dermalog+0002

DERMALOG Identification Systems GmbH

Evaluation of Latent Friction Ridge Technology (ELFT)

Technical performance report of automated latent fingerprint feature extraction and search software.

Last Updated: 07 February 2023

Contents

1 Participation Information **2**

2 Timing Sample **3**

3 Metrics **9**

4 Non-mated Distractor Subjects **11**

5 FBI Laboratory **12**

6 FBI-Provided Solved Dataset #1 **17**

7 Michigan State Police **22**

Not Human Subjects Research
The National Institute of Standards and Technology's Research Protections Office reviewed the protocol for this project and determined it is "not human subjects research" as defined in 15 CFR 27, the Common Rule for the Protection of Human Subjects.

NATIONAL INSTITUTE OF STANDARDS AND TECHNOLOGY
U.S. DEPARTMENT OF COMMERCE

1 Participation Information

1.1 Names

Information in this section is provided by the participant.

- **Participant Name:** DERMALOG Identification Systems GmbH
- **ELFT Identifier:** `dermalog+0002`
- **Search:**
  - **Marketing Name:** DERMALOG
  - **CBEFF Product Owner:** `0x000D`

1.2 Dates

- **Participation Agreement Date:** 23 October 2020
- **First Submission Date:** 20 October 2022 (as version `0001`)
- **Final Submission Date:** 25 October 2022 (as version `0002`)
- **Validation Date:** 25 October 2022
- **Completion Date:** 15 November 2022
- **Report Last Updated Date:** 07 February 2023

1.3 Supplied Libraries and Configurations

Testing was completed using *Ubuntu 20.04.3 LTS*. Files provided by DERMALOG Identification Systems GmbH are listed in Table 1.

Table 1: Information regarding library and configuration files provided as part of `dermalog+0002`.

Filename	MD5 Checksum	Size (MB)
libelft_dermalog_0002.so	`001ab36d5881fa20e157e8c4b94e1a2d`	92.6

2 Timing Sample

A fixed sample of images was randomly and proportionally selected from the ELFT datasets in 2021. The sample is used to assess whether an implementation adheres to the computational speed requirements from the ELFT Test Plan. These values are chosen in such a way that allows the implementation flexibility while allowing NIST to complete the evaluation in a reasonable amount of time. If an implementation exceeds the maximum allowable duration, the participant will be asked to reduce the processing time of their software prior to NIST completing the evaluation. As such, *all* published ELFT submissions conform to the published speed requirements.

2.1 Processor Details

All measurements in this section were performed on a machine equipped with Intel Xeon Gold 6254 Central Processing Units (CPUs). Each CPU features a 3.10 GHz base frequency and 24.75 MB of cache. Timing tests are all **single threaded**—implementations are not permitted to use more than one thread during any function measured here. As such, these values can be used to understand expected scaled performance. NIST testing code embraces the single-threaded nature of implementations to `fork` processes during other non-timed portions of this evaluation, allowing participants to write thread-unsafe code while still using NIST resources to their maximum efficiency. This CPU supports executing several families of processor intrinsic functions, including AVX-512[1].

2.2 Composition

Table 2 shows the quantity of each type of fingerprint image comprising the timing sample dataset.

Table 2: Number of images of each generalized finger position comprising the timing sample dataset.

Image Type	Quantity
Latent	250
Four Finger	476
Full Palm	40
Partial Palm	47
Single Finger	2 784

2.3 Feature Extraction

Features were extracted from all images depicted in Table 2 and stored in templates. If a sample contained EFS data, it was included during this test.

2.3.1 Template Size

Table 3 and Figure 1 show the distribution of file sizes of templates. Failures of any kind reported during template generation result in NIST code writing 0 byte files. These files are excluded from the template size analysis in this section.

[1]The complete set of advertised CPU flags is `fpu`, `vme`, `de`, `pse`, `tsc`, `msr`, `pae`, `mce`, `cx8`, `apic`, `sep`, `mtrr`, `pge`, `mca`, `cmov`, `pat`, `pse36`, `clflush`, `dts`, `acpi`, `mmx`, `fxsr`, `sse`, `sse2`, `ss`, `ht`, `tm`, `pbe`, `syscall`, `nx`, `pdpe1gb`, `rdtscp`, `lm`, `constant_tsc`, `art`, `arch_perfmon`, `pebs`, `bts`, `rep_good`, `nopl`, `xtopology`, `nonstop_tsc`, `cpuid`, `aperfmperf`, `pni`, `pclmulqdq`, `dtes64`, `monitor`, `ds_cpl`, `vmx`, `smx`, `est`, `tm2`, `ssse3`, `sdbg`, `fma`, `cx16`, `xtpr`, `pdcm`, `pcid`, `dca`, `sse4_1`, `sse4_2`, `x2apic`, `movbe`, `popcnt`, `tsc_deadline_timer`, `aes`, `xsave`, `avx`, `f16c`, `rdrand`, `lahf_lm`, `abm`, `3dnowprefetch`, `cpuid_fault`, `epb`, `cat_l3`, `cdp_l3`, `invpcid_single`, `intel_ppin`, `ssbd`, `mba`, `ibrs`, `ibpb`, `stibp`, `ibrs_enhanced`, `tpr_shadow`, `vnmi`, `flexpriority`, `ept`, `vpid`, `ept_ad`, `fsgsbase`, `tsc_adjust`, `bmi1`, `avx2`, `smep`, `bmi2`, `erms`, `invpcid`, `cqm`, `mpx`, `rdt_a`, `avx512f`, `avx512dq`, `rdseed`, `adx`, `smap`, `clflushopt`, `clwb`, `intel_pt`, `avx512cd`, `avx512bw`, `avx512vl`, `xsaveopt`, `xsavec`, `xgetbv1`, `xsaves`, `cqm_llc`, `cqm_occup_llc`, `cqm_mbm_total`, `cqm_mbm_local`, `dtherm`, `ida`, `arat`, `pln`, `pts`, `pku`, `ospke`, `avx512_vnni`, `md_clear`, `flush_l1d`, `arch_capabilities`

Table 3: Template file size summary statistics as seen on the Timing Sample dataset, in kB.

Image Type	Minimum	25%	Median	Mean	75%	Maximum	Failures	Attempts
Latent	0.3	0.7	1.0	1.3	1.6	4.4	2	250
Single Finger	0.4	2.0	2.4	2.4	2.9	5.5	1	2 784
Four Finger	1.5	4.3	4.9	4.9	5.5	8.0	0	476
Partial Palm	NA	NA	NA	NA	NA	NA	47	47
Full Palm	NA	NA	NA	NA	NA	NA	40	40

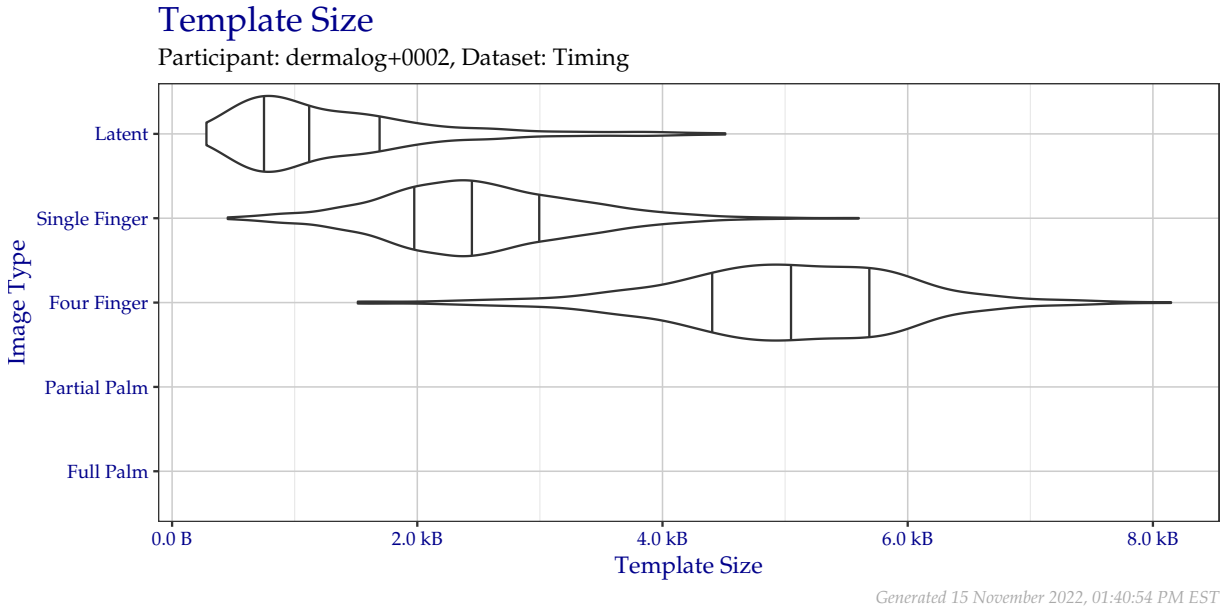

Figure 1: Violin plot of template file sizes as seen on the Timing Sample dataset. Vertical lines from left to right indicate the 25%, 50%, and 75% quantiles respectively.

