UCSD UNDERGRADUATE LAW REVIEW

VELMAR AMADOR-LANKSTER

Facial Recognition in Policing: How Algorithmic Bias Targets People of Color

ABSTRACT. The presence of racially biased facial recognition technology (FRT) in law enforcement presents significant legal and ethical challenges, especially in its disproportionate impact on communities of color. The biases embedded in FRT systems, due to non-diverse training datasets, lead to misidentifications that result in wrongful arrests, detentions, and broader violations of constitutional rights. This exacerbates systemic racial inequalities and further entrenches discriminatory practices within the criminal justice system. The growing reliance on FRT for policing necessitates comprehensive reforms in both technology and law to address these concerns. Therefore, solutions aimed at reducing racial bias must include diversifying training datasets, improving data quality, and incorporating advanced techniques to enhance the accuracy of these systems. However, technological improvements alone are insufficient to resolve deeper racial issues. A holistic approach is required, involving robust legal frameworks for accountability, extensive training for law enforcement, and independent audits of FRT use to ensure its fair and transparent application. While further research into the practical, large-scale implementation of these reforms is necessary, addressing the intertwined technological and legal challenges is essential to address the widespread concerns presented by racially biased FRT in policing and to protect the rights of all individuals.

AUTHOR. Velmar Amador-Lankster is a third-year Political Science and Sociology double major at UC San Diego. After graduation, she plans to attend law school and pursue a career in civil rights law, advocating for those marginalized by the criminal justice system. She would like to thank her family for their unwavering support, Shyan Khan and Eshwari Gundi for their editing contributions, Dr. Christopher Stout for his time and expertise during the review process, and Ansh Mujral for his assistance in refining the technical components of this paper.

INTRODUCTION

Originally developed for use within the larger national security domain, facial recognition technology (FRT) has become integral to policing and widely adopted by local law enforcement agencies. Although law enforcement promotes FRT as a tool for efficient and accurate investigations, its utilization has raised widespread concern about its legality and constitutionality, particularly due to the racial biases embedded in the FRT system. A key element of FRT is that it operates by capturing and comparing facial images to databases for surveillance and identification purposes. However, concerns have surfaced about the system's evident flaws, particularly in the form of "algorithmic bias." This bias leads to inaccurate results in FRT, particularly for individuals with darker skin tones, due to training datasets that predominantly feature white faces. While it is more technically accurate to say the technology struggles to identify darker skin tones, rather than specific racial groups, the outcome remains that people of color– especially those with darker complexions– are disproportionately affected. Racial and ethnic group data consistently show that FRT is inherently less accurate for these populations, illustrating the racial bias embedded in the technology.[1] Framing the issue as an "impact on people of color," as opposed to solely on individuals with darker skin tones, highlights the systemic nature of the problem. The central issue explored in this paper is the practical challenges of FRT's everyday application, rather than its theoretical accuracy, illustrating that in real-world scenarios, the risk of misidentification is notably higher, especially for people of color. The racial bias embedded in facial recognition technology, combined with the complexities of real-world application, significantly increases the risk of misidentification– potentially leading to wrongful arrest or detention of people of color. From a legal perspective, the use of racially biased FRT by police departments could lead to violations of important constitutional rights protected under the Fourth and Fourteenth Amendments.

FRT violates the Fourth Amendment by allowing unreasonable searches and seizures based on inaccurate information derived from a faulty, racially biased FRT.[2] This may lead to police officers arresting or detaining individuals based on data that

---

[1] Hans de Zwart, *Current State of Research: Face Detection Still Has Problems With Darker Faces*, Racism & Tech Blog, Jul. 18, 2023,
https://racismandtechnology.center/2023/07/18/current-state-of-research-face-detection-still-has-problems-with-darker-faces/.

[2] U.S. Const. amend. IV, *Constitution Annotated*, Oct. 12, 2024,
https://constitution.congress.gov/browse/amendment-4/.

disproportionately misidentifies people of color. The Fourth Amendment reads as follows:

> The right of the people to be secure in their persons, houses, papers, and effects, against unreasonable searches and seizures, shall not be violated, and no Warrants shall issue, but upon probable cause, supported by Oath or affirmation, and particularly describing the place to be searched, and the persons or things to be seized.[3]

Furthermore, it infringes upon the Fourteenth Amendment's guarantee of equal protection and due process of law by disproportionately subjecting people of color to the discriminatory impact of FRT, thereby denying them fair and equal treatment under the law.[4] Section 1 of the Fourteenth Amendment reads as follows:

> No State shall make or enforce any law which shall abridge the privileges or immunities of citizens of the United States; nor shall any State deprive any person of life, liberty, or property, without due process of law; nor deny to any person within its jurisdiction the equal protection of the laws.[5]

The consequences of racial biases in FRT for communities of color are particularly evident in its use by law enforcement. When facial recognition images or footage are used by police in criminal investigations, such as in suspect lineups, the technology's inherent biases can result in higher misidentification rates for people of color. As a result, these individuals are more likely to be subjected to wrongful arrests and detentions, which not only unjustly target innocent people and can lead to wrongful convictions and incarceration but also result in long-term personal and societal repercussions. These include acquiring a criminal record that may hinder future employment, limit housing opportunities, and negatively affect other essential aspects of life. This raises serious concerns about potential violations of individuals' Fourth and Fourteenth Amendment rights, demonstrating the need to re-evaluate FRT usage by police and underscoring the necessity for both technological and legal reforms.

I. THE DEVELOPMENT, FUNCTIONALITY, AND INTEGRATION OF FRT

[3] U.S. Const. amend. IV.
[4] U.S. Const. amend. XIV, § 1, *Constitution Annotated*, Oct. 12, 2024, https://constitution.congress.gov/constitution/amendment-14/.
[5] U.S. Const. amend. XIV, § 1.

A. Development and Purpose of FRT

Although the concept of automated facial recognition dates back to the 1960s, full automation of the technology was not achieved until the late 1980s and early 1990s with the introduction of the Eigenfaces method.[6] This development simplified the process of real-time facial recognition by eliminating the manual component. Humans were no longer required to manually locate facial features and input the coordinates into a computer for matching with stored images. Instead, the algorithm could now operate autonomously, recognizing faces with greater efficiency. By eliminating the need for humans to manually identify and input facial features, the process became more efficient and less susceptible to human error. This ability to autonomously recognize faces enabled more accurate and scalable applications, leading to the expanding implementation of facial recognition systems in various fields, including law enforcement and security. In 1993, the United States Department of Defense deployed FRT for national security purposes, advancing the algorithmic technology through its Face Recognition Technology program to standardize benchmark measures for various systems. It was not until 2001 that FRT was first deployed in large-scale security operations, such as at the Super Bowl, to monitor criminal activity among attendees.[7] This period also marked the beginning of the widespread development of FRT systems for government use by private companies. Since the turn of the century, FRT has continued to evolve and develop, largely due to the monumental advances in machine learning and artificial intelligence (AI). Over the past decade, the rapid expansion of FRT has revolutionized security and identification practices, while also sparking major public controversy over its ethical and legal implications.

