CS5242 : Neural Networks and Deep Learning

Lecture 1: Introduction to Deep Learning

Semester 1 2021/22

Xavier Bresson

https://twitter.com/xbresson

Department of Computer Science
National University of Singapore (NUS)

Outline

- AI/DL case studies
- A brief history of AI/DL
- 2012 DL breakthrough
- DL is a universal learning technique
- Terminology AI/ML/DL
- Key ingredients of DL
- AI is the new electricity
- Limitations of AI

Outline

- AI/DL case studies
- A brief history of AI/DL
- 2012 DL breakthrough
- DL is a universal learning technique
- Terminology AI/ML/DL
- Key ingredients of DL
- AI is the new electricity
- Limitations of AI

Deep learning case studies

- AI/DL is more and more present in our daily life:
  - Voice recognition/synthesis
  - AI assistant
  - Machine translation
  - Object recognition
  - Self-driving cars
  - Games
  - Robots
  - Face recognition
  - Healthcare
  - Drug discovery

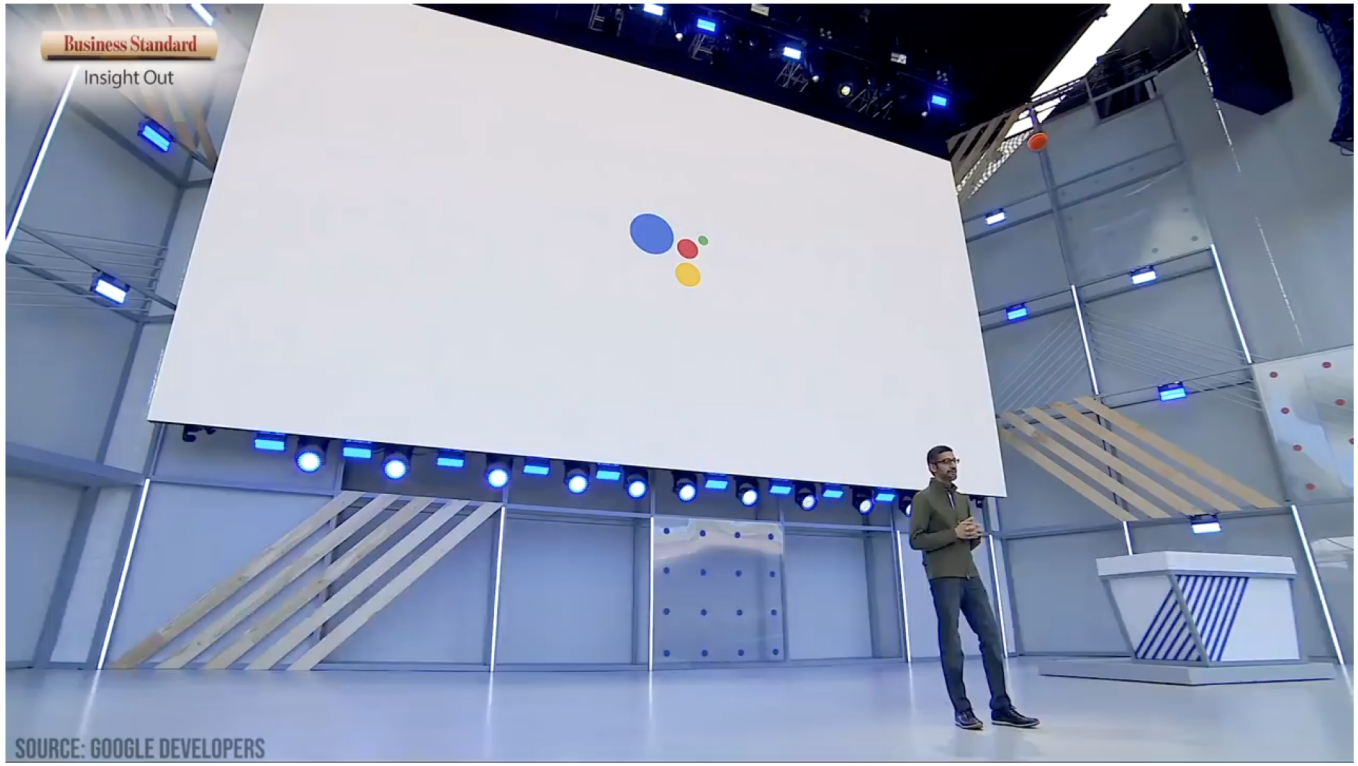
Google search

- Voice recognition - Q&A – voice synthesis : 3 neural networks

Example: What is deep learning?

AI assistant

- Google AI assistant : Voice recognition - Q&A – voice synthesis



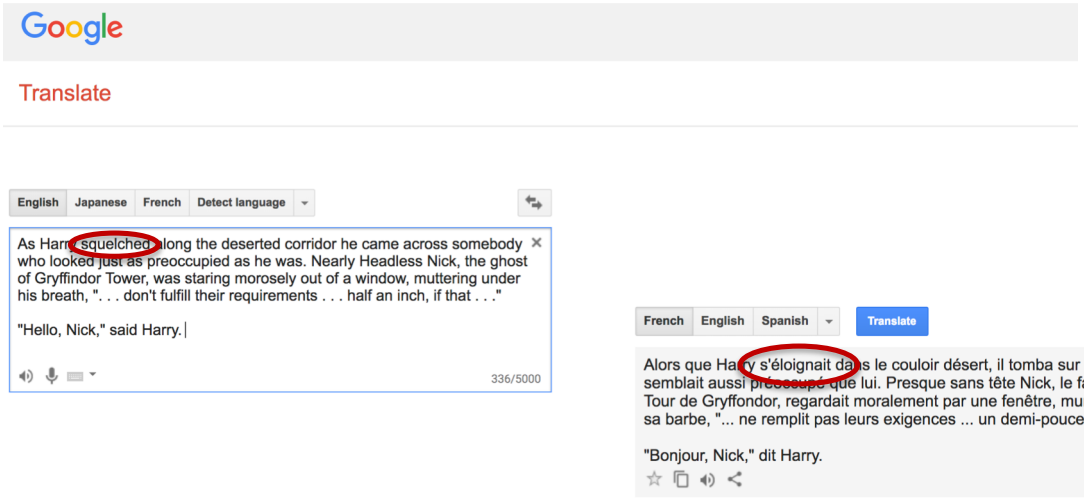
Google's May 2018 I/O developers conference

Turing test passed?

