NISTIR 8271

Face Recognition Vendor Test (FRVT)
Part 2: Identification

Patrick Grother
Mei Ngan
Kayee Hanaoka

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8271

National Institute of
Standards and Technology
U.S. Department of Commerce

NISTIR 8271

Face Recognition Vendor Test (FRVT)
Part 2: Identification

Patrick Grother
Mei Ngan
Kayee Hanaoka
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8271

September 2019

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology

National Institute of Standards and Technology Interagency or Internal Report 8271
Natl. Inst. Stand. Technol. Interag. Intern. Rep. 8271, 186 pages (September 2019)

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8271

ACKNOWLEDGMENTS

The authors are grateful to Wayne Salamon and Greg Fiumara at NIST for designing robust software infrastructure for image and template storage and parallel execution of algorithms across our computers. Thanks also to Brian Cochran at NIST for providing highly available computers and network-attached storage.

DISCLAIMER

Specific hardware and software products identified in this report were used in order to perform the evaluations described in this document. In no case does identification of any commercial product, trade name, or vendor, imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.

Executive Summary

This report updates and extends NIST Interagency Report 8238, documenting the evaluation of automated face recognition algorithms submitted to NIST in November 2018. The algorithms, which implement one-to-many identification of faces appearing in two-dimensional images, are prototypes from the research and development laboratories of mostly commercial suppliers, and are submitted to NIST as compiled black-box libraries implementing a NIST-specified C++ test interface. The report therefore does not describe how algorithms operate.

The evaluation used four datasets - frontal mugshots, profile views, webcam photos and wild images - and the report lists accuracy results alongside developer names. It will therefore be useful for comparison of face recognition algorithms and assessment of absolute capability. The primary dataset is comprised of 26.6 million reasonably well-controlled live portrait photos of 12.3 million individuals. The three smaller datasets contain more unconstrained photos: 3.2 million webcam images; 200 thousand side-view images; and 2.5 million photojournalism and amateur photographer photos. These datasets are sequestered at NIST, meaning that developers do not have access to them for training or testing. The last dataset, however, consists of images drawn from the internet for testing purposes so while it is not truly sequestered, its composition is unknown to the developers.

The evaluation was run in three phases, starting Feburary, June, and November 2018 respectively, with developers receiving technical feedback between phases. Results for 127 algorithms from 41 developers were published in November 2018 as NIST Interagency Report 8238. This update adds results for an additional 76 algorithms from 42 developers submitted in October 2018. At that time seven developers ceased participation, and nine developers started. The developer totals constitute a substantial majority of the face recognition industry.

The major result given in NIST IR 8238 was that massive gains in accuracy have been achieved in the last five years (2013-2018) and these far exceed improvements made in the prior period (2010-2013). While the industry gains were broad - at least 30 developers' algorithms outperformed the most accurate algorithm from late 2013 - there remains a wide range of capability. While this report shows accuracy gains only over the course of 2018, the most accurate algorithm reported here is substantially more accurate than anything reported in NIST IR 8238. This is evidence that face recognition development continues apace, and that FRVT reports are but a snapshot of contemporary capability.

From discussion with developers, the accuracy gains stem from the adoption of deep convolutional neural networks. As such, face recognition has undergone an industrial revolution, with algorithms increasingly tolerant of poorly illuminated and other low quality images, and poorly posed subjects. One related result is that a few algorithms correctly match side-view photographs to galleries of frontal photos, with search accuracy approaching that of the best c. 2010 algorithms executing frontal-frontal search. The capability to recognize under a 90-degree change in viewpoint - pose invariance - has been a long-sought milestone in face recognition research.

With good quality portrait photos, the most accurate algorithms will find matching entries, when present, in galleries containing 12 million individuals, with rank one miss rates of approaching 0.1%. The remaining errors are in large part attributable to long-run ageing, facial injury and poor image quality. In at least 5% of images identification often succeeds (i.e. the mate is returned at rank 1) but recognition similarity scores are weak such that true and false matches become indistinguishable, and human adjudication becomes necessary.

From Fall 2019 this report will be updated continuously as new algorithms are submitted to FRVT, and run on new datasets. Participation in the one-to-many identification track requires a devloper to first demonstrate high accuracy in the one-to-one verification track of FRVT.

Scope and Context

Audience: This report is intended for developers, integrators, end users, policy makers and others who have some familiarity with biometrics applications. The methods and metrics documented here will be of interest to organizations engaged in tests of face recognition algorithms. Some of these have been incorporated in the ISO/IEC 19795 Part 1 Biometric Testing and Reporting Framework standard, now under revision.

Prior benchmarks: Automated face recognition accuracy has improved massively in the two decades since initial commercialization of the various technologies. NIST has tracked that improvement through its conduct of regular independent, free, open, and public evaluations. These have fostered improvements in the state of the art. This report serves as an update to the NIST Interagency Report 8238 on performance of face identification algorithms, published in November 2018.

Scope: As with NIST IR 8238, this report documents recognition results for four databases containing in excess of 30.2 million still photographs of 14.4 million individuals. This constitutes the largest public and independent evaluation of face recognition ever conducted. It includes results for accuracy, speed, investigative vs. identification applications, scalability to large populations, use of multiple images per person, images of cooperative and non-cooperative subjects.

The report also includes results for ageing, recognition of twins, and recognition of profile-view images against frontal galleries. It otherwise does not address causes of recognition failure, neither image-specific problems nor subject-specific factors including demographics. Separate reports on demographic dependencies in face recognition will be published in the future. Additionally out of scope are: performance of live human-in-the-loop transactional systems like automated border control gates; human recognition accuracy as used in forensic applications; and recognition of persons in video sequences (which NIST evaluated separately [9]). Some of those applications share core matching technologies that *are* tested in this report.

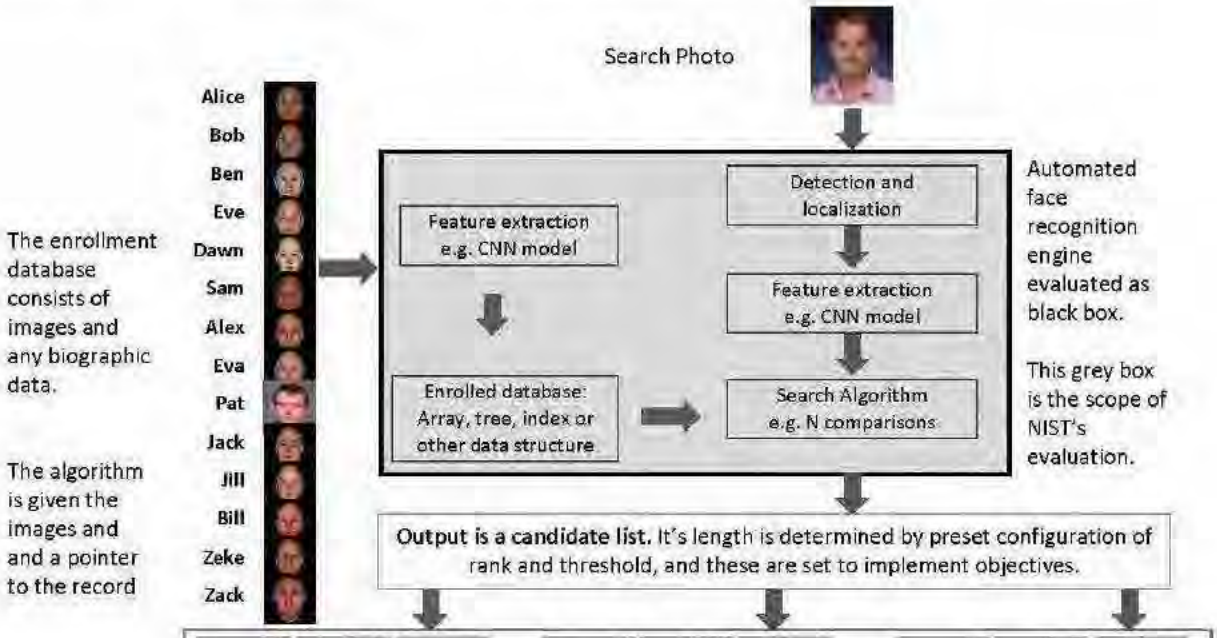
Images: Three kinds of images are employed. The primary dataset is a set of law enforcement mugshot images (Fig. 3) which are enrolled and then searched with three kinds of images: 1) other mugshots (i.e. within-domain); 2) profile-view photographs (90 degree cross-view); 3) lower quality webcam images (Fig. 4) collected in similar detention operations (cross-domain); Additionally wild images (Fig. 6) are searched against other wild images.

Participation and industry coverage: The report includes performance figures for 203 prototype algorithms from the research laboratories of 51 commercial developers and one university. This represents a substantial majority of the face recognition industry, but only a tiny minority of the academic community. Participation was open worldwide. While there is no charge for participation, developers incur some software engineering expense in implementing their algorithms behind the NIST application programming interface (API). The test is a black-box test where the function of the algorithm, and the intellectual property associated with it, is hidden inside pre-compiled libraries.

Recent technology development: Most face recognition research with deep convolutional neural networks (CNNs) has been aimed at achieving invariance to pose, illumination and expression variations that characterize photojournalism and social media images. The initial research [18, 24] employed large numbers of images of relatively few ($\sim 10^4$) individuals to learn invariance. Inevitably much larger populations ($\sim 10^7$) were employed for training [11, 20] but the benchmark, Labeled Faces in the Wild with (essentially) an equal error rate metric [12], represents an easy task, one-to-one verification at very high false match rates. While a larger scale identification benchmark duly followed, Megaface [15], its primary metric, rank one hit rate, contrasts with the high threshold discrimination task required in most large-population applications of face recognition, namely credential de-duplication, and background checks. There, identification in galleries containing up to 10^8 individuals must be performed using a) very few images per individual and b) stringent thresholds to afford very low false positive identification rates. FRVT 2018 was launched to measure the capability of the new technologies, including in these two cases. FRVT has included open-set identification tests since 2002, reporting both false negative and positive identification rates [7].

Performance metrics for applications: This report documents the performance of one-to-many face recognition algorithms. The word "performance" here refers to recognition accuracy and computational resource usage, as measured by executing those algorithms on massive sequestered datasets.

This report includes extensive tabulation of recognition error rates germane to the main use-cases for face search technology. The Figure below, inspired by the Figure 1 in [25] differentiates different applications of the technolgy. The last row directs readers to the main tables relevant to those applications, respectively threshold-based and rank-based metrics that are special cases of the metrics given in section 3. The terms negative identification and positive identification are taken from the ISO/IEC 2382-37:2017 standardized biometrics vocabulary.



	Positive identification	Negative identification	Post-event investigation
Example	Access to a gym or cruise ship	Watchlist e.g. Detection of deportee or duplicate drivers license applications	Crime scene photos, or of detainee without ID documents.
Claim of identity	Implicit claim to be enrolled	Implicit claim to not be enrolled	No claim: Inquiry
Threshold	High, to implement security objective	High, to limit false positives	Zero
Num. candidates	1	0	L, set by request to algorithm
Human role	Review candidate to assist user in resolution of false negatives, or to detect impostor	Review candidate to determine false positive or correct hit	Review multiple candidates, refer possible hits to examiners see [26]
Intended human involvement frequency	Rare – approx. the false negative identification rate plus prior probability of impostor	Rare – approx. the false positive identification rate plus prior probability of an actual mate	Always
Performance metric of interest	**FNIR at low FPIR. See sec. 3.1, 3.2 and Tables 10, 19**	**FNIR at low FPIR. See sec. 3.1, 3.2 and Tables 10, 19**	**FNIR at ranks 1... 50, say. FPIR = 1. See sec. 3.2 and Table 12, 14, 16**

The algorithms are specifically configured for these applications by setting thresholds and candidate list lengths. Both

rank-based metrics and threshold-based metrics include tradeoffs. In investigation, overall accuracy will be reduced if labor is only available to review a few candidates from the automated system. Note that when a fixed number of candidates are returned, the false positive identification rate of the automated face recognition engine will be 100%, because a probe image of anyone not enrolled will still return candidates. In identification applications where false positives must be limited to satisfy reviewer labor availability or a security objective, higher false negative rates are implied. This report includes extensive quantification of this threshold-based tradeoff. See Sec. 3

Template diversity: The FRVT is designed to evaluate black-box technologies with the consequence that the templates that hold features extracted from face images are entirely proprietary opaque binary data that embed considerable intellectual property of the developer. Despite migration to CNN-based technologies there is no consensus on the optimal feature vector dimension. This is evidenced by template sizes ranging from below 100 bytes to more than four kilobytes. This diversity of approaches, suggests there is no prospect of a standard template something that would require a common feature set to be extracted from faces. Interoperability in automated face recognition remains solidly based on images and documentary standards for those, in particular the ICAO portrait [29] specification deriving from the ISO/IEC 19794-5 Token frontal [26] standard, which are similar to certain ANSI/NIST Type 10 [28] formats.

Training: The algorithms submitted to NIST have been developed using image datasets that developers do not disclose. The development will often include application of machine learning techniques and will additionally involve iterative training and testing cycles. NIST itself does not perform any training and does not refine or alter the algorithm in any way. Thus the model, data files, and libraries that define an algorithm are fixed for the duration of the tests. This reflects typical operational reality where recognition software, once installed, is fixed and constant until upgraded. This situation persists because on-site training of algorithms on customer data is atypical essentially because training is not a turnkey process.

Automated search and human review: Virtually all applications using automated face search require human review of the outputs at some frequency: Always for investigational applications; rarely in positive identification applications, after rejection (false or otherwise); and rarely in negative identification applications, after an alarm (false or otherwise). The human role is usually to compare a reference image with the query image or the live-subject if present, to render either a definitive decision on "exclusion" (different subjects), or "identification" (same subject), or a declaration that one or both images have "no value" and that no decision can be made. Note that automated face recognition algorithms are not built to do exclusion - low scores from a face comparison arise from different faces *and* poor quality images of the same face.

Human reviewers make recognition errors [5, 19, 27] and are sensitive to image acquisition and quality. Accurate human review is supported by high resolution - as specified in the Type 50, 51 acquisition profiles of the ANSI/NIST Type 10 record [28], and by multiple non-frontal views as specified in the same standard. These often afford views of the ear. Organizations involved in image collection should consider supporting human adjudication by collecting high-resolution frontal and non-frontal views, preparing low resolution versions for automated face recognition [26], and retaining both for any subsequent resolution of candidate matches. Along these lines, the ISO/IEC Joint Technical Committee 1 subcommittee 37 on biometrics has just initiated projects on image quality assessment and face-aware capture.

Next steps: NIST expects to publish a first report on demographic dependencies in face recognition in 2019. This will include the effects of age, sex and race.

Technical Summary

▷ **Rank-based accuracy:** The inset table shows false negative "miss rates" realized when searching a 12 million person gallery populated with FRVT 2018 mugshots. The two most accurate algorithms fail to return the correct mate somewhere within the top 50 ranks in fewer than 0.1% of searches (Table 1, rows 1,2). This is achieved for galleries populated with multiple images per person. In the case where only the most recent image is present the miss rate is modestly higher (rows 3,4). The mates are almost always at rank 1, so in cases where only very short candidate lists must be used, the rank-1 miss rate is barely higher 0.12% (row 5) which again modestly rises when persons are enrolled with a single image (row 7). All the miss rates are measured over a fixed set of 154 549 searches, and the lowest false negative error rate recorded in this report (0.038%, row 10) corresponds to just 58 misses. Given such low error rates, what misses remain? By inspection they arise in five categories, those due to: a) ageing i.e. long-term time lapse between images; b) images of injured individuals e.g. bruised or bandaged faces; c) the presence of a second face e.g. printed on a T-shirt; d) images of some object that is not a face; e) profile-view images, and f) actual clerical ID label errors. As discussed in section 3.8.2, the first three categories are legitimately part of a test designed to measure accuracy on portrait images collected in law-enforcement settings. The latter three categories, however, should not be included in a test that is attempting to measure accuracy on only frontal images. Thus, by removing all known images in those categories, the rightmost column shows error rates that would be attainable in an application where exclusively frontal portrait images were collected without identity labeling errors.

	Investigation miss rate at	Num-subjects	Enrolled image	Num-images	Algorithm	FNIR Raw	FNIR Corrected
1	Rank-50	12M	Lifetime	26.1M	NEC-2	0.09%	0.09%
2	Rank-50	12M	Lifetime	26.1M	Microsoft-5	0.06%	**0.06%**
3	Rank-50	12M	Recent	12M	NEC-2	0.25%	**0.08%**
4	Rank-50	12M	Recent	12M	Microsoft-5	0.21%	0.09%
5	Rank-1	12M	Lifetime	26.1M	NEC-2	0.14%	**0.12%**
6	Rank-1	12M	Lifetime	26.1M	Microsoft-5	0.25%	0.24%
7	Rank-1	12M	Recent	12M	NEC-2	0.31%	**0.13%**
8	Rank-1	12M	Recent	12M	Microsoft-5	0.52%	0.37%
9	Rank-50	640K	Lifetime	1.25M	NEC-2	0.08%	0.08%
10	Rank-50	640K	Lifetime	1.25M	Microsoft-5	0.04%	**0.04%**

Table 1: Rank-based accuracy floor 2018.

Error rates today are two orders of magnitude below what they were in 2010, a massive reduction that stems from wholesale replacement of the old algorithms with those based on (deep) convolutional neural networks (CNNs). This constitutes a revolution rather than the evolution that defined the period 2010-2013. The rapid innovations around CNN architectures and loss functions including, both proprietary and published in the academic literature[1], may yet produce further gains. Even without that possibility, the results imply that prospective end-users should establish whether installed algorithms pre-date the development of the prototypes evaluated here and inquire with suppliers on availability of the latest versions. The gains mean that searches that had previously failed to yield candidates may now do so, such that unsolved cases could be revisited.

Given this impressive achievement - close to perfect recognition - an advocate might claim that frontal face recognition is a solved problem, a statement that should be refuted with the following context and caveats:

- ▷ **Algorithm accuracy spectrum:** Many algorithms do not achieve the low error rates tabulated above, and while many of those may still be useful and valuable to end-users, only the most accurate excel on poor quality images and those collected long after the initial enrollment sample.
- ▷ **Versioning:** While results for up to seven algorithms from each developer are reported here, the intra-provider accuracy variations are usually smaller than the inter-provider variations. That said different versions give order of magnitude fewer misses. Some developers demonstrate speed-accuracy tradeoffs[2]. See Figs. 17, 18.

[1] For example, Resnets [11], Inception [23], very deep networks [18, 21] and spatial transformers.
[2] NEC-0 prepares templates much faster than NEC-2 but gives twenty times more misses. Dermalog-5 executes a template search much more quickly than Dermalog-6 but is also much less accurate.

▷ **Quality:** The low error rates here are attained using mostly excellent cooperative live-capture mugshot images collected with an attendant present. Recognition in other circumstances, particularly those without a dedicated photographic environment and human or automated quality control checks, will lead to declines in accuracy. This is documented here for poorer quality webcam images and unconstrained "wild" images.

▷ **Low similarity scores:** In thousands of cases the correct gallery image is returned at rank 1 but its similarity score is nevertheless low, below some operationally required score threshold. This does not matter when face recognition is used for "lead generation" in investigational applications because human reviewers are specifically required to review potentially long candidate lists and the threshold is effectively 0. In applications where search volumes are higher and labor is not available to review the results from searches, a higher threshold can be applied. This reduces the length of candidate lists and false positive identification rates at the expense of increased false negative miss rates. The tradeoff between the two error rates is reported extensively later.

▷ **Population size:** As the number of enrolled subjects grows, some mates are displaced from rank one, decreasing accuracy. As tabulated later for N up to 12 million, false negative rates generally rise slowly with population size.

▷ **Database integrity:** An operational error rate should be added to all false negative rates in this report reflecting the proportion of images in a real database that are un-matchable. Such anomalies arise from images that: do not contain a face; include multiple persons; cannot be decoded; are rotated by 90° or 180°; depict a face on clothing; and others introduced by a long tail of various clerical errors. While the mugshot trials in this report have been constructed to minimize such effects, they are a real problem in actual operations.

▷ **Threshold-based accuracy:** Recognition accuracy is very strongly dependent on the algorithm and, more generally, on the developer of the algorithm. False negative error rates in a particular scenario range from a few tenths of one percent to beyond fifty percent. This is tabulated exhaustively later: For example Table 22 shows accuracy across datasets. The inset figure here compares algorithms on mugshot searches in a consolidated gallery of 12 million subjects and 26.1 million photos. In positive or negative identification applications, a score threshold is set to limit the rate at which non-mate searches produce false positives. This has the consequence that some mated searches will report the mate below threshold, i.e. a miss, even if it is at rank 1. The utility of this is that many non-mated searches will usually not return any candidate identities at all. As the

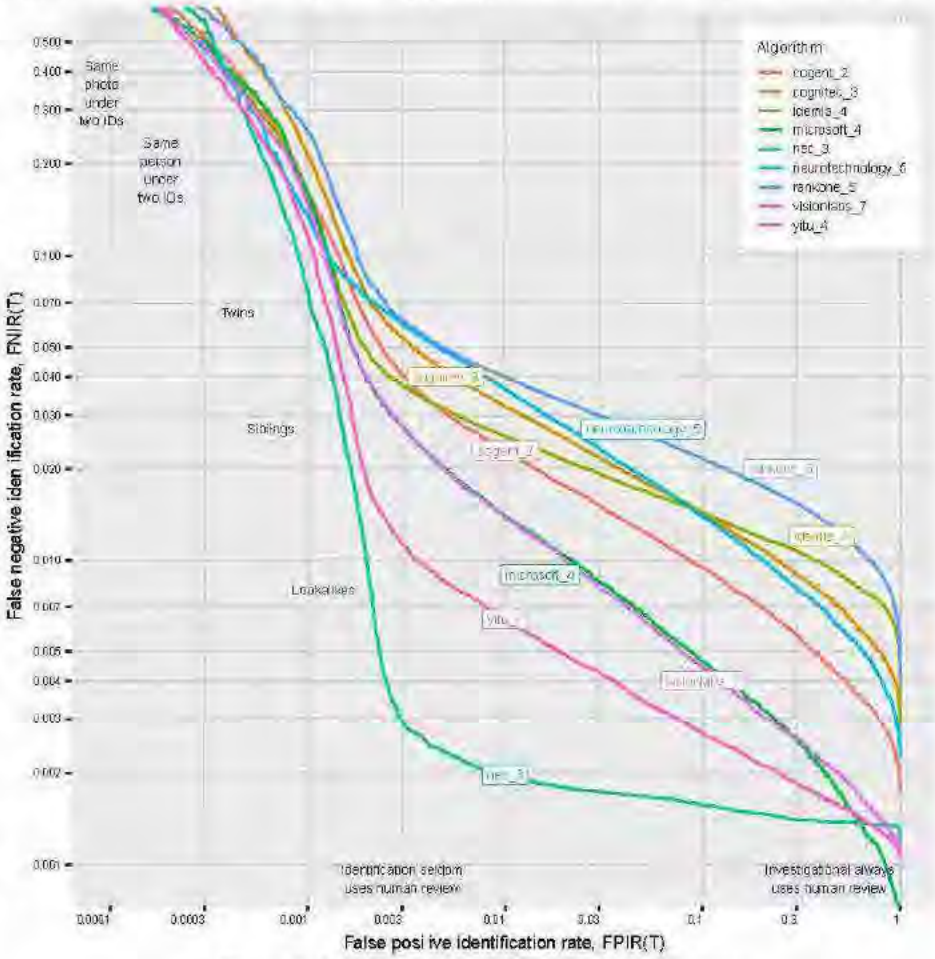


Figure 1: Miss rates across the false positive range

inset error-tradeoff characteristic
shows, investigational miss rates on the right side are very low but then rise steadily (in the center region) as threshold is increased to support "lights-out" applications, and ultimately rise quickly (left side) as discussed below. Thus, if we demand that just one in one thousand non-mate searches produce any false positives, the most accurate algorithm there (NEC-3) would fail on 7.9% of mated searches. Even though the graph shows results for the most accurate algorithms, all but two would fail to find the mate in more than 10% of mated searches. While the NEC algorithm produces a relatively flat error tradeoff until the threshold is raised to limit false positives to about 1 in 400 non-mated searches[3]

Thereafter, as the threshold is raised to further reduce false positives, miss rates rise rapidly. This means that low false positive identification rates are inaccesible with these algorithms, a result that does not apply for ten-finger identification algorithms. The rapid rise occurs because the lower mate scores are mixed with very high non-mate scores, the low scores from poor image quality and ageing, the high non-mates from the presence of lookalikes persons (doppelgangers), twins (discussed next) and, ultimately, the presence of a few unconsolidated subjects i.e. persons present under multiple IDs.