B. Understanding the Functionality of FRT

FRT is designed to identify individuals by analyzing biometric data from facial features, typically taken from images or video footage. This data is subsequently entered into a system where it is automatically encoded and compared against other

[6] Facit AI, *Evolution of Facial Recognition Technology*, Facit AI Insights, Dec. 5, 2024, https://facit.ai/insights/evolution-of-facial-recognition-technology.
[7] John D. Woodward, Jr. et al., *Facial Recognition Technology: A Survey of Policy and Implementation Issues*, RAND Corp., Issue Paper IP-209, 12 (2001), https://www.rand.org/pubs/issue_papers/IP209.html.

images stored in the database to find a match amongst a candidate list of "similar" faces.[8] The database of photos is sourced from facial recognition companies like Clearview AI, which gathers billions of publicly available images from online platforms like YouTube and Facebook, while also allowing customers to upload photos of individuals that are automatically matched with other images and sources within the database.[9] Clearview AI has partnered with more than 3,100 federal and local law enforcement agencies to identify individuals beyond the scope of government databases. As a result of this expansive effort, Clearview AI has been able to amass about 10 billion photos in its database, far surpassing the FBI, which possesses about 640 million images.[10] What makes facial recognition unique is that, unlike other types of biometric data such as DNA and fingerprints, which remain unchanged, FRT must account for various factors that can change over time, including aging, plastic surgery, and other physical alterations.[11] While this presents a challenge, advancements in technology have enhanced FRT's ability to address these factors. Unlike earlier systems, top-tier modern algorithms like Deep Convolutional Neural Networks (DCNNs) are designed to mirror human perception with sensitivity to both individual facial features and their spatial configuration, critical elements for distinguishing individuals even as appearances change.[12] A National Academies Press publication in 2024 titled *Facial Recognition Technology: Current Capabilities, Future Prospects, and Governance*, highlights an experiment showing that modern FRT algorithms can accurately identify individuals in a database of 12 million mugshots, even if their appearance has changed over 2 to 19 years.[13] However, in real-life scenarios where factors like poor image quality or other variables are involved, even the most advanced modern algorithms like DCNNs still face obstacles in accurately identifying individuals. This challenge is more

[8] Clare Stouffer, *How Facial Recognition Software Works*, Norton: Blog, Oct. 5, 2021, https://us.norton.com/blog/iot/how-facial-recognition-software-works.

[9] Kashmir Hill, *The Secretive Company That Might End Privacy as We Know It*, N.Y. Times, Jan. 18, 2020, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html.

[10] Nicol Turner Lee, *Police Surveillance and Facial Recognition: Why Data Privacy Is an Imperative for Communities of Color*, Brookings, Jul. 15, 2021, https://www.brookings.edu/articles/police-surveillance-and-facial-recognition-why-data-privacy-is-an-imperative-for-communities-of-color/.

[11] *INTERPOL: How We Work – Forensics*, https://www.interpol.int/en/How-we-work/Forensics/Facial-Recognition (last visited Jan. 24, 2025).

[12] Virginia E. Strehle et al., *Deep Convolutional Neural Networks Are Sensitive to Face Configuration*, 10 Journal of Vision (2024), https://pmc.ncbi.nlm.nih.gov/articles/PMC11542502/pdf/jovi-24-12-6.pdf.

[13] National Research Council, *Facial Recognition Technology: Current Capabilities, Future Prospects, and Governance* 45 (2024), https://nap.nationalacademies.org/read/27397/chapter/4.

pronounced for middle-tier technology, which exhibits error rates nearly ten times higher when attempting to match photos taken 18 years ago.[14]

The operation of FRT software can be categorized into three primary stages: detection, analysis, and recognition.[15] The first stage focuses on detecting or locating a face within an image or video without yet attempting to identify the individual to which it belongs. At this stage, the FRT algorithms are trained by the system's creator, who inputs a large volume of facial images to help the system learn the intricate details of facial features. The more training the system undergoes, the more accurate it becomes at differentiating subtle variations in facial features. Therefore, the diversity of the training images plays a crucial role in the system's accuracy during both the analysis and recognition stages. Studies reveal that East African, West African, and East Asian groups experience the highest false match rates, while the Eastern European group exhibits the lowest.[16] However, it is important to note that several algorithms developed by Chinese developers have been shown to perform differently from others, particularly in achieving lower false match rates on East Asian faces. This suggests that the design of the algorithms and the selection of training data play a critical role in influencing a model's performance. While achieving equal racial representation in training data, ensuring a more balanced distribution of faces from Caucasian, African, Asian, and other demographic groups, does not automatically eliminate all algorithm biases, the racial composition of the training data significantly influences the racial bias in FRT. Research shows that technology performs less accurately on faces from ethnic groups that were underrepresented during training.[17]

---

[14] William Crumpler, *How Accurate Are Facial Recognition Systems—And Why Does It Matter?*, Center for Strategic and International Studies, Jul. 22, 2020, https://www.csis.org/blogs/strategic-technologies-blog/how-accurate-are-facial-recognition-systems-and-why-does-it.

[15] Thorin Klosowski, *How Facial Recognition Works*, *Wirecutter*, N.Y. Times, Oct. 10, 2022, https://www.nytimes.com/wirecutter/blog/how-facial-recognition-works/.

[16] Kolla, *The Impact of Racial Distribution in Training Data on Face Recognition Algorithms*, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 313, 2023, https://openaccess.thecvf.com/content/WACV2023W/DVPBA/papers/Kolla_The_Impact_of_Racial_Distribution_in_Training_Data_on_Face_WACVW_2023_paper.pdf (last visited Feb. 27, 2025).

[17] Kolla, *The Impact of Racial Distribution in Training Data on Face Recognition Algorithms*, *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 313, 2023, https://openaccess.thecvf.com/content/WACV2023W/DVPBA/papers/Kolla_The_Impact_of_Racial_Distribution_in_Training_Data_on_Face_WACVW_2023_paper.pdf.

The second stage involves the process of mapping the detected face, which refers to the measurement of the distances, shapes, thermal profiles, and nuances of the facial features.[18] This data is converted into a string of numbers or points, commonly known as a faceprint, and added to the recognition database. During the final stage, the system attempts to match the faceprint with existing faceprints in the databases to confirm the identity of the face. To authenticate the face, the system analyzes the faceprint input and compares it to the database, identifying the closest matches and applying confidence scores to evaluate the strength of the match. This reduces the likelihood of errors and improves the overall accuracy of the system.[19]

According to the International Criminal Police Organization, more widely known as INTERPOL, optimal results from FRT require an image with even lighting on the face and a neutral background.[20] While these are the ideal conditions for achieving highly accurate results, real-world deployments of FRT– where uneven lighting, constant movement, and other distractions are common– make it difficult to obtain precise matches. This decrease in accuracy is reflected in data collected by the Facial Recognition Vendor Test, which discovered that the error rate for one leading algorithm rose from 0.1 percent when tested against high-quality mugshots to 9.3 percent when matched to individuals' photos captured in real-world conditions.[21] This reveals that in practical applications, the accuracy of FRT significantly decreases due to various factors that affect its ability to correctly identify individuals from their faceprints. Given these limitations, the use of FRT by law enforcement can disproportionately impact certain demographics, leading to wrongful arrests and detentions.

