Vol 7 | Issue 1 | Pages 100-105

Trends in Artificial Intelligence
ISSN: 2643-6000

Research Article

DOI: 10.36959/643/309

The Camera Doesn't Lie… or Does it? Police Use of Facial Recognition Technology and Racial Bias

*Seppy Pour**

School of Law, University of Sydney, Sydney Australia

Abstract

Law enforcement has transformed drastically by advances in technology. Police, like the rest of society, have been confronted by an array of new challenges and opportunities in an era of rapid technological change. These advances have brought about the potential for significant improvements to operational efficiency and outcomes, particularly during a time when technological advances have similarly given rise to greater demand for police accountability.

This technological boom resulted in the advent of artificial intelligence-led policing. Law enforcement bodies around the world have adopted facial recognition capabilities powered by artificial intelligence and contend that facial recognition technology is an effective tool in preventing, disrupting, investigating, and responding to crime [1]. As the practice has grown, so have criticisms of its use and policing outcomes. Among the many objections is its high rate of inaccuracy, particularly when used in relation to racial minorities.

In this paper, the link between facial recognition technology and racial bias will be reviewed through an analysis of existing research. The review will consider a recent case study in which the use of facial recognition technology by law enforcement was banned by jurisdiction officials. It will also analyse academic research which has attempted to highlight the propensity for racial bias and the circumstances in which such outcomes have arisen. It is intended that this paper will provide a historical account of the research into the intersection of facial recognition technology and racial bias and demonstrate that there is concern for the use of facial recognition technology in the absence of adequate governance mechanisms.

Keywords

Artificial intelligence, Policing, Facial recognition, Crime, Law enforcement, Justice, Race

Context

Facial recognition has garnered significant attention in recent years. As technological capabilities improve, the use of facial recognition as a form of biometrics has also expanded proportionately; applications include password verification, law enforcement investigation, social media functionality, and in retail and advertising [2]. Most uses are generally uncontroversial-self-excluded problem gamblers for example are currently refused entry from gaming venues with facial recognition used, with consent, to identify them [3].

It is necessary for the purposes of this paper to define its key terms. 'Facial analysis' broadly describes any diagnostic method capable of structural analysis of a person's face [4]. 'Facial recognition' is any technology which utilises facial analysis capabilities to identify or verify the identity of a person, typically from a still image or video footage. 'Artificial intelligence' is the underlying algorithmic tool which enables facial recognition technology [5].

The concepts of facial analysis and facial recognition are not novel. Facial recognition technology has existed since the mid-20th century [6]. The original approach designed by Woodrow Bledsoe involved manually calculating the distance between the facial features in a suspect photo and loading these dimensions into a computer for pattern matching against a database of 'mugshots'. As early as 1966, Bledsoe noted the difficulties of such an approach to facial recognition:

This recognition problem is made difficult by the great variability in head rotation and tilt, lighting intensity and angle, facial expression, aging, etc. Some other attempts at facial recognition by machine [sic] have allowed for little or no variability in these quantities… In particular, the correlation is

***Corresponding author:** Seppy Pour, School of Law, University of Sydney, Australia

Accepted: February 01, 2024

Published online: February 03, 2024

Citation: Pour S (2024) The Camera Doesn't Lie… or Does it? Police Use of Facial Recognition Technology and Racial Bias. Trends Artif Intell 7(1):100-105

very low between two pictures of the same person with two different head rotations.

As we will see in the following parts of this paper, the difficulties identified by Bledsoe remain relevant today.

A City-Wide Ban

In May 2019, the city of San Francisco, United States banned by city ordinance the use of non-pre-approved facial recognition technology by law enforcement agencies [7]. The move, which was decided almost unanimously by the city's Board of Supervisors, enshrined San Francisco as the first major city in the US to prohibit the tactic. Since then, other jurisdictions have gone on to ban or limit the use of facial recognition for various uses [8].

