



Privacy Preserving Biometrics

Arun Ross

Professor

Michigan State University

rossarun@cse.msu.edu

<http://www.cse.msu.edu/~rossarun>

Importance of Privacy

- “Privacy is the right to be **let alone**” [Samuel Warren and Louis Brandeis (1890)]
- “Privacy is the claim of individuals, groups, or institutions to **determine for themselves** when, how, and to what extent information about them is communicated to others” [Alan Westin (1970)]
- “Privacy is the right of people to **conceal information** about themselves that others might use to their disadvantage” [Richard Posner (1983)]

The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.

PRIVACY IS DIFFERENT FROM SECURITY

Biometrics

- Automated **recognition** of individuals based on their **biological** and **behavioral** characteristics
- Traits from which **distinguishing, repeatable** features can be extracted

C. G. Brown

Height	1 m 71.6	Head l'gth	19.8	L. Foot	27.1	Circle	Leh	Age	22	Born in	
Eng. Height	5-10 3/4	Head width	16.34	L. Mid. F.	11.2	Periph Z		Apparent Age			
Outs. A	1 m 75.5	Cheek width	14.4	L. Lit. F.	8.7	Color of Left Eye	Leh - Mel	Nativity	<i>Louisville, Ky.</i>		
Trunk	94.9	R. Ear	6.8	L. Fore A.	46.6	Pecul		Occupation	<i>Johnson</i>		

Remarks Incident to Measurement

DESCRIPTIVE

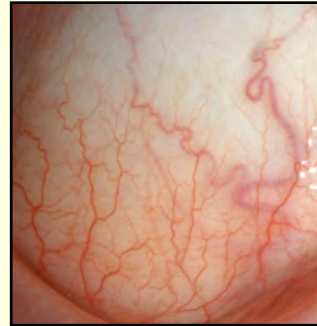
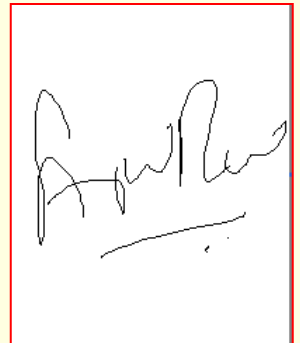
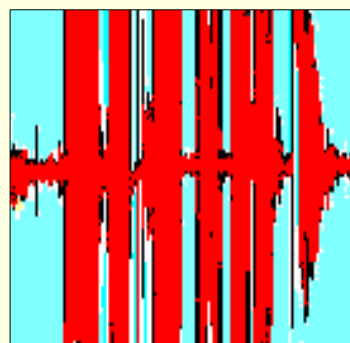
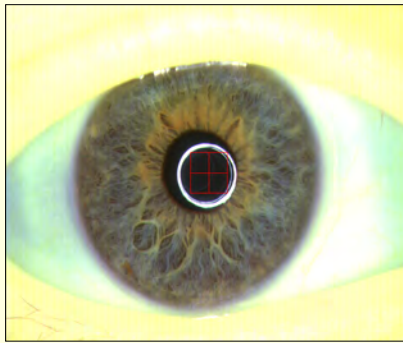
Incl	<i>Reddy</i>	Ridge	<i>None</i>	Beard	<i>Shaved</i>		
Height	<i>M</i>	Base	<i>(Ear) Blue</i>	Hair	<i>Black</i>		
Width	<i>Brn</i>	DIMENSIONS			Complexion	<i>M. Dark</i>	
Pecul		Length	<i>Er</i>	Projection	<i>Er</i>	Weight	<i>165</i>
		Breadth	<i>M</i>	Teeth	<i>Upper front over top</i>	Build	<i>M. Slim</i>
		Pecul		Chin	<i>M. Prom</i>		

BUREAU OF IDENTIFICATION
Department of Police,
Tulane Ave. and Saratoga St.
New Orleans, La.

Measured *Feb 1 1912*
By *Geo. G. Jones*

H.T. F. Rhodes, Alphonse Bertillon: Father of Scientific Detection, Abelard-Schuman, 1956

Biometric Traits



Identity vs Recognition

- Biometrics **does not** explicitly elicit **identity**
- Biometrics **recognizes** the trait of a person

INPUT



Based on a **single** fingerprint image, we **cannot** say this belongs to *Jane Doe*

REFERENCE



We need a **reference** fingerprint image that is known to belong to *Jane Doe* in order to make this assessment

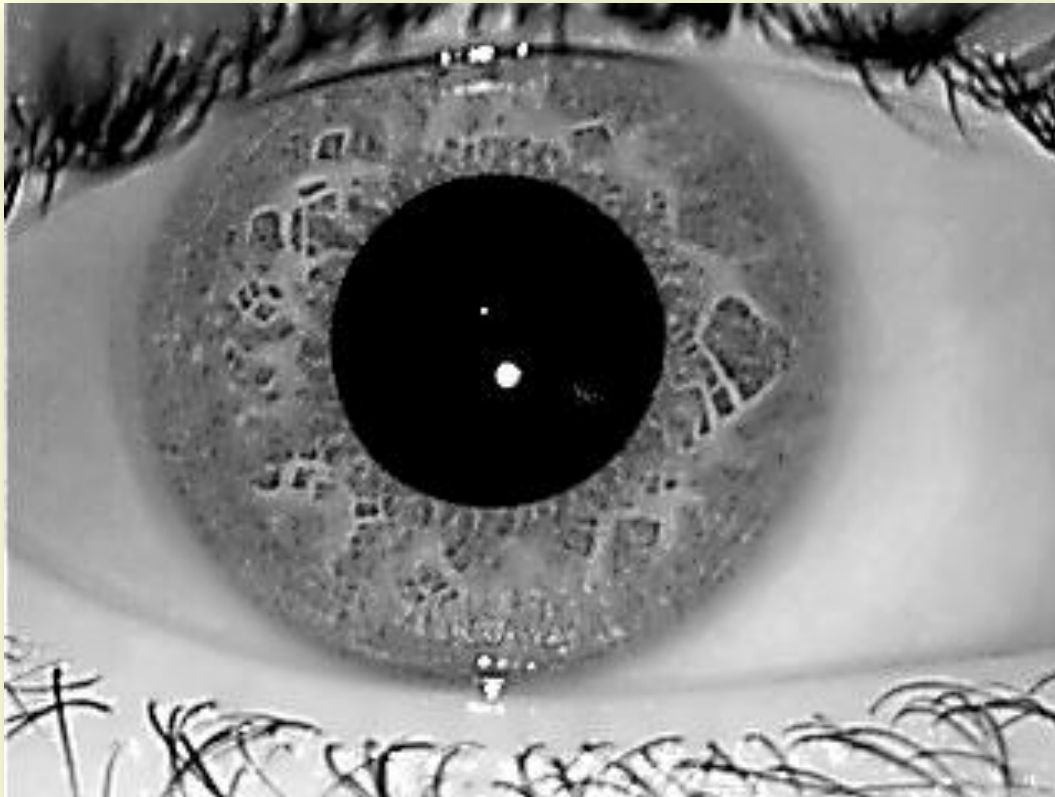
Privacy of Biometric Data

- Age, Gender, Ethnicity, can be **automatically derived** from the face image
- That is, a **trained classifier or a regressor** may be used to automatically deduce certain soft biometric attributes



- Gender: Male
- Age: 25
- Health: Very good
- Eye Sight: Wears glasses
- Ethnicity: Asian Indian

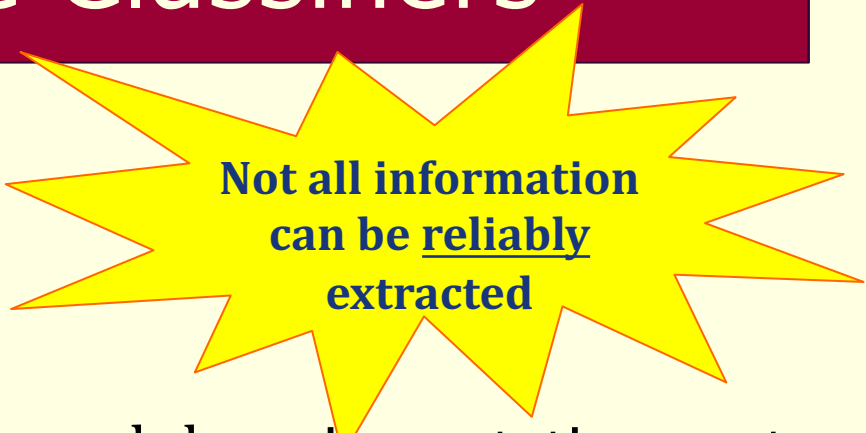
What *else* is revealed in an iris image?



- Viewing the iris as a **textural** entity rather than just a **binary code**

Iris: Attribute Classifiers

- **Biographical:**
Age, Sex, (Race?)
- **Anatomical:**
Distribution of crypts, Wolfflin nodules, pigmentation spots
- **Environmental:**
Sensor, Illumination wavelength, Indoor/Outdoor
- **Pathological:**
Stromal Atrophy
- **Other:**
Pupil dilation level, Contact Lens

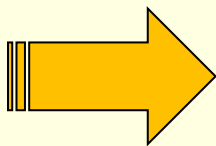
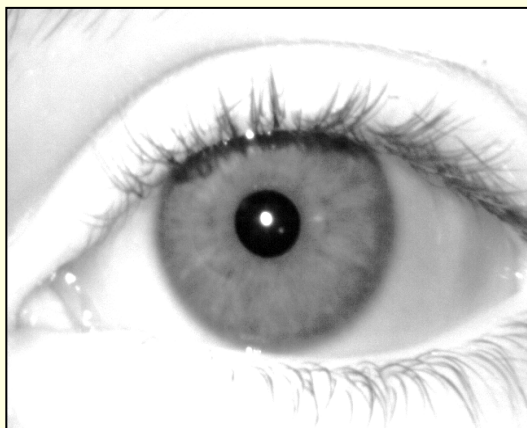


Not all information
can be reliably
extracted



But information
can be aggregated

Biometrics + Forensics



- Subject is a **Male** (90% Confidence), **White** (85% Confidence)
- Image taken using an **Aoptix** camera
- Iris stroma is **plain textured**
- Highly **constricted** pupil suggests **strong ambient illumination**

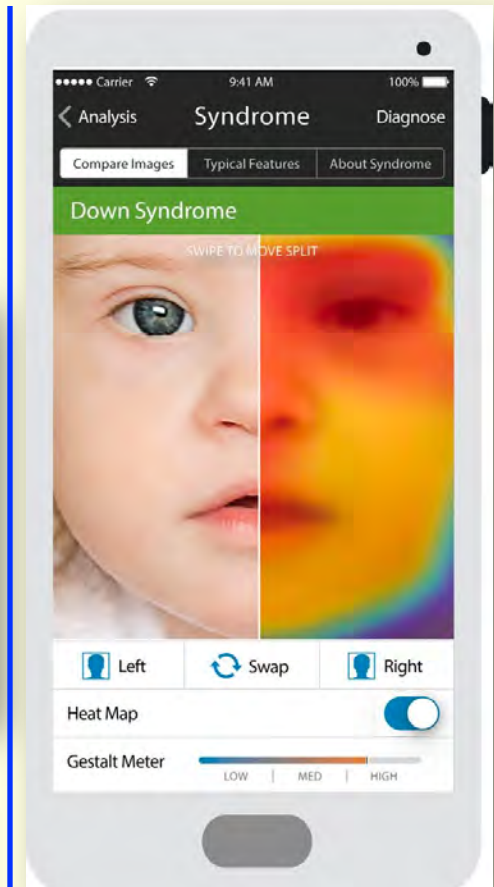
Bridges the gap between human and machine description of data
OR
Compromises privacy?

Face2Gene

MEGAN MOLteni SCIENCE 01.09.17 01:00 PM
**THANKS TO AI, COMPUTERS CAN
NOW SEE YOUR HEALTH
PROBLEMS**

“In hindsight it was all clear to me,” says Gripp, who is chief of the Division of Medical Genetics at A.I. duPont Hospital for Children in Delaware, and had been seeing the patient for years. “But it hadn’t been clear to anyone before.” What had taken Patient Number Two’s doctors 16 years to find took Face2Gene just a few minutes.

Face2Gene is a suite of phenotyping applications that facilitate comprehensive and precise genetic evaluations.



