Failing at Face Value: The Effect of Biased Facial Recognition Technology on Racial Discrimination in Criminal Justice

Carina Wang*

American Heritage High School Palm Beach, Delray Beach 33484, Florida, United States of America

*Corresponding author: Carina Wang, carinawang04@gmail.com

Abstract: Recent years have seen a rise in the development of technological innovations and their implementation in various industries. Specifically, law enforcement agencies across the United States have partnered with technology companies to deploy facial recognition algorithms in the identification and prosecution of criminal suspects. Yet there is concern that law enforcement’s use of facial recognition algorithms based on biased mugshot data pools can lead to criminalizing innocent civilians. Prominent theories including intersection theory, instrumentalization theory, and Alvarado’s theory were analyzed to review arguments that justify concern. We find that intersection theory is supported by empirical evidence that women of color are put at the greatest disadvantage from technological bias; instrumentalization theory is supported by examples of both positive and negative implementations of facial recognition technology, and Alvarado’s theory further suggests the possible reinforcement of existing biases by these poor applications of technology.

Keywords: Racial discrimination; Facial recognition technology

Online publication: October 27, 2022

1. Introduction

Immigration and globalization increased ethnic diversity globally in recent decades, prompting questions about the equal application of justice to populations consisting of people from contrasting backgrounds and cultures. Conceptually, racism has played a role in the US justice system. In practice, the country’s criminal justice system has been criticized for taking away the rights of individuals based on racial or socioeconomic divides. In the digital century, the introduction of technology into policing and criminal justice adds a potential layer of discrimination in law enforcement and governance. Facial recognition use has steadily increased, and it is predicted to double by 2027 (see Figure 1 in the Appendix). Technological innovations are helpful overall, but they can also misconstrue the images and realities of low-income Americans and people of color. Various theories including Alvarado’s theory of ethnicity & racial stereotypes, intersection theory, and instrumentalization theory/critical theory of technology can shed critical light on the modern implementations of facial recognition technology in law enforcement.

2. Types of facial recognition technology

2.1. Feature analysis

From a neurological standpoint, humans recognize other people based on the spatial arrangement of facial features including the eyes, nose, mouth, and chin. This method of facial recognition relies on “the
extraction and measurement of facial features” [1]. Facial recognition algorithms determine the relation of these features via mathematical analysis of angles, distances, and areas of each feature in relation to each other. The eyes are considered the most obvious facial feature with the highest accuracy. Thus, most often, the algorithms first focus on finding the location of the two irises to map out the rest of the facial features.

**Neural Network**
Simulating the way neurons in human brains send signals, facial recognition software train neural networks with vast amounts of data to improve accuracy and efficiency in identification. By analyzing training examples from the photo database, neural networks are then used to identify the face in a new photo that was not part of the original training dataset [2]. For example, a neural network can be trained on a dataset consisting of thousands of pre-labeled images of each individual. The system would “find visual patterns in the images that consistently correlate with particular labels” [3]. After the training process, these networks are able to identify a person in a photo that the neural network has never seen before (see Figure 2 in the Appendix).

**2.2. Holistic Matching**
Rather than extracting features, the holistic face recognition method uses the whole face in the image, creating a vector composed of the gray values of all pixels in the face [4]. For example, the skin texture analysis method uses algorithms to measure lines, pores, and skin texture, developing facial models unique to each individual [5].

**Eigen Faces**
Features on the face (e.g. eyes, nose, and mouth) are pieced together to form an eigenface (see Figure 3 in the Appendix). Once the eigenface is formed for the person in the photo, it is compared to previously created eigenfaces for that same person. The eigenfaces are projected on top of each other, and the distances between the two are calculated. If the distance is within a certain margin, it is concluded that the two images resemble the same person. Similarly, eigenfaces can be used in the case that the “eigenface with the smallest Euclidian distance is the one the person resembles the most” [6].

**2.3. Hybrid**
A combination of Holistic Matching and Feature Analysis, the hybrid method utilizes both recognition processes (see Figure 4 in the Appendix). Features are extracted and analyzed. “The outputs from the individual components are then combined to give the final recognition output” [7].

