Face Detection Disruptions

Josh Honig, Samantha Fleming, Charlotte Prevost, Arrianna Szymczak

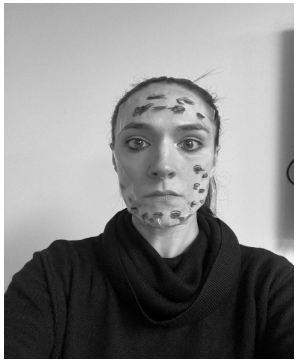
Synthetic Media Attacks

DISGUISE ATTACK - Altering appearance before capture

GOAL ⇒ Make system unable to detect a face

How: Altering or occlude facial landmarks using makeup, clothing, tape, masks, glasses, etc.

SUCCESSFUL DETECTIONS



Tape with pock marks

Tape

Black tape partially obfuscating facial landmarks

Some occlusions do not seem to hinder the system's ability to detect and extract a face...

ERRONEOUS FEATURE EXTRACTION

While some do:

Altering feature landmarks makes the system unable to detect the face properly.

Extractions here show where the system “saw a face.”

UNSUCCESSFUL DETECTION > ROTATION

NO FACE DETECTED

Altering captured image

Using Face-Morph (https://github.com/Azmarie/Face-Morphing/tree/master)

What Face-Morph does

- Detects / aligns faces in images
- Generates corresponding features points between the two images using Dlib's Facial Landmark Detection
- Calculates the triangular mesh with Delaunay Triangulation for each intermediate shape
- Warps the two input images towards the intermediate shape, perform cross-dissolve and obtain intermediate images each frame

GOAL → Make system recognize a fake face.

Process:

- Choose original face that is detected
- Choose which face to morph into
- Morph face
- Extract frames from the morph video
- Check whether morphed face is still close in distance to authentic picture

Pattern-Based Denial of Service Attacks

Goals

- Dynamic generation of “mosaic”
- Use real/synthetic faces, not patterns
- Use genuine face parts
- Trick algorithms other than Viola-Jones Haar Cascade
- Dynamic generation prevents blacklisting of faces

Background Image

We **do** want:

- Confusion
- Noise
- Potential for false detections
- Creepiness

We **don’t** want:

