

What is Computer Vision?



Computer Graphics: Models to Images

Comp. Photography: Images to Images

Computer Vision: Images to Models

Scope of CS 6476

Image Processing
Geometric Reasoning
Recognition
Deep Learning

Machine Learning

Graphics

Computational
Photography

Robotics

Human
Computer
Interaction

Medical
Imaging

Neuroscience

Optics

Computer Vision

Make computers understand images and video **or any visual data.**



What kind of scene?

Where are the cars?

How far is the building?

...

Vision is really hard

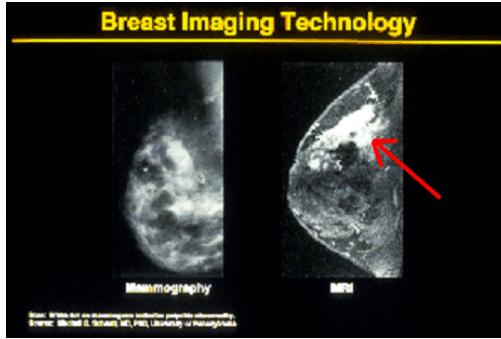
- Vision is an amazing feat of natural intelligence
 - Visual cortex occupies about 50% of Macaque brain
 - One third of human brain devoted to vision (more than anything else)



Why computer vision matters



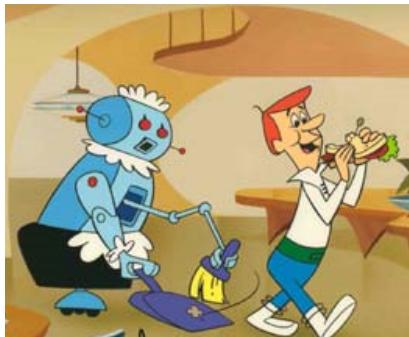
Safety



Health



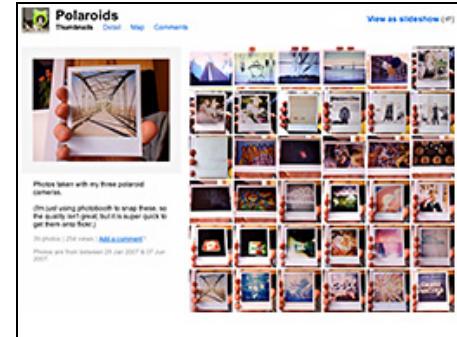
Security



Comfort



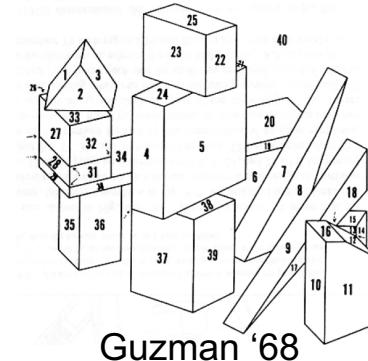
Fun



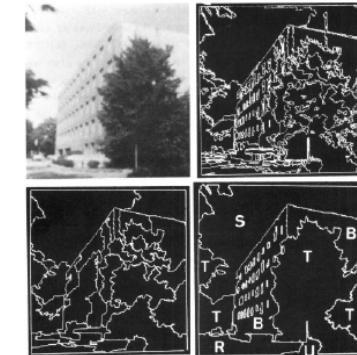
Access

Ridiculously brief history of computer vision

- 1966: Minsky assigns computer vision as an undergrad summer project
- 1960's: interpretation of synthetic worlds
- 1970's: some progress on interpreting selected images
- 1980's: ANNs come and go; shift toward geometry and increased mathematical rigor
- 1990's: face recognition; statistical analysis in vogue
- 2000's: broader recognition; large annotated datasets available; video processing starts
- 2010's: Deep learning with ConvNets
- 2020's: Widespread autonomous vehicles?
- 2030's: robot uprising?



