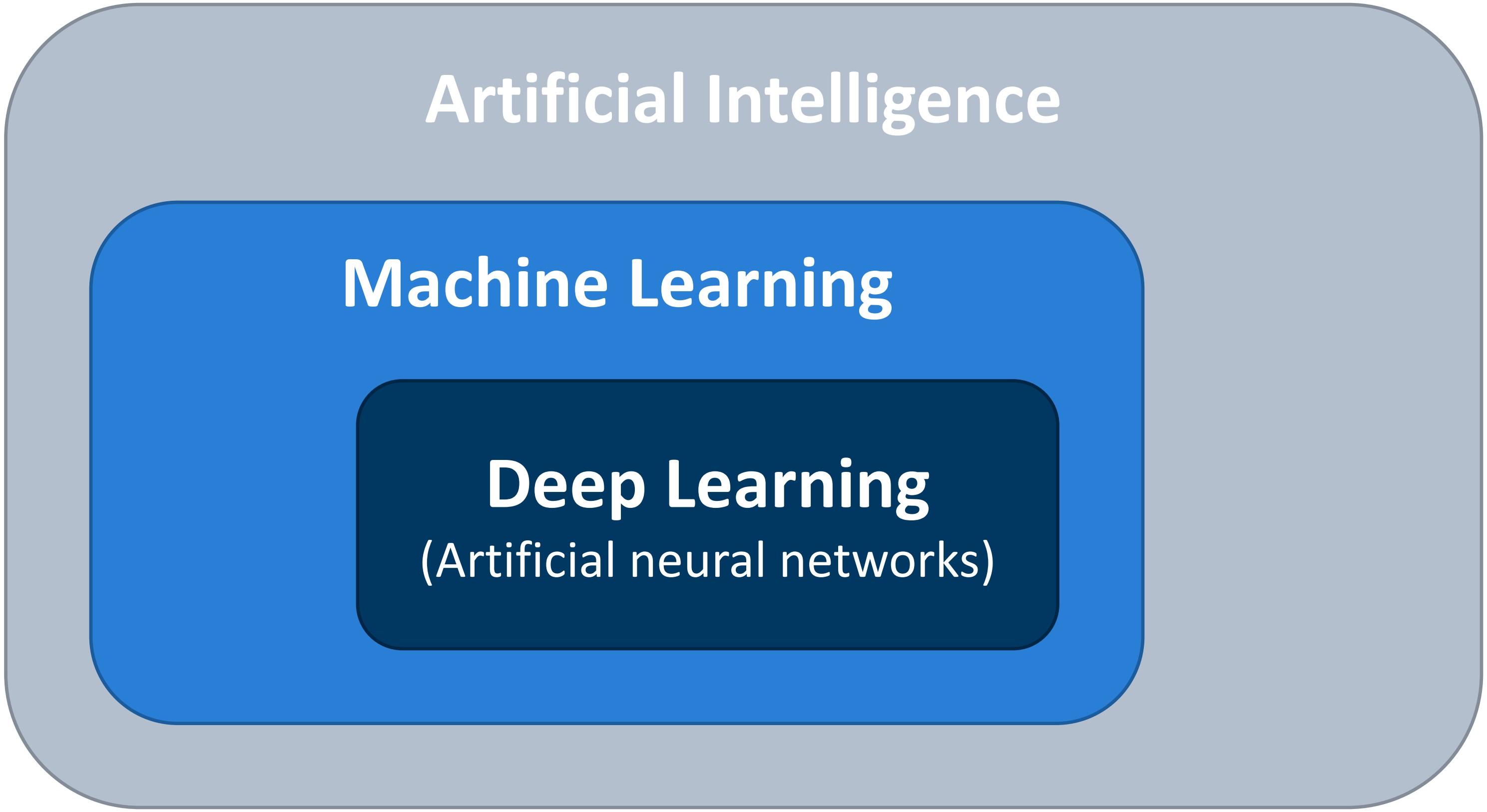
Voice recognition: Siri, Alexa et al.

Speech synthesis

Just over a year ago we presented WaveNet, a new deep neural network for generating raw audio waveforms that is capable of producing better and more realistic-sounding speech than existing techniques. At that time, the model was a research prototype and was too computationally intensive to work in consumer products.

https://deepmind.com/blog/wavenet-launches-google-assistant/

What is Deep Learning?



What is machine learning?

THE SCIENCE OF GETTING COMPUTERS TO ACT WITHOUT BEING EXPLICITLY PROGRAMMED

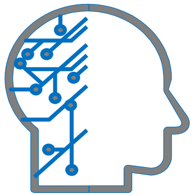
Logical Reasoning vs. Machine Learning

Logical Reasoning

A → X
B → Y
C → Z

Decisions are based on pre-defined rules

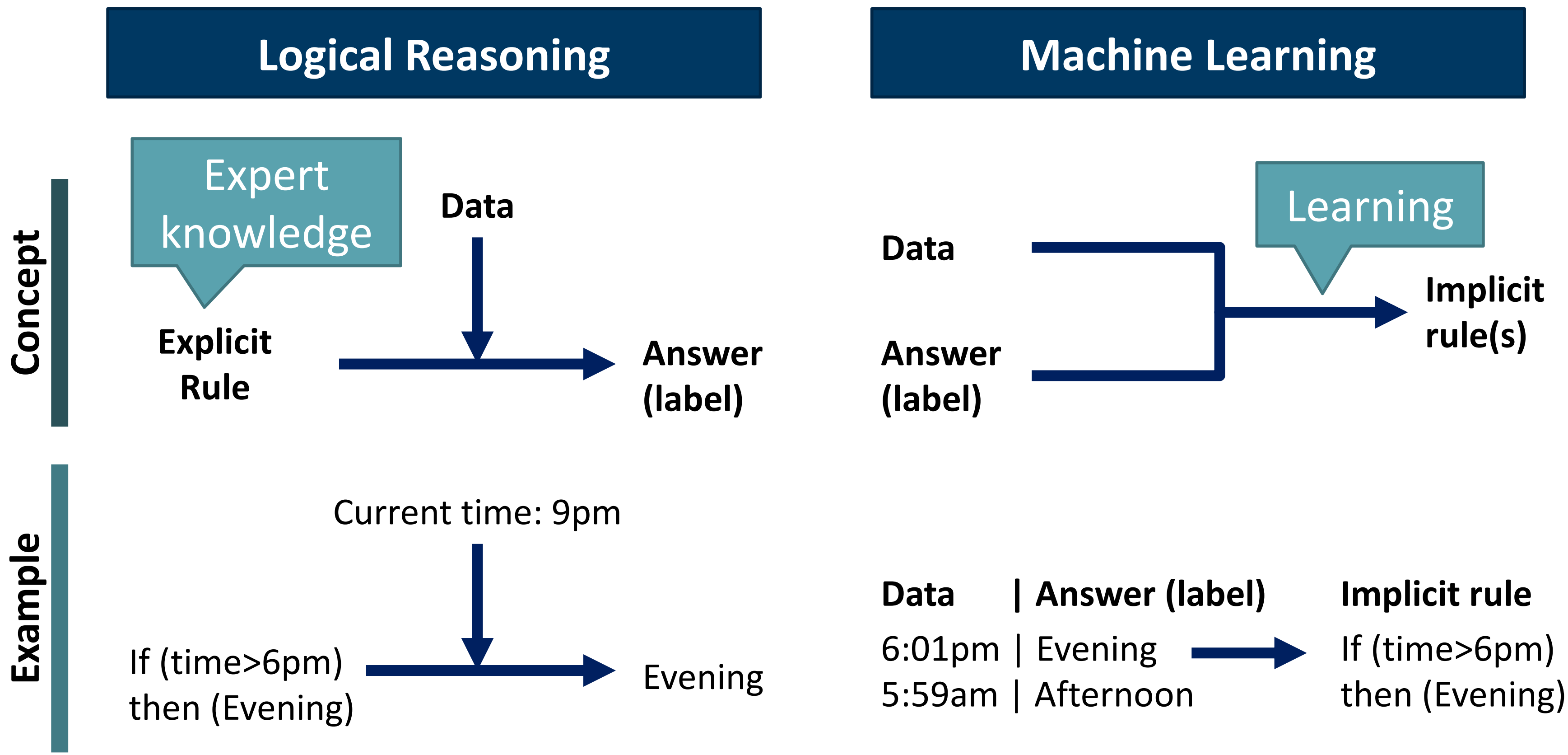
Machine Learning



Decisions are based on a learned mapping of inputs to desired outputs

(without explicitly defining rules)

Logical Reasoning vs. Machine Learning

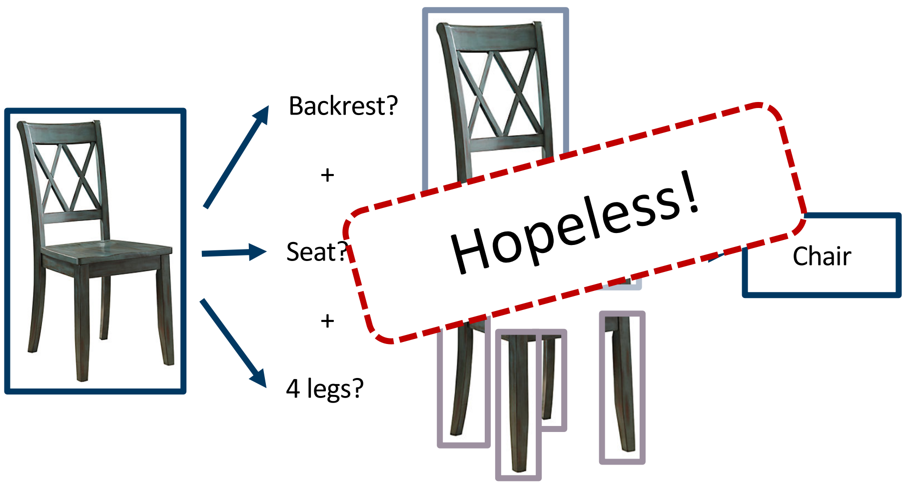


Example: object classification via classic, rule-based AI

Input

Predefined rules

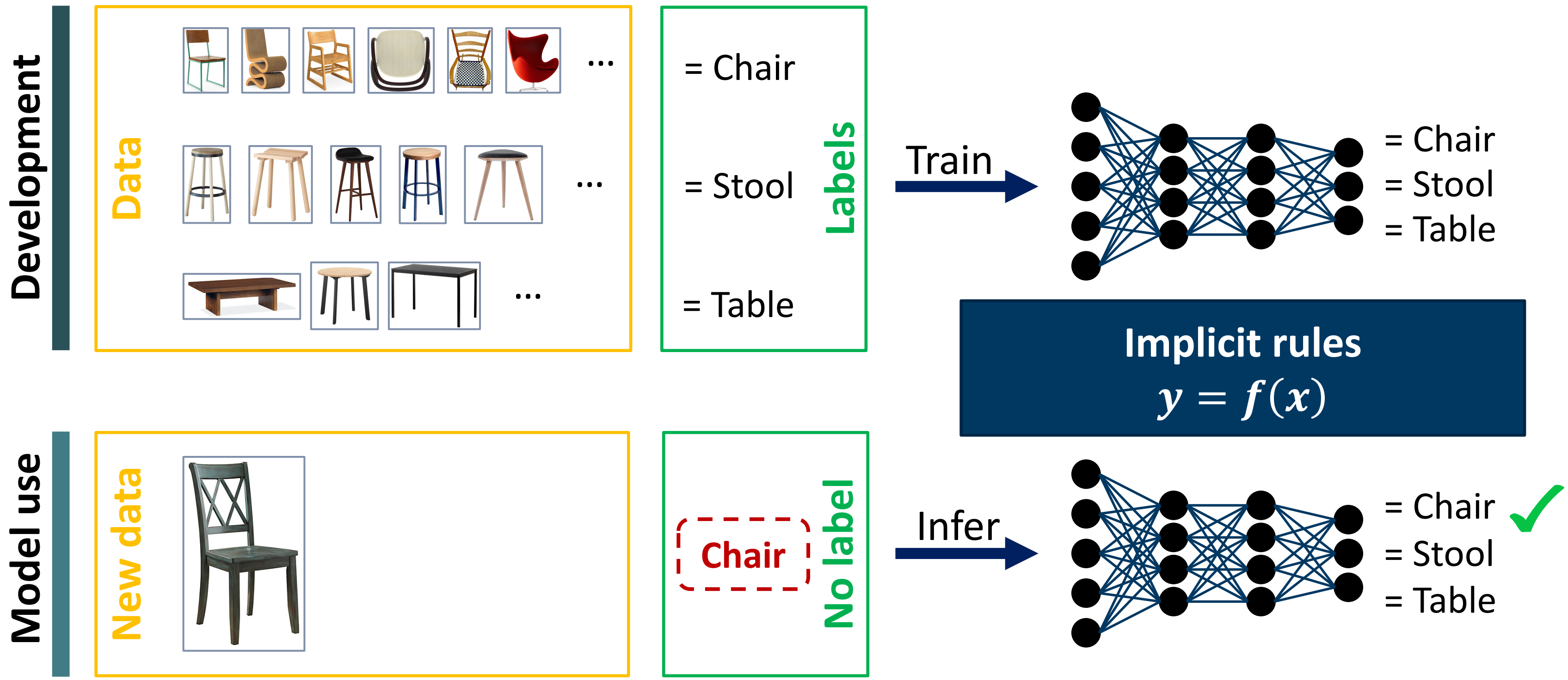
Classification



Reality keeps breaking rules

Viewpoint, occlusion etc.

