EE6363

Deep Learning

Spring, 2021

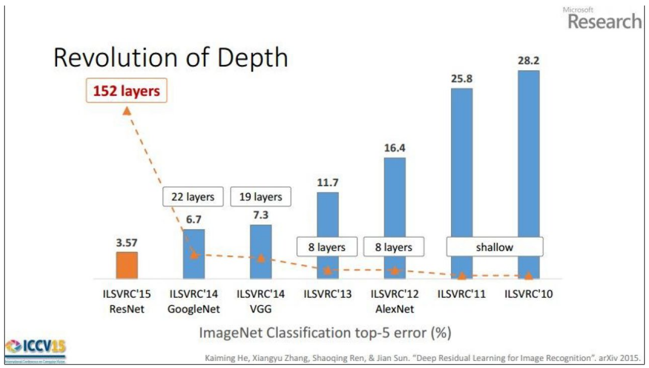
Lecture 1– History and Background of Deep Learning

Instructor: **Yufei Huang**, PhD, *University of Texas at San Antonio*

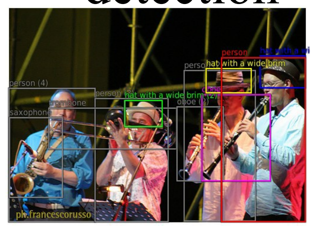
UTSA

Deep learning applications

Image recognition



Object detection



Groundtruth:
person
hat with a wide brim
hat with a wide brim (2)
hat with a wide brim (3)
oboe
oboe (2)
saxophone
trombone
person (2)
person (3)
person (4)

Video game

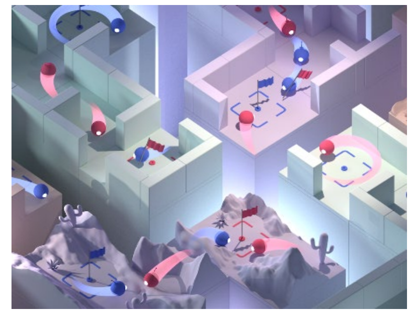
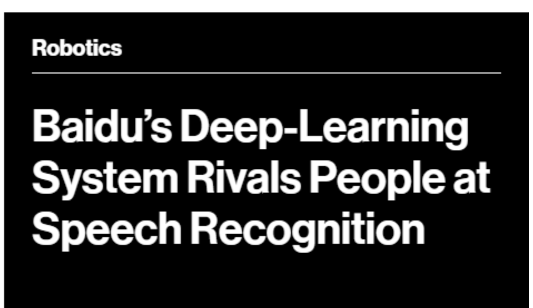


Image generation

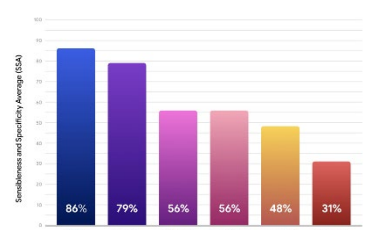
Speech recognition

MIT Technology Review



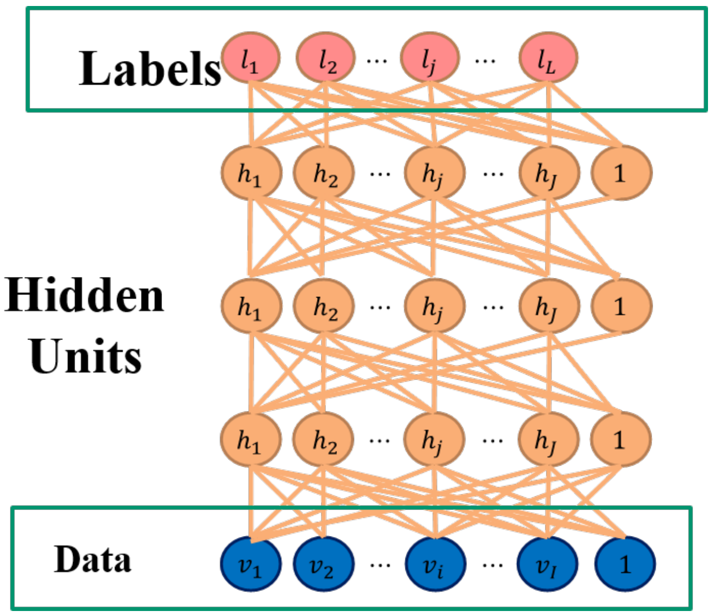
AI GO

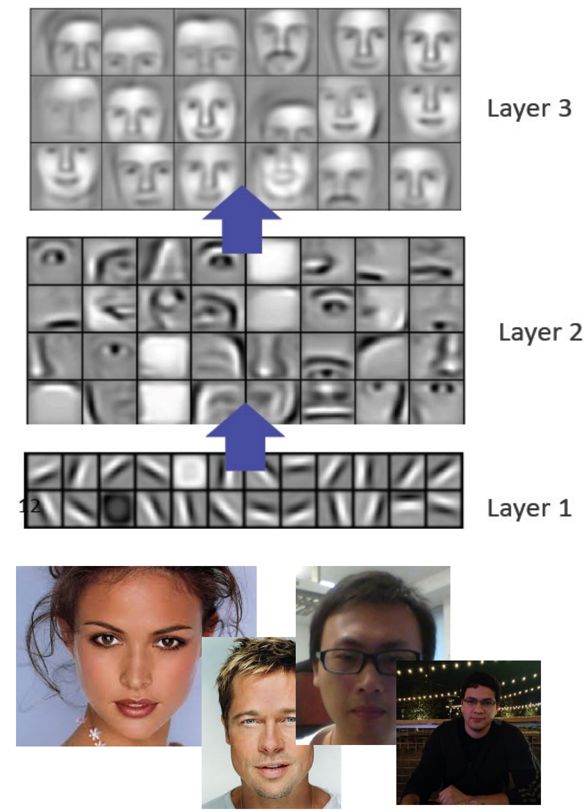
Meena



Medical imaging

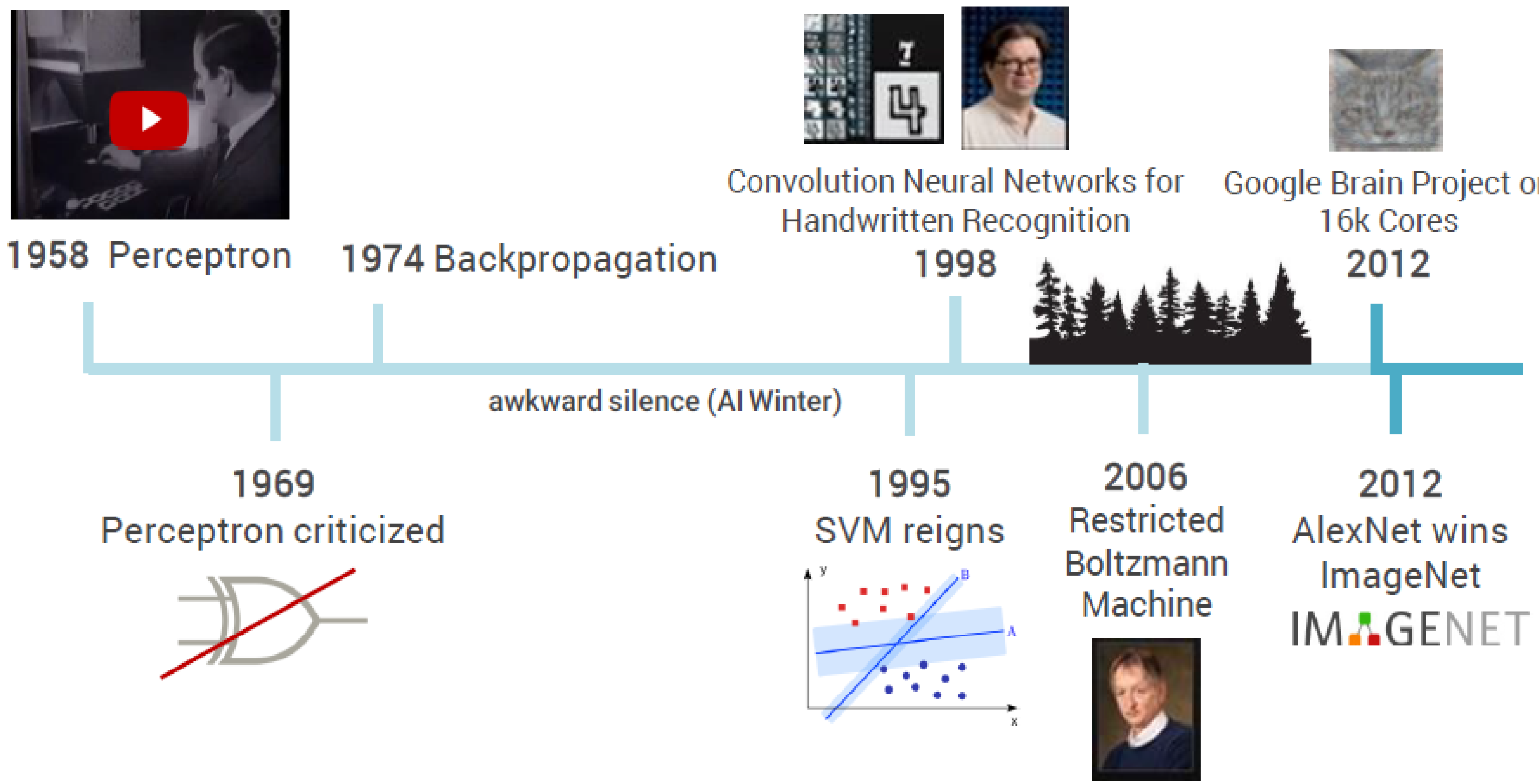
What is Deep Learning? (Deep Neural Networks)



