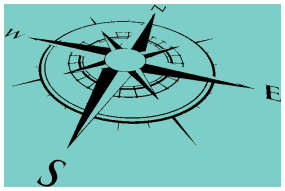
Forensic Technology: Algorithms Used in Federal Law Enforcement

GAO was asked to conduct a technology assessment on the use of forensic algorithms in federal law enforcement. GAO is conducting this assessment in two phases. The first phase describes algorithms being used by federal law enforcement agencies and how these technologies work. The second phase will assess the approaches and challenges related to how federal law enforcement agencies apply these technologies and will identify policy options for addressing these challenges going forward.

What GAO found — Federal law enforcement agencies GAO reviewed are primarily using three types of forensic algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual: probabilistic genotyping, latent print (fingerprint and palm print) analysis, and face recognition. To a lesser extent, agencies also use algorithms to compare iris images, speech, and handwriting. Each type of algorithm uses different characteristics in its assessment. For example, probabilistic genotyping uses statistics to analyze biological samples found during a criminal investigation to assist in comparisons to a known DNA sample taken from a suspect, or to DNA data profiles from a database of known persons. The Federal Bureau of Investigation currently uses probabilistic genotyping and latent fingerprint algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual and face recognition to generate investigative leads. The National Institute of Standards and Technology and other organizations have developed standards to facilitate transmission of data between agencies.

Contents



United States Government Accountability Office (GAO)

Stakeholder(s):

Karen L. Howard, PhD :
Director, Science, Technology Assessment, and Analytics

Members of Congress

The Honorable Eddie Bernice Johnson :
Chairwoman, Committee on Science, Space, and Technology, House of Representatives

The Honorable Frank Lucas :
Ranking Member, Committee on Science, Space, and Technology, House of Representatives

The Honorable Carolyn Maloney :
Chairwoman, Committee on Oversight and Reform, House of Representatives

The Honorable Mark Takano :
House of Representatives

Federal Law Enforcement Agencies

Information Providers :
In conducting this assessment, GAO obtained information from the National Institute of Standards and Technology, the Department of Justice, the Department of Homeland Security, and the Department of Defense; convened an interdisciplinary panel of 16 experts with assistance from the National Academies of Sciences, Engineering, and Medicine; interviewed additional stakeholders, including nonprofit groups and legal experts; and reviewed relevant literature and case law.

National Institute of Standards and Technology

Department of Justice

Department of Homeland Security

Department of Defense

National Academies of Sciences, Engineering, and Medicine

Forensic Algorithms Experts :
Appendix II: Expert Meeting Participation We collaborated with the National Academies of Sciences, Engineering, and Medicine to convene a 1½-day meeting of 16 experts on forensic algorithms used in federal law enforcement. The meeting was held on January 15-16, 2020 in Washington, D.C. Many of these experts provided us with additional assistance throughout our work, including sending additional information for our review or reviewing our draft report for technical accuracy. The experts who participated in this meeting are listed below.

Sarah Chu :
Senior Advisor on Forensic Science Policy, Innocence Project

Michael Coble :
Associate Director of the Center for Human Identification, University of North Texas Health Science Center

Robert English :
Special Counsel, Science and Technology Branch, Federal Bureau of Investigation

Tamara Giwa :
Attorney, Assistant Federal Defender, Federal Defenders of New York

Patrick Grother :
Scientist, Information Technology Laboratory, Information Access Division, Image Group, National Institute of Standards and Technology

William Guthrie :
Division Chief, Statistical Engineering Division, National Institute of Standards and Technology

Karen Kafadar :
Commonwealth Professor and Chair of Statistics, University of Virginia

Dan E. Krane :
Professor and Interim Dean, Wright State University

James Loudermilk :
Senior Director, Innovation and Customer Solutions, IDEMIA National Security Solutions

Anne May :
Biometric Support Center Program Manager, Office of Biometric Identity Management, Department of Homeland Security

Mark Perlin :
Chief Scientific and Executive Officer, Cybergenetics

Peter M. Vallone :
Scientist, Biomolecular Measurement Division, National Institute of Standards and Technology

Kit Walsh :
Senior Staff Attorney, Electronic Frontier Foundation

— continued next page

Stakeholders (continued)

James L. Wayman :
Editor-in-Chief, IET Biometrics Journal

Rebecca Wexler :
Assistant Professor, University of California, Berkeley School of Law

Michael Yates :
Senior Technical Advisor on Biometrics, Science and Technology Branch, Federal Bureau of Investigation

GAO Staff

Karen L. Howard, PhD :
GAO contact — In addition to the contact [the following GAO staff] made key contributions to this report.

