Bias and Fairness in Low Resolution Image Recognition

Sasikanth Kotti, IIT-Jodhpur

Committee :
Dr. Suman Kundu
Dr. Mayank Vatsa

Advisors :
Prof. Mayank Vatsa
Prof. Richa Singh

Outline

- Literature Survey - glimpse
- Motivation & Problem Statement
- Research Contributions
- Low Resolution face recognition - Analysis
- Bias and Fairness in GANs - Analysis
- FairDistillation in GANs - Mitigation

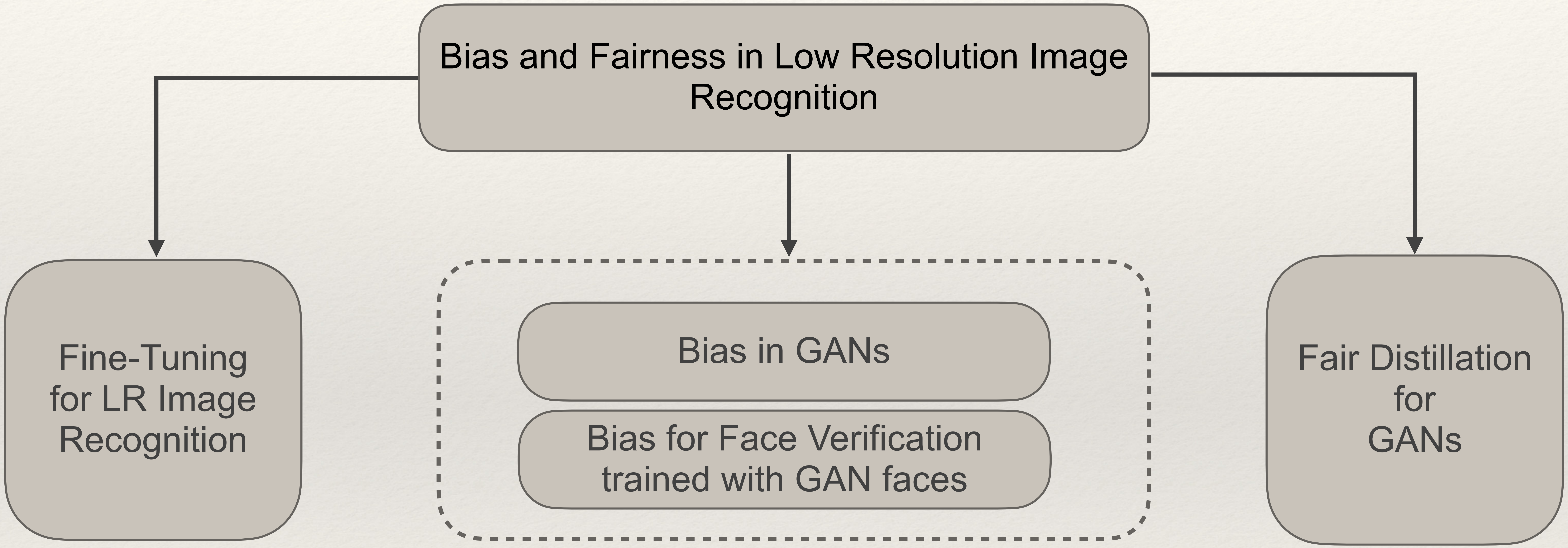
Literature Survey - glimpse

Authors	Summary
Turk et.al., 1991	Foundational method where faces are projected to feature space defined by eigen vectors
X. Wu et.al., 2018	Defines a Light CNN framework to learn representations even from large scale noisy data and introduced Max-Feature-Map (MFM)
Singh et.al., 2021	SOTA for LR/VLR image recognition achieved by novel loss functions namely Derived-Margin softmax loss and Reconstruction-Center (ReCent)
Nagpal et.al., 2019	The work answers questions and presents an in-depth analysis of bias in deep learning based facial recognition systems.
Karakas et.al., 2022	In this work the style space of StyleGAN2 model is used to perform disentangled control of the target attributes for debiasing.
Chang et.al., 2020	Proposed a black box based distillation method for GANs
Celis et.al., 2021	Proposed an optimization framework for learning fair classifiers where protected attributes of a fraction of samples are arbitrarily perturbed.

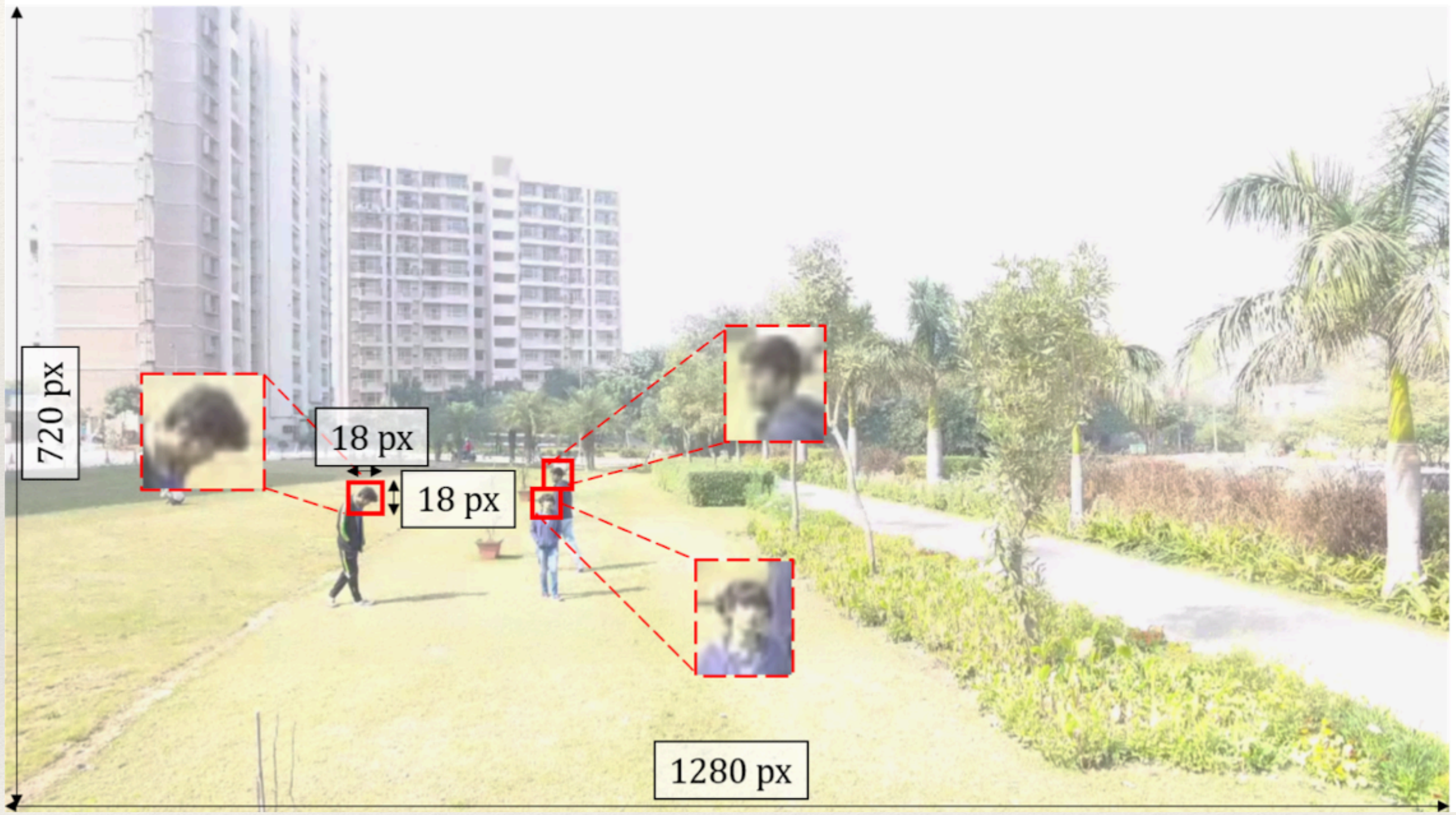
Motivation & Problem Statement

- High Resolution images are not aways available in real world settings
- This is due to sensor limitations, distance from object etc.
- Bias across different sub-groups can cause disparate impact and be detrimental to different groups.
- Bias and Fairness is more relevant in critical applications such as Surveillance, Authentication etc.
- Investigate the suitability of common techniques for low resolution face recognition
- Investigate and evaluate bias and fairness of existing generative models
- Propose methods for mitigating bias in generative models