Identifying People on the Web


- **Faces of Facebook: Privacy in the Age of Augmented Reality (Alessandro Acquisti)**
- Convergence of three technologies:
 - face recognition, cloud computing, online social networks
- Started from an anonymous face in the street
- Ended up with very sensitive information about that person → **data accretion**
- Combined face recognition with the algorithms they developed in 2009 to predict SSNs from public data

Lawful Data Processing (GDPR)

- Ensure fairness and transparency towards individuals
- **Purpose:** the purpose must be known and declared to the individuals when collecting their data
 - Cannot use the collected data for other purposes not compatible with the original purpose
- **Data Minimization:** do not collect data more than that is necessary for the stated purpose
- **Accuracy:** must ensure the data is accurate and up-to-date
- **Storage:** must not store the data for longer than necessary for the stated purpose
- **Integrity and Confidentiality:** must ensure use of appropriate means to safeguard data against unauthorized access, accidental loss, or damage

Controllable Privacy (Image Level)

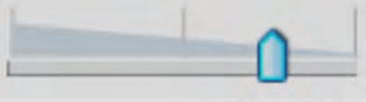
Face Privacy



Input

Output

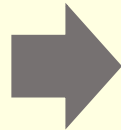
Identity  Age 

Race  Gender 

© Ross/Othman

The image shows a user interface for controlling face privacy. It features two side-by-side portrait photos of a man. The left photo is labeled 'Input' and shows the original face. The right photo is labeled 'Output' and shows the same face with a blurred, less identifiable appearance. Below the photos are four sliders, each with a checkbox and a label: 'Identity' (checked), 'Age' (checked), 'Race' (unchecked), and 'Gender' (unchecked). Each slider has a blue knob indicating its current position. The 'Identity' and 'Age' sliders are positioned towards the right, while 'Race' and 'Gender' are towards the left. The interface is titled 'Face Privacy' at the top.

Controllable Privacy



Vahid	Recognition
Male	Gender
29	Age
White	Ethnicity
Healthy, 197 lb.	Health



Retained

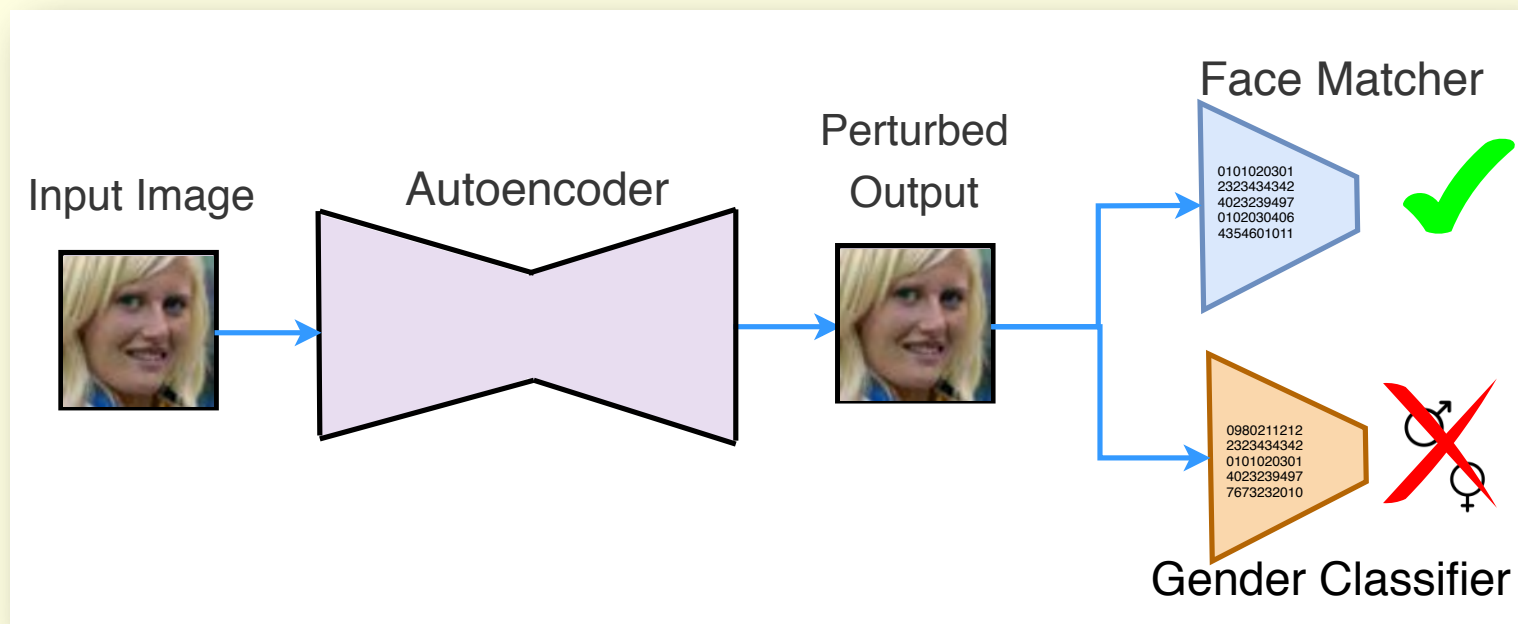


Confounded

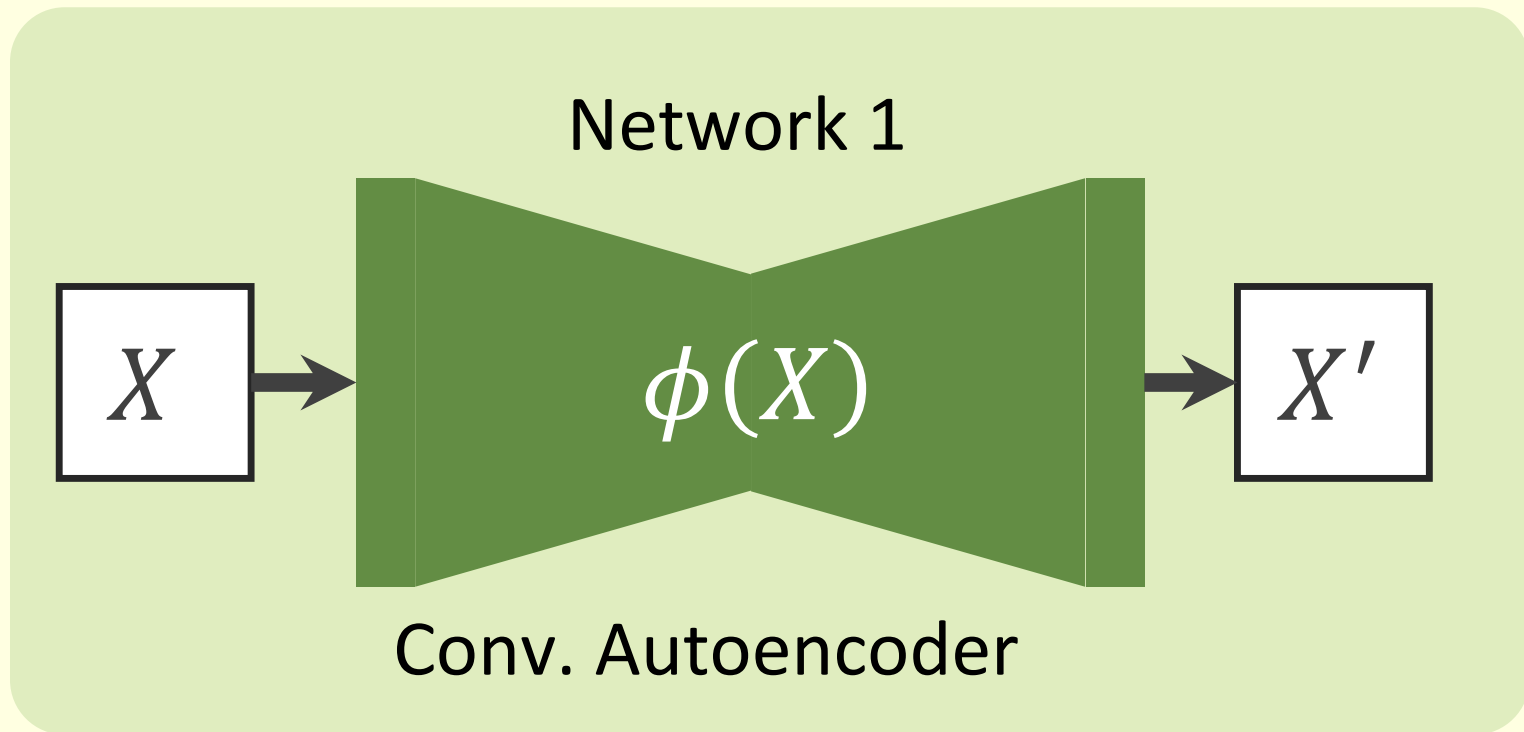
Mirjalili, Raschka, Ross: IJCB 2017, ICB 2018, BTAS 2018, TIP 2020

Semi-Adversarial Networks (SAN)

- Design a transformation model to:
 - Confound gender attribute → gender classifiers will not work
 - Retain recognition capability → face matchers will still work