Sushil Sharma :
Assistant Director

Allen Chan :
Analyst-in-Charge

Mariel Alper

Nora Adkins

Virginia Chanley

Hayden Huang

Eliot Fletcher

Anika McMillon

Eleni Orphanides

Ben Shouse

Vision

Understanding of the use of forensic algorithms in federal law enforcement.

Mission

To conduct a technology assessment on the use of forensic algorithms in federal law enforcement.

1. Law Enforcement

Use forensic algorithms for law enforcement

Federal Law Enforcement Agencies Primarily Use Three Kinds of Forensic Algorithms — Federal law enforcement agencies we reviewed primarily use probabilistic genotyping, latent print, and face recognition algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual. To a more limited extent, agencies also use algorithms to compare iris images, speech, and handwriting.

1.1. Genotypes, Finger Prints & Faces

Use probabilistic genotyping, latent print, and face recognition algorithms

Use of probabilistic genotyping, latent print, and face recognition algorithms — We found that federal law enforcement agencies we reviewed use three main types of forensic algorithms to help assess whether or not evidence collected in a criminal investigation may have originated from an individual: probabilistic genotyping, latent print analysis, and face recognition algorithms.

Stakeholder(s):

DOJ :
DOJ uses forensic algorithms in its criminal investigations.

FBI Laboratory Services :
The FBI's Laboratory Services uses PGS to assist in the interpretation of DNA evidence while investigating criminal cases.

FBI Criminal Justice Information Services :
The FBI's Criminal Justice Information Services—a division that provides tools and services to law enforcement, national security, intelligence community partners, and the general public—has a repository of biometric and criminal history data known as the Next Generation Identification (NGI) System. The FBI uses this system in combination with latent print algorithms and face recognition algorithms to determine whether evidence may have originated from an individual.

Drug Enforcement Administration :
In addition, DOJ's Drug Enforcement Administration uses NGI to help with identifying latent prints.

DHS :
DHS uses forensic algorithms to support homeland security missions and criminal investigations.

Office of Biometric Identity Management (OBIM) :
For example, the Office of Biometric Identity Management (OBIM) uses forensic algorithms in its Automated Biometric Identification System called IDENT for face recognition and latent print analysis.

Transportation Security Administration (TSA) :
The Transportation Security Administration (TSA) uses FBI and OBIM biometric fingerprint algorithms as a part of its civil background investigations.

U.S. Custom and Border Protection (CBP) :
TSA is also testing the U.S. Custom and Border Protection (CBP) biometric facial algorithms to verify passenger identities, for passengers who have opted into the TSA biometric testing program. Since 2017, TSA has conducted a series of pilot tests—some in partnership with CBP—to assess the feasibility of using face recognition to automate traveler identity verification at airport security checkpoints.

DOD :
Finally, DOD also uses forensic algorithms for criminal investigations.

Defense Forensic Science Center (DFSC) :

Within DOD, the Defense Forensic Science Center (DFSC) performs forensic analyses, using latent print analysis and probabilistic genotyping algorithms. An official from a unit of DFSC told us that the agency also submits evidence to other agencies for forensic database searching support, as well as developing its own internal algorithms. For example, DFSC has its own software package for latent print analysis, which can be used to provide statistical support for results of manual comparisons. DFSC recently made this available as open source software. For DNA analysis, DFSC uses software to assist with the separation of mixed DNA profiles—those that contain DNA from more than one individual—and a separate program to assist with certain calculations, such as inferring the biological sex of an individual based on evidence collected during a criminal investigation.

1.1.1. DNA

Evaluate DNA evidence.

Probabilistic genotyping software (PGS) — PGS may be used to evaluate single source or complex DNA evidence. PGS posits two competing hypotheses: A) the DNA evidence is the result of contributions by a person of interest and other unknown, unrelated individuals and B) the DNA evidence is the result of contributions by unknown individuals. It provides a likelihood that the observed data resulted from each scenario, giving a likelihood ratio of hypothesis A versus hypothesis B. PGS is more effective than traditional DNA analysis when the DNA is from two or more individuals or when DNA from some or all of the contributors is present in low quantities. Unlike conventional approaches, PGS can attach a number that statistically measures the strength of the evidence when a DNA sample is from an unknown number of contributors where it is possible that some of the DNA from one or more contributors failed to be detected.

