



High performance. Delivered.

Deep Learning: Software Development Using Deep Neural Networks

TU Dresden

20. November 2017

Uwe Pleban, Ph.D.

Last Update: 20.Nov.2017

Strategy | Consulting | Digital | Technology | Operations

Hungry...for challenges?

Let's have powerful start in the morning and chat about your future!
International consultant(s) talk about a day in life.



Interested? Just send an E-Mail to
Andreas.Range@accenture.com



Subject: "Analytics Breakfast –
Dresden"



High performance. Delivered.

myPPTSelfie

Uwe Pleban, Ph.D.



Principal Director



Accenture GmbH
Campus Kronberg 1
D-61476 Kronberg / Ts.
Germany

uwe.pleban@accenture.com



@uwepleban



Certified Digital SA



BD100: Hackathon
Coach 2016



Milestone Faculty

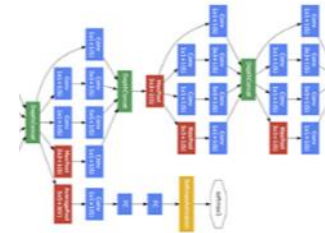


Certified Accenture
Technology SA



Königstein

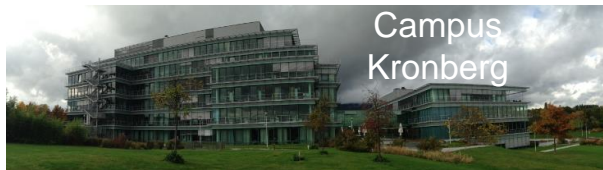
My Interests



$$e^{i\pi} + 1 = 0$$

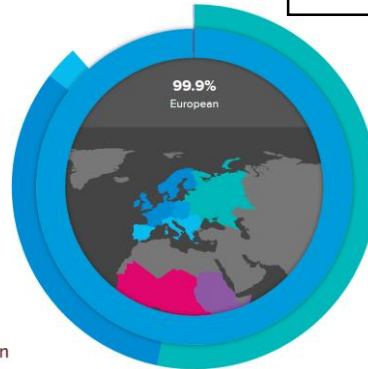


- Principal Director in Digital Delivery – Information and Thought Leadership.
- Solution Designer with special focus on Big Data, Advanced Analytics and AI solutions.
- I also track Exponential Technologies and their societal implications.

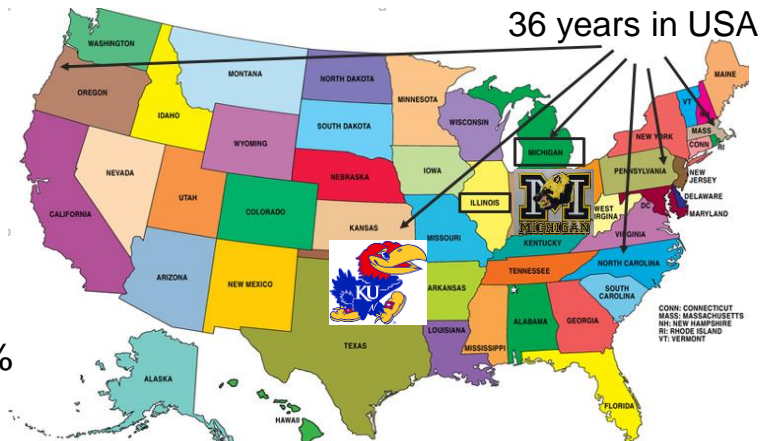


Campus
Kronberg

Word cloud containing terms like: Ann Arbor, Thomson Reuters, column-oriented, RDMS, Teradata, project management, Königstein, Hadoop, Semantics, University of Michigan, Principal Director, Ph.D. Uwe Pleban, University of Kansas, solution design, Deep Learning, Big Data, data analytics, SQL-92, Analytics, campus Kronberg, Academia, Compiler Generation, etc.



Fun Fact: I have 2.9%
Neanderthal DNA
according to 23andMe



36 years in USA

Strengths
Learner | Strategic | Achiever | Futuristic | Analytical



Prolog

Abstract for Talk:

English Original

Deep Learning: Software Development Using Deep Neural Networks

The Science Fiction author Isaac Asimov coined the phrase “Any sufficiently advanced technology is indistinguishable from magic”. While I do not believe in magic, I must admit that the technology behind Deep Learning is quite magical. In this talk we will explore Deep Neural Networks (DNNs), which are the foundation of Deep Learning, a subfield of Machine Learning and thus Artificial Intelligence. Companies like Google have publicly stated that, in essence, they are betting their future on the power of Deep Learning.

DNNs have been applied to image recognition, speech recognition and synthesis, natural language translation, the playing of complex games like Go, and various other complicated tasks, with astounding results. The networks use simple algorithms and encode their knowledge in huge numbers of units or artificial neurons. Each neuron consists of parameters together with a simple activation function. The connections between neurons at various layers of the network drive the flow of information from the input layer via hidden layers to the output layer. In a sense, DNNs use a form of data driven programming rather than the cumbersome and impractical method of using explicitly programmed rules and their combinations. This approach has enabled, for example, the enormous progress with autonomous (self-driving) vehicles, which 10 years ago were in the distant future. At the consumer level, the recently released Google Pixel Buds employ Deep Learning for the real-time translation between 40 languages – the first generation of the “Babel Fish” from the book “The Hitchhiker’s Guide to the Galaxy”.

DNNs For Handwriting and Translation EN→D

Deep Learning Software Development Using Deep Neural Networks

Der Science-Fiction-Autor Isaac Asimov prägte den Satz "Jede hinreichend **fortschrittliche** Technologie ist von **Magie** nicht zu unterscheiden". Während ich nicht an Magie glaube, muss ich zugeben, dass die Technologie, die hinter Deep Learning steht, ziemlich magisch ist. In diesem Vortrag werden wir **Deep** Neuronale Netze (DNNs) untersuchen, die die Grundlage von Deep Learning bilden, einem Teilgebiet des maschinellen Lernens und damit der künstlichen Intelligenz. Unternehmen wie Google haben öffentlich erklärt, dass sie ihre Zukunft im Wesentlichen auf die **Kraft** von Deep Learning setzen.

DNNs wurden auf Bilderkennung, Spracherkennung und Synthese, natürliche Sprachübersetzung, das Spielen von komplexen Spielen wie Go und verschiedene andere komplexe Aufgaben mit erstaunlichen Ergebnissen angewendet. Die Netzwerke verwenden einfache Algorithmen und kodieren ihr Wissen in einer großen Anzahl von Einheiten oder künstlichen Neuronen. Jedes Neuron besteht aus Parametern zusammen mit einer einfachen Aktivierungsfunktion. Die Verbindungen zwischen Neuronen an verschiedenen Ebenen des Netzes treiben den Informationsfluss von der Eingangsschicht über versteckte Schichten zur Ausgabeschicht an. In gewissem Sinne verwenden DNNs eine Form der datengesteuerten Programmierung und nicht die umständliche und unpraktische Methode, explizit programmierte Regeln und deren Kombinationen zu verwenden. Dieser Ansatz hat zum Beispiel den enormen Fortschritt mit autonomen (selbstfahrenden) Fahrzeugen ermöglicht, die vor 10 Jahren **in der fernen Zukunft waren**. Auf der Consumer-Ebene **beschäftigen** die kürzlich veröffentlichten Google Pixel Buds Deep Learning für die Echtzeit-Übersetzung zwischen 40 Sprachen - die erste Generation des "Babel Fish" aus dem Buch "The Hitchhiker's Guide to the Galaxy".

“Fixed Up” Abstract in German

Der Science-Fiction-Autor Isaac Asimov prägte den Satz: "Jede hinreichend fortschrittliche Technologie ist von Magie nicht zu unterscheiden". Während ich nicht an Zauberei glaube, muss ich zugeben, dass die Technologie, die hinter Deep Learning steht, ziemlich magisch ist. In diesem Vortrag werden wir Deep Neuronale Netze (DNNs) untersuchen, welche die Grundlage von Deep Learning bilden, einem Teilgebiet des maschinellen Lernens und damit der künstlichen Intelligenz. Unternehmen wie Google haben öffentlich erklärt, dass sie ihre Zukunft im Wesentlichen auf die Fähigkeiten von Deep Learning aufbauen. DNNs sind auf Bilderkennung, Spracherkennung und Synthese, natürliche Sprachübersetzung, das Spielen von komplexen Spielen wie Go sowie auf unterschiedliche andere komplexe Aufgaben mit erstaunlichen Ergebnissen angewendet worden. Die Netzwerke benutzen einfache Algorithmen und kodieren ihr Wissen in einer großen Anzahl von Einheiten oder künstlichen Neuronen. Jedes Neuron besteht aus Parametern zusammen mit einer einfachen Aktivierungsfunktion. Die Verbindungen zwischen Neuronen auf verschiedenen Ebenen des Netzes treiben den Informationsfluss von der Eingangsschicht über versteckte Schichten zur Ausgabeschicht. In gewissem Sinne verwenden DNNs eine Form der datengesteuerten Programmierung und nicht die umständliche und unpraktische Methode, explizit programmierte Regeln und deren Kombinationen zu verwenden. Dieser Ansatz hat zum Beispiel den enormen Fortschritt bei autonomen (selbstfahrenden) Fahrzeugen ermöglicht, die vor 10 Jahren in ferner Zukunft lagen. Auf der Verbraucher-Ebene beschäftigen die kürzlich veröffentlichten Google Pixel Buds Deep Learning für die Echtzeit-Übersetzung zwischen 40 Sprachen - die erste Generation des "Babel Fish" aus dem Buch "The Hitchhiker's Guide to the Galaxy".

Google Pixel Buds – Available Now for \$159



Be multilingual with Google Translate and Pixel

Pixel Buds can even translate between languages in real time **using Google Translate on Pixel**. It's like you've got your own personal translator with you everywhere you go. Say you're in Little Italy, and you want to order your pasta like a pro. All you have to do is hold down on the right earbud and say, "Help me speak Italian." As you talk, your Pixel phone's speaker will play the translation in Italian out loud. When the waiter responds in Italian, you'll hear the translation through your Pixel Buds. If you're more of a sushi or French food fan, no need to worry—it works in 40 languages.

Fiddly, frustrating, yet sometimes incredible

Pixel Buds are now available for \$159 in the U.S. and are available to pre-order today. They're also coming to Canada, U.K., Germany, Australia and Singapore in November.

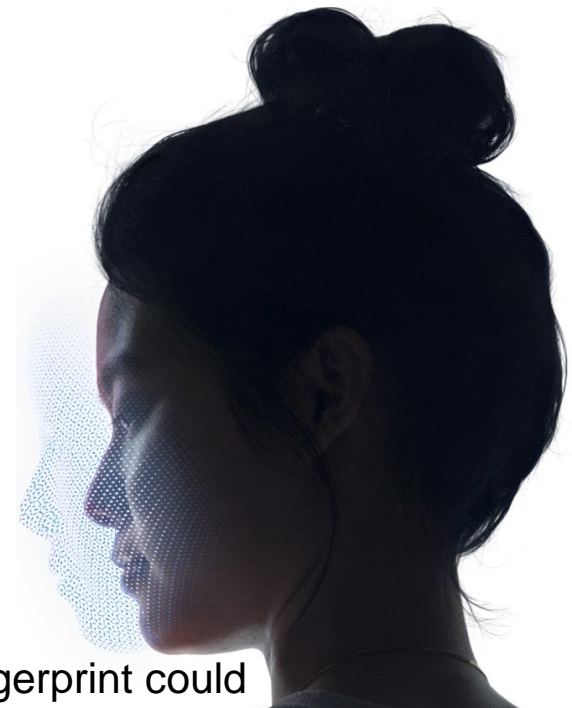
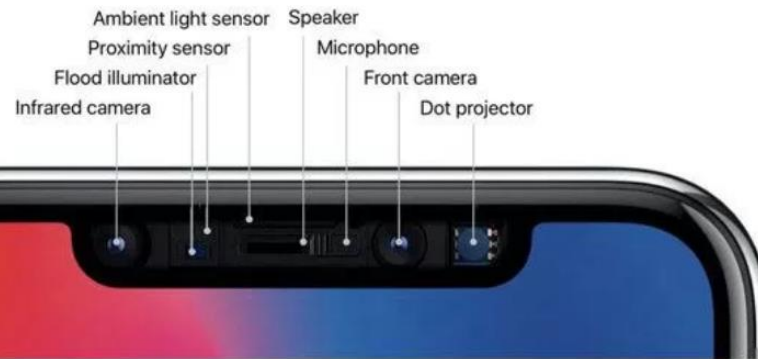
<https://www.blog.google/products/pixel/pixel-buds/>
https://store.google.com/de/product/google_pixel_buds?hl=de-DE

Nov 2017						
MO	DI	MI	DO	FR	SA	SO
44	30	31	1	2	3	4
45	6	7	8	9	10	11
46	13	14	15	16	17	18
47	20	21	22	23	24	25
48	27	28	29	30		

Face Recognition

Face ID on Apple's iPhone X

With the launch of the iPhone X, we've entered a new age of digital security: Your face is the password. Apple's smartphone isn't the first to unlock based on facial recognition—Samsung's Galaxy S8 and Note8 offer it too—but it's a step forward, the first to rely on three-dimensional measurements of a person's face.

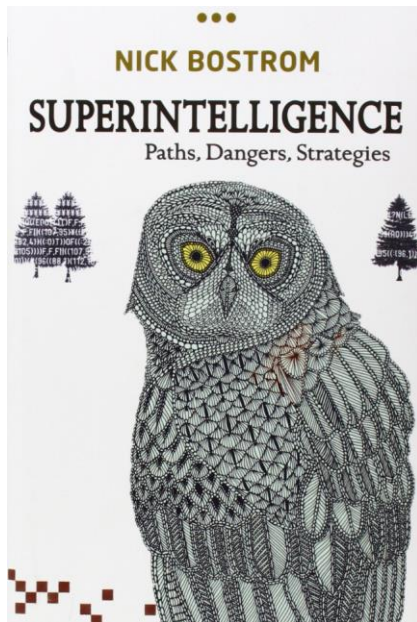


Your Face is Your Passport

- 1 in 50,000: Probability that a random fingerprint could unlock an iPhone with Touch ID
- 1 in 1,000,000: Probability that a random person could unlock an iPhone using Face ID

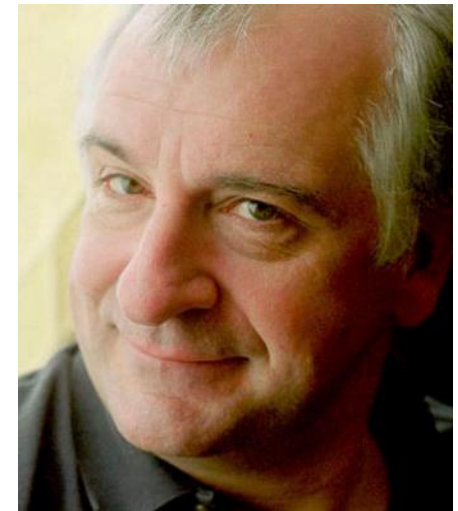
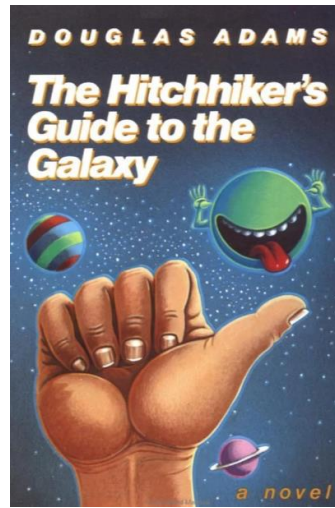
<https://qz.com/email/quartz-obsession/1118751/>

"A lot of cutting edge AI has filtered into general applications, often without being called AI because once something becomes useful enough and common enough it's not labelled AI anymore."



Nick Bostrom, Director of the Future of Humanity Institute at Oxford University

Douglas Adams:




I've come up with a set of rules that describe our reactions to technologies:



1. Anything that is in the world when you're born is normal and ordinary and is just a natural part of the way the world works.
2. Anything that's invented between when you're fifteen and thirty-five is new and exciting and revolutionary and you can probably get a career in it.
3. Anything invented after you're thirty-five is against the natural order of things.

Google Trends – “Deep Learning” Worldwide (1/2)



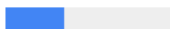
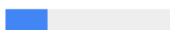
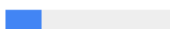
Interest over time 



Interest by region 

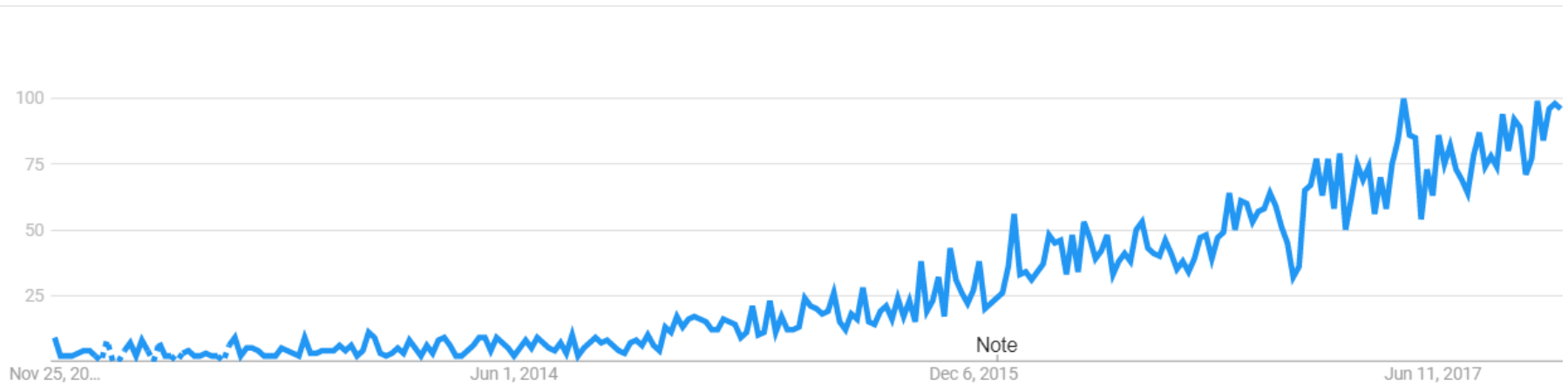
Region  



1	China	100	
2	South Korea	65	
3	Singapore	35	
4	Hong Kong	25	
5	Israel	22	

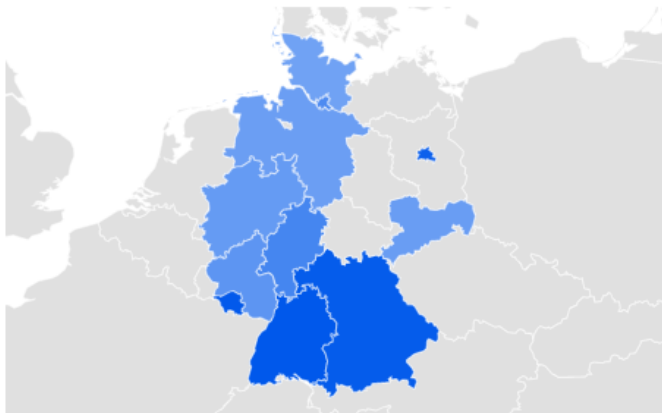
Google Trends – “Deep Learning” Germany (2/2)

Interest over time [?](#)



Interest by subregion [?](#)

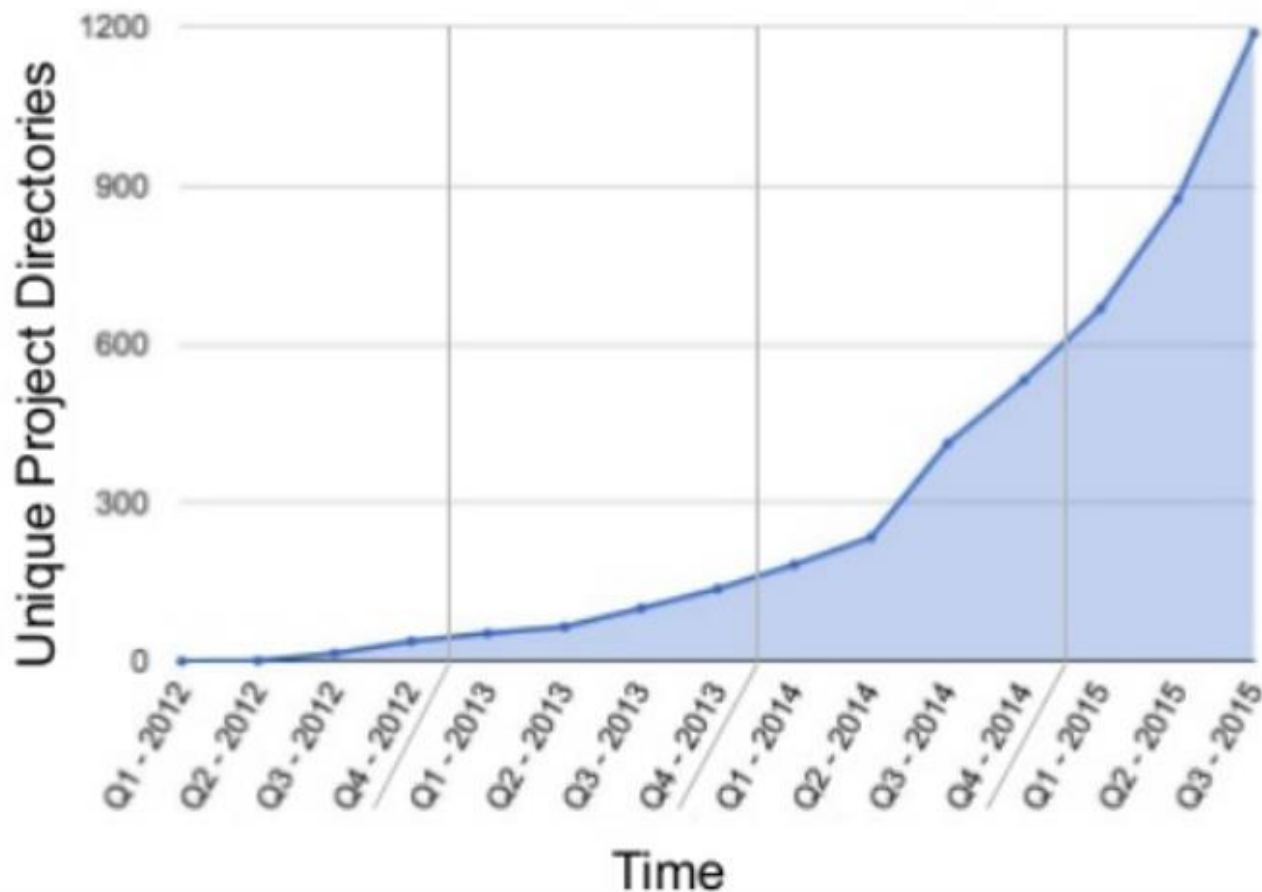
Subregion [▼](#) [↗](#)



1	Baden-Württemberg	100	<div style="width: 100%;"></div>
2	Saarland	99	<div style="width: 99%;"></div>
3	Bavaria	97	<div style="width: 97%;"></div>
4	Berlin	93	<div style="width: 93%;"></div>
5	Hesse	61	<div style="width: 61%;"></div>

Growing Use of Deep Learning at Google

of directories containing model description files



Across many products/areas:

- Android
- Apps
- drug discovery
- Gmail
- Image understanding
- Maps
- Natural language understanding
- Photos
- Robotics research
- Speech
- Translation
- YouTube
- ... many others ...



<https://www.slideshare.net/embeddedvision/largescale-deep-learning-for-building-intelligent-computer-systems-a-keynote-presentation-from-google>

MIT Technology Review: 10 Breakthrough Technologies 2013

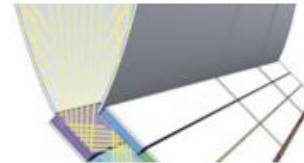
Smart Watches

The designers of the Pebble watch realized that a mobile phone is more useful if you don't have to take it out of your pocket.



Ultra-Efficient Solar Power

Doubling the efficiency of solar devices would completely change the economics of renewable energy. Here is a design that just might make it possible.



Memory Implants

A maverick neuroscientist believes he has deciphered the code by which the brain forms long-term memories.



Prenatal DNA Sequencing

Reading the DNA of fetuses is the next frontier of the genome revolution. Do you really want to know the genetic destiny of your unborn child?



Deep Learning

With massive amounts of computational power, machines can now recognize objects and translate speech in real time. Artificial intelligence is finally getting smart.



MIT Technology Review: 10 Breakthrough Technologies 2017

Technologies Directly Related to Deep Learning

Self-Driving Trucks

Tractor-trailers without a human at the wheel will soon barrel onto highways near you. What will this mean for the nation's 1.7 million truck drivers?



Paying with Your Face

Face-detecting systems in China now authorize payments, provide access to facilities, and track down criminals. Will other countries follow?



Reinforcement Learning

By experimenting, computers are figuring out how to do things that no programmer could teach them.





Introduction



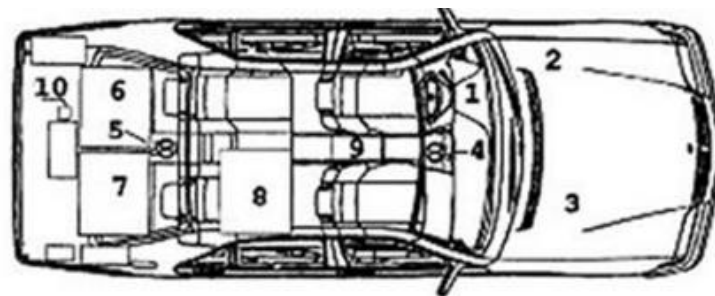
What are
the colors of
this dress?

Autonomous Vehicles

The first truly autonomous car drove **1678 km on various Autobahn segments between Munich to Odense in Denmark and back**. Both longitudinal and lateral guidance were performed autonomously by vision. On highways, the car achieved speeds exceeding 175 km/h. The mean autonomously driven distance without resets was ~9 km; the **longest autonomously driven stretch reached 158 km**. In total, 95% autonomous driving (by distance) was achieved.

