presents

Facial Recognition Technology in Modern Policing

Recommendations and Considerations

2021 Facial Recognition Working Group

Table of Contents

Introduction

The 21st century offers law enforcement an unprecedented opportunity to embrace advanced technologies to keep our communities safe. One of the most valuable of these technologies is facial recognition technology (FRT). FRT has an unprecedented ability to combat criminal activity, identify persons of interest, develop actionable leads, and close cases faster than ever before. Perhaps most importantly, the law enforcement agencies which have embraced this technology have proven its capability of assisting with the ultimate goal of keeping our communities safe.

It is this vision of protecting our communities and preventing future crime that fuels the desire of law enforcement to develop a responsible, appropriate, and effective FRT program. Technology has an ever-increasing impact on our lives. As such, it is critical that law enforcement also have access to and develop programs that leverage these advanced technologies to combat the criminal element. Even more importantly, it is important to recognize the use of such technologies comes with great responsibility.

This product's intent is to assist law enforcement agencies in developing an FRT program that is both effective in its use and responsible in its design. The core principles of such a program are transparency, accountability, and responsibility. Transparency is achieved when the public knows FRT is in use by its police agencies and how it is (and isn't) utilized. Accountability is achieved when the public is aware of its results and impact. The framework outlined in this document will help law enforcement meet the highest standards and achieve its mission of protecting the community while remaining respectful of the individual rights and privacy of the citizens within their respective communities.

There exists great promise for facial recognition technology. This is the primary reason law enforcement is rapidly developing programs that embrace it. Because of the movement towards incorporating the technology into public safety, law enforcement has recognized the need for accountability to the general public, standardized application, and the responsible continued development of policy and practice.

It is important to use this product as a guideline for current best practices in the field. Not every element is necessary for developing a robust program but should be considered to make a program as complete and fitting to the needs of a community as possible. What may work in one community, may not be necessary for another, but this working group recommends considering all options before finalizing an agency's FRT protocols and policies.

To achieve this goal, the Major Cities Chiefs Association (MCCA) assembled representatives of facial recognition programs and agencies nationwide to produce this document. With support from FRT vendors and law enforcement alike, this product is a true partnership. This product represents the working group's recommendations for the procurement, development, operation, and reporting of a well-balanced FRT program.

Much like other powerful technologies, FRT will continue to evolve. This product is intended to provide law enforcement agencies a framework to adapt policies and practices with the evolving technology and will be revisited and updated as needed to appropriately reflect changes in the policing environment.

Executive Summary

In Spring of 2021, the Major Cities Chiefs Association (MCCA) launched an effort to develop a facial recognition technology (FRT) program development guide. This product is intended to be shared with the MCCA's membership to leverage the advanced technology in their crime-fighting efforts. This technology has proven itself to be a powerful tool to combat criminal activity, identify persons of interest, develop actionable leads, and close cases faster.

As more agencies launch FRT programs or acquire software which has FRT capabilities, standard best practices must be discussed regularly in order to ensure this technology is being used with the best intentions which includes the protection of the privacy and civil liberties of citizens. This product discusses many areas of FRT including FRT basics, methodology, program design, program management, qualitative review, and technical evaluation. This document was formulated to discuss all aspects of FRT, yet leaves room for an agency to develop a program which meets its specific needs. Should an agency choose not to utilize all suggestions discussed, it was the intent to also discuss the potential ramifications of omitting certain aspects of FRT best practices.

Stakeholders concerned about FRT generally focus on the following: algorithmic accuracy and bias concerns, threats to privacy (surveillance), chilling effects on first amendment rights, and defendants' rights, or in other words the general lack of disclosure of the use of the technology in their case. It would be irresponsible on the part of law enforcement to ignore these concerns. However, it is our position that all of these can be sufficiently mitigated or eliminated through thoughtful policy, intentional protocols, and responsible program management. While there are documented misuses of FRT results, these would have been avoided had best practices been employed. These cases should serve as a cautionary reference for those wishing to run a successful program. It is important to recognize the overwrought rhetoric and misinformation which is often promoted.

To this end, this report focuses on the building of an FRT program keen on transparency, responsibility, and accountability. Following the recommendations outlined in this product will enhance an agency's ability to straightforwardly address the concerns listed above. Not only does this product serve as a template to design and operate a working FRT program but also serves as a guide to an operational FRT program that is defensible by nature while responsible in its use.

Finally, it should be noted that just like technology itself, the standards surrounding the use of FRT are ever-changing. For this reason, this should be considered more of a living document rather than a finalized product. As the conversation surrounding FRT progresses, so will the guidelines detailed in this product.

Key Recommendations

The widely varying size and scope of MCCA member agencies necessarily requires the key recommendations in this document to be broad in their scope and applicability. Much effort was given to making both the content as a whole and the recommendations below as relevant as possible to all agencies. In general terms, it is the view of the FRT working group that each of the following key recommendations be implemented with the launch of a new FRT program. However, these recommendations are not provided as bright-line requirements for the implementation of FRT at the reader's agency, rather they are meant to serve as important guide posts in any agency's development of a responsible FRT program.

Transparency

- Law enforcement agencies seeking to procure FRT platforms should engage both public and government stakeholders for the purposes of feedback and transparency.
- The documented results of an FRT investigation should be made subject to discovery in the criminal process.
- The eventual outcome of any criminal investigation that utilizes FRT should be captured as part of the agency's data collection process.

Accountability

- Access to an agency's FRT platform should be limited to those members having specialized training in facial identification methods and the application of the technology should be performed by individuals who are not directly involved with a particular investigation.
- Restricting access to an agency's FRT platform to only those members with specialized training in facial identification methods will reduce contextual bias in particular investigations.
- If possible, the application of FRT should be performed by trained individuals who are not directly involved with a particular investigation.
- The identification of a potential lead during an FRT investigation should be documented on a standardized form which requires sufficient detail about the morphological basis of the facial identification process.
- For those agencies wishing to implement the use of FRT, but who are unsure on how to move forward, collaboration with other entities who have already developed robust, responsible programs is recommended.

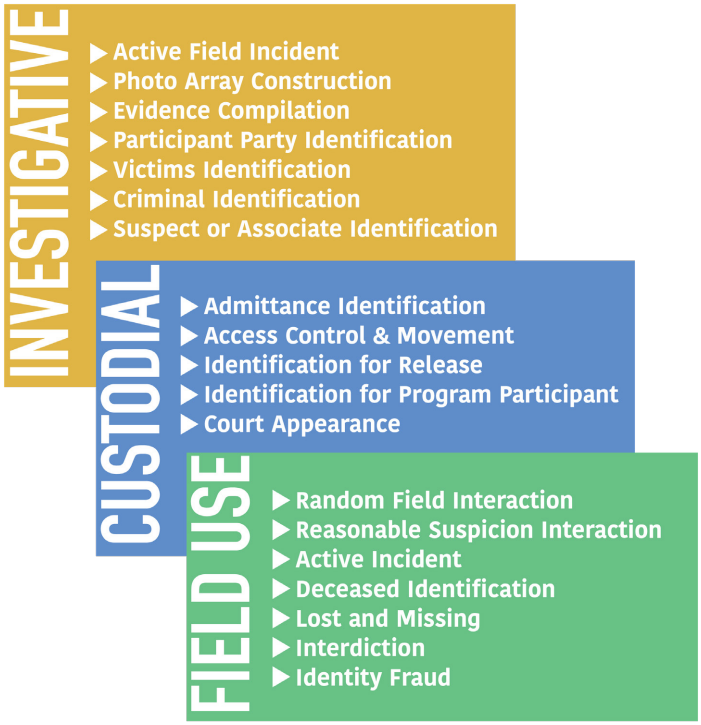
Responsibility

- Careful consideration should be given to the specific and most privacy-conscious approach to what gallery image library is used.
- Agencies should identify an FRT program manager who will be tasked with both the initial deployment and continued oversight and development of the FRT program.
- The results of FRT investigations should be handled as tips/leads only due to the limitations of FRT and the potential consequences of its misuse as outlined in this document.
- FRT examiner training should specifically include familiarization with standardized methods for performing facial identification.
- The initial findings of an FRT investigation should be confirmed by a secondary examiner.

What is Facial Recognition?

What type of program does this product address?

Before we address the recommendations and considerations outlined in this product, we must define the specific type of facial recognition system that is being referenced throughout the majority of this report. There exist three primary applications of FRT platforms: facial verification, field identification, and facial identification. All three applications of FRT serve a purpose and may play a role in law enforcement operations. However, it is the facial identification type of FRT that presently garners widespread public and government concerns. In an effort to lay the foundation of the rest of this document, we must explicitly state the similarities and differences between these three primary applications of FRT.



Facial Verification

This application of FRT employs the use of a computer FRT platform to explicitly confirm a subject's identity. This is the same mode of FRT deployed in many modern-day cellular devices. In law enforcement, this type of FRT can be useful in correctional facilities to grant access to secured areas, confirm inmate identity in a booking or release environment, or confirm identity at border crossings as a few examples.

Field Verification

Field verification is the use of FRT in the field for the purpose of identifying an individual during a live interaction. This mode of FRT is primarily used to attempt to "fill the gaps" in available information such as when a subject lacks formal issued identification or is uncooperative and refuses to give proper identification. This type of FRT can aid in confirming who a subject is claiming to be.

Facial Identification

Facial identification is the most common application of FRT used by law enforcement. It is a direct use of FRT during a law enforcement investigation in which digital imagery of an unknown subject is available. This imagery is uploaded to an FRT platform which has a gallery of previously enrolled images for the specific purpose of identifying unknown subjects. Generally speaking, this is also a two-step process involving a combination of both FRT platform and human involvement with defined roles for both steps of the process.

The first step is the processing of the image through the FRT platform which creates a biometric facial template from the enrolled image and compares it to its gallery of facial templates. The system then returns a list of the enrolled images whose biometric facial template is sufficiently similar to the

submitted probe image's template. The second part of the process, and arguably the most important, involves a trained facial recognition examiner conducting a detailed review of the imagery returned by the system through a manual process of examining the morphological similarities or differences with the unknown subject. The examiner ultimately produces a finding based on the existence of or lack of articulable similarities between a probe and gallery image.

In contrast to the above, there exist FRT platforms that do not require a facial recognition examiner to be involved in the process which simply produces automated findings. These types of systems fail to allow any input or evaluation from an examiner. Although there exist use cases for this type of FRT, this product's recommendations and considerations are not designed for that type of system. Rather, the content found in this product is specific to a two-step FRT program that employs human oversight in the process of facial identification.

