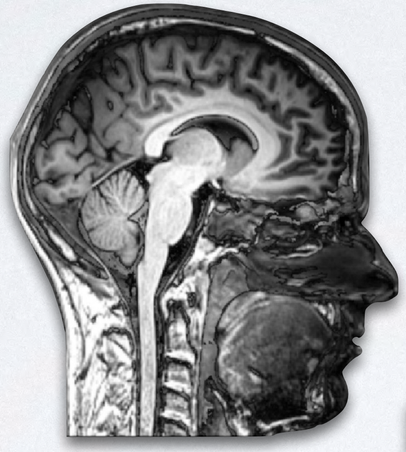
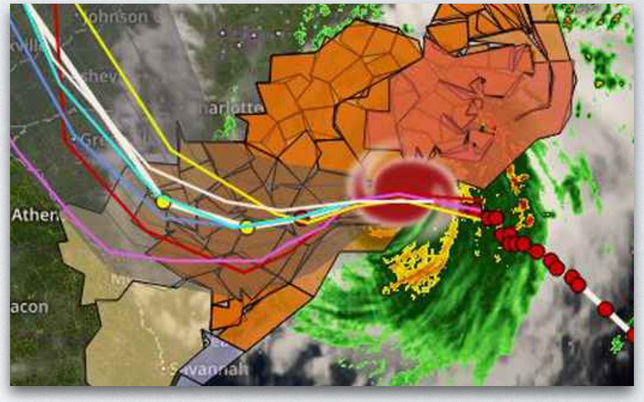
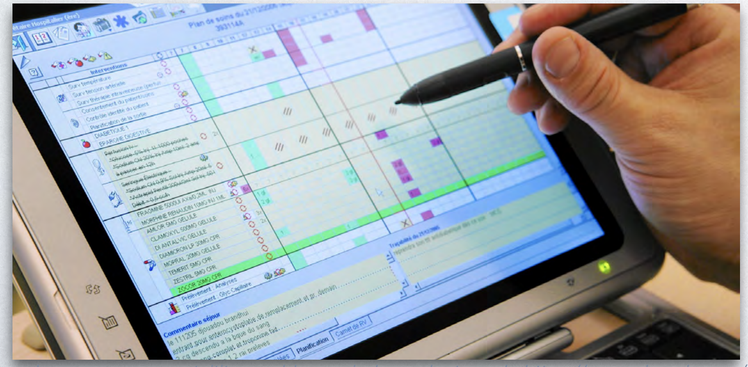
AI: Challenges, Opportunities, and Partnerships with Academia

Rebecca Willett, University of Chicago







archpaper.com, smartcitiesworld.com, independent.co.uk, https://magazine.cim.org/en/technology/the-mechanics-of-a-new-era-in-maintenance-en, https://cvgcares.com/apple-watch-vs-fitbit-which-smartwatch-is-best-for-heart-rate-monitoring/, https://www.technologynetworks.com/tn/product-surveys/2018-lab-water-survey-306836, https://precisionagricultu.re/gps-drones-robots-precision-agriculture-farming/, https://www.grubhub.com/, https://www.forbes.com/sites/harlankrumholz/2018/10/05/the-unfortunate-truth-about-getting-your-medical-records/#1e3c0636cb7b

Transformative potential of AI

- boost health and quality of life
- provide lifelong education and training
- reinvent business innovation and competitiveness
- accelerate scientific discovery and technical innovation
- expand evidence-driven social opportunity and policy
- transform national defense and security

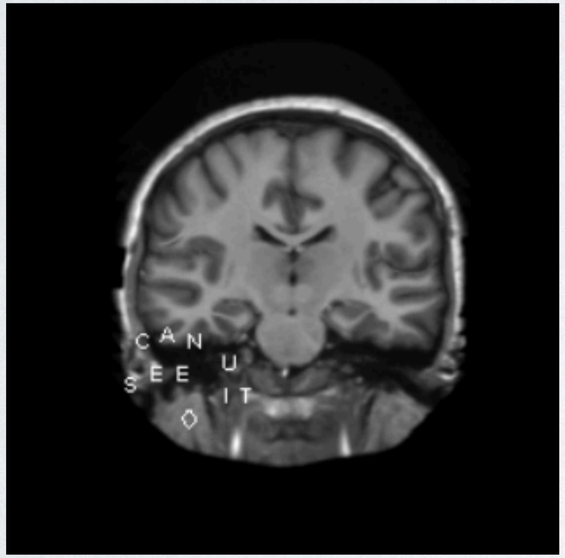
Exploring AI in action

- hunt for subatomic particles
- malware detection
- lip reading
- seasonal forecasting
- automatic translation
- spectrum estimation for wireless communications
- automatic translation
- robot juggling
- epigenetics
- grammar correction
- describe a photo
- renewable energy
- play go
- play chess
- play poker
- help control prosthetic devices
- analyze vast collections of images of galaxies
- facial recognition
- drug discovery
- advanced transportation systems
- compose music
- speech recognition
- mimic artistic style
- design better batteries
- invent recipes
- home automation
- find new disease risk factors
- recommend purchases
- design new materials
- control the power grid
- predictive policing

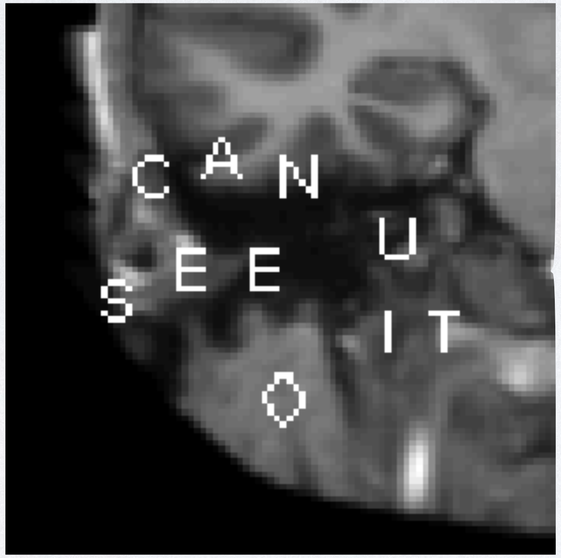
“It's quite obvious that we should stop training radiologists.” — Geoffrey Hinton, an AI luminary, in 2016

“It's quite obvious that we should stop training radiologists.” — Geoffrey Hinton, an AI luminary, in 2016

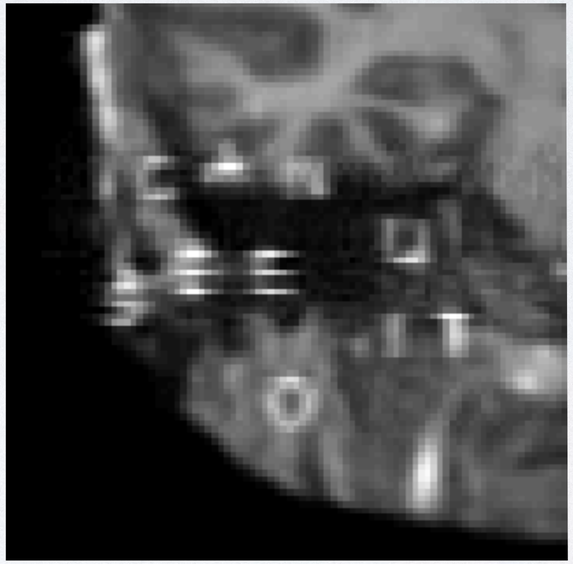
Original



Original (zoomed)