**3. Benefits of facial recognition technology**
Globally, the implementation of advanced facial recognition systems has had positive effects in many cases across various industries.

**3.1. Healthcare**
Facial recognition has been tested in hospitals to streamline patient check-in, lessening the burden on hospital staff while also reducing human clerical errors. A hospital’s facial recognition system can verify a patient’s identity and insurance information to reduce wait time and serve as a security measure by monitoring individuals who enter and leave the hospital. Facial recognition algorithms have been proven successful in diagnosing rare genetic disorders based on identifying slight changes to facial characteristics, then using that information to generate a list of potential diagnoses along with their percent likelihood. For example, Face2Gene (an app that utilizes facial recognition to help doctors make medical diagnoses) has
been used on 250,000 patients and helped in identifying over 7,000 conditions \[8\].

3.2. Security and fraud
Airports across the United States currently utilize facial recognition technology to match passport photos with a database to verify the identity of travelers. In three specific examples, travelers have used fraudulent passports to enter the United States from Brazil, Ghana, and Cameroon. In all of these cases, facial recognition technology alerted U.S. Customs and Border Protection agents that the passport photos did not match the claimed identity. As of June 2020, nearly 300 individuals have been intercepted attempting to enter the U.S. under a fraudulent identity \[9\]. The U.S. Department of Homeland Security expects this technology to be used on 97% of travelers by 2023.

Banks have begun testing solutions that would have customers scan their faces to access ATM services to reduce the likelihood of fraud and hacking associated with the current PIN system \[8\].

3.3. Crime
Within 24 hours after holding a woman at knifepoint, police apprehended and arrested the rapist using facial recognition technology \[10\]. In another example, the New York City Police Department used facial recognition algorithms on security footage to identify a suspected subway terrorist. Within minutes, the system reported hundreds of potential matches. The detectives then sorted through the matches and identified the suspect by the end of the hour. Without facial recognition technology, it would have taken many hours or days to manually sort through videos and images based on witness descriptions of the suspect \[11\].

Spotlight is a tool that utilizes facial recognition algorithms to help find sex trafficking victims in online ads. Reports show that it has been used to help rescue 15,000 children and identify 17,000 traffickers in North America \[9\].

In Detroit, Michigan, a gunman killed three members of the LGBTQ+ community in a targeted attack. The local police department used facial recognition systems on videos from a gas station to identify and prosecute the suspect.

3.4. Finding missing people
In India, facial recognition systems have helped police find 2,930 missing children in the city \[12\].

4. Concerns: the evidence
4.1. Flawed mugshot databases used by facial recognition algorithms
Despite all of the aforementioned advantages, inaccurate surveillance technologies or poor use of those technologies by law enforcement can result in inaccurate rulings in court. Black Americans are more likely to be arrested and incarcerated for minor crimes than White Americans. Consequently, Black people are overrepresented in the mugshot data used by face recognition to make predictions. The result of this system is higher rates of false incarceration for people of color. Face recognition technologies across 189 algorithms are least accurate on women of color, with error rates up to 34% higher than for lighter-skinned males \[13\].

One out of four state and local law enforcement agencies have access to facial recognition technology. Clearview AI is a major startup company that has partnered with over 3,000 law enforcement agencies on all levels, with a photo database significantly larger than government databases (10 billion photos compared to FBI’s 640 million photo database). Similar companies include Vigilant Solutions, ODIN Intelligence, Ayonix, Cognitec, and iOmniscient \[14\]. Most facial recognition software use public images (e.g. mugshot data) to train the algorithms. In fact, Clearview AI specifically worked to acquire all U.S. mugshots from
the past 15 years to build their dataset for training facial recognition models [15]. However, this common way of training facial recognition algorithms is at the core of a biased criminal justice process. Images in the datasets used to train the algorithms are not proportionally representative of the diversity of the United States population. Black Americans are more likely to be arrested for minor crimes compared to White Americans. Thus, Black Americans are overrepresented in mugshot databases used by facial recognition software. This contributes to a cycle where “racist policing strategies lead to disproportionate arrests of Black people, who are then subject to further surveillance” [13].