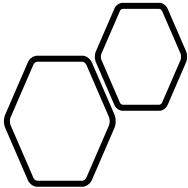
Guzman '68



Ohta Kanade '78



Turk and Pentland '91



Examples of real world
applications

How vision is used now

Optical character recognition (OCR)

Technology to convert scanned docs to text

- If you have a scanner, it probably came with OCR software



Digit recognition, AT&T labs
<http://www.research.att.com/~yann/>



License plate readers
http://en.wikipedia.org/wiki/Automatic_number_plate_recognition



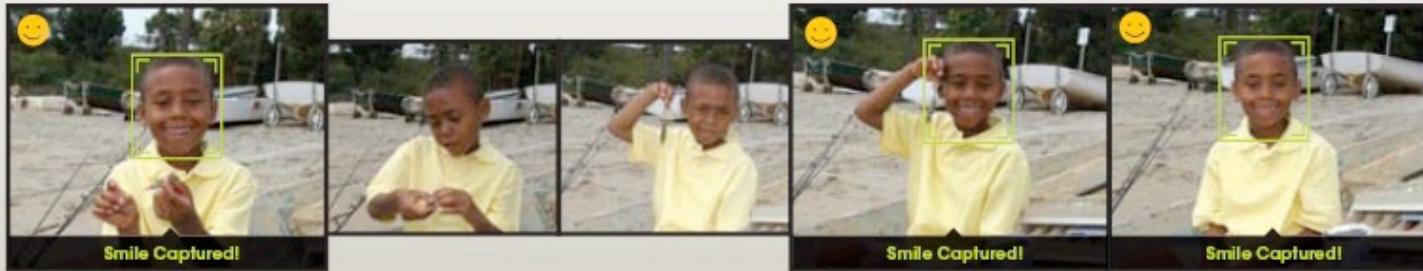
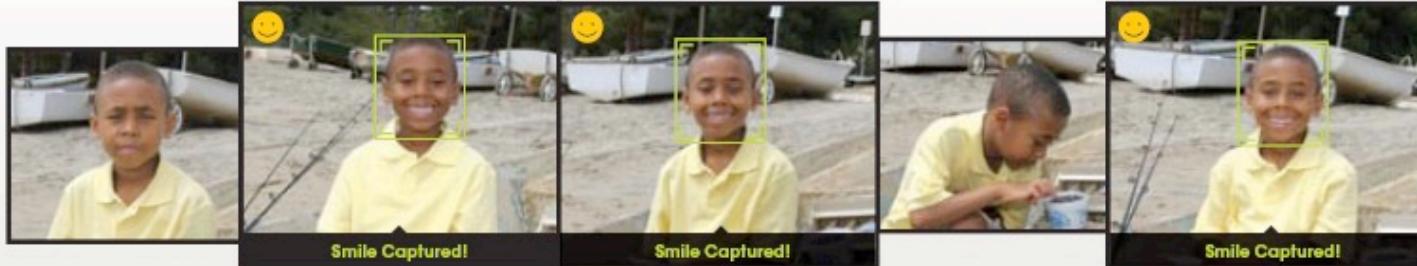
Face detection

Digital cameras detect faces

Smile detection

The Smile Shutter flow

Imagine a camera smart enough to catch every smile! In Smile Shutter Mode, your Cyber-shot® camera can automatically trip the shutter at just the right instant to catch the perfect expression.