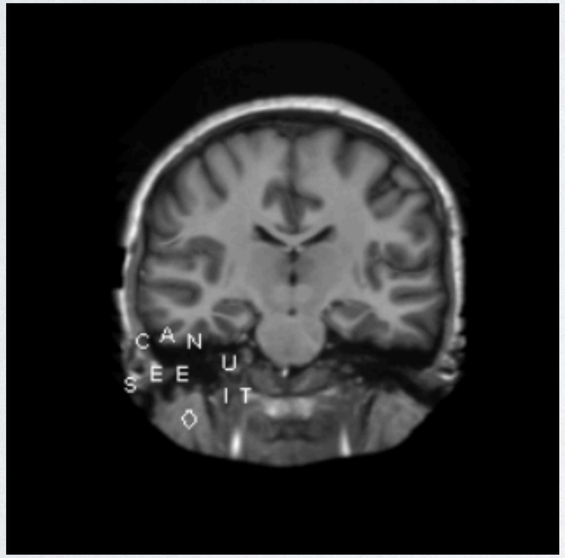
Machine Learning



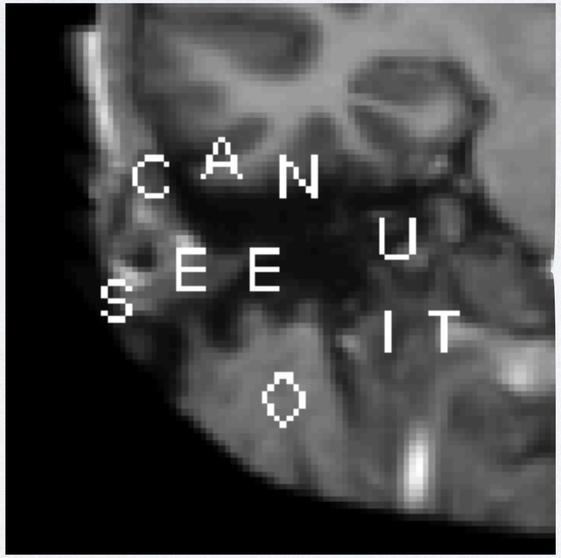
Antun, V., Renna, F., Poon, C., Adcock, B., & Hansen, A. C. (2019).

"It's quite obvious that we should stop training radiologists." — Geoffrey Hinton, an AI luminary, in 2016

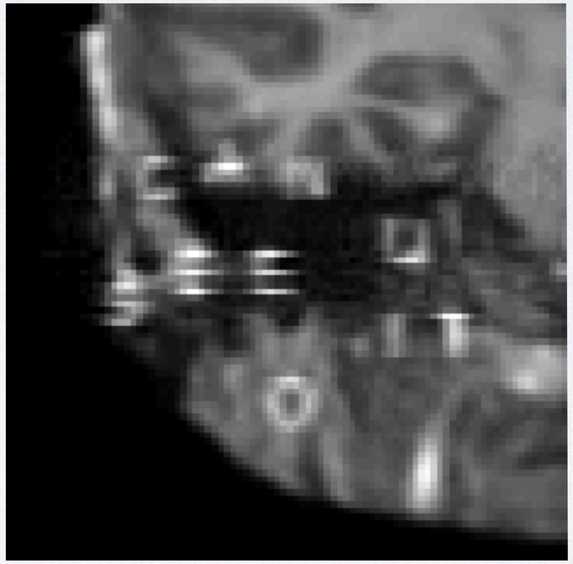
Original



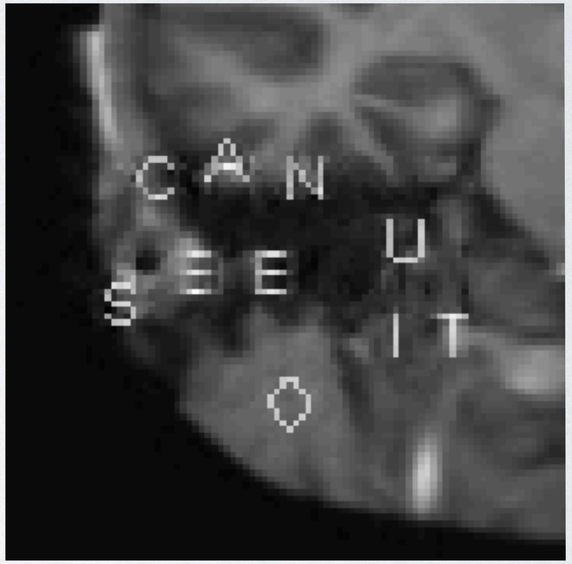
Original (zoomed)



Machine Learning



Classical Method



Antun, V., Renna, F., Poon, C., Adcock, B., & Hansen, A. C. (2019).

Two Viewpoints

AI is Amazing ⇒ Adopt Everywhere!

Focus on deployment and efficiency

More concerned with “could it work?” then “might it fail?”

Trial and error mindset

Benchmark performance → innovation

AI is Curious ⇒ Learn More!

Focus on theory, including accuracy bounds, sampling efficiency, etc.

Concerned with robustness and correctness

The costs of mistakes may be high

Find shortcomings → understanding

Progress depends on both approaches

Robustness and stability

Robustness and stability

No one knows exactly how some methods work:
we cannot predict when outputs will be catastrophically wrong

Explaining and Harnessing Adversarial Examples by Goodfellow et al.

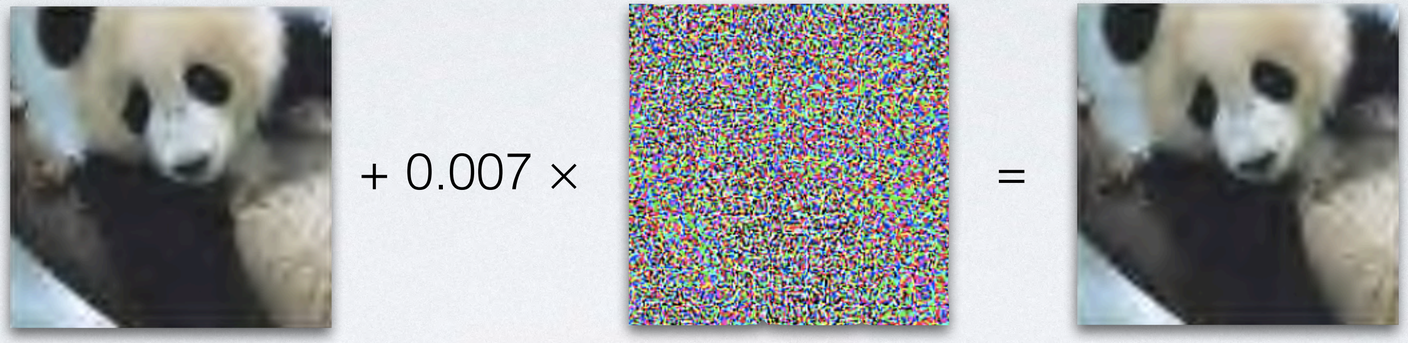
Robustness and stability

“panda”

No one knows exactly how some methods work:
we cannot predict when outputs will be catastrophically wrong

Explaining and Harnessing Adversarial Examples by Goodfellow et al.

