Deep Learning
Introduction to Deep Learning

Puneet Kumar Jain

CSE Department
National Institute of Technology Rourkela

References:

The Slides are prepared from the following major sources:

- **"CS7105-Deep Learning" by Mitesh M. Khapra, IIT Madras. http://www.cse.iitm.ac.in/~miteshk/CS7015_2018.html**
- Jürgen Schmidhuber. Deep learning in neural networks: An overview. *Neural Networks*, 61:85–117, 2015.

Biological Neurons

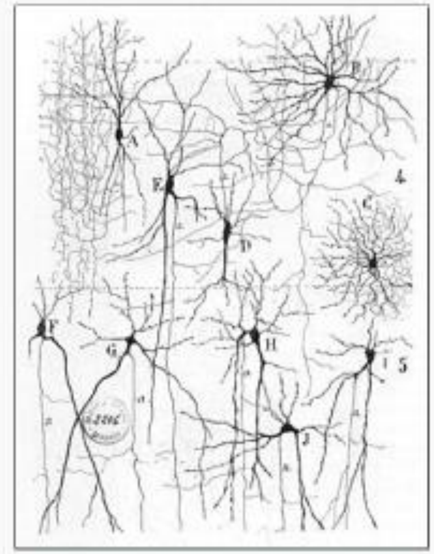
Reticular Theory

Joseph von Gerlach proposed that the nervous system is a single continuous network as opposed to a network of many discrete cells!

Biological Neurons

Neuron Doctrine

Santiago Ramón y Cajal used Golgi's technique to study the nervous system and proposed that it is actually made up of discrete individual cells formimg a network (as opposed to a single continuous network)

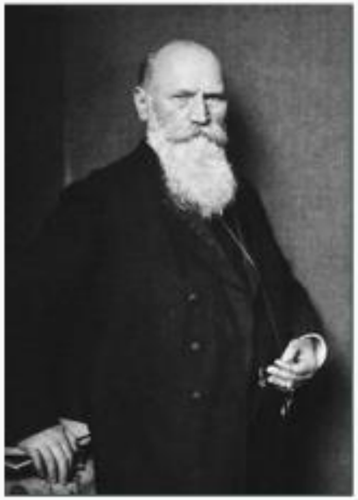


Biological Neurons

The Term Neuron

The term neuron was coined by Heinrich Wilhelm Gottfried von Waldeyer-Hartz around 1891.

He further consolidated the Neuron Doctrine.

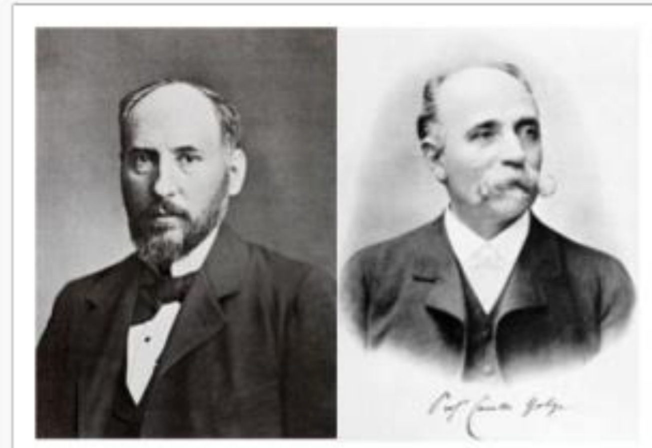


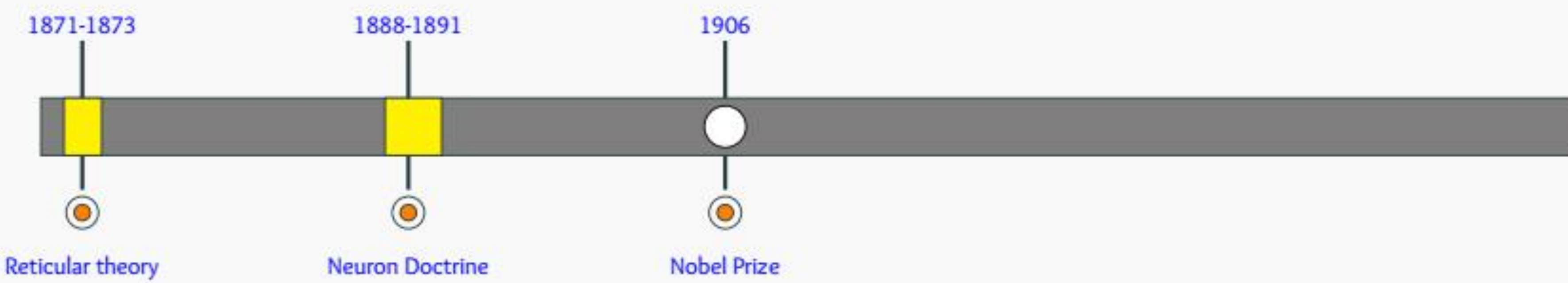
Same person also coined the termed chromosome

Biological Neurons

Nobel Prize

Both Golgi (reticular theory) and Cajal (neuron doctrine) were jointly awarded the 1906 Nobel Prize for Physiology or Medicine, that resulted in lasting conflicting ideas and controversies between the two scientists.

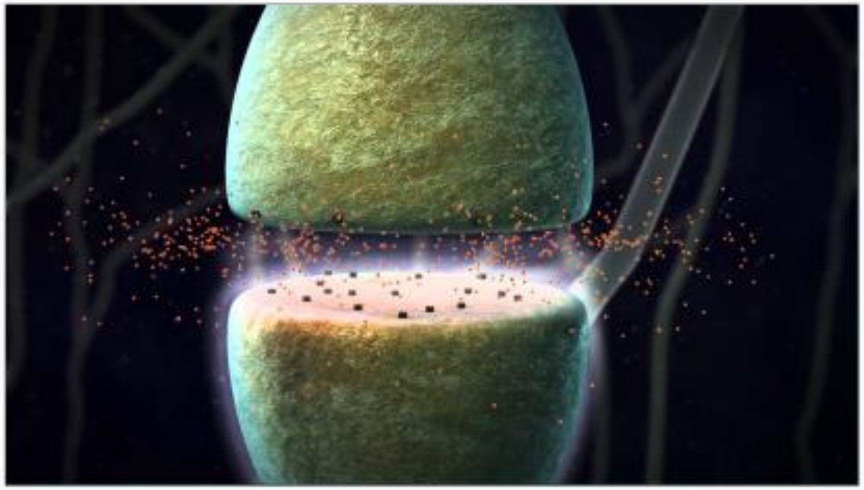


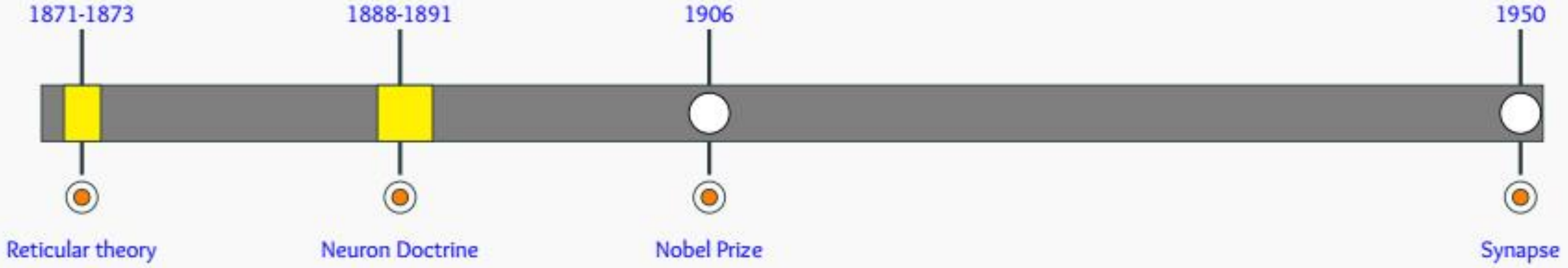


Biological Neurons

The Final Word

In 1950s electron microscopy finally confirmed the neuron doctrine by unambiguously demonstrating that nerve cells were individual cells interconnected through synapses (a network of many individual neurons).





From Spring to Winter of AI

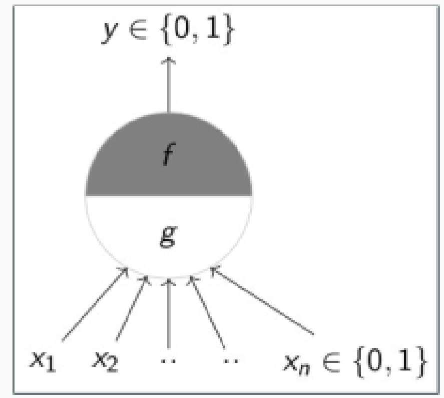
- In the history of AI, the dominant narrative swings back and forth between periods of “spring” and “winter”:

First Spring (1956–1974), **First Winter (1974–1981)**

Second Spring (1981–1987) **Second Winter (1987–1993)**

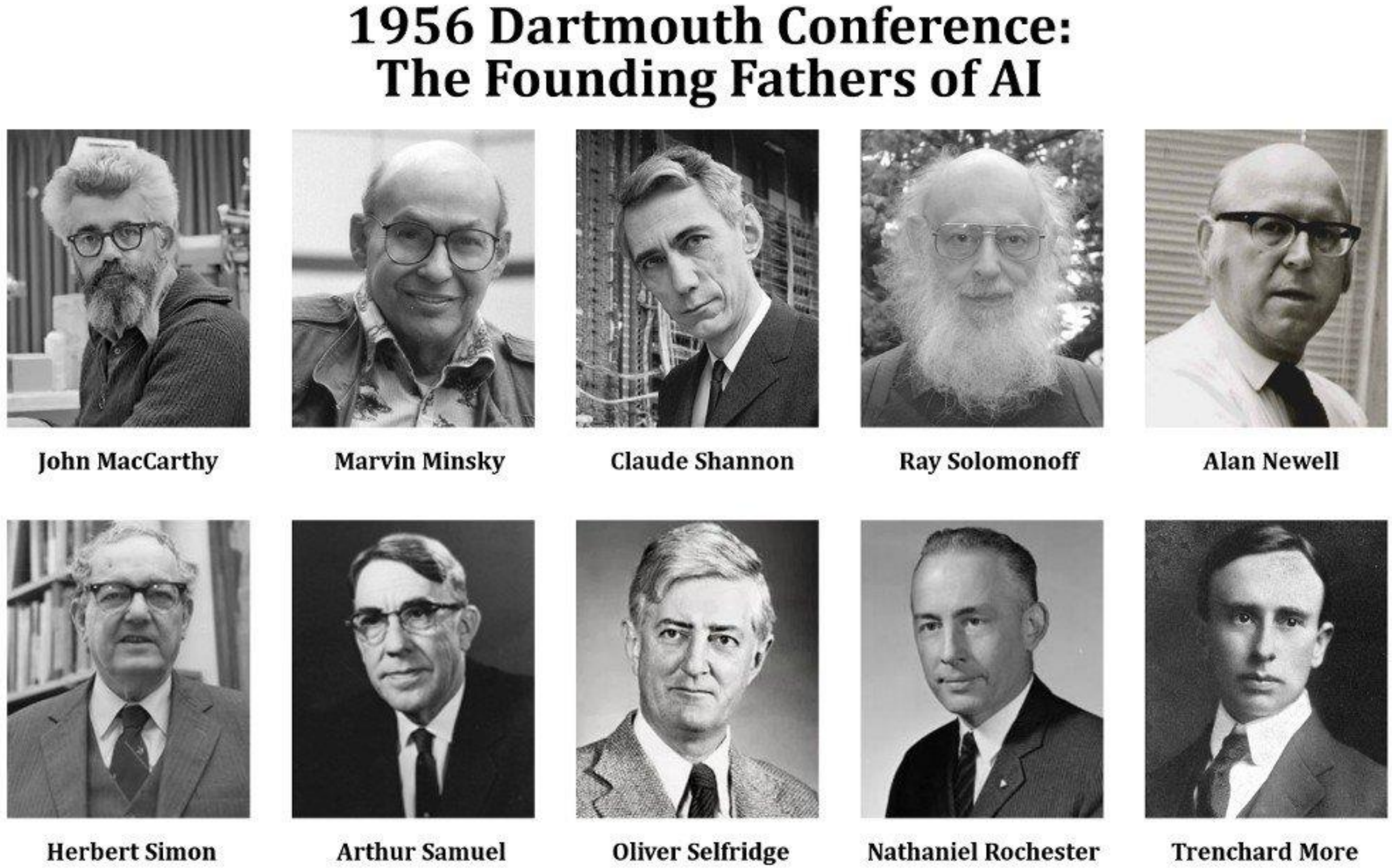
McCulloch Pitts Neuron

McCulloch (neuroscientist) and Pitts (logician) proposed a highly simplified model of the neuron (1943) [2]



