A brief introduction to AI and Deep Learning

Marc Schoenauer
Equipe TAU, INRIA, LRI, UP-Sud et CNRS,
Université Paris-Saclay

Université Franco Italienne, Università Italo Francese
21 Nov. 2018

A brief introduction to AI and Deep Learning

- History and DefinitionS*
- Some Recent Successes*
- Deep Learning
- Limits and Challenges

(*) Many thanks to Bertrand Braunschweig

A brief introduction to AI and Deep Learning

- History and DefinitionS
- Some Recent Successes
- Deep Learning
- Limits and Challenges

History

AI is a recent invention

History

AI is a recent invention

FALSE

History

Before 1956, some visions : Alan Turing, formal neurons, robots — AI as a mean

1956: Dartmouth workshop, first occurence of the term AI — AI as a goal

196x: *Problem solving*, games, natural langage

1968: 2001 a space odyssey, HAL

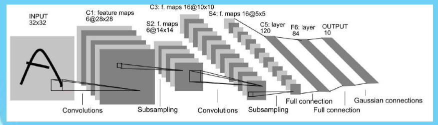
1969: *Perceptrons* (Minsky-Papert), kills research on NNs

1973: Lighthill Report, first AI Winter

198x: Prolog+FGCS; Experts Systems; Checkers (from Samuel to Chinook)

199x: Second AI Winter, but Deep Blue (chess) and first convolutional networks (CNNs)

2000: first Web applications (data)



2010: Deep learning (triumph of CNNs, AlphaGO, ...)

2018+: toward a third AI Winter?

History

Before 1956, some visions : Alan Turing, formal neurons, robots

AI as a mean

Can Machines Think?

The problem is mainly one of programming. [...] brain estimates: 10^{10} to 10^{15} bits. [...] I can produce about a thousand digits of programme lines a day, so that about sixty workers, working steadily through the fifty years, might accomplish the job, if nothing went into the wastepaper basket. Some more expeditious method seems desirable.

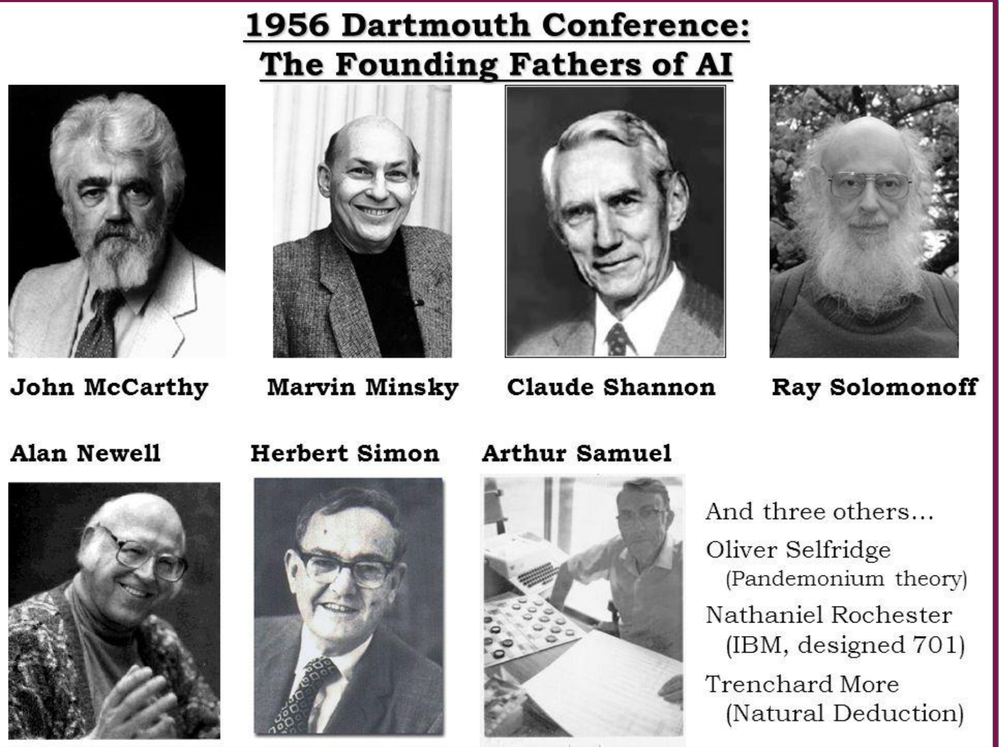
How?

by (...) mimicking education, we should hope to modify the machine until it could be relied on to produce definite reactions to certain commands.
One could carry through the organization of an intelligent machine with only two interfering inputs, one for pleasure or reward, and the other for pain or punishment.

History

Before 1956, some visions : Alan Turing, formal neurons, robots — AI as a mean

1956: Dartmouth workshop, first occurence of the term AI — AI as a goal



We propose a study of artificial intelligence [..]. The study is to proceed on the basis of the conjecture that every aspect of learning or any other feature of intelligence can in principle be so precisely described that a machine can be made to simulate it.

The vision : reasoning is a sequence of logical operations that a computer can reproduce

Goal : A *General Problem Solver*
(aka 2000+ : Artificial General Intelligence)

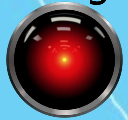
History

Before 1956, some visions : Alan Turing, formal neurons, robots — AI as a mean

1956: Dartmouth workshop, first occurence of the term AI — AI as a goal

196x: *Problem solving*, **games**, natural langage

1968: 2001 a space odyssey, HAL



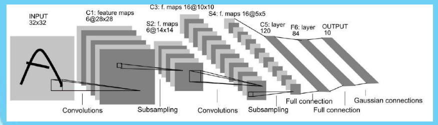
1969: *Perceptrons* (Minsky-Papert), kills research on NNs

1973: Lighthill Report, first AI Winter

198x: Prolog+FGCS; Experts Systems; **Checkers** (from Samuel to Chinook)

199x: Second AI Winter, but Deep Blue (**chess**) and first convolutional networks (CNNs)

2000: first Web applications (data)



2010: Deep learning (triumph of CNNs, Alpha**GO**, ...)

2018+: toward a third AI Winter?

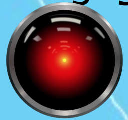
History

Before 1956, some visions : Alan Turing, formal neurons, robots — AI as a mean

1956: Dartmouth workshop, first occurence of the term AI — AI as a goal

196x: *Problem solving*, games, natural langage

1968: 2001 a space odyssey, HAL



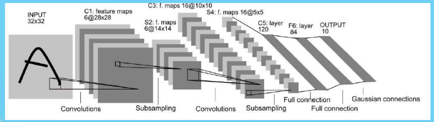
1969: *Perceptrons* (Minsky-Papert), **kills research** on NNs

1973: Lighthill Report, **first AI Winter**

198x: Prolog+FGCS; Experts Systems; Checkers (from Samuel to Chinook)

199x: **Second AI Winter**, but Deep Blue (chess) and first convolutional networks (CNNs)

2000: first Web applications (data)



2010: Deep learning (triumph of CNNs, AlphaGO, ...)

2018+: toward a third AI Winter?

Definition ?

Have machines that accomplish tasks related to (human) intelligence - possibly better than humans.

Definition ?

Have machines that accomplish tasks related to (human) intelligence - possibly better than humans.

BOF…

Definition ?