Stakeholder(s):

Investigators :
Investigators generate profiles from the crime scene evidence sample. Separately, they may also generate profiles from samples taken from one or more persons of interest called a reference sample. This process is used in multiple types of DNA analysis and is not unique to PGS.

1.1.1.1. Extraction

Extract genetic material from both the evidence and reference samples.

The first step in DNA analysis usually involves the extraction of genetic material from both the evidence and reference samples (see fig. 1). Commercially available kits are then used to repeatedly copy specific regions of human DNA that are likely to differ in lengths across individuals. These amplified pieces of DNA are separated by size. The resulting mix of fragment lengths represents a profile, also known as a genotype. The profile is normally represented as a series of peaks on a graph known as an electropherogram... What distinguishes PGS is the steps that follow.

1.1.1.2. Mathematical Model

Use a mathematical model encoded to estimate the likelihoods associated with two competing hypotheses.

In the first of these, investigators use a mathematical model encoded in PGS software to estimate the likelihoods associated with two competing hypotheses, such as hypotheses A and B described above. PGS mathematically compares the crime scene profile with many hypothetical profiles based on various possible genotype combinations. This process allows the software to assess the relative likelihood that various genotype combinations contributed to the crime scene sample. It also allows the software to separate out genotypes of individual contributors. The first steps of PGS do not use a genotype from an individual in question.

1.1.1.3. Variables

Examine many variables simultaneously.

The most sophisticated PGS software models examine many variables simultaneously and can be very computationally intensive. They often do this through a computer simulation that considers a large number of contributor combinations of, for example, two-, three-, and four-person mixtures.

1.1.1.4. Probabilities

Estimate the likelihood that each of the combinations best explains the results.

Using a set of parameters and mathematical modeling of the data, the computer estimates the likelihood that each of these combinations best explains the results.

1.1.1.5. Simulation

Simulate large numbers of possible states of the variables.

For those hypotheses that posit a contribution from a specific individual, the software will simulate large numbers (often hundreds of thousands) of possible states of those variables and return an estimate of the probability that the test results from the evidence sample would appear as they did if the individual had contributed to it.

1.1.1.6. Estimations

Return an estimate of the probability that the test results from the evidence sample would appear as they did if the individual had contributed to it.

1.1.1.7. Alternative Theories

Perform the same simulation using the same model to estimate the probability that the test results from the evidence sample would appear as they did if a different theory of the case was correct.

PGS software will then perform the same simulation using the same model to estimate the probability that the test results from the evidence sample would appear as they did if a different theory of the case was correct. These probability estimates are highly dependent on the models and their variables. However, the ratio of the likelihoods of observing the data under two alternative hypotheses for a single evidence sample can be helpful if they have been estimated by software using the same models and variables. If the ratio of the likelihood of hypothesis A to that of hypothesis B is greater than 1, the test results are more consistent with hypothesis A. Likelihood ratios of less than one suggest that the test results are more consistent with B. (A likelihood ratio is not the probability that the individual's DNA is actually contained in a DNA mixture.)

1.1.1.8. Comparison

Compare the known profile of the individual to the weighted combinations to determine if the individual can be explained as being a contributor or noncontributor.

Once the algorithms have determined the weighted optimal combinations of contributors to the mixture (independent of the profile from the person of interest), investigators compare the known profile of the individual to the weighted combinations to determine if the individual can be explained as being a contributor or noncontributor to the mixture. The result is a number that statistically measures the strength of the evidence, which can be taken into consideration by investigators.

1.1.2. Latent Prints

Compare latent prints with tenprints.

Latent print analysis — A latent print can be a partial or incomplete print left on a surface and then recovered during a criminal investigation. It may be smudged or distorted.

Stakeholder(s):

Investigators :
In latent fingerprint analysis, investigators compare a latent print with a tenprint—a set of prints from all 10 of an individual's fingers, taken under controlled circumstances. Additionally, investigators can compare a latent palm print with known palm prints—a set of four or six prints of known palm data.

Human Examiners :
The latent print is digitally scanned, its details or minutiae marked by a human examiner, and the scan is uploaded into the Automated Fingerprint Identification System (AFIS), which uses multiple algorithms to analyze the print.

Automated Fingerprint Identification System :
The algorithms can improve image quality and read the many minutiae specific to a fingerprint or palm print. The algorithms also compare the layout of minutiae detected in the latent print to those found in a tenprint and palm print database of known individuals. This comparison provides a list of individuals who may be the source of the latent print found during an investigation.

