Introduction to NIST FRVT

Published: March 2023

NIST's Face Recognition Vendor Test (FRVT) is the most respected industry benchmark for companies building and using face recognition technology, and its benefit for the industry is undeniable. However, with the number of tests and complexity of the metrics used, the results can be challenging to parse and understand. This blog post aims to give a high-level look into NIST FRVT and answer some of the frequently asked questions we've seen from partners and peers.

What is NIST FRVT?

The National Institute of Standards and Technology (NIST) is a non-regulatory agency of the U.S. Department of Commerce that develops standards and guidelines for various industries, including biometrics.

The Face Recognition Vendor Test (FRVT) is a set of evaluations conducted by the National Institute of Standards and Technology (NIST) to assess the performance of face recognition algorithms submitted by vendors from around the world. These evaluations aim to provide government agencies, industry, and the research community with unbiased, independent information about the performance of face recognition algorithms, empowering organizations to make informed decisions when selecting a face recognition system.

Background

The NIST Face Recognition Vendor Test (FRVT) program was established in 2000 to evaluate the performance of face recognition algorithms. The program has evolved over the years to include a variety of evaluations. These evaluations are designed to assess the performance of face recognition algorithms under a wide range of conditions, including varying lighting conditions, poses, and image qualities.

Different NIST FRVT tests

NIST FRVT tests are divided into two categories: Ongoing and Special. Ongoing tests are conducted regularly to evaluate the performance of face recognition algorithms submitted by vendors. In contrast, special tests are conducted on an as-needed basis to evaluate specific use cases or scenarios.

In all cases, NIST FRVT is performed on closed “Black box” datasets. No vendors have access to any of the test images, which ensures that submitted software is not trained specifically for NIST FRVT.

Ongoing **NIST FRVT** tests include:

- **FRVT 1:N:** Considered the most challenging face recognition scenario, this test evaluates the performance of face recognition algorithms in a one-to-many matching, where a probe image is compared to multiple gallery images.
- **FRVT 1:1:** Evaluates the performance of face recognition algorithms in a one-to-one matching scenario, where a probe image is compared to a single gallery image. The 1:1 test also includes specific results and metrics for demographic performance, providing accurate insight into how well facial recognition algorithms perform across genders, age groups, and countries of origin.
- **FRVT MORPH:** Evaluates technologies detecting facial morphing (morphed/ blended faces) in still photographs, as well as the resistance of face recognition algorithms against morphing.
- **FRVT QUALITY:** Evaluates image quality algorithms by testing them on large sets of images and comparing their outputs with face recognition outcomes.

- **FRVT PAD:** NIST's newest upcoming test that evaluates the ability of face recognition algorithms to detect presentation attacks or spoofs.

Examples of special NIST FRVT tests include:

- **FRVT Paperless Travel:** Tests face recognition algorithms in a paperless travel application through simulations of face recognition for boarding an aircraft and traversing an airport security point.
- **FRVT Face Mask Effects:** Evaluates the performance of face recognition algorithms on masked faces. This test was created due to the widespread use of facial masks during the COVID-19 pandemic.

Out of the different types of NIST FRVT tests, the 1:1 and 1:N tests are considered to be the most general standard for facial recognition accuracy and performance.

How to read NIST FRVT results?

NIST publishes FRVT results in three formats: leaderboards, reports, and "report cards" for each vendor. The NIST FRVT leaderboards list vendors in a searchable, interactive table, ordered based on the performance with the most commonly used datasets. The report files include more detailed tables with all vendors, and usually show the performance of an algorithm based on the error rates with different standardized datasets, with lower error rates meaning a more accurate algorithm. Each vendor also has an individual scorecard that collects its performance data in a single report.

Vendors can submit their algorithms multiple times to the ongoing tests, and the result tables often include multiple submissions from vendors, with individual report cards for each submission.

To understand the results, a reader must understand a few essential terms:

- **False Negative Identification Rate (FNIR):** Performance metric used in 1:N, showing the rate at which the face recognition system fails to identify a person who is present in the gallery.