Research Contributions



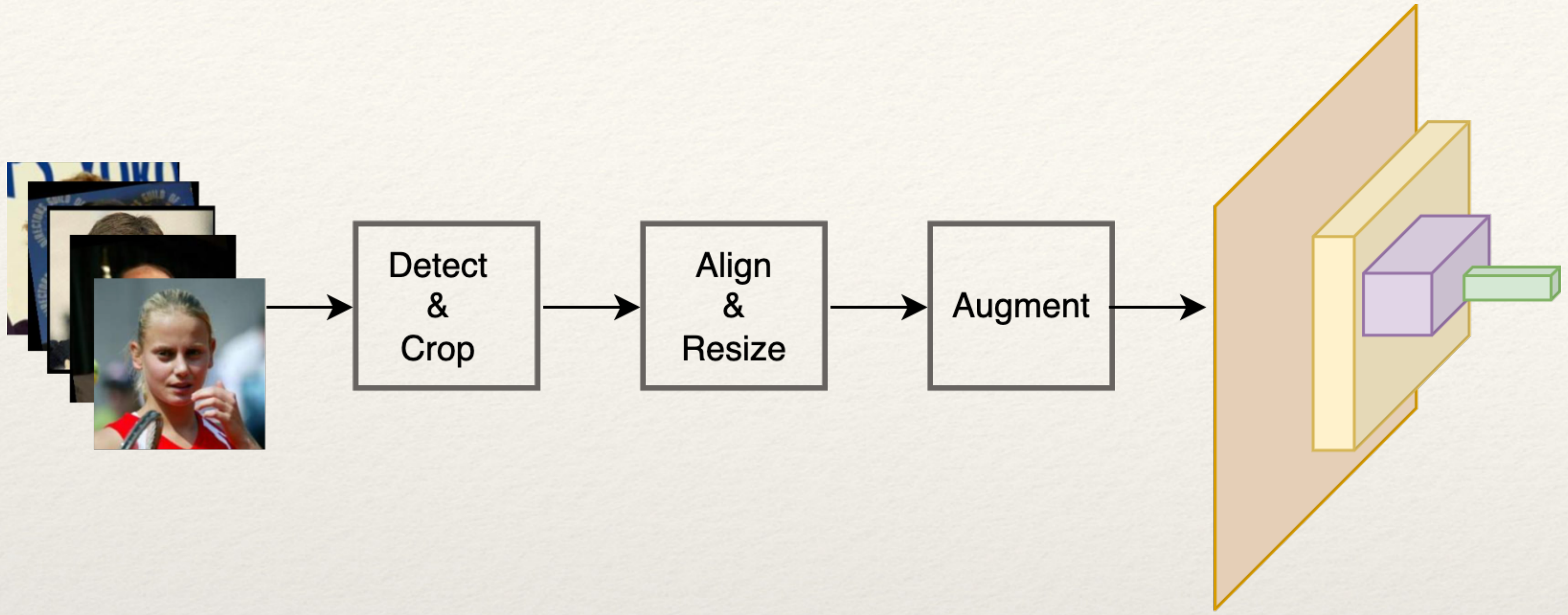
Low Resolution Face Recognition



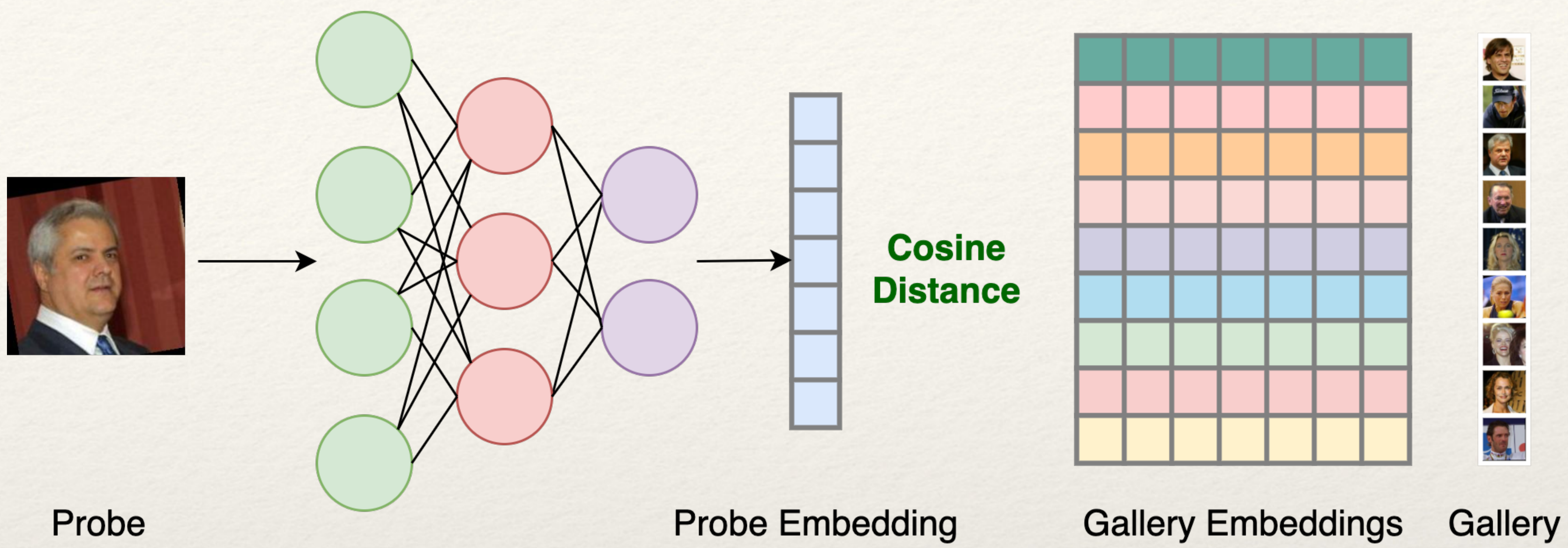
- HR images are not always available
- Information Loss in LR Images
- Vanilla interpolation not helpful
- Generative models as major components
- Bias and fairness due to generative models

Image Source : DeriveNet for (Very) Low Resolution Image Classification, Singh et.al., IEEE T-PAMI 2021

Datasets and Experiments

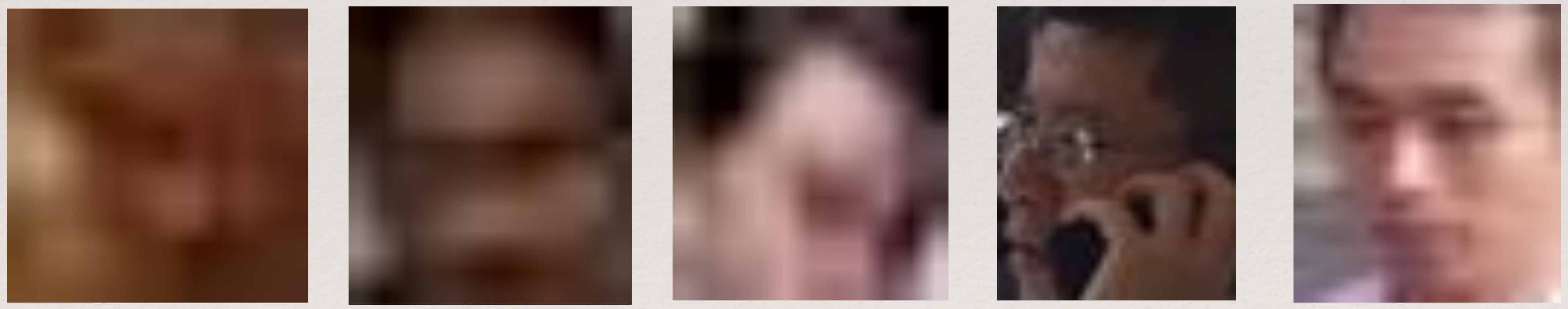


Fine-tuning - Contrastive Loss
(LightCNN29, ArcFace, VGGFace2)



Inference
(LightCNN29, ArcFace, VGGFace2)

LFW

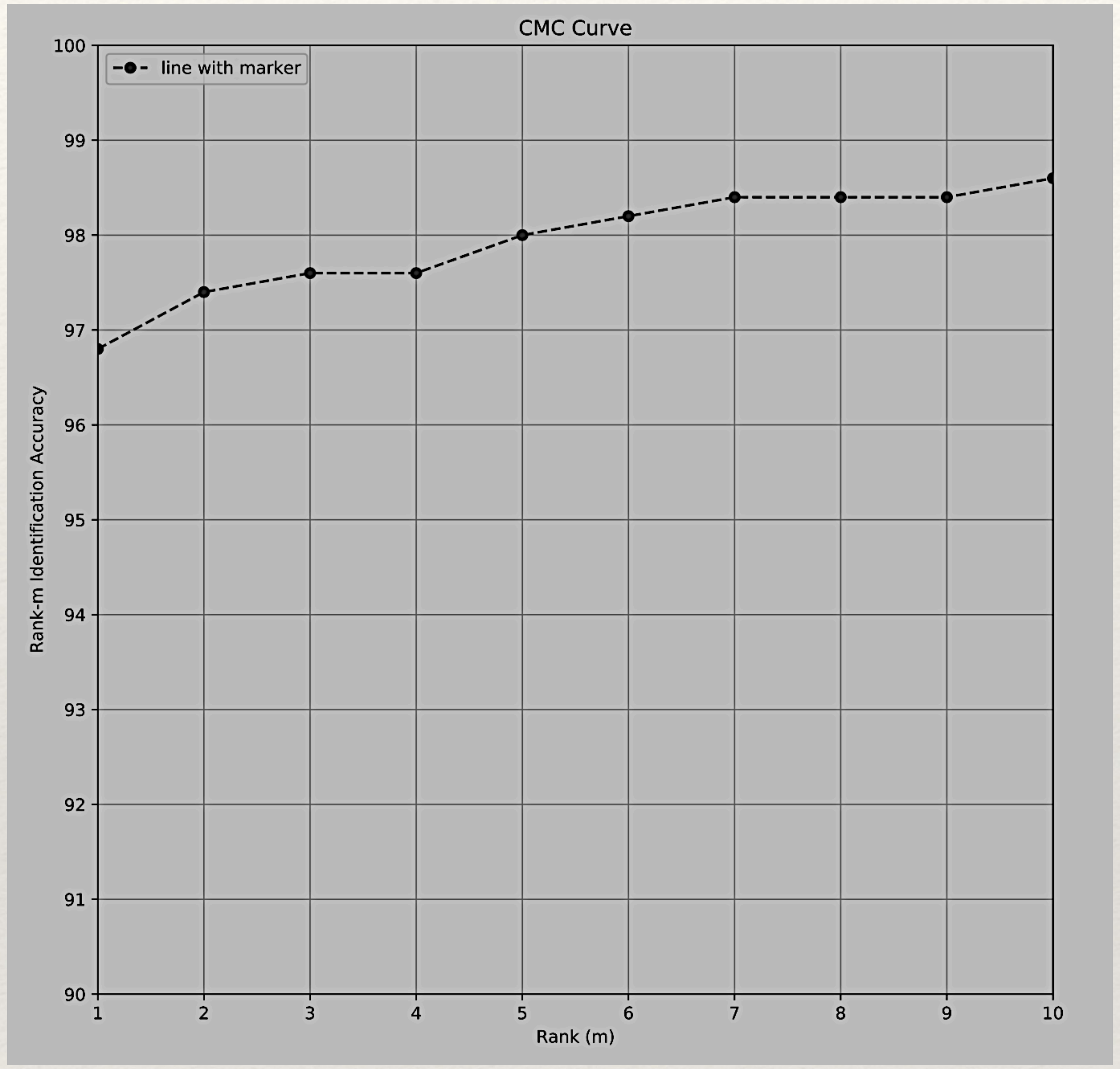


QMUL-SurvFace

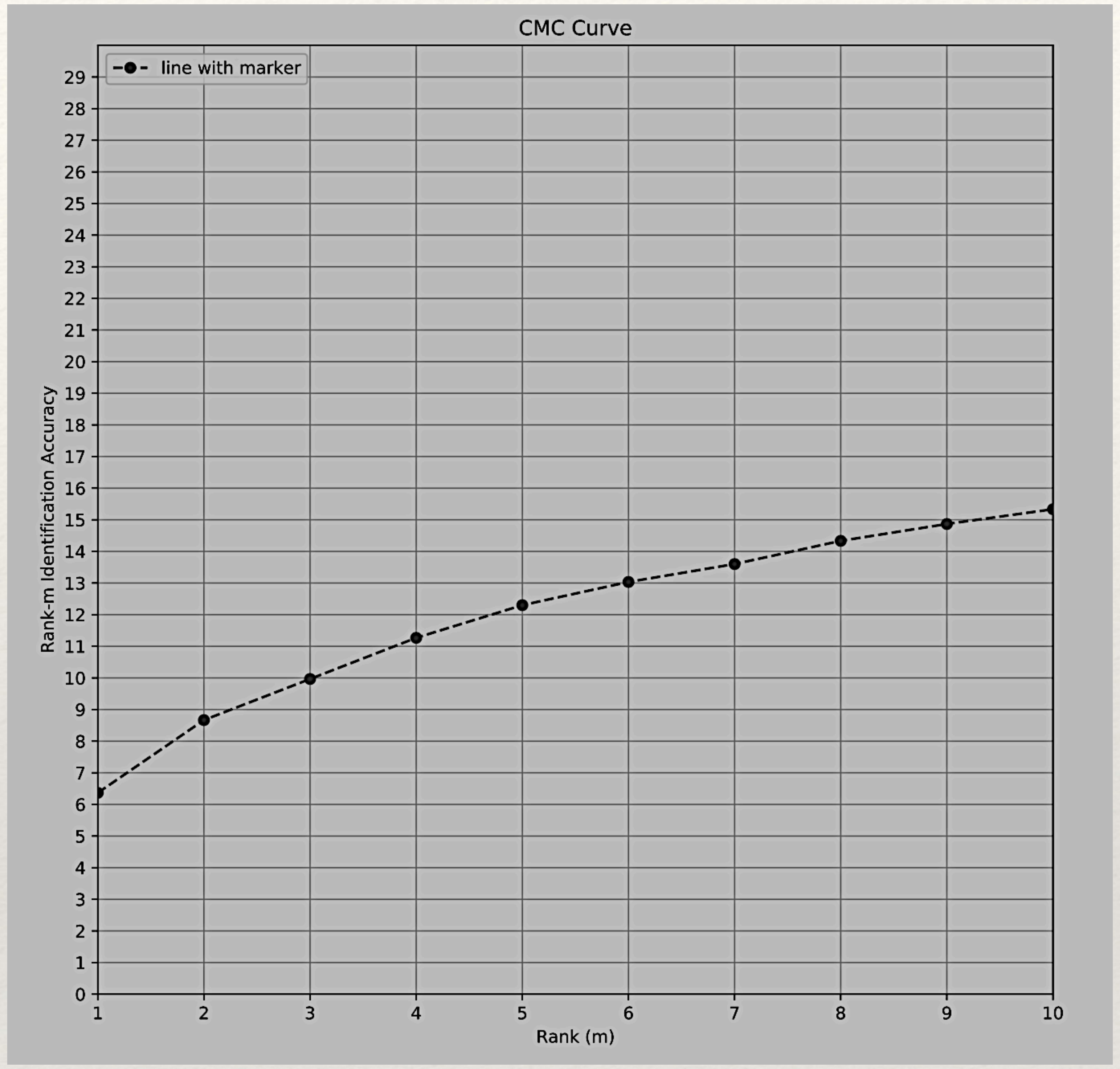
- *Fine-tuning with LFW and QMUL-SurvFace for LightCNN29, ArcFace and VGGFace2*
- *Inference of fine-tuned models for LFW and QMUL-SurFace test split*

Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Huang et.al., 2008
Surveillance Face Recognition Challenge Cheng et.al, 2018

Results - CMC Curves



LightCNN29 - LFW



LightCNN29 - QMUL-SurvFace

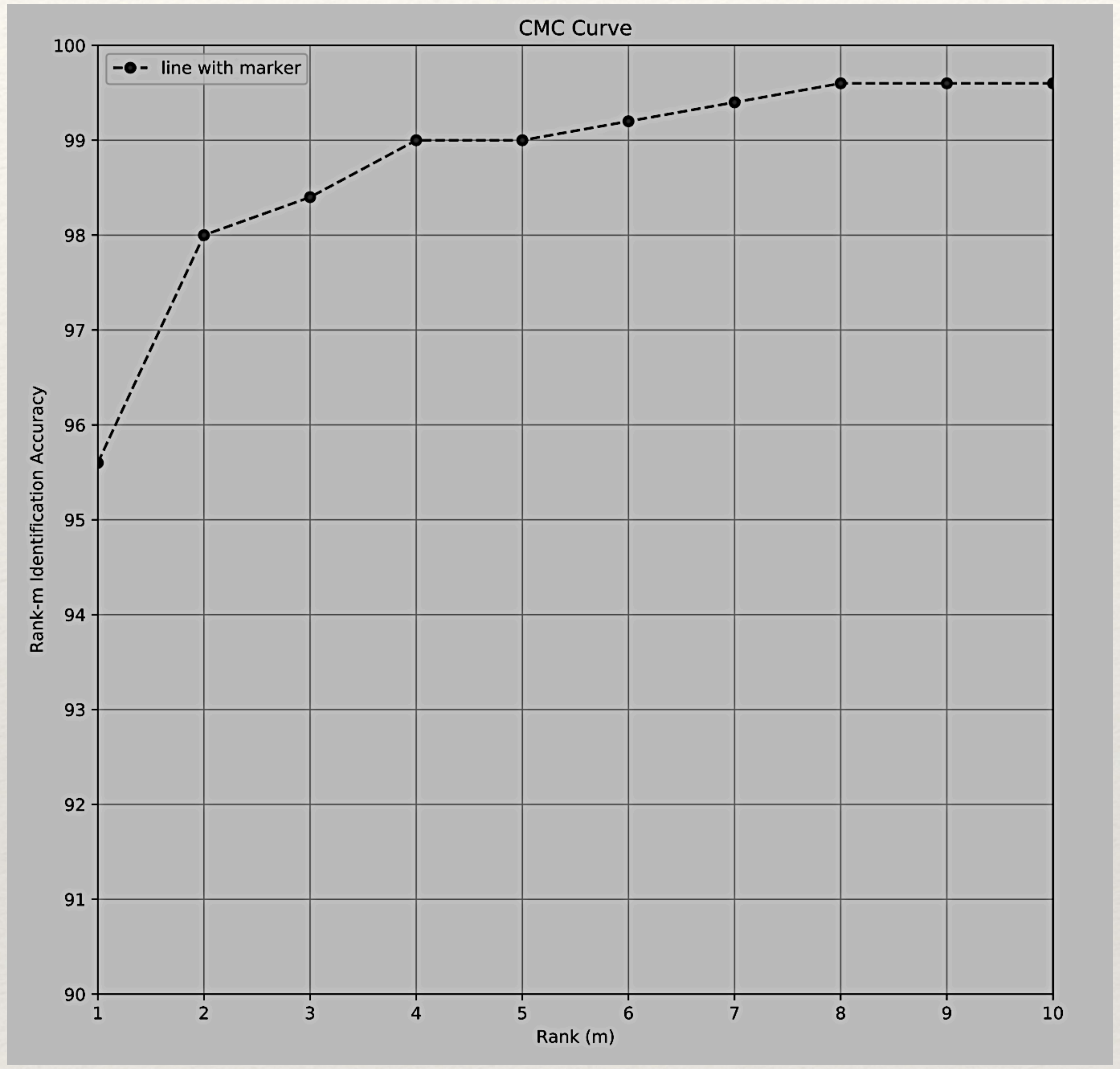
Model	Rank1		Rank10	
Dataset	LFW	SurvFace	LFW	SurvFace
LightCNN29	**96.8**	**6.3**	**98.6**	**15.2**
VGGFace2	95.6	1.6	99.7	5.7
ArcFace	80.7	1.8	89.6	5.8

Identification Accuracy

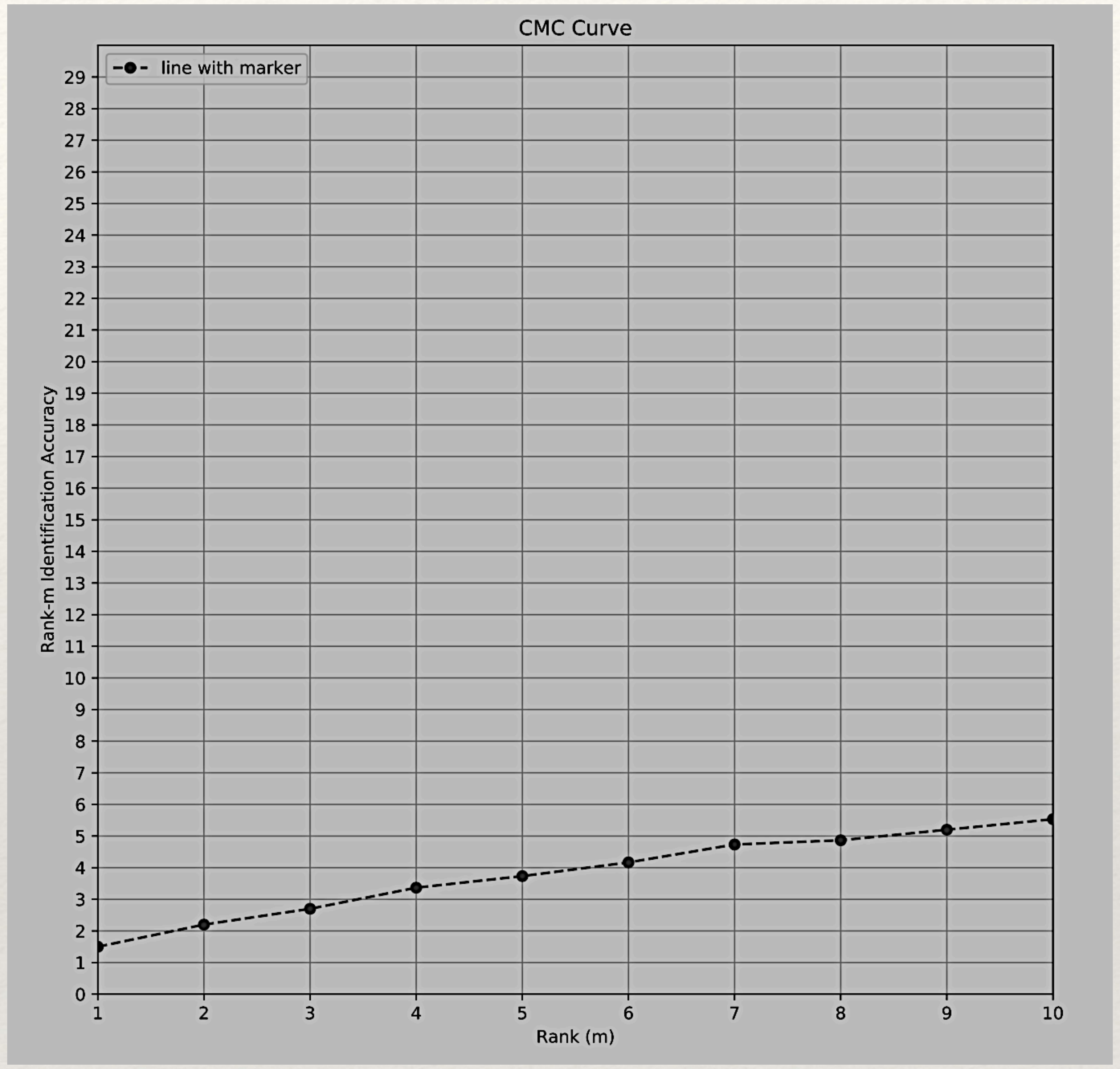
- ***Fine-tuning with LR images for LightCNN29 pre-trained on noisy dataset still degrades performance***

A light CNN for deep face representation with noisy labels Wu, Xiang, et.al., IEEE Transactions on Information Forensics and Security 2018

