



# **Facial Profiling:** Understanding how people and computers “see” racial-ethnic physical characteristics

Dr. Nick Petersen (Sociology & Law)  
Rahul Dass (ULINK Pre-doctoral Fellow, Computer Science)

# N.Y. Teen Blames Apple's Facial Recognition for Wrongful Arrest, Files \$1B Lawsuit

By **Tanasia Kenney** - April 29, 2019

👁 4084 💬 0



Ousmane Bah (left) was arrested but let go after a detective realized he looking nothing like the suspect (right) caught stealing in a from an Apple store. (Photos courtesy of KTVU and Subhan Tariq)

*ICE Used Facial Recognition to Mine State Driver's License Databases*





Police receive an **inaccurate identification** from the software

Police **arrest** an innocent person

Search-Incident-to-Arrest

Police find **contraband**



**FIRED**



Exclusionary Rule?

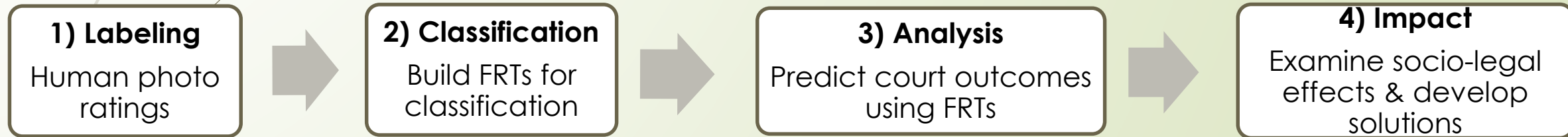


# Facial Recognition Technology (FRT)

- FRTs **detect**, **identify** or **analyze** people's faces
  - E.g., facial recognition software and machine learning models.

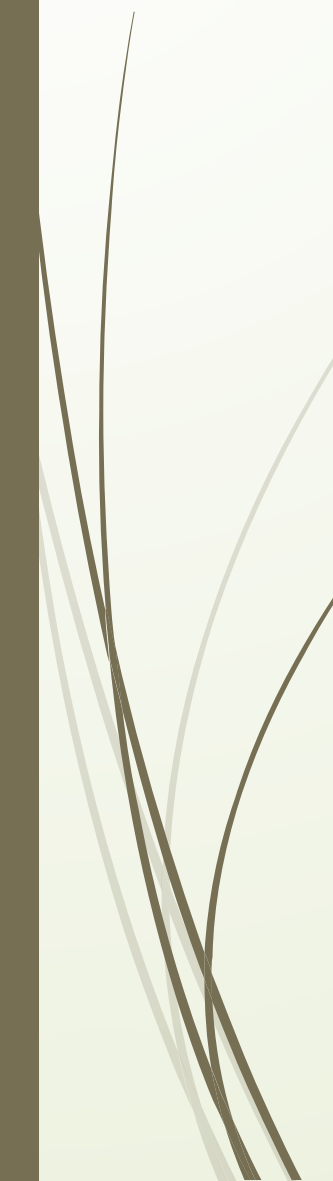


# ML pipeline





# Research Questions

1. How do humans, computers, and the law “see” race/ethnicity?
  2. What drives FRTs to making decisions: Race, ethnicity, skin tone, facial landmarks?
  3. Do FRTs reinforce racial inequalities? If so, how can they be mitigated?
  4. Concerns with FRTs
  5. Social implications
- 



# Data and methods

- Linked mugshots (N=200,000) to Miami court records (race, ethnicity, age, sex etc.)
- Student rated data (N=14,000) providing additional labels (skin tone, Afrocentric features, facial hair, attractiveness etc.)
- Modified SOTA ML models to classify race, and ethnicity by using court and student labels, and comparing ML accuracies
- Investigating if skin tone measures can be “learned” by ML models?
- Discussed socio-legal implications of FRTs

# Student rated data (N=14,000)

- 20+ students rated  $\approx$  14,000 photos
  - Each photo rated by 3 people for reliability
- Ratings as FRT inputs (race, ethnicity, skin tone)

Enter the Picture ID Number here.

What race/ethnicity do you perceive this person to be?

Black Hispanic   Black non-Hispanic   White Hispanic   White non-Hispanic   Other

Rate the degree to which this face has features that are typical of **Hispanic** people.

1 (Not Typical)   2   3   4   5   6   7 (Very typical)

Rate the degree to which this face has features that are typical of **black** people.

1 (Not Typical)   2   3   4   5   6   7 (Very typical)

How would you rate this person's **skin color**?

1 (Very Light)   2   3   4   5   6   7 (Very Dark)

Enter the Picture ID Number here.