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

Machine translation

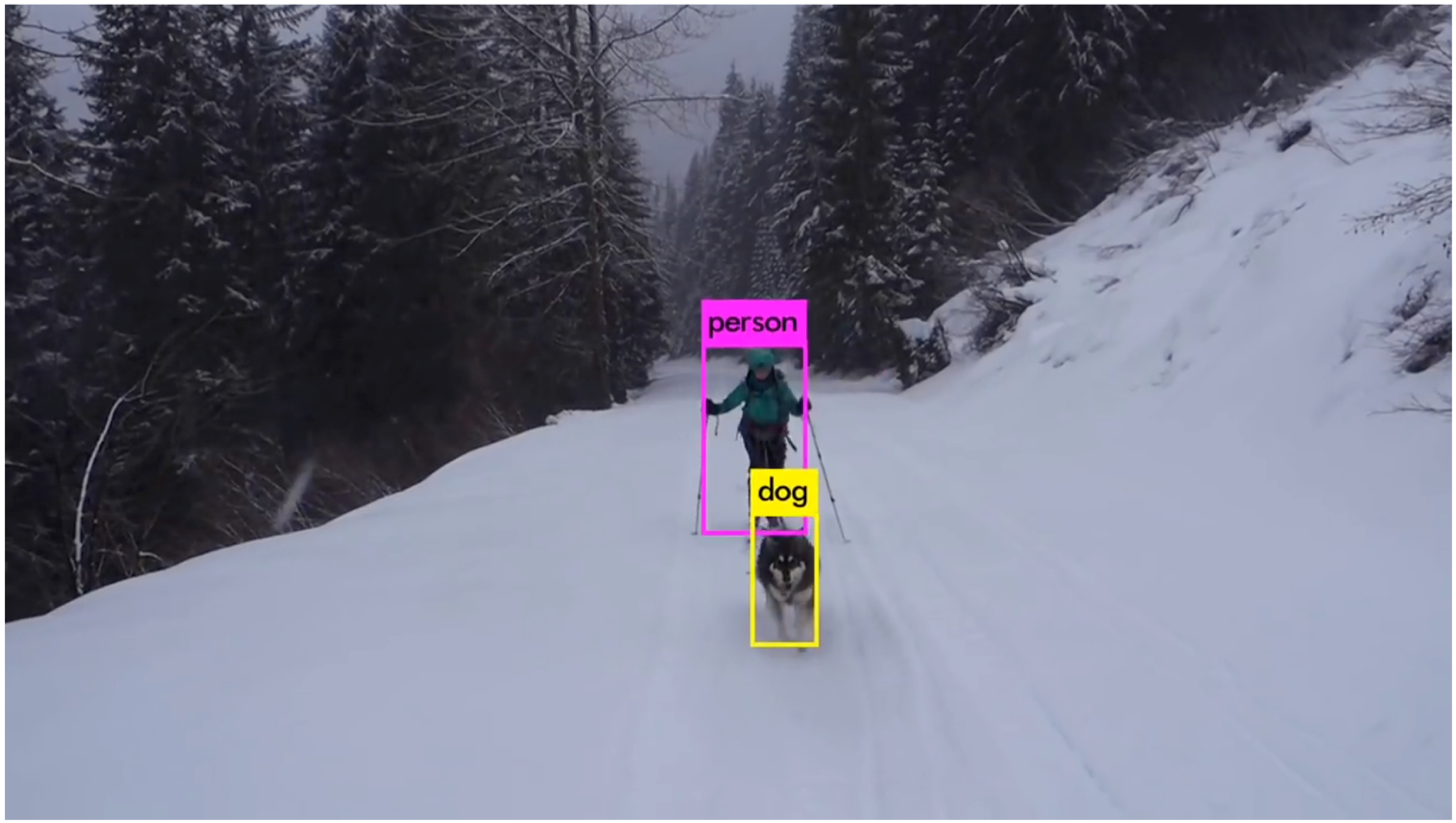
- Harry Potter and The Chamber of Secrets : English ⇒ French ⇒ English



Still far from professional translators, but some translators use it as first draft.

Object recognition

- Real-time object detection and recognition :



YOLO
https://pjreddie.com/darknet/yolo

Self-driving cars

- Requires to solve several computer vision tasks simultaneously :

NVIDIA
2017

Autonomous cars will be the most visible AI application in the coming decade !

Strategy games

- Google DeepMind AlphaGo defeated the world champions at Go, March 2016, May 2017

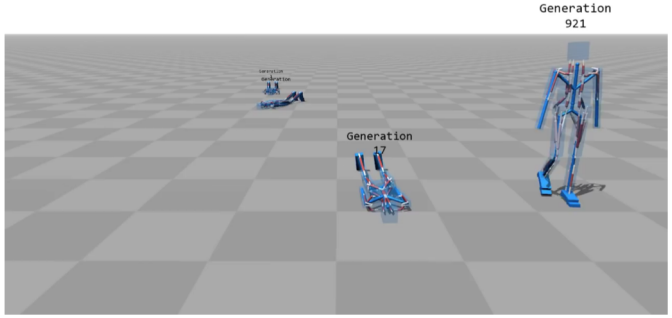
Number of possible board configurations is greater than the number of atoms in the universe

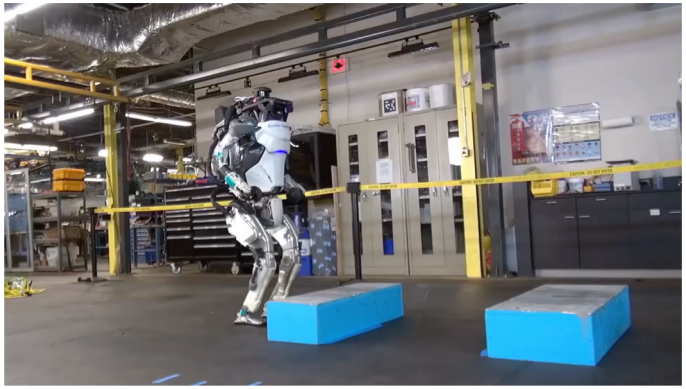
Doom (Facebook)
AI international competition

OpenAI Five: Dota2
The International 2018

Simulated and real-world robots

- Autonomous robots :





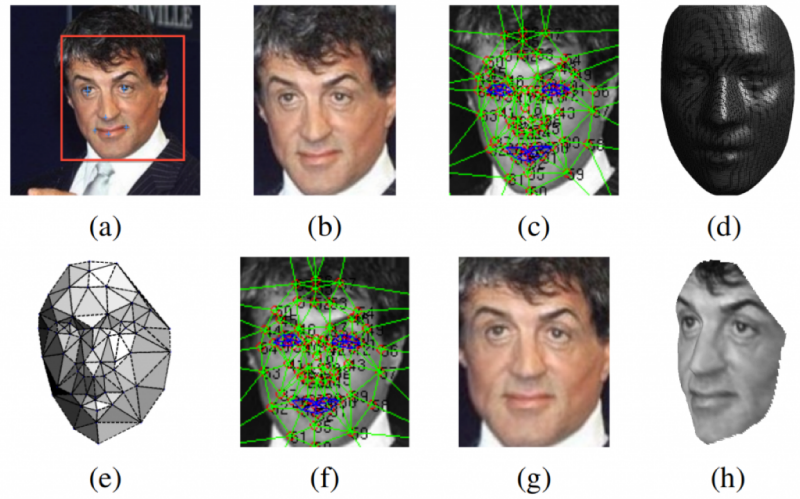
This is not AI-based robots (Boston Dynamics).

This is AI-based robots.

Today, we are far away from autonomous robots !

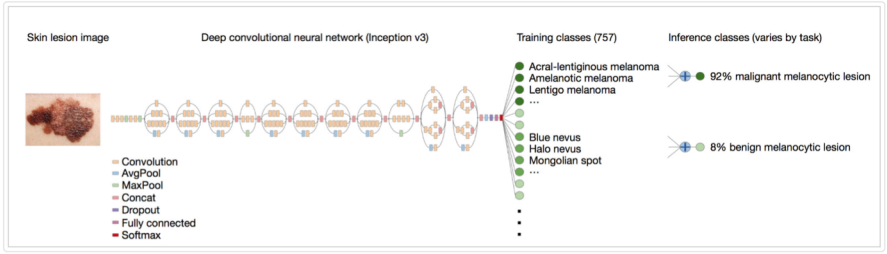
Face recognition

- Facebook face recognition system :



Healthcare

- Dermatology imaging : detection of skin cancer



Dataset: 129K clinical images
consisting of 2,032 different diseases

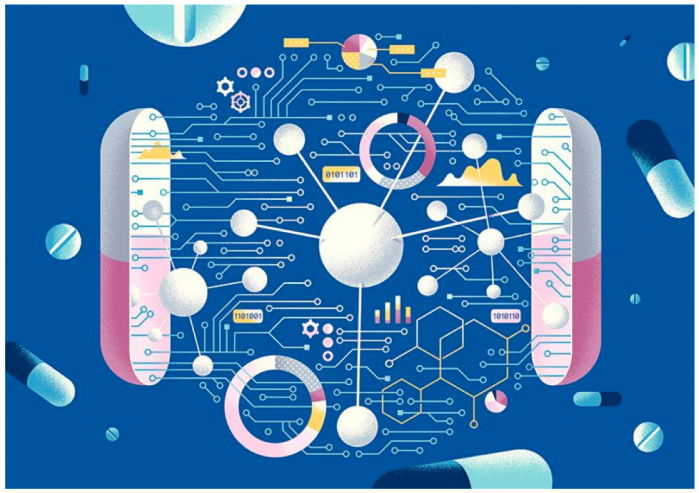
Performance on par or slightly
better than >20 dermatologists,
Feb 2017

Shanghai Huashan Hospital elite radiologists
vs. BioMind, an AI system, July 2018

AI made 87% prediction in 15 min.
Human doctors made 66% prediction in 30 min.