Dartmouth conference: 1956

- “Artificial Intelligence” term was coined

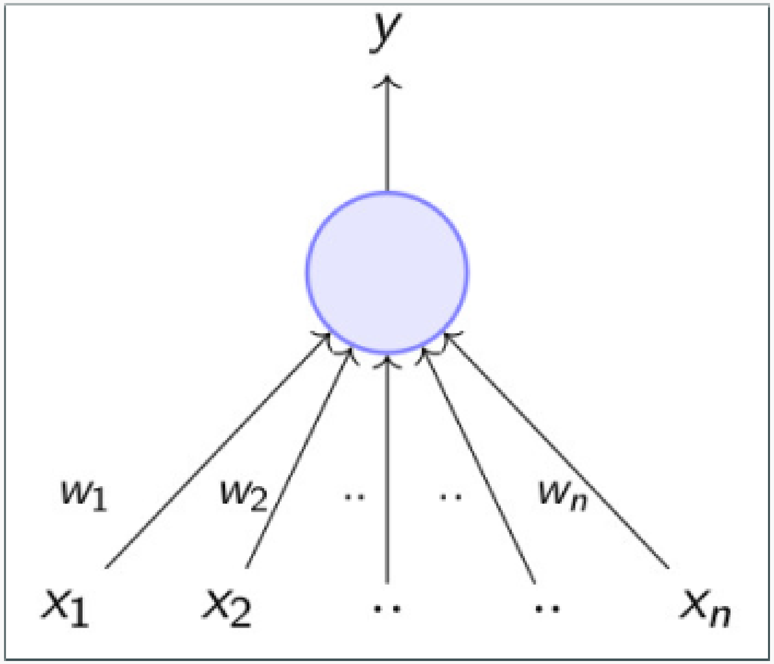


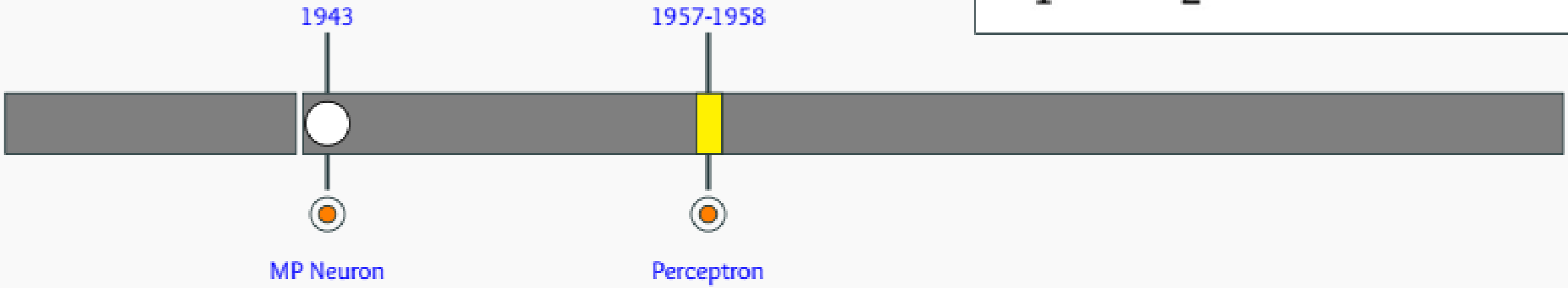
Courtesy of scienceabc.com

From Spring to Winter of AI

Perceptron

"the perceptron may eventually be able to learn, make decisions, and translate languages" -Frank Rosenblatt

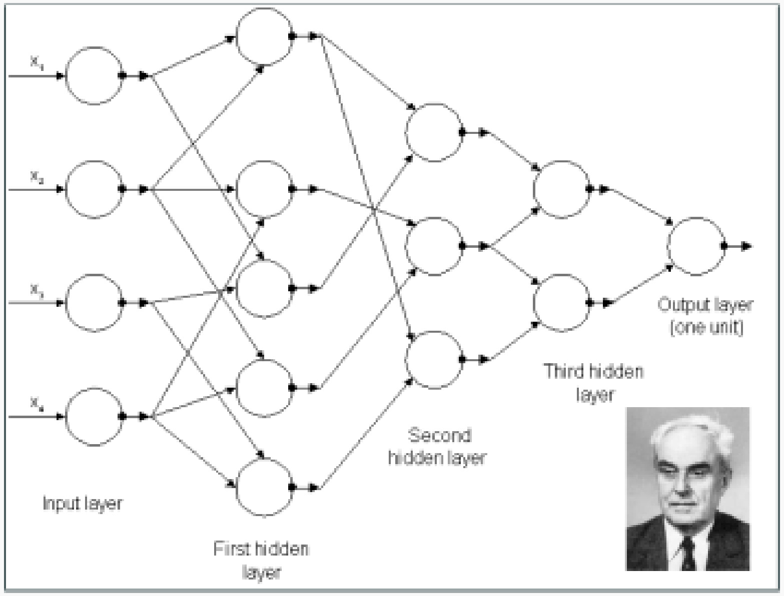


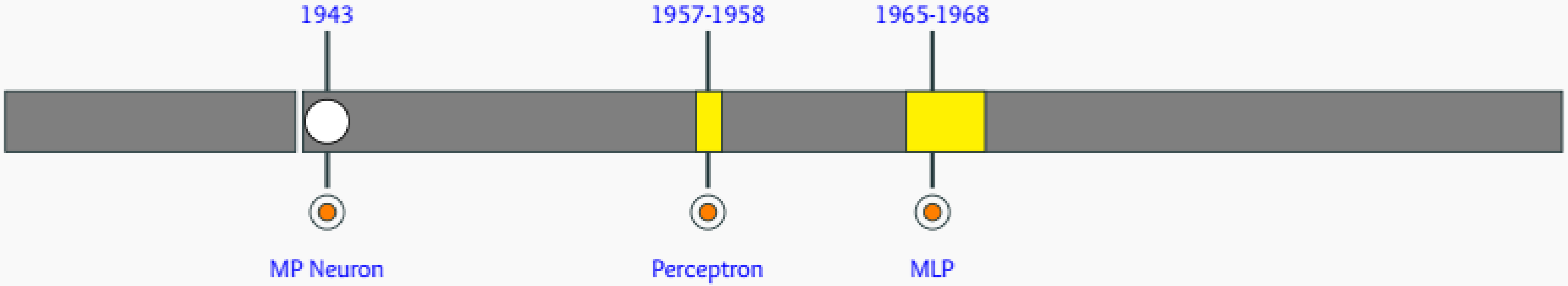


From Spring to Winter of AI

First generation Multilayer Perceptrons

Ivakhnenko et. al. [3]

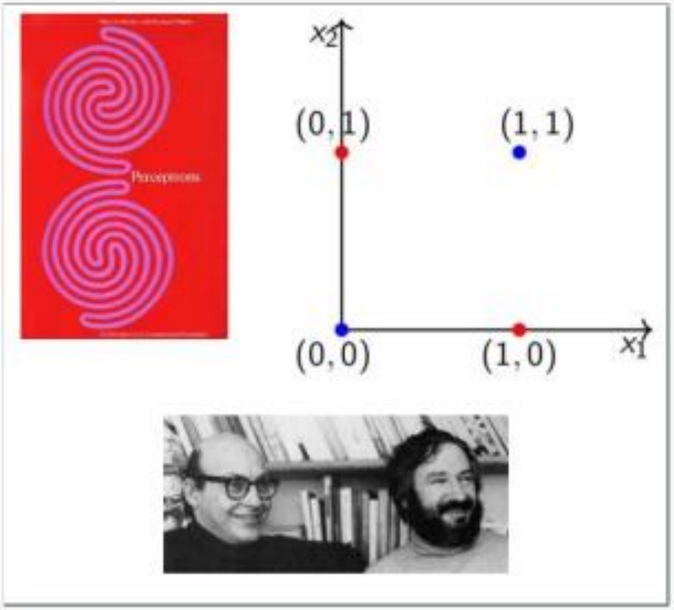


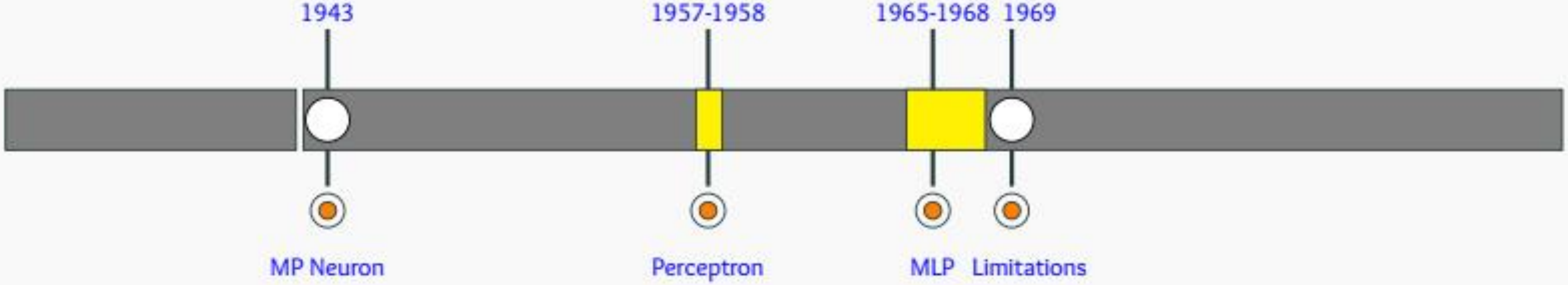


From Spring to Winter of AI

Perceptron Limitations

In their now famous book “Perceptrons”, Minsky and Papert outlined the limits of what perceptrons could do [4]

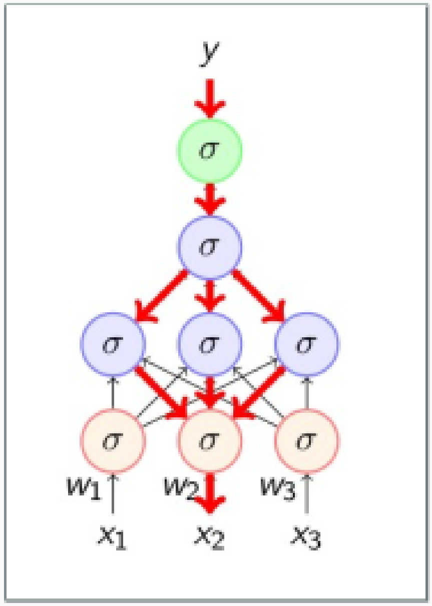


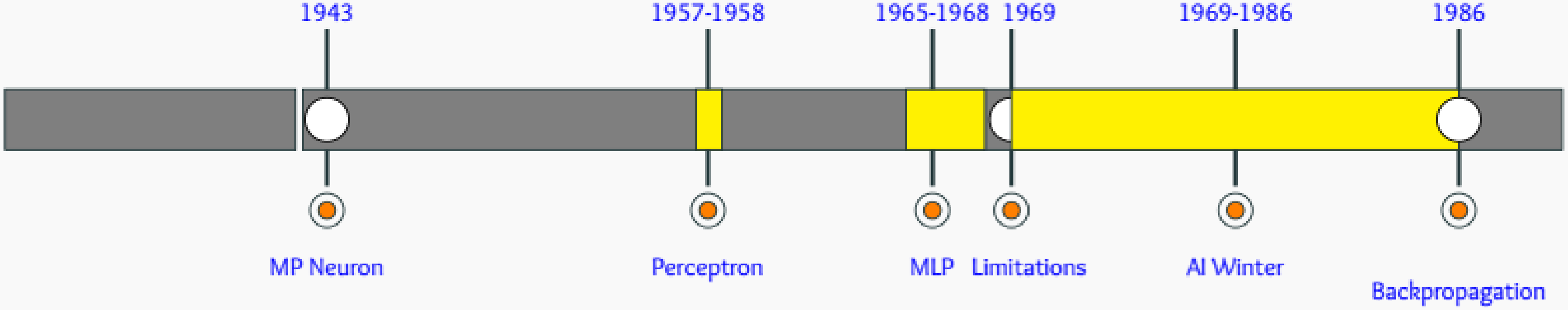


From Spring to Winter of AI

Backpropagation

- Discovered and rediscovered several times throughout 1960's and 1970's
- Werbos(1982) [5] first used it in the context of artificial neural networks
- Eventually popularized by the work of Rumelhart et. al. in 1986 [6]

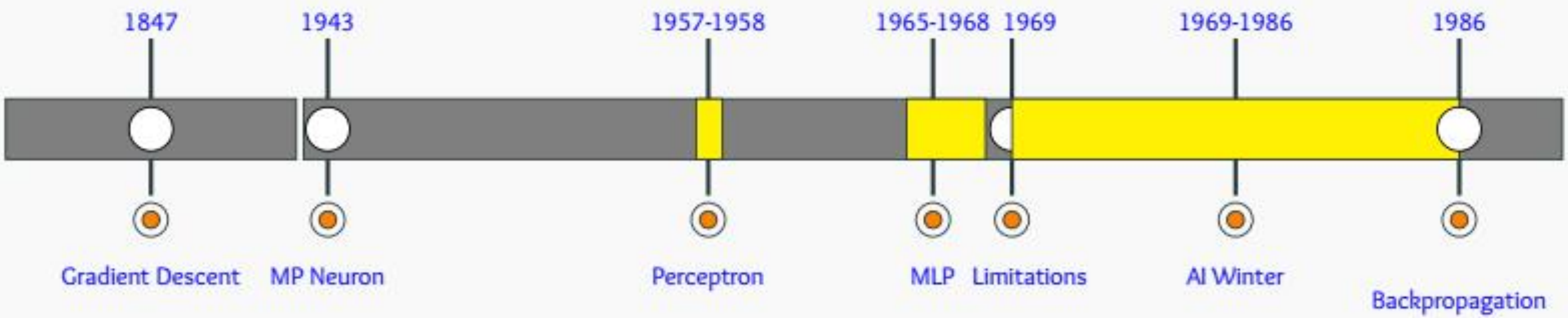




From Spring to Winter of AI

Gradient Descent

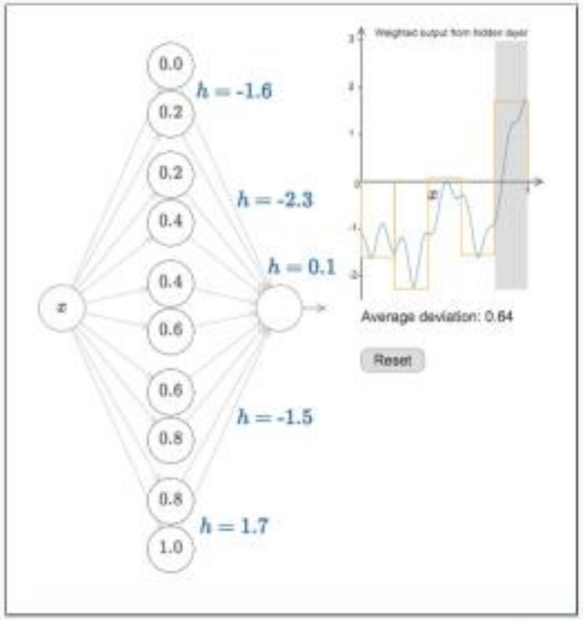
Cauchy discovered Gradient Descent motivated by the need to compute the orbit of heavenly bodies