Latent Print Experts :
An expert independently compares this list of candidates and, based on their own judgment, reaches an identification, exclusion, or inconclusive decision (see fig. 2).

1.1.3. Faces

Compare images of unknown individuals against a database of images of known persons.

Face recognition — If an image of an unknown individual associated with a criminal investigation is available, face recognition could compare it against a database of images of known persons. For example, the image of a known individual is captured under controlled conditions. During comparison to photos of known individuals, a probe photograph (a photo of an unknown individual) is compared against the photos of known individuals in the NGI System that were obtained in controlled conditions. The separate enrollment and matching phases usually depend on multiple algorithms.

1.1.3.1. Detection & Orientation

Detect faces in photos and orient them.

For example, in enrollment an initial algorithm will detect the face in the probe photo and orient it.

1.1.3.2. Templates

Analyze the entire sets of pixels across the images to generate mathematical representations of the faces.

A second algorithm will then analyze the entire set of pixels across the image to generate a mathematical representation of the face. This mathematical representation of the face is called a "template."

1.1.3.3. Comparisons

Compare the probe templates to a gallery database of known templates.

A matching algorithm is then used to compare the probe template to an entire gallery database of known templates.

1.1.1.4. Rankings

Rank a candidate list of faces from the database from most to least similar to the probe photograph.

This process may use an AI technology known as convolutional neural networks. A probe photograph of an individual is digitized into a mathematical language that forms a template. A program runs this information through several algorithms and compares it to a database of known facial images. This results in a candidate list of faces from the database, with a ranking from most to least similar to the probe photograph. Convolutional neural networks may use multiple layers to analyze templates. After the image is filtered through the layers, the resulting mathematical patterns are compared with those extracted similarly from face images in a known database. This comparison method does not use facial features (e.g., eye distance or nose size), but rather uses mathematical aspects of a digitized image. This comparison generates a "similarity score" which is specific to individual algorithms. Once the probe photo has been compared against the entire database, the system will present to the user a candidate list of photos ranked from highest similarity score to lowest.

1.1.1.5. Matches

Generate a list of best-matched photos.

If the algorithm identifies multiple likely candidates, it will generate a list of best-matched photos. The length of this candidate list is determined by the system operator, but typically is between 20 and 100. In contrast, the system could return no candidates if no database photos are found to be sufficiently similar to the probe photo (see fig. 3)

1.2. Other Algorithms

Use of other algorithms — To a lesser extent, federal law enforcement agencies we reviewed also use other algorithms to assess whether or not evidence collected in a criminal investigation may have originated from an individual, such as algorithms for comparing iris images, voice recordings, and handwriting. • Iris recognition algorithms. Iris recognition algorithms compare images of an individual's iris to a database of iris images. • Voice recognition algorithms. • Handwriting recognition algorithms.

In addition to these algorithms that are in use or being tested, agencies are researching and developing additional algorithms they may use in the future. Our expert meeting participants identified gait analysis and genetically variant peptide analysis algorithms as methods being researched.

Stakeholder(s):

DHS :
DHS's OBIM uses iris methods as part of its IDENT system.

FBI :
FBI officials said that the agency has a pilot program to develop iris matching algorithms, which it will soon incorporate into the NGI System as the National Iris System.

U.S. Secret Service :
Officials with the U.S. Secret Service told us that it has the ability to compare a recording of an unknown speaker with one or more recordings of known speakers to help investigators identify the unknown speaker... U.S. Secret Service officials said their agency can use a computer algorithm to compare manually collected digital measurements of hand-writing characteristics to previously collected measurements, some of which may be attributed to a known author.

OBIM :
OBIM is also exploring the use of automatic voice recognition algorithms.

NIST :
According to NIST publications and FBI officials, the two agencies previously collaborated on research on image-based tattoo recognition algorithms.

1.2.1. Irises

Compares iris images associated with an investigation or a person of interest to a database of known iris image patterns.

Iris recognition — Iris recognition compares an iris image associated with an investigation or a person of interest to a database of known iris image patterns.

1.2.1.1. IrisCode

Use handcrafted algorithms to convert digital images into mathematical patterns of the digitized irises.

The iris recognition software uses handcrafted algorithms (as opposed to AI) to convert the digital image into mathematical patterns of the digitized iris, known as an IrisCode.

1.2.1.2. Comparisons

Compare the mathematical patterns of the IrisCode to other IrisCodes of known individuals.