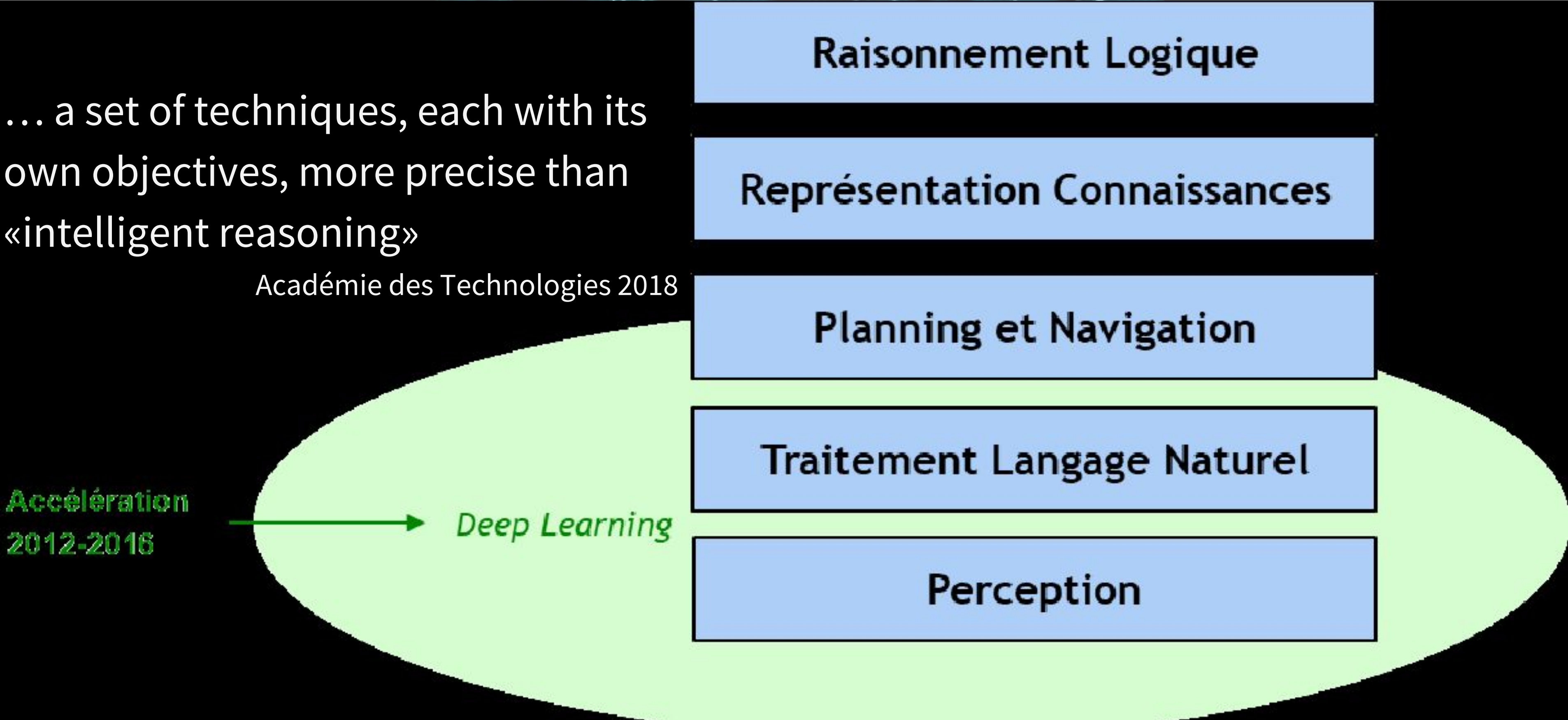
Have machines that accomplish tasks no machine ever did

- Jean-Louis Laurière, 80s
- Philippe Kahn, late 80s
- Gérard Sabah, 2017
  (rapport de l’OPECST)

Definition ?

… a set of techniques, each with its own objectives, more precise than «intelligent reasoning»

Académie des Technologies 2018



A brief introduction to AI and Deep Learning

- History and DefinitionS
- Some Recent Successes
- Deep Learning
- Limits and Challenges

Autonomy and Robotics

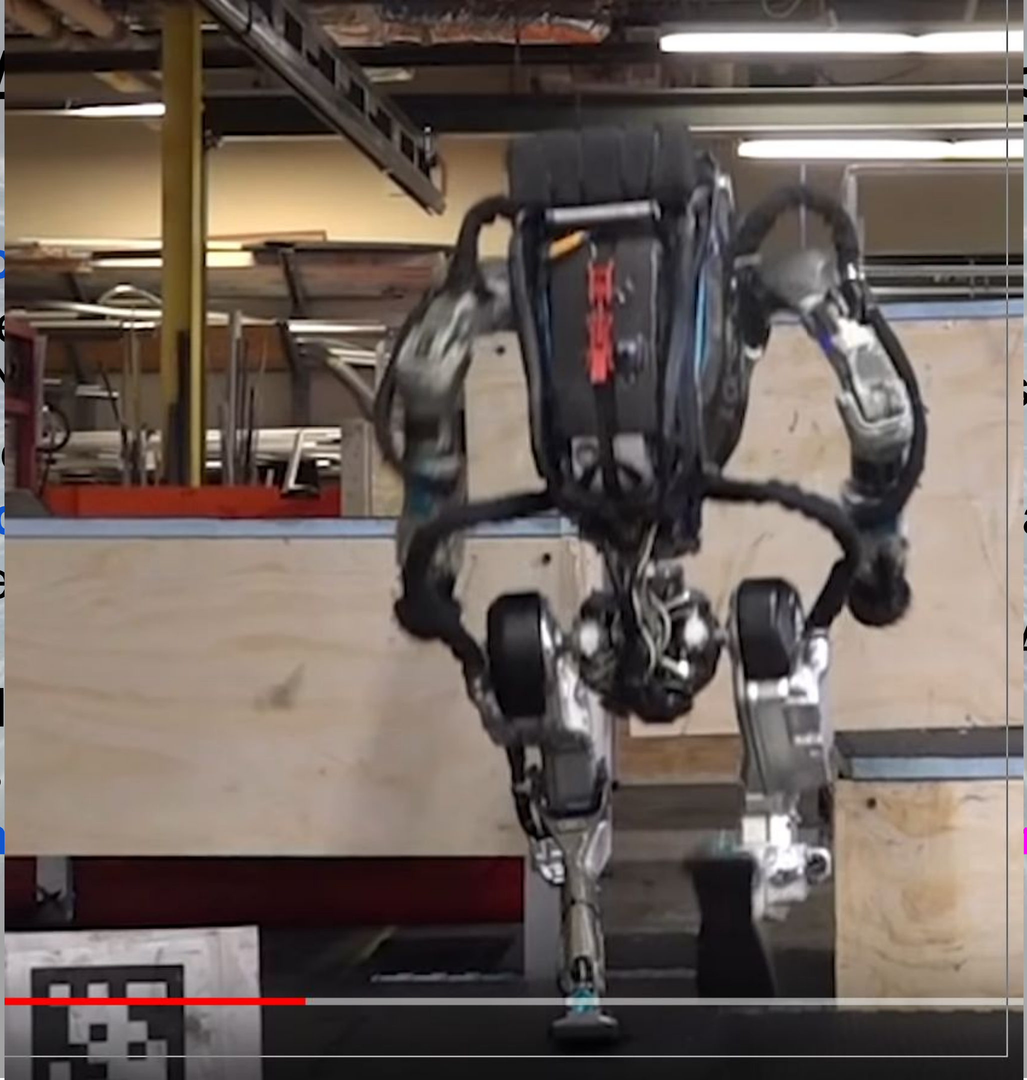
- *DARPA Autonomous Vehicle Challenge 2004-07
  - in the désert, then in urban context
- LeNet (Deep Neural Network) outperforms all challengers from Computer Vision in image recognition 2012-
- *DARPA Rescue Challenge robots who drive, walk in chaotic context, climb stairs, repair broken machines, etc 2015
- **Psibernetix shoots down (in simulation :-) the best US Air Force pilots
  - genetic algorithms and fuzzy logic ... on a Raspberry Pi! 2015
- Intel bought Israeli company MobilEye for 15 billions 2017
- **Boston Dynamics robots better and better performing 1998-

A … s

- *DARPA Auto …
  - in the dése …
- LeNet (Deep N … s from Computer Vision in image …
- *DARPA Resc … haotic context, climb stairs, re …
- **Psibernetix … Air Force pilots
  - genetic alg … Pi! 2015
- Intel bought Isr …
- **Boston Dyna … 1998-

A S

- *DARPA Auto
  - in the dése
- LeNet (Deep N s from Computer Vision in image
- *DARPA Resc aotic context, climb stairs, re
- **Psibernetix Air Force pilots
  - genetic al Pi! 2015
- Intel bought Is
- **Boston Dyn 1998-



Games

- **IBM Watson** beats best human players at Jeopardy 2011
  - NLP + web search + evaluation, 3 seconds on HPC
- **Deepmind** human performances on some (not all) Atari video games with Deep Reinforcement Learning 2013
  - Input: pixels; Output: joystick
- **Deepmind AlphaGo** beats World Champion of GO with a mix of Supervised and Reinforcement Learning 2016-17
- **Deepmind AlphaZéro** beats AlphaGo 100-0 using only Deep Reinforcement Learning and self-plays 2018
  - about 2 stones ahead of best human
  - AlphaZero can also be trained for other games (e.g., chess)
- *__Libratus__ crushes the best Poker players of the world 2017
  - Reinforcement Learning and Bayesian techniques

NLP and and disability support

- **Microsoft Skype Translator** translates several languages in real time with Deep Learning. Similar performances for **Google Translate**, **Pilot**, ...
- **Apple Siri, Microsoft Cortana, Amazon Alexa** personal assistants use Speech Recognition and (some) Automated Reasoning
- ****Google Knowledge Graph** uses semantics to better structure the results of queries
- **Microsoft** translates from Chinese to English as good as human translators
  - with a double Deep Neural Network
- **Ava, RogerVoice** help deaths and hearing-impeached (subtitling, telephone,...)
- **Facebook** can label photos, and describe them to blind people

A brief introduction to AI and Deep Learning

- History and DefinitionS
- Some Recent Successes
- Deep Learning
- Limits and Challenges

Machine Learning

Artificial Intelligence is (Deep) Machine Learning

Machine Learning

Artificial Intelligence is (Deep) Machine Learning

FALSE

Machine Learning

Artificial Intelligence is (Deep) Machine Learning

although ...

