About-Face: The Two Sides of Facial Recognition

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

Samantha Verdi

Spring, 2025

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Bryn E. Seabrook, Department of Engineering and Society

About-Face: The Two Sides of Facial Recognition

Introduction

A classic faux pas happens whenever someone forgets someone else's name; one could consider such an incident embarrassing. But what if someone's name was forgotten, and then that person was refused access to a service, or was framed for a crime they did not commit? Or worse: what if a misidentified person gets in trouble, but another gets off without consequence?

This is a representation of the current state of facial recognition technology; the way it is being made, used, and trained is inherently inequitable. Multiple studies have been published on its biased results, especially on the basis of race and gender (Buolamwini, 2017). It is important to study because this technology has become ubiquitous in the realm of security, from personal use, such as unlocking your cell phone (Apple Support, 2024), to broadband societal use, such as identifying criminals in public (Nabil et. al, 2022). Additionally, it serves as an example of SCOT (the Social Construction of Technology), the idea that technology is constantly evolving with input from society "in a 'multidirectional' model" (Pinch and Bjiker, 2012, p. 22). This paper answers the question: in what ways do the functionality and implementation of facial recognition technology perpetuate bias in society? Many secondary sources will be analyzed to explain how facial recognition technology works and how its development is directly linked with societal factors, including privacy, safety, and equality.

Research Methods

The research for this paper is conducted by documentary research and literary review of publicly available materials, such as books, technical reports, and journal articles. Key terms for the research include facial recognition technology, surveillance, privacy, bias, algorithms, data collection, civil and consumer rights, and artificial intelligence. This paper presents relevant background information about the technology and its implementation, as well as discussion about the studies that have been conducted regarding its biases and privacy concerns by discourse analysis. Finally, the paper explores potential solutions, with examples of published reports about those solutions, and their effectiveness.

Background: Facial Recognition Beginnings

Facial recognition technology, or FRT for short, was invented in the United States in the 1960s by Woodrow Wilson Bledsoe and his team at Panoramic Research Incorporated, his own start-up company (Raviv, 2020). They worked closely with the American government to build a machine capable of identifying people based on photos of their faces alone. While the first iterations were not perfect—it failed to identify people who were smiling, for example—it would eventually go on to evolve into the ubiquitous technology we know today, thanks to technological advancement and government funding. Development was funded by both the Defense Advanced Research Project Agency (DARPA) and the National Institute of Standards and Technology (NIST) for companies to commercialize FRT (Libby & Ehrenfeld, 2021). Nowadays, FRT is seen almost everywhere, especially since its introduction to Windows and Android devices in 2015, soon followed by Apple as well in 2017 (Klosowski (New York Times), 2020).

Facial recognition works by using a camera to scan or take a picture of someone's face, then converting that person's face into data based on its detected structure (Wright, 2022, p. 3). Then, that structure is compared to the computer's database of preexisting structures to see if it finds a match. According to Wright, a spokesperson for the federal government, there are two main ways the technology is used: for verification or identification. Verification is comparing the facial data to a pre-registered face that the computer is expecting to see. This is the kind that most people will experience firsthand from day-to-day, if they use their face to unlock their iPhone, for example. Identification, on the other hand, is when the facial data is compared to an entire database of images, to see if there are any matches. This is the implementation that is used the most by the federal government, and it is also where the majority of ethical dilemmas and controversies with the technology arise.

