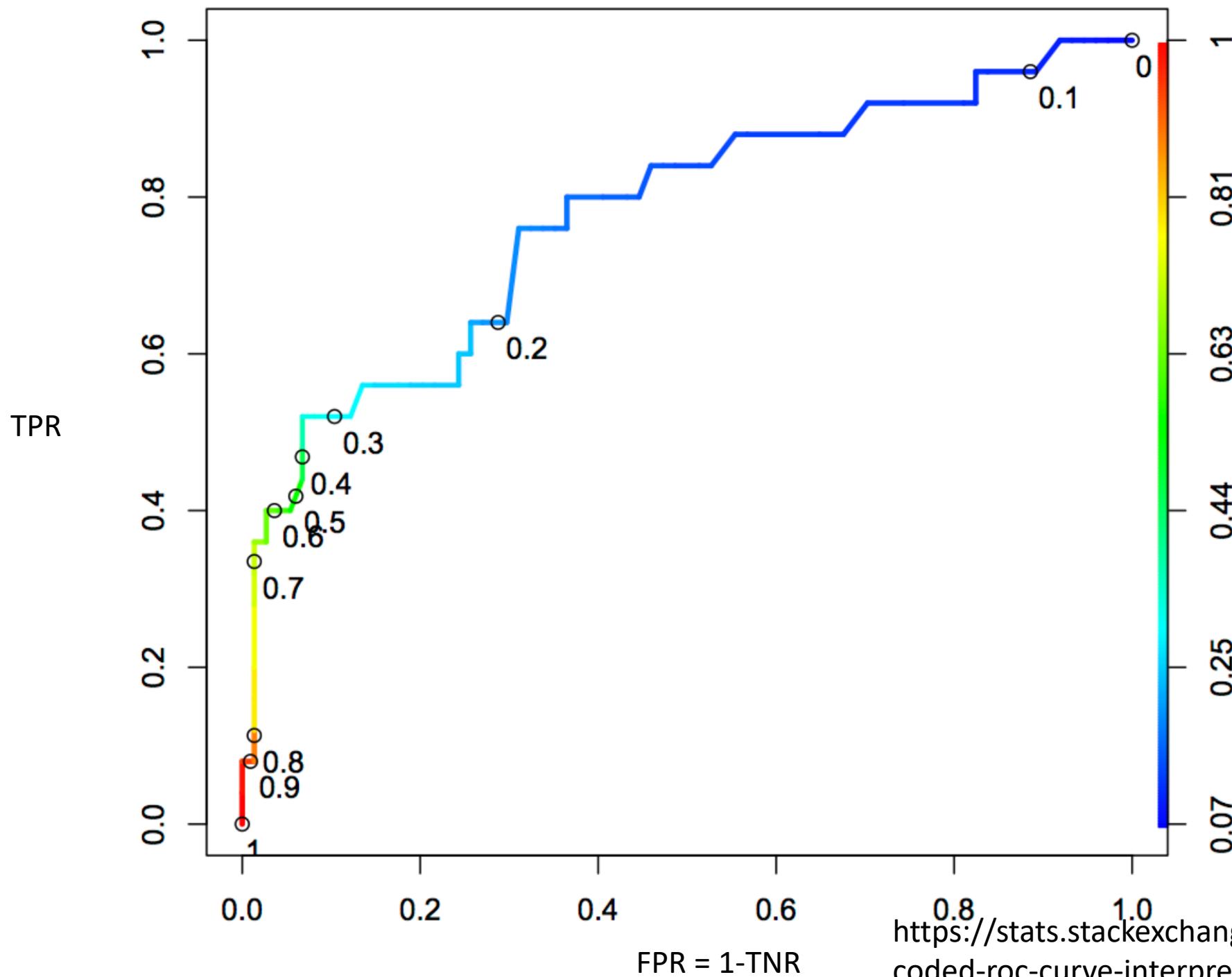


THE UNIVERSITY OF
CHICAGO



<https://stats.stackexchange.com/questions/470696/threshold-coded-roc-curve-interpretation>

Dependency hell...

- While preparing notes in spring 2022, I encountered a Unicode decode error from scikitlearn's LogisticRegressionCV method.
- Since the data was numpy arrays (no Unicode!) I took this to be a bug.
- <https://github.com/scikit-optimize/scikit-optimize/issues/981>



grudloff commented on May 31, 2021

...

Looks like this is an issue of compatibility between `scikit-learn<=0.23.2` and `scipy>=1.6.0` (see this [issue](#)). Maybe the requirements should be raised to require `scikit-learn>0.23.2` ?

- Bug fix in scipy broke scikit-learn
- Can anything be done?

Pinned versions: A deal with the devil



- When your code needs stability, you can insist on the version and subversion number of the operating system, the python engine, and the scientific libraries.
- API of scientific libraries slowly evolves.
“DeprecationWarning” anyone?
- Expect code that specifies pinned versions to work 100% until it breaks, and then it will break catastrophically.
- Pinned version code is more needy than shotgun code; it needs more hard drive space, separate computing environments, containers...



A viral video that appeared to show Obama calling Trump a 'dips---' shows a disturbing new trend called 'deepfakes' [Kaylee Fagan](#) Apr 17, 2018, 3:48 PM

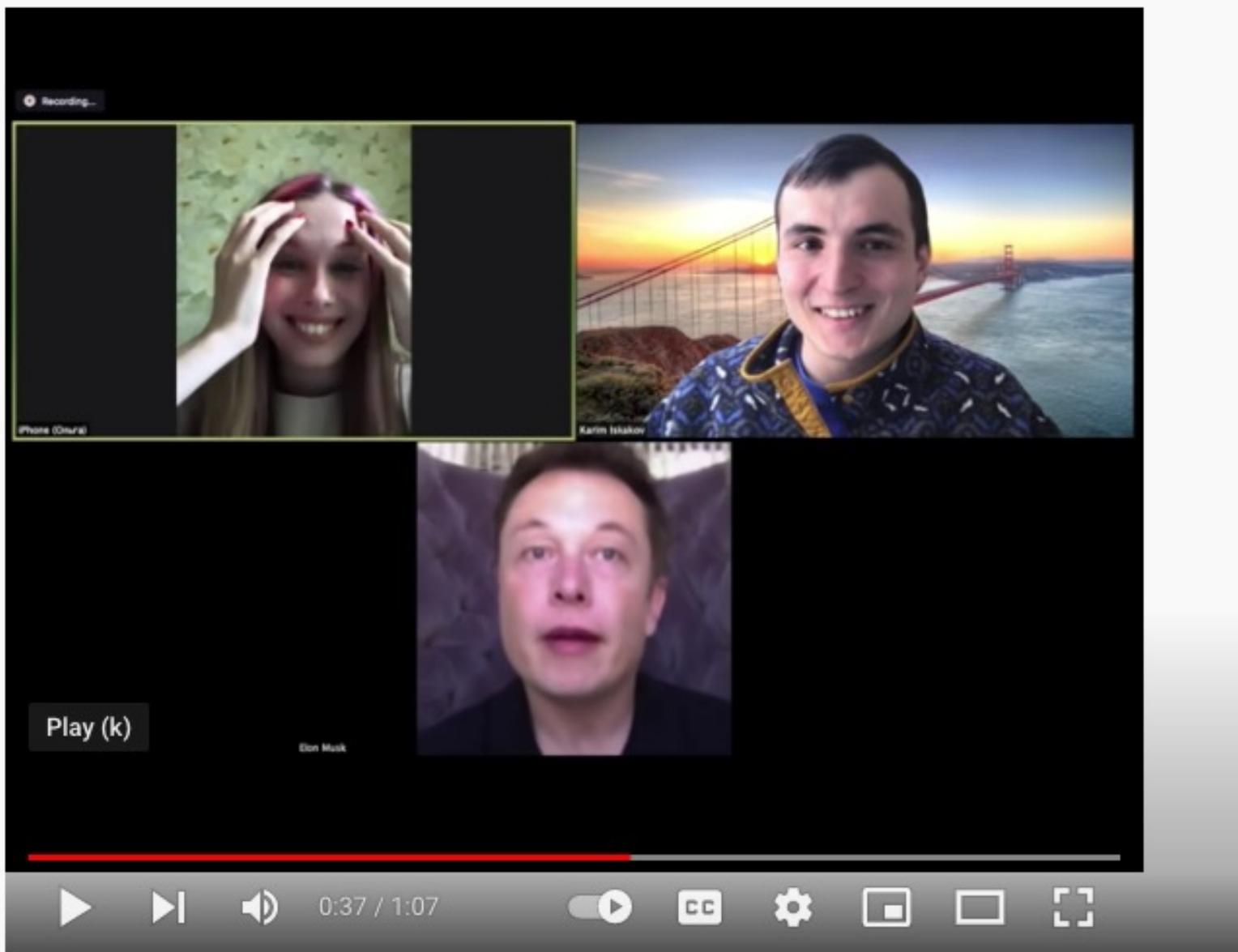
<https://www.businessinsider.com/obama-deepfake-video-insulting-trump-2018-4>