What has changed :

- Data Deluge
- Moore law — or continuation
- New algorithms — or better understanding of old ones

Machine Learning

Learning from examples — recognition tasks

- Supervised — all examples are labelled
- Semi-supervised — some examples are labelled
- Unsupervised — no example is labelled

Reinforcement Learning — sequential decision making

Machine Learning

Learning from examples

- Supervised
- Semi-supervised
- Unsupervised

recognition tasks

all examples are labelled

some examples are labelled

no example is labelled

Reinforcement Learning

sequential decision making

Supervised Learning

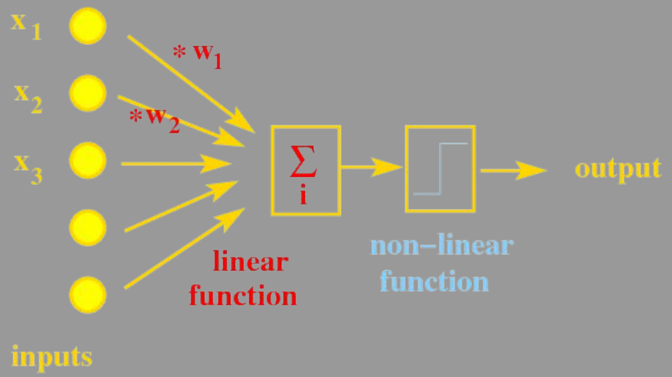
A toy case-study

- One example = (x_1,x_2) + label (red or blue here)
- Goal: a **model** (function of x_1,x_2) that separates the labels
- and allows to correctly label future unlabelled example from (x_1,x_2)

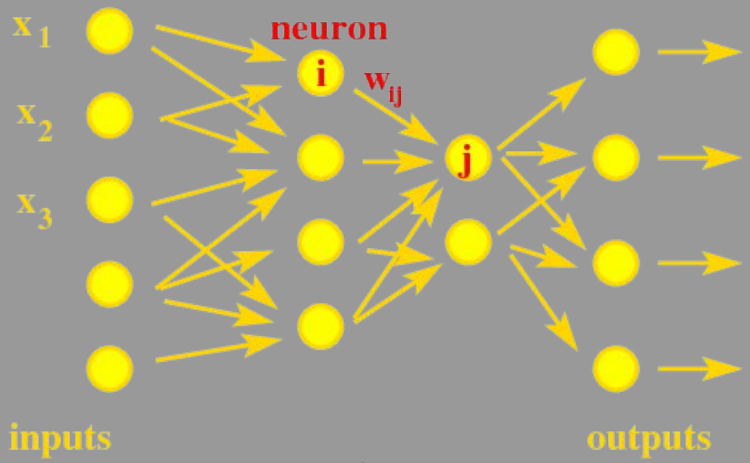
Supervised Learning

A zoology of models

- Polynoms
- Bayésiens Networks
- Decision trees and Random Forrests
- Support Vector Machine (kernel machines)
- **Artificial Neural Networks**



One neuron



A network of neurons

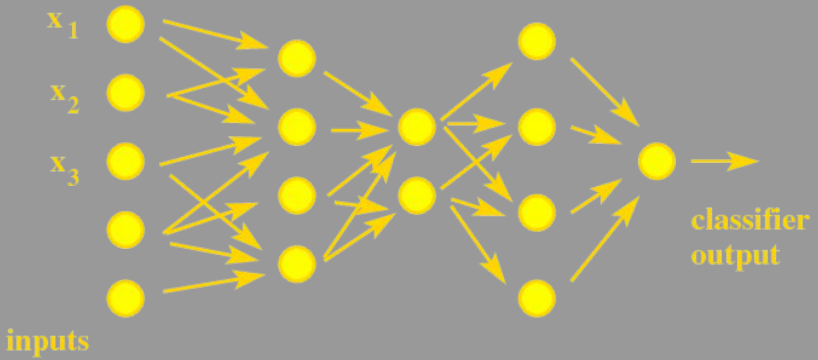
Parameters are the **weights w_{ij}**

Deep Neural Networks

Learning Phase

Back-propagation from the 60s

- Present the examples 1 by 1
  - or mini-batches by mini-batches
- Compute the corresponding error
  - difference between network output and label
- Adjust the weights w_{ij}
  - toward a decrease of the error
- Loop