The mathematical patterns of the IrisCode are compared to other IrisCodes of known individuals.

1.2.1.3. Probabilities

Determine whether the two things being compared are likely to be from the same or different individuals.

Using statistical comparisons, the algorithms determine whether the two things being compared are likely to be from the same or different individuals.

1.2.2. Voices

Isolate and analyze voice samples associated with investigations.

Voice recognition — A voice sample that is associated with an investigation can be isolated and analyzed by voice recognition software.

1.2.2.1. Features

Use forensic algorithms to find abstract, short-term features in voice samples.

In a forensic case, the voice sample from the investigation and a known voice sample are provided to software that uses forensic algorithms to find abstract, short-term features.

1.2.2.2. Scores

Produce numeric scores for the similarity between the investigative and known samples.

These abstract features can be put through further layers of processing and then compared to produce a numeric score giving the similarity between the investigative and known samples.

1.2.2.3. Likelihood

Output a likelihood ratio that speech samples were spoken by the same speaker or different speakers.

The automatic system will output a likelihood ratio (i.e., the likelihood of observing the measured similarity between speech samples assuming that they were spoken by the same speaker or different speakers). These results can be fragile, in the sense of being dependent on confusing factors such as type of microphone, background noise, and transmission channel.

1.2.3. Handwriting

Collect and digitally scan handwriting samples.

Handwriting recognition — As with latent prints, handwriting samples associated with an investigation can be collected and digitally scanned.

1.2.3.1. Measurements

Use forensic algorithms to perform digital measurements of the handwriting features.

The handwriting samples are uploaded into software that uses forensic algorithms to perform digital measurements of the handwriting features that have been manually marked by an expert.

1.2.3.2. Comparison

Compare evidence to writing samples.

Comparisons can be made between evidence and either a known or unknown writing sample. An expert then reviews the results.

2. Data Standards

Use data standards to transmit evidence

Stakeholder(s)

ISO/IEC :
INCITS/ISO/IEC 19794 (parts to be superseded by parts of the ISO/IEC 39794 series) — International — Probabilistic genotyping software, latent prints, face recognition

ANSI :
ANSI/NIST-ITL 1-2011, Update: 2015 — National — Probabilistic genotyping software, latent prints, face recognition — ANSI/NIST

FBI :
Electronic Biometric Transmission Specification — Agency — Latent prints, face recognition

DOD :
Electronic Biometric Transmission Specification — Agency — Latent prints, face recognition

Agencies Use Data Standards to Help Them Transmit Evidence — We found four standards that agencies use to facilitate the transmission of data between agencies for examination by PGS, latent print, and face recognition algorithms. In our review, we found one international standard, one U.S. standard, and two standards specific to a federal agency.

2.1. Biometric Systems

Enable the interoperability and data interchange among biometric applications and systems.

The international standard was developed jointly by the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC) to enable the interoperability and data interchange among biometric applications and systems. It includes guidance for fingerprints, facial images, and DNA data (used for PGS).

Stakeholder(s):

International Organization for Standardization (ISO)

International Electrotechnical Commission (IEC)

2.2. Prints, Faces & DNA

Standardize data for prints, facial images, and DNA.

NIST developed a standard for prints, facial images, and DNA data. The ANSI/NIST standards were developed for federal agencies to specify a common format for data exchange across jurisdictional lines or between dissimilar systems made by different manufacturers. According to NIST officials, these standards were developed with criminal justice in mind.

Stakeholder(s):

NIST

2.3. EBTS

Standardize data for encoding and transmitting biometric image, identification, and arrest data.

The FBI developed a standard for electronically encoding and transmitting biometric image, identification, and arrest data known as the Electronic Biometric Transmission Specification (EBTS). This standard, based on the ANSI/NIST-ITL 1-2011, Update: 2015 standard, applies to the FBI's database of biometric and criminal history information (NGI System) and helps ensure that the data format for prints and facial images matches that of the NGI System.

Stakeholder(s):
FBI

2.4. DOD Biometrics

Standardize the interface with DOD's biometric database.

Similarly, DOD developed the EBTS, based on the ANSI/NIST-ITL 1-2011, Update: 2015 standard, to interface with DOD's biometric database.

Stakeholder(s):
DOD

Administrative Information

Start Date:
End Date:

Publication Date: **2020-05-15**
Source: https://www.gao.gov/assets/710/706849.pdf

Submitter:

Given Name: Owen
Surname: Ambur
Email: Owen.Ambur@verizon.net
Phone: