Intro to Modern Computer Vision

(with an emphasis on self-driving cars)

By Julienne LaChance and Jonathan Lu

Note: We stole a bunch of slides from Drs. Andras Ferencz, Olga Russakovsky, Joshua Hug and Artur Filipowicz

The New York Times Magazine

Home Share 448

The Great A.I. Awakening

How Google used artificial intelligence to transform Google Translate, one of its more popular services — and how machine learning is poised to reinvent computing itself.

BY GIDEON LEWIS-KRAUS DEC. 14, 2016

Motivation

Full-Resolution Residual Networks (FRRNs) for Semantic Image Segmentation in Street Scenes

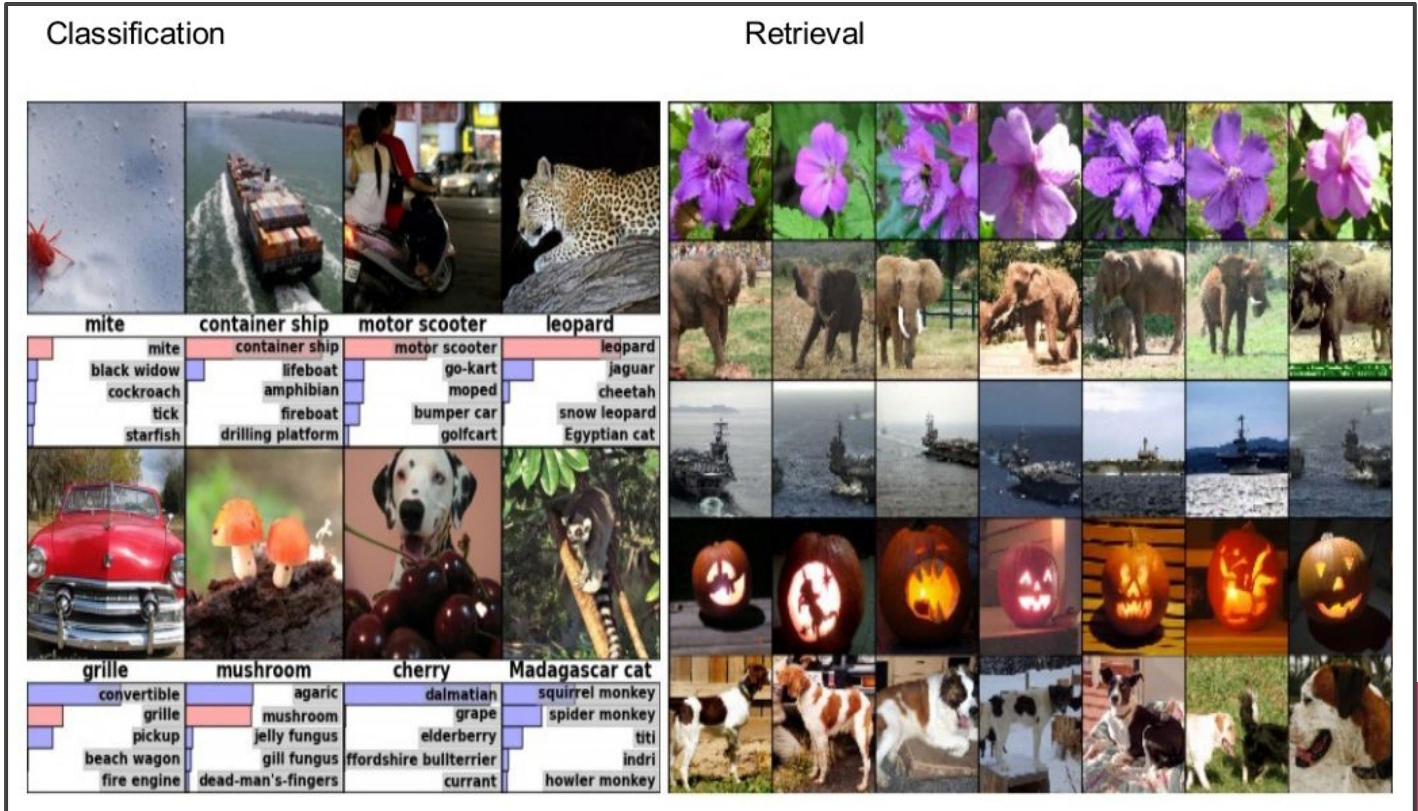
https://www.youtube.com/watch?v=PNzQ4PNZSzc

Overview

- Motivation
- Applications
- A very brief history of neural nets
- Computer vision for social good
- Autonomous Vehicles
- Ethics and Policy of Self-Driving Cars
- Summary

Applications

Image Classification and Retrieval

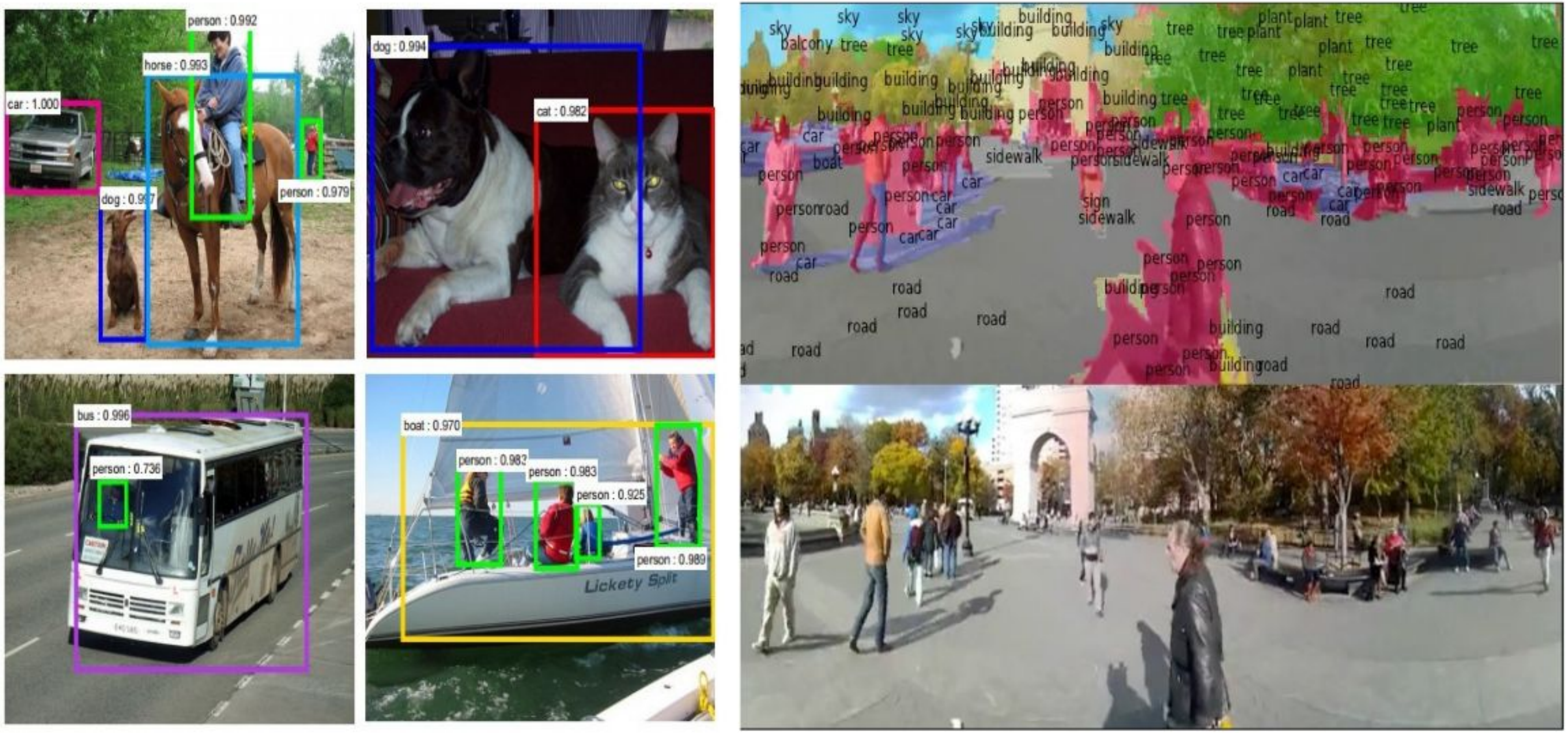


[Krizhevsky 2012]

Object Detection and Segmentation

Detection

Segmentation

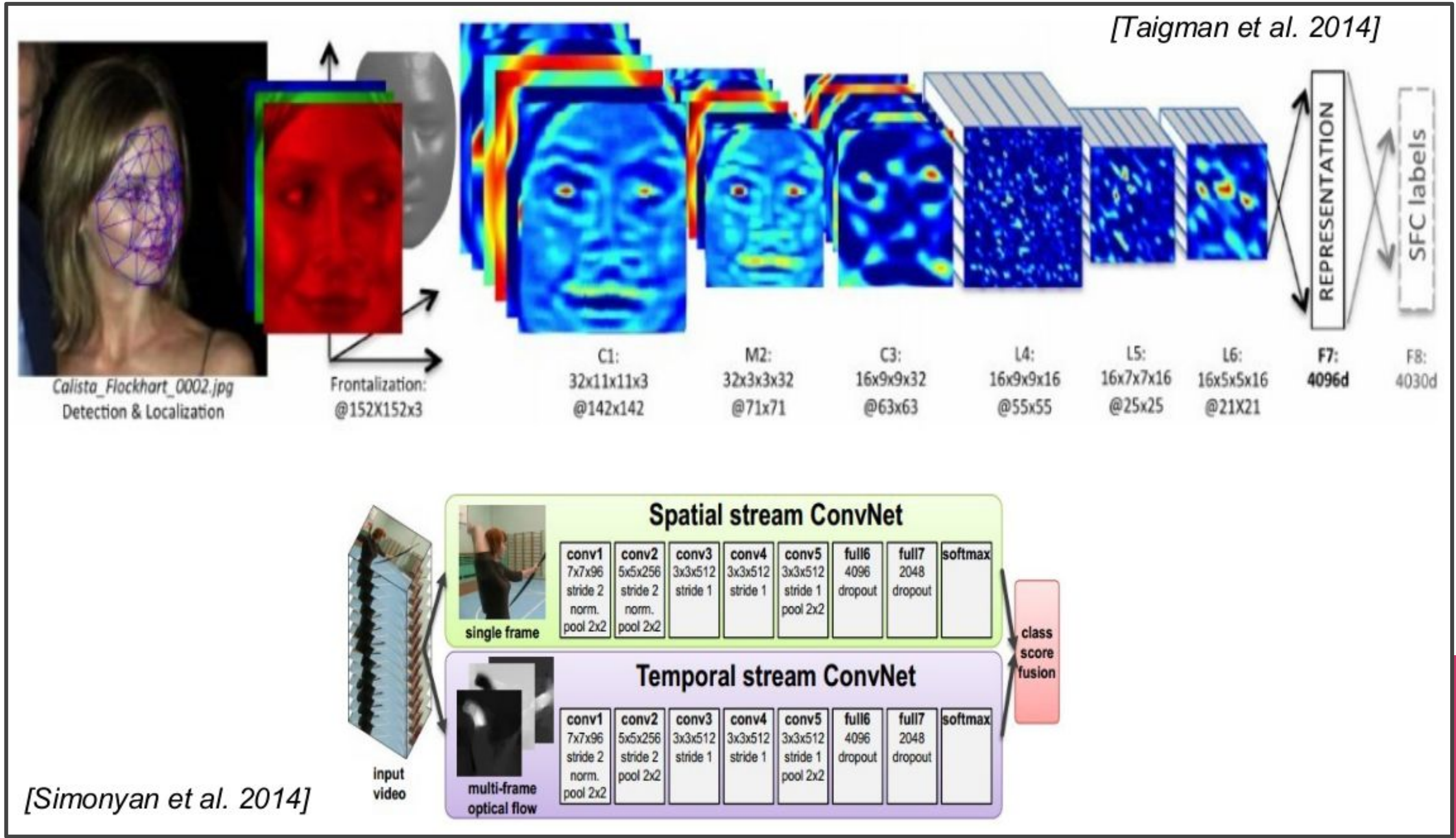


[Faster R-CNN: Ren, He, Girshick, Sun 2015]

[Farabet et al., 2012]