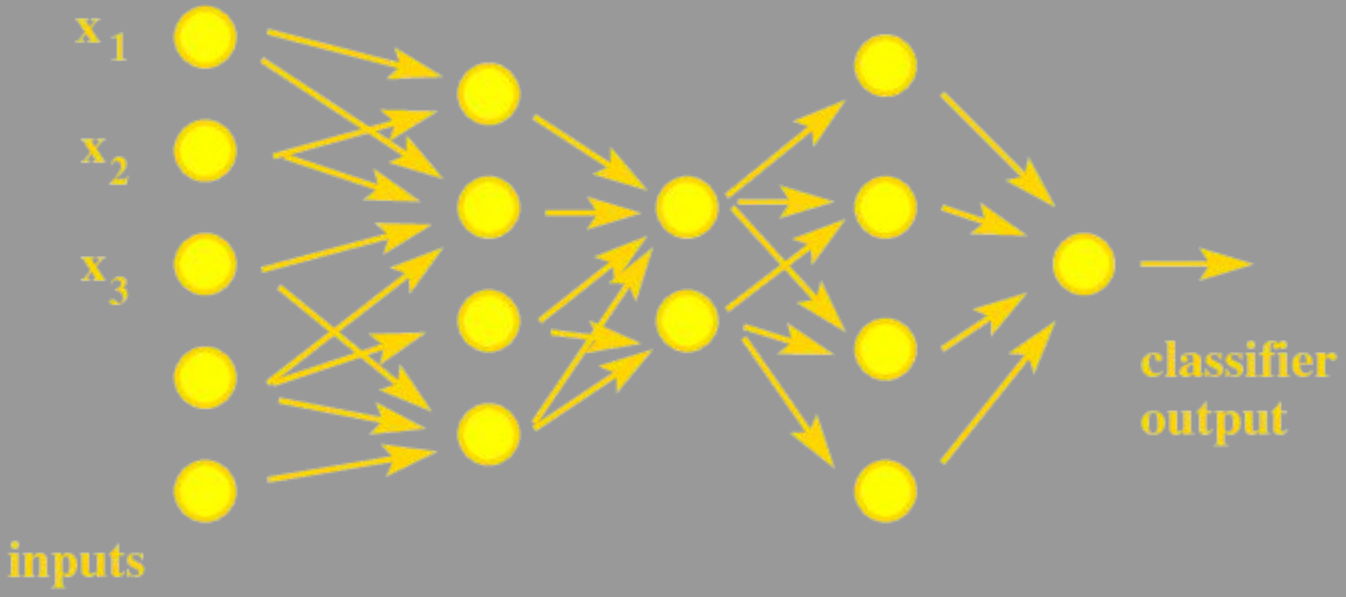


Recognition Phase

Present an unlabelled example, the output of the network is the predicted label

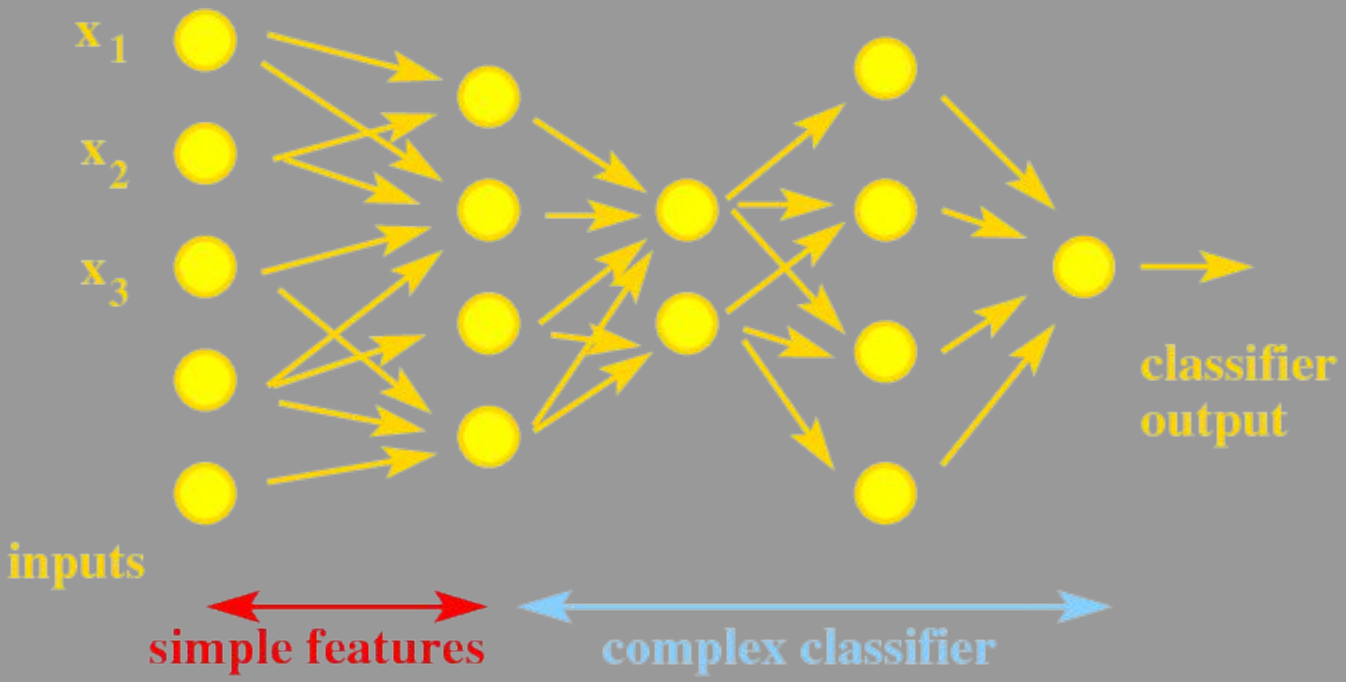
Deep Neural Networks

A Deep (layered) Neural Network is
a sequence of **representations of the data**



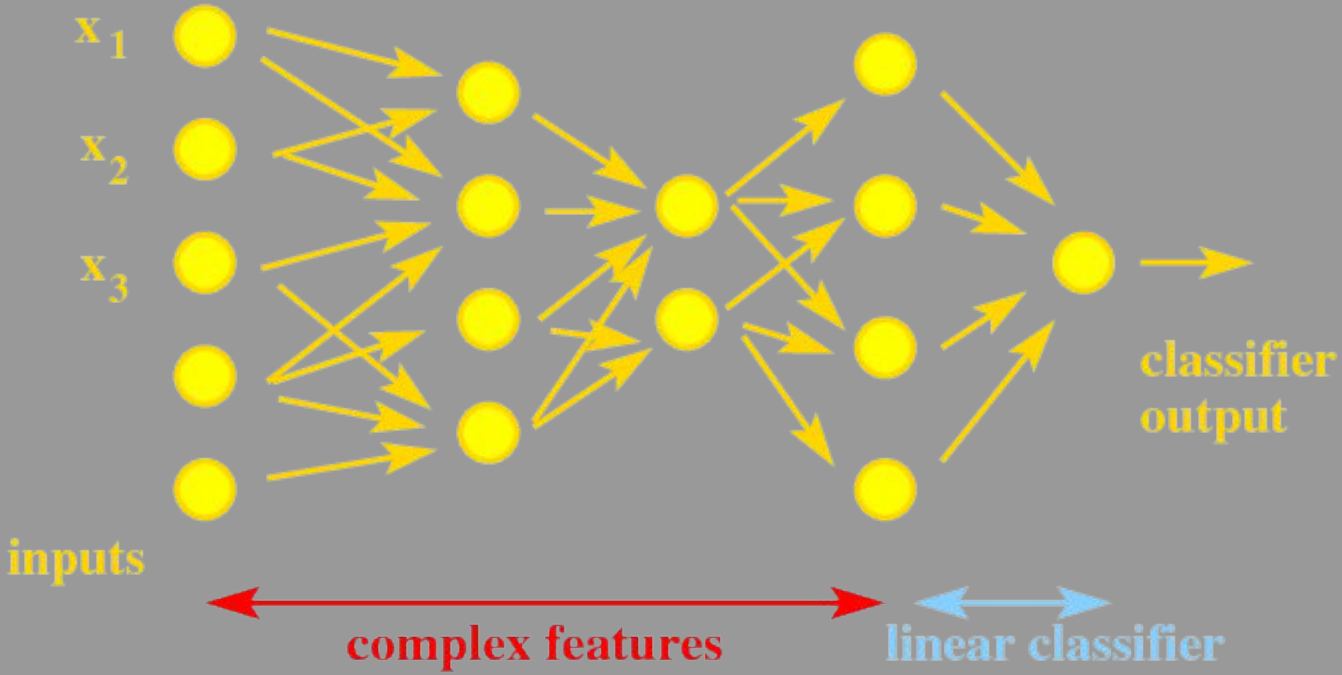
Deep Neural Networks