Caught in the wild (without Cassandra disclaimer)



<https://www.wired.com/story/zelensky-deepfake-facebook-twitter-playbook/>

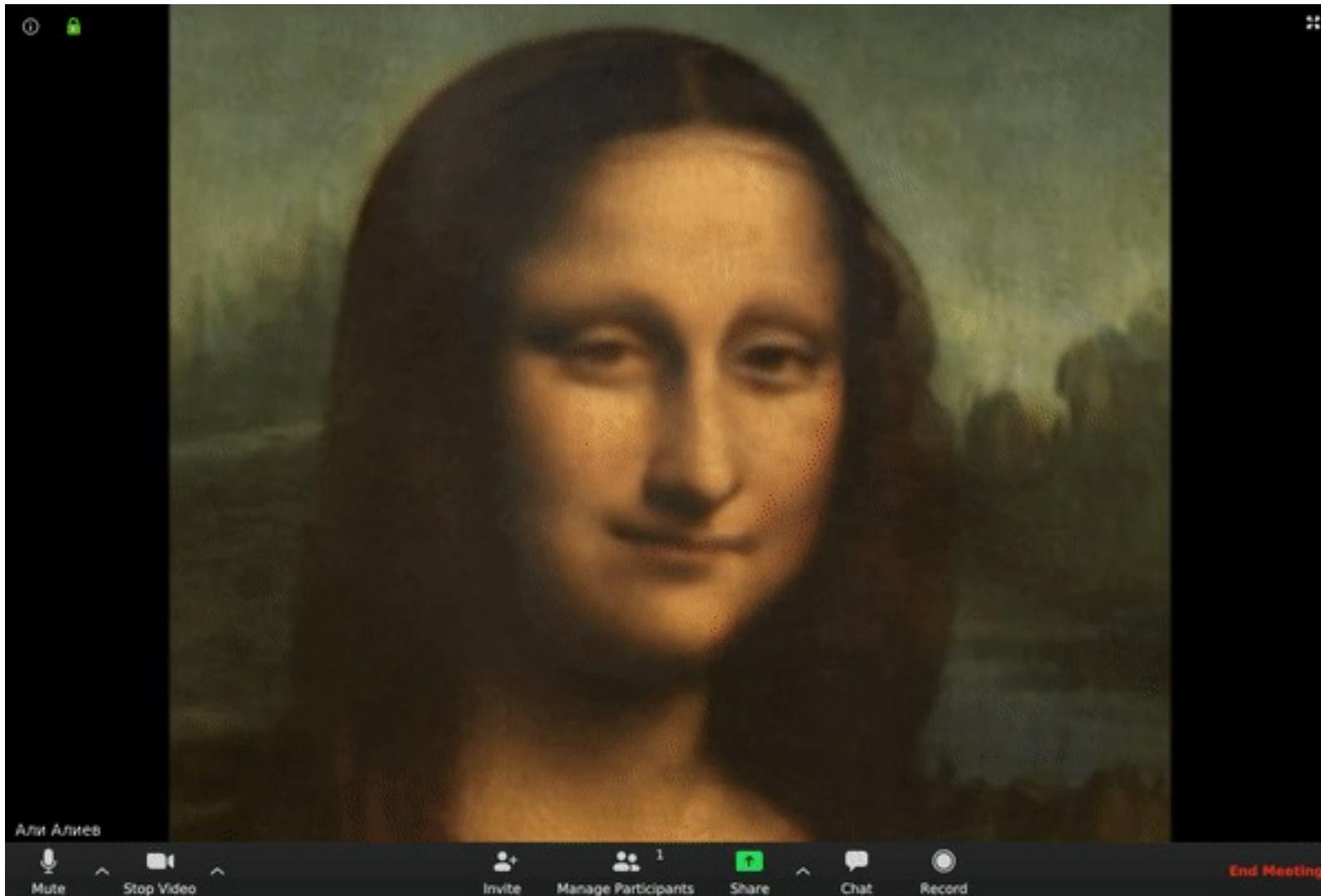
Mar 17, 2022 Tom Simonite A Zelensky Deepfake Was Quickly Defeated. The Next One Might Not Be