Drug discovery

- Learn to generate molecules with special properties - personalized to patients :

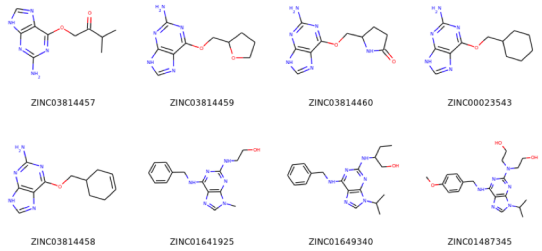


nature
International journal of science

SPOTLIGHT · 30 MAY 2018

How artificial intelligence is changing drug discovery

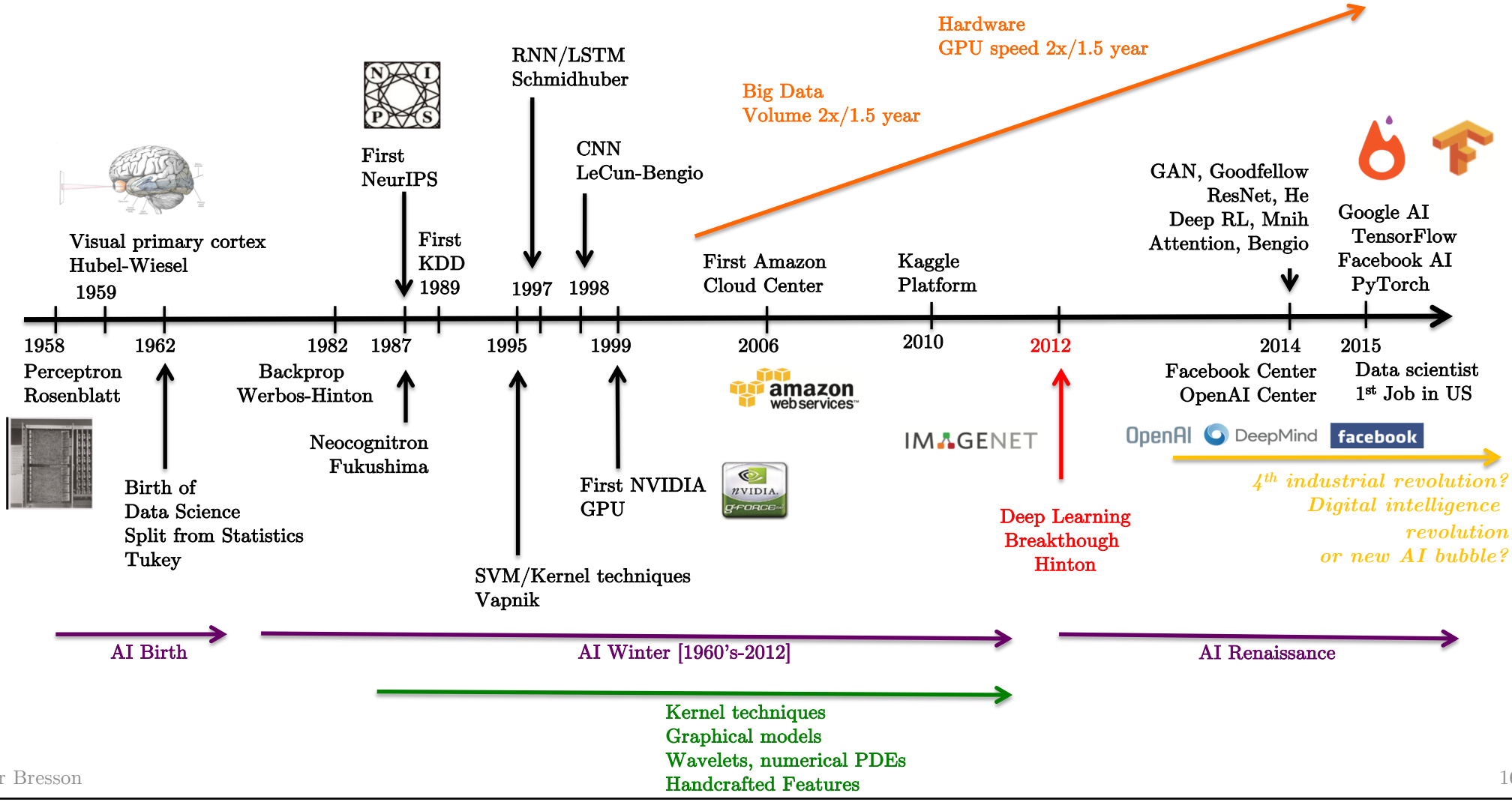
Machine learning and other technologies are expected to make the hunt for new pharmaceuticals quicker, cheaper and more effective.



Outline

- AI/DL case studies
- **A brief history of DL**
- 2012 DL breakthrough
- DL is a universal learning technique
- Terminology AI/ML/DL
- Key ingredients of DL
- AI is the new electricity
- Limitations of AI

A brief history of DL



Outline

- AI/DL case studies
- A brief history of DL
- **2012 DL breakthrough**
- DL is a universal learning technique
- Terminology AI/ML/DL
- Key ingredients of DL
- AI is the new electricity
- Limitations of AI

2012: The year of DL breakthrough

- ImageNet challenge : International Image Classification Challenge
- Goal: Design the best algorithm for image recognition

  (one of the most basic problems in computer vision since 50 years)

Input: An image

Image recognition algorithm

Siamese Cat

Output: A category

2012: The year of DL breakthrough

- ImageNet challenge : International Image Classification Challenge

  - 1.4 million images
  - Collected from Flickr and other search engines.
  - 1000 object categories
  - Each image was manually labelled by a human being : They used a crowdsourcing approach to labeled the pictures with Amazon Mechanical Turk.

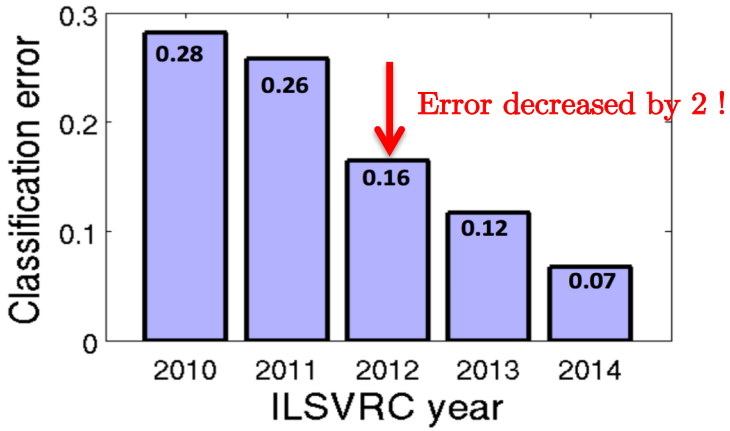
ImageNet dataset

2012: The year of DL breakthrough

- ImageNet challenge :
  - Training set: 1.2 million labelled images (given to the teams)
  - Test set: 0.2 million labelled images (kept by the organizers)
  - Each team use the training set to train their algorithm, then they submit it to the organizers.
  - The organizers evaluate each algorithm on the test set.

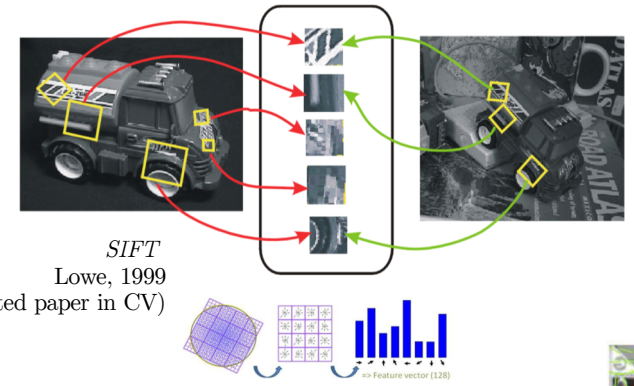
- Observe the plot between 2011 and 2012 :

  Change of paradigm : From 50 years of hand-crafted visual features to learned visual features from massive amount of data.