A Deep (layered) Neural Network is
a sequence of **representations of the data**



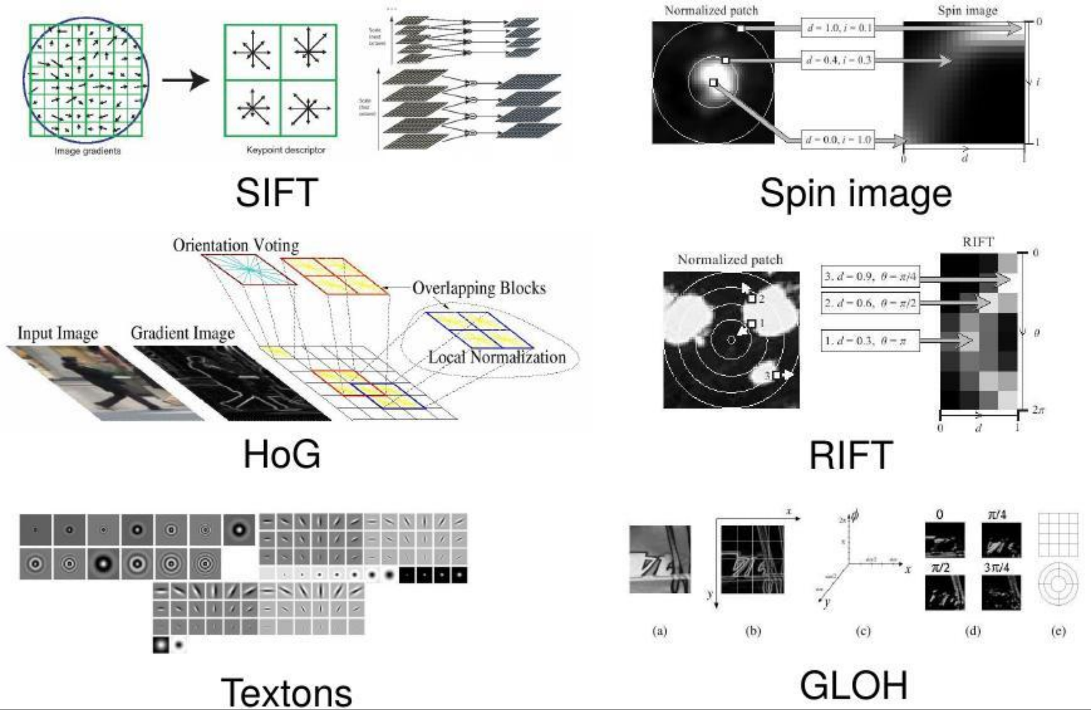
Deep Neural Networks

A Deep (layered) Neural Network is
a sequence of **representations of the data**



Good Old Computer Vision

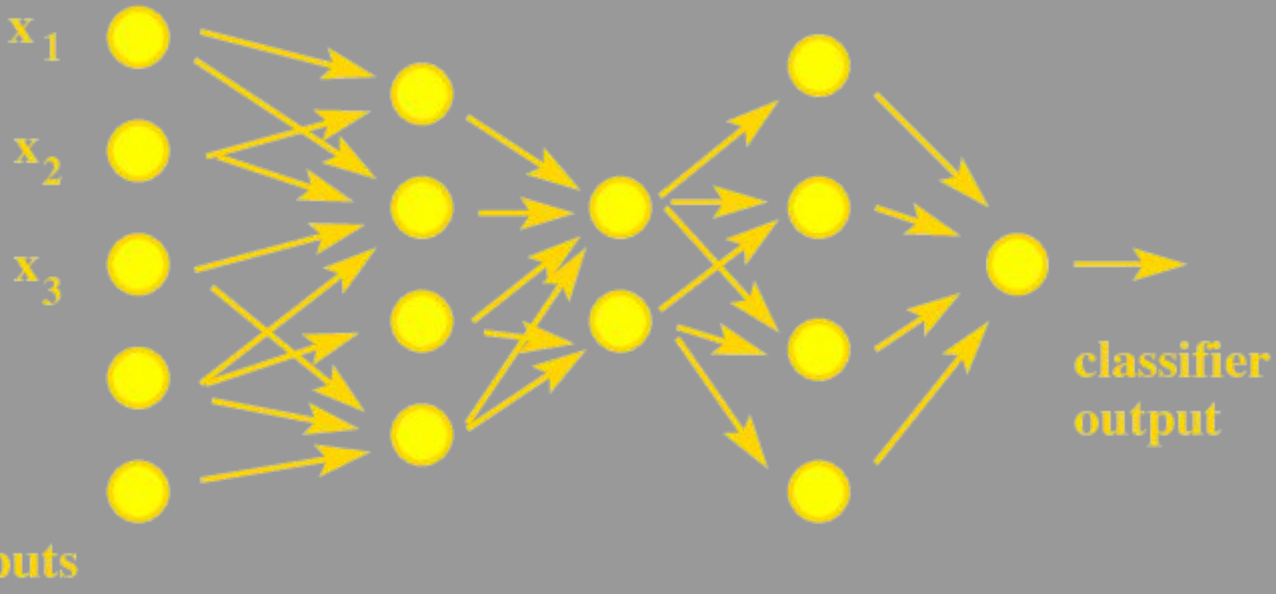
Hand-made features → Learned Classifier → CAT