Elon Musk joined our Zoom call | Avatarify

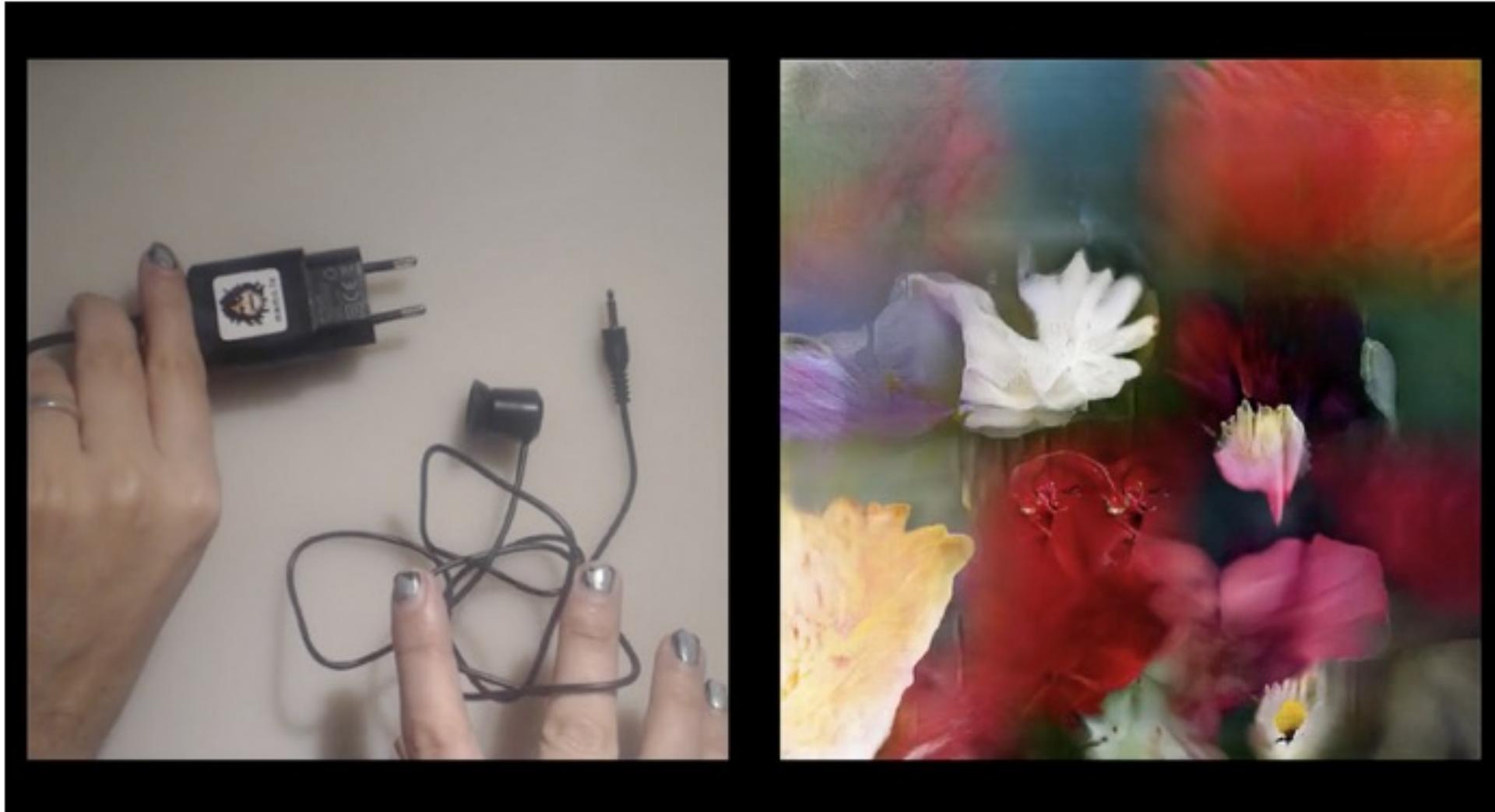
<https://www.youtube.com/watch?v=lONuXGNqLO0>

Avatarify – server for streaming video faceswap



<https://github.com/alievk/avatarify-python>

Learning to see, gloomy Sunday morning

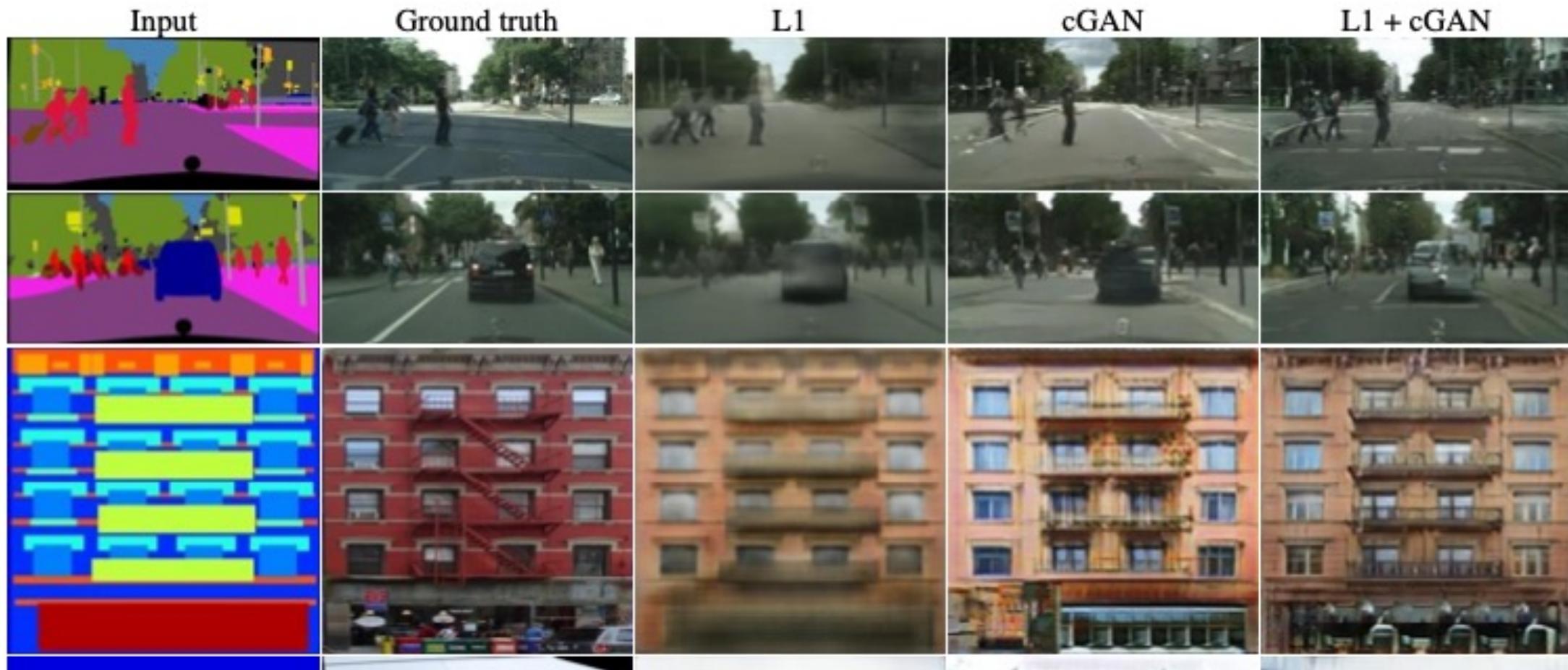


Still from Memo Akten's video [Learning to See](#)