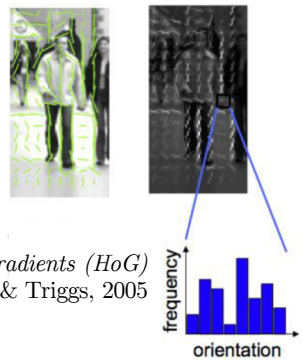
Before 2012

- Paradigm: Handcraft data features and feed them to a classifier.



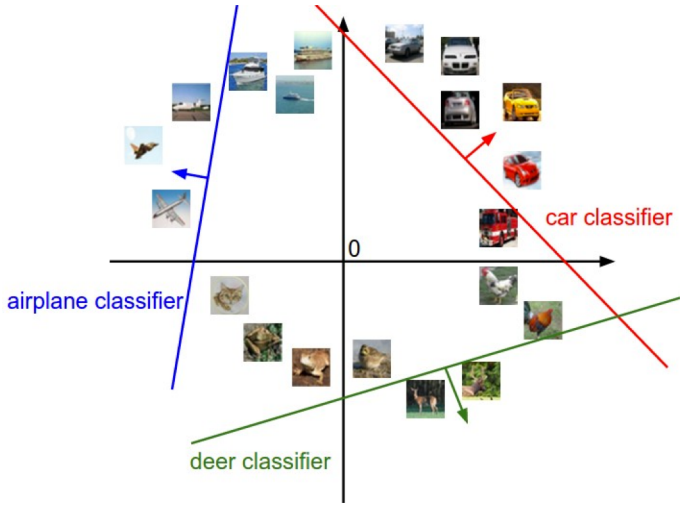
SIFT
Lowe, 1999
(most cited paper in CV)

SIFT features



Histogram of Gradients (HoG)
Dalal & Triggs, 2005

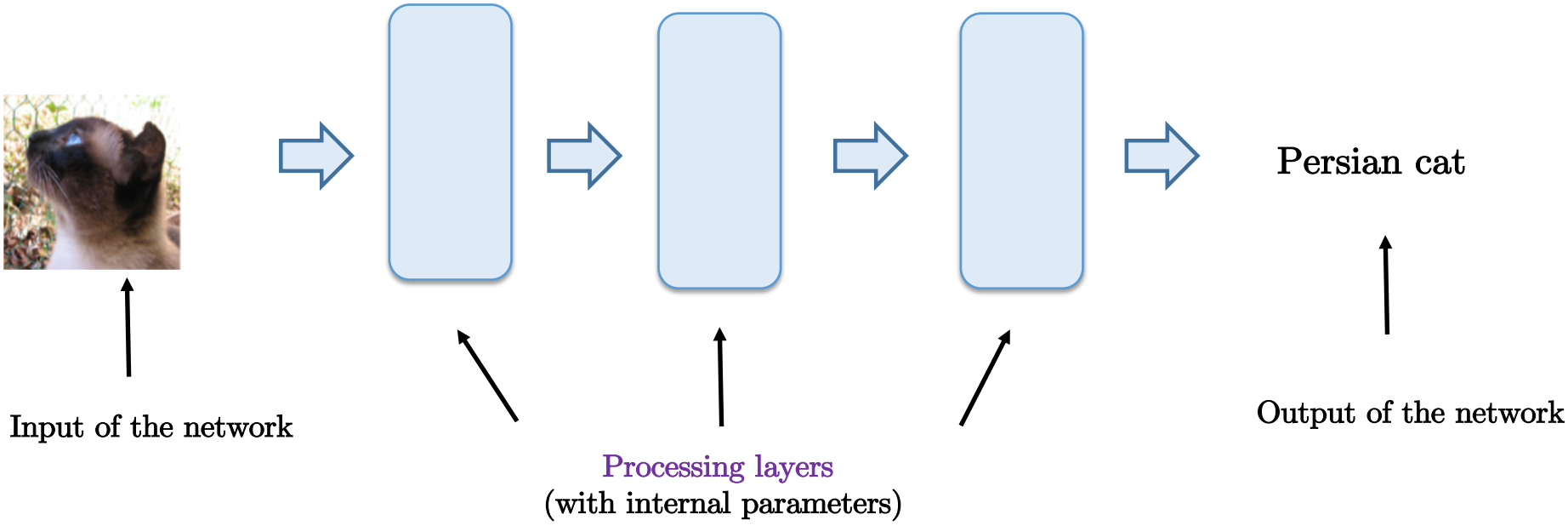
HOG features



SVM classifier

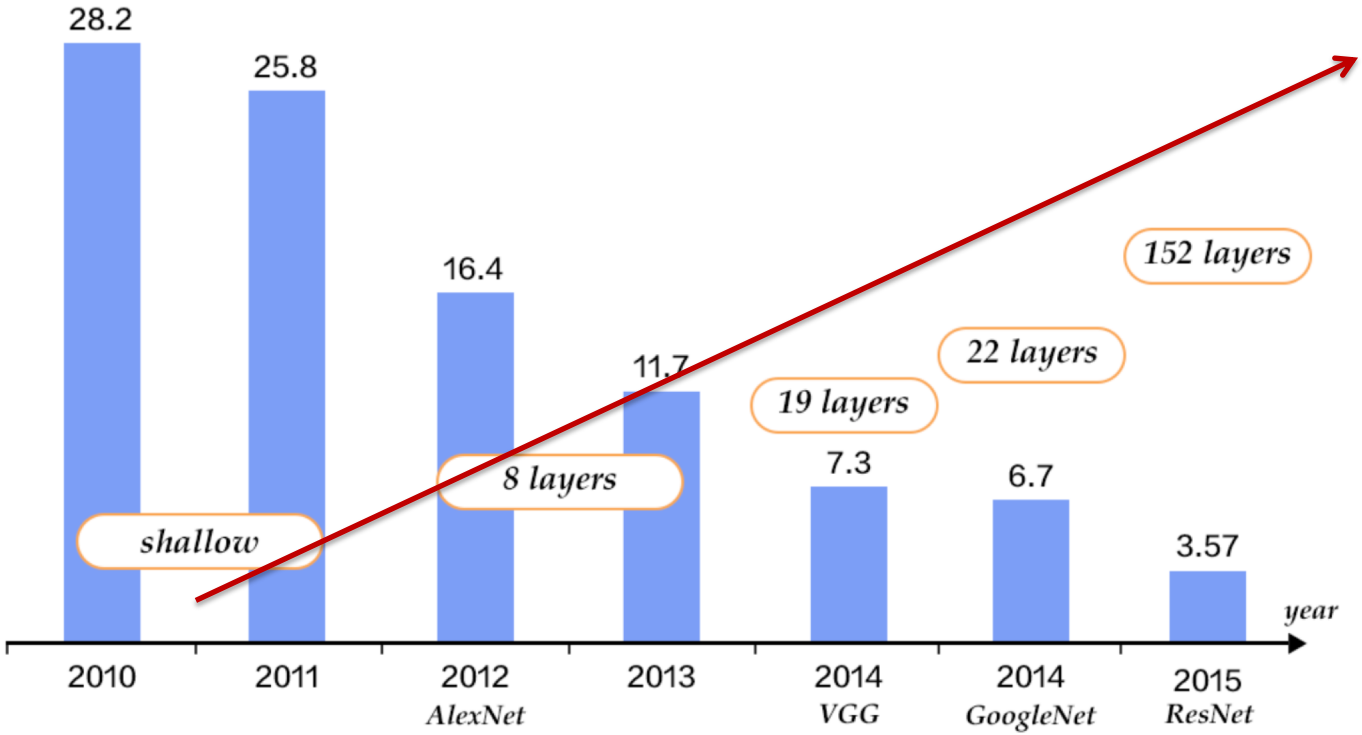
After 2012

- New paradigm: Learn data features and classifier together, a.k.a. end-to-end systems.
- Neural networks are the first class of models that can train end-to-end systems with large learning capacity by using multiple layers.