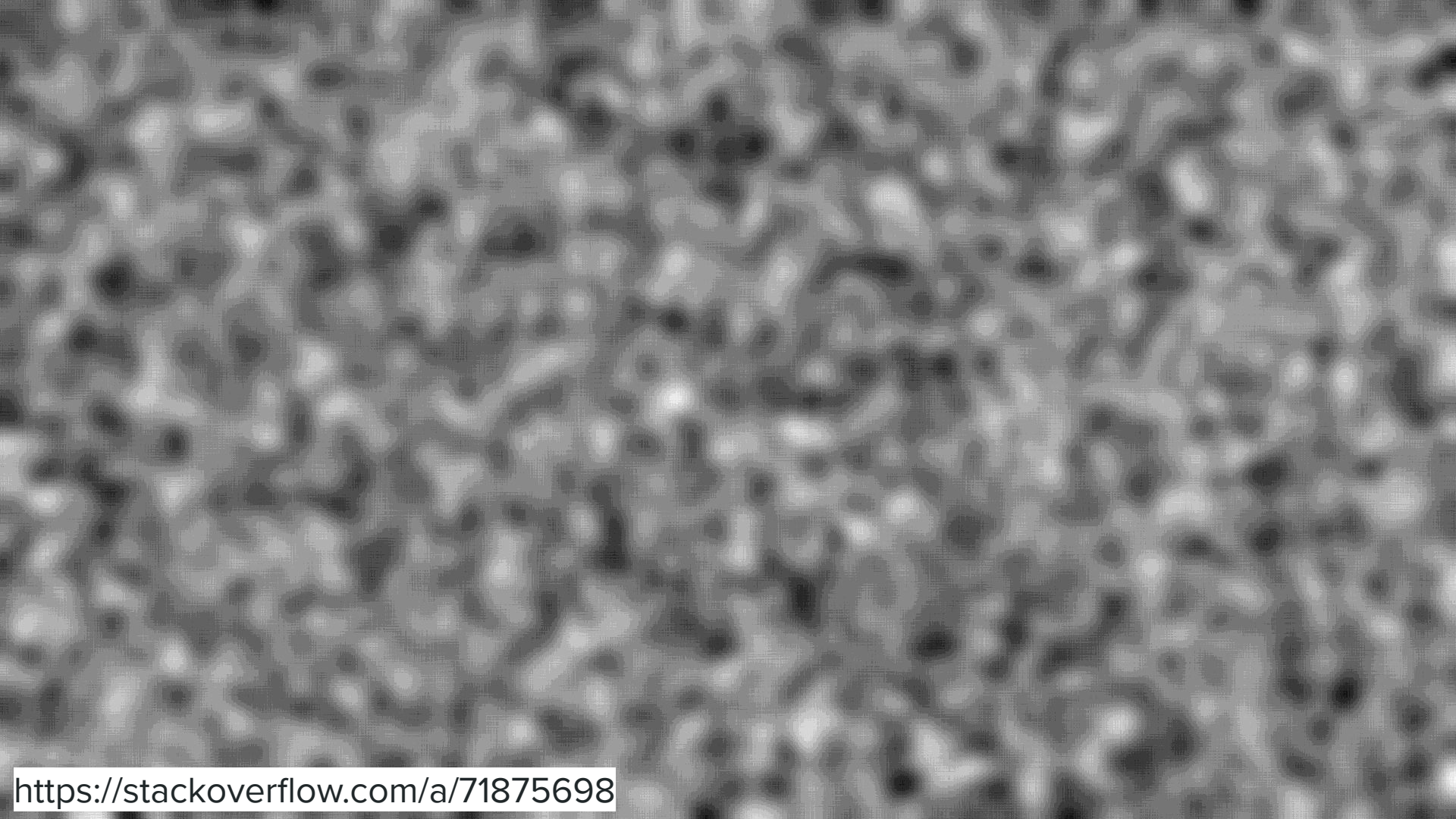
- Consistency
- Detectability
- Boring
- Computationally expensive

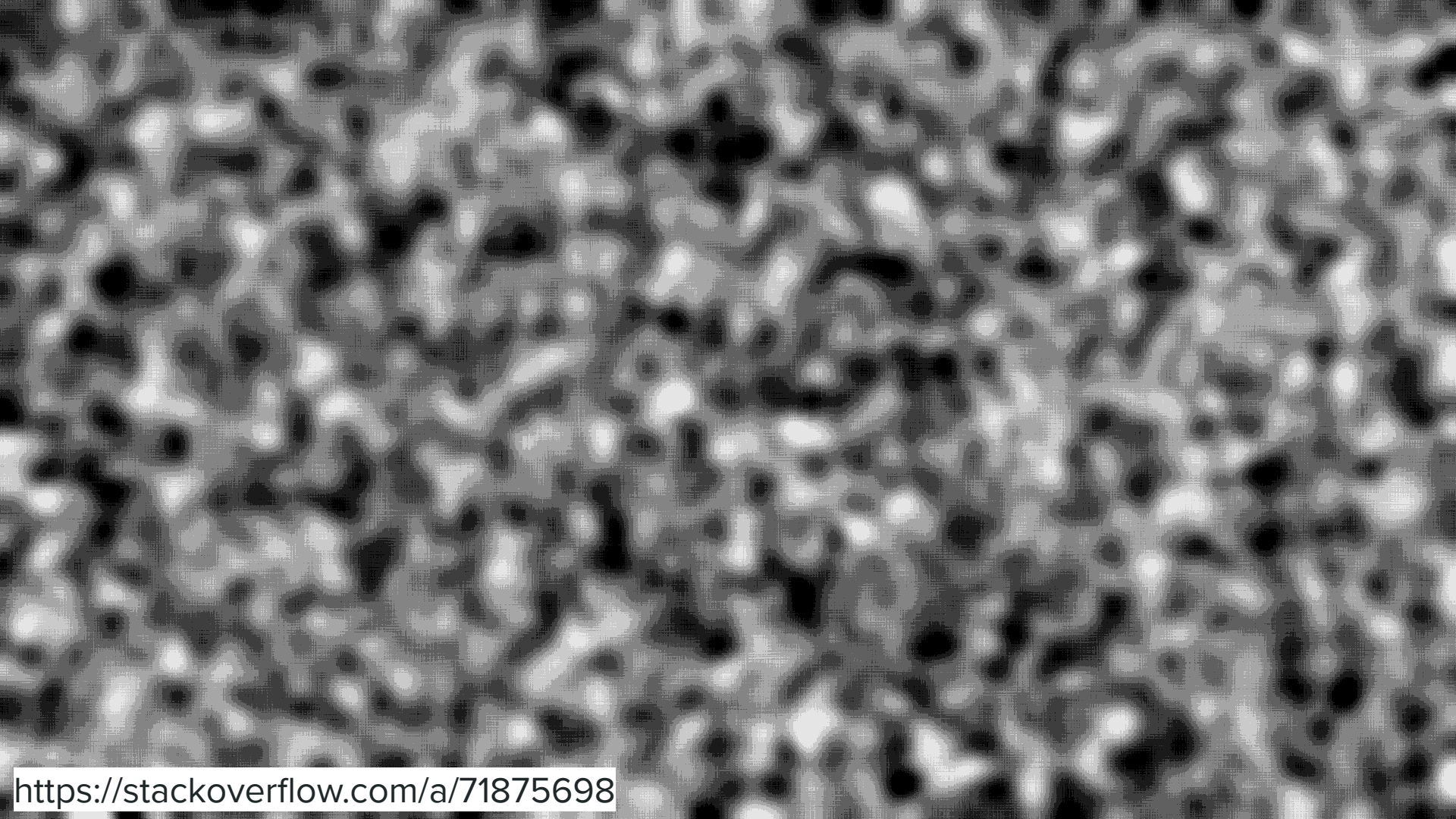
We **don’t care** about:

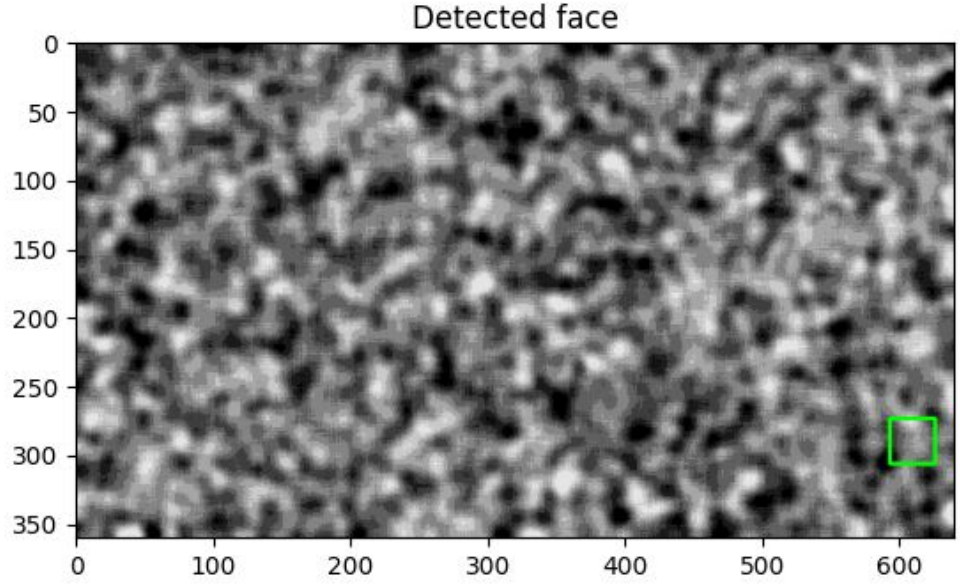
- What’s actually in the image (e.g., nature scene)
- Image size or complexity