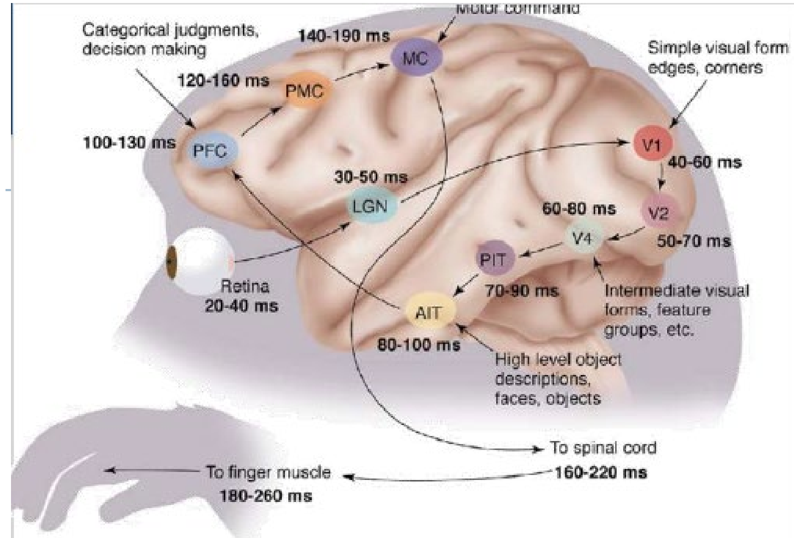


Learn intricate representation of complex data from a huge amount of data

History of DL

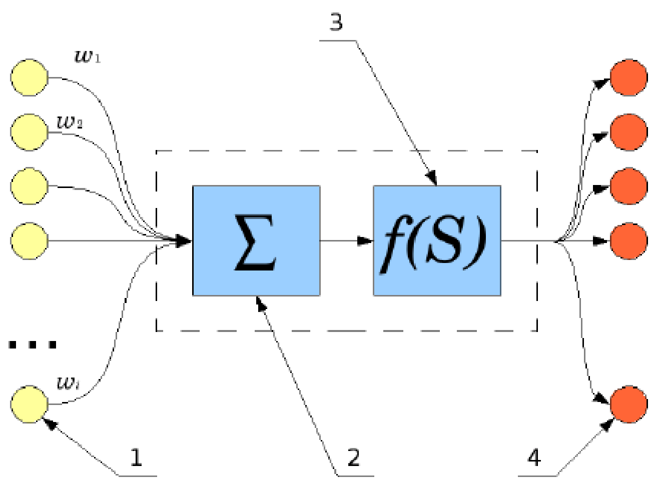


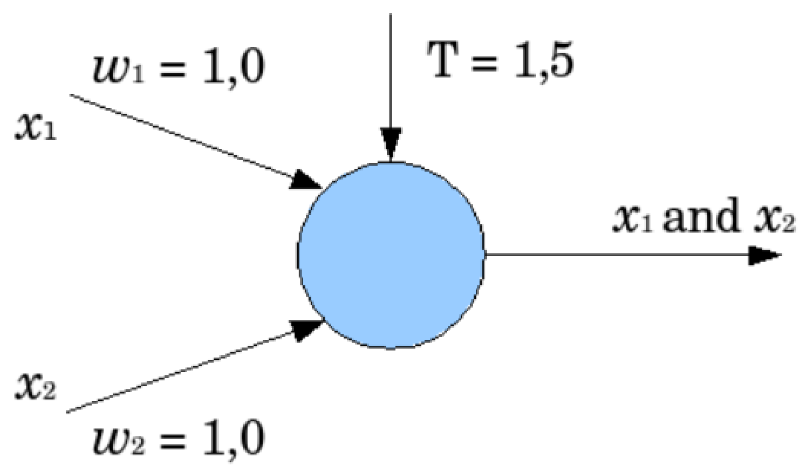
History of DL – Artificial Neuron, 1943



McCulloch and Pitts

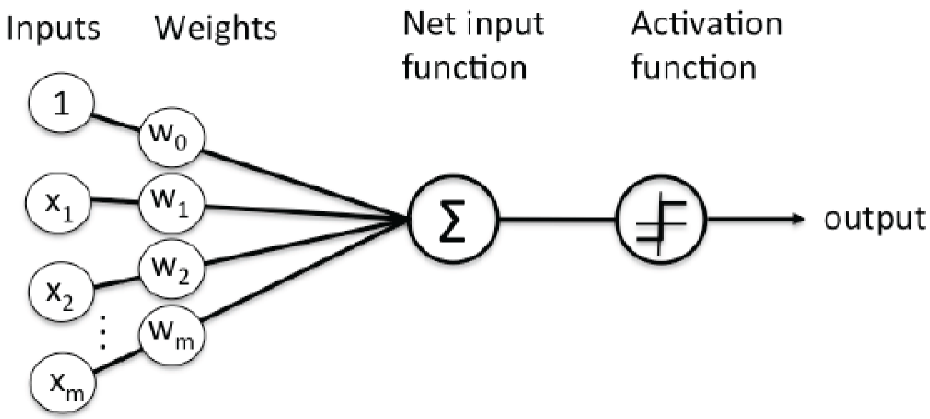
"A Logical Calculus of the Ideas Immanent in Nervous Activity"

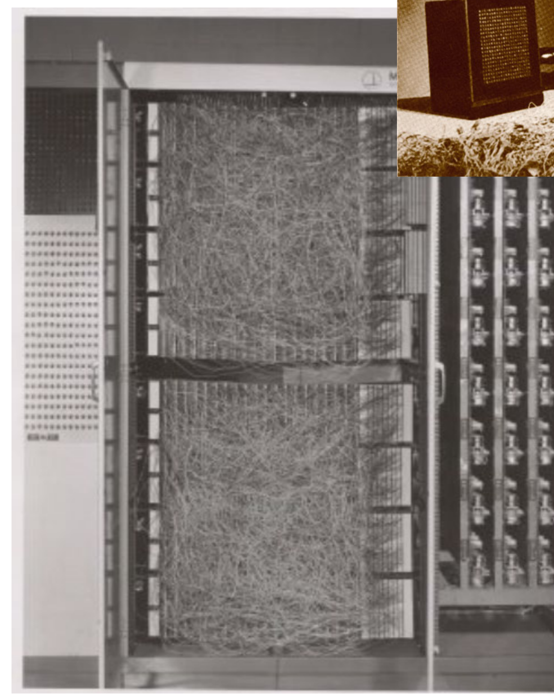




History of DL – Perceptron, 1957

1957 Frank Rosenblatt

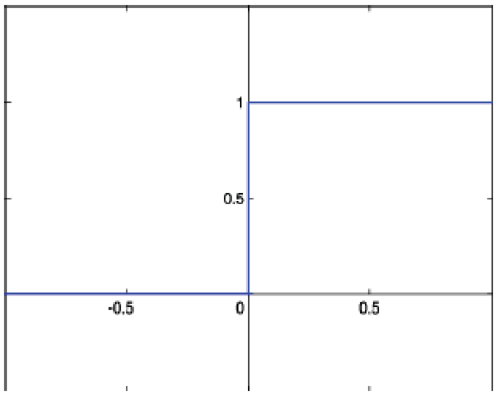


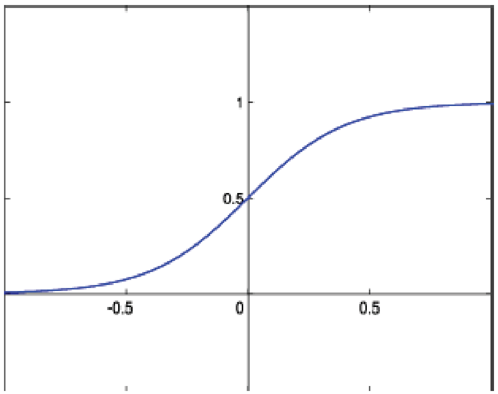


"[The Perceptron is] the embryo of an electronic computer that [the Navy] expects will be able to walk, talk, see, write, reproduce itself and be conscious of its existence."

THE NEW YORK TIMES

History of DL – Backpropagation, 1974





- Measure how **small changes** in weights **affect** output
- Can apply NN to **regression**

(1974) 1986

(Werbos) Rumelhart, Hinton, Williams

"Learning representations by back-propagating errors" (Nature)

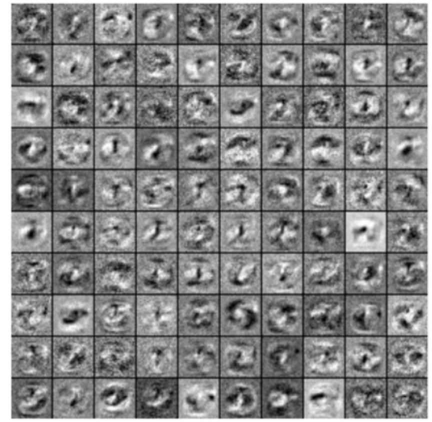
- **Multilayer** neural networks, etc.

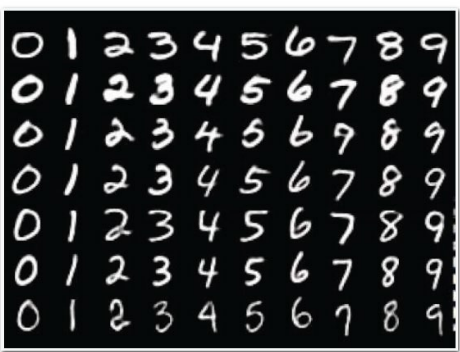
UTSA

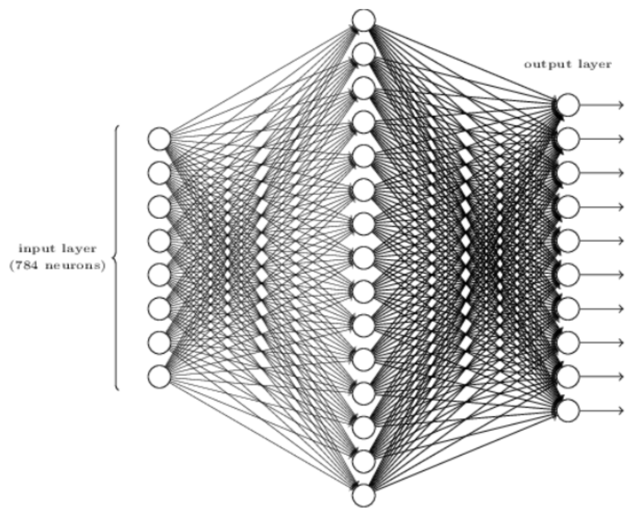
History of DL – NN winter; late 1990 – early 2000

- **Not enough data** (datasets were 1000 times too small)
- **Computers were too slow** (1,000,000 times)
- Not enough attention to network initialization
- Wrong non-linearity