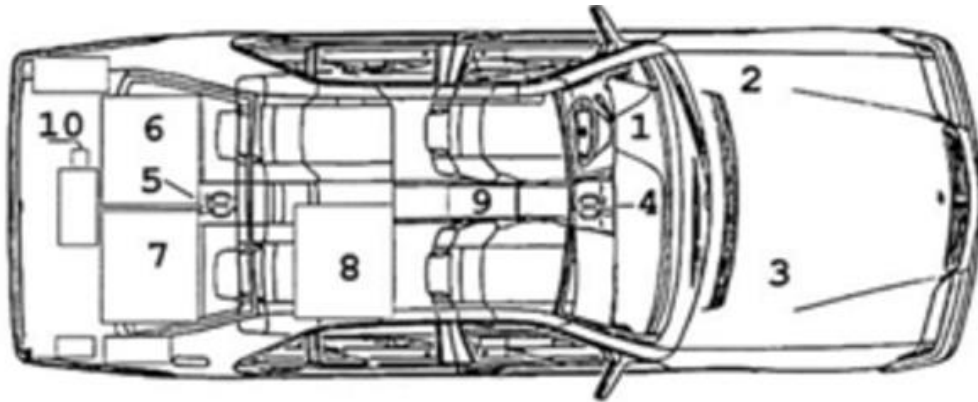
When did the event described above take place?

- A. In the late 1980s
- B. In the fall of 1995
- C. In early September of 2001
- D. During the time of the 2006 FIFA World Cup in Germany

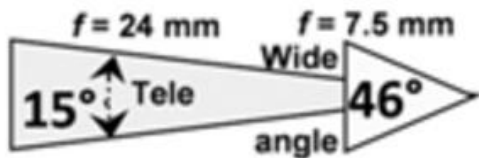


- | | |
|------------------------------|--|
| 1 Torque motor for steering | 6, 8 Transputer system, image processing |
| 2 brake system | 7 processors for gaze & locomotion control |
| 3 electric throttle control | 8 user interface |
| 4 front platform } for 2 CCD | 9 linear accelerometers |
| 5 rear platform } cameras | 10 angular rate sensors |

Autonomous Vehicles



- | | |
|---|--------------------------------------|
| 1 electrical steering motor | 6 Transputer Image Processing system |
| 2 electrical brake control | 7 platform and vehicle controllers |
| 3 electronic throttle | 8 electronics rack, human interface |
| 4 front pointing platform for CCD-cameras | 9 accelerometers (3orthogonal) |
| 5 rear pointing platform | 10 inertial rate sensors |



At distance $L_s \sim 20$ m (~ 60 m),
the resolution is 5 cm/pixel



The driverless car **VaMP** (Versuchsfahrzeug für autonome Mobilität und Rechnersehen) which was developed in the context of the European research project **PROMETHEUS**: (top left) components for autonomous driving; (right) VaMP and view into passenger cabin (lower right); (lower left) bifocal camera arrangement (front) on yaw platform

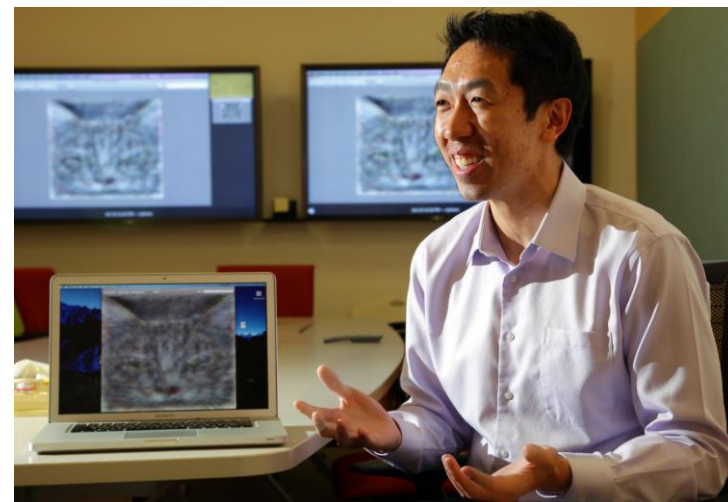
Deep Neural Networks – 2012

In 2012, Google developed a distributed computing infrastructure for training large-scale neural networks.



Which of the following are true?

- A. The computation was spread across 16,000 CPU cores (in Google data centers)
- B. The trained models had more than 1 billion connections
- C. One of the artificial neurons learned to respond strongly to pictures of cats.
- D. All of the above



25 Years of Moore's Law



1990 Compaq Laptop
Cost: \$5,000



2015 MacBook Pro
Cost: \$3,000

To solve the ImageNet challenge, complex neural networks need to be trained.

On a 2015 MacBook Pro, the training time for a particular network may be about 3 weeks. What would have been the equivalent training time on a 1990 Compaq laptop?

- A. 27 months
- B. 27 years
- C. 270 years
- D. 2,700 years
- E. 27,000 years
- F. 270,000 years

Paint Colors

Recently, an open source neural network was trained on a list of about 7,700 Sherwin-Williams paint colors along with their RGB (red, green, blue) color values. The network then invented new paint colors and gave them (hopefully attractive) names.

Which of the color names below were created by the AI software, and which ones are actual names of Sherwin-Williams paint colors?



Coral Gray



Emperor's Silk



Dondarf



Aubusson Blue



Homestar Brown



Jonquil



Duck Egg Blue



Light of Blast



Butter Up



Sticks Red












Fawn Brindle



Stoner Blue

Paint Colors Generated by Char-rnn

 Clardic Fug 112 113 84	 Sticks Red 171 37 34	 Wlittf Bzt 25 64 0
 Snowbonk 201 199 165	 Coral Gray 129 102 100	 Bylfgoam Glosd 229 233 112
 Catbabel 97 93 68	 Rover White 222 222 213	 Gorlpateehcd 63 62 90
 Bunflow 190 174 155	 Corcaunitiol Orange 239 212 202	 Woleebaph Ronder Wily 195 199 199
 Ronching Blue 121 114 125	 Ghasty Pink 231 137 165	 Iroeee CerMowt 222 128 187
 Bank Butt 221 196 199	 Power Gray 151 124 112	 Dondarf 145 151 226
 Caring Tan 171 166 170	 Navel Tan 199 173 140	
 Stargoon 233 191 141	 Bock Coe White 221 215 236	
 Sink 176 138 110	 Horble Gray 178 181 196	
 Stummy Beige 216 200 185	 Homestar Brown 133 104 85	
 Dorkwood 61 63 66	 Snader Brown 144 106 74	
 Flower 178 184 196	 Golder Craam 237 217 177	
 Sand Dan 201 172 143	 Hurky White 232 223 215	
 Grade Bat 48 94 83	 Burf Pink 223 173 179	
 Light Of Blast 175 150 147	 Rose Hork 230 215 198	
 Grass Bat 176 99 108		
 Sindis Poop 204 205 194		
 Dope 219 209 179		
 Testing 156 101 106		
 Stoner Blue 152 165 159		
 Burple Simp 226 181 132		
 Stanky Bean 197 162 171		
 Turdly 190 164 116		

<https://arstechnica.com/information-technology/2017/05/an-ai-invented-a-bunch-of-new-paint-colors-that-are-hilariously-wrong/>
<http://lewisandquark.tumblr.com/>

Bipolar Disorder

Existence enters your entire nation.
A twisted mind reveals becoming manic,
An endless modern ending medication,
Another rotten soul becomes dynamic.

Or under pressure on genetic tests.
Surrounded by controlling my depression,
And only human torture never rests,
Or maybe you expect an easy lesson.

Or something from the cancer heart disease,
And I consider you a friend of mine.
Without a little sign of judgement please,
Deliver me across the borderline.

An altered state of manic episodes,
A journey through the long and winding roads.

Love at First Sight

An early morning on a rainy night,
Relax and make the other people happy,
Or maybe get a little out of sight,
And wander down the streets of Cincinnati.

Death

Being buried under ashes scattered,
Many faces we forgotten own,
About a hundred thousand soldiers gathered,
And I remember standing all alone.

Wave

People picking up electric chronic.
The balance like a giant tidal wave,
Never ever feeling supersonic,
Or reaching any very shallow grave.

What do these poems have in common?

Modern Art: Do You See Anything Out of the Ordinary?

Top ranked by human subjects



<https://www.technologyreview.com/s/608195/machine-creativity-beats-some-modern-art/amp/>

High Resolution Photos of Celebrities: Can You Name Any of These?





The World You & I Live In Today (Part 1) TWYILIT?

Any sufficiently advanced technology is indistinguishable from magic.
Isaac Asimov

Pace of Change Accelerates – S&P 500 Lifespan

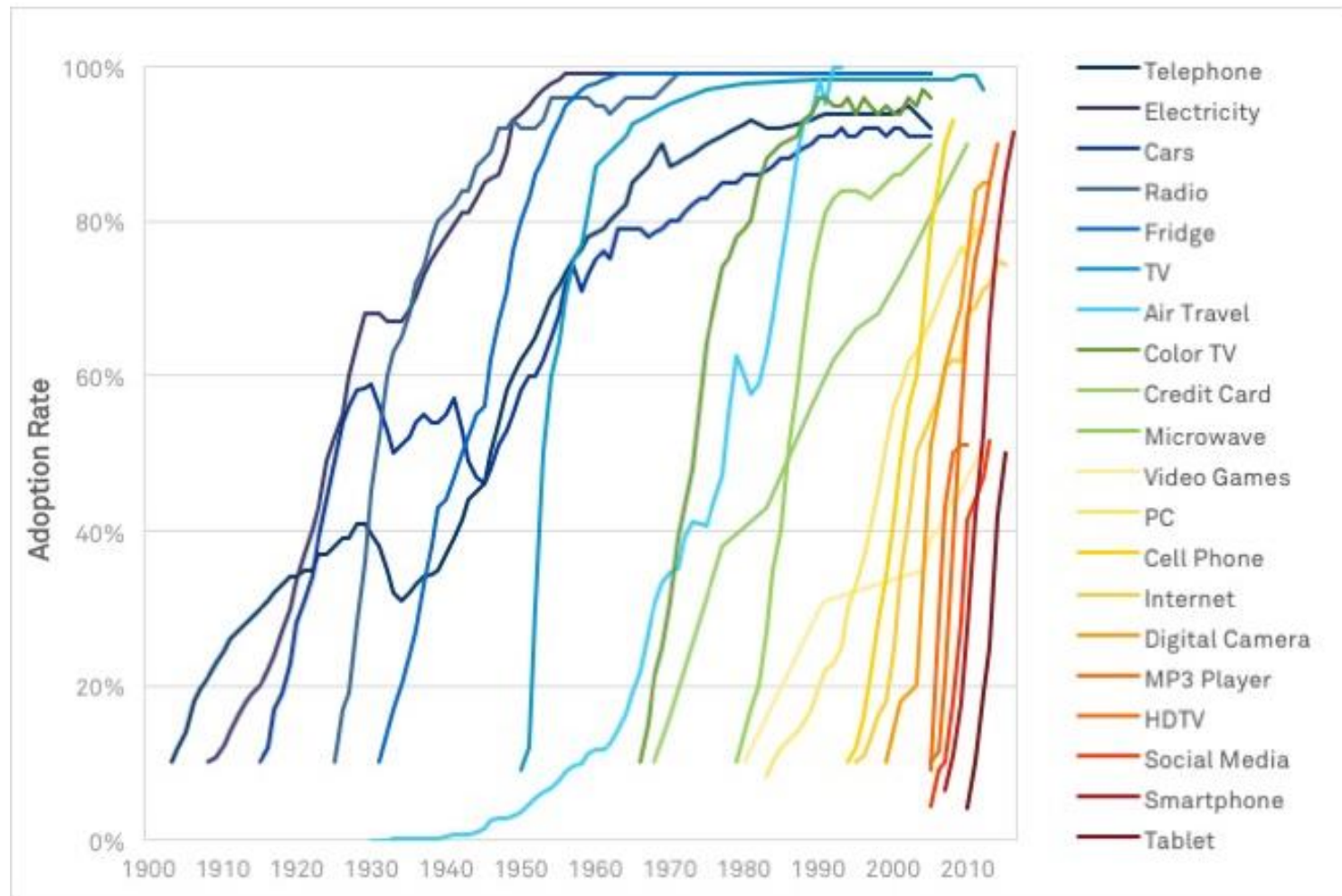


CB Insights @CBinsights - 14 Std.

1/ It's harder than ever for incumbents to stay on top. Witness decline of S&P 500 lifespan [cbinsights.com/research-corporate...](https://cbinsights.com/research-corporate)



Adoption of Technology in the U.S., 1900 to present



Source: Asymco

BLACKROCK®

Source: <http://www.valuwalk.com/2015/12/the-topic-we-should-all-be-paying-attention-to-in-3-charts/>

“Seriously, I don’t like the phrase “Big Data”. I prefer **“Data Science”, which is the automatic (or semi-automatic) extraction of knowledge from data.** That is here to stay, it’s not a fad. The amount of data generated by our digital world is growing exponentially with high rate (at the same rate our hard-drives and communication networks are increasing their capacity). But the amount of human brain power in the world is not increasing nearly as fast. This means that **now or in the near future most of the knowledge in the world will be extracted by machines and reside in machines.** It’s inevitable. An entire industry is building itself around this, and a new academic discipline is emerging.”



-Yann LeCun

Yann LeCun is currently Director of AI Lab at Facebook, and a world renowned Machine Learning researcher

Three Generations of Google AVs



History of Waymo

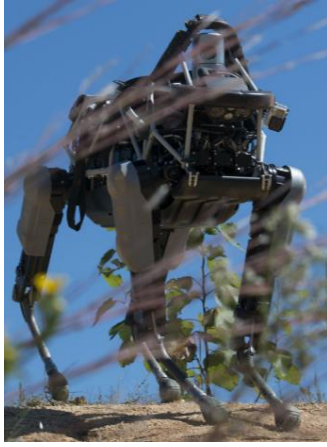
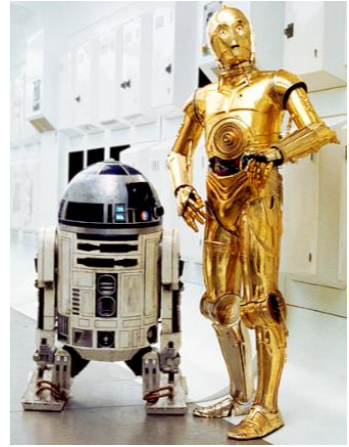
- In 2009, Google started its Self Driving Car Project
- In December 2016, the project spun out from the Google X Lab into a formal Alphabet company called Waymo
- In February 2017, Waymo unveiled an integrated software and hardware platform for AVs, including custom LiDAR and Radar sensor systems
 - Waymo is heavily investing in sensor technology; it has reported that it has cut the costs of LiDAR sensors by 90%
- Now testing autonomous cars in Mountain View (CA), Austin (TX), Kirkland (WA), and Phoenix (AZ)
- **Conducting a public “Early Rider Program” in the Phoenix metro area**
- Partnering with Fiat Chrysler (since May 2016)
 - Currently using 100 Pacifica hybrid minivans
 - Planning to add 500 more minivans to its fleet for the public early ride program in Phoenix



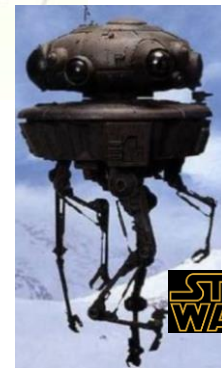
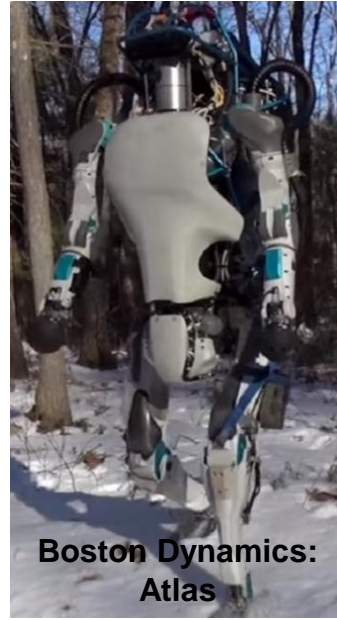
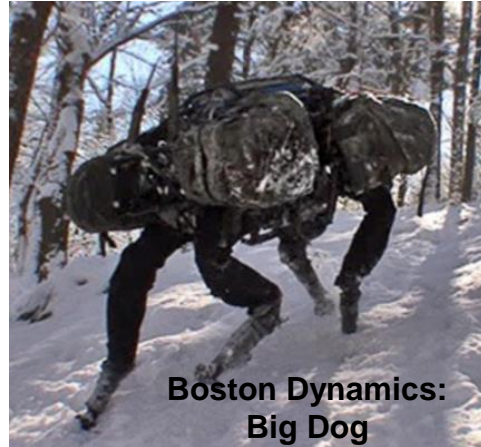
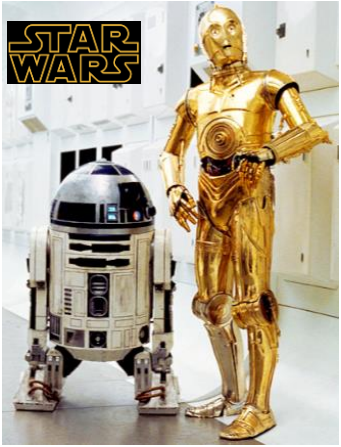
<https://waymo.com/>

<https://www.youtube.com/waymo>

Do You Know Your Star Wars Robots? And the Other Ones?



Do You Know Your Star Wars Robots? And the Other Ones?



Boston Dynamics Robots



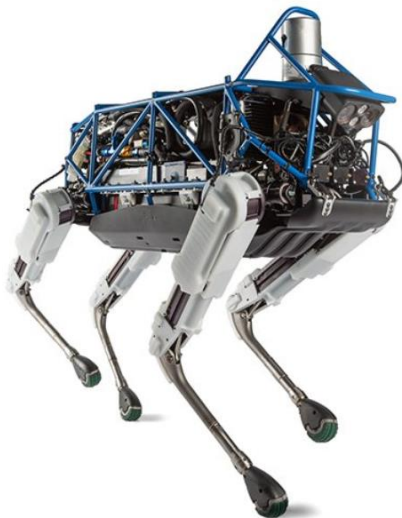
Handle



Spot Mini



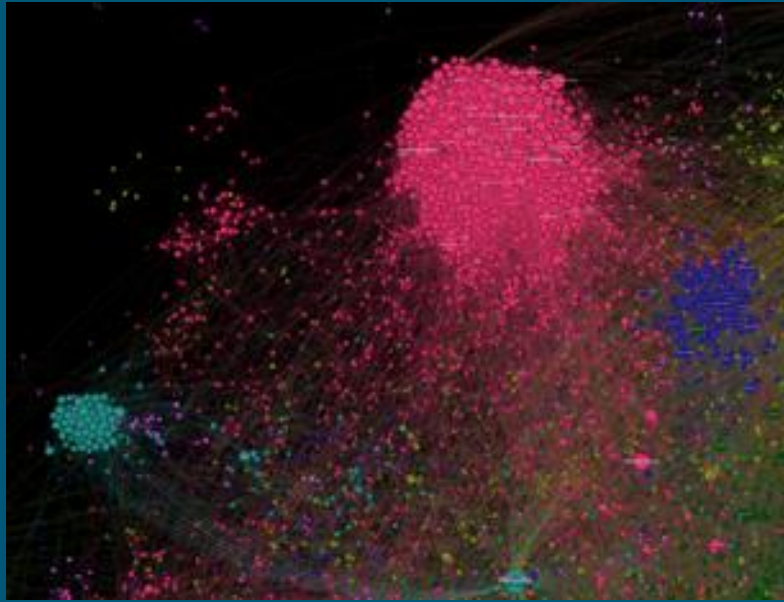
Atlas



Spot

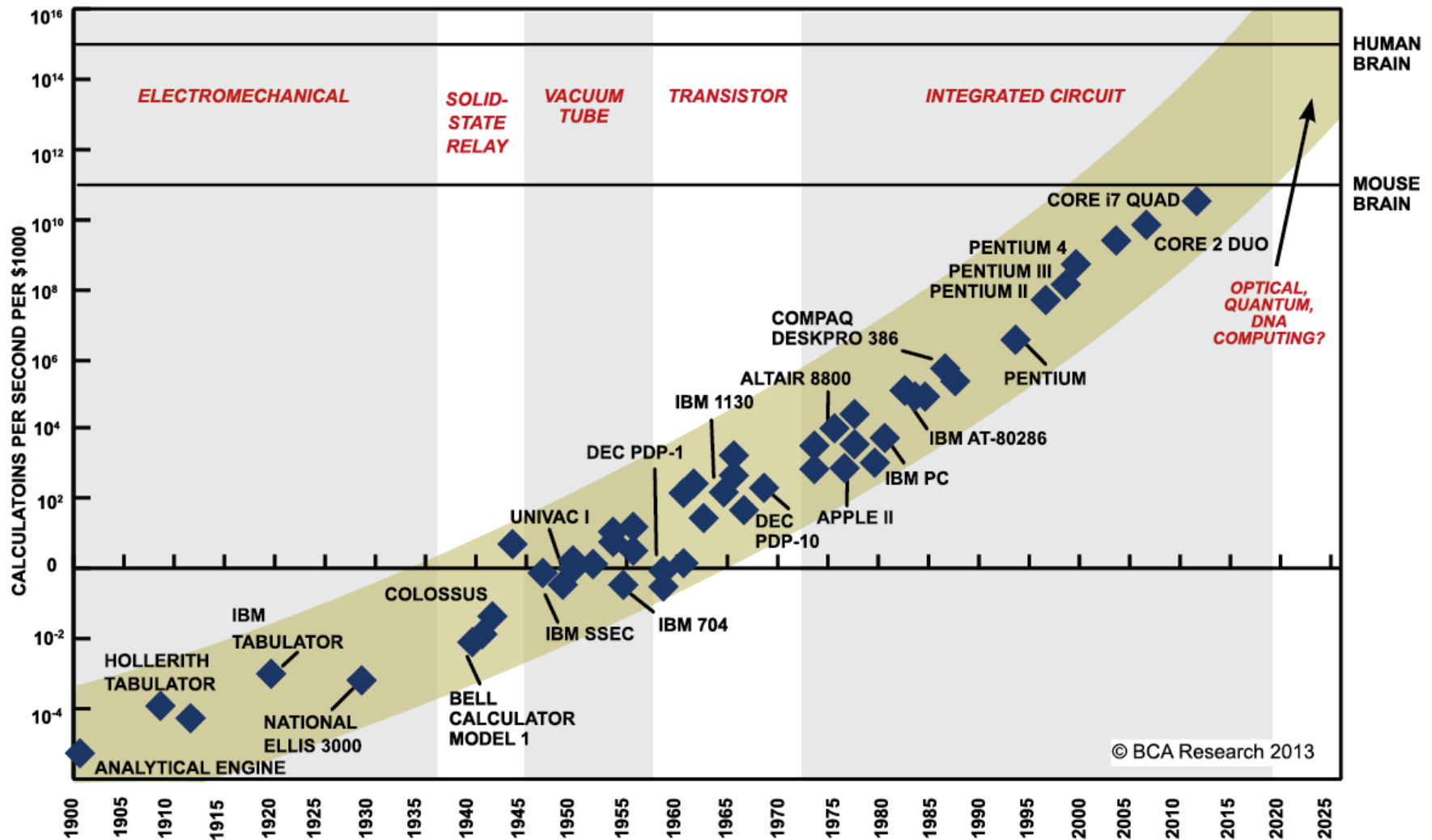


Big Dog



The Laws

Moore's Law – Valid for More Than 100 Years



SOURCE: RAY KURZWEIL, "THE SINGULARITY IS NEAR: WHEN HUMANS TRANSCEND BIOLOGY", P.67, THE VIKING PRESS, 2006. DATAPPOINTS BETWEEN 2000 AND 2012 REPRESENT BCA ESTIMATES.

<http://www.extremetech.com/wp-content/uploads/2015/04/MooresLaw2.png>

Moore's Law – Visualized



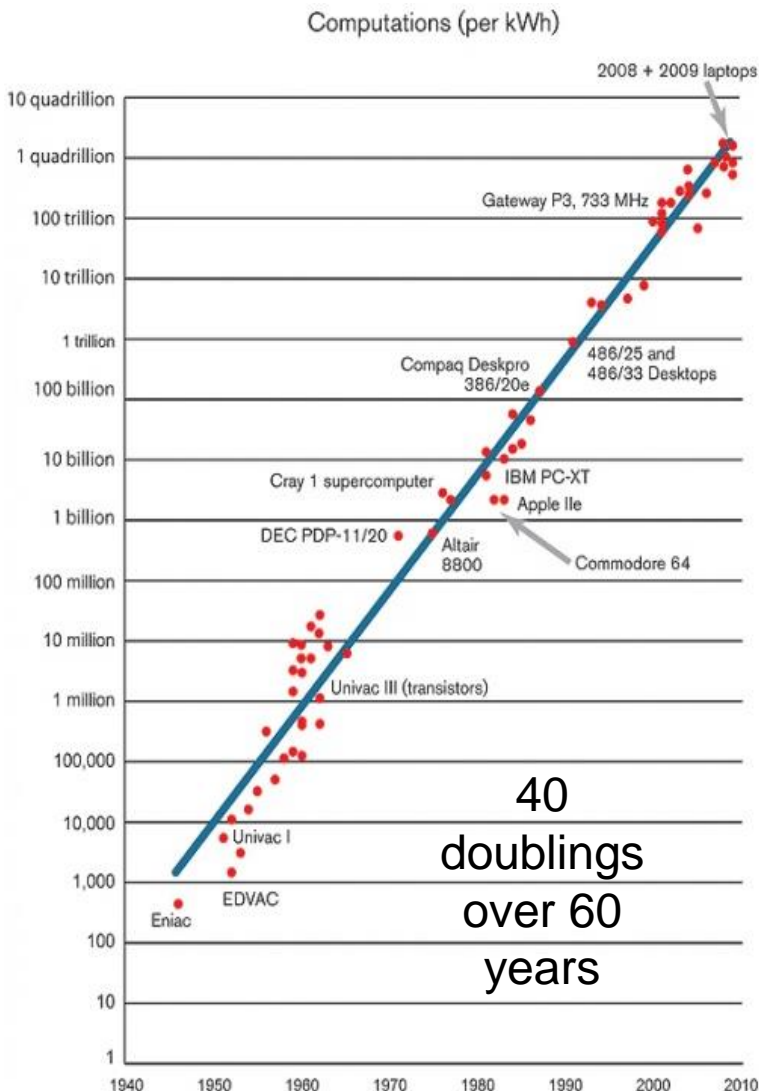
19 years of rendering Lara Croft's face in the Tomb Raider video games



Similarly, for the Wolfenstein video games – 1992 vs. 2014



Koomey's Law



Jonathan Koomey of Stanford U. found that the **electrical efficiency of computing has doubled every 1.6 years since the mid-1940s**

... if a **modern-day MacBook Air** operated at the energy efficiency of computers from 1991, its fully charged battery would last all of **2.5 seconds**.



Some researchers are already building devices that run on „ambient“ energy harvested from light, heat, vibration or TV and radio transmitters.

Moore's and Koomey's Laws in Action – Cray X-MP vs. Apple iPhone



Delip Rao
@deliprao

Follow

A perspective on progress of computing from Bob Mercer's ACL lifetime award talk.