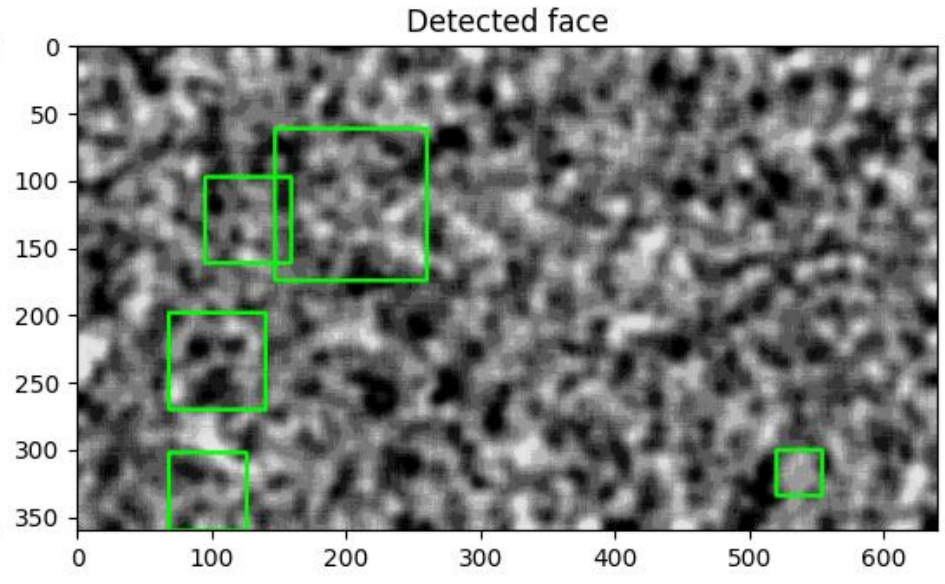
https://stackoverflow.com/a/71875698

https://stackoverflow.com/a/71875698


https://stackoverflow.com/a/71875698


https://stackoverflow.com/a/71875698


Detected face
0
50
100
150
200
250
300
350
0
100
200
300
400
500
600


Detected face
0
50
100
150
200
250
300
350
0
100
200
300
400
500
600

The process

- Get some faces
- Extract the faces
- Perform some image manipulation on the extracted images
- Paste images on top of background
- Run image through face detection algorithm
- Figure out which combination of factors is the best

...but what faces do we use? and what parts of those faces?




JUSTIN XU · UPDATED 3 YEARS AGO

30

New Notebook

Download

Person Face Dataset (thispersondoesnotexist)

Dataset with 10K images of person face generated by thispersondoesnotexist.

Data Card | Code (6) | Discussion (0) | Suggestions (0)

About Dataset

This dataset contains 10K images of 1024×1024 person face generated by This Person Does Not Exist using unofficial "API" by David Lorenzo.