Robustness and stability

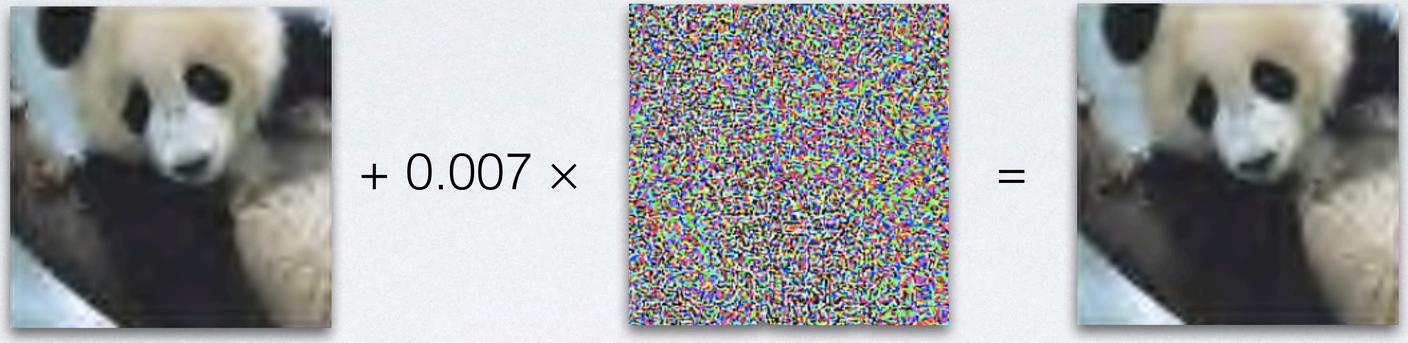


“panda”

No one knows exactly how some methods work:
we cannot predict when outputs will be catastrophically wrong

Explaining and Harnessing Adversarial Examples by Goodfellow et al.

Robustness and stability

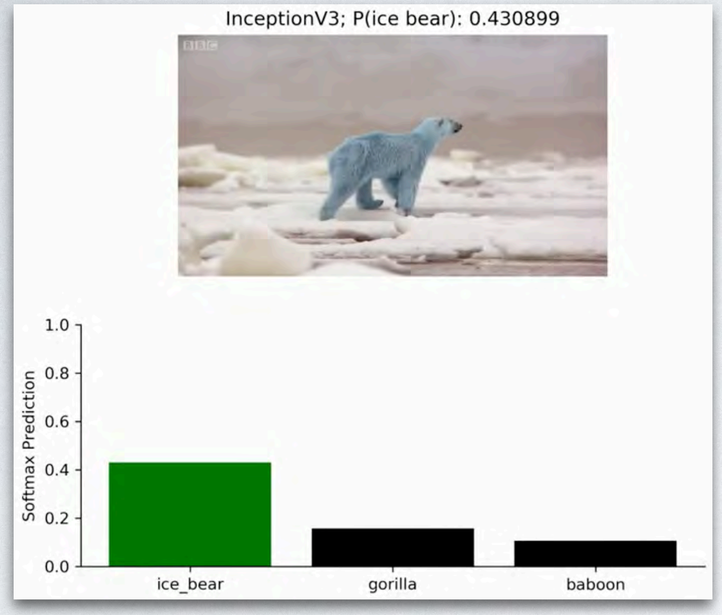


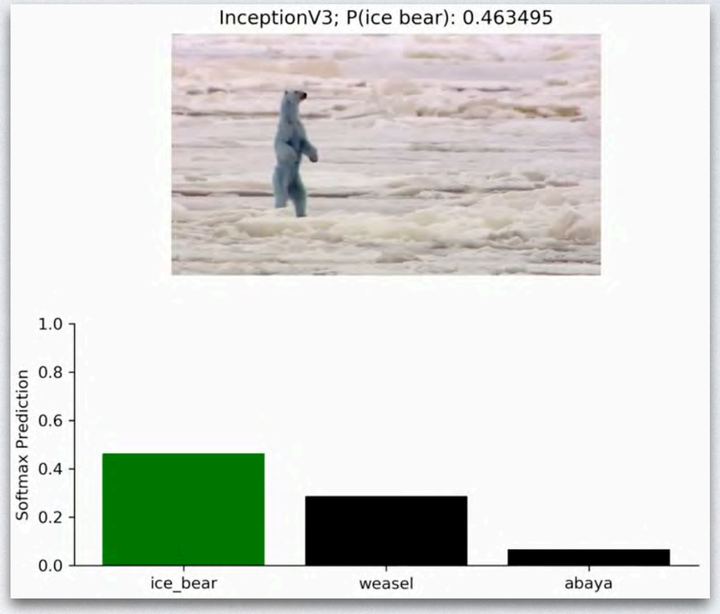
“panda” “gibbon”

No one knows exactly how some methods work:
we cannot predict when outputs will be catastrophically wrong

Explaining and Harnessing Adversarial Examples by Goodfellow et al.