The rationale for the ban can be derived by examining the city ordinance [9]. As one of the earliest examples of an official ban on facial recognition, this rationale will underpin the discussion in this paper. As seen below, each relevant 'general finding' under the ordinance will be used as a discussion point.

The Propensity for Facial Recognition Technology to Endanger Civil Rights and Civil Liberties

The ordinance, dubbed the 'Stop Secret Surveillance' ordinance, listed in its general findings that:

(a) The propensity for facial recognition technology to endanger civil rights and civil liberties substantially outweighs its purported benefits, and the technology will exacerbate racial injustice and threaten our ability to live free of continuous government monitoring.

This finding is not without validity. A seminal analysis by Phillips, et al. found that facial recognition technologies developed in France, Germany and United States correctly identified Caucasian persons more accurately than East Asian persons; those developed in China, Japan and South Korea identified East Asian persons more readily [10]. The authors therefore deduced that the efficacy of a facial recognition algorithm was influenced by the origin of its developers and the demographic makeup of the database on which it operates.

Numerous studies have since bolstered these findings. Klare, et al.'s analysis found that facial recognition algorithms demonstrated lower accuracy on female cohorts (compared to male cohorts), black persons (compared to Caucasian and Hispanic cohorts), and 18-30 year-olds (compared to 30-50 year-olds and 50-70 year-olds) [11]. This discrepancy was consistently observed across six facial recognition technologies, in both those trainable vs untrainable and those commercially available vs. non-commercially available. Facial recognition tests on identical twins have also exhibited lower rates of accuracy when comparing footage taken during dissimilar periods (i.e. two still images or videos taken a long time apart), subjects not maintaining a neutral expression, and in differing light conditions [12].

Similar shortcomings have been identified through work undertaken the National Institute of Standards and Technology (NIST) which conducts voluntary tests of commercially available facial recognition tools every four years [13]. The NIST regularly hosts competitions to compare the efficacy of facial recognition algorithms against human facial reviewers [14]. Phillips and O'Toole's systematic analysis of results across these competitions found that while facial recognition technology was superior to human reviewers for matching frontal faces in still images, human reviewers fared far greater when analysing video footage or complex still images [15]. Academics have therefore surmised that the most accurate facial recognition outcomes are obtained when human reviewers and facial recognition technology is used in tandem [16].

Robust Transparency, Oversight, and Accountability Measures

The ordinance further endorses the view that:

(b) Legally enforceable safeguards, including robust transparency, oversight, and accountability measures, must be in place to protect civil rights and civil liberties before any surveillance technology is deployed [17].

On 23 February 2018, Buolamwini and Gebru presented at the Conference on Fairness, Accountability and Transparency a paper that would go on to become a landmark study on the effectiveness of facial recognition technology. The paper, entitled Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification (Gender Shades), represented the first intersectional analysis of three commercially available facial analysis technologies [18]. The study observed stark differences in the technologies' ability to recognise lighter-toned males, lighter-toned females, darker-toned males, and darker-toned females (the study groups). While there were negligible error rates in relation to lighter-toned males and lighter-toned females across all three algorithms, error rates were observed in relation to darker-toned males and significantly more so in relation to darker-toned females. Accordingly, the authors concluded that all three algorithms performed better on male faces than female faces and on lighter faces than darker faces. Buolamwini and Gebru additionally note that since default camera settings are optimised for lighter skin, accuracy rates are further skewed against darker-toned persons.

The NIST's 2019 quadrennial analysis of facial recognition algorithms confirmed the Gender Shades findings. The authors noted the "recent expansion in the availability, capability, and use of face recognition" and that this expansion had been "accompanied by assertions that demographic dependencies could lead to accuracy variation and potential bias" [19]. The findings included a higher false match rate in relation to faces from East African, West African, and Caribbean person groups compared to those from East Asian or Eastern Europe groups; false match rates were higher for the East Asian person group than the Eastern European group (however this divergence was lower for algorithms developed in China). False matches were consistently more prevalent for women than men

across all algorithms regardless of origin, although the effect was less profound than that based on skin tone. Additionally, false positives were higher for both genders in the 12-20 year-old and 65-and-over age groups regardless of skin tone.