Usability

8.75

License

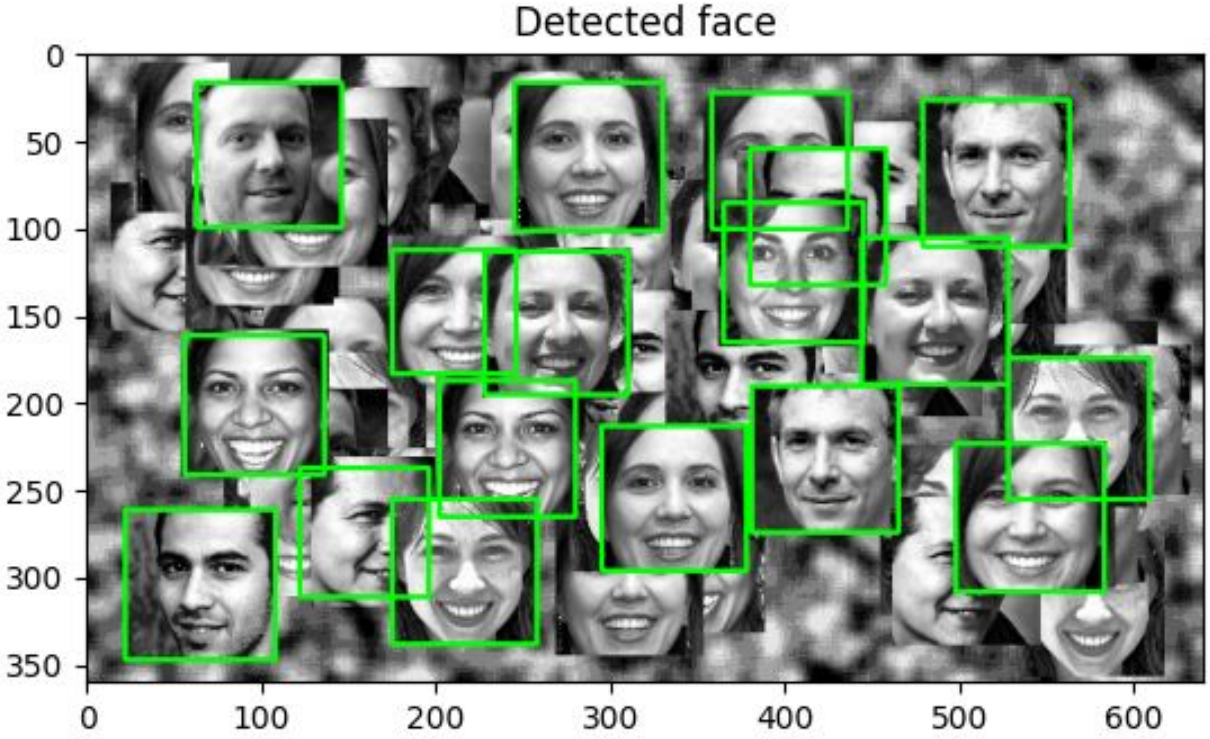
CC0: Public Domain

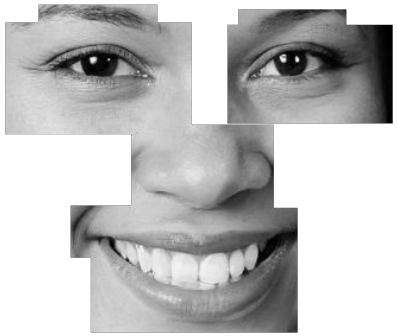
Expected update frequency

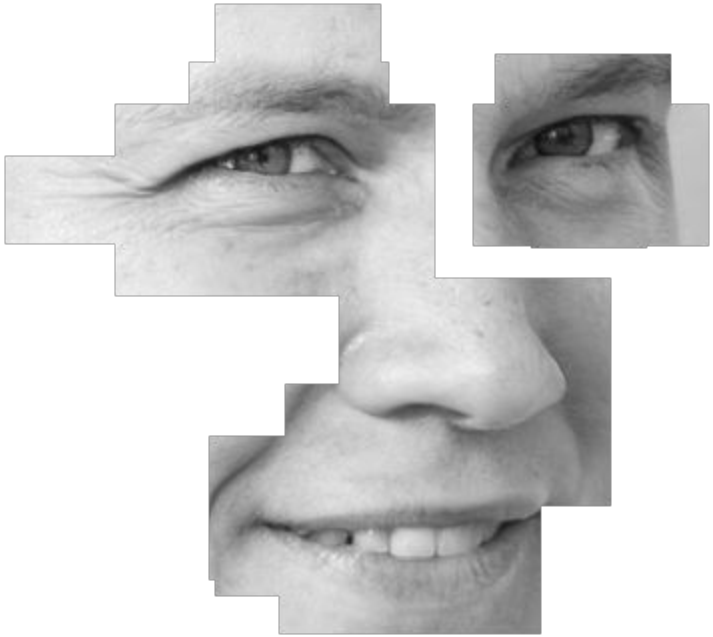
Never

Tags

Image | Computer Vision | People | Adversarial Learning

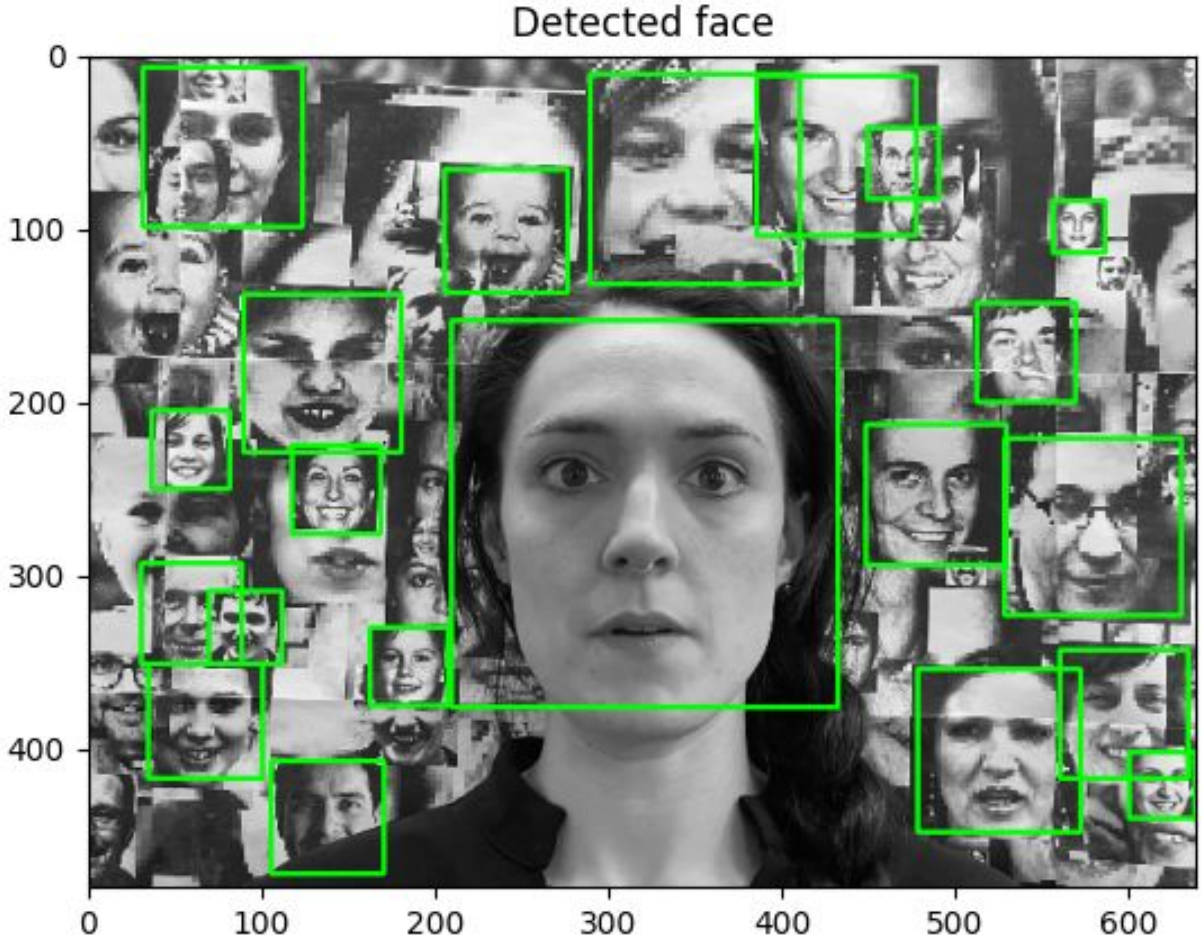

Detected face
0
50
100
150
200
250
300
350
0
100
200
300
400
500
600

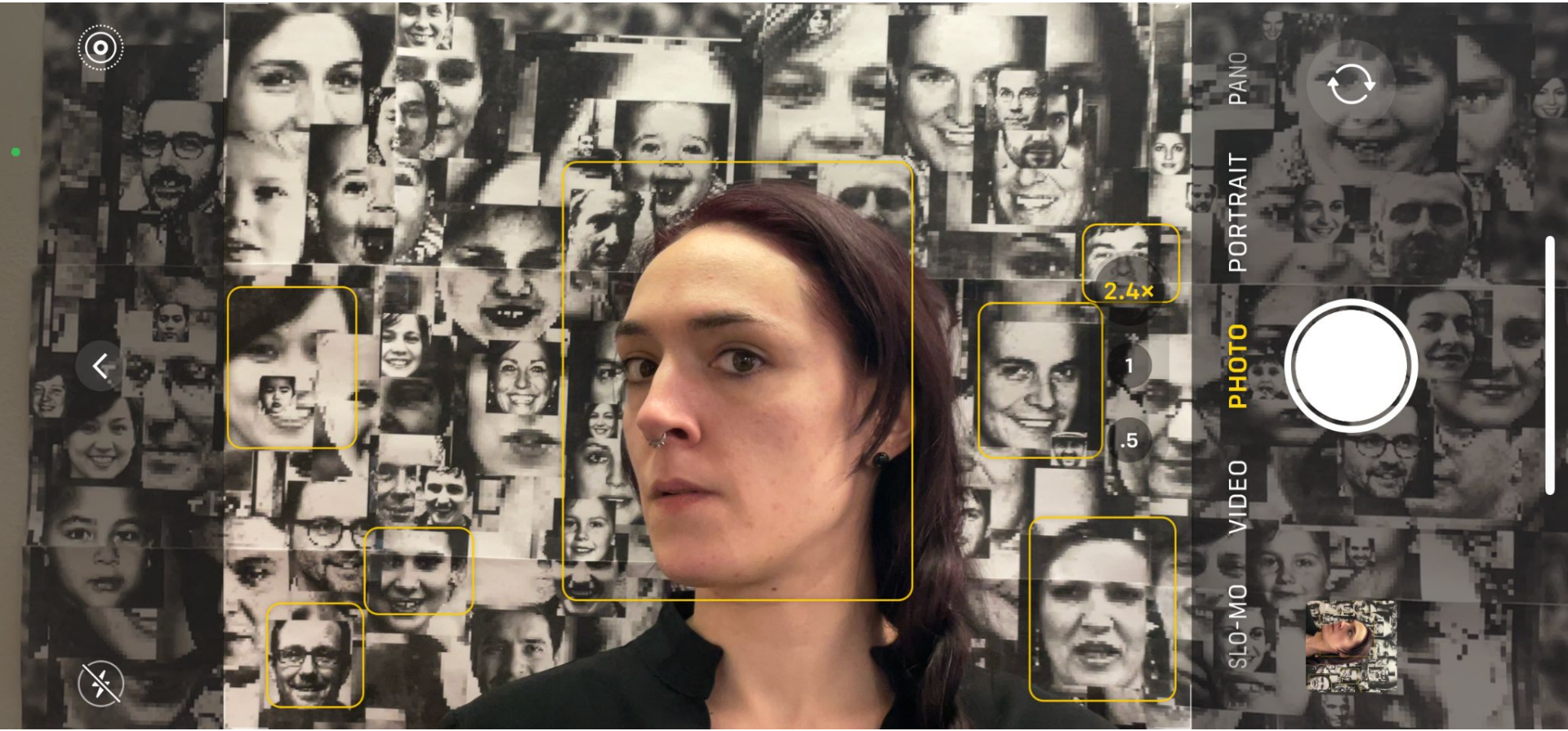


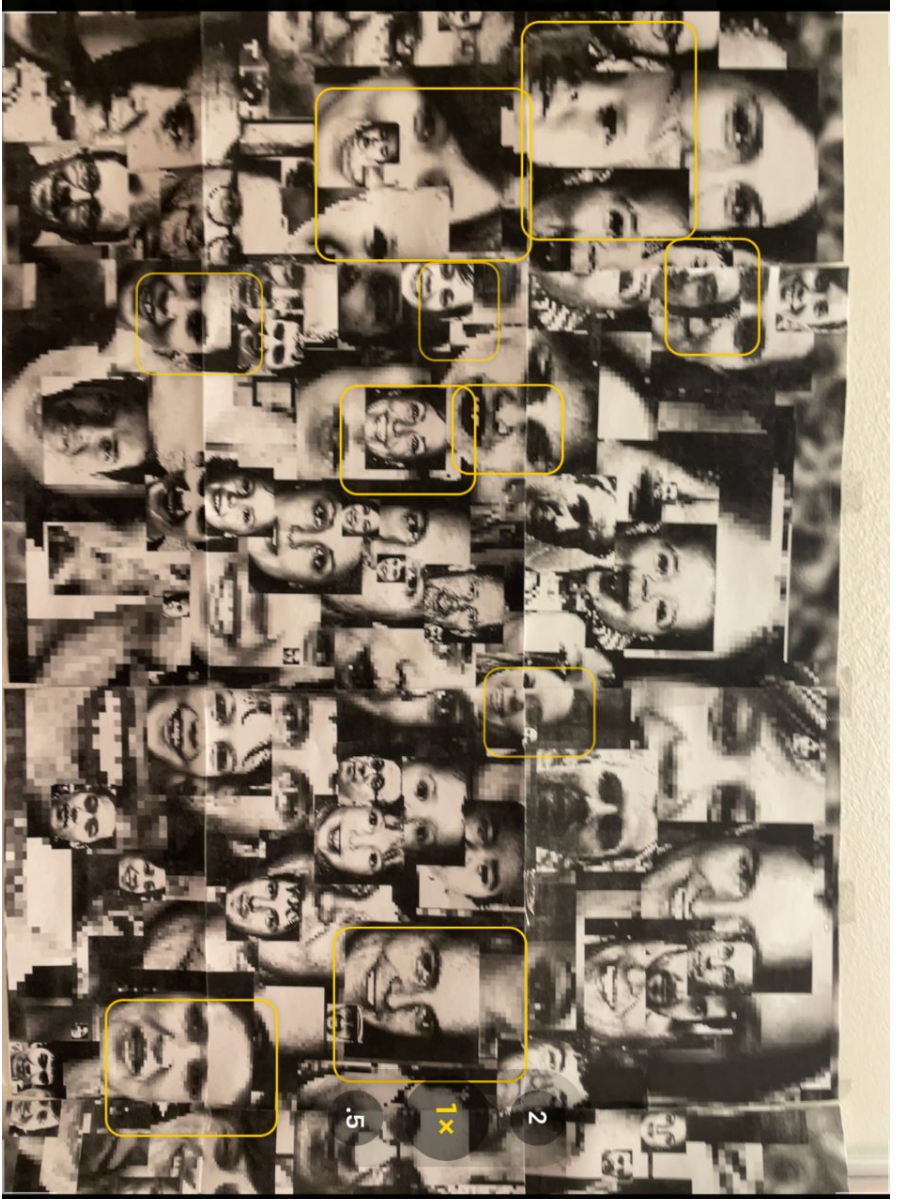



Detected face
0
50
100
150
200
250
300
0
100
200
300
400
500
600


Detected face
0
100
200
300
400
0
100
200
300
400
500
600


Detected face
0
100
200
300
400
0
100
200
300
400
500
600


2.4×
1
.5
SLO-MO
VIDEO
PHOTO
PORTRAIT
PANO



