LOMONOSOV MOSCOW STATE UNIVERSITY

FACULTY OF MECHANICS AND MATHEMATICS

AdvHat: Real-World Adversarial Attack on ArcFace Face ID System

Stepan Komkov[1,2], Aleksandr Petiushko[1,2]

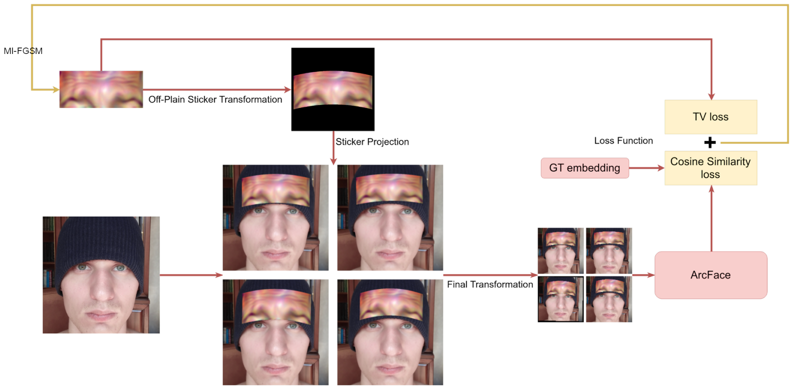
[1]Lomonosov Moscow State University, [2]Huawei Moscow Research Center

HUAWEI

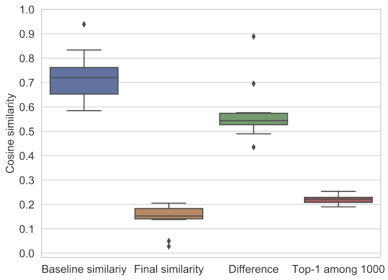
MAIN CONTRIBUTION

A novel easily reproducible technique to attack the best public Face ID system ArcFace in different shooting conditions by printing the special rectangular paper sticker on a common color printer and putting it on the hat

Main pipeline



Simulation results

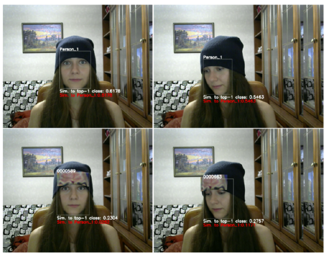


Blue: Anchor VS image w/ hat. **Orange:** Anchor VS image w/ AdvHat. **Green:** “Blue”-”Orange”. **Red:** Max sim to CASIA.

Real-world simulation

Frontal face (advhat: no)
Similarity to origin: 0.61

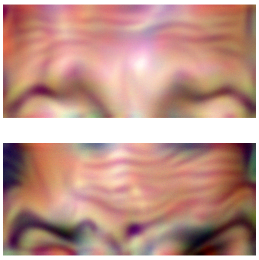
Frontal face (advhat: yes)
Similarity to origin: 0.02
Similarity to other: 0.23



Rotated face (advhat: no)
Similarity to origin: 0.54

Rotated face (advhat: yes)
Similarity to origin: 0.11
Similarity to other: 0.27

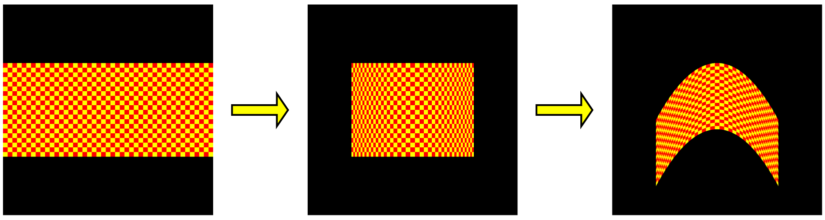
Patch examples



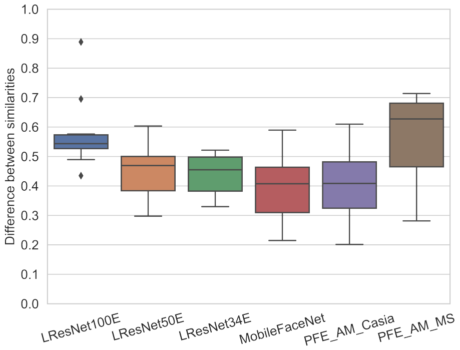
Off-plane transformation

Parabolic transformation in the 3d space: $(x_0, y_0, 0) \rightarrow (x_1, y_0, b \cdot x_1^2)$ so as

$$x_1 = b \cdot \left(|x_0| \cdot \sqrt{x_0^2 + \frac{1}{4 \cdot b^2}} + \frac{1}{4 \cdot b^2} \cdot \ln \left(|x_0| + \sqrt{x_0^2 + \frac{1}{4 \cdot b^2}} \right) - \frac{1}{4 \cdot b^2} \cdot \ln \frac{1}{2 \cdot b} \right)$$



Transferability


Difference between similarities
LResNet100E
LResNet50E
LResNet34E
MobileFaceNet
PFE_AM_Casia
PFE_AM_MS