The deeper the better

- Deep Learning = Neural networks with many layers
- The deeper the more learning capacity :



2018
1000s layers

Outline

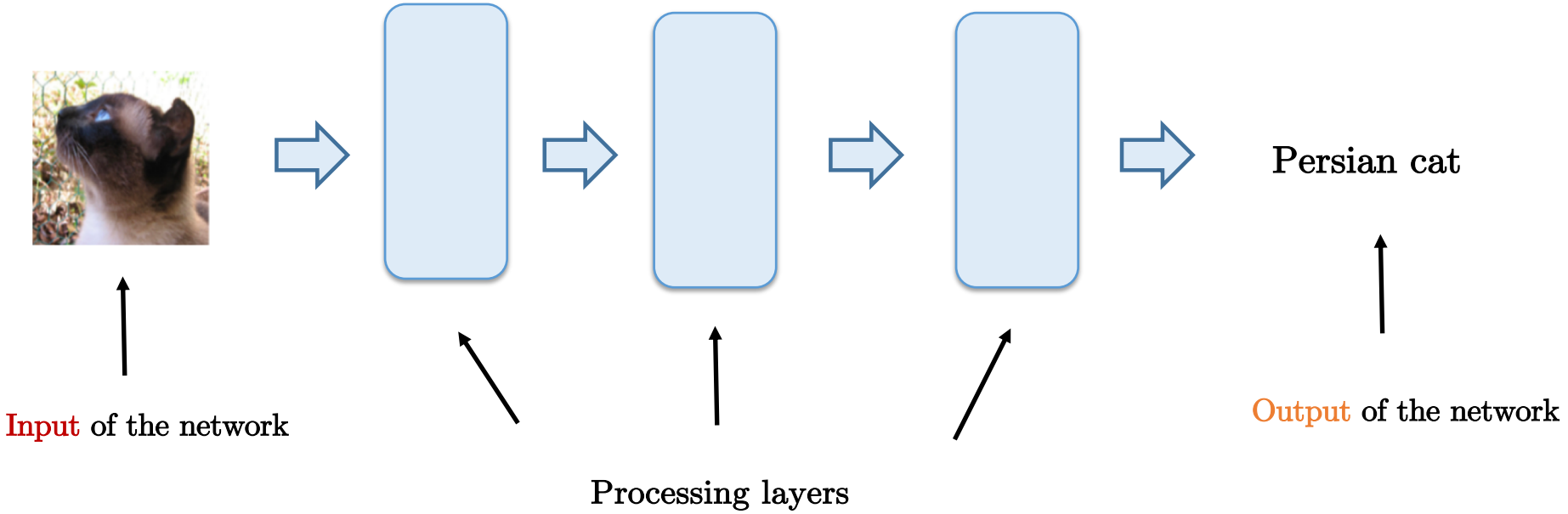
- AI/DL case studies
- A brief history of DL
- 2012 DL breakthrough
- **DL is a universal learning technique**
- Terminology AI/ML/DL
- Key ingredients of DL
- AI is the new electricity
- Limitations of AI

DL/NNs can be used for many tasks

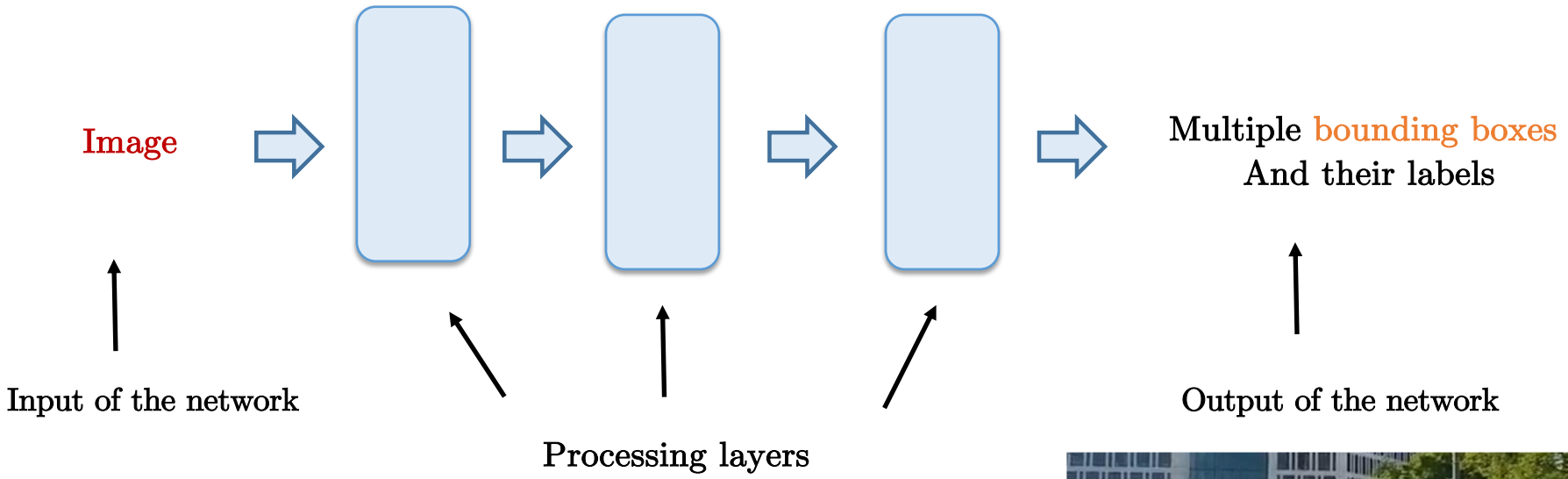
- Neural networks are universal learning techniques that can be applied to very distinct fields like computer vision, natural language processing, physics, chemistry, healthcare, biology, etc.

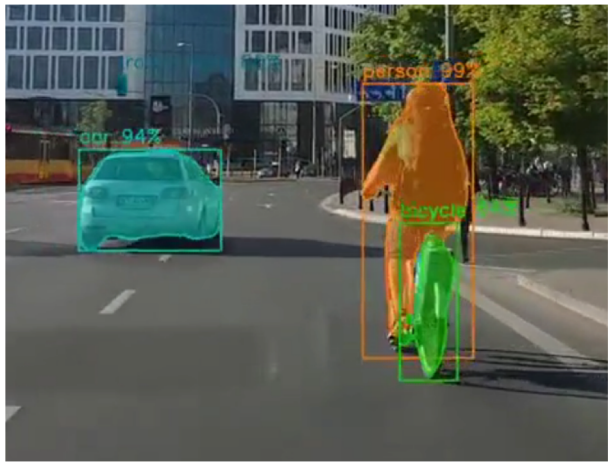
- Tasks where NNs provide the best performance (SOTA):
  - Image recognition
  - Object detection
  - Voice recognition
  - Speech synthesis
  - Translation
  - Recommender systems
  - ...

Image recognition

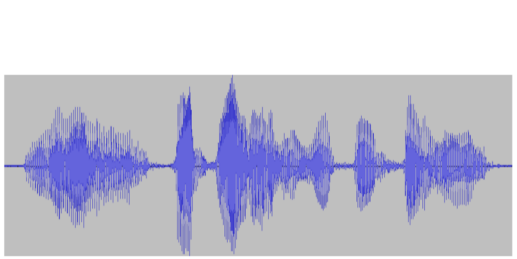


Object detection & recognition

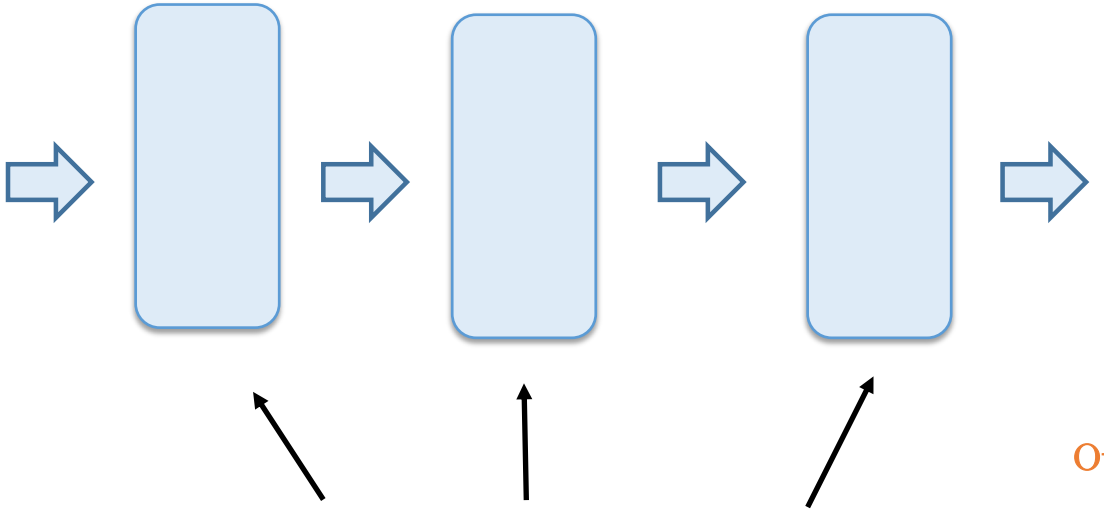




Voice recognition (speech-to-text)



