See how an AI system classifies you based on your selfie

Modern artificial intelligence is often lauded for its growing sophistication, but mostly in doomer terms. If you're on the apocalyptic end of the spectrum, the AI revolution will automate millions of jobs, eliminate the barrier between reality and artifice, and, eventually, force humanity to the brink of extinction. Along the way, maybe we get robot butlers, maybe we're stuffed into embryonic pods and harvested for energy. Who knows.

But it's easy to forget that most AI *right now* is terribly stupid and only useful in narrow, niche domains for which its underlying software has been specifically trained, like playing an ancient Chinese board game or translating text in one language into another.

Ask your standard recognition bot to do something novel, like analyze and label a photograph using only its acquired knowledge, and you'll get some comically nonsensical results. That's the fun behind ImageNet Roulette, a nifty web tool built as part of an ongoing art exhibition on the history of image recognition systems.

ImageNet Roulette will make wondrously random and wildly incorrect assumptions about you
As explained by artist and researcher Trevor Paglen, who created the exhibit *Training Humans* with AI researcher Kate Crawford, the point is not to make a judgement about AI, but to engage with its current form and its complicated academic and commercial history, as grotesque as it might be.

"When we first started conceptualizing this exhibition over two years ago, we wanted to tell a story about the history of

images used to ‘recognize’ humans in computer vision and AI systems. We weren’t interested in either the hyped, marketing version of AI nor the tales of dystopian robot futures,” Crawford told the Fondazione Prada museum in Milan, where Training Humans is featured. “We wanted to engage with the materiality of AI, and to take those everyday images seriously as a part of a rapidly evolving machinic visual culture. That required us to open up the black boxes and look at how these ‘engines of seeing’ currently operate.”

It’s a worthy pursuit and a fascinating project, even if ImageNet Roulette represents the goofier side of it. That’s mostly because ImageNet, a renown training data set AI researchers have relied on for the last decade, is generally bad at recognizing people. It’s mostly an object recognition set, but it has a category for “People” that contains thousands of subcategories, each valiantly trying to help software do the seemingly impossible task of classifying a human being.

And guess what? ImageNet Roulette is super bad at it.

I don’t even smoke! But for some reason, ImageNet Roulette thinks I do. It also appears to believe that I am located in an airplane, although to its credit, open office layouts are only slightly less suffocating than narrow metal tubes suspended tens of thousands of feet in the air.

ImageNet Roulette was put together by developer Leif Ryge working under Paglen, as a way to let the public engage with the art exhibition’s abstract concepts about the inscrutable nature of machine learning systems.

Here’s the behind-the-scenes magic that makes it tick:

> *ImageNet Roulette uses an open source Caffe deep learning framework (produced at UC Berkeley) trained on the images and labels in the “person” categories (which are currently ‘down for maintenance’). Proper nouns and categories with less than 100 pictures were removed.*
>
> *When a user uploads a picture, the application first runs a face detector to locate any faces. If it finds any, it sends them to the Caffe model for classification. The application then returns the original images with a bounding box showing the detected face and the label the classifier has assigned to the image. If no faces are detected, the application sends the entire scene to the Caffe model and returns an image with a label in the upper left corner.*

Part of the project is also to highlight the fundamentally flawed, and therefore human, ways that ImageNet classifies people in “problematic” and “offensive” ways. (One interest example popping up on Twitter is that some men uploading photos appear to be randomly tagged as “rape suspect” for reasons unexplained.) Paglen says this is crucial to one of the themes the project is highlighting, which is the fallibility of AI systems and the prevalence of machine learning bias as a result of its compromised human creators:

> *ImageNet contains a number of problematic, offensive and bizarre categories — all drawn from WordNet. Some use misogynistic or racist terminology. Hence, the results ImageNet Roulette returns will also draw upon those categories. That is by design: we want to shed light on what happens when technical systems are trained on problematic training data. AI classifications of people are rarely made visible to the people being classified. ImageNet Roulette provides a glimpse into that process — and to show the ways things can go wrong.*

> *ImageNet is one of the most significant training sets in the history of AI. A major achievement. The labels come from WordNet, the images were scraped from search engines. The 'Person' category was rarely used or talked about. But it's strange, fascinating, and often offensive.*
>
> *— Kate Crawford (@katecrawford)* *September 16, 2019*

Although ImageNet Roulette is a fun distraction, the underlying message of *Training Humans* is a dark, but vital, one.

"*Training Humans* explores two fundamental issues in particular: how humans are represented, interpreted and codified through training datasets, and how technological systems harvest, label and use this material," reads the exhibition description "As the classifications of humans by AI systems becomes more invasive and complex, their biases and politics become apparent. Within computer vision and AI systems, forms of measurement easily – but surreptitiously – turn into moral judgments."

https://www.theverge.com/tldr/2019/9/16/20869538/imagenet-roulette-ai-classifier-web-tool-object-image-recognition