4.2. Biased policing strategies
In fact, cameras with built-in facial recognition systems were frequently installed in majority-Black areas while rarely installed in predominantly White and Asian neighborhoods. These racially biased surveillance methods contribute to unbalanced mugshot databases, leading to inaccurate facial recognition conclusions. Not only are African Americans more likely to be surveilled, they are also more likely to be stopped by law enforcement and be subjected to facial recognition searches compared to people of other ethnicities [16]. In fact, stop-and-frisk data shows that Black and Latinx people have been pulled over or stopped on the street even if they had done nothing wrong. In 97% of these cases, there is no evidence of any crime, and the individuals are being stopped by law enforcement solely because of skin color [17]. Similarly, another study that analyzed 95 million stops by police found that Black people were more likely to be pulled over than White people, but the disparity decreases at night when it is harder for police to distinguish the race of the driver.

4.3. Challenges with acquiring images/videos to run facial recognition
Oftentimes, law enforcement agencies pull footage from security cameras from public places including stores and gas stations. However, many cameras’ default settings are not optimized to capture darker skin tones, resulting in lower-quality database images of Black Americans [13]. Low-quality images add another layer of uncertainty in facial recognition algorithms, displaying higher levels of inaccurate results.

4.4. Algorithm inaccuracies: statistics
The error rate of facial recognition algorithms rose 9.2% when the images were taken in public compared to high-quality images where the subject is not moving. Error rates increased when the subject was not looking at the camera [8].

4.5. Algorithm inaccuracies based on demographics: statistics
While it is usually incorrect to make a statement about all facial recognition software, the vast majority of facial recognition algorithms display demographic biases backed by empirical evidence. The National Institute of Standards and Technology evaluated most of the industry by analyzing 189 algorithms from 99 different developers. The study made a distinction between “one-to-one” matching that “confirms a photo matches a different photo of the same person in a database” and “one-to-many” matching that determines whether an individual in an image matches anyone in the database. It is acknowledged that the software can make false positives (believing that two different individuals are the same person) and false negatives (not matching two photos of the same person). In application, this means that inaccuracy can work in the direction of both false exoneration and false indictment. American Indians had the highest rate of false positives for one-to-one matching, whereas African American females had the highest rate of false positives for one-to-many matching, showing potential for consequences including false accusations [18]. Asian and African American individuals were up to 100 times more likely to be misidentified compared to White males. Women were more likely to be misidentified than men. Middle-aged White males had the highest
Comparing accuracy levels of different facial recognition softwares

A Gender Shades study analyzed algorithms produced by Microsoft, Face++, and IBM to compare rates of demographic bias. All companies performed the worst on darker females. IBM and Microsoft performed better on light-skinned males, whereas Face++ performed best on darker males. In general, all companies performed better on lighter individuals compared to darker individuals with a difference in error rates between 11.8% and 19.2%. IBM’s algorithm had a 34.3% higher error rate in identifying darker females compared to lighter males. 95.9% of the faces misgendered by Face++ algorithms were of female subjects.

Amazon’s algorithms worked successfully on images of light-skinned males, but misidentified the gender of darker-skinned women 30% of the time. Amazon is marketing their Rekognition system to law enforcement. However, this system incorrectly matched 28 photos of members of Congress. While 20% of Congress members are people of color, they constituted 40% of false matches from the system.

With these high error rates, facial recognition entrenches systemic racism by supercharging the government’s ability to surveil and target marginalized groups, impacting core rights and taking away necessities. Police and other government agencies use these systems to intimidate activists, target immigrants, wrongfully accuse people of crimes, and impede access to needed public resources such as unemployment relief and housing. This use of inaccurate technology leads to self-censorship among communities of color out of fear of retribution, fueling the widespread issue of drowning out the voices of marginalized populations.

4.6. General incarceration rates statistics

38.4% of the inmate population are black people and 93.1% are males. The FBI’s Uniform Crime Reporting Program found that black people were overrepresented among persons arrested for nonfatal violent crimes (33%) and for serious nonfatal violent crimes (36%) relative to their representation in the U.S. population (13%).