Other Uses

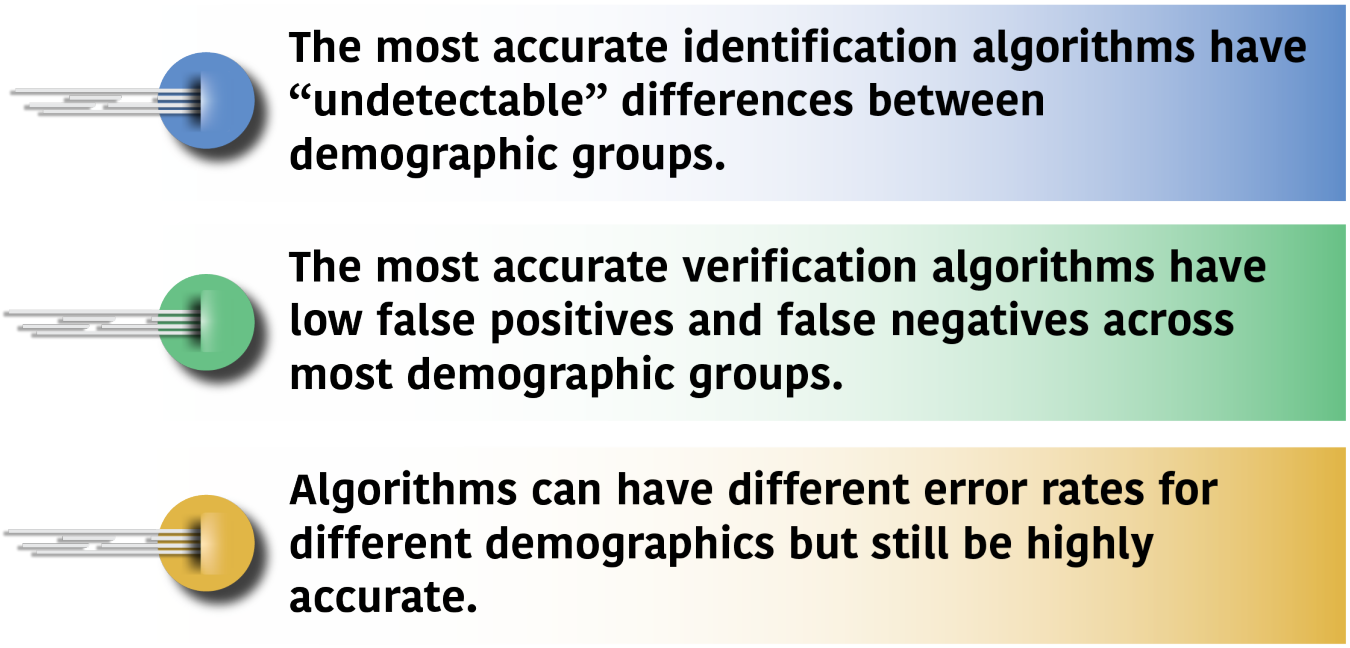
Finally, FRT platforms have the capability of being used as a surveillance tool by identifying persons in real-time using video feeds layered with FRT technology. Known instances of this type of use of FRT can be found in foreign nations and among certain private sector businesses. **MCCA best practices for law enforcement agencies utilizing FRT is to not utilize FRT in this manner except under the most exigent of circumstances or when explicitly permitted under legal authority (e.g., court order).**

Myths and Misconceptions

The above provides a quick snapshot of the real-world uses of facial recognition technology. In contrast, there is a general lack of publicly available information regarding the use of FRT by law enforcement. This has caused popular media to be a primary source for the general public's understanding of FRT. This situation is part of the impetus of this document. Reviewing the widespread myths and misconceptions about FRT in law enforcement is an integral part of any agency understanding the social context the technology currently sits within. At the end of this report in Appendix B, you will find a detailed exploration of many of the common misconceptions about FRT and supporting information to help in the process of educating stakeholders.

In an effort to provide some context on this, we offer the following example. There exists the perception that facial recognition algorithms are inherently biased and perform significantly less reliably among some demographic groups. However, a recent report that analyzed the findings on the National Institute of Standards and Technology (NIST) testing of FRT platforms revealed that: [1]

NIST FRT Algorithm Testing - Findings



1: McLaughlin, Michae; Castro , Michael (2021, September). *The Critics Were Wrong: NIST Data Shows the Best Facial Recognition Algorithms Are Neither Racist Nor Sexist.* ITIF. https://itif.org/publications/2020/01/27/critics-were-wrong-nist-data-shows-best-facial-recognition-algorithms

Methodology

The MCCA Facial Recognition in Modern Policing report serves as a comprehensive document containing recommendations and considerations for the use of FRT. This includes the processes, protocols, procedures, and responsibilities a law enforcement agency should embrace within its use of the valuable technology. This report provides recommended guidelines which are divided into three primary categories including technical considerations, program design, and program management.

Program Design

The Program Design section of this document is intended to provide decision-makers with important guidelines and a road map for designing a responsible FRT program that is both effective in its use, but also mindful of personal privacy and rights of individuals that the program may impact.

Chronological Process

This piece of the report is intended to provide the user an instructional guide on the steps to be taken as well as the correct order in the development of a facial recognition program. The suggestions made can empower a program manager to make responsible decisions in the most appropriate order as an FRT program is being built.

Responsible Procurement

This report will also help guide decision-makers in identifying the most well-respected FRT platforms available to law enforcement. A well-respected FRT platform will be both effective in the algorithmic testing, but also comprehensive in its overall design and supported by responsible ownership. This insight can both help save law enforcement agencies time in determining which FRT programs are available to law enforcement, and which are appropriate to use in terms of their algorithmic capabilities.

SME Training/Development

The third area focuses on a key component of an appropriate facial recognition program: training. Traditionally, law enforcement places great emphasis on the value of best practices training related to all aspects of public safety. Due to the increased concern regarding FRT, a well-designed examiner training and development program is a key piece of the responsible deployment of FRT. There is significant value in developing qualified and expert examiners. This expertise contributes to the quality of FRT investigative findings which ultimately result in increased public safety and confidence in our law enforcement agencies.

Program Management

The Program Management section of this document focuses on the operation and management of an agency's FRT program following the design and launch of the program. Even if an agency carefully considers and implements many of the recommendations in the Program Design portion of this document, each agency's community expectations, environmental challenges, and program management experience will almost certainly require specific adaptations of the recommendations outlined below.

Operational Workflow

The main area of focus in this section is the development of a workflow process related to facial recognition investigations. This piece highlights processes that can be adopted to ensure an organized investigative process which include the investigative request, first and secondary examiner steps, program oversight, and the distribution of findings. Most importantly, the suggested protocols provide a foundation for which an agency can collect comprehensive data points related to their facial recognition program for both operational oversight and reporting purposes.

Program Oversight

One of the essential elements of FRT program management is a carefully crafted oversight strategy. This area addresses program pieces that can be adopted that help both program managers and ultimately, high-level decision-makers evaluate overall FRT program status, value, and impact. It also helps provide the law enforcement agency with statistical reporting figures that allow its use and purpose to be shared with interested parties which may include but are not limited to legislative oversight groups, public oversight committees, and the media in general.

Operational Concerns

Policing agencies that implement an FRT program to improve operational efficiency and outcomes also need to be aware of the legitimate concerns that accompany such an implementation. The use of FRT, the utility of the technology, and the resulting outcomes are of great concern to the general public. A thorough understanding of the concerns of FRT is critical to a responsible program design. More importantly, a program must implement thorough policy and protocol processes that mitigate or eliminate the very real liabilities associated with FRT. This section will address the concerns of FRT and make clear recommendations on how those concerns can be alleviated and or even eliminated through responsible program design.

Technical Evaluation

An important topic covered in this report is a technical evaluation of facial recognition technology. As indicated in the opening of this report, it has become increasingly important to ensure technology adopted by law enforcement meets expectations. More simply stated, the adoption of any new technology requires knowing whether it works and whether it is worth adopting. Another critical question that needs to be answered regarding technical capabilities is if the technology works efficiently enough to ensure the data it provides law enforcement is relevant and accurate. This evaluation is essential in assuring law enforcement and the community alike that the quality of data being produced will contribute to law enforcement's primary goal of keeping our community safe. The technical evaluation in this document is divided into two subtopics.

Program Design

Facial recognition's inherent ability to help identify persons of interest assists law enforcement in solving cases and protecting our communities. However, there are also some real and perceived concerns regarding the use of this emerging technology, and ensuring proper oversight and privacy protection is among the most important. Police agencies have a duty to protect the privacy and civil rights of the community, just as much as it has the responsibility to enforce the law, apprehend criminals, and assist victims. To ensure both objectives are met, it is essential law enforcement agencies take careful steps to develop a facial recognition program that is effective, but appropriate. Taking a thoughtful approach to the application of this technology and design of the program allows for its use in a responsible manner is crucial.

To accomplish this goal, it is recommended that any facial recognition program be mindful of a few key points of interest for both law enforcement and the public alike. Those include proper policies, gallery image sources, controlled access, specialized training, and oversight, each of which are detailed below.

Proper Policies

Ensuring proper policies and procedures are in place for an agency's facial recognition program is critical. FRT policy should direct the purpose, general use, and processes involving facial recognition investigations. These policies will provide the framework by which an agency can simultaneously reap the benefits of the technology and be respectful of the individual privacy and civil rights of the public. Similarly, consistent policies will assist in protecting the integrity of criminal investigations, criminal intelligence, and justice system processes. They should also ensure all deployments of facial recognition will only be for official investigations. Policy and procedure elements should cover the following components:

- Community directed statement on why the agency is establishing an FRT capability
- The purpose of the policy
- General use of the technology
- Clearly defined program oversight roles and responsibilities
- The FRT vendor and system name/type
- Gallery sources
- Who is authorized to access the system
- Any required training for both examiners and officers
- The request and operational and investigative procedures
- The value an investigator should put on a positive FRT finding
- Clear guidance to investigators on being transparent in use and reporting in arrests reports
- Auditing responsibilities

Establishing clearly defined rules and processes on how FRT can be used, how investigations can be conducted, and the value of the findings of an investigation is the best way to minimize the risk of improper use of FRT. These policies should also be made available to the public when required by law or when appropriate.

Self-Restricted Data

Despite the laws in many states permitting the use of public driver's license photos, social media photos, or other publicly available photos for facial recognition gallery use, consideration should be given to the specific and most privacy-conscious approach to what library is used. This could include limiting the facial recognition gallery to photos of persons who have been previously arrested (i.e., mugshots). Although restrictive in nature, this ensures the probe images of suspected criminals are only being compared against images of persons who have previously been involved in criminal investigations in a specific geographic area and whose image has been obtained by law enforcement directly.

Some jurisdictions utilize DMV image databases. In those instances, it is supported by laws that govern the use of such images for FRT. It is also recommended community involvement be a part of that consideration. However, some concerns should be understood before embracing such a strategy. Images collected as a matter of practice in the arrest/booking process may be viewed differently than images in which the subject voluntarily submits to being photographed as a condition of operating a motor vehicle. Second, the inclusion of said galleries can impact the investigative process and place a greater burden on the human review step of the total process due to a larger gallery database.

Specialized Training

One of the primary reasons a two-part FRT examination process is so effective is because of the introduction and role of the trained human examiner. This human element is arguably the most important part of the total process[2]. The value of a technically trained examiner in the practice of facial identification and comparison cannot be overstated. Still, it is similarly important the investigator who is ultimately acting on any intelligence derived from a facial recognition investigation also be appropriately trained. The focus of the recommended training can be broken down into two parts.

"The value of a technically trained examiner in the practice of facial identification and comparison cannot be overstated."

First, it is critical that the examination part of the process is conducted by those who have been specially trained. To be considered an examiner or a subject matter expert in facial recognition technology it is recommended that the examiner complete the following types of training; Many agencies require their examiners to participate in the Federal Bureau of Investigation (FBI) Facial Identification training course.[3] This training focuses on understanding the foundations of facial components, stability of features over time, and other factors such as perspective distortion. It is widely accepted that a morphological analysis approach be used for facial comparison. A morphological analysis is a method of facial comparison in which the features and components of the face are compared. It is based on the evaluation of the correspondence among facial features such as the nose, mouth, ears, eyes, and facial lines and their respective component characteristics. Research has shown that examiners who are trained in the science of face comparison are better at recognizing individuals.

Similarly, it is equally critical that an examiner be trained on, and familiar with, the version of FRT

2: A Government Accountability Office report notes that the process and considerations for face identification are consistent with other forensic identification techniques, specifically latent prints and DNA. (See: Technology Assessment - Forensic Technology – Algorithms Strengthen Forensic Analysis, but Several Factors Can Affect Outcomes, GAO-21-435SP, US Government Accountability Office, July 2021).

3: FISWG, Facial comparison Overview and Methodology Guidelines version 1.0 2019.10.25

software that the agency uses. This training should be supported by vendors and include the operational processes, technical requirements, and a general understanding of how the system functions.

In addition to the initial onboarding training of skills and system familiarity, an examiner should become familiar with all processes and procedures in place for the FRT program. During the onboarding process, it is recommended that examiners in training shadow more experienced FRT examiners to ensure the application and understanding of the skills learned are being executed correctly. Additionally, their work product should be reviewed for accuracy and correct application of the received training. Finally, it is recommended that regular in-house training is completed for skills refreshment, general awareness of FRT case law, and awareness of trends and patterns.