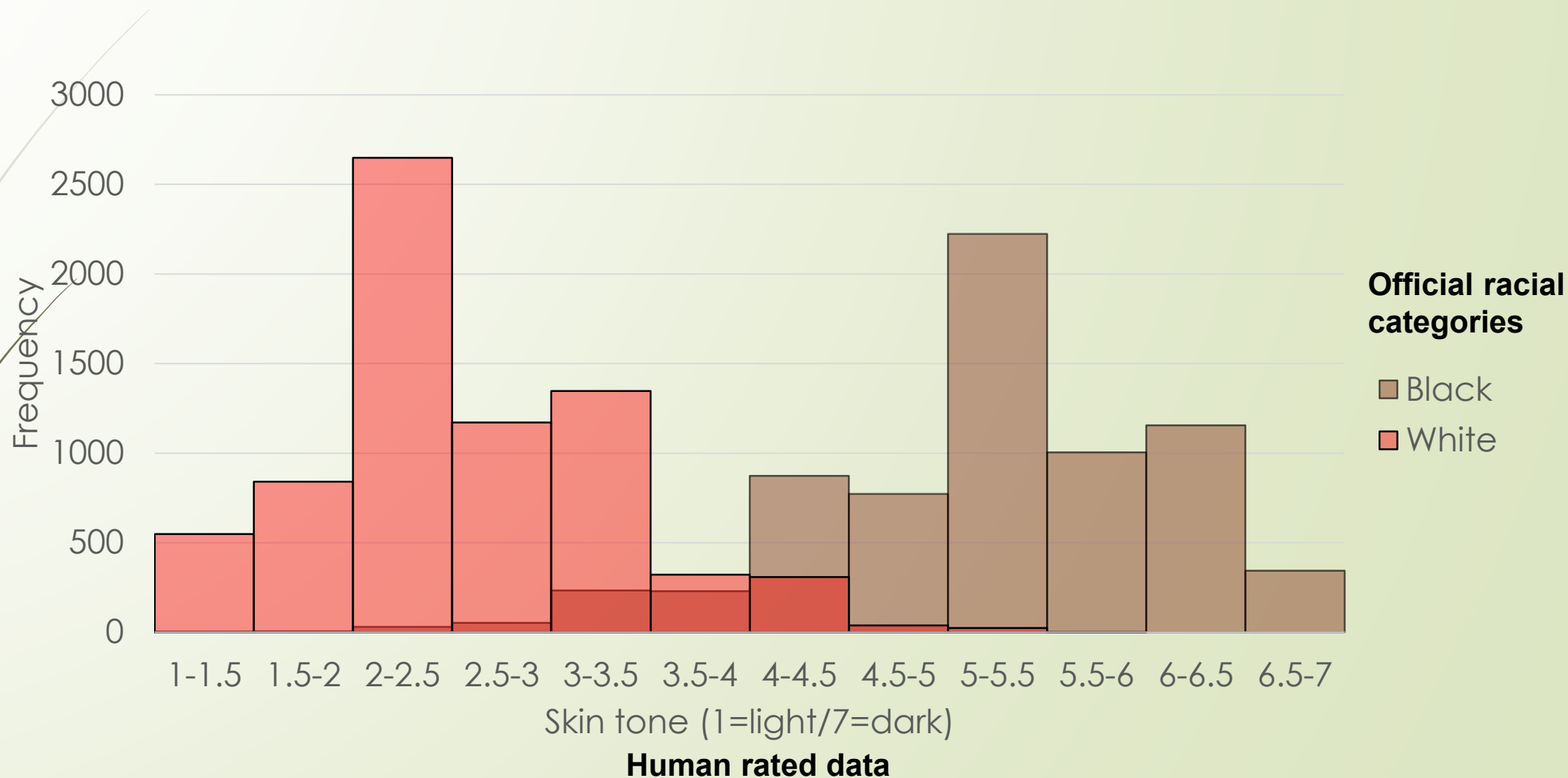
What race/ethnicity do you perceive this person to be?