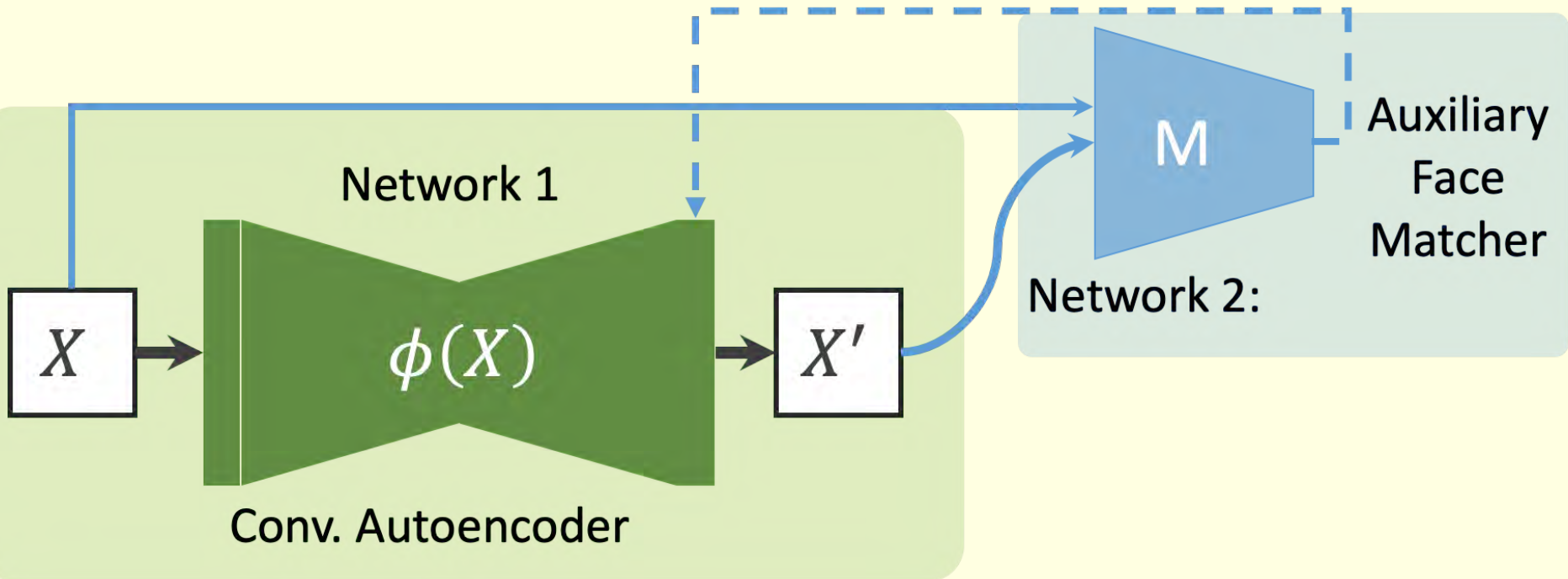


General Architecture of SAN Model

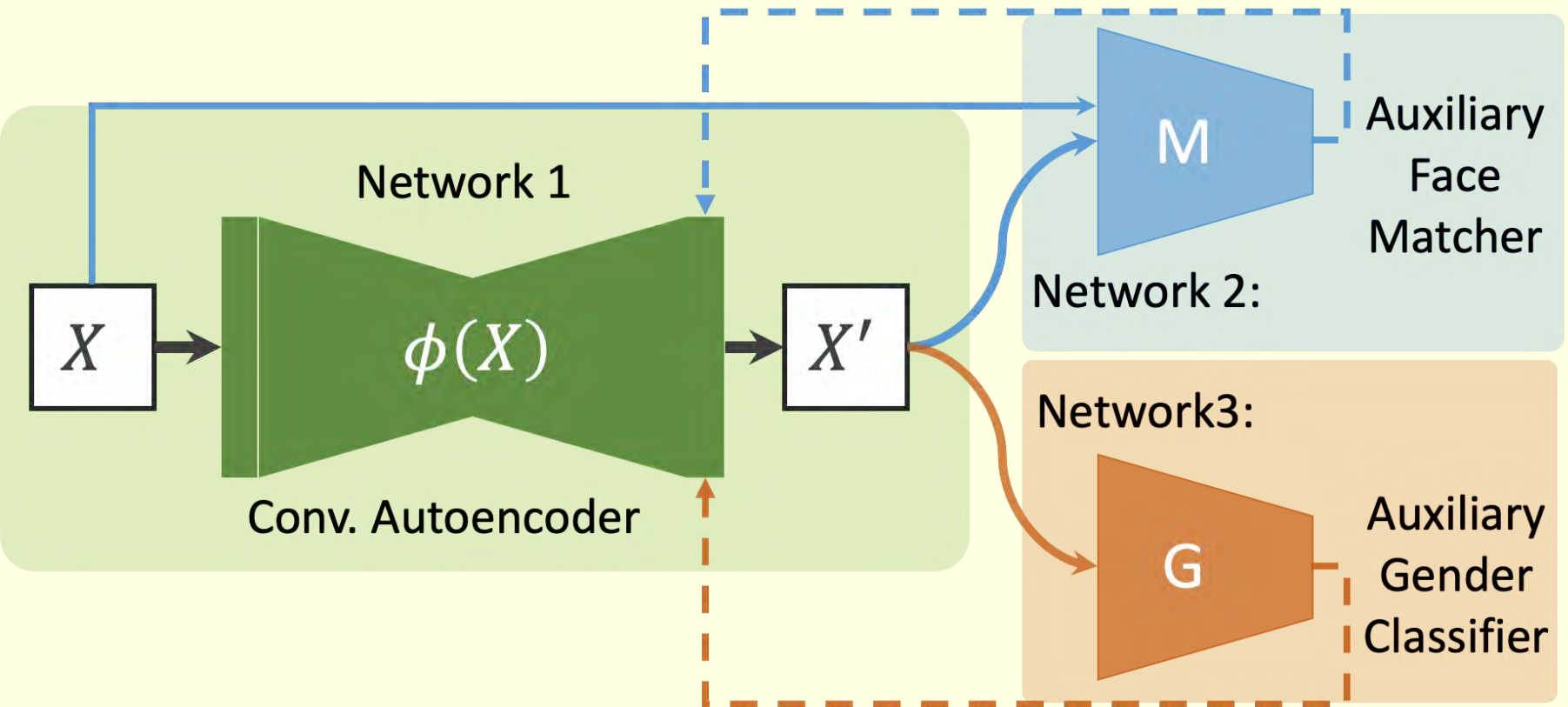


Mirjalili et al., Semi-Adversarial Networks: Convolutional Autoencoders for Imparting Privacy to Face Images, ICB 2018

General Architecture of SAN Model



General Architecture of SAN Model



Loss Functions for Semi-Adversarial Learning

1. Pixel-wise similarity term

$$J_D(X, X'_{SM}) = \sum_{k=1}^N S(X^{(k)}, X'^{(k)}_{SM})$$

- Only used during the pre-training of Autoencoder

2. Loss term related to gender attribute

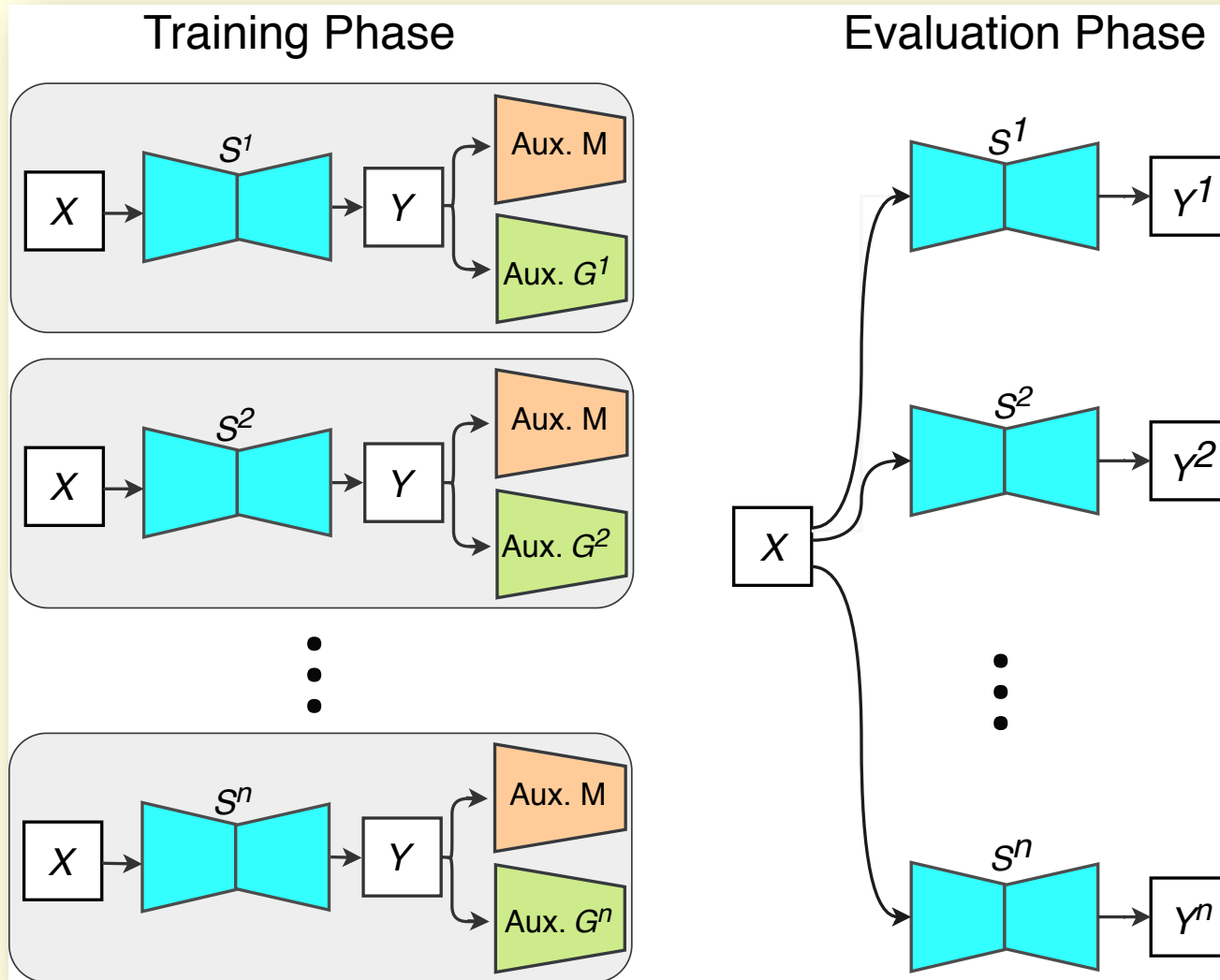
- Correctly predict gender of X'_{SM}
- Flip the gender prediction on X'_{OP}

$$J_G(X, X'_{SM}, X'_{OP}, y; f_G) = S(y, f_G(X'_{SM})) + S(1 - y, f_G(X'_{OP}))$$

3. Loss term related to face identity (recognition)

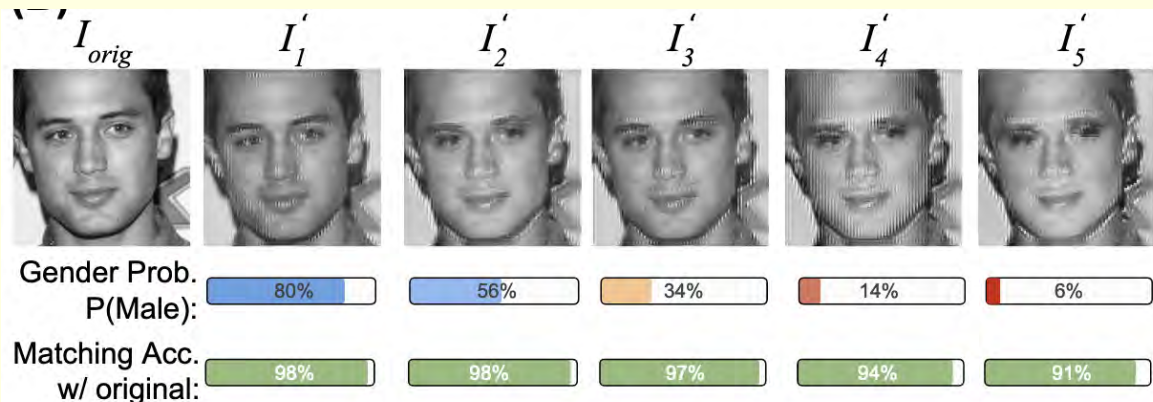
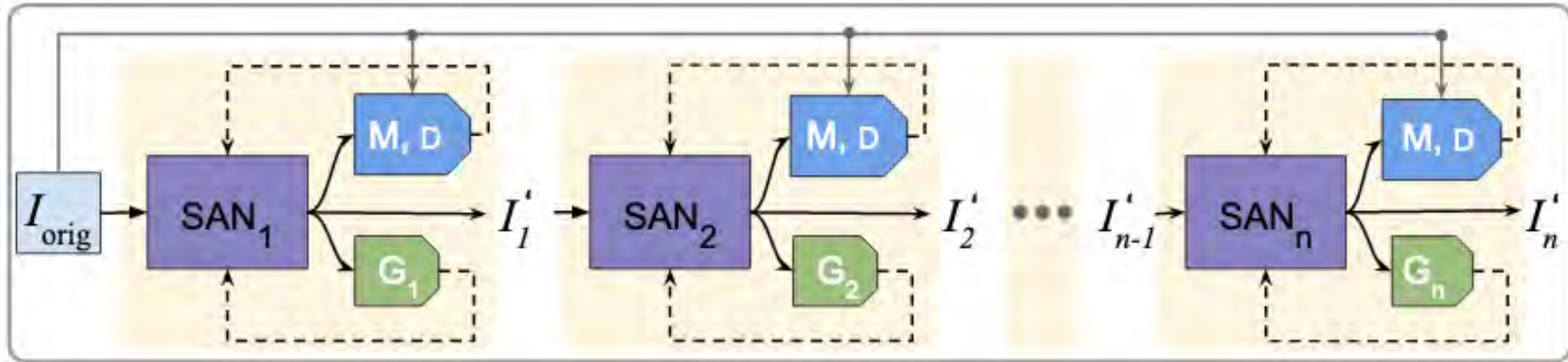
$$J_M(X, X'_{SM}; R_{vgg}) = \left\| R_{vgg}(X'_{SM}) - R_{vgg}(X) \right\|_2^2$$

Ensemble of SANs



V. Mirjalili, S. Raschka, A. Ross, "Gender Privacy: An Ensemble of Semi Adversarial Networks for Confounding Arbitrary Gender Classifiers," BTAS 2018

Sequence of SANs: FlowSAN



V. Mirjalili, S. Raschka, A. Ross, FlowSAN: Privacy-Enhancing Semi-Adversarial Networks to Confound Arbitrary Face-Based Gender Classifiers, "2019

Training Protocol

■ Auxiliary subnetworks

- Auxiliary gender predictor is trained on CelebA dataset, and its parameters are frozen during training of Autoencoder
- Publicly available parameters for VGG are used for the auxiliary face matcher

■ Training the Autoencoder

Step1: pre-training the Autoencoder with two loss terms: pixel-wise similarity + gender term

Step2: replace the pixel-wise similarity term with the matching term based on VGG subnetwork (trained for 20 epochs)

