LEARNING SENSITIVE IMAGES USING GENERATIVE MODELS

Samson Cheung University of Kentucky

H. Wildfeuer, M Nikkah, X. Zhu, W. Tan Innovations Lab, Cisco Systems Inc.

Can we "share data" for learning but preserve privacy?

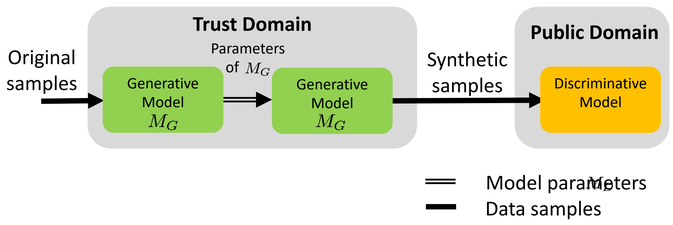
Privacy challenges

- Raw data contains sensitive information (e.g., health, financial): cannot leave premise
- May not trust cloud
- Traditional encryption protects storage and transfer

Desirable attributes

- No changes to central learner
- Extensible to multiple data sources from different entity
- Privacy with minimum performance loss in support of learning tasks.
- Quantifiable privacy protection

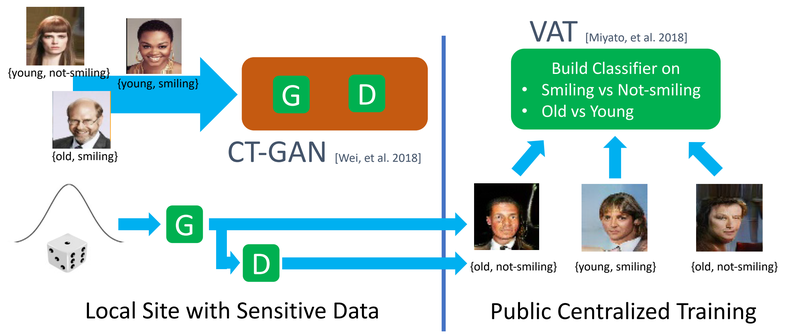
Generative Model Approach



Use GAN trained on sensitive data to generate synthetic surrogates

- Medical domain: Lung disease [Beaulieu-Jones et al., 2017], ICU Time series [Hyland, Esteban & Ratsch, 2017], EHR [Choi et al., 2017]
- This work focus on facial images

Learning Structure & Performance



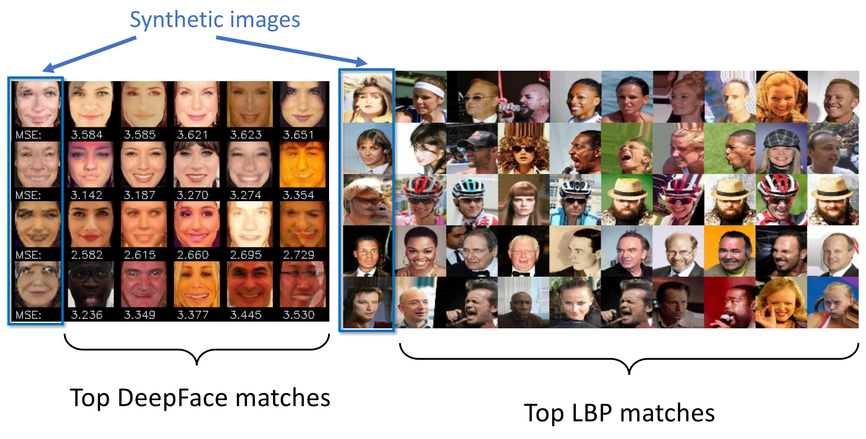
Source	CACD-2000 age	Celeb-A age	Celeb-A smile
original	77.2%	81.2%	92.0%
synthetic	76.5%	78.4%	91.6%

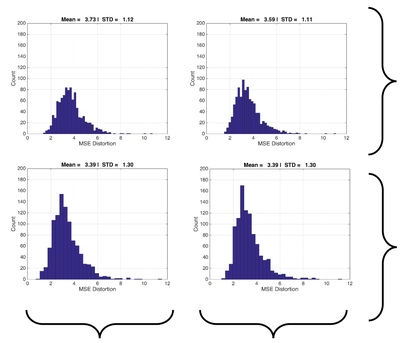
Table 1: Comparison of classifier accuracy when generated images are substituted for originals.

Preserving Privacy

Possible measures

1. "DeepFace" distance [Taigman et al., 2014]
2. Local Binary Pattern [Chen et al., 2014]





Histogram of MSE distance of top-10 matching images from training and test datasets against 1000 queries of GAN-generated images

Categories	GAN vs Train	GAN vs Test
Young	0.537 +/- 0.04	0.510 +/- 0.04
Old	0.491 +/- 0.04	0.495 +/- 0.04
Smiling	0.523 +/- 0.04	0.546 +/- 0.04
Non-smiling	0.511 +/- 0.05	0.545 +/- 0.05

Categories	GAN vs Train	GAN vs Test
Young	0.517 +/- 0.07	0.539 +/- 0.07
Old	0.528 +/- 0.06	0.533 +/- 0.06

Table 2: Statistical comparisons of distributions of minimum distances for Celeb-A and CACD-2000

Key Take-aways

- GAN is a promising approach for protecting facial privacy and preserving learnability in distributed learning.
- Other GAN-based approaches use Differential Privacy, which are shown to be unnecessary in our applications.