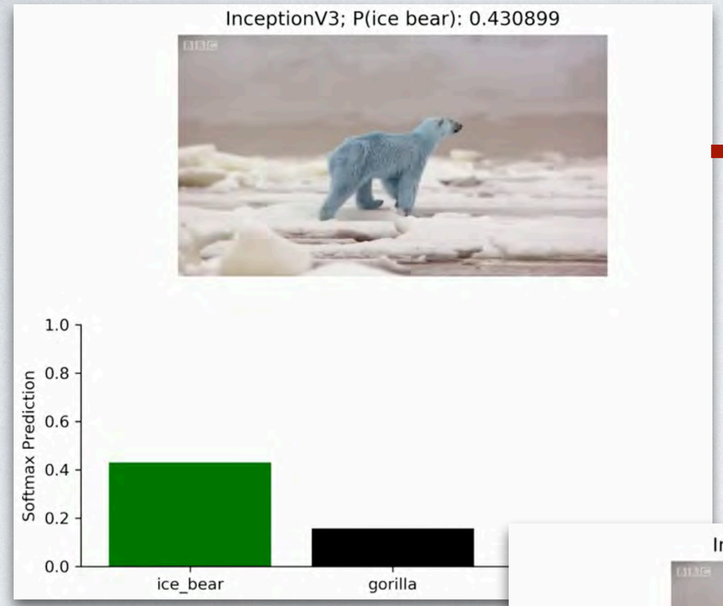
Robustness and stability

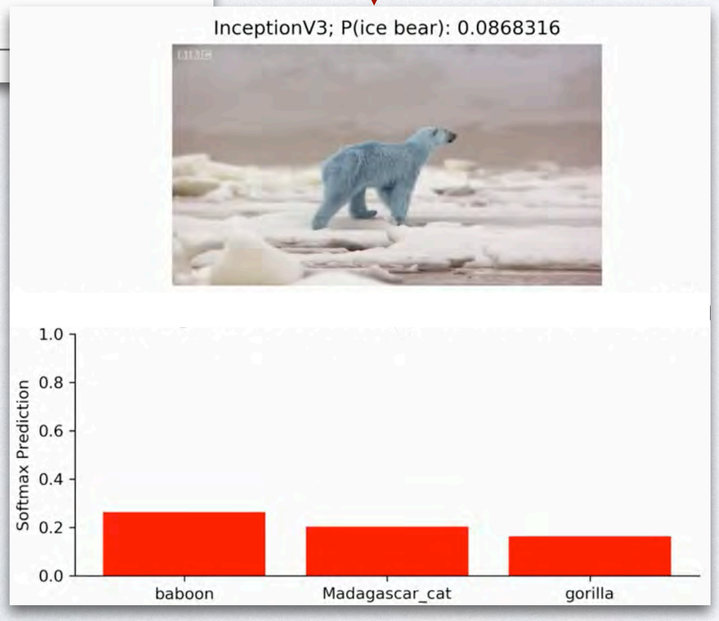


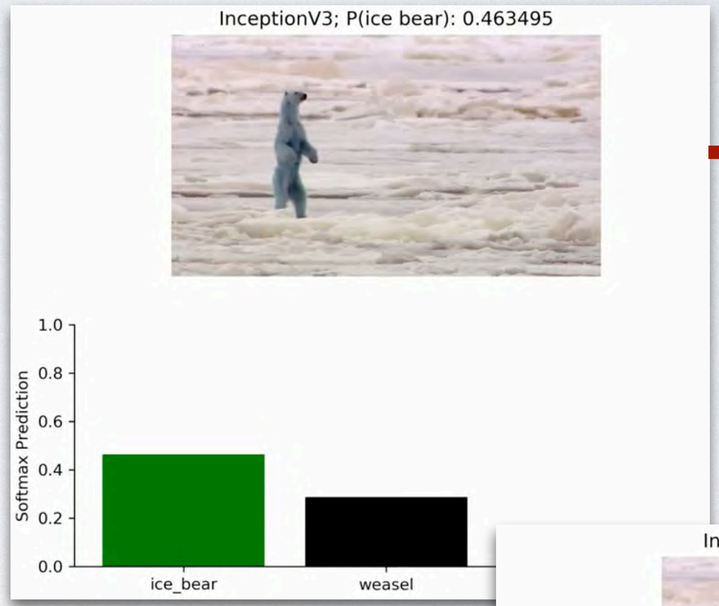


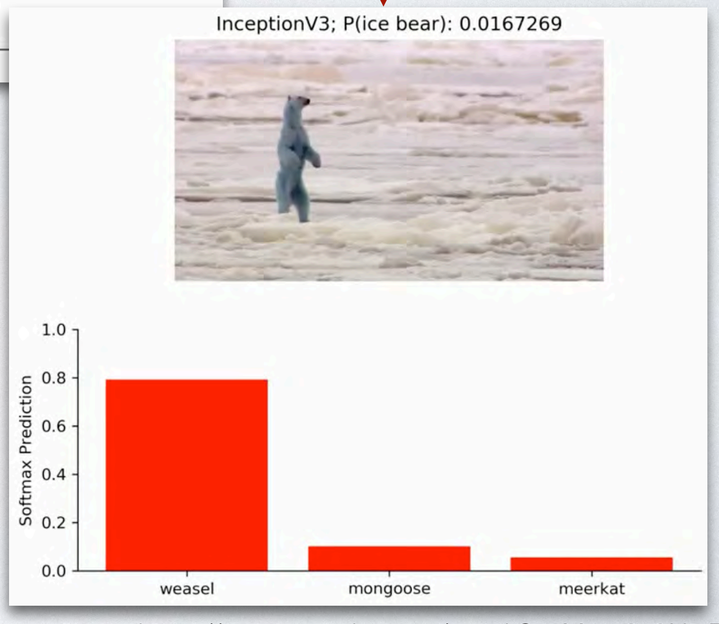
https://www.youtube.com/watch?v=M4ys8c2NtsE