Examples of Inputs and Outputs



Male:
99%



Female:
98%



Male:
97%



Male:
100%



Female:
69%



Male:
99%

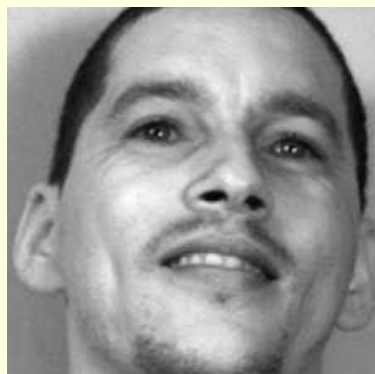


Male:
71%



Female:
58%

Examples of Inputs and Outputs



Male:
98%



Male:
99%



Female:
100%



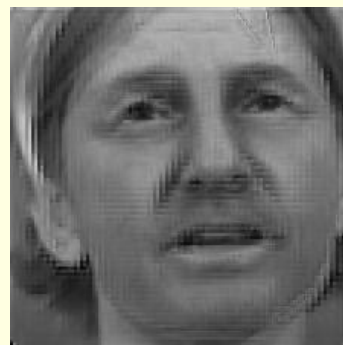
Female:
99%



Female:
79%



Female:
53%



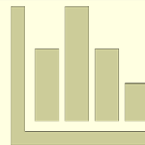
Male:
63%



Male:
67%

Datasets Statistics

Dataset	# Samples	# Subjects	# Male Images	# Female Images
CelebA-train	157,350	--	65,160	92,190
CelebA-test	39,411	--	16,318	23,093
MUCT	3,754	276	1,844	1,910
LFW	12,988	5,658	10,083	2,905
AR-face	3,286	136	1,821	1,465



- CelebA dataset was split into train and test
- CelebA-train was used for training the autoencoder as well as the auxiliary gender predictor

Experimental Design

■ Six unseen gender Classifiers

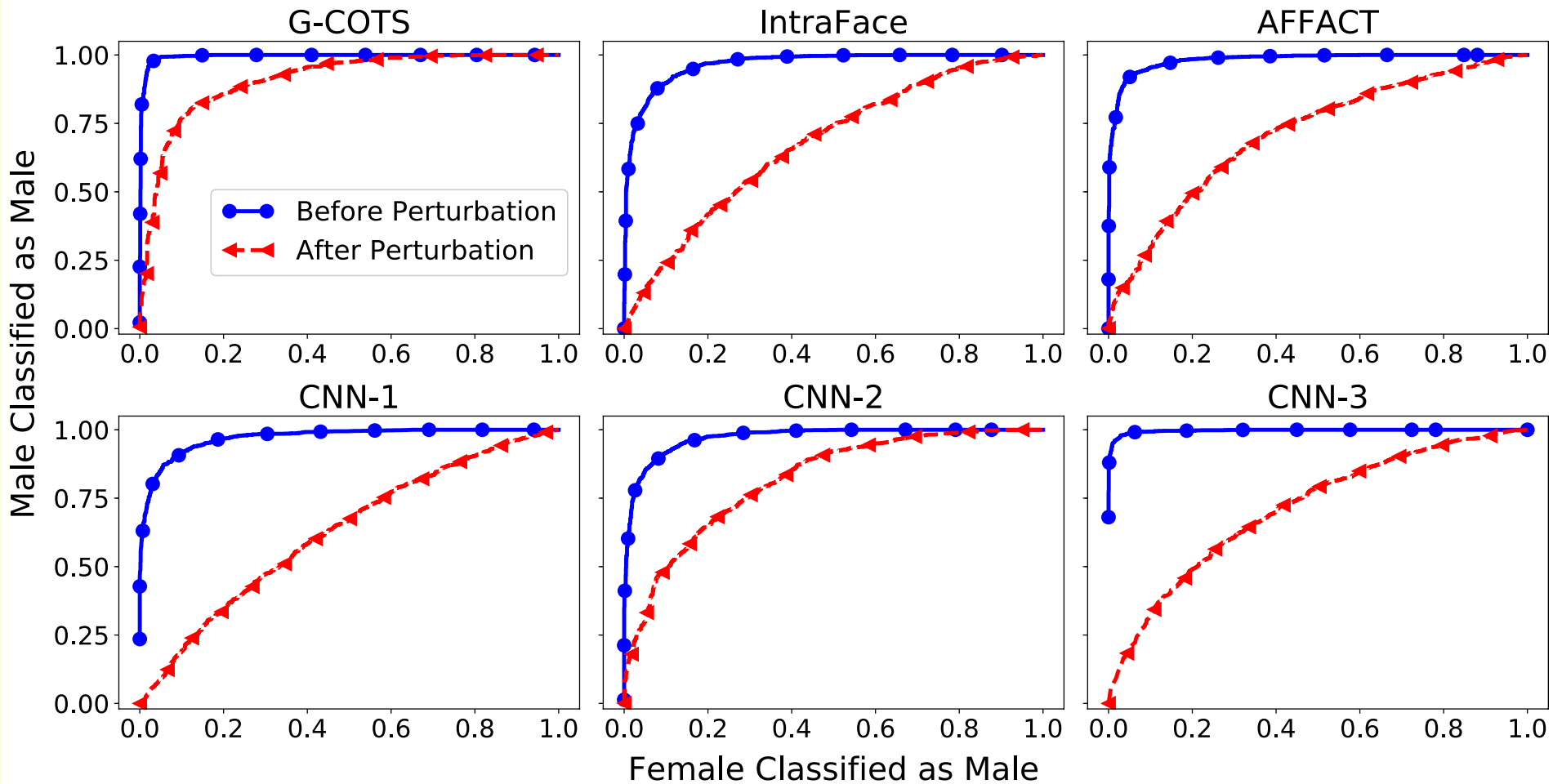
- G-COTS [Commercial]
- IntraFace [De la Torre et al., 2015]
- AFFACT [Günther et al., 2017]
- 3 CNN models [in-house]

■ Four unseen face Matchers

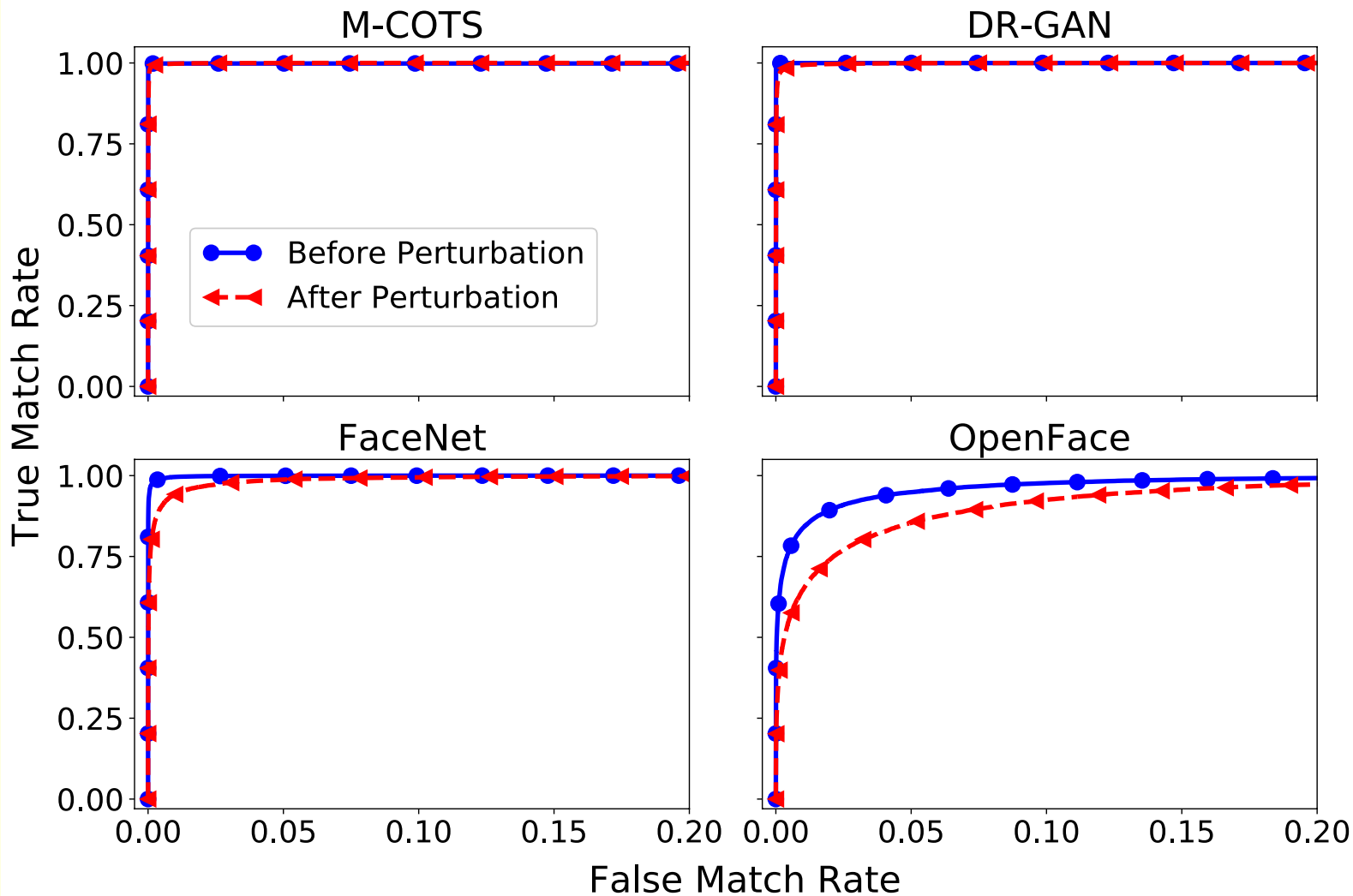
- M-COTS [Commercial]
- DR-GAN [Tran et al., 2017]
- FaceNet [Schroff et al., 2015]
- OpenFace [Amos et al., 2016]

Unseen:
the classifier or face matcher
is not used during training of
the SAN models

Performance Assessment on MUCT dataset: Confound gender classifiers



Performance Assessment on MUCT dataset: Retain Matching Capability



PrivacyNet

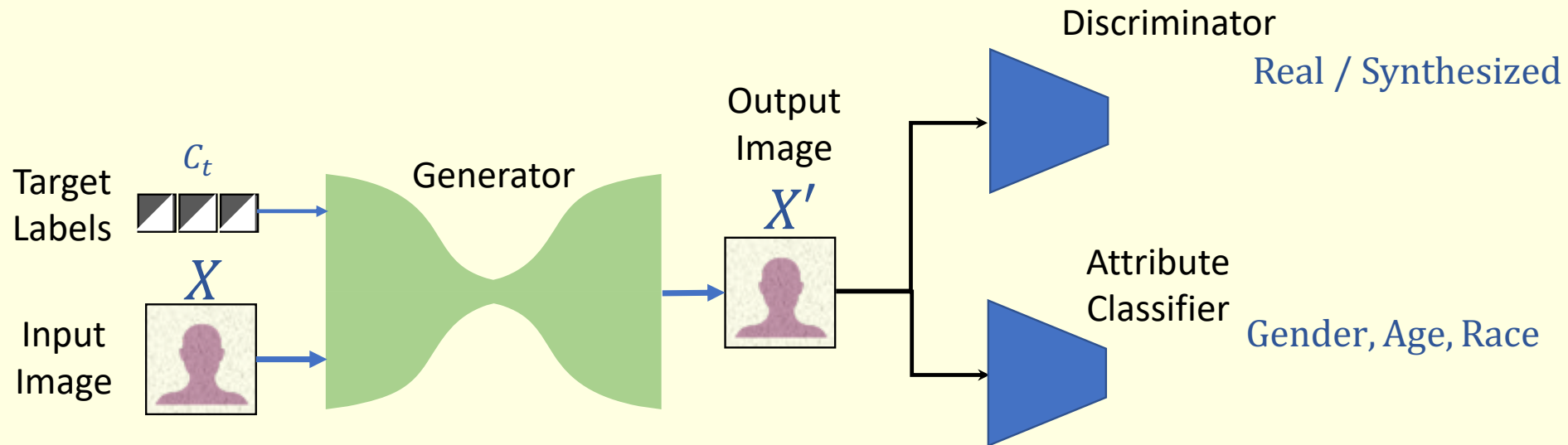
❑ 3 Soft Biometric Attributes and Labels

- Gender – Male; Female
- Age – Young; Middle Age; Old
- Race: African Decent; Caucasian