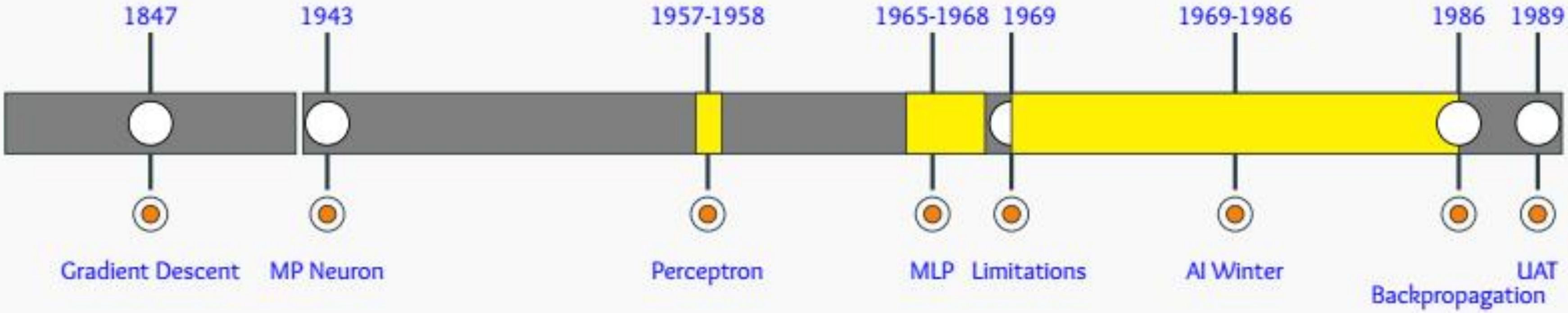


From Spring to Winter of AI

Universal Approximation Theorem

A multilayered network of neurons with a single hidden layer can be used to approximate any continuous function to any desired precision [7]

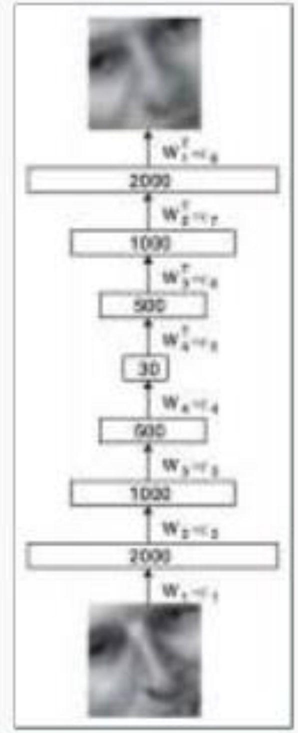




The Deep Revival

Unsupervised Pre-Training

The idea of unsupervised pre-training actually dates back to 1991-1993 (J. Schmidhuber) when it was used to train a “Very Deep Learner”



The Deep Revival

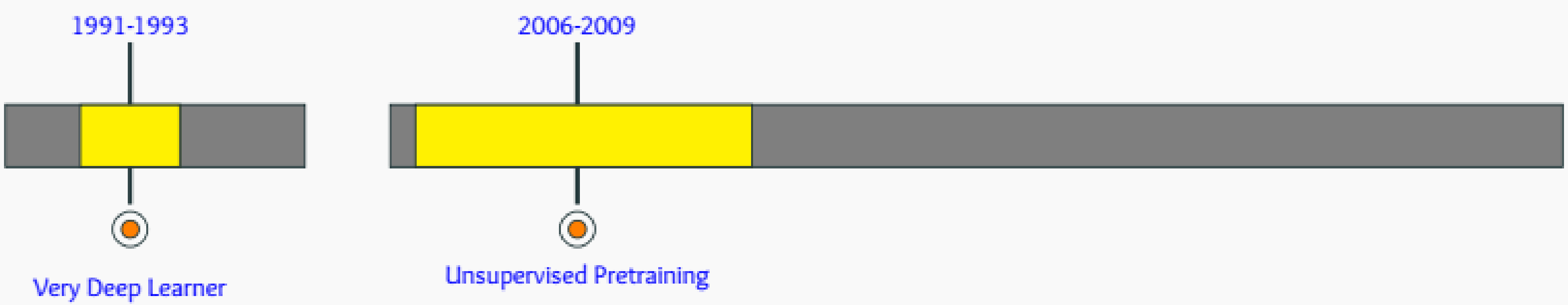
More insights (2007-2009)

Further Investigations into the effectiveness of Unsupervised Pre-training

Greedy Layer-Wise Training of Deep Networks

Why Does Unsupervised Pre-training Help Deep Learning?

Exploring Strategies for Training Deep Neural Networks

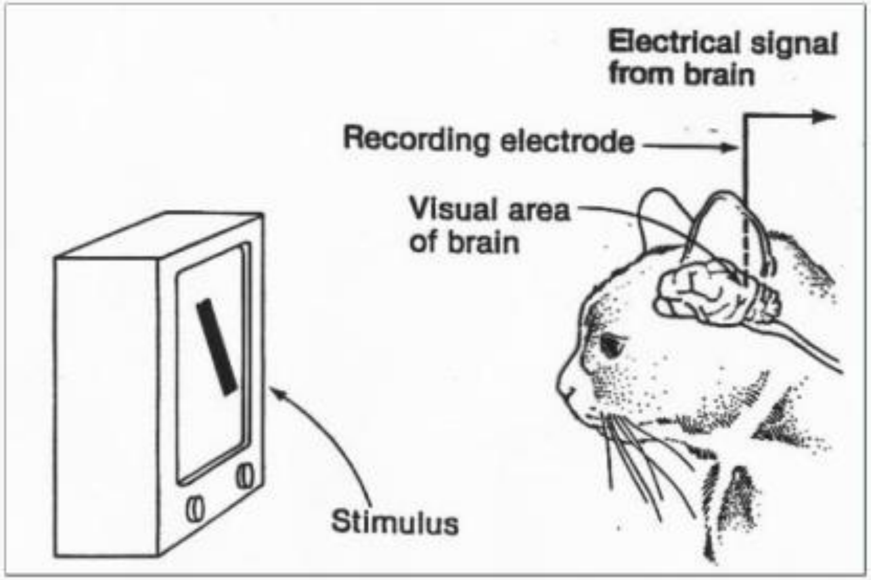


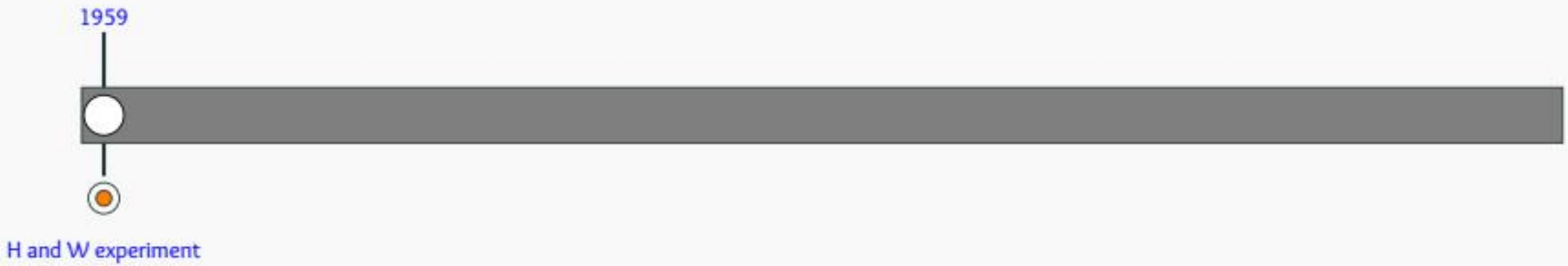
- **Cats to CNN**

From Cats to Convolutional Neural Nets

Hubel and Wiesel Experiment

Experimentally showed that each neuron has a fixed receptive field - i.e. a neuron will fire only in response to a visual stimuli in a specific region in the visual space [18]

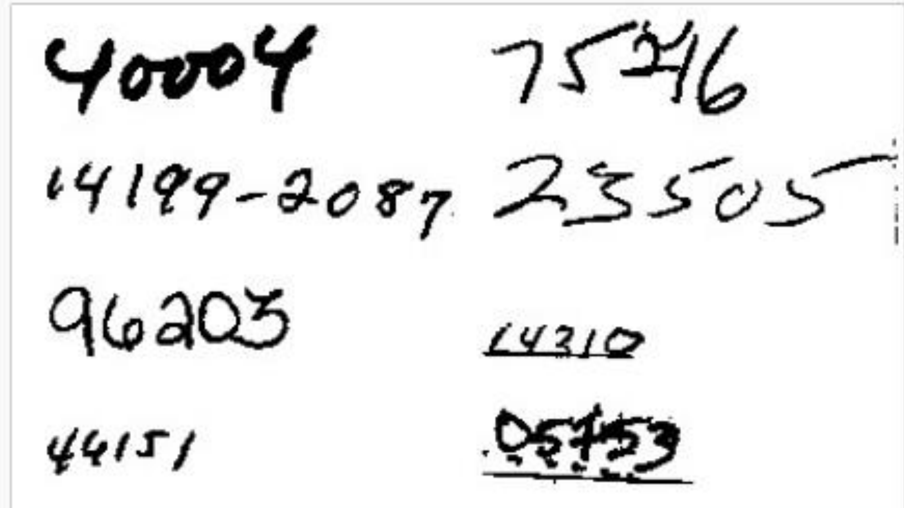


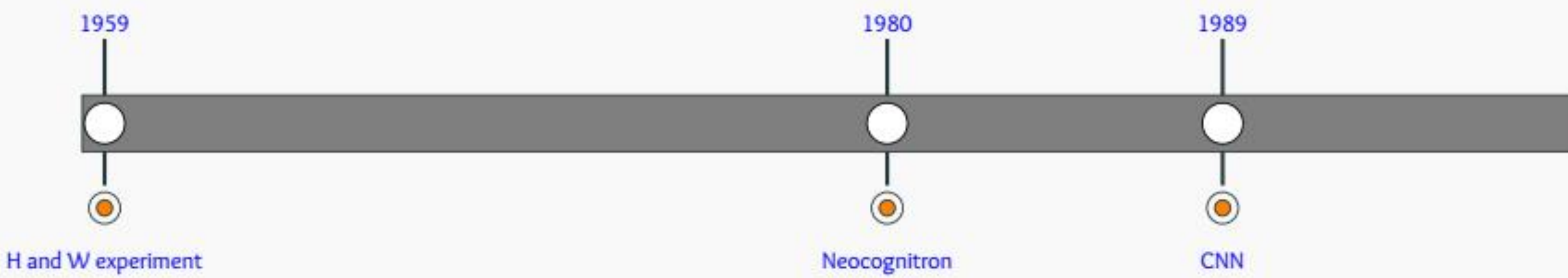


From Cats to Convolutional Neural Nets

Convolutional Neural Network

Handwriting digit recognition using backpropagation over a Convolutional Neural Network (LeCun et. al.) [20]

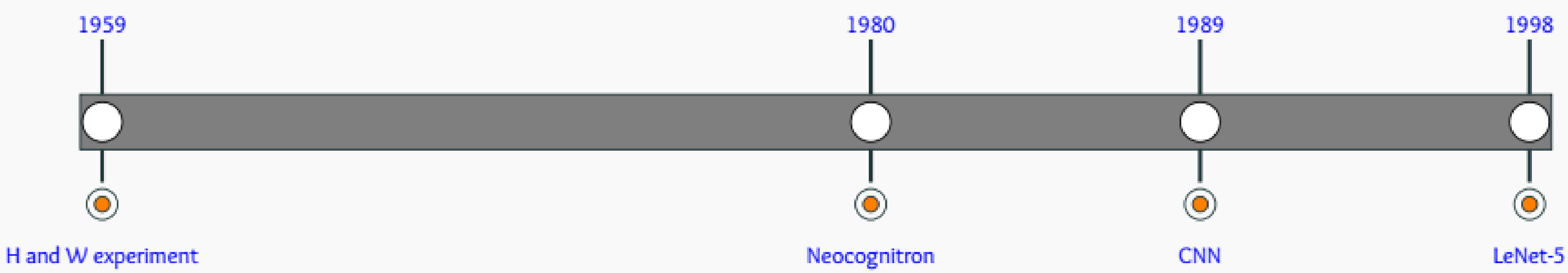




From Cats to Convolutional Neural Nets

LeNet-5

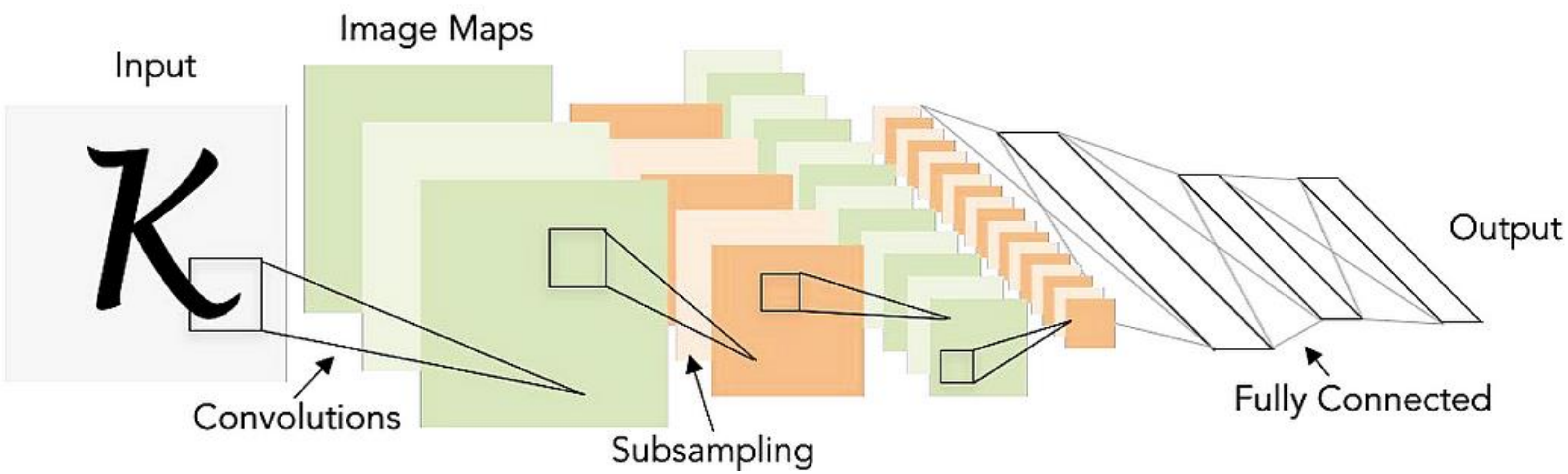
Introduced the (now famous) MNIST dataset (LeCun et. al.) [21]



