CMPE 297-11
Deep Learning

Machine Learning

- *Field of study that gives computers the ability to* ***learn without being explicitly programmed***

By Arthur Samuel (1959)

Samuel’s Checker Program vs. Human (1956)

The IBM 700 Series
Computing Comes to Business

Overview
Transforming the World
Cultural Impacts
The Team
In Their Words

01/02

IBM 701 meets the future president

In 1954, Ronald Reagan, who was a TV personality for General Electric at the time, visited the GE Aircraft Jet Engine Plant in Evendale, Ohio. During this visit, GE manager Herbert Grosch spent a few minutes introducing the future US president to the IBM 701.

The IBM 700 series made scientific calculations and commercial operations easier, but the machines also provided the world with some entertainment.

Playing checkers on the 701
On February 24, 1956, Arthur Samuel’s Checkers program, which was developed for play on the IBM 701, was demonstrated to the public on television. In 1962, self-proclaimed checkers master Robert Nealey played the game on an IBM 7094 computer. The computer won. Other games resulted in losses for the Samuel Checkers program, but it is still considered a milestone for artificial intelligence, and offered the public in the early 1960s an example of the capabilities of an electronic computer.

01/02

IBM Deep Blue vs. Kasparov (1996 & 1997)

Kasparov has Deep Blues after losing

Chess champ: I was rooked

By MICHELE McPHEE, K.C. BAKER and CORKY SIEMASZKO
Daily News Staff Writers

The world's greatest human chess player threw a tantrum and cried foul yesterday after being thrashed by a supercomputer.

It took IBM's Deep Blue just 19 moves to defeat world chess champion Garry Kasparov — a stunning finale to an epic week-long battle of man versus machine.

Not mollified by his $400,000 loser's share, Kasparov stormed off like a sore loser after resigning. He later accused IBM of unfairly programing the high-speed computer to beat him specifically.

He suggested that Deep Blue, which was supposed to play on its own, was coached during the match. He stopped short of saying the computer team cheated.

"I suspect there were things in the match that were well beyond my understanding," Kasparov said. "And when a big corporation with unlimited resources would like to do so, there are many ways to achieve the result, and the result was achieved."

IBM team leader C.J. Tan denied the computer was coached. "Once the clock started, it relied on Deep Blue's system itself," he said.

Kasparov's pal, Michael Khodarkovski, blamed Kasparov's graceless exit on a lack of practice — he said Kasparov had never lost a match.

Kasparov came close to losing to Anatoly Karpov in a 1984-85 championship match that was suspended without a victory on either side.

Kasparov, 34, considered by some chess experts as the greatest player in the history of the game, last year defeated Deep Blue 4-2.

After losing the opening game of the rematch at the Equitable Center in Manhattan, the computer won the second game and fought Kasparov to draws in the next three.

Then yesterday — with a swiftness that stunned the chess world — Deep Blue took advantage of Kasparov's clumsy opening moves and placed him in a no-win situation after less than an hour of play.

Unable to find a way out, Kasparov — playing the black pieces — tipped his king and resigned. He buried his head in his hands and didn't look at IBM's Tan when they shook hands.

The final score was 3½ points for the computer and 2½ points for Kasparov.

Kasparov said he "cracked under the pressure."

"I am ashamed," said Kasparov, who would have won $700,000 if he had beaten the computer.

Patrick Wolff, author of "The Complete Idiot's Guide to Chess," said the world champ "basically cracked."

Kasparov, playing black, used a standard defense known as the "Caro-Kann," forcing white to sacrifice a piece. But for some reason he botched his seventh move and "he became lost," Wolff said.

"This is not a position he wanted to get into," said Ilya Gurevich, a grand master from Manhattan. "It's a pure calculating position where the computer has a big advantage. The computer's strength is tactics."

The computer Kasparov battled was capable of analyzing 200 million positions per second — twice as many positions per second as the IBM model he defeated in Philadelphia last year.

One expert said he was surprised when Kasparov resigned. "It didn't seem lost," said grand master John Fedorowicz of the Bronx, who helped the IBM team prepare its game plan.

At Chess Forum on Thompson St. in Greenwich Village, die-hard chess fans expressed shock at Kasparov's loss.

"This is a historic event," said Mark Wieder, 46, also a computer programer. "The greatest human player of all time lost to a machine."

Chess Forum owner Imad Khachan, 31, said Kasparov was following in the footsteps of other sore losers by suggesting his foe didn't play fair.

"This is not uncommon in chess," he said. "When Viktor Korchnoi was playing Karpov in the '70s, Korchnoi made the accusation that the KGB was sending him telepathic messages to destroy his concentration."

LARRY ZUMWALT

RESIGNED: Chess champion Garry Kasparov is disappointed yesterday after losing to IBM's Deep Blue supercomputer. IBM's team leader C.J. Tan (left) denied the computer had been coached, as Kasparov charged after the historic loss in Manhattan. Kasparov will get $400,000 for his efforts.

Artifical intelligence not black and white

MARK KRIEGEL

FORGET ABOUT THE Garden or the Meadowlands. The real action was outside the Equitable Center on Seventh Ave., where Garry Kasparov, with a name like a hockey player, did battle with Deep Blue, an IBM supercomputer whose name suggests some starlet who did her best work on 42d St. in those halcyon days before Disney.

The scalpers were asking as much as $500 for a $25 seat. "Actually, I'd settle for a couple of hundred," said Ze Ayala. "I have extras."

In the history of New York, there's never been a scalper so hopelessly well mannered as Ze Ayala, Ph.D. Instead of the usual hawker's cry — "Who got tickets?" — Ayala was content to let the business come to him as he burnished his new henna tattoo with a cotton ball doused in lemon juice. "The lemon juice helps the absorption of the dye," he said. The tattoo bore the name of his band, "Flashpot," for whom the tall, long-haired Ayala plays guitar. It should be noted that in lieu of a day job, he works at the Institute of Molecular Evolutionary Genetics at Penn State. His mission wasn't even mercenary; it was professional. He said he wouldn't have sold his own seat for a million bucks.

"My field is artificial intelligence," he said. "And this is the coolest thing to happen in my lifetime."

In such a spirit, the 28-year-old scientist had come to witness the inevitable coolness: Man mangled by Machine.

He couldn't help but root for Kasparov. The sentimental part of him was taken with the charms of obsolescence. But he knew better than to bet against technology.

Kasparov is the best chess player in the world; but unlike Deep Blue, he can also be vain, angry, neurotic, panicked, fearful, in all, human. "Chess is fundamentally psychological," Ayala said. "And that's precisely what Kasparov has working against him."

I asked him how long before Deep Blue is playing lead guitar in his band.

That day will come, he said, and it won't be too long. "You know the band Nine Inch Nails?" he said. "That's all computers. But what you're really asking is: Will a computer be able to write

SEE KRIEGEL PAGE 28

Monday, May 12, 1997 • DAILY NEWS

The New York Daily News edition May 12, 1997 edition is seen.

(THE NEW YORK DAILY NEWS)

HISTORY TECHNOLOGY

Did Deep Blue Beat Kasparov Because of a System Glitch?