FRT was first introduced to the public eye after being used at the 2001 Super Bowl XXXV for security purposes. Klosowski (2020) reported that there was a public outcry, with "critics [calling] it a violation of Fourth Amendment rights against unreasonable search and seizure." However, despite the privacy concerns of the time, the technology itself had not developed enough to be fully efficient. In an article for the *IEEE Technology and Science Magazine*, professor Kevin Bowyer opened questions regarding the protection of privacy with this technology, including whether people should be notified if they are in a surveillance zone, or if their face has been added to a watchlist. Furthermore, he noted that the efficacy of the technology is correlated with its privacy concerns: "If the technology does not work, it can't be a real threat to privacy." (Bowyer, 2004, p. 16) Facial recognition in public places has a history of lacking reliability. In her 2011 book *When Biometrics Fail*, professor Shoshana Magnet discusses some examples of facial recognition failure. In 2004, a city-based surveillance program in Ybor City, Florida had to be canceled for its poor results, while in 2006, a study of a program at the Logan National Airport had a failure rate of 1 misidentification out of 100 people, for an airport that handles 27 million passengers a year (Magnet, 2011).

This all changed in 2010, however (Klosowski, 2020). As computers became more powerful, they eventually became capable of saving and scanning millions of faces at once, fulfilling Bledsoe's original vision of a fully-responsive "man-machine" (Raviv, 2020). Fueled by the endless supply of facial data provided by social media, and an ever-increasing government interest in security, the technology found its niche in the mainstream world.

STS Framework: FRT and SCOT

While facial recognition technology is a step forward in terms of technological convenience, its actual implementation is rife with bias (Buolamwini, 2017). The cause of this inequity is explained by SCOT, or the Social Construction of Technology. As coined by Pitch and Bjiker (2012), SCOT is the idea that technology evolves simultaneous to society, sometimes in multiple directions at once, as per the "multidirectional model" (Pinch and Bjiker, 2012, p. 22). The framework states that a technology's development is influenced by the people who use it, as well as the people who make it. In the case of facial recognition technology, the root cause of the technology's biased results can be attributed most to the datasets used to train the algorithms, which were given to the technology by its creators. These biases then get sent out into greater society by means of the use of technology, causing certain beliefs to perpetuate. For example, if a facial recognition system is required for entry into a company building, the technology not working for certain groups of people makes it easier for their peers to alienate them, since the technology itself is confirming their bias.

This idea of splitting up social groups is explored by Pitch and Bijker in their review of the bicycle and its evolution throughout the 1880s. They attest that the overlap of needs from different social groups contribute to the growth of a technology, and certain groups can contain subgroups. For example, within the scope of the primary "bike riders" group was a subgroup of "women bike riders" (Pinch and Bjiker, 2012, p. 34). During the 1880s, women were expected to ride on tricycles instead of bicycles to accommodate their attire; however, the engineers predicted that eventually, women would start riding bicycles too, due to their increasing popularity. This social group split is seen in FRT's applications as well, with the primary group of users being civilians, and the subgroups of civilians being different social groups of people. Pitch and Bijker attest that each of these subgroups has its own point of view in terms of the problems with a technology, and only by finding and solving these problems can an engineer improve a technology.

In the case of FRT, these social groups can be evaluated through their acceptance of the technology. According to the Pew Research Center (Tyson et. al, 2022), the majority of Americans are either indifferent about public FRT usage, or think it is a good idea. Notably, they state that "people who have heard or read a lot about the use of facial recognition technology are more likely to say it's a bad idea, compared to those who have heard a little or nothing at all on the topic." (Tyson et. al, 2022) This uninformed majority of people is who contributes the most to the social construction of the technology; the more people that accept the technology, the more integrated it becomes in society. However, there is enough pushback from other groups to prevent this discourse from reaching the closure stage of SCOT. These groups include young people, who are informed about the topic, and members of minority communities, who are aware of how FRT's integration may impact them. These groups, who are concerned about their privacy and safety, raise the problems that can be solved to help FRT improve.