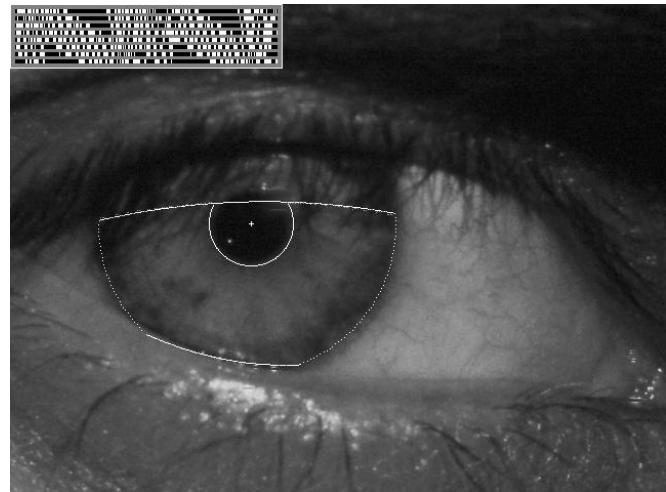
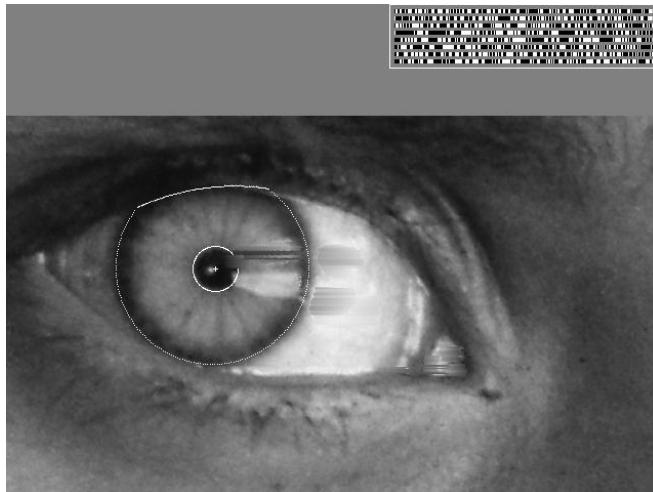


[Sony Cyber-shot® T70 Digital Still Camera](#)

Vision-based biometrics



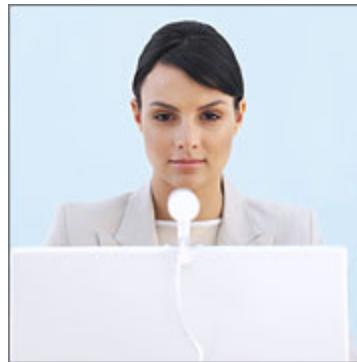
“How the Afghan Girl was Identified by Her Iris Patterns” Read the [story](#)
[wikipedia](#)



Login without a password...