Jennifer Latson @JennieLatson | Feb. 17, 2015

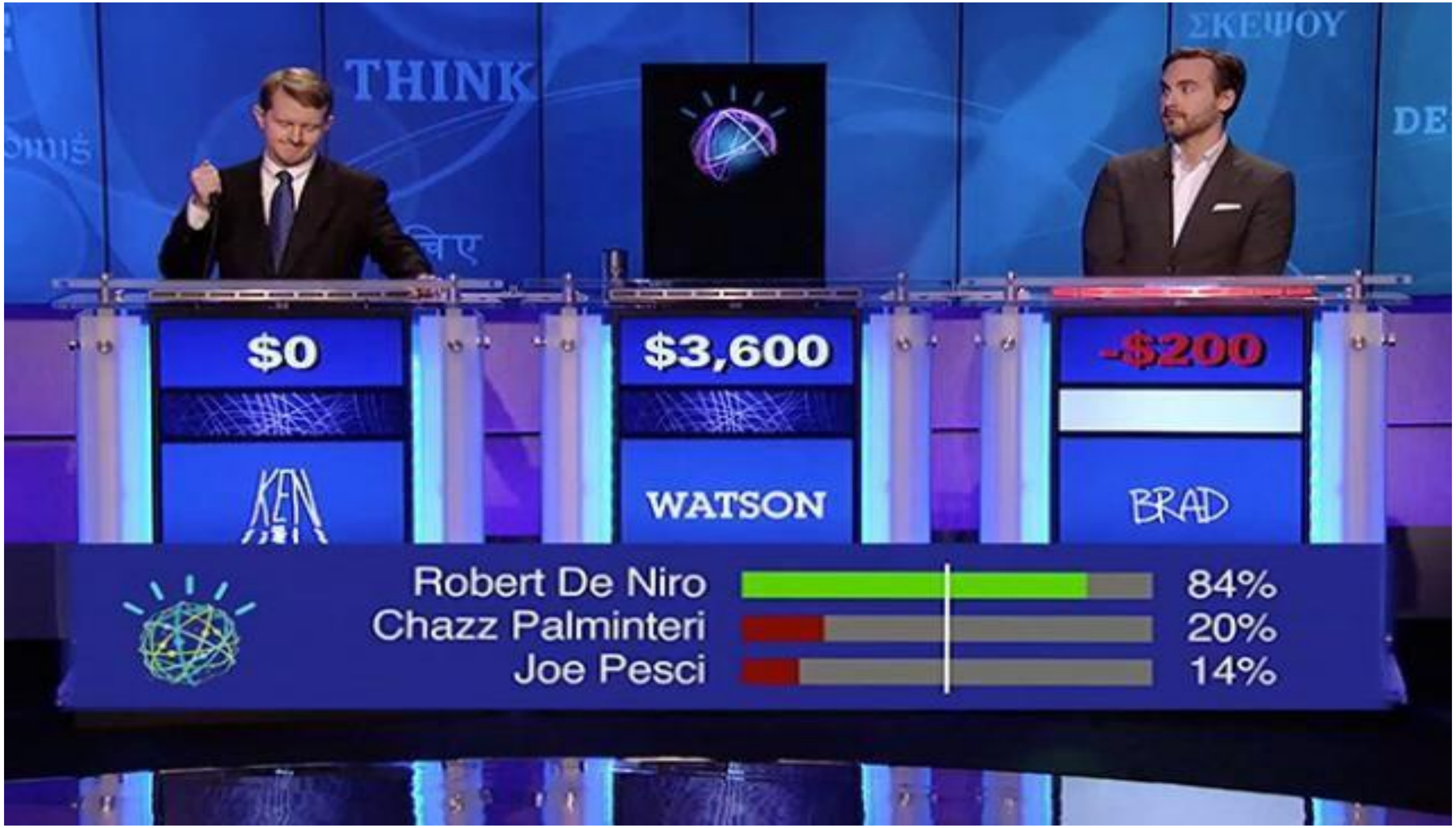
Feb. 17, 1996: Chess champion Garry Kasparov beats the IBM supercomputer "Deep Blue," winning a six-game chess match

Garry Kasparov was not afraid of a computer. When the world chess champion agreed to play a match against Deep Blue, the IBM supercomputer designed to beat him, he was so confident that, according to TIME, he scoffed at an offer to split the $500,000 purse 60-

Ted Thai—The LIFE Picture Collection/Getty

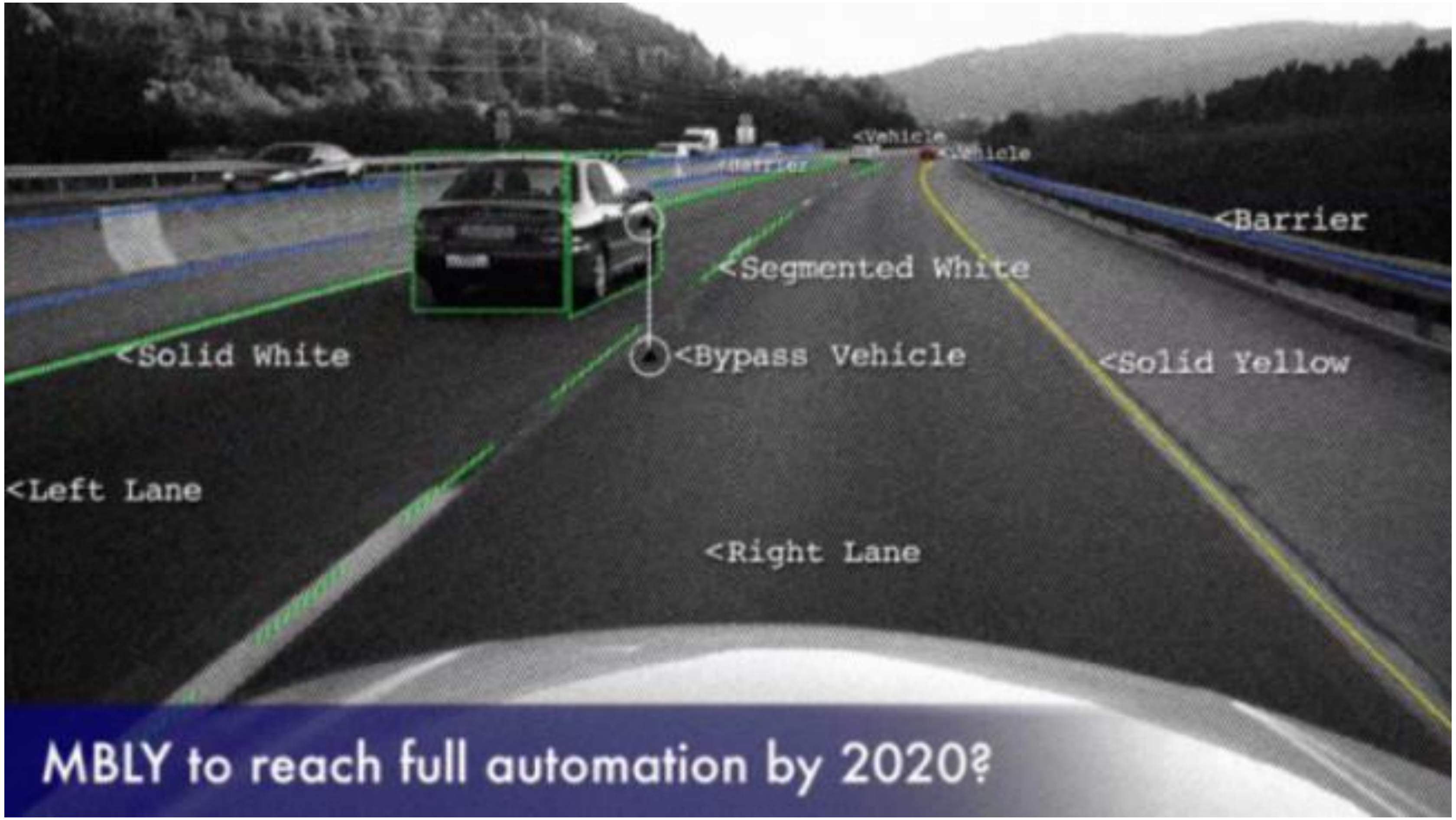
Chess champion Gary Kasparov contemplating a board in 1997, training for his May rematch with a smarter version of Deep Blue

IBM Watson in Jeopardy! (2011)



Fast R-CNN Demo

Mobileye



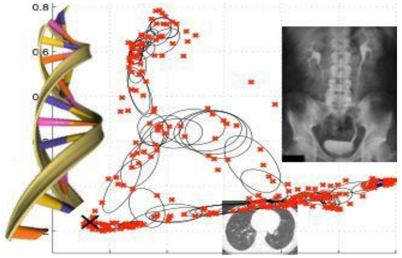
<Barrier
<Segmented White
<Solid White
<Bypass Vehicle
<Solid Yellow
<Left Lane
<Right Lane
MBLY to reach full automation by 2020?