History of DL – Restricted Boltzmann machine 2005



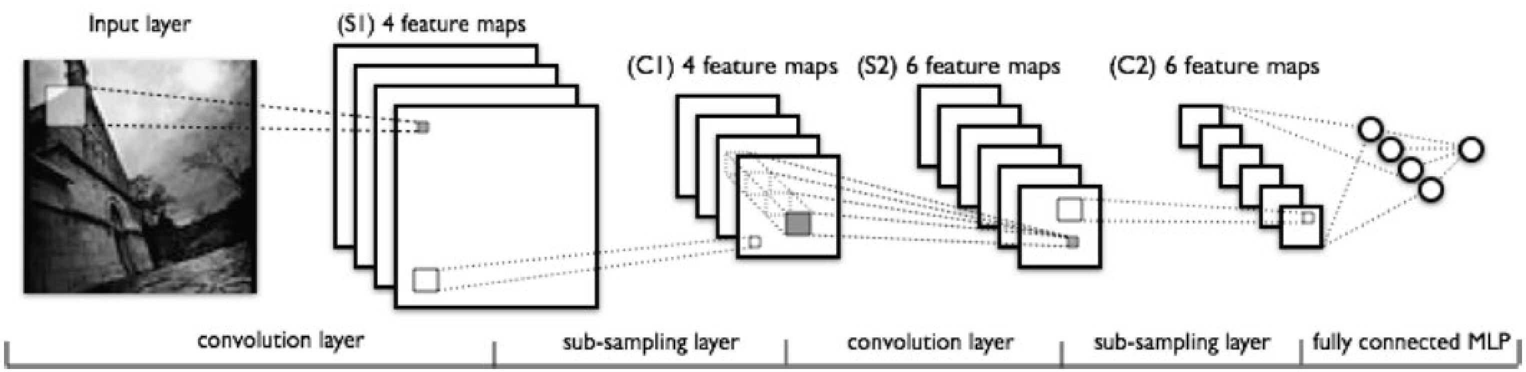




98.2% on the MNIST test set

History of DL – Convolution 1998, 2010

Yann LeCun, NYU & Facebook



99.50% on the MNIST test set

CURRENT BEST: 99.77% by committee of 35 conv. nets

History of DL – ImageNet 2012 -

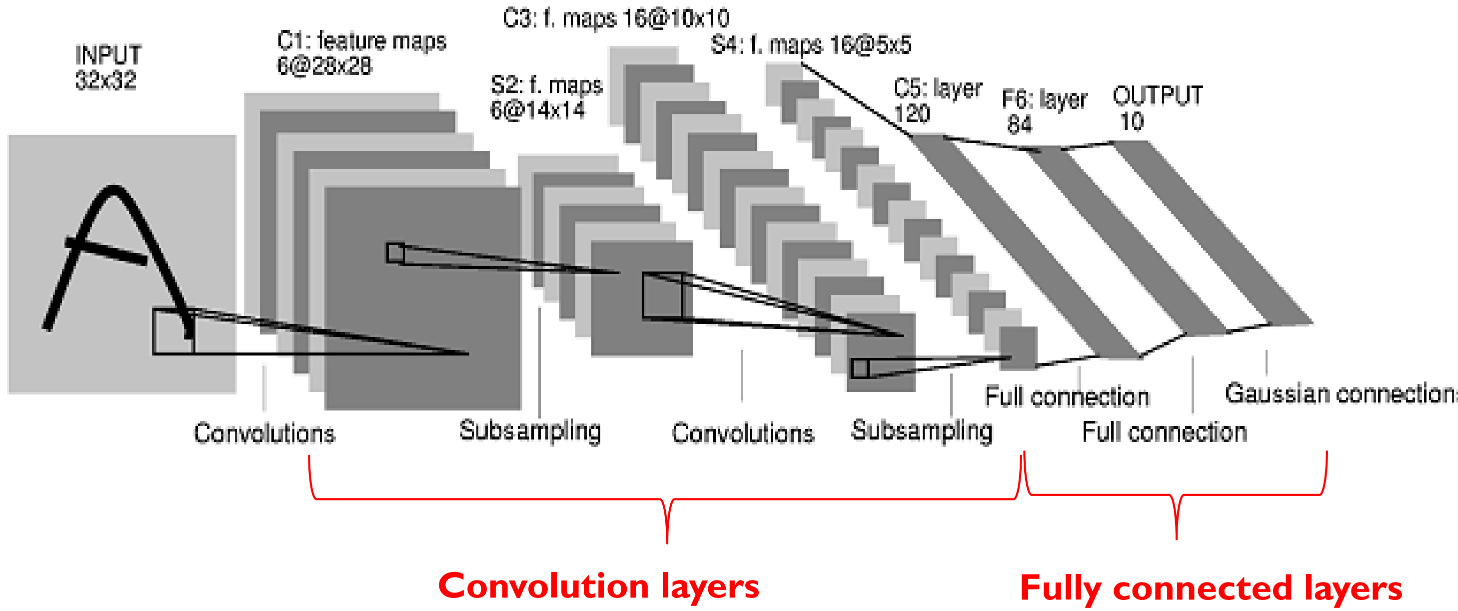
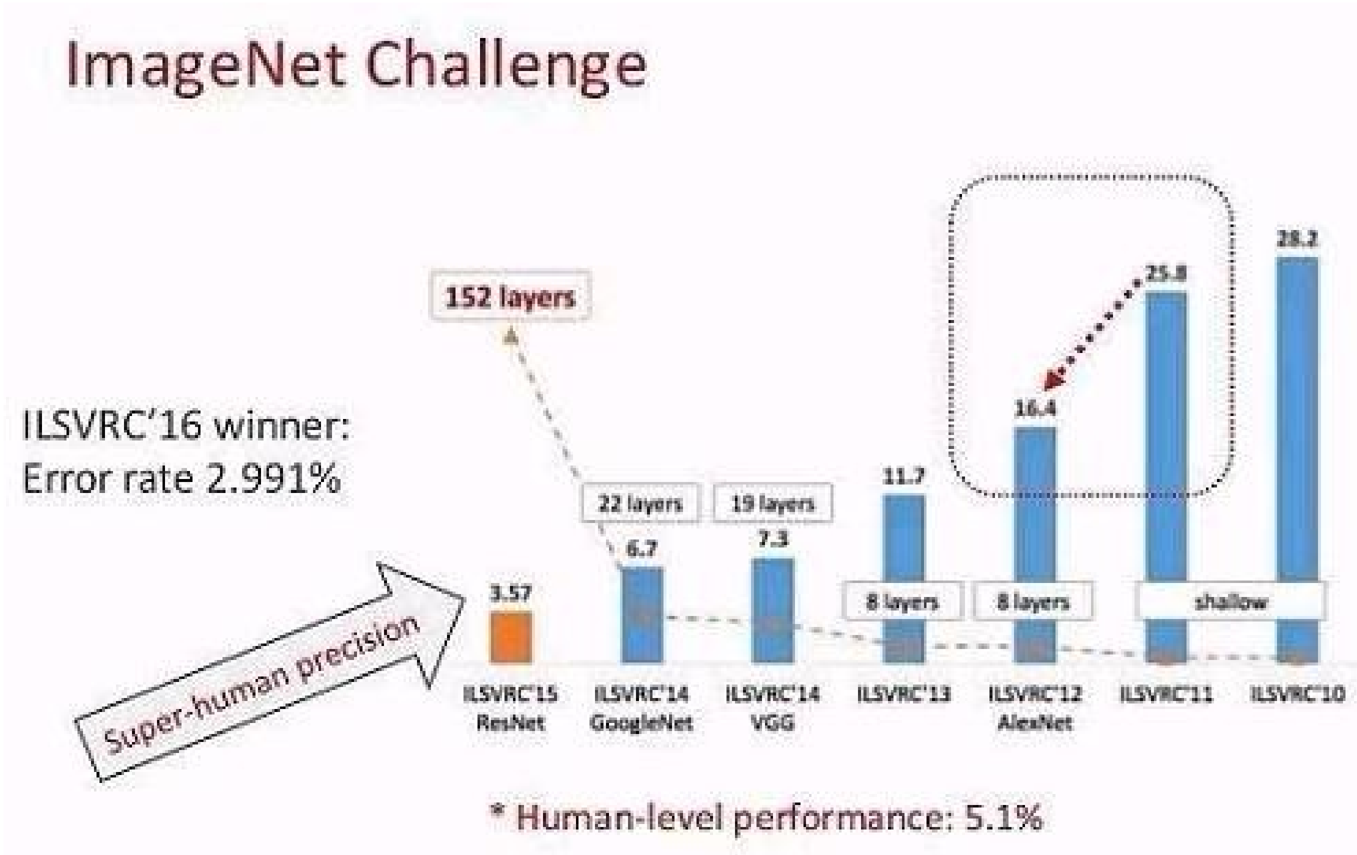


Image recognition

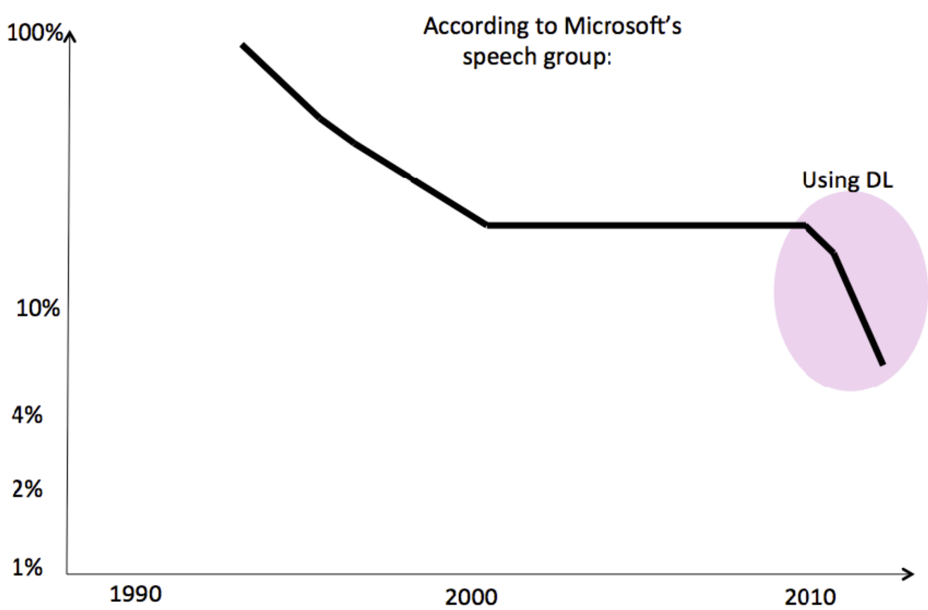
- **<u>AlexNet (2012)</u>**
  - 7 layers;
  - 1000 labels; >1.2M images
  - 17% vs 25.7%
- **GoogLeNet; VGG (2014)**
  - 19 layers (VGG)
  - ~7%