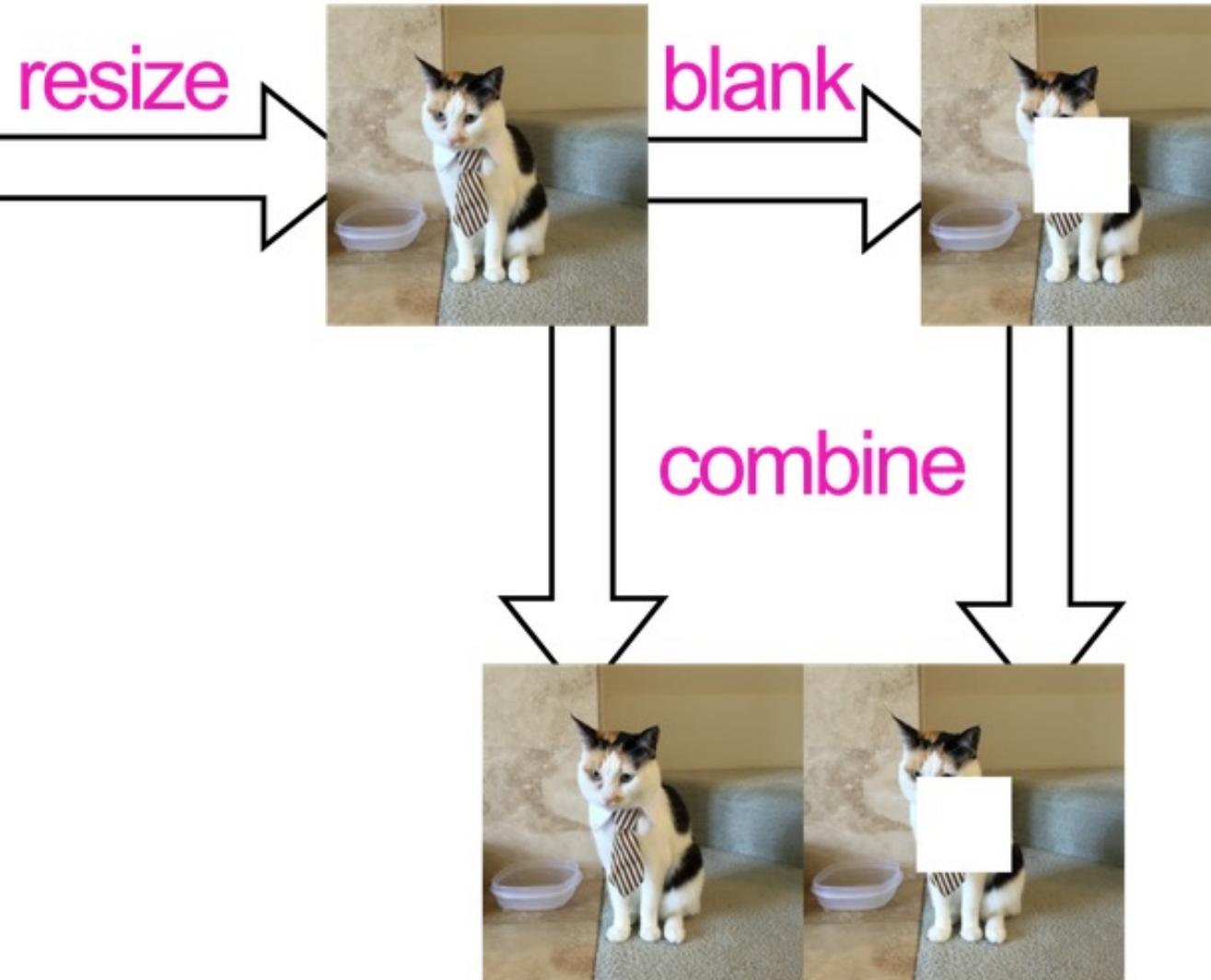
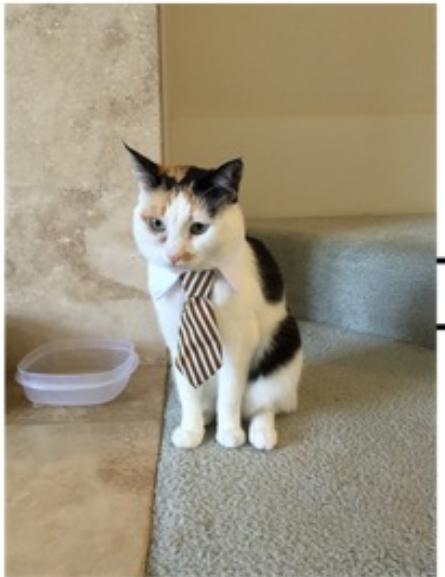
https://youtu.be/UroVey4fJ_g

