Deep-Fake Detector

By: Ryan Joseph

How do can you recognize an AI-generated Image over a real one?

IF YOU CAN'T BEAT THEM JOIN THEM

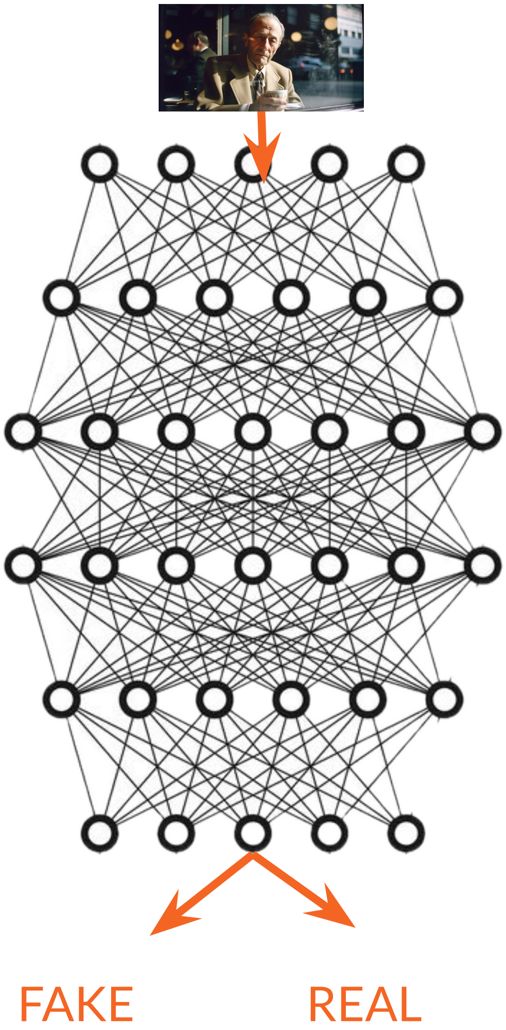
Goal: CREATE AN AI THAT CAN DISTINGUISH BETWEEN AI GENERATED IMAGES AND REAL IMAGES

Scope

I’m Not Jesus You Only Get One Wish

- AI Generated Content
  - Images (yes)
  - Video (no)
  - Text (no)
  - Sound (no)

Methods



Convolutional Neural Network

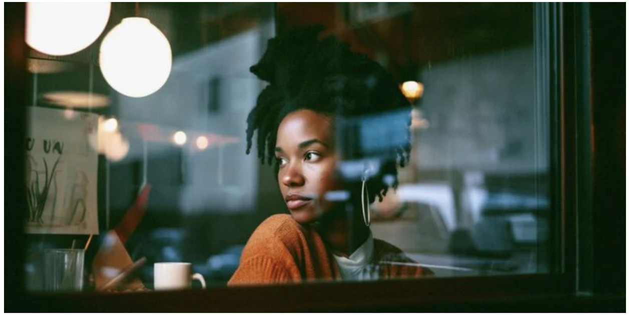
- ResNet-50

Dataset

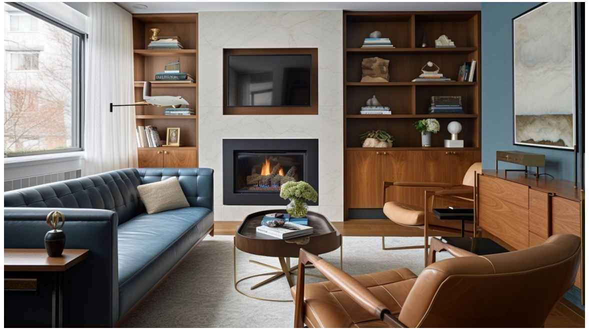
- GenImage (>1,000,000 Fake/Real Images)

You Try…

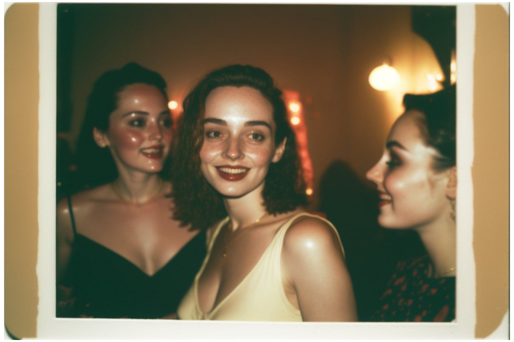
1



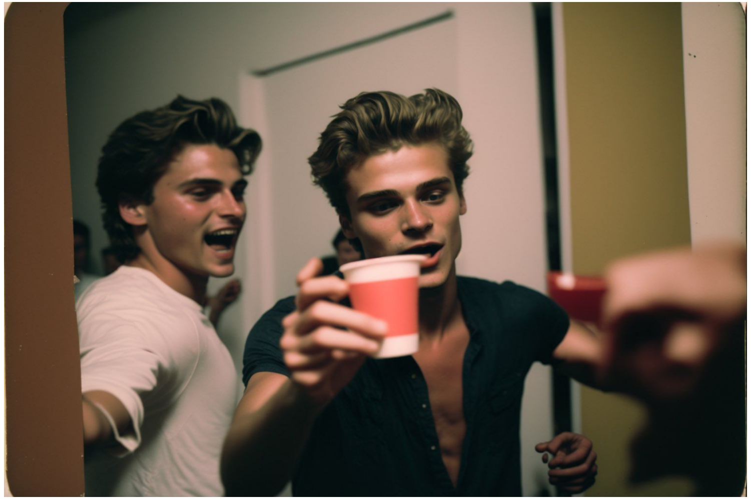
2



3



4



WHICH 2 ARE FAKE AND REAL?

ALL OF THEM WERE FAKE…

TRAINING

NOBODY’S PERFECT

- TRAINED ON M1 MACBOOK
- 20 EPOCHS
- TRAINING ACCURACY: 70%
- VALIDATION ACCURACY: 69%

LET’S GIVE IT A TRY

(DEMO)

For Teacher Purpose

- In the demo I will upload a photo I took of myself so everyone knows it must be real. I will put that photo in my AI model and see if it can rightfully predict its a real image and not AI-generated
- In my demo I will ask the class what AI-generated image to create using the Midjourney AI generator
- I will then re-format the image and put it in my AI-Detector model to see if it can detect it is fake
    - *The entire class will see if my model is worth its salt :)*

Conclusion

It is possible to detect AI generated images

The pace for AI generation > detection

- Detection for this reason will not be a long-term solution in regards to Digital Forensics

Long Term Solution is for AI Image Generation Companies (Midjourney) to create watermarks ingrained

DON’T WORRY I DID THE GRADING ALREADY :)

A project presentation that includes:

- A slide (or more) of project goals that explains what you aim to do and what you want to accomplish ✅

- A slide (or more) of scope statement outlining what is possibly related but not going to be accomplished by the project. This is to protect you from overpromising and not delivering. ✅

- A slide (or more) of methods that you used. ✅

- A slide for introducing the demo if you have one. For projects that did not include programming or demo, your presentation should include in depth explanation of your work, but should not exceed 7-minutes in total. ✅