**Introduction**

Automated face recognition models can be used for tracking activities and relationships of image sharing platform users. Convolutional Neural Networks (CNNs) are previously proposed adversarial perturbation-based approaches which are not practical for real-world applications.

### Practical Requirements:

- **Black-box Attack**: Users do not know about target CNNs.
- **Low Computational Cost**: Users have a few personal images and limited computational resources.
- **Low Storage Cost**: Users do not want to keep a perturbation per image (storage burden).
- **Recoverability**: Users want to recover the original images.
- **Recognizability**: Users want to have recognizable images.
- **Compatibility**: The proposed approach must be practical on all platforms.

### Proposed Schemes

**Universal Ensemble Perturbation (UEP):**

- Uses small CNNs trained only on 10 classes ⇒ Low computational cost
- Trains CNNs locally ⇒ Black-box scheme
- Learns a universal transferable perturbation ⇒ Low storage cost
- Adds perturbation to arc-tangent hyperbolic space of image ⇒ Low loss recovery

\[ z_{\text{perturbed}} = \frac{1}{2} \left[ \text{tanh} \left( \text{arctanh}(z - 0.5) \right) + \beta \times 6 \right] + 0.5 \]

**K-Randomized Transparent Image Overlay (k-RTIO):**

- Semantic-based adversarial perturbation ⇒ Low computational cost
- Uses a secret key and ID of the source image to generate a unique overlay image ⇒ Low storage cost
- Easy to recover ⇒ Reversibility
- No CNNs required for generating perturbations ⇒ Black-box scheme

### Results

- **Dataset**: 1000 images sampled from FaceScrub celebrities’ face dataset
- **Face detection and recognition models**
  1. DeepFace [CVPR2014]
  2. Clarifai.com
  3. Google Vision API

### References


**Contact Information**

Arezoo Rajabi
University of Washington
Email: rajabia@uw.edu

Rakesh Bobba
Oregon State University
Email: rakesh.bobba@oregonstate.edu

**Acknowledgements**

This work has been published in Privacy Enhancing Technologies Symposium (PoPETs) 2021 when Arezoo Rajabi was PhD student at Oregon State University.