Semantic Image Segmentation

Face and Activity Recognition



Human Pose Estimation

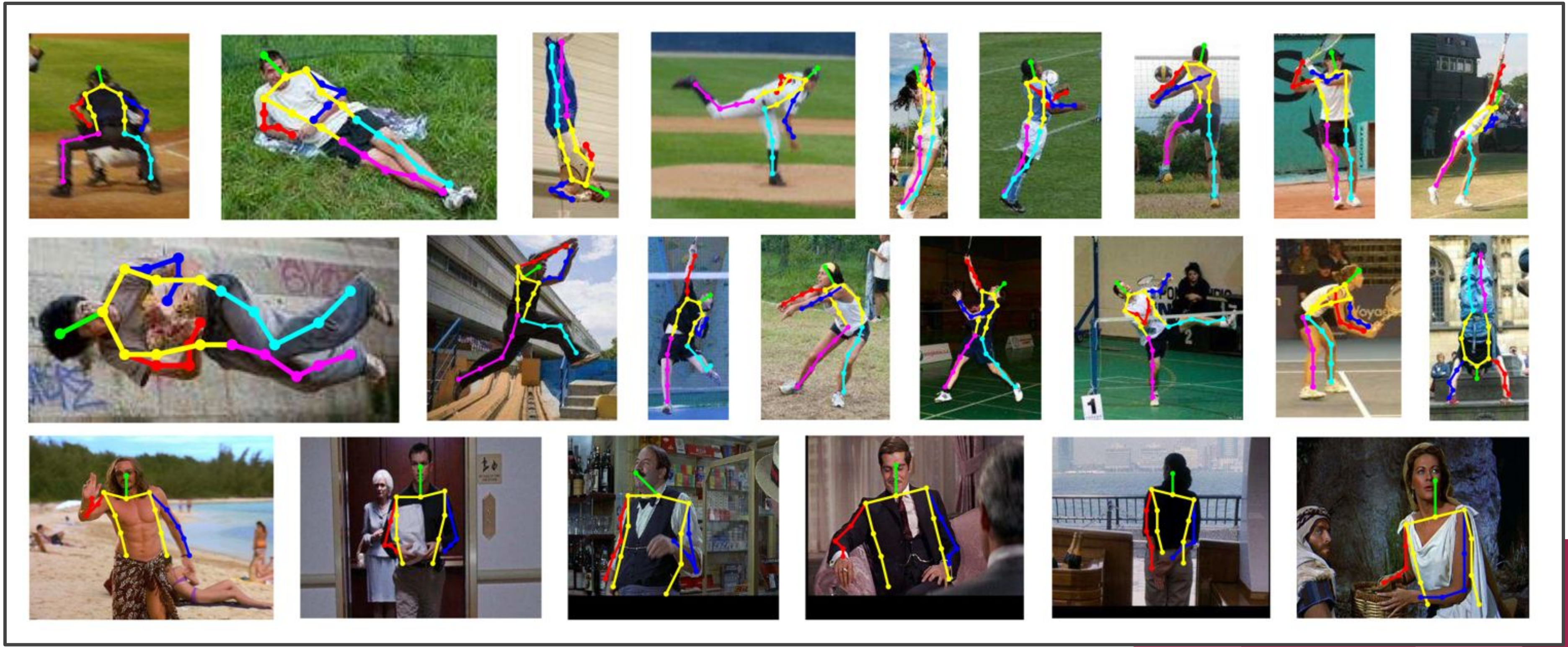


Image Captioning

[Vinyals et al., 2015]

“Deep Art”

reddit.com/r/deepdream

Medical Applications

APRIL 3, 2017

Stanford researchers create deep learning algorithm that could boost drug development

Combining computer science and chemistry, researchers show how an advanced form of machine learning that works off small amounts of data can be used to solve problems in drug discovery.

amazingly subtle information, enabling them
nedical images as well as a doctor. But in
training that involves thousands to trillions
't work all that well in situations where there

Stanford MEDICINE | News Center

FRONT PAGE ALL NEWS TOPICS MULTIMEDIA

Email Share 1.1K Tweet

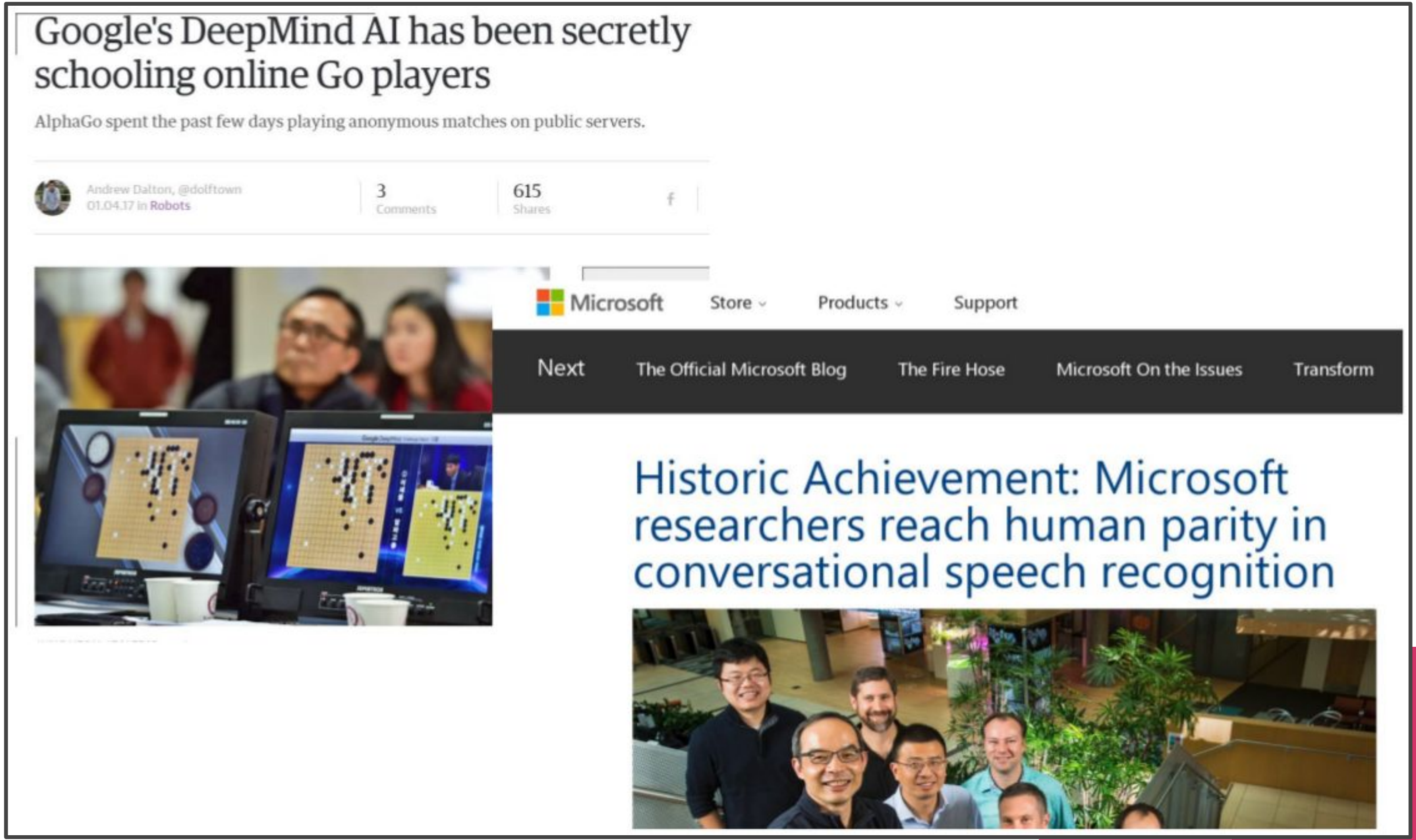
Computers trounce pathologists in predicting lung cancer type, severity

Automating the analysis of slides of lung cancer tissue samples increases the accuracy of tumor classification and patient prognoses, according to a new study.

AUG 16 2016

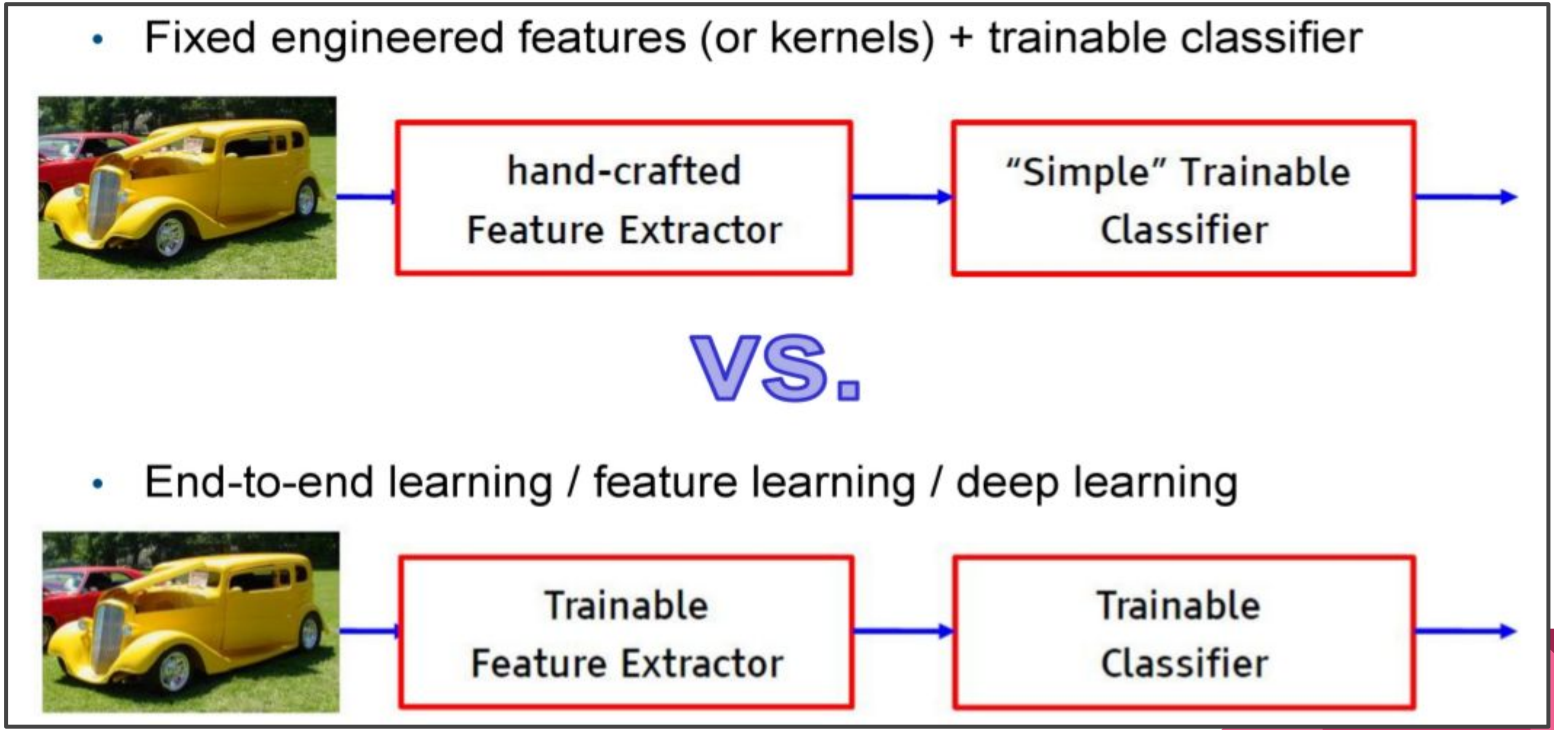
Computers can be trained to be more accurate than pathologists in assessing slides of lung cancer tissues, according to a new study by researchers at the Stanford University School of Medicine.