pix2pix -- CNNs trained on image pairs

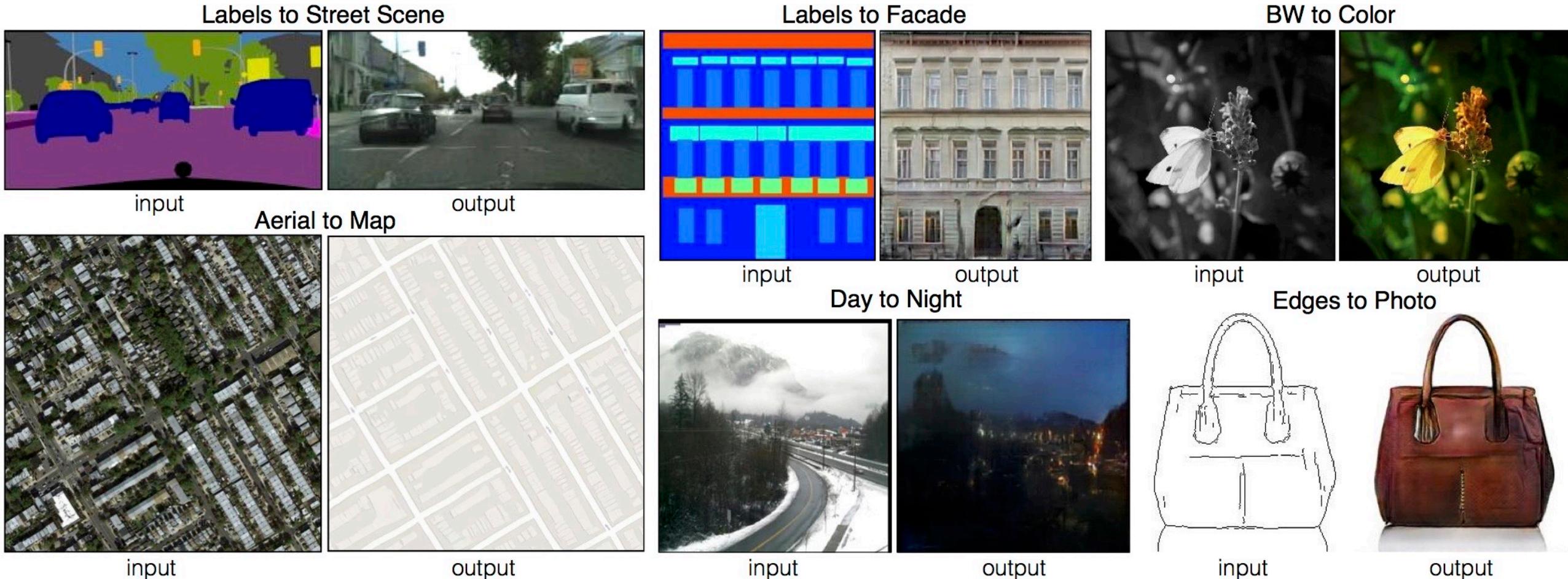


pix2pix -- CNNs trained on image pairs





pix2pix -- CNNs trained on image pairs

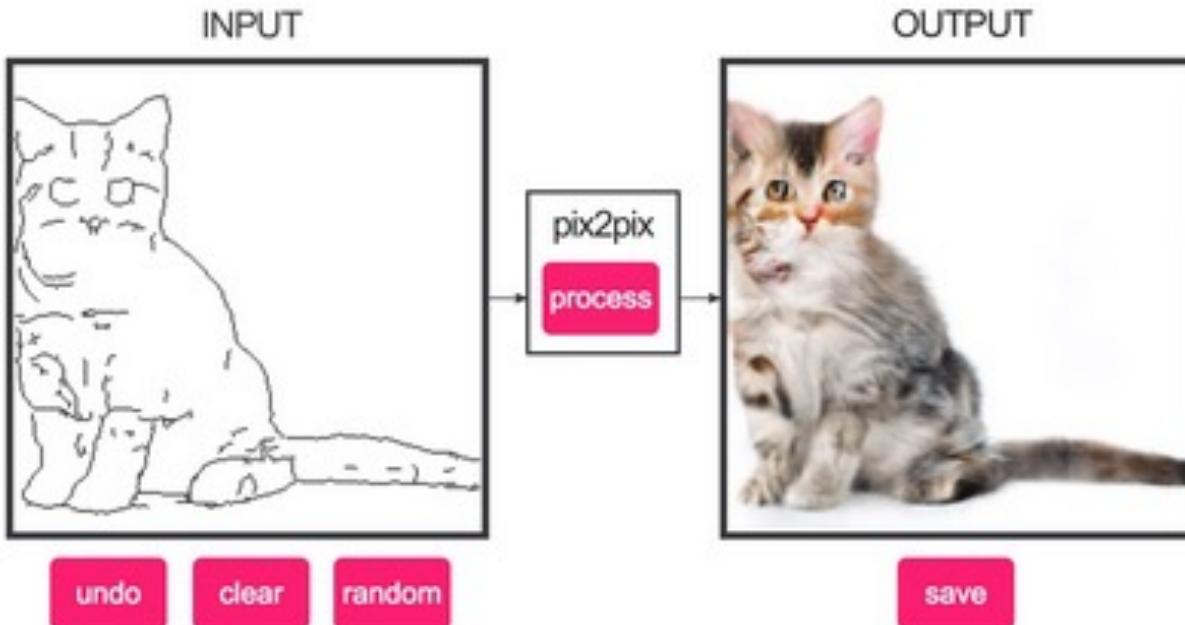


<https://github.com/affinelayer/pix2pix-tensorflow>

Edges2cats

Interactive Demo

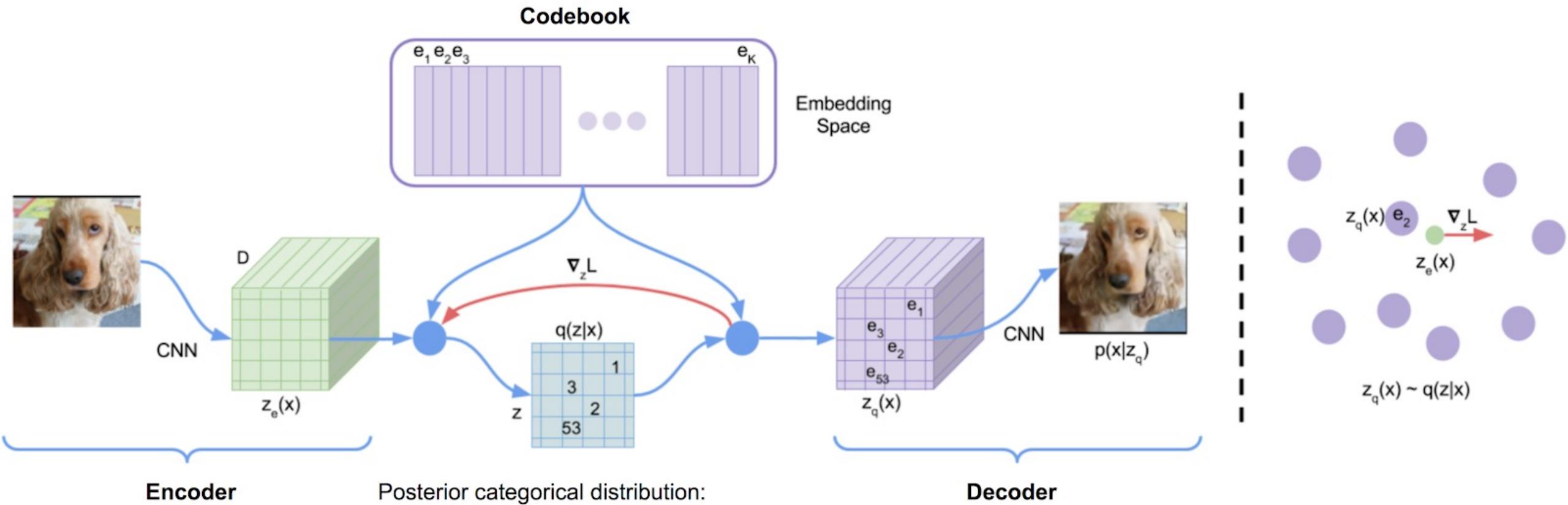
(made by Christopher Hesse)



<https://phillipi.github.io/pix2pix/>

<https://affinelayer.com/pixsrv/>

OpenAI: DALL-E: trained to generate images from text



$$q(\mathbf{z} = \mathbf{e}_k | \mathbf{x}) = \begin{cases} 1 & \text{if } k = \arg \min_i \|\mathbf{z}_e(\mathbf{x}) - \mathbf{e}_i\|_2 \\ 0 & \text{otherwise.} \end{cases}$$

TEXT PROMPT

an illustration of a baby daikon radish in a tutu walking a dog

AI-GENERATED IMAGES



[Edit prompt or view more images ↓](#)

TEXT PROMPT

an armchair in the shape of an avocado [...]