As a result of the academic publication of *Gender Shades* (as well as a website infographing its findings) [20], the three commercial facial recognition algorithms tested by Buolamwini and Gebru were notified of the embedded bias of their respective products. To analyse the impact of publicly demonstrating such shortcomings, Buolamwini and Raji retested the three algorithms roughly seven months following the publication of *Gender Shades* [21]. All three proprietors (the Target Corporations), namely Microsoft, Megvii and IBM, demonstrated an improvement in the accuracy of their algorithms: all three showed less accuracy difference between male/female and lighter tone/darker tone study groups. Unsurprisingly, all three still demonstrated their highest error rates in relation to the darker-toned female subject group. The lowest error rates were still seen in relation to lighter-toned male subject group.

In addition to the existing three algorithms which were retested, Buolamwini and Gebru added to this audit two previously untested facial recognition algorithms: Amazon's 'Rekognition' and Kairos (the Non-Target Corporations). The initial test on the Non-Target Corporations showed that their error rates were considerably higher than the Target Corporations at the time of the analysis but were comparable to the error rates exhibited by the Target Corporations in the Gender Shades study. The Non-Target Corporations mirrored the findings of the Target Corporations - Rekognition and Kairos performed better on male faces than female faces and on lighter-toned faces than darker-toned faces. The Non-Target Corporations showed the worst performance for darker-toned females sub-group.

As concluded by Buolamwini and Raji, it can be deduced that the work undertaken through *Gender Shades* has had the effect of encouraging corporations into prioritising the reduction of racial and gender bias in their products. Moreover, the study's contribution to public discourse and corresponding media attention would presumably have had the added effect of encouraging other product providers to scrutinise their algorithms despite not being directly audited by the authors. Notwithstanding the positive impacts of Gender Shades, public scrutiny of error-prone algorithms is not always enough to motivate improvement, as discussed in the following section.

Intimidate and Oppress Certain Communities and Groups

The Stop Secret Surveillance ordinance unsurprisingly includes in its general findings section:

(c) While surveillance technology may threaten the privacy of all of us, surveillance efforts have historically been used to intimidate and oppress certain communities and groups more than others, including those that are defined by a common race, ethnicity, religion, national origin, income level, sexual orientation, or political perspective (emphasis added) [17].

The potential for false matches discussed in the previous section is particularly unsettling given the highly documented disparities in policing strategies and outcomes.

The link between law enforcement and racial bias is well documented. Racial bias within law enforcement has been highlighted in previous research, particularly among white male police officers [22]. Analysis of common police practices have also revealed disproportionate effects of policing on racial minorities. Disproportionality has been observed at various contact points between police and people of colour, including street stops [23], drug enforcement strategies [24], vehicle pullovers [25], and the quantity and attitudes of officers deployed in communities of colour. In the Australian context, for example, Indigenous Australians are roughly 20 times more likely to be arrested than their non-Indigenous counterparts. This over-representation has been attributed to historical police practices and crime strategies, including when police make the decision to make an arrest and the number of officers deployed to locations with high Indigeneity. Police have therefore historically failed to use their discretion appropriately to reduce Indigenous representation within the criminal justice system.

Robert Julian-Borchak Williams

Wrongful arrests of people of colour based on inaccurate facial recognition analysis are not unheard of. One of the more high-profile incidents involved Robert Julian-Borchak Williams of Michigan, US. Williams was apprehended in January 2020 when Detroit Police, then investigating the theft of five watches worth approximately $US3,800, received a 'match' on an image of a suspect taken from a security video relating to the case. Upon receiving the facial analysis results, two officers attended Williams' home and placed him under arrest. When Williams' wife asked where he was being taken, one of the officers responded, "Google it". At the station, Williams was photographed, fingerprinted, and had his DNA swabbed. He was detained overnight. He was questioned - all based on nothing but an alleged match of his face with that of a suspect in a video. Upon being shown still images from the security footage, Williams strongly denied that he was the man in the picture. He reportedly picked up the still image, held it next to his face and said to the officers, "You think all black men look alike?". Little did Williams know, that in this case, it wasn't the officers who made the identification.