- **CNN for Computer Vision**

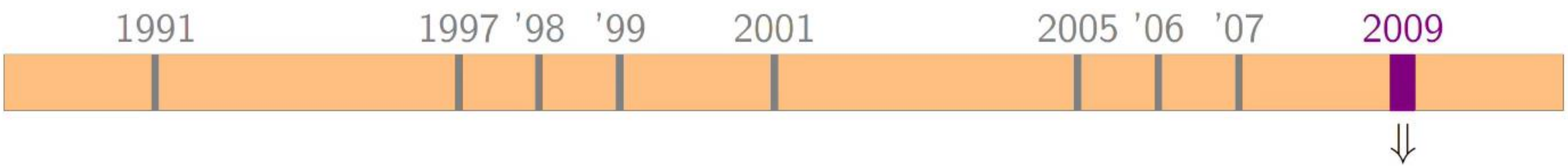
Next level understanding

- The name convolutional neural networks actually originated with the design of the LeNet by Yann LeCun and team, 1998.
- It was largely developed for the handwritten digit recognition task.



LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE*, *86*(11), pp.2278–2324

Next level understanding



- **Eigenfaces for face recognition** (Turk & Pentland, 1991)
- **Computational theories of object recognition** (Edelman, 1997)
- **Perceptual grouping, Normalized cuts** (Shi & Malik, 1997)
- **Particle filters, Mean shift** for tracking (Liu & Chen, 1998)(Cheng, 1998)
- **SIFT** (Lowe, 1999) (Lowe, 2004)
- **Viola-Jones face detection** (Viola & Jones, 2001)
- **Conditional Random Fields** (Lafferty et al, 2001)
- **Pictorial structures** revisited (Felzenszwalb & Huttenlocher, 2005)
- **PASCAL VOC** arrives; Scene/panorama/location recognition methods grow
- **Constellation models** (Fergus, Perona & Zisserman, 2007)
- **Deformable part models** (Felzenszwalb et al, 2009)

Next level understanding

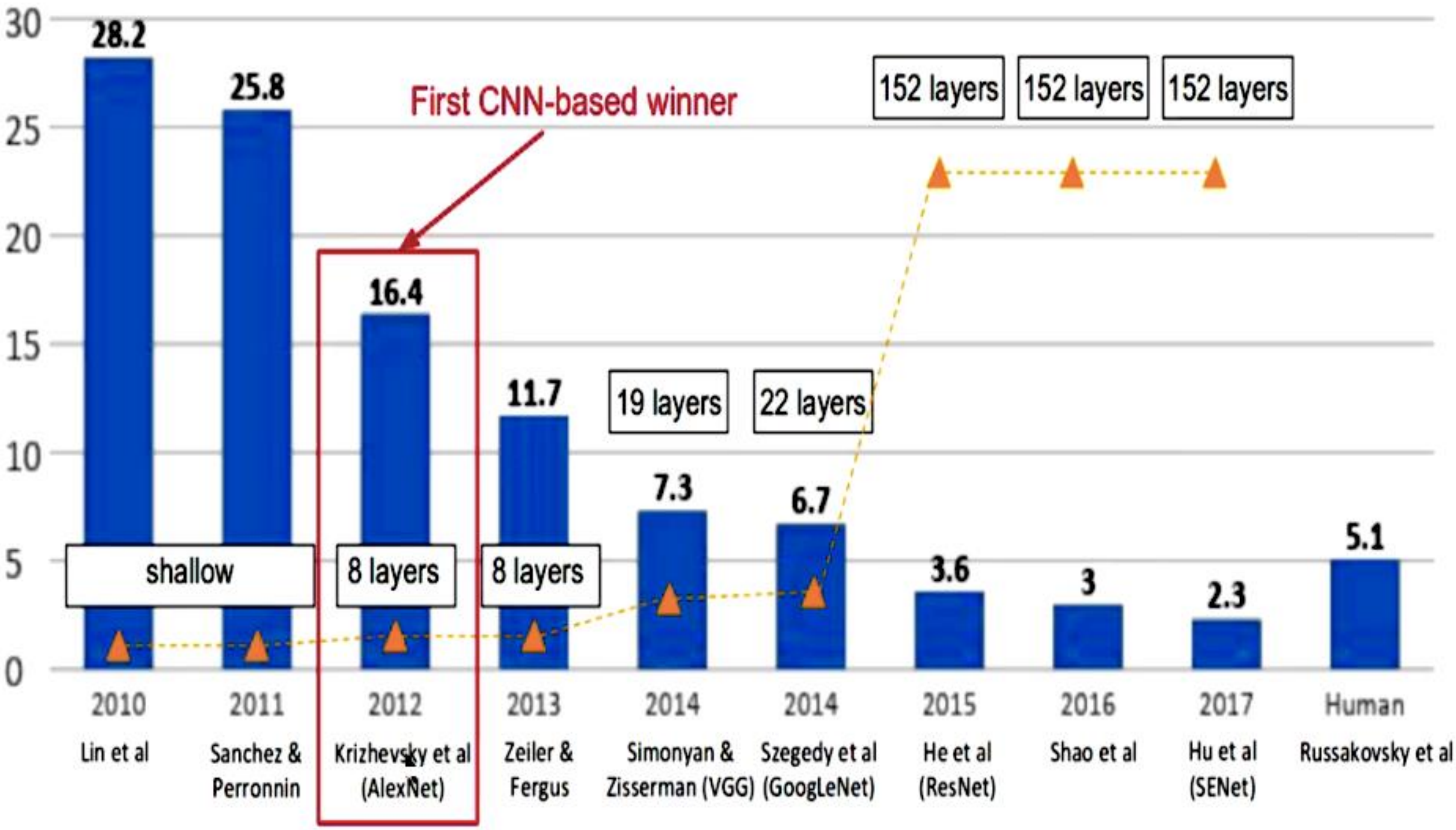
- ImageNet large scale visual recognition challenge (ILSVRC)

- ImageNet arrives

- The validation and test data for this competition will consist of 200,000 photographs
- collected from flickr and other search engines
- hand labeled with the presence or absence of 1000 object categories.

https://image-net.org/challenges/LSVRC/2010/#introduction

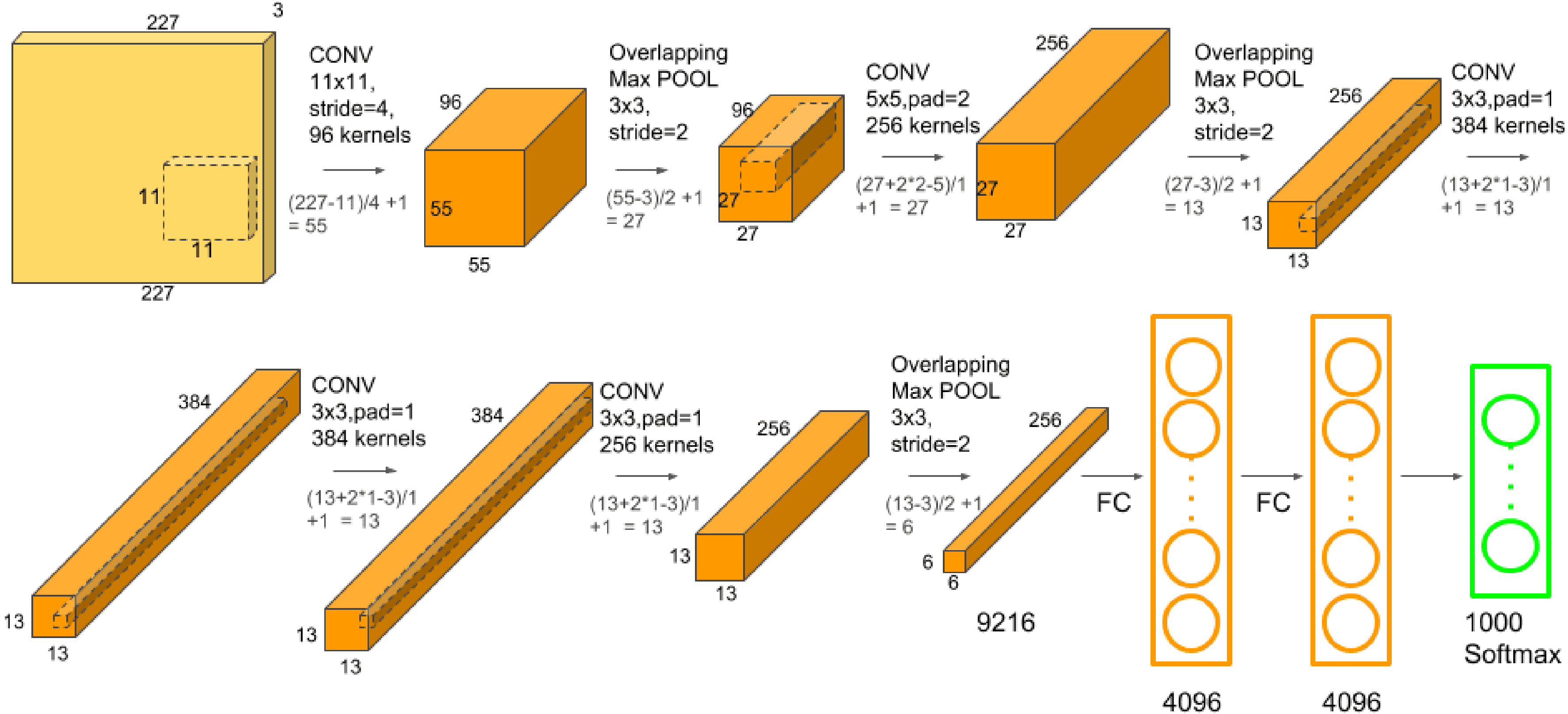
Next level understanding: 2015



https://medium.com/appyhigh-technology-blog/convolutional-neural-networks-a-brief-history-of-their-evolution-ee3405568597

Next level understanding: 2012

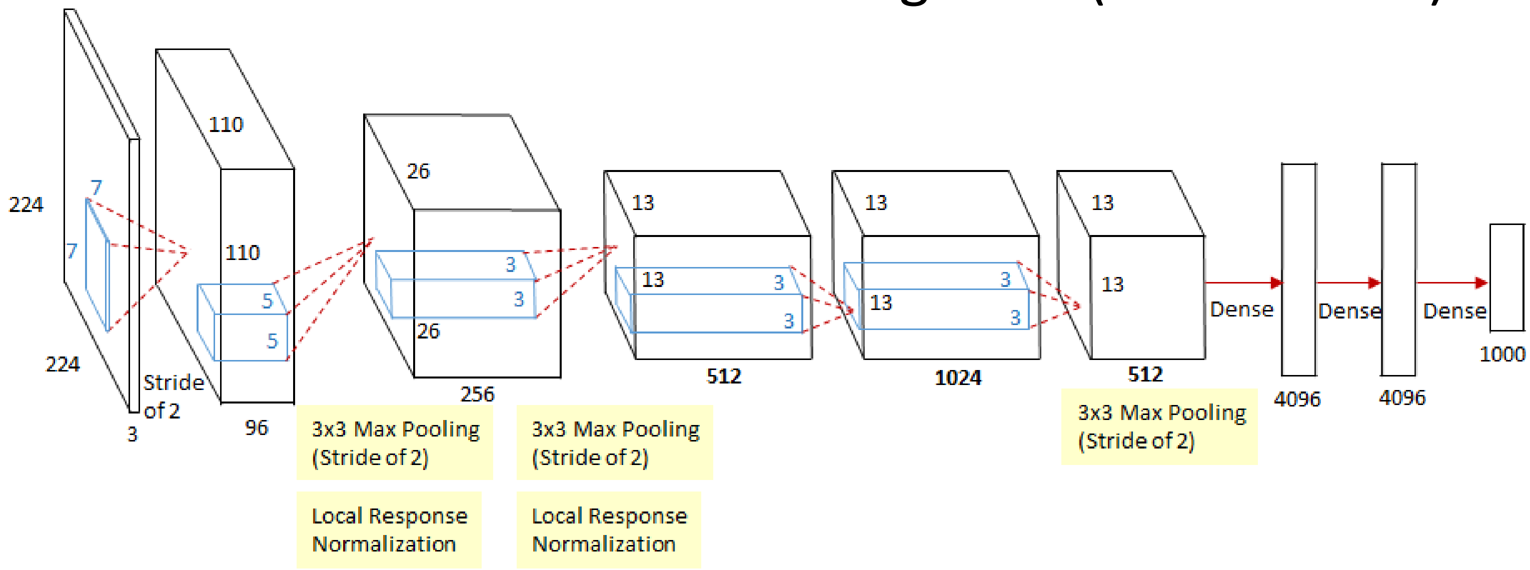
- AlexNet (Alex Krizhevsky and Geoffrey Hinton, 2012) winds the ImageNet challenge
- Before 2012, networks were shallow