4.7. Lack of transparency

One of the greatest concerns regarding facial recognition, besides its demographic biases, is law enforcement’s lack of transparency. The Government Accountability Office criticized the U.S. Customs and Border Protection for “lackluster accuracy audits, poor signage notifying the public the technology is being used, and little information offered to the public on how its systems worked.” From the public’s perspective, a study found that 71.1% of survey respondents reported that they were “very” or “somewhat” concerned about their privacy in regards to facial recognition on video images.

5. Theories

5.1. Alvarado’s theory

Inaccurate technology affects public perception of people of color. A subsect of Alvarado’s theory of ethnicity postulates that stereotypes displayed throughout the media represent ethnic minorities as “dangerous to society,” causing people to blame them for social issues. News outlets often group individuals together based on their race, or under the title of “immigrants”, perpetuating xenophobia. This lack of personalization makes it easier to blame them [ethnic groups] for a range of social problems.
5.2. Intersection theory
Intersection theory, developed in 1989 by Kimberlé Crenshaw, asserts that because racial biases are shaped by other factors (including gender), these attributes must be examined together. The theory claims that racial prejudice can be dissected into many layers of disadvantage. The theory states that: “For example, if we want to understand prejudice, we must understand that the prejudice focused on a white woman because of her gender is very different from the layered prejudice focused on a poor Asian woman, who is affected by stereotypes related to being poor, being a woman, and her ethnic status” [25].

5.3. Instrumentalization theory
Instrumentalization theory (critical theory of technology), proposes that “technology must be analyzed at two levels: one is the level of our original functional relation to reality and the second is the level of design and implementation” [26]. The first level analyzes the technology strictly based on useful properties. The second level contextualizes these properties in relation to existing technologies and societal uses. In the early stages of facial recognition technology (1964), researchers wanted to answer the basic question of whether programming computers are capable of recognizing human faces. In the early 2000s, facial recognition vendor tests began as government agencies evaluated this relatively new technology that was now commercially available.

6. Analyzing theories: evidence
6.1. Alvarado’s theory
Society quite commonly conflates social problems like crime or violence or disease with black and brown people [26]. Oftentimes as a result of the media spreading information about prosecutions that may result from inaccurate technology, people of color are seen as more violent/dangerous in communities. For example, in a study, each of 950+ participants was shown a series of color photographs of white and black male faces of individuals who were all of equal height and weight. Results showed that participants judged the black men to be larger, stronger and more muscular than the white men, even though they were actually the same size. Participants also believed that the black men were more capable of causing harm in a hypothetical altercation and, troublingly, that police would be more justified in using force to subdue them, even if the men were unarmed [27]. On the flip side, individuals who understand the shortfalls of technology misused can be turned away from using technology in the justice system altogether or be disincentivized to use technological innovations out of fear of perpetuating current problems. This raises the dilemma of finding a balance between our growing reliance on technological innovations and ensuring that these services benefit all members of society equally.

6.2. Intersection theory
Examining the validity of intersection theory from the lens of facial recognition in criminal justice, it is evident that race and gender can in fact be layered to result in higher levels of discrimination. Besides, researchers found that facial recognition technology falsely identified Black and Asian faces 10 to 100 times more often than they did white faces. The technologies also falsely identified women more than they did men - making Black women particularly vulnerable to algorithmic bias [28].

6.3. Instrumentalization theory
Facial recognition is not supposed to be used on its own to establish probable cause for an arrest [29]. Yet research shows that many law enforcement agencies have relied almost exclusively on facial recognition systems to make an arrest. In fact, the NYPD noted that it has turned to facial recognition in more than 22,000 cases in the last three years [30]. When misused, these technological innovations end up harming
innocent civilians and exposing vulnerable populations to inapt surveillance systems, while at the same
time building mistrust of the justice system among minority populations. This problem is exacerbated by
the fact that facial recognition software varies across law enforcement agencies, and many agencies have
lower standards of accuracy than the FBI or do not conduct accuracy tests at all. Studies have attempted to
quantify exactly how many arrests/convictions have resulted from inaccurate uses of facial recognition. If
we assume that misidentifications happened in only one out of a thousand searches, or .1% or the time, this
would mean that, in Florida alone, eight people are implicated in a crime they did not commit each month
[31].

