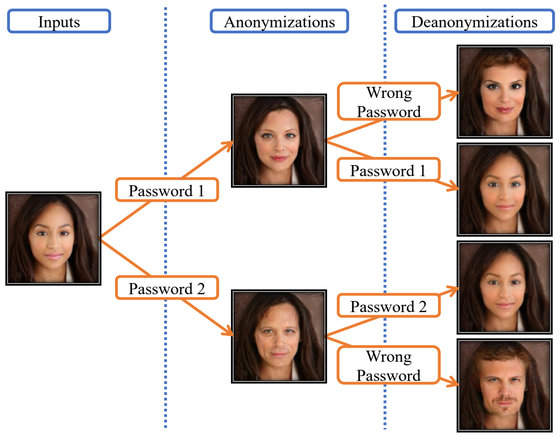
# Password-conditioned Anonymization and Deanonymization with Face Identity Transformers

Xiuye Gu xiuyegu@stanford.edu

## Introduction

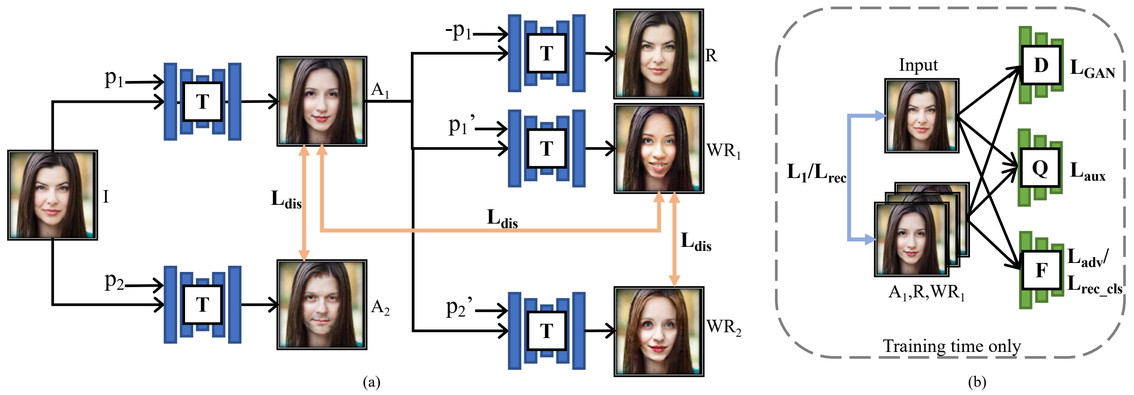
- Cameras are prevalent in our daily lives, and enable many useful systems built upon computer vision technologies such as smart cameras and home robots. However, there is also an increasing societal concern as the captured images/videos may contain privacy-sensitive information (e.g., face identity).
- We propose a novel face identity transformer which enables automated photo-realistic password-based anonymization as well as deanonymization of human faces in visual data.
- Extensive experiments on CASIA, LFW, FFHQ datasets show that our approach enables multimodal password-conditioned face anonymizations and deanonymizations, without sacrificing privacy compared to existing anonymization approaches.




- Our system never stores users' faces (a) on disk, and instead only stores the anonymized faces (b).
- When a user provides a correct recovery password, s/he will get the deanonymized face back (c).
- If a hacker invading their privacy inputs a wrong password, s/he will get a face whose identity is different from the original as well as the anonymized face (c).

Photo-realism is meant to fool the hacker by providing no clues as to whether the real face was recovered.

## Network Architecture

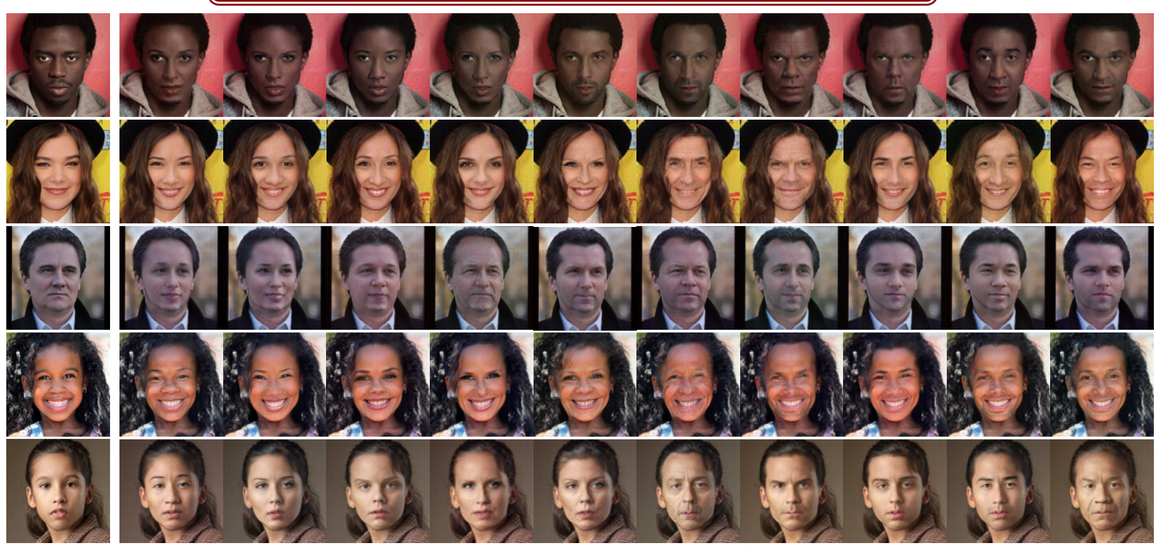



I: Input image, $A_{1,2}$ : Anonymized faces, R: Recovered face, $WR_{1,2}$ : Wrongly Recovered faces.

## AMT Perceptual Studies

- *Anonymization and deanonymization:* when asked "Are they the same person?", turkers reported "yes" **4.7%/100%/0.7%/1.3%** of the time on I vs A / I vs R / I vs WR / A vs WR (low, high, low, low is ideal. 150 pairs per test, 3 turkers per pair).
- *Multimodality:* Turkers reported "yes" **12.2%** and **2.7%** of the time on $A_1$ vs $A_2$ and $WR_1$ vs $WR_2$.
- *Photo-realism:* Turkers label our anonymizations as being more real **30.10%** of the time and label our wrong reconstructions as more real **15.60%** of the time (100 pairs per test, 10 turkers per pair).

## Multimodality Qualitative Results

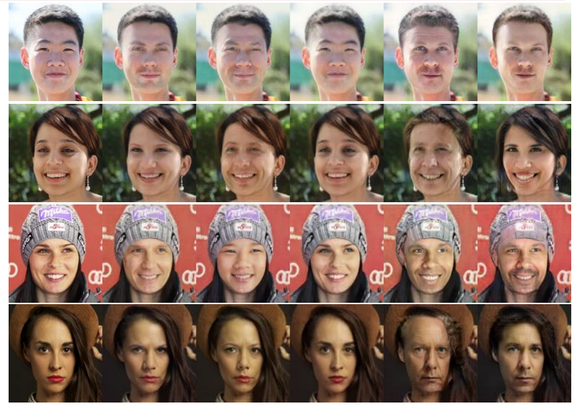


(a) Input (b) Anonymization results conditioned on different passwords

## Hard cases on CASIA Dataset

(a) I (b) $A_1$ (c) $A_2$ (d) R (e) $WR_1$ (f) $WR_2$

## Generalization on FFHQ Dataset



(a) I (b) $A_1$ (c) $A_2$ (d) R (e) $WR_1$ (f) $WR_2$

## Ablation Studies