Beating humans at their own games

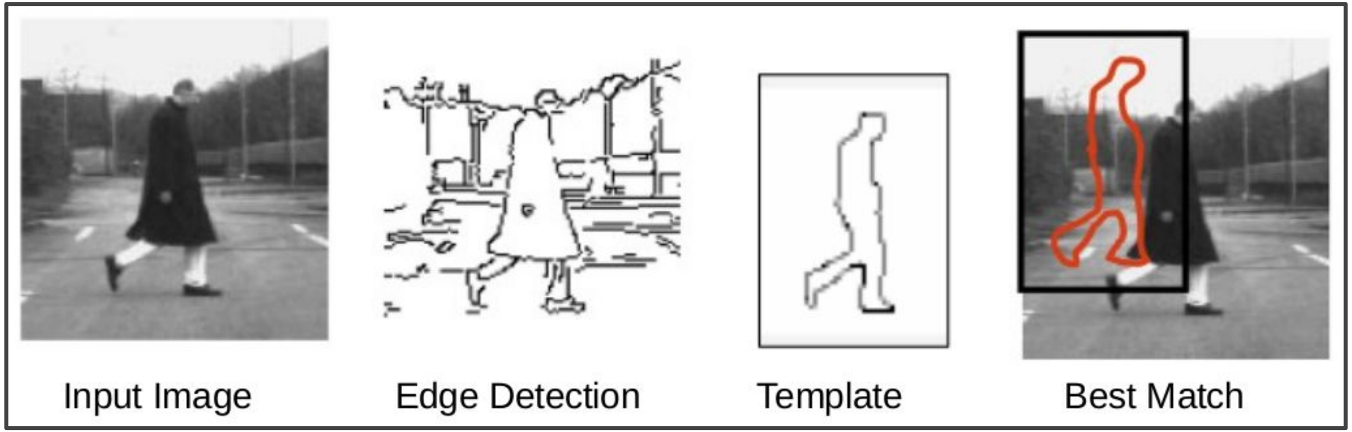


A very brief history of neural nets

Old school machine learning versus today:

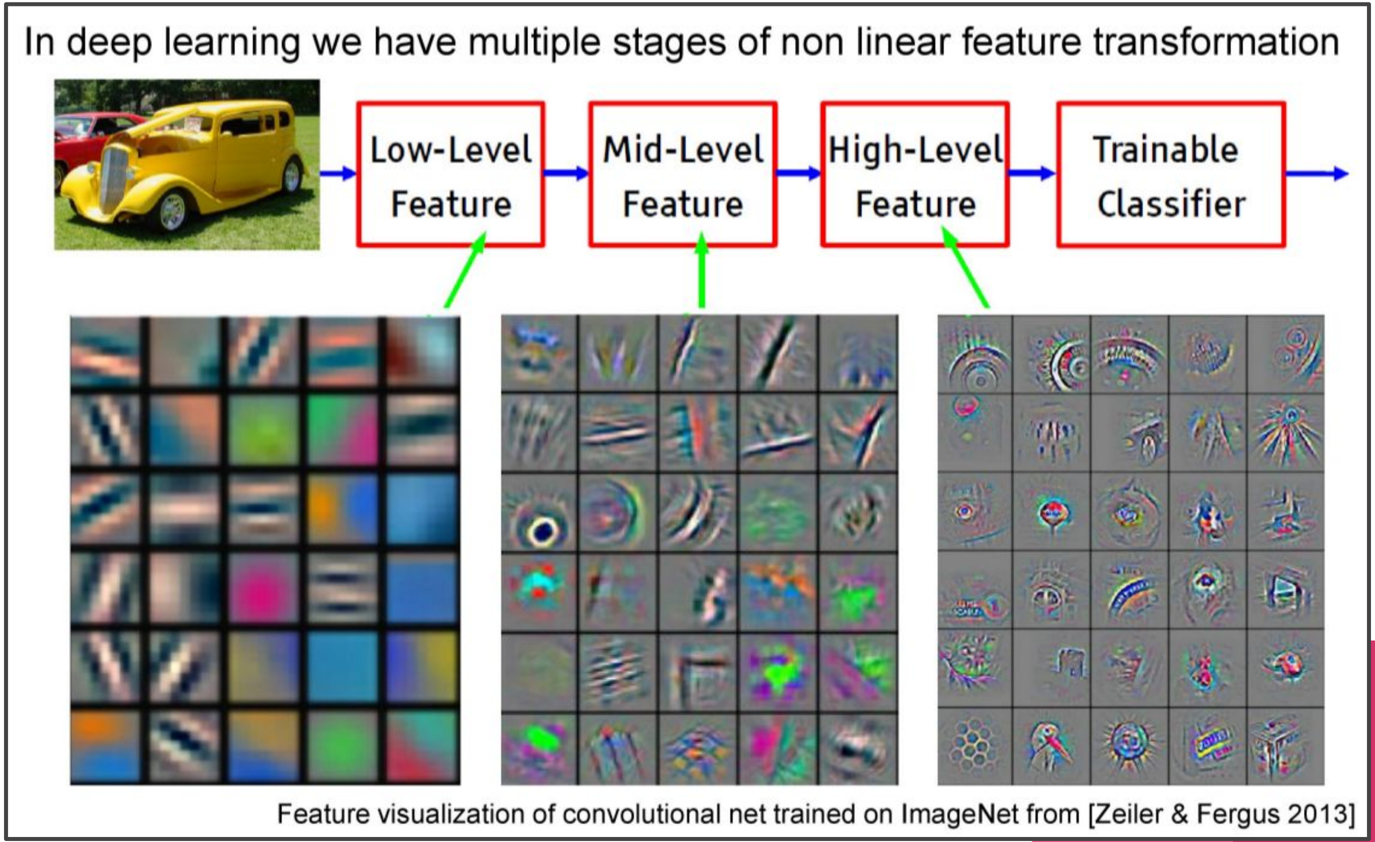


Old school machine learning versus today:

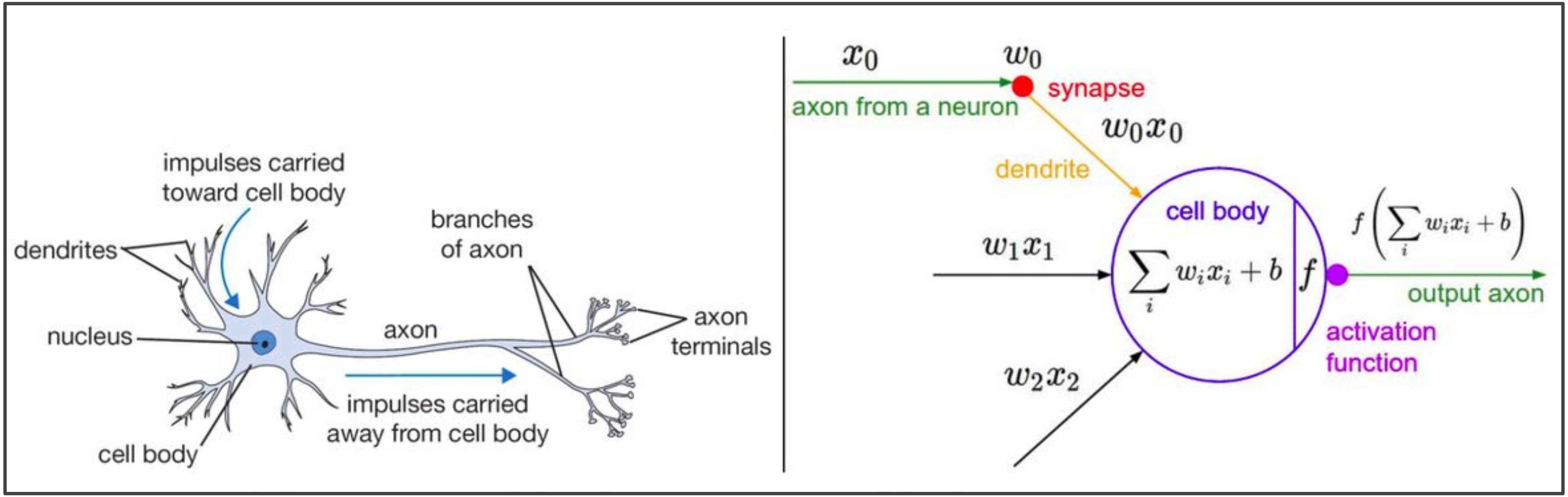


1999: Chamfer matching to do pedestrian detection

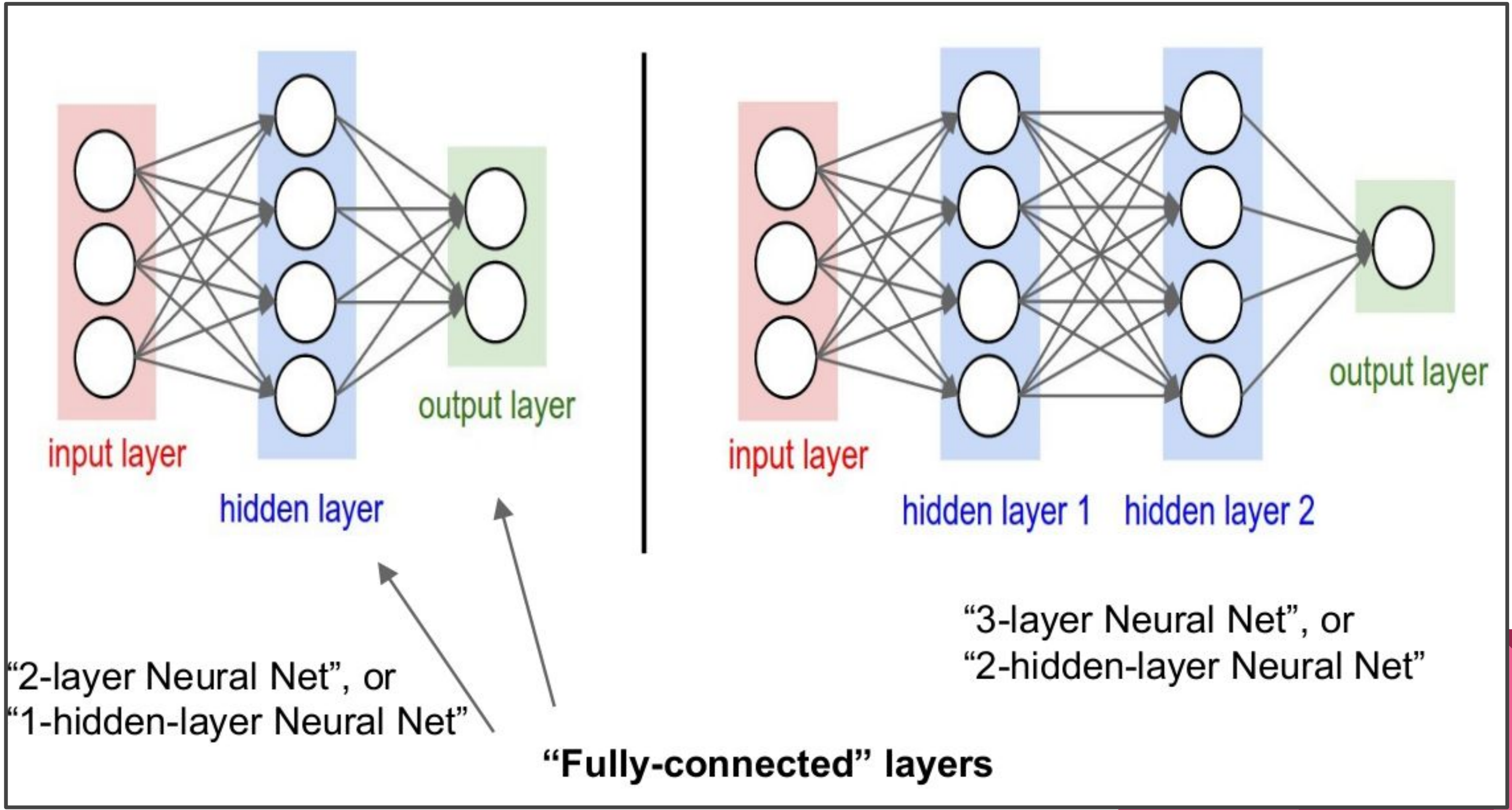
Low level -> high level features



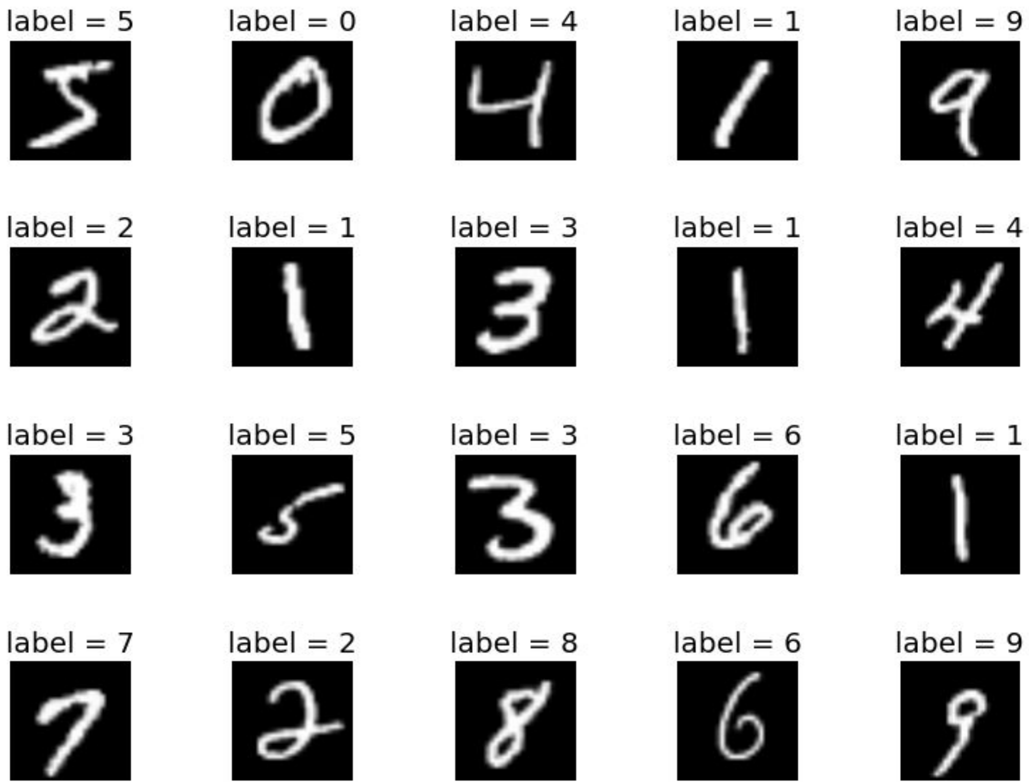
Why “neural” networks?