Robustness and stability





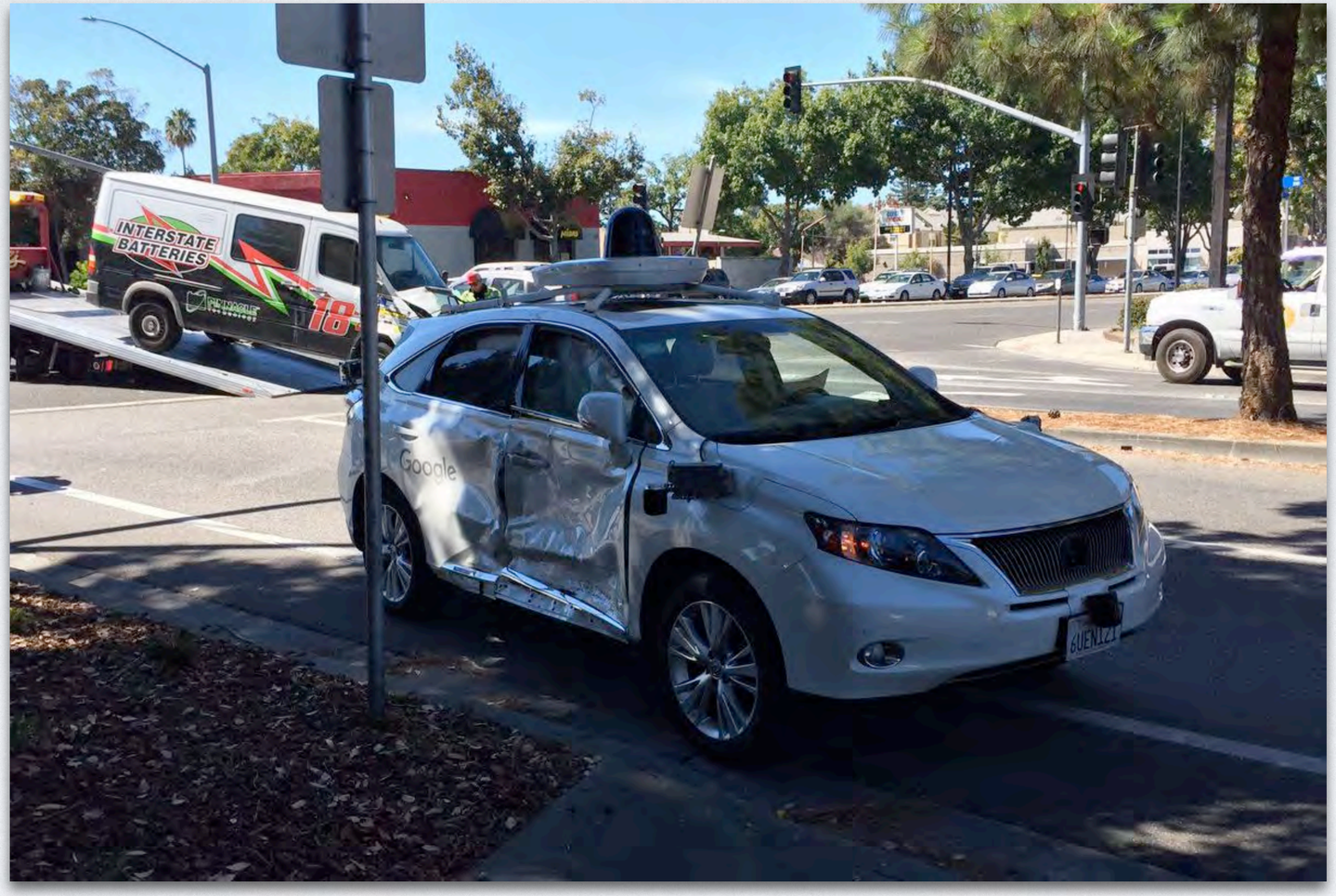




This is not an adversarial example

https://www.youtube.com/watch?v=M4ys8c2NtsE

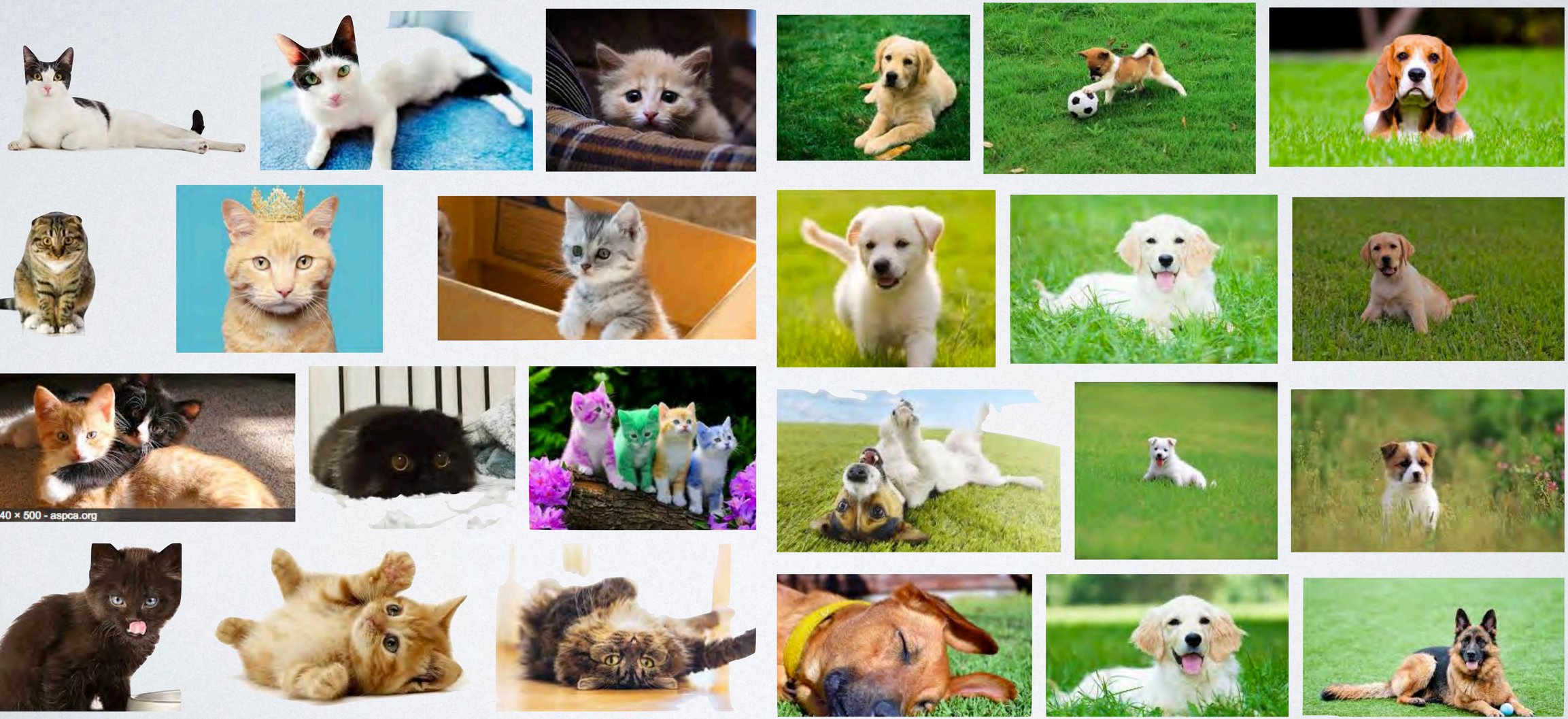
Robustness and stability



