Can Finetuning Overcome Racial Bias In Face Recognition Models?

Linh Tran, Christopher Kanan
University of Rochester

MOTIVATION

Most face datasets are **racially unbalanced**.

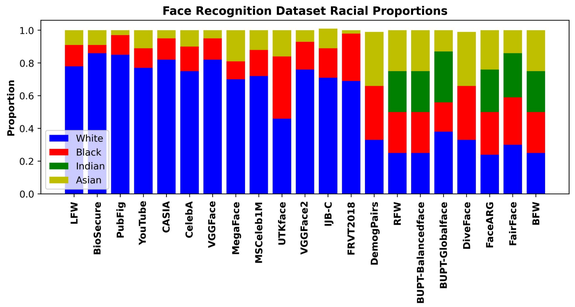


Figure 1. Dataset Racial Distributions in Popular Datasets [1]

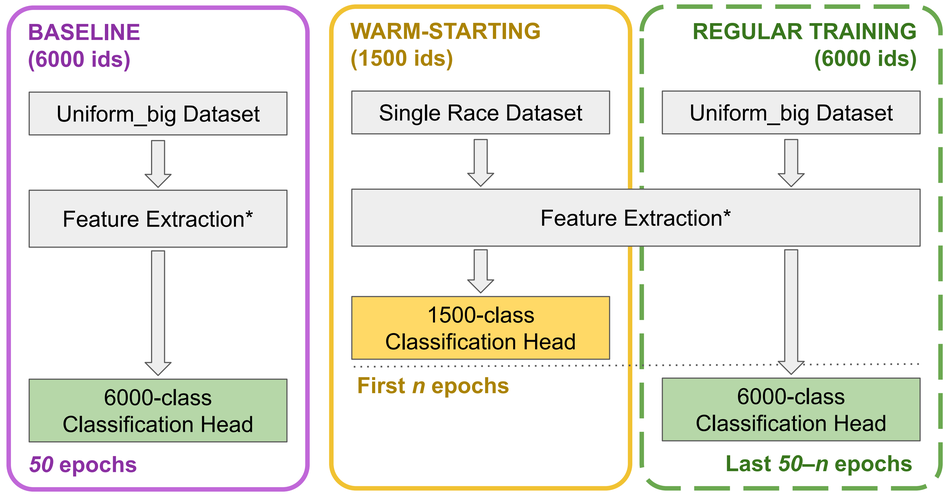
Face recognition models are **racially-biased**, even with **balanced training data** (Table 1).

OVERVIEW

- **Critical period**: the initial training period when model learns exceptionally quickly [2].
- **Warm-starting**: when training, model is introduced to a subset of the training data before seeing all training data [3].

Hypothesis: **racial bias could be mitigated**:
1. Warm-starting the model with subset of data of under-performing races in the first few epochs.
2. Fine-tuning single-race model using uniform data.

WARM-STARTING EXPERIMENT



(*) Feature Extraction Architecture: *VGGm-11* or *ResNet-18*

Results of Warm-starting on **underperforming races: Asian and African**

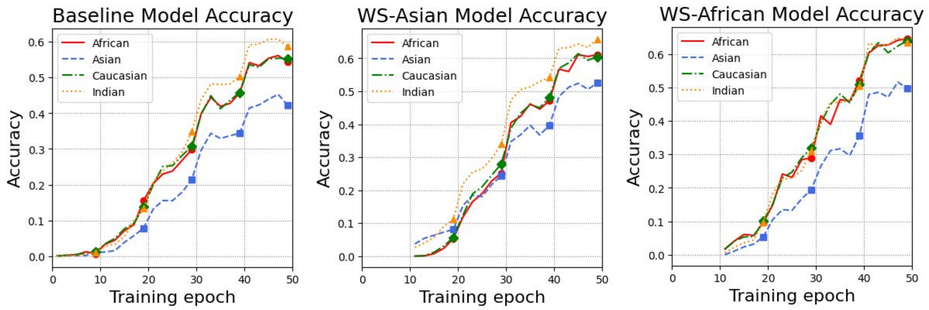


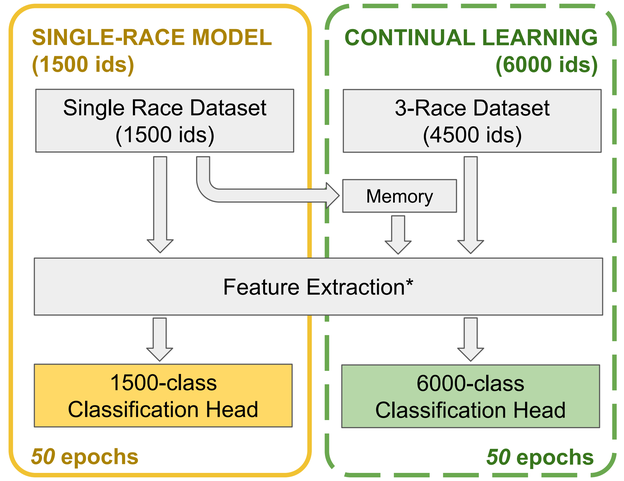
Figure 1. Model's accuracy on different races during training

