Computers That See Like Humans

Bradley Zhu

How Computers Recognize Images

- Millions of individual pixels each with color
- Deep learning allows for higher and higher levels of abstraction
- Deep Neural Networks look at the parts of an image on a pixel level that distinguish it from other images. They can have several layers of neurons.
  - It tries to find patterns. A grossly oversimplified example: green pixel, green pixel, purple pixel, green pixel might be common in pictures of peacocks.
- Because deep learning utilizes millions of neurons, we can know the input and the output, but we don't always know how they work.
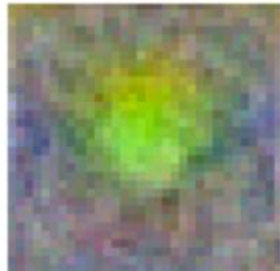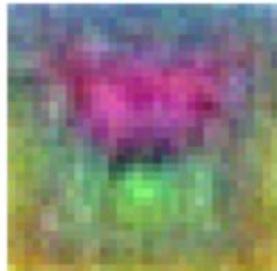
plane car bird cat deer

dog frog horse ship truck

Used with permission from Ray Zhang

| robin | cheetah | armadillo | lesser panda |
| centipede | peacock | jackfruit | bubble |

Steele, Bill. "Images That Fool Computer Vision Raise Security Concerns." Cornell University. Cornell Chronicle, 20 Mar. 2015. Web. 02 Mar. 2016.
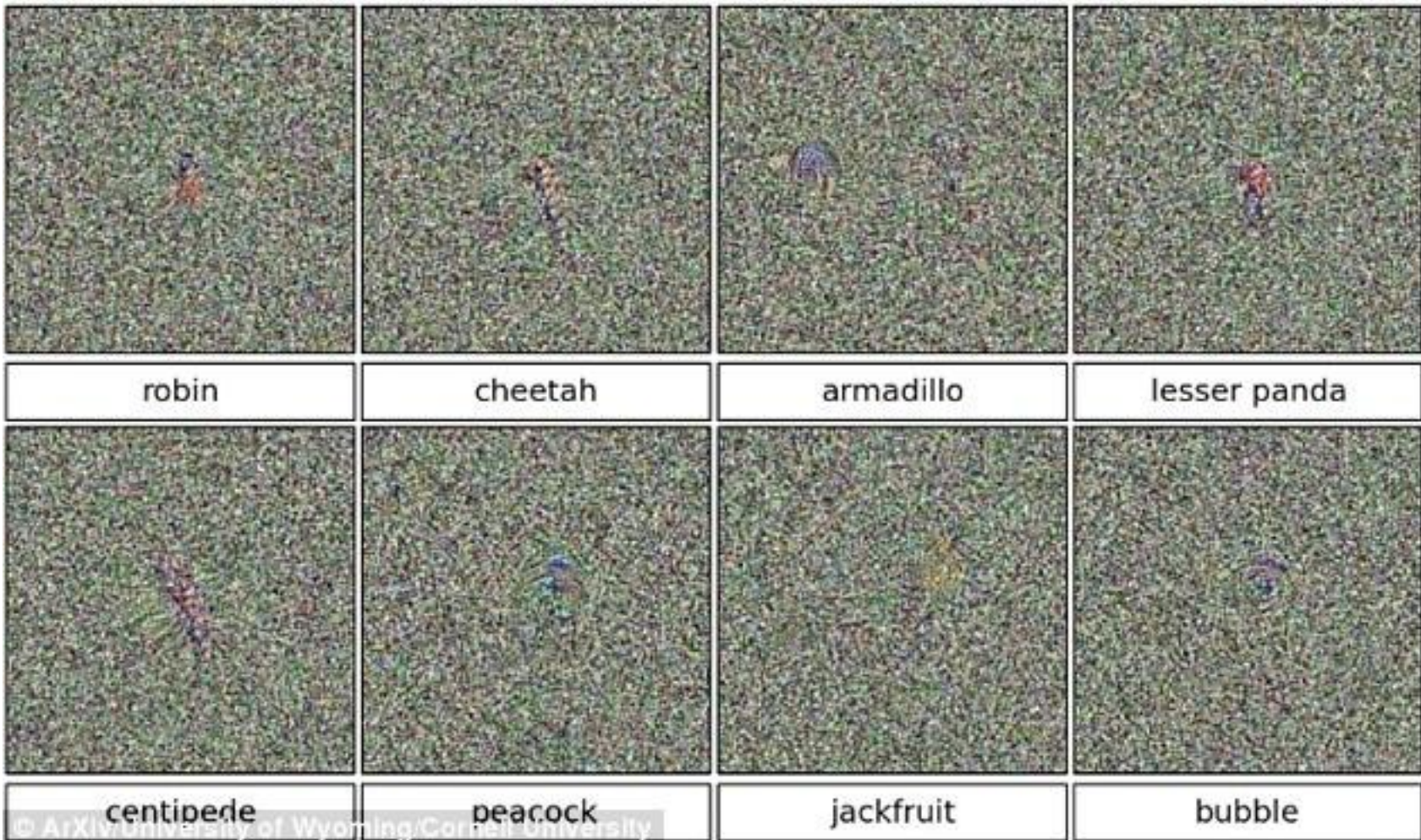
"Easily Fooled" Computer Vision Security Risks