Fingerprint scanners on
many new laptops,
other devices



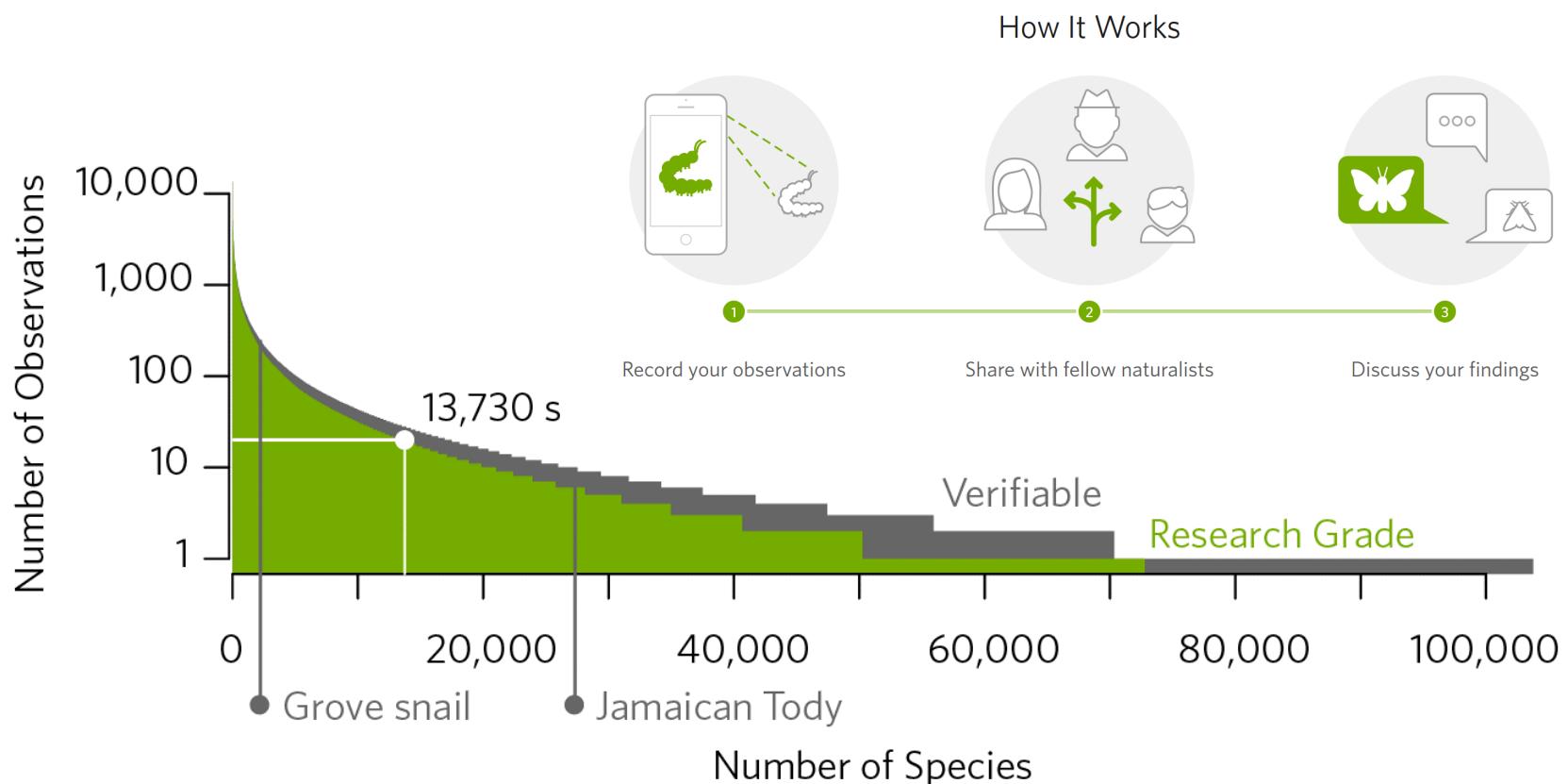
Face recognition systems now
beginning to appear more widely
<http://www.sensiblevision.com/>



Object recognition (in mobile phones)

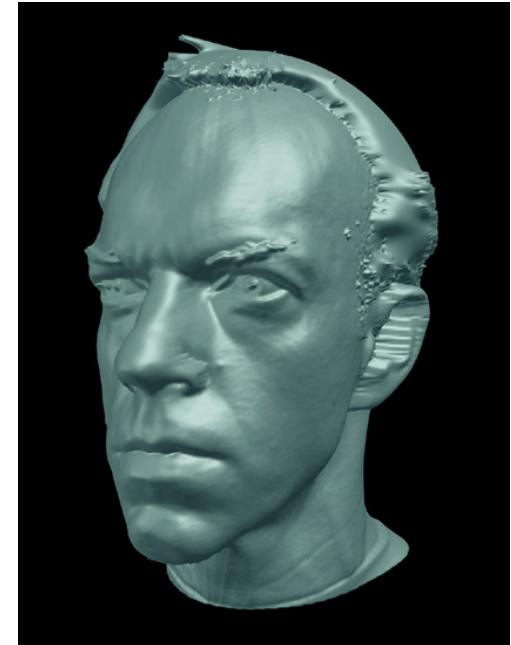
[Point & Find, Nokia](#)
[Google Goggles](#)