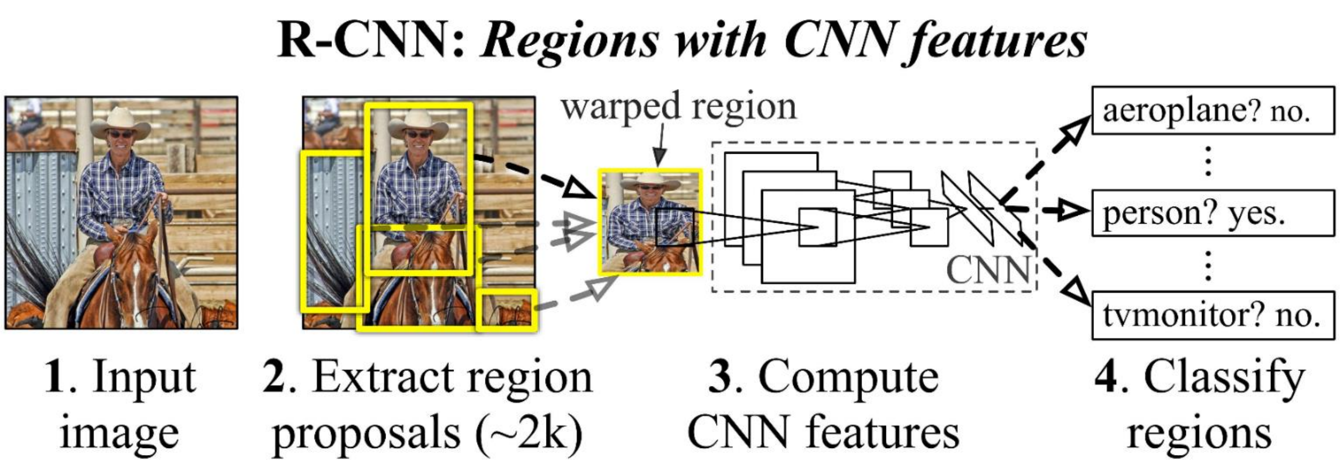


Krizhevsky, A., Sutskever, I. and Hinton, G.E., 2017. ImageNet classification with deep convolutional neural networks. *Communications of the ACM*, *60*(6), pp.84–90.

Next level understanding: 2013

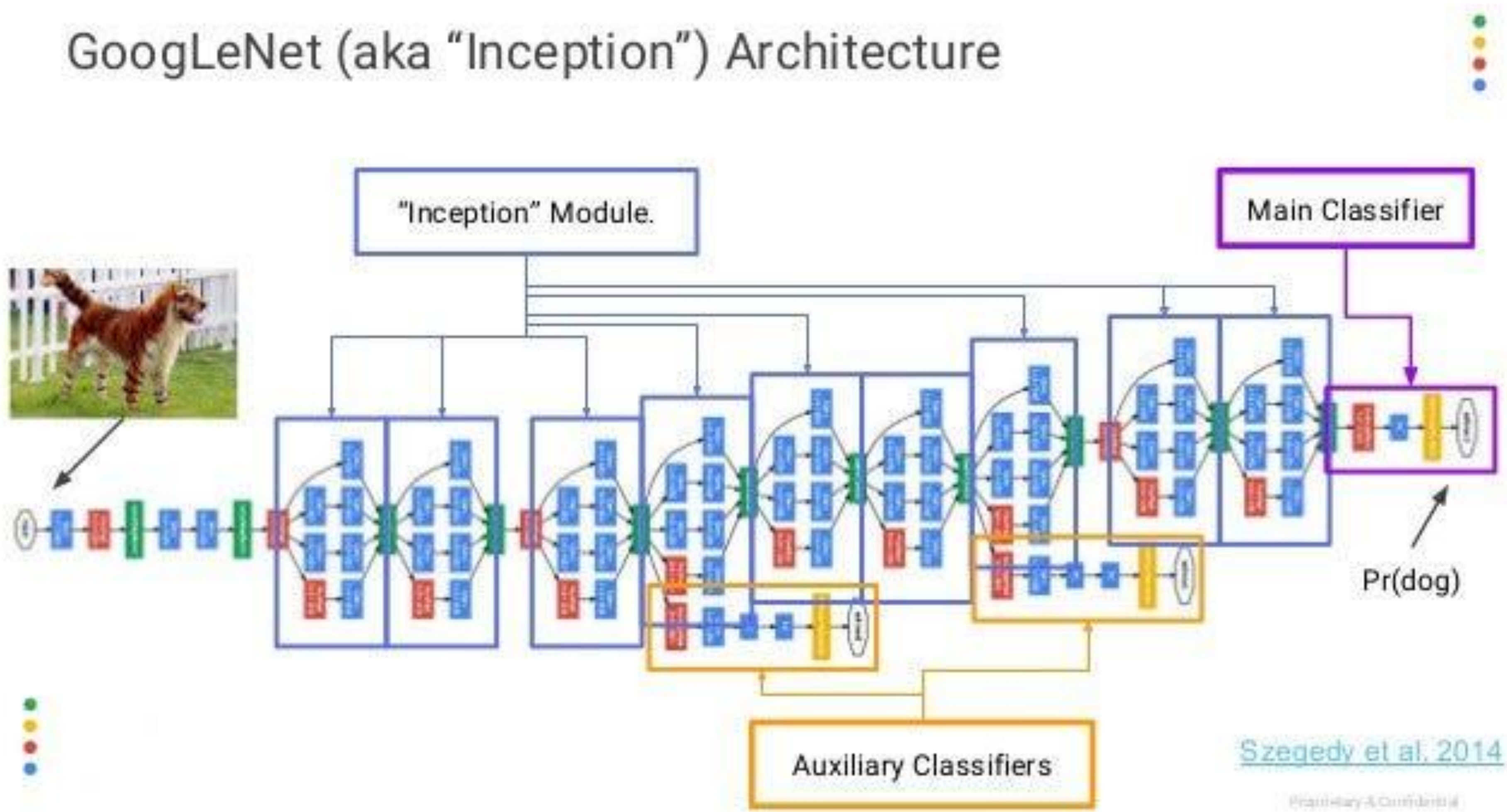
ZFNet: Matthew D Zeiler and Rob Fergus (ILSVRC 2013)





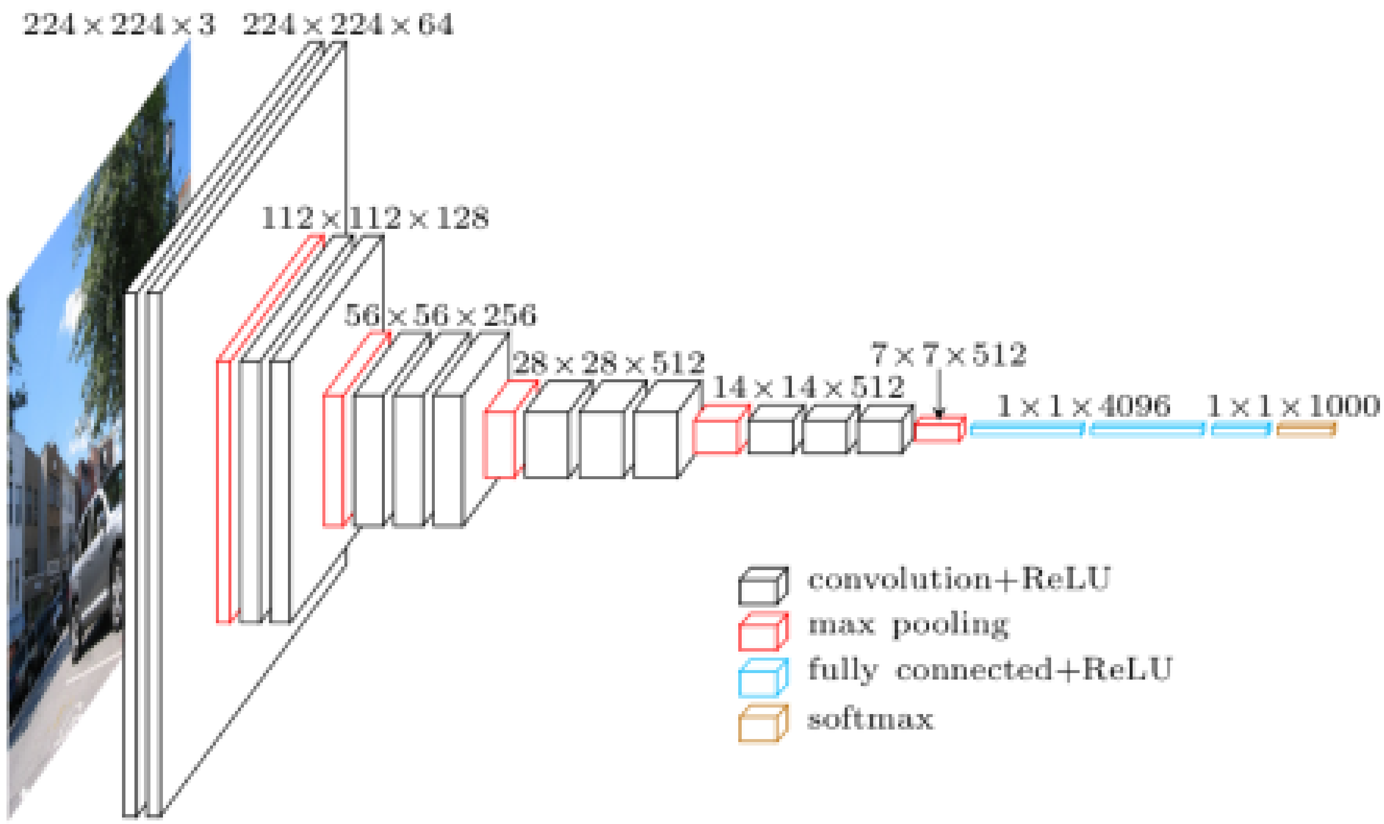
Next level understanding: 2014

- InceptionNet



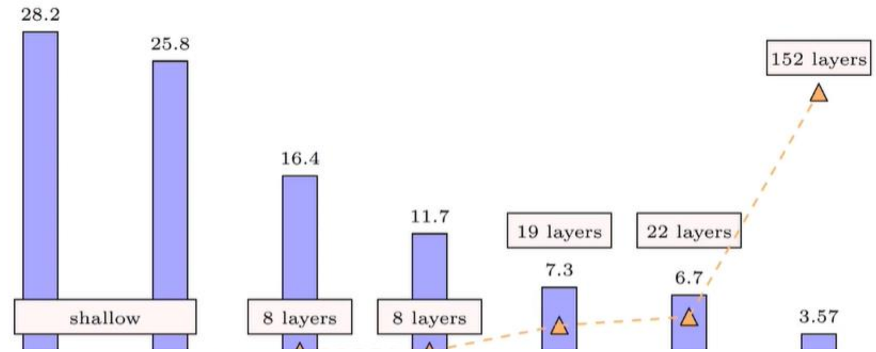
Next level understanding: 2014

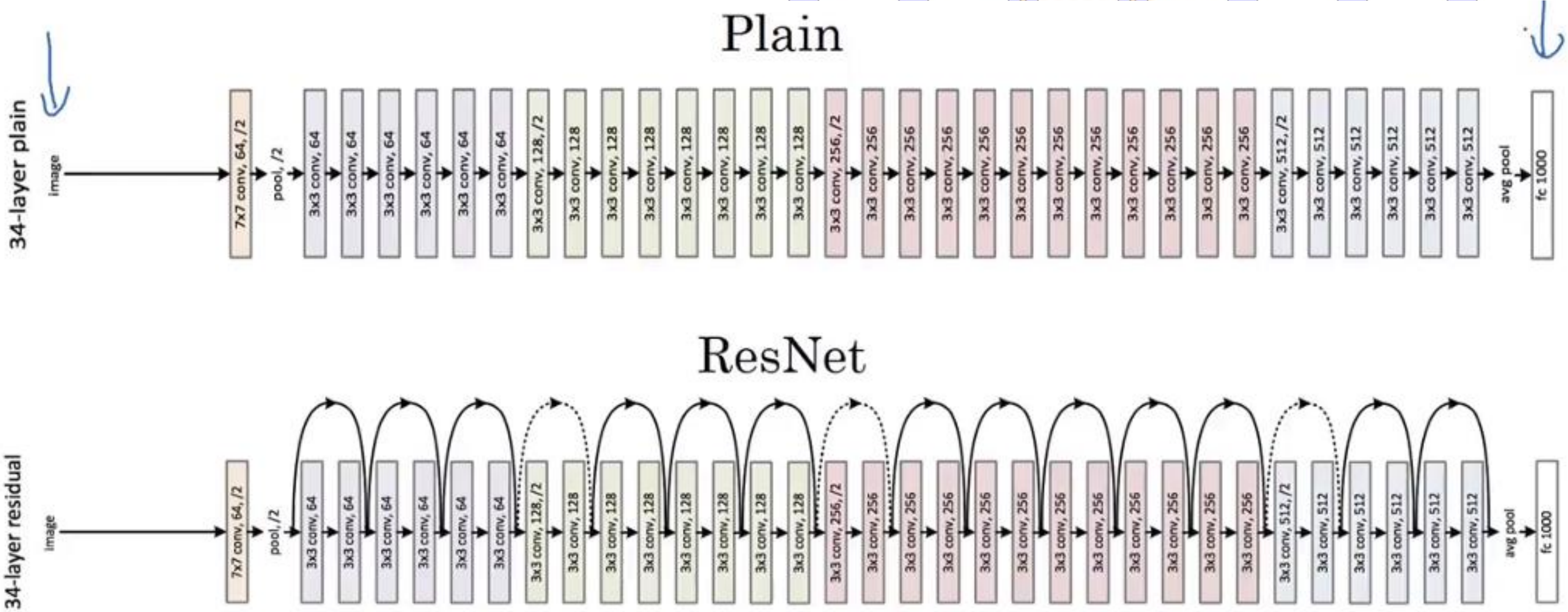
- VGG :invented by Visual Geometry Group (at Oxford University)