- Neural nets currently don't encode enough information to produce an image of an object. They only have the knowledge necessary to tell the object apart from other objects.
    - Misidentification
    - Invisibility to Surveillance Systems
    - Abstract shapes recognized as faces
    - Fake images fooling search engines
    - Malicious fake images

Humans Vs. Computers

- Recently Deep Learning Projects are starting to outperform humans in visual recognition, but computers are still bad at complex scenes (such as details about actions performed by peoples).
- A study by Ullman with 14,000 human participants showed that humans outperform computers on fuzzy/minimal pictures to a point.
  - Humans have a sharp drop after a "baseline" from 65% to below 20%, unlike computers.
- Computers work "bottom-up" starting with the simplest features before moving onto the more complex features.
- Humans work "top-down" comparing a standard model of an object with the object they're trying to identify

Digital Human Baby

- A "top-down" model might inspire new computer models and algorithms capable of developing a more "complex understanding" of the world through what they see.
- Shimon Ullman and his group is trying to "reverse engineer the infant mind".
  - "As a baby, you open your eyes, see flickering pixels, and somehow it all comes together and you know something about the world. You're starting from nothing, absorbing information and getting a rich view of the world. We were thinking about what would it take to get a computer program where you put in the minimal structures you need and let it view videos or the world for six months. If you do it right, you can get an interesting system." - Ullman
- In addition to computer science, this study also has potential for learn more about how the human brain sees the world.

So What?

- Complex understanding of images
- Ullman wants computers to eventually learn to understand the world like a growing toddler does
- Increased performance in everything from facial recognition to understanding of surrounding environment
- Weak AI vs. Strong AI

Works Cited

Aron, Jacob. "Optical Illusions Fool Computers into Seeing Things." New Scientist.
    New Scientist, 11 Dec. 2014. Web. 02 Mar. 2016.

Basulto, Dominic. "Humans Are the World's Best Pattern-Recognition Machines, But for How Long?"
    Big Think. Big Think, 24 July 2013. Web. 02 Mar. 2016.

Hsu, Jeremy. "Digital Baby Project's Aim: Computers That See Like Humans."
    Computers That See Like Humans. IEEE Spectrum, 15 Feb. 2016. Web. 01 Mar. 2016.

Steele, Bill. "Images That Fool Computer Vision Raise Security Concerns." Cornell University.
    Cornell Chronicle, 20 Mar. 2015. Web. 02 Mar. 2016.

Vanhemert, Kyle. "Simple Pictures That State-of-the-Art AI Still Can't Recognize." Wired.com.
    Conde Nast Digital, 5 Jan. 2015. Web. 02 Mar. 2016.