Stories like Williams' are not unique. And to avoid doubt, it should be mentioned that Williams' case and many others like it have occurred in the 2020s, following the publication of Gender Shades and Actionable Audits. These cases all have something else in common: Lack of transparency and oversight. And while individual police departments may choose to implement department-level policies or procedures on the use of facial recognition technologies, there is little regulation at present governing the use of facial recognition technology at state or federal levels in most jurisdictions. The haste with which facial recognition technology has technologically advanced, and subsequently adopted by law enforcement, means that the legal system has failed to keep up.

Academics and human rights activists are not the only ones advocating for better governance of facial recognition technology. Microsoft, one of the early developers of commercially available facial recognition software, has publicly stated that the technology issues "heighten responsibility for tech companies that create these products" and accordingly advocated for "thoughtful government regulation and for the development of norms around acceptable uses" [26]. Microsoft has subsequently gone on to acknowledge the racial bias in facial recognition technology and advocated for improved transparency and the use of human review of facial recognition analysis. Google, whose facial recognition software has previously mixed up images of black people with gorillas, has similarly vowed not to sell its facial recognition technology product until the underlying technology is adequately regulated.

In its address to the public, Microsoft called on the government to regulate the proper use of facial recognition technology, rather than "to ask unelected companies to regulate [the governance space]" because "even if one or several tech companies alter their practices, problems will remain if others do not". The company was remarkably foresighted in this respect. Following the publication of Buolamwini and Raji's *Gender Shades*, a distinct reduction in error rates in the Target Corporation's algorithm error rates could be seen. Unlike the Target Corporations, however, Amazon, one of the two previously untested companies included in Actionable Audits, has had a significantly different response to the claims about the accuracy of its service offering.

On 26 January 2019, Amazon published its response to Actionable Audits (the Response). The response accused Actionable Audits of being misleading and drawing false conclusions. The Response effectively obfuscated Actionable Audits' findings against Rekognition in three ways. Firstly, it argued that 'facial analysis' and 'facial recognition' are "completely different in terms of their underlying technology", suggesting that the entire premise of Actionable Audits was misguided. Secondly, it claimed that Buolamwini and Raji had tested an outdated version of Rekognition that lacked a number of improvements introduced by Amazon in November 2018. Thirdly, the Response claimed that Amazon's own "extensive" internal testing had produced "no significant difference in accuracy with respect to gender [or race] classification". The author claimed that Amazon had undertaken a test of facial recognition involving a "Megaface dataset of 1 million images… and found exactly zero false positive matches".

While the Amazon response concludes, "we are very interested in working with academics in establishing a series of standardized tests for facial analysis and facial recognition", Buolamwini has publicly stated that Amazon's response to her work has been "one of denial, deflection, and delay". For Buolamwini, it is imperative for facial recognition technology to be independently tested by external parties. She states that she "time and time again [finds] that the internal accuracy rates if reported by companies seem to be at odds with external accuracy rates reported by independent third parties". In other words, such impressive results can be wilfully obtained by using low quality baseline benchmarks. Accordingly, developers of facial recognition technology should be compelled to undergo external audit for evaluation in "real-world cases with the results submitted for public scrutiny".

Meaningful Public Input

But in a world where commercial interests largely dictate corporate decision making, should accountability lay with software developers? Arguably so, but as Microsoft suggested in its public address, placing the onus of responsibility on governments would be far more effective in meeting public goals of a healthier dynamic for consumers and developers alike. The Stop Secret Surveillance ordinance addresses community consultation as follows:

(d) Whenever possible, decisions regarding if and how surveillance technologies should be funded, acquired, or used, and whether data from such technologies should be shared, should be made only after meaningful public input has been solicited and given significant weight.