Next level understanding: 2015

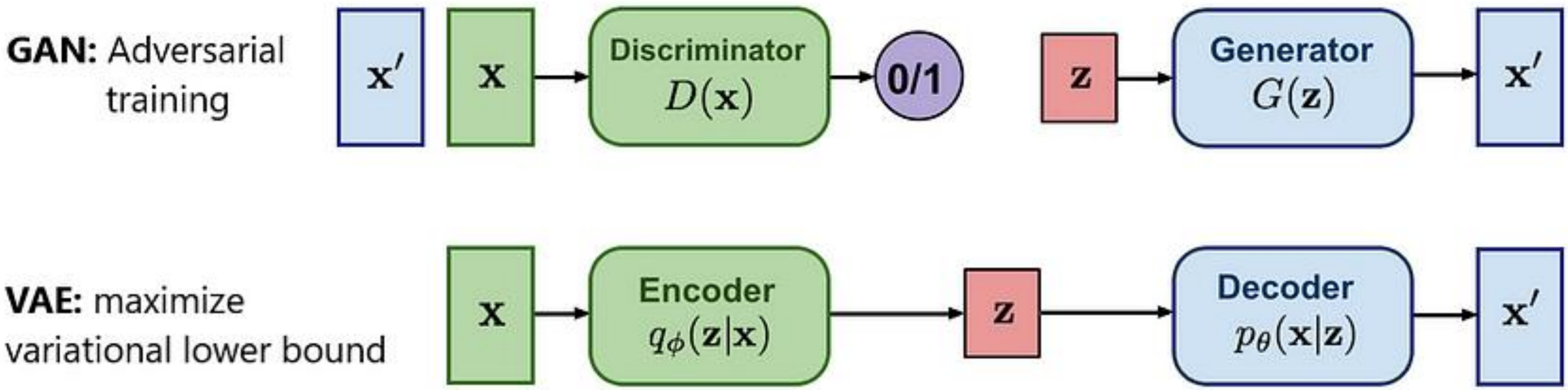
- ResNet: Kaiming He et. al. from Microsoft Research
- CNN somewhat matched human performance





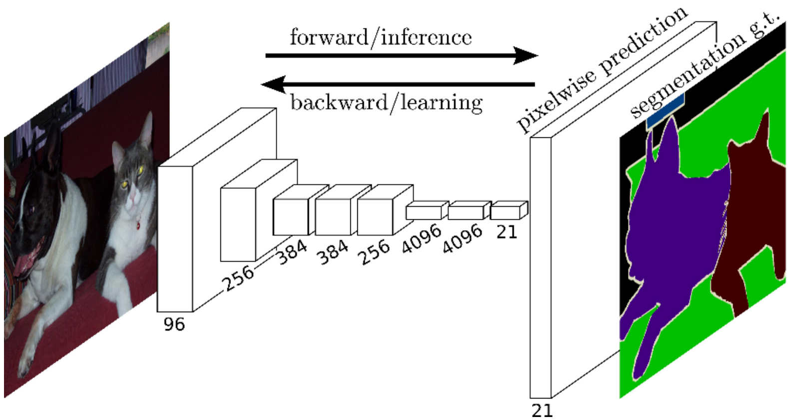
Next level understanding: 2014

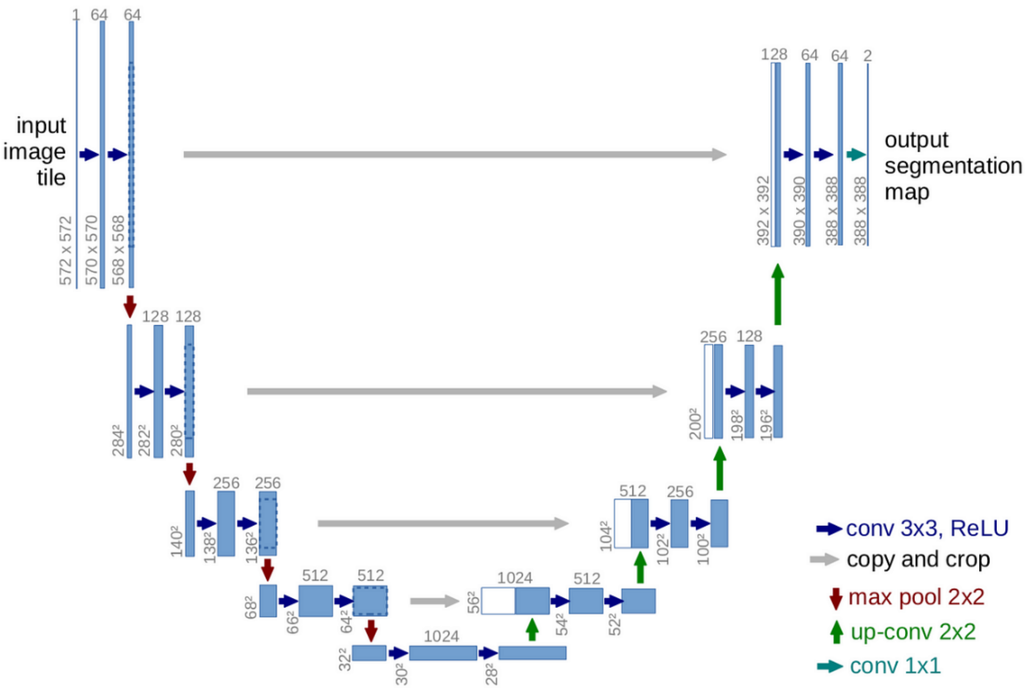
- Deep generative models: GANs, VAEs



Next level understanding: 2015

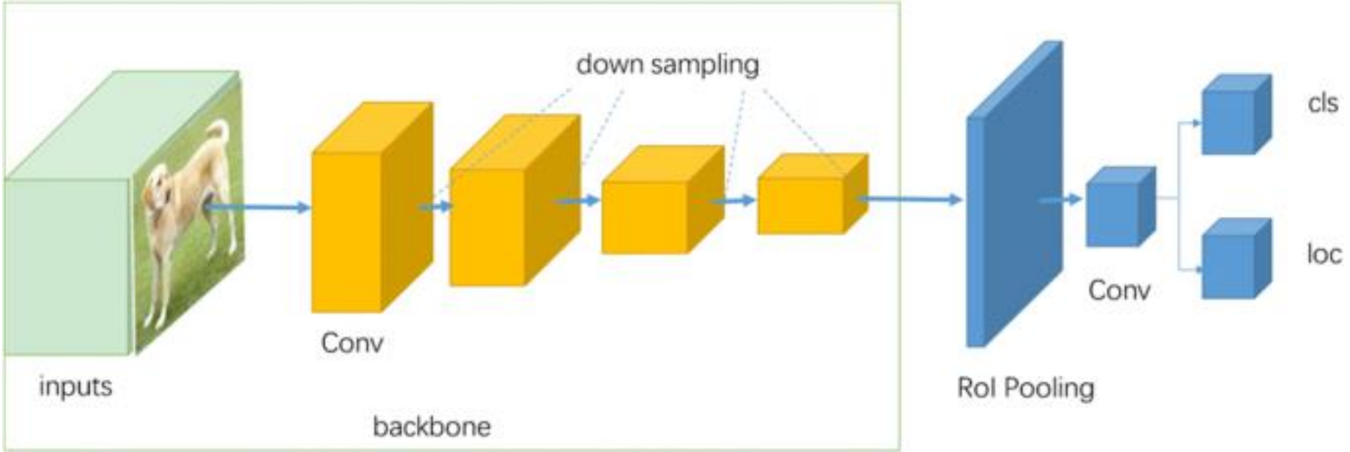
- Segmentation networks: FCN, SegNet and U-Net for semantic segmentation; COCO dataset arrives; VQA dataset arrives

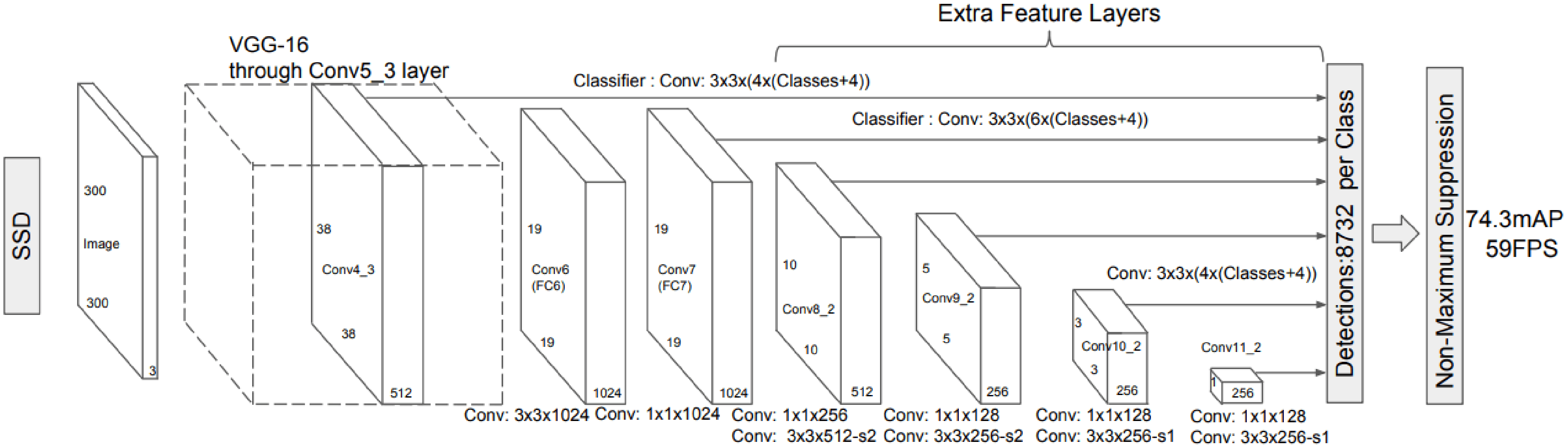




Next level understanding: 2016

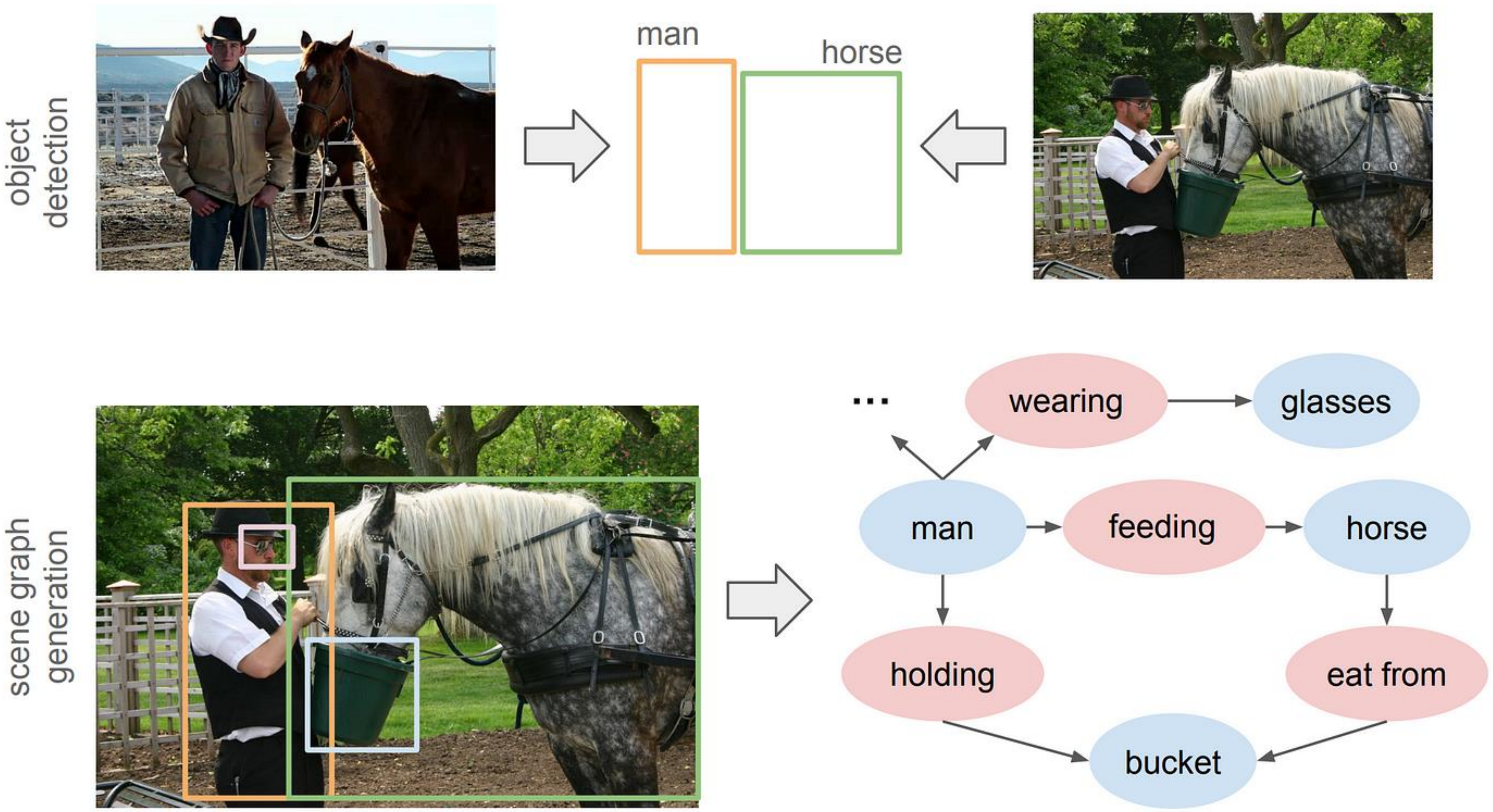
- YOLO and SSD for object detection;
- Cityscapes dataset arrives, Visual Genome dataset arrives