Because examiners using an FRT system can, and often will be, requested to testify in both criminal and civil court cases, expert witness testimony training is also highly recommended. An expert witness is one with expertise in FRT and morphological comparison that far exceeds knowledge levels of a trier of fact. As an expert in FRT, it is important that an examiner be able to effectively communicate findings in a way that courts can best understand and learn strategies for effective direct and cross examination. Expert witness testimony training should also encompass topics that include a general understanding of the court process, roles of prosecutor, defense, jury, and judge and expert witness, as well as the qualification process for expert witnesses.

Additionally, comprehensive training should be provided to any investigator/officer who will be taking any action on the results of facial recognition investigations. Investigators should have a basic understanding of how the system operates, the role of the examiner, and applicable laws and policies that govern the use of FRT in a particular jurisdiction. These key agency members should be aware of the major capabilities and limitations of facial recognition technology broadly as well as the specific implementation of their agency's FRT program. **The limitations of FRT and the potential consequences of its misuse, as outlined in this document, are the main reasons that results of facial recognition examinations should only be used as a tip or lead by investigators.**

Investigators should also understand the importance of documenting the use of FRT technology in case reports as well as the need to disclose its use in applicable arrest documents and the discovery processes.

Limited Access

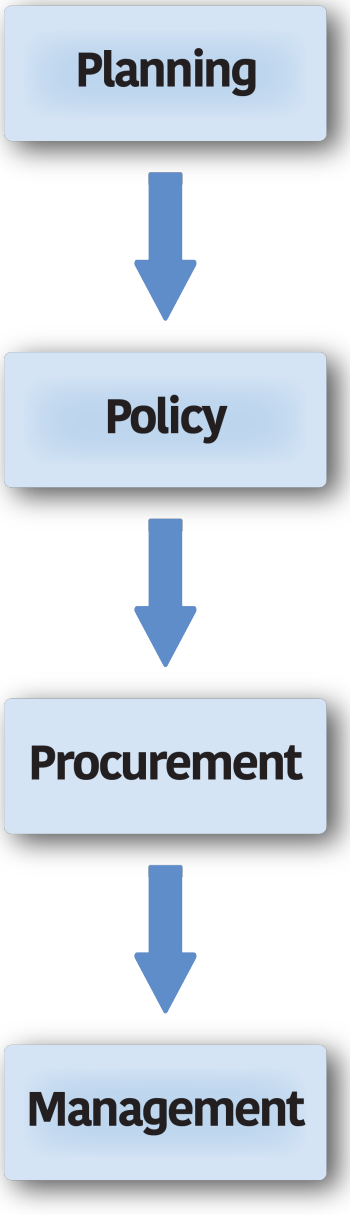
Numerous benefits are gained from the strategic design of centralizing the use and access of FRT to specific entities and restricting access to the technology to only those who have been specially trained in facial recognition. This strategy enhances trust in law enforcement and investigative processes as well as makes the implementation of the other components of a well-balanced program like management, reporting, and responsible use more easily achievable. **Having FRT examiners that are uninvolved with an investigation outside of the FRT component can help ensure FRT findings are consistent with program design and use.** Additionally, it can reduce the concern of investigators becoming overly reliant on the technology.

"Numerous benefits are gained from restricitng access to the technology to only those who have been specially trained in facial recognition."

Another benefit of a centralized program with limited size and access is that the implementation of best practices regarding FRT is more easily achieved. For example, should the vendor release a new version of FRT, only a small group of examiners would require the updated instruction versus training a large group of investigators. An additional benefit is that it streamlines the ability of an agency to collect valuable statistical use. With broad agency access and use, collecting important statistics becomes difficult if not impossible. Furthermore, the management of the program is simplified with a restricted deployment. Ultimately, ensuring processes, protocols, and policies are being strictly followed in a more tightly controlled environment is a goal to strive for.

Program Roadmap

Many aspects of program design have been covered in this section. The following bullet points are presented in a proposed phased order as a basic roadmap for the creation of a new FRT program.



- Establish FRT Goals/Expectations
- Identify Internal Stakeholders
- Engage Private Stakeholders for Feedback
- Identify Acceptable Gallery Sources
- Identify Acceptable Use Cases
- Develop Budget

- Develop Interagency Plan (MOU's) if Applicable
- Integrate Stakeholder Feedback
- Establish Internal Management Personnel
- Define Specific Usage Parameters and Restrictions
- Establish Oversight Measures and Requirements

- Engage with Multiple Vendors in Initial Procurement
- Engage with Client Agencies for Feedback
- Seek input from Citizen/Privacy Advocates
- After Selection of Vendor:
  - Develop Vendor-Specific Training
  - Engage All Stakeholders in "Soft Launch"
  - Monitor Initial Challenges if Any

- Refine/Develop Training, Scale-Up Program
- Finalize Reporting/Auditing Mechanisms
- Implement Continuous Assessment/Management
- Report Successes and Challenges to Stakeholders
- Engage with Partner Agencies Using FRT

Responsible Procurement

When procuring facial recognition technology (FRT) platforms or any other product with privacy and civil liberty implications, law enforcement agencies should seek stakeholder feedback and buy-in during the early stages of the procurement process. Because advanced technology such as FRT may delve into uncharted territory, it is not unreasonable for the public to question law enforcement's intentions and motivations. In January 2020, The New York Times published a piece under the headline, "The Secretive Company That Might End Privacy as We Know It."[4] The article (along with several others published contemporaneously) raised several valid concerns, including but not limited to the following: could rogue officers use the technology to identify and stalk an attractive stranger? Without policies prohibiting such conduct, could law enforcement use the technology to identify organizers of peaceful protests? Could programs that scrape the open web for images contribute to misidentifications? In the absence of model policies for novel or emerging technologies, news and opinion pieces can serve as a starting point in developing policies that address privacy and other concerns surrounding civil liberties. Soliciting stakeholder feedback and early involvement by citizens and privacy advocates at the beginning stages of an FRT program is also crucial.

"Law enforcement agencies should seek FRT software that delivers the most accruate results possible while also being developed by companies whose missions, visions, and values align with their own."

Law enforcement agencies should seek FRT software that delivers the most accurate results possible while also ensuring that the software is developed by companies whose missions, visions, and values align with their own.

Law enforcement agencies should conduct an appropriate degree of market research, which should include a trial of the product and conversations with existing customers to evaluate the product's performance in a real-world setting. While there is a certain degree of calculated risk involved in beginning a trial period before engaging stakeholders for their feedback in policy development and implementation, law enforcement agencies that adopt such an approach will have the ability to highlight initial successes with the technology thereby creating the opportunity to provide tangible examples to their stakeholders of how FRT serves to enhance public safety. Should an agency engage in a pilot program, it is vital to adopt a strict policy on how any information should be used in the course of investigations until a final protocol and all other areas of the program are formalized.

The importance of obtaining stakeholder feedback early in the procurement process cannot be understated. Program managers should make every effort to embrace a spirit of transparency, to accept critical feedback from stakeholders, and to attempt to address concerns concretely. It is sometimes possible to address specific concerns of the public by adopting specific measures in an agency's FRT program policy.

4: Kashmir Hill, "The Secretive Company That Might End Privacy as We Know It," New York Times, January 18, 2020, https://www.nytimes.com/2020/01/18/technology/clearview-privacy-facial-recognition.html

As an example of this type of suggested engagement with interested stakeholders, we offer an example; one major metropolitan police department met with representatives of a civil liberties advocacy organization in which some advocates voiced concerns. With the expressed concerns in mind, the agency incorporated the following into its FRT policy:

- A prohibition on the use of FRT to monitor any person(s) in "real-time"
- A requirement that all FRT leads for a particular case (not just the cases resulting in arrests) be included in official case files and be considered discoverable under applicable rules of evidence
- The inclusion of unambiguous language prohibiting arrests based solely on FRT leads
- An admonishment that end users be aware of the possible existence of algorithmic biases inherent in FRT platforms

Feedback from other stakeholders, such as an agency's procurement department, prosecutor's office, and civilian oversight board (if applicable) are also invaluable. It is also advised to engage with an agency's in-house legal department or corporation counsel on any parameters which need to be considered.

While local procurement ordinances vary across jurisdictions, it is imperative that those ordinances, along with broader guiding principles, are followed when procuring an FRT platform. The National Institute of Governmental Purchasing's (NIGP) Values and Guiding Principles of Public Procurement are one such set of broad principles which should be followed in any procurement process. The Guiding Principles are as follows:

- **Accountability:** Taking ownership and being responsible to all stakeholders for our actions. This value is essential to preserve public trust and protect the public interest.
- **Ethics:** Doing the right thing. This value is essential to deserve the public's trust.
- **Impartiality:** Unbiased decision-making and actions. This value is essential to ensure fairness for the public good.
- **Professionalism:** Upholding high standards of job performance and ethical behavior. This value is essential to balance diverse public interests.
- **Service:** Obligation to assist stakeholders. This value is essential to support the public good.
- **Transparency:** Easily accessible and understandable policies and processes. This value is essential to demonstrate responsible use of public funds.[5]

Obtaining stakeholder feedback and buy-in, conducting diligent market research, adhering to procurement ordinances, along with broader guiding principles and values, and balancing privacy interests with public safety interests will aid law enforcement agencies in the successful procurement of an FRT platform(s).

5: National Institute of Governmental Purchasing: The Institute for Public Procurement. Enduring Beliefs or Ideals Shared by Public Procurement and Stakeholders, Values and Guiding Principles of Public Procurement - https://www.nigp.org/our-profession/values-and-guiding-principles-of-public-procurement

Program Management

We have so far visited several considerations any agency wishing to implement an FRT program should address. The design phase of an FRT program is the essential groundwork necessary to set an agency up for success. However, successful implementation of FRT does not stop at the moment of program launch. Ensuring operational transparency, accountability, and responsibility in the use of the technology requires continuing management and oversight. In this section, we will explore the general workflow of an operational FRT program and several considerations that should be made every step of the way.

Operational Workflow

The operational workflow phases detailed below include a series of elements and considerations which are important to address in the scope of this document. Although the specific structure and size of any agency's FRT program may differ, this section outlines a complete process in a stepwise format that should apply at least in part to any agency's program.

Submission

Submission is the phase of the program where an authorized user has obtained an image of an unknown individual and intends to submit that image for the application of FRT. To support both the overall process of performing an FRT examination and the subsequent auditing and control of the program the following key elements should be implemented in the submission process.

Vendor/Process Isolation

With few exceptions, an agency's access to an FRT is predicated on access to a third-party platform. Most FRT vendors are continuously developing and changing their platforms to meet various customer demands. This can result in widely varying capabilities concerning the capture of the important case and process information. The fast pace of this development thus presents the risk of disrupting an agency's ability to properly manage and audit their FRT programs in the case of over-reliance on a vendor's case management and/or auditing capabilities. Therefore, it is recommended that the FRT program implement solutions for managing the submission of probe imagery and the management of FRT requests/cases outside of the FRT Platform software.

Initial Capture of Case Information

The submission of probe imagery to an FRT program should require sufficient information from the requestor such that the nature and appropriateness of the request can be determined without immediate follow-up. The information required should mirror both the legal and policy stance of the agency concerning FRT.

In addition to allowing the FRT program to determine appropriate use, the initial case information captured will often assist FRT examiners. For example, specific intelligence about the unknown subject may allow an FRT examiner to eliminate or confirm a possible candidate as a potential lead in the case.

Agency members submitting requests should provide adequate contact information such that any follow-up can be completed during the review of the request.

Evaluation

The evaluation phase of the FRT program combines the initial receipt of an FRT request with an overall triaging of the request and probe imagery. The elements of the phase ensure the request adheres to defined protocols and the imagery is appropriate for use on the agency's FRT platform.