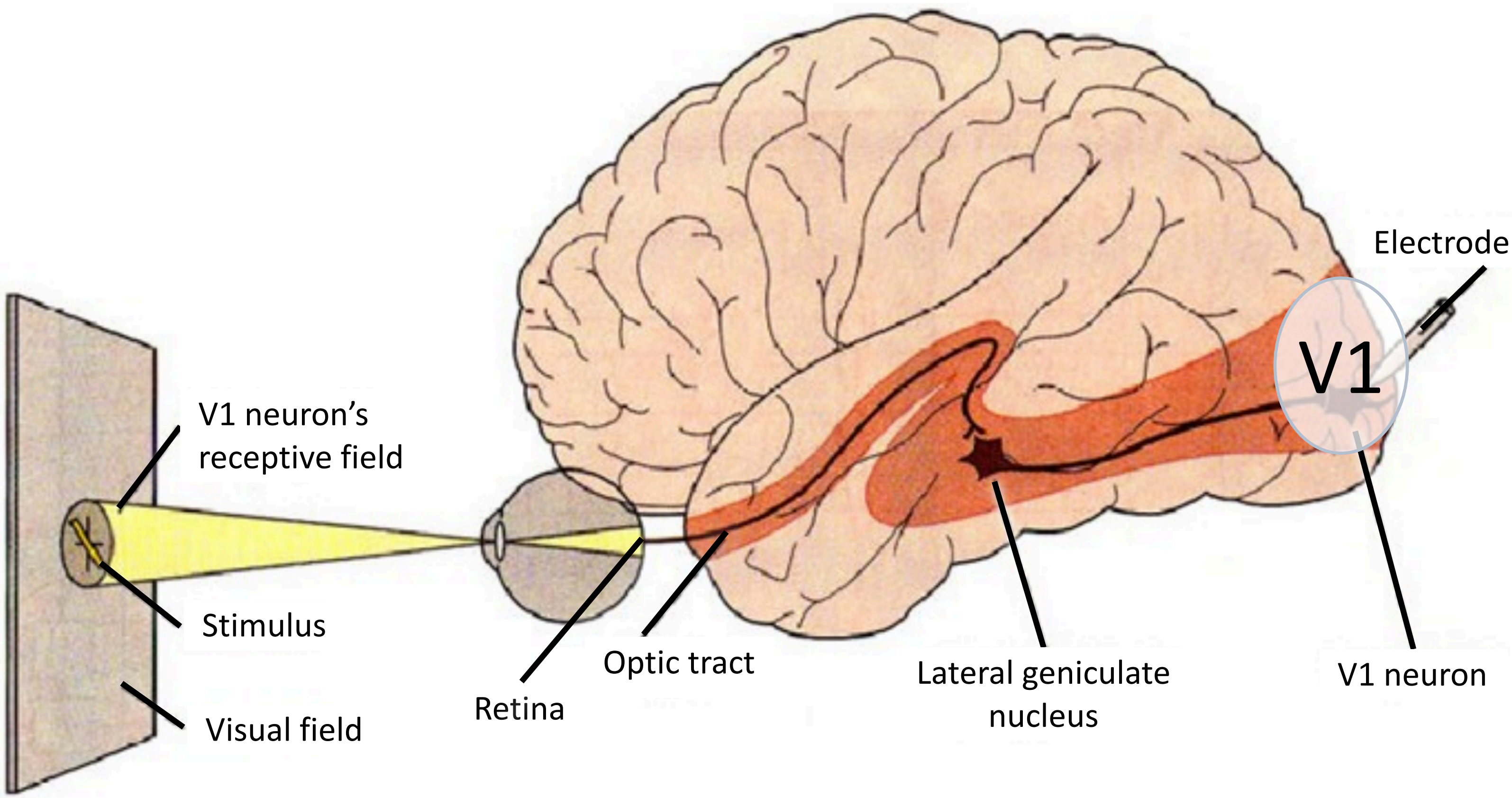
Machine Learning: Learn implicit rules from examples



How do neural networks work?

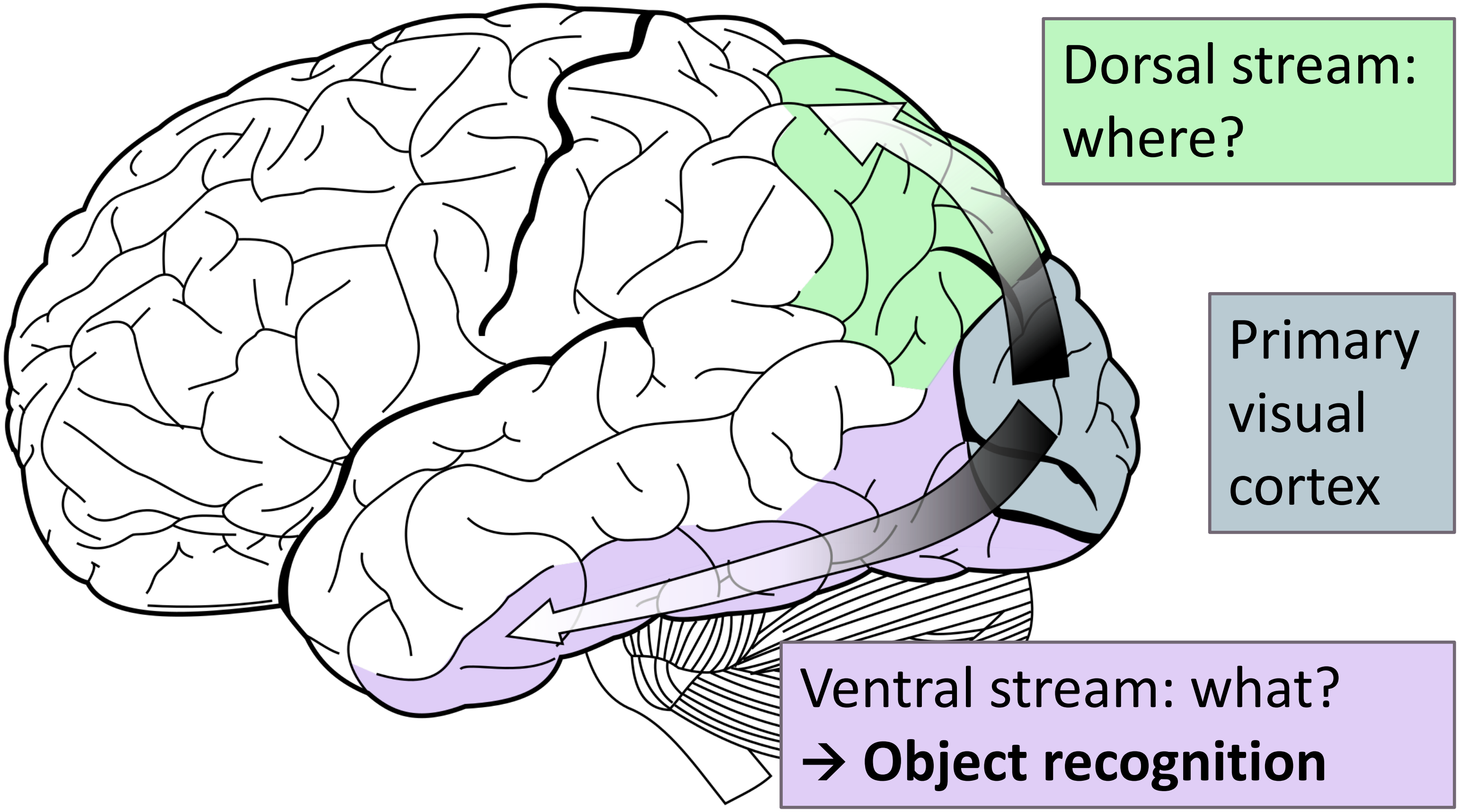
FROM THE BRAIN TO ARTIFICIAL NEURAL NETWORKS

From the eye to the brain



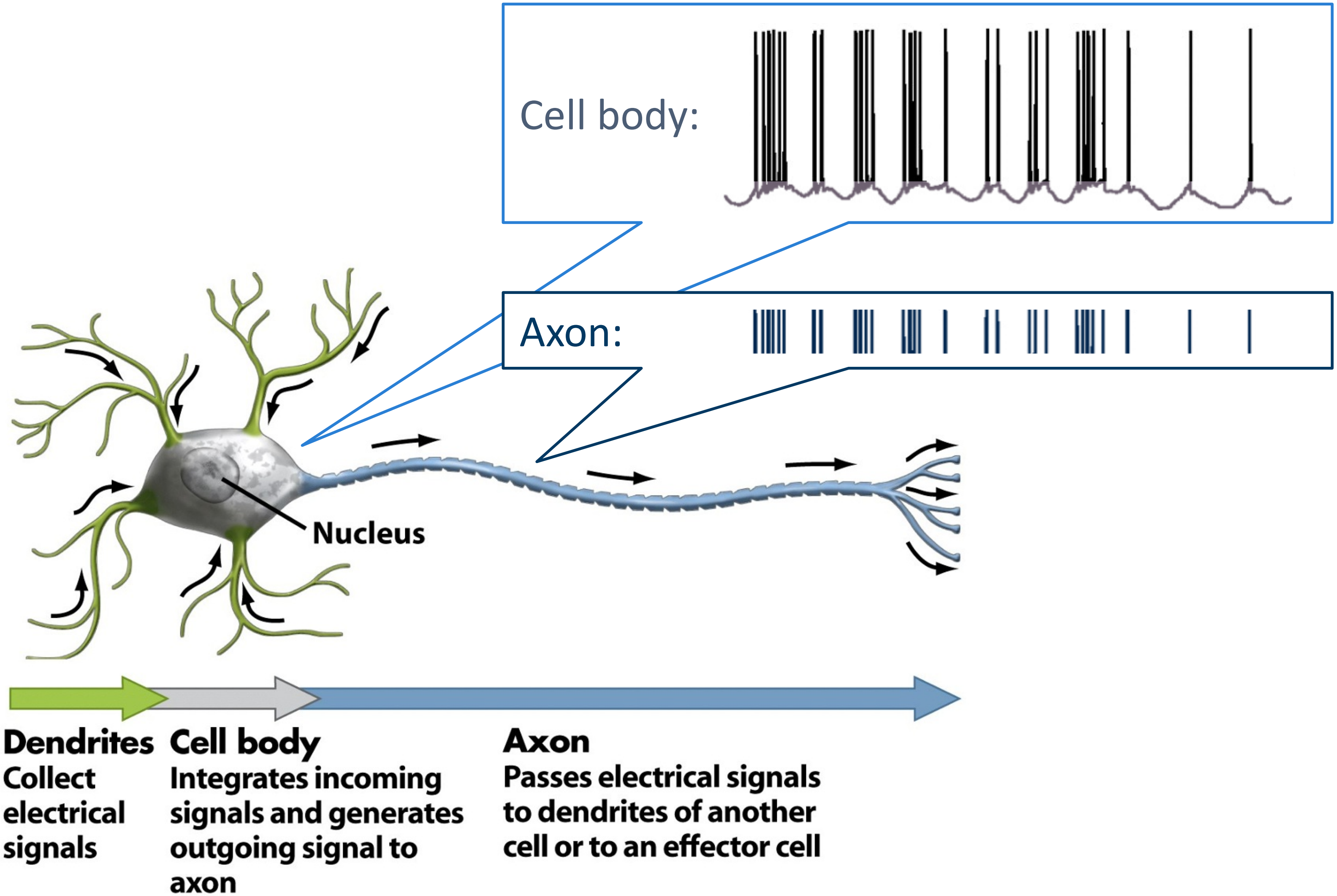
V1 = Primary visual cortex

The visual system: two streams



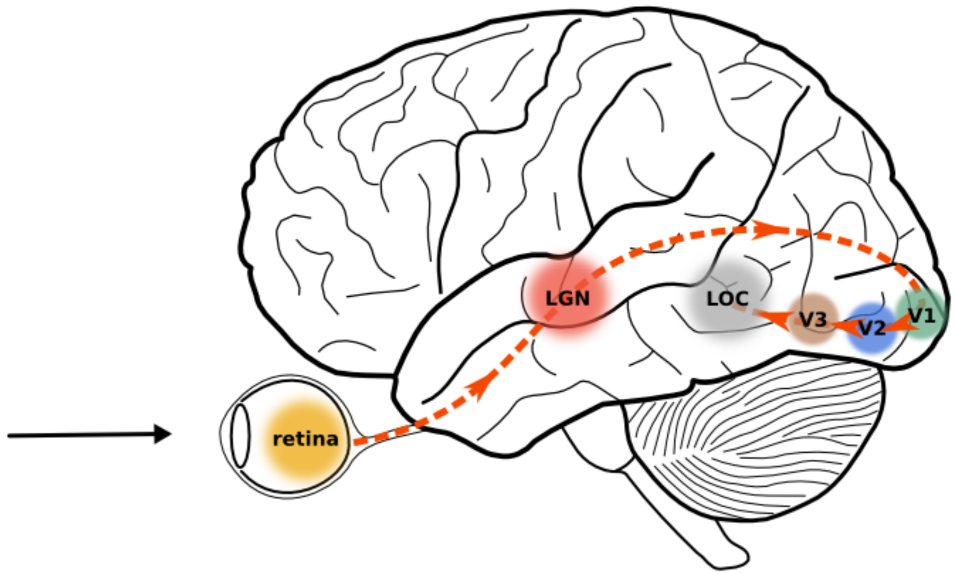
https://en.wikipedia.org/wiki/Two-streams_hypothesis