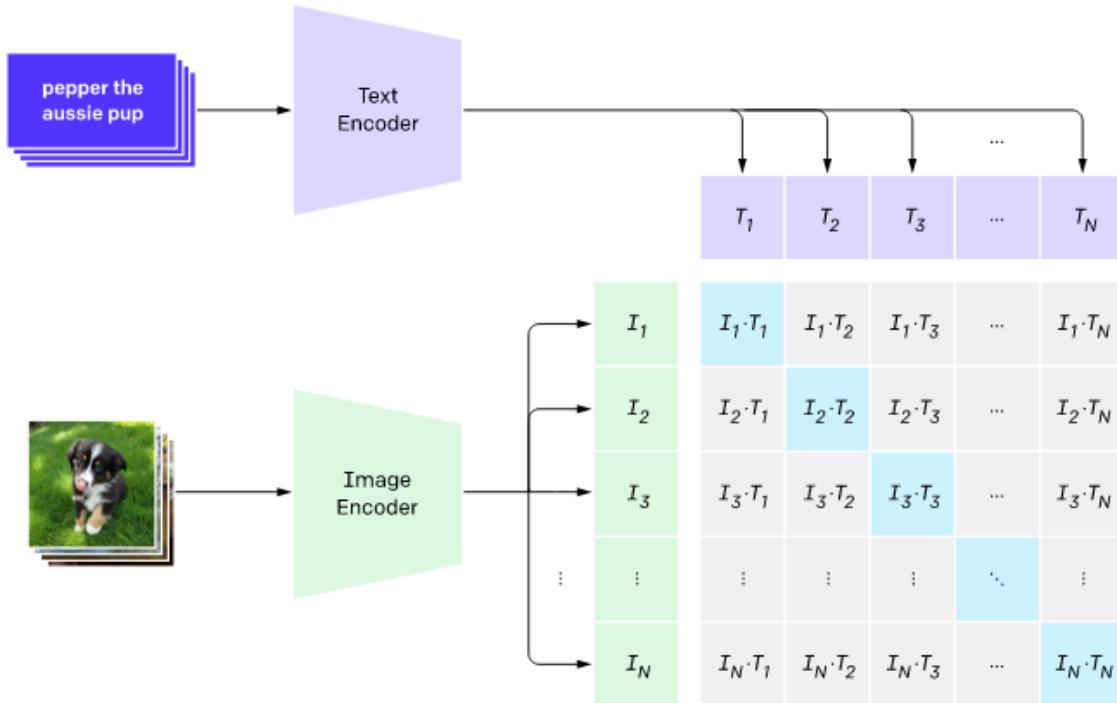
AI-GENERATED IMAGES



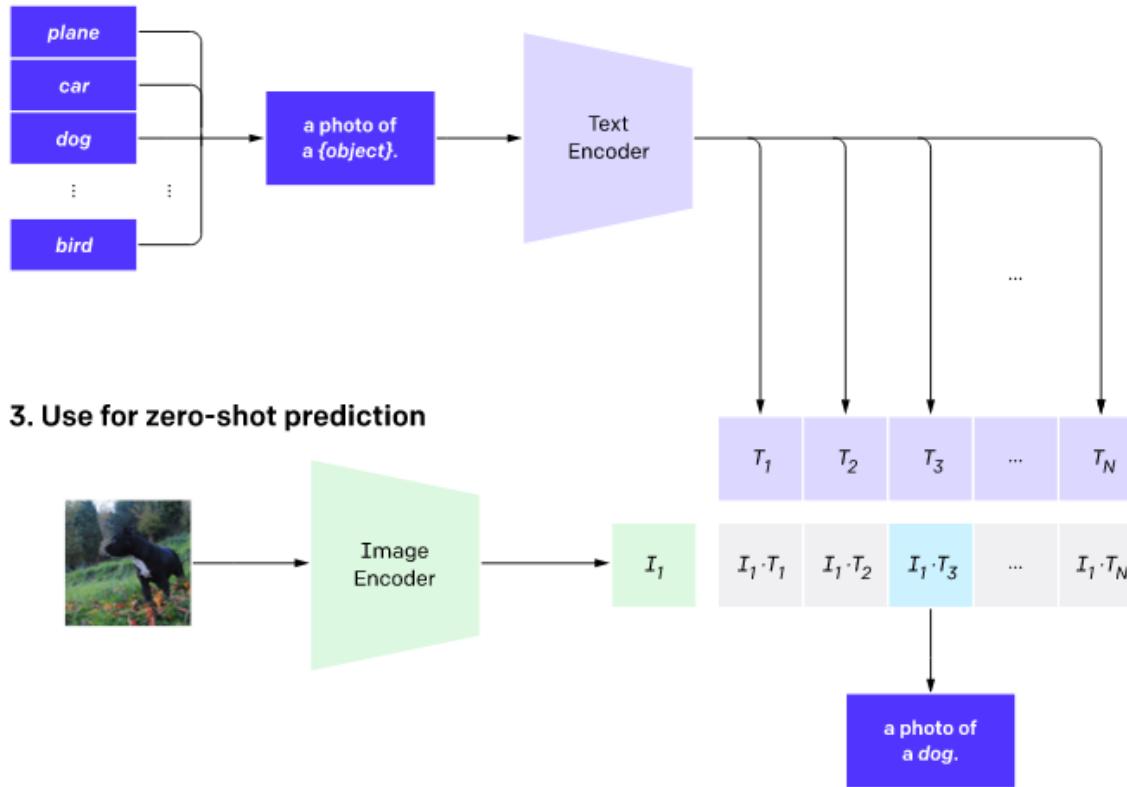
[Edit prompt or view more images ↓](#)

CLIP: another text-to-image framework

1. Contrastive pre-training



2. Create dataset classifier from label text



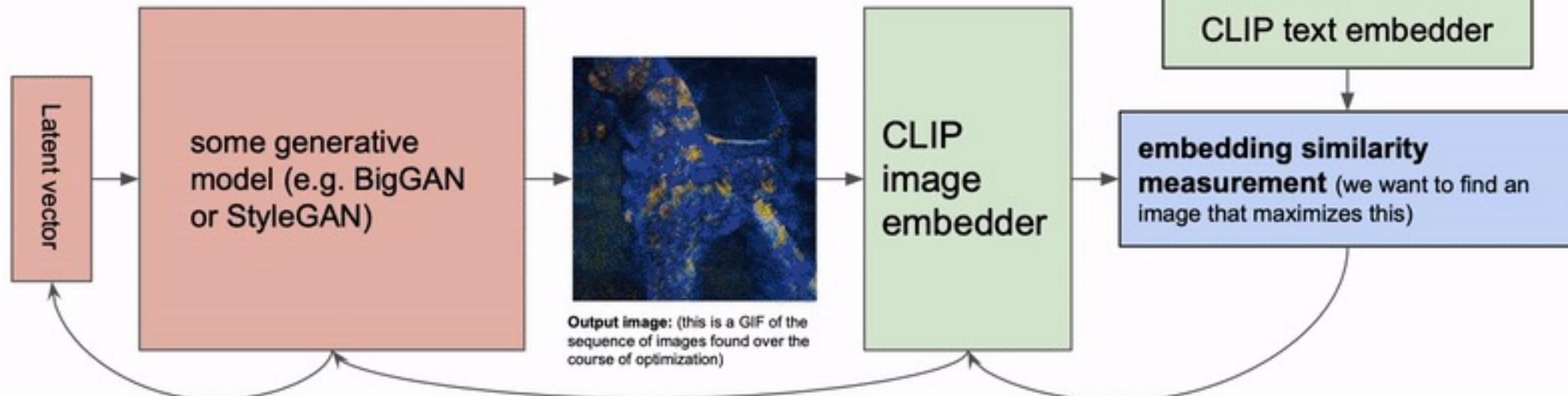
3. Use for zero-shot prediction

CLIP pre-trains an image encoder and a text encoder to predict which images were paired with which texts in our dataset. We then use this behavior to turn CLIP into a zero-shot classifier. We convert all of a dataset's classes into captions such as "a photo of a dog" and predict the class of the caption CLIP estimates best pairs with a given image.

How CLIP Generates Art

Forward Pass:

Push a latent through the generative model to produce an image. Then pass the image to CLIP's image embedder to measure the image's similarity with the text prompt



Backward Pass:

Backpropagate through CLIP and the generative model, all the way back to the latent vector, and then use gradient ascent to update the latent, bringing the image slightly closer to matching with the text prompt.

ART MACHINE: Put in text, get AI art.

<https://is.gd/artmachine> -> https://colab.research.google.com/drive/1n_xrgKDlGQcCF6O-eL3NOd_x4NSqAUjK

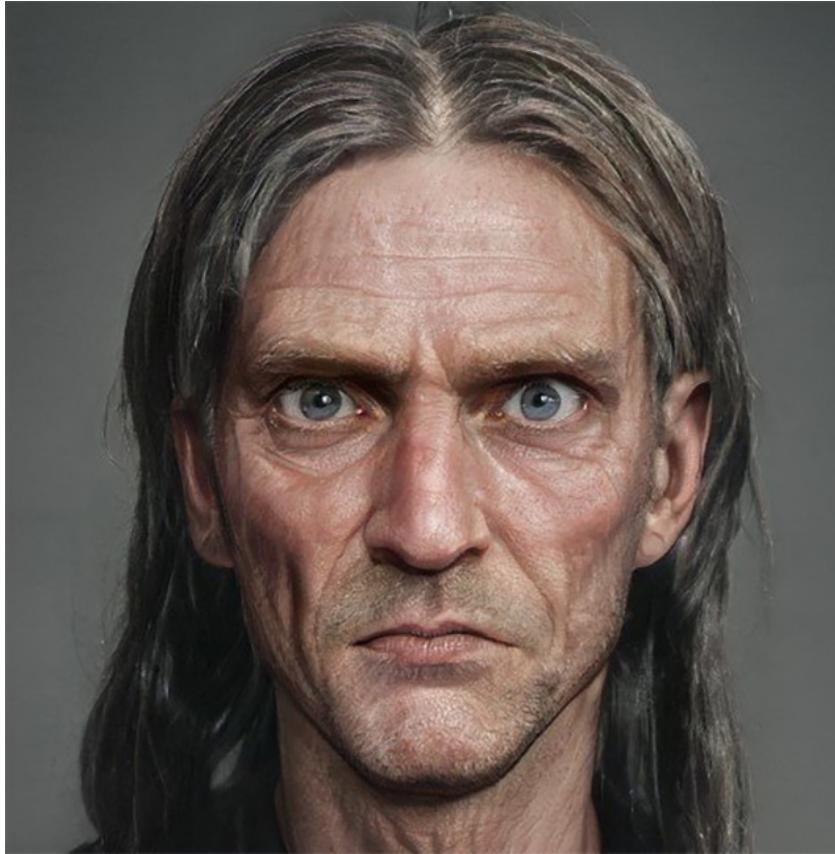
Applications abound



Colorization
Image upscaling



Bellatrix: (mid 40s) long shiny thick black hair, strong jaw, dark hooded eyes, thin lips, gaunt from long term imprisonment, still has vestiges of old beauty