#acl2014



Computational Power

Cray xmp	Apple iPhone
1.05MHz clock	1.3GHz clock
64 Megabytes of RAM	1 Gigabyte of RAM
Up to 32 Gigabytes of disk space	64 Gigabyte storage
Weighed 1.5 tons	Weighs 4 oz.
Consumed 250 Kilowatts	Consumes 0.125 Watts

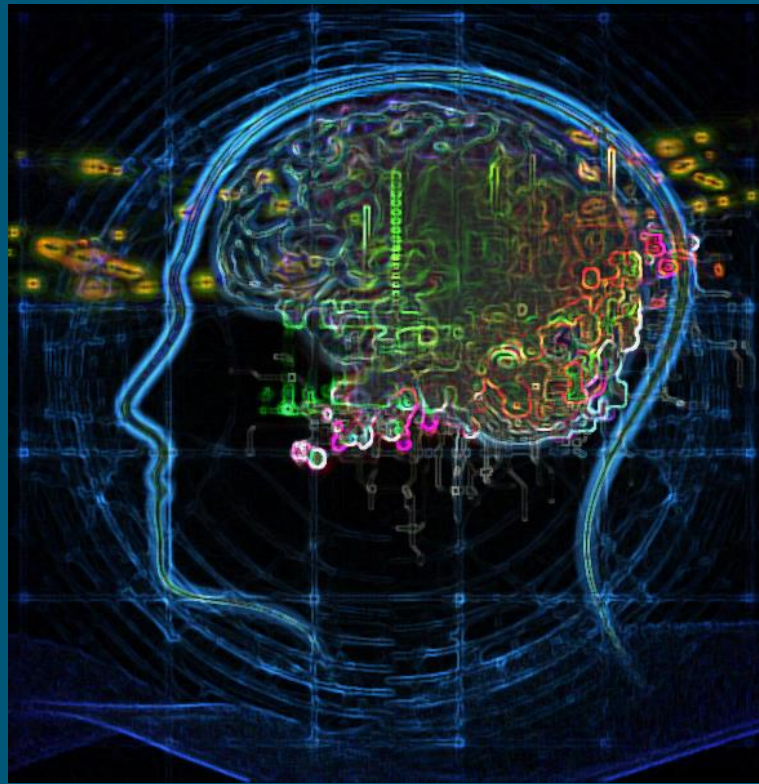
Koomey's Law



A 1984 X-MP/48 cost about US\$15 million *plus the cost of disks.*



64GB²
\$749.00
Available to ship:
In Stock



Artificial/Machine Intelligence

Machine Learning

Deep Learning

Forbes: *Where would you say we are on the continuum of developing true artificial intelligence?*

Hinton: I think we have crossed a very important threshold. Until fairly recently, most people in AI were doing a kind of AI that was inspired by logic. The paradigm for intelligence was logical reasoning and the idea of what an internal representation would look like was it would be some kind of symbolic structure. **That has completely changed with these big neural nets.**

We just think you can have these great big neural nets that learn, and so, **instead of programming, you are just going to get them to learn everything.** For many, many years, people in AI thought that was just fantasy . . .

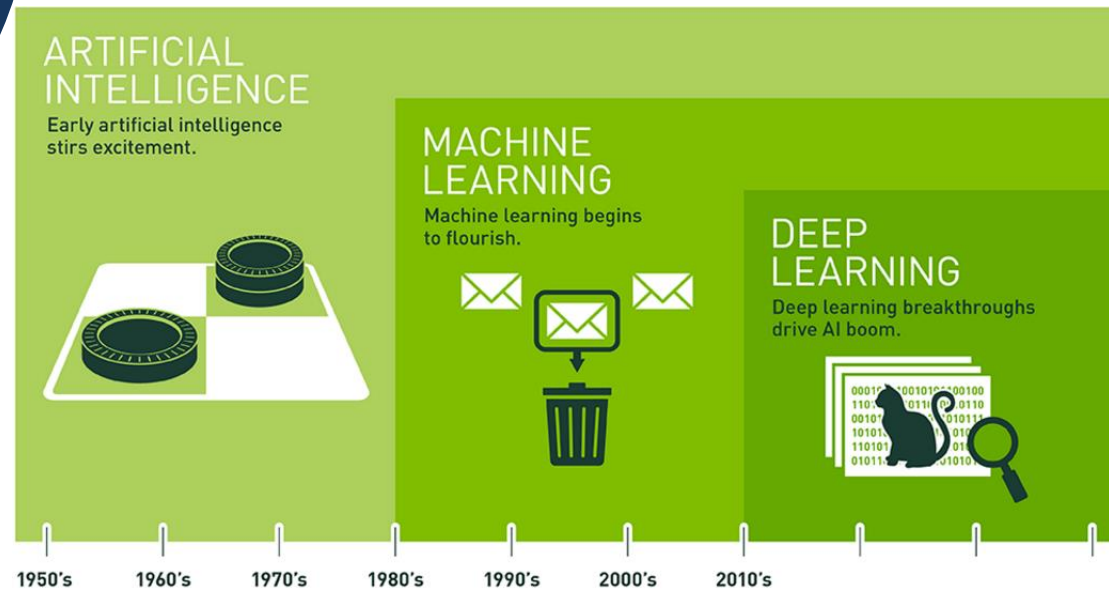
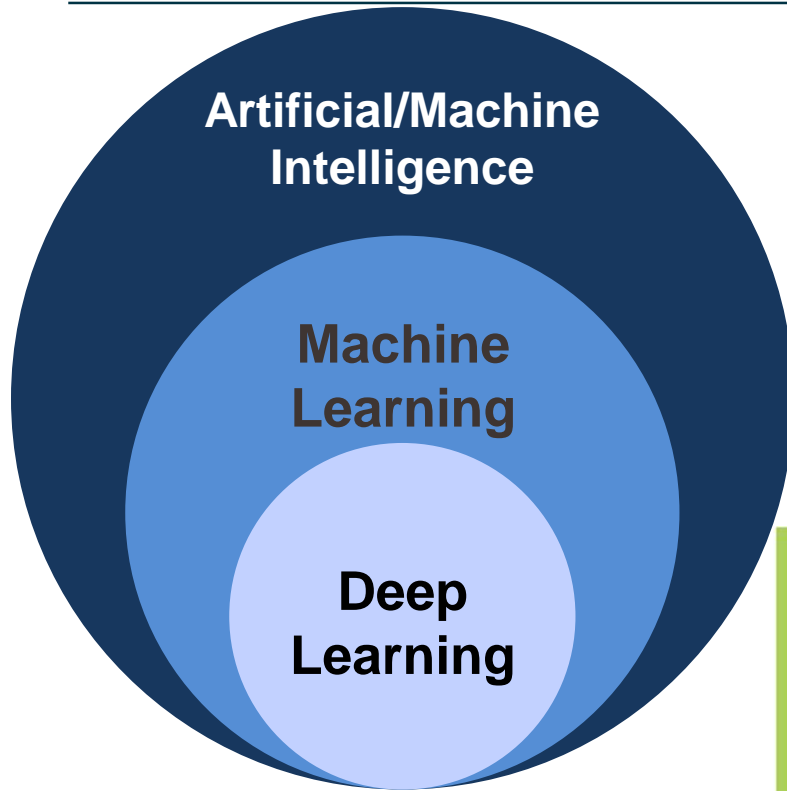
Now, our techniques scale up: you make computation more powerful, we can make you better models; you give us bigger data sets, we can make you better models. That is not true if you program everything. **So this stuff scales better than AI ever did in the past.**



Geoffrey Everest Hinton FRS is a British-born Canadian cognitive psychologist and computer scientist, most noted for his work on artificial neural networks. As of 2015 he divides his time working for Google and University of Toronto.

<https://www.forbes.com/sites/peterhigh/2016/06/20/deep-learning-pioneer-geoff-hinton-helps-shape-googles-drive-to-put-ai-everywhere/#4b3bdfa1693c>

AI ≥ Machine Learning ≥ Deep Learning



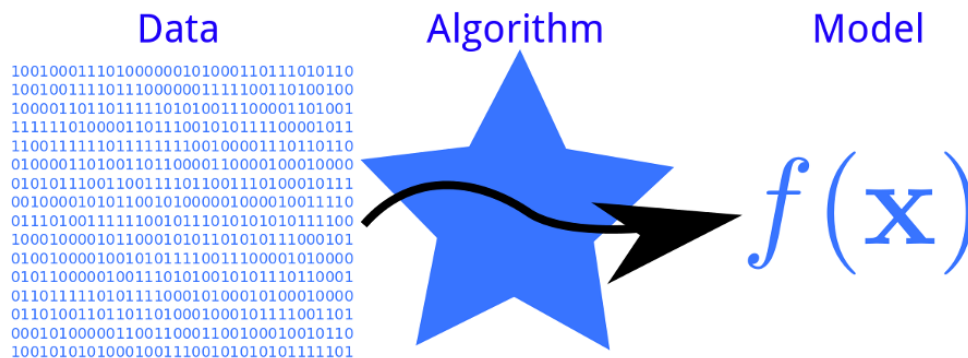
<https://blogs.nvidia.com/blog/2016/07/29/whats-difference-artificial-intelligence-machine-learning-deep-learning-ai/>

Artificial (Machine) Intelligence

- Artificial intelligence (AI) is intelligence exhibited by machines.
- In computer science, the field of AI research defines itself as the study of "intelligent agents": **any device that perceives its environment and takes actions that maximize its chance of success at some goal.**
- The central problems (goals) of AI research include
 - reasoning
 - knowledge representation
 - planning
 - **learning**
 - natural language processing (communication)
 - perception (vision, hearing, ...)
 - ability to move and manipulate objects.

Machine Learning

- Machine learning is a field of artificial intelligence
- Investigates how algorithms can learn from observations and data
 - In contrast to humans explicitly programming in each step of what the software should do.
- Machine learning algorithms find complex relationships in data.
 - Build models based on observations
 - Learn rules required to make optimum decisions or predictions
 - Recognize patterns in data that humans may not be able to recognize
 - Even if humans could recognize those patterns, they might not be consciously aware of how to formulate them in a way that could be programmed in traditional non-ML ways.
- Key aspect of machine learning is supplying a learning algorithm with data.
- The ML algorithm analyzes the data and constructs a model.
- As new input data is presented to the ML algorithm, it is fed into the model, the model processes the input, and outputs a decision or prediction.



Categories of Machine Learning

Supervised learning

- The computer is presented with example inputs (data) and their desired outputs (labels), given by a "teacher", and the goal is to learn a general rule that maps inputs to outputs.
- The data may be strongly or weakly labeled

Unsupervised learning

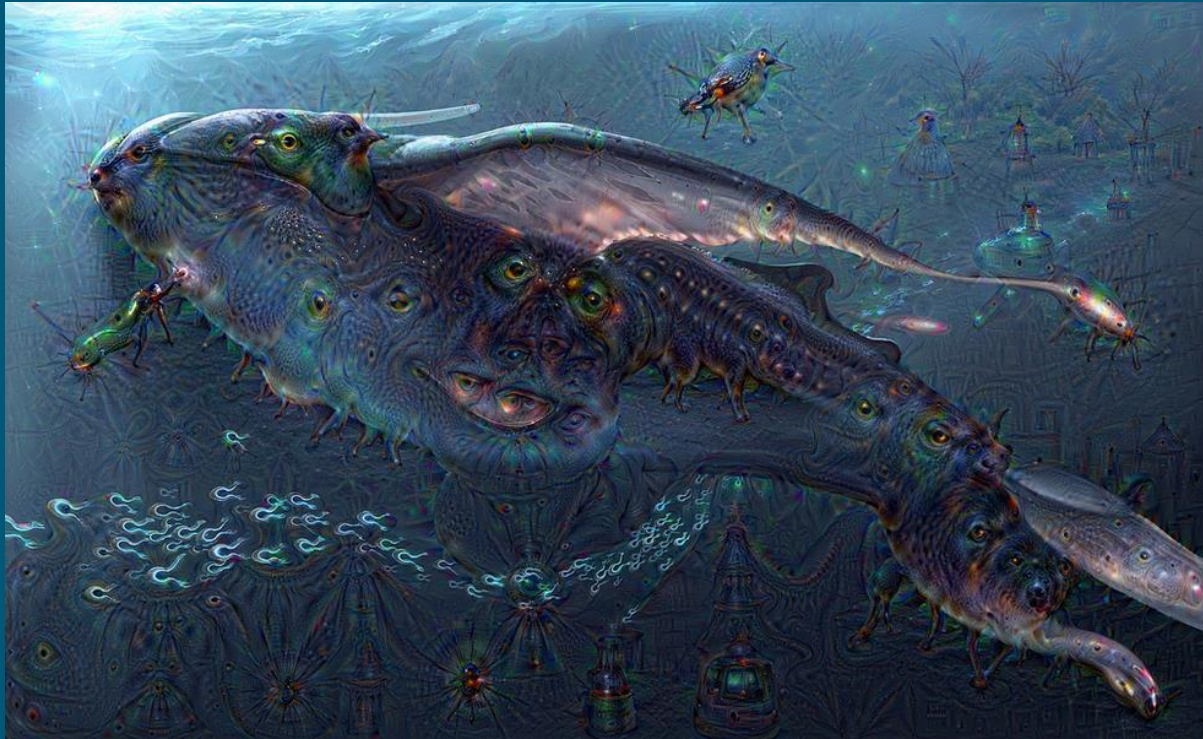
- No labels are given to the learning algorithm, leaving it on its own to find structure in its input.
- Unsupervised learning can be a goal in itself (discovering hidden patterns in data) or a means towards an end (feature learning).

Reinforcement learning

- A computer program interacts with a dynamic environment in which it must perform a certain goal (such as driving a vehicle or playing a game against an opponent).
- The program is provided feedback in terms of rewards and punishments as it navigates its problem space.

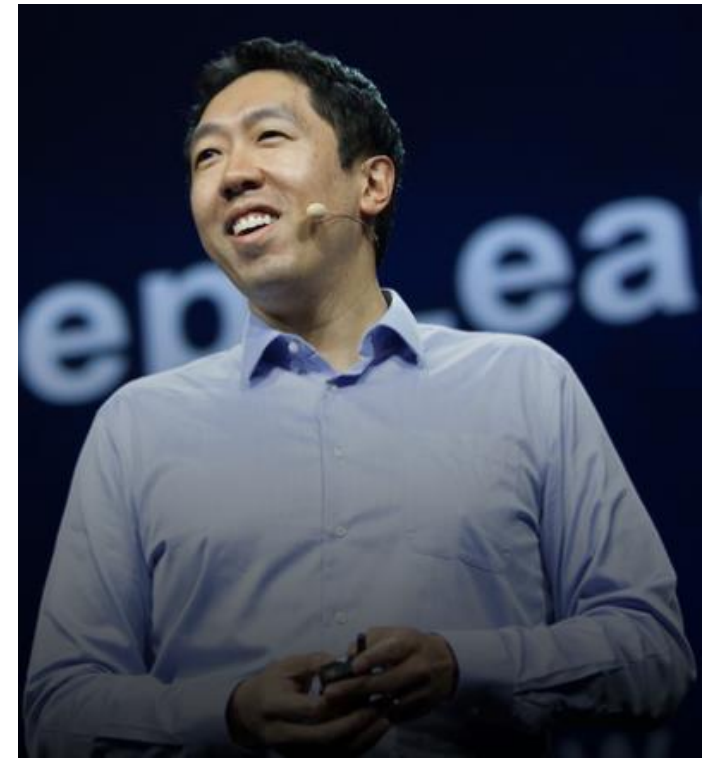
Semi-supervised learning

- Teacher gives an incomplete training signal: a training set with some (often many) of the target outputs missing.



Diving Into Deep Learning

We like to make the analogy that **data is like fuel for a rocket**, and deep neural network models are the rocket engine. **Because we have more data than ever, we are able to build bigger rockets**, with bigger engines that can take us to new places.



Andrew Yan-Tak Ng is a Chinese American computer scientist. He is the former chief scientist at Baidu[§] Research in Silicon Valley, Co-Chairman and Co-founder of Coursera; Adjunct Professor of Computer Science at Stanford University

[§]On March 20, 2017, Ng announced that he was resigning from Baidu.

Machine Learning

Free trial not offered for this course.

The Coursera logo, featuring the word "coursera" in a blue, lowercase, sans-serif font.

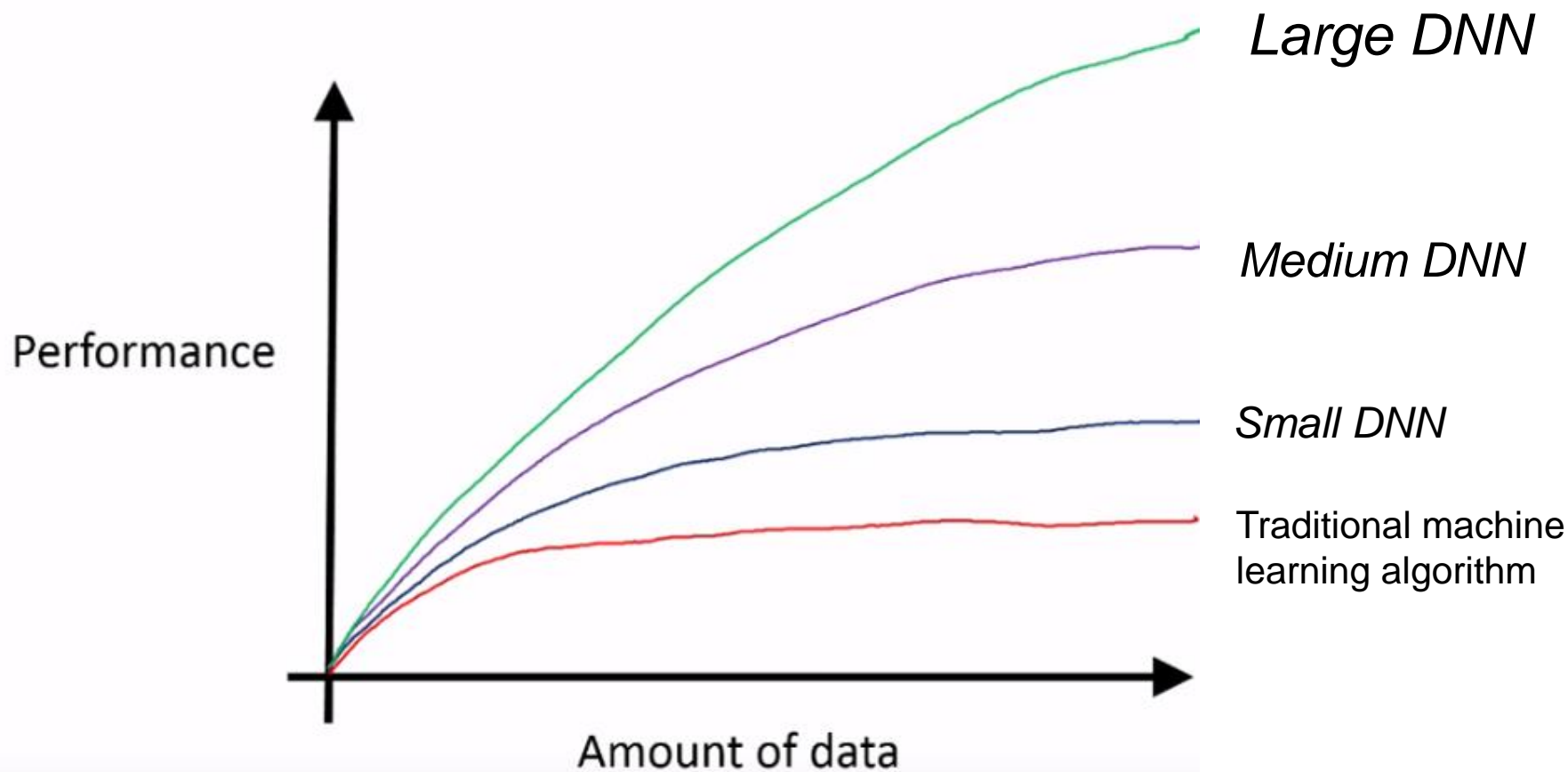
Enroll Now

Starts Mar 20



Rise of Deep Learning

More Data for Training → Improved Performance



Source: <https://www.youtube.com/watch?v=LcfLo7YP8O4>
Andrew Ng, Chief Scientist at Baidu: How Scale is Enabling Deep Learning

DNN = Deep Neural Net

Datasets Over Algorithms – Summary

Year	Breakthroughs in AI	Datasets (First Available)	Algorithms (First Proposed)
1994	Human-level spontaneous speech recognition	Spoken Wall Street Journal articles and other texts (1991)	Hidden Markov Model (1984)
1997	IBM Deep Blue defeated Garry Kasparov	700,000 Grandmaster chess games, aka "The Extended Book" (1991)	Negascout planning algorithm (1983)
2005	Google's Arabic- and Chinese-to-English translation	1.8 trillion tokens from Google Web and News pages (collected in 2005)	Statistical machine translation algorithm (1988)
2011	IBM Watson became the world Jeopardy! champion	8.6 million documents from Wikipedia, Wiktionary, Wikiquote, and Project Gutenberg (updated in 2010)	Mixture-of-Experts algorithm (1991)
2014	Google's GoogLeNet object classification at near-human performance	ImageNet corpus of 1.5 million labeled images and 1,000 object categories (2010)	Convolution neural network algorithm (1989)
2015	Google's Deepmind achieved human parity in playing 29 Atari games by learning general control from video	Arcade Learning Environment dataset of over 50 Atari games (2013)	Q-learning algorithm (1992)
Average No. of Years to Breakthrough:		3 years	18 years

In deep learning, the algorithms used now are versions of the algorithms originally developed in the 1980s/1990s. People were very optimistic about them, but it turns out they didn't work too well. Now we know the reason is they didn't work too well is that we didn't have powerful enough computers, and we didn't have large enough data sets to train them.

Algorithmic Efficiency



In 1997, IBM's Deep Blue beat Kasparov, searching 200m positions per second.

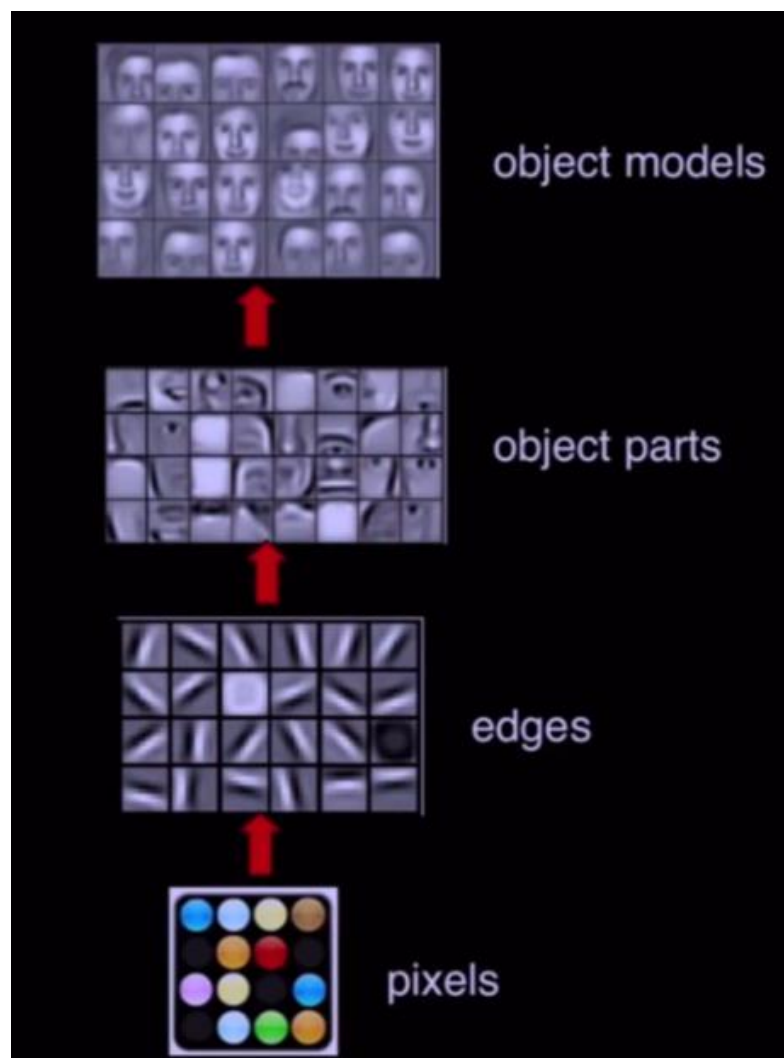


In 2009, a Pocket Fritz running on a mobile phone, won a professional tournament with 9 wins & 1 draw

Searched fewer than 20k positions/s, a **10,000x improvement** in 12 years

Deep Learning Networks

- Attempt to model high-level abstractions in data by using a **deep graph** with **multiple processing layers**, composed of multiple linear and non-linear transformations
- Various deep learning architectures:
 - Deep Neural Networks (DNN) in general
 - CNN: Convolutional Neural Networks
 - RNN: Recurrent Neural Networks
 - . . .
- Replace hodge-podge of handcrafted features with efficient algorithms for unsupervised or semi-supervised feature learning and hierarchical feature extraction.



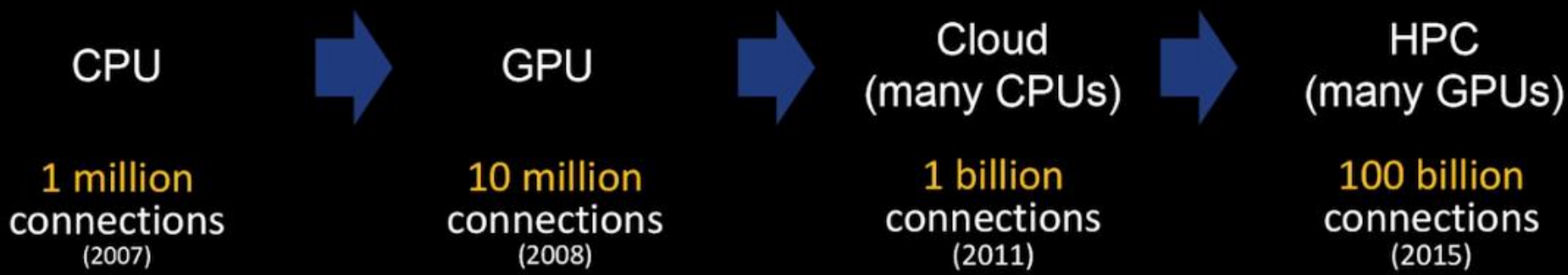
Source: <https://www.youtube.com/watch?v=O0VN0pGgBZM>

Andrew Ng, Chief Scientist at Baidu: Why should you care about deep learning?