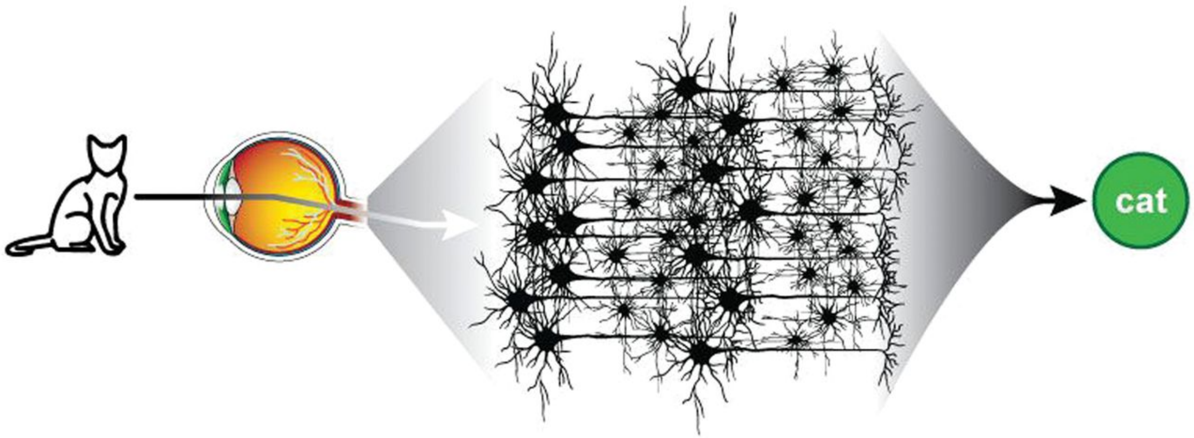
How do neurons communicate?



https://www.thinglink.com/scene/762090261831483394

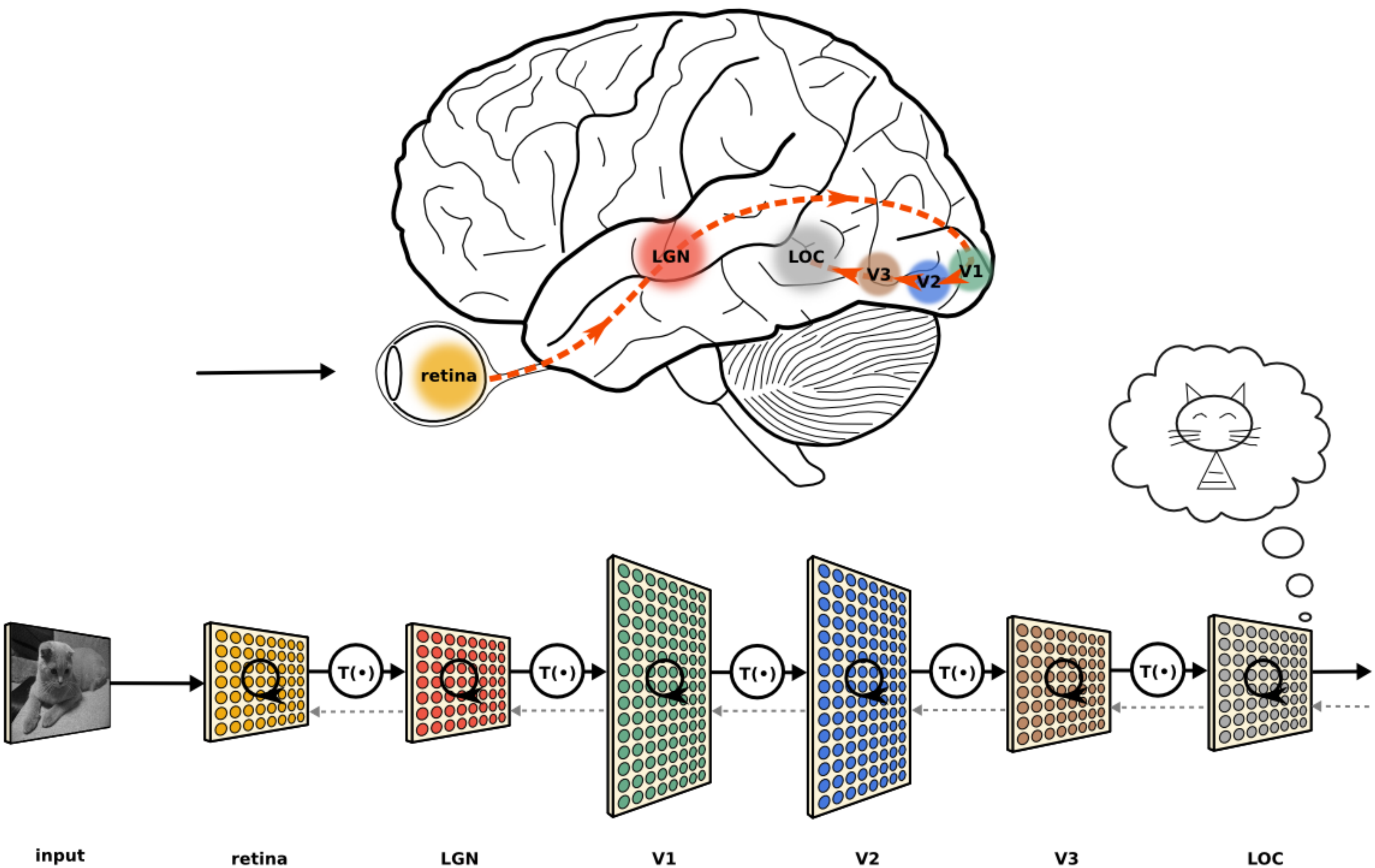
Hierarchical data processing





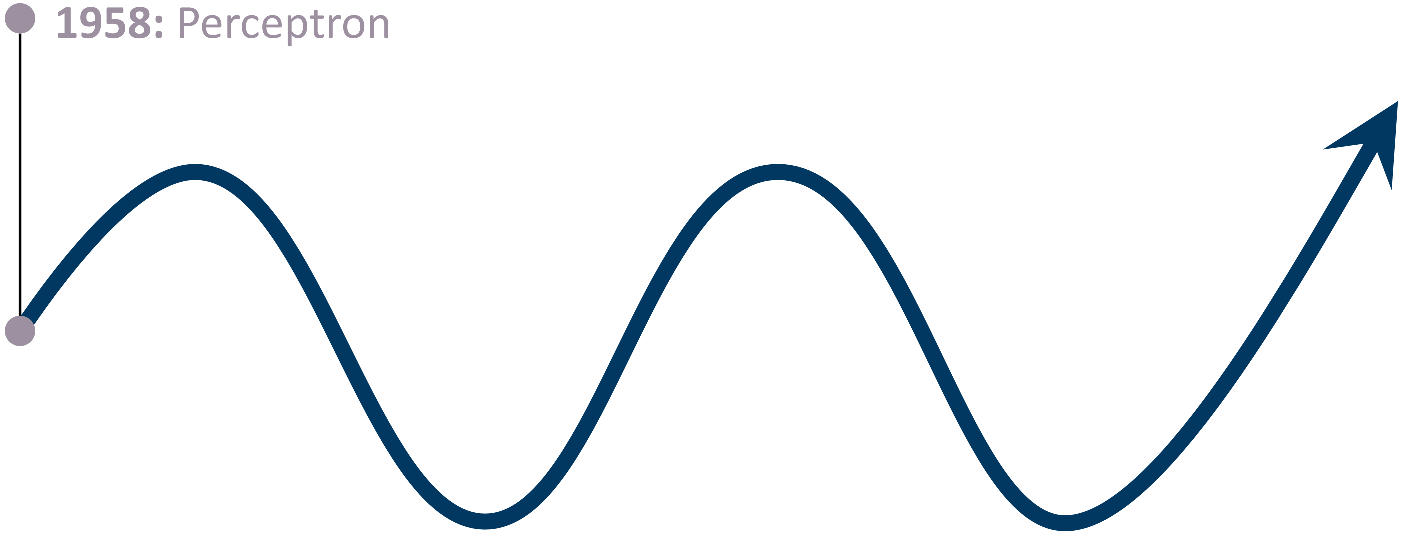
https://neuwritesd.org/2015/10/22/deep-neural-networks-help-us-read-your-mind/
G. Zaharchuk et al. AJNR Am J Neuroradiol doi:10.3174/ajnr.A5543

Hierarchical data processing



https://neuwritesd.org/2015/10/22/deep-neural-networks-help-us-read-your-mind/

Neural networks: summers & winters



The Perceptron (Rosenblatt 1958)

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.

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.

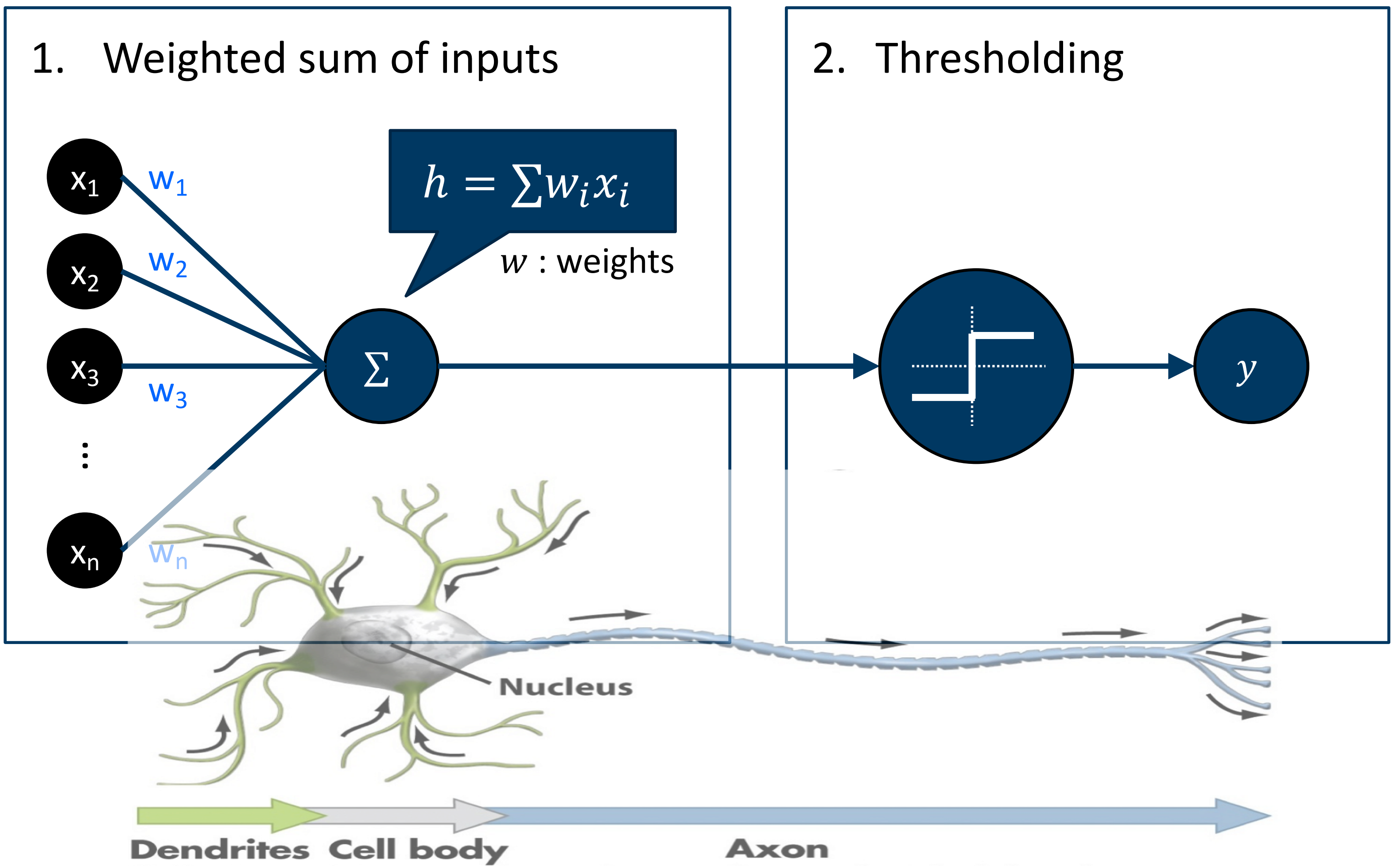
Dr. Rosenblatt said he could explain why the machine learned only in highly technical terms. But he said the computer had undergone a "self-induced change in the wiring diagram."

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.