End-to-end Learning

Features and Classifier are learned together

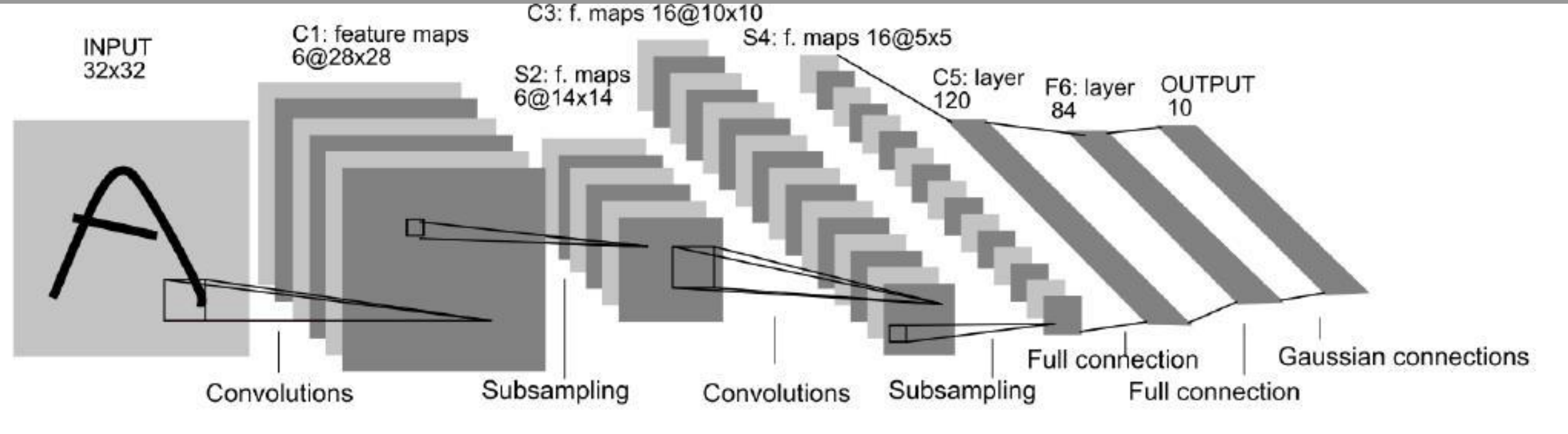
CAT



Convolutional Networks

Features and Classifier are learned together

CHAT

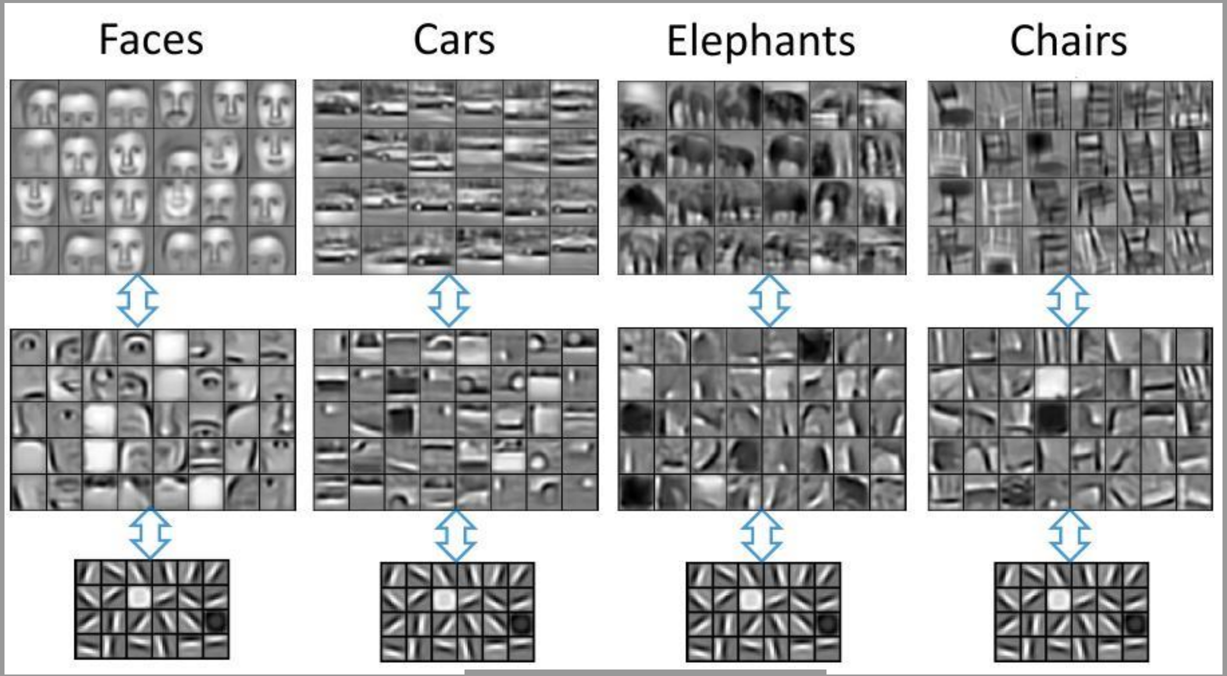


LeNet, LeCun et al., 1998

Convolutional Networks

Features and Classifier are learned together

CHAT

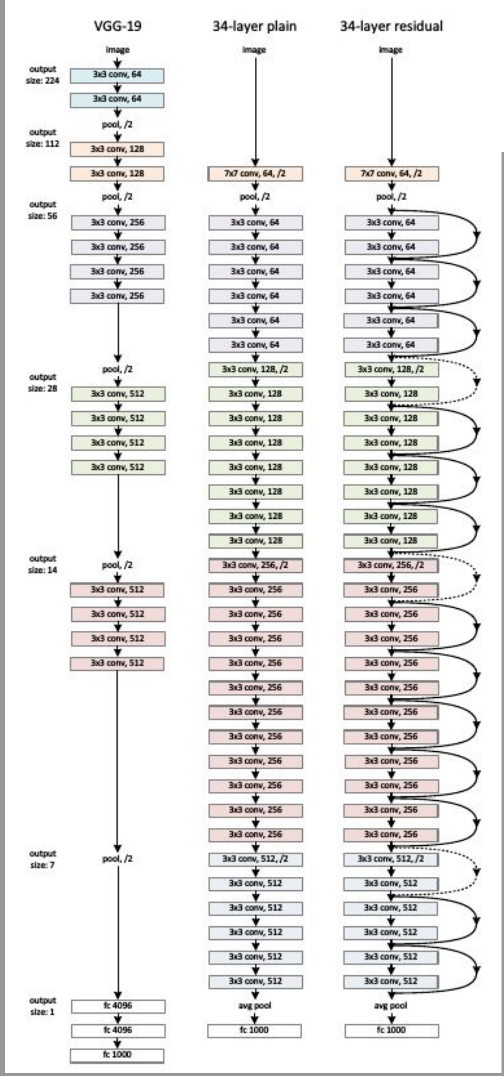


Learned Features

State-of-the-art

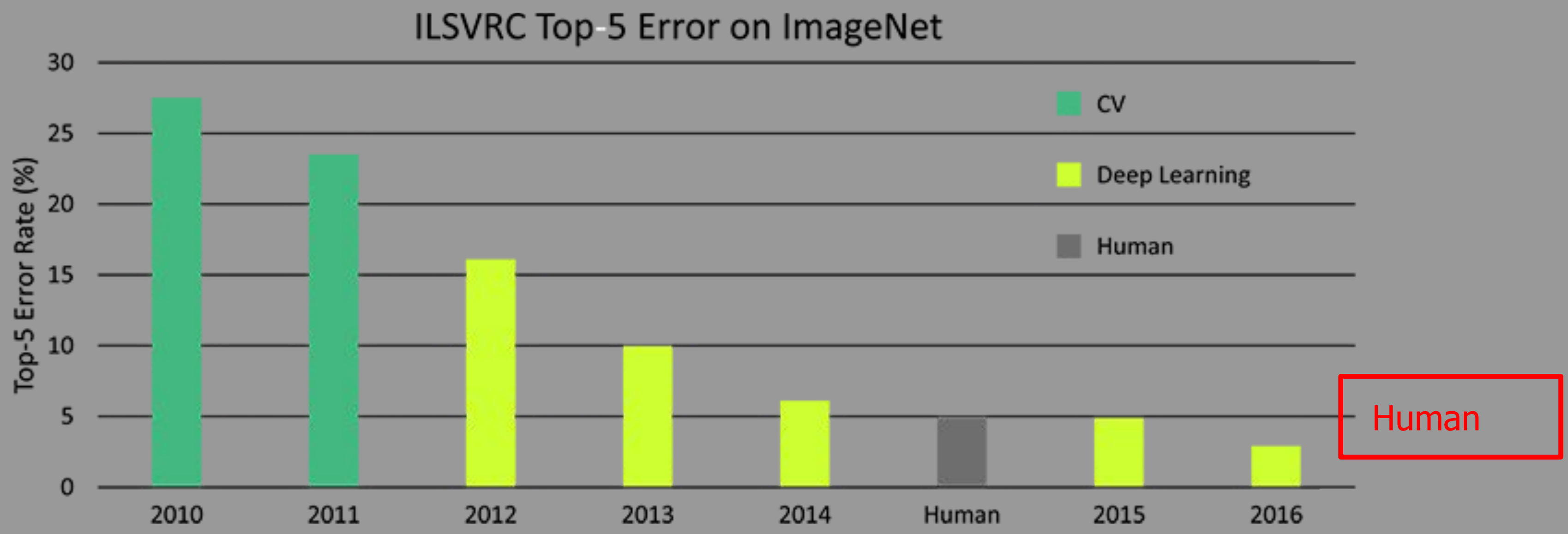
- Many datasets available
  - ImageNet : 14+ M examples, 1000 classes
- (pre-trained) networks with numerous layers
  - up to 152 !
- Millions to billions weights
  - hundreds of GPU mandatory for learning
- Several ‘goodies’
  - Dropout, residual layers, ensembles, ...
- Error on Imagenet: **3.75%** (2016)

He et al., 2015



Deep Supervised Learning

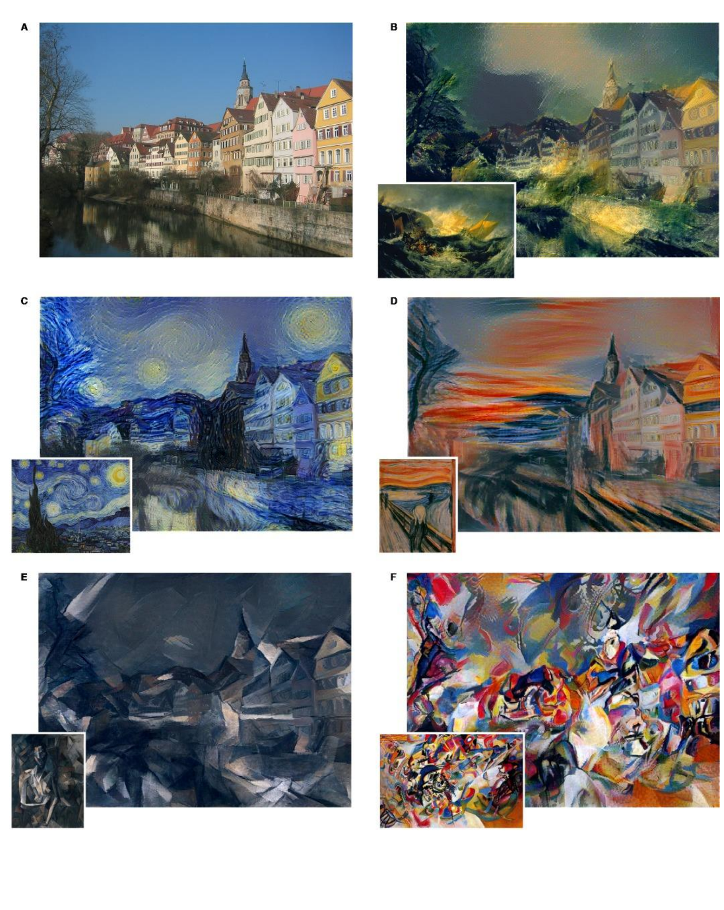
Better than human learning



Deep Supervised Learning

- Outstanding performances
  … in well-defined domains
  - Image recognition
  - Action identification in videos
  - Natural Language Processing
  - Automatic translation
  - Image captioning
- Many unexpected applications, e.g.,
  - domain transfer (DANNs) (see next talk)
  - generative models (GANs)
- Above all, **discovery of latent representations**

But ...



A brief introduction to AI and Deep Learning

- History and DefinitionS
- Some Recent Successes
- Deep Learning
- Limits and Challenges

Limits and Challenges

Beyond performances

- Small Data transfer learning, data augmentation
- Cost
- Validation and certification
- Interpretability and explainability
- Causality
- Transparency and Fairness
- Toward Trustable Good AI

