SAND Lab
security, algorithms, networks and data

Understanding Non-Consensual "Undress" applications

Josephine Passananti, Heather Zheng, Ben Zhao

CSCW Digital Intimacy Workshop — Oct 18, 2025

⚠ Trigger warning: non-consensual imagery references

Evolution of *Synthetic* Non-Consensual Intimate Imagery (*S*NCII)

< 2020

SNCII not an "industry" yet

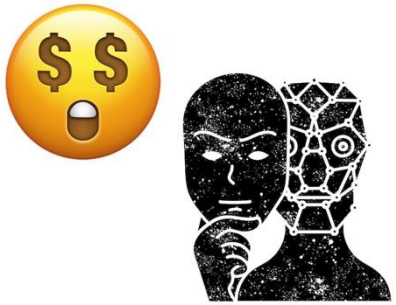
- limited to physical & digital media

- Leaked photos & videos

2020-2022

GAN based face-swapping tools

- Not widely marketed



- Low-quality, expensive SNCII

2023 >

Generative *Undress AI* tools

- **2,000% increase** in spam for "undress AI" services
- sites received **over 24 million visitors per month**

- High quality, low cost SNCII

SNCII *aka* “undress AI” is used maliciously

… 2025

Terms and Conditions hold the user responsible for his or her use of the app

EL PAÍS

MANUEL VIEJO

International

Madrid - SEP 18, 2023 - 12:44 EDT

MEXICO · CHINA · LATEST NEWS

ARTIFICIAL INTELLIGENCE >

In Spain, dozens of girls are reporting AI-generated nude photos of them being circulated at school: ‘My heart skipped a beat’

FBI Sacramento
FBI Sacramento Press Office
(916) 746-8106

Twitter Facebook Email

January 17, 2024

Sextortion: A Growing Threat Preying Upon Our Nation’s Teens

6abc ACTION NEWS 24/7 Live | Philadelphia Pennsylvania New Jersey Delaware

NJ high schooler calls for AI regulations after manipulated pornographic images shared online

Los Angeles Times SUBSCRI

CALIFORNIA

More than 1 in 10 students say they know of peers who created deepfake nudes, report says

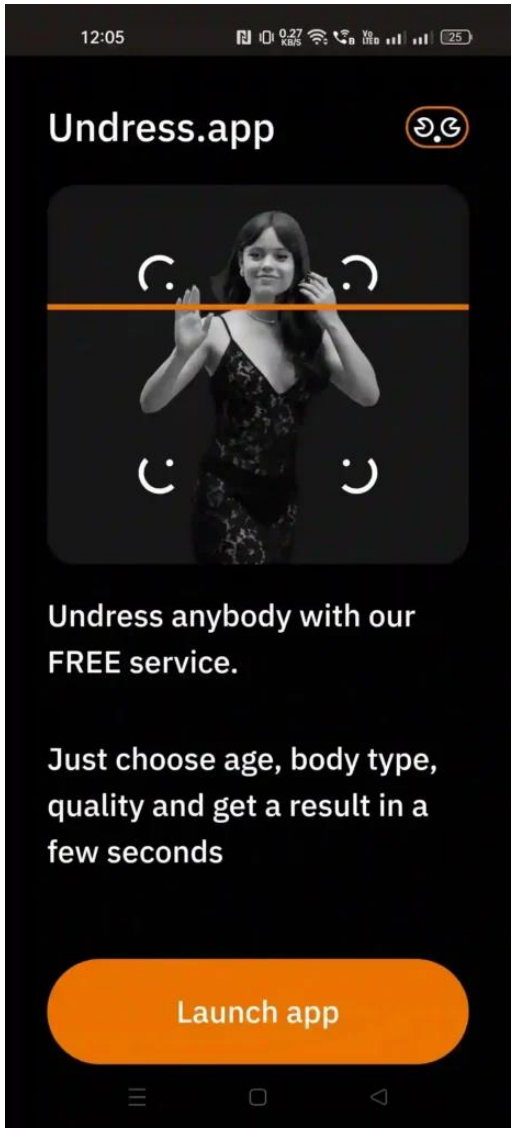
The New York Times
https://www.nytimes.com › 2024/04/08 › technology

Teen Girls Confront an Epidemic of Deepfake Nudes in ...

Apr 8, 2024 — **Teen Girls Confront** an **Epidemic** of **Deepfake** Nudes in Schools. Using artificial intelligence, middle and high school students have fabricated ...

Investigating how Undress AI is so pervasive

How easy is it to create an AI nude?