❑ PrivacyNet

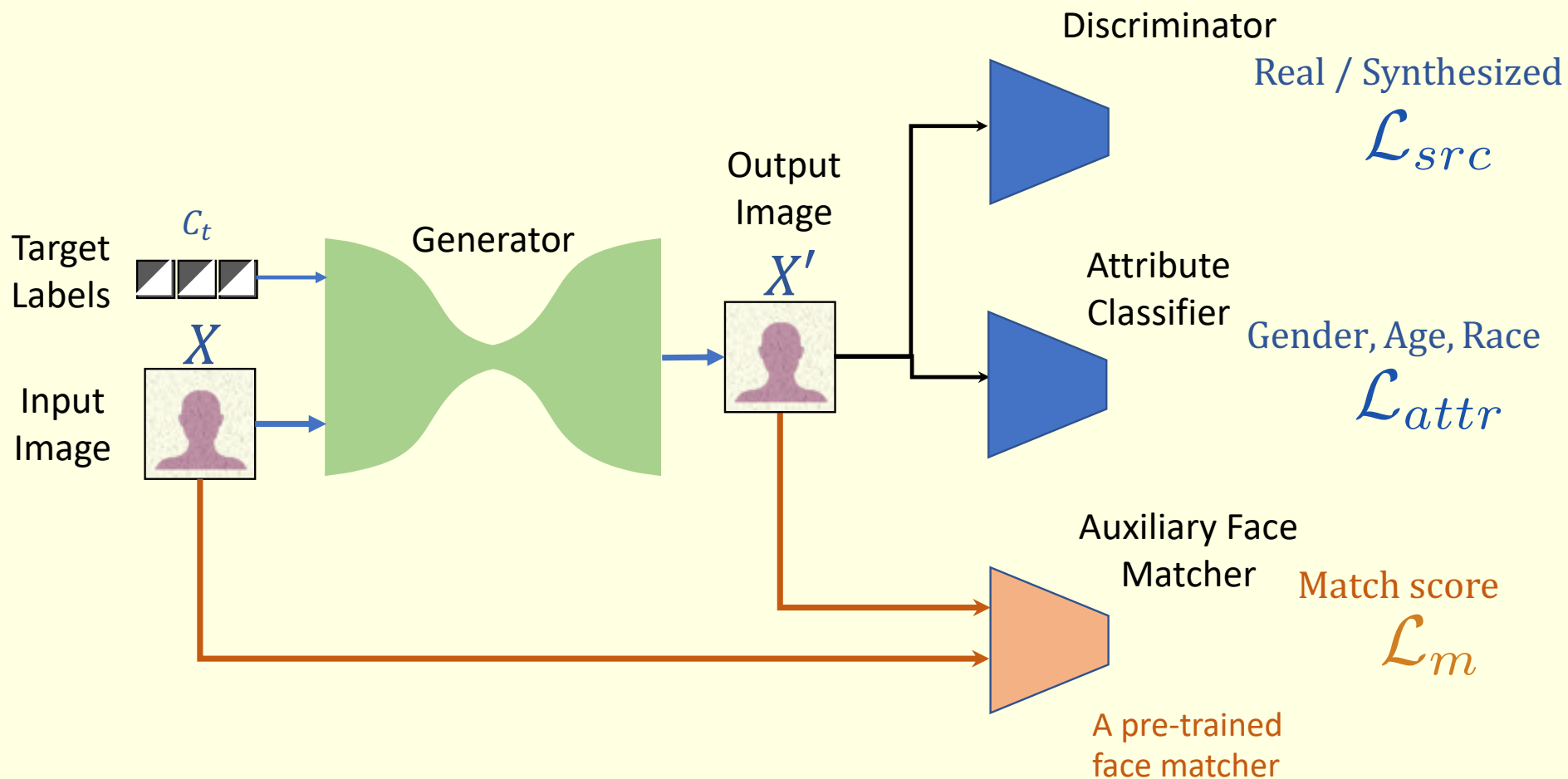
- **Cycle-GAN:**
 - Generator: transforms images to a target label vector
 - Discriminator: distinguishes between real and synthesized images
- **Auxiliary Face Matcher:** ensures that the generated images match against the original images

Multi-attribute Privacy



- ❑ Input image X from original label c_0
- ❑ A regular cycle-GAN that generates output image X' for a given input image X and target label vector c_t .

Multi-attribute Privacy



- ❑ Auxiliary Face Matcher derives the matching-loss term to ensure that the output image X' matches with input X .

PrivacyNet: Loss Functions

- Losses for training the **discriminator** D_{src} and D_{attr} :

1. Source term (real vs. synthesized)

$$\mathcal{L}_{D,src} = \mathbb{E}_X [-\log(D_{src}(X))] + \mathbb{E}_{X,c_t} [-\log(1 - D_{src}(G(X, c_t)))]$$

2. Attribute term

$$\mathcal{L}_{D,attr} = \mathbb{E}_{X,c_0} [-\log(D_{attr}(c_0|X))]$$

- Losses for training the **generator** $G(X, c_t)$:

1. Source term

$$\mathcal{L}_{G,src} = \mathbb{E}_{X,c_t} [\log(D_{src}(G(X, c_t)))]$$

2. Attribute term

$$\mathcal{L}_{G,attr} = \mathbb{E}_{X,c_t} [-\log(D_{attr}(c_t|G(X, c_t)))]$$

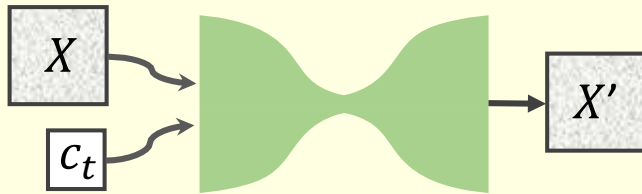
3. Matching term

$$\mathcal{L}_{G,m} = \mathbb{E}_{X,c_t} [\|R(X) - R(G(X, c_t))\|_2^2]$$

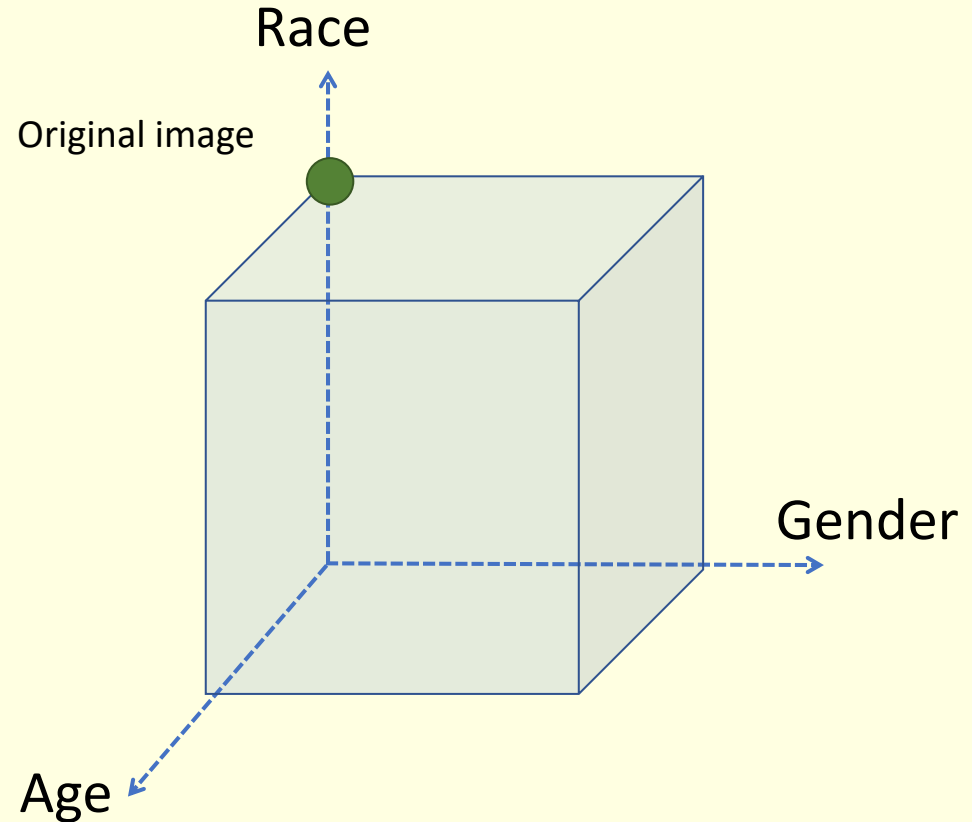
4. Reconstruction loss (cycle-consistency)

$$\mathcal{L}_{rec} = \mathbb{E}_{X,c_0,c_t} [\|X - G(G(X, c_t), c_0)\|_1]$$

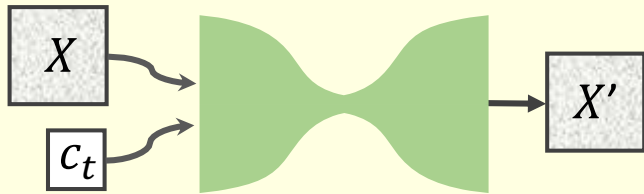
Face Transformation Using PrivacyNet



Original Label c_0 : [0, 1, 1]
[Female, Middle-aged, Caucasian]

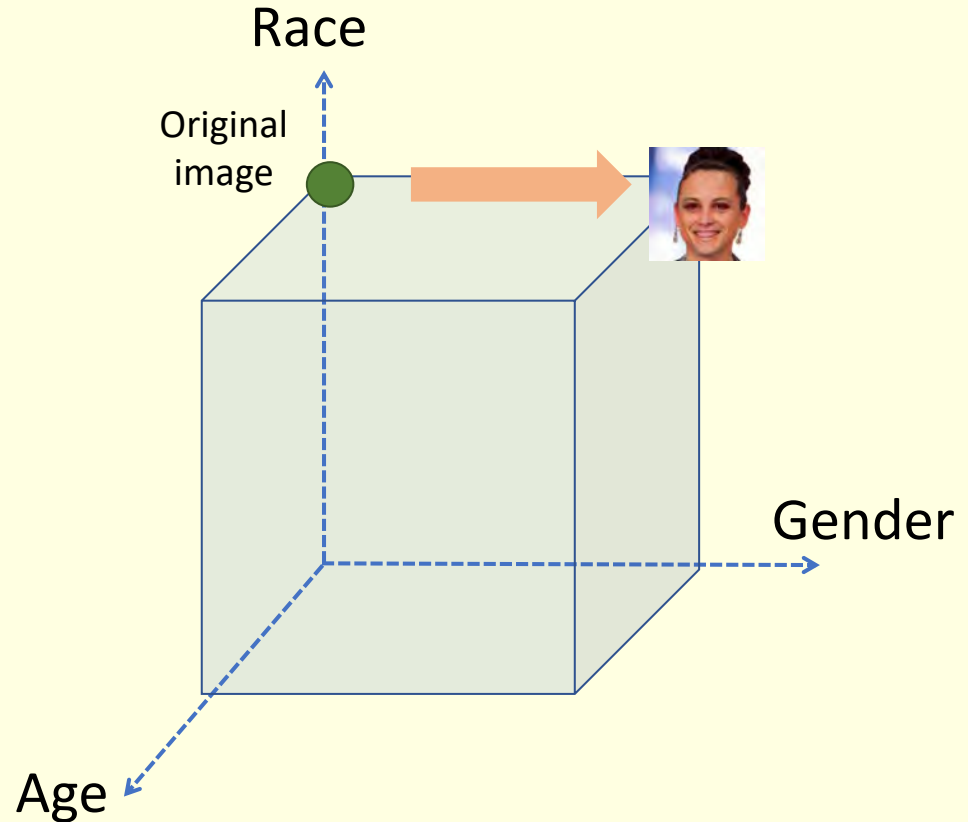


Face Transformation Using PrivacyNet

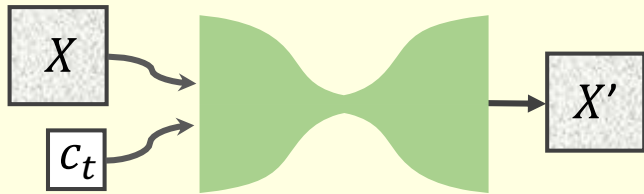


Original Label c_0 : [0, 1, 1]
[Female, Middle-aged, Caucasian]

c_t
[Male, Mid-age, Cauc.]