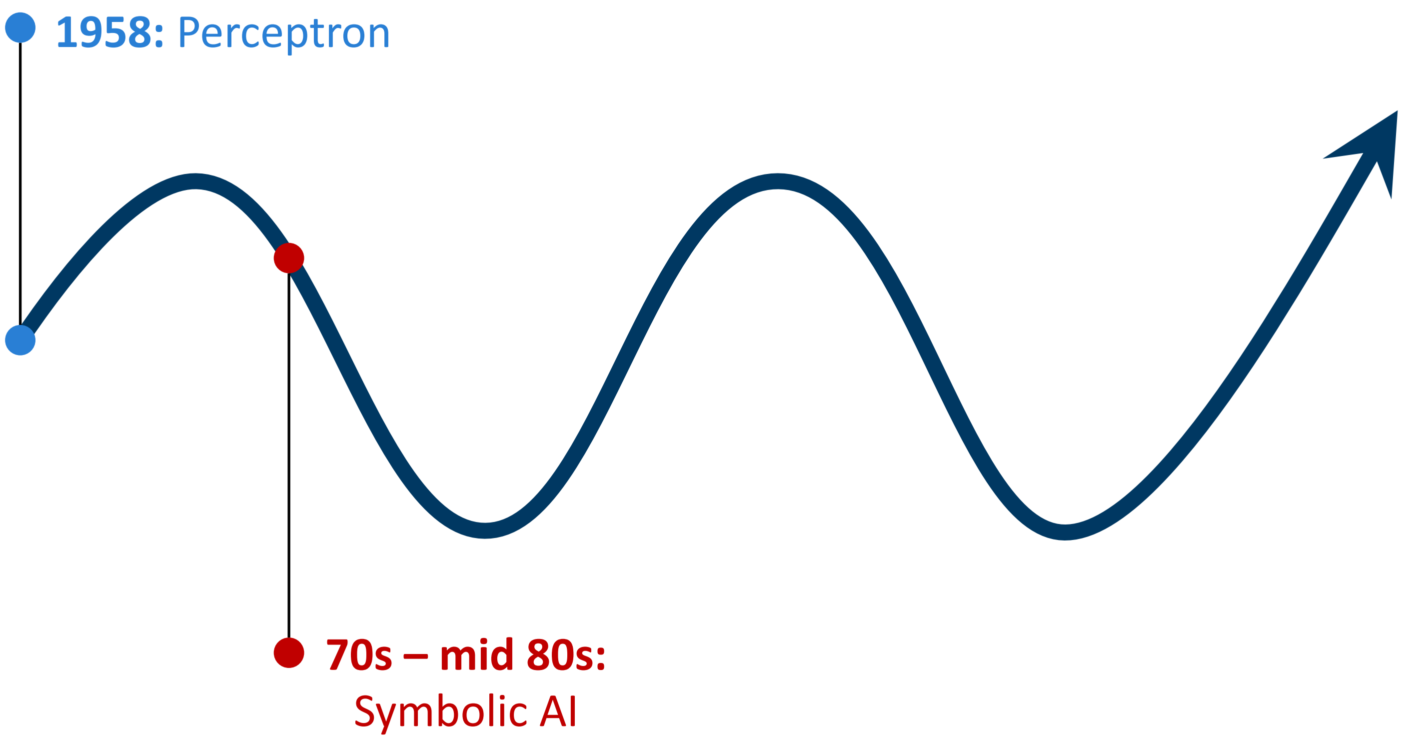
The New York Times

July 7, 1958

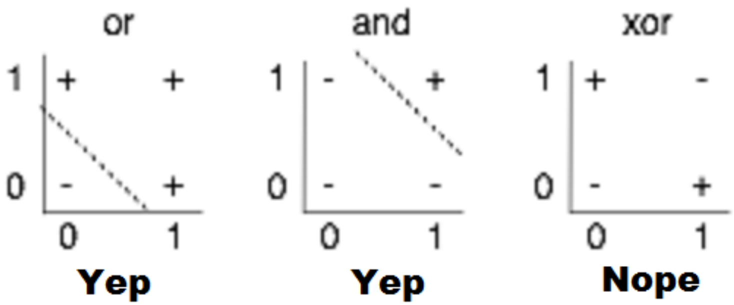
The Perceptron (Rosenblatt 1958)

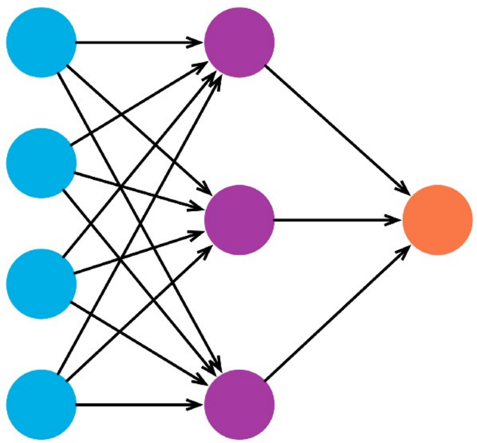


Neural networks: summers & winters



The first AI winter: 1969–1986





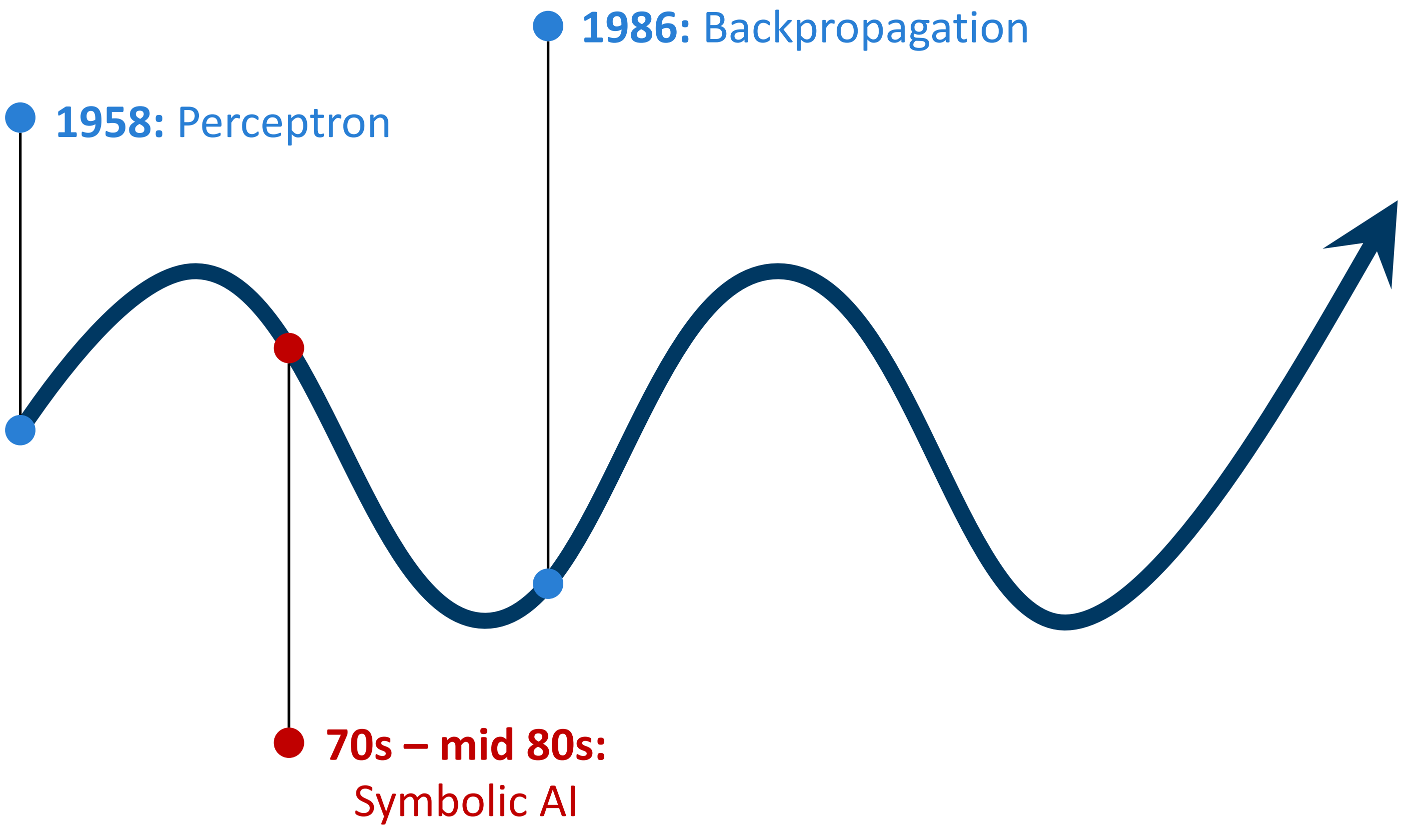
Perceptron

Cannot solve XOR

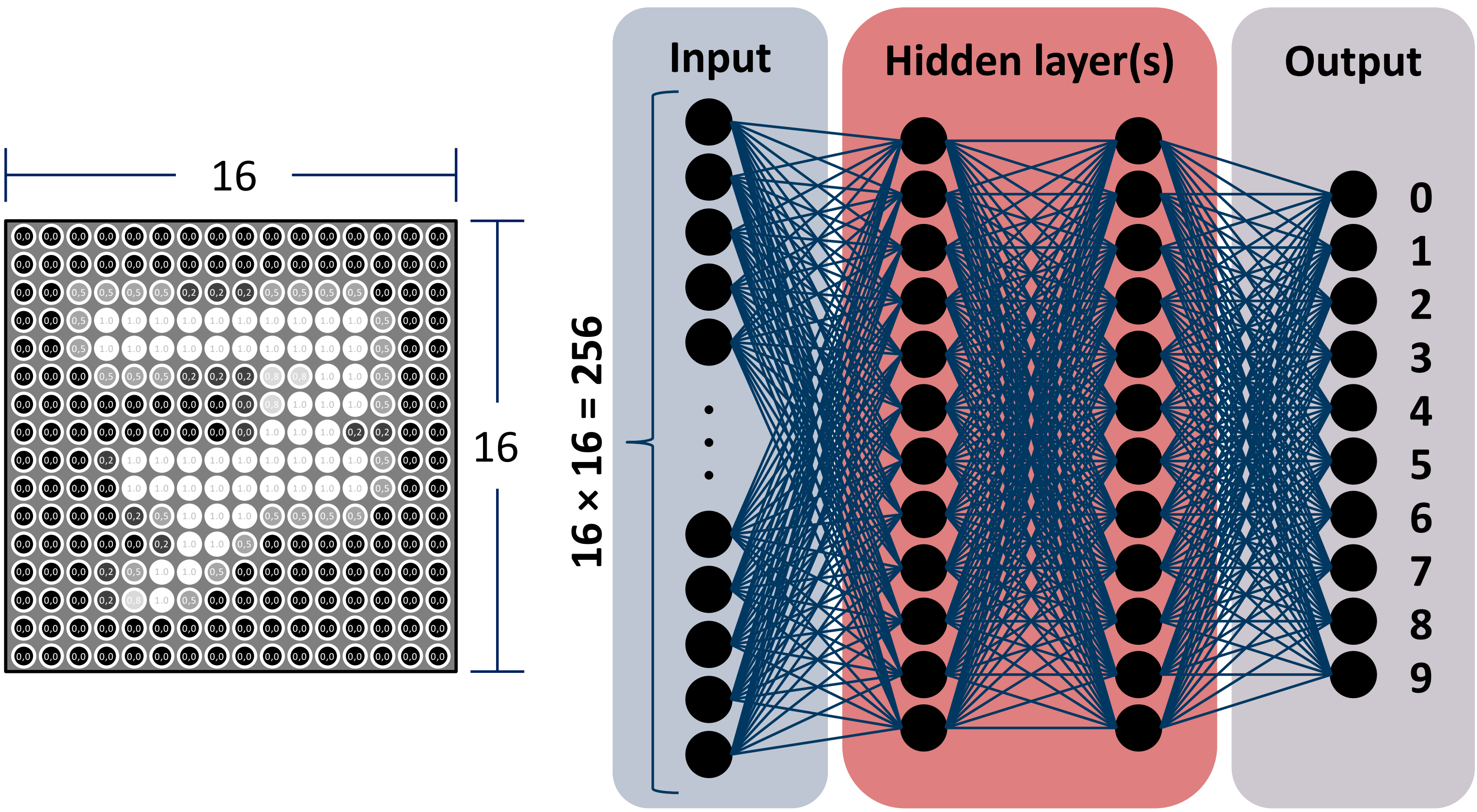
Multi-layer neural net

Problem: how to train?