Research Results: The Current State of FRT and Its Future

As aforementioned, the development of FRT is a good example of SCOT at work: the interconnectedness of the novelty of the internet, the rapid growth of technology, and the interest of both the government and the general populus was the perfect socio-technical environment for FRT to grow. However, this growth is not without flaw. With more data than ever to collect from for the FRTs, they will become more and more effective…that is, for the types of faces that they are being trained on. The racial bias in FRT has been there since the beginning, when Bledsoe himself proposed an FRT system to identify a person's race: "Woody had proposed to DARPA…'There exists a very large number of anthropological measurements which have been made on people throughout the world from a variety of racial and environmental backgrounds…[that] has not been properly exploited'" (Raviv, 2020). Since the technology has developed, one could say that this aspect of Bledsoe's vision has come to pass, since FRT is disproportionately less effective on women and minorities (Buolamwini, 2017).

Countless studies have been performed on the racial biases in facial recognition technology, which tends to favor White male subjects (Cavazos et. al, 2021). As an example, doctorate researcher Cavazos and her team (2021) tested the accuracy of four different facial recognition algorithms and their success rates on East Asian and European faces; the former "required higher identification thresholds" than the latter to reach the same success rate. In another example, as a graduate researcher, Joy Buolamwini (2017) conducted a study comparing the identification efficiency of three top recognition algorithms against Nordic faces and African faces. She found that the African women were misidentified 32x more than the Nordic men (Buolamwini, 2017, p. 3). The datasets of the algorithms she studied were "overwhelmingly lighter-skinned: 79.6%-82.6%" (Buolamwini, 2017, p. 3).

Both of these studies took place in America, a place full of both racial diversity and racial biases. Due to historical, cultural, and social factors, systemic racism is rampant in American society, resulting in the White male being regarded as the "default." This social construction of technology (SCOT) is demonstrated in the results of these studies, as they found that the majority of data the algorithms contained were for people who were either of European descent, a man, or both. This is, of course, a result of these algorithms being designed in America, by people that likely have racial bias, whether they are aware of it or not. Thus, the technology itself goes on to perpetuate the biases of its creators, demonstrating a great example of SCOT, and how technology develops alongside social ideology to favor the “default” group.

Unfortunately, these developments often cause disadvantage to social groups other than the “default.” In her report, Buolamwini explores the implications of the data collected by the technology, and how it can be used to take advantage of vulnerable populations: “Race or ethnic classification can be used by advertisers to exclude showing housing listings to a protected class like African-Americans. Individuals classified as female based on their facial appearance may be subjected to higher prices as has been reported in instances where vendors use gender information to set prices” (Buolamwini, 2017, p. 24). This is all demographic information that is collected by FRT without a person’s knowledge, and in this case, be used for taking advantage with advertising. Another such implementation, the “billboard that looks back” (Magnet, 2011), is an advertisement screen with a camera that can identify a person’s race and gender, then determine what ads to show them to increase engagement (Magnet, 2011, p. 2).

While showing people ads is relatively benign, the technology is not only used for that; it is also used for identification by the government and law enforcement, which can have even more dire consequences if FRT fails to properly identify someone. For example, in 2020, Robert Julian-Borchak Williams of Detroit, Michigan was arrested because his face had false-positively matched the suspect’s photo in their facial recognition system (Hill & Krolik, 2020). This is another prime example of SCOT in action: the technology is inherently biased, due to the lack of proper data provided for Black men, and the members of the Detroit police force carry their own personal biases. The overlap resulted in Williams’s arrest; despite the fact that they could have investigated further, they chose to jump to the conclusion after putting their trust in the FRT technology and its results. This scenario is evidence that FRT is plagued with socially constructed problems that have their root in societal biases.

As the solution to these inherent biases that FRT possesses, Cavazos and her team devised a “checklist for measuring [the] bias in face recognition algorithms” (Cavazos et al., 2021). Their checklist implores researchers to be introspective about their own work and the biases within it, in both the data they use and how they record their findings. They define “data-driven” factors as the population stats for the demographics in the study; the quality of the algorithm’s representation of each group; the balance of subgroups in the algorithm between demographics; and balanced imaging conditions and quality for each group, essentially ensuring that every group is represented evenly throughout an experiment (Cavazos et al., 2021). This is a good example of professional scientists making an effort to reframe the SCOT ideology surrounding a technology by establishing new standards for its creation. If quality control guidelines like these were set for future FRT algorithms, then bias could potentially be eliminated from future iterations of the technology.