Case Review

Any agency members that have been trained and authorized to perform an FRT Examination should be considered subject matter experts in the application of the technology. As such, Case Review is the final phase where the legality and procedural appropriateness of the use of FRT can be ensured.

Imagery Management

Due to the innumerable potential sources of probe imagery, requestors sometimes submit probe imagery in obscure formats or embedded in digital documents and/or video. This is sometimes due to the ability of investigators to access digital devices or having to rely on a multi-step process to obtain and/or isolate imagery or video themselves.

As such, it is incumbent upon FRT examiners to be proficient in the interpretation of various file formats and their use as well as the methods needed to isolate facial probe imagery from whatever file or format submitted.

Imagery Review

Following the isolation and management of FRT probe imagery, the FRT examiner must perform a review of the imagery. This review relies upon the examiner's training in FRT techniques and their familiarity with the FRT vendor's product capabilities to determine an image's viability for use in FRT.

The goal of the image review should be to determine whether the submitted imagery is viable for the application of FRT. This process need not produce an objective measure or value to determine the viability of probe imagery. Some parts of this review are necessarily subjective.

It is recommended that this review process be completed before searching for or obtaining any results from the specific FRT platform. Completing this holistic review of the submitted probe imagery before upload to the FRT platform provides an additional layer of objectivity to the process. Many FRT platforms will return possible results on extremely poor imagery. This can lead to situations where an FRT examiner is presented with possible candidates who can strongly resemble the person pictured in an extremely poor-quality probe image. This resemblance has the potential to create an element of confirmation bias in FRT examiners where the quality of the discrete morphological comparisons in reporting potential leads can suffer.

Once the FRT examiner has determined the imagery is appropriate for FRT, it can then be uploaded into the FRT platform.

Examination

The examination phase of the FRT program is the phase during which the results returned from the specific FRT platform are assessed by the FRT examiner. The steps before this process and those that follow are designed to reduce the overall reliance upon the FRT algorithm in the law enforcement/public safety decision-making process. The goal of the examination phase is to locate a candidate image within the returns from the FRT algorithm which can be adequately confirmed as a potential lead. The consistent application of established facial morphological analysis between the probe image and the returned results is the cornerstone of producing these facial recognition leads.

Performing Initial Search

The search process in most FRT platforms have configuration options that relate to the operation of the platform and the number/type of search results that are returned. These options can limit and/or refine the gallery images that the probe imagery is compared to. The higher number or increasingly specific search parameters that are defined will necessarily limit the size of the search gallery.

The application of specific filters on physical descriptors should be limited to the amount of specificity that is known about the unknown person. For example, age ranges should only be limited by the extremes in age that are apparent in the probe imagery. This will allow the reasonable elimination of persons that the probe imagery is compared to. However, care must be taken when considering the use of metadata (such as age) when the reliability of that data – for both the unknown person and those included in the database – may be questionable. The goal is to reduce the statistical possibility of misidentification while not negatively impacting the possibility for a potential lead.

FRT platforms may provide configuration options related to the number of possible candidates that are returned. This configuration option should be weighted to the highest number to make the comparison of probe imagery to each candidate practicable in terms of time.

Initial Assessment of Candidates

In a review of a pool of possible candidates returned by an FRT platform, a trained FRT examiner will often be able to eliminate some candidates as potential leads immediately. The initial assessment of the pool of candidates should be focused on this "first pass" elimination.

Detailed Comparison

Once the pool of possible candidates has been reduced to those candidates that cannot be quickly eliminated, a detailed comparison should be performed on each remaining candidate. This detailed comparison should involve a systematic one-to-one comparison of the probe imagery to the remaining possible candidates. Most FRT platforms provide digital tools and techniques to assist in this process.

The goal of the detailed comparison is to locate a series of discrete morphological similarities between the probe image and a single candidate image. This set of similarities should be composed of commonly defined elements that are ideally from two or three different portions of the face. As an example, identifying morphological similarities between the nose, ear, and mouth is more desirable than identifying multiple similarities between only the ears of the imagery.

When sufficient detail is present in both the probe image and candidate image such that a morphological similarity can be articulated with specific language, that similarity might be referred to as a "lock." Such "locks" are what is necessary in the human assessment and confirmation of a candidate as a potential lead.

Three possible scenarios exist as an outcome of the detailed comparison step:

- One potential lead
- No potential lead
- Multiple potential leads
  - Although likely extremely rare, the technical possibility of two different individuals presenting sufficient locks with the probe imagery may occur. For the purposes of this document and the considerations presented, this scenario should be treated as a "no potential lead".

Facial Recognition Lead Production

When in the event one potential lead is located in the candidate pool during the detailed comparison, the findings of the detailed comparison should be documented in a standardized form. This form can be produced by the agency or can be a technical tool provided by the FRT platform. The standardization of the lead documentation provides a level of objectivity to the process by defining the level of detail and the specific information that will be provided as a return to the requestor.

Review

Prior to completing the FRT investigation and before returning the findings of the primary FRT examiner, a review process should be implemented by the FRT program. The goal of this review process is to provide an additional level of consistency and control with respect to the application of standardized training practices.

Secondary Review

A secondary review should consist of a separate FRT examiner reviewing the findings of the primary examiner. This can be a review of the facial recognition lead report alone, or a complete secondary search of the probe imagery within the FRT platform. During the secondary review, the reviewing FRT examiner should report either a concurrence with the provided results or rejection of the provided results. The secondary reviewer should be able to provide specific and articulable reasons for not agreeing with the provided results.

The outcome of the secondary review should be documented in the FRT request and subsequently reviewed by program managers. The cause for any lack of concurrence with results should be analyzed by program managers and the circumstances of the disagreement should be reviewed with the primary FRT examiner. The number and nature of FRT investigations with disagreements over the results should be monitored over time.

Management Review

A management review should be implemented on as many FRT investigations as practicable. This review is intended to ensure the proper application of the technology generally as well as adherence to all applicable policies and procedures. The management review should periodically include the auditing of the FRT platform for each reviewed FRT investigation. **This management review should be completed prior to returning results on the FRT investigation.**

Distribution

The results of any application of FRT should be documented and retained for historical reference. This should include those FRT examinations that return no results. The return of facial recognition lead reports to requestors should be standardized and it is suggested to also include the following key pieces of information:

- The agency's stance concerning the use of FRT information.
- Agency policy on disclosure of the use of FRT on arrest documents.
- Any significant applicable laws on FRT.
- Any agency's policy requirements on how investigators are to proceed with the investigation (i.e., not rely strictly on a photo line-up to establish probable cause.)

The distribution phase is also an ideal opportunity to remind investigators that the collection of data related to the results of the investigation after the use of FRT is valuable. As previously mentioned, a comprehensive auditing and reporting process provides a greater foundation for an accurate evaluation of the technology.

Auditing and Reporting

The documentation of an agency's FRT program processes is a critical part of a transparent and accountable FRT program. The existence of an FRT request in a criminal case, the outcome of the subsequent FRT investigation, as well as the eventual use of the findings in the criminal case are the main data points that should be captured.

This documented data ultimately provides a statistical snapshot of the impact the technology has within the greater mission of responding to and preventing crime. **A comprehensive auditing and reporting process provides a foundation for a true evaluation of the technology and how it improves policing outcomes for both the agency and the community.**

FRT data collection serves three primary purposes. First, data collection on every FRT investigation conducted should be made to understand the purpose of the investigation. Second, data collection on the part of an examiner including all important details of the examination phase helps to ensure quality control measures are being met. Finally, a detailed analysis of the data and any metrics designed to measure the data help identify program deficiencies that may include training, timeliness, process hurdles, etc.

This data can and should be converted into measurements that provide clear information on FRT use, investigative results, and compliance with the policy. Similarly, it provides an FRT program an opportunity to share success cases that denote the value of the technology, the contribution of the examiners, and the impact of the technology in the broad crime-fighting mission.

An added benefit of maintaining statistical usage data is that it potentially can identify deficiencies that

need to be addressed. Deficiencies can normally be alleviated through training. However, that need for mitigation may not ever become known without data and metrics that paint a full picture.

Program Oversight

Strong program oversight enhances the ability to develop proper policies, perform relevant training, and ensure responsible use. Established oversight allows an agency to properly manage any existing or emerging operational concerns. Therefore, it is recommended that agencies have a clearly defined program management component of their FRT program.

To ensure the program is operating responsibly, it is the recommendation of this report that a program manager is identified and assigned to operate an FRT program. This individual should be well versed in the use of the technology, have a sound understanding of the analysis of facial identification as a human examiner, and should be empowered to ensure the program is following protocols, policy, applicable law, and the expectations of the agency.

A well-defined program management structure promotes community trust that the regulations, protocols, and policies governing the use of FRT are in place and being followed. Similarly, it provides the community a point of contact to provide feedback, education, and transparency. It also enhances the ability of an agency to work closely with their FRT platform vendor. Even well-defined operational processes and procedures can mean little if there isn't a responsible party who ensures those procedures are being executed in their intended manner.

It is recommended that the program manager collect and report regularly on the use of the technology. A snapshot of a program can provide not just heads of agencies but also concerned members of the public insight into how FRT is being used to affect crime.

A program manager should be expected to stay current on changes and emerging trends of FRT and be expected to ensure that protocols and department policies are current and properly address any changes or advancements in the technology broadly.

A program manager should be expected to ensure the FRT examiners are properly trained. That responsibility may include the initial training of the technology and also in the discipline of facial identification and comparison. A manager should also ensure that regular on-gong training is being fulfilled.

Finally, having a program manager affords both an agency and the FRT vendor a clear point of contact. This can enhance the communication between agency and vendor which is desirable for both parties. On the part of law enforcement, it affords the agency a better opportunity to share with the vendor any challenges that they may experience. For example, an agency may identify an area in the system processes that needs enhancement or possibly identify a deficit in the reporting and auditing capabilities of the system. In both examples, good communication between the agency and vendor would enhance the ability of both parties to resolve any issues that arise and to improve upon the FRT platform.

In conclusion, it is our recommendation that a program manager be an integral part of the program design from the program's inception. Additionally, the program manager should be tasked with ensuring the program strictly follows all policy, protocols, and design pieces that enhance responsibility and accountability.

Operational Concerns

When developing a FRT program, it is imperative that law enforcement agencies consider the legitimate operational concerns that exist with the use of FRT. Some of the concerns when utilizing FRT include threats to privacy, violations of civil rights and personal freedoms, and potential data theft by both authorized and unauthorized users of the platform.

In order to address these concerns, agencies should have policies and safeguards in place to prevent such misuse of the system. Additionally, agencies using FRT should have adequate data storage capabilities and should have clearly established record retention and purging policies with regard to the images stored within the FRT platform.

"The documentation of an agency's FRT program processes is a critical part of a transparent and accountable FRT program."

Another consideration is the challenge of the collection of data that demonstrates the efficacy and or the ultimate results produced from FRT examinations. This is specifically due to the fact the technology is used as an investigative tool and a potential lead may not result in a successful arrest and prosecution of the identified subject. Since it is common for FRT examiners to be removed from the investigation, FRT programs often encounter difficulty in collecting valuable data regarding the final outcomes of the investigations. Highly respected existing FRT programs have countered this concern through data collection starting at the investigative inception. It is suggested that follow-up be made after the results are returned to the investigators to determine case status. **It is advised that an agency put in place a system or process to the extent possible to capture the end results of an investigation that utilized FRT.**

Qualitative Review

In order to support a complete picture of any one product, it can be important to research both past successes of an agency using a prospective FRT product and to analyze the measured impacts and successes of FRT technology broadly. In this section, we explore a specific FRT program and its statistics, as well as highlight some success stories directly attributable to the technology.

Working FRT Program Statistics

The Integrated Justice Information Systems (IJIS) Institute and the International Association of Chiefs of Police (IACP) collaborated to produce a Law Enforcement Facial Recognition Use Case Catalog that lays out 19 common and beneficial use cases of Facial Recognition Technology in Law Enforcement . They also support the use cases with some real-world examples of each. In this section, we supplement this excellent reference with an example of an established (multi-year) program in a major US city that operates in accordance with the processes and policies outlined in this document.