Input of the network

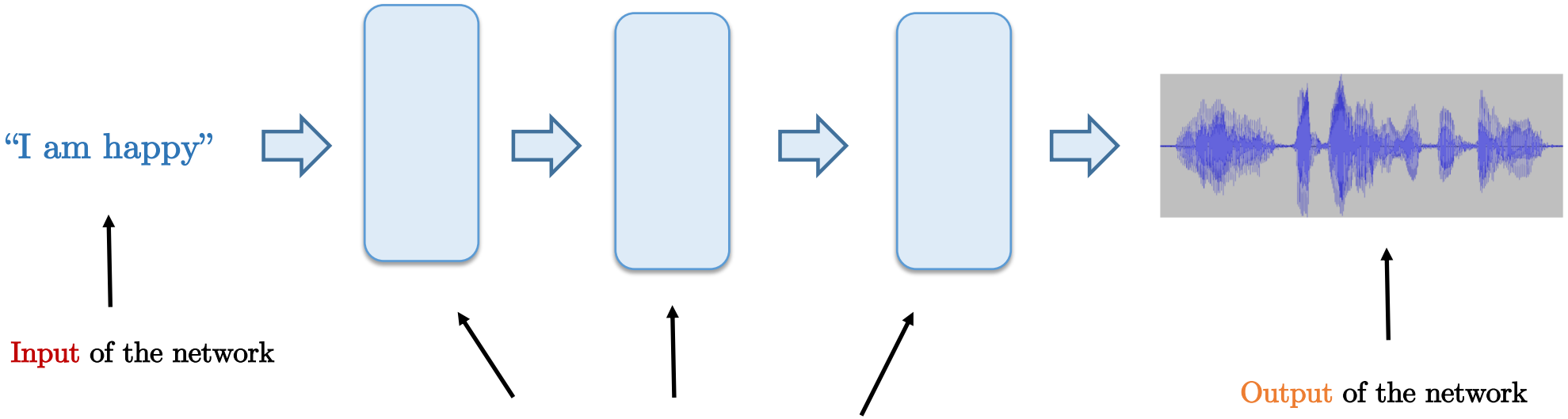


Processing layers

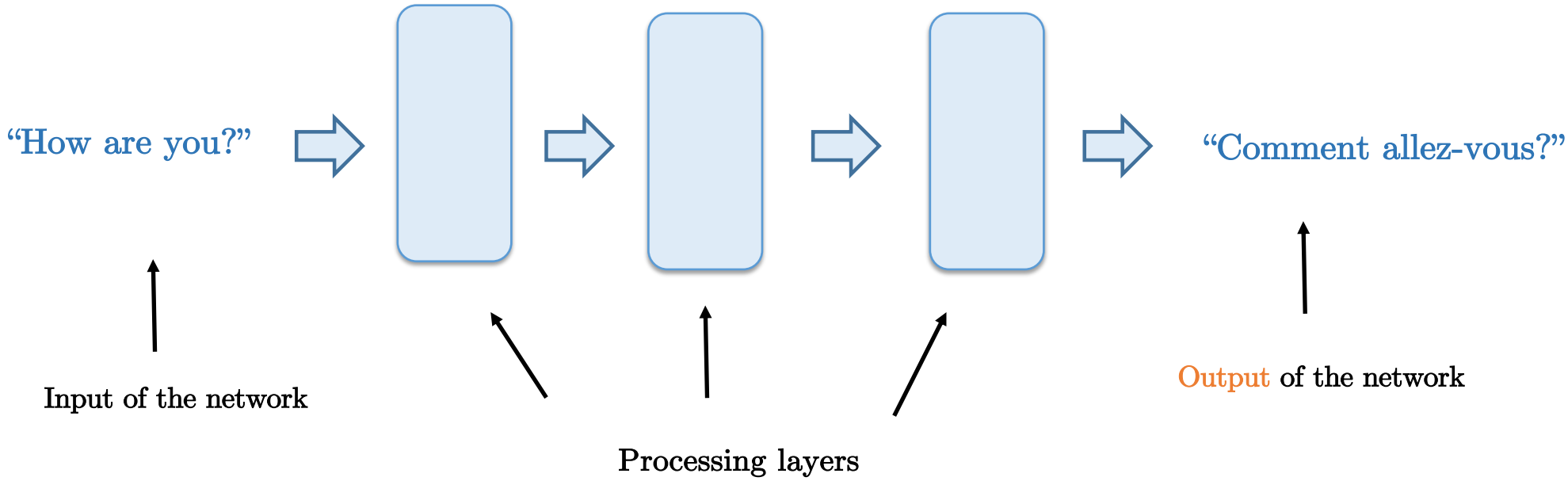
"What is your name?"

Output of the network

Speech Synthesis (text-to-speech)

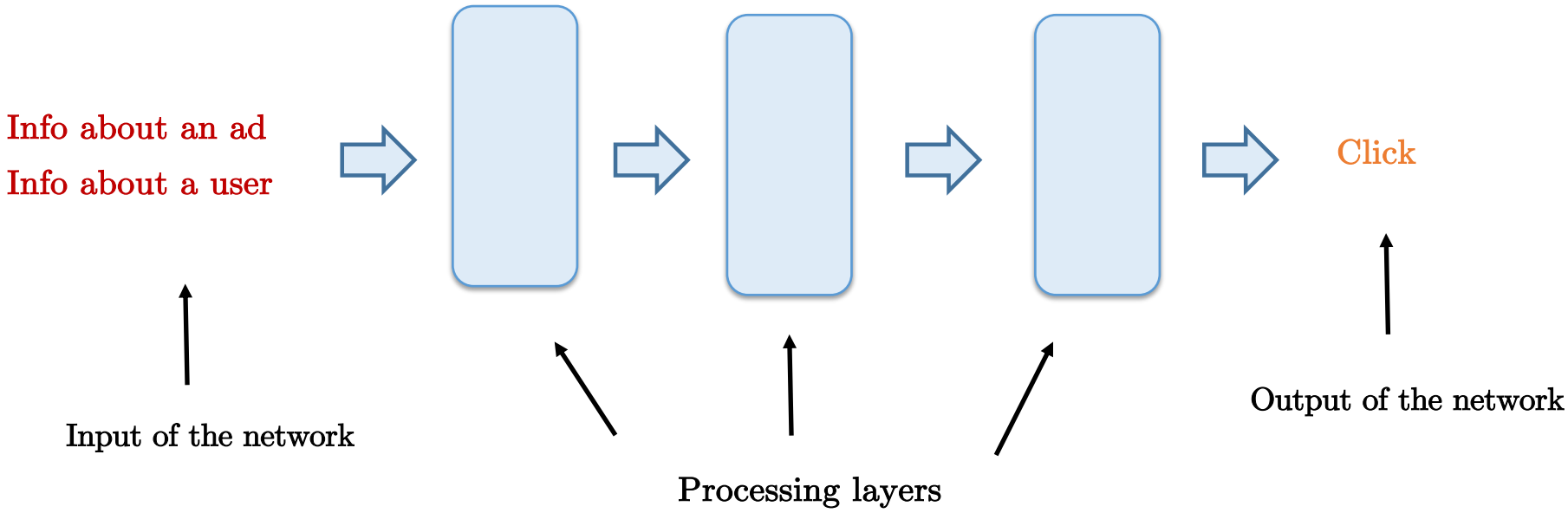


Translation (sequence-to-sequence)



Online advertisement

- Will this user click on this ad?