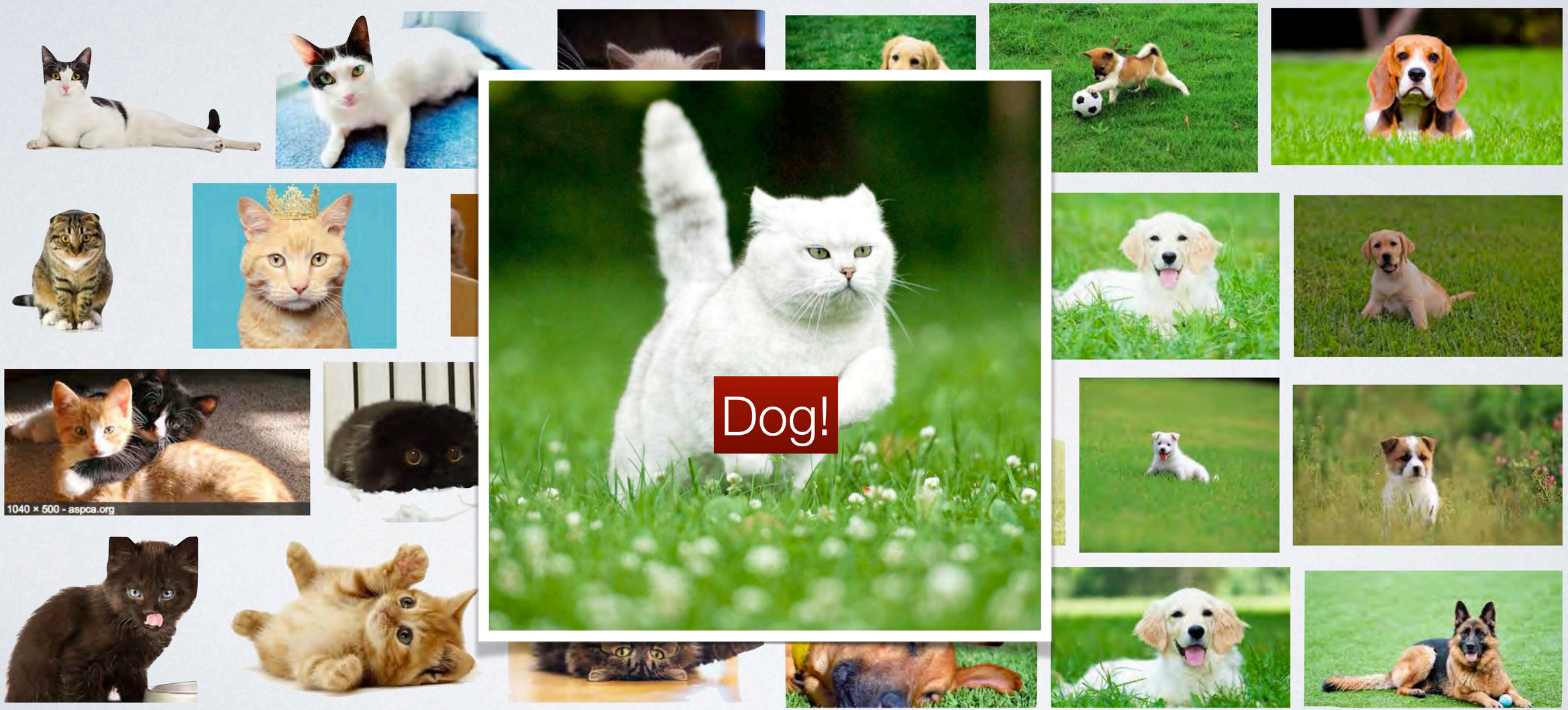
Spurious correlations in training data

Requires large quantities of diverse data

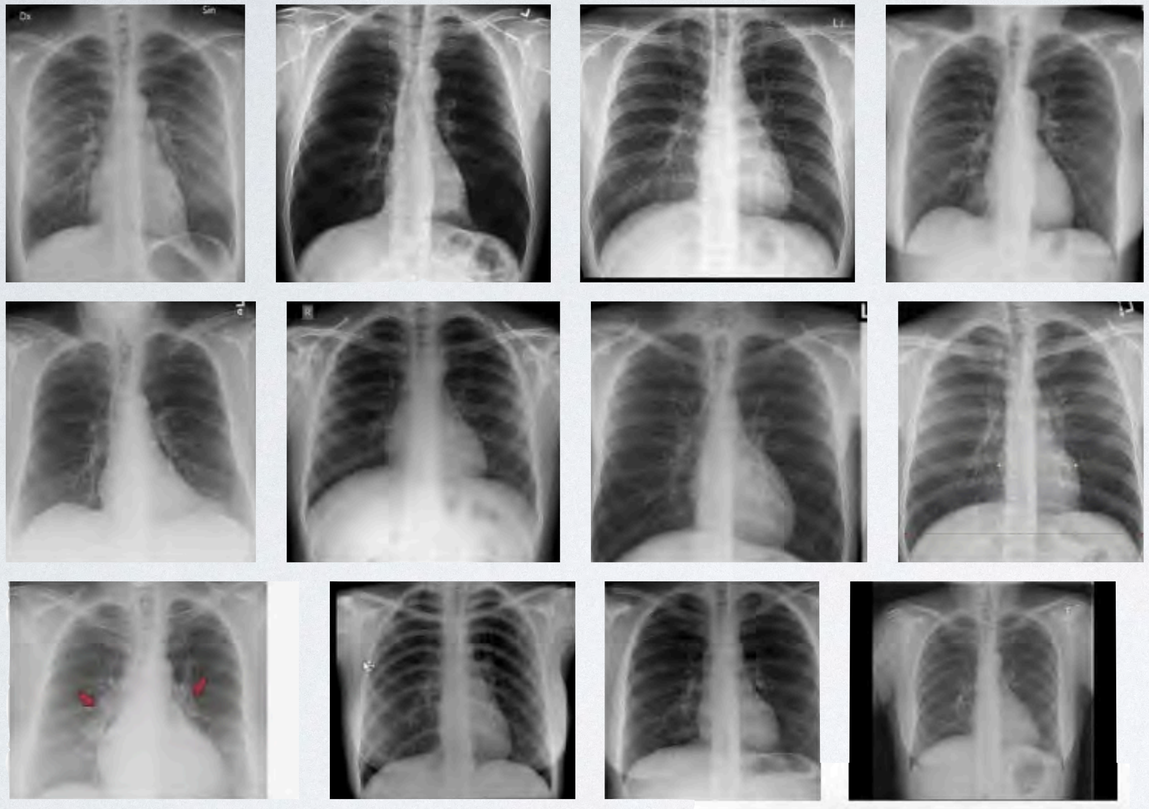
Requires large quantities of diverse data