- **Deep Residual Net (2015)**
  - 152 layers
  - 3.57% > human recognition

History of DL –
ImageNet 2012 -



DL in speech recognition



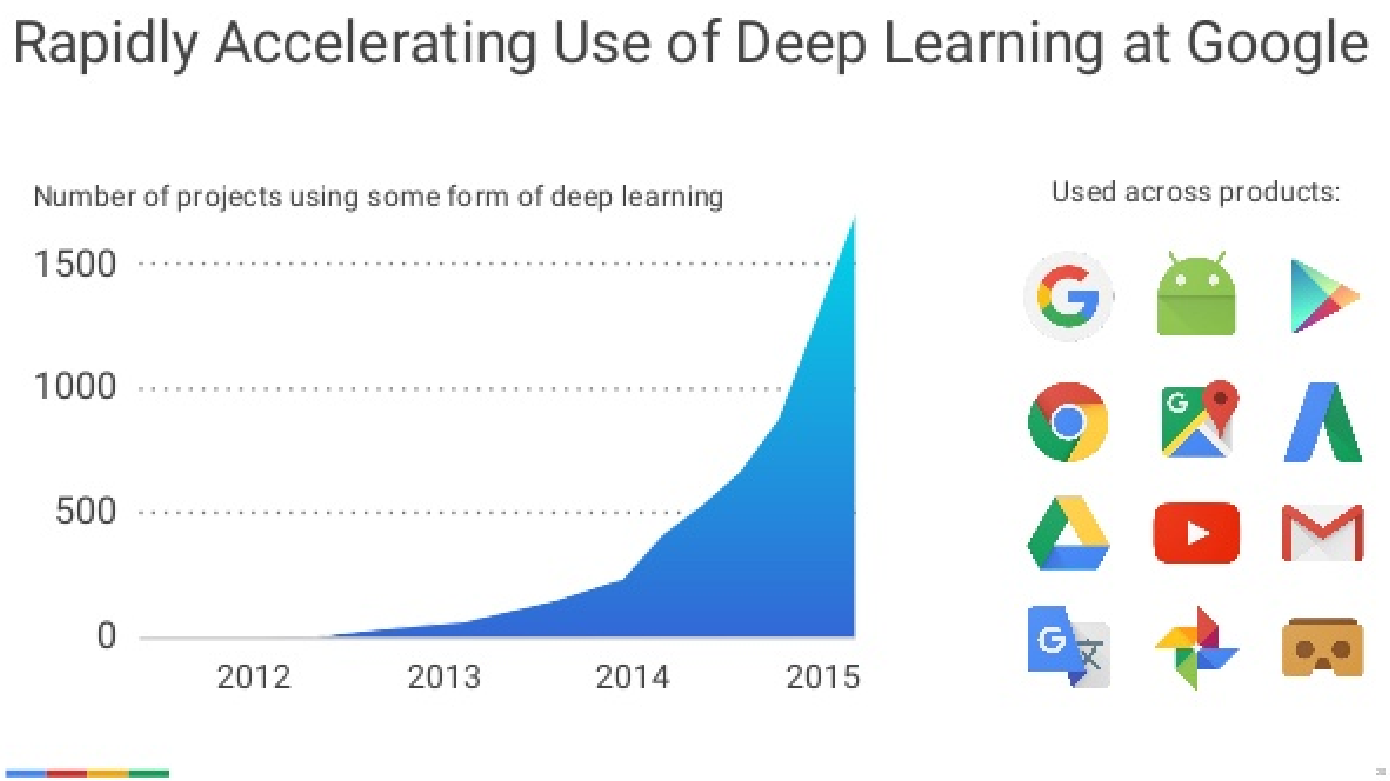
Deep Learning in Speech Recognition

BRIEF

Microsoft achieves 'human parity' in speech recognition system

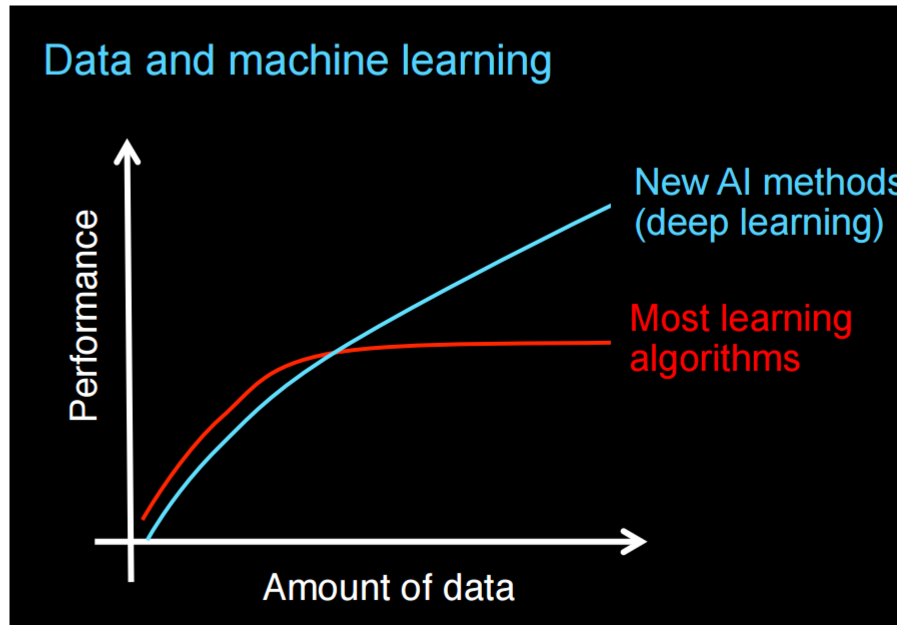
WER 5.1%

What about now?

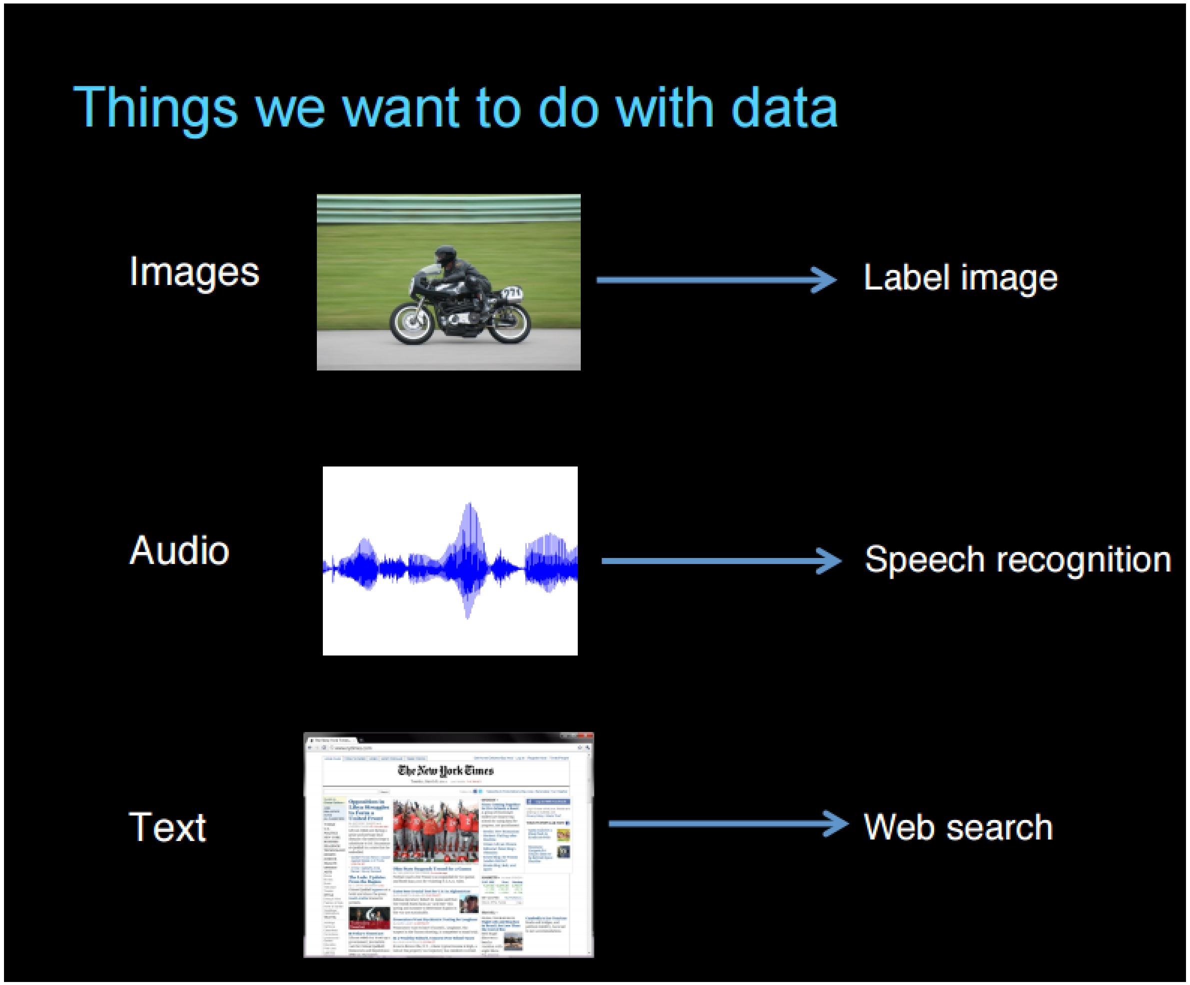


What are Deep Learning applications?

- Deep learning Applications
  - Vision
  - Audio
  - Language
  - BCI (EEG,…)
  - Bioinformatics
  - Art
  - Finance

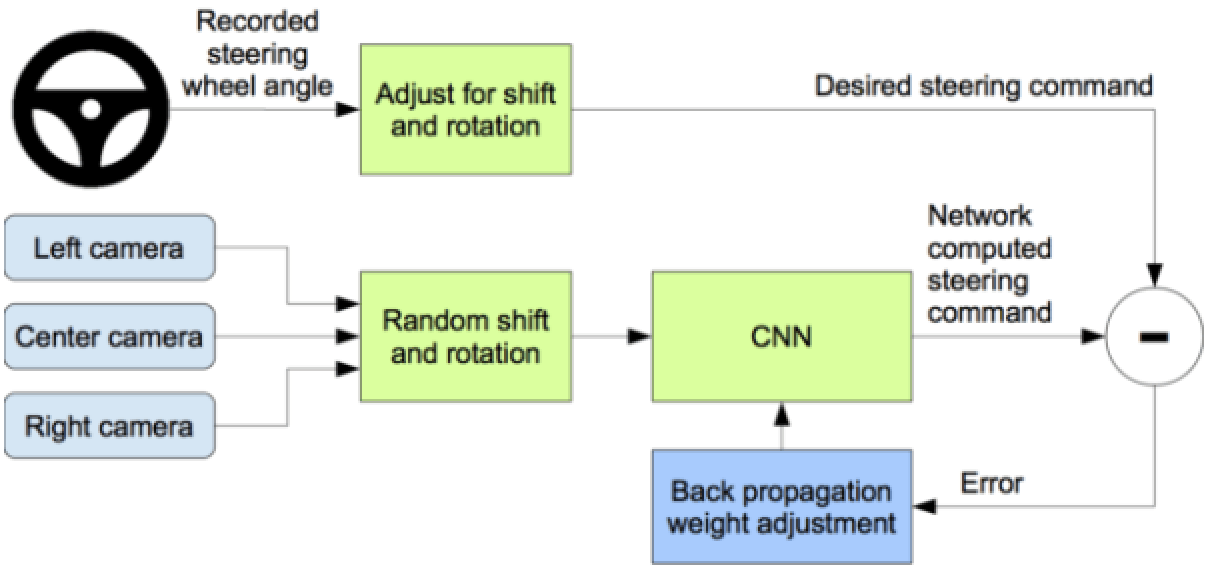


What are DL applications?



