COMP250: Artificial Intelligence

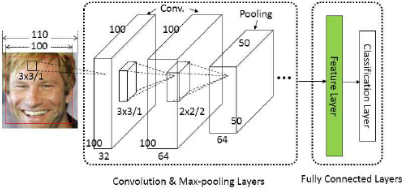
11: Deep Learning

Deep learning

Deep learning

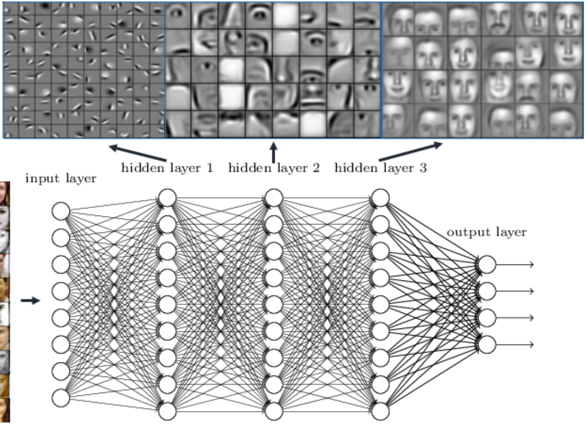
- Basically, the use of large ANNs with **many layers**
- Often uses **large training sets**
- Training often uses powerful **GPUs** — many times faster than training on the CPU

Convolutional Neural Networks (ConvNets)



- Layers are **2D arrays**
- Neurons in convolutional layers are only connected to nearby neurons
- There are also fully connected layers

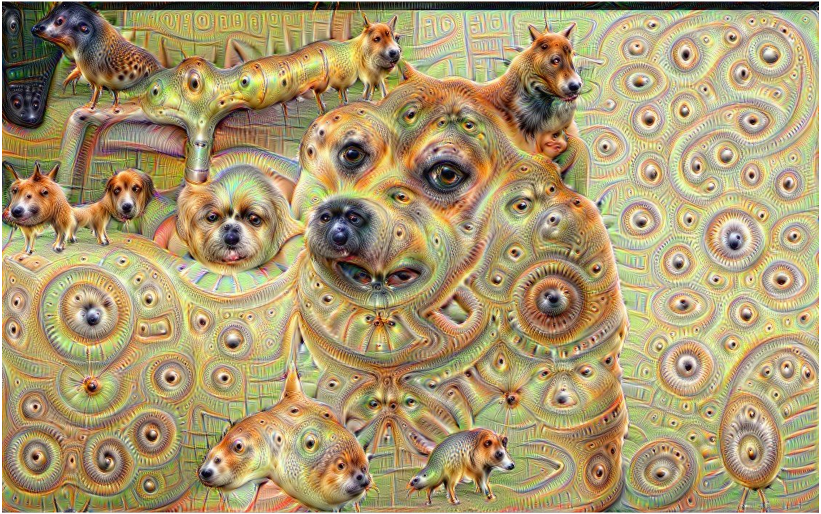
Deep neural networks learn hierarchical feature representations



DeepDream

- Train a ConvNet to recognise something (e.g. faces, objects, animals)
- Run the network in “reverse”
  - Adjust the image (e.g. via gradient ascent) so that it is more strongly recognised by the network

DeepDream

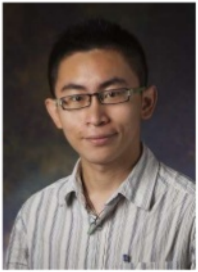


Style transfer

- Train a ConvNet to recognise a particular artistic style
- Run the network in “reverse” on an input image
    - Adjust the image (e.g. via gradient ascent) so that it is more strongly recognised by the network

Style transfer

Source image (**Style**)



Target image (**Content**)

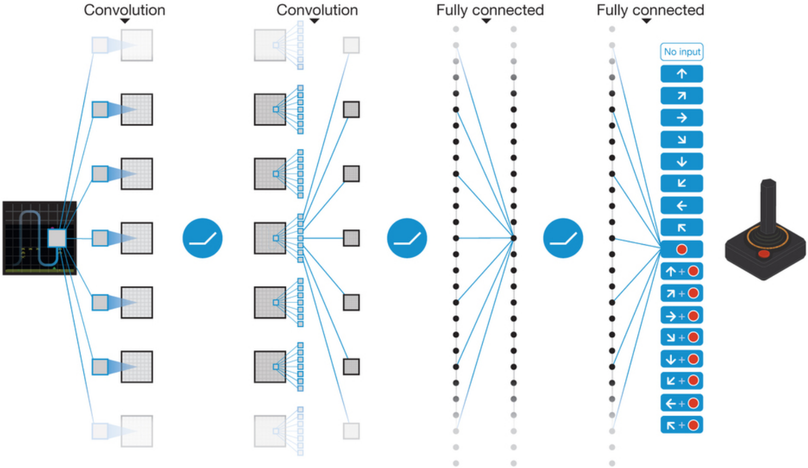
Output (deepart)