The program is operated by the police department of a large US city and has been operational for several years. The enrollment database used by the department to conduct investigations consists solely of their own booking images so they are high-quality images. The following statistics and examples are from Full-Year 2020:

- Over 1000 facial recognition investigation requests were received by the program and serviced by a dedicated facial recognition team of approximately 20 trained examiners.
- Of the requests received, 49.8% were determined to be suitable for examination: The facial recognition examiners and tools determined that about half of the images submitted were not of sufficient quality to conduct further analysis for a potential lead.
- Of the requests that were processed, 71.7% resulted in a potential lead (or 35.7% of the total submissions): Many of the investigations yielded leads that were pursued in combination with other information and investigative efforts.
- Many Positive Outcomes: Multiple crimes (many serial) were solved across areas including Robbery/ Theft, Larceny, Homicide, Financial/Fraud, Sexual Assault / Trafficking, Drugs and Explosives.
- Zero Wrongfully Arrested: There are no known instances (in 2020 or the history of the department) of individuals being wrongfully accused due to FRT.

Real-World Success Stories

To supplement the technical evaluation, this report also includes use case examples based on success stories. These excerpts are from real law enforcement events involving facial recognition. Collectively, they provide a snapshot of the value of FRT and its positive influence on public safety concerns. Also included is a review of an active agency's FRT program along with a hypothetical comparison of NIST's findings of an FRT program.

Sex Crime - Facial Recognition Request Assists with a Priority Sexual Assault Case: a request was submitted by sex crimes detectives who were on the scene of a sexual assault involving a juvenile victim. Within an hour of receiving the request, facial recognition technology was able to identify a likely candidate of the suspect involved. The information was shared with the lead detectives who were then able to positively identify the subject through additional investigative methods.

Financial Crime - Facial Recognition Helps with a Financial Crimes Investigation: Detectives submitted a

facial recognition investigation request seeking to identify a female responsible for withdrawing money from a victim's bank account at different ATM locations. The investigation led to a potential lead and her identifiers were shared with the investigating detectives. They later confirmed that this female had been involved in several other cases for obtaining money under false pretenses and other financial crimes.

Grand Larceny - Facial Recognition Assists with a Grand Larceny Auto Investigation: Detectives submitted a facial recognition investigation request seeking assistance with the identification of a subject suspected of stealing multiple vehicles from a rental car company. This had been part of an ongoing series with multiple related events. The facial recognition investigation yielded a likely candidate, and this information was given back to the detectives who were able to confirm that the subject identified was the suspect involved.

Counter-Terrorism - Facial Recognition Helps Identify a Person of Interest in Explosive Device Investigation: A request for facial recognition to assist in identifying a person of interest in a case involving a believed explosive device was submitted. Surveillance footage captured an image of a driver of a vehicle who was believed to be responsible. A facial recognition investigation identified a potential lead. Detectives contacted the subject and confirmed the lead was accurate. Detectives obtained additional evidence and probable cause to continue their investigation, leading to an arrest warrant being obtained.

Narcotics - Investigation Helps Identify Drug Suspects Involved in Shooting: Patrol detectives were seeking assistance identifying two suspects who were involved in a drug transaction. Those same suspects were also responsible for firing several shots at a victim. Social media account information was provided to assist with the investigation. Once two potential leads were identified, the subjects' information was relayed back to detectives who confirmed the identity of both subjects.

Robbery - Ending a Violent Robbery Series Using Facial Recognition: Robbery detectives submitted a request to assist with a facial recognition investigation for a suspect responsible for multiple business robberies. On several related events, the suspect would enter the business, batter the clerk, and then take cash out of the register. A potential lead was identified through the investigation and detectives were able to positively identify the suspect via additional investigative means including several victims who all identified the same suspect through a photo lineup.

Homicide - Facial Recognition Technology Helped with a Homicide Investigation: Detectives requested a facial recognition investigation for a violent assault/battery incident. The victim was knocked unconscious during the incident and hospitalized for several weeks with a brain injury. Later, he died as a result of those injuries. The facial recognition investigation assisted detectives with the suspect's identity. The suspect was arrested on the appropriate charges.

Technical Evaluation

As noted above, facial recognition may be used in different scenarios. It may be used to search a database to generate a lead, or it may be used to verify the identity of a user before granting a privilege, such as unlocking a mobile phone or allowing a border crossing. How well a given facial recognition algorithm works in verification is easier to understand than in the search scenario even though law enforcement overwhelmingly uses search more than verification. To further enhance the evaluation of FRT technology in real-world scenarios, a qualitative review is also presented which focuses on use cases and statistical impacts.

Algorithm Evaluation

At the heart of any FRT platform is the underlying algorithm used to perform both analysis and comparison of digital face imagery. Generally speaking, any computer program which is designed to solve a problem based on its inputs is an algorithm. The code written in Microsoft Excel which sorts an arbitrary list of numbers is an algorithm. The computer routine in our smartphones that decides the best route to a destination is an algorithm.

Most algorithms can be assessed against a provable successful output; such is the case with sorting where the "quality" of the algorithm comes down to the speed at which it can perform the sort. This is because the output of a sorting algorithm can be easily checked for accuracy. Other algorithms, such as an algorithm routing a road trip to a list of destinations cannot be easily checked for optimal solutions, therefore the "quality" of these algorithms must be assessed over both the speed and the quality of the outputs of the algorithm.

In the case of facial recognition algorithms, the size of the image galleries and the utilization of one probe image in most cases rarely causes the duration of the algorithm to be a factor. Hence, assessing the quality of an FRT algorithm necessarily involves analyzing the accuracy of the output of the algorithm. In this section, the various use cases of FRT algorithms and how the accuracy can be assessed are explored.

Accuracy of Verification Algorithms

In a phone unlock verification case, there are only two possible inputs – either the owner of the device is trying to get in, or an imposter is trying to get in. The system works "perfectly" if the owner always gets in and all imposters are kept out. The attached diagram depicts the range of outcomes.

	System 'says'	
Ground Truth	**Owner**	**Imposter**
Owner ("Match")	**UNLOCKS (True Positive)**	**DOES NOT UNLOCK (False Negative)**
Imposter ("No Match")	**UNLOCKS (False Positive)**	**DOES NOT UNLOCK (True Negative)**

In a facial recognition verification use case like this, the system acquires a "live" facial image, extracts a template from that image, and compares it against a facial template stored on the device. If the match score exceeds an established threshold value, the system will unlock the phone because a "match" has occurred. If the match score does not exceed the threshold score, the request to unlock is rejected. The system is making a "decision" based on a comparison of the threshold score with the score generated by the live image versus the template.

Although the template stored on the phone is fixed, the "live" face image will change from an attempt to attempt, due to changes in pose, illumination, and expression. The threshold score is a critical aspect of this process – if the threshold is set too low, then there is a better chance that an imposter can unlock a phone, but if the threshold is too high, an owner might not be able to get into their phone, because the system cannot accommodate normal changes in the face. Therefore, mobile phone security settings are likely to incorporate lower thresholds than other security systems, such as in a border crossing, because phone sellers don't want customers to get frustrated with overzealous security settings.

How does a threshold get determined and what does that mean for accuracy? Facial recognition algorithm developers rely upon "match score distributions" to help identify threshold scores. A match score distribution allows one to plot the match scores generated for a large number of true matches and non-matches. In a "perfect" system, all true matches would generate scores above all non-match scores. Refer to Figure #1 below. In this case, a threshold score could be set at a value in between the non-match (red) and true match (green) scores and accuracy would be perfect. Every true match would be declared a "match" and every non-match would be declared a "non-match." Numerically, we would describe the "true match" accuracy rate as 100%, with a false match rate of 0%. "False Match" and "False Positive" represent the same thing.

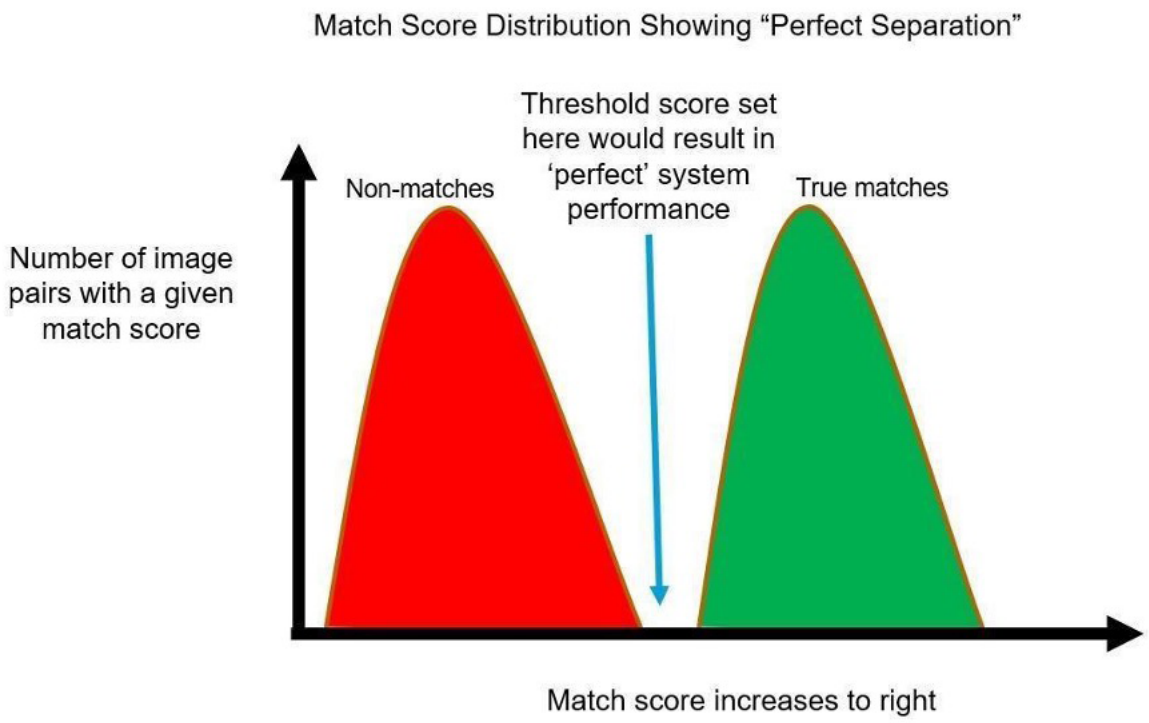


Figure #1

In reality, no biometric system is this perfect. Changes in pose, illumination, and expression, among other factors, can reduce the match score generated for a true match pair, while twins and other "look-alikes" can lead to non-match pairs with high scores, as depicted in Figure #2 below.