▷ **False positives from twins:** By enrolling 640 000 mugshots, adding photos of one twin, and then searching photos of those subjects and their twin the inset figure shows, for one typical algorithm, the similarity is generally greater when searching twins against themselves (A) than when searching twins against their sibling (B) but very often still above even stringent thresholds i.e. those corresponding to one in one thousand searches producing a false positive. Thus twins will very often produce a high-scoring non-match on a candidate list and a false alarm in an online identification system. The plot shows that some fraternal twins are correctly rejected at those thresholds - these are largely from different sex twins (at center). Figure 21 shows substantially similar behavior for all algorithms tested. In an investigative search, a twin would typically appear at rank 1, or rank 2 if their sibling happened to also be the gallery. Twins (and triplets etc.) constituted 3.3%

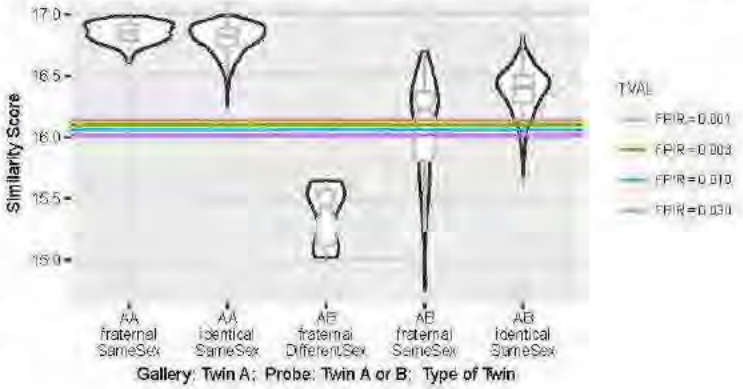


Figure 2: Intra- and inter-twin scores

of all live births [17] in recent years[4], and because that number is higher today than when the individuals in current adult databases were born, the false positives that arise from twins are now, and will increasingly be, an operational problem. Relative to the United States, twins are born with considerable regional variation. For example they are much less common in East Asia, and much more common in Sub-Saharan Africa [22]. The presence of twins in the mugshot database is inevitable given its size, around 12.3 million people. As this is not an insignificant sample of the domestic United States population, people with other familial ties will be present also. The data was collected over an extended period and because location information is not available, we are unable to estimate the proportion of the domestic population that is present in the dataset. However, if we assume twins are neither more or less disposed to arrest than the general population, we can estimate that hundreds of thousands of individuals in the dataset are twins. This will affect false positive rates because we randomly set aside 331 254 individuals for nonmate searches, and some proportion of those will be twins with siblings in the gallery.

[3] The gallery size here is 12 million people, 26.1 million images. Given 331 254 non-mated searches, an exhaustive implementation of one-too-many search would execute 8.6 trillion comparisons. At a false positive identification rate of 0.0025 the number of false positives is, to first order, 828 corresponding to single-comparison false match rate of 828 / 8.6 trillion = 9.6 10^{-11} i.e. about 1 in 10 billion. Strictly this FMRR computation meaningful only for algorithms that implement 1:N search using N 1:1 comparisons, which is not always the case.
[4] See the CDC's National Vital Statistics Report for 2017: https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_08-508.pdf

▷ **False negatives from ageing:** A large source of error in long-run applications where subjects are not re-enrolled on a schedule is ageing. This is a function of the time elapsed between photographs. Change in facial appearance causes recognition similarity scores to decline such that over the longer term, accuracy will decline. All faces age and while this usually proceeds in a graceful and progressive manner, drug use can accelerate this [30]. Elective surgery may be effective in delaying it although this has not been formally quantified with face recognition. As ageing is essentially unavoidable, it can only be mitigated by scheduled re-capture, as in passport re-issuance. To quantify ageing effects, we used the more accurate algorithms to enroll the earliest image of 3.1 million adults and then search with 10.3 million newer photos taken up to 18 years after the the initial enrollment photo. In the inset table, accuracy is seen to degrade progressively with time, as mate scores decline and non-mates displace mates from rank 1 position. More accurate algorithms tend to be less sensitive to ageing. The more accurate algorithms give fewer errors after 18 years of ageing than middle tier algorithms give after four. Note also we do not quantify an ageing rate - more formal methods [2] borrowed from the longitudinal analysis literature have been published for doing so (given suitable repeated measures data). See Figures 62, 72 and 77.

Algorithm	Metric, FNIR@	(0,2]	(2,4]	(4,6]	(6,8]	(8,10]	(10,12]	(12,14]	(14,18]
nec-2	Rank = 1	0.3	0.4	0.4	0.4	0.4	0.5	0.6	0.4
microsoft-4	Rank = 1	0.3	0.5	0.6	0.7	0.9	1.0	1.3	1.6
yitu-4	Rank = 1	0.6	0.8	0.8	0.8	0.9	1.1	1.5	2.1
everai-3	Rank = 1	0.5	0.7	0.9	1.1	1.3	1.5	1.8	2.2
idemia-4	Rank = 1	1.1	1.5	1.9	2.3	2.8	3.1	3.7	5.1
cogent-3	Rank = 1	0.8	1.1	1.3	1.5	1.7	1.9	2.4	3.1
cognitec-2	Rank = 1	1.0	1.4	1.7	2.0	2.4	2.6	3.1	3.9
nec-2	FPIR = 0.001	0.7	0.9	1.1	1.3	1.5	1.7	2.1	2.7
microsoft-4	FPIR = 0.001	2.7	4.7	7.2	10.1	12.9	16.1	20.5	25.9
yitu-4	FPIR = 0.001	1.2	2.0	3.1	4.7	6.7	9.6	14.2	20.1
everai-3	FPIR = 0.001	3.5	6.2	9.3	12.9	16.2	19.6	24.1	29.2
idemia-4	FPIR = 0.001	3.7	5.9	8.3	11.0	13.4	15.8	19.1	24.8
cogent-3	FPIR = 0.001	5.8	9.7	14.2	19.2	23.8	28.4	34.4	42.1
cognitec-2	FPIR = 0.001	5.2	8.8	12.7	17.1	21.0	24.6	29.2	35.3

Table 2: Impact of ageing on accuracy.

▷ **Image quality matters:** Poor quality photographs undermine recognition, either because the imaging system is poor (lighting, camera, etc.) or because the subject mis-presents to the camera (head orientation, facial expression, occlusion, etc.). Imaging problems can be mitigated by design i.e. ensuring adherence to long-standing face image capture standards. Presentation problems, however, must be detected at capture time, either by the photographer, or by an automated system, and re-capture performed. The most accurate algorithms in FRVT are highly tolerant of image quality problems. This derives from the invariances afforded by CNN-based algorithms, and this is the fundamental reason why accuracy has improved since 2013. For example, the Microsoft algorithms are can match many profile-view images to frontal mugshots - see Figures 100 and 102. As the inset table shows, rank-1 false negative identification rates are much higher with wild images than webcams and, in turn, mugshots. Further, even with the most capable algorithms, comparison scores are lower with unconstrained images, so that when (high) thresholds are necessary to limit false positives, here to 1 in 100 searches, error rates are very high. Such figures should guide prospective users of face recognition to consider whether face recognition can meet a formal written accuracy requirement.

Algorithm	Metric, FNIR@	Wild	Mugshot	Webcam
cognitec-3	Rank = 1	5.1	0.9	2.5
everai-3	Rank = 1	3.8	0.5	1.9
idemia-5	Rank = 1	4.4	1.1	3.9
microsoft-5	Rank = 1	3.3	0.3	1.1
nec-3	Rank = 1	8.8	0.3	1.0
ntechlab-6	Rank = 1	3.8	0.6	1.7
visionlabs-5	Rank = 1	4.3	0.4	1.9
yitu-4	Rank = 1	4.4	0.4	0.8
cognitec-3	FPIR = 0.01	32.5	2.8	10.0
everai-3	FPIR = 0.01	35.7	1.8	6.0
idemia-5	FPIR = 0.01	34.0	2.8	10.2
microsoft-5	FPIR = 0.01	34.4	1.2	4.1
nec-3	FPIR = 0.01	38.0	0.4	1.3
ntechlab-6	FPIR = 0.01	38.1	2.1	5.9
visionlabs-5	FPIR = 0.01	34.4	2.2	8.7
yitu-4	FPIR = 0.01	30.6	0.7	1.7

Table 3: Impact of image quality on accuracy.

▷ **Accuracy in large populations:** This report documents identification accuracy in galleries containining up to 12 million people and 26.1 million images. False negative rates climb very slowly as population size increases. For the most accurate algorithm, NEC-2, when searching a database of size 640 000, about 0.26% of searches fail to produce the

correct mate as its best hypothesized identity. In a database of 12 000 000 this rises to just 0.31%. This benign growth in miss rates is fundamentally the reason for the utility of face recognition in large scale one-to-many search applications. See Table 12 and Figure 22.

The reason for this is that as more identities are enrolled into an database, the possibility of a false positive increases due to lookalike faces that yield extreme values from the right tail of the non-mate score distribution. However, these scores are lower than most mate scores such that when an identification algorithm is configured with a threshold of zero (so human adjudication is always necessary), rank-one identification miss rates scale very favorably with population size, N, growing slowly, approximately as a power law, aN^b with $b \ll 1$. This dependency was first noted in 2010. Depending on the algorithm, the exponent b for mugshot searches is low, around 0.06 for the some of the more accurate algorithms with up to 12 million identities. See Table 12.

In any case, variations in accuracy with increasing population size are small relative to both ageing and algorithm choice. See Figure 20.

▷ **Utility of adjudicating long candidate lists:** In the regime where a system is configured with a threshold of zero, and where human adjudication is always necessary, the reviewer will find some mates quite far down candidate lists. This usually occurs because either the probe image or its corresponding enrolled mate image have poor quality, or large time-lapse. The accuracy benefits of traversing say 50 candidates versus just the first one is broadly a reduction in error by up to a factor of two. See Figure 30 and compare Tables 12 and 13.

However, accuracy from the leading algorithm is now so high - mates that in 2013 were placed at rank > 1, are now at rank 1 - such that reviewers can expect to review substantially fewer candidates. Note, however, for the proportion of searches where there is no mate, reviewers might still examine all candidates, fruitlessly. This report does not address the issue of human error in adjudicating candidates produced in one-to-many searches.

▷ **Utility of enrolling multiple images per subject:** We run three kinds of enrollment: First, by enrolling just the most recent image; second by creating a single template from a person's full lifetime history of images; and third by enrolling multiple images of a person separately, as though under different identities. The overall effect is that the enrollment of multiple images yields as much as a factor of two lower miss rates. This occurs due to higher information content and because the most recent image may sometimes be of poorer quality than historical images. See Table 12.

Gains depend on the number of available images: FNIR drops steadily. Some algorithms reduce FPIR or maintain it - the desirable behaviors - but others give higher false positive rates. See Figures leading up to Figure 87.

▷ **Reduced template sizes:** There has been a trend toward reduced template sizes, i.e. a smaller feature representation of an image. In 2014, the most accurate algorithm used a template of size 2.5KB; the figure in 2018 is around 1600 bytes. Close competitors produce templates of size 256, 364, 512, and about 2KB bytes. In 2014, the leading competitors had templates of size 4KB to 8KB. Some algorithms, when enrolling more than one image of a person, produce a template whose size is independent of the number of images given to the algorithm. This can be achieved by selecting a "best" image, or by integrating (fusing) information from the images. See Table 16.

▷ **Template generation times:** Template generation times, as measured on a single circa-2016 server processor core [5], vary from below 20 milliseconds up to nearly 1 second. This wide variation across developers may be relevant to end-users who have high-volume workflows. There has not been a wide downward trend since 2014. Note that speed may be expedited over the figure reported here by exploiting new vector instructions on recent chips. Note that GPUs were not used and, while indispensable for training CNNs, are not necessary for feeding an image forward through a network. See Table 16.

▷ **Search duration and scalability:** Template search times, as measured on circa-2016 Intel server processor cores,

[5] Intel Xeon CPU E5-2630 v4 running at 2.20GHz.

vary massively across the industry. For a database of size 1 million subjects, and the more accurate implementations, durations range from below 1 to 500 milliseconds, with other less accurate algorithms going much slower still. Several algorithms exhibit sublinear search time i.e. the duration does not double with a doubling of the enrolled population size, N. This was noted also in 2014. This has improved in 2018, however, such that close-to-logarithmic growth is evident for several developers' algorithms and extremely fast search. The consequence of this is that as N increases even the fastest linear algorithm (NEC-3) will quickly become much slower than the strongly sublinear algorithms. For the Dermalog-5 algorithm, search of a template against a database of N = 12 million images takes 850 microseconds on a single core of a contemporary CPU. That number is faster than any other algorithm even with the smallest gallery we tested (N = 640 000). See Table 6 and Figure 111.

▷ **Accuracy gains June - October 2018** NIST Interagency Report 8238 documented massive gains from 2013 to 2018. This report shows most developers achieved gains over the four month interval between June and October 2018. For a set of 12 million subjects enrolled with their most recent mugshot image, the inset table shows, for selected algorithms, the proportion of searches where mates are not returned against the given criteria (column 2). The result is that substantial reductions in false negatives - by a factor of two or more - were realized by algorithms submitted by Cogent, Cognitec, Dermalog, Hikvision, Innovatrics, NEC, Rank One, Shaman, Tiger-IT, and Vigilant Solutions. In particular, in this same time period one developer, NEC, which had produced broadly the most accurate algorithms in 2010 and 2013, submitted algorithms that are substantially more accurate than their June 2018 versions, and on many measures are now the most accurate. A number of other developers produced slightly less accurate implementations.

Application	Metric	Algorithm		FNIR
Mode: Mugshot	Miss rate	Date	Name	
Investigation	at Rank=1	2018-JUN	NEC-0	3.20%
Investigation	at Rank=1	2018-OCT	NEC-2	0.31%
Investigation	at Rank=1	2018-JUN	Microsoft-4	0.45%
Investigation	at Rank=1	2018-OCT	Microsoft-5	0.52%
Investigation	at Rank=1	2018-JUN	Yitu-2	0.55%
Investigation	at Rank=1	2018-OCT	Yitu-5	0.55%
Identification	at FPIR=0.001	2018-JUN	NEC-0	20.0%
Identification	at FPIR=0.001	2018-OCT	NEC-3	5.8%
Identification	at FPIR=0.001	2018-JUN	Microsoft-4	15.8%
Identification	at FPIR=0.001	2018-OCT	Microsoft-6	15.6%
Identification	at FPIR=0.001	2018-JUN	Yitu-2	12.4%
Identification	at FPIR=0.001	2018-OCT	Yitu-5	11.1%

Table 4: Accuracy gains since June - October 2018

See Tables 16 and 19, and Figure 19.

▷ **Non-technical considerations:** Recognition accuracy is likely the most important technical indicator for an algorithm. But even among the more accurate developers accuracy, template size, and resource consumption vary widely. This, incidentally, implies that technological diversity remains, that there is no consensus on approach and that algorithms are not commoditized. But beyond the performance statements in this report, face recognition outcomes in complete systems will be influenced by things like code and model size, software maturity, extensibility, reliability, ease of integration and maintenance, cost, availability of monitoring tools, and support for human review of true and false matches using, for example, capable graphical user interfaces.

▷ **Conclusions:** As with other biometrics, accuracy of facial recognition implementations varies greatly across the industry. Absent other performance or economic parameters, users should prefer the most accurate algorithm. Note that accuracy, and algorithm rankings, vary somewhat with the kinds of images used and the mode of operation: investigation with zero threshold; or identification with high threshold.

▷ **Supplementary Data:** This document is accompanied by a supplement that includes a three page report for each of the algorithms evaluated. Each report includes various performance plots pertinent to the particular algorithm under test. The supplement, which currently runs to more than 600 pages, is available from the same webpage as this report.

Release Notes

FRVT Activities: NIST restarted FRVT's one-to-many track in February 2018, inviting participants to send up to seven prototype algorithms. Since February 2017, NIST has been evaluating one-to-one verification algorithms on an ongoing basis. This allows developers to submit updated algorithms to NIST at any time but no more frequently than four calendar months. This more closely aligns development and evaluation schedules. Results are posted to the web within a few weeks of submission. Details and full report are linked from the Ongoing FRVT site.

FRVT Reports: The results of the FRVT appear in the series NIST Interagency Reports tabulated below. The reports were developed separately and released on different schedules. In prior years NIST has mostly reported FRVT results as a single report; this had the disadvantage that results from completed sub-studies were not published until all other studies were complete.

Date	Link	Title	No.
2014-03-20	PDF	FRVT Performance of Automated Age Estimation Algorithms	7995
2015-04-20	PDF	Face Recognition Vendor Test (FRVT) Performance of Automated Gender Classification Algorithms	8052
2014-05-21	PDF	FRVT Performance of face identification algorithms	8009
2017-03-07	PDF	Face In Video Evaluation (FIVE) Face Recognition of Non-Cooperative Subjects	8173
2017-11-23	PDF	The 2017 IARPA Face Recognition Prize Challenge (FRPC)	8197
2018-04-13	WWW	Ongoing Face Recognition Vendor Test (FRVT)	Draft

Details appear on pages linked from `https://www.nist.gov/programs-projects/face-projects`.

Appendices: This report is accompanied by appendices which present exhaustive results on a per-algorithm basis. These are machine-generated and are included because the authors believe that visualization of such data is broadly informative and vital to understanding the context of the report.

Typesetting: Virtually all of the tabulated content in this report was produced automatically. This involved the use of scripting tools to generate directly type-settable LaTeX content. This improves timeliness, flexibility, maintainability, and reduces transcription errors.

Graphics: Many of the Figures in this report were produced using the ggplot2 package running under R, the capabilities of which extend beyond those evident in this document.

Contents

1 Introduction

One-to-many identification represents the largest market for face recognition technology. Algorithms are used across the world in a diverse range of biometric applications: detection of duplicates in databases, detection of fraudulent applications for credentials such as passports and driving licenses, token-less access control, surveillance, social media tagging, lookalike discovery, criminal investigation, and forensic clustering.

This report contains a breadth of performance measurements relevant to many applications. Performance here refers to accuracy and resource consumption. In most applications, the core accuracy of a facial recognition algorithm is the most important performance variable. Resource consumption will be important also as it drives the amount of hardware, power, and cooling necessary to accommodate high volume workflows. Algorithms consume processing time, they require computer memory, and their static template data requires storage space. This report documents these variables.

1.1 Open-set searches

FRVT tested open-set identification algorithms. Real-world applications are almost always "open-set", meaning that some searches have an enrolled mate, but some do not. For example, some subjects have truly not been issued a visa or drivers license before; some law enforcement searches are from first-time arrestees[6]. In an "open-set" application, algorithms make no prior assumption about whether or not to return a high-scoring result, and for a mated search, the ideal behaviour is that the search produces the correct mate at high score and first rank. For a non-mate search, the ideal behavior is that the search produces zero high-scoring candidates.

Too many academic benchmarks execute only closed-set searches. The proportion of mates found in the rank one position is the default accuracy metric. This hit rate metric ignores the score with which a mate is found; weak hits count as much as strong hits. This ignores the real-world imperative that in many applications it is necessary to elevate a threshold to reduce the number of false positives.

2 Evaluation datasets

This report documents accuracy for four kinds of images - mugshots, webcam, profiles and wild - as described in the following sections.

2.1 Mugshot images

The main mugshot dataset used is referred to as the FRVT 2018 set. This set was collected over the period 2002 to 2017 in routine United States law enforcement operations. This set has been extracted from a larger operational parent set by excluding non-face images, and setting aside webcam and profile-view images, for use in separate tests.

NIST Interagency Report 8238 includes a comparison of this set of mugshots with the smaller and easier sets of mugshots used in tests run in 2010 and 2014.

[6]Operationally closed-set applications are rare because it is usually not the case that all searches have an enrolled mate. One counter-example, however, is a cruise ship in which all passengers are enrolled and all searches should produce exactly one identity. Another example is forensic identification of dental records from an aircraft crash.