Source: https://en.wikipedia.org/wiki/Deep_learning

Deep Learning Powered by Computing Advances

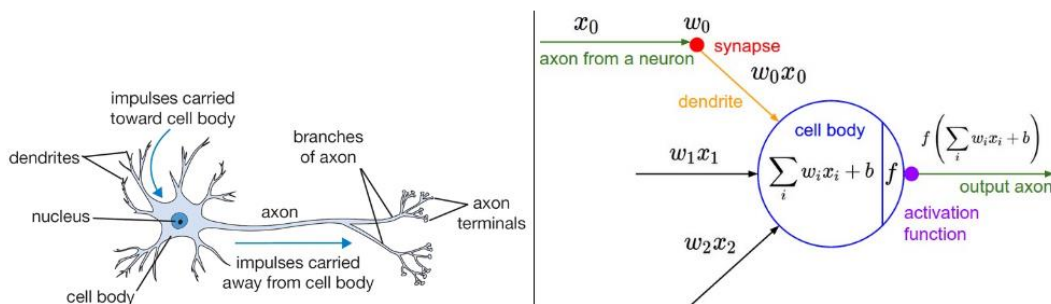
Rocket engines: Deep Learning driven by scale



Source: <https://www.youtube.com/watch?v=O0VN0pGgBZM>

Andrew Ng, Chief Scientist at Baidu: Why should you care about deep learning?

ANN – Artificial Neural Network



A cartoon drawing of a biological neuron (left) and its mathematical model (right).

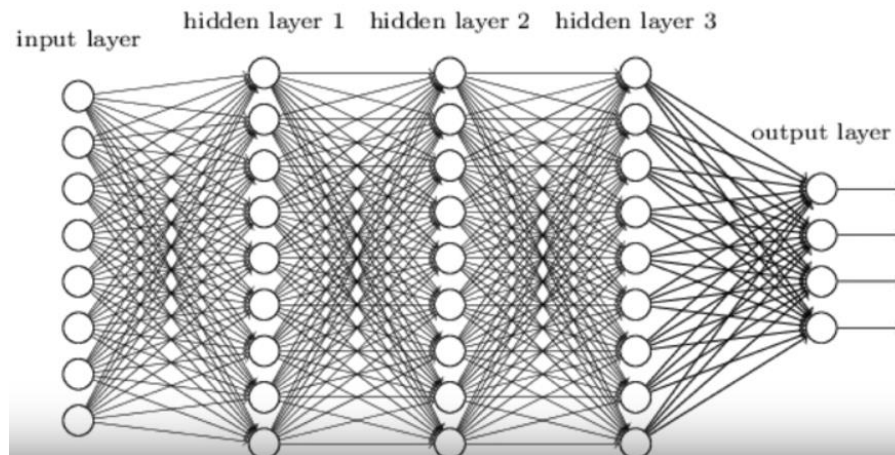
Artificial Neuron (Unit)

- applies weights to inputs (dot product)
- passes the result through a non-linear activation function
- feeds its output forward to all connected neurons of next layer

(Artificial) Neural Networks are inspired by our understanding of the biology of our brains – all those interconnections between the neurons. But, unlike a biological brain where any neuron can connect to any other neuron within a certain physical distance, these artificial neural networks have discrete layers, connections, and directions of data propagation.

Deep Neural Networks

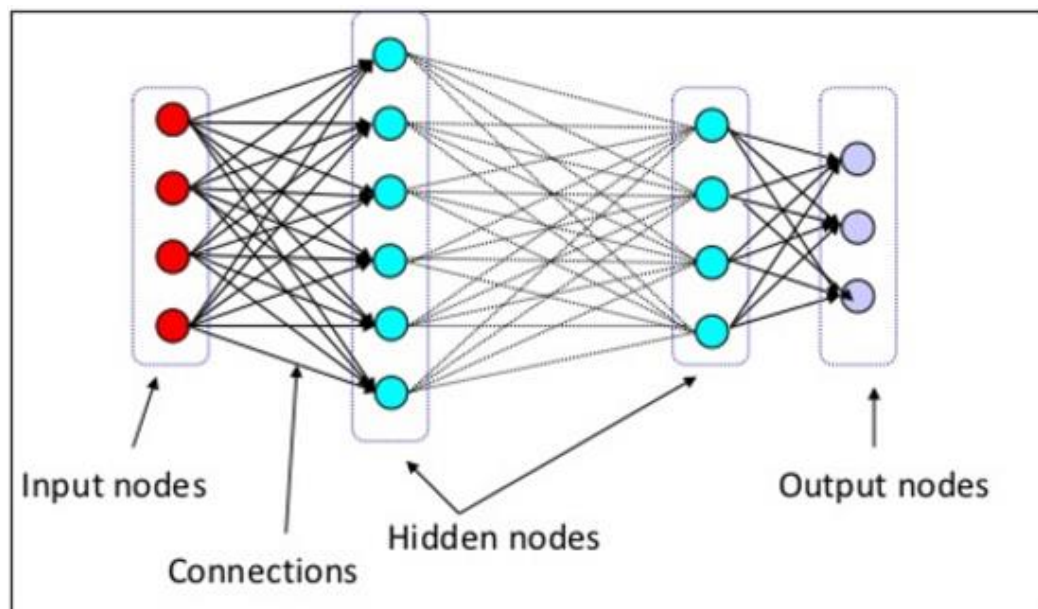
- Neural network with hidden layers is called a Deep Neural Net (DNN)
- Classic DNNs are fully connected
 - Neurons connect across adjacent layers, never within a layer
- Numerous architectural variations exist



<http://cs231n.github.io/neural-networks-1/>

<https://www.slideshare.net/hammadwan/deep-neural-networks>

Deep (Layered) Networks (~1985)



$$\begin{aligned} \text{Output: } y_i &= f(w_i^1 x_1 \mid w_i^2 x_2 \mid w_i^3 x_3 \mid \dots \mid w_i^m x_m) \\ &= f\left(\sum_j w_i^j x_j\right) \end{aligned}$$

Neuron = Node = Unit

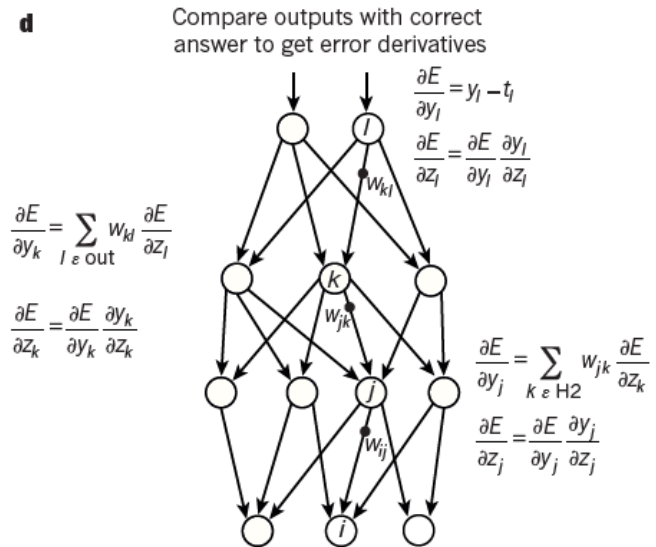
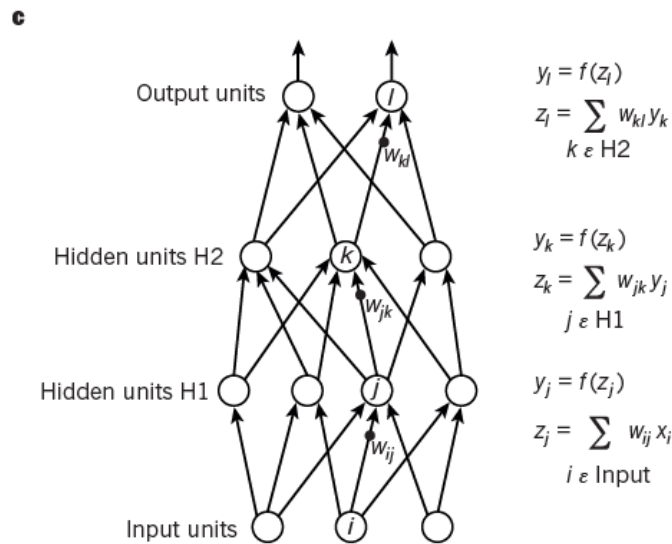
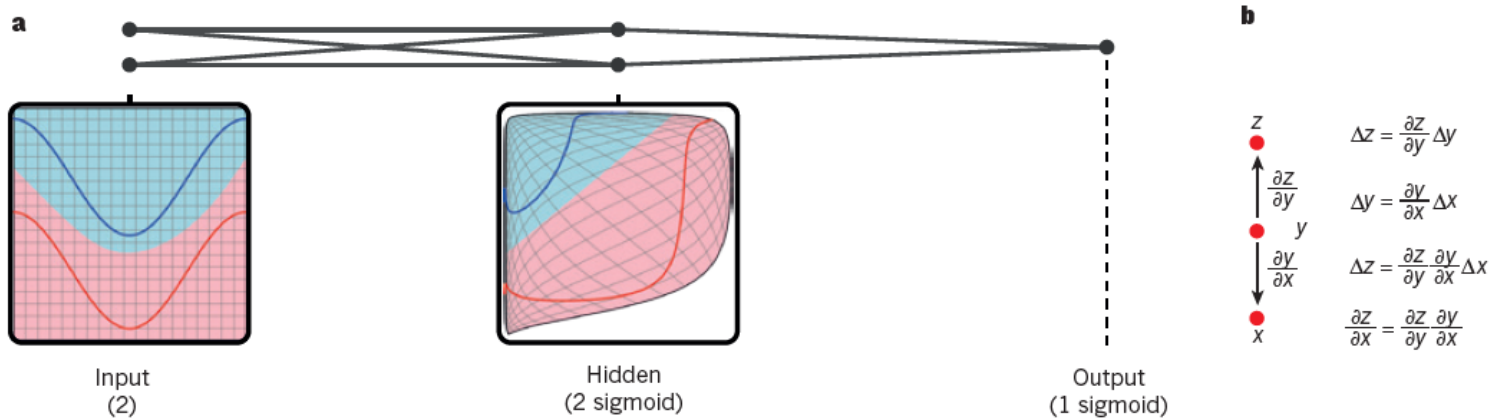
Mathematical insight:

Neural Networks with at least one hidden layer are universal approximators.

Given any continuous function $\mathbf{f}(\mathbf{x})$ and some $\epsilon > 0$ there exists a Neural Network $\mathbf{g}(\mathbf{x})$ with one hidden layer (with a reasonable choice of non-linearity for activation, e.g. the sigmoid function) such that $\forall \mathbf{x}, |\mathbf{f}(\mathbf{x}) - \mathbf{g}(\mathbf{x})| < \epsilon$. In other words, the neural network can approximate any continuous function.

<https://www.slideshare.net/hammawan/deep-neural-networks>

Key Advance: Backpropagation



The Magic of DNNs:

Basic Ingredients

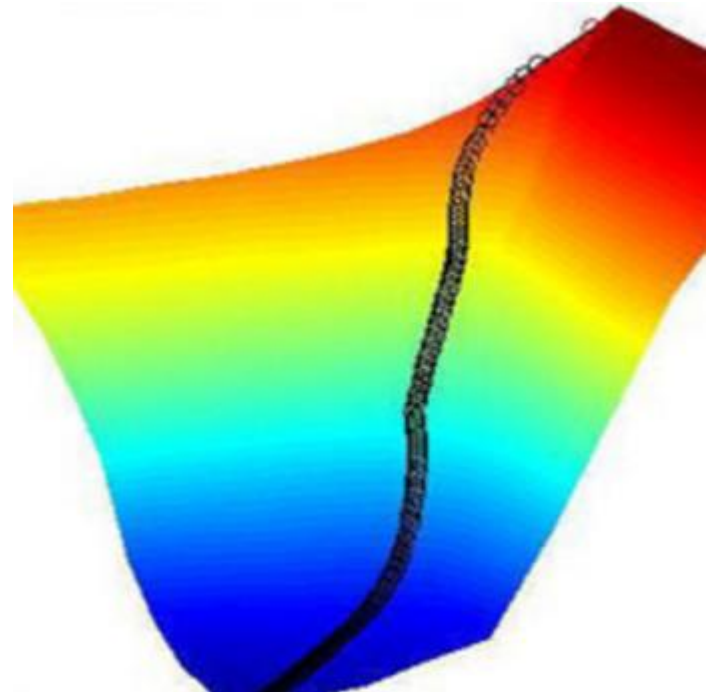
Ingredients:

- Overall architecture: number of layers (depth), number and width of filters/kernels, width of each layer
- Inputs: tensors of integers or single precision floating point numbers
- Outputs: tensors of integers or single precision floating point numbers
- Network internal data: single precision floating point numbers
 - Learnable parameters: tensors of weights and biases at each layer
 - Hyper-parameters: learning rate, mini-batch size, number of epochs (iterations) for training, weight initialization, drop-out rate, ...
- Operations
 - Dot products: floating point multiplication and addition
 - ReLU (Rectified Linear Unit): activation function $f(x) = \max(0, x)$
 - Max pooling: $\max(\text{pixel-region})$
- Algorithmic steps of SGD (stochastic gradient descent)
 - Initialization of weights and biases with random floating point numbers
 - Feed forward mechanism to compute each activation layer
 - Determine symbolic partial derivatives (built into programming framework)
 - Backward propagation algorithm
- Fast parallel hardware – GPUs or TPUs

Stochastic Gradient Descent (SGD)

Stochastic gradient descent (SGD), also known as incremental gradient descent, is a stochastic approximation of the gradient descent optimization method for minimizing an objective function that is written as a sum of differentiable functions.

In other words, SGD tries to find minima or maxima by iteration.



The Magic of DNNs: Learning Algorithm

Basic supervised learning algorithm for training the network from scratch

- Initialize the hyper-parameters
 - Initialize the network nodes/neurons/units (learnable parameters) with properly chosen random numbers
 - **Iterate**
 - pick some training example (*input*, *label*)
 - run the DNN on input [feed forward mechanism]
 - adjust weights at each layer to make the computed output closer to the desired output (*label*) [backpropagation]
- until done**

Andrej Karpathy

Software 2.0

Neural networks are not just another classifier, they represent the beginning of a fundamental shift in how we write software.

They are Software 2.0.



Software 1.0

- Classical software stack consists of explicit instructions to the computer written by a programmer.
- By writing each line of code, the programmer is identifying a specific point in program space with some desirable behavior.
- Programming languages: Java, C++, Python, ...
- Programming paradigms: procedural, functional, object-oriented, ...

Andrej Karpathy

Software 2.0

Neural networks are not just another classifier, they represent the beginning of a fundamental shift in how we write software.

They are Software 2.0.

Software 2.0

- Software 2.0 is written in neural network weights.
- No human is involved in writing this code because there are a lot of weights (typical networks might have millions), and coding directly in weights is kind of hard (I tried).
- Instead, we specify some constraints on the behavior of a desirable program (e.g., a dataset of input output pairs of examples) and use the computational resources at our disposal to search the program space for a program that satisfies the constraints.

Andrej Karpathy

Software 2.0

Software 2.0

- It turns out that a large portion of real-world problems have the property that it is significantly easier to collect the data than to explicitly write the program.
- **A large portion of programmers of tomorrow do not maintain complex software repositories, write intricate programs, or analyze their running times. They collect, clean, manipulate, label, analyze and visualize data that feeds neural networks.**
- **Software 2.0 is not going to replace 1.0 (indeed, a large amount of 1.0 infrastructure is needed for training and inference to “compile” 2.0 code), but it is going to take over increasingly large portions of what Software 1.0 is responsible for today.**

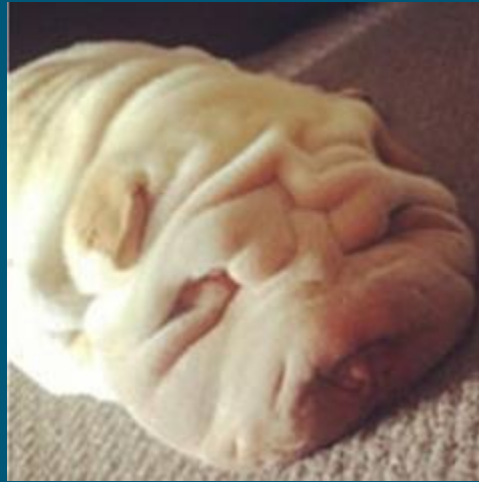


Image Recognition is Hard ...







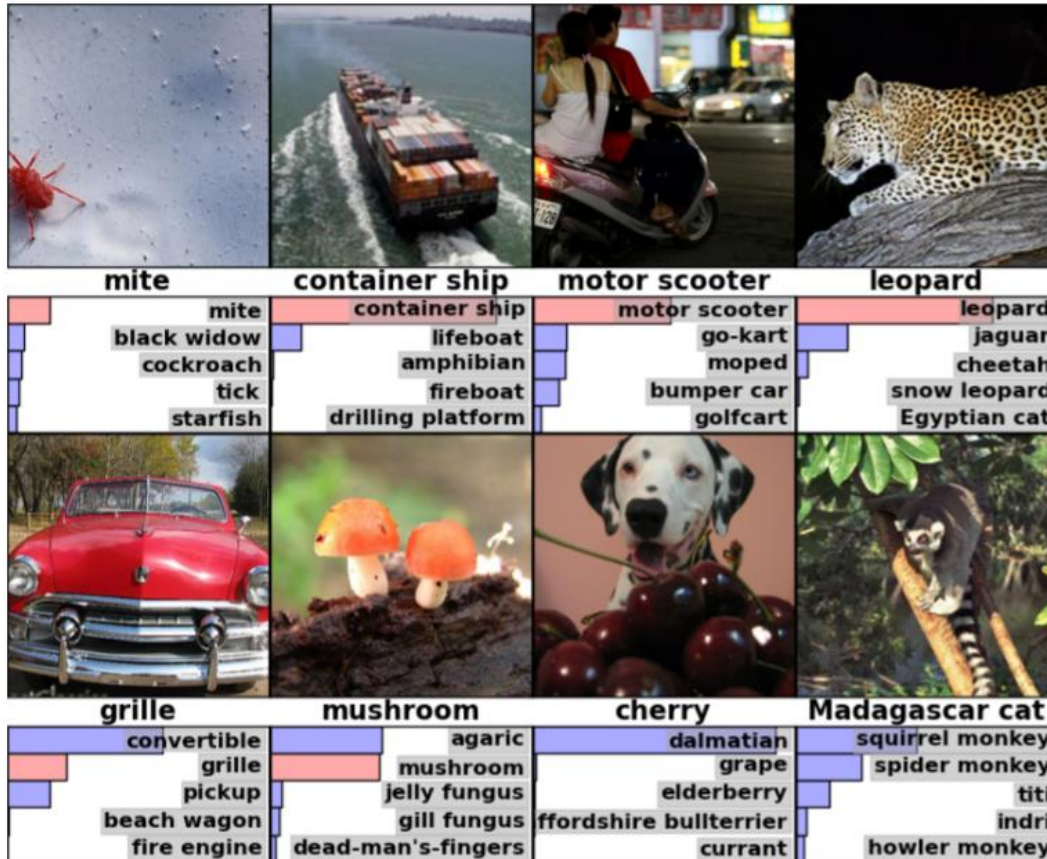




Deep Learning for Computer Vision

ImageNet Challenge (ILSVRC)

ImageNet consists of over 15 million labeled high-resolution images in over 22,000 categories



Eight ILSVRC-2010 test images and the five labels considered most probable by our model. The correct label is written under each image, and the probability assigned to the correct label is also shown with a red bar (if it happens to be in the top 5)

<http://www.cs.toronto.edu/~fritz/absps/imagenet.pdf>

Challenge:

- Given an image, predict in which one of 1,000 categories it falls

Since 2010, the annual ImageNet Large Scale Visual Recognition Challenge (ILSVRC) has been conducted as a competition where research teams submit programs that classify and detect objects and scenes.

Breakthrough with AlexNet: Convolutional Neural Net for the 2012 ILSVRC

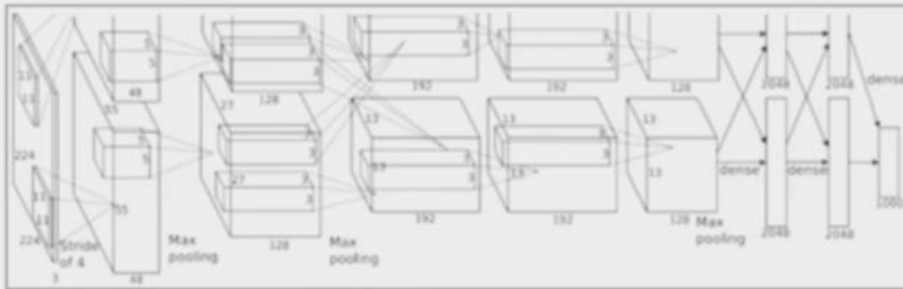
ImageNet Classification with Deep
Convolutional Neural Networks
[Krizhevsky, Sutskever, **Hinton**, 2012]



IMAGENET

Deng et al.
Russakovsky et al.

“AlexNet”



NVIDIA et al.

- Network architecture similar to LeNet from 1998
- Much bigger than LeNet; trained on much bigger dataset; used GPUs for training
- First time that Deep Neural Networks were noticed by the computer vision community and adopted for larger images
- AlexNet won the ImageNet challenge: ILSVRC (ImageNet Large Scale Visual Recognition Challenge)

The Magic of ConvNets

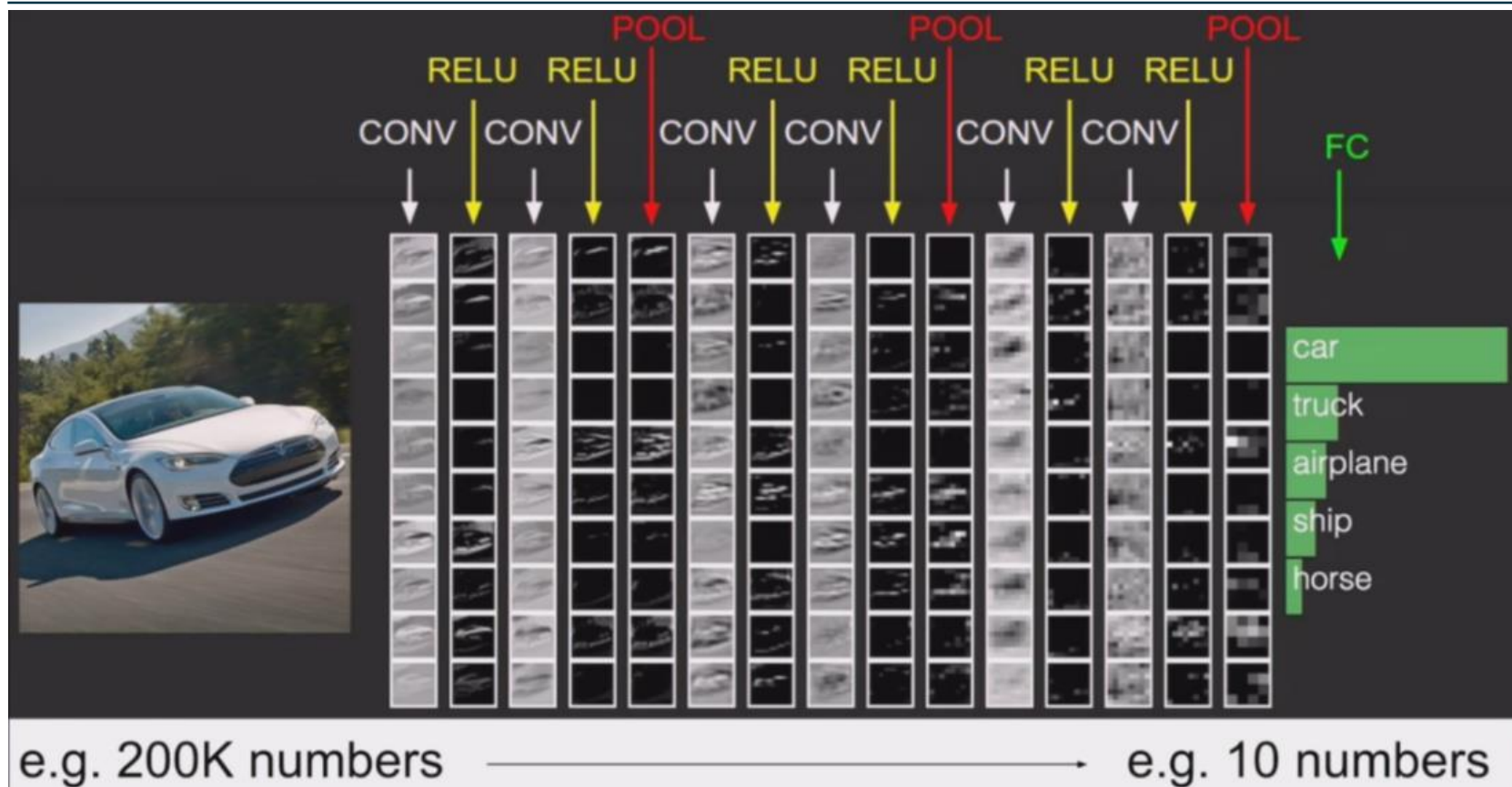


Only two basic operations are involved throughout:

1. Dot products $\mathbf{w}^T \mathbf{x}$
2. Max operations $\max(\cdot)$

parameters
(~10M of them)

High Level Structure of a ConvNet



- Use of convolutional layers is computationally much more efficient than using fully connected layers everywhere (far fewer weights/parameters and multiplications)
- The ReLU (Rectified Linear Units) layers apply simple non-linearities
- The Pool layers apply max() operations – similar to image downsampling
- The final fully connected layer(s) contain(s) neurons that connect to the entire input image

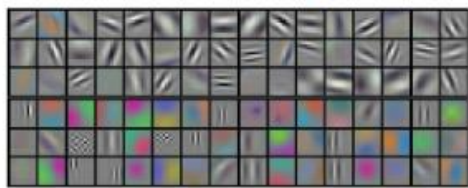
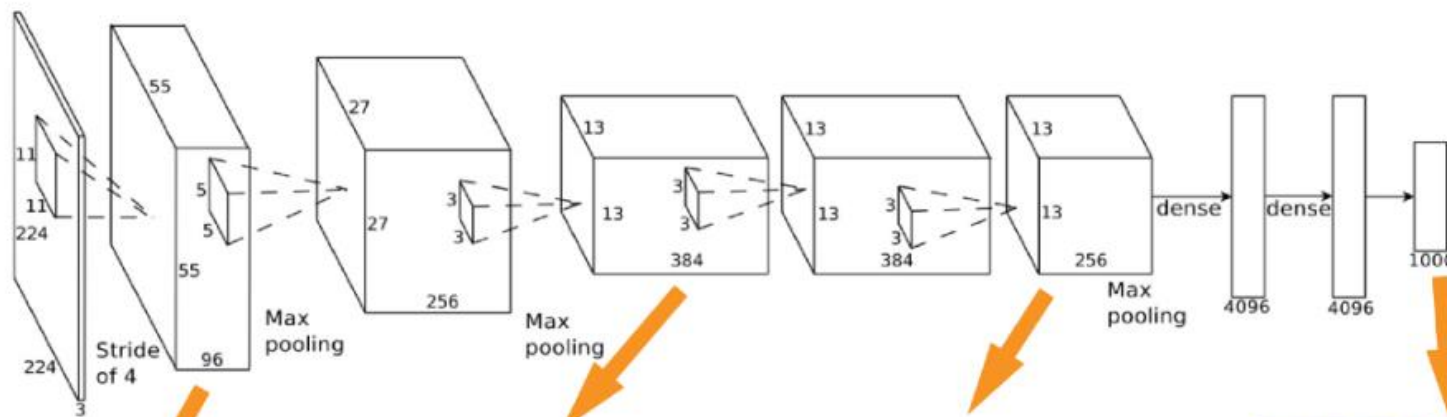
Deep Visualization Toolbox

Allows you to interactively peek inside a CNN to see how it works.
Great Fun!