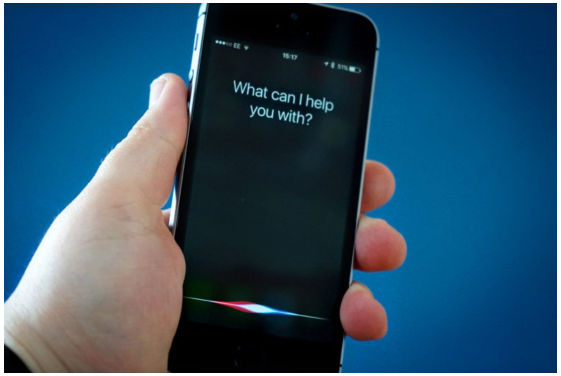
DL for robot control

Deep Learning for Self-Driving Cars



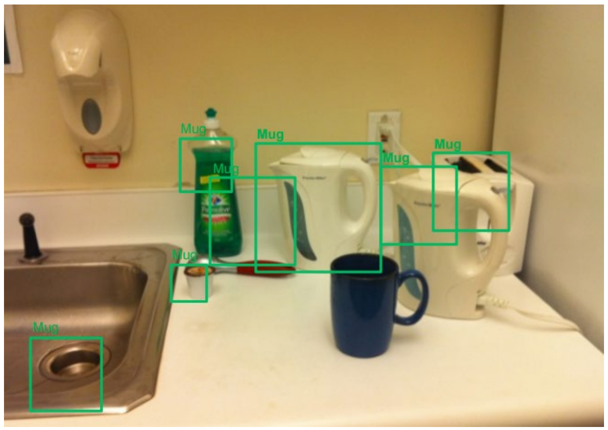
DL for speech recognition

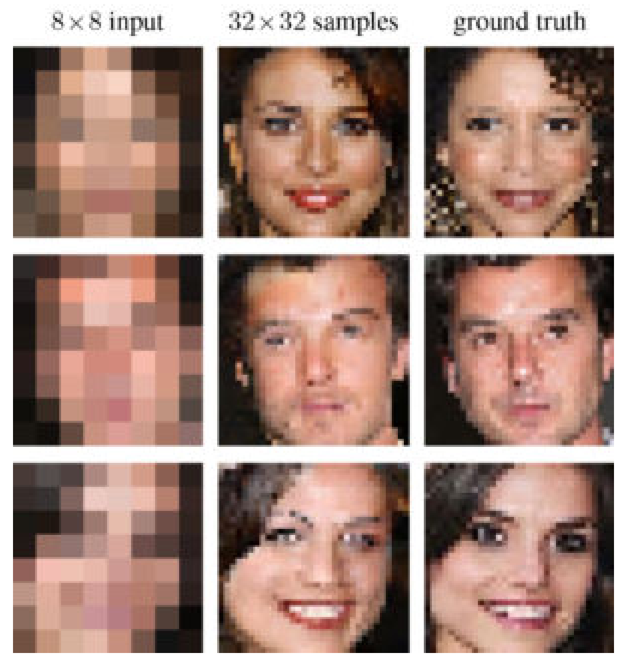
- Siri



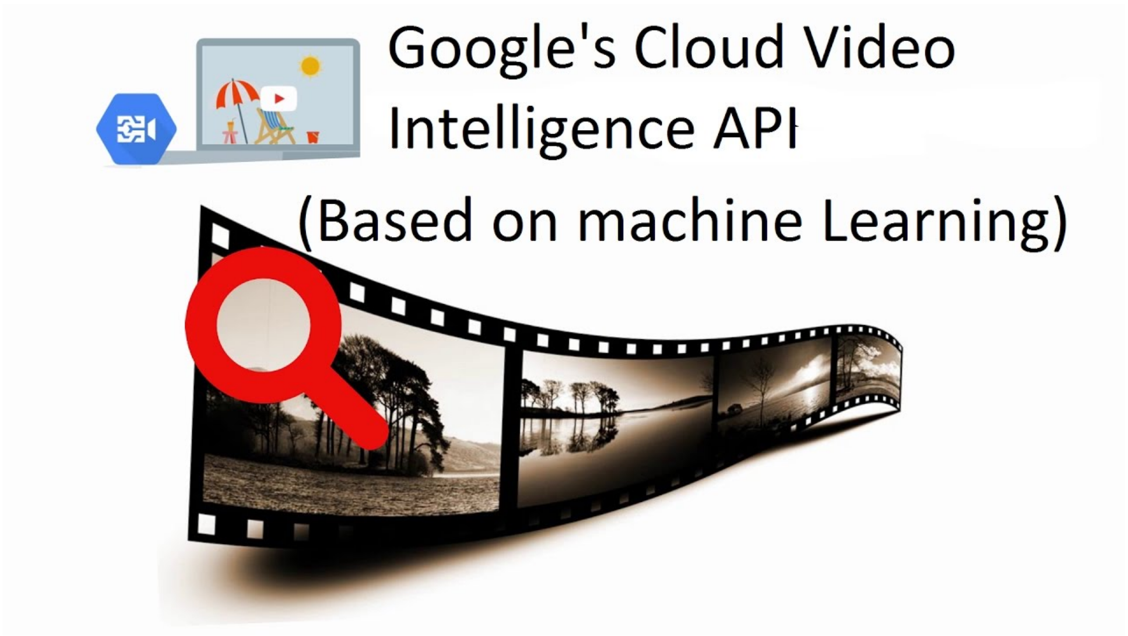
UTSA

DL for object recognition and image enhancement





DL in video search



DL for language processing

Google's Neural Machine Translation

...reduced translation errors by an average of 60% when compared to the prior Google Translate technology

DL for content recommendation

Biomedine

nature methods
Techniques for life scientists and chemists

Menu

home ▸ archive ▸ issue ▸ brief communication ▸ full text

NATURE METHODS | BRIEF COMMUNICATION

Predicting effects of noncodi
deep learning–based sequer

Jian Zhou & Olga G Troyanskaya

Affiliations | Contributions | Corresponding autho

Nature Methods **12**, 931–934 (2015) | doi:10.1038/r

nature biotechnology

Menu

home ▸ archive ▸ issue ▸ computational biology ▸ analysis ▸ full text

NATURE BIOTECHNOLOGY | COMPUTATIONAL BIOLOGY | ANALYSIS

日本語要約

Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning

Babak Alipanahi, Andrew Delong, Matthew T Weirauch & Brendan J Frey

Affiliations | Contributions | Corresponding author

Nature Biotechnology **33**, 831–838 (2015) | doi:10.1038/nbt.3300

Deep Learning in Physics: Searching for Exotic Particles

nature COMMUNICATIONS

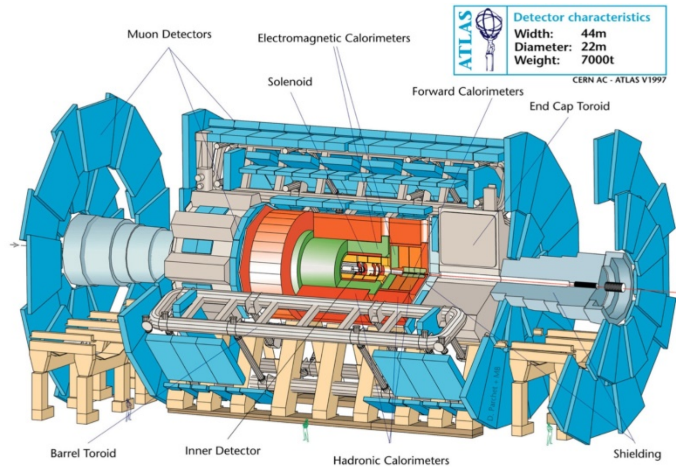
ARTICLE

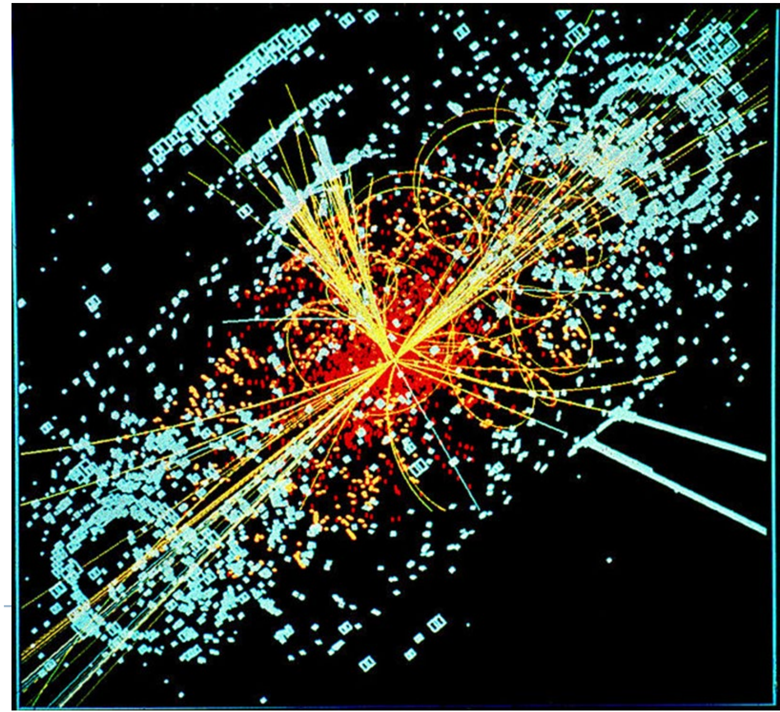
Received 19 Feb 2014 | Accepted 4 Jun 2014 | Published 2 Jul 2014 DOI: 10.1038/ncomms5308

Searching for exotic particles in high-energy physics with deep learning

P. Baldi[1], P. Sadowski[1] & D. Whiteson[2]

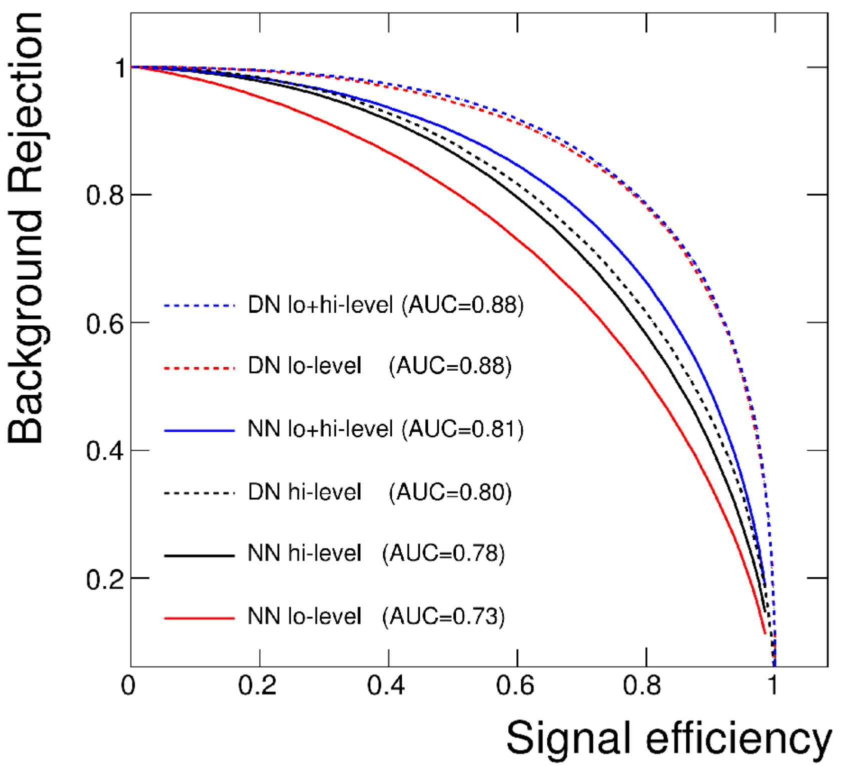
Collisions at high-energy particle colliders are a traditionally fruitful source of exotic particle discoveries. Finding these rare particles requires solving difficult signal-versus-background classification problems, hence machine-learning approaches are often used. Standard approaches have relied on 'shallow' machine-learning models that have a limited capacity to learn complex nonlinear functions of the inputs, and rely on a painstaking search through manually constructed nonlinear features. Progress on this problem has slowed, as a variety of techniques have shown equivalent performance. Recent advances in the field of deep learning make it possible to learn more complex functions and better discriminate between signal and background classes. Here, using benchmark data sets, we show that deep-learning methods need no manually constructed inputs and yet improve the classification metric by as much as 8% over the best current approaches. This demonstrates that deep-learning approaches can improve the power of collider searches for exotic particles.