AlphaGo vs. LEe, Sedol (2016)

Machine Learning

amazonfamily
Exclusive Deals

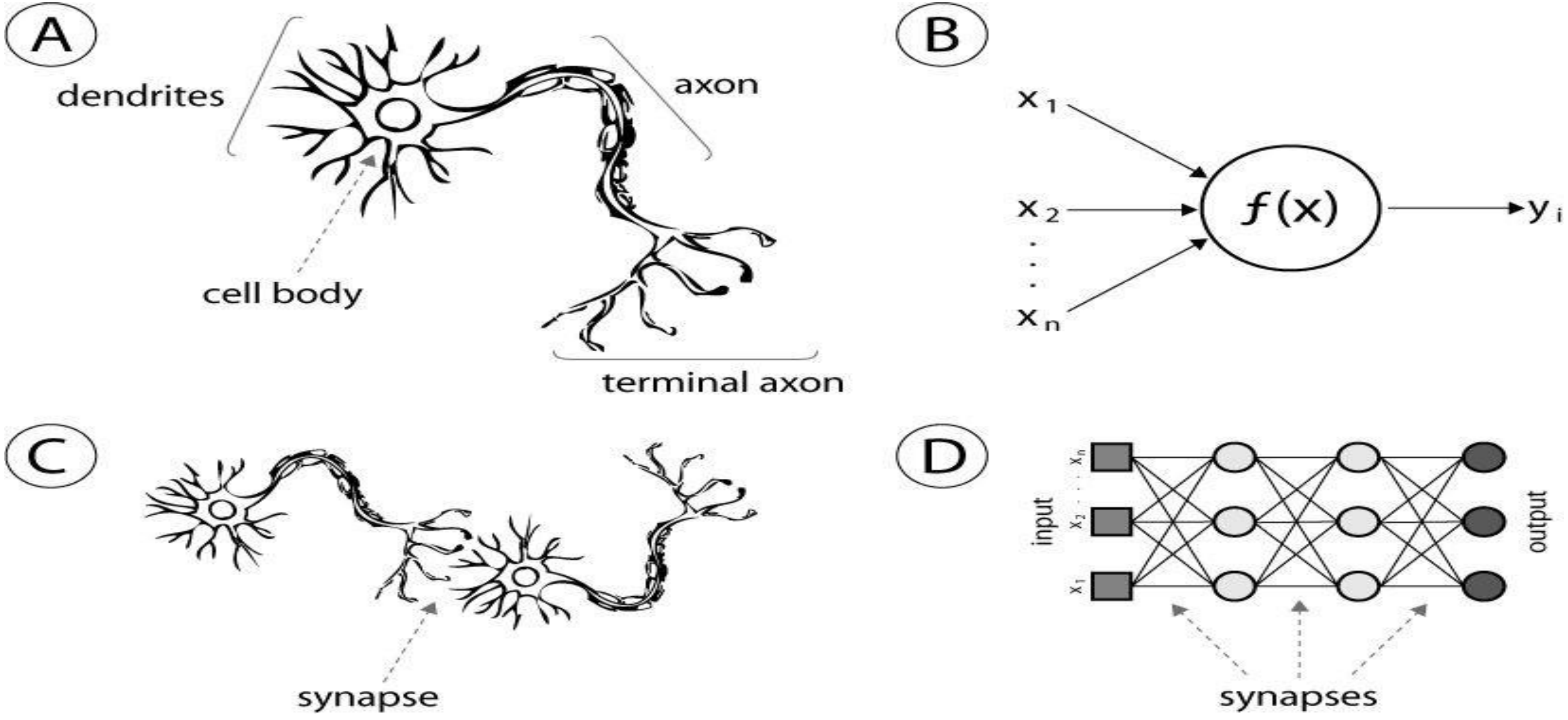
FiveThirtyEight



Neural Network

- *Collection of connected units (neurons), inspired by the biological neural networks that constitute human brains*

Neural Networks?

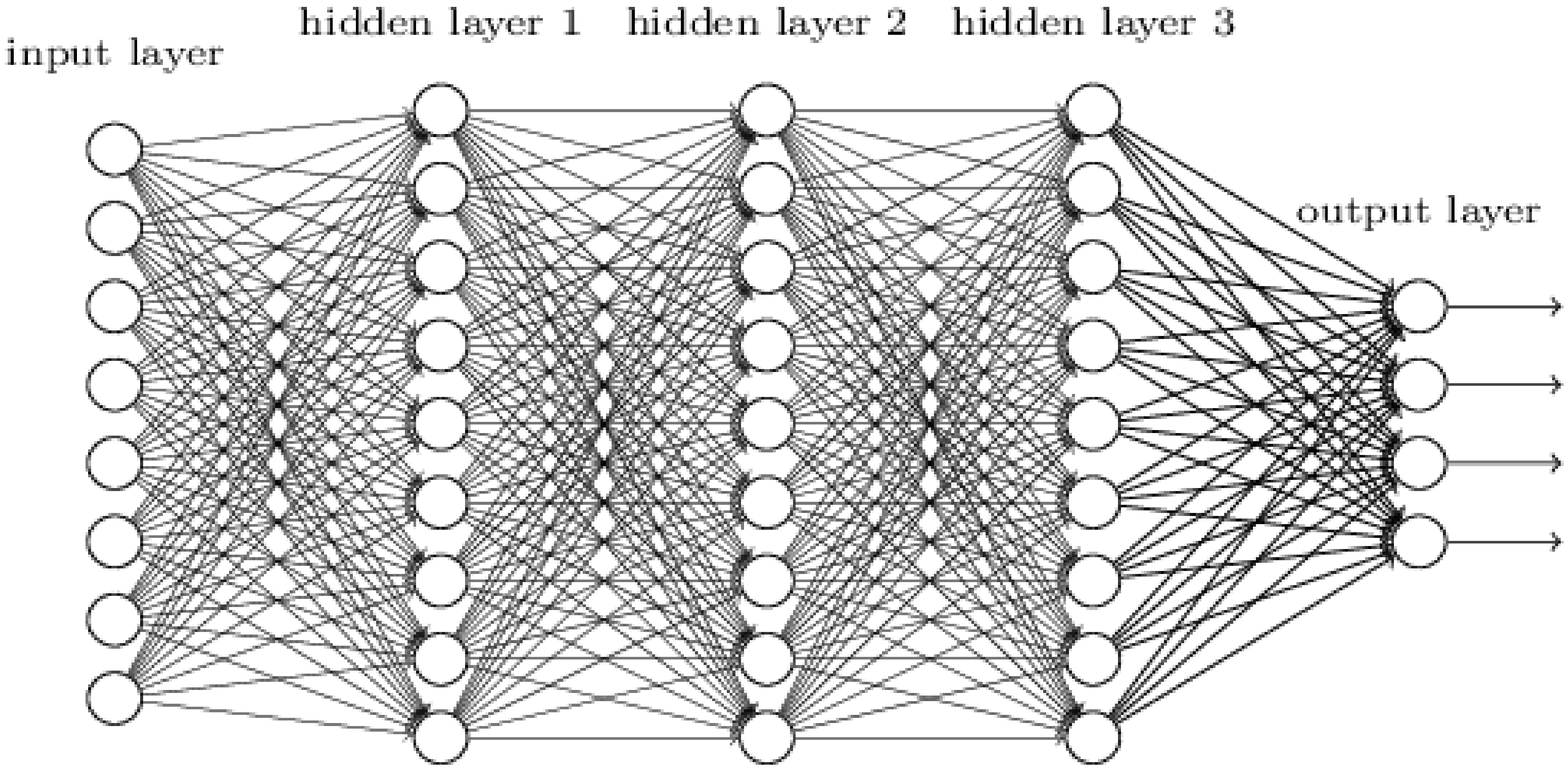


Softmax Regression (Multi-Class)