Results - CMC Curves



VGGFace2 - LFW



VGGFace2 - QMUL-SurvFace

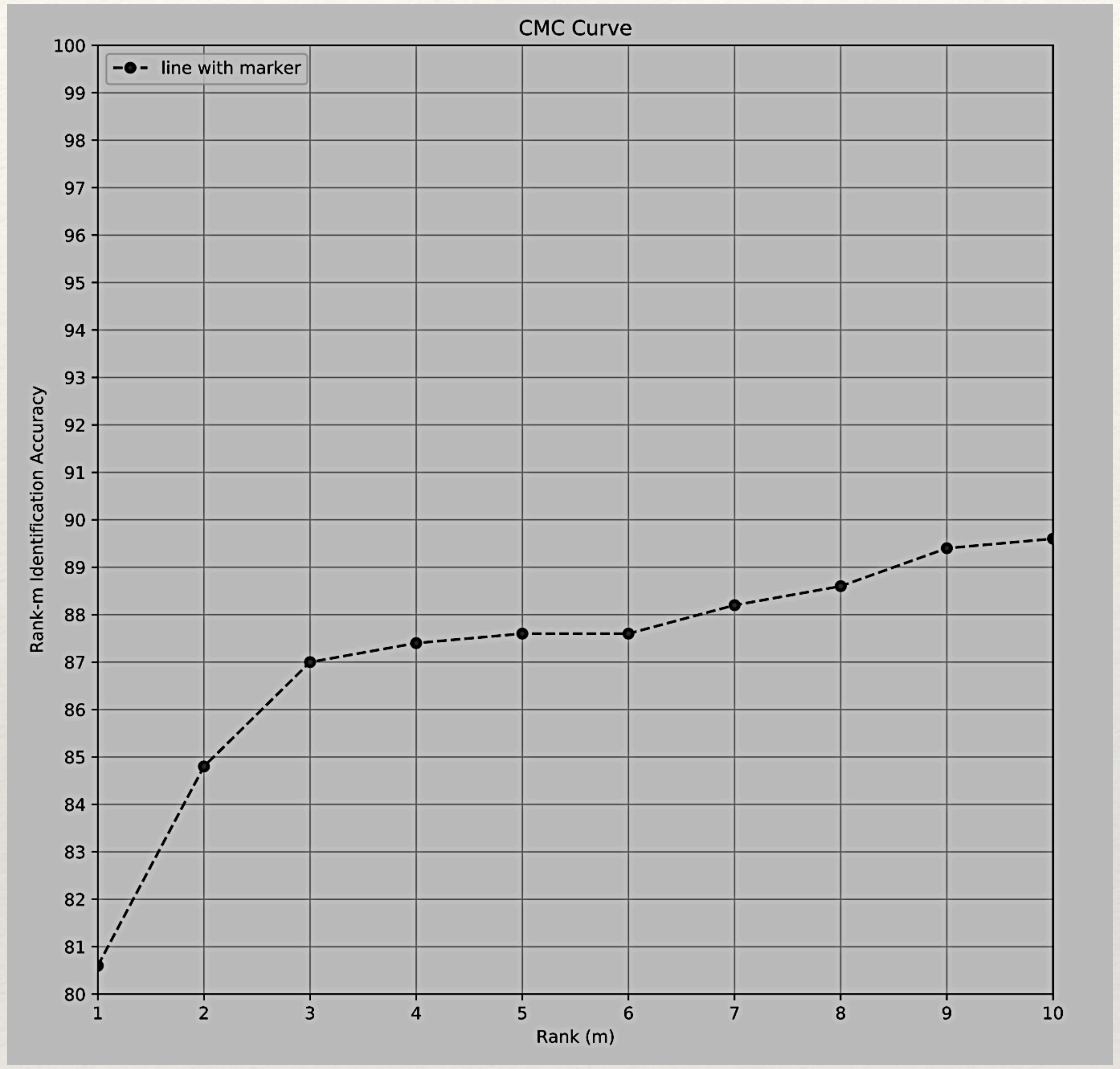
Model	Rank1		Rank10	
Dataset	LFW	SurvFace	LFW	SurvFace
LightCNN29	96.8	6.3	98.6	15.2
VGGFace2	**95.6**	**1.6**	**99.7**	**5.7**
ArcFace	80.7	1.8	89.6	5.8

Identification Accuracy

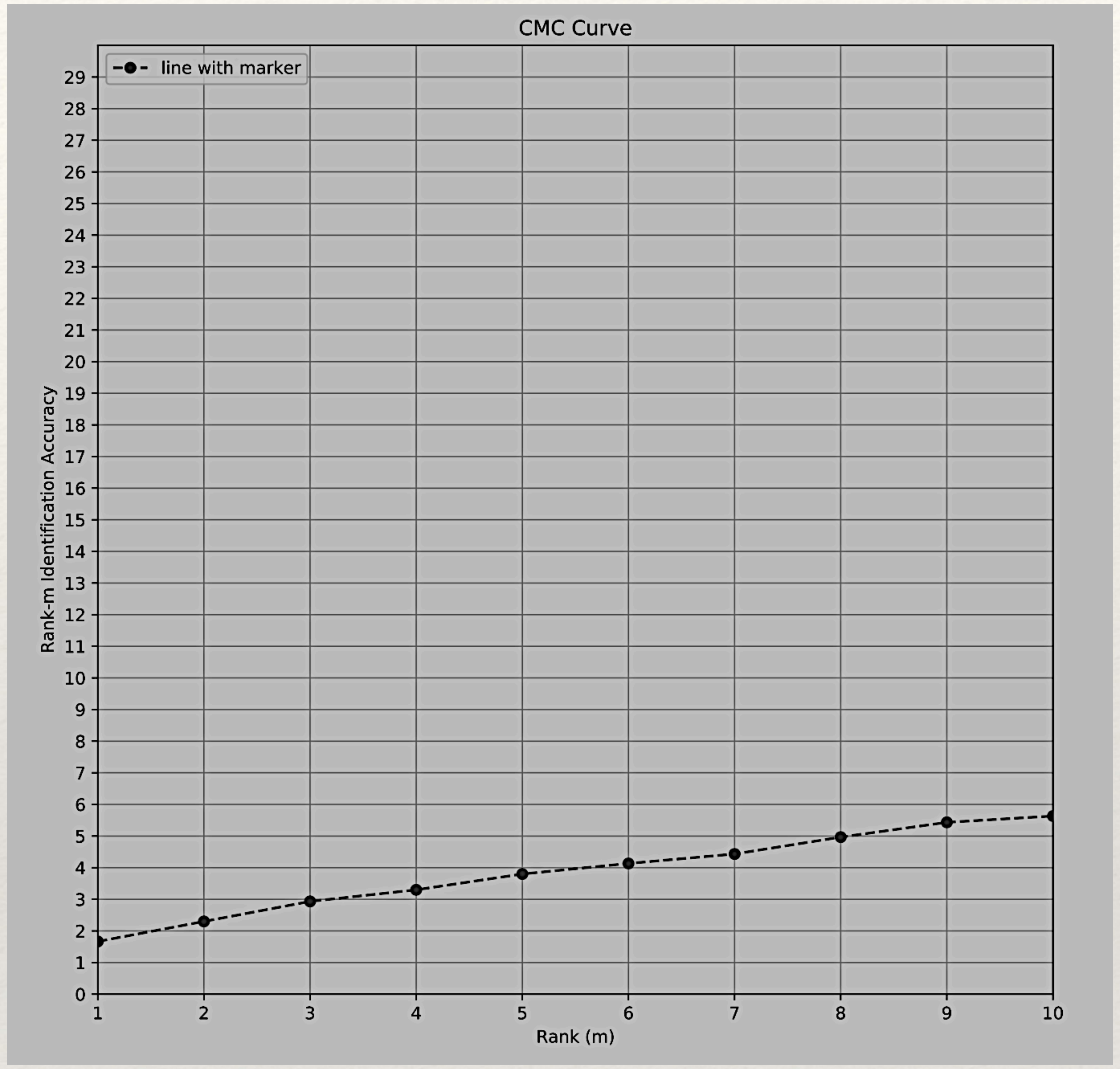
- ***Fine-tuning with LR images for VGGFace2 pre-trained model degrades performance***

VGGface2: A dataset for recognising faces across pose and age. Cao, Qiong et.al., FG 2018

Results - CMC Curves



ArcFace - LFW



ArcFace - QMUL-SurvFace

Model	Rank1		Rank10	
Dataset	LFW	SurvFace	LFW	SurvFace
LightCNN29	96.8	6.3	98.6	15.2
VGGFace2	95.6	1.6	99.7	5.7
ArcFace	**80.7**	**1.8**	**89.6**	**5.8**

Identification Accuracy

- ***Fine-tuning with LR images for ArcFace which is SOTA model degrades performance***

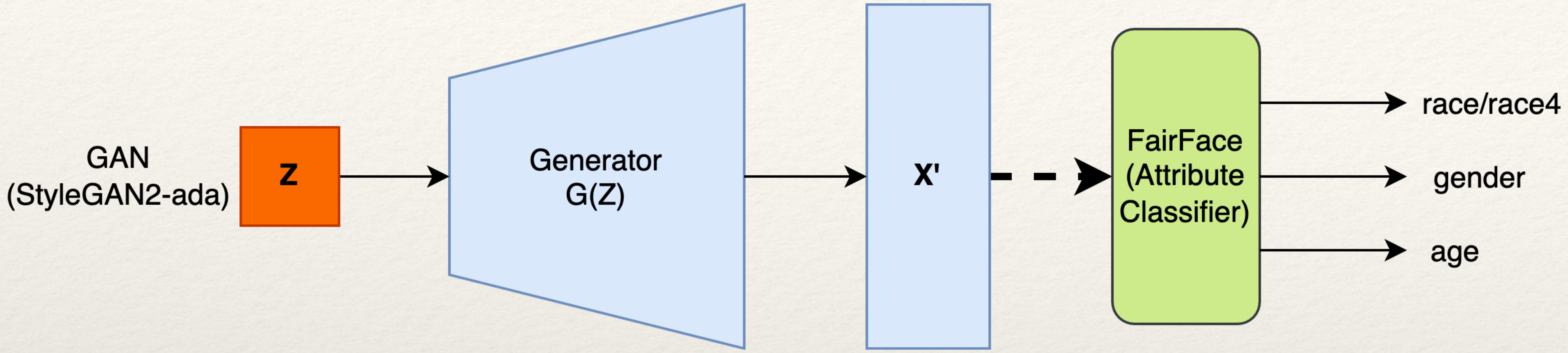
Arcface: Additive angular margin loss for deep face recognition Deng, Jiankang et.al., CVPR 2019

Summary and Outcomes

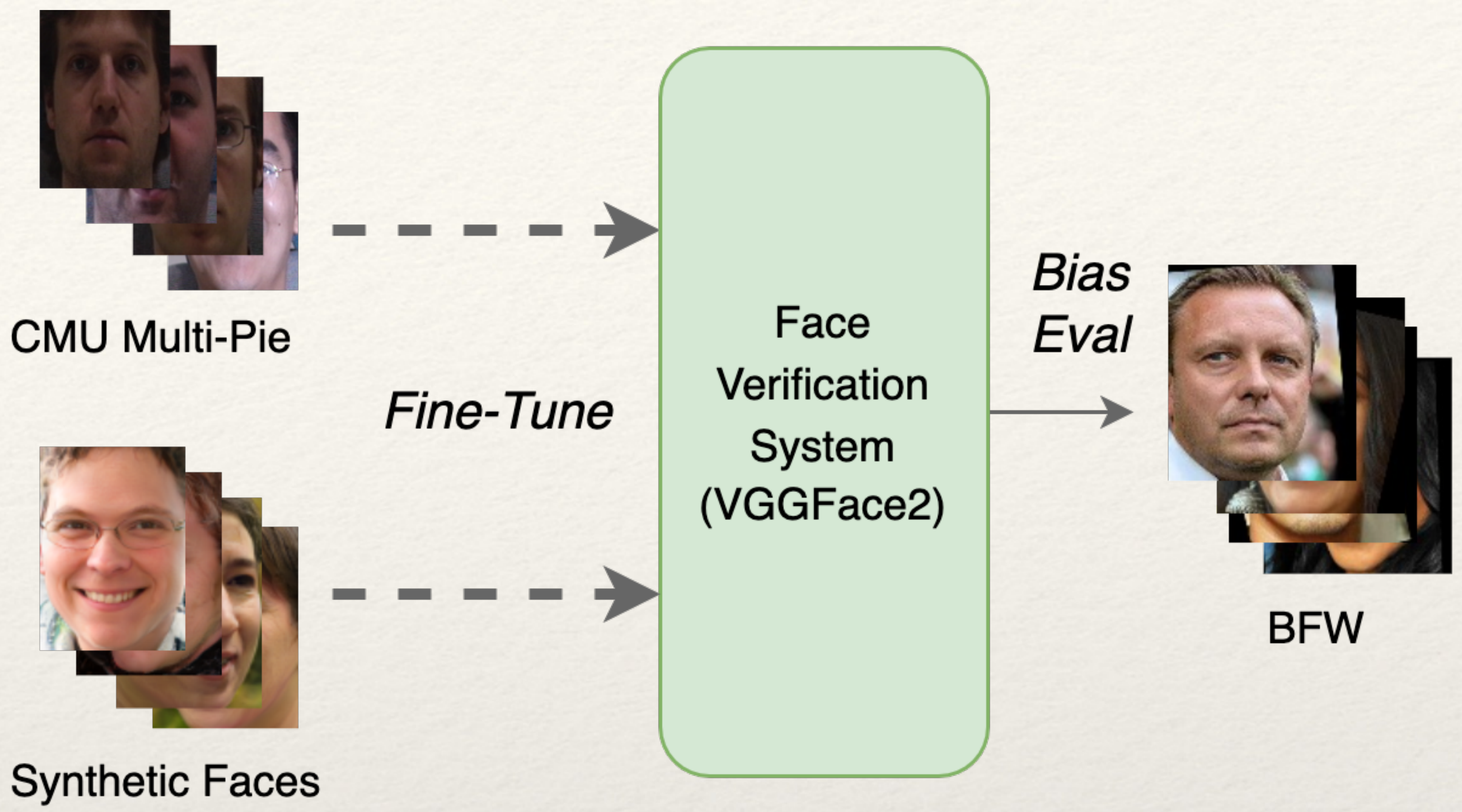
- Fine-tuning with High Resolution faces, gives good performance
- Fine-tuning with Low Resolution faces, gives worse performance
- Existing architectures cannot be adopted as-is for Low Resolution face recognition
- Architectures with GAN components may be unfair and biased