UTSA

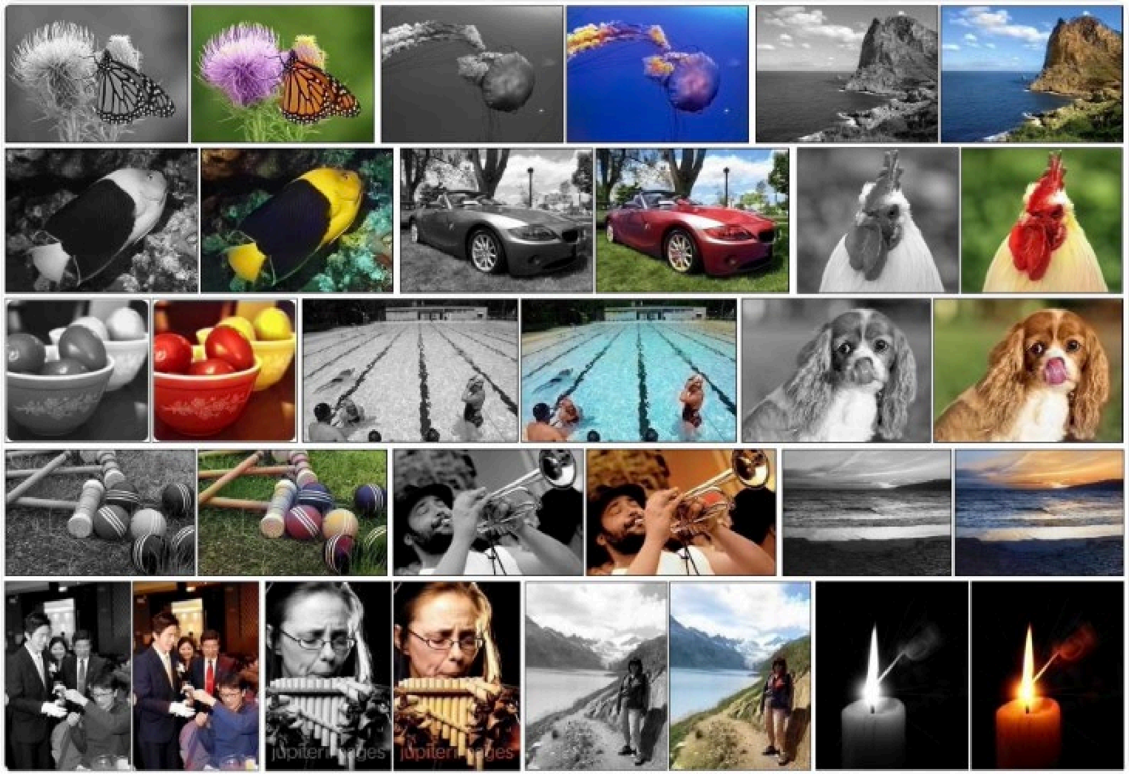
Deep Learning in Physics: Searching for Exotic Particles



	AUC		
Technique	Low-level	High-level	Complete
BDT	0.73	0.78	0.81
NN	0.733 (0.007)	0.777 (0.001)	0.816 (0.004)
DN	0.880 (0.001)	0.800 (< 0.001)	0.885 (0.002)

Deep network improves AUC by 8%

Automatic Colorization of Black and White Images



Papers

Deep Colorization ,2015

Colorful Image Colorization ,2016

Learning Representations for Automatic Colorization ,2016

Image Colorization with Deep Convolutional Neural Networks ,2016

Automatically Adding Sounds To Silent Movies

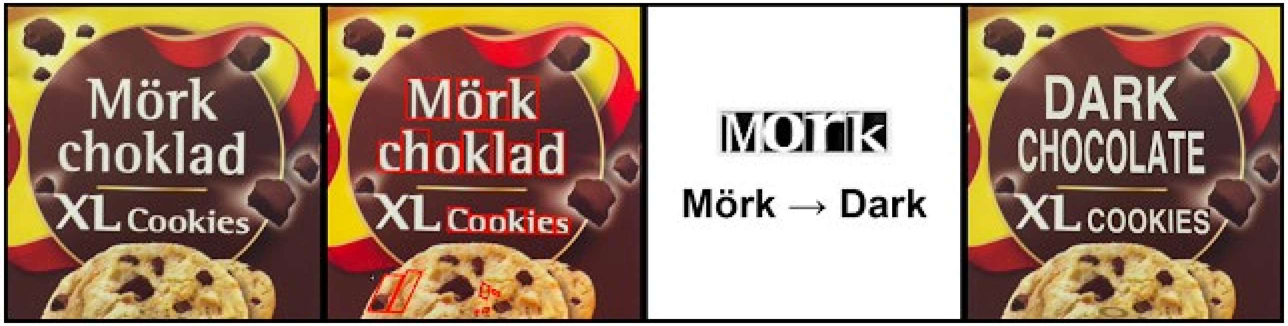
- In this task the system must synthesize sounds to match a silent video.
- The system is trained using 1000 examples of video with sound of a drum stick striking different surfaces and creating different sounds. A deep learning model associates the video frames with a database of pre-rerecorded sounds in order to select a sound to play that best matches what is happening in the scene.
- The system was then evaluated using a turing-test like setup where humans had to determine which video had the real or the fake (synthesized) sounds.
- A very cool application of both convolutional neural networks and LSTM recurrent neural networks.

Paper

Visually Indicated Sounds ,2015

Automatic Machine Translation

- As you would expect, convolutional neural networks are used to identify images that have letters and where the letters are in the scene. Once identified, they can be turned into text, translated and the image recreated with the translated text. This is often called instant visual translation.



Papers

[Sequence to Sequence Learning with Neural Networks](#) ,2014

[Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation](#),2014

[Deep Neural Networks in Machine Translation: An Overview](#) ,2015

Object Classification and Detection in Photographs

- This task requires the classification of objects within a photograph as one of a set of previously known objects.
- State-of-the-art results have been achieved on benchmark examples of this problem using very large convolutional neural networks. A breakthrough in this problem by Alex Krizhevsky et al. results on the ImageNet classification problem called AlexNet.

Papers

ImageNet Classification with Deep Convolutional Neural Networks ,2012
Some Improvements on Deep Convolutional Neural Network Based Image Classification,2013
Scalable Object Detection using Deep Neural Networks ,2013
Deep Neural Networks for Object Detection ,2013

Others…

- **Automatic Handwriting Generation**
- This is a task where given a corpus of handwriting examples, generate new handwriting for a given word or phrase.

Generating Sequences With Recurrent Neural Networks , 2013

- **Automatic Text Generation**
- This is an interesting task, where a corpus of text is learned and from this model new text is generated, word-by-word or character-by-character.

Papers

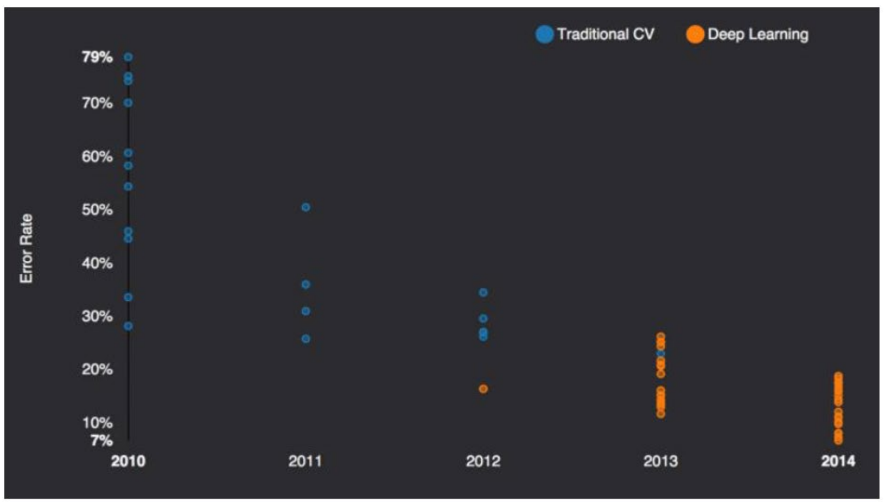
Generating Text with Recurrent Neural Networks ,2011
Generating Sequences With Recurrent Neural Networks ,2013

What's unique about DL?

What's unique about DL?

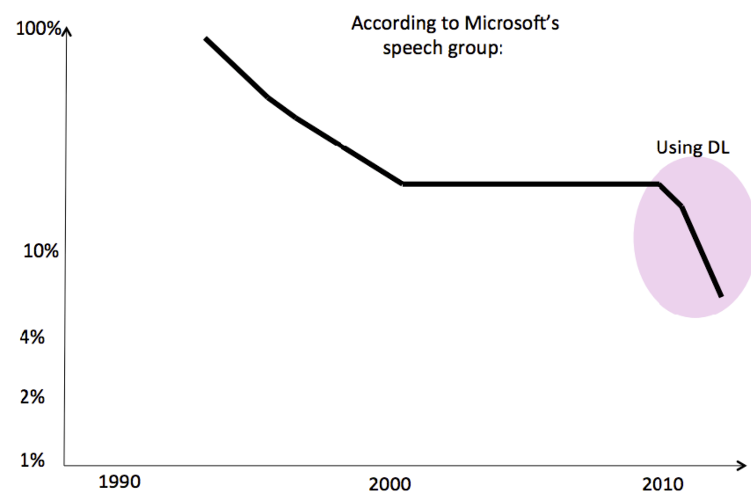
Maximize AI performance

Computer vision



ImageNet: The "computer vision World Cup"