Argus Filch: (?age?) bulging eyes, pale eyes, sunken cheeks, pasty-faced, (missing bald patch)

Why is this happening?

- The math is decades old, but the cool applications are in the past 10 years.
- The math will endure, matplotlib, ggplot, torch, tensorflow may fall out of style
- Cheaper compute, cheaper storage, cheaper network, more data
- Lower barrier to entry (notebooks, collab) means many more people working with ML/AI tools
- Look at the amount of art! AI isn't just for giant corporate conglomerates anymore.

With great power comes great responsibility

- <https://groups.csail.mit.edu/vision/TinyImages/>



June 29 <http://localhost:8890/mediawiki-1.35.2/index.php/Special>AllPages>

It has been brought to our attention [1] that the Tiny Images dataset contains some derogatory terms as categories and offensive images. This was a consequence of the automated data collection procedure that relied on nouns from WordNet. We are greatly concerned by this and apologize to those who may have been affected.

The dataset is too large (80 million images) and the images are so small (32 x 32 pixels) that it can be difficult for people to visually recognize its content. Therefore, manual inspection, even if feasible, will not guarantee that offensive images can be completely removed.

We therefore have decided to formally withdraw the dataset. It has been taken offline and it will not be put back online. We ask the community to refrain from using it in future and also delete any existing copies of the dataset that may have been downloaded.

How it was constructed: The dataset was created in 2006 and contains 53,464 different nouns, directly copied from Wordnet. Those terms were then used to automatically download images of the corresponding noun from Internet search engines at the time (using the available filters at the time) to collect the 80 million images (at tiny 32x32 resolution; the original high-res versions were never stored).

Why it is important to withdraw the dataset: biases, offensive and prejudicial images, and derogatory terminology alienates an important part of our community -- precisely those that we are making efforts to



Imagenet roulette (art project)
<https://excavating.ai/>

ImageNet contains 2,833 subcategories under the top-level category “Person.”

As we go further into the depths of ImageNet’s Person categories, the classifications of humans within it take a sharp and dark turn. There are categories for Bad Person, Call Girl, Drug Addict, Closet Queen, Convict, Crazy, Failure, Flop, Fucker, Hypocrite, Jezebel, Kleptomaniac, Loser, Melancholic, Nonperson, Pervert, Prima Donna, Schizophrenic, Second-Rater, Spinster, Streetwalker, Stud, Tosser, Unskilled Person, Wanton, Waverer, and Wimp. There are many racist slurs and misogynistic terms.

Image of journalist Izzy Stephen tagged with a subset of Imagenet’s unflattering person labels.

<https://medium.com/@isobel.stephen/what-can-we-learn-from-ai-art-4b0a52476dd9>



Jacky Alciné
@jackyalcine



Google Photos, y'all fucked up. My friend's not a gorilla.

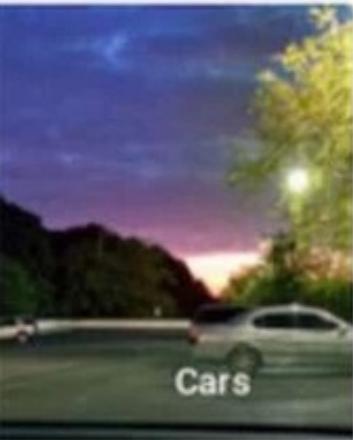


Image databases need careful review..