http://vision.stanford.edu/teaching/cs231n-demos/linear-classify

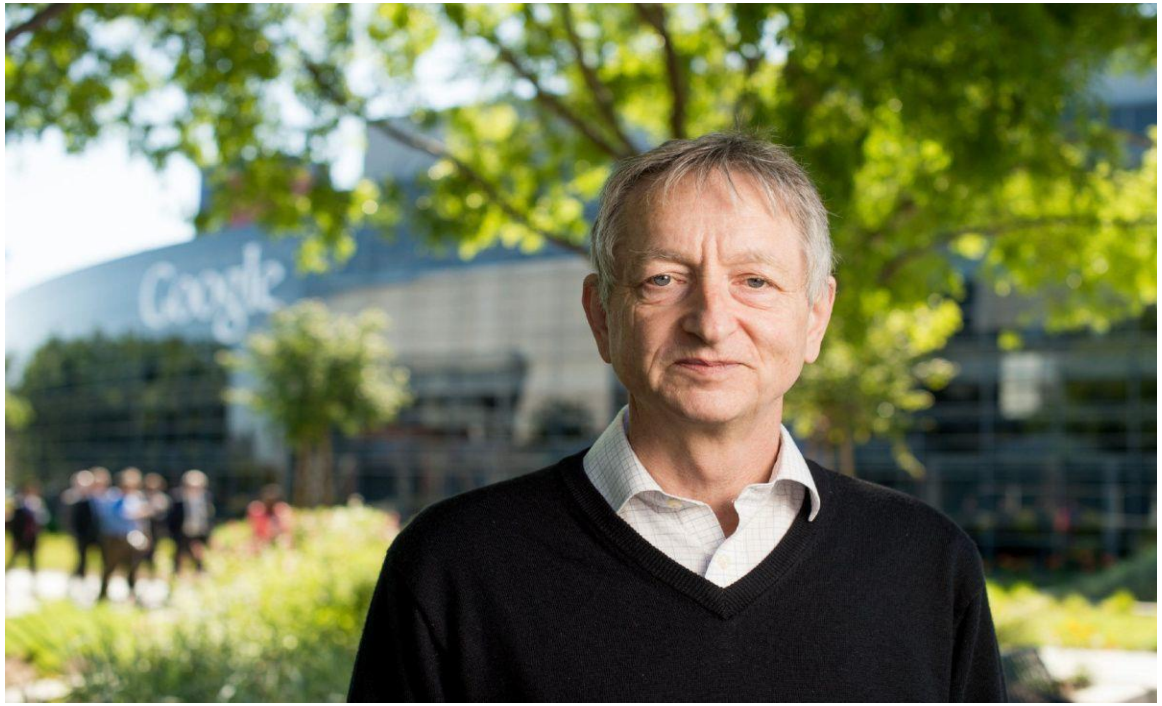
By Andrej Karpahty

Backpropagation for training



Geoffrey Hinton

- Pre-training for deep neural networks



LETTER — Communicated by Yann Le Cun

A Fast Learning Algorithm for Deep Belief Nets

Geoffrey E. Hinton
hinton@cs.toronto.edu
Simon Osindero
osindero@cs.toronto.edu
Department of Computer Science, University of Toronto, Toronto, Canada M5S 3G4

Yee-Whye Teh
tehyw@comp.nus.edu.sg
Department of Computer Science, National University of Singapore, Singapore 117543

We show how to use "complementary priors" to eliminate the explaining-away effects that make inference difficult in densely connected belief nets that have many hidden layers. Using complementary priors, we derive a fast, greedy algorithm that can learn deep, directed belief networks one layer at a time, provided the top two layers form an undirected associative memory. The fast, greedy algorithm is used to initialize a slower learning procedure that fine-tunes the weights using a contrastive version of the wake-sleep algorithm. After fine-tuning, a network with three hidden layers forms a very good generative model of the joint distribution of handwritten digit images and their labels. This generative model gives better digit classification than the best discriminative learning algorithms. The low-dimensional manifolds on which the digits lie are modeled by long ravines in the free-energy landscape of the top-level associative memory, and it is easy to explore these ravines by using the directed connections to display what the associative memory has in mind.

backpropagation,
boltzmann machines

Geoff Hinton
Google

convolution

Yann Lecun
Facebook

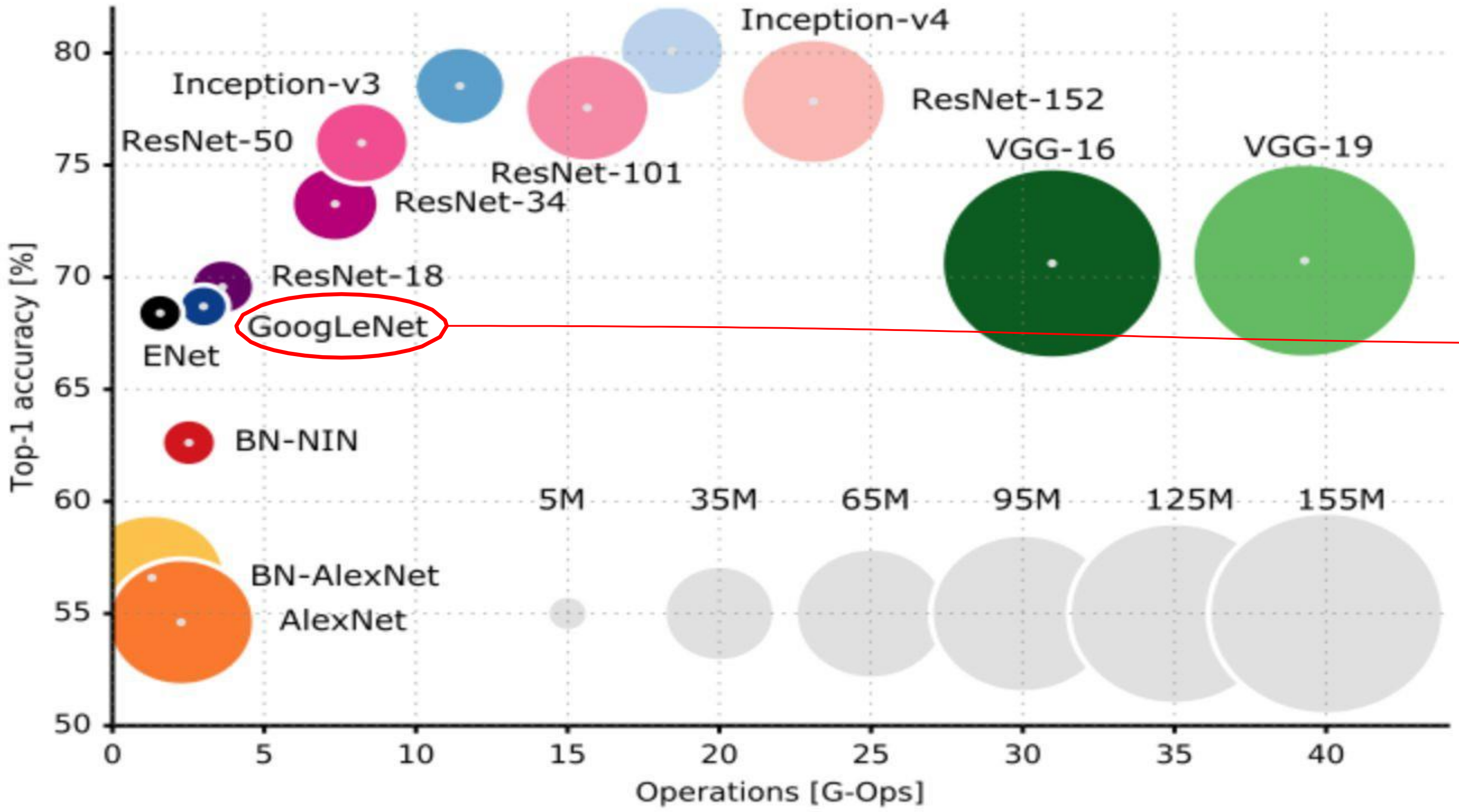
stacked auto-
encoders

Yoshua Bengio
U. of Montreal

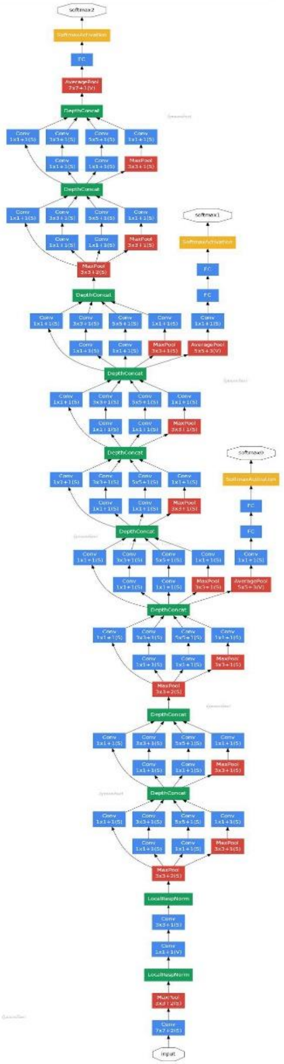
GPU utilization

Andrew Ng
Baidu

How Deep?