It must however be understood that this adequate governance is twofold. Firstly, the design and implementation of overseeing legislation must be undertaken by the legislature. In the Australian context, this would be the responsibility of both the state and territory governments (to the extent that facial recognition use relates to policing, corrections, and any other state jurisdictional matters) and the federal government. Regulations would require extensive consultation with industry groups, law enforcement officials, and civil rights groups to develop a well-balanced governance framework with adequate safeguards. The Australian Human Rights Commission has twice called for a halt to the government's use of facial recognition technology until adequate protections are put in place.

Secondly, day-to-day accountability of facial recognition must become a reality before it can be validly used by law enforcement, particularly in the legal context, because accountability goes beyond simply auditing a facial recognition algorithm. This is because the use of facial recognition by law enforcement agencies is also largely opaque and unaccounted for. Consider how the use of facial recognition would occur in practice: an officer may 'identify' person as a suspect by virtue of an algorithm presenting the person as a 'match'. The officer makes an arrest based on that intelligence and, in the absence of any indisputable evidence to the contrary, that individual becomes a suspect and investigatory resources are largely put towards proving that individual's culpability. Because police tactics are typically not submitted as evidence in criminal trials, the technique is not subject to the same level of public or judicial scrutiny as other forms of intelligence and evidence gathering. The practice is therefore largely undocumented and untested in the public domain.

If one were to apply this impediment to the Robert Julian-Borchak Williams case study, we realise that Williams could not have reasonably known that his identification was made by a computer algorithm. In fact, according to the American Civil Liberties Union (ACLU) who represented Williams in his criminal case, Williams was only made aware of the use of

facial recognition thanks to a sarcastic comment from one of the investigating officers that "the computer must have gotten it wrong". Williams also had an alibi at the time of the alleged offence, but this was largely ignored. The ACLU raised procedural fairness issues concerning the inability to challenge facial recognition results, both through a lack of knowledge on how particular algorithms work and because most defendants do not even know that a match was the basis for police suspicion. While facial recognition results are sometimes likened to the results of fingerprinting of DNA analysis, facial recognition technology is largely unexamined thus far and can be undertaken in secret.

Further raising roadblocks in practice is police departments resisting attempts by civil rights and privacy advocates to bring to light to how and when facial recognition technology is used. In one high-profile example, a black man in Florida was convicted of supplying drugs to undercover officers after facial recognition was used to identify him [27]. Detectives in this case received multiple potential matches from a photo of the suspect, but the defendant was never given access to photos of the other 'matches'. Despite multiple appeals arguing that this undermined the defendant's right to receive evidence favourable to the accused, the photos have not been turned over. In New York, the New York Police Department (NYPD) went to great legal lengths to claw back documentation relating to its use of facial recognition technology that it had accidentally disclosed to the Georgetown Center [sic] on Privacy and Technology due to an administrative error. The NYPD was successful in its legal action but was criticised for "inconsistently and selectively disclosing information" and further described as being "mystifying [in] what [it] was trying to keep from the public".

Conclusion

The Stop Secret Surveillance ordinance imposed in the San Francisco area represents a progressive movement against the many unknowns of facial recognition technology. Analysis of the ordinance demonstrates a clear concern for the lack of checks and balances of the use of facial recognition technology, its demonstrated tendency to act with racial bias, its potential for abuse, and the lack of public consultation prior to its adoption. Despite the many unknowns of facial recognition, there is adequate evidence to suggest that these are valid concerns.

This paper reviewed the most influential studies to date on the accuracy rates of facial recognition technologies. These studies consistently demonstrated a propensity for algorithms to mirror the biases of the datasets on which they are trained, including any in-built racial or gender biases. Beyond the academic analyses, multiple examples of misidentifications of black citizens in the US, with related commentary from human rights and civil liberties groups, suggests that these concerns are translating into real world injustices.