Bias and Fairness GANs

Datasets and Experiments

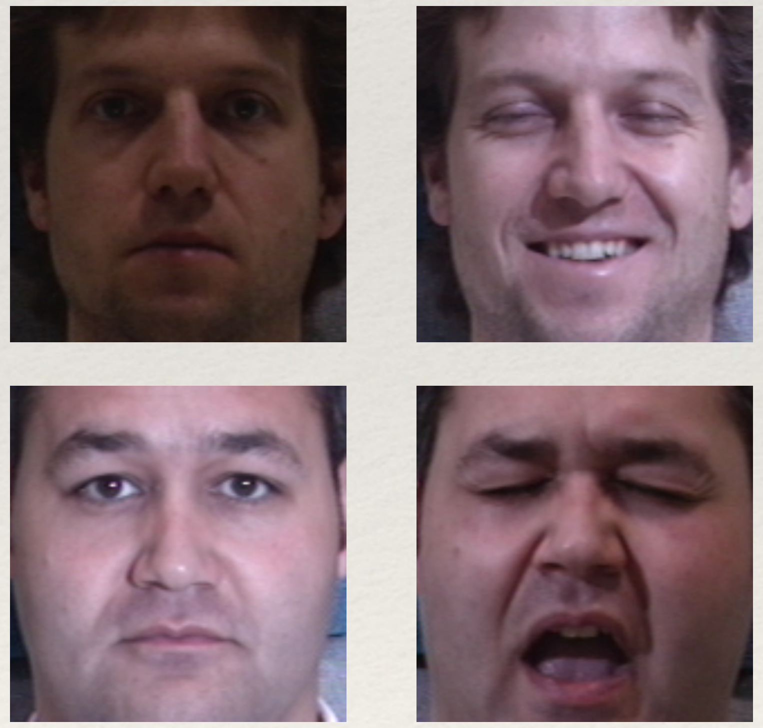


GAN Bias Estimator



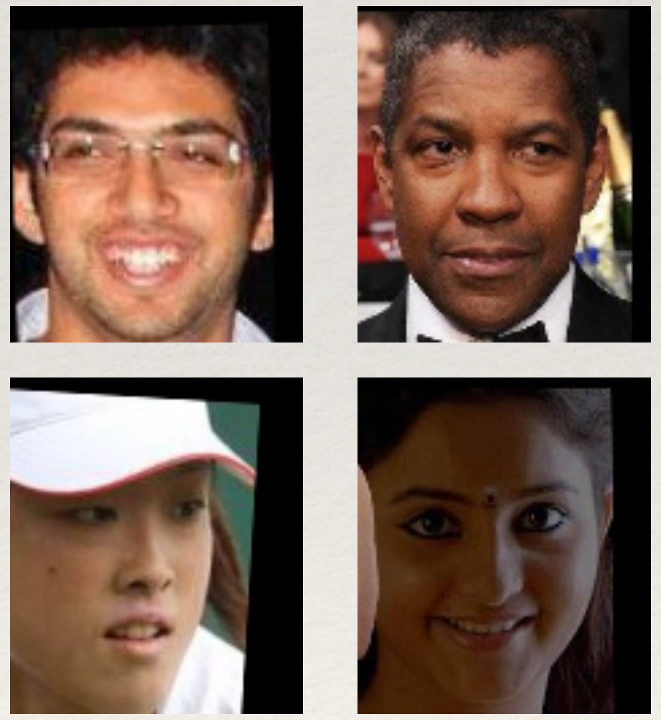
Bias Estimation in Face Verification System

FFHQ



CMU Multi-PIE

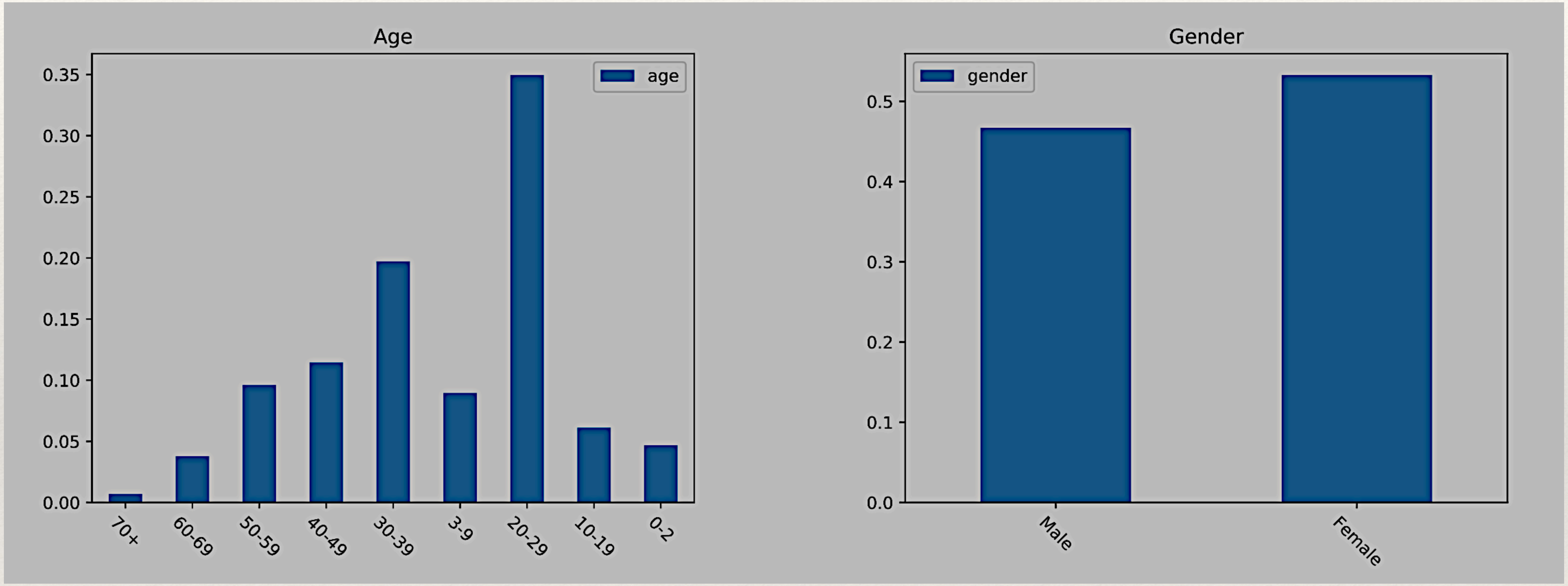
DiscoFaceGAN



BFW

Analyzing and improving the image quality of stylegan, Karras, Tao et.al., CVPR 2020
Training generative adversarial networks with limited data, Karras, Tao et.al., Neurips 2020
Disentangled and Controllable Face Image Generation via 3D Imitative-Contrastive Learning, Deng, Yu, et.al., CVPR 2020

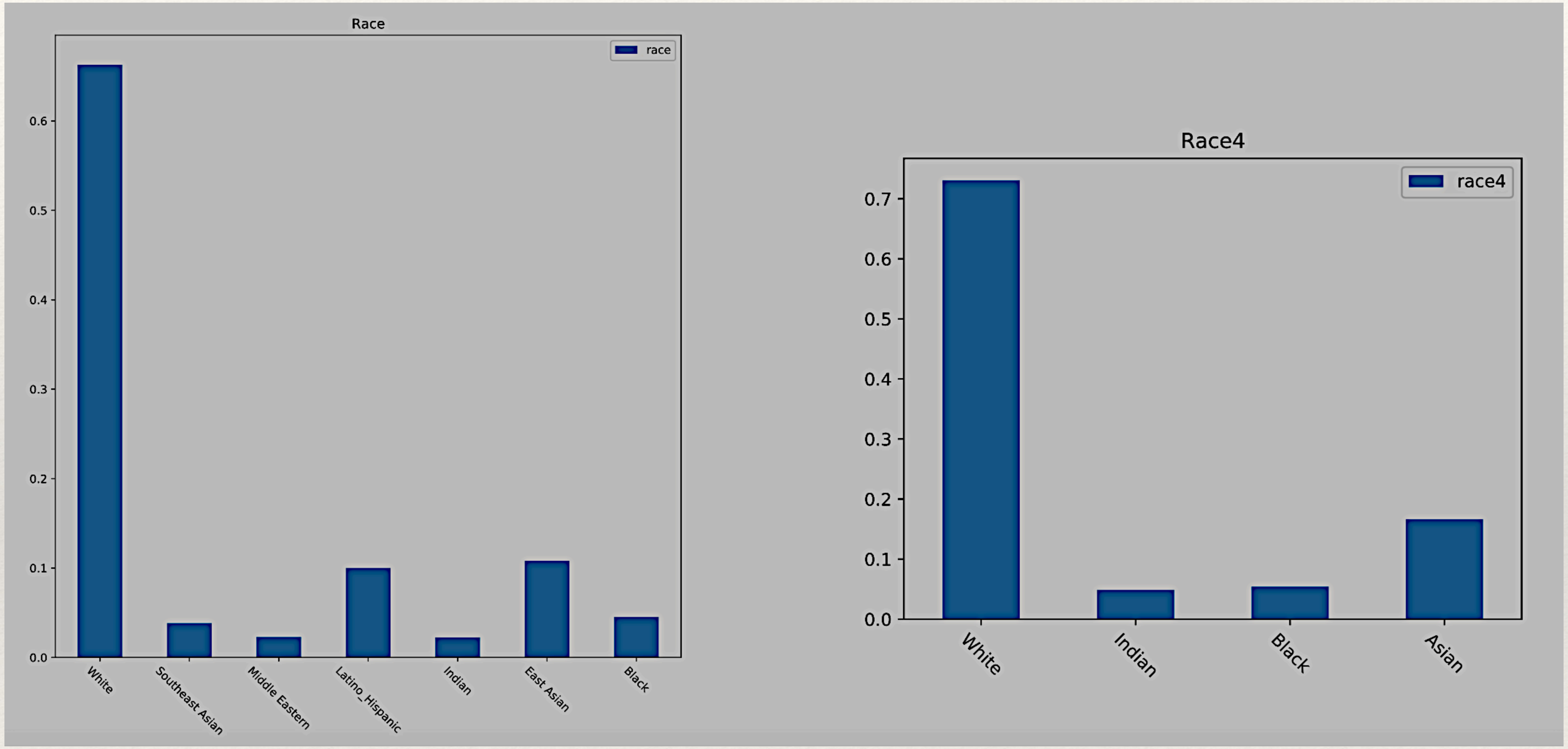
Results - Experiment1



- ***GANs Biased towards age group "20-29"***

A style-based generator architecture for generative adversarial networks. Karras et.al., CVPR 2019