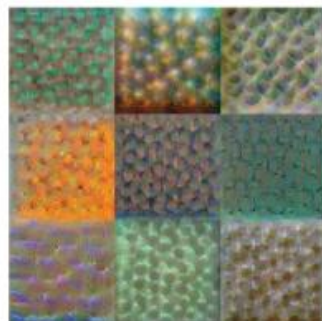


<https://www.youtube.com/watch?v=AgkfIQ4IGaM>
<http://yosinski.com/deepvis>

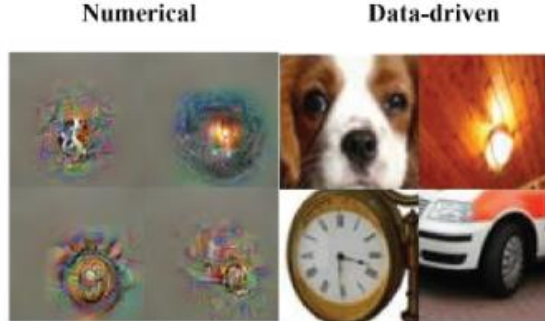
Peeking Into AlexNet



Conv 1: Edge+Blob



Conv 3: Texture



Conv 5: Object Parts

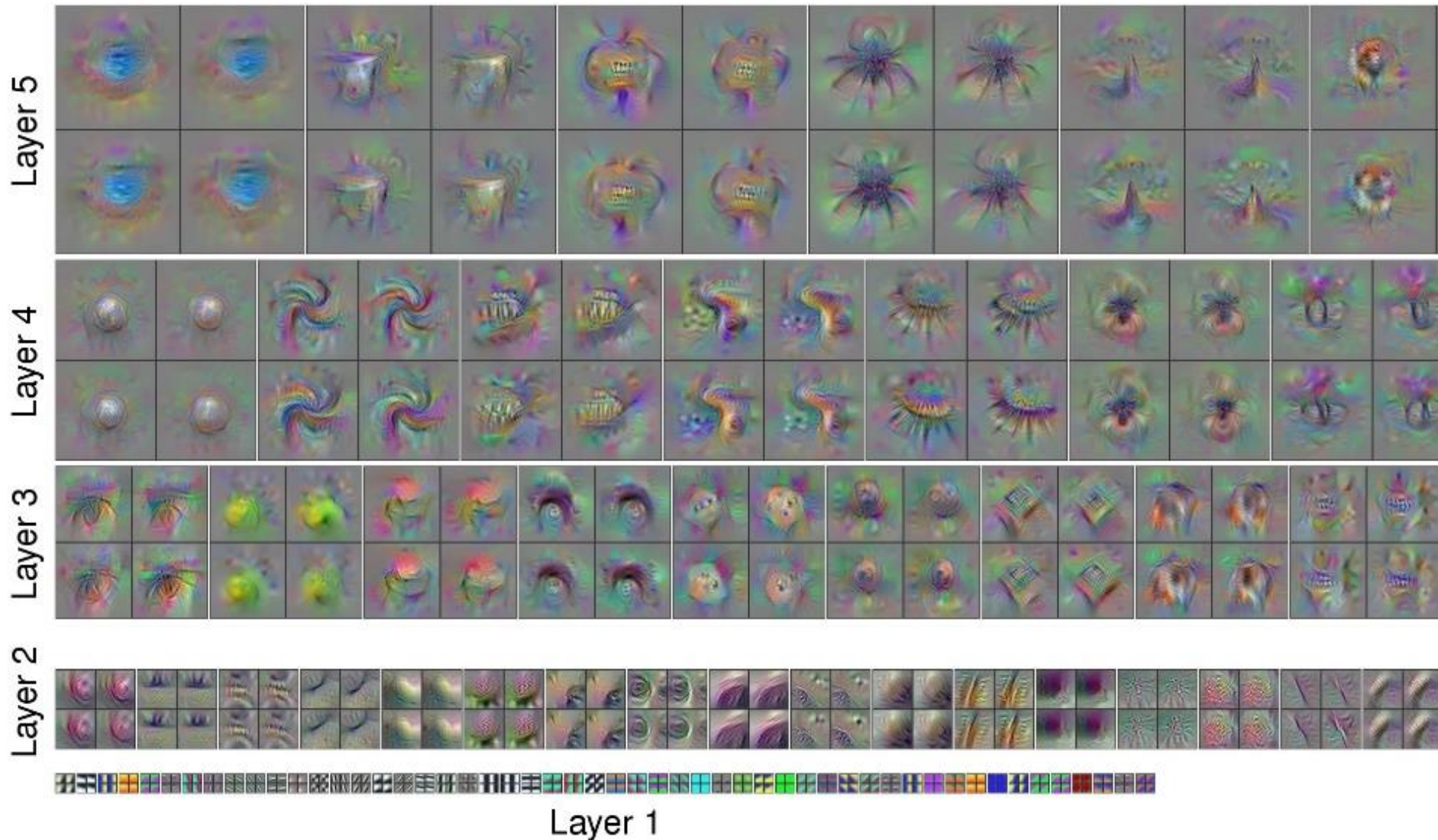


Fc8: Object Classes

What AlexNet sees at every step

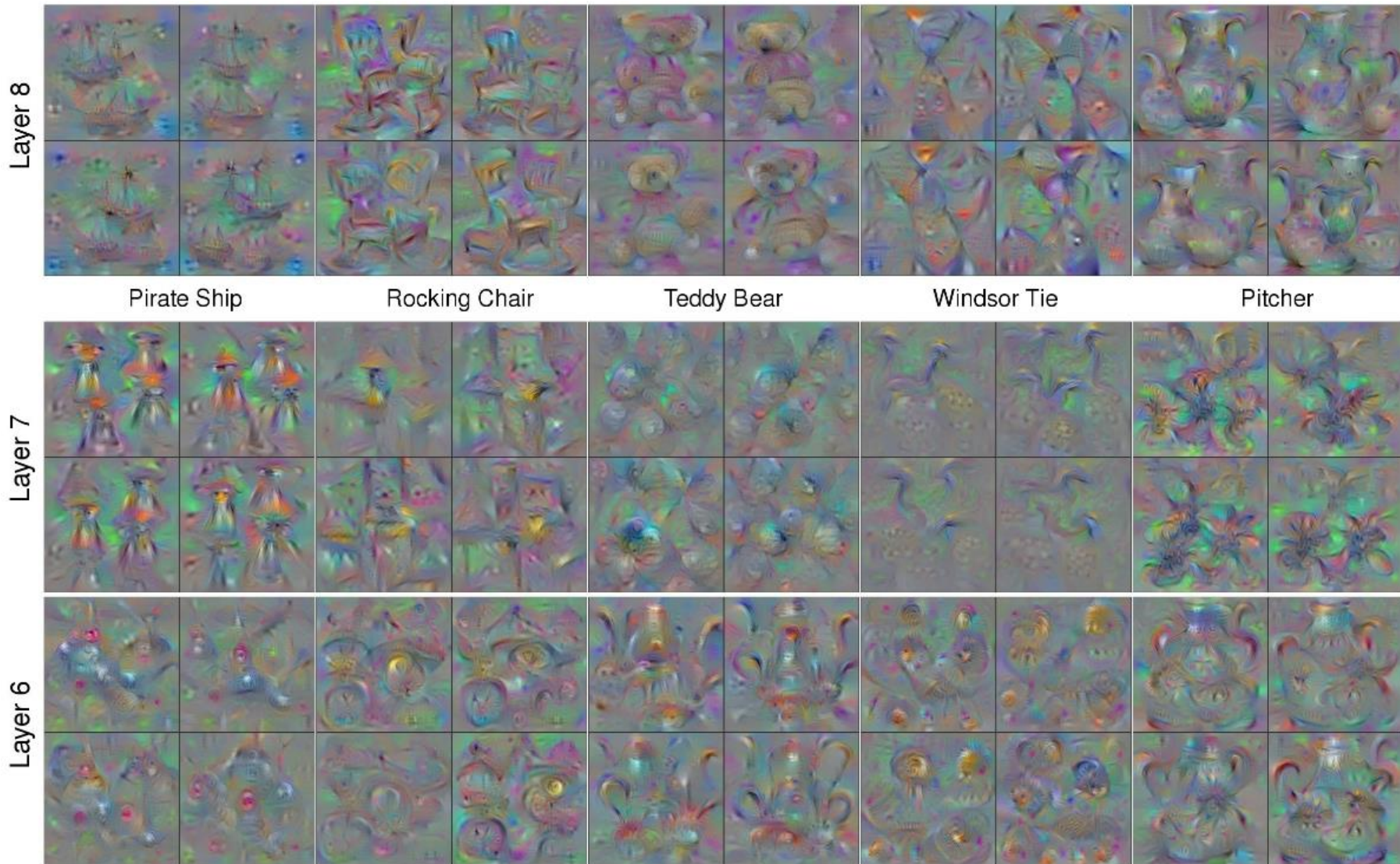
“Conv n” means “Convolutional Layer n”

What CNNs “See” at Different Layers



<http://yosinski.com/deepvis>

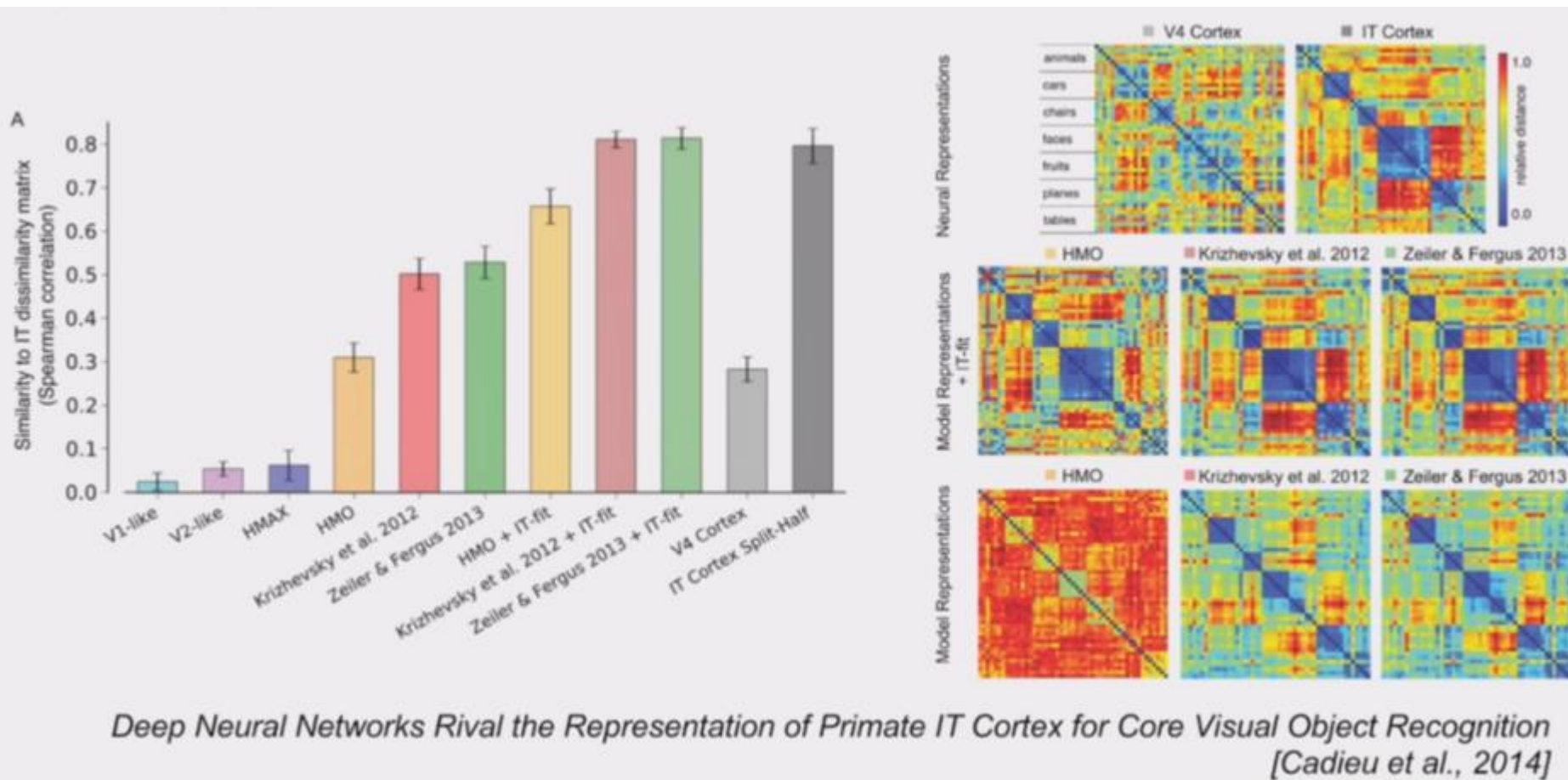
What CNNs “See” at Different Layers



Summary

- One can recognize important features at different scales
 - Edges L1
 - Corners L2
 - Wheels L3
 - Eyes L3
 - Shoulders L4
 - Faces L5
 - Handles L6
 - Bottles, etc. L7
- Complexity increases in higher-layer features as they combine simpler features from lower layers.
- The variation of patterns also increases in higher layers
 - increasingly invariant, abstract representations are learned.
- In particular, the jump from Layer 5 (the last convolution layer) to Layer 6 (the first fully-connected layer) brings about a large increase in variation.

ConvNets and the Visual Cortex



- Architectures appear to be converging toward a structure that may resemble the way the visual cortex in primates works for object recognition

Application: Image Captioning



A man holding a tennis racquet on a tennis court.



Two pizzas sitting on top of a stove top oven



A group of young people playing a game of Frisbee

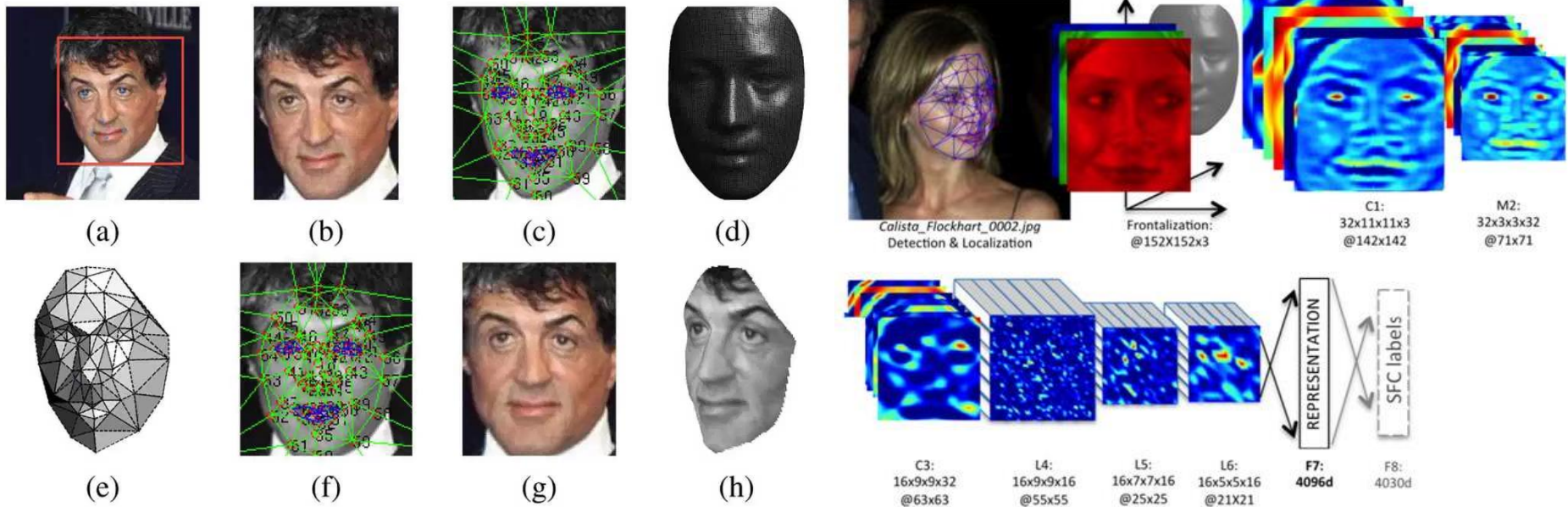


A man flying through the air while riding a snowboard

In cases like the last one, it becomes apparent that the DNN really has no understanding of our world.

Deep Face – Facebook Research 2014

- This deep network involves more than 120 million parameters
- Trained on the largest facial dataset to-date, an identity labeled dataset of 4 million facial images belonging to more than 4,000 identities.
- **Accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset**
- Reduced error of current state of the art by > 27%, closely approaching human-level performance



Style Transfer

Original



Style Guide



Result





Generated Image

Input



Style



Visual Attribute Transfer through Deep Image Analogy

<https://arxiv.org/pdf/1705.01088.pdf>

Visual Attribute Transfer through Deep Image Analogy

Style Guide

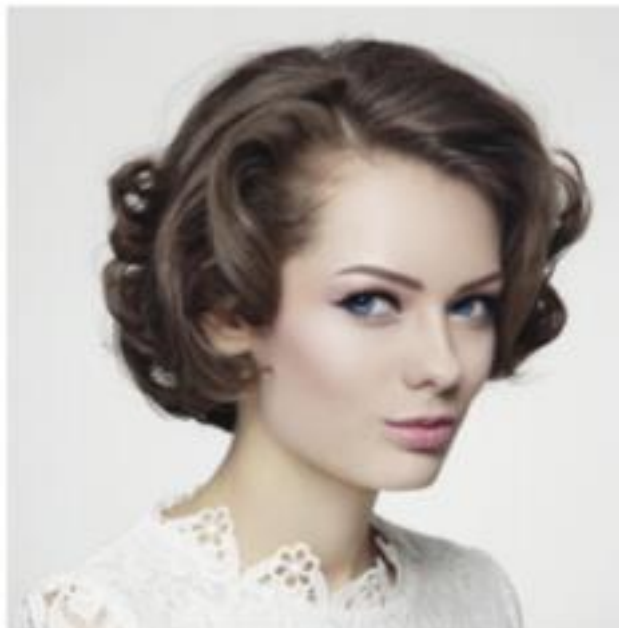


Input



Generated Image

Input



Input



Deep Dream Generator

Generated “Deep Dream”[§]

Input



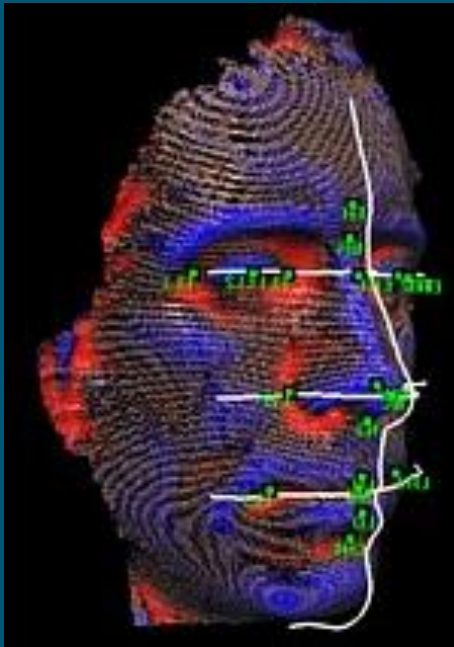
ConvNets and Other Types of DNNs: Powerful Tools Used Everywhere

Machine Vision Applications

- Photo search
- **Face recognition and verification**
- Recognition of house numbers
- Visual perception in autonomous vehicles (self driving cars), robots, drones
- Recognition of Chinese ideograms
- Analysis of radiological images for cancerous tumours, heart disease, pneumonia, etc.
- **Satellite image analysis – sweet water reservoirs**
- Recognition of galaxies
- Image captioning
- **Lip reading**

Other Applications

- Speech recognition
- Speech generation – WaveNet from DeepMind
- **Music generation**: classical music, pop songs
- Poetry generation: Haikus, Sonnets
- **Art**: style transfer, DeepDream hallucinations
- **Game playing**: Go, Atari games
- Drug design
- Genomic analysis



Applications

Image Recognition

Application: Image Captioning



A man holding a tennis racquet on a tennis court.



Two pizzas sitting on top of a stove top oven



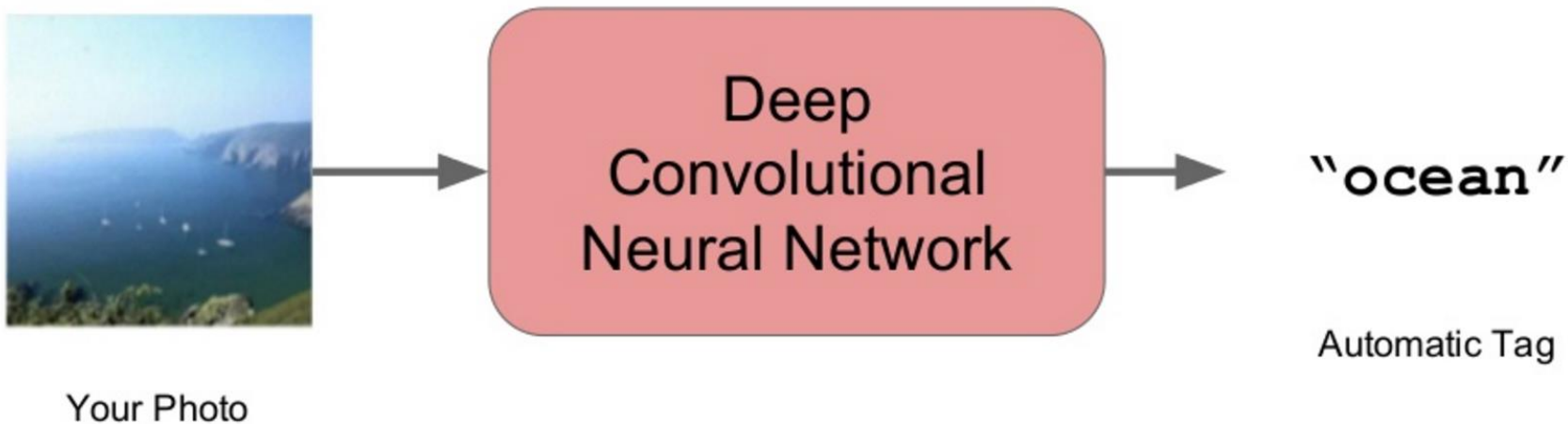
A group of young people playing a game of Frisbee



A man flying through the air while riding a snowboard

Application: Automatic Tagging of Photos

Google Photos Search

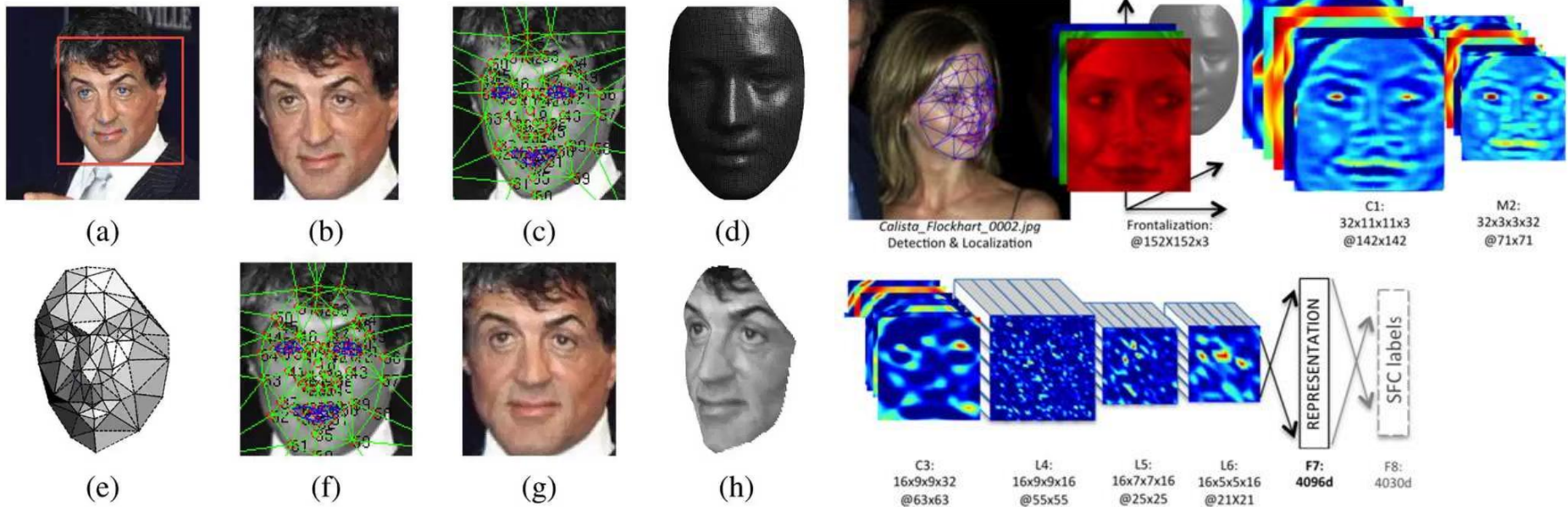


Search personal photos without tags.