C. Integration of FRT into Policing Practices and Its Benefits for Law Enforcement

FRT has become a common tool for law enforcement officers at the federal, state, and municipal levels.[22] This technology has been cited as instrumental in tracking

[18] Klosowski, *supra* note 15.
[19] Getsafeandsound, *Facial Recognition System*, https://getsafeandsound.com/blog/facial-recognition-system/(last visited Oct. 4, 2024).
[20] *INTERPOL: How We Work – Forensics*.
[21] Crumpler, *supra* note 14.
[22] Clare Garvie et al., *The Perpetual Line-Up*, Center on Privacy & Technology, Georgetown Law, Oct. 18, 2016, https://www.perpetuallineup.org/.

suspects, finding missing persons, solving cold cases, and performing critical tasks.[23] FRT has been integrated within existing police databases, such as mugshots and driver's license records, and expanded through collaborations with private companies like Clearview AI to access additional data sources. Proponents of this technology argue that FRT has revolutionized how police officers solve crimes and locate suspects, significantly speeding up the process of identification and tracking with unprecedented precision. They assert that before FRT, identifying suspects captured on surveillance cameras was challenging and tedious, especially without strong leads.[24] For example, in 2019, the New York City Police Department used FRT to identify a suspect involved in a potential subway bombing. Detectives applied the technology to images of a man who was suspected of leaving rice cookers in the Fulton Street subway station. FRT returned hundreds of potential matches, and after multiple stages of review and verification, the suspect was identified within one hour.[25] In addition to helping law enforcement in time-sensitive situations, FRT has also played a crucial role in exonerating innocent individuals. For example, in 2022, a Florida man who had been falsely accused of vehicular homicide was exonerated after FRT identified and located a key witness to the fatal crash. The witness confirmed that the man was the passenger, not the driver, of the vehicle involved in the accident.[26] Furthermore, in addition to tracking suspects and solving crimes, FRT can assist in real-time identification and provide valuable investigative leads from large volumes of data. By streamlining key investigative processes, FRT technology has significantly improved resource allocation, allowing law enforcement agencies to reduce the time and resources previously spent manually investigating leads.[27]

A 2017 study by Georgetown Law's Center on Privacy and Technology found

---

[23] Abigail Gaetz, *Understanding the Debate on Facial Recognition Technology in Policing: Pros, Cons, and Privacy Concerns*, The Alliance for Citizen Engagement, Sept. 6, 2024, https://ace-usa.org/blog/research/understanding-the-debate-on-facial-recognition-technology-in-policing-pros-cons-and-privacy-concerns/.

[24] *Id.*

[25] Craig McCarthy, *How NYPD's Facial Recognition Software ID'd Subway 'Rice Cooker Kook'*, N.Y. Post, Aug. 25, 2019, https://nypost.com/2019/08/25/how-nypds-facial-recognition-software-ided-subway-rice-cooker-kook.

[26] Jake Parker, *Facial Recognition Success Stories Showcase Positive Use Cases of the Technology*, Security Industry Association, Jul. 16, 2020, https://www.securityindustry.org/2020/07/16/facial-recognition-success-stories-showcase-positive-use-cases-of-the-technology/.

[27] Gaetz, *supra* note 23.

that, as of 2013, approximately 25.6 percent of state and local law enforcement agencies had access to FRT or could use it after requesting permission, completing training, or signing an agreement.[28] This statistic highlights how widespread FRT has become, with roughly one in four state and local law enforcement agencies utilizing the technology, each with varying levels of implementation. For instance, in August 2019, the Chula Vista Police Department in San Diego County was authorized to use drones that covered over 33 percent of the city's geographic region, with the Federal Aviation Administration approving an expansion to permit drones to cover 100 percent of the city's territory as of March 2021.[29] This technology has primarily been deployed to provide live video feeds to police officers before they arrive at the scene of a 911 call, which has played a critical role in allowing officers to assess the situation from a safe distance, improving response strategy. As of March 2021, this program utilized drones in more than 5,400 missions and assisted in more than 650 arrests, a development that underscores the expanding reliance on FRT via drones by police departments. FRT has become increasingly widespread across law enforcement agencies, making the need to ensure its accuracy more critical than ever.

D. The Key Issue with FRT

The advancement of FRT has expanded government surveillance in urban spaces intended for public use, such as sidewalks and parking lots. Despite its numerous benefits for law enforcement, the use of FRT undeniably presents significant issues. One concerning problem is that this technology disproportionately misidentifies people of color, stemming from both technological and algorithmic flaws. As a result, already vulnerable demographics within the criminal justice system are further exposed to potential issues with FRT and the resulting consequences.

FRT has disproportionately high rates of inaccuracy when identifying and matching people of color, a profound systemic flaw due to a lack of diverse photos used to train the system. Even IJB-A data constructed by the National Institute of Standards and Technology (NIST), which is intended to be more representative of geographical diversity compared to the typical non-diverse datasets, still contains facial images that

[28] Garvie et al., *supra* note 22.

[29] Faine Greenwood, *Chula Vista Police Have Been Using Drones to Watch You*, Slate, May 12, 2021, https://slate.com/technology/2021/05/chula-vista-police-drone-program.html.

are 79.6 percent lighter skin tones.[30] A study conducted in 2018 found that while the FRT error rate for white men was 0.8 percent, the error rate for darker-skinned women was around 34.7 percent, revealing a major discrepancy in the accuracy of FRT that increases with intersectionality. NIST has reported that individuals of Asian, African, and American Indian descent– particularly African American women– experience higher rates of false-positive identifications.[31] Researchers have attributed this disparity to several factors, one of which is the "own-race bias."[32] In the context of FRT, this phenomenon suggests that database systems– developed predominantly by white male engineers– are trained primarily using white faces.[33] As a result, the technology is more accurate at detecting and recognizing white faces. However, FRT systems struggle to recognize faces from other racial or ethnic groups, resulting in significant biases in recognition accuracy.

System discrepancy can be traced back to the earliest FRT, which were primarily trained on sample sets composed of white men. Limited exposure to non-white faces during training leads to significantly higher error rates when FRT systems attempt to identify the faces of people of color. In the context of law enforcement, inaccuracies in identification can have profound consequences that reinforce existing racial disparities in the criminal justice system, targeting and discriminating against people of color.[34] A study conducted by the NIST in 2019 found that image quality is a key factor in facial recognition errors, with racial disparities playing a significant role.[35] The study reveals that while certain image quality issues, such as poor camera angles and blurriness, affect all racial groups similarly, other problems, like underexposure, have a disproportionate impact on individuals with darker skin tones. This problem can be traced to the historical calibration of camera technology, which has primarily been designed for lighter skin tones. As a result, the legacy of camera design has prioritized the accurate recognition of people with lighter skin tones, further exacerbating the racial bias caused

[30] Beth Findley, *Why Racial Bias Is Prevalent in Facial Recognition Technology*, JOLT Digest, Nov. 3, 2020, https://jolt.law.harvard.edu/digest/why-racial-bias-is-prevalent-in-facial-recognition-technology.
[31] T.J. Benedict, *The Computer Got It Wrong: Facial Recognition Technology and Establishing Probable Cause to Arrest*, 79 Wash. & Lee L. Rev. 849 (2022).
[32] Hoo Keat Wong et al., *The Own-Race Bias for Face Recognition in a Multiracial Society*, Front Psychol., Mar. 6, 2020, https://pubmed.ncbi.nlm.nih.gov/32210861/.
[33] David McDaniels, *Racial Disparities in Facial Recognition Technology*, ICAAD, Jun. 14, 2021, https://icaad.ngo/2021/06/14/racial-disparities-in-facial-recognition-technology/.
[34] Wong et al., *supra* note 32.
[35] Findley, *supra* note 30.

by the non-diverse training of facial recognition systems and leading to less accurate outcomes for people of color.