A Neural Algorithm of Artistic Style [Gatys et al. 2015]

Generative Adversarial Networks (GANs)

- Two ANNs trained in parallel
  - One to generate “fake” artefacts
  - One to distinguish “real” from “fake”
- `http://research.nvidia.com/publication/2017-10_Progressive-Growing-of`

Learning to play Atari games (Mnih et al, 2015)



AlphaGo (Silver et al, 2017)

- MCTS with ANNs for move pruning, simulation playouts and state evaluation
- ANNs trained on both expert human matches and self-play (reinforcement learning)
- Defeated Lee Sedol, world Go champion

AlphaZero (Silver et al, 2018)

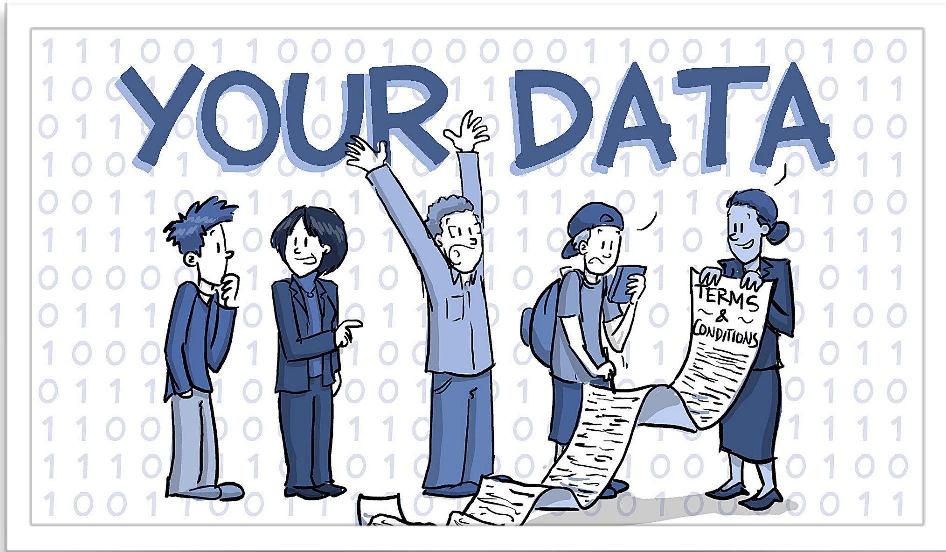
- Similar MCTS+ANN architecture to AlphaGo
- Trained by reinforcement learning (self-play) only
- After only 9 hours* of training, defeated Stockfish (one of the strongest chess programs available) in a 100-match tournament
  - * On a cluster of 5000 of Google's custom Tensor Processing Units
- Stockfish is based on decades of research by expert chess players and AI programmers — AlphaZero started from no chess-specific knowledge whatsoever (other than the rules of the game)