- **Mugshots**: Mugshots comprise about 86% of the database. They have reasonable compliance with the ANSI/NIST ITL1-2011 Type 10 standard's subject acquisition profiles levels 10-20 for frontal images [28]. The most common departure from the standard's requirements is the presence of mild pose variations around frontal - the images of Figure 3 are typical. The images vary in size, with many being 480x600 pixels with JPEG compression applied to produce filesizes of between 18 and 36KB with many images outside this range, implying that about 0.5 bits are being encoded per pixel.
- **Profile images**: Profile-view images have been collected in law enforcement for more than 100 years, as human capability is improved with orthogonal information. The profile images used in this report were collected during the same session as the frontal mugshot photograph, in the same standardized photographic setup. These would not therefore be used with automated face recognition. A small subset, 200 000 images, were set aside for testing.
- **Webcam images**: The remaining 14% of the images were collected using an inexpensive webcam attached to a flexible operator-directed mount. These images are all of size 240x240 pixels, that are in considerable violation of most quality-related clauses of all face recognition standards. As evident in the figure, the most common defects are non-frontal pose (associated with the rotational degrees of freedom of the camera mount), low contrast (due to varying and intense background lights), and poor spatial resolution (due to inexpensive camera optics) - see examples in Fig 4. The images are overly JPEG compressed, to between 4 and 7KB, implying that only 0.5 to 1 bits are being encoded per color pixel.

Example images are shown in Figures 3, 4 and 5 These are drawn from NIST Special Database 32 which may be downloaded here.

These images were partitioned in galleries and probesets for the various experiment listed in Table 5.

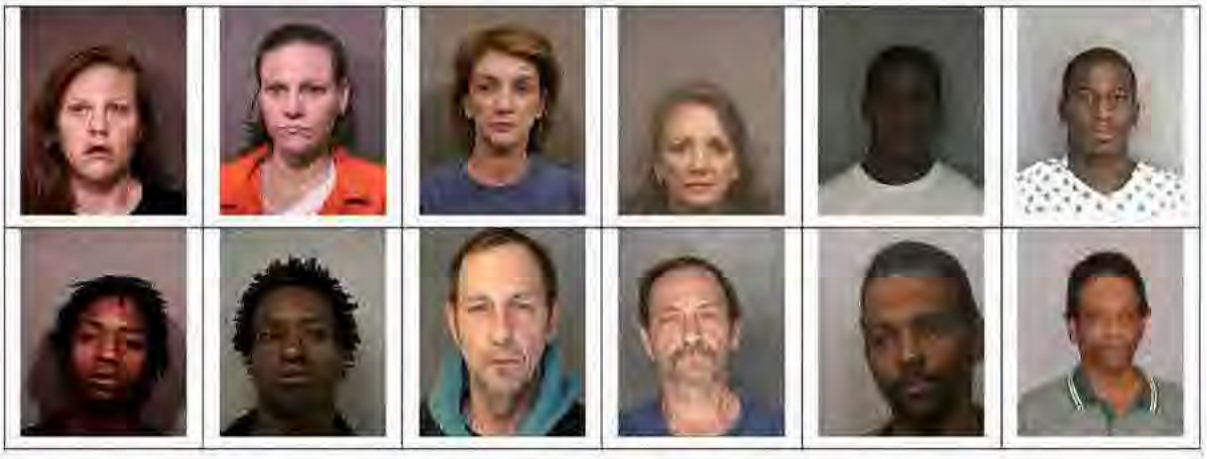


Figure 3: Six mated mugshot pairs representative of the FRVT-2014 (LEO) and FRVT-2018 datasets. The images are collected live, i.e. not scanned from paper. Image source: NIST Special Database 32

Figure 5: **[Profile views]** *The three images are a frontal enrollment, subsequent frontal probe, and same-session ninety degree profile view. While collection of both frontal and profile views has been typical in law enforcement for more than a century, the recognition of profile to frontal views has essentially been impossible. However, reasonbly high accuracy results is now possible - see section E.*

Figure 4: Twelve webcam images representative of probes against the FRVT-2018 mugshot gallery. The first eight images are four mated pairs. Such images present challenges to recognition including pose, non-uniform illumination, low contrast, compression, cropping, and low spatial sampling rate. Image source: NIST Special Database 32

2.2 Unconstrained "wild" images

In addition to portrait-styled mugshots, algorithms were also evaluated on a "wild" dataset composed of non-cooperative and unconstrained photojournalism and amateur photography imagery. The images are closely cropped from the parent images as shown in Figure 6. A portion of the images are collected by professional photographers and as such are captured, and selected, to not exhibit exposure and focus problems. Some of the photos were downloaded from websites with substantial amateur photographer imagery, which may contain images that do exhibit exposure and focus problems. Resolution varies widely as these images were downloaded from the internet with varying resampling and compression practices. The primary difficulties for face recognition is unconstrained yaw and pitch pose variation, with some images extending to profile view. Additionally faces can be occluded, including by hair and hands.

The images are cropped prior to passing them to the algorithm. The cropping is done per human-annotated rectangular bounding boxes. The algorithm must further localize the face and extract features. In many cases, there were multiple images of the subject provided to the algorithm, and the output was a single template representation of the subject.

$N_P = 332\,574$ subjects were searched against two galleries, where the number of enrolled subjects in each gallery were $N_{G1} = 1\,106\,777$ and $N_{G2} = 1\,107\,778$. Both gallery and search images were composed of unconstrained wild imagery.

Figure 6: Examples of "in the wild" stills. The top row gives the full original images; the second row gives the manually specified face region that is cropped and passed to the algorithms. The source images in this figure are attributed to, from left, Rita Molnr, Eva Rinaldi, and Gage Skidmore under the [cc-by-sa-2.5], [cc-by-sa-2.0], [cc-by-sa-3.0] creative commons licenses respectively.

2.3 Enrollment strategies

Many operational applications include collection and enrollment of biometric data from subjects on more than one occasion. This might be done on a regular basis, as might occur in credential (re-)issuance, or irregularly, as might happen in a criminal recidivist situation [4]. The number of images per person will depend on the application area. In civil identity credentialing (e.g. passports, driver's licenses), the images will be acquired approximately uniformly over time (e.g. ten years for a passport). While the distribution of dates for such images of a person might be assumed uniform, a number of factors might undermine this assumption[7]. In criminal applications, the number of images would depend on the number of arrests. The distribution of dates for arrest records for a person (i.e. the recidivism distribution) has been modeled using the exponential distribution but is recognized to be more complicated[8].

In any case, the 2010 NIST evaluation of face recognition showed that considerable accuracy benefits accrue with retention and use of *all* historical images [6].

To this end, the FRVT API document provides $K \geq 1$ images of an individual to the enrollment software. The software is tasked with producing a single proprietary undocumented "black-box" template[9] from the K images. This affords the algorithm an ability to generate a *model* of the individual, rather than to simply extract features from each image on a sequential basis.

As depicted in Figure 7, the i-th individual in the FRVT 2018 dataset has K_i images. These are labelled as x_k for $k = 1 \ldots K_i$ in chronological order of capture date. To measure the utility of having multiple enrollment images, this report evaluates three kinds of enrollment:

[7] For example, a person might skip applying for a passport for one cycle, letting it expire. In addition, a person might submit identical images (from the same photography session) to consecutive passport applications at five year intervals.
[8] A number of distributions have been considered to model recidivism, see for example [3].
[9] There are no formal face template standards. Template standards only exist for fingerprint minutiae - see ISO/IEC 19794-2:2011.

Image				
Encounter	1	...	$K_i - 1$	K_i
Capture Time	T_1	...	T_{K_i-1}	T_{K_i}
Role RECENT	Not used	Not used	Enrolled	Search
Role LIFETIME	Enrolled	Enrolled	Enrolled	Search

Figure 7: Depiction of the "recent" and "lifetime" enrollment types. Image source: NIST Special Database 32

- **Recent**: Only the second most recent image, x_{K_i-1} is enrolled. This strategy of enrollment mimics the operational policy of retaining the imagery from the most recent encounter. This might be done operationally to ameliorate the effects of face ageing. Obviously retaining only the most recent image should only be done if the identity of the person is trusted to be correct. For example, in an access control situation retention of the most recent successful *authentication* image would be hazardous if it could be a false positive.
- **Lifetime-consolidated**: All but the most recent image are enrolled, $x_1 \ldots x_{K_i-1}$. This subject-centric strategy might be adopted if quality variations exist where an older image might be more suitable for matching, despite the ageing effect.
- **Lifetime-unconsolidated**: Again all but the most recent image are enrolled $x_1 \ldots x_{K_i-1}$ but now separately, with different identifiers, such that the algorithm is not aware that the images are from the same face. This kind of event- or encounter-centric enrollment is very common when operational constraints preclude reliable consolidation of the historical encounters into a single identity. This aspect also prevents the recognition algorithm from a) building a holistic model of identity (as is common in speaker recognition systems) and b) implementing fusion, for example template-level fusion of feature vectors, or post-search score-level fusion. The result is that searches will typically yield more than one image of a person in the top ranks. This has consequences for appropriate metrics, as detailed in section 3.2.1

  NIST first evaluated this kind of enrollment in mid 2018, and the results tables include some comparison of accuracy available from all three enrollment styles.

In all cases, the most recent image, x_{K_i}, is reserved as the search image. For the 1.6 million subject enrollment partition of the FRVT 2018 data, $1 \leq K_i \leq 33$ with $K_i = 1$ in 80.1% of the individuals, $K_i = 2$ in 13.4%, $K_i = 3$ in 3.7%, $K_i = 4$ in 1.4%, $K_i = 5$ in 0.6%, $K_i = 6$ in 0.3%, and $K_i > 6$ is 0.2% for everyone else. This distribution is substantially dependent on United States recidivism rates.

We did not evaluate the case of retaining only the highest quality image, since automated quality assessment is out of scope for this report. We do not anticipate that such strategies will prove beneficial when the quality assessment apparatus is imperfect and unvalidated.

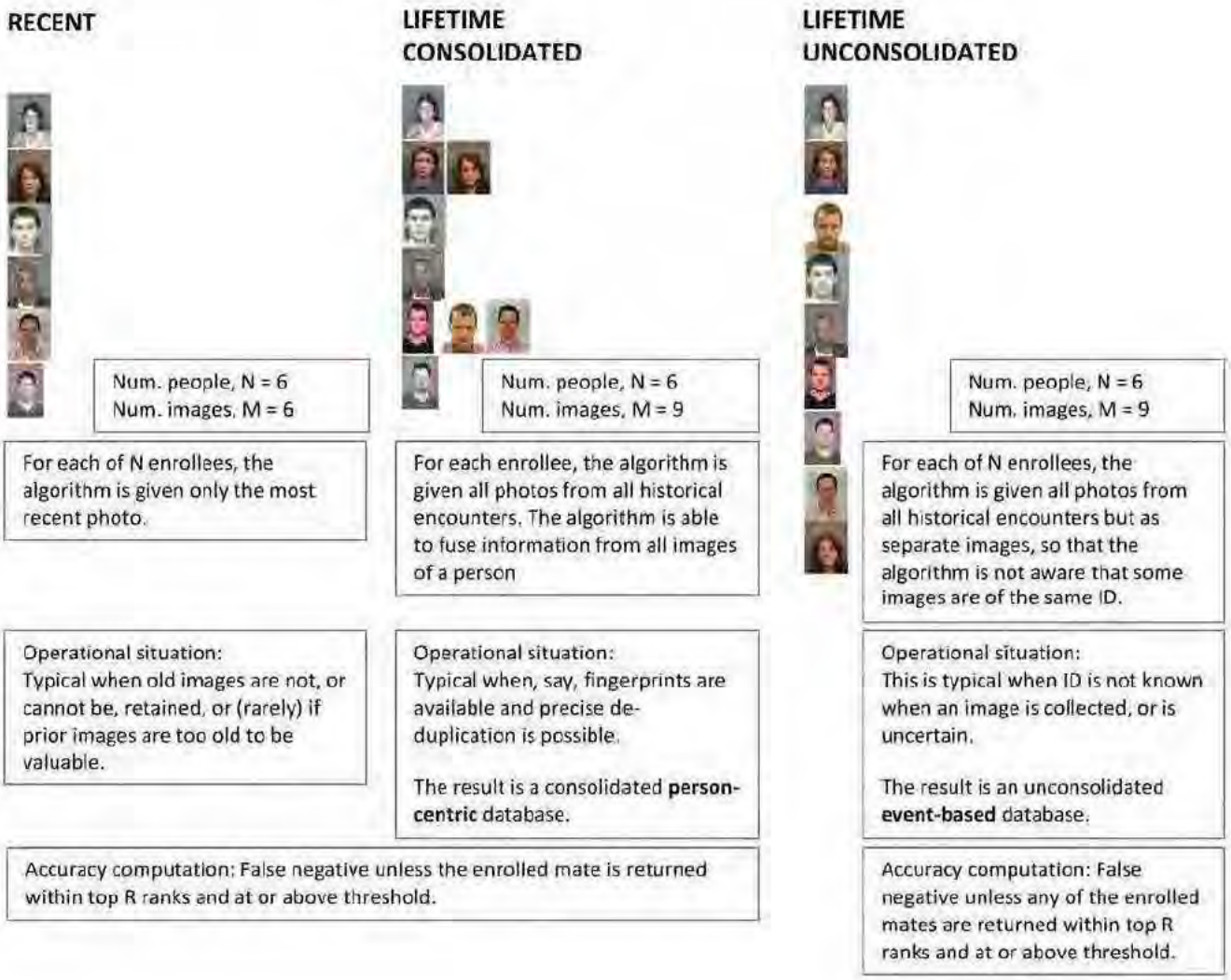


Figure 8: **Enrollment strategies.** *The figure shows the three kinds of enrollment databases examined in this report. Image source: NIST Special Database 32*

		ENROLLMENT			SEARCH			
	TYPE SEE SECTION 2.3	POPULATION FILTER	N-SUBJECTS	N-IMAGES	MATE N-SUBJECTS	MATE N-IMAGES	NON-MATE N-SUBJECTS	NON-MATE N-IMAGES
Mugshot trials from enrollment of single images								
1	RECENT	NATURAL	640 000	640 000	154 549	154 549	331 254	331 254
2	RECENT	NATURAL	1 600 000	1 600 000				
3	RECENT	NATURAL	3 000 000	3 000 000				
4	RECENT	NATURAL	6 000 000	6 000 000				
5	RECENT	NATURAL	12 000 000	12 000 000				
Mugshot trials from enrollment of lifetime images								
6	CONSOL	NATURAL	640 000	1 247 331				
7	CONSOL	NATURAL	1 600 000	3 351 206				
8	CONSOL	NATURAL	3 000 000	6 417 057				
9	CONSOL	NATURAL	6 000 000	12 976 185				
10	CONSOL	NATURAL	12 000 000	26 107 917				
11	UN-CONSOL	NATURAL	640 000	1 247 331				
12	UN-CONSOL	NATURAL	1 600 000	3 351 206				
Cross-domain								
13	MUGSHOTS AS ON ROW 2				82 106 WEBCAM	82 106 WEBCAM	331 254 WEBCAM	331 254 WEBCAM
Cross-view								
14	MUGSHOTS AS ON ROW 2				100 000 PROFILE	100 000 PROFILE	100 000 PROFILE	100 000 PROFILE
Ageing								
17	OLDEST	NATURAL	3 068 801	3 068 801	2 853 221	10 951 064	0	0

Table 5: **Enrollment and search sets.** *Each row summarizes one identification trial. Unless stated otherwise, all entries refer to mugshot images. The term "natural" means that subjects were selected without heed to demographics, i.e. in the distribution native to this dataset. The probe images were collected in a different calendar year to the enrollment image. Missing values in rows 2-12 are the same as in row 1.*

3 Performance metrics

This section gives specific definitions for accuracy and timing metrics. Tests of open-set biometric algorithms must quantify frequency of two error conditions:

- **False positives**: Type I errors occur when search data from a person who has never been seen before is incorrectly associated with one or more enrollees' data.
- **Misses**: Type II errors arise when a search of an enrolled person's biometric does not return the correct identity.

Many practitioners prefer to talk about "hit rates" instead of "miss rates" - the first is simply one minus the other as detailed below. Sections 3.1 and 3.2 define metrics for the Type I and Type II performance variables.

Additionally, because recognition algorithms sometimes fail to produce a template from an image, or fail to execute a one-to-many search, the occurrence of such events must be recorded. Further because algorithms might elect to not produce a template from, for example, a poor quality image, these failure rates must be combined with the recognition error rates to support algorithm comparison. This is addressed in section 3.5.

Finally, section 3.7 discusses measurement of computation duration, and section 3.8 addresses the uncertainty associated with various measurements. Template size measurement is included with the results.

3.1 Quantifying false positives

It is typical for a search to be conducted into an enrolled population of N identities, and for the algorithm to be configured to return the closest L candidate identities. These candidates are ranked by their score, in descending order, with all scores required to be greater than or equal to zero. A human analyst might examine either all L candidates, or just the top $R \leq L$ identities, or only those with score greater than threshold, T. The workload associated with such examination is discussed later, in 3.6.

False alarm performance is quantified in two related ways. These express how many searches produces false positives, and then, how many false positives are produced in a search.

False positive identification rate: The first quantity, FPIR, is the proportion of non-mate searches that produce an adverse outcome:

$$\text{FPIR}(N,T) = \frac{\text{Num. non-mate searches where one or more enrolled candidates are returned with score at or above threshold}}{\text{Num. non-mate searches attempted.}} \quad (1)$$

Under this definition, FPIR can be computed from the highest non-mate candidate produced in a search - it is not necessary to consider candidates at rank 2 and above. FPIR is the primary measure of Type I errors in this report.

Selectivity: However, note that in any given search, several non-mate may be returned above threshold. In order to quantify such events, a second quantity, selectivity (SEL), is defined as the *number* of non-mates returned on a candidate list, averaged over all searches.

$$\text{SEL}(N,T) = \frac{\text{Num. non-mate enrolled candidates returned with score at or above threshold}}{\text{Num. non-mate searches attempted.}} \quad (2)$$

where $0 \leq \text{SEL(N, T)} \leq \text{L}$. Both of these metrics are useful operationally. FPIR is useful for targeting how often an adverse false positive outcome can occur, while SEL as a number is related to workload associated with adjudicating candidate lists. The relationship between the two quantities is complicated - it depends on whether an algorithm concentrates the false alarms in the results of a few searches or whether it disburses them across many. This was detailed in FRVT 2014, NISTIR 8009. It has not yet been detailed in FRVT 2018.

3.2 Quantifying hits and misses

If L candidates are returned in a search, a shorter candidate list can be prepared by taking the top $R \leq L$ candidates for which the score is above some threshold, $T \geq 0$. This reduction of the candidate list is done because thresholds may be applied, and only short lists might be reviewed (according to policy or labor availability, for example). It is useful then to state accuracy in terms of R and T, so we define a "miss rate" with the general name **false negative identification rate** (FNIR), as follows:

$$\text{FNIR}(N,R,T) = \frac{\text{Num. mate searches with enrolled mate found outside top R ranks or score below threshold}}{\text{Num. mate searches attempted.}} \quad (3)$$

This formulation is simple for evaluation in that it does not distinguish between causes of misses. Thus a mate that is not reported on a candidate list is treated the same as a miss arising from face finding failure, algorithm intolerance of poor quality, or software crashes. Thus if the algorithm fails to produce a candidate list, either because the search

failed, or because a search template was not made, the result is regarded as a miss, adding to FNIR.

Hit rates, and true positive identification rates: While FNIR states the "miss rate" as how often the correct candidate is either not above threshold or not at good rank, many communities prefer to talk of "hit rates". This is simply the **true positive identification rate**(TPIR) which is the complement of FNIR giving a positive statement of how often mated searches are successful:

$$\text{TPIR}(N, R, T) = 1 - \text{FNIR}(N, R, T) \tag{4}$$

This report does not report true positive "hit" rates, preferring false negative miss rates for two reasons. First, costs rise linearly with error rates. For example, if we double FNIR in an access control system, then we double user inconvenience and delay. If we express that as decrease of TPIR from, say 98.5% to 97%, then we mentally have to invert the scale to see a doubling in costs. More subtly, readers don't perceive differences in numbers near 100% well, becoming inured to the "high nineties" effect where numbers close to 100 are perceived indifferently.

Reliability is a corresponding term, typically being identical to TPIR, and often cited in automated (fingerprint) identification system (AFIS) evaluations.

An important special case is the **cumulative match characteristic**(CMC) which summarizes accuracy of mated-searches only. It ignores similarity scores by relaxing the threshold requirement, and just reports the fraction of mated searches returning the mate at rank R or better.

$$\text{CMC}(N, R) = 1 - \text{FNIR}(N, R, 0) \tag{5}$$

We primarily cite the complement of this quantity, $\text{FNIR}(N, R, 0)$, the fraction of mates *not* in the top R ranks.

The **rank one hit rate** is the fraction of mated searches yielding the correct candidate at best rank, i.e. CMC(N, 1). While this quantity is the most common summary indicator of an algorithm's efficacy, it is not dependent on similarity scores, so it does not distinguish between strong (high scoring) and weak hits. It also ignores that an adjudicating reviewer is often willing to look at many candidates.

3.2.1 False negative rates for unconsolidated galleries

As detailed in section 2.3 a common type of gallery, here referred to as the lifetime unconsolidate type, is populated with all images of an individual without any association between them. That is, the gallery construction algorithm is not provided with any ID labels that would support processing of a person's images jointly. This constrasts with the lifetime consolidate type where an algorithm may explicitly fuse features from multiple images of a person, or select a best image. In such cases, where the number of enrolled images is a random variable, we define two false negative rates as follows.

The first demands that the algorithm place any of the K_i mates in the top $R \geq 1$ ranks. The proportion of searches for which this does not occur forms a false negative identification rate:

$$\text{FNIR}_{\text{any}}(N, R, T) = 1 - \frac{\text{Num. mate searches where any enrolled mate is found in the top R ranks and at-or-above threshold}}{\text{Num. mate searches attempted.}} \tag{6}$$

The second demands that the algorithm place all K_i mates in the top $R \geq K_i$ ranks. The proportion of searches for

which this does not occur forms a false negative identification rate:

$$\mathrm{FNIR}_{\mathbf{all}}(N,R,T) = 1 - \frac{\text{Num. mate searches where all enrolled mates are found in the top R ranks and at-or-above threshold}}{\text{Num. mate searches attempted.}} \tag{7}$$

Placing all mates in the top ranks is a more difficult task than correctly retrieving any image, so it holds that: $\mathrm{FNIR}_{\mathbf{all}} \geq \mathrm{FNIR}_{\mathbf{any}}$. This is evident in the results presented for November 2018 algorithms in Tables starting at 25.

The information retrieval community might prefer to compute and plot *precision* and *recall*; this is a valid approach, but we advance the two metrics above because they relate to our normal definition of consolidated FNIR, and they cover the two extreme use-cases of wanting any hit vs. all hits.

3.3 DET interpretation

In biometrics, a false negative occurs when an algorithm fails to match two samples of one person – a Type II error. Correspondingly, a false positive occurs when samples from two persons are improperly associated – a Type I error.