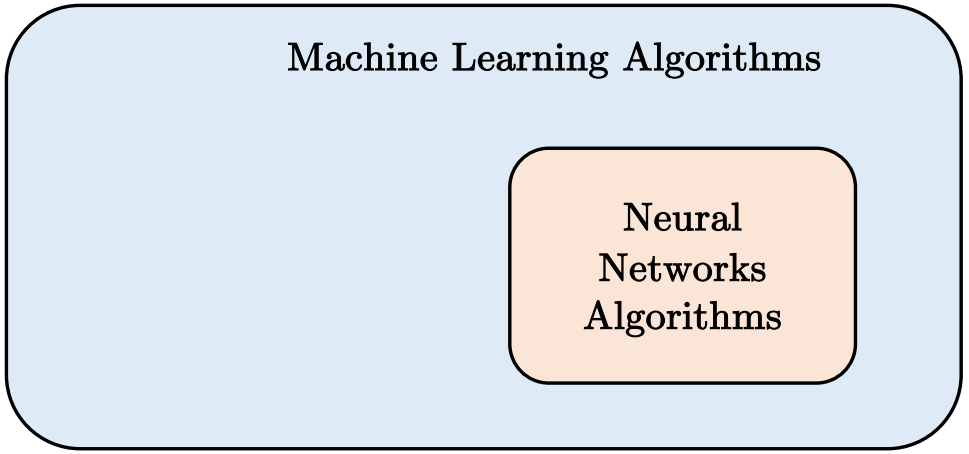


Outline

- AI/DL case studies
- A brief history of DL
- 2012 DL breakthrough
- DL is a universal learning technique
- **Terminology AI/ML/DL**
- Key ingredients of DL
- AI is the new electricity
- Limitations of AI

AI/ML/DL

- A.I. : Artificial intelligence, a vague term that can mean many things (no definition exists).
- Machine learning : Techniques that solve tasks from data learning.
- Deep Learning : Neural networks with many layers
- Nowadays, A.I. advances are exclusively based on deep learning algorithms.



Outline

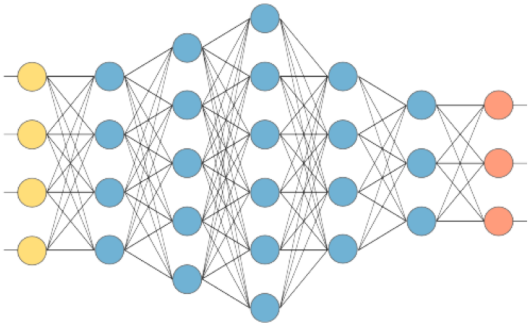
- AI/DL case studies
- A brief history of DL
- 2012 DL breakthrough
- DL is a universal learning technique
- Terminology AI/ML/DL
- **Key ingredients of DL**
- AI is the new electricity
- Limitations of AI

Key successful ingredients

- Deep Learning success = Big data + GPUs + Neural Networks

+

+



Big data

- We live in a digital society:
  - Eric Schmidt : "Every 2 days we create as much information as we did up to 2003."
  - Moore's law of data : Volume of data double every 1.5 years.

- Some datasets:
  - ImageNet : 1.4 millions images
  - Baidu faces : Baidu use a training set of 200 millions images for their face recognition system
  - Translation : Europarl parallel corpus
    - Extracted from the proceedings of the European parliament.
    - 2 million sentences translated in all of the languages of the EU
  - Self-driving cars : Uber's Self-Driving Cars Hit 2 Million Miles

Big data

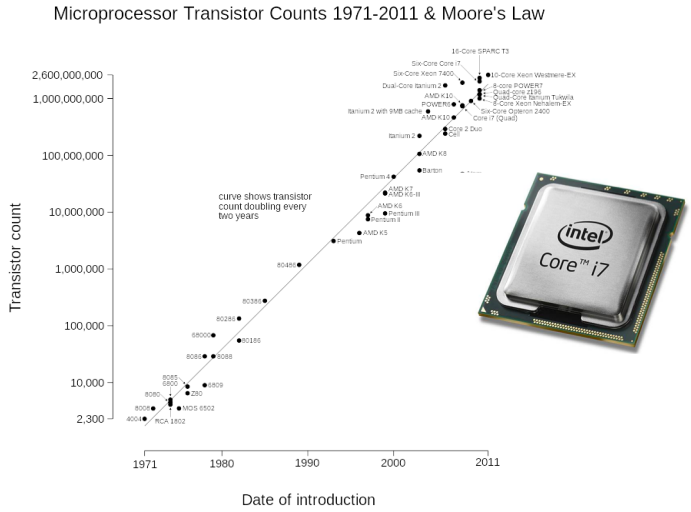
- How much data to get to super-human performances?
  - Rule of thumb: $n=d/\varepsilon$, d number of data features, ε is the accuracy
  - Example: ImageNet $d=256^2 \approx 1M$ and $\varepsilon \approx 1 \Rightarrow n=1.4M$ and 16% error (human error is 5%)
- Collecting data is easy but labeling data is challenging:
  - Labeling is time consuming (each data to a class).
  - It requires humans (learning ability bounded by human intelligence and bias).
  - Some important data is not accessible (e.g. medical data, nuclear meltdown).
- Deep learning is designed for labeled data (supervised learning). It is greedy in (labeled) data.
- Deep learning for unlabeled data is an open problem (unsupervised learning) $\Rightarrow$ Next AI revolution !

Rise of computational power

- Moore's law : 100x faster every decade
  - CPU, GPU, TPU, Cloud computing

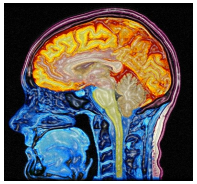
1st GPU in 1999

Amazon cloud in 2006



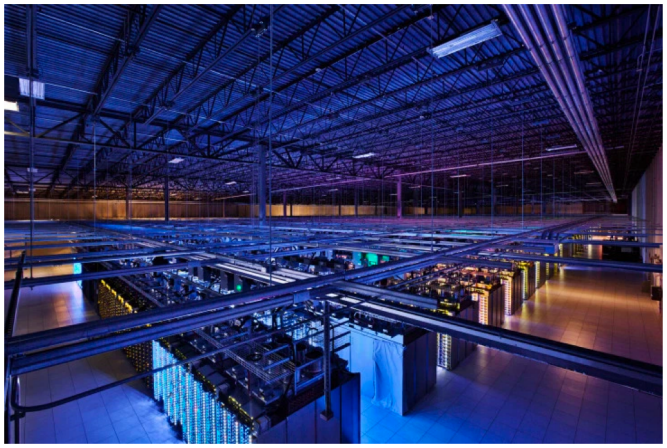
- Brain has 10^{14} synapses/10^3 Tflops = Computers in 2030 (100 times faster than today)