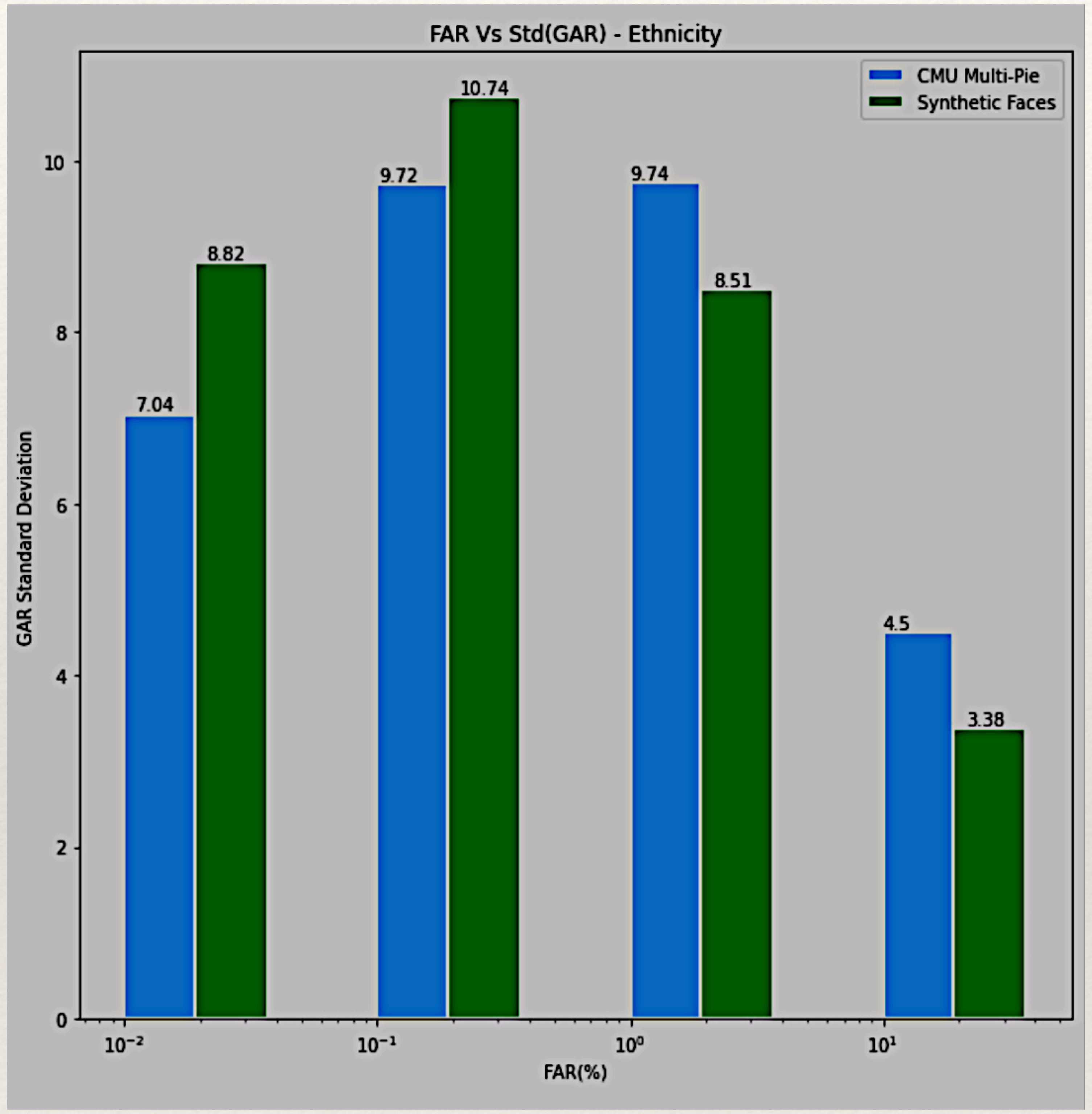
Results - Experiment1

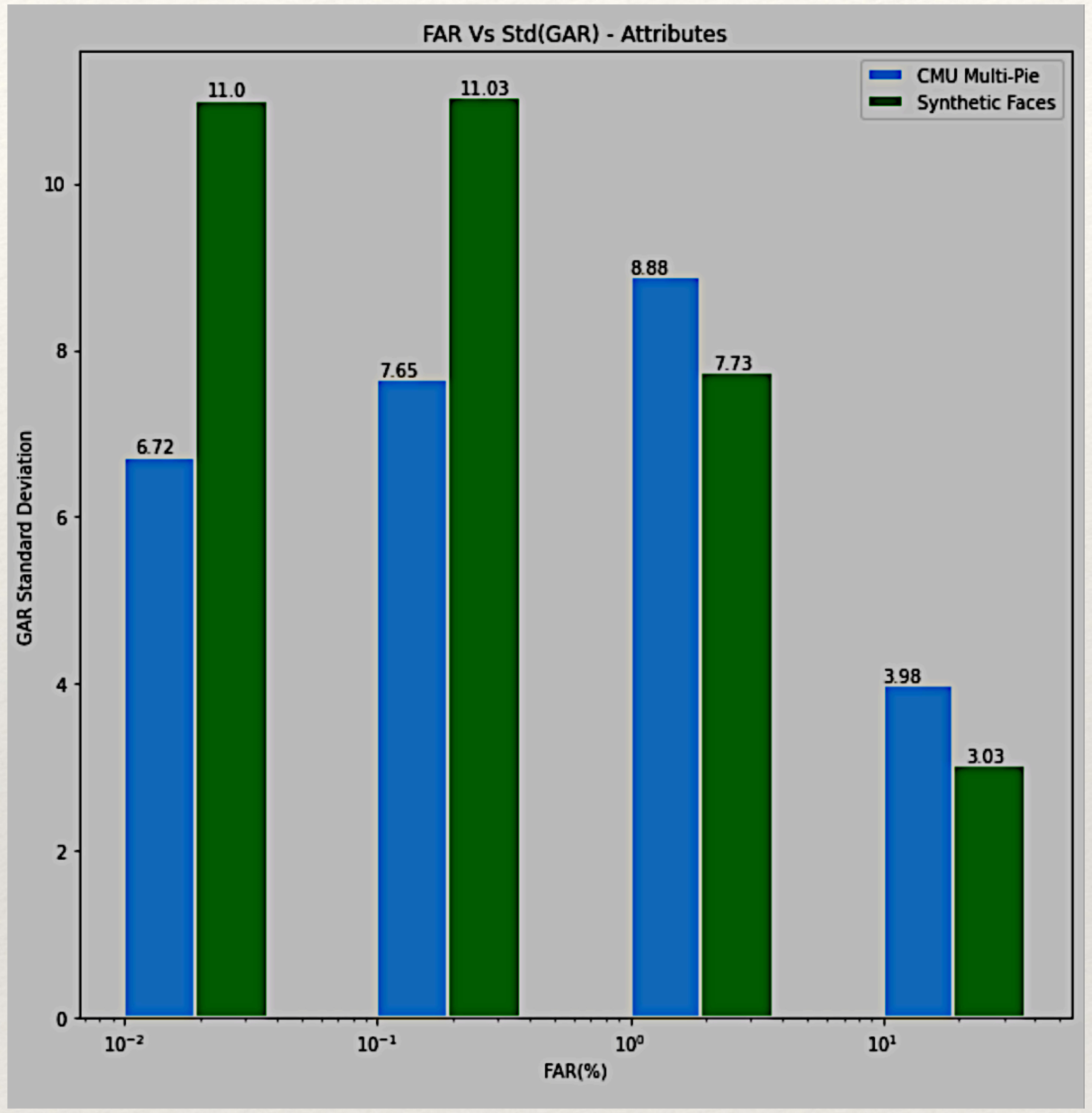


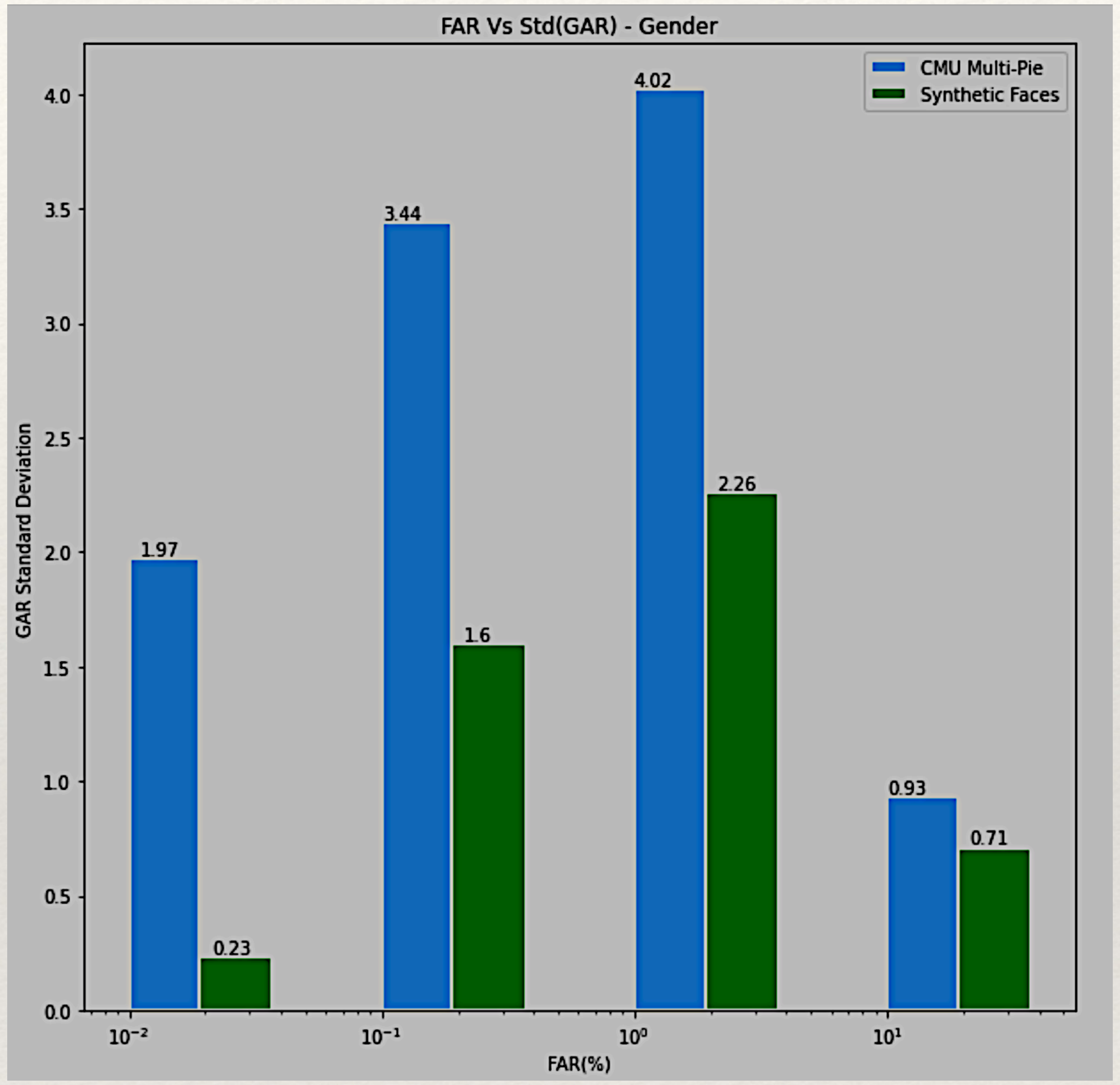
- ***GANs are biased towards "white" faces***