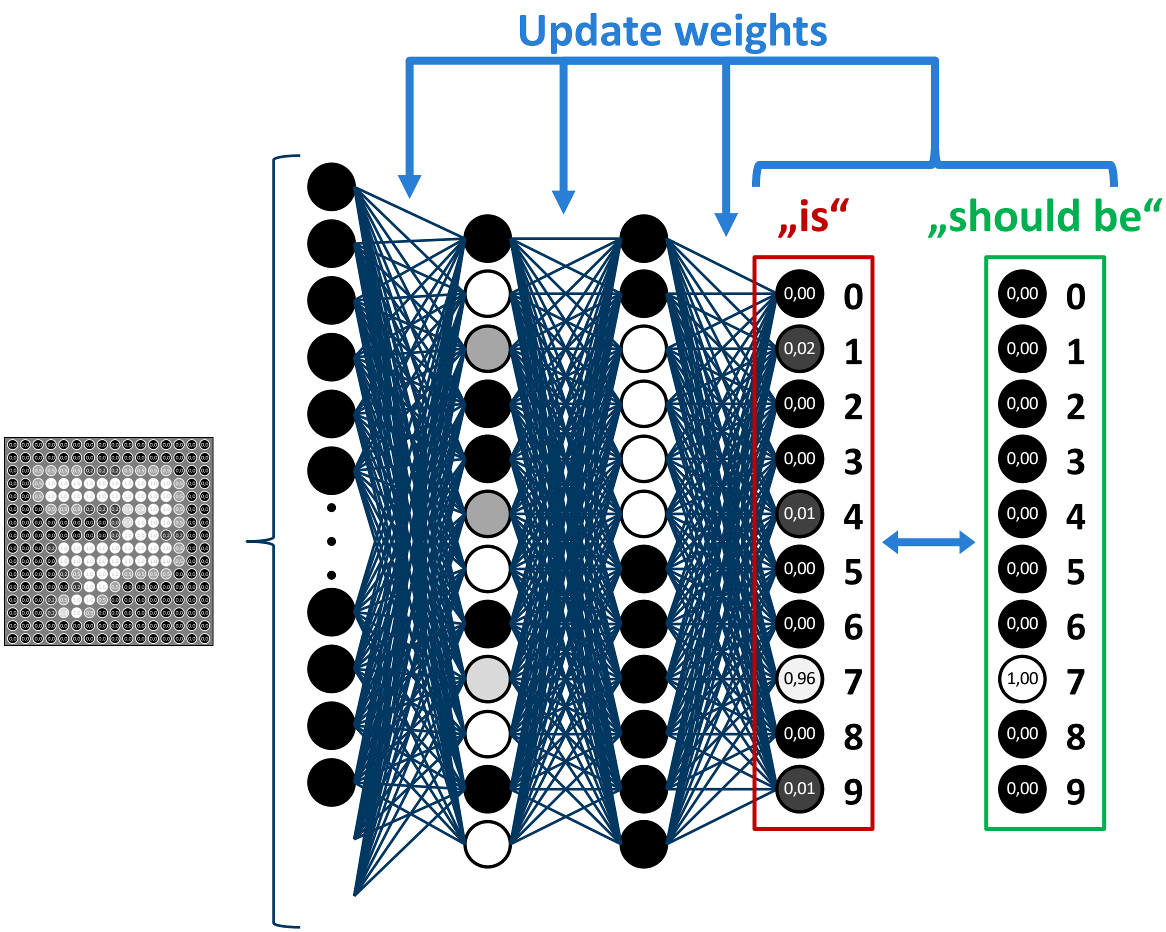
Neural networks: summers & winters



Multi-layer neural networks

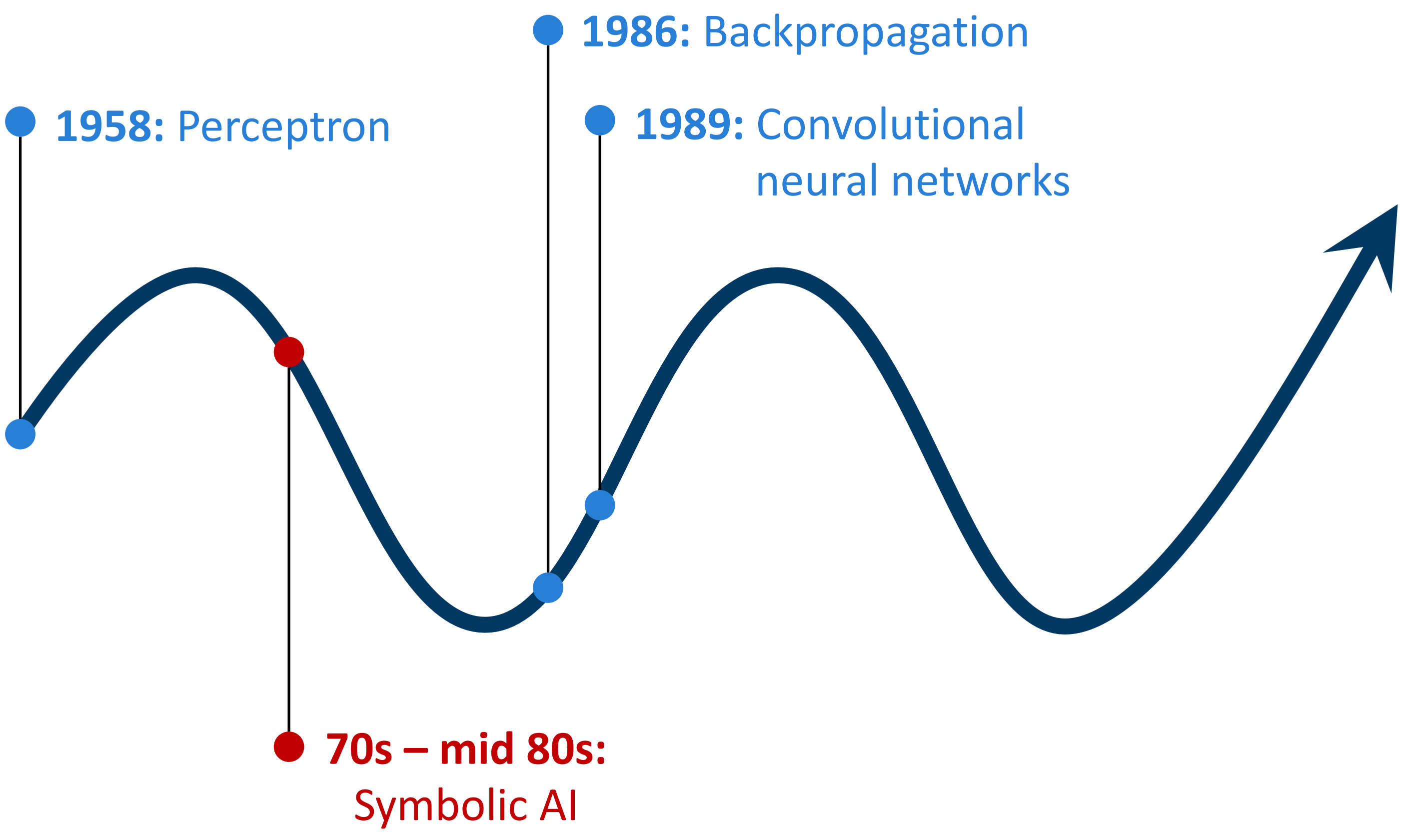


Training through backpropagation

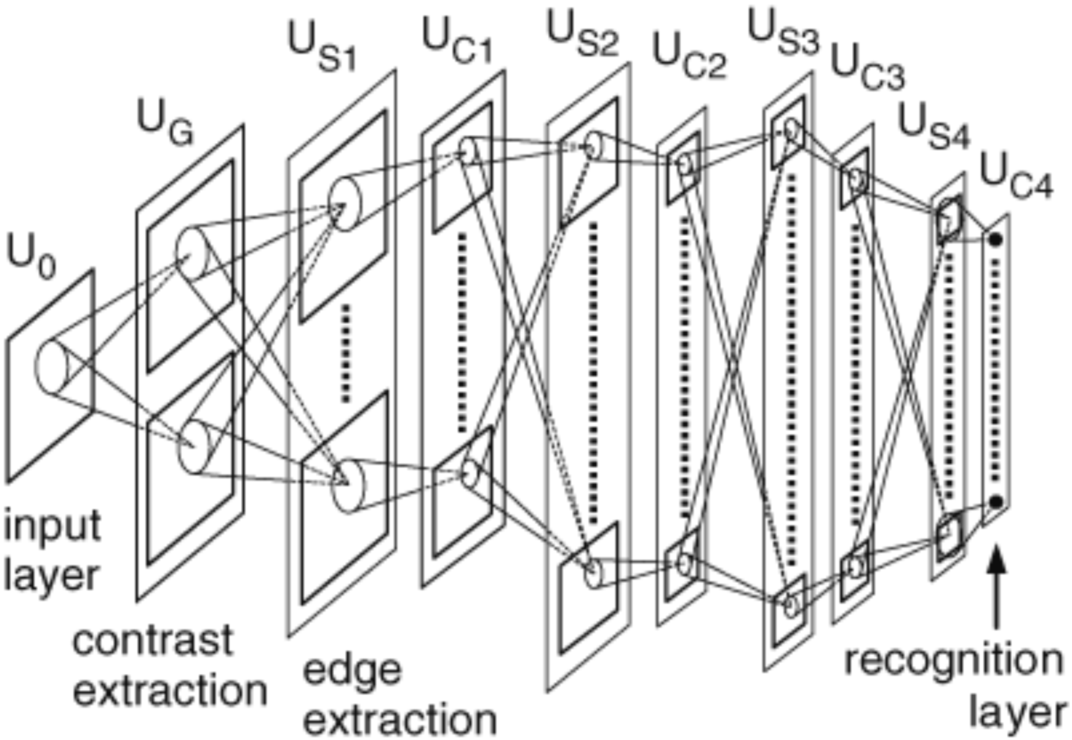


Kelley 1960, ARS Journal
Popularized by Rumelhart, Hinton, Williams 1986, Nature

Neural networks: summers & winters

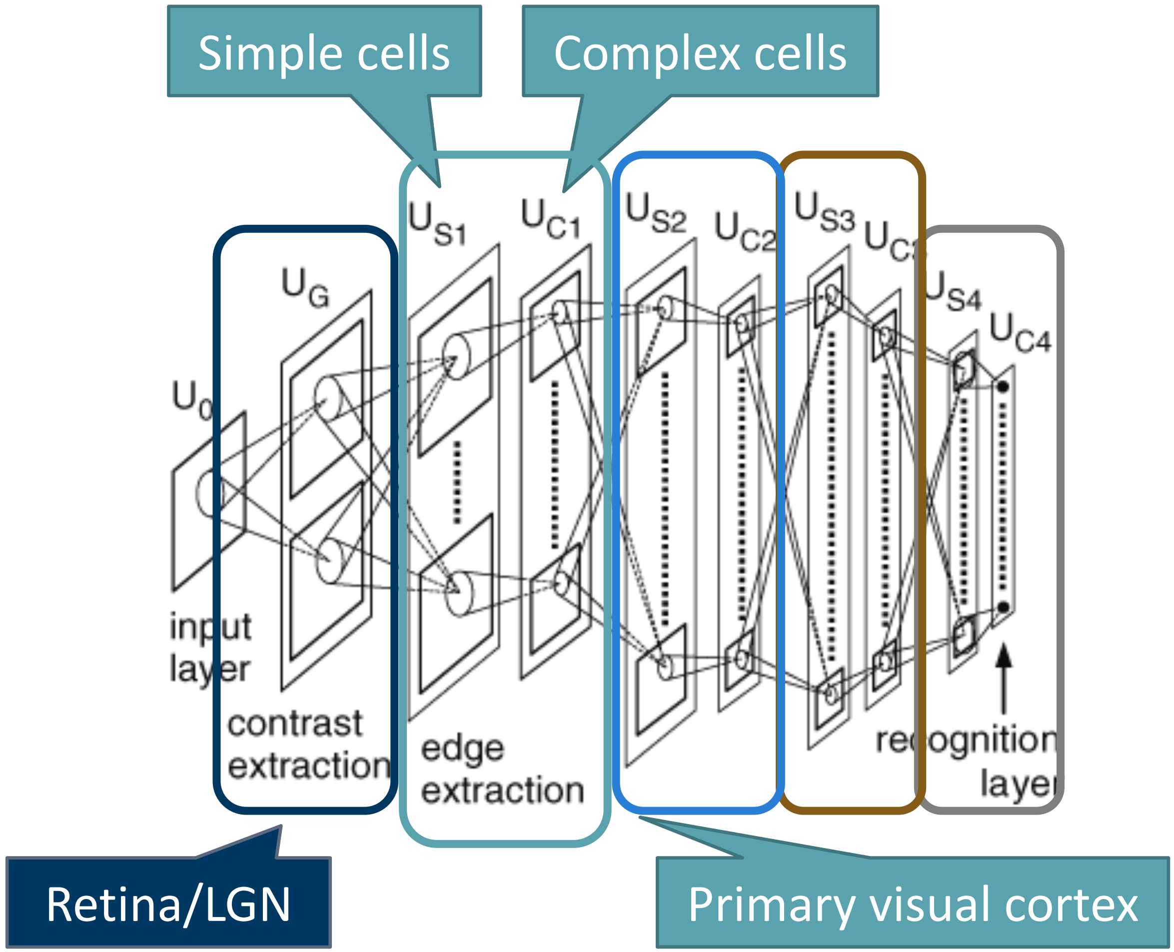


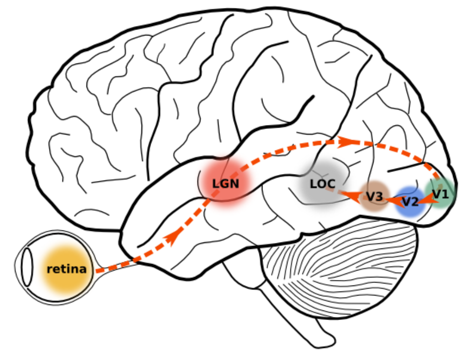
The Neocognitron (Fukushima 1980)



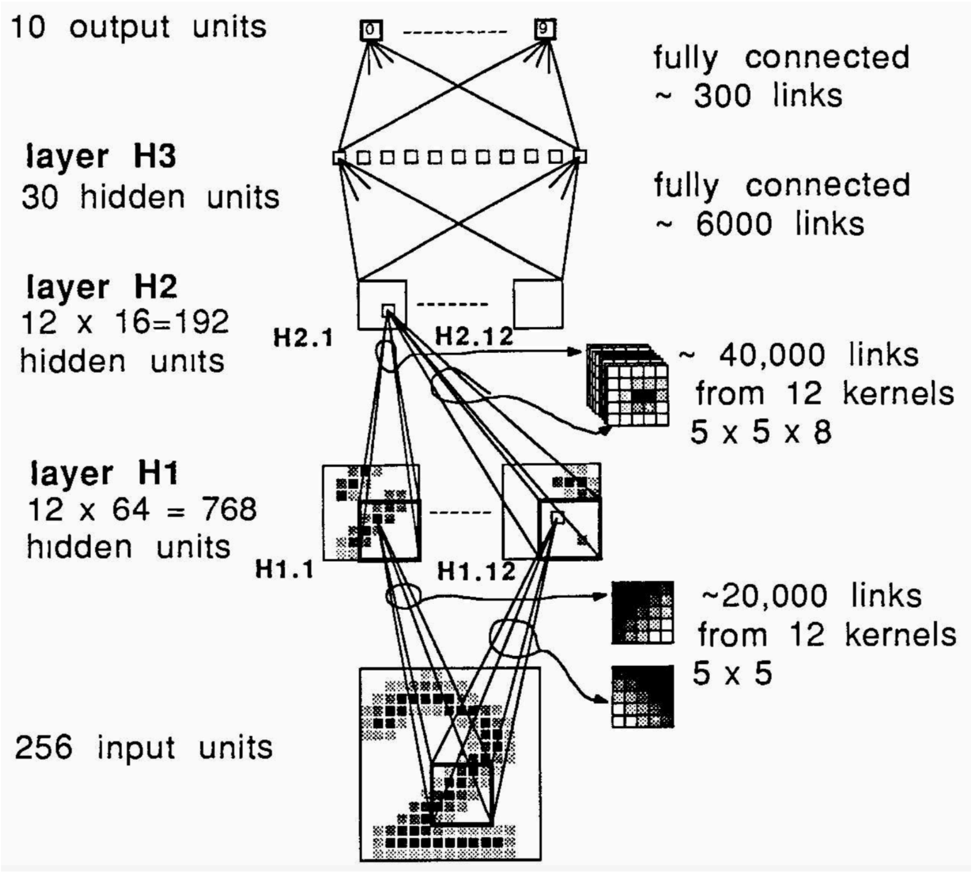
K. Fukushima (1980): "Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position", *Biological Cybernetics*.