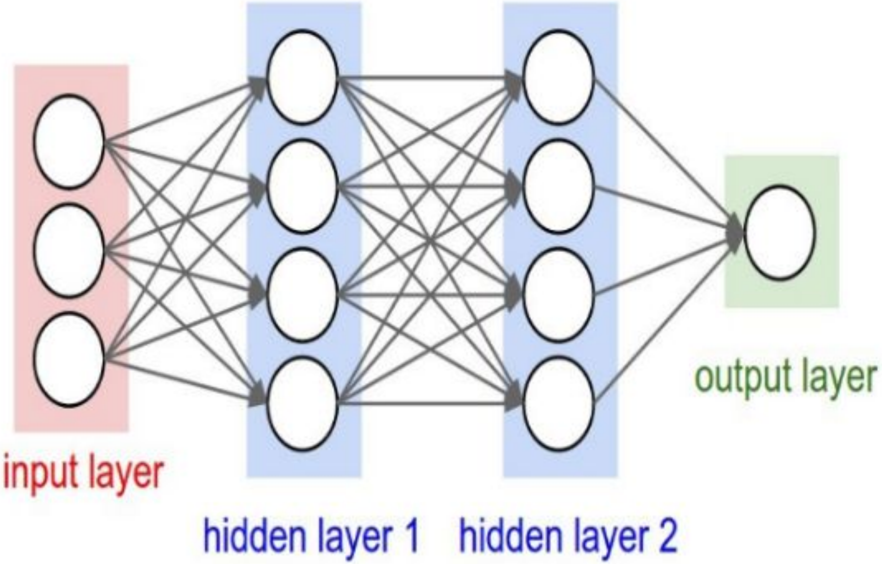
Neural networks



Example: MNIST dataset

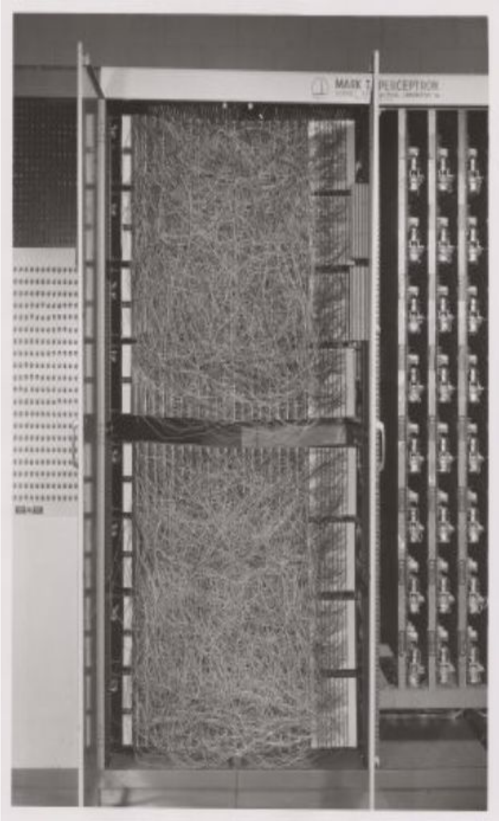


Example: MNIST dataset



3

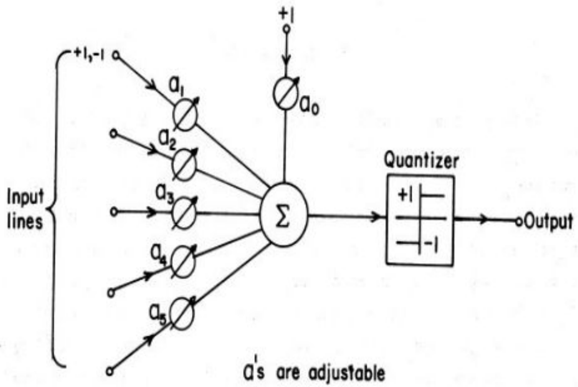
The first artificial neural networks

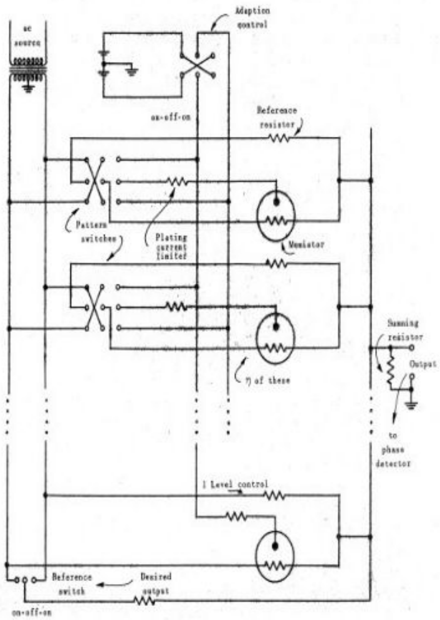


- 1957: The Mark I Perceptron machine was the first implementation of the perceptron algorithm.
- The machine was connected to a camera that used 20×20 cadmium sulfide photocells to produce a 400-pixel image.
- Recognized letters of the alphabet.
- Update rule:

$$w_i(t+1) = w_i(t) + \alpha(d_j - y_j(t))x_{j,i}$$

The first artificial neural networks

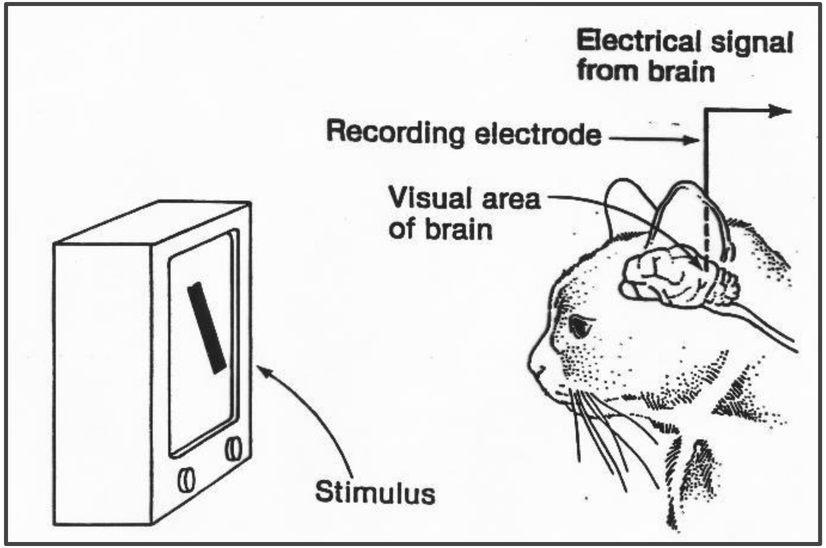


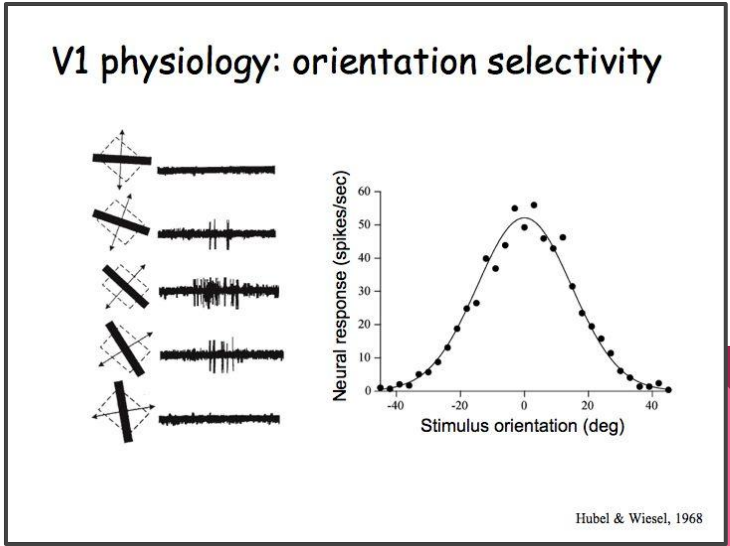


- 1960 Adaline/Madaline: (Many Adaptive Linear Neuron)

Processing in the visual system

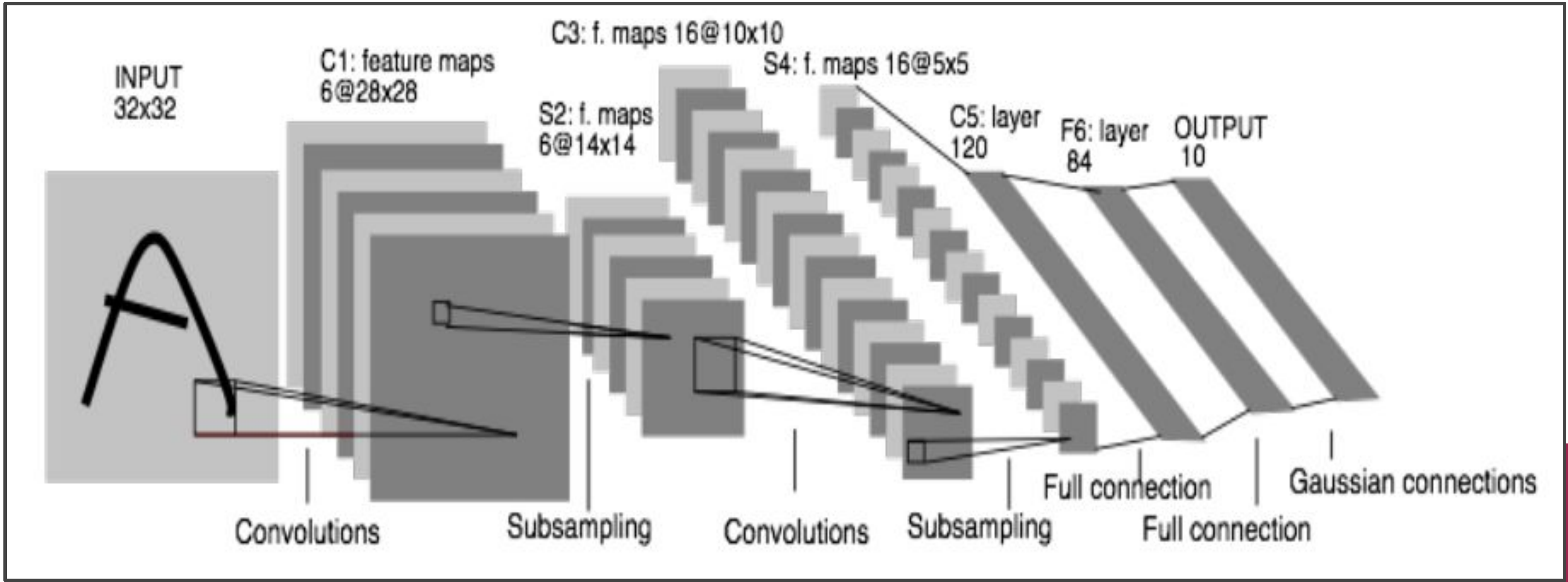
- 1959: first Hubel and Wiesel experiments
  - The team inserted a microelectrode into the primary visual cortex of an anesthetized cat, then projected patterns of light and dark on a screen in front of the cat. Neuronal responses to the patterns were recorded.
  - Visual parcels in the neo-cortex found to generate edge detectors, motion detectors, stereoscopic depth detectors and color detector.