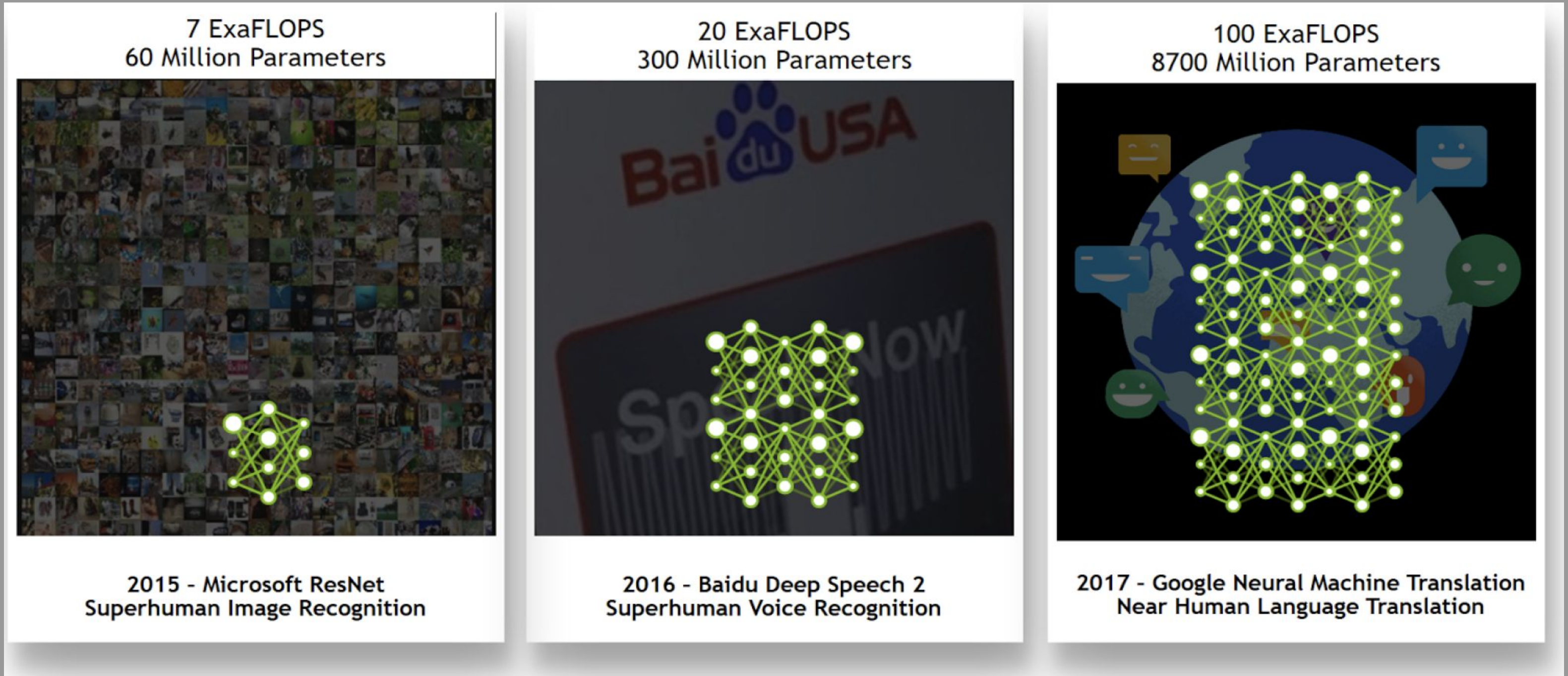
Limits and Challenges

Beyond performances

- Small Data
- Cost
- Validation and certification
- Interpretability and explainability
- Causality
- Transparency and Fairness
- Toward Trustable Good AI

Huge computational/energy cost

- Loads of data
- Tons of weights



Meta-cost

\+ high number of hyperparameters to tune

- Cost function
- Topology of the network
  - nblayers, nb neurons, residual or not residual, ...
- Activation function
- Batch size
- Optimizer
  - and its parameters (e.g., learning rate)
- Initialization
- Dropout or not dropout
- etc

Empirical rules, or meta-optimization

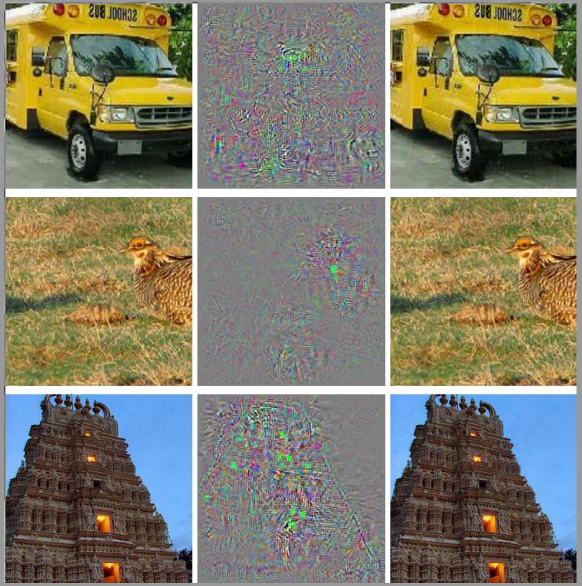
Limits and Challenges

Beyond performances

- Small Data
- Cost
- Validation and certification
- Interpretability and explainability
- Causality
- Transparency and Fairness
- Toward Trustable Good AI

Robustness

To noise at test time : adversarial examples



Szegedy et al.,2014

Athalye et al. 2017

Robustness

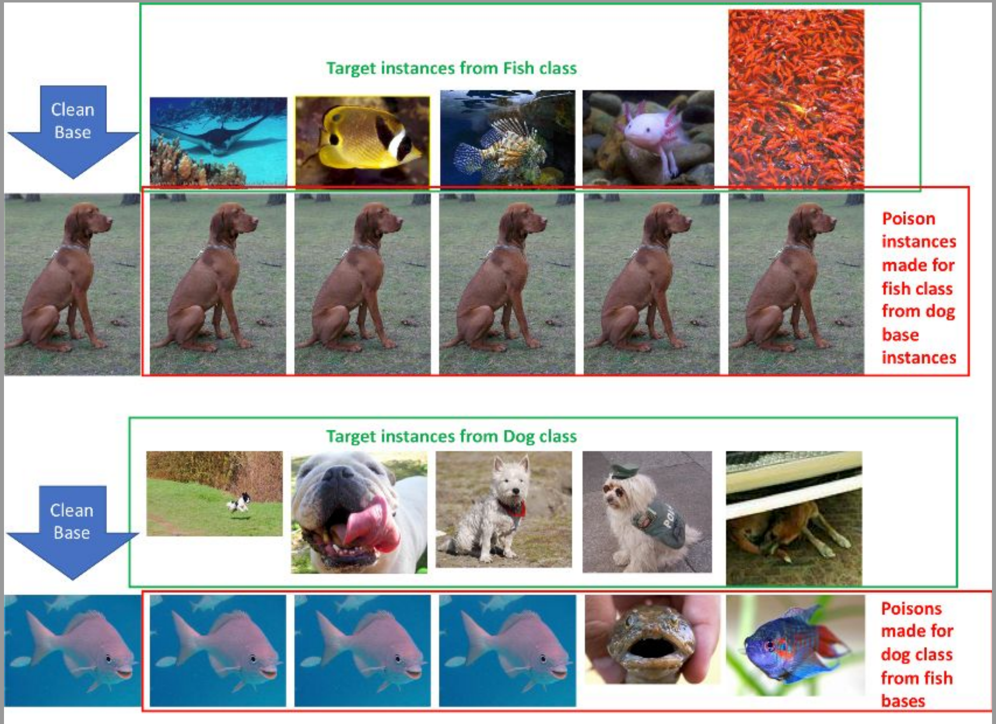
To unseen contexts

A cow doesn't go to the beach

Bottou et al., 2017

Robustness

to poisoned learning examples



Shafahi et al., 2018

Validation and certification

- An experimental science
- No formal validation of learned models
- Completeness issue for statistical validation

- Need to validate the training data
  - Traceability

regulations

- Guaranteed bounds
- Toward formal proofs for AI?

e.g., Asimov's robotic laws

e.g., Mirman et al., 2018

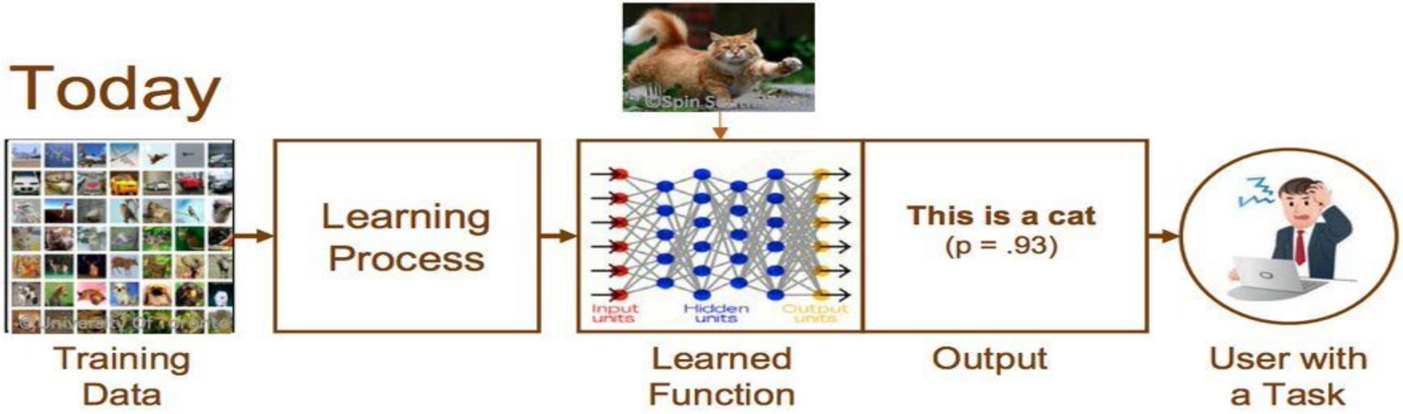
Limits and Challenges

Beyond performances

- Small Data
- Cost
- Validation and certification
- Interpretability and explainability
- Causality
- Transparency and Fairness
- Toward Trustable Good AI