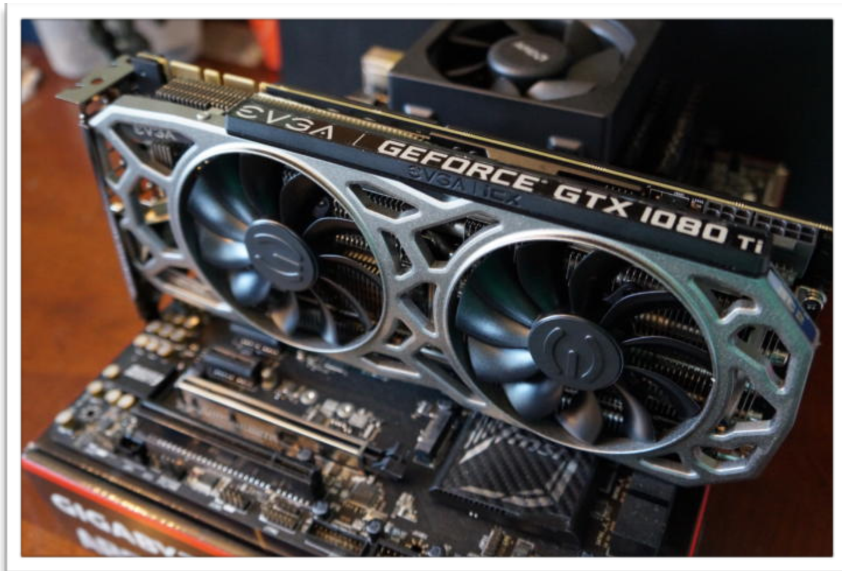
Deep learning for PCG

https://www.youtube.com/watch?v=3wcpLwvBTYo

AI and Society

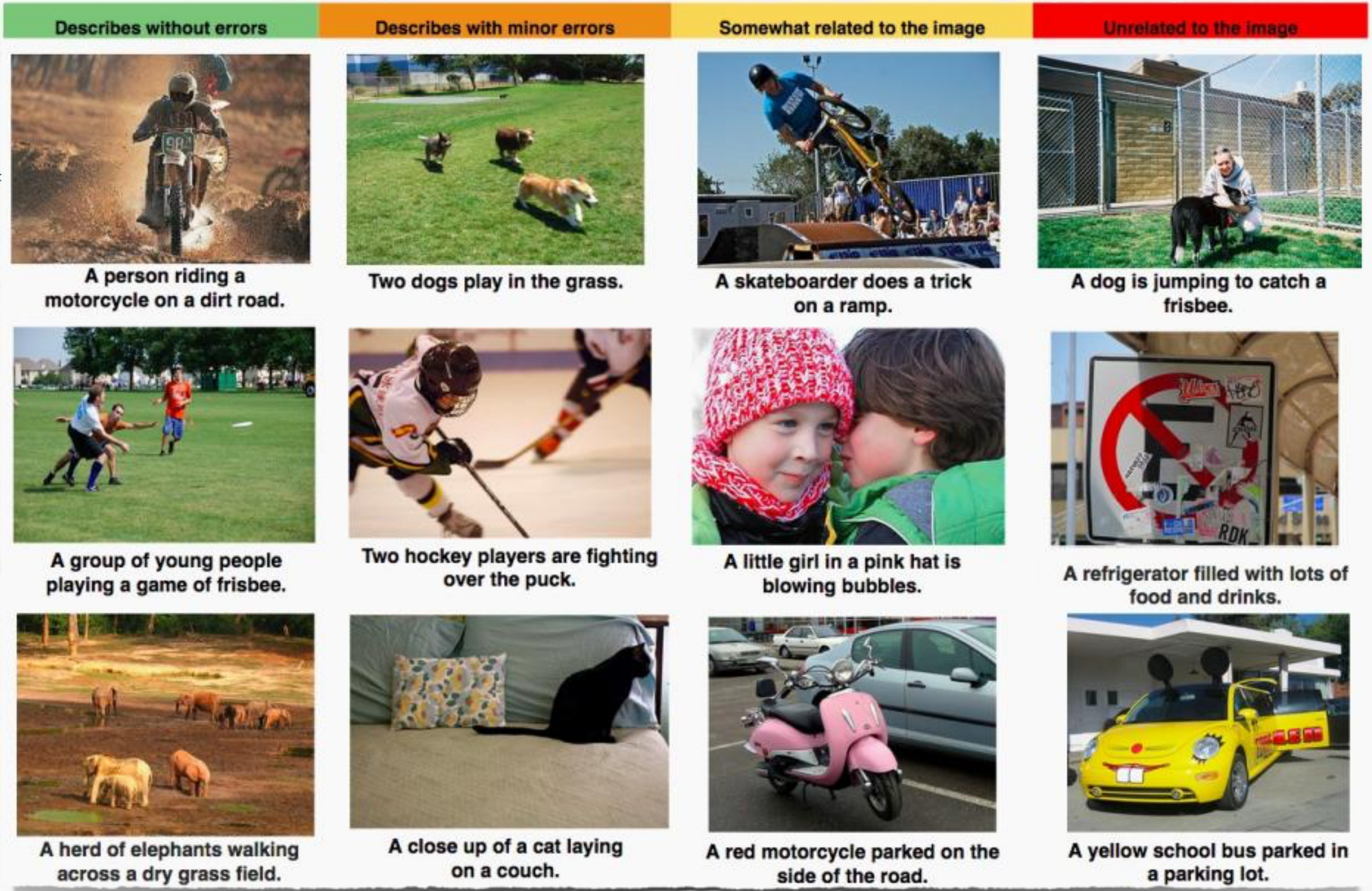
Why now?





What is machine learning good for?

- Discovering patterns in data
- Making inferences or predictions based on these patterns


Describes without errors
Describes with minor errors
Somewhat related to the image
Unrelated to the image
A person riding a motorcycle on a dirt road.
Two dogs play in the grass.
A skateboarder does a trick on a ramp.
A dog is jumping to catch a frisbee.
A group of young people playing a game of frisbee.
Two hockey players are fighting over the puck.
A little girl in a pink hat is blowing bubbles.
A refrigerator filled with lots of food and drinks.
A herd of elephants walking across a dry grass field.
A close up of a cat laying on a couch.
A red motorcycle parked on the side of the road.
A yellow school bus parked in a parking lot.