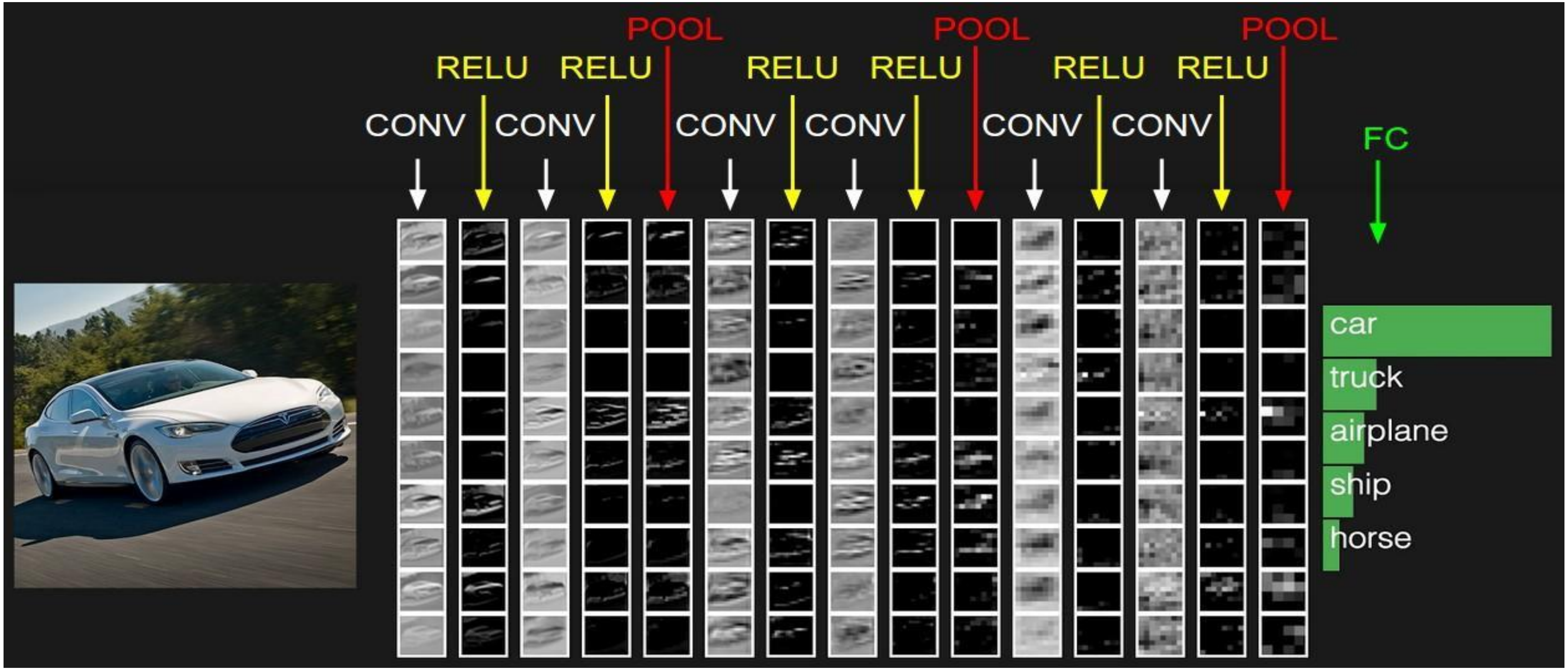
Top-1 accuracy [%]
Operations [G-Ops]
80
75
70
65
60
55
50
0
5
10
15
20
25
30
35
40
Inception-v4
Inception-v3
ResNet-152
ResNet-50
ResNet-101
ResNet-34
VGG-16
VGG-19
ResNet-18
GoogLeNet
ENet
BN-NIN
5M
35M
65M
95M
125M
155M
BN-AlexNet
AlexNet



ML, NN & DL Summary

- Machine Learning
  - Field of study that gives computers the ability to learn without being explicitly programmed
- Neural network
  - One of ML classifiers, being trained by backpropagation
  - Extension of logistic & softmax regression
- Deep Learning
  - Framework of using neural networks with deep (many) layers
  - Most of state-of-the-art ML-based products utilized DL!

Convolutional Neural Network (CNN)



Convolutional Neural Network (CNN)