Ultimately, the disproportionate impact of FRT is not only a result of inherent flaws in the algorithms but is further exacerbated by substandard policing practices. A key factor contributing to the impact of racial bias within the technology is law enforcement's improper use and overreliance on FRT. There is a lack of due diligence in handling FRT images and video footage, particularly concerning the use of unreliable images for identification in criminal investigations. Humans tend to place an overwhelming amount of trust in automated technology, a concept known as automation bias.[36] Automation bias occurs when individuals using the technology over-rely on computer output as a shortcut, replacing the necessary steps of critical information gathering and processing.[37] Even when people are aware of the inherent flaws in these systems, they remain hesitant to question the outputs of the technology (i.e., facial matches), a behavior known as automation-induced complacency.[38] This is problematic because it may cause individuals to blindly accept inaccurate data rather than take the necessary action to verify the information in a follow-up investigation. It has been found that, at times, individuals may even override their correct judgments in favor of inaccurate recommendations from decision support systems like FRT. This is particularly relevant when considering the overreliance on FRT by police, as it emphasizes the potential for automation bias to influence decision-making in criminal investigations. This can have significant consequences, such as inaccurate identifications that lead to wrongful arrests. Some law enforcement agencies have acknowledged this issue and recommended that FRT matches should be treated as anonymous tips rather than definitive evidence.[39] This highlights the importance of police departments conducting thorough follow-up investigations to confirm accurate identification, with FRT matches serving as one component of the process, rather than the conclusion.[40]

---

[36] Benedict, *supra* note 31, at 860.

[37] Kate Goddard et al., *Automation bias: a systematic review of frequency, effect mediators, and mitigators,* Journal of the American Informatics Association, Jun. 16, 2011, https://pmc.ncbi.nlm.nih.gov/articles/PMC3240751/.

[38] *Id.*

[39] U.S. Comm. on Civil Rights., The Civil Rights Implications of the Federal Use of Facial Recognition Technology 89 (Sept. 2024).

[40] Matthew Guariglia, *Police Use of Face Recognition Continues to Wrack Up Real-World Harms*, Electronic Frontier Foundation, Jan. 15, 2025, https://www.eff.org/deeplinks/2025/01/police-use-face-recognition-continues-wrack-real-world-harms.

Nonetheless, this critical practice has not been adopted by all police departments, resulting in continued wrongful arrests and detentions due to inadequate investigations following FRT matches. As emphasized by former Detroit Police Chief James Craig, if the city's police officers only relied on FRT without any other support, it would result in misidentifications 96 percent of the time.[41] This shows how serious issues may arise when police departments rely exclusively on FRT as the sole method for identifying or surveilling a suspect. The presence of automation bias in law enforcement, when police officers place excessive trust in automated systems, demonstrates the urgent need for stringent regulations and oversight, especially when these technologies are applied in high-stakes situations such as criminal investigations.

II. THE CONSTITUTIONAL IMPLICATIONS OF USING RACIALLY BIASED FRT IN LAW ENFORCEMENT

The widespread use of FRT by police departments, coupled with its racially biased algorithms, raises significant legal concerns regarding its impact on constitutional protections, particularly concerning unreasonable searches and seizures, as well as equal protection and due process.

A. Constitutional Concerns: Violation of the Fourth Amendment

The Fourth Amendment of the United States Constitution safeguards people's rights against unreasonable searches and seizures of their persons, houses, papers, and effects.[42] This amendment mandates that probable cause must be established to obtain search and seizure warrants. The requirement of probable cause states that there must be a reasonable basis for believing that a crime has been committed or that evidence related to a crime can be found in the place to be searched.[43] This requirement is rooted in the notion that citizens are entitled to protection from unwarranted or arbitrary government intrusion into their privacy.

However, the concept of reasonable basis, which forms the foundation of the legal standard of probable cause, is subjective. This is due to the absence of a clearly defined,

[41] Christina Swarns, *When Artificial Intelligence Gets It Wrong*, Innocence Project, Sept. 19, 2023, https://innocenceproject.org/news/when-artificial-intelligence-gets-it-wrong/.
[42] U.S. Const. amend. IV, *Constitution Annotated*, Oct. 12, 2024, https://constitution.congress.gov/browse/amendment-4/.
[43] Legal Information Institute, *Probable Cause*, Cornell Law School, https://www.law.cornell.edu/wex/probable_cause (last visited Feb. 24, 2025).

widely agreed-upon definition of what exactly constitutes a reasonable basis. As a result, this lack of clarity allows for manipulation and inconsistency in applying what constitutes a reasonable basis under probable cause, as what one individual may consider reasonable, another might dispute, especially when the facts are ambiguous. In the context of FRT, one officer might view a blurry, unreliable FRT match as sufficient to form a reasonable basis for probable cause, while another officer might question the basis of such a match. This subjectivity and lack of clear definition create a legal gray area, where the absence of a solid standard for reasonable basis undermines consistency in how probable cause is applied.

Given this ambiguity, this paper asserts that the use of FRT undermines the standard of probable cause, as it leads to arrests based on unreliable evidence, lacking a reasonable basis to suspect criminal involvement. The responsibility for establishing probable cause rests with the police officers, who must present enough evidence to justify actions like obtaining a search warrant or making an arrest. Because FRT disproportionately misidentifies people of color at significantly higher rates than white individuals, potential matches lack the accuracy and impartiality necessary to establish a reasonable belief that a specific person has committed a crime. Since the technology has not proven to be accurate across all demographics, an FRT match alone should not justify arresting someone. The reliance by police officers on FRT outputs, despite blurry and unreliable images, to establish probable cause in warrant applications highlights a failure to uphold their duty of providing reliable and accurate evidence. This results in a violation of an individual's right to protection from unreasonable searches and seizures.

To properly use FRT matches in establishing probable cause, significant improvements in the algorithmic accuracy of the technology are needed. Additionally, law enforcement must ensure that any FRT matches used in warrant applications are of the highest quality and supported by thorough investigations to verify the suspect's identity. Unless these conditions are met, many argue that the reliance solely on FRT as evidence for probable cause constitutes an unreasonable search or seizure, as individuals are subjected to wrongful arrest based on inaccurate or racially biased technology.

Ultimately, while the legal definition of probable cause– especially its focus on reasonable basis– remains ambiguous, it is crucial to recognize that any technology that is demonstrably racially biased in its application should not be solely relied upon to establish probable cause. Without additional corroborating evidence, unjust outcomes that disproportionately affect marginalized communities will become commonplace.

Therefore, the potential for discrimination and injustice must outweigh the need for technological convenience in law enforcement practices.

B. Constitutional Concerns: Violation of the Fourteenth Amendment

The Fourteenth Amendment of the United States Constitution protects individuals from the government's arbitrary deprivation of their rights to life, liberty, and property without due process of law.[44] This amendment was enacted to prevent the government from acting arbitrarily, meaning it cannot deprive any individual of these rights without a fair and reasonable basis and due legal process. Furthermore, this amendment ensures that all individuals are afforded equal protection under the law.