2.3.2 Template Creation Duration

Table 4 and Figure 2 show the distribution template creation durations in seconds. Failures of any kind reported during template generation result in NIST code writing 0 byte files, but only after the template creation method returns. These times are included in the template creation duration analysis in this section.

Table 4: Duration of template creation in seconds for images from the Timing Sample dataset.

Image Type	Minimum	25%	Median	Mean	75%	Maximum	Failures	Attempts
Latent	0.0	0.3	0.4	0.8	0.7	14.5	2	250
Single Finger	0.0	0.3	0.4	0.4	0.5	0.7	1	2 784
Four Finger	0.1	0.5	0.5	0.5	0.6	0.9	0	476
Partial Palm	0.7	3.3	4.2	4.2	5.0	12.1	47	47
Full Palm	7.7	9.4	10.0	10.1	10.8	13.0	40	40

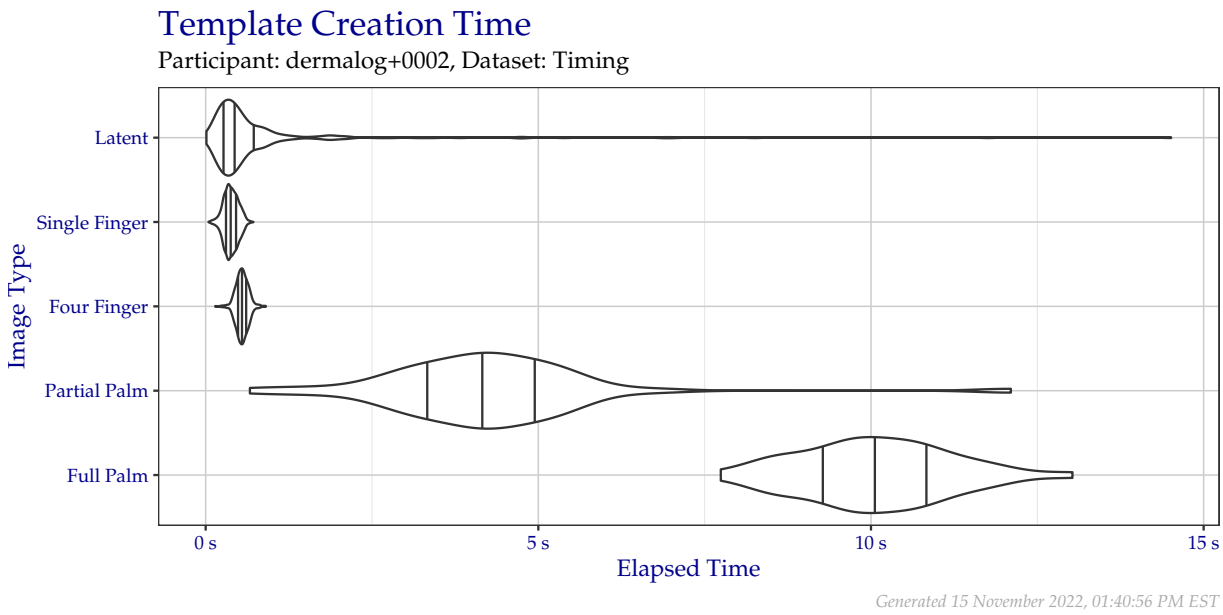

Figure 2: Violin plot of the duration of template creation in seconds for images from the Timing Sample dataset. Vertical lines from left to right indicate the 25%, 50%, and 75% quantiles respectively.

2.3.3 Template Creation Memory Consumption

Figure 3 shows the amount of RAM consumed by the single testing process as a function of time during the template creation procedure, including RAM consumed by the NIST testing apparatus.

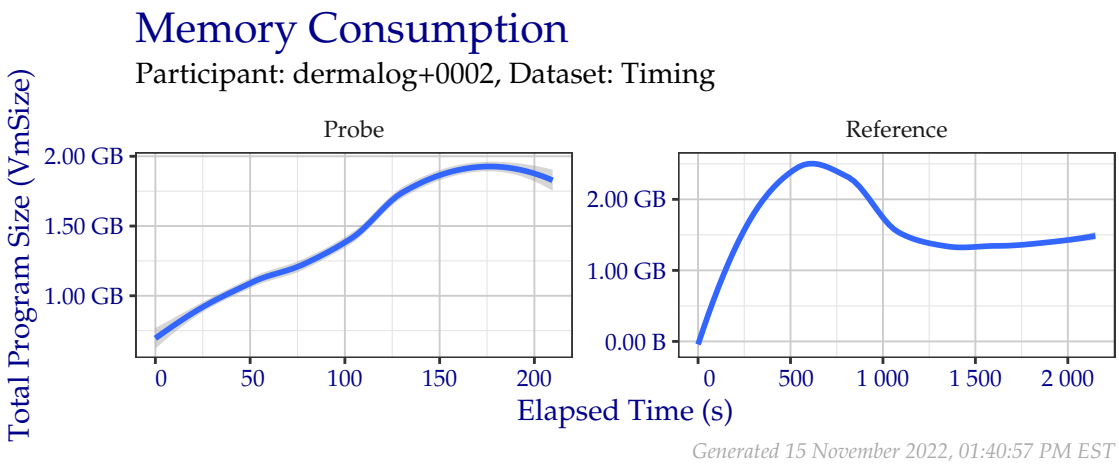

Figure 3: Amount of RAM used while creating templates in the Timing Sample dataset.

2.4 Enrollment Database

Reference templates are combined into a participant-defined database structure for optimal searching. The required storage for the Timing Sample enrollment database with plain impression distractors was **35.9 GB**, and the required storage for the Timing Sample enrollment database with rolled impression distractors was **66.5 GB**. Each database consisted of the same ≈ 1 600 000 non-mated subjects. Each subject had at least one, but typically ten, distal phalanx captures to enroll. ≈ 150 000 had one or more palm captures.

2.5 Search

Out of the latent templates generated in Table 2, a fixed random sample of 100 of the resulting latent templates were searched against the enrollment databases described in Subsection 2.4. The results presented in Subsection 2.5 are based on the measurements made on or during those 100 searches.

2.5.1 Search Duration

Table 5 and Figure 4 show the amount of time elapsed during searches of the fixed search probe set when searching against the enrollment databases described in Subsection 2.4. While unsuccessful searches expend operator time, they are not included in this metric, because search failures typically occur instantaneously (e.g., a template indicates that a probe was of too poor quality to search), which can artificially lower the average search time.

ELFT defines maximum average search durations for participants based on the number of subjects in the enrollment database. Due to the potential for extended runtimes, NIST may choose to allow some discretion in the enforcement of maximum search durations during times of high demand for compute resources. For example, if a maximum average search duration was 4 hours, but after completing all searches, the average search duration was 4.5 hours, it may be prudent to continue the evaluation, since a resubmission may require regeneration of millions of templates and several thousand repeated searches.

Table 5: Search time durations of the search probe set from the Timing Sample dataset, in seconds.