Facial analysis and facial recognition technology has existed since the 1960s, but its current form is a relatively new concept. Its design, capabilities, accuracy, versatility, and uses are rapidly evolving. With this evolution has come many questions and concerns regarding the risks of this type of technology. Notwithstanding the tremendous law enforcement value it provides, these questions are valid and must be addressed before it should be used for any purpose which could result in the deprivation of an individual's liberty.

References

1. Facial Recognition. NSW Police Force, (Web Page).
2. Parmar DN, Mehta BB (2014) Face recognition methods & applications. International Journal of Computer Applications in Technology 4: 84-86.
3. Kidd J (2022) NSW Pubs and clubs to install facial recognition technology to help stop self-excluded gamblers, ABC News.
4. Buolamwini J (2019) Response: Racial and gender bias in amazon rekognition - commercial AI system for analyzing faces. Hackernoon.
5. Ford TW (2021) It's time to address facial recognition, the most troubling law enforcement AI tool, Bulletin of the Atomic Scientists.
6. Woodrow Bledsoe Originates of Automated Facial Recognition, History of Information (Web Page).
7. Fausset KCR, Kovaleski SF (2019) San Francisco Bans Facial Recognition Technology, The New York Times.
8. Hamilton B, Berry K (2021) Portland becomes first jurisdiction to ban certain uses of facial recognition by private businesses, Davis Wright Tremaine LLP.
9. (2019) Acquisition of Surveillance Technology.
10. Phillips PJ, Jiang F, Narvekar A, et al. (2011) An other-race effect for face recognition algorithms. ACM Transactions on Applied Perception 8: 1-9.
11. Klare B, Member IEEE, Burge MJ, et al. (2012) Face recognition performance: Role of demographic information. IEEE Transactions on Information Forensics and Security 7: 1789-1797.
12. Phillips PJ, Flynn PJ, Bowyer KW, et al. (2011) Distinguishing identical twins by face recognition (Conference Paper, International Conference on Automatic Face and Gesture Recognition, 21 March 2011).
13. Face Recognition Vendor Test (FRVT) (2020) National Institute of Standards and Technologies.
14. National Institute of Standards and Technologies (2019) Face Challenges.
15. Phillips PJ, O'Toole AJ (2014) Comparison of human and computer performance across face recognition experiments. Image and Vision Computing 32: 74-81.
16. Phillips PJ, Yates AN, Hu Y, et al. (2018) Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. Psychological and Cognitive Sciences 115: 6171-6176.
17. Acquisition of Surveillance Technology (n 20).
18. Buolamwini J, Gebru T (2018) Gender shades: Intersectional accuracy disparities in commercial gender classification. (Conference Paper, Conference on Fairness, Accountability, and Transparency, 23 February 2018).
19. Grother P, Ngan M, Hanaoka K (2019) Face Recognition Vendor

Test (FRVT) Part 3: Demographic Effects. National Institute of Standards and Technology 1.

20. Buolamwini J, et al. Gender Shades, MIT Media Lab (Web Page).
21. Buolamwini J, Raji ID (2019) Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products (Conference Paper, AAAI/ACM Conference on AI, Ethics, and Society, January 2019).
22. Roscigno VJ, Preito-Hodge K (2021) Racist Cops, Vested "Blue" Interests, or Both? Evidence from Four Decades of the General Social Survey. Socius: Sociological Research for a Dynamic World 7: 1.
23. Fagan J, Braga AA, Brunson RK, et al. (2016) Stops and stares: Street Stops, surveillance, and race in the new policing. Fordham Urban Law Journal 43: 539.
24. Mitchell O, Caudy M (2015) Examining racial disparities in drug arrests. Justice Quarterly 32: 288-313.
25. Lundman RJ, Kaufman RL (2003) Driving while black: Effects of race, ethnicity, and gender on citizen self-reports of traffic stops and police actions. Criminology 41: 195.
26. Smith B (2018) Facial recognition technology: The need for public regulation and corporate responsibility. Microsoft on the Issues.
27. Conarck B (2019) Court denies facial recognition evidence appeal. The Florida Times-Union.

DOI: 10.36959/643/309