Google Research Blog - June 2013

Deep Face – Facebook Research 2014

- This deep network involves more than 120 million parameters
- Trained on the largest facial dataset to-date, an identity labeled dataset of four million facial images belonging to more than 4,000 identities.
- Accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset
- Reduced error of current state of the art by > 27%, closely approaching human-level performance

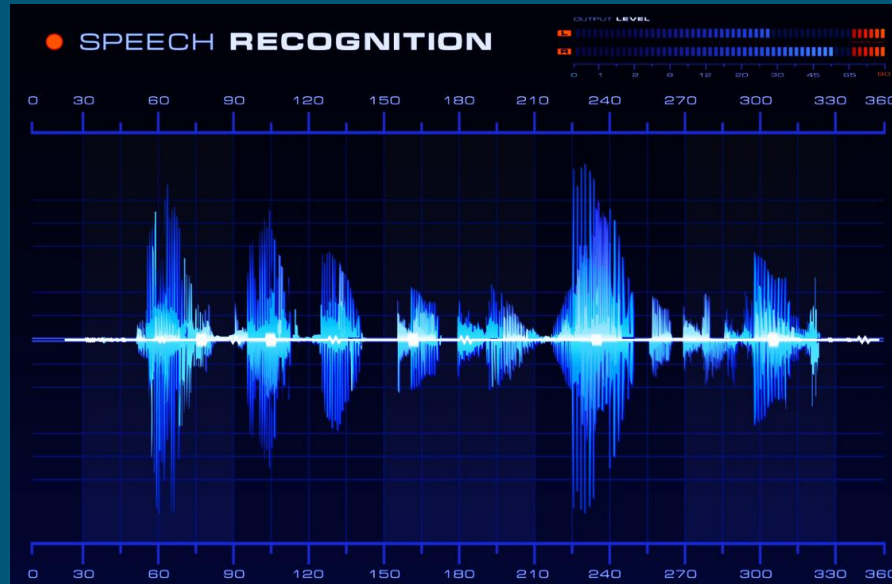


April 2016						
MO	DI	MI	DO	FR	SA	SO
28	29	30	31	1	2	3
4	5	6	7	8	9	10
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28	29	30	

Twitter: Content Analysis of Live Video Streams

- Twitter's AI team, known as **Cortex**, has developed an algorithm that can instantly recognize what's happening in a live feed.
- Twitter effectively built a custom supercomputer made entirely of graphics processing units (GPUs) to perform the video classification and serve up the results.
- The Cortex team has ambitions to develop a sophisticated recommendation system to help filter and curate all sorts of content shared through the service, based on a user's previous activity
- This is being tested on **Periscope**, an app owned by Twitter that lets users transmit live video from their smartphones.
- The team is using an approach known as deep learning to recognize the activity in clips.
- Such a tool could be useful for advertising, by algorithmically matching ads to the contents of videos as they are filmed and broadcast. As more and more video moves online, in fact, the algorithm could help Twitter tailor ads to such content a lot more efficiently.





Applications

Voice/Speech Recognition

Person to Machine (P2M) Voice Interaction Adoption Keys = 99% Accuracy in Understanding & Meaning + Low Latency

*As speech recognition accuracy goes from say 95% to 99%, all of us in the room will go from barely using it today to using it all the time. Most people underestimate the difference between 95% and 99% accuracy – **99% is a game changer...***

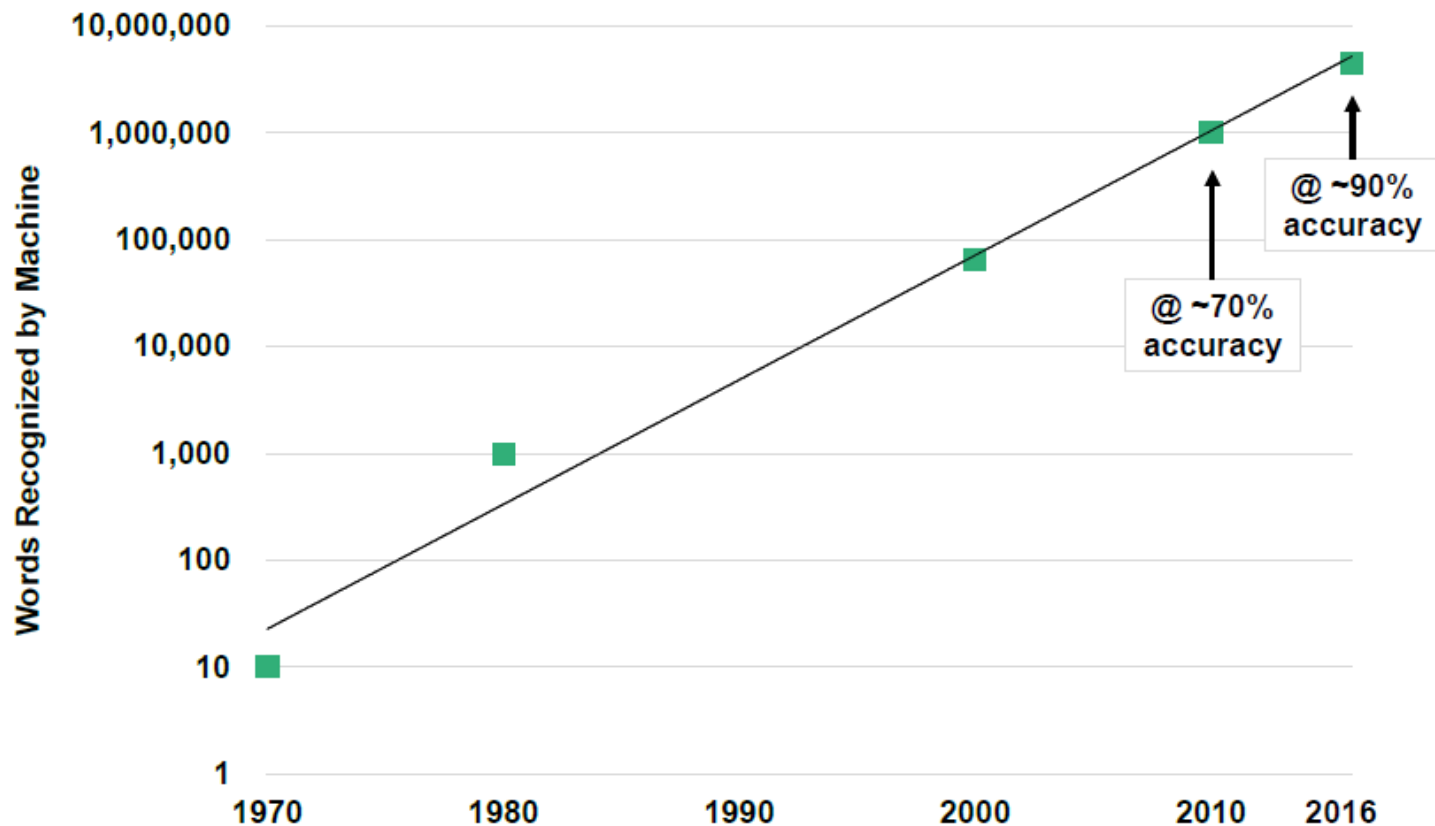
No one wants to wait 10 seconds for a response. Accuracy, followed by latency, are the two key metrics for a production speech system...

● **ANDREW NG, CHIEF SCIENTIST AT BAIDU**

Machine Speech Recognition @ Human Level Recognition for... Voice Search in Low Noise Environment, per Google

Next Frontier = Recognition in heavy background noise in far-field & across diverse speaker characteristics (accents, pitch...)

Words Recognized by Machine (per Google), 1970 – 2016

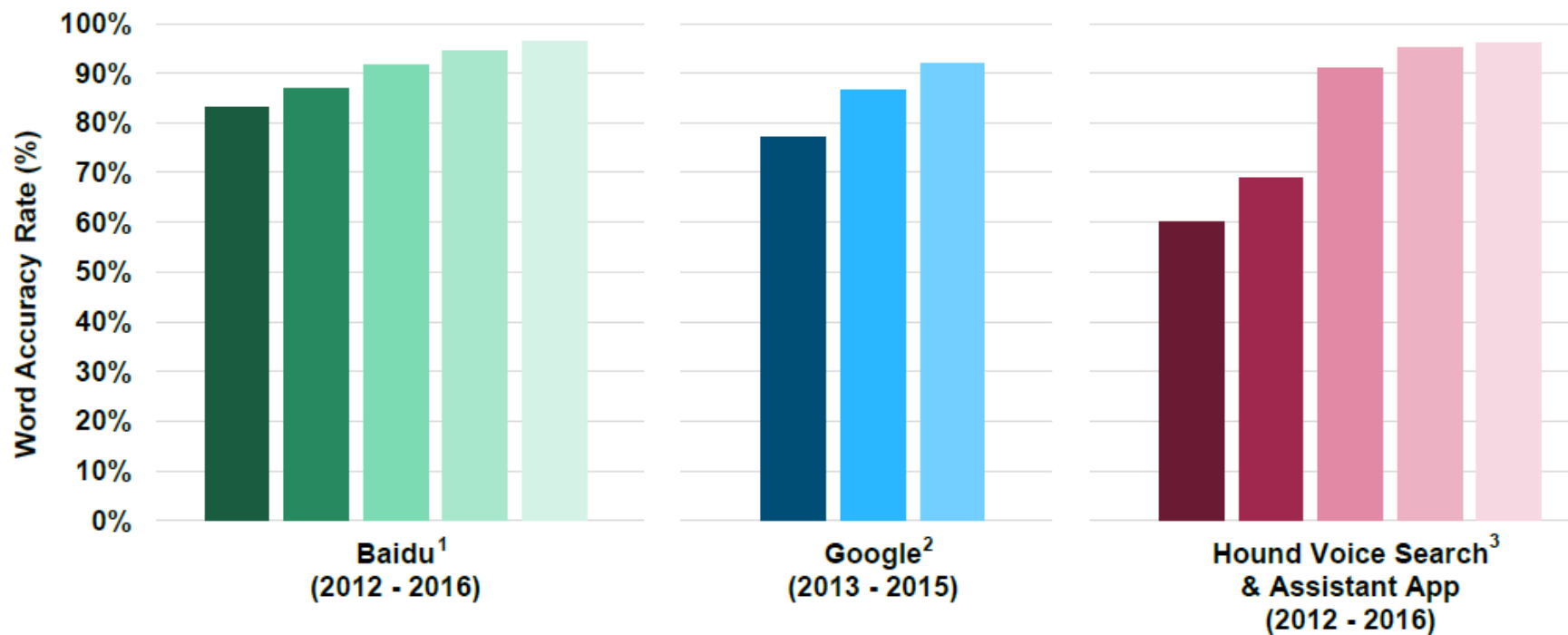


Source: KPCB – Internet Trends Report 2016 <http://www.kpcb.com/internet-trends>

Voice Word Accuracy Rates Improving Rapidly... +90% Accuracy for Major Platforms

Word Accuracy Rates by Platform*, 2012 – 2016

**Word accuracy rate definitions are unique to each company...see footnotes for more details*

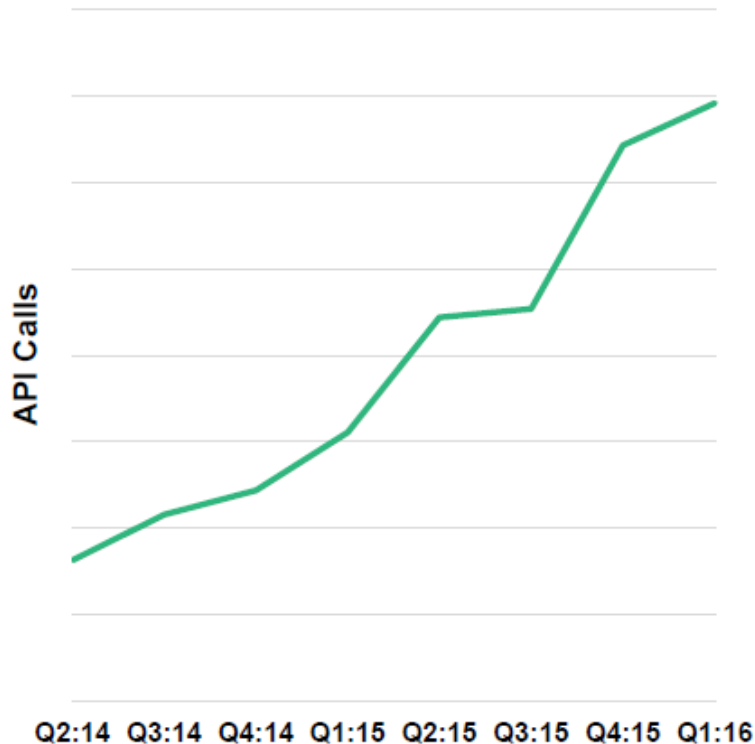


Source: KPCB – Internet Trends Report 2016 <http://www.kpcb.com/internet-trends>

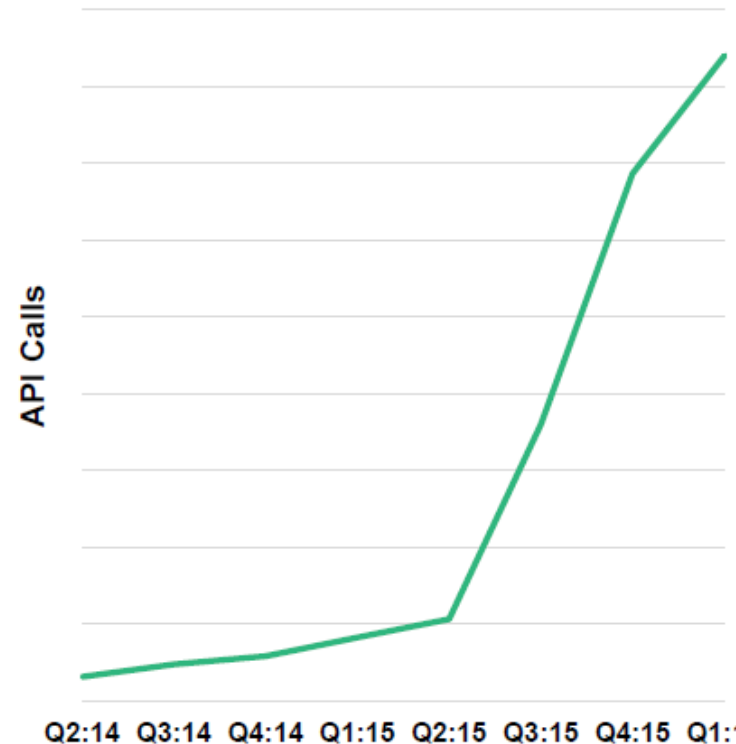
Baidu Voice = Input Growth >4x...Output >26x, Since Q2:14

Usage across all Baidu products growing rapidly...typing Chinese on small cellphone keyboard even more difficult than typing English...Text-to-Speech supplements speech recognition & key component of man-machine communications using voice

**Baidu Speech Recognition Daily Usage by API Calls,
Global, 2014 – 2016¹**



**Baidu Text to Speech (TTS) Daily Usage by API Calls,
Global, 2014 – 2016²**



Source: KPCB – Internet Trends Report 2016 <http://www.kpcb.com/internet-trends>

Dezember 2015						
MO	DI	MI	DO	FR	SA	SO
49	30	1	2	3	4	5 6
50	7	8	9	10	11	12 13
51	14	15	16	17	18	19 20
52	21	22	23	24	25	26 27
53	28	29	30	31		

Speech Recognition of English and Mandarin

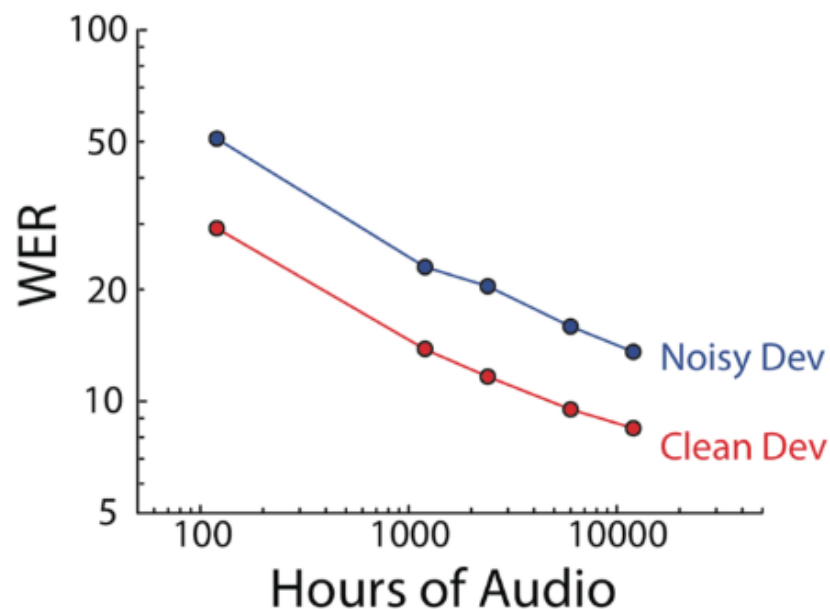
Deep Speech 2, is especially significant in how it relies entirely on machine learning for translation. Whereas older voice-recognition systems include many handcrafted components to aid audio processing and transcription, the Baidu system learned to recognize words from scratch, simply by listening to thousands of hours of transcribed audio.

Baidu app for smartphones lets users search by voice, and also includes a voice-controlled personal assistant called Duer. Voice queries are more popular in China because it is more time-consuming to input text, and because some people do not know how to use Pinyin, the phonetic system for transcribing Mandarin using Latin characters.



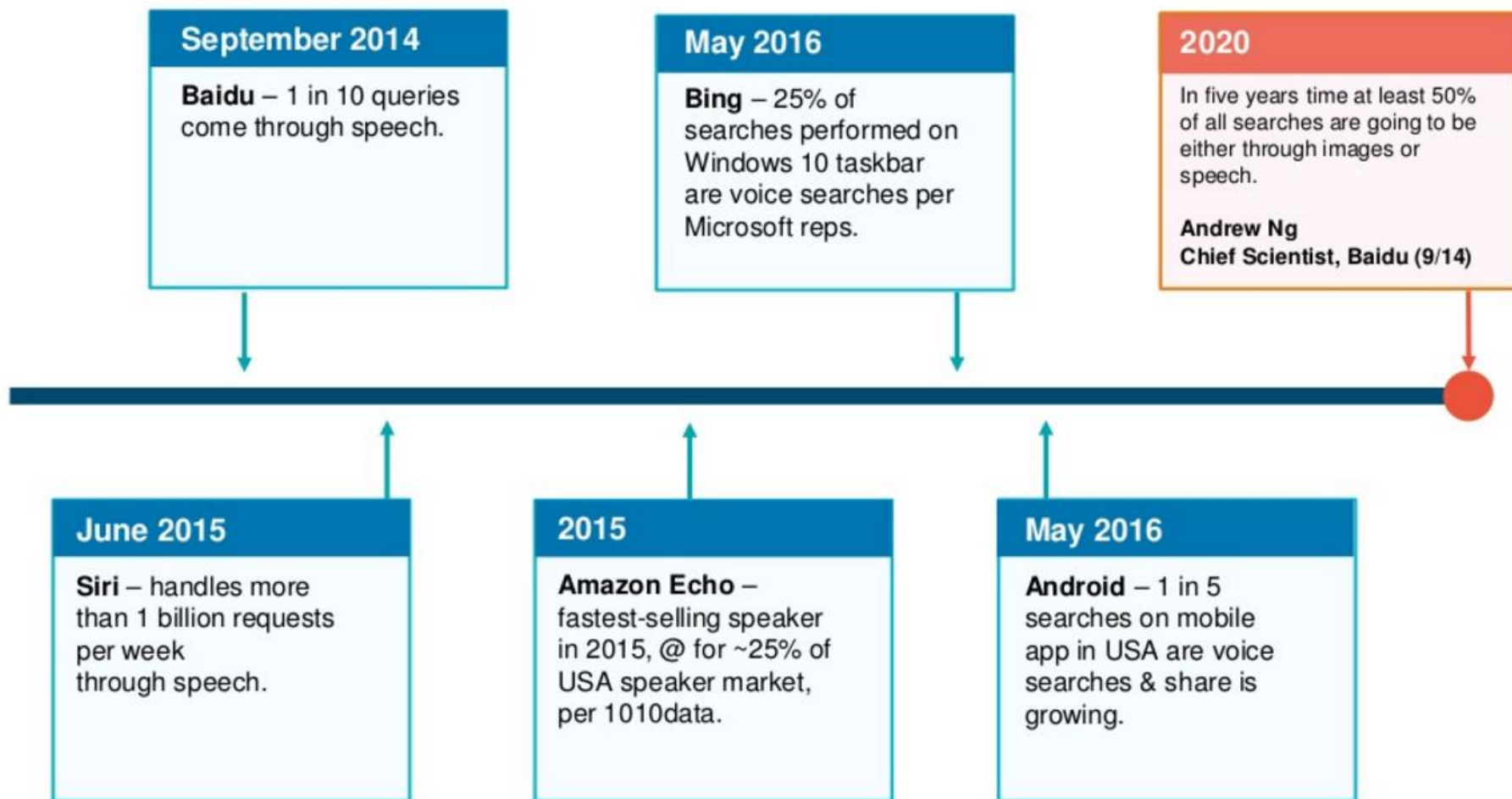
Baidu Speech/Voice Recognition

- Baidu's research in speech recognition shows that using 10x more data can lower relative error rates by 40 percent. Figuring out how to train models on ever larger datasets is therefore a key part of progress in AI.
- As datasets grow, high performance computing (HPC) becomes more important in order to train the models.
- Today, training one of Baidu's speech recognition models consumes 20 billion billion math operations (**20 exaflops**), and that number continues to increase.
- There are many areas, including *precision agriculture*, *consumer finance* and *medicine*, where Baidu sees a clear opportunity for AI to have a big impact.

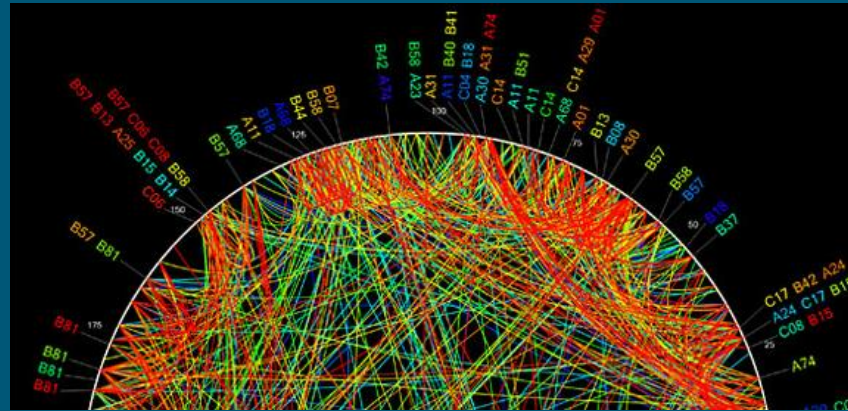


Baidu's speech recognition research has shown big reductions in word error rate (WER) using 10x more data.

Voice = Gaining Search Share... USA Android @ 20%...Baidu @ 10%...Bing Taskbar @ 25%



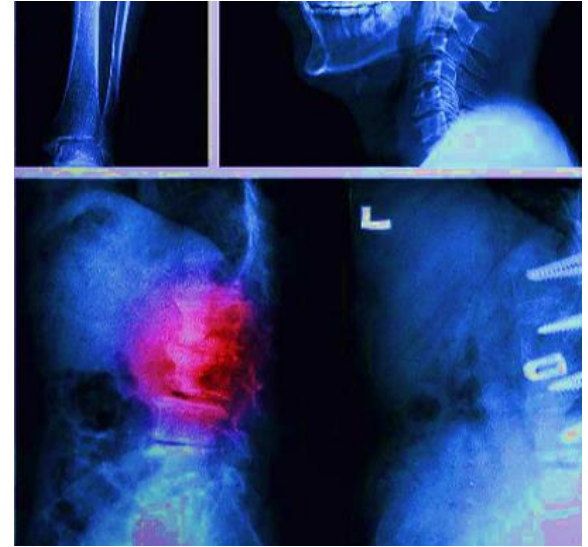
Source: KPCB – Internet Trends Report 2016 <http://www.kpcb.com/internet-trends>



Applications Medicine

Deep Learning and Medical Imaging Data (1/2)

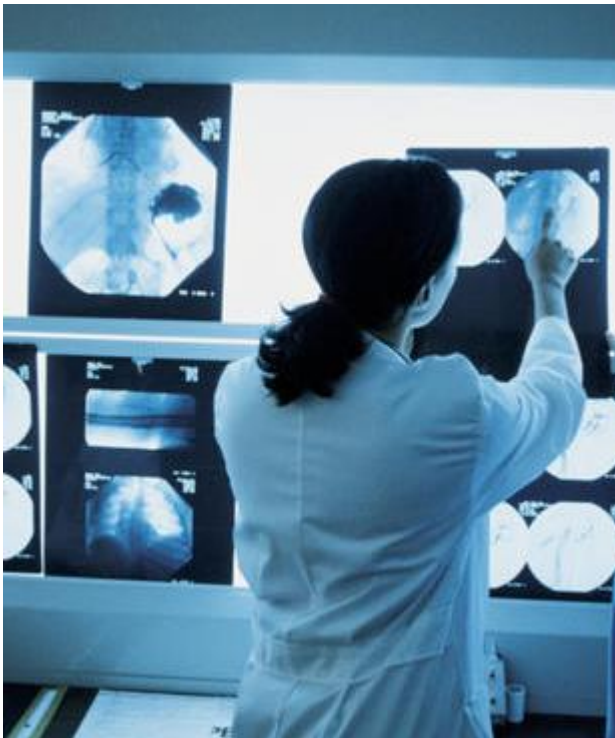
- Enlitic has developed deep learning networks that analyze medical imaging data such as X-rays and MRIs.
- Their networks increase diagnostic accuracy in less time and at reduced cost when compared to traditional diagnostic methods.
- Enlitic's software also allows **comparison of an individual patient's radiological data with millions of other patients who received the same diagnosis** in order to identify and track treatment outcomes for the most similar cases.
- Deep learning can readily handle a broad spectrum of diseases in the entire body, and all imaging modalities (x-rays, CT scans, etc.).



- Contextualizes the imaging data by comparing it to large datasets of past images, and by analyzing ancillary clinical data, including clinical reports and laboratory studies.
- Enlitic is using deep learning to **automatically screen for hundreds of specific diseases within your medical imaging data**, while being cheaper, faster, and more accurate than the manual methods used today

Deep Learning and Medical Imaging Data (2/2)

- The deep learning algorithm can **increase the accuracy of a radiologist's interpretation by 50-70%** and is **50,000 times faster**.
- Enlitic is trying to market their technology as something that “enables” radiologists as opposed to replacing them.