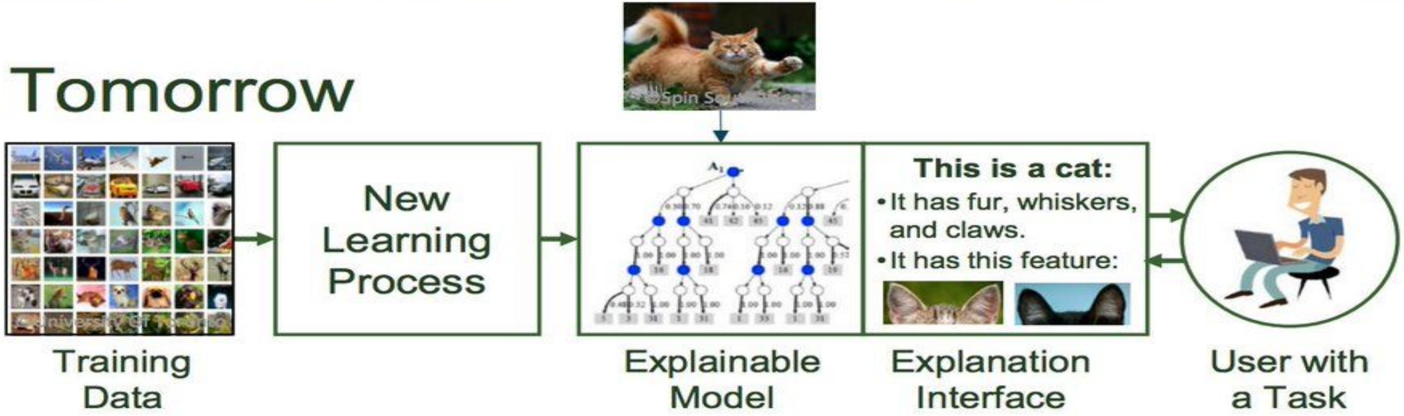
Explainable AI – What Are We Trying To Do?

Today



- Why did you do that?
- Why not something else?
- When do you succeed?
- When do you fail?
- When can I trust you?
- How do I correct an error?

Tomorrow



- I understand why
- I understand why not
- I know when you'll succeed
- I know when you'll fail
- I know when to trust you
- I know why you erred

Interpretability and explainability

Learned models are black boxes

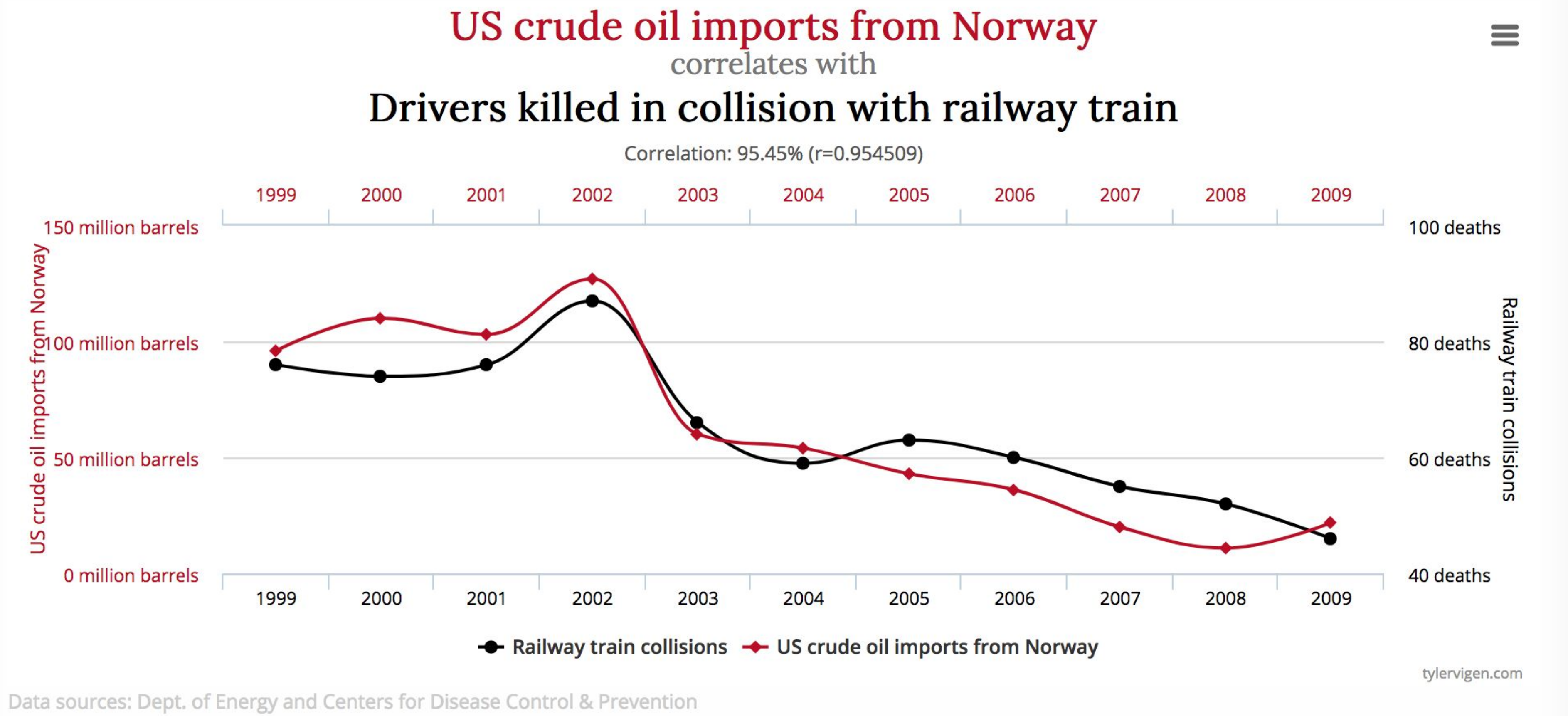
- Ill-defined and subjectives concepts
- Depends on the type of model
  - moderately: decision trees are ok
  - … not random forests
- **Debate**
  - How much are you ready to lose in accuracy?
  - Cite the nearest known examples e.g., influence fns, Koh & Liang, 2017
  - Well, we trust our doctor, don't we …
- Symbolic to the rescue?

Limits and Challenges

Beyond performances

- Small Data
- Cost
- Validation and certification
- Interpretability and explainability
- Causality
- Transparency and Fairness
- Toward Trustable Good AI

Correlation vs causality



Correlation vs causality

Supervised learning doesn't make a difference

- "What if" scenarios needed for decision making
- Causality usually from common sense
- Difficult to learn from data

- ~OK for pairs of variables (several challenges 2008+)
- Still an open question in general

Limits and Challenges

Beyond performances

- Small Data
- Cost
- Validation and certification
- Interpretability and explainability
- Causality
- Transparency and Fairness
- Toward Trustable Good AI

Transparency and Fairness

Mandatory for societal acceptance

- Open Source not sufficient
    - Open Data
    - controlled experiments
- Recognized labels
- Discrimination Impact Assessment?

auditability by law

e.g., FDU, Maathics (Toulouse)

See also the **TransAlgo** platform

Toward Trustable Good AI

Combat fear-mongering

- Scientific and legal advances
  - Human in control
  - Accountability

- Ethical rules — enforce European humanist values
  - Design
    - Public debate, CCNE-bis, … — CERNA, COERLE, …
  - Control
    - Citizen crowd control, independent institution, …

Without trust, societal AI winter ahead

Collaboration, not Competition

Questions ?