Face Transformation Using PrivacyNet



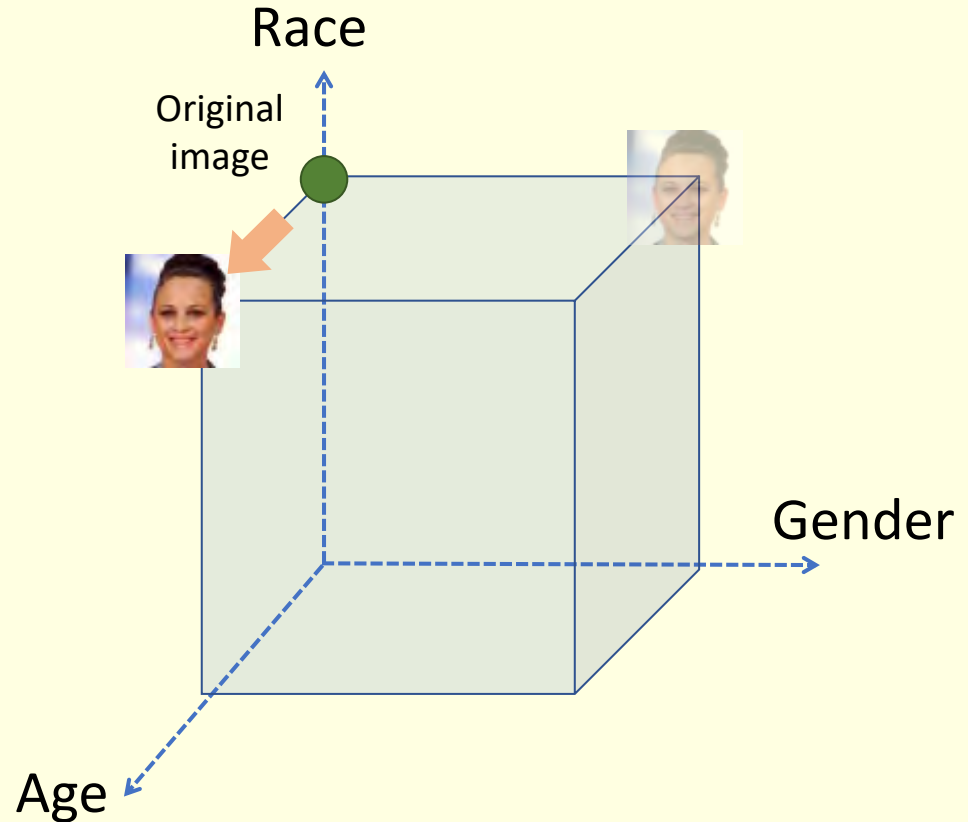
c_t

[Male, Mid-age, Cauc.]

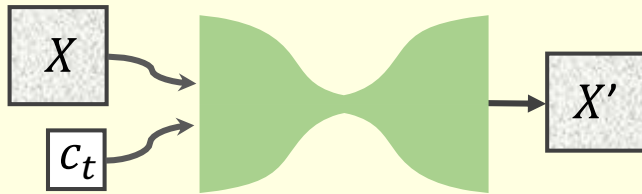
[Female, **Old**, Cauc.]

Original Label c_0 : [0, 1, 1]

[Female, Middle-aged, Caucasian]



Face Transformation Using PrivacyNet



c_t

[Male, Mid-age, Cauc.]

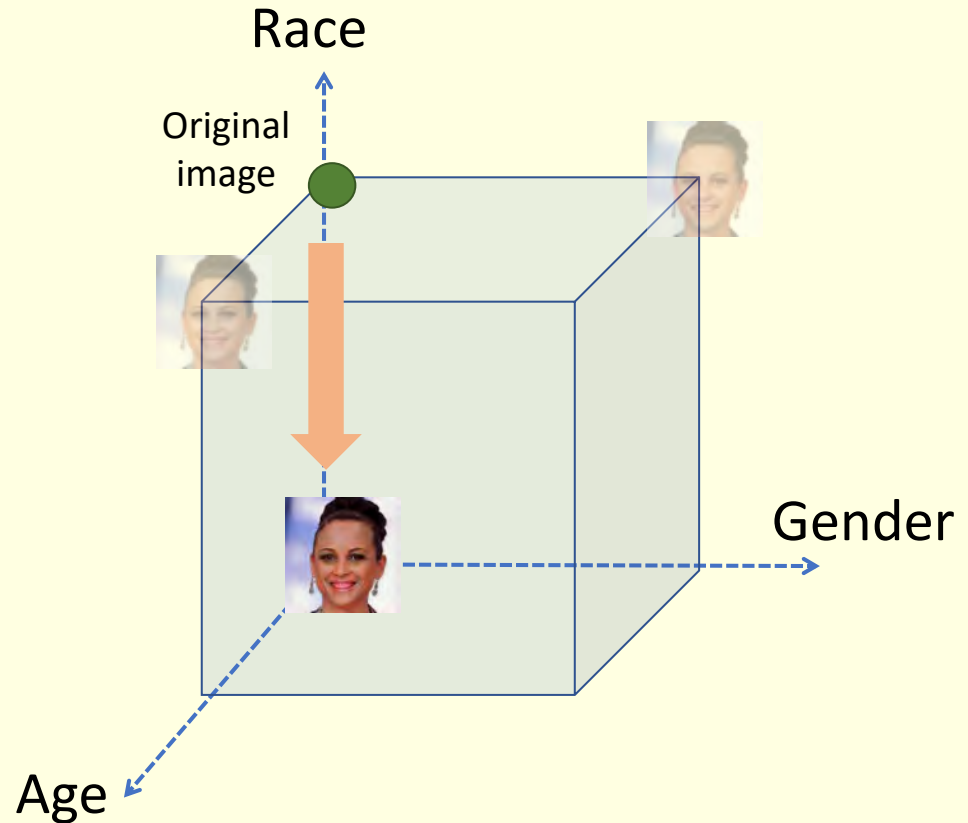
[Female, Old, Cauc.]

[Female, Mid-age, **Afric.**]

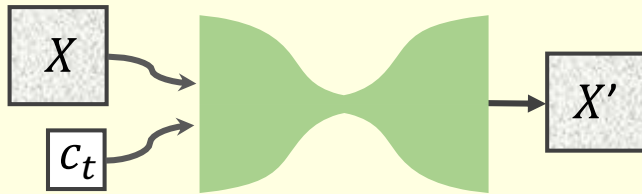


Original Label c_0 : [0, 1, 1]

[Female, Middle-aged, Caucasian]



Face Transformation Using PrivacyNet



c_t

[Male, Mid-age, Cauc.]

[Female, Old, Cauc.]

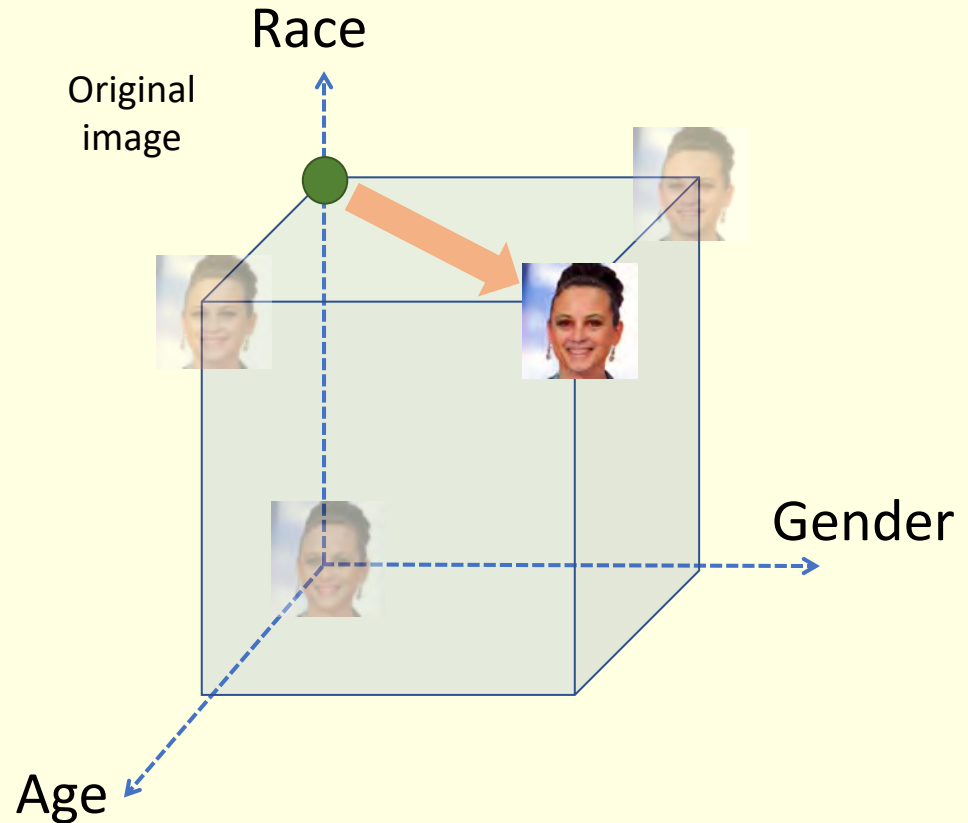
[Female, Mid-age, Afric.]

[Male, Old, Cauc.]

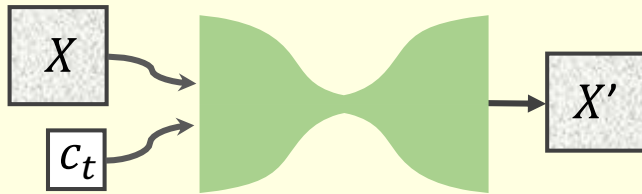


Original Label c_0 : [0, 1, 1]

[Female, Middle-aged, Caucasian]



Face Transformation Using PrivacyNet



c_t

[Male, Mid-age, Cauc.]

[Female, Old, Cauc.]

[Female, Mid-age, Afric.]

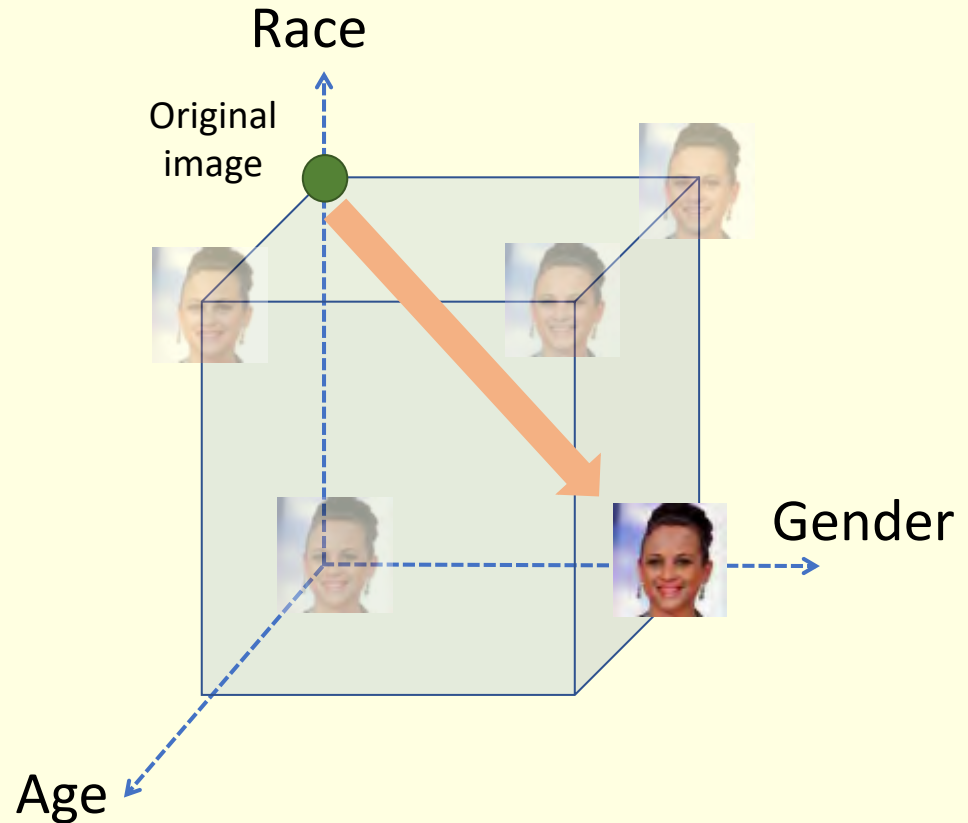
[Male, Old, Cauc.]