- **False Positive Identification Rate (FPIR):** Performance metric used in 1:N, showing the rate at which the face recognition system incorrectly identifies a person as being present in the gallery when they are not.
- **False Match Rate (FMR):** Performance metric used in 1:1, showing the rate at which the face recognition system incorrectly matches two individuals.
- **False non-match rate (FNMR):** Performance metric used in 1:1, showing the rate at which the face recognition system fails to match two individuals who are the same person.
- **Gallery & probe types:** Types of face images used for the testing. Some common types of image types are:
    - **Visa images** represent images taken as a part of a visa application process. They are typically high-resolution, frontal-facing images used to evaluate face recognition algorithms' performance on high-quality photos.
    - **Border images** represent photos taken at border crossings, such as passport control or customs. They are typically lower resolution and may be taken at an angle, making them more challenging for face recognition algorithms to process.
    - **Mugshot images** represent photos taken by law enforcement agencies, typically after an arrest. They are typically frontal-facing and may be taken under controlled lighting conditions.
    - **Mugshot-Webcam** simulates the real world scenario of using a high quality enrollment photo and poor quality probe image in a law enforcement use case.
    - **Visa-Border** combination is typically used as the default ranking in leaderboards, and it simulates the real world scenario of a visa submission and an identity check at the border.
    - **Visa-Kiosk** combination simulates the real world scenario of a visa submission and an identity check at a self-serve kiosk.
    - **Border ΔT ⩾ 10 YRS:** This is a subset of border images where the image was taken at least 10 years apart from the visa image, this is used to evaluate the performance of face recognition algorithms on pictures of the same person taken at different ages.

It's crucial to understand the relationship between false positive and false negative matches to understand the performance of a face recognition algorithm. In an access control scenario, false positives are a security risk, as a false positive can mean a person's face is matched with another person's identity.

This is why the false positive rate is typically set at a very low threshold in NIST FRVT. In the same scenario, while not a security risk, false negatives increase inconvenience and processing time, as a person's face is incorrectly marked as not matching their identity.

Necessary security thresholds differ between use cases, and even though a lower error rate is generally better, a balance should be struck between security and user experience.

Important considerations

NIST FRVT evaluations are vital for assessing face recognition algorithms' performance. The variations of the tests used databases, and probe image types, can provide an invaluable asset for companies comparing face recognition vendors and their algorithms. To best utilize NIST FRVT results, companies should consider how their use case compares with the different NIST FRVT testing types, probe images, and security thresholds.

It is also essential to understand that not all vendors offer their NIST-tested algorithms to partners or customers for commercial use but might have separate algorithms for NIST testing and operational use. Organizations comparing face recognition vendors should always ask their partners if their NIST-tested algorithm is the same that's commercially available to partners.

Lastly, it's important to note that NIST FRVT occurs under specific conditions, which may or may not reflect the real-world scenarios in which face recognition systems are used. In operational use, the performance of algorithms can be affected by factors such as lighting conditions, the quality of the camera, the network connection, and the processing power of the device, which might not be fully captured by NIST FRVT evaluations. These conditions may be better evaluated by other industry benchmarks such as the Department of Homeland Security's Biometric Technology Rally, or internal use case-specific testing.

The DHS Biometric Technology Rally is an evaluation program conducted by the United States Department of Homeland Security (DHS) to assess the performance of biometric technology, including face recognition, fingerprint recognition, and iris recognition.

The Biometric Technology Rally program is designed to provide government agencies with information about the performance of biometric technology under various conditions, including different lighting and weather conditions, and to help inform decisions about deploying biometric technology. For more information, see the **Maryland Test Facility website**.

Conclusion

NIST FRVT evaluations provide invaluable information on the performance and accuracy of facial recognition algorithms and vendors. While the results can sometimes be challenging to understand, the variety in the testing programs provides crucial performance data for organizations comparing face recognition vendors for different use cases.

For more information or to schedule a demo, please contact us at paravision.ai/HID.