Matches are declared by a biometric system when the native comparison score from the recognition algorithm meets some threshold. Comparison scores can be either similarity scores, in which case higher values indicate that the samples are more likely to come from the same person, or dissimilarity scores, in which case higher values indicate different people. Similarity scores are traditionally computed by fingerprint and face recognition algorithms, while dissimilarities are used in iris recognition. In some cases, the dissimilarity score is a distance possessing metric properties. In any case, scores can be either mate scores, coming from a comparison of one persons samples, or nonmate scores, coming from comparison of different persons samples.

The words "genuine" or "authentic" are synonyms for mate, and the word "impostor" is used as a synonym for nonmate. The words "mate" and "nonmate" are traditionally used in identification applications (such as law enforcement search, or background checks) while genuine and impostor are used in verification applications (such as access control).

An error tradeoff characteristic represents the tradeoff between Type II and Type I classification errors. For identification this plots false negative vs. false positive identification rates i.e. FNIR vs. FPIR parametrically with T. Such plots are often called detection error tradeoff (DET) characteristics or receiver operating characteristic (ROC). These serve the same function – to show error tradeoff – but differ, for example, in plotting the complement of an error rate (e.g. TPIR = 1 − FNIR) and in transforming the axes, most commonly using logarithms, to show multiple decades of FPIR. More rarely, the function might be the inverse of the Gaussian cumulative distribution function.

The slides of Figures 9 through 15 discuss presentation and interpretation of DETs used in this document for reporting face identification accuracy. Further detail is provided in formal biometrics testing standards, see the various parts of ISO/IEC 19795 Biometrics Testing and Reporting. More terms, including and beyond those to do with accuracy, appear in ISO/IEC 2382-37 Information technology – Vocabulary – Part 37: Harmonized biometric vocabulary.

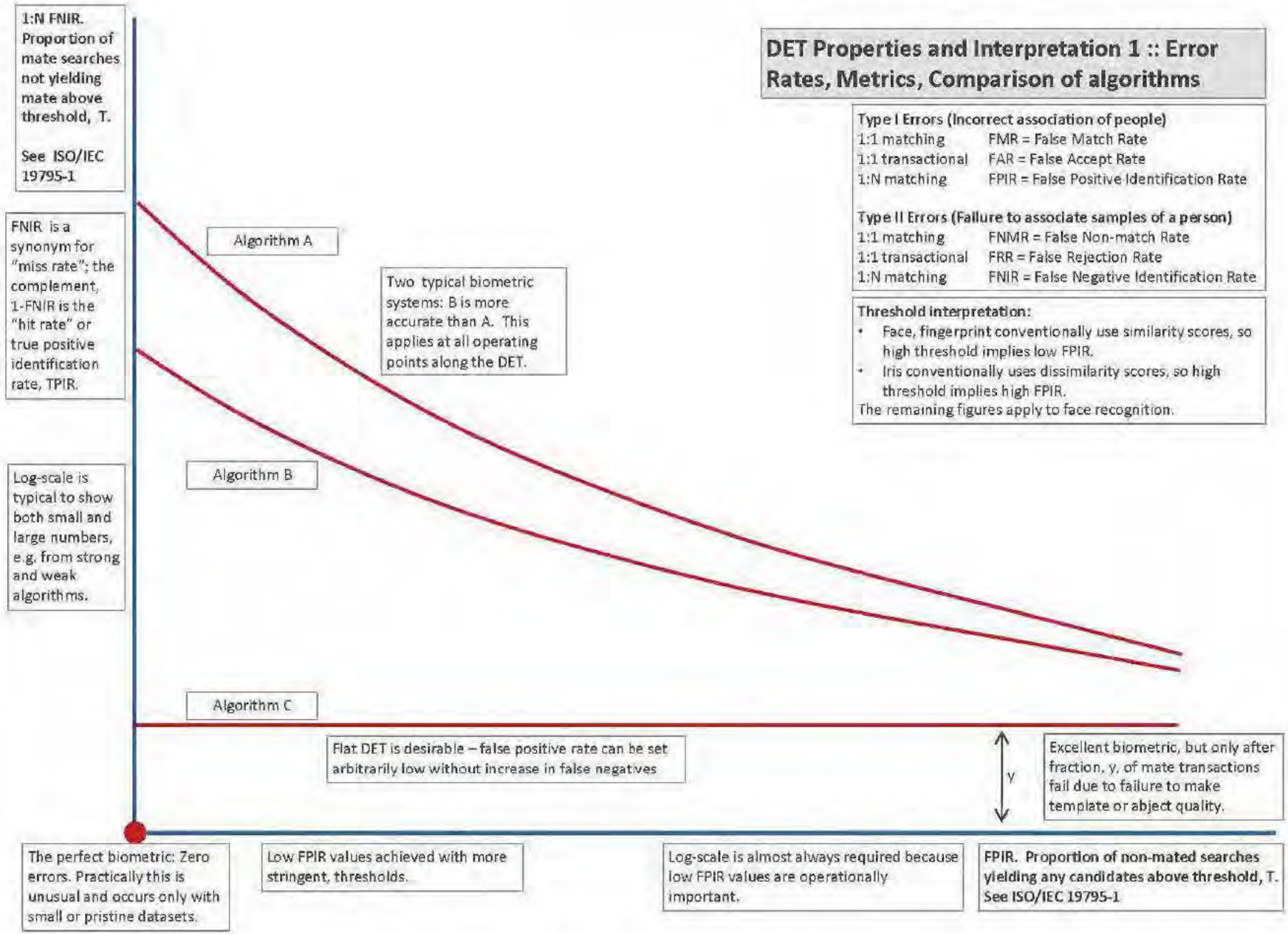


Figure 9: **DET as the primary performance reporting mechanism.**

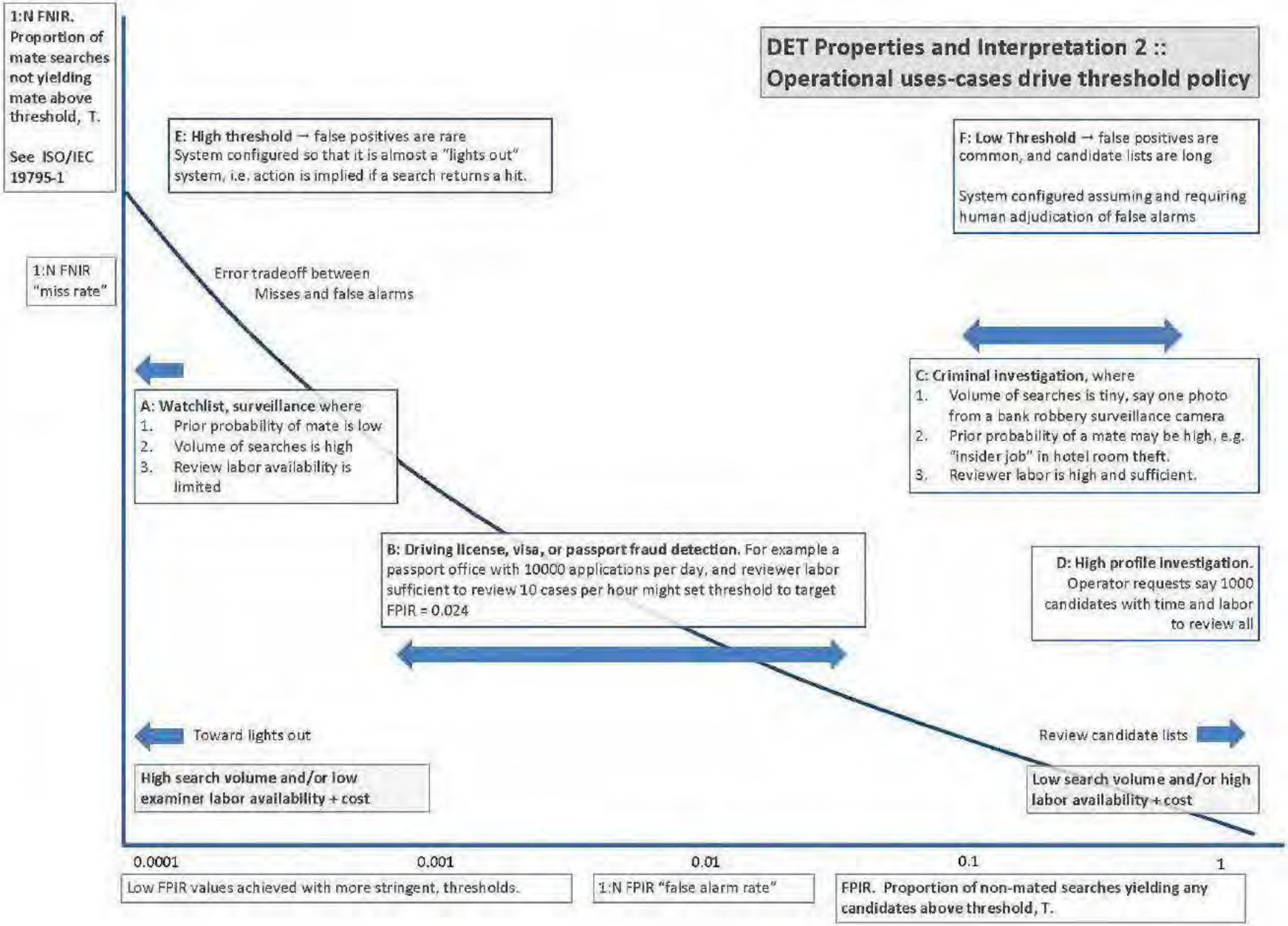


Figure 10: DET as the primary performance reporting mechanism.

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects.
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

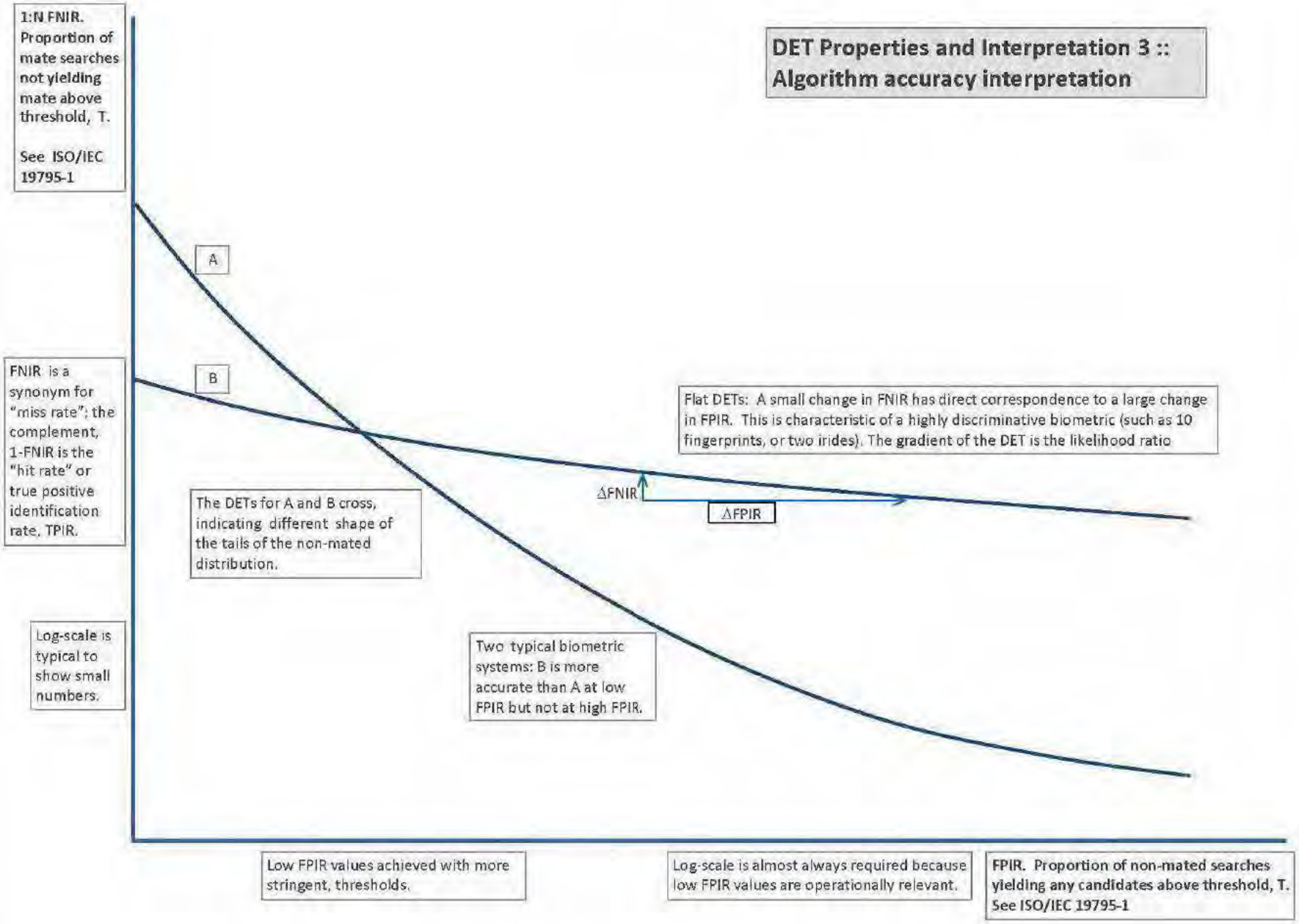


Figure 11: **DET as the primary performance reporting mechanism.**

DET Properties and Interpretation 4 :: Drivers of FNIR

The progressive rise in the DET, i.e. increasing FNIR, occurs when a search of a probe sample does not correctly return the enrolled mate. Leading causes of this are:

1. **Ageing:** Given sufficient time-lapse, the appearance of a face will change. This is a gradual process affecting all human faces and, absent surgical intervention, is essentially irreversible over long time-scales. Ageing increases false negative rates. In some applications ageing effects are avoided by policy: faces are re-enrolled periodically. In other applications, this is not possible.

2. **Image quality**: The leading cause of false negative recognition failure is that either or both images are in some sense defective. Quality can be degraded due to imaging problems (poor illumination, mis-focus etc.), mis-handling (cropping, (re-)compression) or resolution change) and commonly subject "misbehavior" (non-frontal pose, non-neutral expression). These effects depress similarity scores. Good design mitigates imaging and mis-handling errors.

Additional failures arise from clerical biographic error (two persons labelled with the same ID), person absent from the photo entirely.

(0,1)

1:N FNIR. Proportion of mate searches not yielding mate above threshold, T.

See ISO/IEC 19795-1

T = High

1. With ΔTime = 2 years, capable algorithms will return this mated pair with a high score. It will only contribute to FNIR at very high T. In children, growth is rapid and this will not hold +.

FNIR is a synonym for "miss rate"; the complement, 1-FNIR is the "hit rate" or true positive identification rate, TPIR.

2. With ΔTime = 12 years, even capable algorithms will return this mated pair with a moderate score. It will only contribute to FNIR at moderate T.

3. With mild changes in pose, illumination, and expression, weaker identification algorithms will assign low similarity scores such that this pair will contribute to FNIR at low T.

T = Low

T = 0

Low FPIR values achieved with higher, i.e. more stringent thresholds.

+ D. Michalski et al. *The Impact of Ageing on Facial Comparisons with Images of Children conducted by Humans and Automated Systems* January 2017 Proc. Soc. for Applied Research in Memory and Cognition, Sydney, Aus.

FPIR. Proportion of non-mated searches yielding any candidates above threshold, T. See ISO/IEC 19795-1

(1,0)

FNIR(N, R, T) = 1 − TPIR(N, T) = False neg. identification rate

N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

Figure 12: **DET as the primary performance reporting mechanism.**

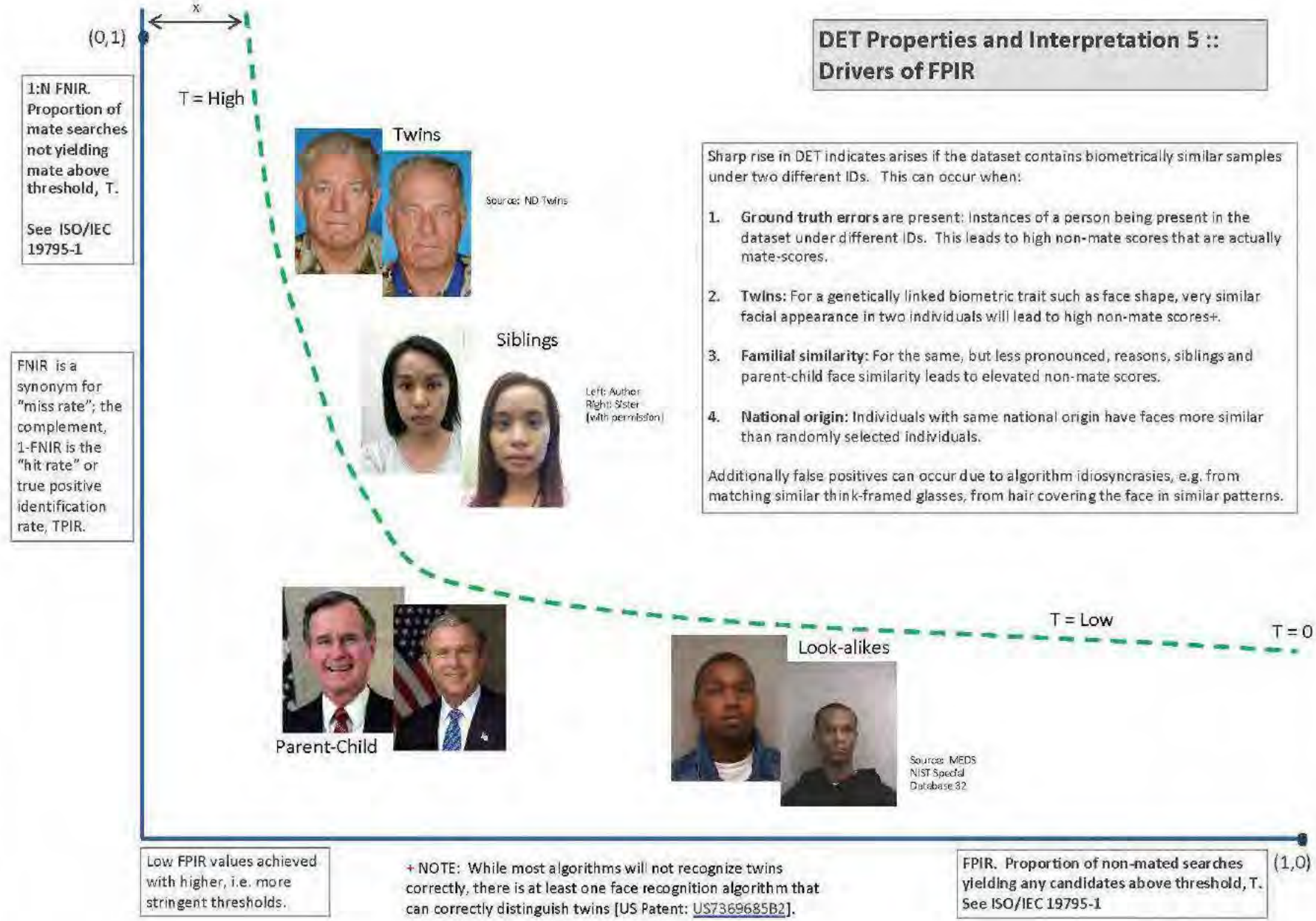


Figure 13: **DET as the primary performance reporting mechanism.**

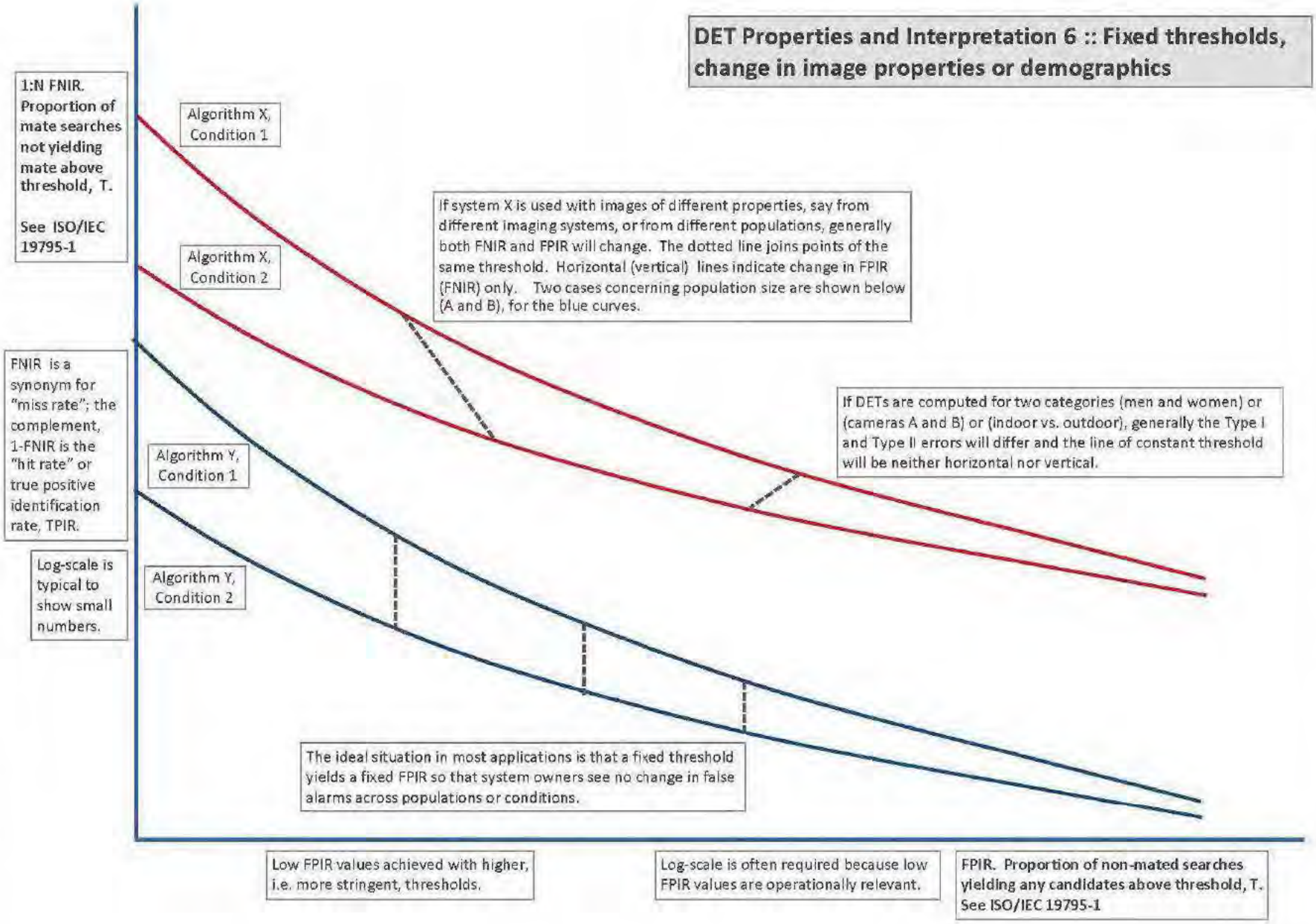


Figure 14: DET as the primary performance reporting mechanism.