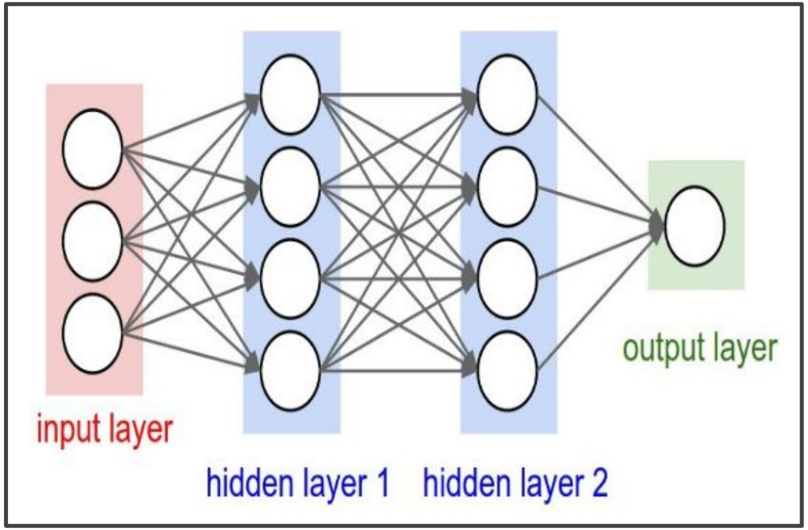


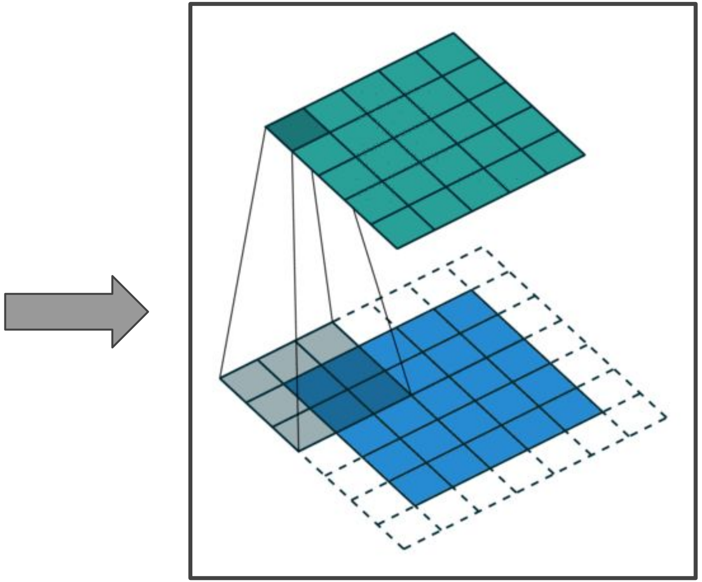
The first artificial neural networks

- 1986: First time back-propagation becomes popular (for training neural networks efficiently)
- 1998: LeCun develops convolutional neural networks for handwritten digit classification



Convolutional neural networks





- Traditional neural networks are composed of ***fully-connected layers***, in which neurons are connected to every neuron in the previous layer. But this is impractical for use in images.
- Instead, we utilize a ***convolution*** operation, tiling regions with the same set of shared weights. The features learned are effectively filters on the input regions.

Modern artificial neural networks

- 2012: First strong modern results for image classification, followed by a “revolution of depth”.

And even deeper…
ResNet has 152 layers!

What really prompted the modern age of AI?

1. The prevalence of household GPUs (thanks, gamers)

2. The development of massive, high-quality data sets

Computer Vision for Social Good

Discussion Question:

What are some ways computer vision algorithms can be used for social good?

Applications of Computer Vision

- Classification/Retrieval
- Object Detection/Segmentation
- Semantic Image Segmentation
- Face and Activity Recognition, Pose Estimation
- Image/Video Captioning

Object Detection + Classification:
Traffic Video Analysis for Assessing Safety Risks

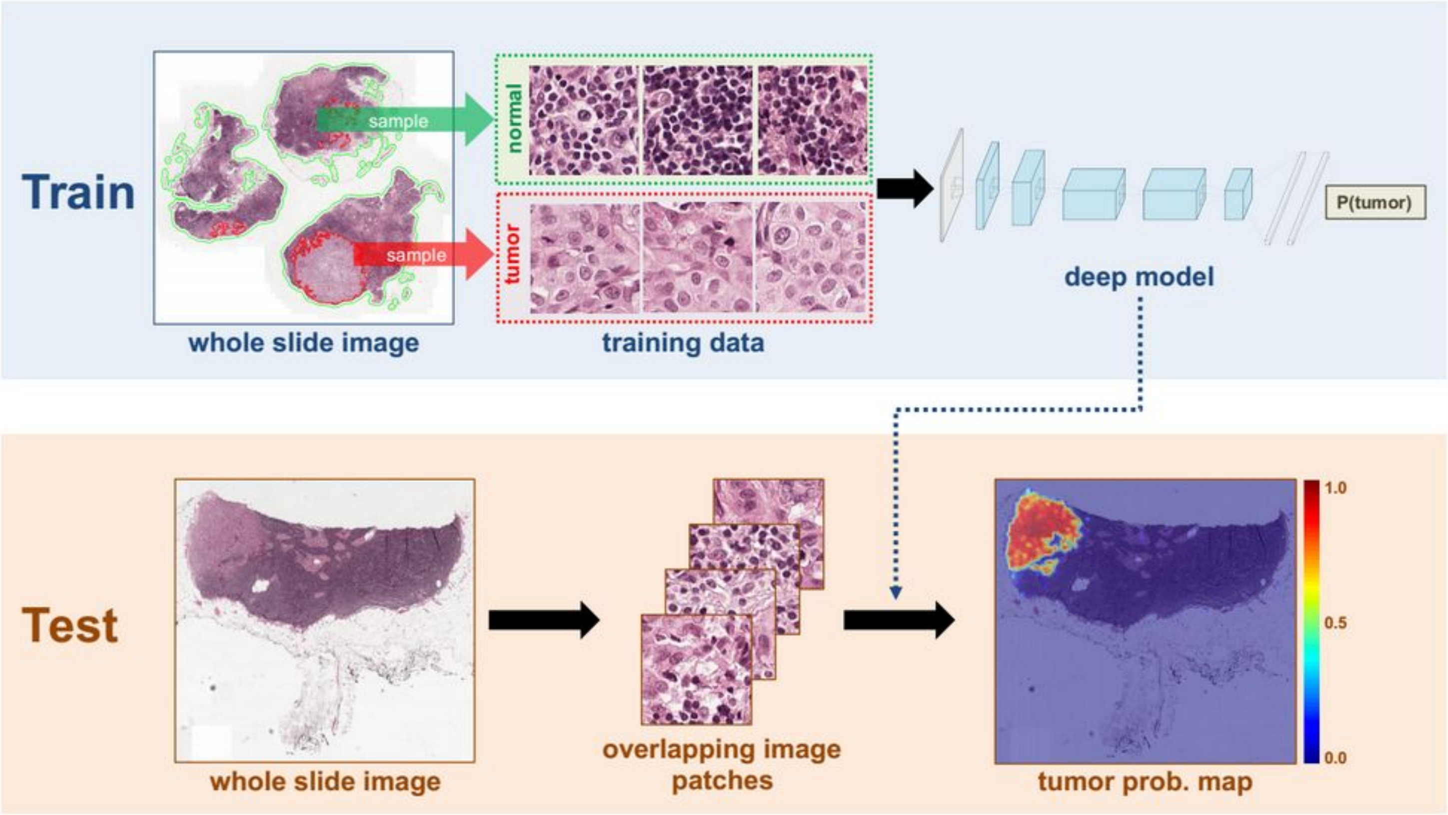
- Nearly 2,000 people die annually as a result of being involved in traffic-related accidents in Jakarta, Indonesia.

- The city government has invested resources in thousands of traffic cameras to help identify potential short-term (e.g. vendor carts in a hazardous location) and long-term (e.g. poorly engineered intersections) safety risks.

https://dssg.uchicago.edu/project/improving-traffic-safety-through-video-analysis/

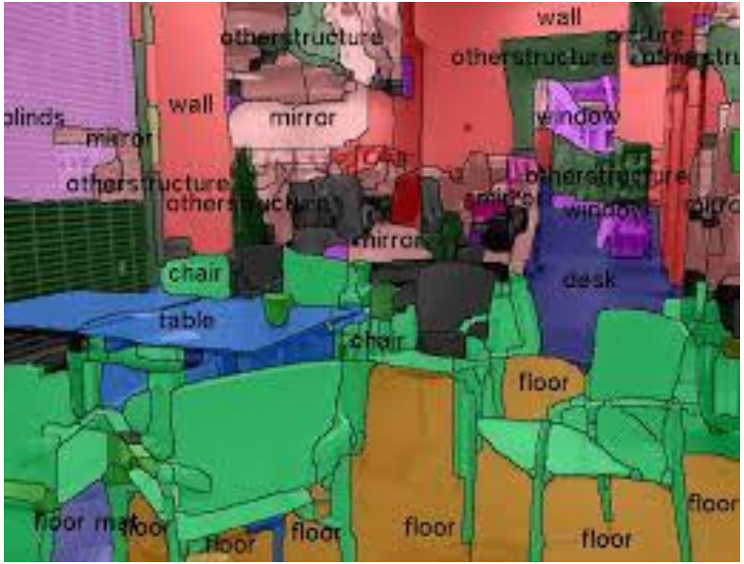
Traffic Counting and Classification

Object Detection/Segmentation: Breast Cancer Diagnoses

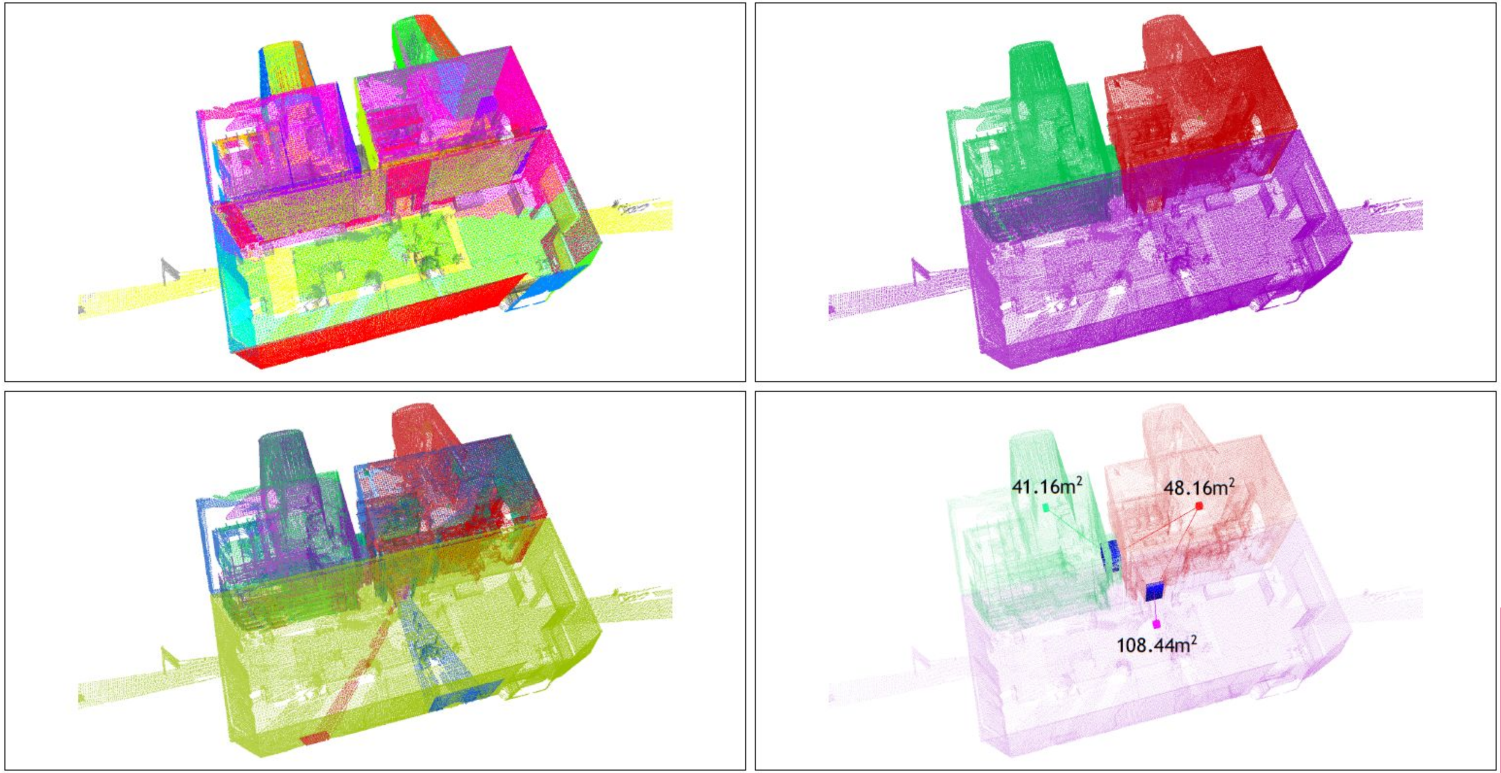


Semantic Segmentation: Indoor Scene Understanding

- If we wish to design better indoor spaces for handicap accessibility, first we need to understand what stuff is present in our current living and working spaces.