Black Hispanic  
Black non-Hispanic  
White Hispanic  
White non-Hispanic  
Other

Rate the degree to which this face



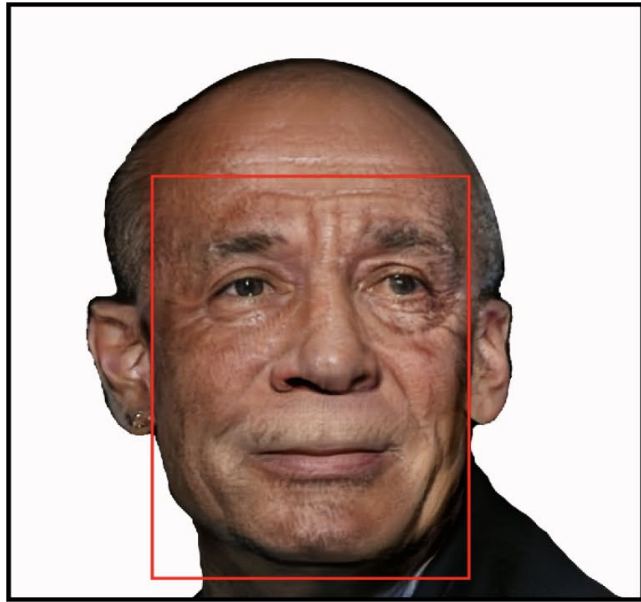
# Skin tone by official racial categories for student rated data (N=14,000)



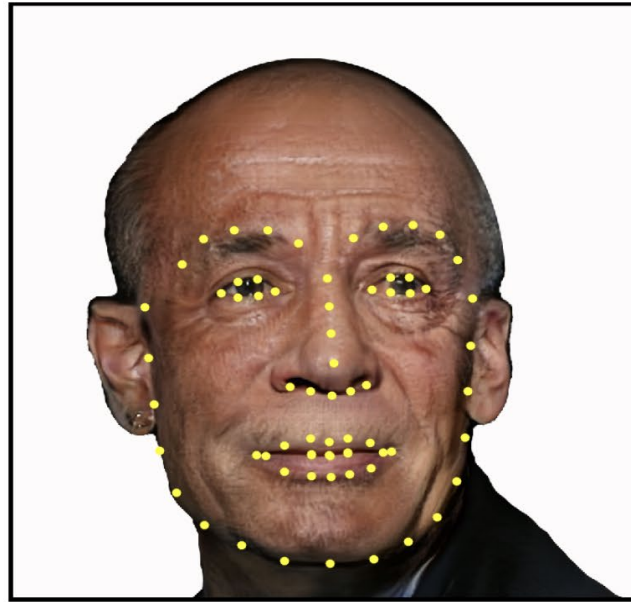
# Outline methods for ML

- **SOTA ML libraries:** Keras + TensorFlow; fastai + PyTorch
- **Models tested:** VGG-16, VGG-Face, ResNet-50
- **CS Data Center:** 2 x Nvidia Tesla P100 GPUs
- **Phase 1: feeding raw images**  
(subject to extraneous features such background and clothing)
- **Phase 2: preprocessed images using OpenFace**  
(face detection + landmark detection + face alignment + cropping)

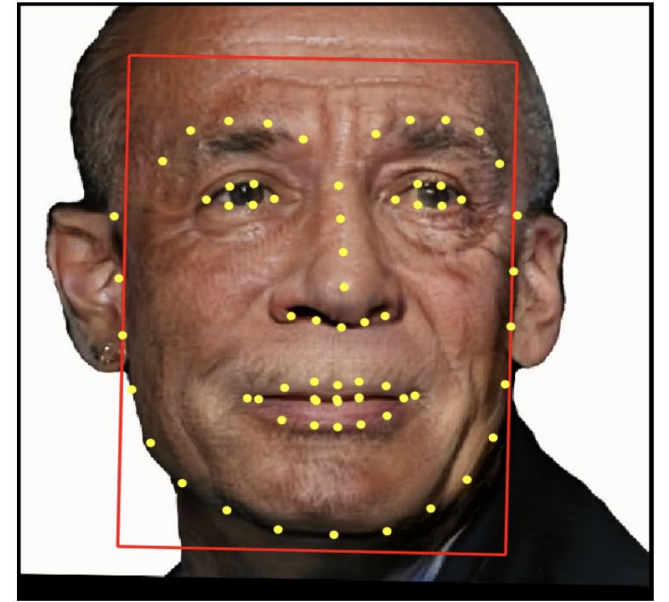
# Example: Data Preprocessing



Face Detection



Landmark Detection



Cropping and Resizing

Source: Merler et al., *Diversity in Faces*, 2019 (IBM)



# Classification accuracies for race vs. race-ethnicity

## ❑ Race (White vs. Black)

- ❑ Controlled (**2K samples**) = 91.72%
- ❑ Imbalanced (**200K samples**) = 97.21%

## ❑ 4-Race/Ethnicity groups:

- ❑ Controlled (**4K samples**) = 70.71%
- ❑ Imbalanced (**200K samples**) = 80.93%