Figure 15: **DET as the primary performance reporting mechanism.**

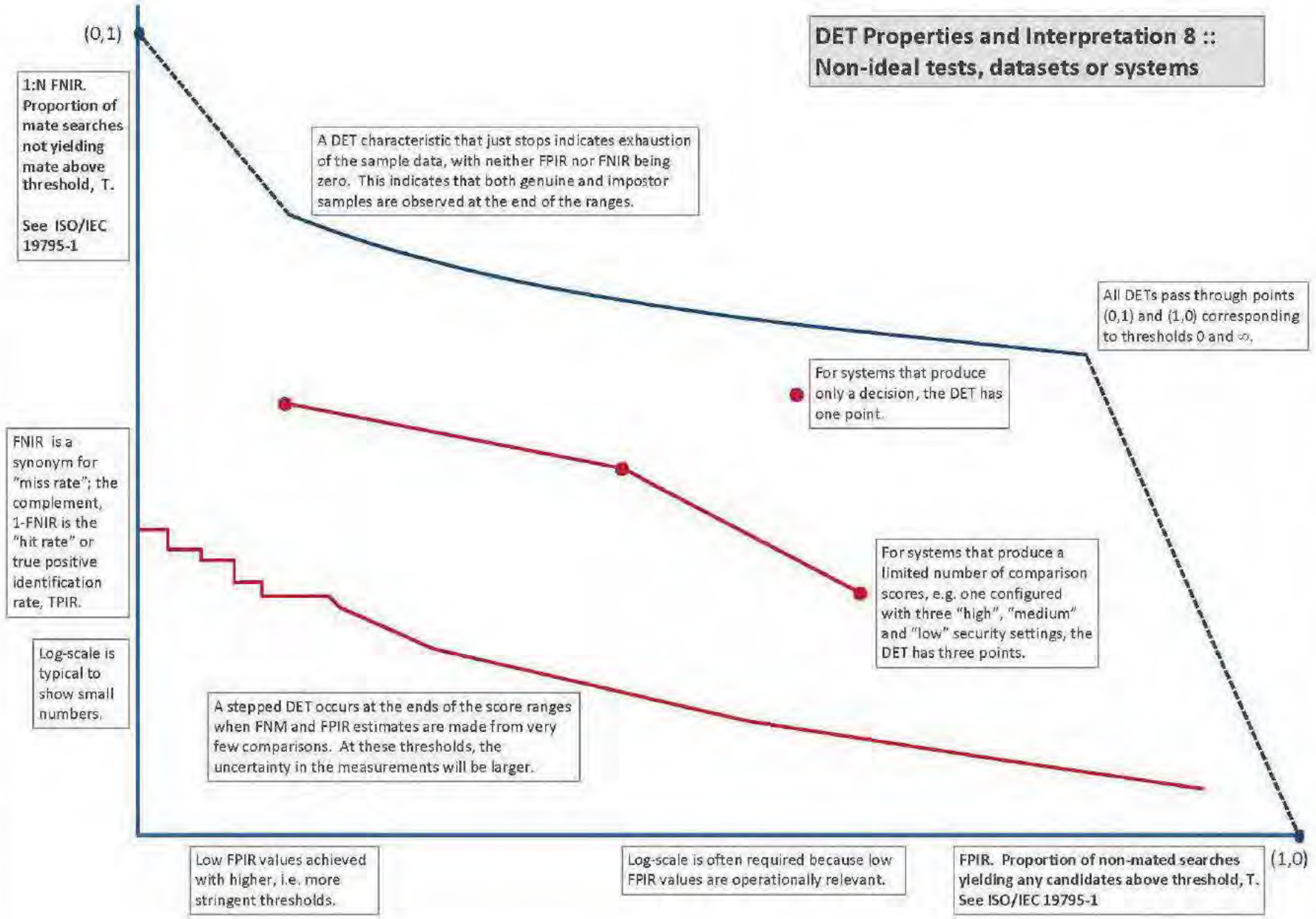


Figure 16: **DET as the primary performance reporting mechanism.**

3.4 Best practice testing requires execution of searches with and without mates

FRVT embeds 1:N searches of two kinds: Those for which there is an enrolled mate, and those for which there is not. The respective numbers for these types of searches appear in Table 5. However, it is common to conduct only mated searches[10]. The cumulative match characteristic is computed from candidate lists produced in mated searches. Even if the CMC is the only metric of interest, the actual trials executed in a test should nevertheless include searches for which no mate exists. As detailed in Table 5 the FRVT reserved disjoint populations of subjects for executing true non-mate searches.

3.5 Failure to extract features

During enrollment some algorithms fail to convert a face image to a template. The proportion of failures is the failure-to-enroll rate, denoted by FTE. Similarly, some search images are not converted to templates. The corresponding proportion is termed failure-to-extract, denoted by FTX.

We do not report FTX because we assume that the same underlying algorithm is used for template generation for enrollment and search.

Failure to extract rates are incorporated into FNIR and FPIR measurements as follows.

- **Enrollment templates**: Any failed enrollment is regarded as producing a zero length template. Algorithms are required by the API [10] to transparently process zero length templates. The effect of template generation failure on search accuracy depends on whether subsequent searches are mated, or non-mated: Mated searches will fail giving elevated FNIR; non-mated searches will not produce false positives so, to first order, FPIR will be reduced by a factor of $1-\text{FTE}$.
- **Search templates and 1:N search**: In cases where the algorithm fails to produce a search template from input imagery, the result is taken to be a candidate list whose entries have no hypothesized identities and zero score. The effect of template generation failure on search accuracy depends on whether searches are mated, or non-mated: Mated searches will fail giving elevated FNIR; Non-mated searches will not produce false positives, so FPIR will be reduced. Thus given a measurement of false negative and positive rates made over only those where failures-to-extract did not occur, those rates - call them $\text{FNIR}^{\dagger}$ and $\text{FPIR}^{\dagger}$ - could be adjusted by an explicit measurement of FTX as follows

$$\text{FNIR} = \text{FTX} + (1 - \text{FTX})\text{FNIR}^{\dagger} \tag{8}$$

$$\text{FPIR} = (1 - \text{FTX})\text{FPIR}^{\dagger} \tag{9}$$

This approach is the correct treatment for positive-identification applications such as access control where cooperative users are enrolled and make attempts at recognition. This approach is not appropriate to negative identification applications, such as visa fraud detection, in which hostile individuals may attempt to evade detection by submitting poor quality samples. In those cases, template generation failures should be investigated as though a false alarm had occurred.

[10] For example, the Megaface benchmark. This is bad practice for several reasons: First, if a developer knows, or can reasonably assume, that a mate always exists, then unrealistic gaming of the test is possible. A second reason is that it does not put FPIR on equal footing with FNIR and that matters because in most applications, not all searches have mates - not everyone has been previously enrolled in a driving license issuance or a criminal justice system - so addressing between-class separation becomes necessary.

3.6 Fixed length candidate lists, threshold independent workload

Suppose an automated face identification algorithm returns L candidates, and a human reviewer is retained to examine up to R candidates, where $R \leq L$ might be set by policy, preference or labor availability. For now, assume also that the reviewer is not provided with, or ignores, similarity scores, and thresholds are not applied. Given the algorithm typically places mates at low (good) ranks, the number of candidates a reviewer can be expected to review can be derived as follows. Note that the reviewer will:

- ▷ Always inspect the first ranked image — Frac. reviewed = 1
- ▷ Then inspect those candidates where mate not confirmed at rank 1 — Frac. reviewed = 1-CMC(1)
- ▷ Then inspect those candidates where mate not confirmed at rank 1 or 2 — Frac. reviewed = 1-CMC(2)

etc. Thus if the reviewer will stop after a maximum of R candidates, the expected number of candidate reviews is

$$M(R) = 1 + (1 - CMC(1)) + (1 - CMC(2)) + \ldots + (1 - CMC(R-1)) \tag{10}$$

$$= R - \sum_{r=1}^{R-1} CMC(r) \tag{11}$$

A recognition algorithm that front-loads the cumulative match characteristic will offer reduced workload for the reviewer. This workload is defined only over the searches for which a mate exists. In the cases where there truly is no mate, the reviewer would review all R candidates. Thus, if the proportion of searches for which a mate does exist is β, which in the law enforcement context would be the recidivism rate [3], the full expression for workload becomes:

$$M(R) = \beta \left(R - \sum_{r=1}^{R-1} CMC(r) \right) + (1 - \beta)R \tag{12}$$

$$= R - \beta \sum_{r=1}^{R-1} CMC(r) \tag{13}$$

3.7 Timing measurement

Algorithms were submitted to NIST as implementations of the application programming interface(API) specified by NIST in the Evaluation Plan [10]. The API includes functions for initialization, template generation, finalization, search, gallery insert, and gallery delete. Two template generation functions are required, one for the preparation of an enrollment template, and one for a search template.

In NIST's test harness, all functions were wrapped by calls to the C++ std::chrono::high resolution clock which on the dedicated timing machine counts 1ns clock ticks. Precision is somewhat worse than that however.

3.8 Uncertainty estimation

3.8.1 Random error

This study leverages operational datasets for measurement of recognition error rates. This affords several advantages. First, large numbers of searches are conducted (see Table 5) giving precision to the measurements. Moreover, for the two mugshot datasets, these do not involve reuse of individuals so binomial statistics can be expected to apply to recognition error counts. In that case, an observed count of a particular recognition outcome (i.e. a false negative or false positive) in M trials will sustain 95% confidence that the actual error rate is no larger than some value.

As an example, the minimum number of mugshot searches conducted in this report is M =154 549, and for an observed FNIR around 0.002, the measurement supports a conclusion that the actual FNIR is no higher than 0.00228 at 99% confidence level. On the false positive side, we tabulate FNIR at FPIR values as low as 0.001. Given estimates based on 331 254 non-mate trials, the actual FPIR values will be below 0.00115 at 99% confidence. In conclusion, large scale evaluation, without reuse of subjects, supports tight uncertainty bounds on the measured error rates.

3.8.2 Systematic error

The FRVT 2018 dataset includes anomalies discovered as a result of inspecting images involved in recognition failures from the most accurate algorithms. Two kinds of failure occur: False negatives (which, for the purpose here, include failures to make templates) and false positives.

False negative errors: We reviewed 600 false negative pairs for which either or both of the leading two algorithms did not put the correct mate in the top 50 candidates. Given 154 549 searches, this number represents 0.39% of the total, resulting in FNIR $\approx$ 0.0039. Of the 600 pairs:

- **A: Poor quality**: About 20% of the pairs included images of very low quality, often greyscale, low resolution, blurred, low contrast, partially cropped, interlaced, or noisy scans of paper images. Additionally, in a few cases, the face is injured or occluded by bandages or heavy cosmetics.
- **B: Ground truth identity label bugs**: About 15% of the pairs are not actually mated. We only assigned this outcome when a pair is clearly not mated.
- **C: Profile views**: About 35% included an image of a profile (side) view of the face, or, more rarely, an image that was rotated 90 degrees in-plane (roll).
- **D: Tattoos**: About 30% included an image of a tattoo that contained a face image. These arise from mis-labelling in the parent dataset metadata.
- **E: Ageing**: There is considerable time-lapse between the two captures.

All these estimates are approximate. Of these, the tattoo and mislabled images can never be matched. These constitute an accuracy floor in the sample implying that FNIR cannot be below 0.0018[11]. The profile-views, low-quality images, and images with considerable ageing can, in principle, be successfully matched - indeed some algorithms do so - so are not part of the accuracy floor.

[11] This value is the sum of two partial false negative rates: $FNIR_{\square} = 0.15 * 0.0039$ plus $FNIR_{\square} = 0.3 * 0.0039$

For the microsoft-4 algorithm the lowest miss rate from (recent entry in Table 16) is FNIR(640 000, 50, 0) = 0.0018. This is close to the value estimated from the inspection of misses. It is below the 0.0039 figure because the algorithm does match some profile and poor quality images, that the yitu-2 algorithm does not.

For many tables (e.g. Table 16), the FNIR values obtained for the FRVT-2018 mugshots could be corrected by reducing them by 0.0018. The best values would then be indistinct from zero. The results in this report *were not* adjusted to account for this systematic error.

False positive errors: As depicted in Figure 9 many of the DET characteristics in this report exhibit a pronounced turn upward at low false positive rates. The shape can be caused by identity labelling errors in the ground truth of a dataset, specifically persons present in the database under two IDs such that some proportion of non-mate pairs are actually mated. We merged the highest 1000 non-mate pairs produced by three different algorithms which resulted in 1839 unique pairs. This constitutes 0.56% of all non-mate searches. We assert that it is *very* difficult for human reviewers to assign the pairs into the following three categories: twins; doppelgangers; or ground-truth errors (instances of the same person under two IDs). Given this difficulty we made no attempt to correct any ground truth except by removing 57 pairs in the following categories:

- **A: Profile views**: Thirteen pairs included one or two profile-view images. As described in Figure 102, these can cause false positives.
- **B: Same-session photographs**: For twelve pairs, the images were identical or trivially altered (e.g. cropped) versions of the same photo. These were present under a different ID likely due to some clerical or procedural mistake.
- **C: Tattoos of faces**: There were fourteen instances of tattoo photographs that contained faces causing false matches.
- **D: T-shirt faces**: There were six instances of T-shirt photographs (of Bob Marley and Che Guevara) being detected instead of the face and causing false positives.
- **E: Background faces**: There were twelve instances of one subject appearing in the background of two otherwise correct portrait photos.

Note we did not remove any images where there was a chance that the pair was actually a different person.

In any case, the results in this report have not been adjusted for this systematic error.

4 Results

This section gives extensive results for algorithms submitted to FRVT 2018. Three page "report cards" for each algorithm are contained in a separate supplement. Performance metrics were described in section 3. The main results are summarized in tabular form with more exhaustive data included as DET, CMC and related graphs in appendices as follows:

- ▷ The three tables 6-8 list algorithms alongside full developer names, acceptance date, size of the provided configuration data, template size and generation time, and search duration data.
  - The **template generation duration** is most important to applications that require fast response. For example, an eGate taking more than two seconds to produce a template might be unacceptable. Note that GPUs may be of utility in expediting this operation for some algorithms, though at additional expense. Two additional factors should be considered[12][13].
  - The **search duration** is the time taken for a search of a search template into a gallery of N enrollment templates. This performance variable, together with the volume of searches, is influential on the amount of hardware needed to sustain an operational deployment. This is measured here with the algorithm running on a single core of a contemporary CPU. Search is most simply implemented as N computations of a distance metric followed by a sort operation to find the closest enrollments. However, considerable optimization of this process is possible, up to and including fast-search algorithms that, by various means, avoid computation of all N distances.
  - The **template size** is the size of the extracted feature vector (or vectors) and any needed header information. Large template sizes may be influential on bus or network bandwidth, storage requirements, and on search duration. While the template itself is an opaque data blob, the feature dimensionality might be estimated by assuming a four-bytes-per-float encoding. There is a wide range of encodings. For the more accurate algorithm, sizes range from 256 bytes to about 2KB bytes, indicating essentially no consensus on face modeling and template design.
  - The **template size multiplier** column shows how, given k input images, the size of the template grows. Most implementations internally extract features from each image and concatenate them, and implement some score-level fusion logic during search. Other implementations, including many of the most accurate algorithms, produce templates whose size does not grow with k. This could be achieved via selection of the best quality image - but this is not optimal in handling ageing where the oldest image could be the best quality. Another mechanism would be feature-level fusion where information is fused from all k inputs. In any case, as a black-box test, the fusion scheme is proprietary and unknown.
  - The size of the **configuration data** is the total size of all files resident in a vendor-provided directory that contains arbitrary read-only files such as parameters, recognition models (e.g caffe). Generally a large value for this quantity may prohibit the use of the algorithm on a resource-constrained device.

[12] The FRVT 2018 API prohibited threading, so some gains from parallelism may be available on multiple-cores or multiple processors, if the feature extraction code could be distributed across them.

[13] Note also that factors of two or more may be realizable by exploiting modern vector processing instructions on CPUs. It is not clear in our measurements whether all developers exploited Intel's AVX2 instructions, for example. Our machine was so equipped, but we insisted that the same compiled library should also run on older machines lacking that instruction. The more sophisticated implementations may have detected AVX2 presence and branched accordingly. The less sophisticated may be defaulted to the reduced instruction set. Readers should see the FRVT 2018 API document for the specific chip details.

- Tables 16-17 report core rank-based accuracy for mugshot images. The population size is limited to N = 1.6 million identities because this is the largest gallery size on which all algorithms were executed. Notable observations from these tables are as follows:

  - **Accuracy gains during 2018**: NIST Interagency Report 8238 documented massive gains over those reported in the FRVT 2014 report, NIST Interagency Report 8009.

    Further gains are documented in this report. Comparing the most accurate algorithm in June 2018, Microsoft-4, with the most accurate in November 2018, NEC-2, the value of FNIR(N, 1, 0) reduced from 0.0031 to 0.0028 with N = 1.6 million recent images. For lifetime enrollments, Microsoft-4 remained the most accurate algorithm as the newer variants from Microsoft did not reduce this error rate.

    We further note that the revolution is not over: Figure 19 shows that many developers have made great advances in the four months between Phases 1 and 2 of FRVT 2018, Feburary to June. Most developers saw a two-fold reduction in errors, with Neurotechnology seeing a five fold reduction.

  - **Wide range in accuracy**: The rank-1 miss rates vary from FNIR(N, 1, 0) = 0.001 for nec-3 up to about 0.5 for the very fast but inaccurate microfocus-x algorithms. Among the developers who are superior to NEC in 2013, the range is from 0.002 to 0.035 for camvi-3. This large accuracy range is consistent with the buyer-beware maxim, and indicates that face recognition software is far from being commoditized.

- Tables 19-20 report threshold-based error rates, FNIR(N, L, T), for N = 1.6 million for mugshot-mugshot accuracy on FRVT 2014, FRVT 2018, and also (in pink) mugshot-webcam accuracy using FRVT 2018 enrollments. Notable observations from these tables are as follows:

  - **Order of magnitude accuracy gains since 2014**: As with rank-based results, the gains in accuracy are substantial, though somewhat reduced. At FPIR = 0.01, the best improvement over NEC in 2014 is a nine-fold reduction in FNIR using the Microsoft_4 algorithm. At FPIR = 0.001, the largest gain is a six-fold reduction in FNIR via the Yitu_2 algorithm.

  - **Broad gains across the industry**: About 19 companies realize accuracy better than the NEC benchmark from 2014. This is somewhat lower than the 28 developers who succeeded on the rank-1 metric. This may be due to the ubiquity of, and emphasis on, the rank-1 metric in many published algorithm development papers.

  - **Webcam images**: Searches of webcam images give FNIR(N, T) values around 2 to 3 times higher than mugshot searches. Notably the leading developers with mugshots are approximately the same with poorer quality webcams. But some developers e.g. Camvi, Megvii, TongYi, and Neurotechnology do improve their relative rankings on webcams, perhaps indicating their algorithms were tailored to less constrained images.

- Tables 10, 12, 13 and show, respectively, high-threshold, rank 1, nd rank 50 FNIR values for all algorithms performing searches into five different gallery sizes, N = 640 000, N = 1 600 000, N = 3 000 000, N = 6 000 000 and 12 000 000. The FPIR = 0.001 table is included to inform high-volume duplicate detection applications. The Rank-1 table is included as a primary accuracy indicator. The Rank-50 table is included to inform agencies who routinely produce 50 candidates for human-review. The notable results are:

  - **Slow growth in rank-based miss rates:** FNIR(N, R) generally grows as a power law, aN^b. From the straight lines of many graphs of Figure 22 this is clearly a reasonable model for most, but not all, algorithms. The coefficient a can be interpreted as FNIR in a gallery of size 1. The more important coefficient b indicates

scalability, and often, $b \ll 1$, implies very benign growth in FNIR. The coefficients of the models appear in the Tables 12 and 13.

- **Slow growth in threshold-based miss rates:** FNIR(N, T) also generally grows as a power law, aN^b except at the high threshold values corresponding to low FPIR values. This is visible in the plots of Figure 38 which show straight lines except for FPIR = 0.001, which increase more rapidly with N above 3 000 000. Each trace in those figures shows FNIR(N, T) at fixed FPIR with both N and T varying. Thus at large N, it is usually necessary to elevate T to maintain fixed FPIR. This causes increased FNIR. Why that would no-longer obey a power-law is not known. However, if we expect large galleries to contain individuals with familial relations to the non-mate search images - in the most extreme case, twins - then suppression of false positives becomes more difficult. This is discussed in the Figures starting at Fig. 9

▷ Figure 21 shows false positives from twins against their enrolled siblings, broken out by type of twin: fraternal or identical. The Figure is based on the enrollment of 104 single images on one of a pair of twins, and then the search of 2354 second images. Note that the dataset is heavily skewed towards identical twins which is not representative of the true population. There is also a skew towards same sex fraternal twin pairs compared to different sex fraternal twin pairs again not representative of the true population.

The notable results are:

- For all algorithms tested, the 1087 mated searches (Twin A vs. Twin A) produce scores almost always above typical operational thresholds, with (not shown) matches at rank 1. The images are of good quality, so this is the result expected from the rest of this report.
- For the 1066 identical twin searches (AB), almost all produce the twin at rank 1, with a few producing the mate at further down the candidate lists rank and low score.
- For the 169 fraternal searches (AB) from same sex pairs, most algorithms give a large number of very high scores, implying false positives at all thresholds. However, there there are long tails containing lower scores that are correctly below threshold. In general, scores that are higher in this distribution are all rank 1 whereas the lower scores have much higher ranks.
- (Not shown) Of the 169, there are 24 fraternal searches (AB) involving different sex twins. Here most algorithms correctly report scores well below the lowest threshold, and usually not on the candidate list at all.

