Topic:

- Facial recognition has gotten faster, cheaper, and more accurate in recent years. What are the ethical implications of companies using this technology and their security systems to track consumer activity? Specifically, how could racial biases contained in facial recognition models create ethical problems when the technology is implemented?
- Research side: Redo (ish) previous iterations of facial recognition racial bias studies. Compare and contrast results; are the differences concerning?

Initial Notes:

- Specific setting as our motivation: customer service
- Keep your analysis on open source models
  - survey/analysis of outcomes from existing facial recognition models
- Algorithmic bias on specific attributes
  - How this bias can impact the application setting around which we motivated our study?
- Paper should focus on one big "ethical concern"
  - Algorithmic biases → racial biases
    - How could things deepen (racial) bias through their interpretations
- Paper:
  - First have some original experiments

Paper:

- First have some original experiments
  - Gender shades
  - Traffic intersections using FRT
- Introduce our experiments
- Tying out experiments to initial motivation (customer service)
- Primary application: customer service and security in commercial settings (retail stores, airports, stadiums, etc)

Introduction

- What is facial recognition technology (FRT)?
  - What it is and why its growing in popularity
- Customer service as the application setting – Explain how FRT is being used for:
  - Security and surveillance: loss prevention, shoplifting detection
  - Customer identification: loyalty programs, personalized service

Technical Background

- How FRT works
- Overview of OpenCV, DeepFace, and Face Recognition (3 open source models we evaluate)

Original Experiments(Literature review):

1) Gender Shades:

Algorithmic biases can manifest across many attributes, but they are particularly pronounced in the realm of race. This raises serious concerns about the ethical implications of FRT, and how consumers may be affected. One of the most influential studies to highlight these disparities is the *Gender Shades* Study by Buolamwini and Gebru (2018). Researchers from MIT and Stanford University conducted a study evaluating three leading gender classification systems – Microsoft, IBM, and Face++ – using the Pilot Parliaments Benchmark, a dataset intentionally balanced across gender and skin tone. In their results, they found that the three program's error rates had racial bias concerns: while light-skinned men never had an error rate higher than 0.8%, dark-skinned females had error rates as high as 34.7% in one case. This huge difference in accuracy not only points to systemic bias, but also the intersectionality of race and gender and resulting algorithmic unfairness. The researchers attributed these disparities to the unbalanced training data these algorithms had been trained on, which overwhelmingly included and favored light-skinned individuals. Thus when deploying FRT in high-stakes environments, such as customer service, it is essential to address the misidentification that could lead to unfair surveillance, exclusion, and more consequences.

Algorithmic bias:
- Concept: the systematic and repeatable errors in computer systems that create unfair outcomes, particularly in marginalized groups
- Examples:
    - Gender bias in hiring algorithms
    - Age bias in credit lending
    - Socioeconomic bias in insurance risk assessments
    - **Facial recognition:** mention biases in accuracy, lighting sensitivity, and demographic disparities
- Transition to racial bias
    - (explain) While algorithmic biases can occur along many attributes, **racial bias in facial recognition** has particularly sever implications for privacy, safety and civil rights

Application setting motivation
- Customer service & security – how racial bias in FRT affects real-world settings like:
    - Retail stores: higher rates of false positives for theft among certain demographics
    - Airports and stadiums: denial of entry or heightened surveillance based on facial analysis

Real-World implications

While the original experiments and our results illuminate the algorithmic biases and disparities in FRT, real-world incidents showcase the tangible consequences of these biases.

In law enforcement, facial recognition misidentifications have led to wrongful arrests of Black individuals. For instance, in January 2020, Robert Williams, a Black man from Detroit, was wrongfully arrested after an FRT system misidentified him as a shoplifting suspect[1]. Despite clear discrepancies between Williams and the suspect in surveillance footage, human oversight could not make up for the algorithm's error; Williams spent 30 hours in jail and posted a $1,000 bond before this grave mistake was corrected. The heavy reliance on technology in law enforcement often leads to uncritical acceptance of its results, even when those results are flawed. When facial recognition systems are racially biased, their deployment in high-stakes contexts like policing is especially dangerous. Algorithmic errors can lead to real-world harm, with human oversight often failing to catch these mistakes.

The risks associated with algorithmic biases in FRT don't necessarily stop at misidentifications; they can be compounded by corporate policies that use these technologies for selective enforcement. One example is Madison Square Garden's controversial use of FRT in 2022 to enforce an "exclusion list" that bans lawyers whose firms are engaged in litigation against the company[2]. Lawyers were denied entry from attending events at venues of MSG Entertainment, including Radio City Music Hall and MSG itself. Lawyer Kelly Conlon was denied entry to MSG after being identified through facial recognition software. Although she had no direct involvement in litigation against the company, just her association with her firm was enough to trigger the ban. This case exemplifies how with any algorithmic biases in FRT, corporate policies and exclusionary practices can magnify them. This sets the precedent that FRT is not unbiasedly used for security, identification, etc, but instead can also be leveraged as a selective and controlling technology. In customer service settings particularly, the consequences could be very detrimental. If businesses adopt similar policies of exclusion, algorithmic biases could be used to justify this discriminatory practice. The corporate leveraging of FRT highlights the need for increased oversight to prevent the deepening of biases in these models.

---

[1] https://mit-serc.pubpub.org/pub/bias-in-machine/release/1
[2] https://www.nytimes.com/2022/12/22/nyregion/madison-square-garden-facial-recognition.html