Distractor Imp.	Mated?	Min	25%	Median	Mean	75%	Maximum	Failures	Searches
Plain	False	321	5 547	8 275	10 019	14 393	24 895	1	100
Plain	True	322	5 557	8 285	10 036	14 419	24 973	1	100
Roll	False	572	9 718	15 214	18 080	26 363	46 534	1	100
Roll	True	574	9 742	15 261	18 138	26 466	46 670	1	100

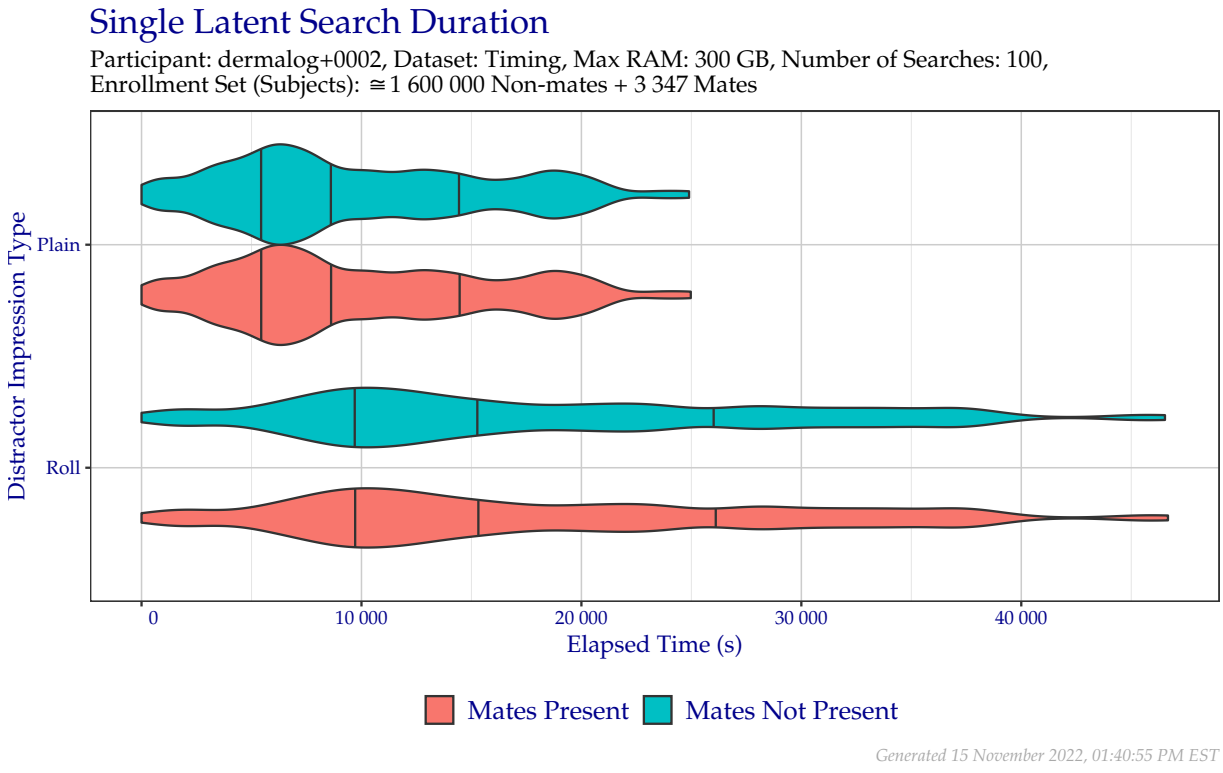

Figure 4: Violin plot of search time durations of the search probe set from the Timing Sample dataset. Vertical lines from left to right indicate the 25%, 50%, and 75% quantiles respectively.

2.5.2 Search Memory Consumption

Figure 5 shows the amount of RAM consumed by the single testing process as a function of time during the search procedure, including RAM consumed by the NIST testing apparatus. Implementations were permitted to use up to 300 GB to store templates. Note the different scales on each panel—implementations that do not change the contents of RAM may not show variation.

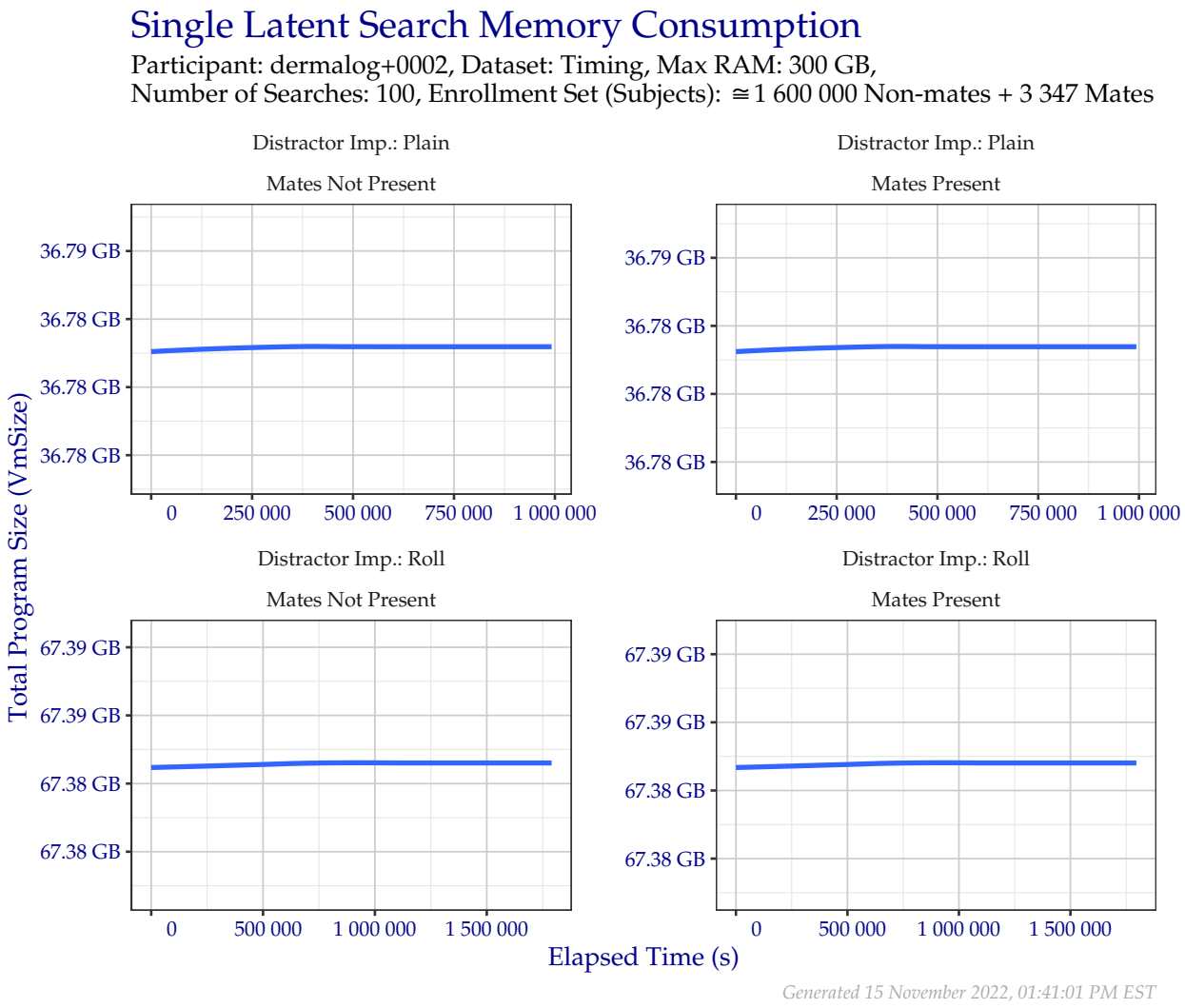

Figure 5: Amount of RAM used while searching templates in the Timing Sample dataset.

3 Metrics

3.1 Location

When a metric depicts search accuracy in this document, it is reported in terms of Location: Region, Hand, and Subject.

- **Region**: The correct region of the correct subject was returned.
  - For search probes sourced from a distal phalanx (i.e., a "latent fingerprint"), the correct finger position 1–10 shall be returned.
  - For search probes sourced from a palm or a non-distal phalanx, the most localized region shall be returned. Some palm regions may be interchangeable based on the exemplars provided (e.g., a palm probe's source could reasonably be seen in a lower palm, hypothenar, and writer's palm exemplar). Credit is given for **Region** in this case.
- **Hand**: A friction ridge position from anywhere on the correct hand of the correct subject is returned. This is designed to aid in diagnosing segmentation error.
- **Subject**: Any finger position from the correct subject is returned. This is designed to reward the situation where an implementation cannot ascertain the most localized region from the set of exemplars enrolled.