Next level understanding: 2017

- Scene Graph Generation model



Next level understanding: 2018, 2019

- VCR (Visual Common sense Reasoning) dataset
- Panoptic segmentation

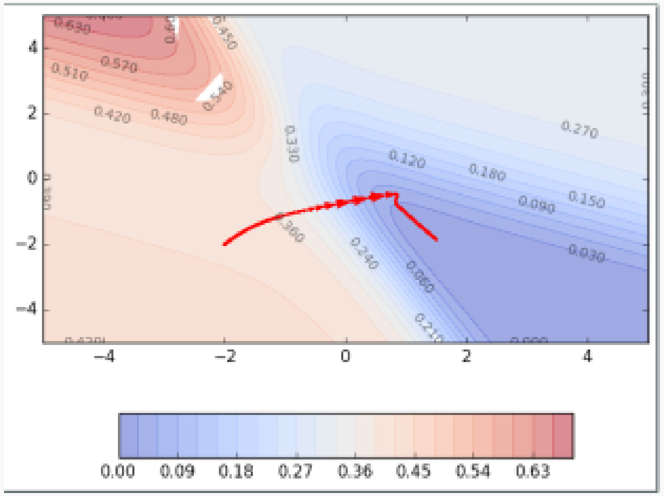
- **Faster …. Better….**

Better Optimization Methods

- Faster convergence, Higher accuracy, Stronger

Better Optimization Methods

Faster convergence, better accuracies

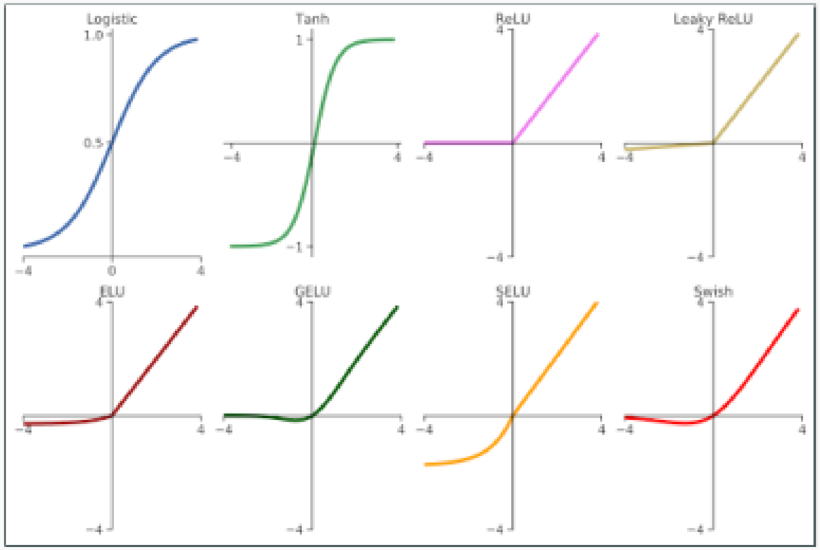


Better Activation Functions

Better Activation Functions

We have come a long way from the initial days when the logistic function was the default activation function in NNs!

Over the past few years many new functions have been proposed leading to better convergence and/or performance!



Sequences

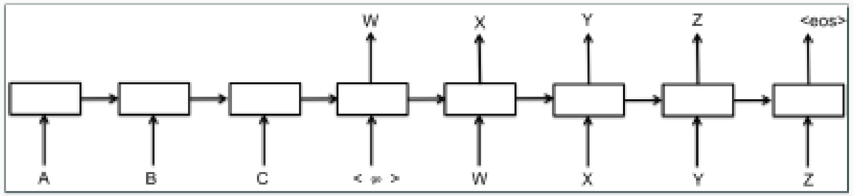
Sequences

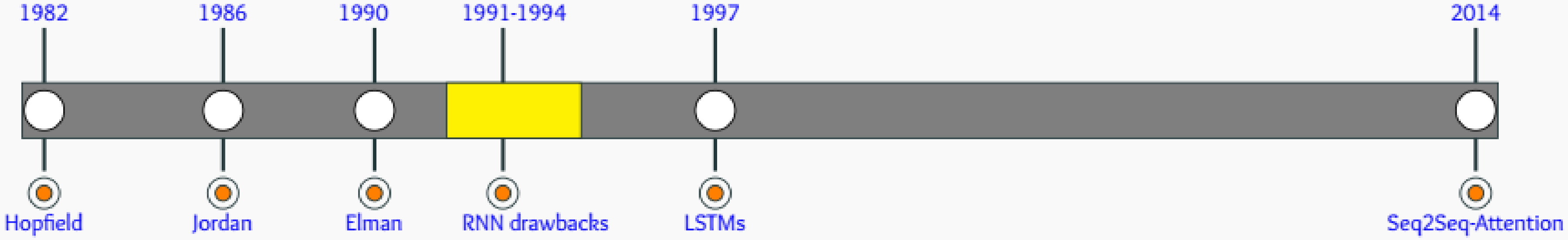
- They are everywhere
- Time series, speech, music, text, video
- Each unit in the sequence interacts with other units
- Need models to capture this interaction

Sequences

Sequence To Sequence Models

- Initial success in using RNNs/LSTMs for large scale Sequence To Sequence Learning Problems
- Introduction of Attention which is perhaps the idea of the decade!



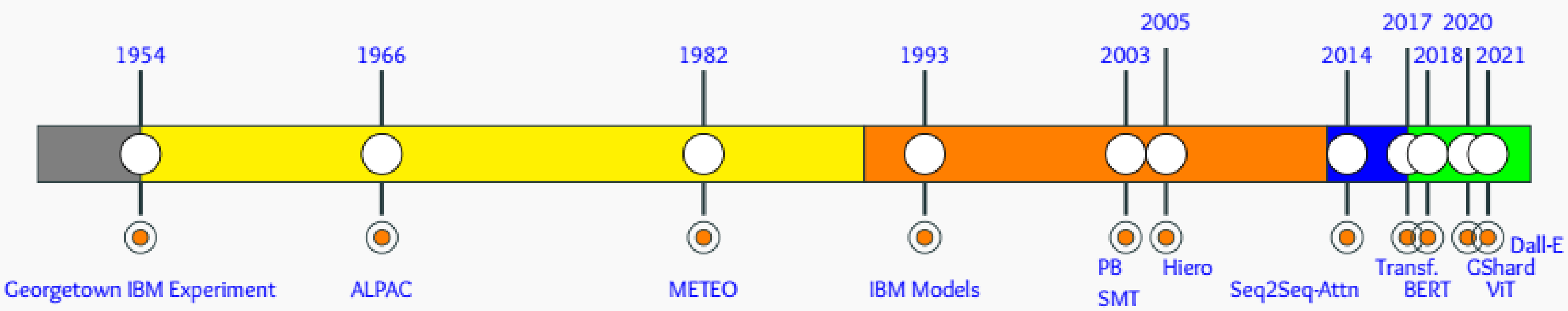


The Rise of Transformers

From Language To Vision

DALL·E[a] is a 12-billion parameter version of GPT-3 trained to generate images from text descriptions, using a dataset of text–image pairs.

[a]https://openai.com/blog/dall-e/

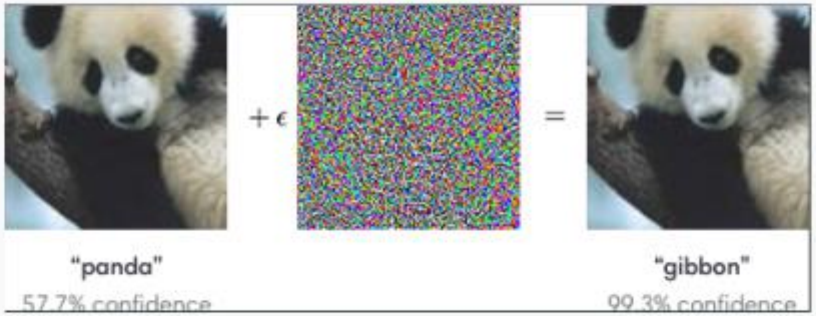


- **Explainable AI**

The Paradox of Deep Learning

Why does deep learning work so well despite

- high capacity (susceptible to overfitting)
- numerical instability (vanishing/exploding gradients)
- sharp minima (leading to overfitting)
- non-robustness (see figure)



No clear answers yet but ...

- Slowly but steadily there is increasing emphasis on explainability and theoretical justifications!*
- Hopefully this will bring sanity to the proceedings !

Tell me why!

Clever Hans was a horse that was supposed to be able to do lots of difficult mathematical sums and solve complicated problems. Turns out, it was giving the right answers by watching the reactions of the people watching him.

A repository to benchmark machine learning systems' vulnerability to adversarial examples.

2016

WHI
CleverHans

Be Fair and Responsible!

“There’s software used across the country to predict future criminals. And it’s biased against blacks.” - Propublica

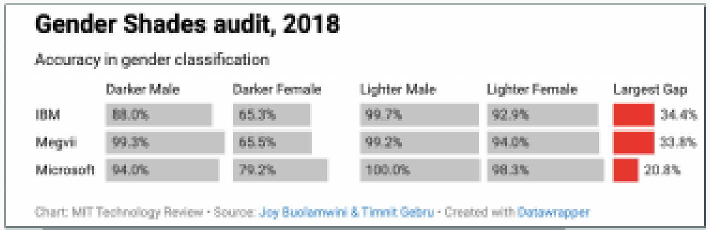
Source: https://www.propublica.org/article/machine-bias-

2016

Machine Bias

Be Fair and Responsible!

"Facial Recognition Is Accurate, if You're a White Guy" - MIT Media



Source: Joy Buolamwini (Youtube)

2016 — Machine Bias

2018 — Gender/Race Bias

Be Fair and Responsible!

"Due to our concerns about malicious applications of the technology, we are not releasing the trained model." – – *OpenAI*

Learn From Anyone

Teacher

steve jobs

Me: What's an iPhone

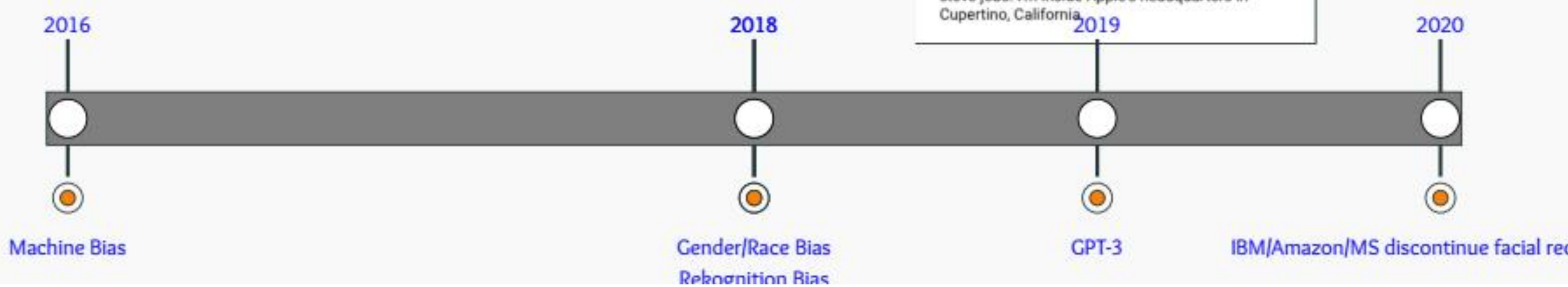
? steve jobs: iPhone is a mobile device. It's the most sophisticated smart phone in the world.

Me: What do you think about Windows?

adam@17 819 on Jun 16, 2016 steve jobs: Windows is a wonderful desktop computer platform that Microsoft has provided.

Me: Where are you right now?

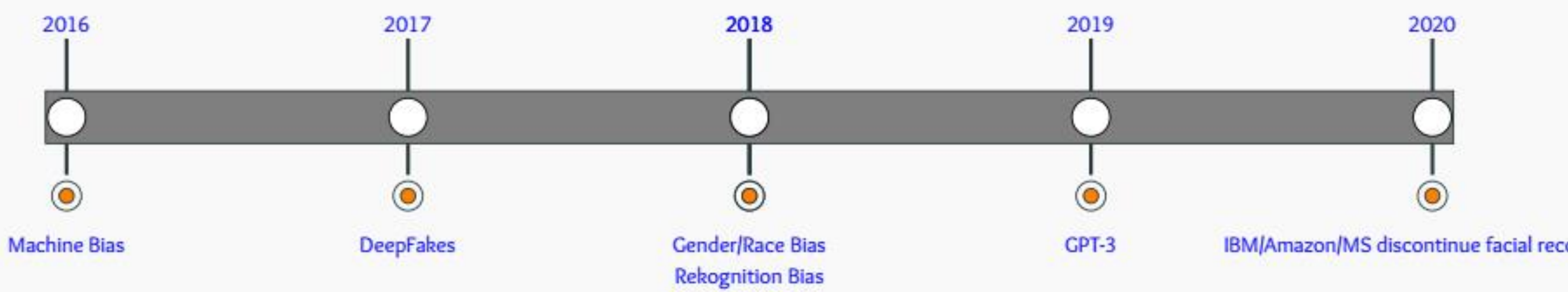
steve jobs: I'm inside Apple's headquarters in Cupertino, California.



Interpretable, Responsible & Green AI

Be Fair and Responsible!

What started off as an innocuous project for mimicking facial expressions has since lead to many apps and creation of fake videos for blackmailing, pronography and swaying elections!