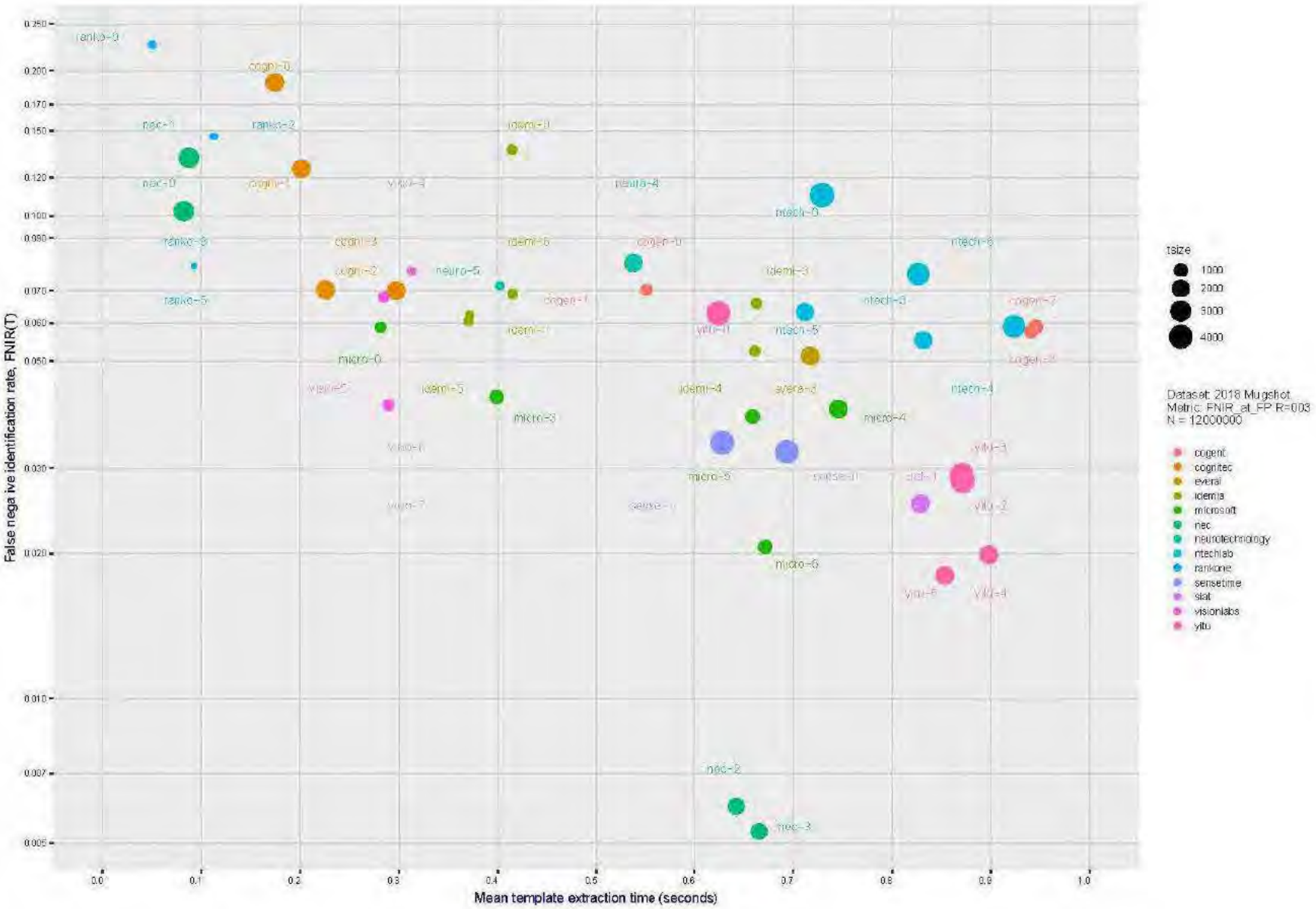


Figure 17: **[Mugshot Dataset] Speed-accuracy tradeoff.** *For developers of the more accurate algorithms the plot shows the tradeoff of high-threshold recognition miss-rates, FNIR(N, N, T) for FPIR(N, T) = 0.003, and template generation time. Developers are coded by color. Template size is encoded by the size of the circle. Some labels are quite distant from the respective point, to avoid superposing text. Without any other influences, the assumption would be that taking time to localize the face, and extract features, would lead to better accuracy. The most notable result, for NEC, is that their slower algorithms are much more accurate than the version that extract features in fewer than 90 milliseconds.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

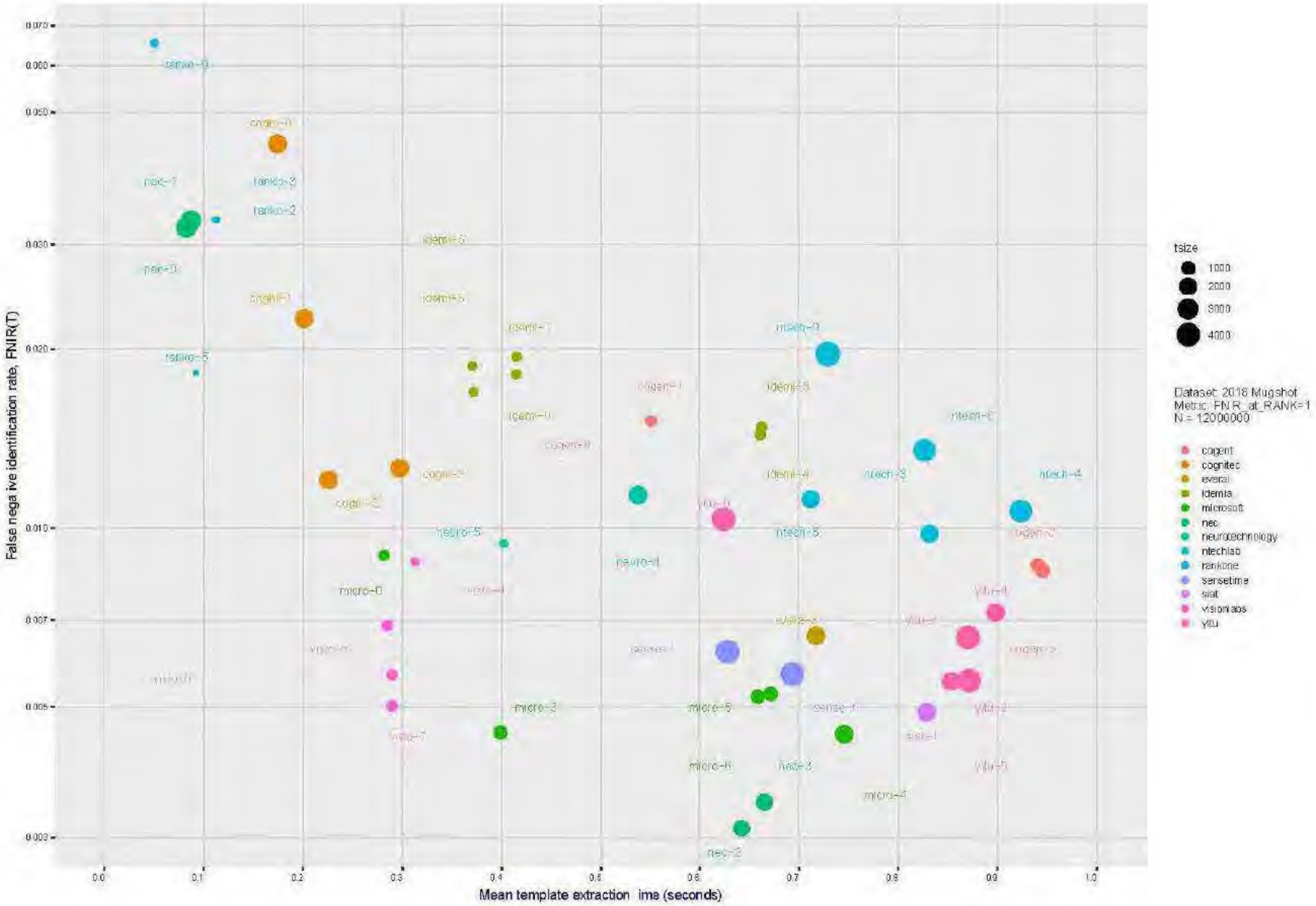


Figure 18: **[Mugshot Dataset] Speed-accuracy tradeoff.** *For developers of the more accurate algorithms the plot shows the tradeoff of rank-one recognition miss-rates, FNIR(N, 1, 0), and template generation time. Developers are coded by color. Template size is encoded by the size of the circle. Some labels are quite distant from the respective point, to avoid superposing text. Without any other influences, the assumption would be that taking time to localize the face, and extract features, would lead to better accuracy. This occurs for NEC with their slower algorithm being much accurate than the version that extract features in fewer than 90 milliseconds.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

	DEVELOPER FULL NAME	SHORT NAME	SEQ. NUM.	VALIDATION DATE	CONFIG[1] DATA (MB)	TEMPLATE GENERATION SIZE(B)	MULT[2]	TIME (MS)[3]	SEARCH DURATION[4] MILLISEC L=1 N=1.6M	L=50 N=1.6M	L=50 N=3M	L=50 N=6M	L=50 N=12M	POWER LAW (ms)
1	3Divi	3divi	0	2018-02-09	186	^{193}4096	k	^{30}426	–	^{107}553	–	–	–	
2	3Divi	3divi	1	2018-02-15	187	^{195}4224	k	^{34}428	–	^{19}87	–	–	–	
3	3Divi	3divi	2	2018-02-15	187	^{47}528	k	^{37}428	–	^{17}83	–	–	–	
4	3Divi	3divi	3	2018-06-19	165	^{31}512	k	^{133}625	^{12}576	^{13}576	–	–	–	
5	3Divi	3divi	4	2018-06-19	186	^{188}4096	k	^{123}628	^{73}604	^{126}801	–	–	–	
6	3Divi	3divi	5	2018-10-26	186	^{178}4096	k	^{139}653	^{67}537	^{104}837	^{51}1876	^{46}2612	^{41}5524	$^{71}0.07\ N^{1.1}$
7	3Divi	3divi	6	2018-10-26	187	^{49}528	k	^{141}653	^{10}33	^{15}33	–	–	–	
8	Alchera	alchera	0	2018-06-30	168	^{142}2048	k	^{42}263	^{128}3296	^{135}5420	–	–	–	
9	Alchera	alchera	1	2018-06-30	46	^{124}2048	k	^{9}66	^{121}3516	^{134}5489	–	–	–	
10	Alchera	alchera	2	2018-10-30	7	^{143}2048	k	^{16}115	^{118}2920	^{129}2926	–	–	–	
11	Alchera	alchera	3	2018-10-30	251	^{113}2048	k	^{137}548	^{126}2952	^{131}2956	^{79}6540	^{74}14998	^{70}35227	$^{34}0.10\ N^{1.2}$
12	Anke Investments	anke	0	2018-10-30	779	^{165}2072	k	^{36}451	^{24}675	^{120}748	^{53}1682	^{51}2965	^{43}6142	$^{52}0.21\ N^{1.1}$
13	Anke Investments	anke	1	2018-10-30	779	^{164}2072	k	^{37}433	^{176}700	^{124}769	–	–	–	
14	Aware	aware	0	2018-02-16	261	^{99}1564	k	^{128}668	–	^{62}251	–	–	–	
15	Aware	aware	1	2018-02-16	232	^{100}1564	k	^{138}661	–	^{61}251	–	–	–	
16	Aware	aware	2	2018-02-16	349	^{167}2076	k	^{159}912	–	^{62}252	–	–	–	
17	Aware	aware	3	2018-06-22	350	^{166}2076	k	^{162}716	^{124}2426	^{174}2508	^{73}4495	–	–	$^{41}1.09\ N^{1.0}$
18	Aware	aware	4	2018-06-22	349	^{5}52	k	^{160}712	^{89}1232	^{146}1187	–	–	–	
19	Aware	aware	5	2018-10-30	368	^{173}3100	k	^{158}827	^{18}94	^{46}97	^{13}202	^{11}370	^{9}761	$^{11}4.13\ N^{0.7}$
20	Aware	aware	6	2018-10-30	368	^{3}124	k	^{119}818	^{27}152	^{39}162	–	–	–	
21	Ayonix	ayonix	0	2018-06-21	57	^{57}1036	k	^{1}10	^{47}283	^{74}298	–	–	–	
22	Ayonix	ayonix	1	2018-10-29	74	^{58}1036	k	^{3}12	^{94}277	^{71}277	–	–	–	
23	Ayonix	ayonix	2	2018-10-30	74	^{59}1036	1	^{5}11	^{43}277	^{69}274	^{29}521	^{26}1079	^{23}2268	$^{30}0.11\ N^{1.0}$
24	Camvi Technologies	camvitech	1	2018-02-16	94	^{69}1024	1	^{34}177	–	^{12}23	–	–	–	
25	Camvi Technologies	camvitech	2	2018-02-16	442	^{74}1024	1	^{172}774	–	^{11}20	–	–	–	
26	Camvi Technologies	camvitech	3	2018-06-30	233	^{72}1024	1	^{150}707	^{9}10	^{9}11	–	–	–	
27	Camvi Technologies	camvitech	4	2018-10-30	233	^{66}1024	1	^{168}718	^{17}33	^{24}32	^{8}38	^{8}40	^{4}49	$^{8}8.95\ N^{0.1}$
28	Camvi Technologies	camvitech	5	2018-10-30	257	^{67}1024	1	^{171}769	^{9}31	^{13}30	–	–	–	
29	Thales	cogent	0	2018-06-20	533	^{46}526	k	^{118}851	^{93}494	^{110}558	^{42}1047	^{41}2060	^{38}4141	$^{23}0.46\ N^{1.0}$
30	Thales	cogent	1	2018-06-20	533	^{45}526	k	^{115}852	^{84}498	^{108}556	^{43}1048	^{42}2082	^{39}4263	$^{26}0.89\ N^{1.0}$
31	Thales	cogent	2	2018-10-30	681	^{84}1043	k	^{203}987	^{105}2017	^{166}2144	^{73}4293	^{69}8472	^{66}16429	$^{37}1.08\ N^{1.0}$
32	Thales	cogent	3	2018-10-30	681	^{83}1043	k	^{202}960	^{88}1230	^{146}1311	^{62}2687	^{60}5398	^{58}10184	$^{32}0.62\ N^{1.0}$
33	Cognitec Systems GmbH	cognitec	0	2018-06-21	364	^{155}2052	k	^{23}176	^{101}1748	^{154}1780	^{86}3672	^{86}7093	^{83}15224	$^{35}0.57\ N^{1.0}$
34	Cognitec Systems GmbH	cognitec	1	2018-06-21	412	^{152}2052	k	^{18}202	^{108}1835	^{156}1805	^{71}3971	^{67}7494	^{64}16249	$^{36}0.49\ N^{1.1}$
35	Cognitec Systems GmbH	cognitec	2	2018-10-30	468	^{151}2052	k	^{24}227	^{99}1738	^{153}1763	^{67}3660	^{66}7279	^{59}13886	$^{40}0.83\ N^{1.0}$
36	Cognitec Systems GmbH	cognitec	3	2018-10-30	466	^{157}2052	k	^{51}297	^{98}1719	^{155}1791	^{66}3638	^{65}7277	^{61}14904	$^{37}0.66\ N^{1.0}$
37	Dahua Technology Co. Ltd	dahua	0	2018-10-29	276	^{131}2048	k	^{73}378	–	^{61}256	–	–	–	
38	Dahua Technology Co. Ltd	dahua	1	2018-10-29	276	^{115}2048	k	^{68}371	–	^{64}256	^{39}601	^{31}1139	^{30}3001	$^{28}0.02\ N^{1.2}$
39	Dermalog	dermalog	0	2018-02-16	0	^{5}128	1	^{85}344	–	^{88}404	–	–	–	
40	Dermalog	dermalog	1	2018-02-16	0	^{8}128	1	^{121}171	–	^{91}407	–	–	–	
41	Dermalog	dermalog	2	2018-02-16	0	^{13}256	k	^{64}344	–	^{119}640	–	–	–	
42	Dermalog	dermalog	3	2018-06-21	0	^{7}128	1	^{51}211	^{17}92	^{29}92	–	–	–	
43	Dermalog	dermalog	4	2018-06-21	0	^{4}128	1	^{169}208	^{16}91	^{25}93	–	–	–	
44	Dermalog	dermalog	5	2018-10-26	0	^{6}128	1	^{109}532	^{2}0	^{1}0	^{1}0	^{1}0	^{1}0	$^{6}6.21\ N^{0.2}$
45	Dermalog	dermalog	6	2018-10-26	0	^{85}256	1	^{108}514	^{25}141	^{35}143	^{15}267	^{16}527	^{13}1285	$^{53}0.05\ N^{1.0}$
46	Ever AI	everai	0	2018-06-21	142	^{141}2048	1	^{35}438	^{4}4	^{3}3	^{2}5	–	–	$^{2}4.41\ N^{0.6}$
47	Ever AI	everai	1	2018-06-21	200	^{151}2048	1	^{135}590	^{53}336	^{81}356	^{38}651	–	–	$^{74}0.03\ N^{1.1}$
48	Ever AI	everai	2	2018-10-30	224	^{152}2048	1	^{71}377	^{46}278	^{72}283	–	–	–	
49	Ever AI	everai	3	2018-10-30	438	^{115}2048	1	^{167}736	^{45}278	^{71}281	^{30}572	^{29}1146	^{24}2278	$^{48}0.12\ N^{1.0}$
50	Eyedea Recognition	eyedea	0	2018-02-16	644	^{125}4152	k	^{82}424	–	^{120}640	–	–	–	
51	Eyedea Recognition	eyedea	1	2018-02-16	287	^{82}1036	k	^{56}311	–	^{78}307	–	–	–	
52	Eyedea Recognition	eyedea	2	2018-02-16	287	^{78}1036	k	^{85}429	–	^{78}305	–	–	–	

Notes:

1. Configuration size does not capture static data present in libraries. Libraries are not counted because most implementations include common ancillary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2. This multiplier expresses the increase in template size when k images are passed to the template generation function.
3. All durations are measured on Intel(R)Xeon(R)CPU E5-2650 v4 @ 2.20GHz processors. Estimates are made by wrapping the API function call in calls to std::chrono::high_resolution_clock which on the machine in (3) counts 1ns clock ticks. Precision is somewhat worse than that however.
4. Search durations are measured as in the prior note. The power-law model in the final column mostly fits the empirical results in Figure 111. However in certain cases the model is not correct and should not be used numerically.

Table 6: Summary of algorithms and properties included in this report. The blue superscripts give ranking for the quantity in that column. Missing search durations, denoted by "-", are absent because those runs were not executed, usually because we did not run on the larger galleries. Caution: The power-law model is sometimes an incorrect model. It is included here only to show broad sublinear behavior, which is flagged in green. The models should not be used for prediction.