3.2 Cumulative Match Characteristic (CMC)

The Cumulative Match Characteristic (CMC) plots in this document show the false negative identification rate (FNIR) without respect for similarity score when searching probes against a enrollment database where a single mated identity for each search probe was present.

- ≈ 1 600 000 non-mated subjects were enrolled.
  - All subjects had at least one, but typically ten, distal phalanx captures to enroll. ≈ 150 000 had one or more palm captures to enroll.
  - Two different combinations of non-mates were searched in separate enrollment databases. While both contain the identical subjects, one set contains only plain impressions and the other contains only rolled impressions.
- The requested size of the candidate list was always 100 subjects.
- All possible Extended Feature Set (EFS) data was provided when "Image + EFS" is listed for probes. The type of EFS data present varies for each sample in each dataset. Initial experiments show nominal (if any) change when EFS data was provided alongside exemplars.
- Probe impression type was always "Unknown Finger" or "Unknown Palm," as appropriate. Future studies may show results using the impression type "Unknown Friction Ridge" for both types of probes.
- The metric *hit rate* is equivalent to 1 − miss rate, or 1 − FNIR. For example, an FNIR of 0.1 indicates a hit rate of 0.9 (i.e., 90%).

3.3 Detection Error Tradeoff (DET)

The Detection Error Tradeoff (DET) plots in this document show the tradeoff between false positive and false negative identification rates when searching probes against a enrollment database where a single mated identity for each search probe was present.

- ≈ 1 600 000 non-mated subjects were enrolled.
  - All subjects had at least one, but typically ten, distal phalanx captures to enroll. ≈ 150 000 had one or more palm captures to enroll.
  - Two different combinations of non-mates were searched in separate enrollment databases. While both contain the identical subjects, one set contains only plain impressions and the other contains only rolled impressions.
  - Non-mated similarity scores come from all ranks when searching probes against an enrollment dataset without any mated subjects enrolled.
- The requested size of the candidate list was always 100 subjects.

  - Mated similarity scores come from the correct location appearing at *any* rank.
- All possible EFS data was provided when "Image + EFS" is listed for probes. The type of EFS data present varies for each sample in each dataset. Initial experiments show nominal (if any) change when EFS data was provided alongside exemplars.
- Probe impression type was always "Unknown Finger" or "Unknown Palm," as appropriate. Future studies may show results using the impression type "Unknown Friction Ridge" for both types of probes.

4 Non-mated Distractor Subjects

When searching probes in each of the subsequent sections, the non-mated distractor subjects that comprised the majority of each enrollment database remained the same. The results of Section 4 are based off of these distractor subjects.

4.1 Failures

Table 6 shows the number of failures to create reference templates for non-mated distractor subjects.

Table 6: Number of failures to create reference templates.

Distal Phalanx Impression Type	Failures	≈ Attempts
Plain	5 665	1 600 000
Roll	5 530	1 600 000

5 FBI Laboratory

The results of Section 5 are based on searches of the sequestered dataset *FBI Laboratory*. This dataset consists of 49 operational latent distal phalanx probes. Members of the FBI manually annotated the probe images and confirmed the ground truth mate. All probes searched were a single sample depicting a region from a distal phalanx. EFS data provided with the probe image *may* include:

- Pattern classification
- Minutia locations (unconfirmed source)

5.1 Failures

Table 7 shows the number of failures to create templates. Table 8 shows the number of failures to produce a candidate list.

Table 7: Number of failures to create templates.

Image Type	Content	Failures	Attempts
Exemplar	Image	0	38
Probe	Image + EFS	0	49
Probe	Image	0	49

Table 8: Number of failures to produce a candidate list. This number includes any failures to create a probe template from Table 7.

Distractor Imp.	Probe Content	Failures	Attempts
Plain	Image	0	49
Roll	Image	0	49
Plain	Image + EFS	0	49
Roll	Image + EFS	0	49

5.2 CMC Plots

The CMC plots in Figure 6 show the FNIR of dermalog+0002 when searching FBI Laboratory against enrollment database where a single mated identity for each search probe was present. The plots are faceted by the distractor impression type and whether probe EFS data was provided. Tabular versions of FNIR at select ranks can be viewed in Table 9.

5.3 CMC Table

The values in Table 9 correspond to Figure 6.

Table 9: Region FNIR values from CMC plotted in Figure 6.

Distractor Imp.	Probe Content	Rank 1	Rank ≤ 50	Rank ≤ 100
Plain	Image	0.5306	0.2857	0.2857
Roll	Image	0.4490	0.3061	0.2653
Plain	Image + EFS	0.7551	0.4898	0.4898
Roll	Image + EFS	0.6531	0.4490	0.4490

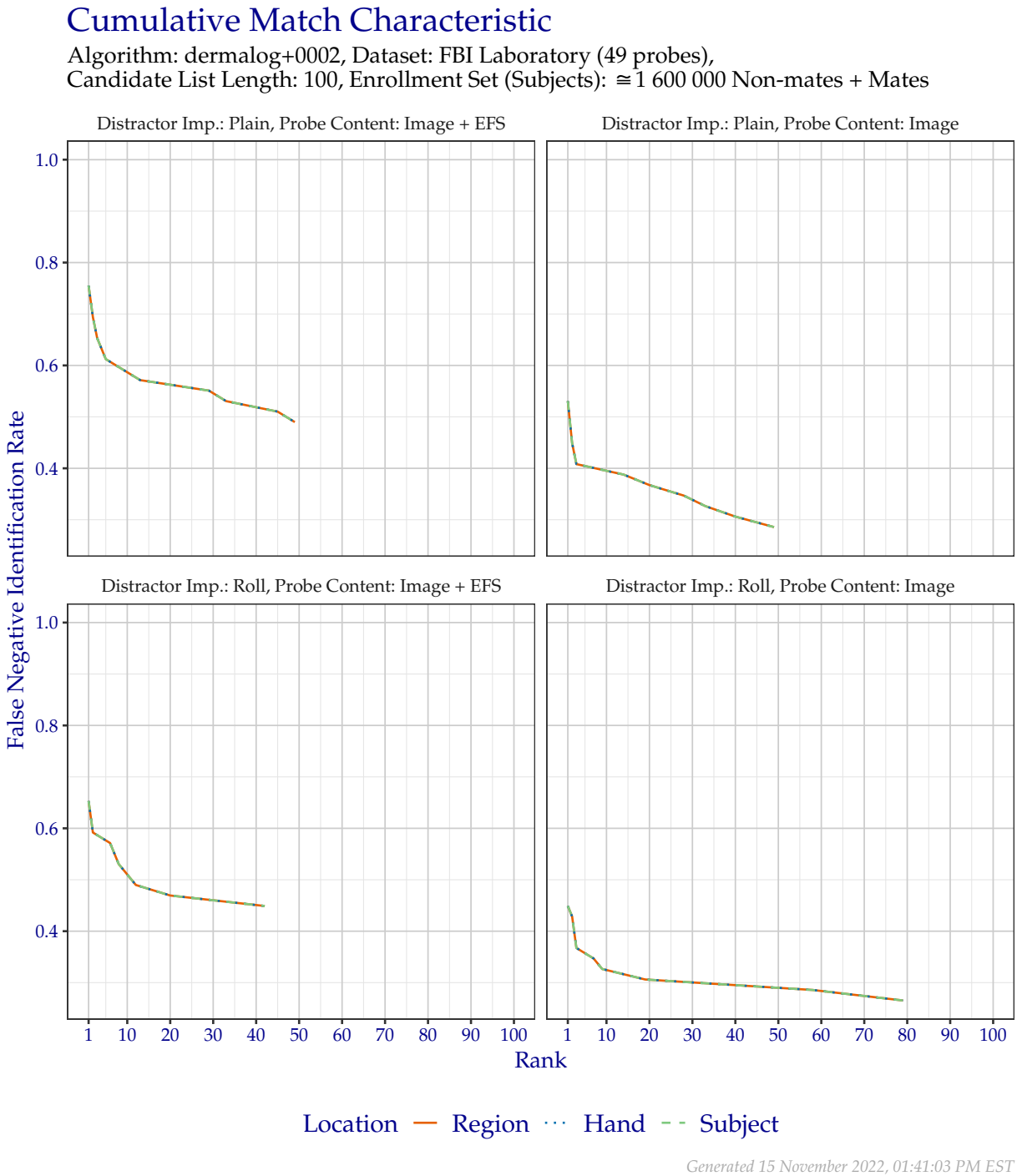

Figure 6: CMC when searching FBI Laboratory probes, faceted by distractor impression type and whether probe EFS data was provided.