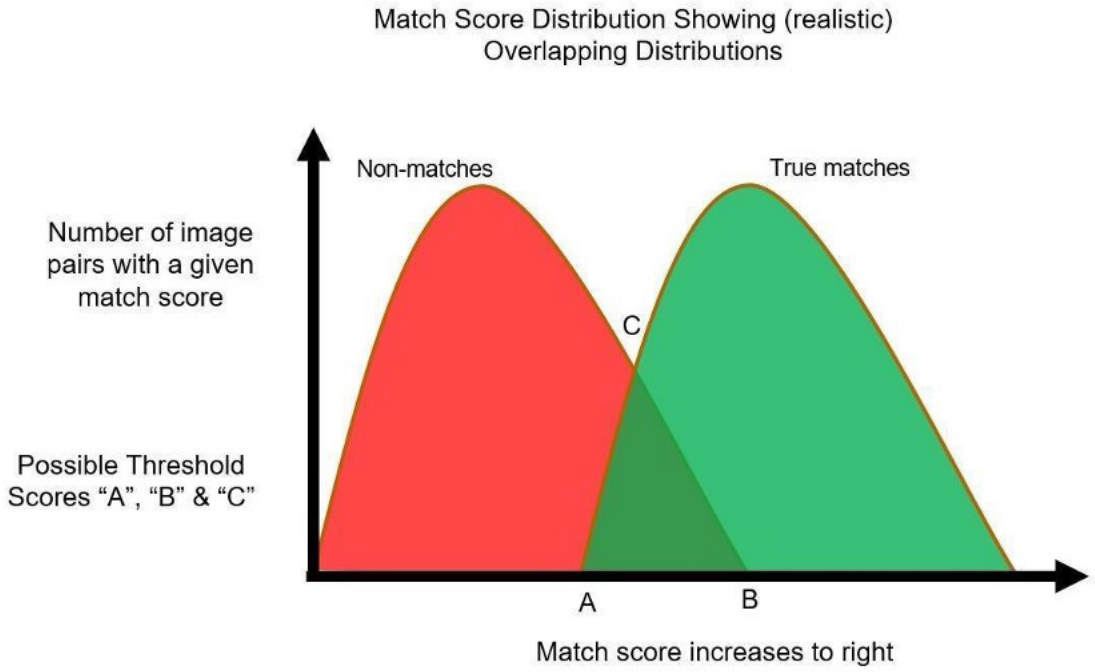


Figure #2

With a match score distribution as depicted in Figure #2, different thresholds will lead to different outcomes. If the threshold score is set at "A," then all true matches will be declared as such. In the mobile phone example, the user would always get in, but at the risk that others could also get in – and that would be declared a "false match." If we were to determine that the total number of non-matches (red) represented above score "A" is about 20% of the total number of true matches (green), then for any given comparison that resulted in a score above "A", the distribution implies that an imposter might get in once every five comparisons.

If the threshold score were set at "B," then the user could have some assurance that no one else could get into their phone, but the owner might be rejected more often than is desired – approximately 40% of the time if you throw out all true match pairs represented to the left of "B."

The threshold score "C" represents a score that occurs just as often for matches as it does for non-matches. According to this distribution, either decision – "match" or "non-match" is just as likely to be correct for that specific score.

The match score distribution in Figure #2 is closer to reflecting the real-world scenario for most biometric algorithms today than the "perfect" distribution in Figure #1. As the score increases from position "C", the probability increases that a given pair of images is an actual match, while below that score it is more likely to be a non-match.

In order to provide an "apples-to-apples" comparison of the accuracy of different algorithms, it is common practice to report the percentage of true matches that are missed (the "False Non-Match Rate" or FNMR) for all scores at or above the score above which a false match has a fixed rate. At the time of this writing, according to the NIST FRVT website, the very best algorithm for 1:1 verification had an FNMR of 0.0022 (22-in-10,000) using a threshold set for a False Match Rate (FMR) of 1-in-100,000. In terms of "hits", this would mean the algorithm had a true match rate of 99.78%.

Prior to 2018, NIST frequently used a fixed false match rate of 1-in-1,000 to define the threshold score for a given algorithm. The reduction of fixed false match rate to 1-in-100,000 (or sometimes 1-in-a-million) is an indication of how good the algorithms have become.

Accuracy of 1:N Search Algorithms

The primary law enforcement use of facial recognition is not for verification, but lead development through 1-to-many (1:N) searches. When a probe is searched against a gallery using a facial recognition system, the system returns a list of the highest-scoring candidates, which are then reviewed by trained personnel to see if a viable lead is present. The "decision" of whether a viable lead has been identified is not made by the algorithm, but by the reviewer. In other words, the system is not declaring a "match" so there can be no "false match" that would otherwise be wholly attributed to the system.

Technical Considerations

Assessing the Efficacy of Facial Recognition Platforms

The Face Recognition Vendor Test (FRTV) program run by the National Institute of Standards and Technology (NIST) is currently the only publicly available option where products may be voluntarily submitted for accuracy characterization. This is an excellent reference and resource; however, some interpolation is required to map their published results to law enforcement use cases and conditions.

Generally, there are two major aspects associated with assessing the efficacy of a particular product: Accuracy and Demographic Bias.

Accuracy

NIST's Facial Recognition Vendor Test program assesses voluntarily submitted algorithms for Identification (1:N) accuracy in various combinations of image types; however, they do not have a test series that assesses algorithms with the lowest quality probe input images (e.g., 'wild' images in their terminology are the lowest quality images) which are typically more challenging. The typical usage for law enforcement involves probe images that are not posed (e.g., sourced from video surveillance cameras, smartphones, etc.) being searched against an enrollment database that is composed of high quality, posed images (e.g., booking photos). The closest approximation to this in the 1:N NIST test results is believed to be input images sourced from the 'webcam' case which is the lowest quality image used. The verification tests (1:1) results produced by NIST do include a test case against 'wild' images so these results also provide an indication of a product's performance; however, the wild images may not be curated in a way that allows for accurate testing of demographic variation.

The typical concern with face identification accuracy is a false positive identification (incorrectly matching a probe image with an entry in the enrollment database). As explained above, there is a tradeoff between the minimization of false positives and false negatives (failing to match a probe image with an entry in the enrollment database). Increased accuracy of one results in diminished accuracy of the other. In their testing, NIST holds the maximum false positive rate constant (typically fractions of a percent but this threshold varies across different test cases) and reports the false negative rate that results from those conditions. The better performing algorithms generate low false negatives while still operating at or below the prescribed false positive rate. In general, the highest quality products (upper 25%) are exceptionally accurate in that they produce fractions to up to only 2-3 percentage points of false negatives depending upon the test configuration while maintaining a false positive rate in the fractions of a percent.

According to a report[6] that assessed the accuracy of humans of different training and innate talents, the accuracy of quality FRT products exceeds that of humans; however, process and human analysis is very important to the overall outcome. The report further found that the most accurate result (100% in their test case) was achieved through the combination of a trained human examiner supported by FRT technology which is better than the technology or humans individually could achieve.

Demographic Bias

This relates to inconsistency of product's accuracy across images of individuals of different race, gender and/or ethnicity. This was a pronounced problem years ago, but once it was identified, there has been considerable technical progress in improving this condition. NIST published a specific analysis of product sensitivity to demographic effects that found very small to undetectable differences in result accuracy due to demographic effects for high performing products. The Information Technology & Innovation Foundation, among others, analyzed the results and concluded the following :

- The most accurate identification algorithms have "undetectable" differences between demographic groups
- The most accurate verification algorithms have low false positives and false negatives across most demographic groups

6: Proceedings of the National Academy of Sciences. (2021, September). Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. https://www.pnas.org/content/115/24/6171

- Algorithms can have different error rates for different demographics and still be highly accurate

An important consideration in facial identification is the enrollment database itself. In principle, the database can be composed of anything that the agency is authorized to collect that provides a useful backdrop for its identification use cases. Variables associated with the enrollment database that may influence the program include:

- **Size:** The number of enrolled images may have implications on the sheer accuracy of the search results and/or performance (processing time). These effects are not significant with reasonably sized databases and are largely insignificant for most applications, but this may provide an important assessment criterion of technology partners.
- **Quality:** Image quality plays an important role in the effectiveness of the tool (as well as the effectiveness of any human analysis). Quality images with structured pose, lighting, and resolution provide the best results and typically this type of image (captured in controlled environments) is what is used by law enforcement.
- **Accuracy:** The correctness and completeness of the metadata attached to images in the enrollment database are essential to ensure proper identification as well as to facilitate match verification via contextual information.
- **Source:** This plays a key role in the aforementioned attributes of size, quality, and accuracy but also is an important component of how the program is perceived relative to privacy and surveillance of society.

The source and composition of the enrollment database is a function of other policy decisions for the program; however, the process described above is independent of that since it is largely founded on human analysis and adjudication.

A technology product plays two important functions in the FRT identification process:

1. It performs the actual FRT search and match assessment through comparison of templates generated from the probe image and all images in the enrollment database.
2. It enforces access control, records all the examination steps for audit, and documents the process to produce the comprehensive facial recognition lead report (output).

As with any outcome that leverages technology, especially for critical applications such as FRT, there must be a trusted partnership between the user and technology provider (a technology provider may be a vendor, another agency, or indigenous).

FRT technology providers should be expected to:

- Ensure that their products are regularly characterized for accuracy under realistic conditions.
  - Submit their products to objective benchmarking and disclose their results. The Face Recognition Vendor Test (FRTV) program run by the National Institute of Standards and Technology (NIST) is currently the only available public option but it requires some interpolation to map their results to law enforcement use cases and conditions. It is important to ensure that the submitted algorithm(s) translate to the product (vs. a specialized implementation that is tuned to a particular case or that may not scale).
  - Agencies may be able to assess the overall accuracy as well as susceptibility to demographic bias themselves through other testing regimens and indeed they should have representative cases (typical probe images and enrollment database) that they can use to assess and continually revalidate providers.
- Be transparent about their product's operation, efficacy under relevant conditions, as well as how they test and continuously improve their products.

- If artificial intelligence and machine learning is the basis of their algorithm, the partner should disclose the source of their training data, their legal entitlement to that data as well as how they continue to curate the training data set.
- Have thoughtful controls built into their products such that users can be authenticated/authorized, users can be properly constrained (aka 'guard rails') in their application of the technology (e.g., not employ it outside of other workflows), and that usage can be audited for compliance with policies.
- Offer comprehensive training on their product that clearly articulates its intended usage as well as any limitations or constraints.
- Demonstrate efficacy and internal processes that ensure responsibility, accountability, and ethics in the contemplation, design, development, cybersecurity, deployment, and support of products that contain sensitive technologies such as FRT and that manage sensitive data.
- Demonstrate a track record of ethical and responsible technology.
- Have methods to systematically obtain feedback and data from the system in operation to assess performance and troubleshoot problems. The partner should have well-defined, privacy-respecting mechanisms and procedures in place leaving all control of any data that is disclosed with the program administrator.

If the product is consumed as-a-service and hosted by a technology provider (vs. an on-prem dedicated deployment), additional considerations apply since in this case the provider stores and processes the enrollment database, as well as any searches, conducted. These considerations include:

- Providers should have the necessary facilities and certifications to process this type of data (e.g., CJIS and FedRAMP compliant data centers and processes).
- Providers should be able to offer a DPIA (Data Privacy Impact Assessment) for their program.
- Providers should have a clear data policy.
- Contractually they should not impose terms that give them entitlement to any agency data (i.e., the agency is the controller of all data that they submit or that is generated on their behalf during processing).

Conclusion

As stated in the introduction, this product has been designed to be continually updated with the best practices, design mechanics, and up-to-date information when needed. From the information provided in this first version of this document, it is clear that facial recognition technology is a complex and ever-evolving tool for law enforcement which must be revisited often to ensure it is being used in a way to assist law enforcement in solving crime, but also protecting the public from misuse and privacy violations.

We are still in the beginning stages of knowing the vast uses and ways to use facial recognition technology to aid in investigations. It is necessary to keep a dialogue going with all various stakeholder groups - law enforcement at the federal, state, and local level, privacy and civil liberty advocates, private sector specialists, and local communities.

While this product exclusively explores the use of two-part verification facial recognition technology for facial identification, some agencies may choose to employ a different model. It is within the right of an agency to employ a FRT program which is most fitting for their needs and the desire of a community, but with any choice there are advantages and disadvantages, and it is suggested that all agencies wishing to deploy any type of facial recognition technology use the tenets discussed in this document to make the best choice for their agency.

For those agencies wishing to implement the use of FRT, but who are unsure on how to move forward, collaboration with other entities who have already developed robust, responsible programs is recommended. The sharing of best practices in crime-fighting technology among law enforcement has a history of beneficial impacts. Law enforcement has an obligation to be good stewards of information and policies and they must be willing to share this with fellow law enforcement agencies wishing to begin using FRT.