Face recognition: too bias, or not too bias?, Robinson, Joseph et.al., CVPRW 2021

Results - Experiment2







- ***Face Verification models trained or fine-tuned with Synthetic faces exhibit bias for ”race” attribute***

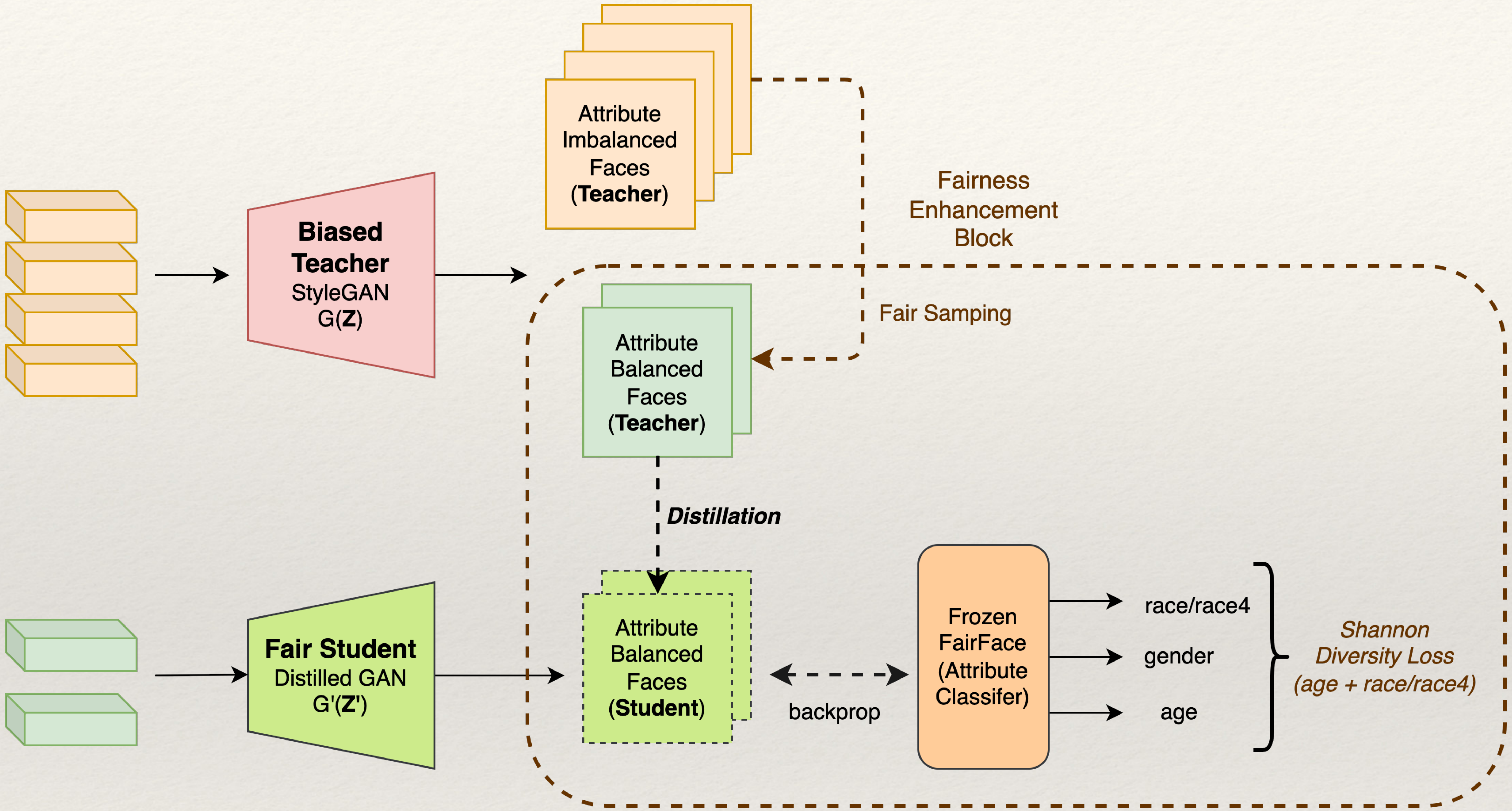
Muti-pie, Gross, Ralph et.al., Image and vision computing 2010
Fairface: Face attribute dataset for balanced race, gender and age for bias measurement and mitigation, Karkkainen et.al., WACV 2021

Summary and Outcomes

- GANs are biased towards age group ”20-29” and ”White” faces
- Face Verification models trained or fine-tuned with Synthetic faces exhibit bias for “race” attribute
- Face Verification models trained or fine-tuned with Synthetic faces doesn’t exhibit any bias for ”gender” attribute
- At high FAR rates no bias (low DoB_{fv}) is observed (Hypothesis : Biases masked by high false acceptances)

Fair Distillation GANs

Fair Distillation - GANs



- Fair Sampling : Image sampling guided by Fairface classifier attribute sub-groups

- Shannon Diversity Loss :

$$L_{SD} = \sum_{attribs} \left(1 - \frac{\Sigma_i^{\,k} p_i log(p_i)}{log(k)}\right)$$

- Architecture for mitigating bias via Fair Distillation of GANs

Tinygan: Distilling bigger for conditional image generation, Chang et.al., ACCV 2020