Semantic Segmentation: Accessibility Design

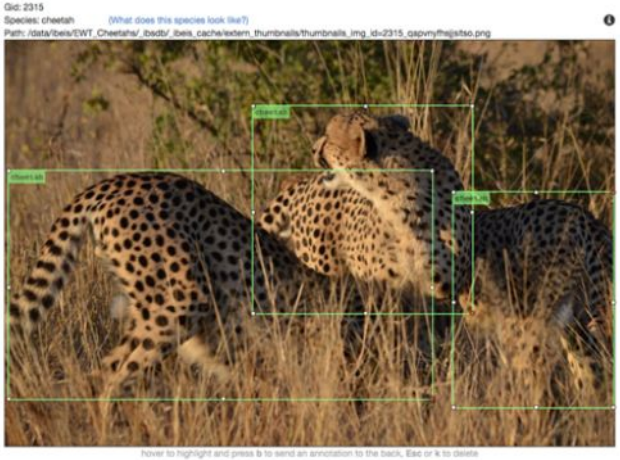


Identifying Individuals / Motion Tracking

- The IBEIS Project: how to use safari pictures (taken by tourists!) to help ecologists study large herds of animals.
- Deep learning is used to:
  - Find animals in each photo, and
  - Identify individual animals based on their markings.
- Helps develop new methods to combat extinction.

Image
Based
Ecological
Information
System
IBEIS

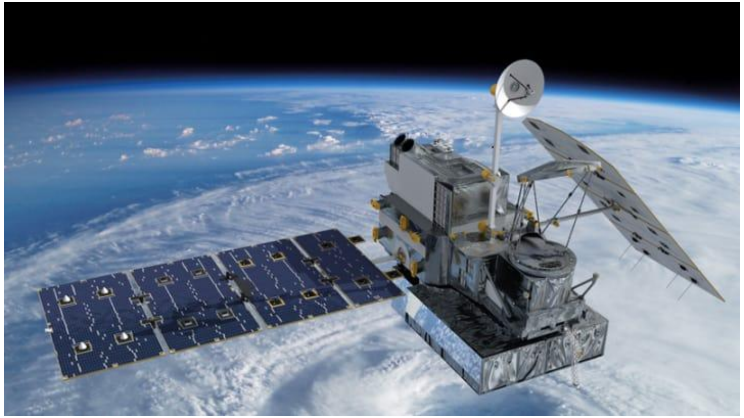

HOW MANY
ZEBRAS & GIRAFFES
ARE THERE AT THE NAIROBI NATIONAL PARK?
Join in the 'Great Zebra and Giraffe Count' & help Scientists find out!
All you have to do is grab a camera (not a cellphone), drive around Nairobi Park for an hour or so on the 1st or 2nd of March and take as many pictures of zebras and giraffes as you can.
Scientists working with the Kenya Wildlife Service will do the rest! They will use IBEIS (Image-Based Ecological Information System) software to find all the zebras and giraffes in the photos and to recognize individual animals. IBEIS will help carry out a 'sight-resight' analysis that will estimate the overall numbers of the animals in the park.
SO:
Form a team
Pre-Register
Show up, get briefed, get set up
Be assigned a route and drive it
Stop and take the pictures
Share your card
Receive a thank you gift
And stay tuned for the analysis will be emailed to you.
Image Based Ecological Information System
IBEIS
And best of all, have fun and help scientists learn about the zebras and giraffe of Nairobi National Park


Gid: 2315
Species: cheetah
(What does this species look like?)
Path: /data/ibeis/EWT_Cheetahs/_ibsdb/_ibeis_cache/extern_thumbnails/thumbnails_img_id=2315_qapvnyfhejsitso.png
hover to highlight and press b to send an annotation to the back, Esc or k to delete

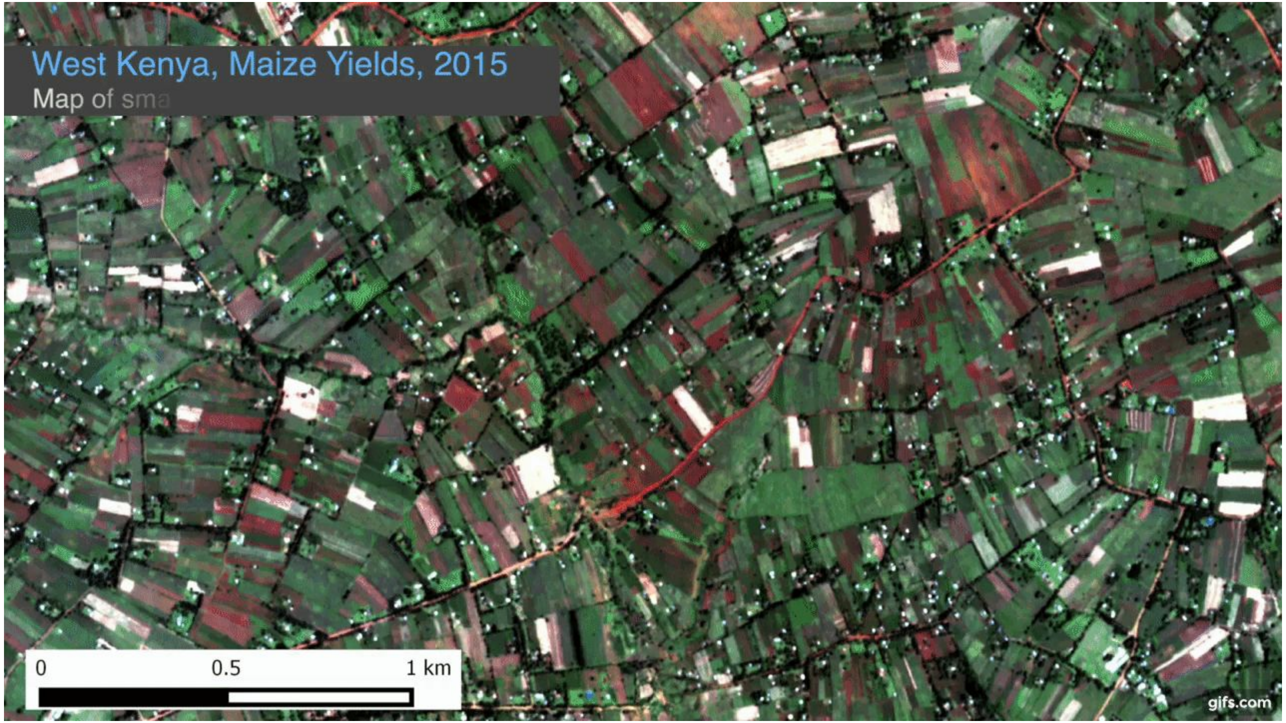

qnid4474 : qaid5864
aid2420 : 0.0078
nid4474 : aid6521 : 0.126
aid290 : 0.0

Computer Vision for Policy Design + Planning

- How can we use existing data to improve civil planning?
- One idea: using satellite images to measure farm yields, and learn strategies to support small-scale farmers.



Policy Design + Planning



Computer Vision for Social Good

Takeaway:

Extremely powerful tools can be built by cleverly applying modern techniques for tasks like object detection, classification, and so on.

It's up to you to find creative new ways to solve old problems.

Autonomous Vehicles

LIDAR: How autonomous vehicles “see”

- LIDAR works sort of like radar: a device on top of the vehicle emits pulses of infrared light (invisible to the human eye) and measures how long they take to come back after hitting nearby objects .

So you might look something like this...

Some systems are augmented with cameras

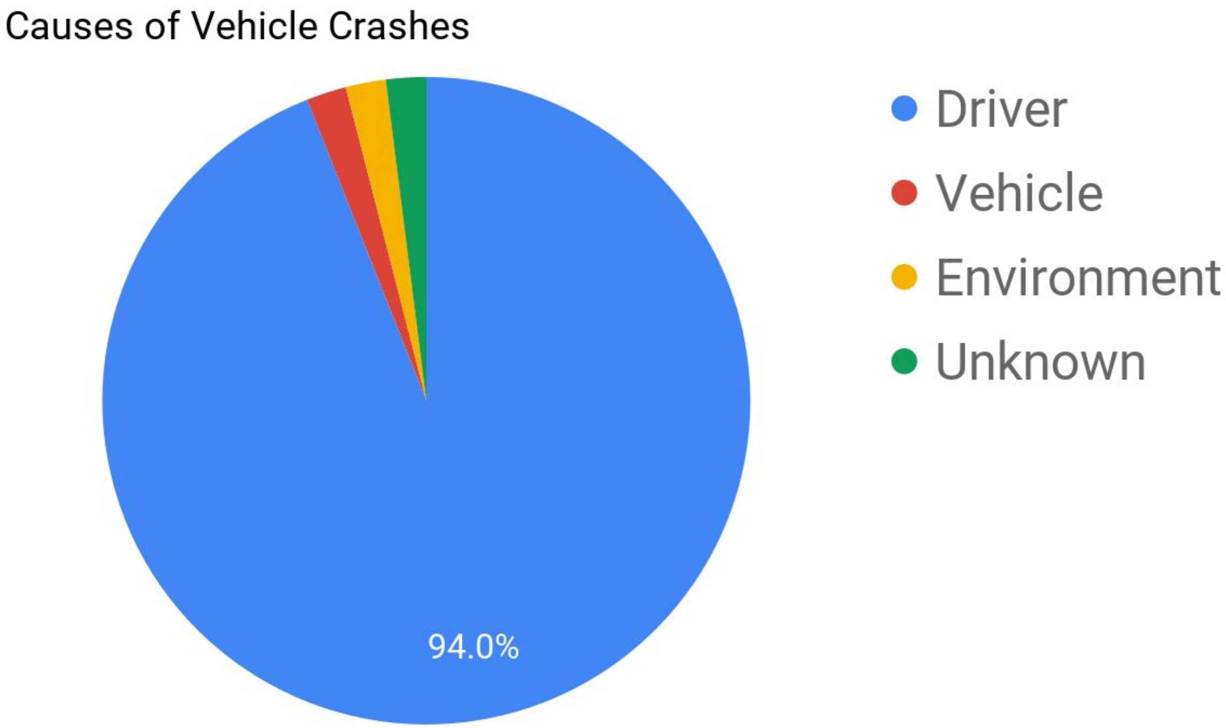
- Visual input is more intuitive to human developers.
  So it makes it easier to do things like lane detection.

Why have self-driving cars?