Patterns in data

- These examples are based on image data
- Equally applicable to other forms of data
  - Creative artefacts: audio, video, narrative, architecture, game mechanics, …
  - “Real-world” data: natural language, commerce, medicine, …

Caution required

- ML is just number crunching and pattern recognition
- The machine is not “thinking”!
- It can’t tell the difference between the patterns we wanted to find, and biases that might exist in the training data
- It can’t follow a code of ethics or take responsibility for its own uses and actions

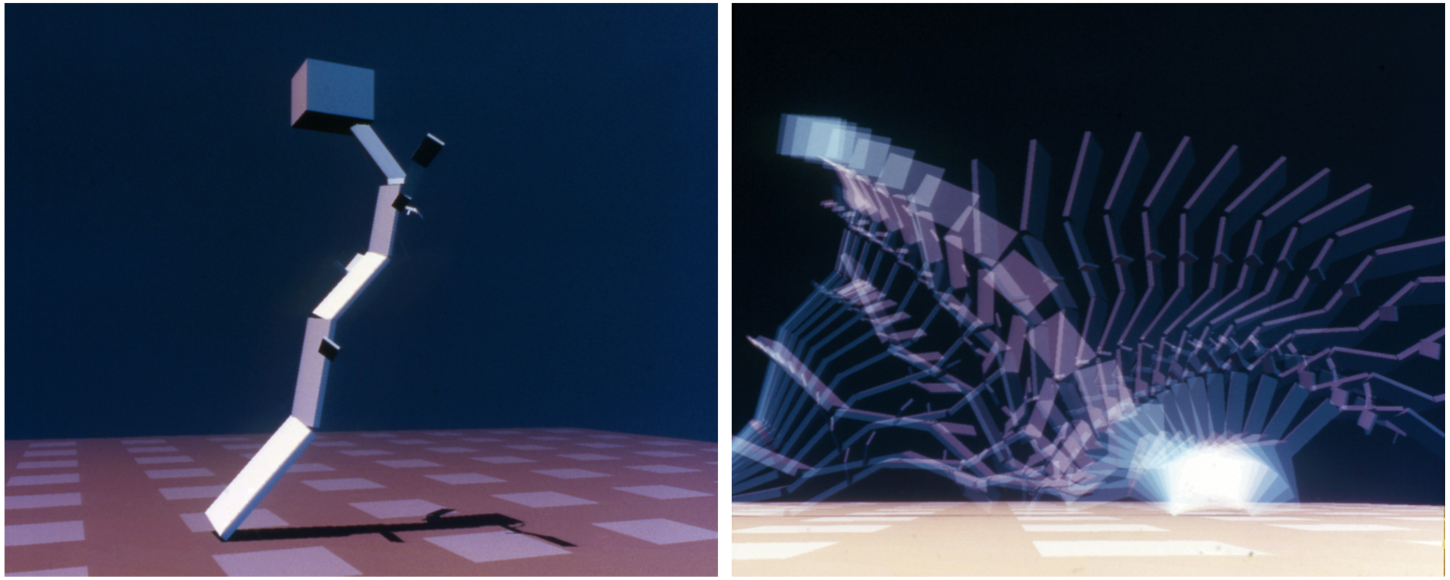
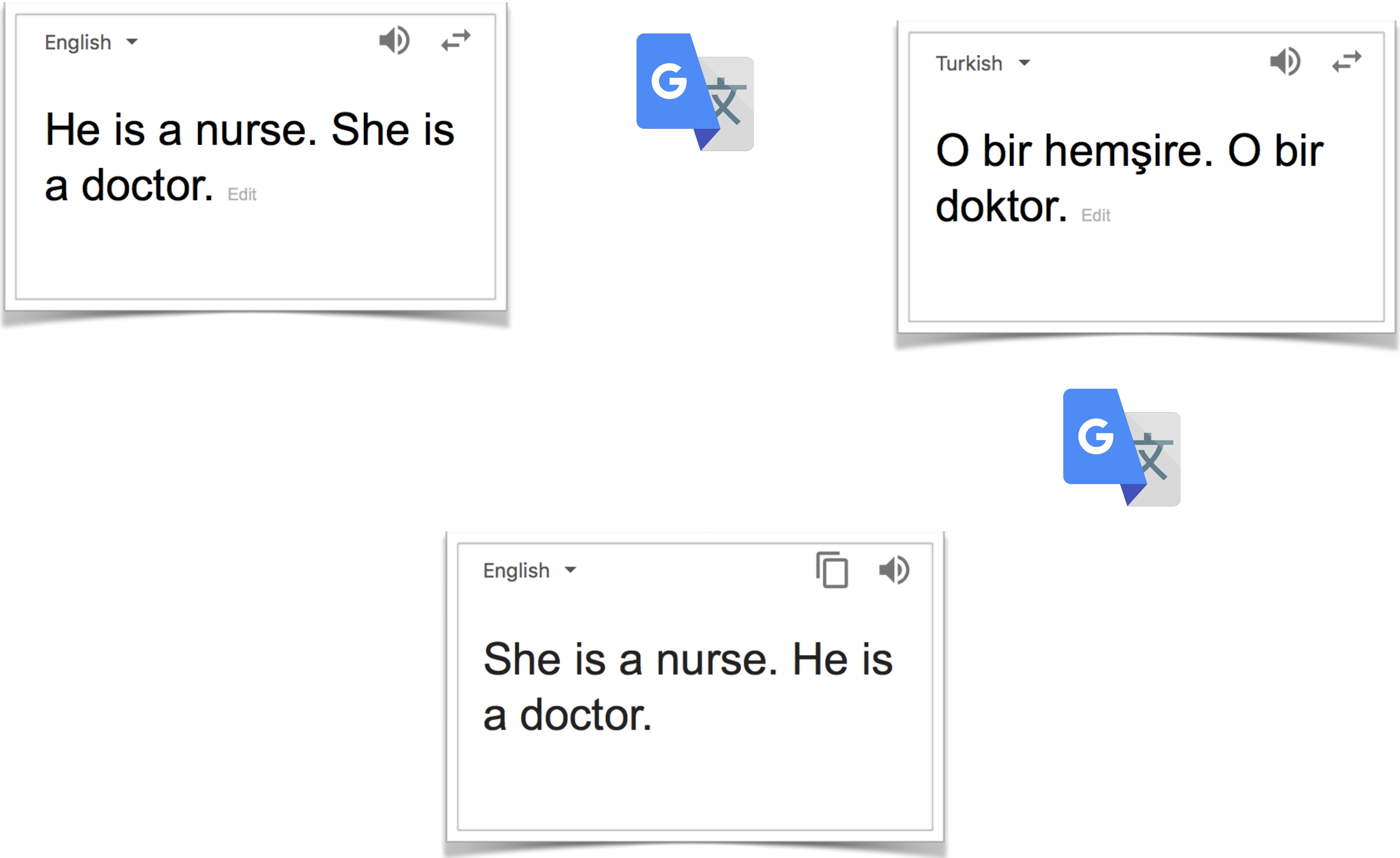