The widespread use of FRT has revealed significant flaws in its algorithms and technology, leading to disproportionate consequences for the arrests and detention of individuals from certain demographic groups. Some have argued that FRT violates the Fourteenth Amendment's Equal Protection Clause because it is inherently biased and unfairly targets marginalized communities. Algorithmic misidentification by FRT can cause law enforcement to disproportionately impact certain racial and ethnic groups. Under existing legal precedents, to establish a claim of racially biased law enforcement, it must be demonstrated that the enforcement action resulted in discriminatory effects and was carried out with discriminatory intent.[45] Although it may seem like a straightforward process, applying this framework to automated, algorithm-driven systems like FRT is challenging, as demonstrating discriminatory intent within a non-human entity is difficult.[46] Additionally, if the technology remains fundamentally prone to racial bias and inaccuracy, it could lead to the unjust deprivation of rights and liberties, particularly for people of color, through misidentification and wrongful arrests. If an individual is arrested or detained based on biased technology, their due process rights are violated as they are subjected to an unfair legal process.

The use of FRT in criminal investigations dates back to 2001 and has been critical in many cases, but this is rarely disclosed to defendants during legal proceedings. If defendants are not informed that FRT has been used as evidence, they will not have the opportunity to challenge the reliability of the technology. Thus, this lack of

---

[44] U.S. Const. amend. XIV, § 1, *Constitution Annotated*, Oct. 12, 2024, https://constitution.congress.gov/constitution/amendment-14/.
[45] Kelsey Y. Santamaria, *Facial Recognition Technology: Current and Planned Uses by Federal Agencies*, Congressional Research Service, Sept. 24, 2020, https://www.congress.gov/crs-product/R46541.
[46] *Id*.

transparency could lead to unfair legal outcomes, especially if the technology has flaws or biases that disproportionately affect certain groups. For instance, over the past four years, the Miami Police Department has conducted 2,500 FRT searches in criminal investigations, which have resulted in at least 186 arrests and more than 50 convictions.[47] However, only a mere 7 percent of those individuals involved were notified that FRT had been used in their cases.[48] In the rare instance that defendants are informed of FRT use, it is uncommon for error rates or racial biases in the technology to be discussed in court. Furthermore, prosecutors usually refrain from introducing FRT evidence at trial, which deprives defendants of the opportunity to contest on evidentiary grounds.[49] In such cases, withholding information about the use of this technology deprives defendants of the opportunity to challenge the accuracy of the results, which can be perceived as a failure to uphold defendants' due process rights to a fair trial.[50]

III. THE EXISTING LEGAL FRAMEWORK ON RACIAL BIAS IN FRT

A. Existing Legal Standards

The constitutional concerns raised by local law enforcement's use of FRT are further complicated by the absence of clearly defined national regulatory guidelines. According to the US Commission on Civil Rights, there are currently no federal laws or regulations that explicitly authorize or limit FRT use by the federal government.[51] As a result, there is a lack of unified federal oversight, leaving the management of FRT up to the discretion of city and state authorities. Some states, like California, have taken measures to govern the use of FRT by police departments. In October 2019, California enacted a 3-year moratorium on the use of FRT in police body cameras, a

[47] Joel R. McConvey, *Police Use FRT in Exactly the Ways Critics Fear: Washington Post*, Biometric Update, Oct. 7, 2024, https://www.biometricupdate.com/202410/police-use-frt-in-exactly-the-ways-critics-fear-washington-post.

[48] *Id.*

[49] Benedict, *supra* note 31, at 859.

[50] National Association of Criminal Defense Lawyers, *Defense Use of Facial Recognition Technology: Advisory*, https://www.nacdl.org/getattachment/94f8d9f8-eba4-44f7-8763-47a2f3038e8a/defenseusefacialrecognitionadvisory.pdf.

[51] U.S. Commission on Civil Rights., The Civil Rights Implications of the Federal Use of Facial Recognition Technology 99 (Sept. 2024).

law that took effect in January 2020.[52] In 2023, California Assemblymember Philip Ting introduced AB 642, a bill mandating that law enforcement agencies design regulations for FRT, publish annual reports, and ensure systematic deletions of certain photos every six months (e.g., those of individuals not charged with crimes or minors).[53] The purpose was to prevent the indefinite retention of personal images, which could lead to unnecessary privacy violations or surveillance overreach. AB 642 would permit the use of FRT in specific cases with a NIST certification that the system met a minimum accuracy rate.[54] Although AB 642 would allow the use of FRT only when strict accuracy standards are met, there were widespread concerns about the ethics of allowing a racially biased system to be used at all.[55] These groups argued that this bill further marginalizes certain demographics, particularly those already vulnerable to being targeted by the criminal justice system. Due to these concerns, AB 642 did not pass, although the proposal raised important concerns that should be considered in future regulatory legislation. Furthermore, cities have enacted measures, specifically in the form of bans, to address the concerns associated with police use of FRT. In May 2019, San Francisco became the first city in the United States to ban the use of FRT by city agencies, including the police department.[56] Shortly after, in June 2019, the City Council of Somerville, Massachusetts, voted unanimously to implement a similar ban on FRT use by city agencies, including law enforcement.[57] In July 2019, the City Council of Oakland, California, followed suit with a unanimous vote to prohibit the use of FRT by city agencies, including the police department.[58]

B. Relevant Case Law

Several legal cases have highlighted significant inaccuracies of FRT, revealing how its flawed algorithms often lead to false identifications. These errors have disproportionately affected marginalized communities, with instances of wrongful arrests and detentions. These cases underscore the potential dangers of overrelying on

---

[52] *State Facial Recognition Policy*, Electronic Privacy Information Center, https://epic.org/state-facial-recognition-policy/ (last visited Feb. 8, 2025).
[53] C.A. AB642, 2023 Assem. Comm., No. 642 (Ca.2023).
[54] *Id.*
[55] *Id*.
[56] *State Facial Recognition Policy*, *supra* note 52.
[57] *Id.*
[58] *Id.*

FRT in law enforcement, especially when the technology fails to accurately recognize individuals from diverse racial and ethnic backgrounds.

In *Parks v. McCormac*, an African American man named Nijeer Parks was wrongfully arrested in a shoplifting case in 2019.[59] The Woodbridge police department acted on an unreliable facial recognition match and failed to conduct a proper investigation before wrongfully arresting Parks. As a result, Parks was detained for ten days, although he was nowhere near the crime scene at the time of the incident. In Parks' situation, when police officers attempted to identify a suspect, they uploaded a blurry image from a fake driver's license that the suspected shoplifter presented. Subsequently, the FRT system flagged Parks' old arrest photo as a possible match. However, when a Woodbridge sergeant sought an arrest warrant for Parks, he asserted that the person in the blurry image of the fake driver's license was the "same person" as the individual in Parks' old arrest photo.[60] Instead of cross-referencing this potentially unreliable FRT match, due to the shoplifter's blurry *and* fake license, the sergeant did not take the necessary steps to verify the individual's identity. Moreover, the sergeant demonstrated confidence in the FRT match despite its unreliability, wrongly accusing Parks of being the suspect. As seen in this case, the sergeant failed to establish the probable cause required by law, leading to the case being classified as malicious prosecution – a legal claim where an individual is wrongfully subjected to legal action without proper grounds or evidence and with an improper motive.