While that may make their value proposition less threatening, let's face facts here. The average radiologist's salary is \$286,000 a year. In the U.S. there is one radiologist per 10,000 people which points to an estimated number of 31,800 radiologists in total. That means that around \$9 billion dollars a year is being spent on radiologists. You can try to dance around the facts as much as you like, but the truth is, these are exactly the types of service jobs that deep learning is predicted to replace. That \$9 billion is where Enlitic plans to make their money based on a business model that makes money by sharing a cut of the cost savings realized. **It's hard to recommend radiology as a career option when you look at this scenario.**

Enlitic's Technology in Action (1/2)


In initial benchmarking test against the publicly-available LIDC dataset, Enlitic technology detected lung cancer nodules in chest CT images 50% more accurately than an expert panel of radiologists



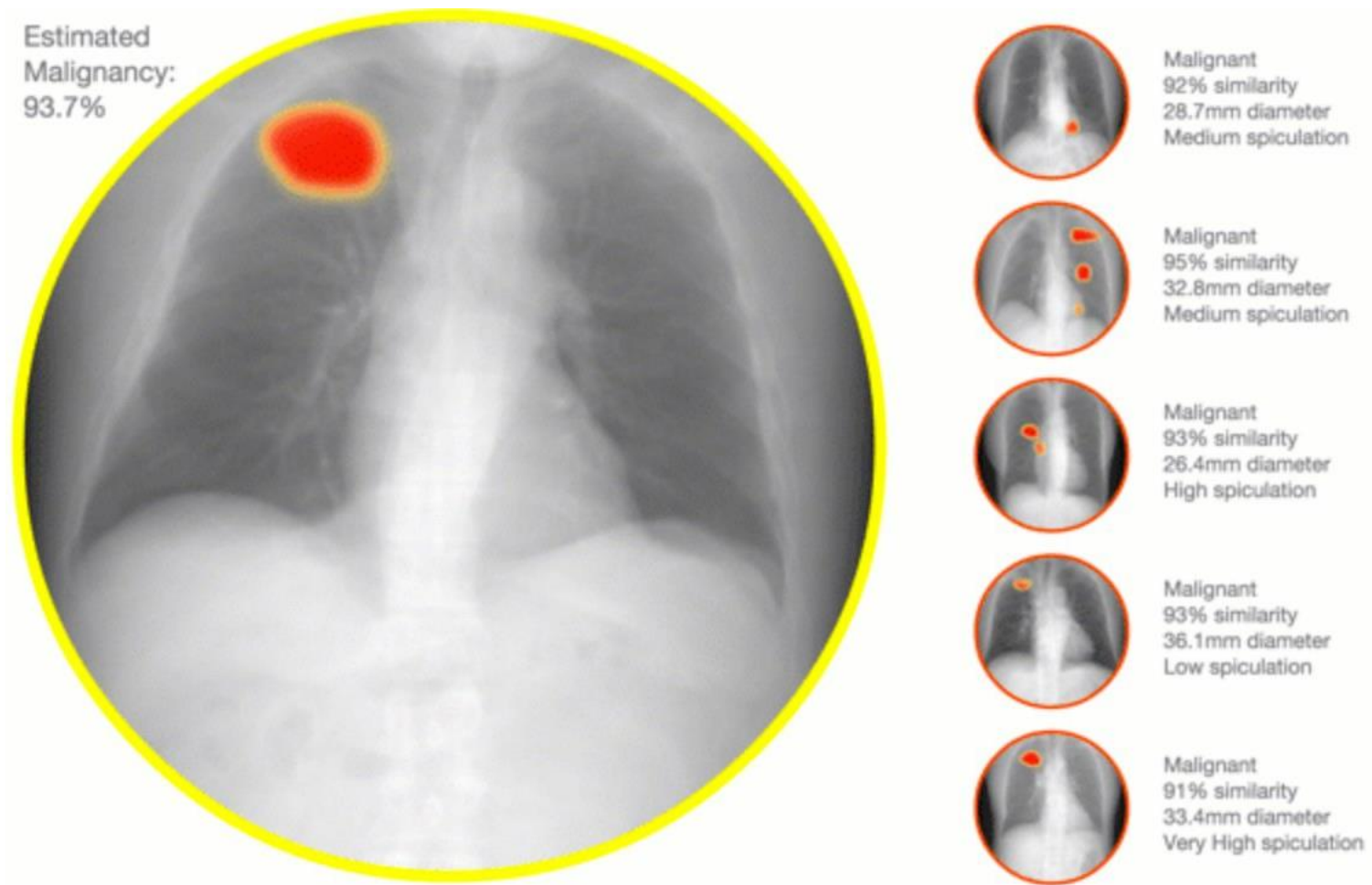
In initial benchmarking tests, Enlitic's deep learning tool regularly detected tiny fractures as small as 0.01% of the total x-ray image



Enlitic's deep learning tool is designed to simultaneously support many diseases



Enlitic's Technology in Action (1/2)



Screen shot of a scan showing a tumor in a patient's lungs, with the other smaller shots of tumors to the right that are were examples from which the computer learned to detect the growth.

Jeremy Howard, CEO of Enlitic

... in **medicine**, ... applications include **radiology** (looking at the inside of the body using X-rays or MRIs), **pathology** (looking at tissue through a microscope) and **dermatology** (looking at pictures of skin).

It takes literally decades for humans to see enough examples of things so that they can accurately pick up what's going on, for example, in a MRI.

Computers, on the other hand, can look at 50 million MRIs and understand every kind of disease as it appears in every kind of person at every stage of time, and **can, therefore, be as good as the best radiologist in every single sub-specialty.**

They can allow the physician, or even a nurse in some remote province of China, to deeply understand what's going on in a medical image.

Healthcare, in general, is an almost \$10 trillion industry—that's \$3 trillion in the US—possibly the largest industry in the world.

The fact that we can now use deep learning to understand medical data in this way is going to be totally world changing.





Google's DeepMind

DeepMind and Demis Hassabis

- Founded in London in 2010; Demis Hassabis was one of the co-founders
- Backed by some of the most successful technology entrepreneurs in the world.
- Acquired by Google in 2014, now part of the Alphabet group.
- Continue to be based in London, alongside some of the country's leading academic, cultural and scientific organizations in the King's Cross Knowledge Quarter.



His work systematically connecting imagination with episodic memory for the first time was included in the top 10 scientific breakthroughs of 2007 in the annual list compiled by Science.

	Demis Hassabis
Born	27 July 1976 (age 40) London, England
Nationality	British
Fields	Machine learning Neuroscience ^[1]
Institutions	DeepMind Technologies Bullfrog Productions Lionhead Studios University College London Computer Laboratory, University of Cambridge
Education	Christ's College, London (state comprehensive school)
Alma mater	University of Cambridge (BA) University College London (PhD)
Notable awards	FRSA Pentamind World Champion



AlphaGo Beats Top Ranked Go Grandmaster 4–1

“...this test bodes well for AI’s potential in solving other problems. AlphaGo has the ability to look “globally” across a board—and find solutions that humans either have been trained not to play or would not consider. This has huge potential for using AlphaGo-like technology to find solutions that humans don’t necessarily see in other areas ...”

Demis Hassabis, CEO and Co-Founder of DeepMind



<https://googleblog.blogspot.de/2016/03/what-we-learned-in-seoul-with-alphago.html/>



AlphaGo Beats Top Ranked Go Grandmaster 4–1



12.Mar.2016

Google's AI Takes Historic Match Against Go Champ With Third Straight Win

13.Mar.2016


Go Grandmaster Lee Sedol Grabs Consolation Win Against Google's AI

15.Mar.2016






Google's AI Wins Fifth And Final Game Against Go Genius Lee Sedol

Google's AlphaGo gets 'divine' Go ranking

AlphaGo was given an honorary "ninth dan" professional ranking, equivalent to that held by Lee Sedol who has 18 international titles to his name and is widely considered one of the greatest Go players of the modern era.

 **Google DeepMind**
Challenge Match
8 - 15 March 2016

FINAL SCORES

Match	Black	White	Result
1	Lee Sedol	 AlphaGo	W + Res
2	 AlphaGo	Lee Sedol	B + Res
3	Lee Sedol	 AlphaGo	W + Res
4	 AlphaGo	Lee Sedol	W + Res
5	Lee Sedol	 AlphaGo	W + Res

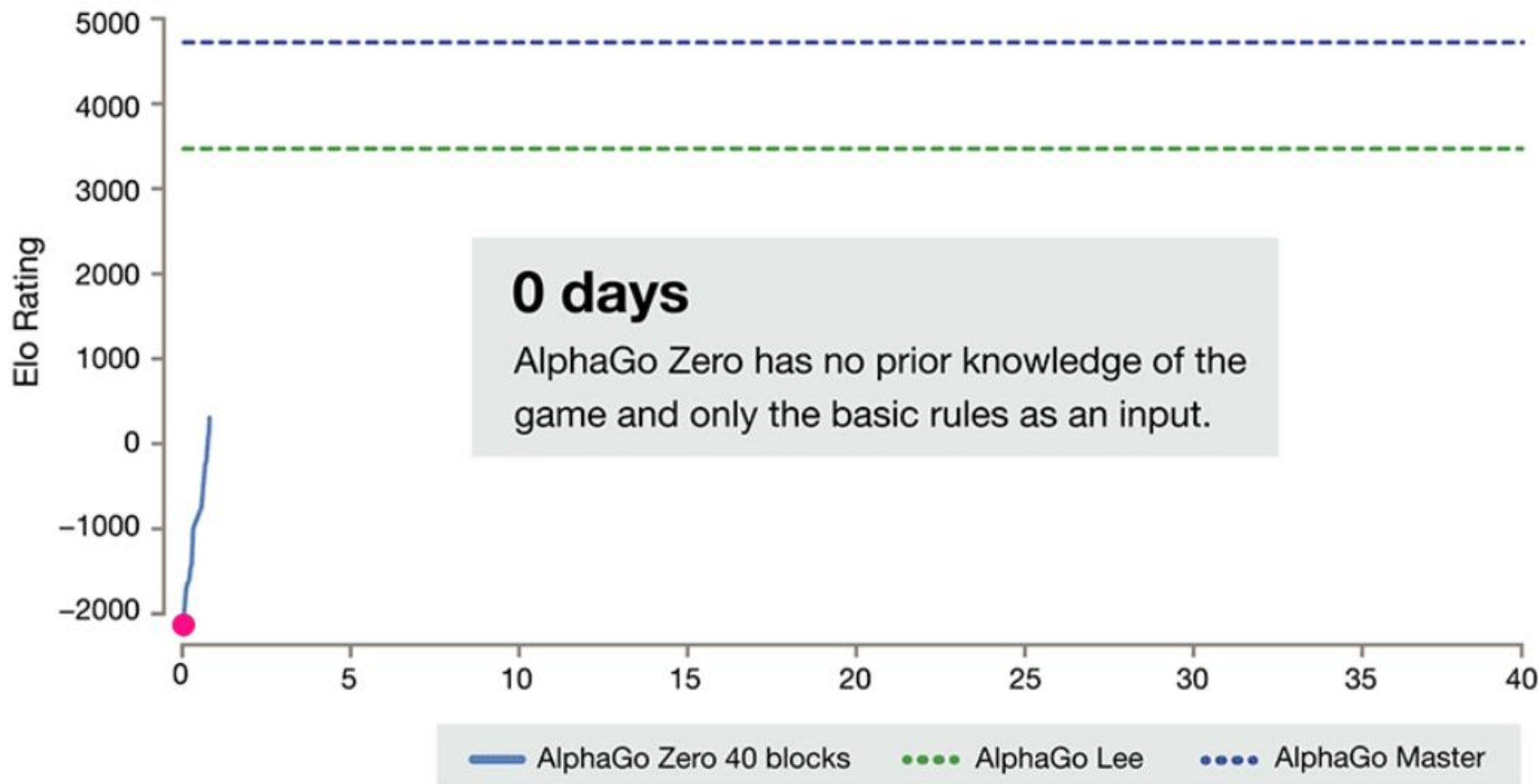
AlphaGo Retires from Competitive Go After Defeating World Number One 3-0

Mai 2017						
MO	DI	MI	DO	FR	SA	SO
17	24	25	26	27	28	29 30
18	1	2	3	4	5	6 7
19	8	9	10	11	12	13 14
20	15	16	17	18	19	20 21
21	22	23	24	25	26	27 28
22	29	30	31			



<https://www.theverge.com/2017/5/27/15704088/alphago-ke-jie-game-3-result-retires-future>
<https://techcrunch.com/2017/05/24/alphago-beats-planets-best-human-go-player-ke-jie/>

AlphaGo Zero



<https://deepmind.com/blog/alphago-zero-learning-scratch/>

AlphaGo Zero



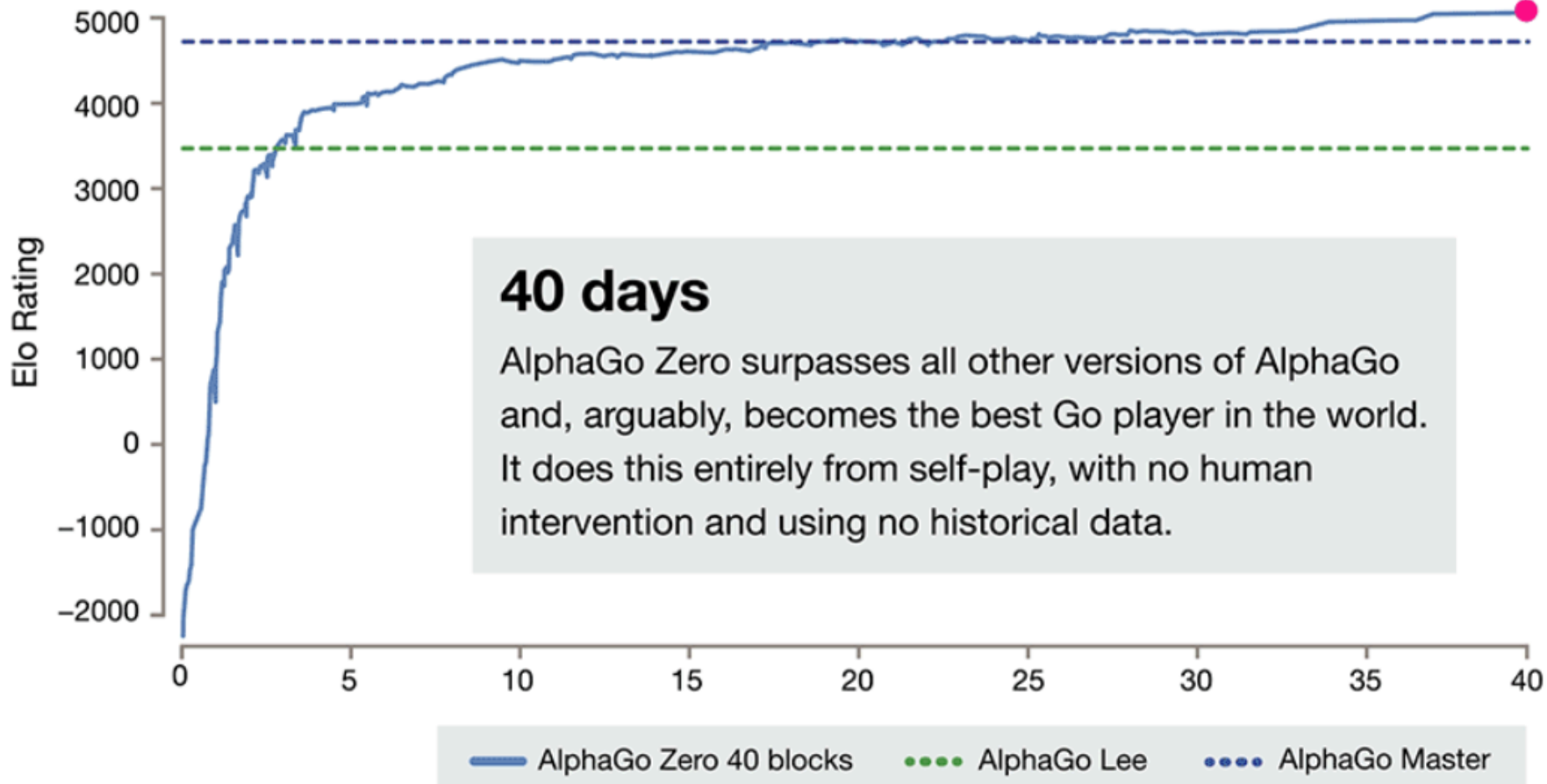
<https://deepmind.com/blog/alphago-zero-learning-scratch/>

AlphaGo Zero



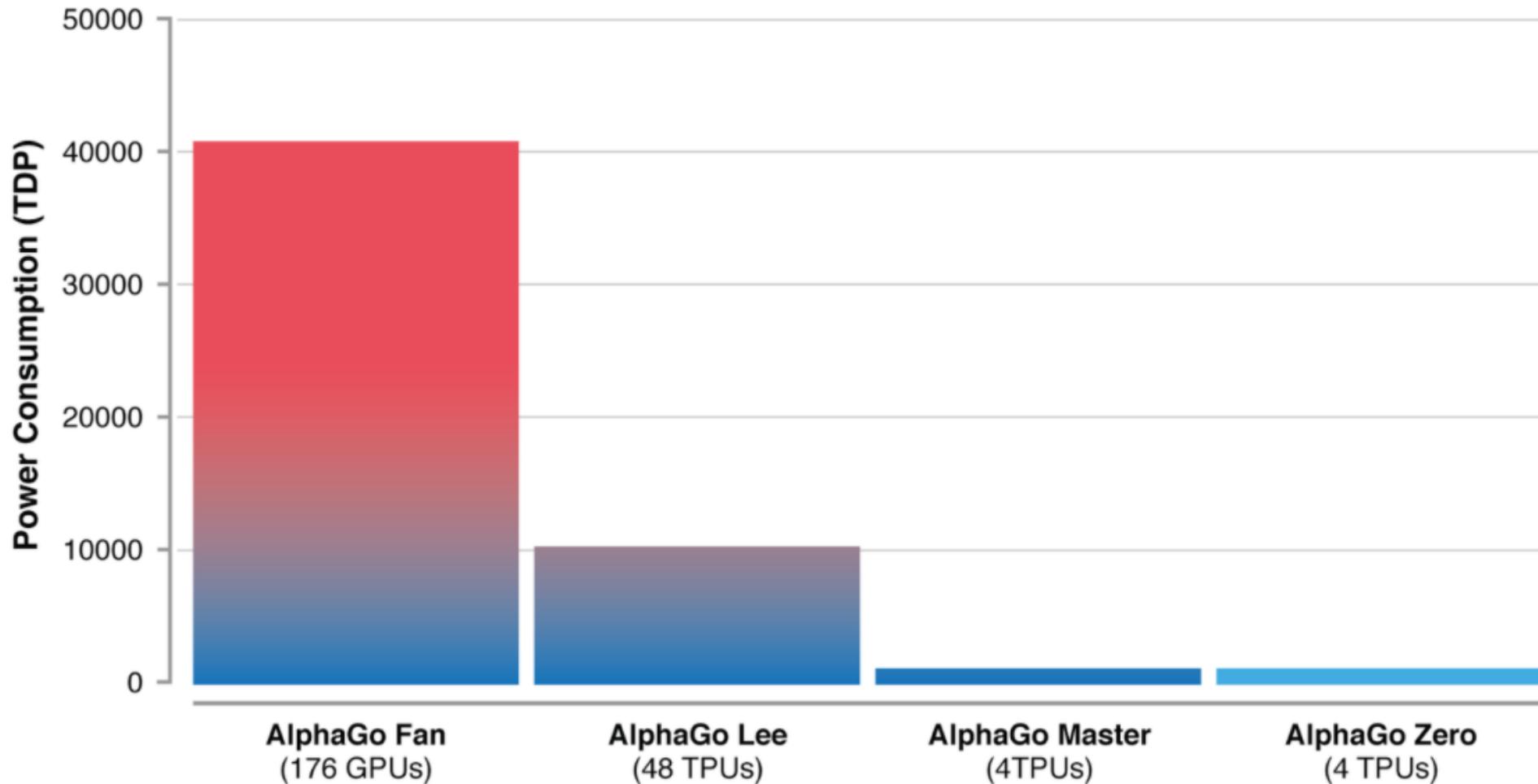
<https://deepmind.com/blog/alphago-zero-learning-scratch/>

AlphaGo Zero



<https://deepmind.com/blog/alphago-zero-learning-scratch/>

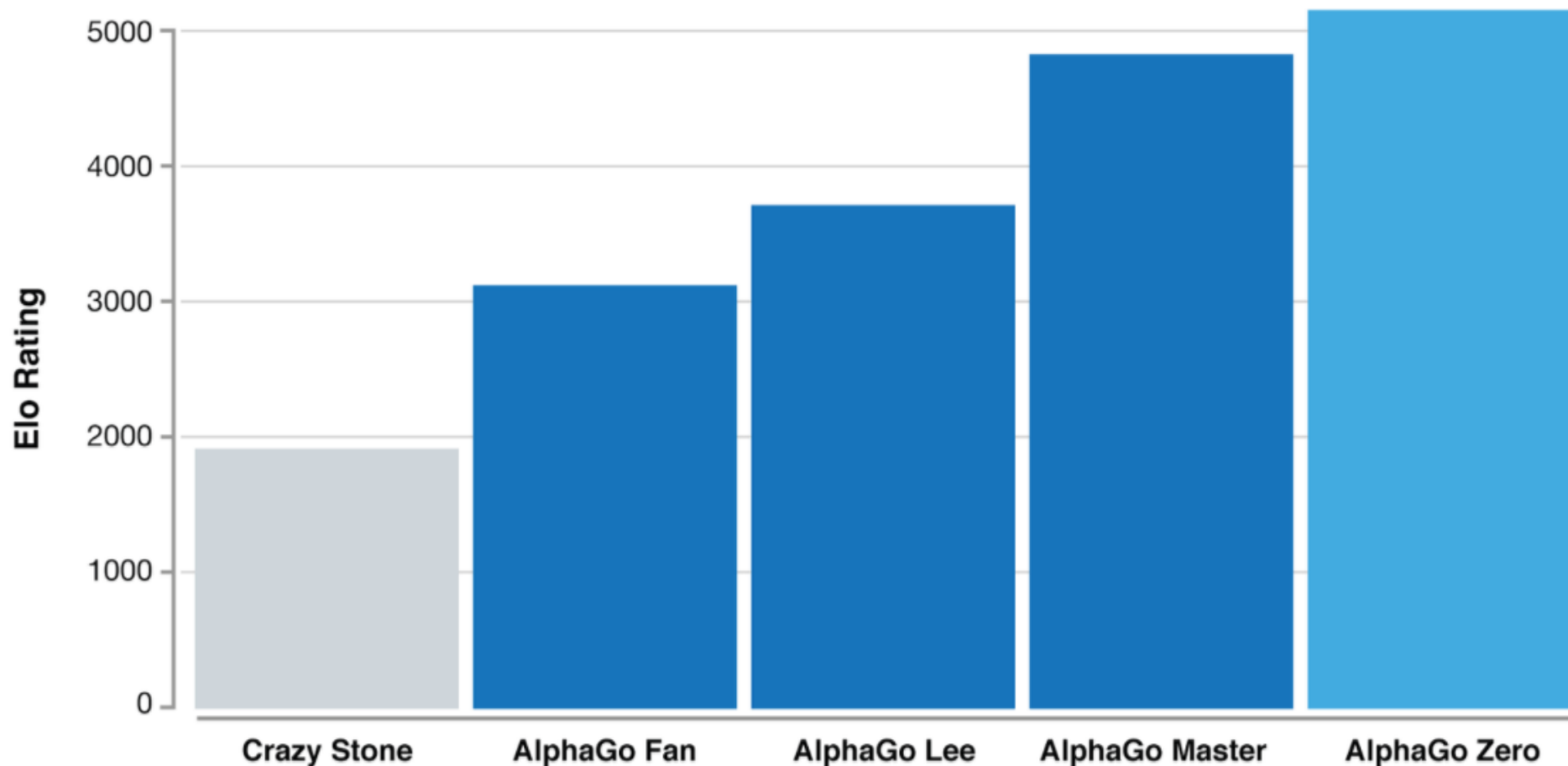
AlphaGo Zero



AlphaGo has become progressively more efficient thanks to hardware gains and more recently algorithmic advances

<https://deepmind.com/blog/alphago-zero-learning-scratch/>

AlphaGo Zero



Elo ratings - a measure of the relative skill levels of players in competitive games such as Go - show how AlphaGo has become progressively stronger during its development

<https://deepmind.com/blog/alphago-zero-learning-scratch/>

AlphaGo Zero – Summary

1. Beat the previous version of AlphaGo (Final score: 100–0).
2. Learned to perform this task from scratch, without learning from previous human knowledge (i.e. recorded game play).
3. Reached world champion level Go playing in just 3 days of training.
4. Did so with an order of magnitude less computational resources (4 TPUs vs 48 TPUs).
5. Accomplished this with much less training data (3.9 million games vs 30 millions games).

DeepMind's "Watch, Listen, Attend, and Spell" Neural Network

DeepMind's Neural Net can now lip read better than humans after watching thousands of hours of TV

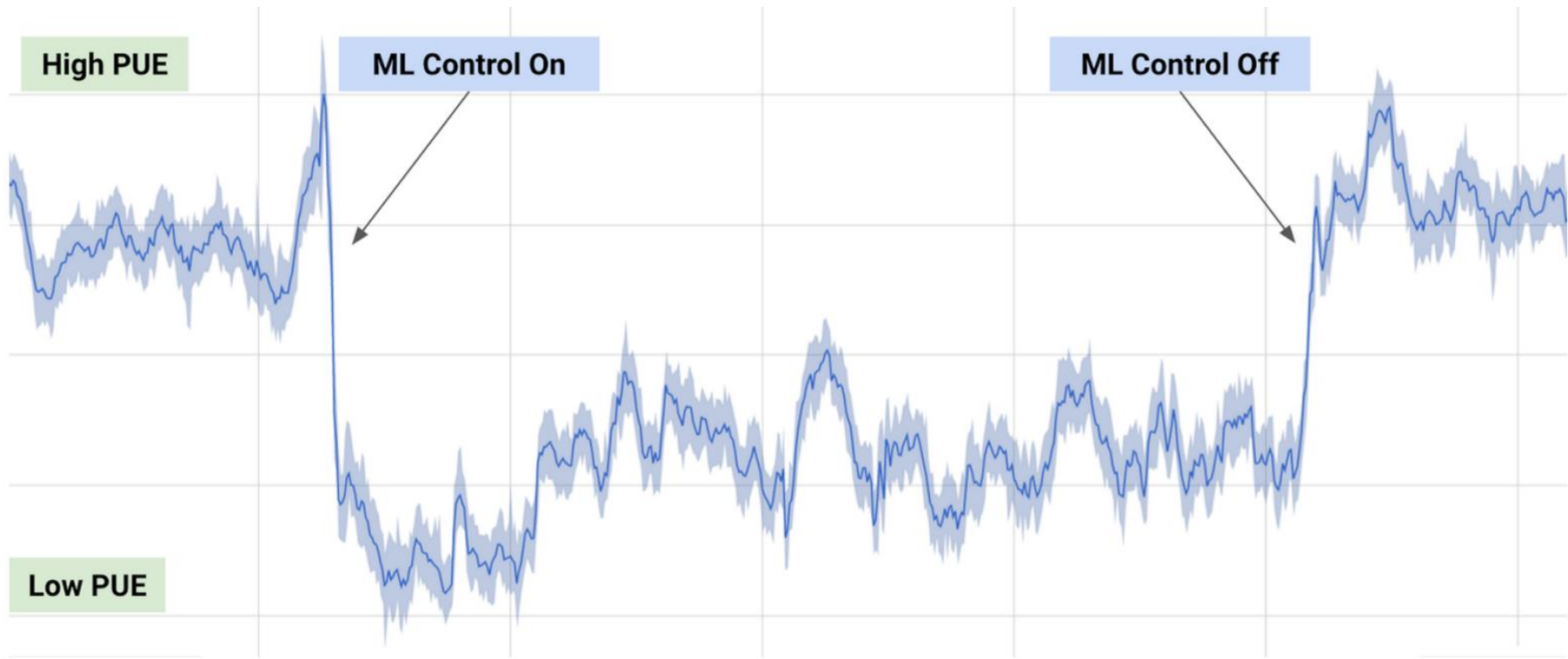


- Using more than 5,000 of hours of TV footage from the BBC, DeepMind scientists trained a **neural network to annotate video footage with 46.8 percent accuracy**.
- Videos included 118,000 difference sentences and some 17,500 unique words
- Tested on the same footage, a **professional human lip-reader was only able to get the right word 12.4 percent** of the time.
- Potential applications:
 - Help hearing-impaired people understand conversations.
 - Annotate silent films
 - Control digital assistants like Siri or Alexa by just mouthing words to a camera

Application: Optimizing Energy Consumption in Data Centers (1/3)

- By applying DeepMind's machine learning to our own Google data centers, we've managed to reduce the amount of energy we use for cooling by up to 40 percent.
- In any large scale energy-consuming environment, this would be a huge improvement. Given how sophisticated Google's data centers are already, it's a phenomenal step forward.
- Using a system of neural networks trained on different operating scenarios and parameters within our data centers, we created a more efficient and adaptive framework to understand data center dynamics and optimize efficiency.
- We accomplished this by taking the historical data that had already been collected by thousands of sensors within the data center -- data such as temperatures, power, pump speeds, setpoints, etc. -- and using it to train an ensemble of deep neural networks.
- We **trained the neural networks on the average future PUE** (Power Usage Effectiveness), which is defined as the ratio of the total building energy usage to the IT energy usage. We then trained **two additional ensembles of deep neural networks to predict the future temperature and pressure of the data center over the next hour**. The purpose of these predictions is to simulate the recommended actions from the PUE model, to ensure that we do not go beyond any operating constraints.