5.4 DET Plots

The DET plots in Figure 7 show the false positive and false negative identification rate tradeoffs of dermalog+0002 when searching FBI Laboratory against enrollment database where a single mated identity for each search probe was present. The plots are faceted by the distractor impression type and whether probe EFS data was provided. Tabular versions of FNIR at select FPIR can be viewed in Table 10.

5.5 DET Table

The values in Table 10 correspond to Figure 7.

Table 10: Region FNIR values corresponding to FPIR plotted in Figure 7.

Distractor Imp.	Probe Content	FPIR = 0.01	FPIR = 0.02	FPIR = 0.1
Plain	Image	0.3714	0.3143	0.2286
Roll	Image	0.2778	0.2500	0.0833
Plain	Image + EFS	0.6800	0.6000	0.4800
Roll	Image + EFS	0.7407	0.6296	0.5185

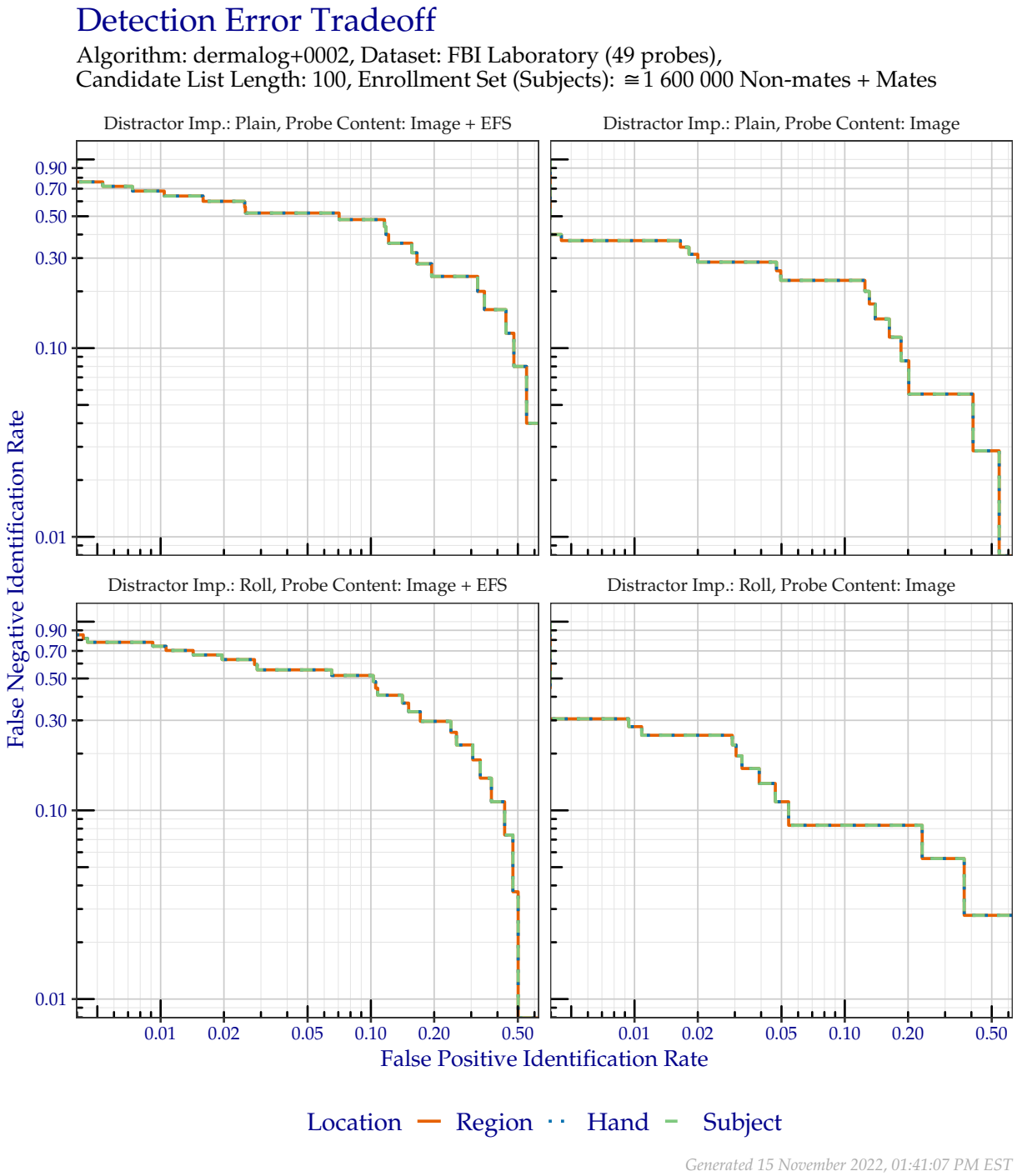

Figure 7: DET when searching FBI Laboratory probes, faceted by the distractor impression type and whether probe EFS data was provided.

6 FBI-Provided Solved Dataset #1

The results of Section 6 are based on searches of the sequestered dataset *FBI-Provided Solved Dataset #1*. This dataset consists of 516 operational probes collected from a particular type of crime. Members of the FBI manually annotated the probe images and confirmed the ground truth mate. All probes searched were a single sample depicting a region from a distal phalanx. EFS data provided with the probe image *may* include:

- Pattern classification
- Core locations (unconfirmed source)
- Delta locations (unconfirmed source)
- Minutia locations (unconfirmed source)

6.1 Failures

Table 11 shows the number of failures to create templates. Table 12 shows the number of failures to produce a candidate list.

Table 11: Number of failures to create templates.

Image Type	Content	Failures	Attempts
Exemplar	Image	0	173
Probe	Image + EFS	0	516
Probe	Image	0	516

Table 12: Number of failures to produce a candidate list. This number includes any failures to create a probe template from Table 11.

Distractor Imp.	Probe Content	Failures	Attempts
Plain	Image	2	516
Roll	Image	2	516
Plain	Image + EFS	1	516
Roll	Image + EFS	1	516

6.2 CMC Plots

The CMC plots in Figure 8 show the FNIR of dermalog+0002 when searching FBI-Provided Solved Dataset 1 against enrollment database where a single mated identity for each search probe was present. The plots are faceted by the distractor impression type, mated impression type, and whether probe EFS data was provided. Tabular versions of FNIR at select ranks can be viewed in Table 13.

6.3 CMC Table

The values in Table 13 correspond to Figure 8.

Table 13: Region FNIR values from CMC plotted in Figure 8.

Distractor Imp.	Probe Content	Rank 1	Rank ≤ 50	Rank ≤ 100
Plain	Image	0.3508	0.2326	0.2112
Roll	Image	0.3275	0.2112	0.1977
Plain	Image + EFS	0.5000	0.3527	0.3159
Roll	Image + EFS	0.4884	0.3178	0.2849