iNaturalist



https://www.inaturalist.org/pages/computer_vision_demo

Special effects: shape capture



The Matrix movies, ESC Entertainment, XYZRGB, NRC

Special effects: motion capture



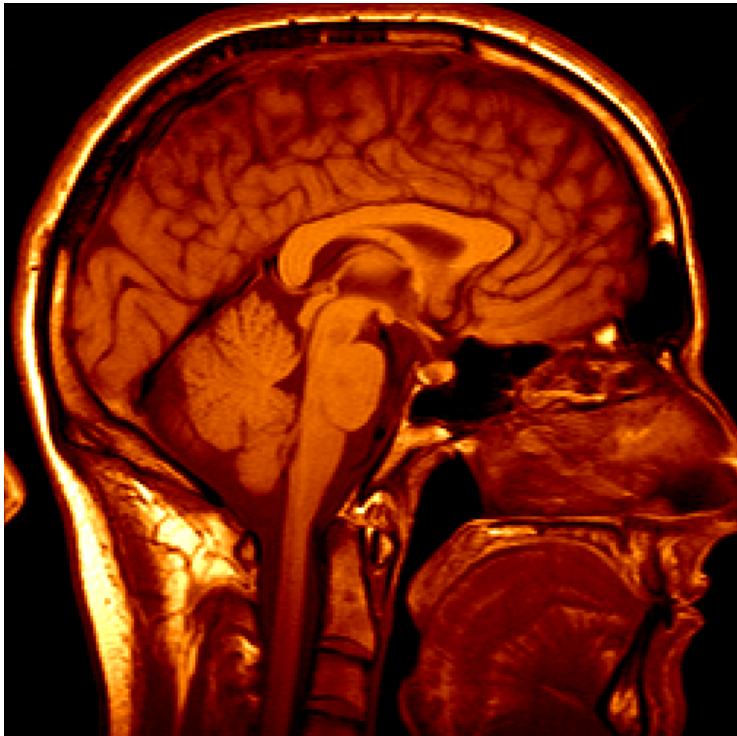
Pirates of the Caribbean, Industrial Light and Magic

Sports



Sportvision first down line

Medical imaging



3D imaging
MRI, CT



Image guided surgery
Grimson et al., MIT

Smart cars

Slide content courtesy of Amnon Shashua

The image shows a screenshot of the Mobileye website. At the top, there are navigation tabs: 'manufacturer products' (with a right arrow), 'consumer products' (with left and right arrows), and 'News'. Below this is a banner with the text 'Our Vision. Your Safety.' and an overhead view of a car with three cameras labeled: 'rear looking camera', 'forward looking camera', and 'side looking camera'. Below the banner are three main sections: 'EyeQ Vision on a Chip' (with an image of a chip), 'Vision Applications' (with an image of a person walking across a crosswalk), and 'AWS Advance Warning System' (with an image of a dashboard display). To the right, there is a 'News' sidebar with a list of articles and a 'Events' sidebar with links to 'Mobileye at Equip Auto, Paris, France' and 'Mobileye at SEMA, Las Vegas, NV'.

- > **EyeQ** Vision on a Chip
- > **Vision Applications**
Road, Vehicle, Pedestrian Protection and more
- > **AWS** Advance Warning System

News

- > [Mobileye Advanced Technologies Power Volvo Cars World First Collision Warning With Auto Brake System](#)
- > [Volvo: New Collision Warning with Auto Brake Helps Prevent Rear-end](#)

Events

- > [Mobileye at Equip Auto, Paris, France](#)
- > [Mobileye at SEMA, Las Vegas, NV](#)

- Mobileye
 - Market Capitalization: 11 Billion dollars
 - Bought by Intel for 15 Billion dollars

Google cars



Oct 9, 2010. "[Google Cars Drive Themselves, in Traffic](#)". *The New York Times*. John Markoff

June 24, 2011. "[Nevada state law paves the way for driverless cars](#)". *Financial Post*. Christine Dobby

Aug 9, 2011, "[Human error blamed after Google's driverless car sparks five-vehicle crash](#)". *The Star (Toronto)*