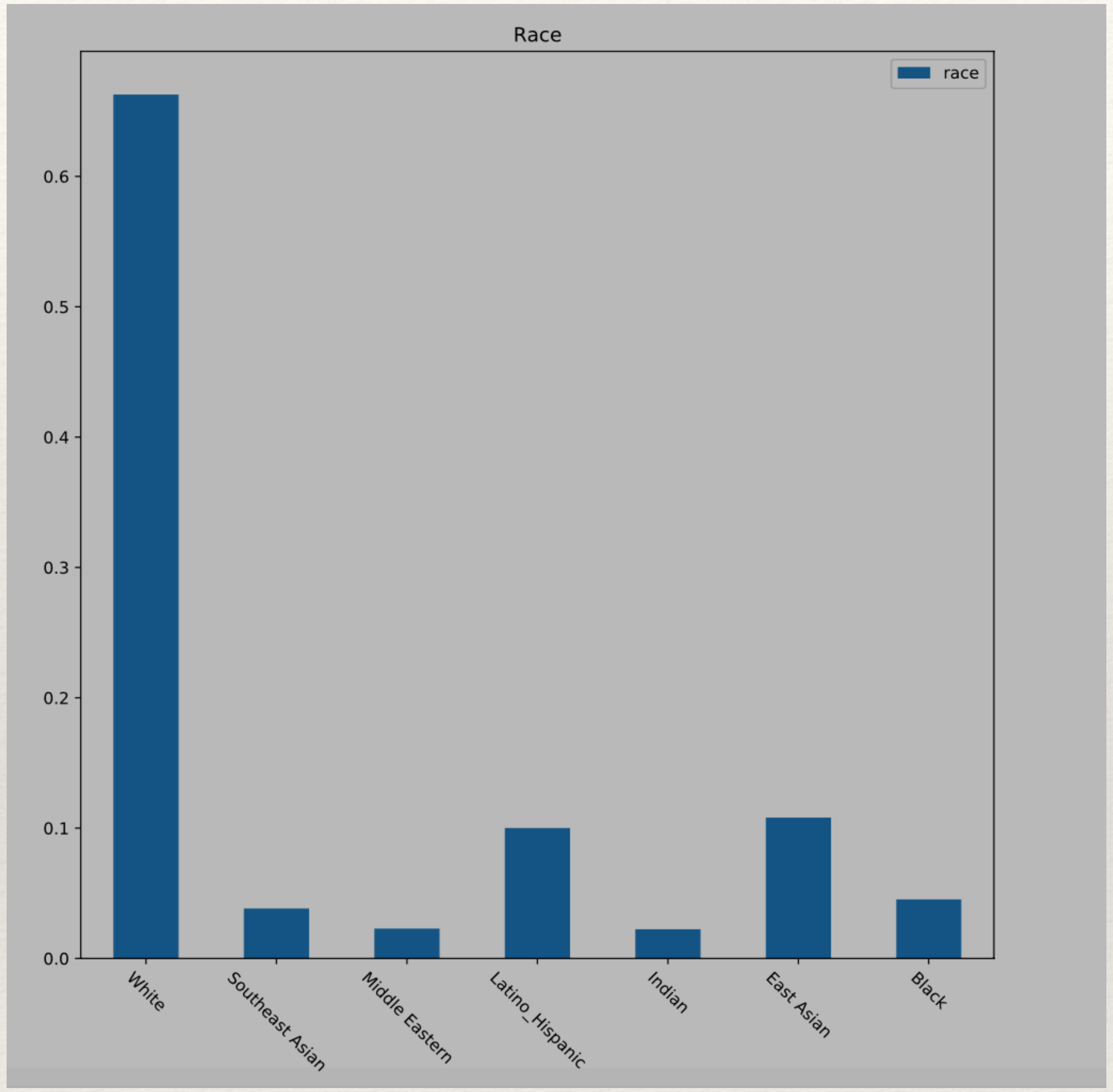
Results

Teacher GAN

Student GAN

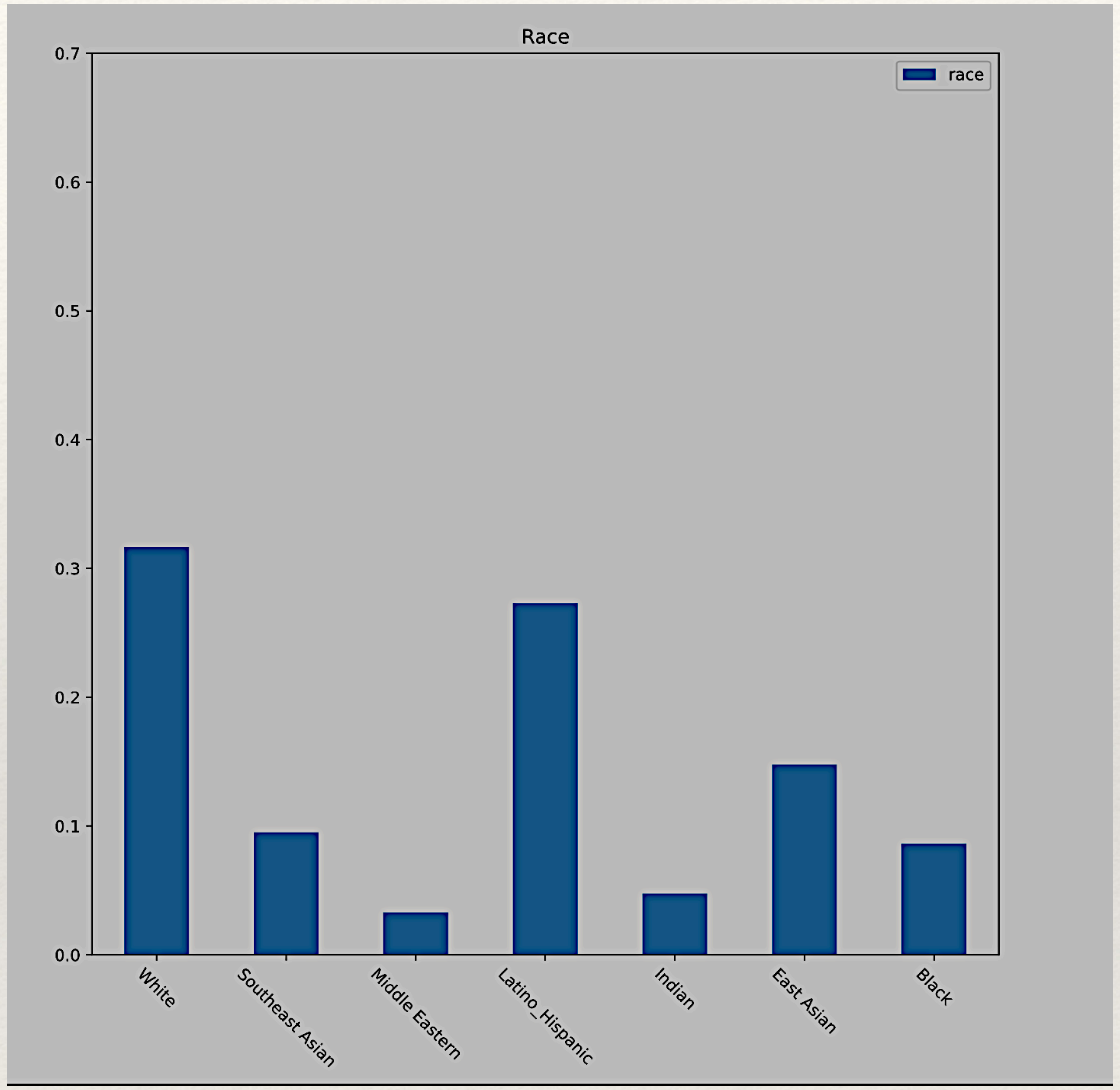
- ***Higher Quality faces from Teacher GAN***
- ***Lower Quality faces from Student GAN***

Results - Race



Teacher GAN

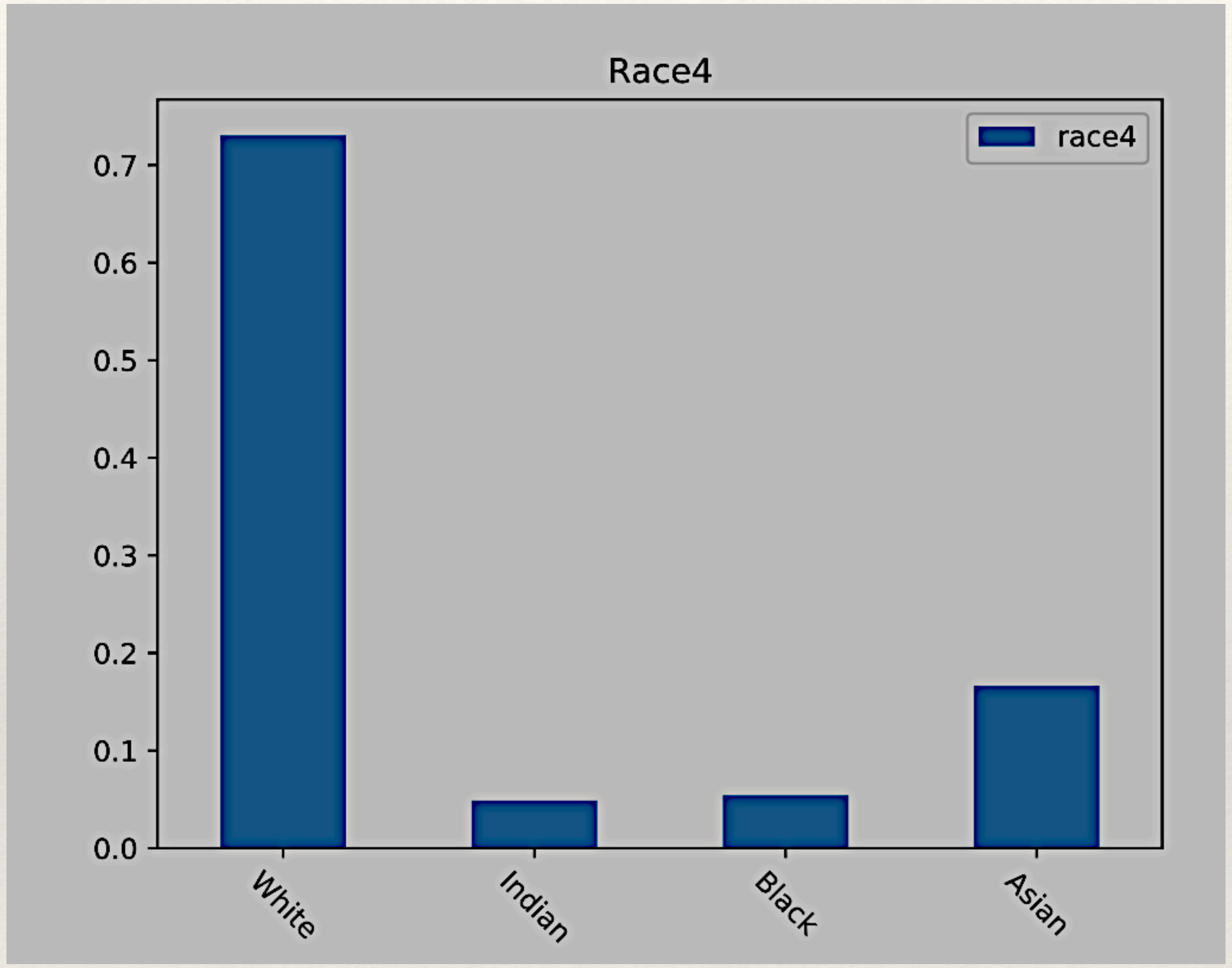
Vs



Student GAN

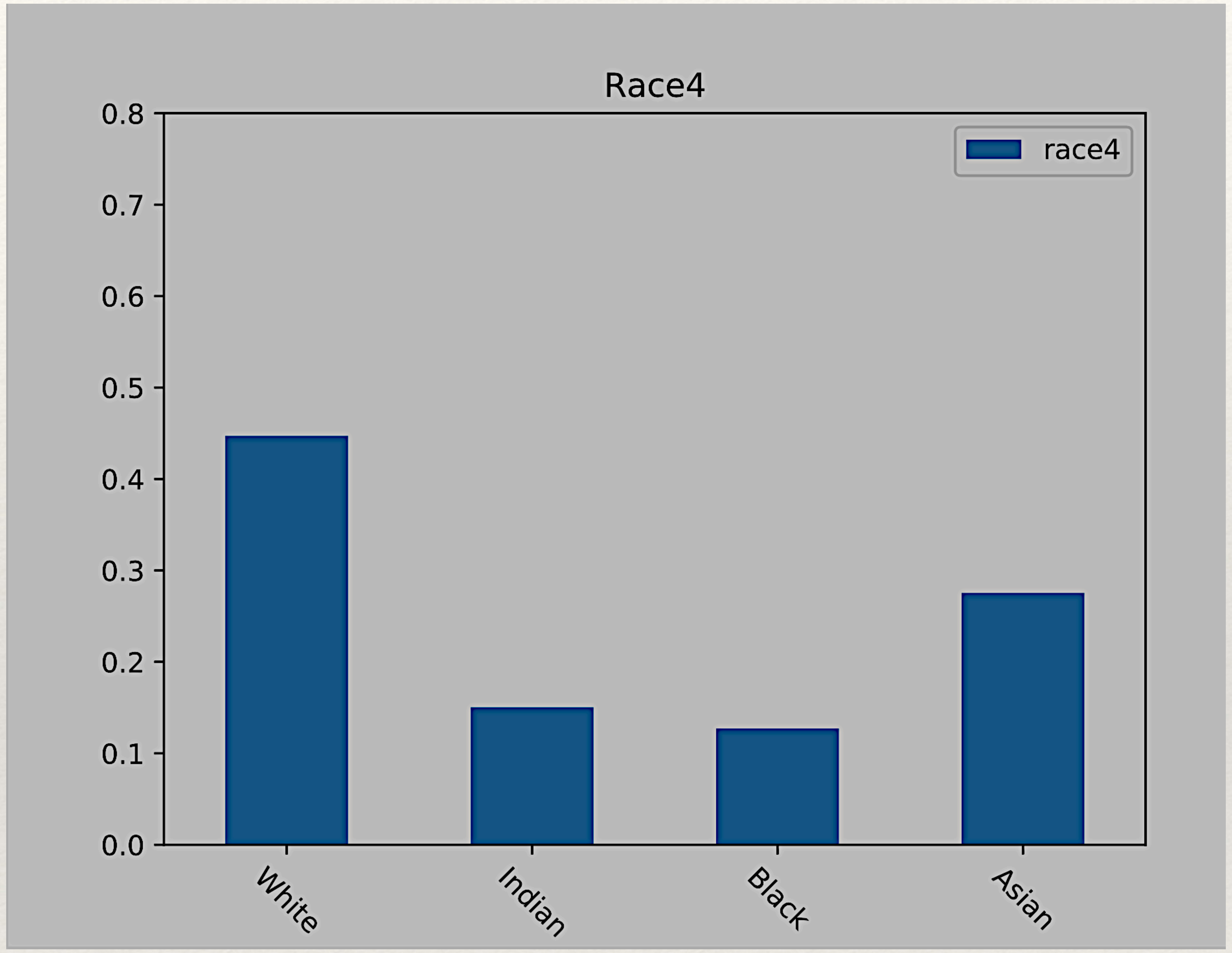
StudentGAN is fairer than Teacher GAN for Race

Results - Race4



Teacher GAN

Vs



Student GAN

StudentGAN is fairer than Teacher GAN for Race4

Summary and Outcomes

- Fair distillation improved imbalance/fairness for the Student GAN especially for Race/Race4 attribute
- Improvement of imbalance/fairness with respect to age attribute across sub-groups is not substantial
- The quality of faces from Student GAN is lower than the quality from Teacher GAN

Publications and Outcomes

- On Biased Behavior of GANs for Face Verification - Responsible Computer Vision workshop, ECCV 2022
- Accepted to Research Week with Google - 2022 and Research Week with Google - 2023
- Fortunate and thankful to get the opportunity to perform High Quality Research and develop love for Research

Conclusion & Future Directions

- Fine-tuning is insufficient for adopting existing architectures for low resolution face recognition
- SOTA architectures for LR face recognition may be unfair and biased
- GANs are biased towards different sub-groups and can impact downstream models
- FairDistillation for debiased Student GANs
- Better loss formulations and architectural improvements to prevent bias in all categories of generative models
- Bias across different Computer Vision tasks such as Object Detection, Object Segmentation and Others

Thank you