Control	None	Race	Epoch	Race & Epoch
P-Values (Baseline vs WS-African)	0.8874	0.8462	0.3022	0.1157
P-Values (Baseline vs WS-Asian)	0.0408*	0.0412*	0.0482*	0.0485*

**statistically significant at 5%*

Table 2. Relative Accuracy to Caucasian by Epoch and Race

CONTINUAL LEARNING EXPERIMENT



(*) Feature Extraction Architecture: *VGGm-11* or *ResNet-18*

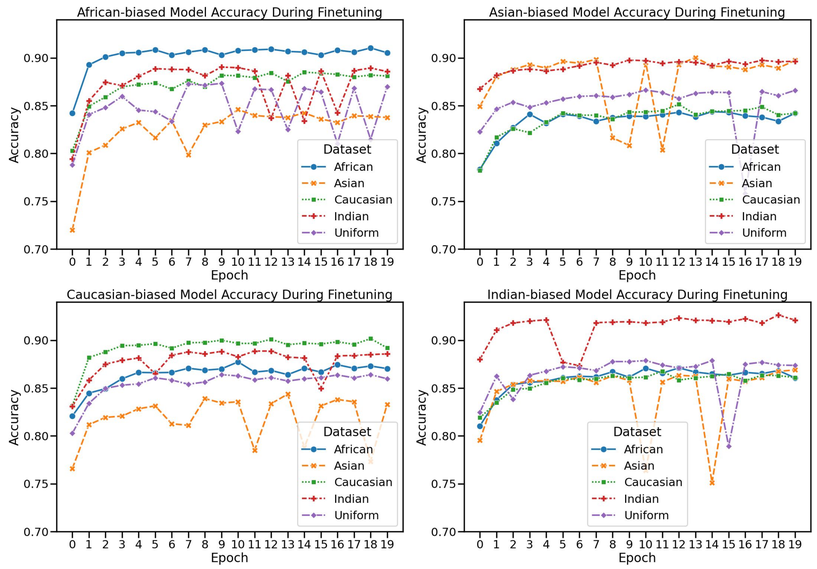


Figure 2. Model's accuracy on different races during fine-tuning

DATASETS

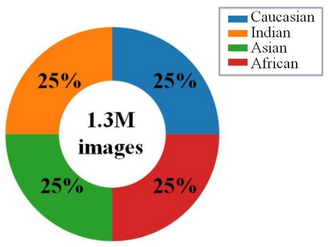
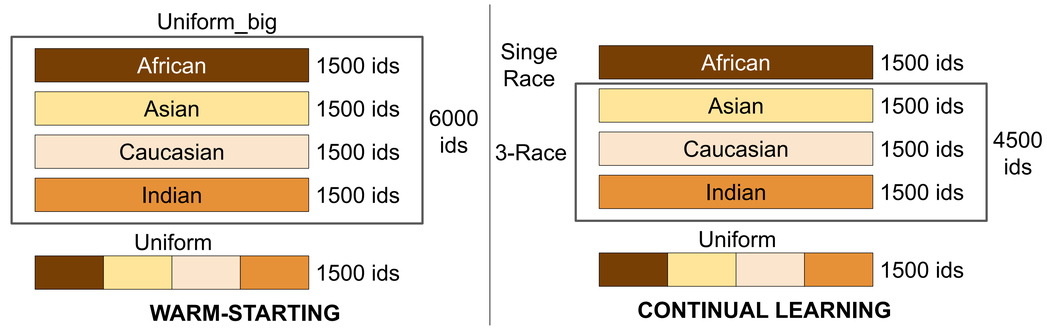


Fig 2. BUPT-BalancedFace [5]

Training:
BUPT-BalancedFace [5]:
- 1.3M images
- 7000 identities/race

Testing:
RFW: Racial Faces in The Wild [6]:
- 3000 identities/race



REFERENCES

1. Sumsion A, Torrie S, Lee D-J, Sun Z. Surveying Racial Bias in Facial Recognition: Balancing Datasets and Algorithmic Enhancements. Electronics. 2024; 13(12):2317. https://doi.org/10.3390/electronics13122317
2. Alessandro Achille, Matteo Rovere, and Stefano Soatto. Critical learning periods in deep networks. 2018.261
3. Jordan T. Ash and Ryan P. Adams. On warm-starting neural network training, 2020. URL https://arxiv.org/abs/1910.08475.266
4. German I. Parisi, Ronald Kemker, Jose L. Part, Christopher Kanan, and Stefan Wermter. Continual lifelong learning with neural networks: A review. Neural Networks, 113:54–71, 2019. ISSN 0893-6080. doi: https://doi.org/10.1016/j.neunet.2019.01.012. URL https://www.sciencedirect.com/science/article/pii/S0893608019300231.298
5. Mei Wang, Weihong Deng. Mitigating Bias in Face Recognition using Skewness-Aware Reinforcement Learning. CVPR2020.
6. Mei Wang, Weihong Deng, Jiani Hu, Xunqiang Tao, Yaohai Huang. Racial Faces in the Wild: Reducing Racial Bias by Information Maximization Adaptation Network. ICCV2019.