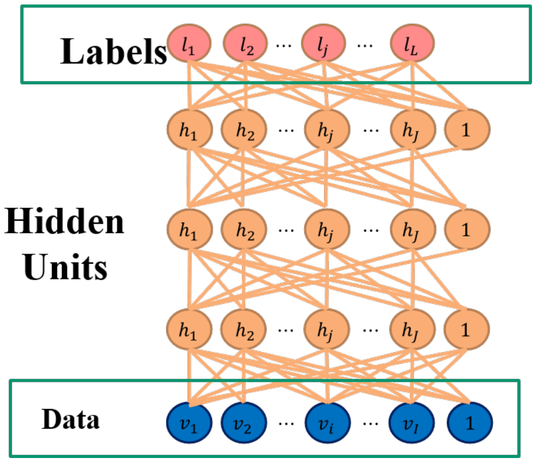
Speech recognition

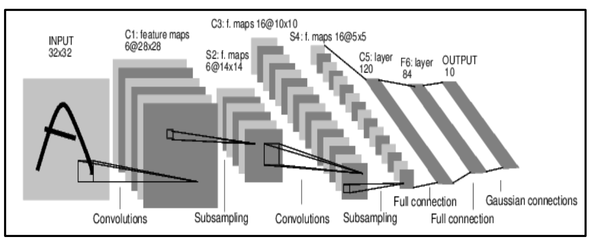


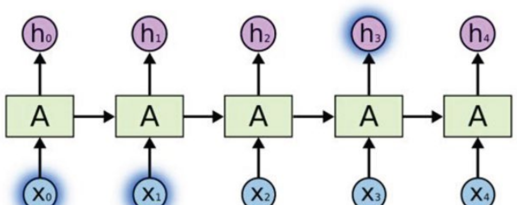
Deep Learning in Speech Recognition

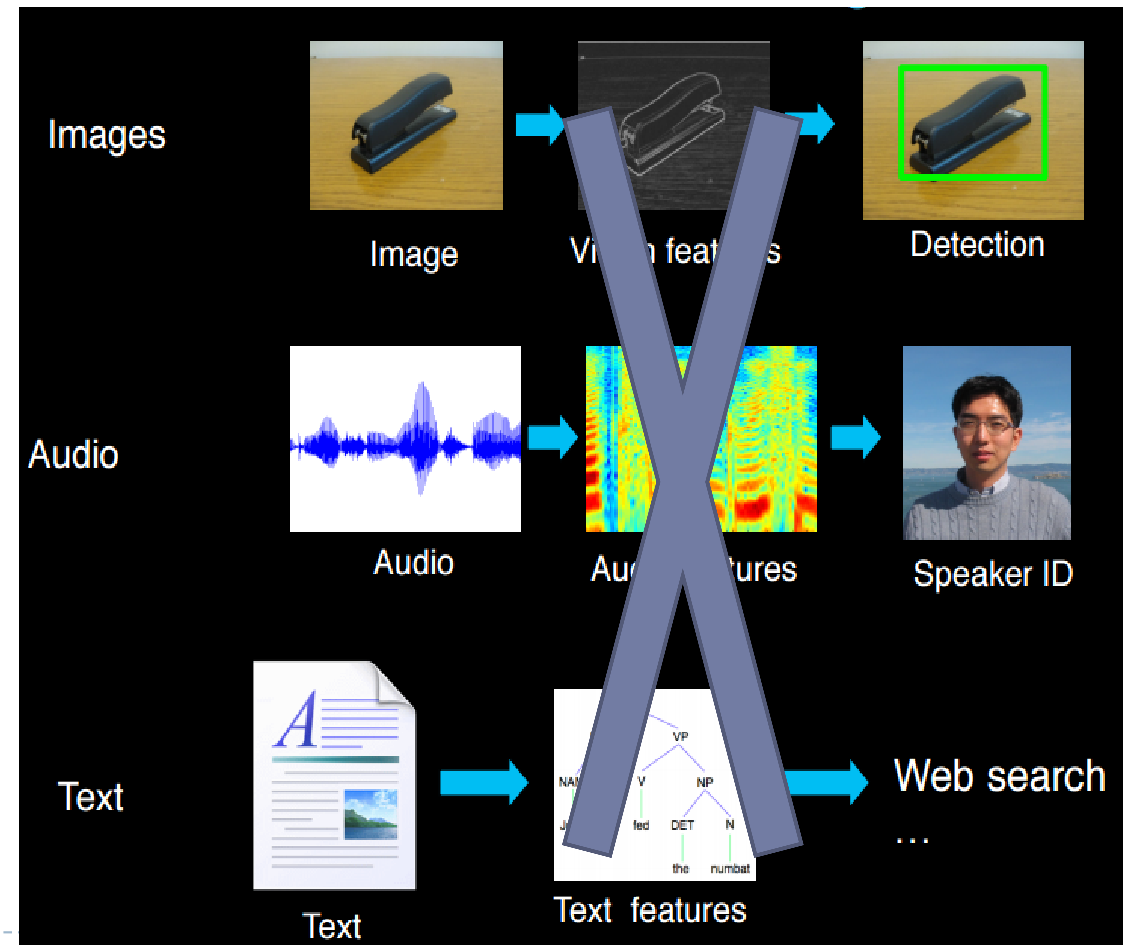
What's unique about DL?

Deeply simples AI systems

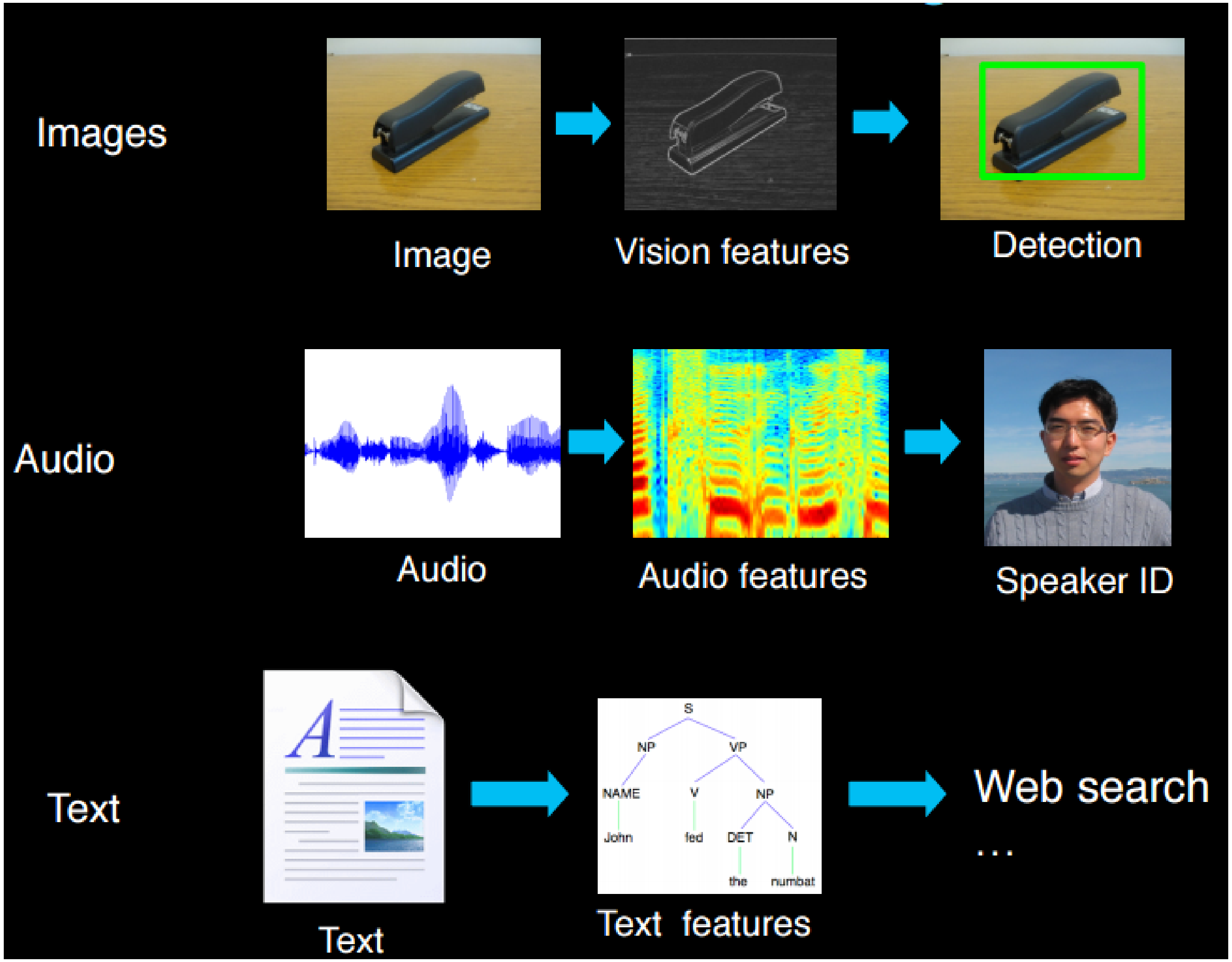




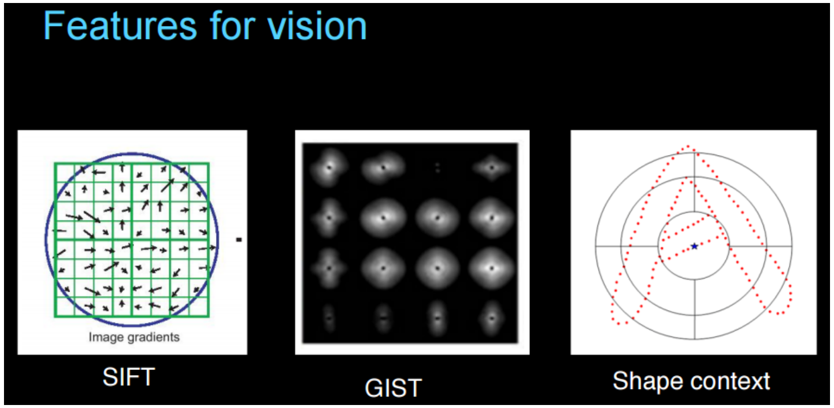


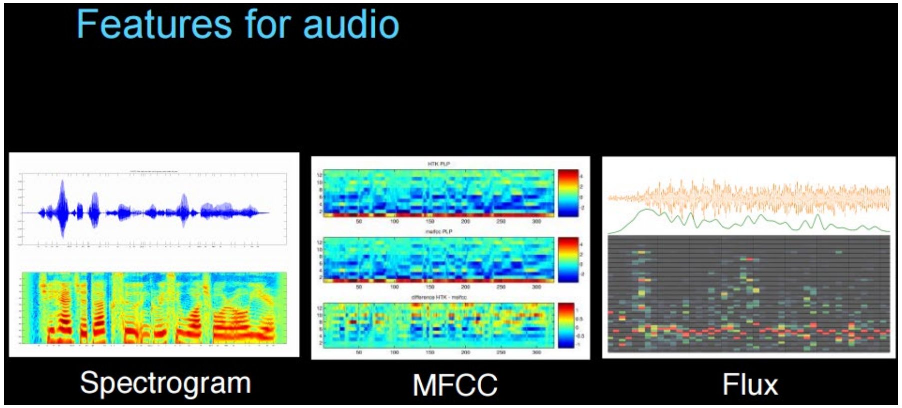


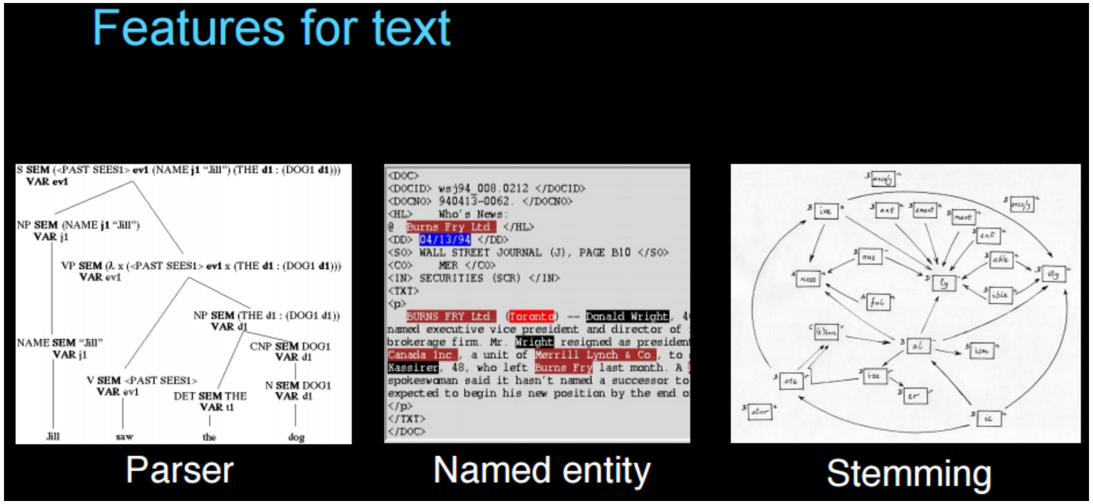
Hand crafted features for traditional machine learning?