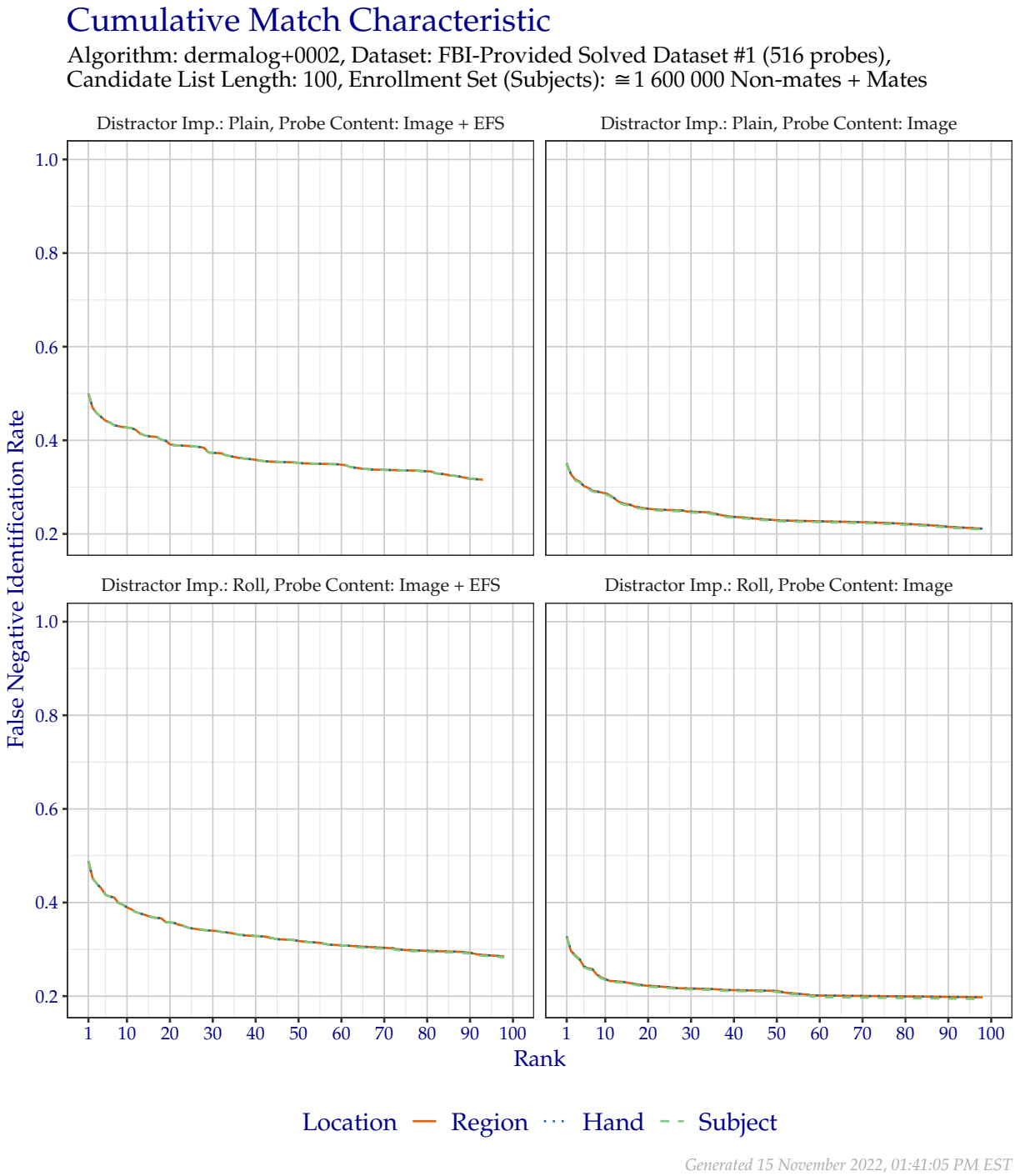

Figure 8: CMC when searching FBI-Provided Solved Dataset 1 probes, faceted by distractor impression type, mated impression type, and whether probe EFS data was provided.

6.4 DET Plots

The DET plots in Figure 9 show the false positive and false negative identification rate tradeoffs of dermalog+0002 when searching FBI-Provided Solved Dataset 1 against enrollment database where a single mated identity for each search probe was present. The plots are faceted by the distractor impression type, mated impression type, and whether probe EFS data was provided. Tabular versions of FNIR at select FPIR can be viewed in Table 14.

6.5 DET Table

The values in Table 14 correspond to Figure 9.

Table 14: Region FNIR values corresponding to FPIR plotted in Figure 9.

Distractor Imp.	Probe Content	FPIR = 0.01	FPIR = 0.02	FPIR = 0.1
Plain	Image	0.2187	0.2015	0.1622
Roll	Image	0.2150	0.1957	0.1570
Plain	Image + EFS	0.4023	0.3626	0.2606
Roll	Image + EFS	0.3930	0.3496	0.2846

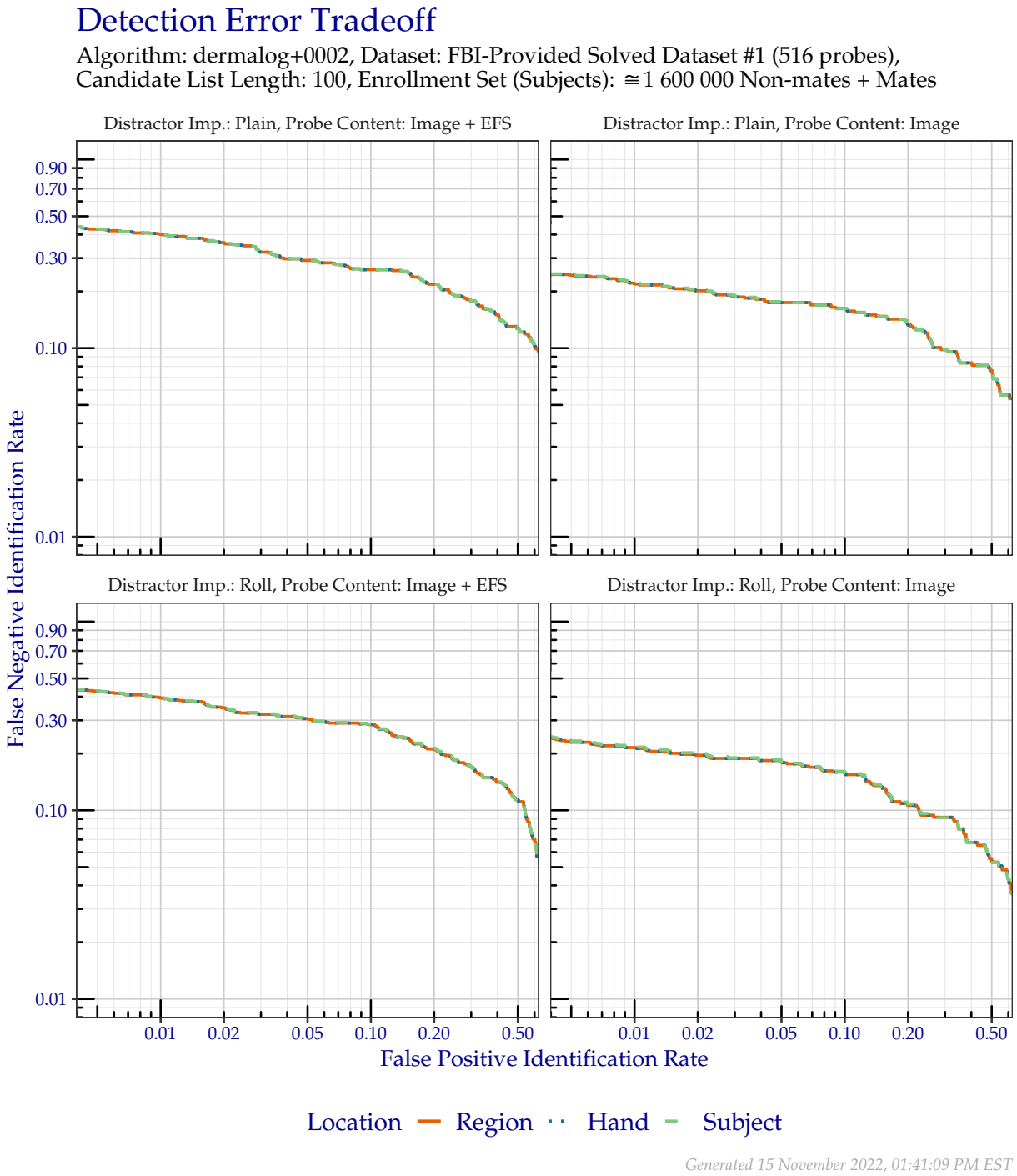

Figure 9: DET when searching FBI-Provided Solved Dataset 1 probes, faceted by the distractor impression type, mated impression type, and whether probe EFS data was provided

7 Michigan State Police

The results of Section 7 are based on searches of the sequestered dataset *Michigan State Police*. This dataset consist of of 2 174 operational latent probes. No EFS data was provided for probes or mated exemplars.

All probes searched were a single friction ridge sample from somewhere on the hand. Because the ELFT API indicates to implementations whether an image comes from the distal or palm region, analysis is separated between the two.

Note: While NIST biometric technology evaluations typically use sequestered law enforcement data, a literature search indicates that this collection of data may have been supplied to other research organizations that are not subject to the same strict sequestration policies as NIST.