7. Appeals process
Defendants have the constitutional right to “probe” the accuracy of the facial recognition systems used to
accuse them before being convicted [32]. The Supreme Court ruled in Brady v. Maryland that prosecutors
must provide defendants and jurors access to “potentially exculpatory evidence,” which would include
information regarding the workings and results of facial recognition algorithms [33].

However, few defense attorneys actually challenge the accuracy of the system, and not all courts agree
that defendants have this right. A Florida appellate court ruled in 2019 that a convicted individual did not
have the right to view the results of the facial recognition test that led to his arrest, even though the algorithm
only had a one-star confidence level (extremely low) for producing the correct match.

8. Conclusion
Law enforcement agencies on the local, state, and federal level have partnered with tech companies to
deploy facial recognition technology in their local communities. The technologies use different techniques,
including “feature analysis,” “holistic matching,” and ‘hybrid methods.” To various degrees, all these
approaches, when used in law enforcement, can reinforce racial stereotypes across the country. This is
because the databases used to train these facial recognition software are biased in their over-representation
of minorities. This results in inaccurate rulings in court and wrongful arrests, perpetuating existing racial
stereotypes. Theories of intersectionality suggest that women of color are especially vulnerable.
Instrumentalization theory proposes that the intrinsic effects and the technological capabilities of facial
recognition must be separated from the biases induced by the specific ways in which these technologies are
used by the government. Finally, Alvarado’s theory of ethnicity shows the impact of biased technology on
skewing incarceration demographics, with secondary implications on public perception and stereotyping of
minorities. Facial recognition and other technologies used in criminal justice can result in biases against
people of color and perpetuate racist ideologies. As society becomes more technologically advanced, it is
increasingly important to monitor and ensure the accurate uses of these technologies, to avoid introducing
new societal divides along racial lines.

Disclosure statement
The author declares no conflict of interest.

References
[2] IBM Cloud Education, 2020, Neural Networks, IBM Cloud Learn Hub, viewed July 26, 2022,


[31] Garvie C, 2020, The Untold Number of People Implicated in Crimes They Didn’t Commit Because of Face Recognition, ACLU, June 20, 2020, viewed 26 July 2022, https://www.aclu.org/news/privacy-
technology/the-untold-number-of-people-implicated-in-crimes-they-didnt-commit-because-of-face-
recognition#:\=text=We%20have%20no%20idea%20how,\=every%20month%E2%80%94in%20Florida%20alone

Them a Chance to Challenge the Tech, ACLU, March 12, 2019, viewed August 29, 2022,
https://www.aclu.org/blog/privacy-technology/surveillance-technologies/florida-using-facial-
recognition-convict-people.

[33] Aaron Mak, 2019, Facing Facts: A Case in Florida Demonstrates the Problems with Using Facial
Recognition to Identify Suspects in Low-Stakes Crimes, Slate, Jan 25, 2019, viewed August 29, 2022,

[34] FT Reporters, 2021, In Charts: Facial Recognition Technology? And How Much Do We Trust It?,
Financial Times, May 16, 2021, viewed August 29, 2022, https://www.ft.com/content/f6a9548a-a235-
414e-b5e5-3e262e386722

Environmental Science and Pollution Research, 27: 9589–9603. https://doi.org/10.1007/s11356-019-
07574-w

Independent component with Score-based Fusion Process. Computer Vision and Pattern Recognition,

Publisher’s note

Bio-Byword Scientific Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.
Appendix

North America’s facial recognition market expected to double in size by 2027
Technology market size ($m)

Figure 1. Growing use of facial recognition \[34\]

[Image of a bar chart showing the growth of facial recognition technology market size from 2016 to 2027.]

Diagram of Artificial Neural Networks \[35\]

[Diagram of an artificial neural network with input, hidden, and output layers.]

Flowchart of the algorithm of the Eigenfaces Method \[6\]

[Flowchart showing the steps of the Eigenfaces Method: start, read training set of N x N images, resize image dimensions to N x 1, select training set of N x M dimensions, find average face, subtract from the faces in the training set, create matrix A, calculate covariance matrix, find the minimum Euclidian distance or image unrecognizable, output: image with the minimum Euclidian distance or image unrecognizable.]
Figure 4. Hybrid Facial Recognition \cite{36}