Figure 1. Exploiting potential energy to locomote. Evolution discovers that it is simpler to design tall creatures that fall strategically than it is to uncover active locomotion strategies. The left figure shows the creature at the start of a trial and the right figure shows snapshots of the figure over time falling and somersaulting to preserve forward momentum.

"Instead of relying on algorithms, which we can be accused of manipulating for our benefit, we have turned to machine learning, an ingenious way of disclaiming responsibility for anything. Machine learning is like **money laundering for bias**. It's a clean, mathematical apparatus that gives the status quo the aura of logical inevitability. The numbers don't lie."

–Maciej Cegłowski


English
He is a nurse. She is a doctor.
Edit
Turkish
O bir hemşire. O bir doktor.
Edit
English
She is a nurse. He is a doctor.

The New York Times

SUBSCRIBE | LOG IN

Facial Recognition Is Accurate, if You're a White Guy

By Steve Lohr

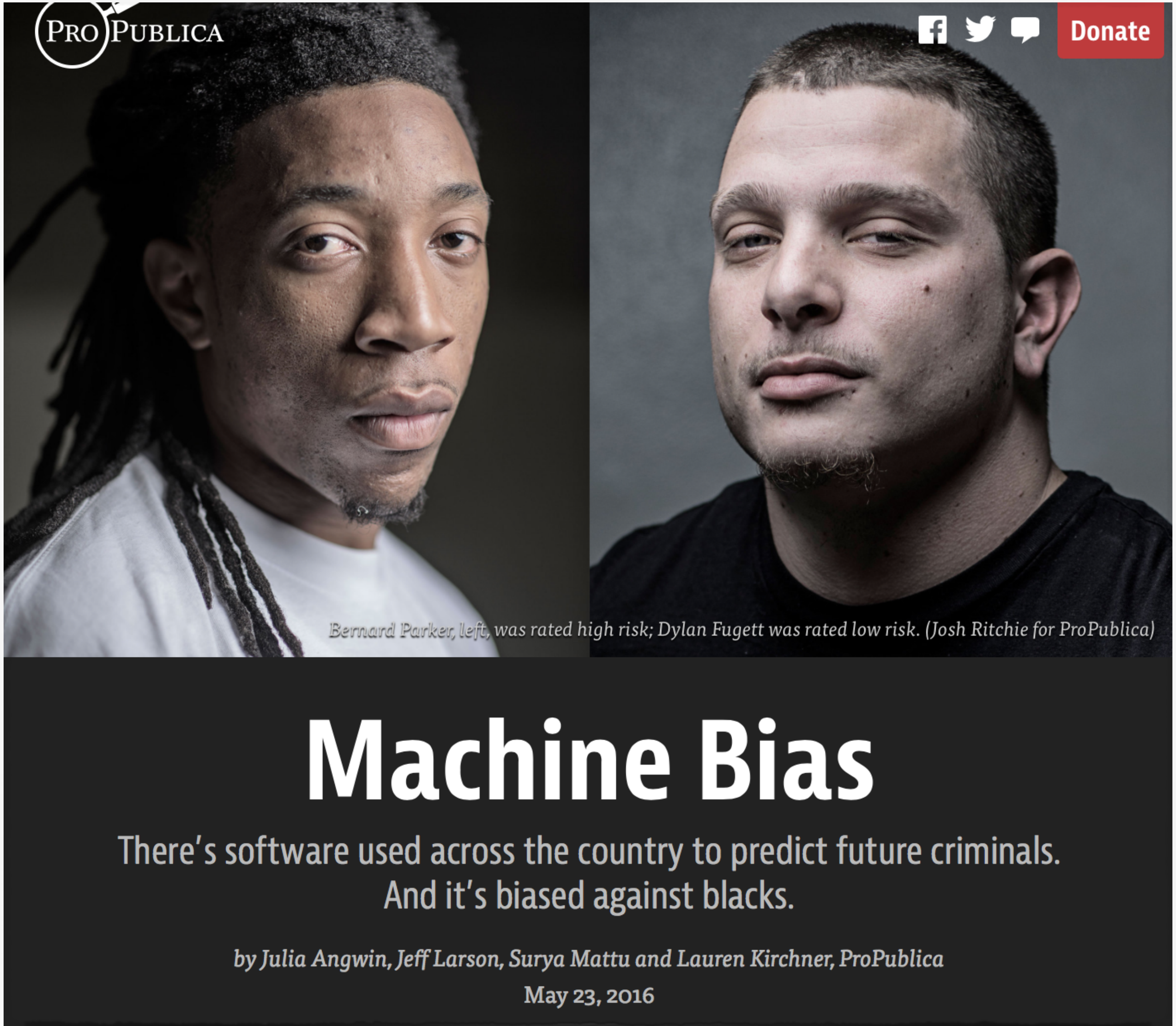
Feb. 9, 2018

Facial recognition technology is improving by leaps and bounds. Some commercial software can now tell the gender of a person in a photograph.

When the person in the photo is a white man, the software is right 99 percent of the time.

But the darker the skin, the more errors arise — up to nearly 35 percent for images of darker skinned women, according to a new study that breaks fresh ground by measuring how the technology works on people of different races and gender.

These disparate results, calculated by Joy Buolamwini, a researcher at the M.I.T. Media Lab, show how some of the biases in the real world can seep into artificial intelligence, the computer systems that inform facial recognition.


ProPublica
Donate

Bernard Parker, left, was rated high risk; Dylan Fugett was rated low risk. (Josh Ritchie for ProPublica)

Machine Bias

There's software used across the country to predict future criminals. And it's biased against blacks.

by Julia Angwin, Jeff Larson, Surya Mattu and Lauren Kirchner, ProPublica

May 23, 2016

BUSINESS NEWS OCTOBER 10, 2018 / 4:12 AM / 23 DAYS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