Optimizing Energy Consumption in Data Centers (2/3)



Optimizing Energy Consumption in Data Centers (3/3)

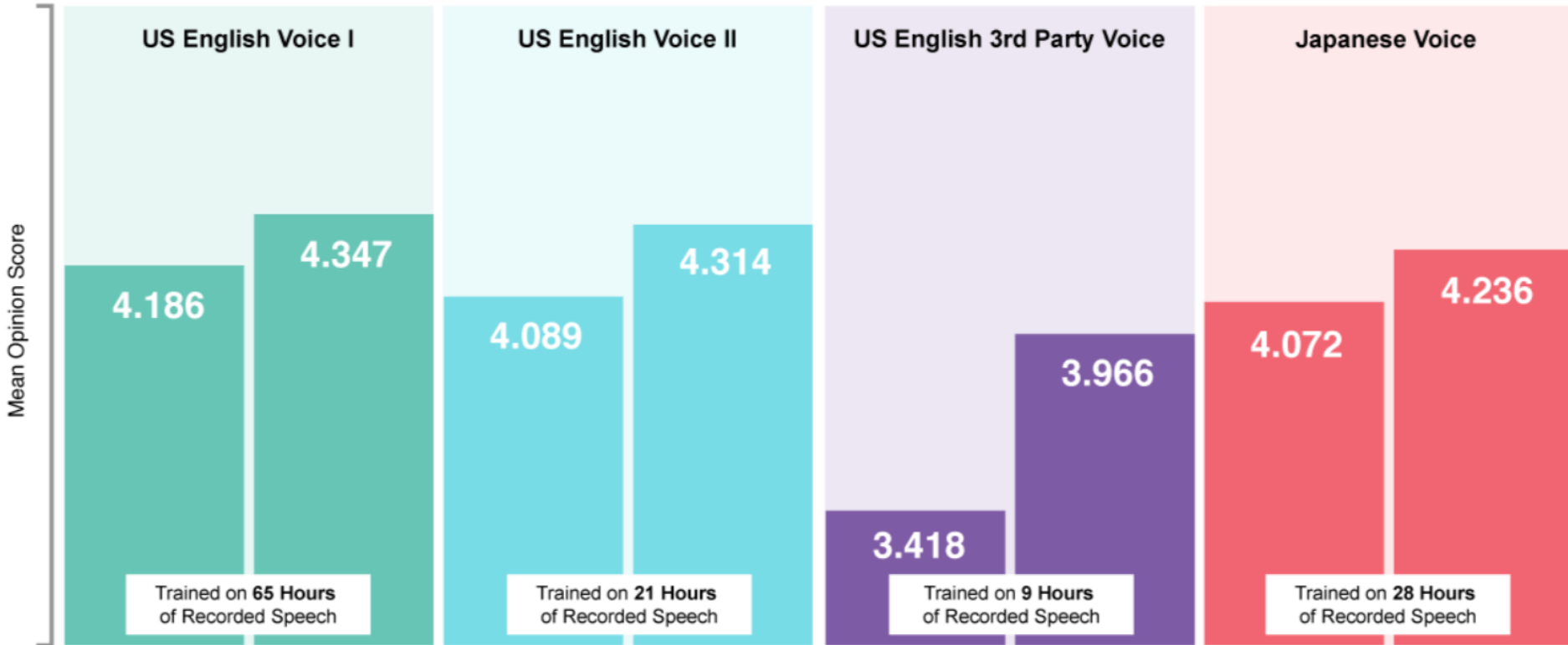
- DeepMind's machine learning system was able to **consistently achieve a 40 percent reduction in the amount of energy used for cooling.**
- This **equates to a 15 percent reduction in overall PUE overhead** after accounting for electrical losses and other non-cooling inefficiencies.
- It also produced the lowest PUE the site had ever seen.
- Possible applications include
 - improving power plant conversion efficiency (getting more energy from the same unit of input)
 - reducing semiconductor manufacturing energy and water usage
 - helping manufacturing facilities increase throughput.
- DeepMind are planning to roll out this system more broadly and will share how they did it in an upcoming publication, so that other data center and industrial system operators – and ultimately the environment – can benefit from this major step forward.
- DeepMind is in discussions with the UK government about optimizing energy distribution within the UK energy grid!

WaveNet

- In the Fall of 2016, DeepMind introduced a prototype version of WaveNet, a convolutional neural network, for generating English and Japanese speech
- In October 2017, the production version was described in a blog post
- Main features:
 - WaveNet is being used to generate the Google Assistant voices for US English and Japanese across all platforms
 - Using the new WaveNet model results in a range of more natural sounding voices for the Assistant
 - WaveNet creates individual waveforms from scratch, one sample at a time, with 16,000 samples per second and seamless transitions between individual sounds
 - The new, improved WaveNet model still generates a raw waveform but at speeds **1,000 times faster** than the original model, meaning it requires just 50 milliseconds to create one second of speech
 - The model is also higher-fidelity, capable of creating waveforms with **24,000 samples a second**
 - The resolution of each sample has been increased from 8 bits to **16 bits**, the same resolution used in compact discs
 - The new US English voice I gets a mean-opinion-score (MOS) of 4.347 on a scale of 1-5, where even human speech is rated at just 4.667
 - It runs on Google's latest TPU (Tensor Processing Unit) cloud infrastructure

WaveNet – Performance

Mean Opinion Scores



The blog post contains audio samples of the three voices. The contrast between the old “concatenative Text-To-Speech synthesis” and WaveNet is striking.

<https://deepmind.com/blog/wavenet-launches-google-assistant/>



The World You & I Live In Today (Part II)

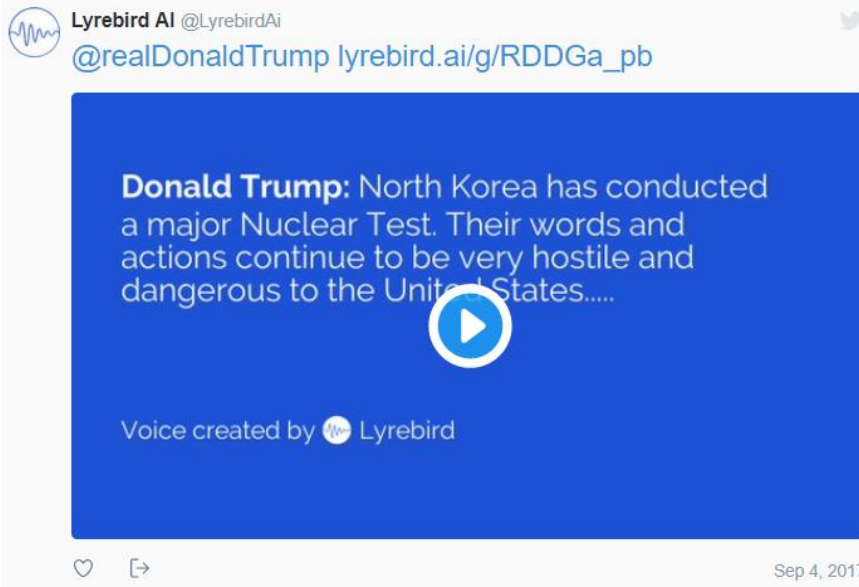
TWYILIT – Twilight!

The future is already here – it is just not very evenly distributed.
William Gibson

Welcome to the beta version of Lyrebird

Lyrebird allows you to create a digital voice that sounds like you with only one minute of audio.

Create your digital voice




Lyrebird AI @LyrebirdAi
@realDonaldTrump lyrebird.ai/g/RDDGa_pb

Donald Trump: North Korea has conducted a major Nuclear Test. Their words and actions continue to be very hostile and dangerous to the United States.....

Voice created by Lyrebird

Sep 4, 2017



Lyrebird AI @LyrebirdAi
@BarackObama lyrebird.ai/g/Wbt3UHVr

Barack Obama: Michelle and I are thinking of the victims and their families in Barcelona. Americans will always stand with our Spanish friends. Un abrazo.

Voice created by Lyrebird

Sep 4, 2017

- Will provide an open API accessible to anyone

<https://techcrunch.com/2017/04/25/lyrebird-is-a-voice-mimic-for-the-fake-news-era/>
<https://lyrebird.ai/demo>

Alternative Face v1.1

Françoise Hardy Imitates Kellyanne Conway

https://www.youtube.com/watch?v=af_9LXhcebY

	MO	DI	MI	DO	FR	SA	SO
5			1	2	3	4	5
6	6	7	8	9	10	11	12
7	13	14	15	16	17	18	19
8	20	21	22	23	24	25	26
9	27	28	1	2	3	4	5
10	6	7	8	9	10	11	12



Françoise Hardy is a French singer and songwriter. Kellyanne Conway is the current Counselor to U.S. president Donald Trump. The audio is from a famous interview with Ms. Conway on 22 January 2017, in which she mentions “alternative facts”.

J.S. Bach Chorale: Wer nur den lieben Gott lässt walten



Dez 2016

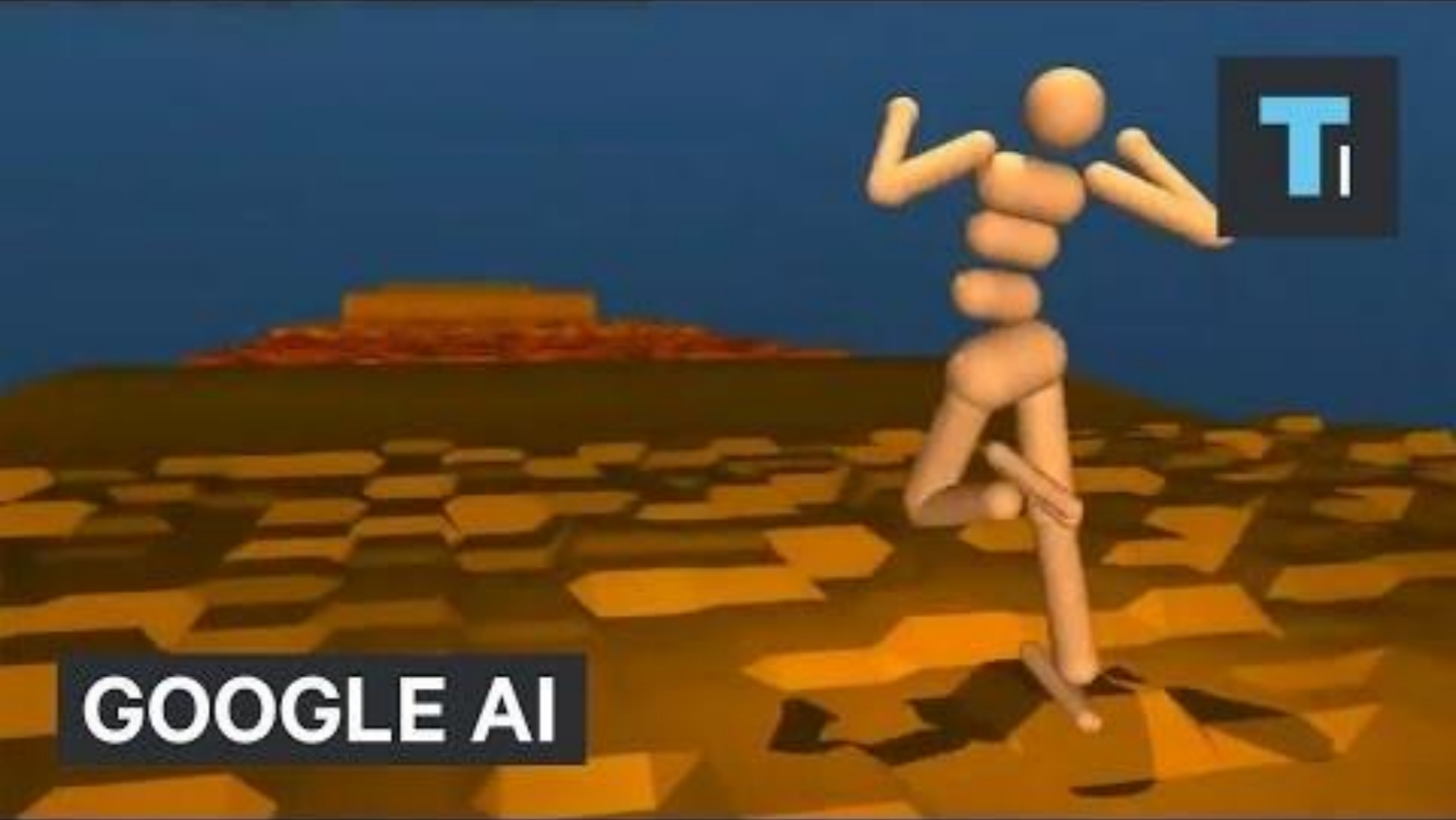
	MO	DI	MI	DO	FR	SA	SO
48				1	2	3	4
49	5	6	7	8	9	10	11
50	12	13	14	15	16	17	18
51	19	20	21	22	23	24	25
52	26	27	28	29	30	31	1
1	2	3	4	5	6	7	8



<https://youtu.be/QiBM7-5hA6o>

“Wer nur den lieben Gott lässt walten” by Georg Neumark, 1641

Google's DeepMind AI just taught itself to walk



GOOGLE AI

<https://www.youtube.com/watch?v=gn4nRCC9TwQ>

Drone Autonomously Avoiding Obstacles at 30 MPH



https://www.youtube.com/watch?v=_qah8olzCwk

Boston Dynamics Atlas – Humanoid Robot



<https://www.youtube.com/watch?v=rVlhMGQgDkY>

Boston Dynamics Robots – Handle



Boston Dynamics Robots – Spot Mini





<https://www.youtube.com/watch?v=fRj34o4hN4I>

Boston Dynamics Robots – Happy Holidays



https://www.youtube.com/watch?v=RDZu04v7_hc



References

References – General

Ars Technica – www.arstechnica.com

Big Data News – www.bigdatanews.com

Data Science Central – www.datasciencecentral.com

Exponential Finance

- On YouTube, search for “exponential finance 2014” – you will find 37 fascinating videos on emerging and disruptive technologies, the concept of exponential organizations, and numerous panel discussions.

KDnuggets – www.kdnuggets.com

KPCB (Kleiner, Perkins, Caufield, and Byers) – www.kpcb.com

- Internet Trends Report 2013 / 2014 / 2015 / 2016

MIT’s Technology Review – www.technologyreview.com

Next BIG Future – www.nextbigfuture.com

O’Reilly Radar – www.radar.oreilly.com

Singularity Hub – www.singularityhub.com

Slideshare – www.slideshare.net

Techcrunch – www.techcrunch.com

References – Newsletters

Ars Technica – civis@arstechnica.com

Azeem Azhar: The Exponential View – azeem@azhar.co.uk

KurzweilAI.Net – newsletter@kurzweilai.net

Peter Diamandis – peter@diamandis.com

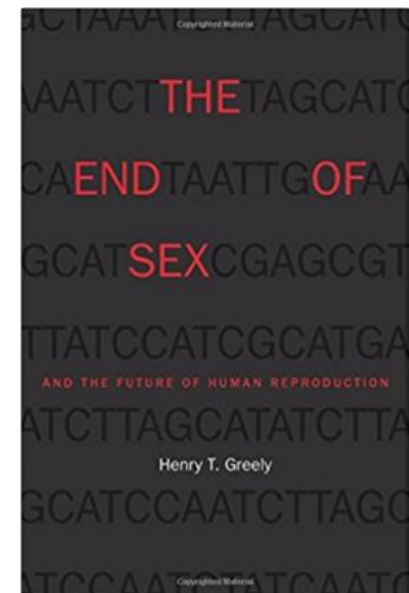
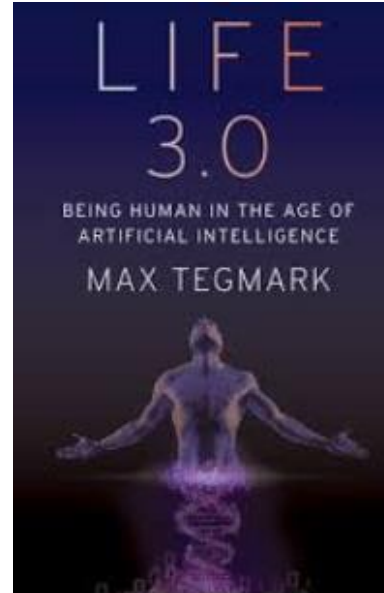
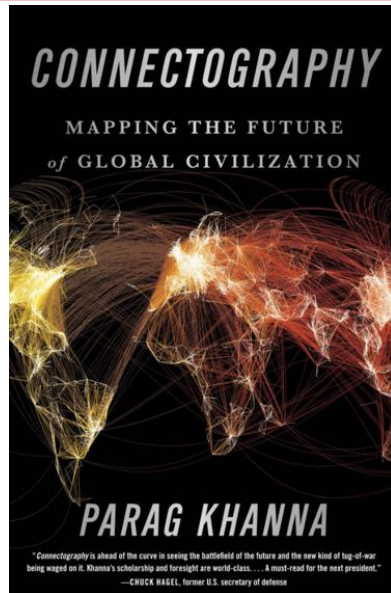
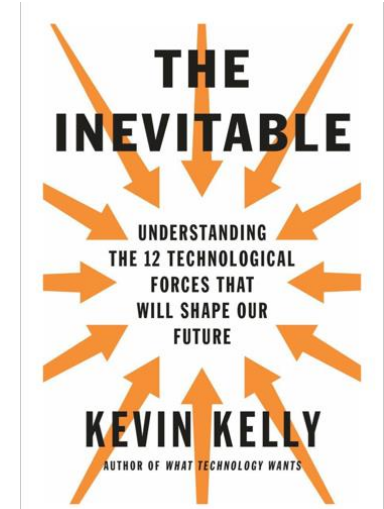
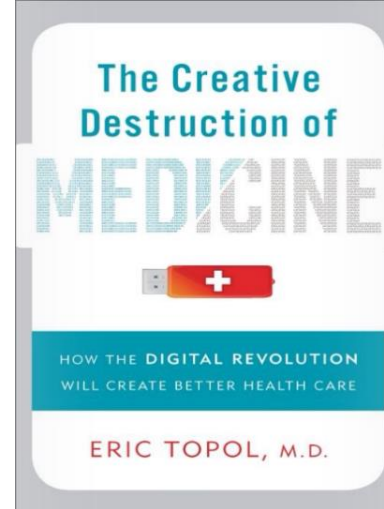
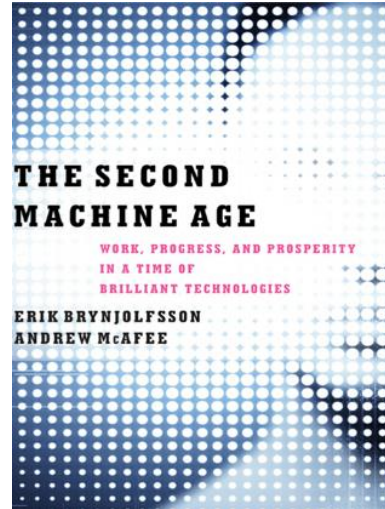
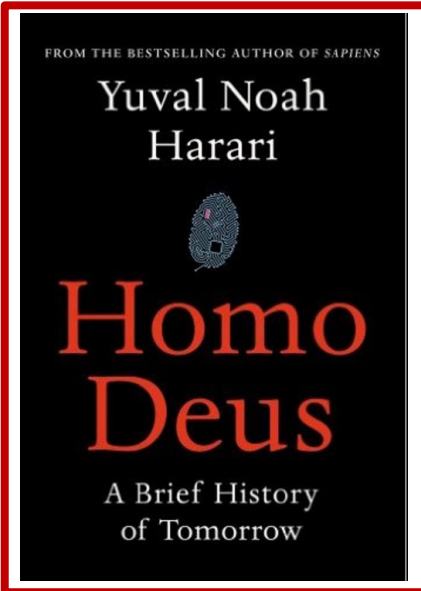
Quartz Daily Brief – hi@qz.com

Transmission – oliver@transmission.ai

WIRED Awake – newsletter@wired.co.uk

Interesting Books

Also available in German





Appendix

YouTube Video: Deep Learning for Computer Vision

Deep Learning for Computer Vision

Andrej Karpathy

10:15 – 11:45 am, Sep 24, 2016

Bay Area Deep Learning School 2016

▶ ▶| 🔊 1:27:36 / 10:33:45

Length: ~85 minutes

<https://www.youtube.com/watch?v=eyovmAt0Ux0>

YouTube Video: Bay Area Deep Learning School 2016 – Day 1







Streamed live on Sep 24, 2016

Day 1 of Bay Area Deep Learning School featuring speakers Hugo Larochelle, Andrej Karpathy, Richard Socher, Sherry Moore, Ruslan Salakhutdinov and Andrew Ng. Detailed schedule is at <http://www.bayareadlschool.org/schedule>

<https://www.youtube.com/watch?v=eyovmAtUx0>

MOOCs – Massively Open Online Courses

Overview	<h2>Machine Learning</h2> <p>About this course: Machine learning is the science of getting computers to act without being explicitly programmed. In the past decade, machine learning has given us self-driving cars, practical speech recognition, effective web search, and a vastly improved understanding of the human genome. Machine learning is so pervasive today that you probably use it dozens of times a day without knowing it. Many</p> <p>▼ More</p> <p>Created by: Stanford University</p> <div></div> <p>Taught by: Andrew Ng, Associate Professor, Stanford University; Chief Scientist, Baidu; Chairman and Co-founder, Coursera</p> <p></p>
Syllabus	
FAQs	
Creators	
Ratings and Reviews	
<h2>Machine Learning</h2> <p>Free trial not offered for this course.</p> <p>Enroll Now Starts Mar 20</p> <p>Financial Aid is available for learners who cannot afford the fee. Learn more and apply.</p>	
<p> Language</p>	English, Subtitles: Spanish, Hindi, Japanese, Chinese (Simplified)

<https://www.coursera.org/learn/machine-learning>

CS 231n – Stanford University

Convolutional Neural Networks for Visual Recognition



Course Instructors



Fei-Fei Li



Andrej Karpathy



Justin Johnson

- Lecture videos on YouTube
- Slides
- Notes
- Assignments

Fast.ai Course Part 1 – Highly Pragmatic



HOME

ABOUT

GETTING STARTED

LESSONS ▾

PART 2

CONTACT

BLOG

Welcome to fast.ai's 7 week course, **Practical Deep Learning For Coders, Part 1**, taught by Jeremy Howard (Kaggle's #1 competitor 2 years running, and founder of Enlitic). Learn how to build state of the art models without needing graduate-level math—but also without dumbing anything down. Oh and one other thing... it's totally free!

When you're done here, head over to part 2, **Cutting Edge Deep Learning for Coders**, to continue your learning.

"fast.ai... can actually get smart, motivated students to the point of being able to create industrial-grade ML deployments"



Harvard Business Review

The Business of Artificial Intelligence

<http://course.fast.ai/>

Fast.ai Course Part 2



PART 2

LESSONS ▾

CONTACT

BLOG

PART 1

Welcome to fast.ai's second 7 week course, **Cutting Edge Deep Learning For Coders, Part 2**, where you'll learn the latest developments in deep learning, how to read and implement new academic papers, and how to solve challenging end-to-end problems such as natural language translation.

"fast.ai... can actually get smart, motivated students to the point of being able to create industrial-grade ML deployments"



Harvard Business Review

The Business of Artificial Intelligence

This course assumes you are familiar with part 1, [Practical Deep Learning for Coders](#), so head over there if you haven't completed that course, or are not already familiar with current deep learning best practices. We will be assuming familiarity with everything from part 1, such as: CNNs (including resnets), RNNs (including LSTM and GRU), SGD/Adam/etc, batch normalization, data augmentation, Keras, and numpy. Like in part 1, there are around 20 hours of lessons, and you should plan to spend around **10 hours a week** for 7 weeks to complete the material. The course is based on lessons recorded at [The Data Institute at USF](#).



UNIVERSITY OF
SAN FRANCISCO

<http://course.fast.ai/part2.html>

TensorFlow for Poets

TensorFlow For Poets

<https://codelabs.developers.google.com/codelabs/tensorflow-for-poets/#0>



**ON THE ROAD TO SUCCESS,
THERE ARE NO SHORTCUTS.**

JOIN OUR TEAM
600-669-0322
shaffertrucking.com



**OUR MOST
VALUABLE
RESOURCE SITS
63 FEET AHEAD.**

1,916 RETWEETS 1,975 FAVORITES



Hungry...for challenges?

Let's have powerful start in the morning and chat about your future!
International consultant(s) talk about a day in life.



Interested? Just send an E-Mail to
Andreas.Range@accenture.com



Subject: "Analytics Breakfast –
Dresden"



High performance. Delivered.