[**Male**, Mid-age, **Afric.**]

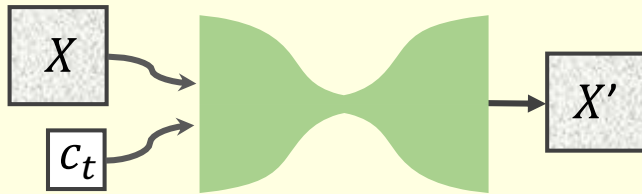


Original Label c_0 : [0, 1, 1]

[Female, Middle-aged, Caucasian]



Face Transformation Using PrivacyNet

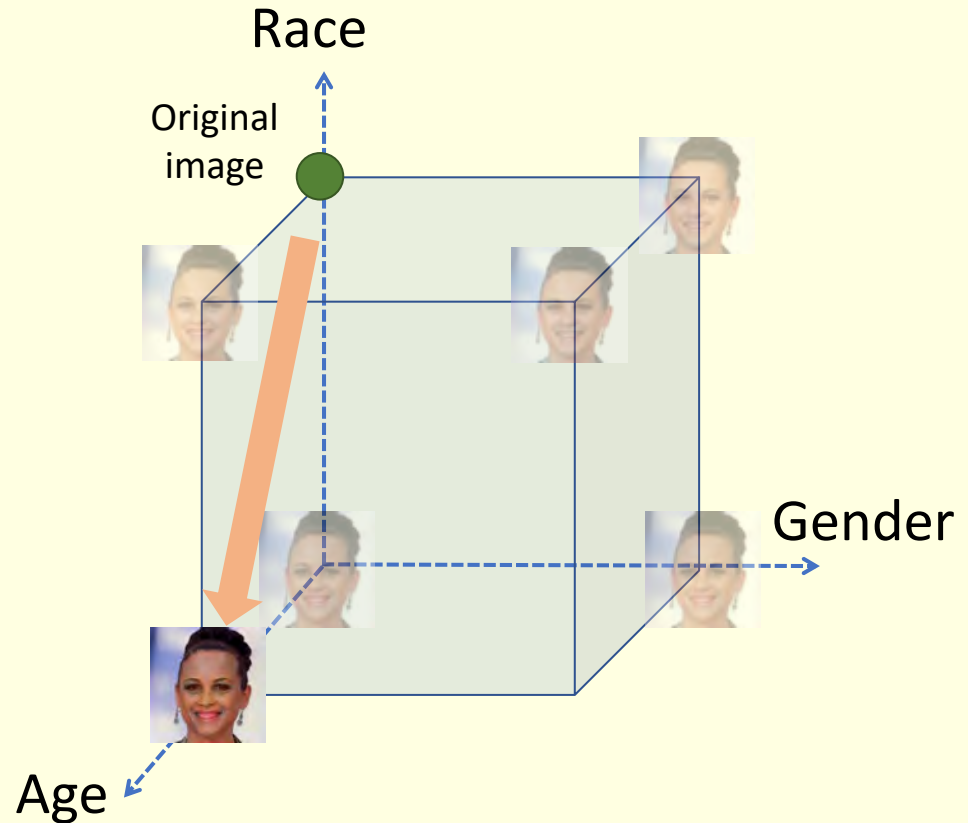


c_t

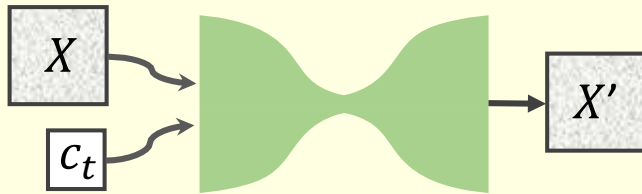
- [Male, Mid-age, Cauc.]
- [Female, Old, Cauc.]
- [Female, Mid-age, Afric.]
- [Male, Old, Cauc.]
- [Male, Mid-age, Afric.]
- [Female, **Old**, **Afric.**]

Original Label c_0 : [0, 1, 1]

[Female, Middle-aged, Caucasian]



Face Transformation Using PrivacyNet

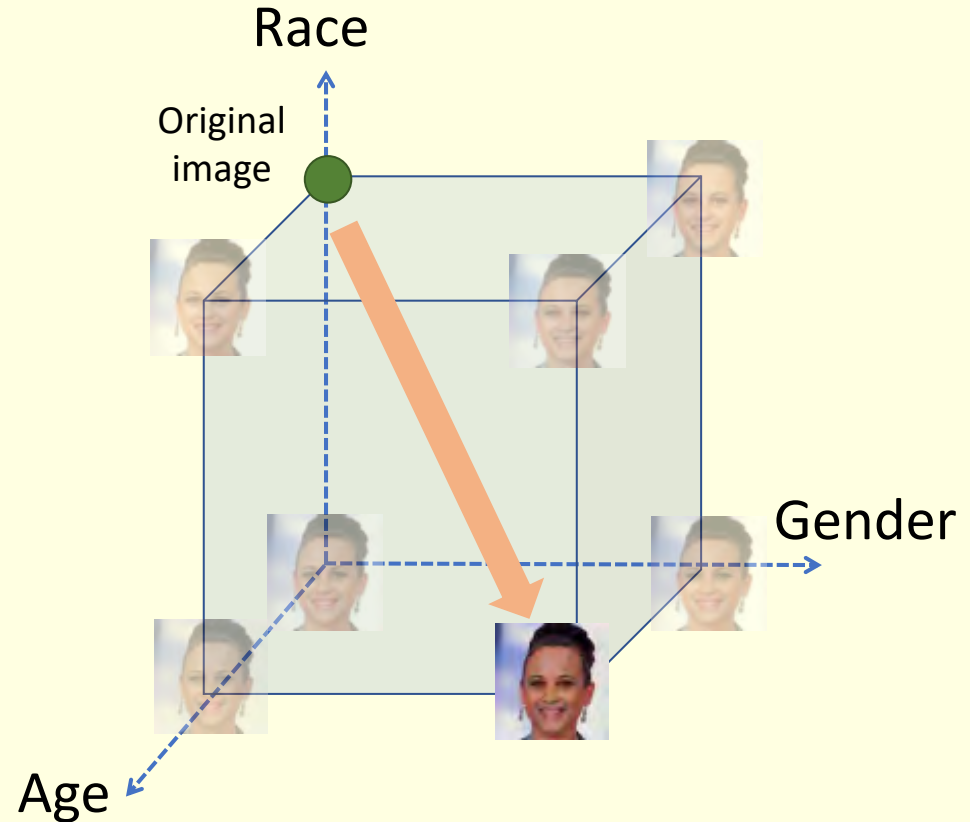


c_t

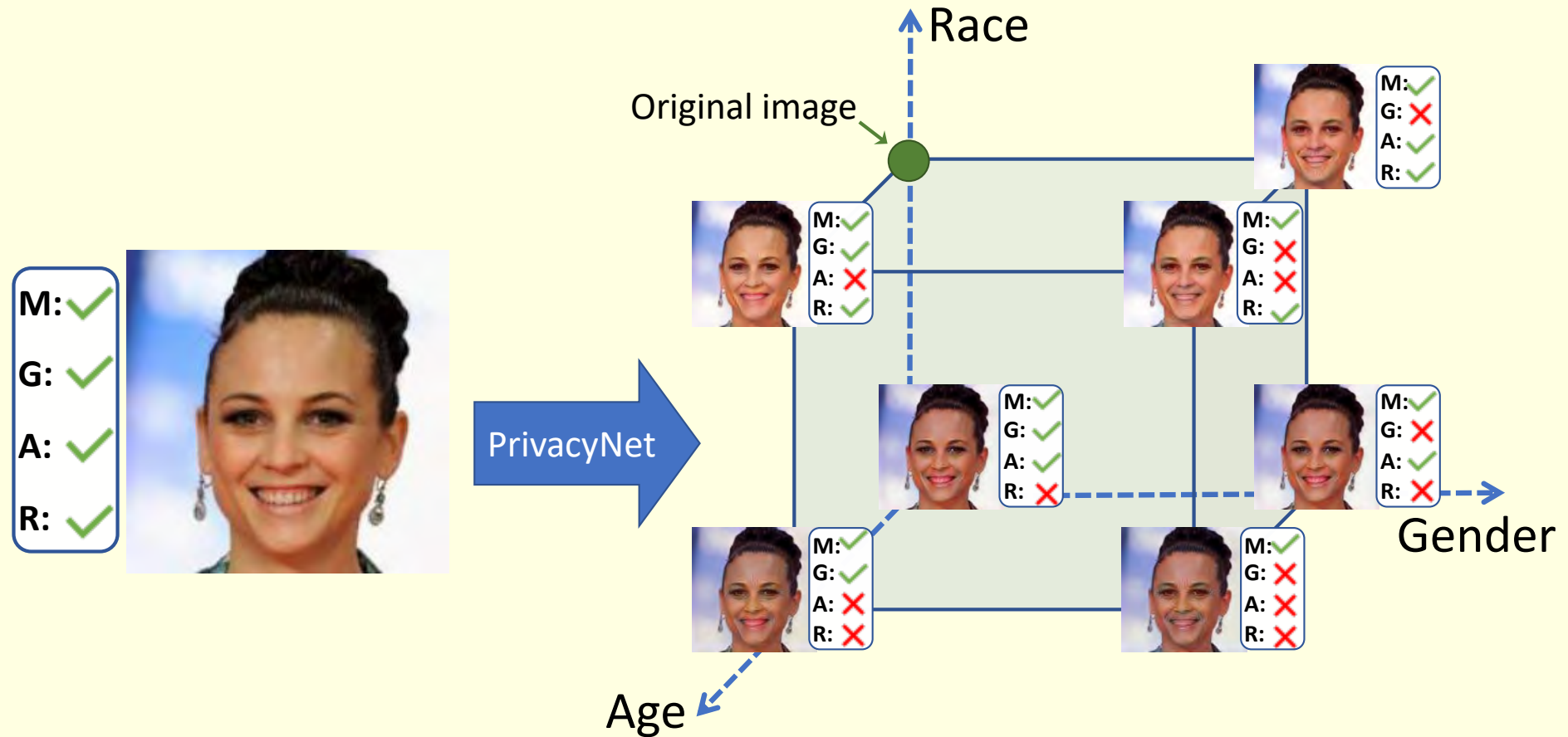
- [Male, Mid-age, Cauc.]
- [Female, Old, Cauc.]
- [Female, Mid-age, Afric.]
- [Male, Old, Cauc.]
- [Male, Mid-age, Afric.]
- [Female, Old, Afric.]
- [Male, Old, Afric.]**

Original Label c_0 : [0, 1, 1]

[Female, Middle-aged, Caucasian]



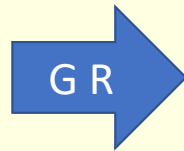
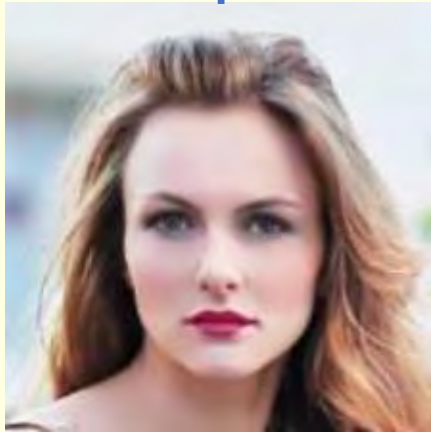
Face Transformation Using PrivacyNet



Mirjalili et al., PrivacyNet: Semi-Adversarial Networks for Multi-attribute Face Privacy, IEEE TIP 2020

Face Transformation Using PrivacyNet

Original Image



PrivacyNet



GAN



Match Scores with original image:

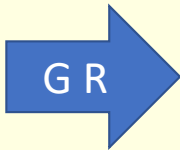
M-COTS: 0.99
DR-GAN: 0.93
SE-ResNet-50: 0.72

0.41 ↓
0.69 ↓
0.14 ↓

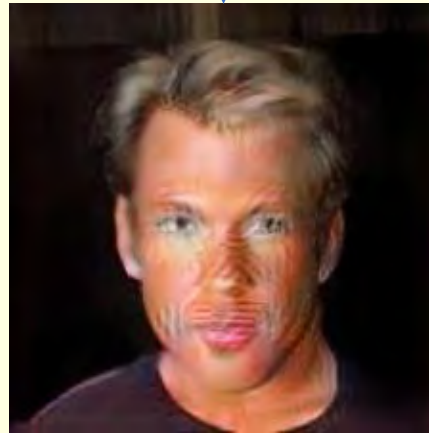
Mirjalili et al., PrivacyNet: Semi-Adversarial Networks for Multi-attribute Face Privacy, IEEE TIP 2020