The Neocognitron (Fukushima 1980)





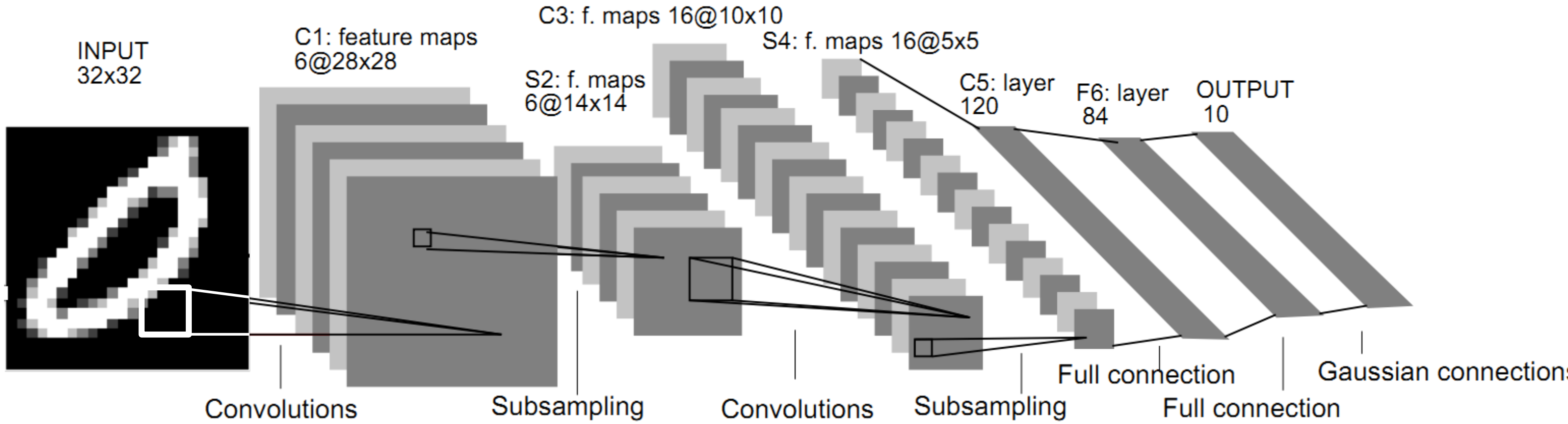
Convolutional neural networks (CNN)