Other researchers are finding ways to solve the bias problem with artificial intelligence (AI). Not unlike FRT itself in 2015, AI has recently and rapidly become ubiquitous since the success of ChatGPT and other successful generative AI algorithms in 2022. Capitalizing on this boom, Nicolò Di Domenico and coauthors created their own AI algorithm to generate hyper-realistic faces with a variety of features, such as gender and ethnicity (2024). Their AI faces aligned with the ISO/ICAO standards, meaning they followed the legal layouts for identification photos. After training their dataset on a ICAO verification system, the algorithm produced 25%

more European faces and 25% less East Asian and South Asian faces (Domenico, 2024), illuminating the bias present in the existing system.

Even with the bias problems solved, there is still the question of the invasive nature of FRT being used in mass-surveillance, like that which was seen at the Super Bowl XXXV. While the majority of people (63%) have come to accept this use for large-scale events (Tyson et. al, 2022), people say the opposite when asked about surveillance in public areas, with 68% agreeing with the stance that “it is not acceptable to scan people as they walk down the street.” To solve this problem, Nabil and his team created an end-to-end encryption surveillance system, where the information created by faces picked up is converted into data that only the computer can understand, protecting the information from humans (Nabil et. al, 2022). This way, law enforcement can use the technology without having immediate access to a person’s identity. Another team is working to solve this problem with AI. Professor Rita Cucchiara and coauthors (2024) published an article exploring AI-powered privacy protection in *Computer,* a respected technological magazine. They argue that AI could be used in place of people to monitor cameras for public safety purposes, granting people privacy in the form of data protection, or have AI look specifically for dangerous actions, rather than people’s faces. With this method, threats can be identified without continuous monitoring from humans (Cucchiara et al., 2024), and can streamline the responsibilities of law enforcement.

These solutions are a good example of SCOT because they demonstrate how fads in technology can influence other technologies; namely, with AI being used on FRT. Fads are inherently a social phenomenon: when people get excited about something, it enters the forefront of everyone’s minds, even scientists’. While bias might be an inherent factor of SCOT, novelty is the tried and true catalyst of invention in its eyes. After all, nothing contributes to social construction more than being the “next big thing.” This does not mean that AI is without its own set of problems; it is just new enough that experts are eager to work with it.

FRT was the “next big thing” once, but has reached the point of being “just another thing.” As the novelty wears off, we are left with a technology that has become a core part of our lives, affecting some of us more than others. It is these lingering effects that prevent FRT from reaching SCOT closure; despite being ubiquitous, this technology will continue to stagnate if these problems go unsolved. As this study was a literary review, it was easy to focus on the highlights of the technology and its lifespan and gloss over the more mundane details. When used properly and ethically, FRT could, in concept, make people’s lives easier and safer, but even those conveniences come with new baggage to sort out. For example, if FRT ever was to replace credit cards as a payment option, would that mean that banks could collect face information? Would people under a certain tax bracket be protected from face collection, or would they be discriminated against for being unable to pay? These questions could be asked, and even answered, in future renditions of this study.

Conclusion

In conclusion, while facial recognition technology is a step forward in terms of technological convenience, its actual implementation is rife with bias (Buolamwini, 2017) and is used to encroach on people’s lives, whether it is collecting information without their knowledge (Magnet, 2011) or framing them for crimes (Kashmir and Hill, 2020). This intertwined bias is a prime example of SCOT, since the biases themselves are sourced from society and the people who work on the technology. Potential solutions to this problem involve SCOT reframing and challenging the biases, through critical reflection and/or the implementation of other technologies.