- Yann LeCun's LeNet5

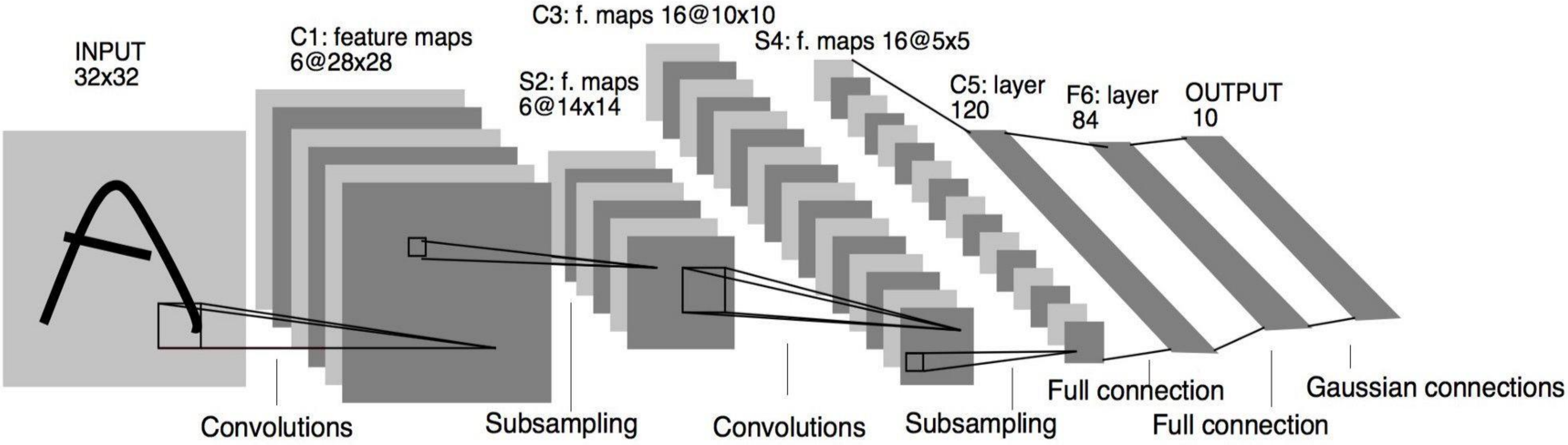


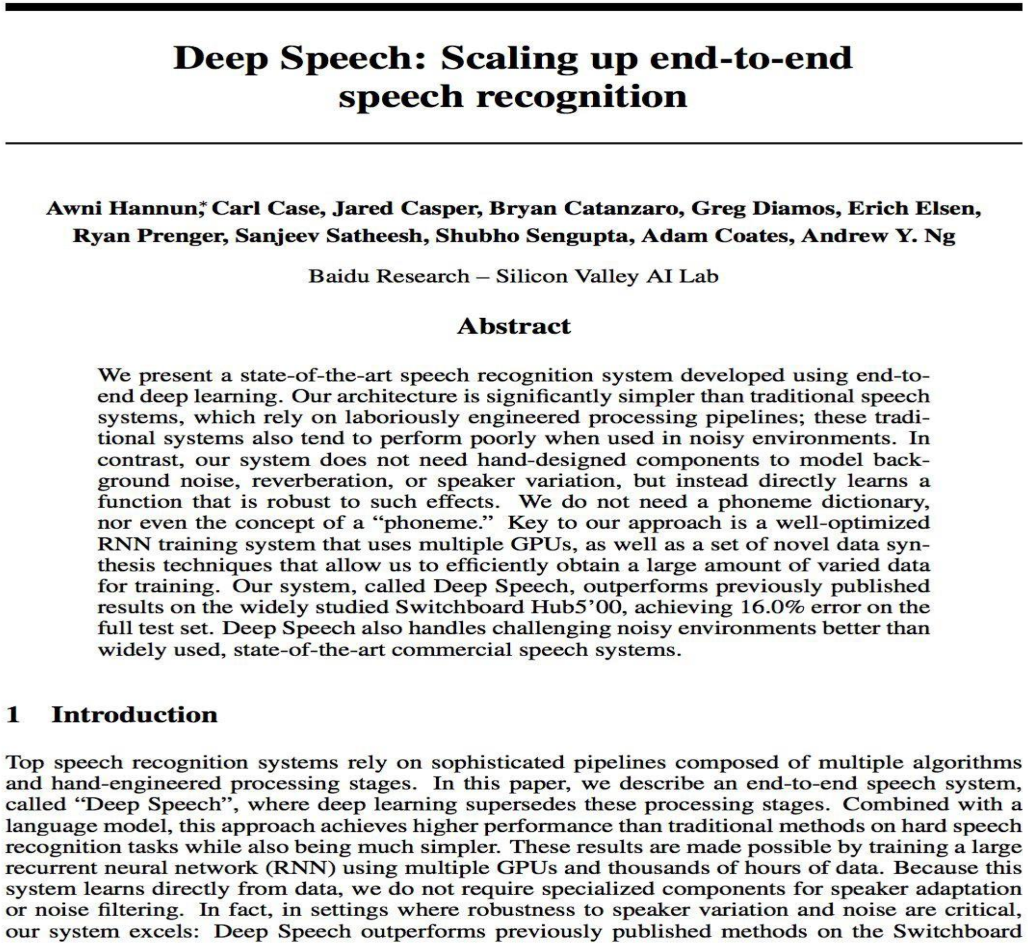
Image Classification

	Top 5 error
Imagenet 2011 winner (not CNN)	25.7%
Imagenet 2012 winner	16.4% (Krizhesvky et al.)
Imagenet 2013 winner	11.7% (Zeiler/Clarifai)
Imagenet 2014 winner	6.7% (GoogLeNet)
Human: Andrej Karpathy	5.1%
Baidu Arxiv paper: 3 Jan '15	6.0%
MS Research Arxiv paper: 6 Feb '15	4.9%
Google Arxiv paper: 2 Mar '15	4.8%

CNN in TensorFlow

https://www.tensorflow.org/tutorials/layers

CNN in Speech Recognition



Deep Speech: Scaling up end-to-end speech recognition

Awni Hannun*, Carl Case, Jared Casper, Bryan Catanzaro, Greg Diamos, Erich Elsen, Ryan Prenger, Sanjeev Satheesh, Shubho Sengupta, Adam Coates, Andrew Y. Ng

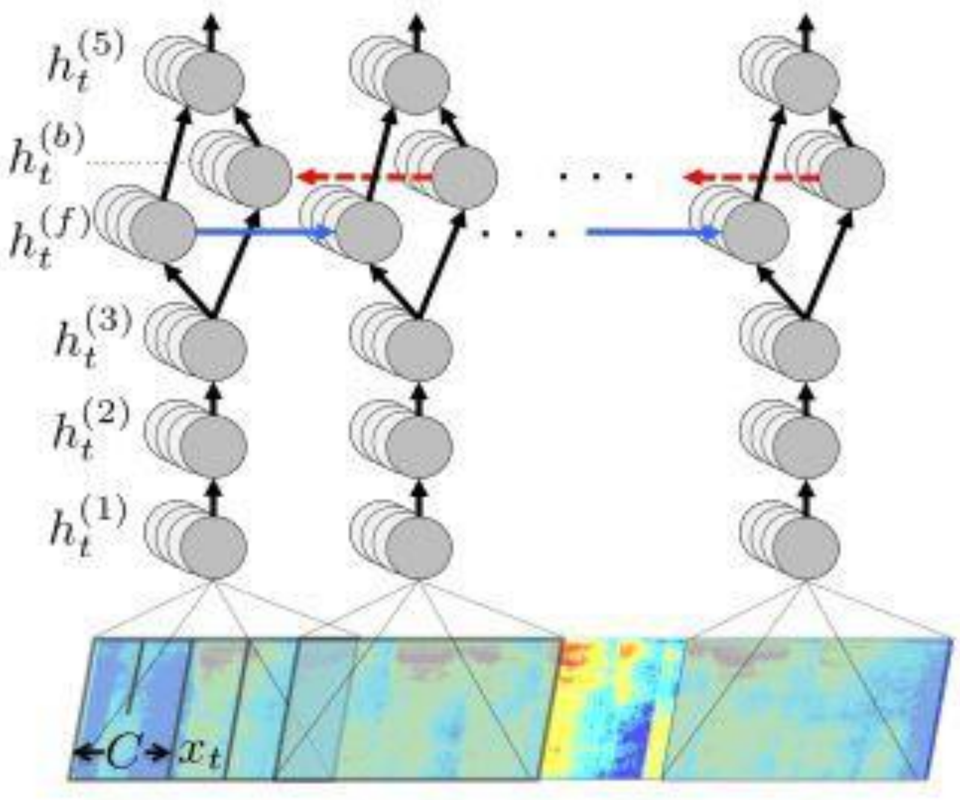
Baidu Research – Silicon Valley AI Lab

Abstract

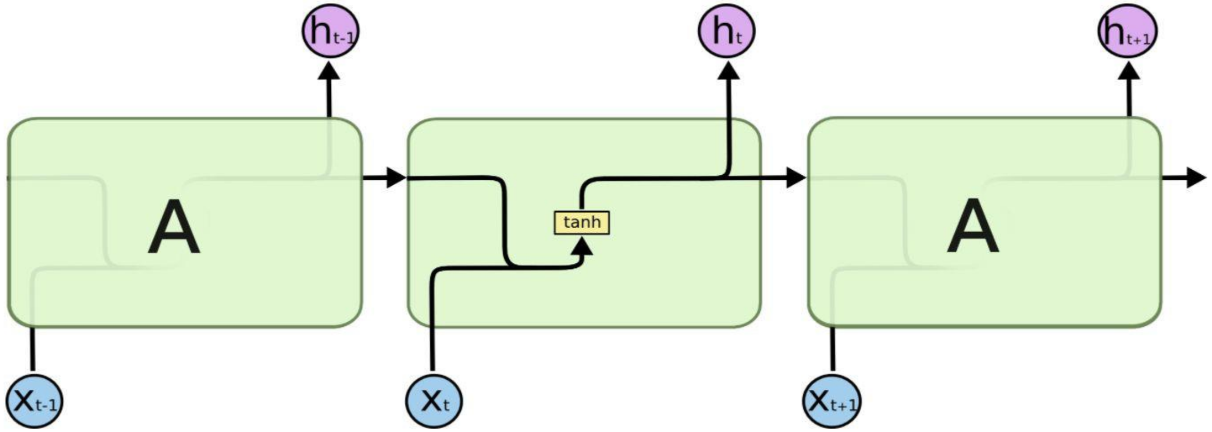
We present a state-of-the-art speech recognition system developed using end-to-end deep learning. Our architecture is significantly simpler than traditional speech systems, which rely on laboriously engineered processing pipelines; these traditional systems also tend to perform poorly when used in noisy environments. In contrast, our system does not need hand-designed components to model background noise, reverberation, or speaker variation, but instead directly learns a function that is robust to such effects. We do not need a phoneme dictionary, nor even the concept of a "phoneme." Key to our approach is a well-optimized RNN training system that uses multiple GPUs, as well as a set of novel data synthesis techniques that allow us to efficiently obtain a large amount of varied data for training. Our system, called Deep Speech, outperforms previously published results on the widely studied Switchboard Hub5'00, achieving 16.0% error on the full test set. Deep Speech also handles challenging noisy environments better than widely used, state-of-the-art commercial speech systems.

1 Introduction

Top speech recognition systems rely on sophisticated pipelines composed of multiple algorithms and hand-engineered processing stages. In this paper, we describe an end-to-end speech system, called "Deep Speech", where deep learning supersedes these processing stages. Combined with a language model, this approach achieves higher performance than traditional methods on hard speech recognition tasks while also being much simpler. These results are made possible by training a large recurrent neural network (RNN) using multiple GPUs and thousands of hours of data. Because this system learns directly from data, we do not require specialized components for speaker adaptation or noise filtering. In fact, in settings where robustness to speaker variation and noise are critical, our system excels: Deep Speech outperforms previously published methods on the Switchboard