7.1 Failures

Table 15 shows the number of failures to create templates. Table 16 shows the number of failures to produce a candidate list.

Table 15: Number of failures to create templates.

Image Type	Content	Distal Failures	Palm Failures	Attempts
Exemplar	Image	0	88	1 365
Probe	Image	21	161	2 174

Table 16: Number of failures to produce a candidate list. This number includes any failures to create a probe template from Table 15.

Distractor Imp.	Probe Content	Distal Failures	Palm Failures	Attempts
Plain	Image	29	161	2 174
Roll	Image	29	161	2 174

7.2 Distal Region CMC

The CMC in Figure 10 shows results from *only* the distal phalanx probes from Michigan State Police.

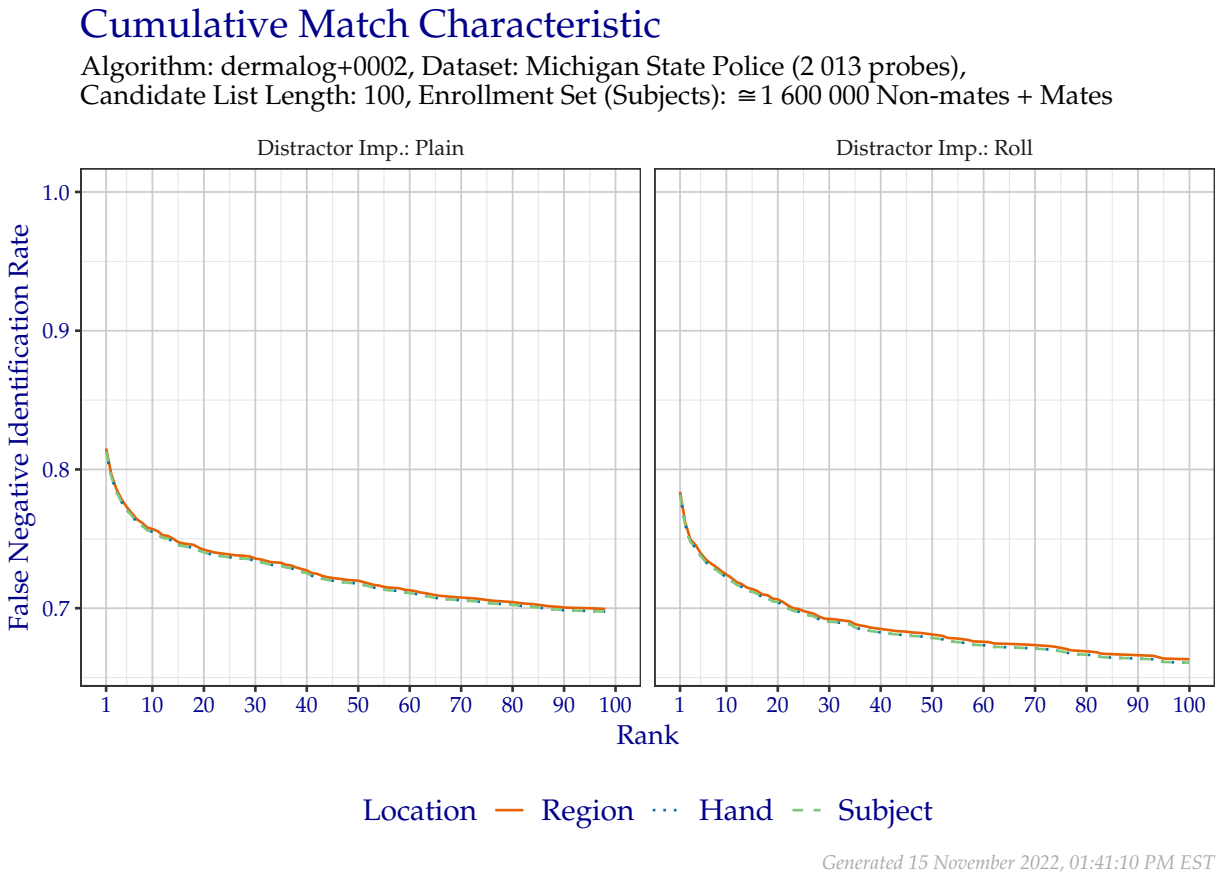

Figure 10: CMC when searching Michigan State Police distal phalanx probes, faceted by distractor impression type.

The values in Table 17 correspond to Figure 10.

Table 17: Region FNIR values from CMC plotted in Figure 10.

Distractor Imp.	Probe Content	Rank 1	Rank ≤ 50	Rank ≤ 100
Plain	Image	0.8152	0.7198	0.6995
Roll	Image	0.7839	0.6811	0.6632

7.3 Palm Region CMC

The CMC in Figure 11 shows results from *only* the palm probes from Michigan State Police.

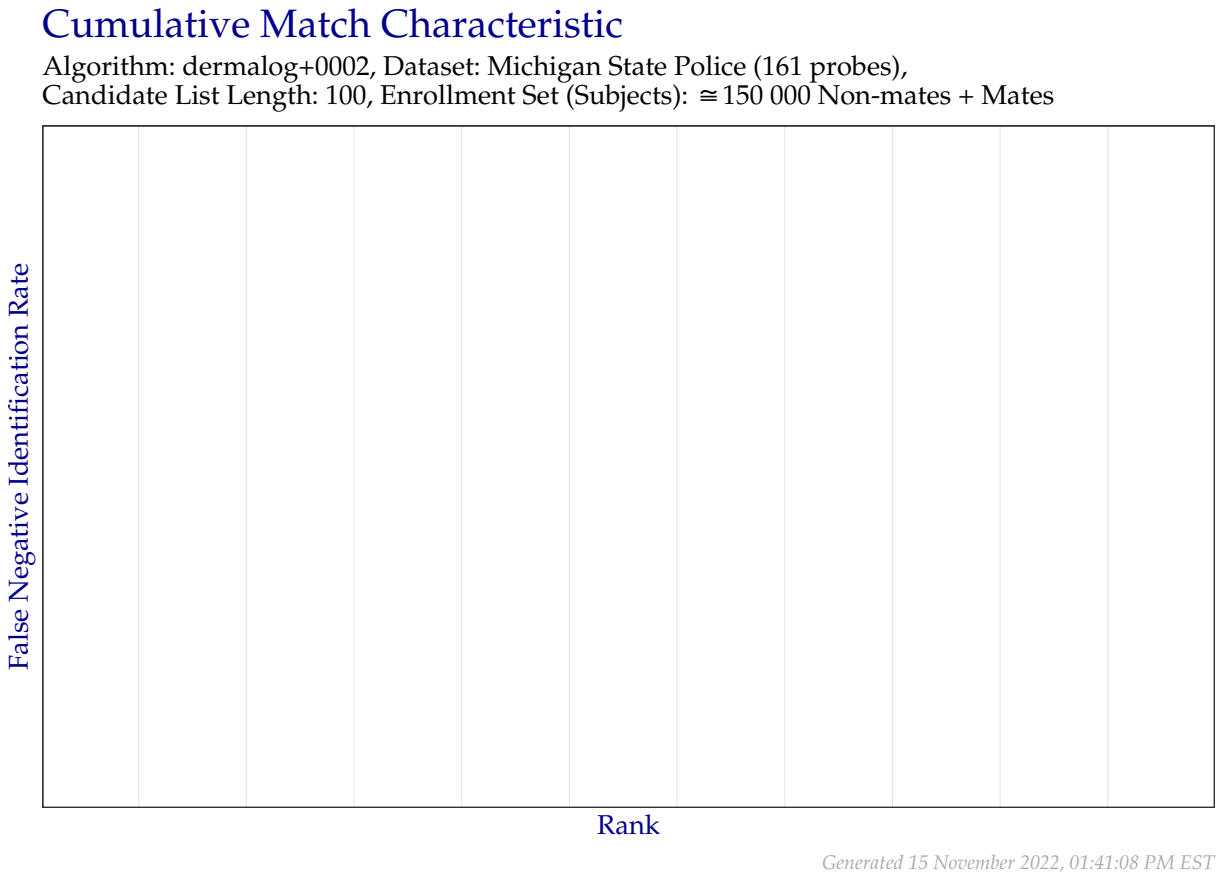

Figure 11: CMC when searching Michigan State Police palm probes, faceted by distractor impression type.

The values in Table 18 correspond to Figure 11.