In the case *Oliver v. City of Detroit*, a lawsuit was filed in October 2020 against Detroit Police Officer Busa and others for failing to properly investigate a larceny case. The officers relied exclusively on faulty FRT identification, which led to an inaccurate identification of the suspect.[61] This resulted in the wrongful arrest and detention of Michael Oliver, an African American man, who was jailed for three days before being released on account of misidentification.[62] Oliver's attorney argued that his constitutional rights were violated under 42 US Code §1983, which allows individuals

[59] *Parks v. McCormac*, No. 2:21-cv-04021 (D.N.J. filed Mar 3, 2021).
[60] Dana DiFilippo, *Lawsuit Seen as Crucial Test of Police Use of Facial Recognition Technology*, N.J. Monitor, Feb. 1, 2024, https://newjerseymonitor.com/2024/02/01/lawsuit-seen-as-crucial-test-of-police-use-of-facial-recognition-technology/.
[61] *Oliver v. Bussa,* No. 2:20-cv-12711 (E.D. Mich. filed Oct. 20, 2020).
[62] *Incident 439: Detroit Police Wrongfully Arrested Black Man Due to Faulty Facial Recognition*, AI Incident Database, https://incidentdatabase.ai/cite/439/ (last visited Mar. 4, 2025).

to seek redress for violations of their rights by government officials, including due process infringement, wrongful seizure, and racial discrimination.[63]

Furthermore, in *Williams v. City of Detroit*, an African American man named Robert Williams was wrongfully arrested in Detroit as a result of a false FRT match.[64] Despite that Williams was nowhere near the crime scene, his photo was included in a lineup to match the blurry surveillance image of the suspect. He was misidentified as the suspect and subsequently arrested. The American Civil Liberties Union filed a civil rights lawsuit on Williams' behalf in April 2021, arguing that the Detroit police lacked the necessary policies and training for the proper use of FRT, leading to his wrongful arrest. This case ultimately reached a landmark settlement in June 2024, prompting the Detroit police to implement stricter policies on the training and use of FRT. These new policies include barring officers from using FRT match results as the sole basis for witness identification procedures and preventing arrest warrants from being issued based solely on FRT match outcomes.[65]

Lastly, in *Woodruff v. City of Detroit*, Porcha Woodruff, an 8-month pregnant African American woman, was wrongfully arrested and detained for 11 hours by six Detroit Police Department officers in February 2022.[66] Her arrest was based on the resemblance between her 2015 mugshot and an FRT image of a suspect wanted for robbery and carjacking. During her detention, Woodruff endured intense pain and contractions, demonstrating a clear lack of care and consideration from the police who detained her. The most troubling aspect of the case was that she was unnecessarily subjected to pain and distress due to faulty FRT output. Although she was initially charged and held on a $100,000 bond, all charges were dropped a month later due to a lack of evidence linking her to the crime. In response to this incident, Woodruff has filed a lawsuit against the City of Detroit for wrongful arrest, which is still ongoing as of February 2025.[67] She is the sixth known African American to be falsely accused

[63] 42 U.S. Code § 1983.

[64] *Williams v. City of Detroit*, No .22-1344 (6th Cir. Dec. 2, 2022).

[65] Paige Gross, *Facial Recognition in Policing Is Getting State-by-State Guardrails,* The Highland County Press, Feb. 20, 2025,
https://highlandcountypress.com/headlines/facial-recognition-policing-getting-state-state-guardrails#gsc.tab=0.

[66] *Woodruff v. City of Detroit, City of*, No. 5:23-cv-11886 (E.D. Mich. filed, Aug. 30, 2025).

[67] George Hunter, *Detroit Police Department Falsely Arrested Woman After Faulty Facial Recognition Hit Lawsuit*, The Detroit News, Feb. 24, 2025,
https://www.detroitnews.com/story/news/local/detroit-city/2025/02/24/detroit-police-department-falsely-arrested-woman-after-faulty-facial-recognition-hit-lawsuit/80004419007/.

based on an inaccurate FRT image and the first woman to take legal action in such a case.[68]

IV. REFORM RECOMMENDATIONS: TECHNOLOGICAL AND LEGAL APPROACHES

Given the widespread use of FRT by local law enforcement and its disproportionate impact on people of color, it is essential to address this issue from various perspectives: reforming the technology itself and ensuring accountability for its use by law enforcement agencies.

A. Technological Reform: Improving FRT Algorithms

First, it is essential to address the root cause of the issue: FRT systems are built and trained using non-diverse datasets that are primarily comprised of white faces. This results in a significant disparity in accuracy, with the technology performing more accurately for white individuals than for people of color. Therefore, to eliminate this discrepancy, it is important to implement a new approach that diversifies the faceprints used to train FRT systems. In 2024, the University of Southern California emerged as a leading pioneer in the diversification effort, with computer science researchers introducing an innovative method called quality-diversity (QD) algorithms.[69] QD algorithms are designed to explore a wide range of high-performing solutions to a given problem.[70] These algorithms have proven effective in addressing gaps in real-world data, particularly in areas with underrepresented skin tones, intersectional groups, or genders. This algorithmic method is significantly more streamlined than traditional methods, having the capability to generate a dataset of 50,000 images within 17 hours. What makes this new algorithmic approach particularly intriguing is its potential to not only train entirely new FRT systems but also re-train existing ones. For example, the QD algorithm can generate a diverse and more balanced dataset from the outset when training new FRT systems, having the capacity to tackle bias at its root. Additionally, it can re-train existing systems by using its capacity to generate synthetic data that fills in diversity gaps and balances content from different demographics,

[68] Kashmir Hill, *Eight Months Pregnant and Arrested After False Facial Recognition Match*, N.Y. Times, Aug. 6, 2023, https://www.nytimes.com/2023/08/06/business/facial-recognition-false-arrest.html.
[69] Caitlin Dawson & Anna Hsu, *Diversifying Data to Beat Bias in AI,* USC Viterbi School of Engineering, Feb. 21, 2024, https://viterbischool.usc.edu/news/2024/02/diversifying-data-to-beat-bias/.
[70] Jean-Baptiste Mouret, *Quality Diversity Algorithms,* Inria, https://members.loria.fr/jbmouret/qd.html (last visited Feb. 6, 2025).

effectively "repairing" biased classifiers. This results in a reduction of the bias originally present in the FRT dataset, enhancing the fairness and accuracy of the FRT's classifications and improving its ability to identify people of color.

The researchers began by using Quality-Diversity Generative Sampling (QDGS) with StyleGAN2, an advanced deep-learning model designed to generate highly realistic images. Their goal was to create synthetic facial image datasets containing 15,000 to 50,000 images to enhance diversity in skin tones, genders, ages, and other factors.[71] They also created two random datasets to serve as comparisons to the facial images generated by the StyleGAN2 model, with these datasets sampled from a wide range of variables that influence their appearance.[72] Next, to ensure the images were diverse, they employed an objective function that upheld three key goals: (1) a text description to verify the images met specific quality standards (e.g., ensuring diversity), (2) an image margin to encourage variety in the photos, and (3) a regularization factor to keep the images within reasonable limits. After the images were generated, the researchers employed a method known as K-means clustering, which grouped similar images. Subsequently, they created several variations of each image, including altering facial expressions and rotating the faces. After the training phase, the model was tested on numerous datasets to examine whether it could recognize faces from all demographics and skin tones. The ultimate goal was to enhance FRT systems such that they recognize a diverse range of facial features, particularly for individuals with darker skin tones who have historically been underrepresented in traditional datasets.