Facial recognition technology is being used daily to aid law enforcement in capturing the most violent criminals in our country and bringing closure for victims. It has been proven to be highly successful in solving various types of crimes afflicting our communities when used with the highest degree of responsibility, transparency, and accountable management.

The Major Cities Chiefs Association will continue to monitor this growing and developing technology as well as how it is being used across its member's jurisdictions and release updated versions of this product when appropriate. The MCCA will continue to engage all stakeholders with an interest in this evolving sector and communicate the latest information with its membership for member agencies to adjust their best practices as needed.

Acknowledgements

This product was put together in partnership and through the efforts, knowledge, and expertise of the following individuals. The Major Cities Chiefs Association is grateful for the work done by this working group for all their time and effort toward bettering the operations of law enforcement. This list is not exhaustive, not all persons with significant contributions to this product can be listed. A special thanks goes out to the technical advisors from vendors and law enforcement.

Armando R. Aguilar
Assistant Chief of Police
Miami Police Department, Florida

●

Bill Steinmetz
Lieutenant
Las Vegas Metropolitan Police Department, Nevada

●

Christian P. Quinn
Senior Director of Government Affairs
Brooks Bawden Moore LLC

●

Ivonne D. Valdes
Sergeant
Miami Police Department, Florida

●

Jim Lowery
Deputy Chief
Arlington Police Department, Texas

●

Kelcy McArthur
Statewide Network of Agency Photos Unit Manager
Michigan State Police

●

Kelly Bluth
Detective
Las Vegas Metropolitan Police Department, Nevada

Krystal Howard
Departmental Manager
Michigan State Police

●

Laura Cooper
Executive Director
Major Cities Chiefs Association

●

Megan Noland
Director of Special Projects
Major Cities Chiefs Association

●

Paden Weber
Officer
Las Vegas Metropolitan Police Department, Nevada

●

Patricia Williams
Associate Director
Major Cities Chiefs Association

●

Patrick T. Quinn
Lieutenant
Chicago Police Department, Illinois

●

Resource Guide and Further Reading

There exists an extensive amount of resources applicable to the content in this product. Some of the most relavant resources for continued reading, research, and application are listed below for convenience.

- Major Cities Chiefs Association - http://www.majorcitieschiefs.com
- National Institute of Standards and Technology (NIST) - Biometrics -https://www.nist.gov/biometrics
- Congressional Research Service - Federal Law Enforcement Use of Facial Recognition Technology (Oct. 27, 2020)- https://crsreports.congress.gov/product/pdf/R/R46586
- Congressional Research Service - Facial Recognition Technology and Law Enforcement: Select Constitutional Considerations (Sept. 24, 2020) - https://crsreports.congress.gov/product/pdf/R/R46541
- ITIF Report - The Critics Were Wrong: NIST Data Shows the Best Facial Recognition Algorithms Are Neither Racist Nor Sexist - https://itif.org/publications/2020/01/27/critics-were-wrong-nist-data-shows-best-facial-recognition-algorithms
- House Judiciary Committee Testimony - Mr. Barry Friedman - Jacob D. Fuchsberg Professor of Law and Faculty Director, Policing Project, New York University School of Law (Jul. 13, 2021) - https://docs.house.gov/meetings/JU/JU08/20210713/113906/HMTG-117-JU08-Wstate-FriedmanB-20210713.pdf

Appendix A

Terminology and Definitions

Algorithm – an algorithm is a set of rules that instruct a computer on how to accomplish a task or solve a problem. A facial recognition algorithm is the software implementation of techniques used to verify or determine an individual's identity by processing a video frame or a digital image in which the individual's face is visible. Typically, the algorithm compares facial features in an image to faces contained within a database or gallery.

Candidate List - in facial identification, a rank-ordered list generated from a facial recognition search. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Examiner - in facial identification, a forensic face examiner is a trained specialist at analyzing images of faces and comparing them for identification purposes to determine if the images are of the same individual.

Facial Identification - (1) the process of determining the identity of an unknown person from a photo database, known as the enrollment database or gallery to answer the question, "Can this unknown person be matched to any image enrolled in the database?" It is often referred to as one-to-many matching (1:N) because it compares a probe image to all images in the enrollment database. (Source: Accuracy and Bias of Face Recognition Technology and Law Enforcement Use: Factual Background, NYU Policing Project, 2021.) (2) the process of searching a probe into a gallery (Source: Face Recognition Vendor Test (FRVT) Part 3 Demographic Effects, NISTIR-8280, by NIST, 2019.)

Facial Recognition - in facial identification: (1) by automated systems, the automated searching of a facial image as a probe in a facial recognition system (one-to-many), typically resulting in a group (candidate list) of facial images being returned to a human operator in ranked order based on system-evaluated similarity; (2) by humans, the mental process by which an observer identifies a person as being one they have seen before. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Facial Verification - (1) the process of authenticating a person's asserted identity by comparing two image templates to answer the question, "Are these two images the same person?" It is also referred to as one-to-one (1:1) matching because a probe image—the image inputted into a face recognition system—is only compared to one other stored image (Source: Accuracy and Bias of Face Recognition Technology and Law Enforcement Use: Factual Background, NYU Policing Project, 2021.) (2) the process of comparing two samples to determine if they belong to the same person or not. (Source: Face Recognition Vendor Test (FRVT) Part 3 Demographic Effects, NISTIR-8280, by NIST, 2019.)

Facial Comparison (in facial identification) - a manual process to identify similarities or dissimilarities between two (or more) facial images or facial image(s) and a live subject for the purpose of determining if they represent the same person or different person. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Facial Identification (FI) - the discipline of image-based comparisons of human facial features. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Facial Recognition (FR) - see face recognition. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Facial Review - in facial identification, an adjudication of a candidate list. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

False Negative - in facial identification, is when the face recognition system fails to match a person's face to an image that is, in fact, contained in a database. In other words, the system will erroneously return zero results in response to a query. Also referred to as 'false non-match'. (Source: Electronic Frontiers Foundation – Face Recognition)

False Positive - in facial identification, is when the face recognition system does match a person's face to an image in a database, but that match is actually incorrect. Also referred to as 'false match'. (Source: Electronic Frontiers Foundation – Face Recognition)

Gallery - in facial identification, an FR system's database, which typically contains all known-person biometric references (samples or templates, or both). Also referred to as 'enrollment database' (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Morphological Analysis - in facial identification, direct comparison of class and individual facial characteristics without explicit measurement. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Probe - in facial identification, a facial image or template searched against the gallery in a facial recognition (FR) system. (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Third-party Imagery - in facial identification, images used in facial recognition (FR), or facial identification (FI) that were not captured by the agency performing the comparison (for example, family snapshots of a missing person). (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Uncontrolled Image or Ad-hoc Image - in facial identification, an image not captured in accordance with facial identification/facial recognition (FI/FR) standards or guidelines (for example, a surveillance image). (Source: Standard Terminology Relating to Forensic Science, ASTM 1732-19, by ASTM International, 2020)

Appendix B

Misconceptions and Reality

Myth:

Most community members oppose law enforcement using facial recognition.

Reality:

Despite how vigorously privacy advocates and other special interest groups have recently opposed FRT, there is little public support for banning or substantially limiting the responsible use of FRT by law enforcement. Polling conducted by NetChoice in 2020 revealed most Americans would prefer state and local governments work with law enforcement to improve the use of facial recognition (83%) rather than banning the technology (17%).

Similar findings were found in 2020, by the polling firm, Schoen Cooperman Research, when they conducted a comprehensive nationwide poll regarding Americans' opinions on FRT. Their survey found that 68% of Americans believe facial recognition can make society safer, 70% feel it is sufficiently accurate in identifying people of all races and ethnicities, and 66% believe law enforcement's use of facial recognition is appropriate.

The PEW Research center also conducted extensive research on the topic in 2019. They found that when the public was asked about their confidence that different entities will use facial recognition tools responsibly, they express much greater trust in law enforcement agencies than in advertisers or technology companies. A majority of U.S. adults (56%) trust law enforcement agencies at least somewhat to use facial recognition technologies responsibly, with 17% indicating that they trust these agencies a great deal to use facial recognition.

By contrast, only around one-third of U.S. adults trust technology companies to use FRT responsibly, and just 18% trust advertisers with these technologies. Trust in law enforcement does vary based on demographic factors such as race, age, and political affiliation. White adults express higher levels of trust in the use of FRT by law enforcement than black adults (43%). Generally, older adults have greater trust in law enforcement's use of FRT than persons who are 18-29 years old.

These findings do not indicate that most community members don't want police leveraging FRT technology to keep their communities safe. It does affirm that there is a need to engage every community in a transparent and intentional way, particularly those that have historically been marginalized or over-policed.

Reference(s):

More Than Half of U.S. Adults Trust Law Enforcement to Use Facial Recognition Responsibly | Pew Research Center
https://www.securityindustry.org/report/u-s-public-opinion-research-on-the-support-of-facial-recognition/
https://netchoice.org/media-press/americans-want-facial-recognition-use-by-law-enforcement-improved-but-not-banned/

Myth:

Many states and local communities are already banning facial recognition from use because of the dangers it poses.

Reality:

Despite the headlines, legislation introduced to ban or significantly limit the use of FRT actually have had very little support. Proposed bills failed to advance, or were completely rejected by legislatures in at least 17 states during the 2020 and 2021 sessions including California, Colorado, Hawaii, Kentucky, Maine, Maryland, Massachusetts, Michigan, Minnesota, Montana, Nebraska, New Hampshire, New Jersey, New York, Oregon, South Carolina, and Washington.

Vermont and Virginia are the only states that have banned law enforcement's use of FRT, and in both states, the laws have already resulted in unintended consequences. Haphazard outright bans fail to consider exemptions that allow for the use of other tools that rely on digital face comparison technology not necessarily used to investigate specific suspects. Such tools are critical in child sexual exploitation and human trafficking investigations. Similar tools are used in digital forensics to categorize media and parse digital evidence files. Vermont passed a bill in May, just months after enacting the law, to exempt such software from the ban.

The new Virginia law, which went into effect July 1, 2021, eliminated a regional program in operation since 2017 which was utilized in over 14,000 instances, with no reported misidentifications. Among its many successes, the program is credited with helping to exonerate an innocent man accused of a violent crime in Virginia, identify a veteran posting suicidal messages online, arrest multiple bank robbery suspects, close down an organized identity theft ring, and further numerous investigations into gun trafficking and other violent crimes.

Massachusetts established defined use conditions applicable only to law enforcement, through its police reform measure. While a broad public sector ban was initially considered in Maine, the legislature passed an amended measure in June that allows for law enforcement use under narrow conditions. Washington State's 2020 law establishing conditions for public sector applications of facial recognition went into effect on July 1, 2021.

In March, Utah enacted comprehensive policy safeguards for government applications. The measure, supported both by the Utah Department of Public Safety as well as the American Civil Liberties Union, establishes requirements for public-sector and law enforcement use, including conditions for access to identity records held by the state, and transparency requirements for new public sector applications of facial recognition technology.

At the municipal level, only three localities in the United States enacted broad bans of FRT in 2021: Minneapolis and King County, Washington, passed bans on government use, while the Baltimore City Council recently approved an expansive ban that restricts government, personal, and commercial use.

Reference(s):
Most State Legislatures Have Rejected Bans and Severe Restrictions on Facial Recognition | Security Industry Association
http://www.mainelegislature.org/legis/bills/getPDF.asp?paper=HP1174&item=3&snum=130
https://le.utah.gov/~2021/bills/static/SB0034.html

Myth:

Authorizing the use of FRT is too great a risk as it will likely lead to over-utilization.

Reality:

FRT is often discussed as an all-or-nothing option, with critics calling for increased bans and moratoriums. This narrative fails to consider the opportunities that communities have to leverage this powerful tool to keep people safe and give victims justice, while still adhering to democratic policing principles. Effective policy creation, where both community members and law enforcement collaborate to define program guardrails and best practices can preserve equity for all community members. Mutually beneficial outcomes can still be realized without regarding privacy and public safety as mutually exclusive concepts. Communities have the ability to define program elements such as appropriate use cases, nature of image repository, data governance, information sharing, and minimum standards regarding both the technology adopted and the personnel authorized to use it.