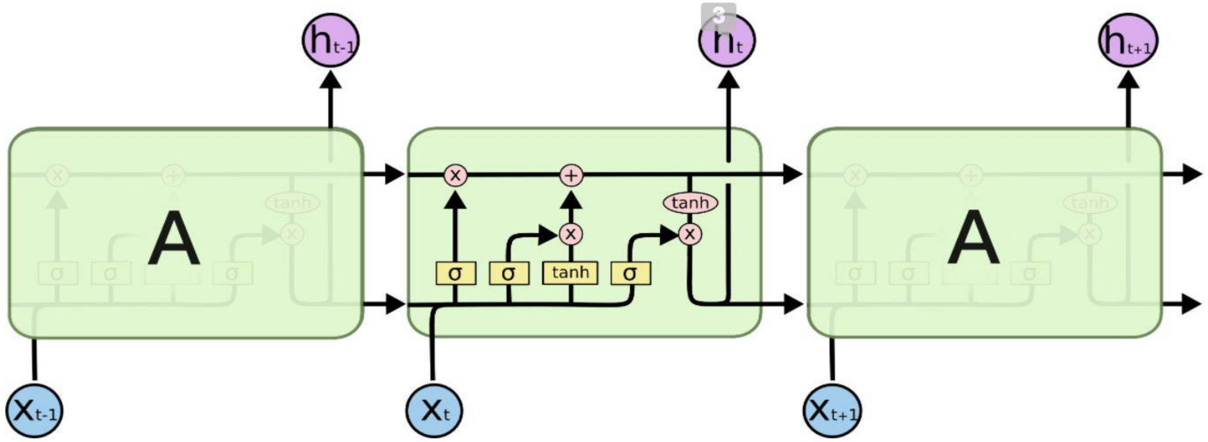


RNN and LSTM

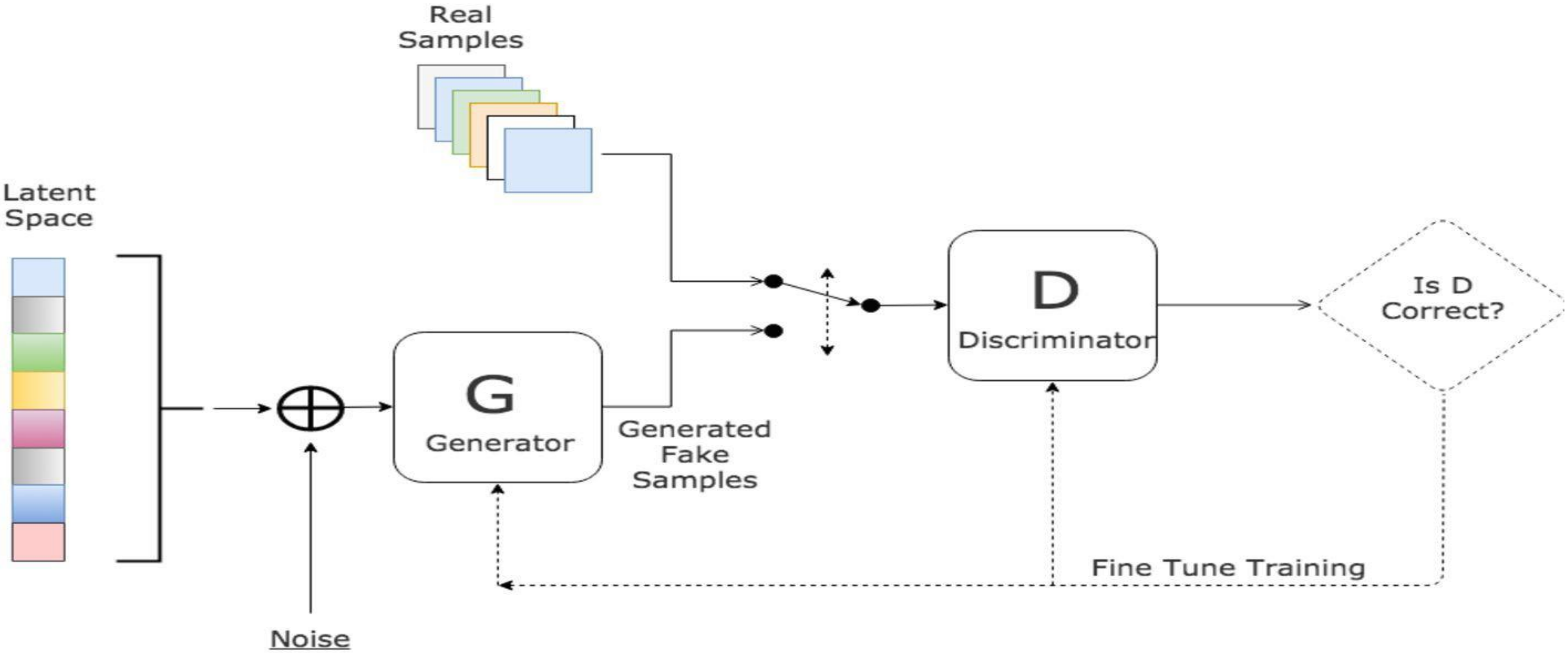


- Dependency across Time

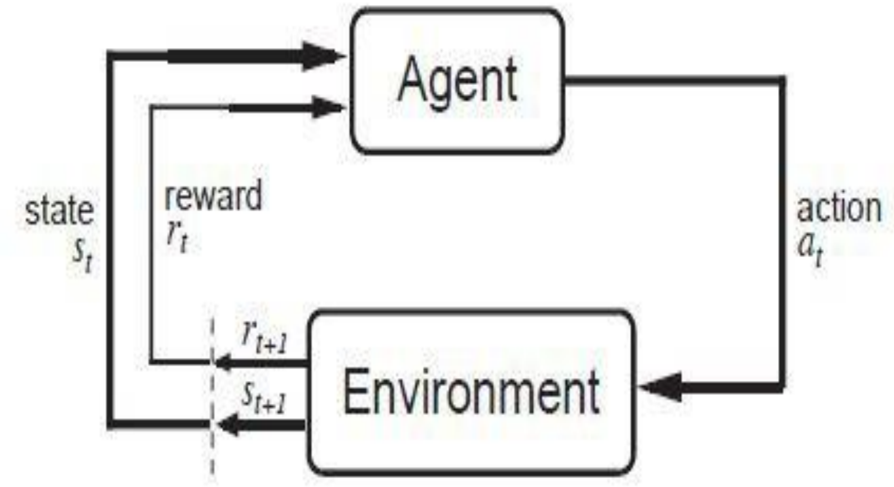
- Popular for Sequence Labeling (e.g., Speech Recognition)
- Recurrent Neural Network's Variant



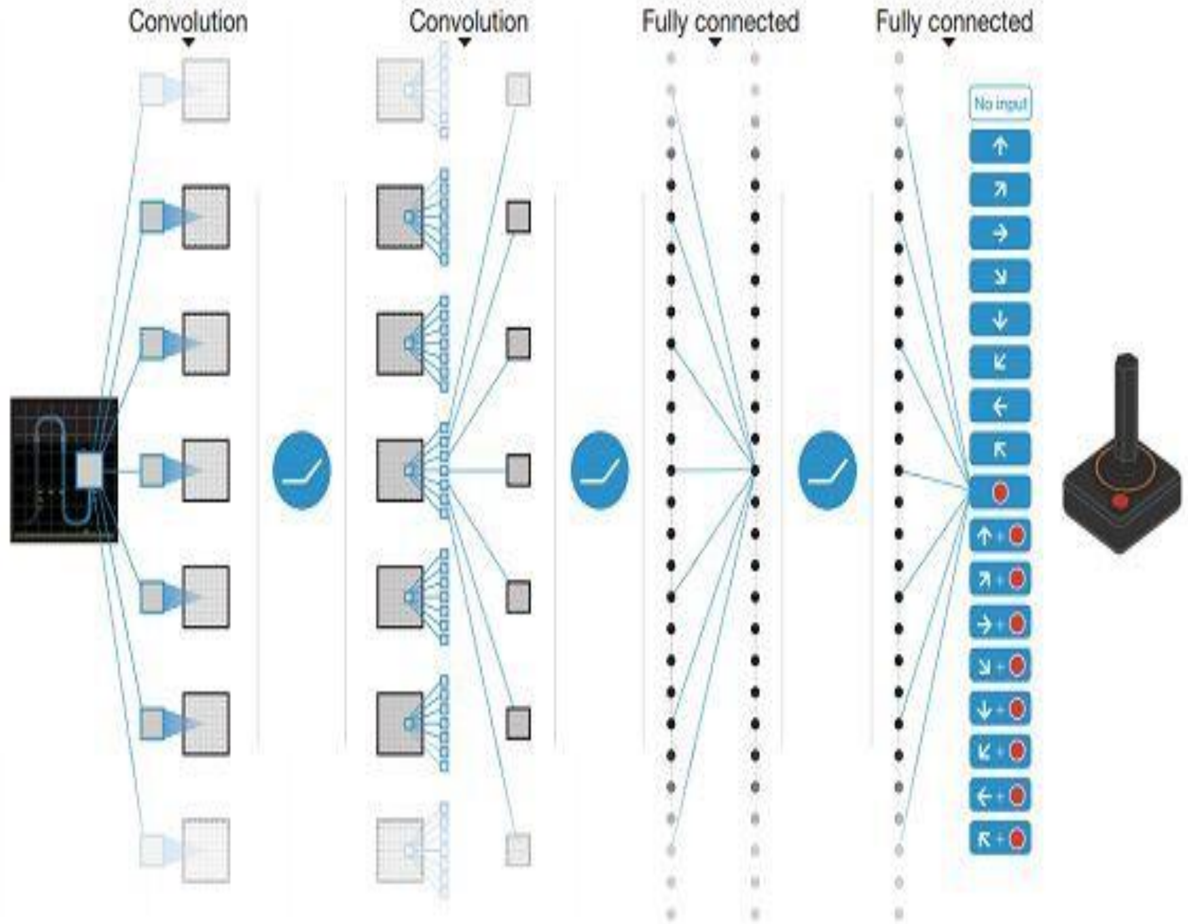
Generative Adversarial Network (GAN)