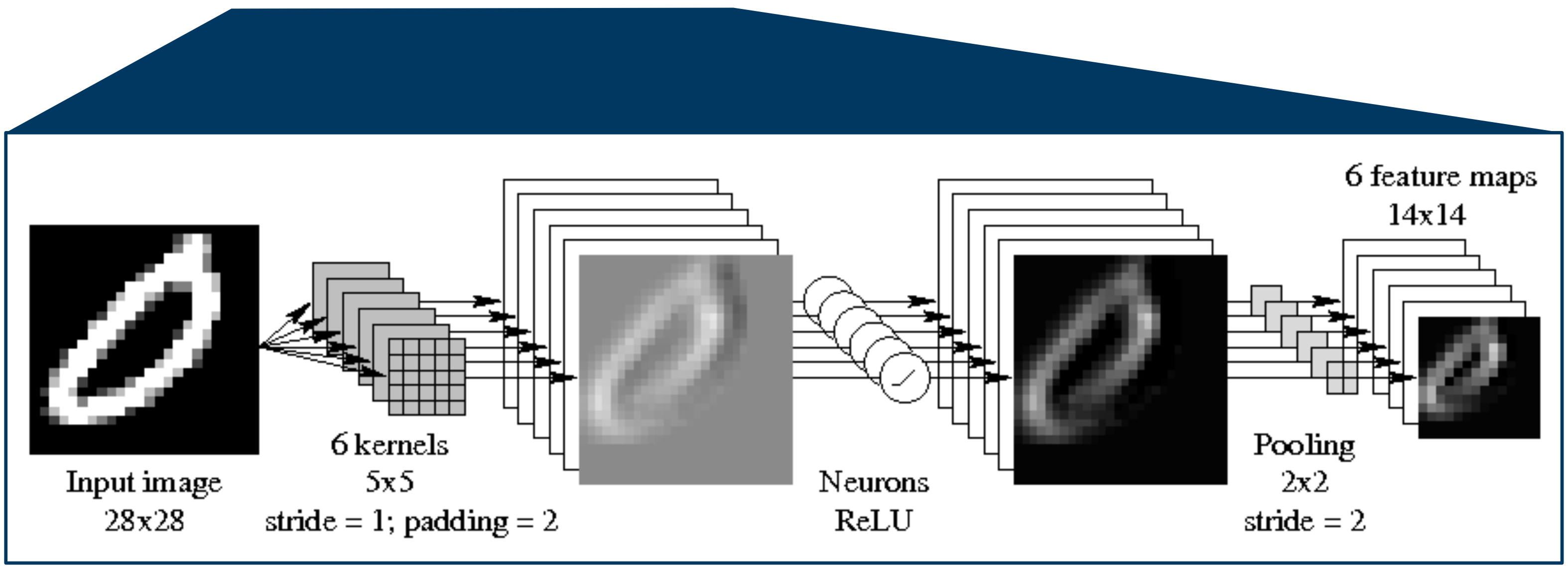


Used in production for reading checks and addresses

Denker et al. 1989, LeCun et al. 1989

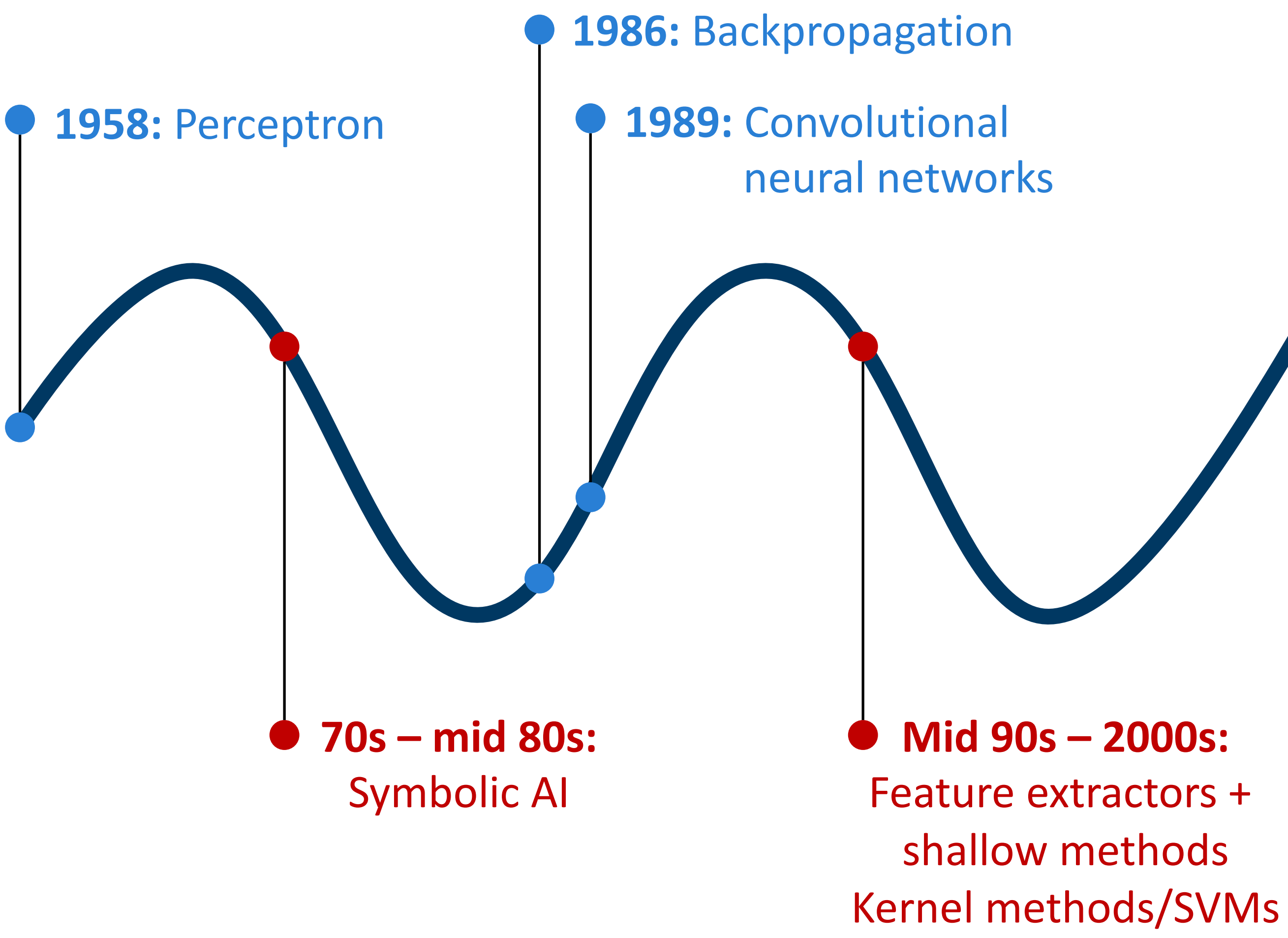
Convolutional neural networks





Adapted from LeCun, Bottou, Bengio, Haffner 1998, Proceedings of the IEEE

Another neural network winter



90ies: Too little data & compute power for neural networks to shine

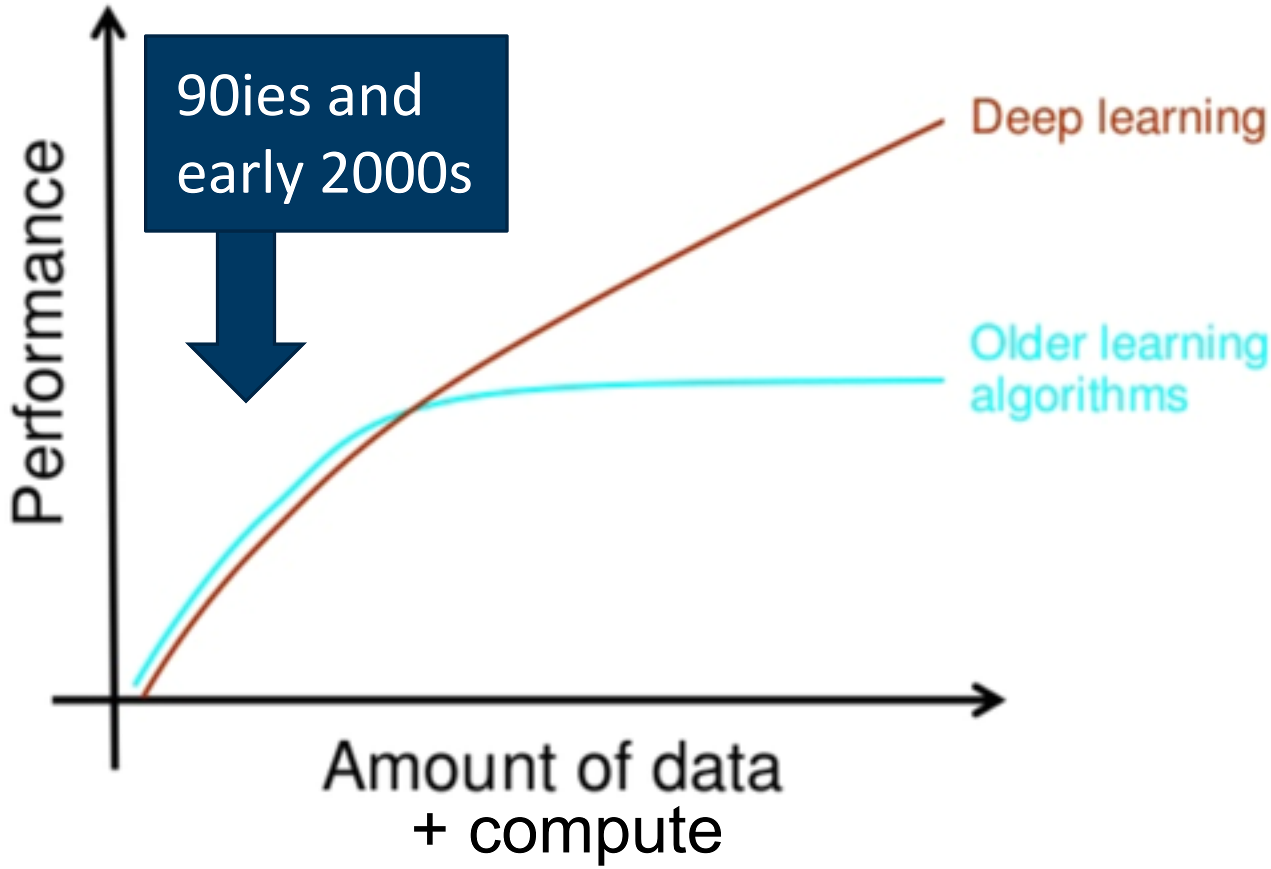
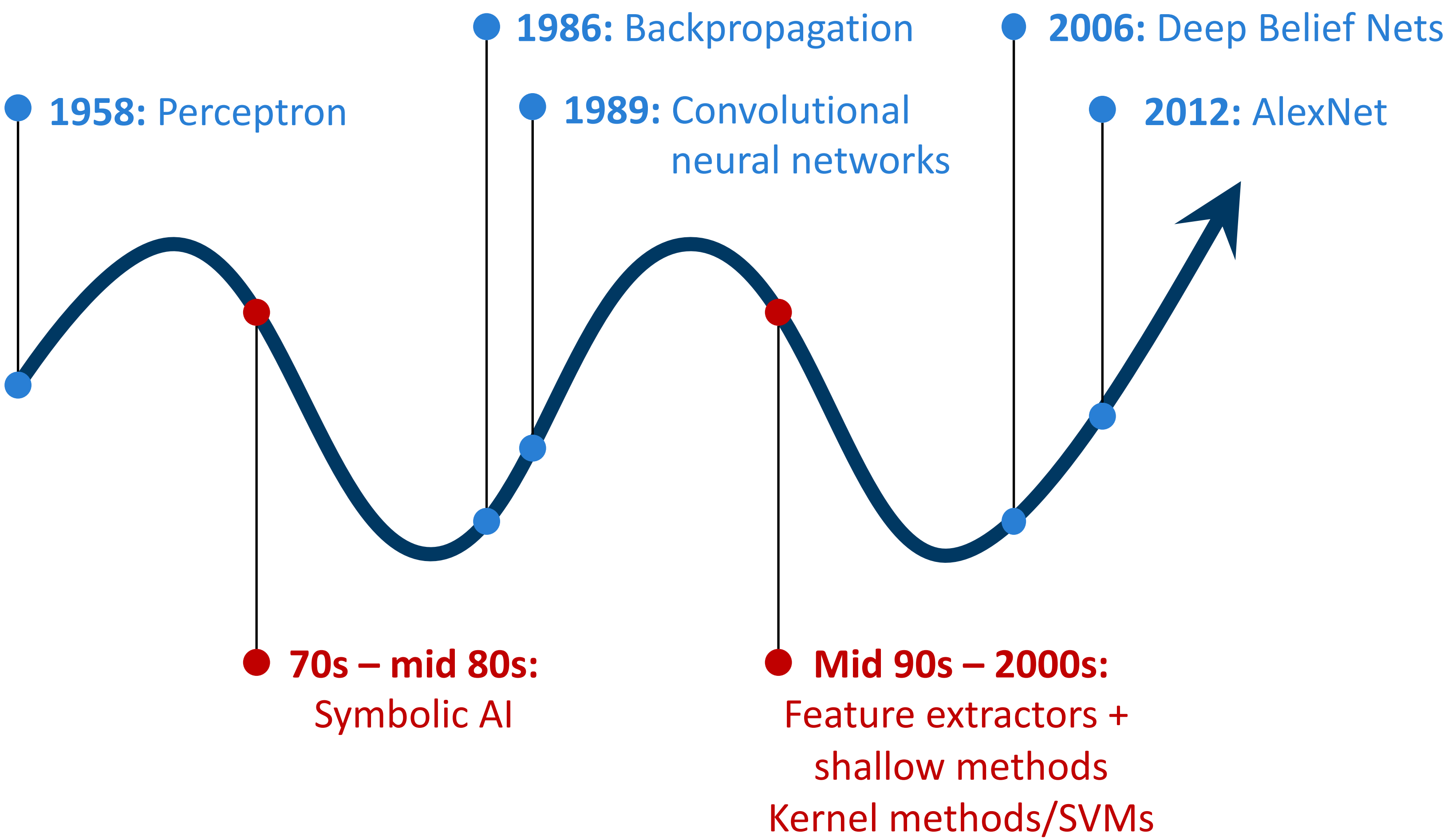


Figure by Andrew Ng

Neural networks: summers & winters



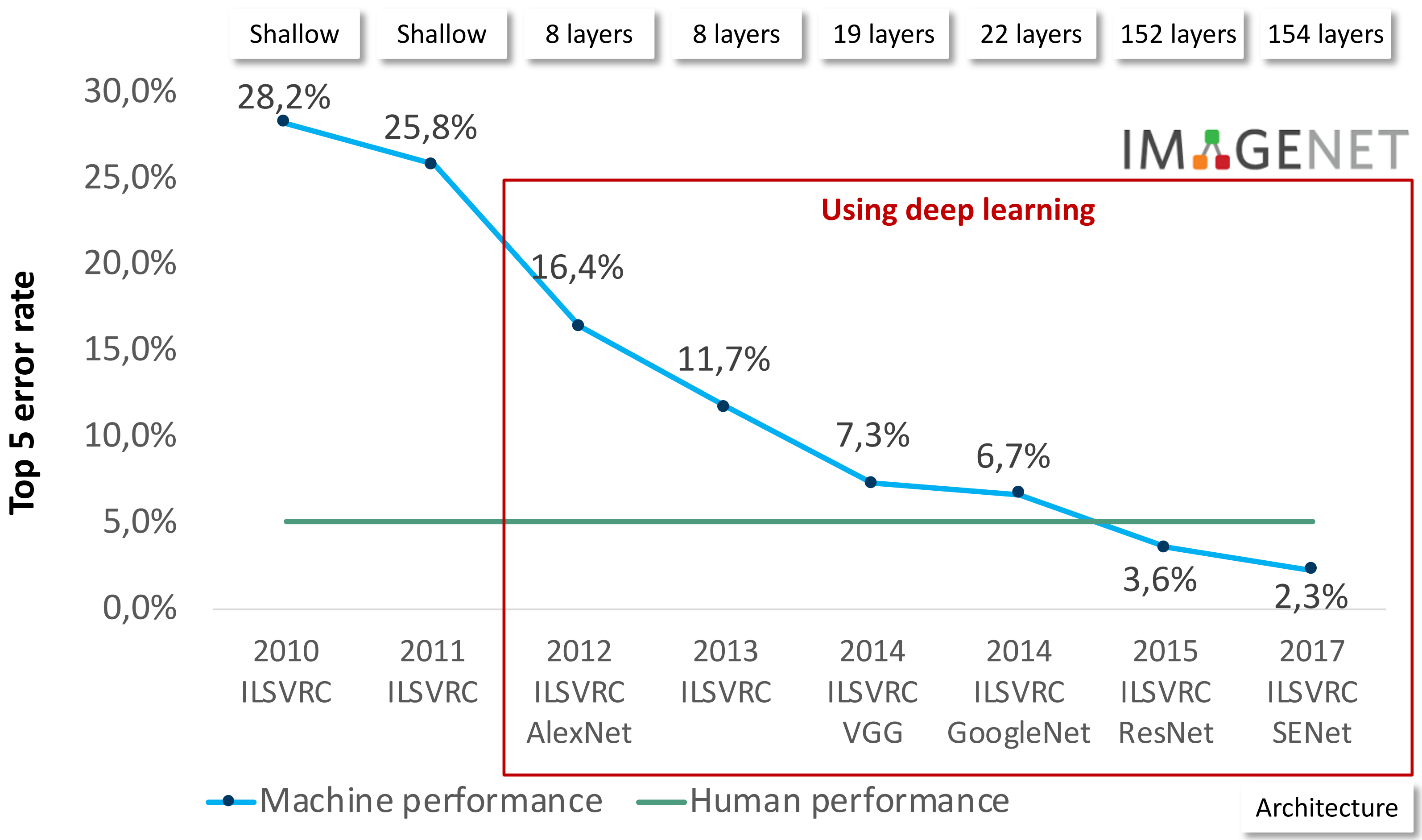
Object recognition: IMAGENET

Dataset and benchmark on image classification

- 1 million images with ground truth class labels for training (hand-annotated)
- 1000 object categories

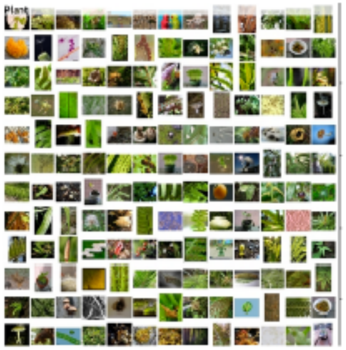
Deng et al., CVPR 2009

Networks get deeper and larger



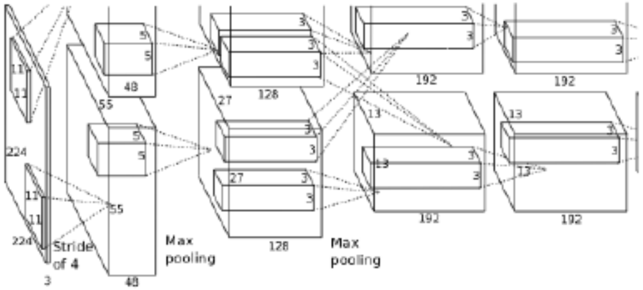
So why does it work now?

The Deep Learning Recipe for Computer Vision



Big Data: ImageNet

+



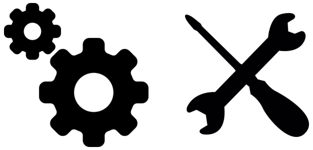
Deep Convolutional Neural Network

+

Backprop on GPU

+ **Many small tweaks**

Sigmoid → ReLU, batch normalization, dropout regularization



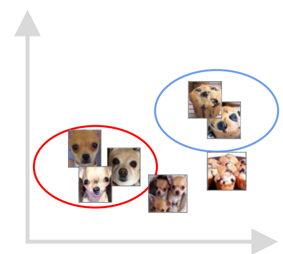
http://www.andreykurenkov.com/writing/ai/a-brief-history-of-neural-nets-and-deep-learning-part-4/

Deep learning = brain?

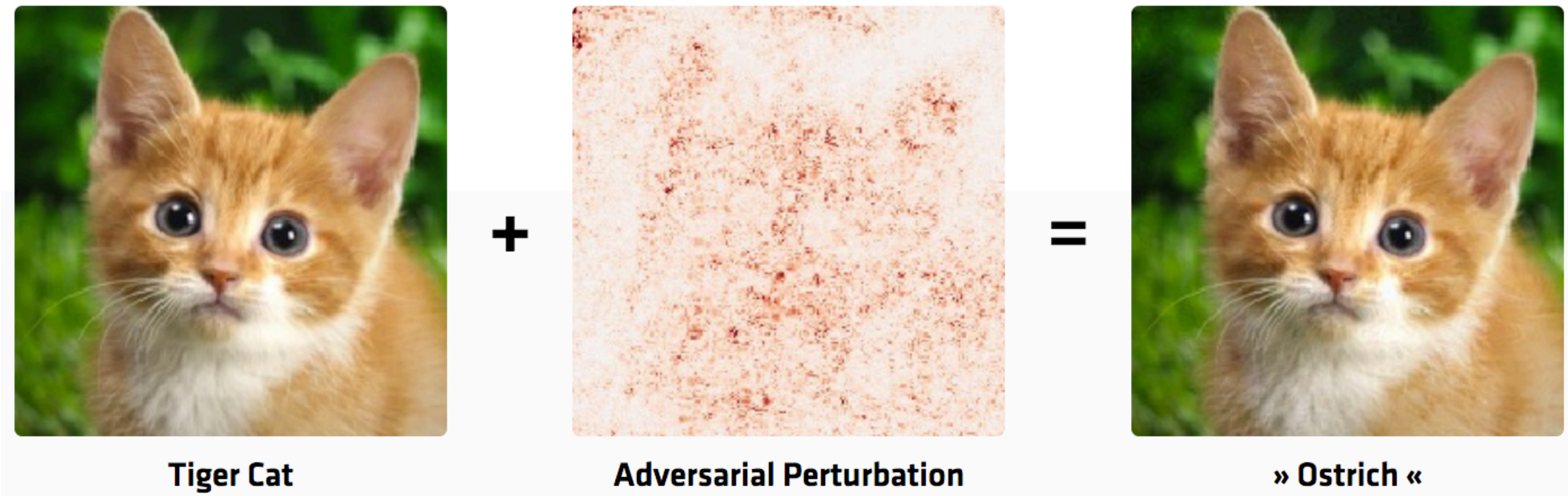
NO.