Interactive Games: Kinect

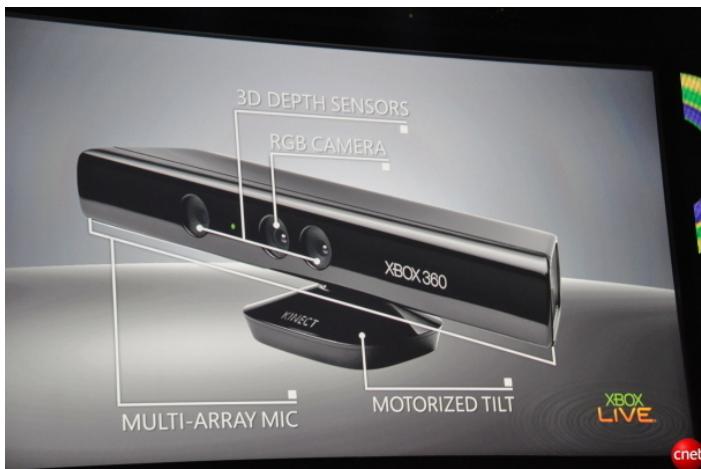
- Object Recognition:

<http://www.youtube.com/watch?feature=iv&v=fQ59dXOo63o>

- Mario: <http://www.youtube.com/watch?v=8CTJL5IUjHg>

- 3D: <http://www.youtube.com/watch?v=7QrnwoO1-8A>

- Robot: <http://www.youtube.com/watch?v=w8BmgtMKFbY>



Augmented Reality and Virtual Reality



Magic Leap, Oculus, Hololens, etc.

Industrial robots



Vision-guided robots position nut runners on wheels

Vision in space



[NASA'S Mars Exploration Rover Spirit](#) captured this westward view from atop a low plateau where Spirit spent the closing months of 2007.

Vision systems (JPL) used for several tasks

- Panorama stitching
- 3D terrain modeling
- Obstacle detection, position tracking
- For more, read “[Computer Vision on Mars](#)” by Matthies et al.

Amazon Prime Air



<https://www.amazon.com/b?node=8037720011>

State of the art today?

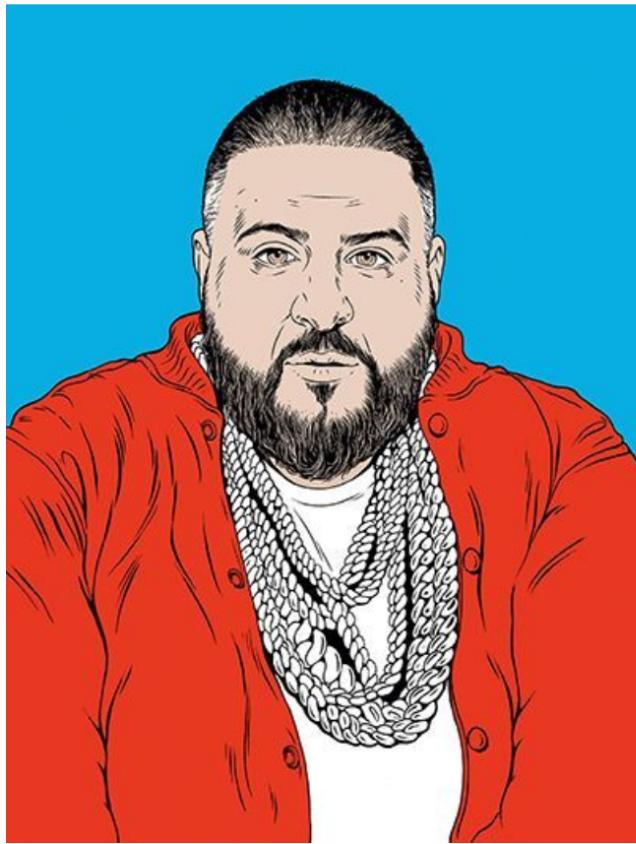
With enough training data, computer vision nearly matches human vision at most recognition tasks

Deep learning has been an enormous disruption to the field. More and more techniques are being “deepified”.

WIRED

100

WHO'S SHAPING THE DIGITAL WORLD?