Myth:

Facial recognition violates basic privacy protections and its utilization will lead to mass surveillance initiatives as seen in other countries, infringing upon people's basic right to assemble and exercise their First Amendment rights.

Reality:

FRT in the United States is used primarily as a back-end investigative tool after a crime has already occurred. FRT platforms as used by law enforcement in the United States, are generally not configured to interface with any other systems to perform real-time analysis using live video surveillance, such as security cameras, drone footage, body-worn cameras, in-car-video, or other sources that would potentially enable the real-time tracking of a specific individual via their facial features.

FRT is utilized by law enforcement in a very limited scope, and usually for specific use cases. Most uses of FRT with a nexus to protests or demonstrations involve efforts to identify a specific actor who committed a crime in the presence of others who were demonstrating. Communities can adopt policies, to include restrictions barring FRT from being used to gather intelligence related to First Amendment-protected speech, associations, or activity. Exceptions can be memorialized in policy authorizing the use of FRT if a crime is committed during such activity. Furthermore, the use of FRT can be specifically limited in scope to investigate only a specific criminal actor.

Myth:

The use of FRT has already been the reason numerous people have been unjustly arrested.

Reality:

While the thought of any innocent person being arrested by police is extremely troubling, FRT has actually been linked to very few cases involving the arrest of the wrong subject. Relative to the scope of utilization, these instances can be characterized statistically as extremely rare. Police leaders and subject matter experts with specific knowledge of the events attribute the outcomes in these cases more to flawed human processes, such as insufficient investigative follow-up, as opposed to flawed technology.

The risk of misidentification due to FRT can be effectively mitigated with certain use policies, including a mandated requirement for trained human review, and adopting the stance that FRT findings constitute

solely an investigative lead. Facial recognition findings should always be based on a two-part process involving both the technology and a person. Appropriately implemented, FRT enhances and accelerates human decision-making, but should never replace it.

If an investigator cuts corners, the remedy to such issues is holding them accountable for their actions and omissions if they fail to appropriately corroborate an investigative lead before taking enforcement action. These circumstances can be mitigated by crafting sound policies and effectively training personnel on concepts such as cognitive bias and potential demographic performance variations of FRT platforms. Many agencies declare in both policy, and on their report of investigative findings that the results do not substantiate a positive identification, independently establish probable cause, or otherwise merit the conclusion that a person is guilty of any criminal act.

Unfortunately, there is no identification technology that is absolutely perfect. However, with a tested accurate platform and trained examiner operating within appropriate policies, FRT can serve as a tool of precision that is far better than many alternatives. Due to a host of factors, FRT is expected to be perfect or better than perfect despite it already having an arguably more reliable record than other forms of identifying unknown subjects, especially eyewitness input, such as photo line-ups, suspect "show-ups", and similar unscientific processes.

Founded in 1992, the Innocence Project seeks to exonerate persons wrongly convicted and to reform the criminal justice system to prevent future injustices. Their research substantiates that relying solely on eyewitness identification is the leading cause of wrongful convictions:

Eyewitness misidentification is a consistent and outsized contributor to wrongful convictions. Nationally, 69% of DNA exonerations — 252 out of 367 cases — have involved eyewitness misidentification, making it the leading contributing cause of these wrongful convictions. Further, the National Registry of Exonerations has identified at least 450 non-DNA-based exonerations involving eyewitness misidentification.

It is not that people have ill intentions or are inherently inattentive. Human memory is complex and memories themselves are "fragile." People perceive events from a single perspective and their own recollection of events can be incomplete or mistaken. Furthermore, over time memories usually deteriorate and can even be altered as we attempt to recall things, especially when influenced by new information. Generally, people tend to be even less accurate when attempting to identify members of a race that is different than their own. Excluding FRT as a tool available to law enforcement leaves investigators with limited, unscientific subjective options to make critical decisions such as who should be arrested.

Reference(s):
https://innocenceproject.org/how-eyewitness-misidentification-can-send-innocent-people-to-prison/

Myth:

Facial recognition algorithms have been found to be consistently unreliable and are still too inaccurate to be used by law enforcement. Facial recognition technology consistently misidentifies women and people of color and only exacerbates the constant surveillance and criminalization that marginalized community members already face.

Reality:

There is understandably no greater concern related to law enforcement's utilization of facial recognition than the fear that it may exacerbate historic challenges to social justice or further strain the relationship between marginalized communities and the police officers who serve them. Even if FRT is sufficiently technically proficient, the perception that it is not, especially in the case of communities of color, can erode public trust and undermine police legitimacy.

Therefore, it is imperative that law enforcement strives to be transparent in the adoption of FRT, how it works, and how it will be specifically utilized. The promise of FRT is that it may serve as a tool of precision to home in on a specific suspect, based on a comparative scientific process, rather than conducting blanket sweeps of a geographic area, imposing on persons who are only guilty of wearing the same clothes or fitting the physical description of a suspect, as reported by a victim or witness.

In 2019, a Montgomery County, Maryland bank was robbed. Detectives arrived on scene and collected images of the suspect from the bank's security system. A potential suspect was quickly developed using FRT. Based on the lead, police responded to a location associated with the subject where they observed him outside still wearing the same clothes that he had on during the robbery. Ideally, this is an example of how FRT can serve not only to locate a dangerous subject at large but also to minimize the impact of enforcement efforts on the broader community.

The perception of facial recognition being biased began with early versions of the technology showing inconsistent accuracy rates across different demographics such as age, gender, and skin color. In the past, these inconsistencies stemmed from multiple causes, including a lack of demographically representative data used to train algorithms, cosmetic appearance modifications, and the physics of light reflection which can specifically contribute to technical challenges in detecting skin tone variations in digital images. Today, the general accuracy of facial recognition technology has improved substantially and like most technology, rapidly continues to more so every year.

A 2018 report is often cited as showing a 35% error rate for FRT in the identification of black women. In fact, those researchers tested older "face gender classification technologies" which are not used for identification by law enforcement. FRT compares two or more images for similarities to help identify a specific person based on their unique facial morphological features. By discussing these technologies interchangeably and citing outdated research that doesn't pertain to current FRT, many publications and even media reports have attributed exaggerated rates of racial disparity to FRT that are misleading. A review of impartial sources examining the current state of the science is imperative. The National Institute of Standards and Technology (NIST) is part of the U.S. Department of Commerce and is highly regarded as an authority to provide impartial scientific assessments. NIST has assessed the accuracy of facial recognition algorithms across different demographic groups and continually updates their findings. The most recent update was issued January 19, 2021, and can be accessed via: https://pages.nist.gov/frvt/reports/11/frvt_11_report.pdf

NIST assessed the false-positive and false-negative rates of facial recognition algorithms using various types of digital images. Several key findings include:

- The most accurate of the algorithms tested have such minute differences between racial/demographic groups that they are considered "undetectable."
- Many substantially outperform non-scientific methods of identification traditionally used by law enforcement.
- The most accurate algorithms have low false positives and false negatives across most demographic groups.
- Algorithms can have different error rates for different demographics but still be highly accurate. An algorithm's rate of demographic variance can be dependent upon system thresholds applied by an end-user.
- Lower performing algorithms do show measurable differences in performance, a critical issue that must be addressed through continual accuracy improvements and exclusion from the use by law enforcement.

The Critics Were Wrong: NIST Data Shows the Best Facial Recognition Algorithms Are Neither Racist Nor Sexist | ITIF

The soundest alternative to banning FRT is adopting appropriate regulations mandating that only thoroughly-trained image analysis algorithms, meeting certain accuracy thresholds be utilized by law enforcement and that assessment by independent testers be funded to ensure continuous improvement of the technology so that only the most effective tools are deployed in the field.

Reference(s):

- https://www2.montgomerycountymd.gov/mcgportalapps/Press_Detail_Pol.aspx?Item_ID=31818
- Motorola Solutions, Facial Recognition Technology, Advancing Public Safety Capabilities, 2021
- https://www.securityindustry.org/2021/07/23/what-science-really-says-about-facial-recognition-accuracy-and-bias-concerns/
- https://www.ibia.org/download/datasets/5124/NIST%20Report%20on%20Facial%20Recognition-%20A%20Game%20Changer.pdf
- https://pages.nist.gov/frvt/reports/11/frvt_11_report.pdf
- https://nvlpubs.nist.gov/nistpubs/ir/2019/NIST.IR.8280.pdf#page=6

Myth:

FRT lends itself to use for certain types of offenses more than others, such as crimes where there is a digital image of a suspect e.g., bank robberies, assaults, and theft on public transportation, burglary, identity theft, etc. The utilization of FRT will lead to over-enforcement of these types of crimes because of the increased opportunities available to leverage facial recognition technology's capabilities.

Reality:

While perhaps well-intentioned, assertions of this nature confuse general criminal justice reform initiatives that desire to see fewer people facing penalties for criminal activity, with the functional equivalent of hindering the police so they don't have the capability to investigate and solve more crime. In essence, this concern acknowledges that FRT works, but seeks to limit the frequency that people are arrested and prosecuted for crimes, and therefore introduced into the criminal justice system. Similar flawed logic is offered by suggesting FRT should only be used to investigate the most egregious crimes or in exigent situations.**Myth:**
Mandating that a search warrant be required for law enforcement to utilize FRT is an effective way to allow the technology to be used while maintaining some degree of oversight.

Reality:

Most successful law enforcement facial recognition programs rely, at least in part, on the utilization of arrest photos as an element of the image repository to check probe photos against. Law enforcement agencies or their partners are often not only the originator of the images but the custodian of them. Requiring law enforcement to obtain a search warrant before conducting an FRT query of these images is procedurally ambiguous. In essence, it necessitates that police departments serve search warrants on themselves or partner law enforcement agencies.

This practice is legally unnecessary and likely to hamper collaboration and/or delay critically important investigations, potentially jeopardizing public safety. Additionally, to obtain a search warrant, one must articulate specific things such as:

- A reasonable description of the person or place be searched and the items to be searched for;
- Facts constituting probable cause supporting the issuance of a warrant;
- And and explanation as to how the thing to be searched for constitutes evidence of the commission of the said offense.

The search warrant requirement would impose barriers to law enforcement agencies from accessing records owned by them, that would otherwise be made readily available to most members of the community upon request.

Myth:

Law enforcement's use of FRT will lead to the creation of digital databases that store unique biometric identifying information, which will be vulnerable to data breaches or misuse.

Reality:

FRT, like any other government database, must rely upon best practices for cyber-security and all users must employ good cyber hygiene to safeguard systems. Most FRT programs do not involve the creation of new stand-alone platforms with unique cyber vulnerabilities. FRT programs are built on existing platforms akin to automated fingerprint systems or other electronic record management systems used by law enforcement.

Specific practices related to data governance, access, retention, and purging procedures are recommended for FRT programs including but not limited to:

- Dedicated role-based oversight of FRT programs to ensure compliance with applicable laws, regulations, standards, and policies related to data governance.
- Identification of an authorizing official to control access to the system and ensure that end-users meet all requirements required by law or articulated in policy prior to being given access.
- Utilizing credentialed, role-based access criteria with least privilege enforcement as appropriate, within all FRT platform access points.
- Ensuring that user accounts and associated authorizations are validated regularly and maintained in a secure "need-to-know" status - deleting any accounts found to be inactive.
- Enforcing the use of multi-factor authentication access controls.
- Ensuring protocols are followed to purge facial recognition data (including probe images) in accordance with an adopted retention policy.
- Conducting and documenting random audits to ensure user compliance and system functionality.
- Maintaining current, supported operating systems and frequently patching systems based on the manufacturer's recommendations to safeguard against new malware, viruses, and other vulnerabilities.
- Suspending or rescinding access if a user is found to be in non-compliance with cyber policies.