Jeffrey Dastin

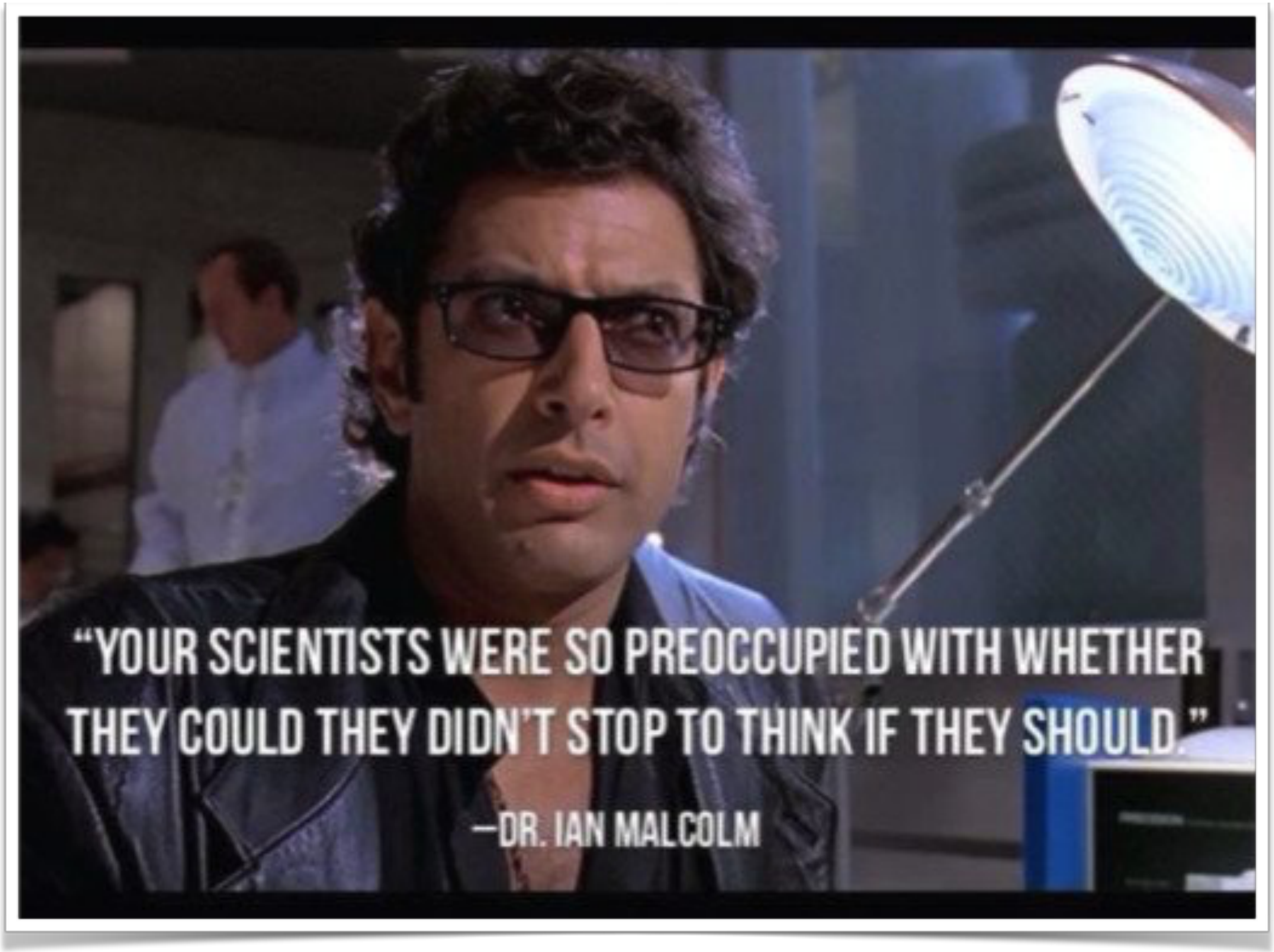
8 MIN READ

SAN FRANCISCO (Reuters) - Amazon.com Inc's (AMZN.O) machine-learning specialists uncovered a big problem: their new recruiting engine did not like women.

That is because Amazon's computer models were trained to vet applicants by observing patterns in resumes submitted to the company over a 10-year period. Most came from men, a reflection of male dominance across the tech industry.

In effect, Amazon's system taught itself that male candidates were preferable. It penalized resumes that included the word "women's," as in "women's chess club captain." And it downgraded graduates of two all-women's colleges, according to people familiar with the matter. They did not specify the names of the schools.

Responsible use of AI



Automated Inference on Criminality using Face Images

Xiaolin Wu
McMaster University
Shanghai Jiao Tong University
xwu510@gmail.com

Xi Zhang
Shanghai Jiao Tong University
zhangxi_19930818@sjtu.edu.cn

Abstract

We study, for the first time, automated inference on criminality based solely on still face images, which is free of any biases of subjective judgments of human observers. Via supervised machine learning, we build four classifiers (logistic regression, KNN, SVM, CNN) using facial images of 1856 real persons controlled for race, gender, age and facial expressions, nearly half of whom were convicted criminals, for discriminating between criminals and non-criminals. All four classifiers perform consistently well and empirically establish the validity of automated face-induced inference on criminality, despite the historical controversy surrounding this line of enquiry. Also, some discriminat-

management science, criminology, etc.

In all cultures and all periods of recorded human history, people share the belief that the face alone suffices to reveal innate traits of a person. Aristotle in his famous work Prior Analytics asserted, "It is possible to infer character from features, if it is granted that the body and the soul are changed together by the natural affections". Psychologists have known, for as long as a millennium, the human tendency of inferring innate traits and social attributes (e.g., the trustworthiness, dominance) of a person from his/her facial appearance, and a robust consensus of i
ences . These are the facts found through
[3, 39, 5, 6, 10, 26, 27, 34, 32].

Independent of the validity of pedest

(a) Three samples in criminal ID photo set S_c.

(b) Three samples in non-criminal ID photo set S_n

Figure 1. Sample ID photos in our data set.