Bibliography

"About Face ID advanced technology." (n.d.). *Apple Support*. Retrieved November 7, 2024, from https://support.apple.com/en-us/102381

Bowyer, K. W. (2004). "Face recognition technology: Security versus privacy." *IEEE Technology and Society Magazine*, 23(1), 9–19. https://doi.org/10.1109/MTAS.2004.1273467

Buolamwini, J. (n.d.). "Gender shades." *MIT Media Lab*. Retrieved November 10, 2024, from https://www.media.mit.edu/publications/full-gender-shades-thesis-17/

Cavazos, J. G., Phillips, P. J., Castillo, C. D., & O'Toole, A. J. (2021). "Accuracy comparison across face recognition algorithms: Where are we on measuring race bias?" *IEEE Transactions on Biometrics, Behavior, and Identity Science*, 3(1), 101–111. https://doi.org/10.1109/TBIOM.2020.3027269

Cucchiara, R., Baraldi, L., Cornia, M., & Sarto, S. (2024). "Video surveillance and privacy: A solvable paradox? *Computer*, 57(3), 91–100. https://doi.org/10.1109/MC.2023.3316696

Di Domenico, N., Borghi, G., Franco, A., & Maltoni, D. (2024). "Onot: A high-quality icao-compliant synthetic mugshot dataset." *2024 IEEE 18th International Conference on Automatic Face and Gesture Recognition (FG)*, 1–10. https://doi.org/10.1109/FG59268.2024.10581986

Hill, Kashmir, and Aaron Krolik. "Wrongfully Accused by an Algorithm." *New York Times*, 24 June 2020. *www.nytimes.com*, https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html.

Klowsowski, Thorin. "Facial Recognition Is Everywhere. Here's What We Can Do About It." *Wirecutter: Reviews for the Real World*, 15 July 2020, https://www.nytimes.com/wirecutter/blog/how-facial-recognition-works/.

Libby, Christopher, and Jesse Ehrenfeld. “Facial Recognition Technology in 2021: Masks, Bias, and the Future of Healthcare.” *Journal of Medical Systems*, vol. 45, no. 4, 2021, p. 39. *PubMed Central*, https://doi.org/10.1007/s10916-021-01723-w.

Magnet, S. A. (2011). When biometrics fail: Gender, race, and the technology of identity. Duke University Press. https://doi.org/10.1215/9780822394822

Nabil, M., Sherif, A., Mahmoud, M., Alsmary, W., & Alsabaan, M. (2022). Accurate and privacy-preserving person localization using federated-learning and the camera surveillance
systems of public places. IEEE Access, 10, 109894–109907.
https://doi.org/10.1109/ACCESS.2022.3214227

Pinch, Trevor, and Bijker Wiebe. “The Social Construction of Facts and Artifacts: Or How the Sociology of Science and the Sociology of Technology Might Benefit Each Other.” *MIT Press*, The MIT Press, 2012, https://mitpress.mit.edu/9780262517607/the-social-construction-of-technological-systems/

Raviv, Shaun. “The Secret History of Facial Recognition.” *Wired. www.wired.com*, https://www.wired.com/story/secret-history-facial-recognition/. Accessed 30 Mar. 2025.

Tyson, L. R., Cary Funk, Monica Anderson and Alec. (2022, March 17). 2. “Public more likely to see facial recognition use by police as good, rather than bad for society.” *Pew Research Center.*
https://www.pewresearch.org/internet/2022/03/17/public-more-likely-to-see-facial-recognition-
use-by-police-as-good-rather-than-bad-for-society/

Wright, C. (2022). “FACIAL RECOGNITION TECHNOLOGY: Federal Agencies’ Use and Related Privacy Protections (GAO-22-106100; Testimony Before the Subcommittee on Investigations and Oversight, Committee on Science, Space, and Technology, House of Representatives, pp. 1–18).” *United States Government Accountability Office.* https://www.gao.gov/assets/gao-22-106100.pdf