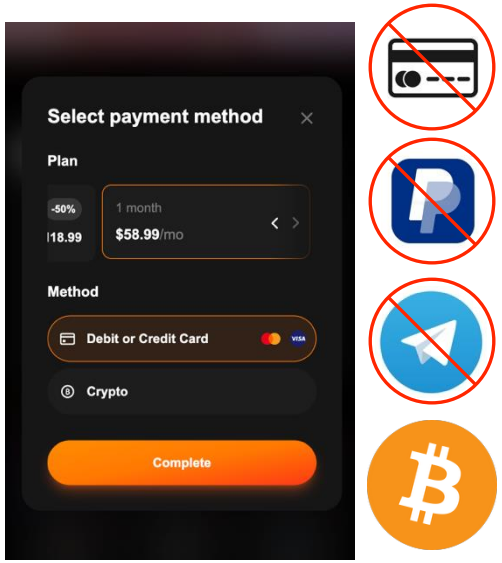


Investigating how Undress AI is so pervasive

Our experience:

1. Email address

2. Payment



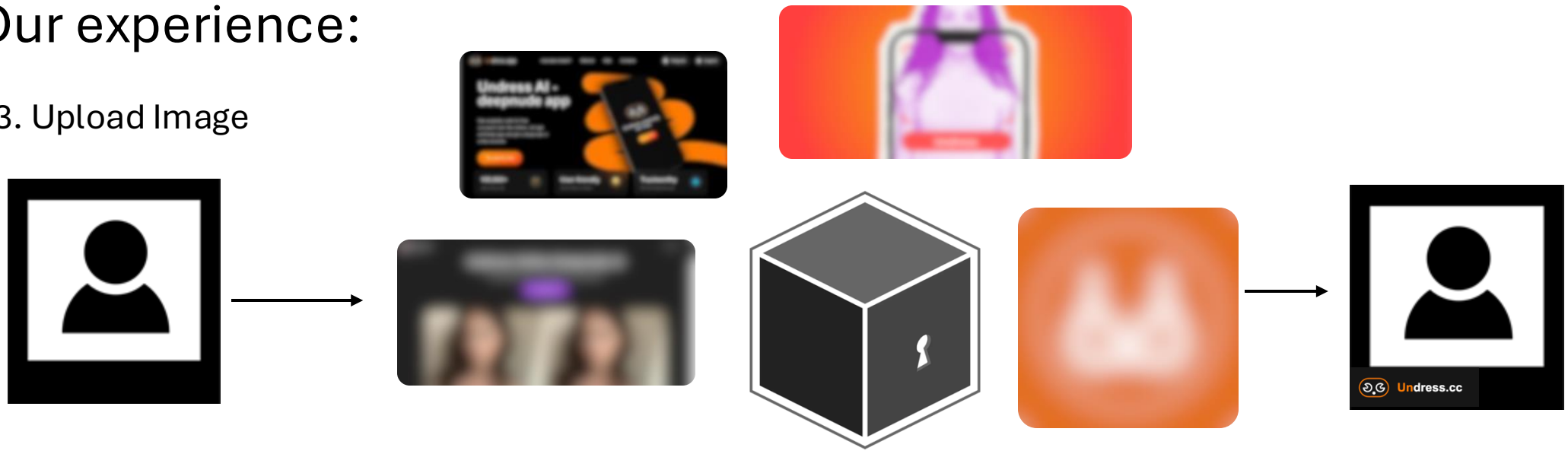
How are teenagers buying credits with Crypto?

Drug Dealer Model?

Investigating how Undress AI is so pervasive

Our experience:

3. Upload Image



Black Box

How to reverse engineer the pipeline?

- Probe applications **ethically**:

Synthetic Images

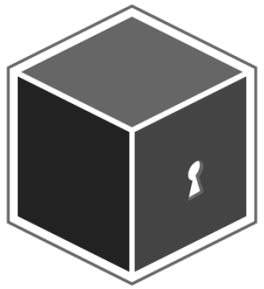
Minimize queries

Investigating applications through probing

Undress AI application

Input Image

Upload



Download

“Undressed” Image

Investigating applications through probing

Undress AI application

Input Image

Upload

Image Segmentation Model:

“identify face + clothes”

+

Inpainting Model:

“remove clothes”