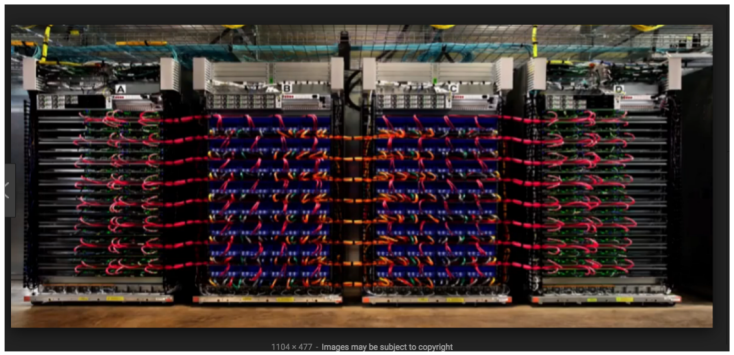
2030
=



Deep learning infrastructure

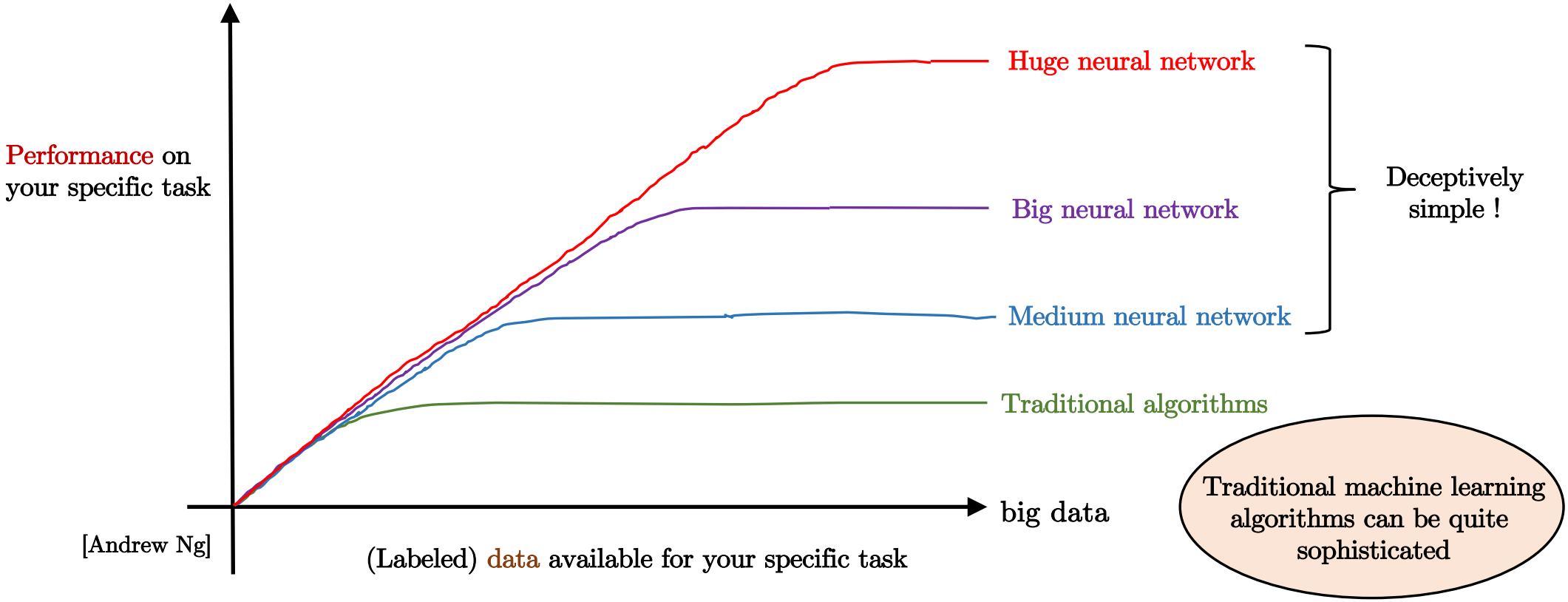
- Google Tensor Processing Unit (TPU) :
  - 2013: Google projects that they would need to double their datacenter to meet demand for speech recognition.
  - The TPU was designed, verified, built and deployed in data centers in just 15 months.





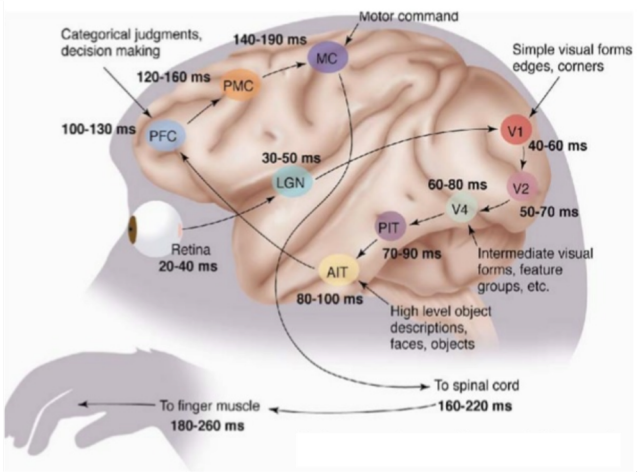
Machine learning algorithms

- Only deep learning systems are able to absorb (exploding) big data :

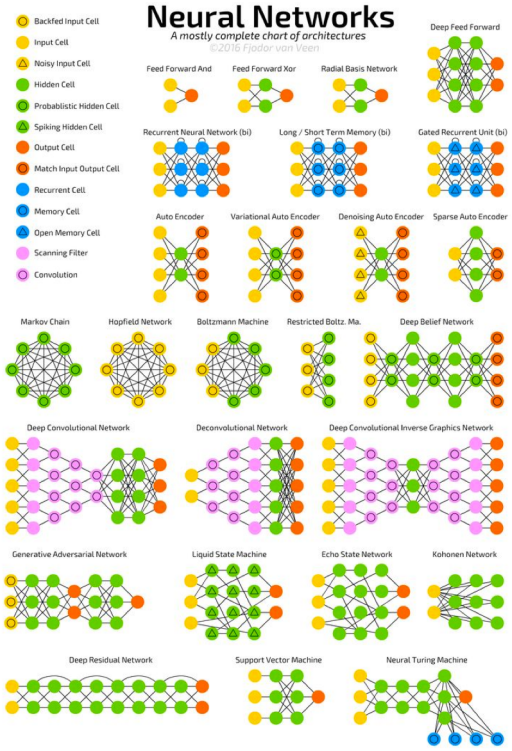


Deep learning architectures

- Most research works focus on the best architecture that solves a specific task (there is no DL model that can solve several tasks like the human brain).



Biological neural network



Artificial neural networks

- Most successful architectures:
  - Convolutional Neural Networks: Computer Vision, Video
  - Recurrent Neural Networks: NLP, Machine translation, speech-to-text


G. Hinton


Y. LeCun


Y. Bengio


Y. Schimdhuber

Turin Award 2019 !

Outline

- AI/DL case studies
- A brief history of DL
- 2012 DL breakthrough
- DL is a universal learning technique
- Terminology AI/ML/DL
- Key ingredients of DL
- **AI is the new electricity**
- Limitations of AI

AI is the new electricity

- "AI is the new electricity" - Andrew Ng

  Electricity has changed the whole industry, sector after sector like transport, manufacturing, agriculture, etc ⇒ AI has the same potential.

- Industrial revolutions :
  - 1st industrial revolution (machine revolution) [1760-1840]: From hand-tools to first machines (agriculture)
  - 2nd industrial revolution (electrical revolution) [1870-1914]: From mechanical machines to electricity
  - 3rd industrial revolution (digital revolution) [1947-2012]: From analogue electronic technology to digital (internet, smartphones)
  - 4th industrial revolution (smart digital revolution) [2012-]: From digital data to augmented intelligence

Outline

- AI/DL case studies
- A brief history of DL
- 2012 DL breakthrough
- DL is a universal learning technique
- Terminology AI/ML/DL
- Key ingredients of DL
- AI is the new electricity
- **Limitations of AI**

Current state of AI

- Modern AI breakthrough started in 2012 and there is no coming back – AI will be part of any machine systems.
- AI has a few impressive results but AI is in its infancy – immature engineering field (solutions are empirical, not robust, not scalable, not transferable).
- It will take (at least) a decade to develop a solid software, a hardware environment and a mathematical framework (statistics, optimization, etc) for AI.
- AI is a buzzword – Machine learning and data science are better names.
- Success is limited to perceptual tasks (vision, hearing) – no breakthrough results in cognitive tasks (reasoning, planning, common sense).

AI vs IA

- AI (Artificial Intelligence):
  - Terminator, Hollywood movies – Machines take over the world.
  - No sense
  - Definition of intelligence is unclear.

M. Jordan

- IA (Intelligence Augmentation):
  - Lots of sense – under development (will have a huge impact in our lives)
  - Better search engines, recommendation systems, natural language translation, Q&A's.
  - Augment people skills.

What is possible and not yet possible?

- Computer vision
  - Possible: Labeling of objects in visual scenes
  - Not yet possible: Predict what happens next, common-sense understanding of visual scenes
- Speech analysis
  - Possible: Speech-to-text and text-to-speech in many languages
  - Not yet possible: Common-sense understanding of auditory scenes
- Natural language processing
  - Possible: Adequate translation and Q&A (chatbot)
  - Not yet possible: Semantic/abstraction understanding of language, human-level dialog (Turing test)
- Robotics
  - Possible: Industrial programmed robots
  - Not yet possible: Robots that interact with humans and are autonomous.

Questions?