DJ Khaled

Credit [Louise Zergaeng Pomeroy](#)

73. DJ Khaled

Snapchat icon; DJ and producer

Louisiana-born Khaled Mohamed Khaled, aka DJ Khaled, cut his musical chops in the early 00s as a host for Miami urban music radio WEDR. He proceeded to build a solid if not dazzling career as a mixtape DJ and music producer (he founded his label We The Best Music Group in 2008, and was appointed president of Def Jam South in 2009).

69. Geoffrey Hinton

Psychologist, computer scientist; researcher, Google Toronto

British-born Hinton has been dubbed the "godfather of deep learning". The Cambridge-educated cognitive psychologist and computer scientist started being an ardent believer in the potential of neural networks and deep learning in the 80s, when those technologies enjoyed little support in the wider AI community.

But he soldiered on: in 2004, with support from the Canadian Institute for Advanced Research, he launched a University of Toronto programme in neural computation and adaptive perception, where, with a group of researchers, he carried on investigating how to create computers that could behave like brains.

Hinton's work – in particular his algorithms that train multilayered neural networks – caught the attention of tech giants in Silicon Valley, which realised how deep learning could be applied to voice recognition, predictive search and machine vision.

The spike in interest prompted him to launch a free course on neural networks on e-learning platform Coursera in 2012. Today, 68-year-old Hinton is chair of machine learning at the University of Toronto and moonlights at Google, where he has been using deep learning to help build internet tools since 2013.

63. Yann Lecun

Director of AI research, Facebook, Menlo Park

LeCun is a leading expert in deep learning and heads up what, for Facebook, could be a hugely significant source of revenue: understanding its user's intentions.

62. Richard Branson

Founder, Virgin Group, London

Branson saw his personal fortune grow £550 million when Alaska Air bought Virgin America for \$2.6 billion in April. He is pressing on with civilian space travel with [Virgin Galactic](#).

61. Taylor Swift

Entertainer, Los Angeles





Credit [Google DeepMind](#)

[Google](#)-backed startup DeepMind Technologies has built an [artificial intelligence](#) agent that can learn to successfully play 49 classic Atari games by itself, with minimal input.



The story of AlphaGo so far

AlphaGo is the first computer program to defeat a professional human Go player, the first program to defeat a Go world champion, and arguably the strongest Go player in history.

AlphaGo's first formal match was against the reigning 3-times European Champion, Mr Fan Hui, in October 2015. Its 5-0 win was the first ever against a Go professional, and the results were published in full technical detail in the international journal, [Nature](#). AlphaGo then went on to compete against legendary player Mr Lee Sedol, winner of 18 world titles and widely considered to be the greatest player of the past decade.

AlphaGo's 4-1 victory in Seoul, South Korea, in March 2016 was watched by over 200 million people worldwide. It was a landmark achievement that experts agreed was a decade ahead of its time, and earned AlphaGo a 9 dan professional ranking (the highest certification) - the first time a computer Go player had ever received the accolade.

During the games, AlphaGo played a handful of [highly inventive winning moves](#), several of which - including move 37 in game two - were so surprising they overturned hundreds of years of received wisdom, and have since been examined extensively by players of all levels. In the course of winning, AlphaGo somehow taught the world completely new knowledge about perhaps the most studied and contemplated game in history.

Since then, AlphaGo has continued to surprise and amaze. In January 2017, an improved AlphaGo version was revealed as the online player "Master" which achieved [60 straight wins in online fast time-control games](#) against top international Go players.

In May 2017, Alpha Go took part in The Future of Go Summit in the birthplace of Go, China, to delve deeper into the mysteries of Go in a spirit of mutual collaboration with the country's top players. You can read more about the five day summit [here](#).



BigGAN

2018 Arxiv, ICLR 2019



Large Scale GAN Training for High Fidelity Natural Image Synthesis

Andrew Brock, Jeff Donahue, Karen Simonyan

This person does not exist