Deep Reinforcement LEarning



AlphaGo



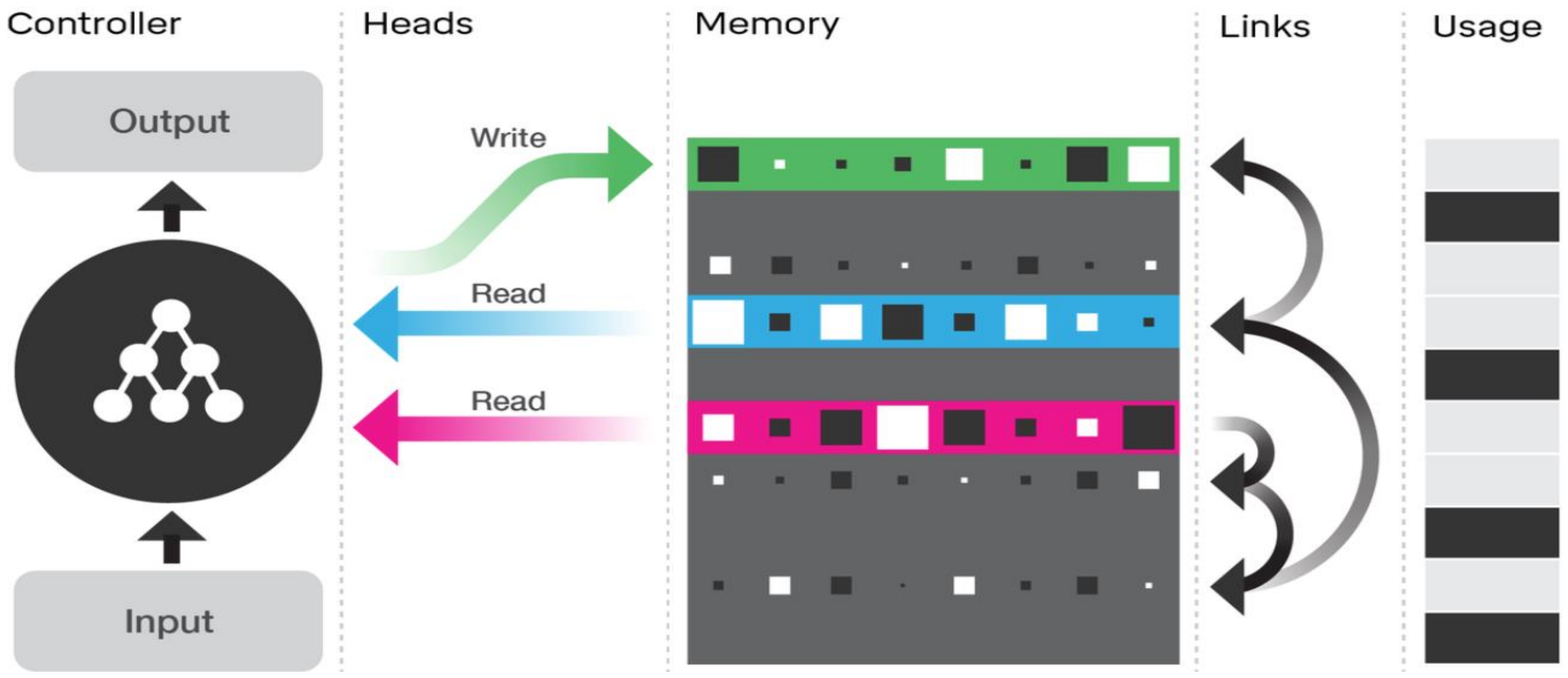

0000
0000
3


000 5 1

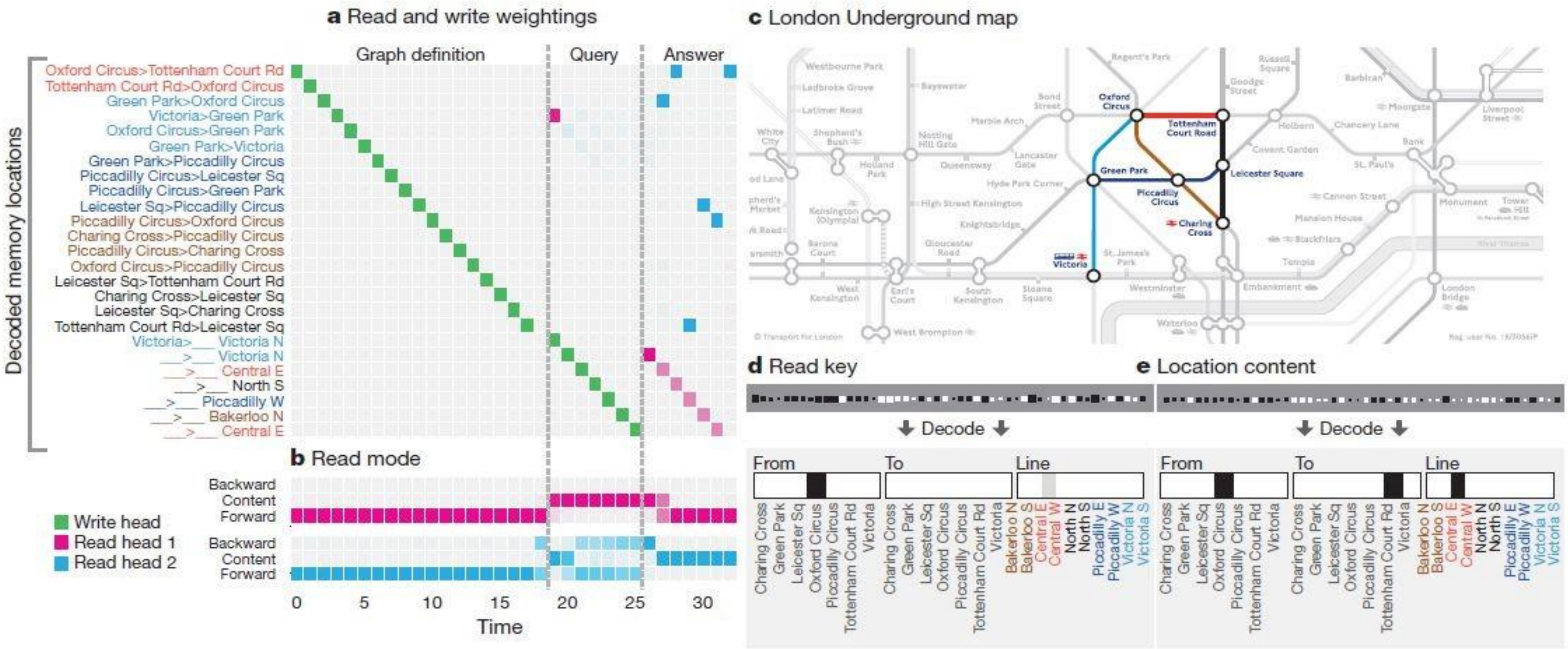
DNN Frameworks

Language	Library
C++	Caffe
Julia	MOCHA
Lua	Torch
Python	Caffe
Python	Theano
Python/C++	Tensorflow
Python	Karas

Differentiable Neural Computer (DNC) 2017



Answer Questions



Autonomous Robots

Any Questions?