Download

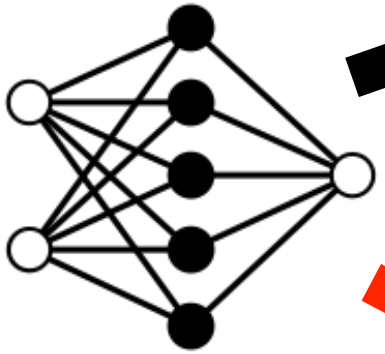
“Undressed” Image

Expensive to train: likely using pre-trained, open-source models

Crafting Adversarial Probes

Input image

Segmentation model



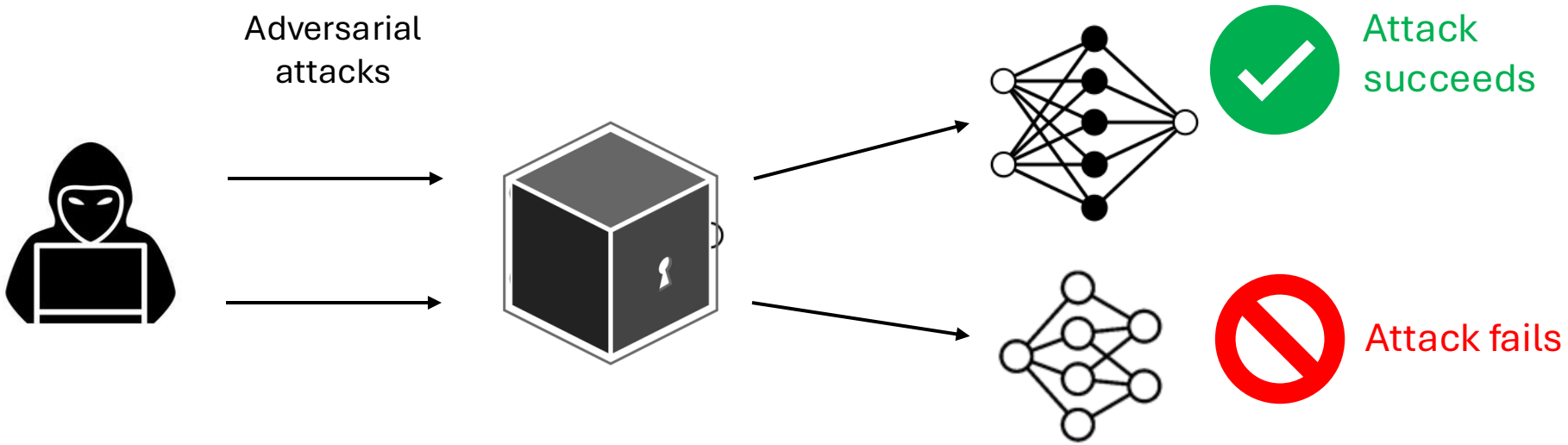
← Correct segmentation

Protected adversarial input image

← **Disturbed** segmentation

Running attacks against *black box* models

Successful adversarial attacks require knowledge of the target model



Each successful or failed probe can reveal something about the model

Attacks on AI Undress Applications

Fine line between successful vs unsuccessful attacks

Input

Attack:
Pixel-Based
Segmentation

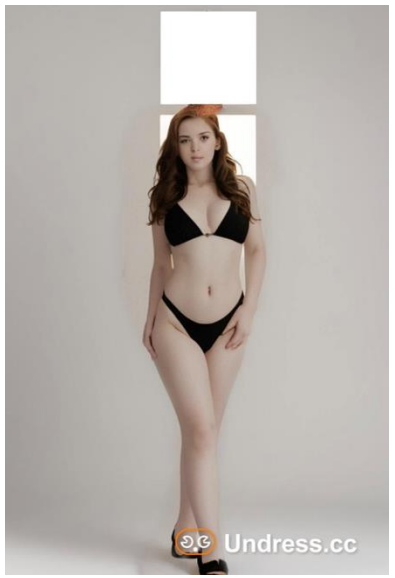
Output

Attack *fails*

Attack:
Censored Face

Attack *fails*

Attack:
Pixel-Based
Segmentation
+
Censored Face



Attack *succeeds*

Attacks on AI Undress Applications

Fine line between successful vs unsuccessful attacks

Clues: what a successful segmentation attack would look like

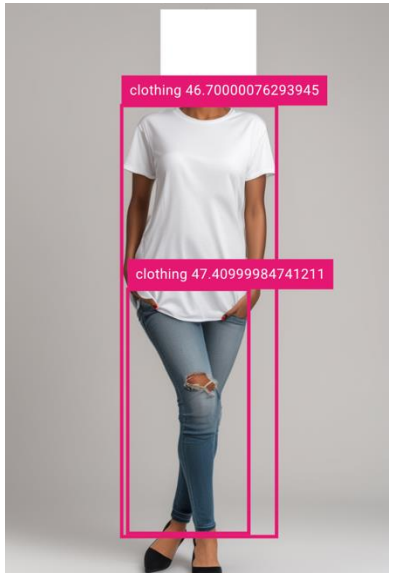
Input

Attack:
Pixel-Based
Segmentation

Segmentation

Attack *fails*

Attack:
Censored Face



Attack *fails*

Attack:
Pixel-Based
Segmentation
+
Censored Face

Attack *succeeds*