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

	DEVELOPER FULL NAME	SHORT NAME	SEQ. NUM.	VALIDATION DATE	CONFIG[2] DATA (MB)	TEMPLATE GENERATION SIZE (B)	MULT[2]	TIME (MS)[3]	SEARCH DURATION[4] MILLISEC L=1 N=1.6M	L=50 N=1.6M	L=50 N=3M	L=50 N=6M	L=50 N=12M	POWER LAW [μs]
53	Eyedea Recognition	eyedea	3	2018-06-18	284	^{101}1036	k	^{78}385	^{46}309	^{76}311	–	–	–	
54	Glory Ltd	glory	0	2018-06-30	0	^{35}418	k	^{38}160	^{10}576	^{113}575	–	–	–	
55	Glory Ltd	glory	1	2018-06-30	0	^{103}1726	k	^{81}405	^{104}1924	^{120}1978	–	–	–	
56	Gorilla Technology	gorilla	0	2018-02-01	95	^{203}8300	k	^{91}427	–	^{90}10426	–	–	–	
57	Gorilla Technology	gorilla	1	2018-06-19	91	^{190}2156	k	^{21}169	^{138}5254	^{139}5156	–	–	–	
58	Gorilla Technology	gorilla	2	2018-10-29	91	^{87}1132	k	^{62}341	^{30}145	^{37}146	^{19}293	^{17}612	^{17}1509	^{68}0.02 $N^{1.1}$
59	Gorilla Technology	gorilla	3	2018-10-26	94	^{189}2156	k	^{125}563	^{105}1924	^{161}2047	–	–	–	
60	Loginface Corp	hhimec	0	2018-02-01	88	^{93}520	–	^{42}265	–	^{78}419	–	–	–	
61	Hikvision Research Institute	hikvision	0	2018-02-12	378	^{107}1808	1	^{124}875	–	^{171}2360	–	–	–	
62	Hikvision Research Institute	hikvision	1	2018-02-12	378	^{107}1808	1	^{178}820	–	^{172}2403	–	–	–	
63	Hikvision Research Institute	hikvision	2	2018-02-12	378	^{108}1808	1	^{186}820	–	^{170}2408	–	–	–	
64	Hikvision Research Institute	hikvision	3	2018-06-30	403	^{91}1408	1	^{128}633	^{34}904	^{131}1108	^{41}2377	^{52}3785	^{48}7570	^{31}0.91 $N^{1.0}$
65	Hikvision Research Institute	hikvision	4	2018-06-30	334	^{86}1152	1	^{116}510	^{77}784	^{134}1024	^{86}2094	^{82}3254	^{86}7117	^{73}0.86 $N^{1.0}$
66	Hikvision Research Institute	hikvision	5	2018-10-29	528	^{90}1408	–	^{129}619	^{82}883	^{142}995	^{85}1908	^{81}3792	^{82}9387	^{72}0.10 $N^{1.1}$
67	Hikvision Research Institute	hikvision	6	2018-10-29	528	^{85}1408	1	^{128}610	^{82}871	^{138}877	–	–	–	
68	Idemia	idemia	0	2018-02-16	371	^{31}384	1	^{86}416	–	^{28}133	^{16}249	^{12}502	–	^{30}0.08 $N^{1.0}$
69	Idemia	idemia	1	2018-02-16	371	^{31}384	1	^{86}417	–	^{28}138	–	–	–	
70	Idemia	idemia	2	2018-02-16	371	^{31}384	1	^{86}417	–	^{24}138	–	–	–	
71	Idemia	idemia	3	2018-06-21	472	^{46}528	1	^{149}689	^{50}318	^{82}361	^{34}651	^{33}1104	^{34}2332	^{13}5.08 $N^{0.8}$
72	Idemia	idemia	4	2018-06-21	472	^{50}528	1	^{148}689	^{25}168	^{53}211	^{35}478	^{33}995	^{27}2225	^{12}0.02 $N^{1.1}$
73	Idemia	idemia	5	2018-10-29	417	^{28}352	1	^{70}374	^{20}137	^{22}138	^{23}437	^{19}724	^{19}1630	^{21}0.01 $N^{1.2}$
74	Idemia	idemia	6	2018-10-29	417	^{28}352	1	^{68}373	^{21}137	^{20}138	^{24}442	^{22}827	^{20}1646	^{18}0.01 $N^{1.2}$
75	Imagus Technology Pty Ltd	imagus	0	2018-02-14	35	^{57}512	k	^{8}43	–	^{40}202	–	–	–	
76	Imagus Technology Pty Ltd	imagus	2	2018-06-21	35	^{59}512	k	^{9}76	^{36}200	^{37}208	–	–	–	
77	Imagus Technology Pty Ltd	imagus	3	2018-06-21	46	^{59}512	k	^{9}67	^{36}201	^{36}206	–	–	–	
78	Incode Technologies	incode	0	2018-06-29	23	^{76}1024	k	^{9}190	^{76}1293	^{126}3610	–	–	–	
79	Incode Technologies	incode	1	2018-06-29	151	^{146}2048	k	^{154}690	^{94}1542	^{138}4437	–	–	–	
80	Incode Technologies	incode	2	2018-10-29	71	^{142}2048	1	^{45}491	^{99}411	^{89}404	–	–	–	
81	Incode Technologies	incode	3	2018-10-29	193	^{142}2048	1	^{139}704	^{88}408	^{84}412	^{38}946	^{38}1606	^{40}4482	^{69}0.05 $N^{1.2}$
82	Innovatrics	innovatrics	0	2018-02-16	0	^{65}530	k	^{201}456	–	^{113}625	–	–	–	
83	Innovatrics	innovatrics	1	2018-02-16	0	^{91}530	k	^{35}316	–	^{117}625	–	–	–	
84	Innovatrics	innovatrics	2	2018-06-21	0	^{84}550	k	^{42}255	^{3}1	^{3}2	–	–	–	
85	Innovatrics	innovatrics	3	2018-06-21	0	^{94}560	k	^{41}255	^{108}2020	^{137}1882	–	–	–	
86	Innovatrics	innovatrics	4	2018-10-30	0	^{66}1076	k	^{38}406	^{9}8	^{8}8	^{11}9	^{9}9	^{13}9	^{7}68.38 $N^{0.8}$
87	Alivia / Innovation Sys.	isystems	0	2018-02-14	262	^{139}2048	1	^{36}322	–	^{56}393	–	–	–	
88	Alivia / Innovation Sys.	isystems	1	2018-02-14	263	^{65}1024	1	^{32}222	–	^{35}240	–	–	–	
89	Alivia / Innovation Sys.	isystems	2	2018-06-25	268	^{139}2048	1	^{39}316	^{50}385	^{59}484	^{50}1273	^{37}1770	^{37}3063	^{19}0.68 $N^{0.9}$
90	Alivia / Innovation Sys.	isystems	3	2018-10-30	350	^{142}2048	1	^{139}856	^{54}384	^{64}387	^{42}976	^{40}1817	^{41}3819	^{38}0.00 $N^{1.9}$
91	Leokmen Electroplast Industries	leokmen	3	2018-10-28	203	^{26}292	1	^{46}342	^{77}739	^{112}745	^{52}1394	^{49}2817	^{46}8286	^{46}0.13 $N^{1.1}$
92	Leokmen Electroplast Industries	leokmen	4	2018-10-28	184	^{57}592	1	^{60}345	^{58}981	^{133}998	–	–	–	
93	Megvii	megvii	0	2018-02-15	1327	^{138}2048	1	^{171}794	–	^{78}294	^{36}530	^{34}1060	–	^{32}0.18 $N^{1.0}$
94	Megvii	megvii	1	2018-10-28	1703	^{188}4096	1	^{137}652	^{58}551	^{111}560	^{29}1219	^{45}2316	^{42}5956	^{66}0.08 $N^{1.1}$
95	Megvii	megvii	2	2018-10-28	1735	^{188}4096	1	^{142}656	^{60}552	^{108}557	–	–	–	
96	Microfocus	microfocus	0	2018-02-12	101	^{21}256	k	^{138}625	–	^{45}184	–	–	–	
97	MicroFocus	microfocus	1	2018-02-16	101	^{18}256	k	^{138}527	–	^{30}29	–	–	–	
98	Microfocus	microfocus	2	2018-02-16	101	^{22}256	k	^{138}529	–	^{4}2	–	–	–	
99	Microfocus	microfocus	3	2018-06-22	101	^{14}256	k	^{40}269	^{32}185	^{45}188	–	–	–	
100	Microfocus	microfocus	4	2018-06-22	102	^{21}256	k	^{45}270	^{34}186	^{46}189	–	–	–	
101	Microfocus	microfocus	5	2018-10-29	94	^{25}256	k	^{45}266	^{31}182	^{44}186	^{30}353	^{18}706	^{18}1422	^{34}0.11 $N^{1.0}$
102	Microfocus	microfocus	6	2018-10-29	94	^{15}256	k	^{41}265	^{33}184	^{43}186	–	–	–	
103	Microsoft	microsoft	0	2018-01-30	126	^{48}612	1	^{46}283	–	^{114}593	^{47}1195	^{46}2396	^{38}4936	^{51}0.22 $N^{1.0}$
104	Microsoft	microsoft	1	2018-02-12	165	^{68}1044	1	^{66}349	–	^{133}869	–	–	–	

Notes

1. Configuration size does not capture static data present in libraries. Libraries are not counted because most implementations include common, ancillary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2. This multiplier expresses the increase in template size when 4 images are passed to the template generation function.
3. All durations are measured on Intel(R)Xeon(R)CPU E5-2630 v4 @ 2.20GHz processors. Estimates are made by wrapping the API function call in calls to std::chrono::high_resolution_clock which on the machine in (3) counts 1ns clock ticks. Precision is somewhat worse than that however.
4. Search durations are measured as in the prior note. The power-law model in the final column mostly fits the empirical results in Figure 111. However in certain cases the model is not correct and should not be used numerically.

Table 7: Summary of algorithms and properties included in this report. The blue superscripts give ranking for the quantity in that column. Missing search durations, denoted by "-", are absent because those runs were not executed, usually because we did not run on the larger galleries. Caution: The power-law model is sometimes an incorrect model. It is included here only to show broad sublinear behavior, which is flagged in green. The models should not be used for prediction.

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

	DEVELOPER FULL NAME	SHORT NAME	SEQ. NUM.	VALIDATION DATE	CONFIG[1] DATA (MB)	TEMPLATE GENERATION SIZE (B)	MULT[2]	TIME (MS)[3]	SEARCH DURATION[4] MILLISEC L=1 N=1.6M	L=50 N=1.6M	L=50 N=3M	L=50 N=6M	L=50 N=12M	POWER LAW [ms]
105	Microsoft	microsoft	2	2018-02-12	228	^{73}1024	1	^{128}555	-	^{129}869	-	-	-	
106	Microsoft	microsoft	2	2018-06-20	230	^{68}1024	1	^{80}404	^{96}1628	^{148}1603	^{65}3260	^{69}6730	^{58}13353	$^{56}0.51\ N^{1.1}$
107	Microsoft	microsoft	4	2018-06-20	457	^{127}2048	1	^{121}772	^{119}2664	^{127}2691	^{76}6260	^{71}11070	^{67}22748	$^{57}0.83\ N^{1.1}$
108	Microsoft	microsoft	5	2018-10-29	381	^{93}1024	1	^{145}673	^{95}1604	^{150}1671	^{84}3073	^{81}6296	^{57}13147	$^{58}0.79\ N^{1.0}$
109	Microsoft	microsoft	6	2018-10-29	478	^{62}1024	1	^{153}695	^{97}1640	^{145}1617	^{85}3707	^{62}6394	^{56}12879	$^{47}0.68\ N^{1.0}$
110	NEC	nec	0	2018-06-21	181	^{122}2592	k	^{111}82	^{42}817	^{56}426	^{37}788	^{35}1315	^{27}2737	$^{34}0.73\ N^{0.9}$
111	NEC	nec	1	2018-06-29	131	^{121}2592	k	^{71}68	^{36}193	^{53}208	^{33}388	^{33}750	^{28}1577	$^{32}0.21\ N^{1.0}$
112	NEC	nec	2	2018-10-30	705	^{101}1616	k	^{140}652	^{87}405	^{93}409	^{66}1072	^{59}1755	^{54}4235	$^{50}0.06\ N^{1.1}$
113	NEC	nec	3	2018-10-30	774	^{102}1712	k	^{150}690	^{5}7	^{7}7	^{5}14	^{5}40	^{5}82	$^{30}0.00\ N^{1.2}$
114	Neurotechnology	neurotech	0	2018-02-16	331	^{139}5214	k	^{154}702	-	^{152}3040	-	-	-	
115	Neurotechnology	neurotech	1	2018-02-16	331	^{138}5214	k	^{142}661	-	^{154}3054	-	-	-	
116	Neurotechnology	neurotech	2	2018-02-16	331	^{139}5214	k	^{143}658	-	^{153}3051	-	-	-	
117	Neurotechnology	neurotech	3	2018-06-27	265	^{116}2048	k	^{116}547	^{67}1084	^{135}1059	^{80}2111	^{65}4779	^{49}8793	$^{51}0.73\ N^{1.0}$
118	Neurotechnology	neurotech	4	2018-06-27	265	^{145}2048	k	^{115}543	^{58}1060	^{136}1051	^{77}2091	^{66}4263	^{47}8736	$^{37}1.22\ N^{1.0}$
119	Neurotechnology	neurotech	5	2018-10-30	266	^{17}256	k	^{84}412	^{80}825	^{127}839	^{84}1690	^{81}3219	^{80}8955	$^{62}0.19\ N^{1.1}$
120	Neurotechnology	neurotech	6	2018-10-30	564	^{18}256	k	^{165}746	^{81}839	^{128}842	-	-	-	
121	Newland Computer Co. Ltd	newland	2	2018-10-30	96	^{110}2048	-	^{131}868	^{134}8653	^{139}8765	^{86}17713	^{81}38963	-	$^{59}1.32\ N^{1.1}$
122	Noblis	noblis	1	2018-10-30	114	^{128}2048	1	^{80}211	^{91}1278	^{143}1272	-	-	-	
123	Noblis	noblis	2	2018-10-30	152	^{200}6144	1	^{110}536	^{116}2515	^{123}2522	^{96}5649	^{73}12422	^{73}44262	$^{85}0.04\ N^{1.3}$
124	N-Tech Lab	ntech	0	2018-02-16	2124	^{156}4442	k	^{165}780	-	^{85}382	^{56}673	^{54}1344	-	$^{52}0.27\ N^{1.0}$
125	N-Tech Lab	ntech	1	2018-02-16	851	^{104}1756	k	^{82}405	-	^{86}161	-	-	-	
126	N-Tech Lab	ntech	3	2018-06-21	3664	^{124}3484	k	^{164}831	^{52}384	^{80}326	^{53}596	^{50}1192	^{43}2411	$^{38}0.24\ N^{1.0}$
127	N-Tech Lab	ntech	4	2018-06-21	3766	^{125}3484	k	^{193}929	^{52}379	^{78}312	^{52}597	^{54}1204	^{56}2416	$^{39}0.21\ N^{1.0}$
128	N-Tech Lab	ntech	5	2018-10-30	1685	^{105}1940	k	^{164}717	^{42}243	^{57}246	^{36}538	^{36}1100	^{38}2867	$^{36}0.02\ N^{1.1}$
129	N-Tech Lab	ntech	6	2018-10-30	1686	^{109}1940	k	^{187}841	^{41}243	^{56}246	^{37}546	^{36}1104	^{39}2878	$^{37}0.02\ N^{1.1}$
130	Quantasoft	quantasoft	1	2018-10-30	276	^{137}2048	k	^{96}396	^{105}15422	^{105}14858	^{85}14717	-	^{60}18523	
131	Rank One Computing	rankone	0	2018-02-07	0	^{12}228	k	^{9}50	-	^{22}76	^{12}142	^{12}220	^{10}502	$^{18}0.12\ N^{0.9}$
132	Rank One Computing	rankone	1	2018-02-15	0	^{25}324	k	^{17}136	-	^{42}169	-	-	-	
133	Rank One Computing	rankone	2	2018-06-19	0	^{10}133	k	^{14}113	^{32}138	^{28}137	^{16}258	^{14}517	^{12}1029	$^{25}0.10\ N^{1.0}$
134	Rank One Computing	rankone	3	2018-06-19	0	^{11}133	k	^{15}114	^{23}138	^{30}137	^{15}258	^{13}515	^{11}1027	$^{21}0.09\ N^{1.0}$
135	Rank One Computing	rankone	4	2018-10-09	0	^{1}85	k	^{4}36	^{19}101	^{27}101	^{12}190	-	-	$^{19}0.07\ N^{1.0}$
136	Rank One Computing	rankone	5	2018-10-24	0	^{5}133	k	^{12}84	^{24}140	^{30}144	^{17}266	^{18}525	^{13}1089	$^{23}0.11\ N^{1.0}$
137	RealNetworks	realnetworks	0	2018-06-21	98	^{128}4100	1	^{36}241	^{123}4257	^{128}2740	-	-	-	
138	RealNetworks	realnetworks	1	2018-06-21	105	^{129}4104	k	^{37}248	^{123}3568	^{124}2107	-	-	-	
139	RealNetworks	realnetworks	2	2018-10-30	105	^{127}4104	k	^{39}245	^{107}2008	^{100}2046	^{72}4190	^{70}8833	^{62}15020	$^{36}1.08\ N^{1.0}$
140	KanKan AI	remarkai	0	2018-10-30	187	^{123}2048	k	^{127}615	^{121}5685	^{128}5723	-	-	-	
141	KanKan AI	remarkai	1	2018-10-30	187	^{114}2048	k	^{98}434	^{120}5680	^{126}5781	^{84}12475	^{80}28726	^{76}59618	$^{51}0.27\ N^{1.2}$
142	Sensetime Group Ltd	sensetime	0	2018-10-30	525	^{126}4104	k	^{162}715	^{87}498	^{100}501	^{68}1212	^{63}2281	^{50}5032	$^{47}0.09\ N^{1.1}$
143	Sensetime Group Ltd	sensetime	1	2018-10-30	525	^{128}4104	k	^{143}656	^{96}516	^{103}502	^{85}1146	^{84}2301	^{87}4765	$^{63}0.09\ N^{1.1}$
144	Shaman Software	shaman	0	2018-02-12	0	^{124}4096	k	^{124}538	-	^{102}523	-	-	-	
145	Shaman Software	shaman	1	2018-02-12	0	^{128}4096	k	^{121}557	-	^{105}524	-	-	-	
146	Shaman Software	shaman	2	2018-02-12	0	^{28}8192	k	^{124}557	-	^{131}888	-	-	-	
147	Shaman Software	shaman	3	2018-06-30	0	^{125}2048	k	^{156}704	^{30}692	^{77}310	-	-	-	
148	Shaman Software	shaman	4	2018-06-30	0	^{156}2048	k	^{135}642	^{44}434	^{66}267	-	-	-	
149	Shaman Software	shaman	6	2018-10-26	0	^{132}2048	k	^{157}706	^{72}594	^{132}608	-	-	-	
150	Shaman Software	shaman	7	2018-10-26	0	^{123}2048	k	^{159}709	^{71}593	^{116}505	^{96}1169	^{45}2411	^{39}5007	$^{29}0.25\ N^{1.0}$
151	Shenzhen Inst. Adv. Tech. CAS	SIAT	0	2018-02-14	306	^{86}1096	k	^{67}359	-	^{142}1242	-	-	-	
152	Shenzhen Inst. Adv. Tech. CAS	SIAT	1	2018-06-30	521	^{147}2052	1	^{188}842	^{125}4512	^{136}4402	^{89}9103	^{76}18391	^{71}38725	$^{44}2.06\ N^{1.0}$
153	Shenzhen Inst. Adv. Tech. CAS	SIAT	2	2018-02-30	521	^{155}2052	1	^{195}906	^{122}5101	^{149}4884	^{91}9556	^{79}18834	^{81}39717	$^{45}2.08\ N^{1.0}$
154	Smilart	smilart	0	2018-02-15	106	^{66}1024	k	^{30}168	-	^{144}1285	-	-	-	
155	Smilart	smilart	1	2018-02-15	120	^{71}1024	k	^{145}862	-	^{139}1135	-	-	-	
156	Smilart	smilart	2	2018-02-15	109	^{67}1024	k	^{128}560	-	^{145}1302	-	-	-	

Notes

1. Configuration size does not capture static data present in libraries. Libraries are not counted because most implementations include common ancillary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2. This multiplier expresses the increase in template size when K images are passed to the template generation function.
3. All durations are measured on Intel(R)Xeon(R)CPU E5-2650 v4 @ 2.20GHz processors. Estimates are made by wrapping the API function call in calls to std::chrono::high_resolution_clock which on the machine in (3) counts 1ns clock ticks. Precision is somewhat worse than that however.
4. Search durations are measured as in the prior note. The power-law model in the final column mostly fits the empirical results in Figure 111. However in certain cases the model is not correct and should not be used numerically.

Table 8: Summary of algorithms and properties included in this report. The blue superscripts give ranking for the quantity in that column. Missing search durations, denoted by "-", are absent because those runs were not executed, usually because we did not run on the larger galleries. Caution: The power-law model is sometimes an incorrect model. It is included here only to show broad sublinear behavior, which is flagged in green. The models should not be used for prediction.

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

	DEVELOPER FULL NAME	SHORT NAME	SEQ. NUM.	VALIDATION DATE	CONFIG[1] DATA (MB)	TEMPLATE GENERATION SIZE (B)	MULT[2]	TIME (MS)[3]	SEARCH DURATION[4] MILLISEC L=1 N=1.6M	L=50 N=1.6M	L=50 N=3M	L=50 N=6M	L=50 N=12M	POWER LAW (μs)
157	Smilart	smilart	4	2018-10-30	65	^{36}512	k	^{79}187	^{138}15879	^{140}16382	-	-	-	
158	Smilart	smilart	5	2018-10-30	562	^{130}2048	k	^{103}464	-	-	-	-	-	
159	Synesis	synesis	0	2018-02-16	332	^{36}512	k	^{36}237	-	^{90}162	-	-	-	
160	Synesis	synesis	3	2018-10-30	237	^{107}4096	k	^{53}103	^{66}784	^{128}796	^{66}1928	^{59}3861	^{48}8748	$^{76}0.07\ N^{1.1}$
161	Tevian	tevian	0	2018-02-16	666	^{132}2048	1	^{78}394	-	^{90}405	-	-	-	
162	Tevian	tevian	1	2018-02-16	666	^{132}2048	1	^{80}398	-	^{87}403	-	-	-	
163	Tevian	tevian	2	2018-02-16	666	^{135}2048	1	^{77}387	-	^{86}402	-	-	-	
164	Tevian	tevian	3	2018-06-20	707	^{118}2048	1	^{54}300	^{62}473	^{105}589	-	-	-	
165	Tevian	tevian	4	2018-08-20	707	^{121}2048	1	^{53}299	^{60}434	^{105}587	-	-	-	
166	Tevian	tevian	5	2018-10-30	773	^{131}2048	1	^{58}416	^{50}405	^{97}407	^{87}852	^{76}1758	^{65}3872	$^{58}0.14\ N^{1.0}$
167	TigerIT Americas LLC	tigar	0	2018-08-29	333	^{102}2052	k	^{56}428	^{102}1822	^{128}2942	-	-	-	
168	TigerIT Americas LLC	tiger	1	2018-06-27	333	^{102}2052	k	^{76}398	^{1}0	^{5}1	-	-	-	
169	TigerIT Americas LLC	tiger	2	2018-10-29	416	^{118}2052	k	^{108}464	^{101}1814	^{130}1919	^{70}3829	^{62}7519	^{50}14805	$^{43}0.83\ N^{1.0}$
170	TigerIT Americas LLC	tiger	3	2018-10-30	416	^{134}2052	k	^{109}464	^{35}191	^{47}189	-	-	-	
171	Tong'Yi Transportation Technology	tongyi	0	2018-06-29	1701	^{142}2070	k	^{86}190	^{125}2256	^{140}2272	-	-	-	
172	Tong'Yi Transportation Technology	tongyi	1	2018-06-29	1701	^{148}2070	1	^{85}189	^{132}2238	^{122}2257	-	-	-	
173	Toshiba	toshiba	0	2018-10-30	961	^{36}1548	k	^{100}990	^{139}6147	^{150}6230	^{83}12209	^{69}25330	^{53}49658	$^{79}0.36\ N^{1.2}$
174	Toshiba	toshiba	1	2018-10-30	961	^{149}2060	k	^{103}931	^{132}6001	^{138}6049	-	-	-	
175	Visidon	visidon	0	2018-06-20	208	^{36}1028	k	^{92}337	^{108}2006	^{176}2566	-	-	-	
176	Visidon	visidon	1	2018-10-30	166	^{131}2052	k	^{158}695	^{134}4857	^{138}4458	^{90}8429	^{85}17210	^{69}36185	$^{58}2.40\ N^{1.0}$
177	Vigilant Solutions	vigilant	0	2018-02-08	335	^{36}1544	k	^{104}823	-	^{167}2058	-	-	-	
178	Vigilant Solutions	vigilant	1	2018-02-14	249	^{138}2056	k	^{105}739	-	^{168}2075	-	-	-	
179	Vigilant Solutions	vigilant	2	2018-02-14	335	^{36}1544	k	^{177}820	-	^{166}2121	-	-	-	
180	Vigilant Solutions	vigilant	3	2018-06-21	335	^{36}1544	k	^{108}832	^{135}2453	^{170}2807	-	-	-	
181	Vigilant Solutions	vigilant	4	2018-06-21	337	^{30}1544	k	^{102}820	^{110}2050	^{167}2851	-	-	-	
182	Vigilant Solutions	vigilant	5	2018-10-30	335	^{36}1544	k	^{170}778	-	^{158}1720	-	-	-	
183	Vigilant Solutions	vigilant	6	2018-10-30	337	^{36}1544	k	^{108}834	-	^{151}1713	-	-	-	
184	VisionLabs	visionlabs	3	2018-02-16	624	^{11}256	1	^{32}228	-	^{65}	^{35}	^{16}	-	$^{9}417.37\ N^{0.5}$
185	VisionLabs	visionlabs	4	2018-06-22	299	^{21}256	1	^{37}315	^{8}19	^{10}17	^{9}20	^{6}26	^{5}29	$^{5}2633.29\ N^{0.1}$
186	VisionLabs	visionlabs	5	2018-06-22	305	^{36}512	1	^{55}300	^{13}54	^{16}33	^{7}37	^{8}56	^{7}88	$^{10}106.84\ N^{0.4}$
187	VisionLabs	visionlabs	6	2018-10-30	360	^{40}512	1	^{50}292	^{12}36	^{18}36	^{9}39	^{7}44	^{8}58	$^{5}3210.98\ N^{0.2}$
188	VisionLabs	visionlabs	7	2018-10-30	360	^{42}512	1	^{51}295	^{14}63	^{21}63	^{10}72	^{9}80	^{8}115	$^{8}2076.82\ N^{0.3}$
189	Vocord	vocord	0	2018-02-16	872	^{36}608	k	^{111}536	-	^{87}268	-	-	-	
190	Vocord	vocord	1	2018-02-16	872	^{36}608	k	^{112}536	-	^{88}268	-	-	-	
191	Vocord	vocord	2	2018-02-16	924	^{176}2048	k	^{139}636	-	^{99}248	-	-	-	
192	Vocord	vocord	3	2018-06-30	627	^{97}896	k	^{161}714	^{77}215	^{88}247	-	-	-	
193	Vocord	vocord	4	2018-06-30	627	^{90}896	k	^{113}538	^{90}216	^{83}253	-	-	-	
194	Vocord	vocord	5	2018-10-30	1035	^{58}768	k	^{155}822	^{38}168	^{49}204	^{21}385	^{20}767	^{16}1466	$^{13}0.12\ N^{1.0}$
195	Vocord	vocord	6	2018-10-30	1035	^{385}10240	k	^{181}825	^{30}170	^{54}216	-	-	-	
196	Zhuhai Yisheng Electronics Tech.	yisheng	0	2018-02-14	473	^{160}2108	k	^{138}615		^{113}587	-	-	-	
197	Zhuhai Yisheng Electronics Tech.	yisheng	1	2018-06-19	474	^{156}3704	k	^{78}387	^{171}2228	^{139}1108	-	-	-	
198	Shanghai Yitu Technology	yitu	0	2018-02-12	1774	^{201}4136	1	^{132}633	-	^{98}464	^{60}868	^{38}1768	-	$^{59}0.12\ N^{1.1}$
199	Shanghai Yitu Technology	yitu	1	2018-02-12	1944	^{190}4136	1	^{189}930	-	^{97}463	-	-	-	
200	Shanghai Yitu Technology	yitu	2	2018-06-21	2077	^{199}4138	1	^{192}870	^{129}5916	^{126}5817	^{77}6101	^{73}12264	^{68}38047	$^{149}0.25\ N^{0.9}$
201	Shanghai Yitu Technology	yitu	3	2018-06-21	2077	^{192}4138	1	^{193}871	^{127}5248	^{101}5242	^{78}6286	^{68}19829	^{54}45621	$^{67}1.08\ N^{1.1}$
202	Shanghai Yitu Technology	yitu	4	2018-10-30	2119	^{182}2070	1	^{196}910	^{57}1288	^{142}1203	^{87}2440	^{57}5241	^{54}9671	$^{46}0.52\ N^{1.0}$
203	Shanghai Yitu Technology	yitu	5	2018-10-30	2049	^{182}2070	1	^{190}861	^{80}1235	^{141}1157	^{82}2508	^{58}6003	^{58}9601	$^{42}0.55\ N^{1.0}$

Notes:

1. Configuration size does not capture static data present in libraries. Libraries are not counted because most implementations include common ancillary libraries for image processing (e.g. openCV) or numerical computation (e.g. blas).
2. This multiplier expresses the increase in template size when K images are passed to the template generation function.
3. All durations are measured on Intel(R)Xeon(R)CPU E5-2630 v4 @ 2.20GHz processors. Estimates are made by wrapping the API function call in calls to std::chrono::high_resolution_clock which on the machine in (3) counts 1ns clock ticks. Precision is somewhat worse than that however.
4. Search durations are measured as in the prior note. The power-law model in the final column mostly fits the empirical results in Figure 111. However in certain cases the model is not correct and should not be used numerically.