The results of the study were promising. The researchers were able to increase the proportion of images in the database with dark skin tones from 9.4 percent to 25.2 percent. Additionally, for models trained on biased datasets, QDGS increased the recognition accuracy for dark-skinned faces from 88.08 percent to 88.94 percent. While this 0.86 percent improvement may not seem numerically significant, it represents a meaningful step toward gradually narrowing the performance gap between light-skinned and dark-skinned faces. On the other hand, this 0.86 percent improvement indicates that while progress is being made in recognition accuracy for people of color, more research and effort are needed to achieve more substantial quantitative improvements. Nevertheless, based on this data, the researchers concluded that QDGS achieved the highest average accuracy across facial recognition

[71] Allen Chang et al., *Quality-Diversity Generative Sampling for Learning with Synthetic Data*, Cornell University, Feb. 27, 2024, https://arxiv.org/abs/2312.14369.
[72] *Id.*

benchmarks, supporting the use of balanced synthetic datasets to enhance model performance.

While QDGS holds potential as one of the methods to eliminate racial bias, it is important to be mindful of its limitations. One key constraint of QDGS is that its effectiveness relies significantly on the quality of the user-defined language prompts. Therefore, if these prompts are not precise or sufficiently detailed, they may unintentionally perpetuate existing biases or miss important aspects of diversity. Another barrier to widespread implementation is that generating large-scale, balanced synthetic datasets with QDGS can be resource-intensive and costly, especially as the complexity of the desired attributes and diversity increases. This can make it challenging to implement on large datasets or in real-time applications. However, rather than viewing these challenges as a reason to abandon QDGS, it is important to recognize that they simply highlight the need for further research to make the algorithm more practical for large-scale implementation.

B. Legal Reform: Implementing the Accountability, Training, and Auditing (ATA) Framework

In addition to improving the technology itself, it is equally important to ensure that FRT is used responsibly by law enforcement. It is not enough to simply reform the technology; there must also be advocacy for its responsible use and reliance. As highlighted throughout this paper, FRT should never be the sole resource relied upon by police for investigations. Even after the technology is reformed to improve accuracy in identifying darker skin tones, users must continue to exercise caution to prevent automation bias. Therefore, I recommend implementing a three-pronged framework focused on legal reform: Accountability, Training, and Auditing (ATA).

The first prong focuses on accountability, ensuring that algorithm developers and police departments are held responsible for their creation and implementation of FRT technology. A major issue with FRT use by police is the absence of clear, stringent accountability frameworks, leading to misuse and ineffective application. Establishing a legal framework will provide guidelines and expectations for developers creating FRT algorithms, as well as for police departments and officers using the technology in investigations. To achieve this, the framework must first require developers to meet specific statistical thresholds, demonstrating that (1) the training datasets are balanced across demographics, and (2) the technology achieves a high overall accuracy rate when deployed in public settings.

To begin, the FRT model must be trained with a balanced dataset containing a wide range of images, with an equal representation of white people and people of color.[73] Ideally, the dataset should be balanced with equal representation of white people and people of color (i.e., 50/50), though perfect equality is difficult to achieve. An imbalance of up to 10 percent (i.e., 55/45) may still be acceptable, as it still ensures the dataset remains sufficiently diverse for the model to be trained to accurately generalize between the two groups. However, as the imbalance increases, such as a 20 percent imbalance (i.e., 60/40), the model may become biased, prioritizing the classification of the majority group over the minority group. In other words, the greater the imbalance in the dataset, the more the system will favor learning from the majority group. For instance, if there are more images of white individuals than people of color, the model will have more exposure to white features, leading to better performance in white identifications and increased error rates for underrepresented classes. It is important to recognize that the ideal 50/50 dataset split is *not* intended to represent actual US population demographics. Rather, it serves as a corrective mechanism to ensure optimal model performance. The "overrepresentation" of people of color in training data (i.e., while people of color make up 39.9 percent of the US population as of 2019,[74] they represent 50 percent of the dataset) is necessary to combat the accuracy discrepancies within the classes in FRT. This approach will not create reverse bias, an issue where facial recognition systems perform worse for white individuals due to the "overrepresentation" of people of color in the training data. Rather, it aims to equalize performance across groups by addressing the disparities where FRT has historically underperformed on marginalized populations. Additionally, the term "people of color" includes a wide range of ethnic and racial groups, each with unique facial features. A truly balanced dataset must reflect this diversity, not just between white individuals and people of color. A current challenge in data curation is sourcing large quantities of representative data, particularly when utilizing institutional datasets that may be biased. New algorithms, like QDGS, present promising solutions that allow for algorithmic correction of sourcing biases by generating balanced synthetic data using

[73] *Introduction to Balanced and Imbalanced Datasets in Machine Learning*, Encord, Nov. 11, 2022, https://encord.com/blog/an-introduction-to-balanced-and-imbalanced-datasets-in-machine-learning/.
[74] William H. Frey, *New Census Data Shows the Nation Is Diversifying Even Faster Than Predicted,* Brookings, Jul. 1, 2020, https://www.brookings.edu/articles/new-census-data-shows-the-nation-is-diversifying-even-faster-than-predicted/.

prompt-guided sampling.[75] For example, rather than generating a mix of dark-skinned, feminine, and older faces, QDGS produces the creation of specific images, such as a "realistic, detailed image of a dark-skinned, feminine individual in their 60s with short hair." Applying this process across a variety of attribute combinations to create balanced, diverse, and unbiased datasets, tools like QDGS demonstrate that fairness in FRT can be designed through intentional dataset construction.

Next, the framework must also require the technology to achieve a high accuracy rate in public settings, with overall accuracy exceeding 95 percent. Additionally, the true positive rate for people of color should match that of white people, while minimizing the true negative rates for both groups. This means that the FRT system must prove to be equally accurate for people of color and white people, both in terms of identifying individuals correctly (i.e., true positives) and avoiding false classifications (i.e., true negatives). Thus, if FRT systems are expected to meet the standard of having an overall accuracy above 95 percent and demonstrate equal accuracy for both groups, it will effectively rule out racial bias as a factor in FRT inaccuracies. Any disparities in error rates would instead be attributed to external factors, such as uneven lighting or poor image quality.

To develop algorithms that meet these accuracy thresholds, substantial effort must go into the data collection process, which is crucial for training the models. However, data collection is both expensive and complex, as it requires a large, diverse dataset with millions of images from various settings, including different times of day, environments, and angles. Therefore, collecting sufficient data to properly train FRT algorithms with balanced datasets is challenging, as it requires substantial resources to either purchase data from large companies, such as Clearview AI, or manually gather photos from the internet. The deliberate collection of diverse photos– capturing a wide range of facial features across different demographic groups– forms the foundation for developing a racially balanced FRT algorithm. While progress is being made to determine the viability of this process, it has not yet been fully realized. Further research is needed to explore the practicality of large-scale data collection and make the process more accessible. Nonetheless, as the data collection process becomes more widely implementable and financially feasible, it will be crucial to create accuracy standards for algorithm developers. If developers cannot demonstrate that their algorithms meet these requirements, they should not be legally permitted to release this technology for use.

[75] Chang et al., *supra* note 71.

The second prong emphasizes training, making sure that FRT is used responsibly and professionally by police departments. Establishing a robust and transparent framework that holds police departments and officers accountable for any improper use of this technology is critical.[76] A legal framework must be established that requires police departments to provide in-depth training for officers on the proper use of FRT. To allow flexibility in how departments manage their training programs, a general oversight requirement should mandate that all officers complete training and pass exams on the proper use of FRT within a specified time frame – ideally within 6 to 8 months of the legal framework's enactment. This approach gives departments the flexibility to determine the pace and structure of the training, while also making it clear that if they wish to use this technology, they must comply with the framework's training completion timeline. Upon completion of the training, officers must be required to demonstrate an adequate understanding of the material by passing an exam that tests their knowledge of the content. If departments fail to meet these requirements within the given time frame, they should be prohibited from accessing the technology.