Face Transformation Using PrivacyNet

Original Image



PrivacyNet



GAN



Match Scores with original image:

M-COTS: 0.99
DR-GAN: 0.93
SE-ResNet-50: 0.81

0.39 ↓
0.74 ↓
0.31 ↓

Mirjalili et al., PrivacyNet: Semi-Adversarial Networks for Multi-attribute Face Privacy, IEEE TIP 2020

Modifying Gender

Original



Male

Female

Male

Male

Female

Female

Male

Female

Female

Male



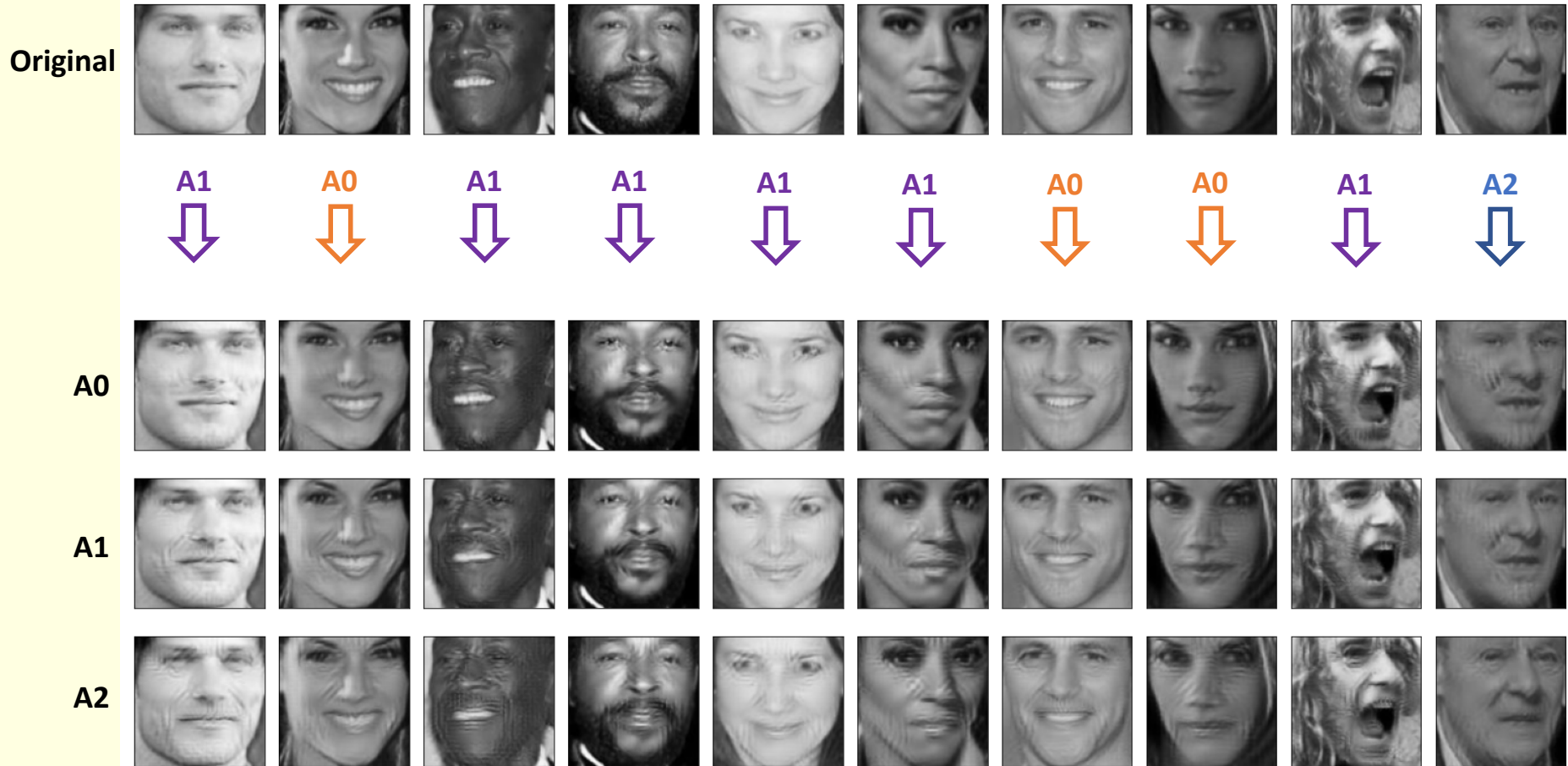
PrivacyNet



Baseline-GAN



Modifying Age



Evaluation: Datasets

Number of samples

Dataset	Total	Gender		Age			Race		
		Male	Female	Young	Middle-age	Old	African	White	Other
CelebA	202599	84434	118165*	79848	91373*	16337	11119	142225*	49255
LFW	13293	4281	1465	--	--	--	4490	510	751
MORPH	55608	47057*	8551	25009	26614	3985	42897*	10736	1975
MUCT	3754	1844	1910	1326	1807	620	1030	1480	1244
RaFD	1608	1008*	600	1276*	332	0	0	1608*	0
UTK	24104	12582	11522	12980*	6068	5056	4558	10222*	9324

* Highlighted numbers show over-represented classes

Evaluation: Datasets

Number of subjects

Dataset	Total	Gender		Race		
		Male	Female	African	White	Other
CelebA	10177	4582	5595	553	7018*	2606
LFW	5749	4274*	1461	510	4482*	748
MORPH	13673	11512*	2161	10358*	2728	587
MUCT	276	131	145	79	106*	91
RaFD	67	42*	25	0	67*	0
UTK	--	--	--	--	--	--

* Highlighted numbers show over-represented classes

Face Matchers

- Commercial face matcher:

- Rank-One (state-of-the-art commercial matcher)
[<https://www.rankone.io/>]

- Open-source face matchers:

- ArcFace

[ArcFace: Additive Angular Margin Loss for Deep Face Recognition, Deng et al., 2019]

- DR-GAN

[Disentangled Representation Learning GAN for Pose-Invariant Face Recognition, Tran et al., 2017]

- SE-ResNet-50:

[VGGFace2: A dataset for recognising face across pose and age, Cao et al., 2018]

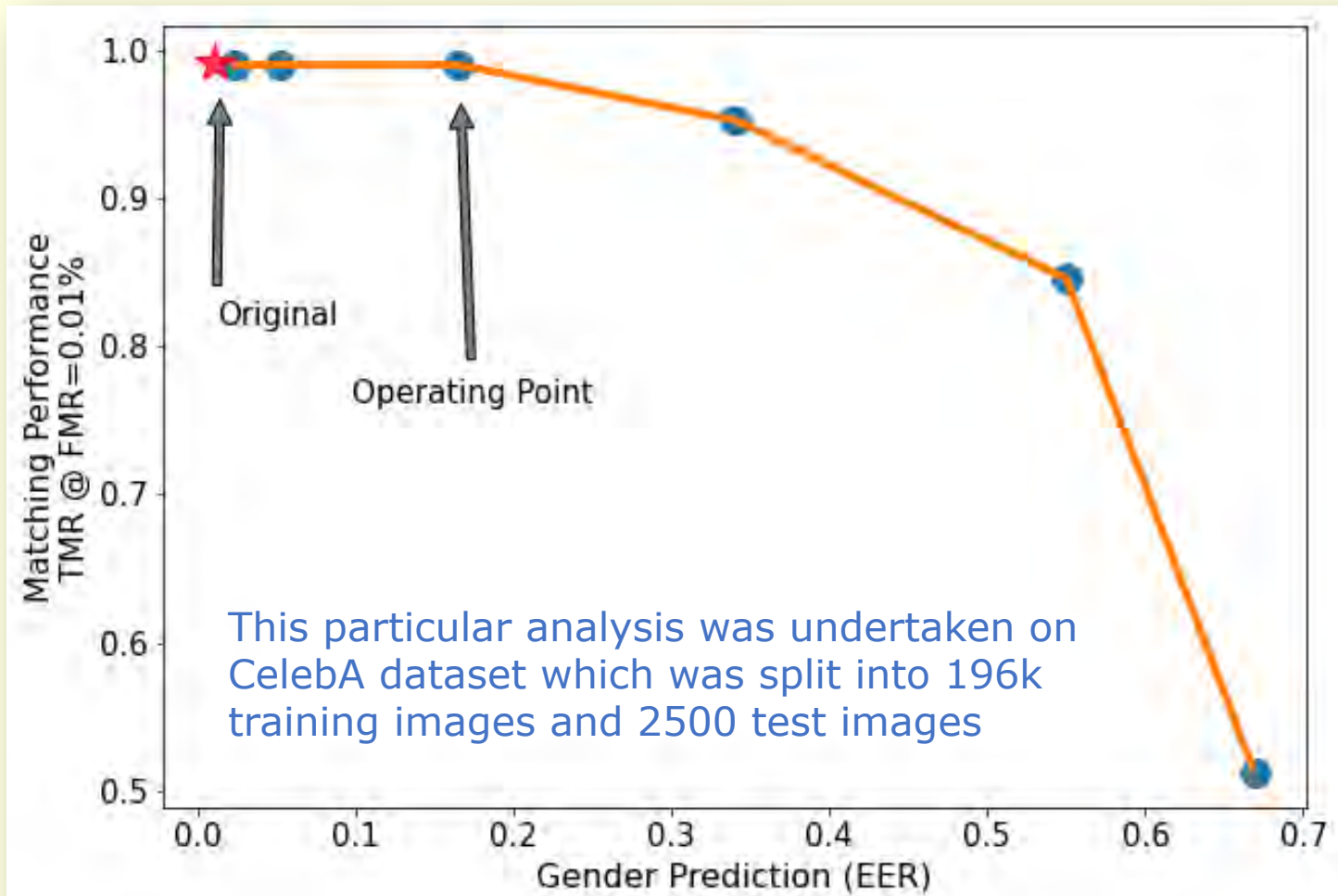
Attribute Classifiers

- RankOne Computing → gender, age, race
 - [<https://www.rankone.io/>]
- IntraFace → gender, race
 - [IntraFace, De la Torre, 2016]
- AFFACT → gender, race
 - [AFFACT - Alignment-Free Facial Attribute Classification Technique, Günther, 2017]

Attribute Classifiers

- RankOne Computing → gender, age, race
 - [<https://www.rankone.io/>]
- IntraFace → gender, race
 - [IntraFace, De la Torre, 2016]
- AFFACT → gender, race
 - [AFFACT - Alignment-Free Facial Attribute Classification Technique, Günther, 2017]

Experimental Results



Summary

- **Semi-Adversarial Network**

- Perturbing one classifier while retaining the performance of other

- **Results confirm that**

- Automatic demographic prediction is confounded → providing demographic privacy to face images
- Matching utility is still retained
- Extending to multiple attributes: gender, age, race/ethnicity
- Differential privacy: some attributes preserved; others confounded
- Visual realism of images

Privacy Enhancing Technology

- Preserving the **privacy** of a user's stored biometric data
 - Regulate **cross-linking** across applications
 - Regulate **gleaning** additional information from biometric data (e.g., medical condition)

Need to

- Define Privacy and Privacy Metrics
- Guarantee Privacy
- Develop Differential Privacy Schemes

Related Papers



- V. Mirjalili, S. Raschka, A. Ross, "**PrivacyNet: Semi-Adversarial Networks for Multi-attribute Face Privacy**," IEEE TIP 2020.
- V. Mirjalili, S. Raschka, A. Ross, "**FlowSAN: Privacy-Enhancing Semi-Adversarial Networks to Confound Arbitrary Face-Based Gender Classifiers**," IEEE Access, 2019.
- V. Mirjalili, S. Raschka, A. Ross, "**Gender Privacy: An Ensemble of Semi Adversarial Networks for Confounding Arbitrary Gender Classifiers**," BTAS 2018
- V. Mirjalili, S. Raschka, A. Namboodiri, A. Ross, "**Semi-Adversarial Networks: Convolutional Autoencoders for Imparting Privacy to Face Images**," ICB 2018
- V. Mirjalili and A. Ross, "**Soft Biometric Privacy: Retaining Biometric Utility of Face Images while Perturbing Gender**," IJCB 2017
- A. Othman and A. Ross, "**Privacy of Facial Soft Biometrics: Suppressing Gender But Retaining Identity**," ECCVW 2014