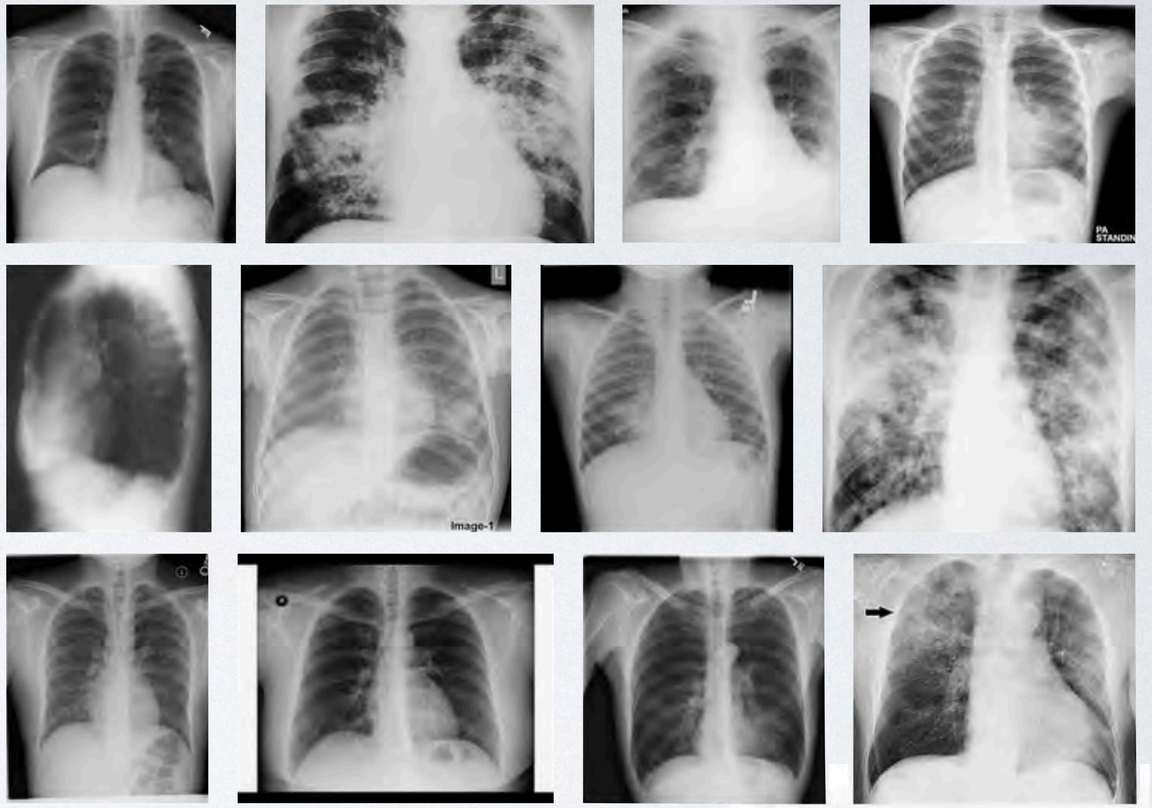


Requires large quantities of diverse data

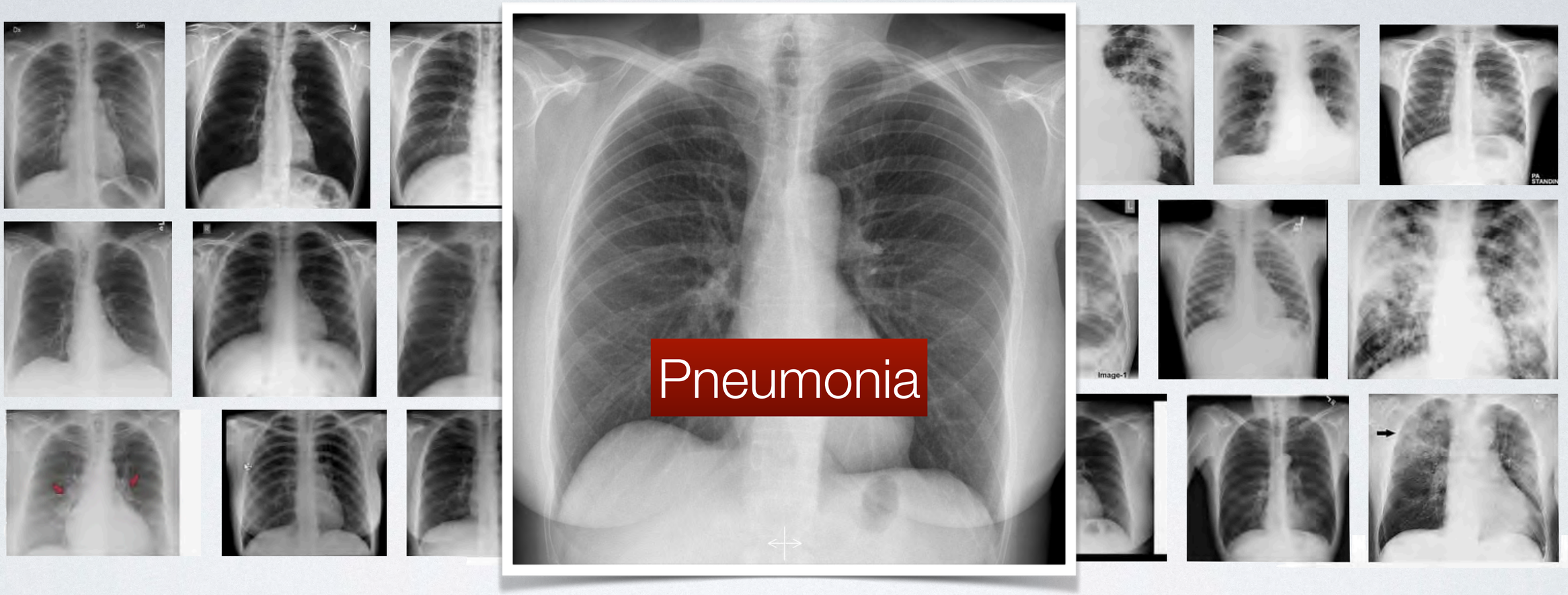


Requires large quantities of diverse data

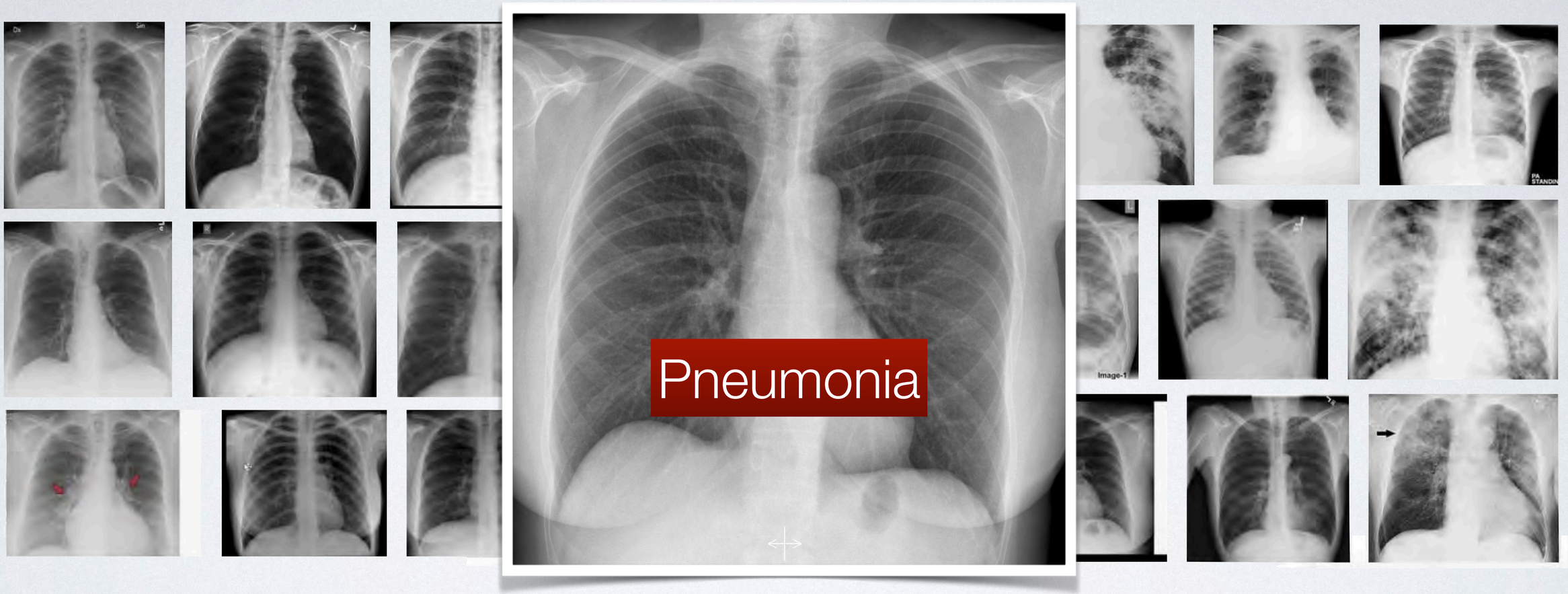




Requires large quantities of diverse data



Requires large quantities of diverse data



"CNNs could rely on subtle differences in acquisition protocol, image processing, or distribution pipeline (e.g., image compression) and overlook pathology." *– Zech et al, 2018*

Requires large quantities of diverse data

Is our data sufficiently diverse for making decisions on healthcare or finance?