Another key element of holding police departments accountable is the establishment of clear and enforceable guidelines that outline when and how FRT can be used and how outputs (i.e., image matches) should be applied. This framework must establish stringent criteria for when FRT is appropriate, such as in high-risk, time-sensitive cases like criminal investigations, missing persons cases, or imminent threats. Conversely, the framework should discourage, or even prohibit, the use of FRT in routine surveillance or minor incidents that do not warrant such technology, reserving its application for more urgent or complex situations. However, given the widespread use of FRT across police departments, it may be challenging to eliminate its use in routine instances, as many departments have become increasingly reliant on the technology for various operations. Beyond establishing *when* the technology should be used, the framework must set clear guidelines on *how* it can be used. Specifically, it must ensure that FRT outputs are never relied upon as the sole or deciding factor in confirming a suspect's identity. FRT should be considered a supplementary tool that assists in investigations, but should not replace officers' professional responsibility to

[76] Courtney Levin, *Your Guide to Developing a Facial Recognition Technology Program,* Police1, Nov. 21, 2024,
https://www.police1.com/police-products/police-technology/police-software/facial-recognition/your-guide-to-developing-a-facial-recognition-technology-program.

follow up on the technology's results. Police must be required to use FRT image matches in conjunction with other forms of evidence, such as alibi verification, forensic evidence, eyewitness testimony, and other methods. This approach will help mitigate the overreliance on FRT, which has contributed to issues stemming from the technology's inherent flaws. Although FRT technology is expected to advance, it is still crucial that police remain accountable for its use. Technology will never be entirely free of flaws, and human oversight will always be necessary. Furthermore, there must be significant penalties for non-compliance. If police departments fail to follow the guidelines for proper use and reliance on FRT, they must face legal actions, fines, disciplinary measures, or even legally binding restrictions on their future access to the technology. If officers violate these guidelines, particularly if they are repeat offenders, they should face disciplinary consequences, such as mandatory retraining, suspension, or even termination. The primary purpose of these penalties is to serve as a deterrent, ensuring that both departments and individual officers understand the consequences of violating the accountability framework in their use of this technology.

The third and final prong highlights the importance of auditing as a key mechanism for ensuring police compliance with guidelines related to the use of FRT.[77] To ensure impartiality and neutrality in the auditing process, audits should be conducted by independent third-party organizations–such as specialized firms or institutions with expertise in FRT–that have no vested interest in the outcomes. Contracting external, non-governmental auditors enhances transparency, builds public trust, and ensures that the audits are not influenced by political agendas or policy interests. The role of these independent auditors is to conduct comprehensive assessments of how police departments are using FRT, including the instances in which the technology has been applied, the accuracy of its outputs, and how these outputs are utilized in investigations. Ultimately, the primary goal of the auditors is to determine whether FRT has a disproportionate impact on certain racial communities. To achieve this, the auditing process must incorporate several key elements. First, auditors must assess whether FRT is being used exclusively in cases where it is authorized and following established guidelines. The technology must not be used excessively or inappropriately in situations where it is unnecessary, such as routine surveillance. However, this is contingent on the guidelines outlined in earlier prongs, particularly the second prong, which emphasizes the need for clear criteria in determining when the technology should be used. Next, auditors must ensure that accuracy standards are

[77] Allan Y. Jiao, *The Landscape of Police Auditing in the USA: The State of the Art*, 18 Oxford Academic 113 (Nov. 15, 2024), https://doi.org/10.1093/police/paae113.

met, with a focus on maximizing and balancing true positives for both people of color and white individuals, while minimizing true negatives for both groups. Moreover, once the audit is completed, the findings must be made accessible as public official records. These findings should be presented clearly and understandably to citizens, oversight bodies, and policymakers. By doing so, auditors can promote transparency and foster trust between police departments and the communities they serve. Furthermore, after completing the audit, auditors should offer practical recommendations to resolve any identified issues. These suggestions might involve updating policies or practices to prevent misuse or recommending enhancements to training programs if they are deemed insufficient in ensuring the proper and responsible use of FRT by officers. Most importantly, audits should not be a one-time occurrence but rather a periodic and ongoing process. Regular audits must be conducted to ensure police departments remain in compliance with evolving standards. These audits should also assess the effectiveness of any corrective actions implemented following previous audits and recommendations. Additionally, if an audit uncovers serious violations of the guidelines or a failure by police departments to make required corrective changes, clear and enforceable penalties must be imposed on both the departments and individual officers. Potential penalties could include fines, restrictions on FRT access, or even suspension of officers found in breach of the guidelines. However, a key limitation is that the effectiveness of auditing hinges on the consistency of auditing standards across auditors, institutions, and jurisdictions. Additionally, it is important to recognize that auditors do not possess legally binding authority to enforce actions. While they can offer recommendations and identify areas for improvement, the responsibility for enforcing and implementing these changes ultimately rests on the willingness of police departments to take action. Therefore, if auditors discover misconduct or misuse of FRT, or if departments refuse to take corrective actions, a government entity with enforcement authority must step in to ensure compliance and, if needed, impose penalties for non-compliance.

Ultimately, implementing the ATA framework is a critical step toward ensuring the ethical and professional use of FRT by law enforcement. By establishing clear guidelines, providing training, and conducting audits to ensure accountability, both developers and police departments can be held responsible for creating and using the technology, respectively. While further research is needed, this framework presents a comprehensive approach to creating a legal and operational environment where FRT can be used by law enforcement without infringing on individuals' rights or exacerbating systemic biases.

CONCLUSION

The use of racially biased FRT by law enforcement raises significant legal and ethical concerns, particularly regarding its impact on people of color. The inherent biases in FRT systems, resulting from a lack of diverse training datasets, lead to disproportionate misidentifications of people of color and wrongful arrests. This flawed technology not only infringes upon individuals' constitutional rights– violating the Fourth Amendment's protection against unreasonable searches and seizures and the Fourteenth Amendment's guarantee of equal protection under the law– but also exacerbates the systemic inequalities already entrenched in the criminal justice system. In addition to the flaws within the technology itself, the overreliance on FRT by police, especially when there is insufficient follow-up verification of its outputs, increases the risk of algorithmic injustice and deepens racial disparities in law enforcement.

Both technological and legal-related issues underscore the urgency for reform on both fronts, not only to improve FRT algorithms but also to ensure proper usage of the technology. Therefore, it is essential to diversify training datasets through the implementation of methods such as QDGS algorithms, which can help reduce racial biases and improve accuracy in identifying individuals across various racial and ethnic backgrounds. However, technological reform alone is inadequate to address this complex issue. To ensure the ethical and responsible use of FRT, particularly by law enforcement, it is crucial to adopt the AIA framework, which sets clear guidelines for holding both algorithm developers and police departments accountable, ensures proper training for users, and mandates periodic audits to verify compliance with usage standards. Together, these reforms can significantly contribute to a more equitable and just application of FRT, ensuring it is applied fairly to all individuals, regardless of skin color. Moreover, incorporating diverse voices in the development of FRT ensures that the technology aligns with the values and needs of all communities, with experts dedicated to promoting justice, reducing bias, and advancing equity. With ongoing research into the implementation and large-scale adoption of potential solutions, addressing both technological and legal issues will be key in resolving the legal implications of racially biased FRT in law enforcement.