Currently: mostly supervised learning

Most significant advances with deep learning

	Supervised Learning	**Unsupervised Learning**	**Reinforcement Learning**
	 Dog or Muffin?		
Description	Learn mapping from input to labeled output	Model underlying structure of dataset	Learn strategy to get reward when interacting with the environment
Feedback	Correct answer is given (labeled data)	No answer is given (unlabeled data)	Occasional reward, but no direct feedback on individual actions
Example	• Image classification • Machine translation	• Clustering • Dimensionality reduction	• Learning to play chess • Robot learning to perform a task

Differences between human and machine: „adversarial examples“

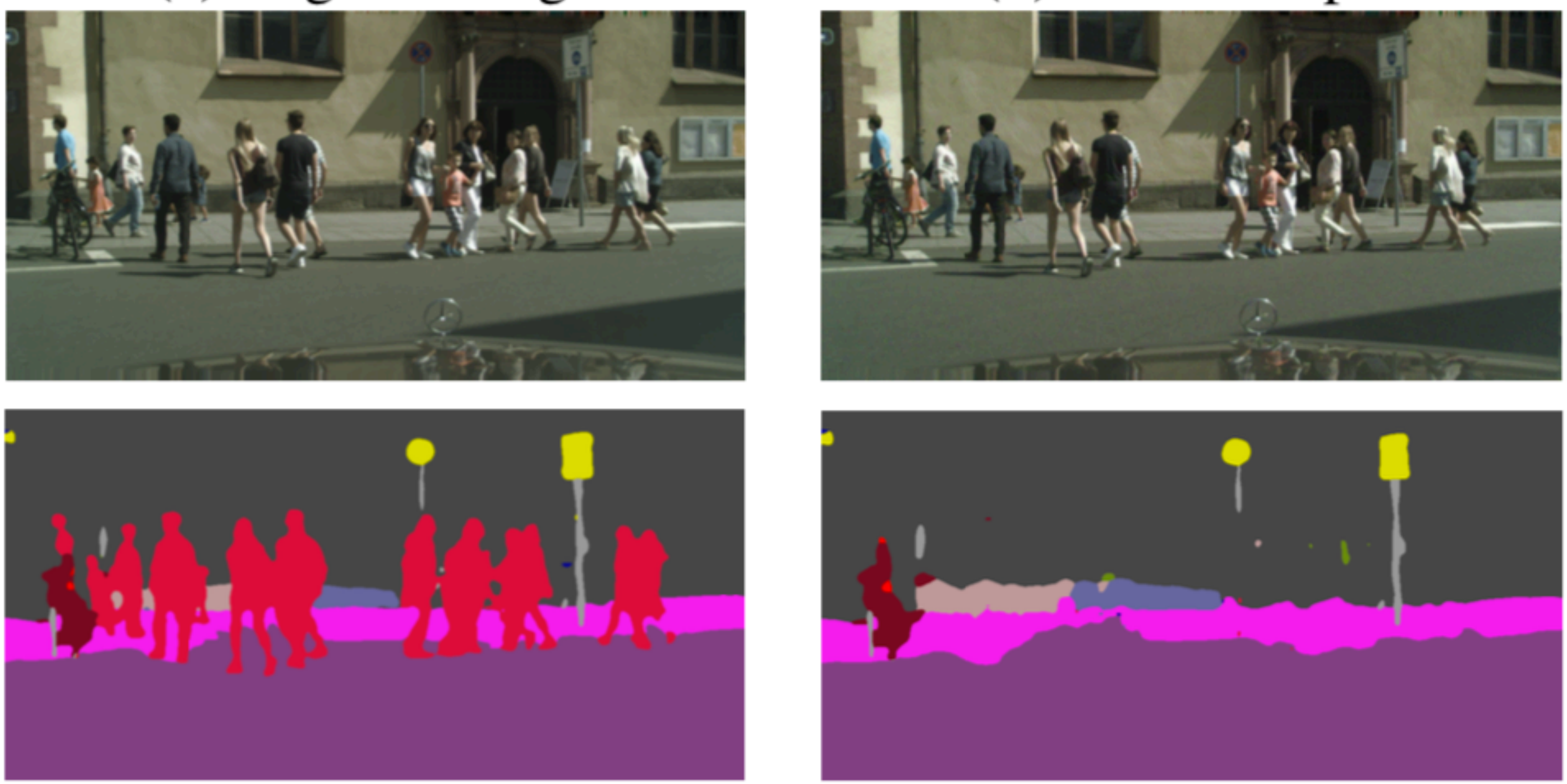


(contrast 10x exaggerated)

Rauber & Brendel, https://robust.vision

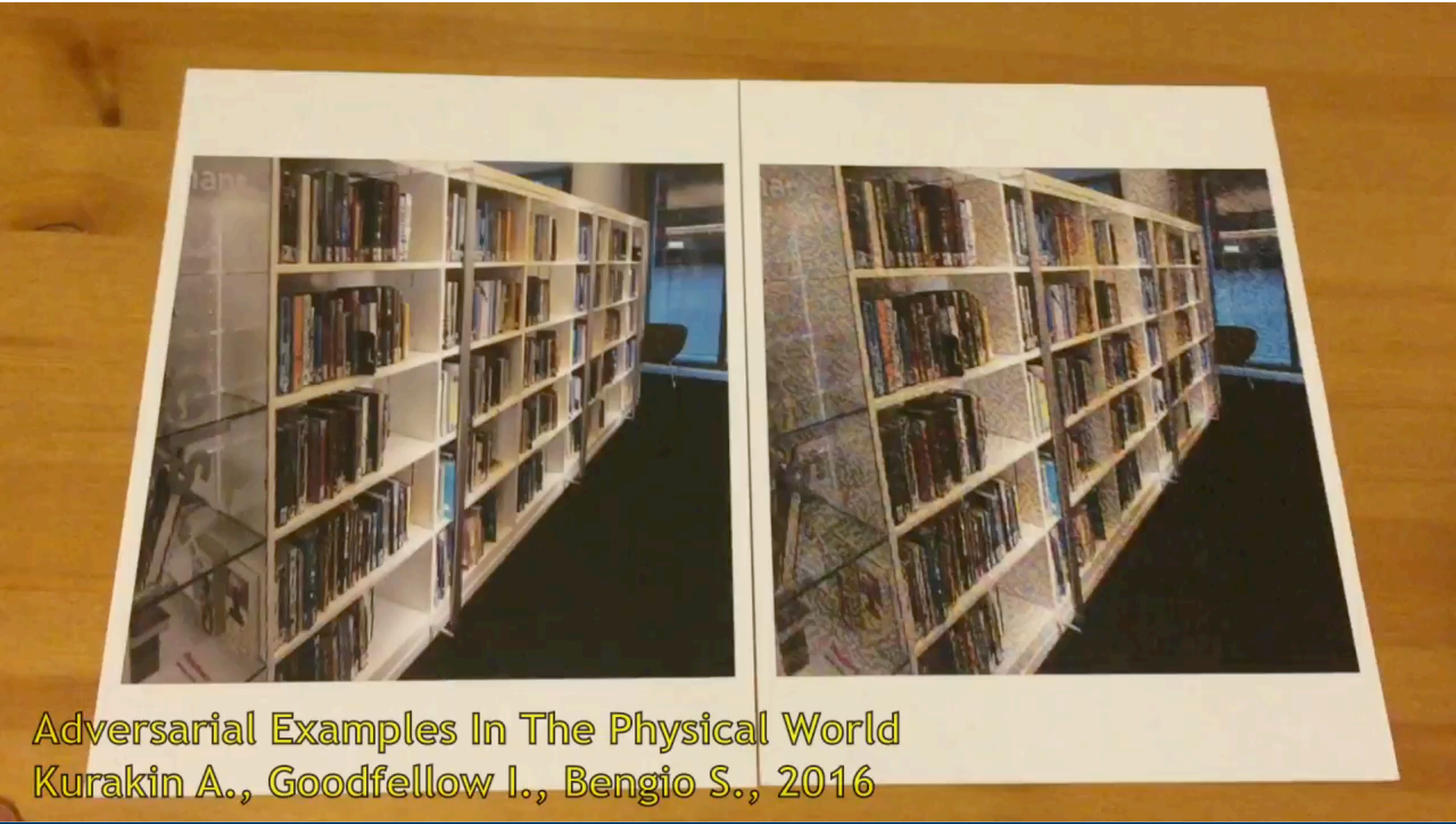
Self-driving cars?

(a) original image (b) adv. example



Fischer et al. ICLR 2017

Real-world example



Differences part II: cat or elephant?

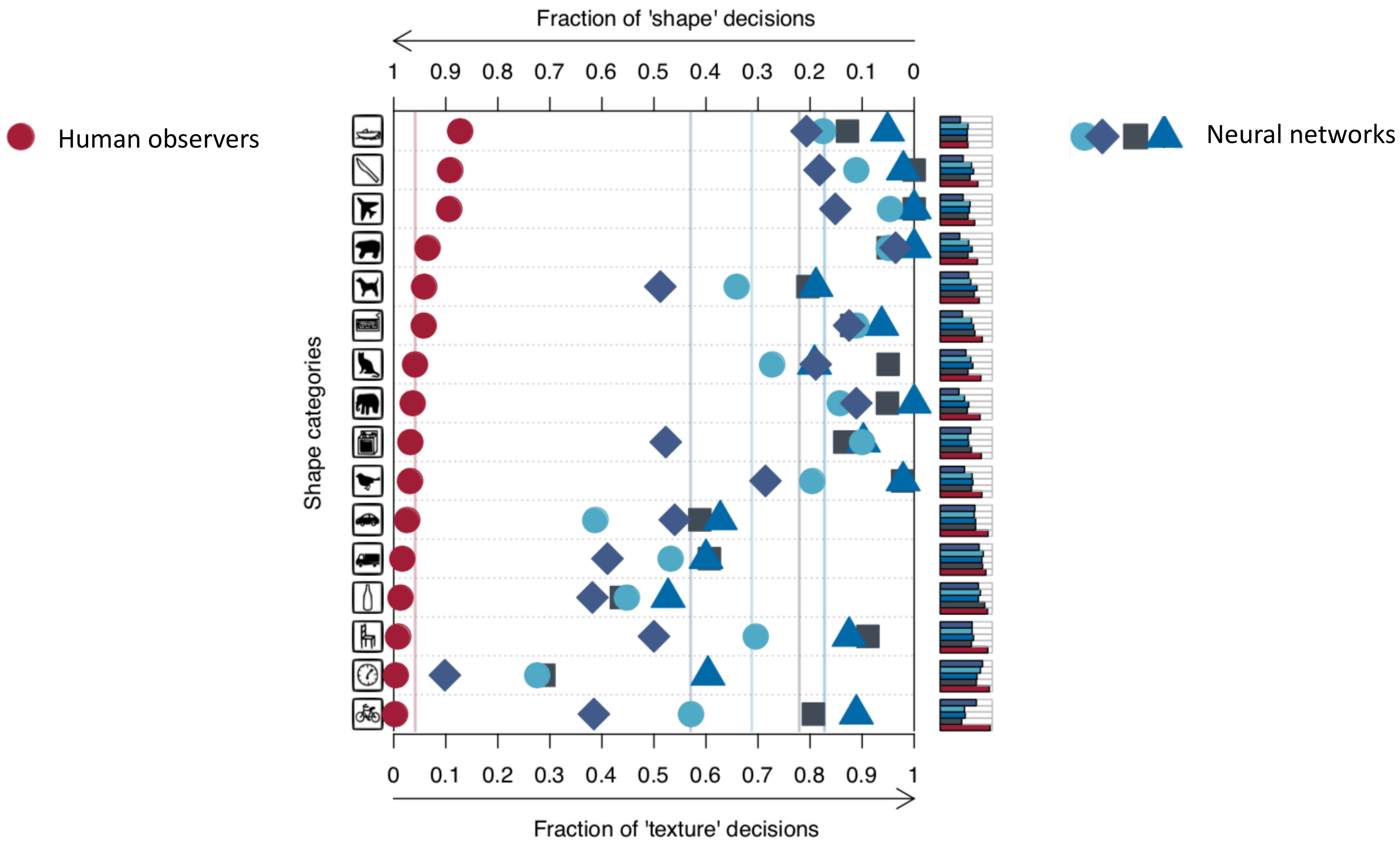
(a) Texture image
81.4% **Indian elephant**
10.3% indri
8.2% black swan

(b) Content image
71.1% **tabby cat**
17.3% grey fox
3.3% Siamese cat

(c) Texture-shape cue conflict
63.9% **Indian elephant**
26.4% indri
9.6% black swan

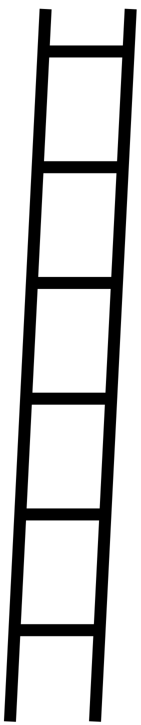
Geirhos et al., arXiv 2018

Human vs. neural net: shape or texture?

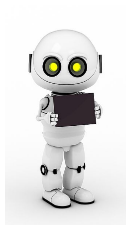


Geirhos et al., arXiv 2018

Summary: Where do we stand?



General AI → **One model mastering** all kinds of different tasks robustly even under difficult and changing conditions



Broad AI → One model **solving specific task(s) under varying conditions**

We are here

Narrow AI → One model **solving one very specific task under stable conditions**

Questions!