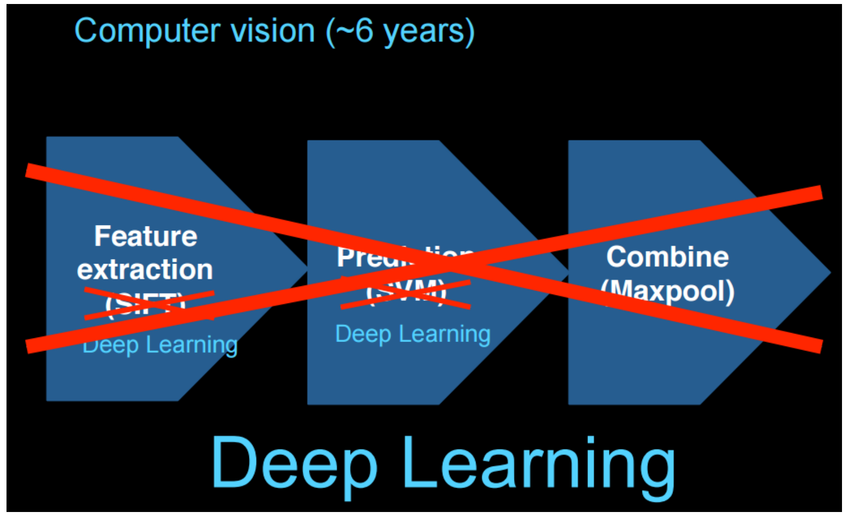
Hand crafted features for traditional machine learning?

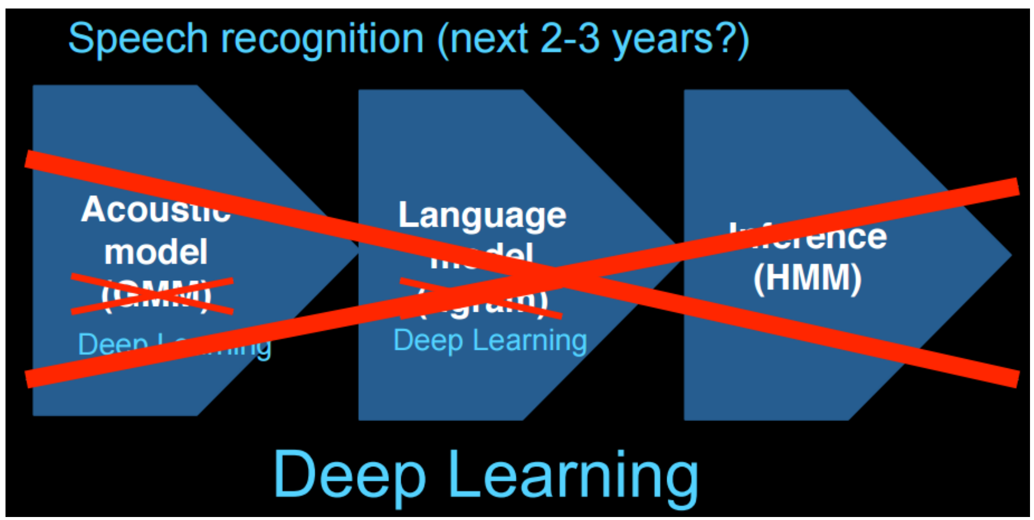






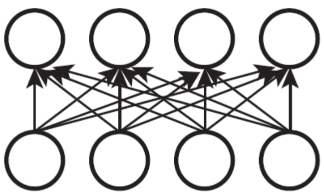
Deep learning trend?



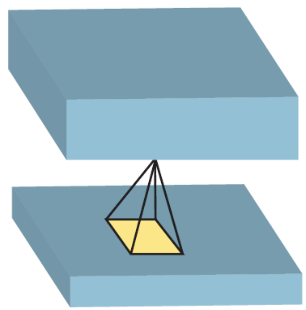


Deep learning modules

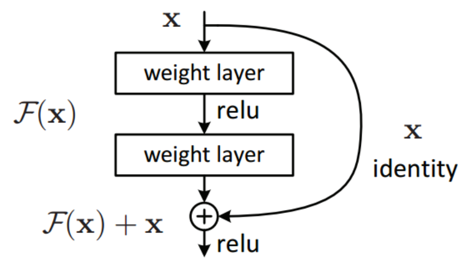
Dense layers



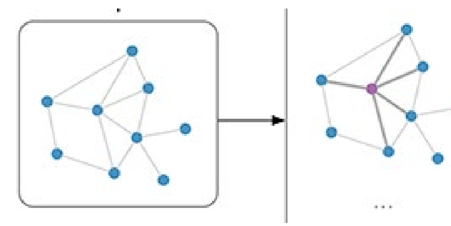
Convolutional layer



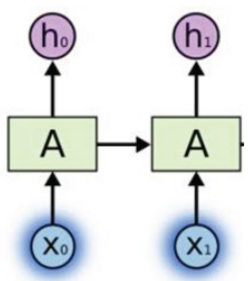
Residual layer



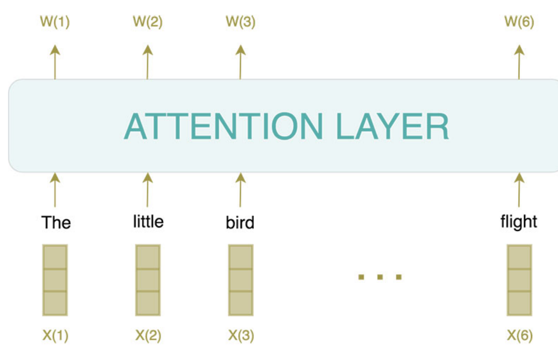
Graph convolution



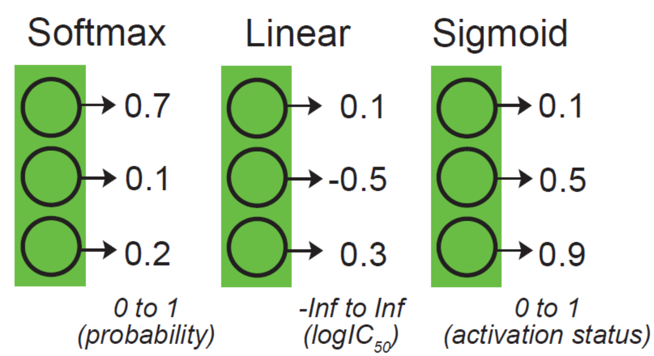
Recurrent



Attention layer

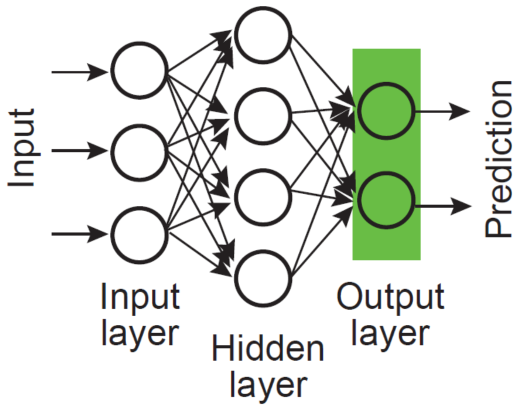


Prediction layers

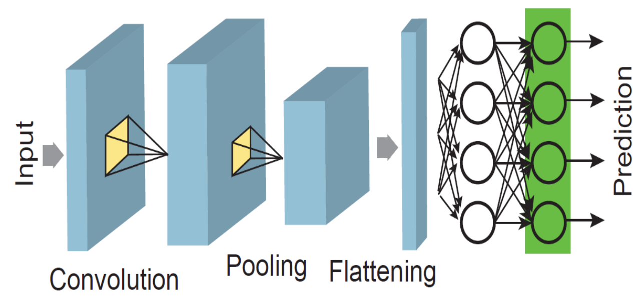


Supervised deep learning models

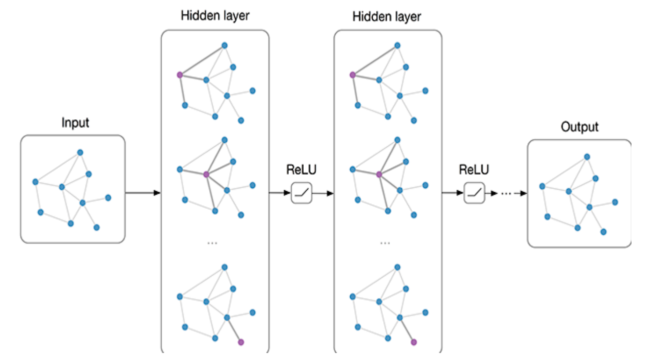
DNN



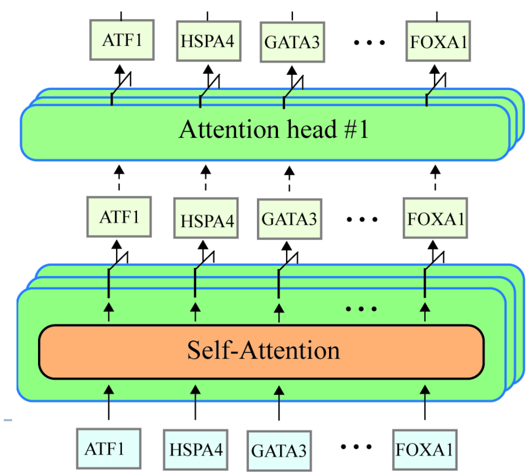
Convolutional neural networks (CNN)



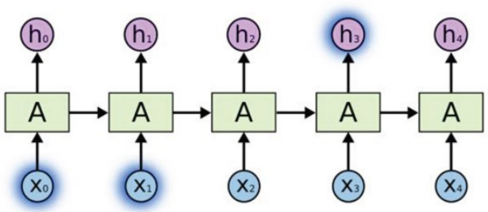
Graph CNN



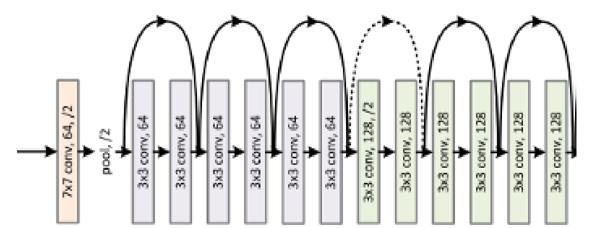
Transformer



RNN (LSTM, GRU)



ResNet



Big players - academics

Geoffrey Hinton: University of Toronto & Google

Yann LeCun: New York University & Facebook

Andrew Ng: Stanford & Baidu

Yoshua Bengio: University of Montreal

Jürgen Schmidhuber: Swiss AI Lab & NNAISENSE

Big players
- Companies

Big players
- Startups