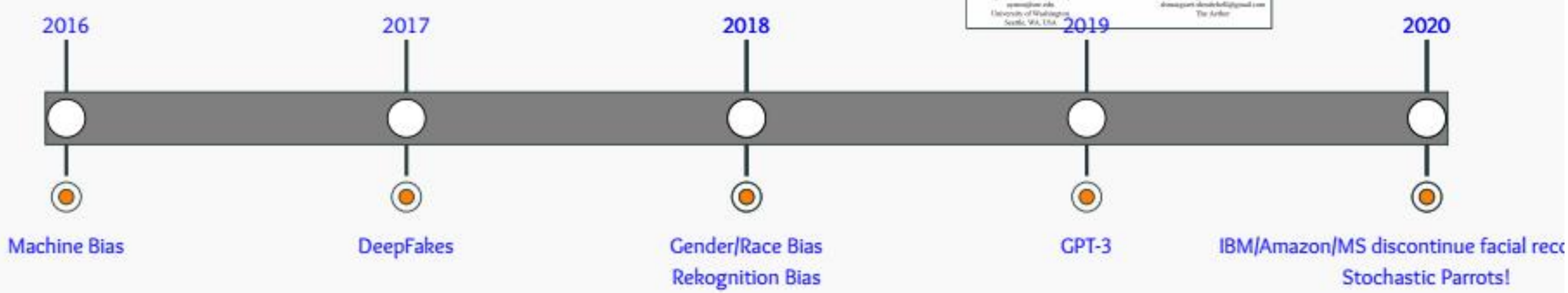


Be Fair and Responsible!

"Models are only as good as the data. Be responsible while curating data." – *Bender et. al.*


On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?

Emily M. Bender — Timnit Gebru — Angelina McMillan-Major — Shmargaret Shmitchell

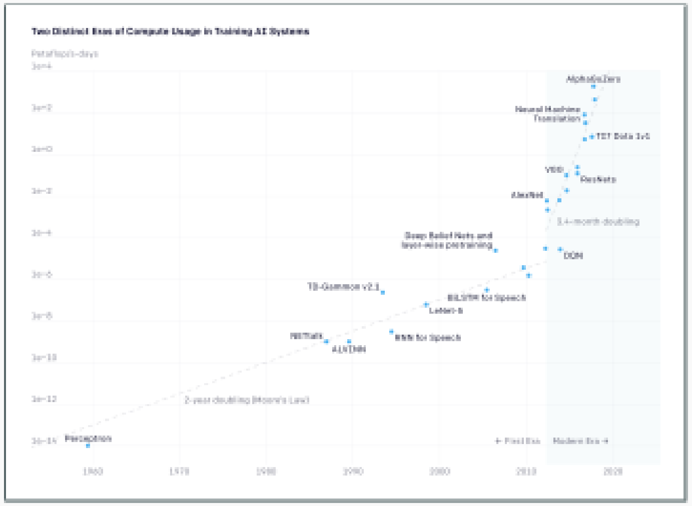


Interpretable, Responsible & Green AI

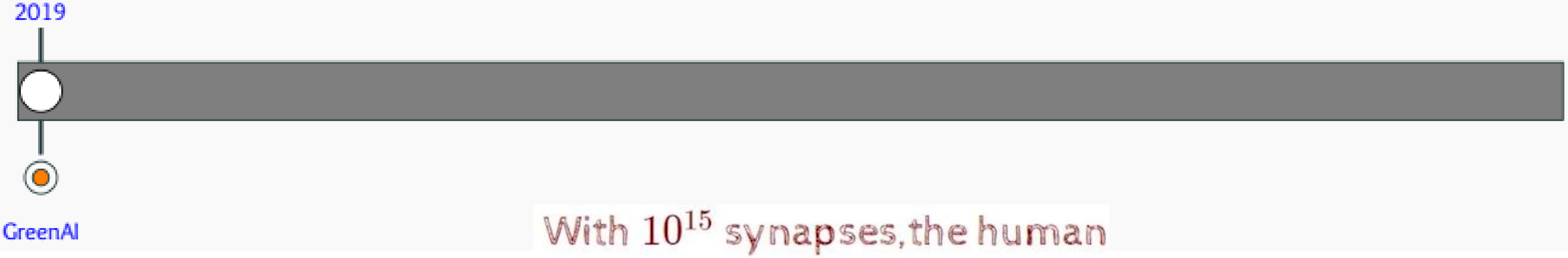
Push for Green AI

The computations required for deep learning research have been doubling every few months, resulting in an estimated 300,000x increase from 2012 to 2018 – *AllenAI*

Ironically, deep learning was inspired by the human brain, which is remarkably energy efficient.



https://openai.com/blog/ai-and-compute/

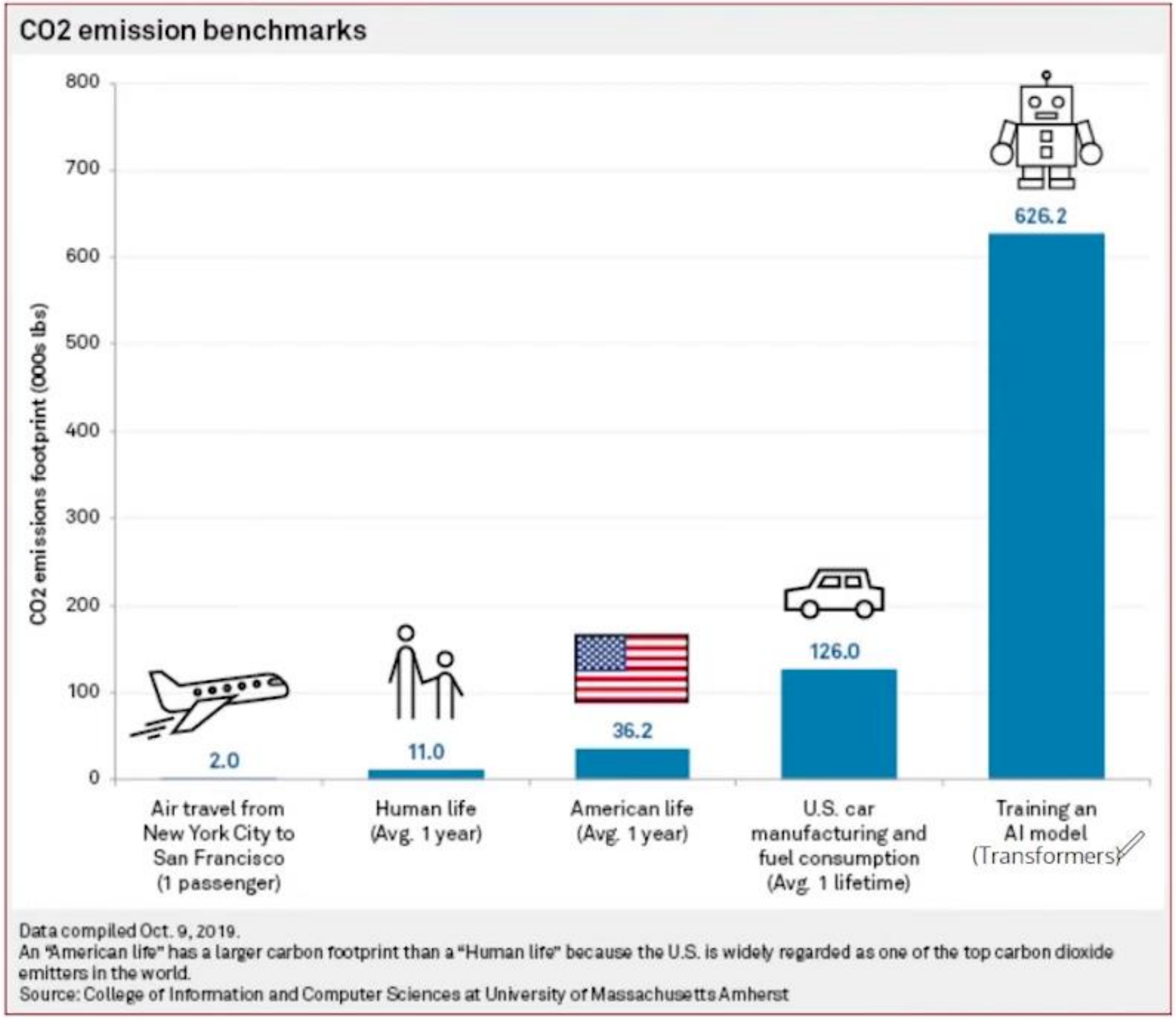


With 10^{15} synapses, the human brain consumes only 15 watts of power [Ref]

Interpretable, Responsible & Green AI

Push for Green AI

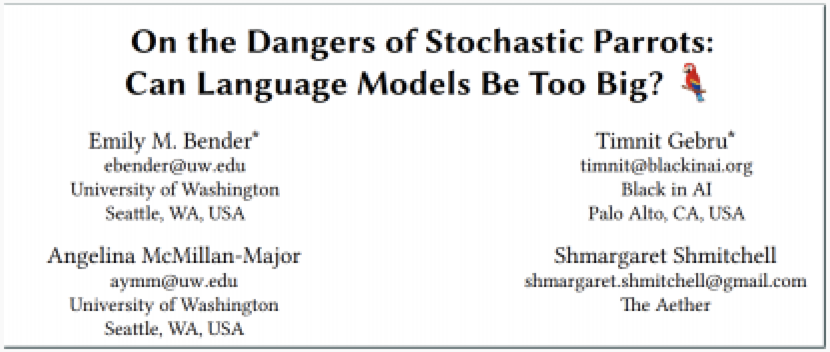
Call for energy and policy considerations for Deep Learning



Interpretable, Responsible & Green AI

Push for Green AI

"Is it fair that the residents of the Maldives (likely to be underwater by 2100) or the 800,000 people in Sudan affected by drastic floods pay the environmental price of training and deploying ever larger English LMs, when similar large-scale models aren't being produced for Dhivehi or Sudanese Arabic?" – *Bender et. al.*



On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?

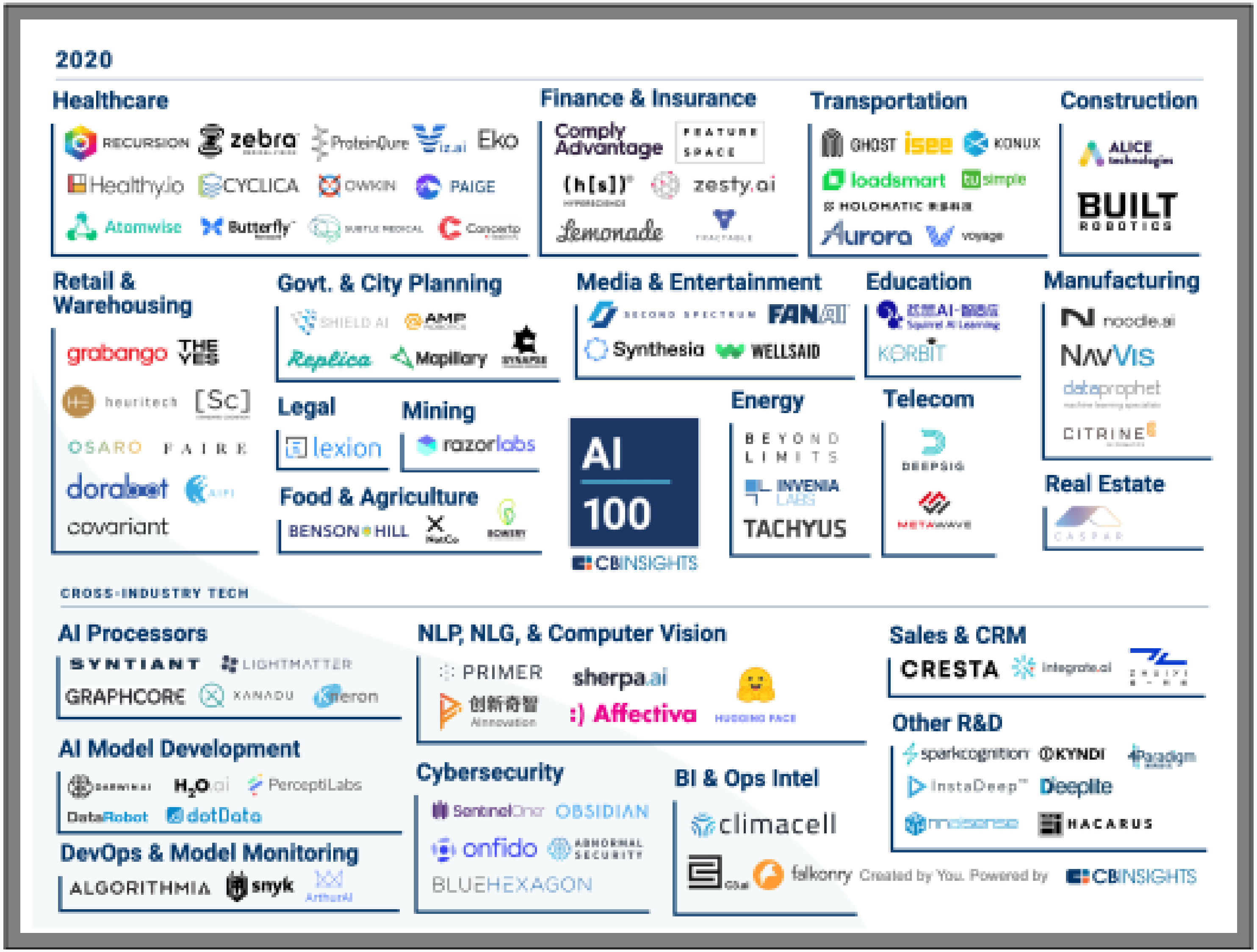
Emily M. Bender*
ebender@uw.edu
University of Washington
Seattle, WA, USA

Timnit Gebru*
timnit@blackinai.org
Black in AI
Palo Alto, CA, USA

Angelina McMillan-Major
aymm@uw.edu
University of Washington
Seattle, WA, USA

Shmargaret Shmitchell
shmargaret.shmitchell@gmail.com
The Aether

AI Revolutionize Scientific Research



Source: https://www.cbinsights.com/research/artificial-intelligence-top-startups/

- **End of topic**