Drug addict, junkie, junky

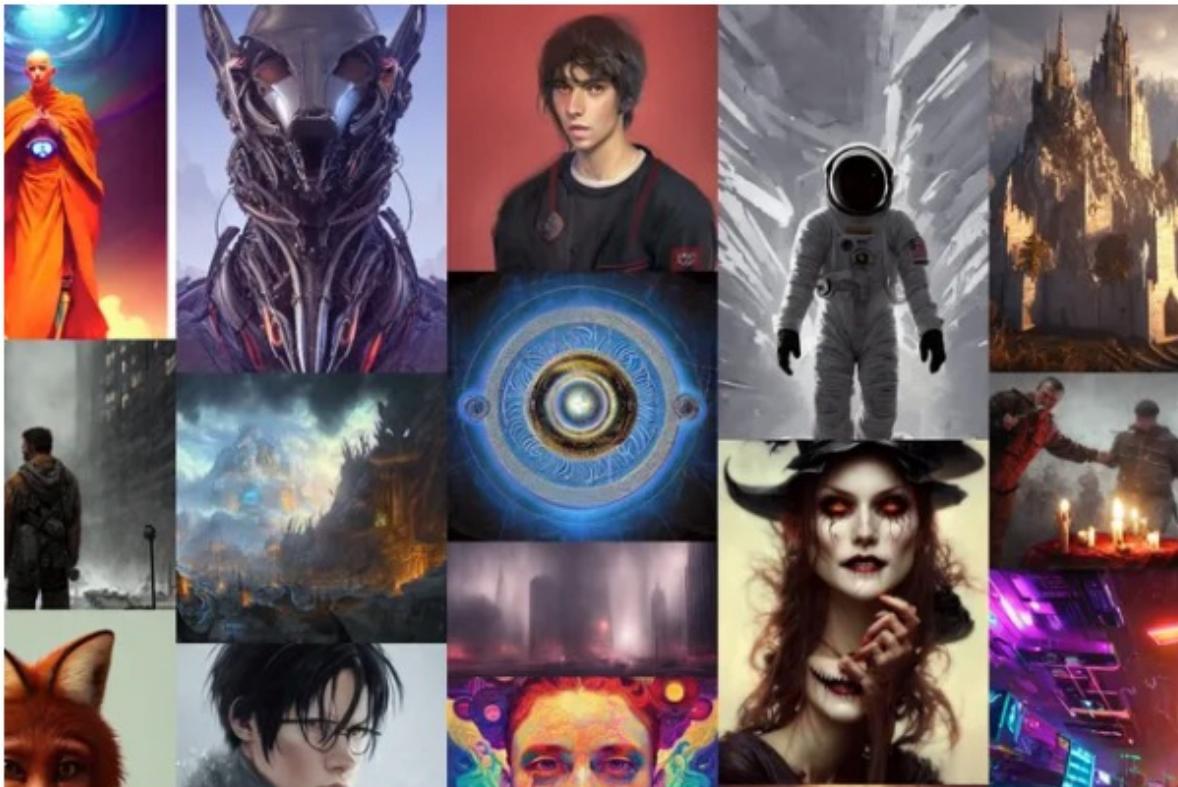
A narcotics addict

192 pictures 69.42% Popularity Percentile Wordnet IDs

Treemap Visualization Images of the Synset Downloads

- dissenter, dissident, protester, objector, c
- nonconformist, recusant (5)
 - bohemian (0)
 - maverick, rebel (0)
 - enfant terrible (0)
 - heretic (0)
 - beatnik, beat (0)
- NIMBY (0)
- conscientious objector, CO (0)
- political dissident (0)
- drug user, substance abuser, user (15)
 - addict (8)
 - speed freak (0)
 - caffeine addict, caffeine addict (0)
 - + drug addict, junkie, junky (5)
 - opium addict, opium taker (0)
 - crack addict, binger (0)
 - withdrawer (0)
 - cocaine addict (0)
 - heroin addict (0)
 - tripper (0)
 - head (4)
 - agnostic, doubter (0)
 - greeter, saluter, welcomer (0)
 - percher (0)
 - gentile (0)
 - laugher (1)
 - baldhead, baldpate, baldy (0)
 - advocate, advocator, proponent, exponent (0)
 - nonreligious person (14)
 - abjurer (0)
 - Pisces, Fish (0)
 - iunior (0)

AI art tools Stable Diffusion and Midjourney targeted with copyright lawsuit



A collage of AI-generated images created using Stable Diffusion. Image: [The Verge via Lexica](#)

/ The suit claims generative AI art tools violate copyright law by scraping artists' work from the web without their consent.

By [JAMES VINCENT](#)

Jan 16, 2023 at 5:28 AM CST | □ [29 Comments](#) / [29 New](#)



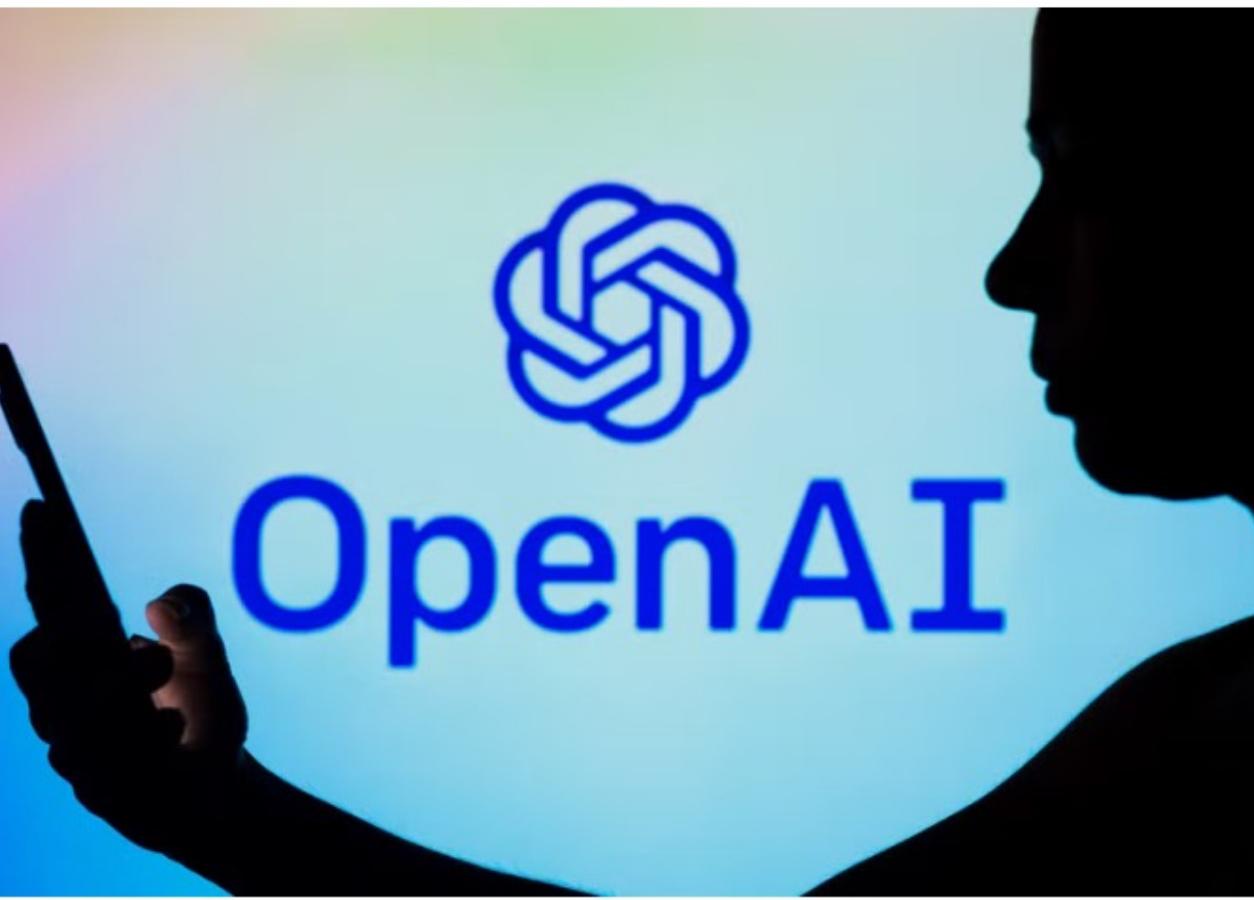
A trio of artists have launched a lawsuit against Stability AI and Midjourney, creators of AI art generators Stable Diffusion and

Finally, an A.I. Chatbot That Reliably Passes “the Nazi Test”

OpenAI's new ChatGPT is scary-good, crazy-fun, and—so far—not particularly evil.

LEX KANTROWITZ

DEC 02, 2022 • 11:30 AM



Google bans deepfake-generating AI from Colab

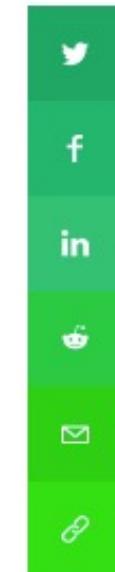
Kyle Wiggers @kyle_l_wiggers / 4:49 PM CDT • June 1, 2022

Comment



 **Image Credits:** Alex Tai/SOPA Images/LightRocket / Getty Images

Google has banned the training of AI systems that can be used to generate deepfakes on its Google Colaboratory platform. The [updated terms of use](#), [spotted](#) over the [weekend](#) by Unite ai and



<https://www.bleepingcomputer.com/news/google/goog-quietly-bans-deepfake-training-projects-on-colab/>

The complete list of disallowed projects are listed below:

- file hosting, media serving, or other web service offerings not related to interactive computing with Colab
- downloading torrents or engaging in peer-to-peer file sharing
- using a remote desktop or SSH
- connecting to remote proxies
- mining cryptocurrency
- running denial-of-service attacks
- password cracking
- using multiple accounts to work around access or resource usage restrictions
- creating deepfakes

How much is technical?

- What is the difference between a geek and a dork?
- Programmers in corporate environments don't struggle with syntax, documentation, troubleshooting. They struggle with managing relationships.
- When do you ask for help? Who do you ask? (Type I errors and Type II errors both cut into productivity)
- In academic environments, specialization of knowledge (stovepiping) means elite researchers just collaborate to get the expertise they need.



'geek'



'nerd'

Try to use the spells in the spellbook for good?



Improvements for DATA221 in the spring?