Table 18: Region FNIR values from CMC plotted in Figure 11.

Distractor Imp.	Probe Content	Mated Content	Rank	FNIR
Plain	Image	Image	NA	NA
Roll	Image	Image	NA	NA

7.4 Distal Region DET

The DET in Figure 12 shows results from *only* the distal phalanx probes from Michigan State Police.

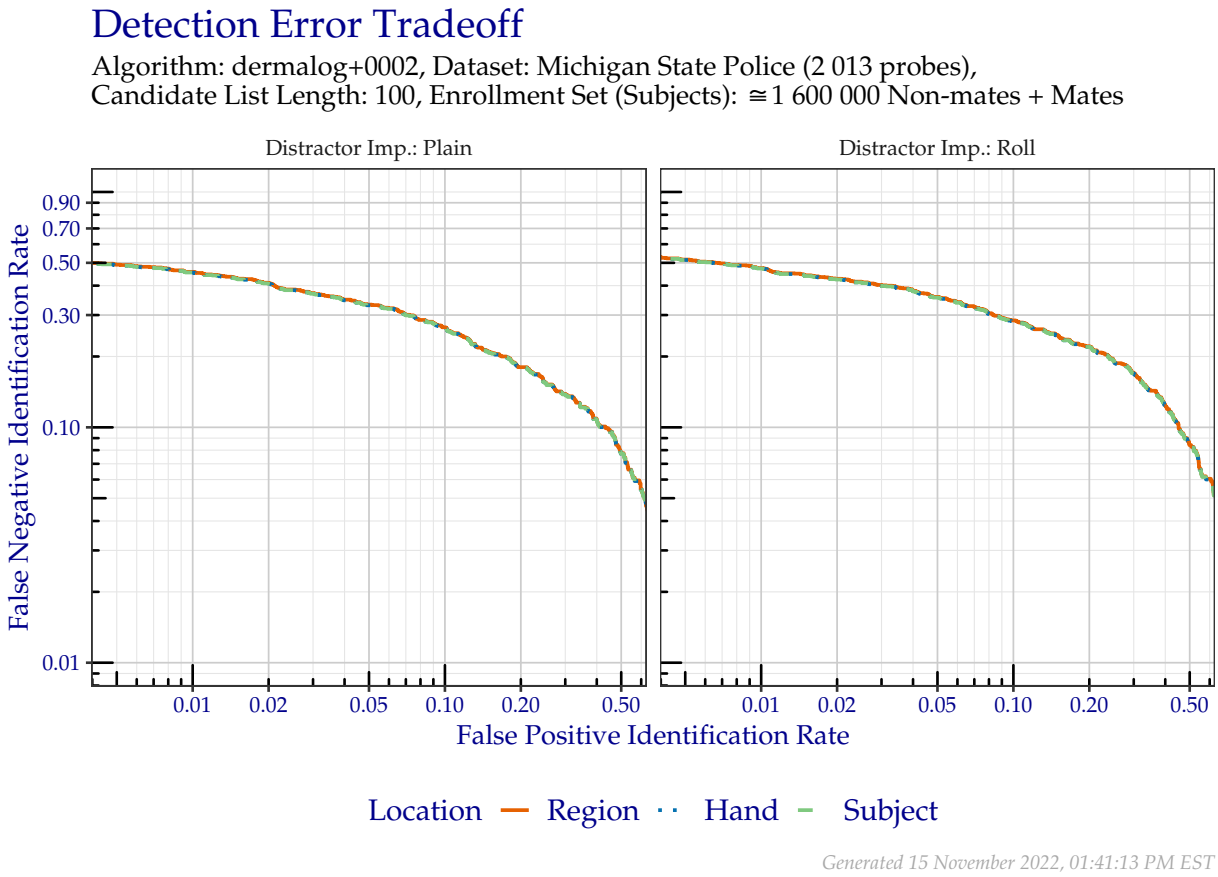

Figure 12: DET when searching Michigan State Police distal phalanx probes, faceted by distractor impression type.

The values in Table 19 correspond to Figure 12.

Table 19: Region FNIR values corresponding to FPIR plotted in Figure 12.

Distractor Imp.	Probe Content	FPIR = 0.01	FPIR = 0.02	FPIR = 0.1
Plain	Image	0.4579	0.4099	0.2661
Roll	Image	0.4749	0.4277	0.2847

7.5 Palm Region DET

The DET in Figure 13 shows results from *only* the palm probes from Michigan State Police.

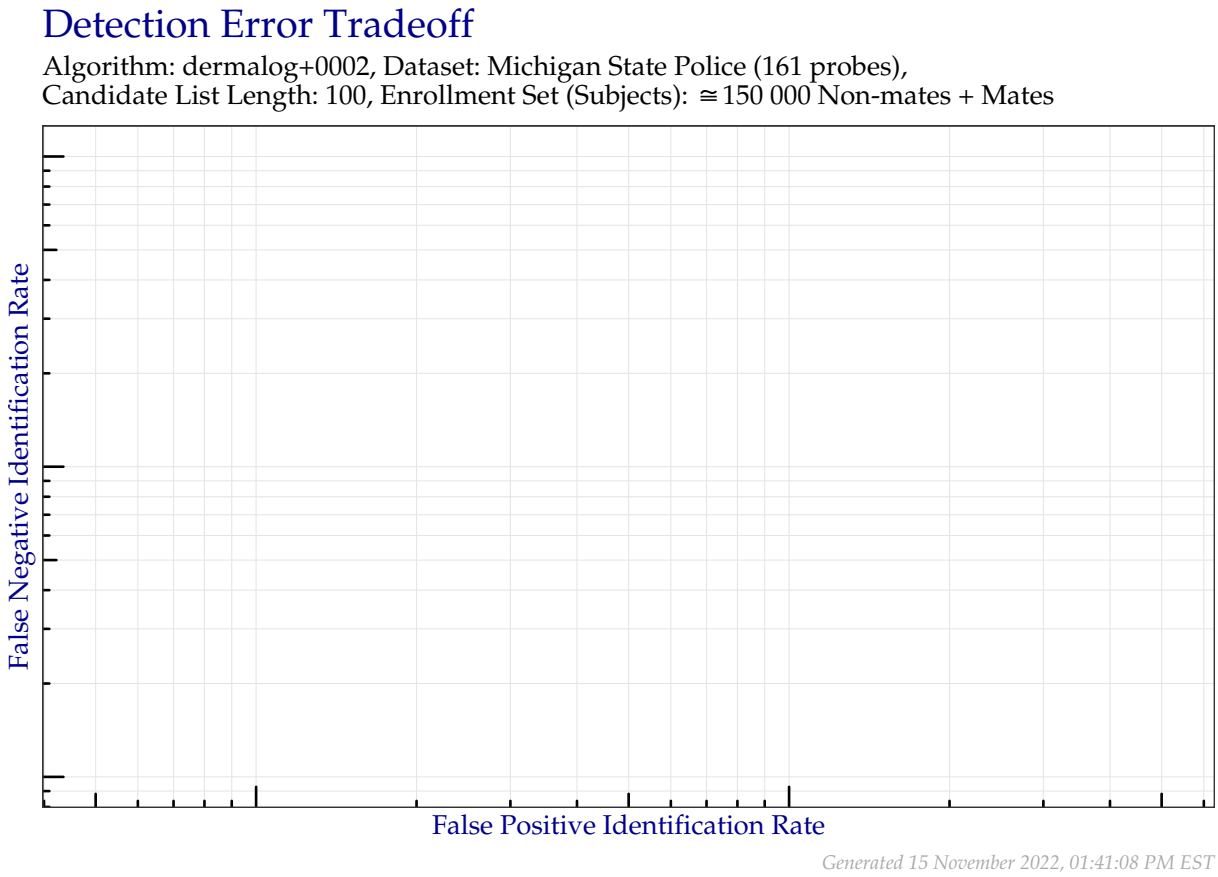

Figure 13: DET when searching Michigan State Police palm probes, faceted by distractor impression type.

The values in Table 20 correspond to Figure 13.

Table 20: Region FNIR values corresponding to FPIR plotted in Figure 13.

Distractor Imp.	Probe Content	FPIR = 0.1	FPIR = 0.01	FPIR = 0.02
Plain	Image	NA	NA	NA
Roll	Image	NA	NA	NA