SLO-MO VIDEO PHOTO PORTRAIT PANO

Face Alteration Attacks / Poisoning

Glaze

-Developed by University of Chicago

-Intended use is for art to be cloaked and prevent ai from training on art style

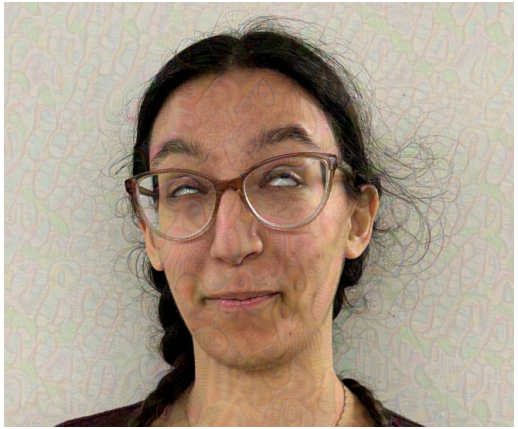
-uses calculations to edit pixels

Hypothesis

- Will the edited pixels confuse facial detection?
- How effective are the strength models on glaze?

Results

- Glazing an image did not affect facial detection
- Did it affect distance?
- We saw that in some images that were compared that some had higher distances on the default setting than compared to the high setting with glaze intensity
- Overall, there was some increase in distance between low, default, and high glaze intensity



FIN