"CNNs could rely on subtle differences in acquisition protocol, image processing, or distribution pipeline (e.g., image compression) and overlook pathology." *– Zech et al, 2018*

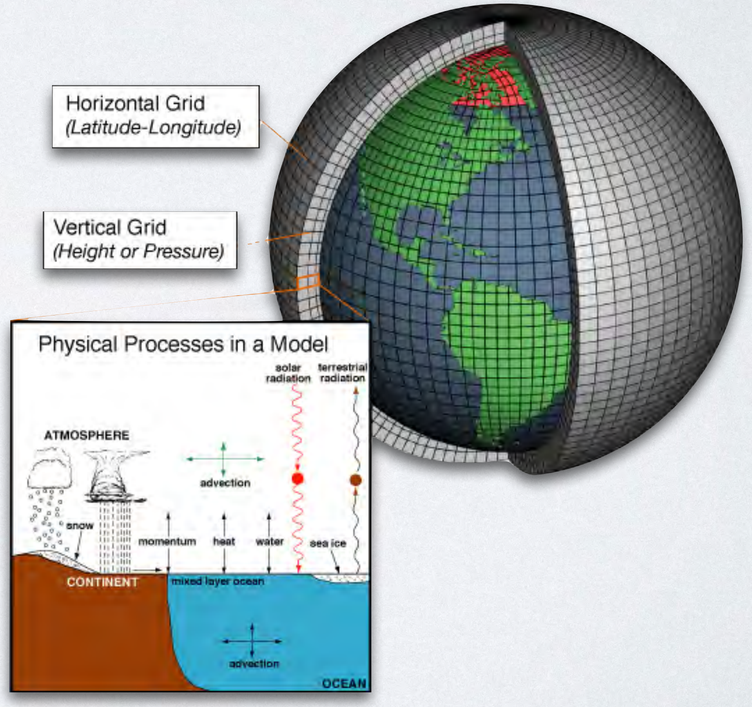
Leveraging models of physics and humans

"Worried about superintelligent robots rising up and attacking us? Close your doors."

— *Gary Marcus and Ernest Davis*

https://what-if.xkcd.com/5/

Physical models may guide and stabilize many AI methods

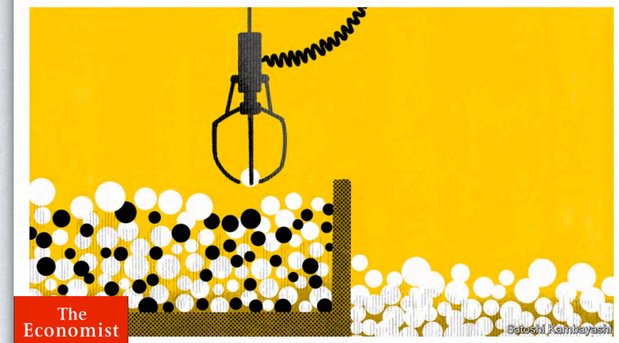


https://magazine.cim.org/en/technology/the-mechanics-of-a-new-era-in-maintenance-en,, https://www.technologynetworks.com/tn/product-surveys/2018-lab-water-survey-306836, https://en.wikipedia.org/wiki/General_circulation_model

Buttonwood

The benefits of better credit-risk models will be spread unevenly

Machine-learning models show the disquieting effect of finer judgments



Even if we understand AI, *human-AI interactions* may be unpredictable

THE IRISH TIMES

Roma attacked in Paris after fake videos circulate on social media

Propagated online, the urban myth of marauding Roma has taken on a life of its own

Wed, Mar 27, 2019, 18:34

Lara Marlowe Paris Correspondent

A man holds a child in a camp of the Roma community that was attacked in Bobigny, near Paris, Photograph: Kenzo Tribouillard/AFP/Getty

Forbes

What Happens When Self-Driving Cars Kill People?

Ron Schmelzer Contributor
COGNITIVE WORLD Contributor Group

PHOTO BY YENDER FONSECA FROM PEXELS

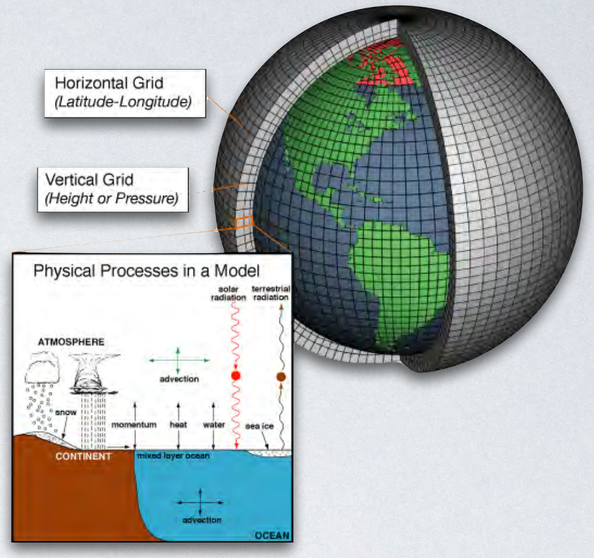
Strangelove redux: US experts propose having AI control nuclear weapons

By Matt Field, August 30, 2019

A US missile test. Photo via Wikimedia Commons. Public Domain.

Some key challenges in modern AI

- How can we integrate training data, physical models, and human judgement?
- Will my method work “in the wild”? What risks do I face at deployment?
- Do I have “enough” diverse data? Can I quantify uncertainty and interpret models?
- What model should I use when? When is deep learning the right tool?



These questions are not beyond reach

Radhakrishnan, Yang, Belkin, Uhler, 2019

Recent theory & practice:

- explains memorization by in deep learning
- gives insight into what makes neural networks special
- provides robustness to adversarial inputs
- helps make systems faster and training easier

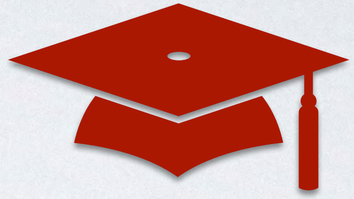
Leadership requires sustained investment in both applications and foundations of AI with partnerships among academia, government, and industry.

Partnerships between academia, industry, government

Understands full spectrum of operating modes

Transitions academic advances to consumers

Industry



Academia

Foundational contributions

Understanding, predicting, and preventing failure modes

Training next generation of AI experts

Government

Advances high-performance scalability of academic research

Oversight and regulatory responsibilities