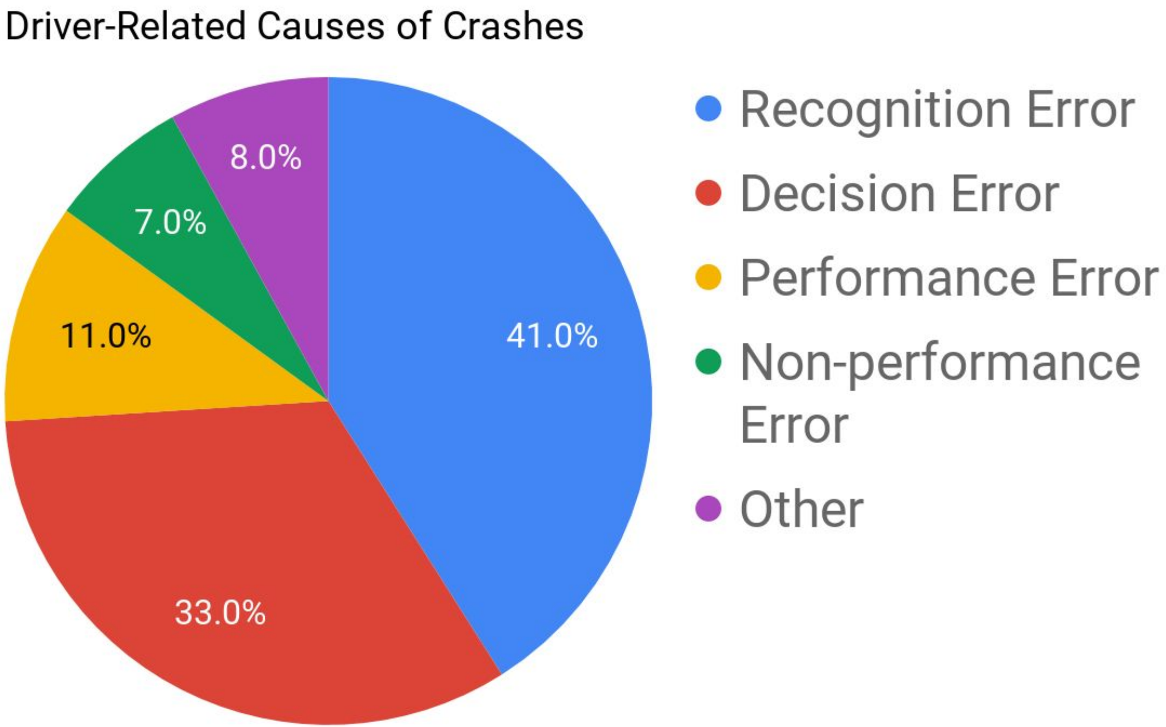
Why have them? Safety.

- 35,092/yr die in car crashes
- Leading Cause of death for Teens
  - 16,375/year! (of ~40 mill)
- **90%** of Crashes could be prevented!

Why have them? Safety.



Why have them? Safety.



Why have them? Quality of Life.

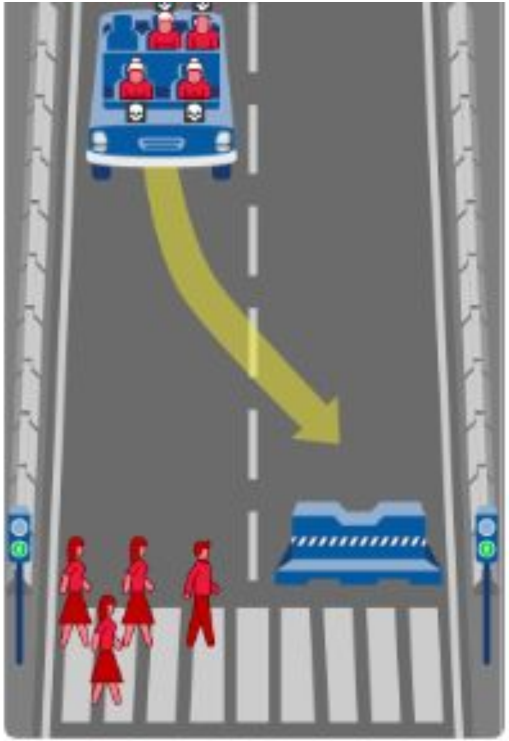
- Commute, Traffic
- Mobility, if you can't drive!
  - Elderly, Disabled, Young
  - But must be <u>accessible</u>!
- Savings!
  - $1.3 Trillion for U.S. Economy (8% GDP)

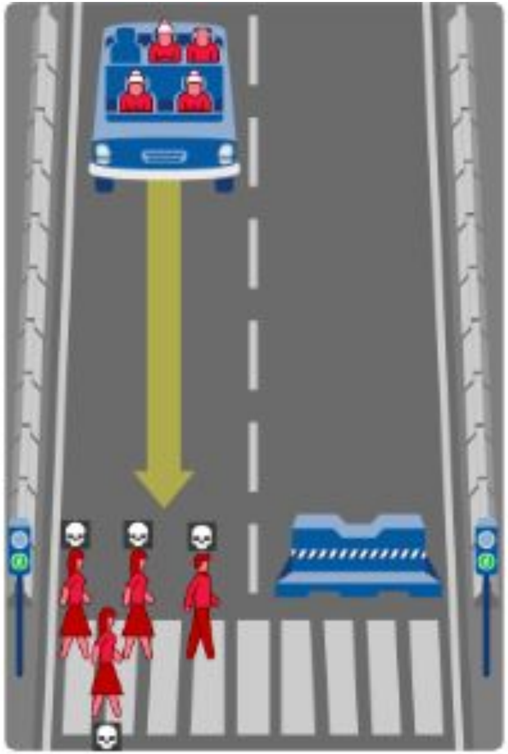
Where are we now?

- Small Steps
  - Automatic Parking
  - Autopilot
  - Lane-keep assist
- 2018 Uber Crash
- Projections: 2018-2030

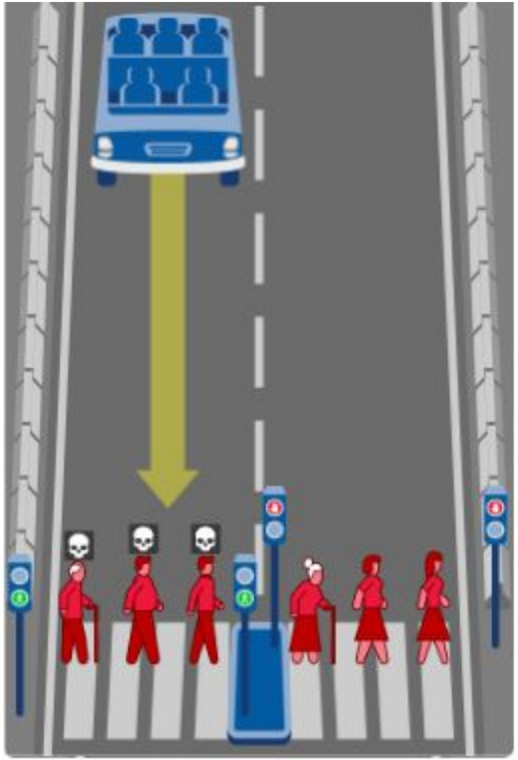
How will these impact society?

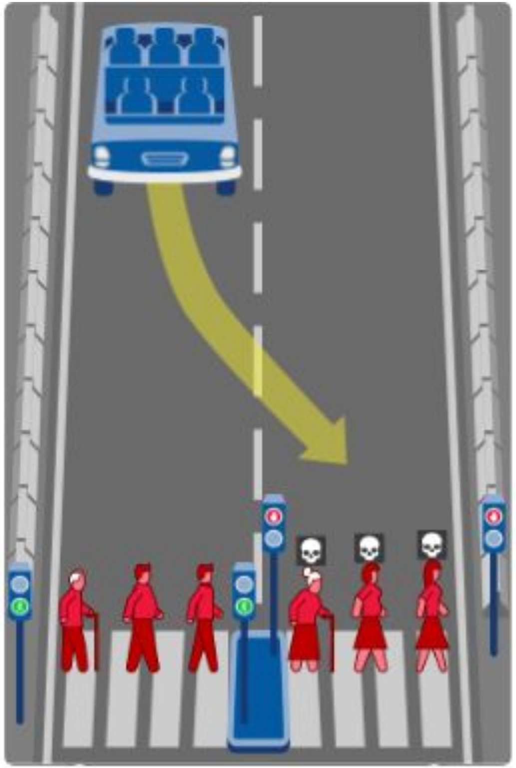
What should the car do?





What should the car do?





Technology Replaces Jobs

Which of these can be automated? What else?

US Labor Pool

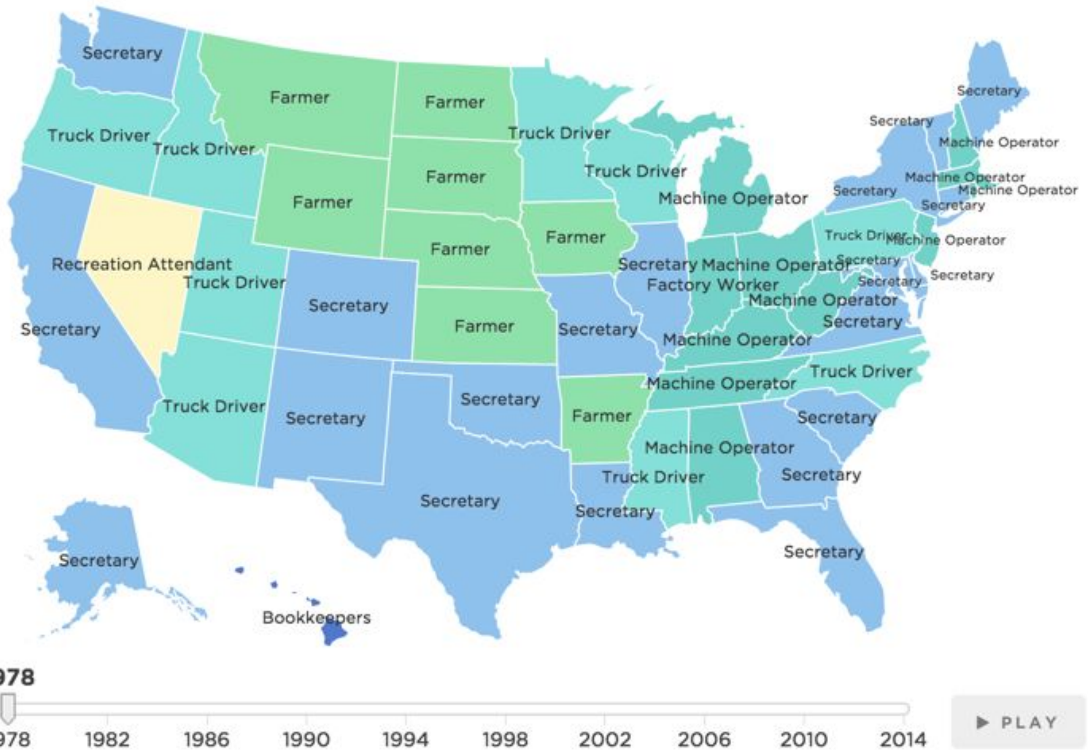
Of the following, what was the most common job in 2014?

1. Machine Operator
2. Farmer
3. Nurse
4. Truck Driver
5. Secretary
6. Teacher

FOLLOWUP

In how many states is this the most common job?

What are the most common jobs?



http://www.npr.org/sections/money/2015/02/05/382664837/map-the-most-common-job-in-every-state

US Labor Pool

Of the following, what was the most common job in 2014?

1. Machine Operator
2. Farmer
3. Nurse
4. **<u>Truck Driver</u>**
5. Secretary
6. Teacher

FOLLOWUP

In how many states is this the most common job? **<u>29</u>**

3.5 million truck drivers (~3% of workforce) unemployed.
How do you handle this?

Possibilities

- Tax AI
- Job Training Programs
- Universal Basic Income
- Auction Permits for Self-driving Vehicles

Summary

- Motivation
- Applications
- A very brief history of neural nets
- Self-Driving Cars
- Ethics and Policy Questions

Questions?

References

- Slides 2; 6-14; 16-24: Olga Russakovsky and Andras Ferencz. “Introduction to Computer Vision” lecture slides: http://www.cs.princeton.edu/courses/archive/fall17/cos429/outline.html
- Slides 27-29: Olga Russakovsky. “Advanced Topics in Computer Science- Visual Recognition” lecture slides: https://registrar.princeton.edu/course-offerings/course_details.xml?courseid=002129&term=1184
- Slides 30-36: Antonio Torralba and Alyosha Efros. “Unbiased Look at Dataset Bias”: http://people.csail.mit.edu/torralba/research/bias/
- Slide 37: Moritz Hardt. “Fairness in Machine Learning” lecture slides: https://fairmlclass.github.io/
- Slides 41-58: Riley Simmons-Edler and Berthy Feng. Visual Recognition seminar slides: https://docs.google.com/presentation/d/1Nw7f63wNaoRWTa6j77rGwQjSULTAIcnKqQQ7fveVbAk/edit#slide=id.p

Backup Slides

Datasets and Bias

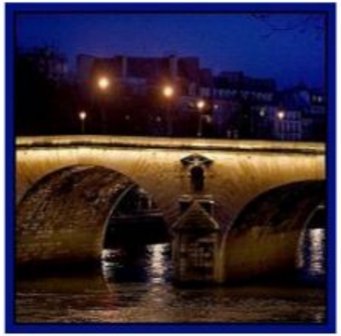
But we must be careful!