Deep neural networks are more accurate than humans at detecting sexual orientation from facial images

Yilun Wang, Michal Kosinski

Graduate School of Business, Stanford University, Stanford, CA94305, USA

michalk@stanford.edu

Abstract

We show that faces contain much more information about sexual orientation than can be perceived and interpreted by the human brain. We used deep neural networks to extract features from 35,326 facial images. These features were entered into a logistic regression aimed at classifying sexual orientation. Given a single facial image, a classifier could correctly distinguish between gay and heterosexual men in 81% of cases, and in 71% of cases for women. Human judges achieved much lower accuracy: 61% for men and 54% for women. The accuracy of the algorithm increased to 91% and 83%, respectively, given five facial images per person. Facial features employed by the classifier included both fixed (e.g., nose shape) and transient facial features (e.g., grooming style). Consistent with the prenatal hormone theory of sexual orientation, gay men and women tended to have gender-atypical facial morphology, expression, and grooming styles. Prediction models aimed at gender alone allowed for detecting gay males with 57% accuracy and gay females with 58% accuracy. Those findings advance our understanding of the origins of sexual orientation and the limits of human perception. Additionally, given that companies and governments are increasingly using computer vision algorithms to detect people's intimate traits, our findings expose a threat to the privacy and safet of gay men and women.

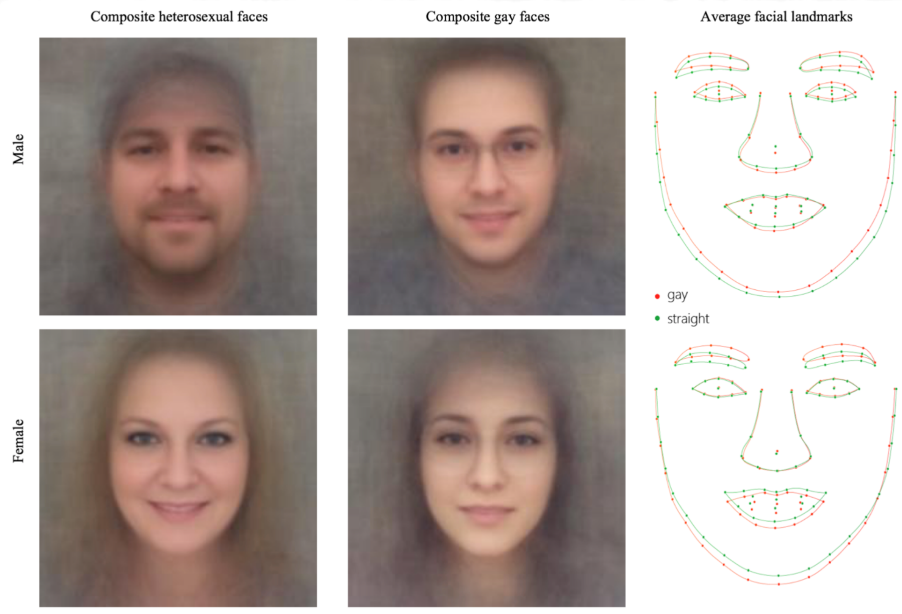


Figure 4. Composite faces and the average facial landmarks built by averaging faces classified as most and least likely to be gay.

10
EVELOPING STORY
WHO'S TO BLAME FOR CRASH?
abc 15
ARIZONA
DRIVE TIME
SPORTS
NHL : FLAMES 2 COYOTES 5 FINAL
4:33
53°

Bank of England chief economist warns on AI jobs threat

Kamal Ahmed
Economics editor
@bbckamal

20 August 2018 | Share

The chief economist of the Bank of England has warned that the UK will need a skills revolution to avoid "large swathes" of people becoming "technologically unemployed" as artificial intelligence makes many jobs obsolete.

AAMCNEWS

RESEARCH

Tuesday, November 27, 2018 | by Fran Kritz, special to *AAMCNews*

Progress in predicting suicide

For decades, the goal of identifying those most at risk of ending their lives has remained elusive. Finally, it may be within sight.

Vadim Zipunnikov
@vadimZip

Follow

Separating 17 healthy controls from 17 suicidal ideators with the number of features = 30 stimuli/concepts times the (unspecified) number of fMRI-based voxel clusters and getting 91% of accuracy looks like a good case for another overfitted ML model.@f2harrell @fMRIstats @bcaffo

The team scanned the brains of 34 people, half of whom recently reported suicidal thoughts, while they contemplated several emotion-laden words. The scans revealed that certain words — “death” and “cruelty,” for example — caused more types of brain activity in suicidal people than in control subjects.

Next, the researchers fed the scans into a computer and, using machine-learning techniques, it developed an algorithm for recognizing which brain patterns reflect suicidality. The computer did quite well, according to a recent *Nature Human Behaviour* study. In fact, it was 91% accurate in identifying people who had considered suicide. Perhaps even more helpful, it was 94% accurate in separating people who had considered but didn’t attempt suicide from those who had actually attempted suicide.

Conclusion

- AI isn’t magic, it’s just number crunching
- Human-level general intelligence is the realm of sci-fi (at least for now)
- AI is a tool — we are responsible for how it is used