Non-Hispanic White



Non-Hispanic Black



Hispanic White



Hispanic Black

# Classification accuracies for skin tone

## ❑ 7-skin tone measures (Likert scale)

- ❑ Controlled (**399 samples**) = 63.97%
- ❑ Unbalanced (**14K samples**) = 64.39%

## ❑ Comments:

- Insufficient data to "learn" skin tone
- Students' data had 32-unique values, thus rounding issue

## ❑ Possible steps to improve skin tone classification:

1. Convert problem from image classification to image regression
2. Create a "Deep Learning System"
  - Combining images + metadata (race, ethnicity, landmark measures)

# Next steps...

## Assessing Dataset/Algorithmic Biases in SOTA ML models

- Several learning paradigms: (re-)training and (re-)testing models based on single/multi-classification tasks
- Evaluating our models on other benchmark datasets such as CelebA, PPB, UKT-Face
- Assessing different ML metrics for **balanced vs. unbalanced** datasets



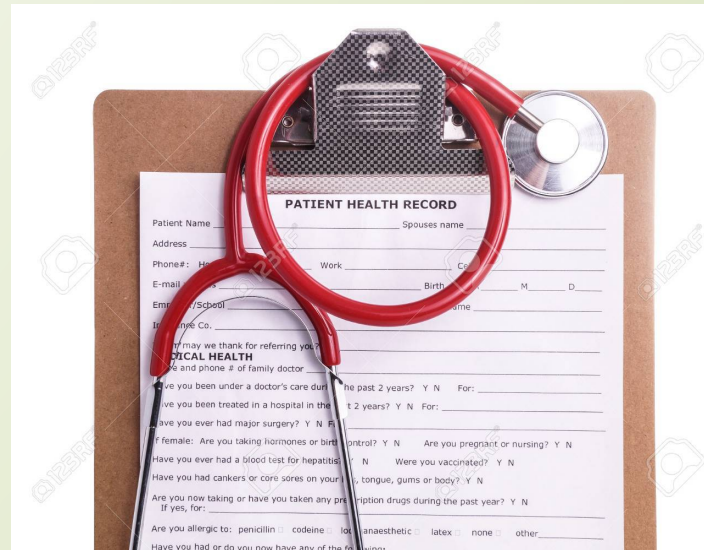
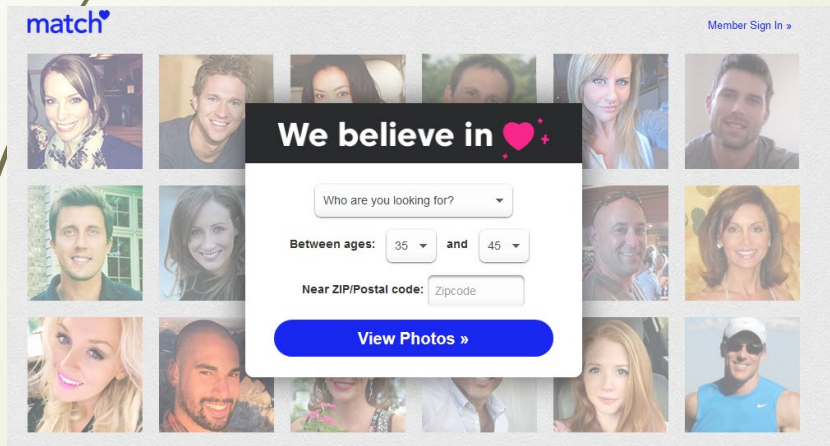
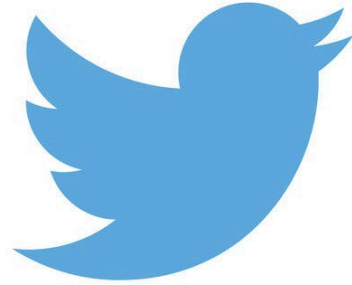


# Concerns with FRTs

- Maintaining individual's privacy
  - Data storage and accessibility concerns
  - Lack of policies for “proper” use of FRTs
  
- Black box vs interpretable AI?
  - Images are fed to FRTs and classifications are outputted
    - 1) too complicated function for humans to understand
    - 2) simple model but decision details unavailable
  
  - If “high-stakes” decisions were to be made, do we want inspect + interpret FRTs’ predictions by human experts

# Applications

- criminal justice
- social media
- online dating
- medical studies



# Thanks to...



## **Team members:**

Dr. Marisa Omori (Sociology)  
Dr. Tamara Rice Lave (Law)  
Dr. Ubbo Visser (Computer Science)  
Dr. Cameron Riopelle (Library)

## **Provost's Research Office:**

ULINK Phase I funding and Fellowship