Do datasets provide a *good* representation of the world?

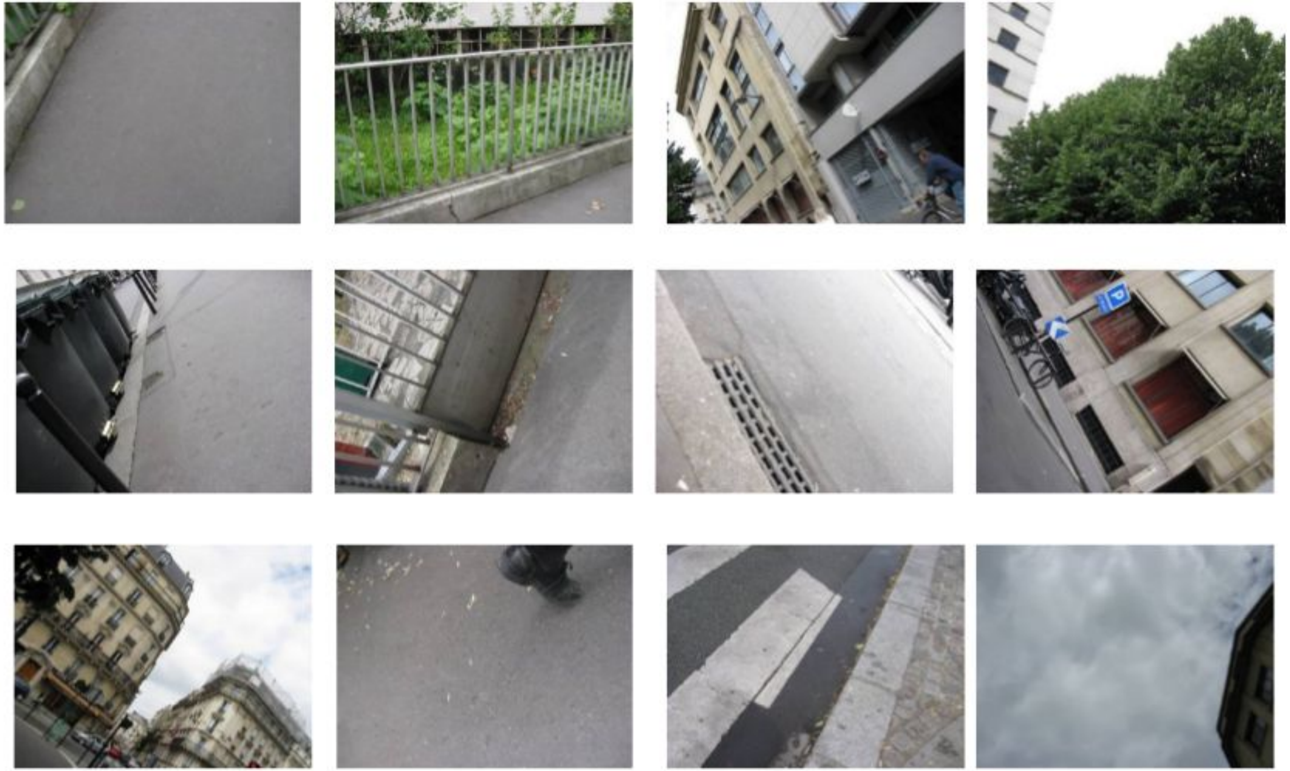
Visual Data is Inherently Biased

- Internet is a tremendous repository of visual data (Flickr, YouTube, Picassa, etc)
- But it's <u>not</u> random samples of visual world

Flickr Paris



Sampled Alyosha Efros's Paris



Sampling Bias

- People like to take pictures on vacation

Photographer Bias

- People want their pictures to be recognizable and/or interesting

vs.

Social Bias

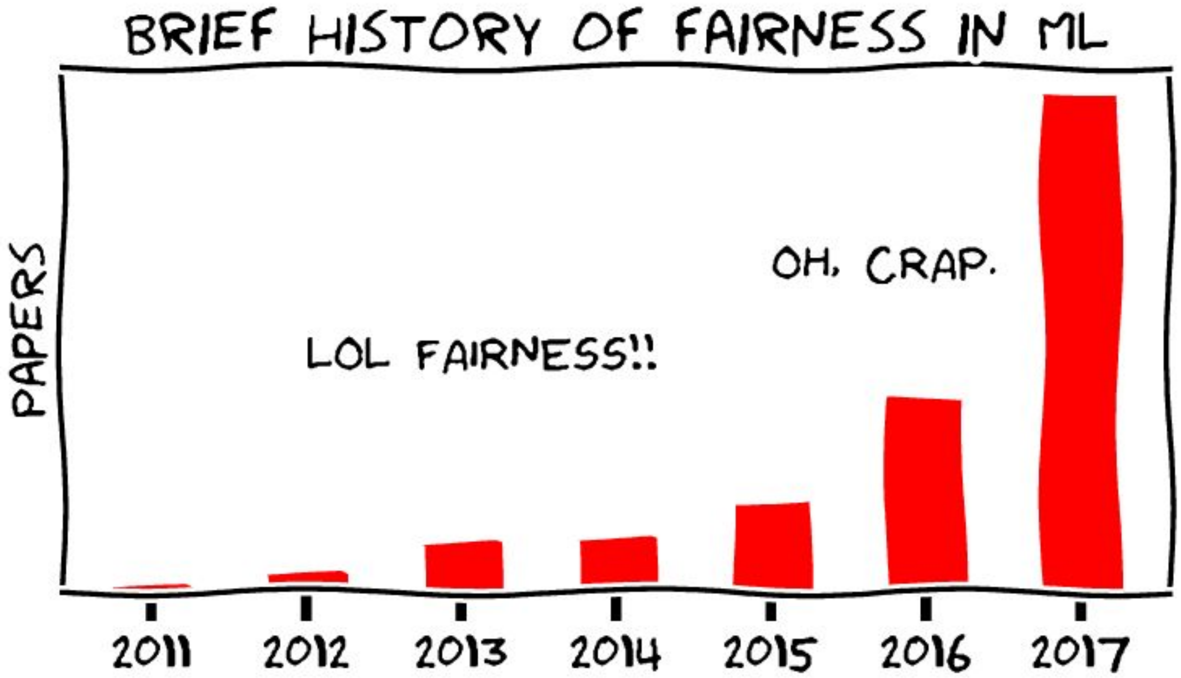
Little Leaguer

Kids with Santa

The Graduate

Newlyweds

“100 Special Moments” by Jason Salavon


BRIEF HISTORY OF FAIRNESS IN ML
PAPERS
LOL FAIRNESS!!
OH, CRAP.
2011
2012
2013
2014
2015
2016
2017

Where does bias come from?

Collection:

- Demographic, geographic, behavioral, temporal biases,

Measurement:

- What do we choose to measure? How do we measure (e.g., *grit*)?

Pre-existing biases

- gender roles in text and images, racial stereotypes in language

To learn more about bias in AI:

- ?


LEVEL 0
There are no autonomous features.

LEVEL 0

There are no autonomous features.

LEVEL 1

These cars can handle one task at a time, like automatic braking.

LEVEL 0

There are no autonomous features.

LEVEL 1

These cars can handle one task at a time, like automatic braking.

LEVEL 2

These cars would have at least two automated functions.

LEVEL 0

There are no autonomous features.

LEVEL 1

These cars can handle one task at a time, like automatic braking.

LEVEL 2

These cars would have at least two automated functions.

LEVEL 3

These cars handle "dynamic driving tasks" but might still need intervention.

SOURCE: SAE International

LEVEL 0

There are no autonomous features.

LEVEL 1

These cars can handle one task at a time, like automatic braking.

LEVEL 2

These cars would have at least two automated functions.

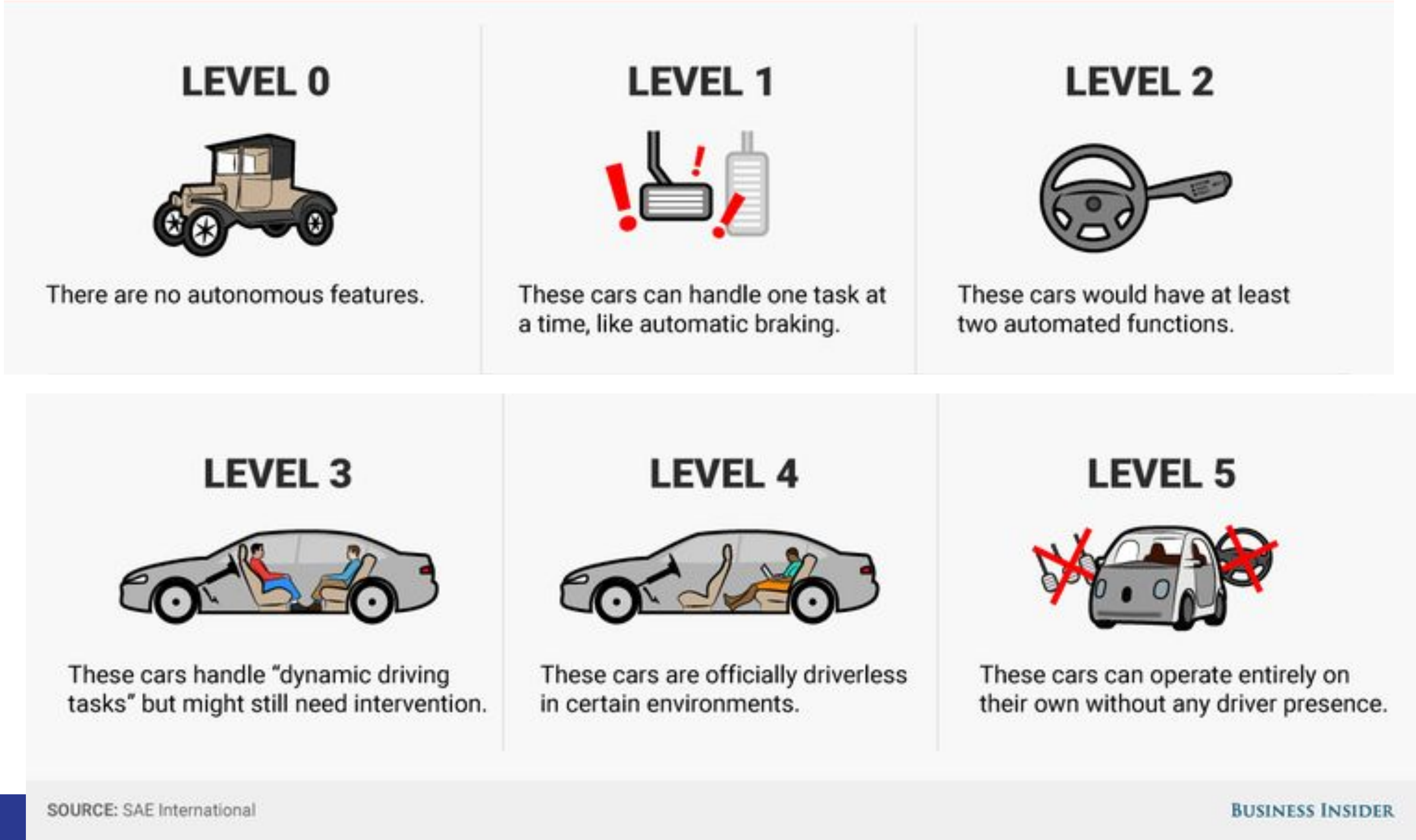
LEVEL 3

These cars handle "dynamic driving tasks" but might still need intervention.

LEVEL 4

These cars are officially driverless in certain environments.

SOURCE: SAE International


LEVEL 0
There are no autonomous features.
LEVEL 1
These cars can handle one task at a time, like automatic braking.
LEVEL 2
These cars would have at least two automated functions.
LEVEL 3
These cars handle "dynamic driving tasks" but might still need intervention.
LEVEL 4
These cars are officially driverless in certain environments.
LEVEL 5
These cars can operate entirely on their own without any driver presence.
SOURCE: SAE International
BUSINESS INSIDER

Technology: More Wealth-- but More Inequality