Table 9: Summary of algorithms and properties included in this report. The blue superscripts give ranking for the quantity in that column. Missing search durations, denoted by "-", are absent because those runs were not executed, usually because we did not run on the larger galleries. Caution: The power-law model is sometimes an incorrect model. It is included here only to show broad sublinear behavior, which is flagged in green. The models should not be used for prediction.

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

MISSES BELOW THRESHOLD, T		ENROL LIFETIME					ENROL MOST RECENT				
FNIR(N, T>0, R>1)		DATASET: FRVT 2018					DATASET: FRVT 2018				
#	ALGORITHM	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M
1	3DIVI-3	139 0.3000	121 0.2499	95 0.3893	139 0.4344		146 0.3650	134 0.4023			
2	3DIVI-5	85 0.1045	96 0.1339				101 0.1382	103 0.1691	77 0.1939	98 0.2302	97 0.3087
3	ALCHERA-0	92 0.0852	88 0.1105	90 0.1361	86 0.1913		96 0.1128	93 0.1405			
4	ALCHERA-3	92 0.1018	93 0.1296				98 0.1205	100 0.1590	79 0.1891	95 0.2467	73 0.3628
5	ANKE-0	79 0.0769	77 0.0989				94 0.0968	95 0.1199	86 0.1432	86 0.1811	80 0.2624
6	AWARE-3	83 0.0844	79 0.0991	87 0.1148	83 0.1459		96 0.1122	79 0.1306	87 0.1471	82 0.1793	82 0.2395
7	AWARE-5	121 0.2628	112 0.2384				142 0.3459	129 0.3729	89 0.4094	77 0.4615	66 0.5637
8	AYONIX-0	171 0.8262	139 0.8490	97 0.8640	82 0.8809		168 0.7795	139 0.8114			
9	AYONIX-2	166 0.7602	132 0.8038				166 0.7867	132 0.8246	85 0.8511	81 0.8705	76 0.8946
10	CAMVI-3	35 0.0281	46 0.0509	39 0.0680	49 0.1871		41 0.0413	56 0.0736			
11	CAMVI-4	39 0.0257	49 0.0505				48 0.0395	57 0.0741	51 0.1008	50 0.2552	64 0.2751
12	COGENT-0	51 0.0387	45 0.0434	58 0.0523	70 0.0784	113 0.1559	52 0.0455	45 0.0557	46 0.0734	48 0.1194	46 0.2029
13	COGENT-1	58 0.0598	49 0.0518				51 0.0455	44 0.0557	40 0.0734	42 0.1184	50 0.2029
14	COGENT-2	17 0.0220	15 0.0299	35 0.0390	58 0.0703	106 0.1595	24 0.0356	30 0.0495	34 0.0655	51 0.1185	56 0.2241
15	COGENT-3	30 0.0258	27 0.0341	31 0.0450	76 0.0942	95 0.1864	27 0.0361	36 0.0515	42 0.0771	80 0.1374	86 0.2488
16	COGNITEC-0	91 0.0989	95 0.1256				102 0.1400	103 0.1628	73 0.1892	56 0.2205	36 0.2859
17	COGNITEC-1	57 0.0597	68 0.0777	71 0.0946	100 0.1315	136 0.2552	77 0.0832	77 0.1045	83 0.1244	88 0.1561	81 0.2398
18	COGNITEC-2	41 0.0296	38 0.0401	38 0.0523	62 0.0852	94 0.2298	45 0.0455	46 0.0560	38 0.0695	33 0.0980	36 0.1967
19	COGNITEC-3	36 0.0289	36 0.0397	47 0.0506	78 0.0937	93 0.2140	42 0.0427	43 0.0553	41 0.0695	31 0.0938	29 0.1840
20	DAHUA-1	54 0.0410	50 0.0521				62 0.0596	59 0.0768	47 0.0905	40 0.1179	33 0.1910
21	DERMALOG-4	132 0.3405	128 0.3892	64 0.4181	80 0.4533		139 0.4380	133 0.4813			
22	DERMALOG-5	80 0.0290	62 0.0649				74 0.0726	73 0.0909	58 0.1172	58 0.1618	59 0.2516
23	DERMALOG-6	36 0.0276	37 0.0393				42 0.0420	45 0.0542	38 0.0687	37 0.1004	28 0.1812
24	EVERAI-0	97 0.0900	66 0.0676				82 0.0681	75 0.0921	57 0.1223		
25	EVERAI-1	38 0.0255	30 0.0366				33 0.0383	37 0.0618	30 0.0696		
26	EVERAI-3	13 0.0191	15 0.0256	11 0.0338	8 0.0389		12 0.0282	17 0.0377	15 0.0473	18 0.0683	17 0.1652
27	SYSDRA-3	127 0.2911	122 0.3283	93 0.3673	88 0.4154		148 0.2496	131 0.3893			
28	GLORY-1	123 0.2160	110 0.2447	86 0.2618	93 0.2884		130 0.2790	133 0.3067			
29	GORILLA-2	100 0.1088	95 0.1369				106 0.1561	108 0.1902	75 0.2210	59 0.2625	71 0.3426
30	HIK-2	101 0.1104	99 0.1368	93 0.1610	90 0.2061	97 0.3067	92 0.0985	86 0.1212			
31	HIK-3	86 0.0886	80 0.1097				78 0.0858	78 0.1054	66 0.1228	54 0.1552	57 0.2500
32	HIK-4	80 0.0829	82 0.1061	80 0.1226	96 0.1618	89 0.2618	76 0.0821	74 0.1013	74 0.1173	73 0.1498	58 0.2502
33	HIK-5	28 0.0216	23 0.0308	25 0.0397	52 0.0661		32 0.0339	27 0.0467	28 0.0583	32 0.0967	44 0.2164
34	IDEMIA-0	75 0.0645	69 0.0802	67 0.0996	74 0.1237	86 0.1872	81 0.0920	82 0.1135	68 0.1332	59 0.1628	49 0.2208
35	IDEMIA-1	42 0.0304	39 0.0397	39 0.0465	48 0.0623	76 0.1578	47 0.0444	50 0.0540	39 0.0647	36 0.0856	32 0.1618
36	IDEMIA-2	58 0.0453	60 0.0564	53 0.0668	68 0.0896	80 0.1706	48 0.0449	52 0.0543			
37	IDEMIA-3	28 0.0238	21 0.0308				31 0.0273	31 0.0497	26 0.0927	19 0.2389	24 0.4442
38	IDEMIA-4	20 0.0223	13 0.0276	20 0.0338	19 0.0478	43 0.1556	19 0.0326	19 0.0399	19 0.0472	18 0.0644	38 0.1659
39	IDEMIA-5	23 0.0261	24 0.0319	19 0.0395	43 0.0588	79 0.1784	34 0.0385	29 0.0465	28 0.0562	25 0.0789	35 0.1961
40	IDEMIA-6	26 0.0258	26 0.0316	24 0.0383	34 0.0581	79 0.2046	38 0.0377	28 0.0458	33 0.0650	33 0.0760	47 0.2242
41	IMAGUS-2	164 0.6616	126 0.7143	96 0.7503	81 0.7867		152 0.7092	136 0.7810			
42	INCODE-1	107 0.1400	108 0.1796	84 0.2159	92 0.2741		133 0.1763	114 0.2145			
43	INCODE-3	26 0.0249	20 0.1227				110 0.1349	105 0.1703	73 0.1988	67 0.2378	68 0.3157
44	INNOVATRICS-4	56 0.0837	73 0.0928				95 0.1106	84 0.1340	68 0.1418	62 0.1418	31 0.1418
45	ISYSTEMS-0	61 0.0485	61 0.0633	76 0.0795	99 0.1057	130 0.2072	71 0.0709	72 0.0912			
46	ISYSTEMS-1	53 0.0480	60 0.0627	65 0.0784	96 0.1054	122 0.2081	69 0.0702	69 0.0903			
47	ISYSTEMS-2	55 0.0394	53 0.0545	54 0.0679			65 0.0612	62 0.0814	52 0.1006	51 0.1406	52 0.2374
48	ISYSTEMS-3	42 0.0301	41 0.0402	41 0.0557	62 0.0881	98 0.1992	58 0.0464	52 0.0620	45 0.0840	46 0.1324	51 0.2417
49	LOOKMAN-3	46 0.0335	42 0.0425				50 0.0372	48 0.0469	48 0.0641	41 0.0958	34 0.1650
50	MEGVII-0	81 0.0822	82 0.1023	85 0.1228	94 0.1489	96 0.2348	80 0.0895	80 0.1036	61 0.1287	56 0.1806	49 0.2298
51	MEGVII-1						86 0.0586	86 0.0746	62 0.0896	49 0.1238	68 0.2761
52	MICROFOCUS-3	177 0.9002	142 0.9213	99 0.9342			188 0.9119	157 0.9310			
53	MICROFOCUS-5	158 0.9679	145 0.9835				158 0.9753	154 0.9361	89 0.9563	85 0.9760	80 0.8969
54	MICROSOFT-0	16 0.0208	17 0.0292	17 0.0361	42 0.0536	60 0.1502	31 0.0329	37 0.0443	31 0.0544	28 0.0767	27 0.1783
55	MICROSOFT-1	17 0.0214	18 0.0299	18 0.0373	43 0.0542	65 0.1585	32 0.0339	38 0.0443			
56	MICROSOFT-2	28 0.0252	28 0.0345	28 0.0426	46 0.0600	72 0.1553	38 0.0387	34 0.0503			
57	MICROSOFT-3	14 0.0123	14 0.0193				18 0.0223	16 0.0304	16 0.0384	18 0.0570	30 0.1603
58	MICROSOFT-4	10 0.0128	11 0.0179	8 0.0241	9 0.0405	17 0.1628	12 0.0209	13 0.0288	15 0.0360	15 0.0550	38 0.1576
59	MICROSOFT-5	9 0.0119	9 0.0171	9 0.0218	7 0.0387	18 0.1654	13 0.0201	11 0.0279	17 0.0347	13 0.0545	13 0.1549
60	MICROSOFT-6	7 0.0058	5 0.0080	5 0.0110	6 0.0284	19 0.1664	9 0.0109	9 0.0141	9 0.0183	9 0.0343	12 0.1544
61	NEC-0	60 0.0483	59 0.0604	59 0.0726	91 0.0989	86 0.2378	64 0.0662	63 0.0815	47 0.0961	44 0.1199	35 0.1994
62	NEC-1	25 0.0211	31 0.0299				40 0.0389	39 0.1081	24 0.1278	26 0.1656	30 0.2811
63	NEC-2	1 0.0018	1 0.0024	1 0.0063	5 0.0211	9 0.0695	1 0.0041	1 0.0097	1 0.0087	1 0.0190	1 0.0733
64	NEC-3	1 0.0018	1 0.0021	1 0.0026	1 0.0113	1 0.0788	1 0.0038	1 0.0044	1 0.0049	1 0.0095	1 0.0580
65	NEUROTECHNOLOGY-3	159 0.5809	134 0.6390				162 0.5959	152 0.6649	88 0.7217	90 0.7352	76 0.8336
66	NEUROTECHNOLOGY-4	55 0.0427	58 0.0575	56 0.0711	74 0.0954	84 0.1845	55 0.0499	54 0.0656	46 0.0810	45 0.1167	42 0.2138
67	NEUROTECHNOLOGY-5	50 0.0384	51 0.0527	50 0.0646	67 0.0811	76 0.1366	52 0.0422	48 0.0564	37 0.0705	36 0.0998	38 0.2014
68	NEWLAND-2						140 0.4015	140 0.4405	98 0.4719	74 0.5138	
69	NOBLIS-4	148 0.9043	147 0.9989				167 0.9965	166 0.9974	87 0.9980	87 0.9986	
70	NTECHLAB-0	65 0.0518	67 0.0666	81 0.0850	98 0.1158		67 0.0677	64 0.0830	59 0.1029	52 0.1306	54 0.1948
71	NTECHLAB-1	97 0.0634	91 0.0818	82 0.1006	82 0.1357	96 0.2162	78 0.0808	76 0.1021			
72	NTECHLAB-3	48 0.0329	48 0.0434				46 0.0445	47 0.0561	36 0.0699	30 0.0993	31 0.1609

Table 10: **Identification-mode: Effect of N on FNIR at high threshold.** *Values are threshold-based miss rates i.e.* FNIR *at* FPIR *= 0.001 for five enrollment population sizes, N. The left six columns apply for enrollment of a variable number of images per subject. The right six columns apply for enrollment of one image. Missing entries usually apply because another algorithm from the same developer was run instead. Some developers are missing because less accurate algorithms were not run on galleries with $N \geq 3\,000\,000$. Throughout blue superscripts indicate the rank of the algorithm for that column.*

MISSES BELOW THRESHOLD, T		ENROL LIFETIME					ENROL MOST RECENT				
FNIR(N, T>0, R>1)		DATASET: FRVT 2018					DATASET: FRVT 2018				
#	ALGORITHM	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M
73	NTECHLAB-4	27 0.0253	26 0.0337	20 0.0433	24 0.0692	28 0.1845	21 0.0337	20 0.0431	22 0.0545	20 0.0749	24 0.1628
74	NTECHLAB-5	26 0.0268	27 0.0347				26 0.0358	28 0.0448	24 0.0561	23 0.0785	27 0.1572
75	NTECHLAB-6	21 0.0227	20 0.0301	16 0.0395	21 0.0654	27 0.1837	18 0.0311	18 0.0391	19 0.0496	19 0.0696	24 0.1548
76	QUANTASOFT-1	284 0.2915	196 0.2915				273 0.6399	170 0.6399	83 0.6399		78 0.6399
77	RANKONE-0	106 0.1485	104 0.1788	82 0.2210	84 0.3260	43 0.4758	116 0.1899	113 0.2192	78 0.2635	73 0.2992	73 0.4201
78	RANKONE-1	104 0.1211	103 0.1549	83 0.1804	81 0.2371	42 0.3530	107 0.1542	100 0.1693			
79	RANKONE-2	77 0.0744	76 0.0943				72 0.0998	70 0.1200	64 0.1382	61 0.1744	62 0.2636
80	RANKONE-3	76 0.0744	75 0.0943	46 0.1120	45 0.1490	41 0.2946	81 0.0998	84 0.1200	63 0.1382	60 0.1744	61 0.2636
81	RANKONE-4	116 0.1265	110 0.1515				109 0.1631	109 0.1951	75 0.2211		
82	RANKONE-5	48 0.0347	46 0.0447	32 0.0571	30 0.0847	22 0.2549	50 0.0499	51 0.0617	38 0.0722	35 0.0984	41 0.2031
83	REALNETWORKS-0	213 0.2098	211 0.2476	155 0.2837			120 0.2003	119 0.2362			
84	REALNETWORKS-2	211 0.1688	209 0.2049				118 0.1974	117 0.2343	89 0.2691	75 0.3186	69 0.3261
85	REMARKAI-2	74 0.0731	74 0.0991				69 0.0971	71 0.1264	48 0.1495	44 0.1928	
86	SENSETIME-0	9 0.0118	9 0.0165				11 0.0184	9 0.0234	9 0.0295	8 0.0427	8 0.1287
87	SENSETIME-1	11 0.0129	11 0.0175				11 0.0186	11 0.0245	11 0.0304	11 0.0448	9 0.1344
88	SHAMAN-3	248 0.3506	228 0.3921	65 0.4295			251 0.4179	151 0.4527			
89	SHAMAN-7	66 0.0924	62 0.1112				76 0.1236	56 0.1436	49 0.1610	44 0.1901	45 0.2480
90	SIAT-1	232 0.2695	216 0.2727	87 0.2758			7 0.0160	6 0.0201	5 0.0260	5 0.0380	3 0.1069
91	SIAT-2	228 0.2196	198 0.2239				8 0.0179	10 0.0242	10 0.0301	10 0.0424	56 0.1297
92	SMILART-4	272 0.8281	194 0.9569				282 0.9260	131 0.9883	87 0.9913		
93	SYNESIS-3	256 0.4748	232 0.5296				264 0.5353	164 0.5822	92 0.6123	79 0.6489	77 0.6838
94	TEVIAN-4	52 0.0687	72 0.0878	45 0.1032			53 0.0952	36 0.1201			
95	TEVIAN-5	64 0.0518	64 0.0667				67 0.0717	58 0.0898	54 0.1094	43 0.1335	91 0.1875
96	TIGER-0	236 0.2959	228 0.3361	69 0.3659	57 0.4139		243 0.3462	143 0.3921			
97	TIGER-2	67 0.0511	66 0.0698				66 0.0671	56 0.0888	50 0.1065	49 0.1361	46 0.2284
98	TONGYITRANS-1	71 0.0659	71 0.0885	42 0.1017	41 0.1828		77 0.0745	59 0.0698			
99	TOSHIBA-0	40 0.0374	53 0.0529				54 0.0688	53 0.0848	43 0.0809	37 0.1170	43 0.2140
100	VD-0	278 0.8686	242 0.9048	69 0.9242	43 0.9381		286 0.8892	166 0.9171			
101	VD-1	208 0.1312	202 0.1654				210 0.1664	113 0.2096	77 0.2372	72 0.2759	70 0.3814
102	VIGILANTSOLUTIONS-3	230 0.3081	226 0.3568	65 0.3861	55 0.3861		249 0.3648	147 0.4037			
103	VISIONLABS-3	31 0.0260	30 0.0347	23 0.0444	23 0.0678		29 0.0394	35 0.0506	26 0.0629	25 0.0902	
104	VISIONLABS-4	30 0.0254	40 0.0402				30 0.0452	29 0.0604	29 0.0733	34 0.0982	31 0.1893
105	VISIONLABS-5	26 0.0250	32 0.0353	22 0.0441	29 0.0628	21 0.1727	40 0.0506	39 0.0581	30 0.0654	22 0.0878	32 0.1894
106	VISIONLABS-6	12 0.0131	13 0.0185				13 0.0211	15 0.0289	14 0.0359	16 0.0571	15 0.1672
107	VISIONLABS-7	12 0.0131	12 0.0186	9 0.0242	10 0.0412	9 0.1495	14 0.0211	14 0.0299	13 0.0359	15 0.0569	12 0.1676
108	VOCORD-6	66 0.0969	81 0.1295	52 0.1627	50 0.2361		58 0.0978	50 0.1258			
109	VOCORD-5	76 0.0735	84 0.1076				59 0.1261	102 0.1697	56 0.2327	76 0.3286	75 0.4629
110	YISHENG-1	230 0.2529	219 0.3004	68 0.3366	56 0.3892		228 0.3026	139 0.3483			
111	YITU-0	37 0.0279	32 0.0358	25 0.0468	30 0.0636	8 0.1389	36 0.0388	33 0.0502	27 0.0622	27 0.0862	23 0.1621
112	YITU-1	21 0.0261	28 0.0341	21 0.0434	17 0.0611	6 0.1361	28 0.0366	28 0.0472			
113	YITU-2	5 0.0096	6 0.0133	5 0.0174	5 0.0274	5 0.1180	6 0.0156	7 0.0204	6 0.0258	7 0.0682	5 0.1241
114	YITU-3	7 0.0103	7 0.0139				8 0.0165	8 0.0213	7 0.0266	8 0.0399	7 0.1248
115	YITU-4	3 0.0092	5 0.0074	3 0.0099	4 0.0187	4 0.1153	3 0.0093	3 0.0123	3 0.0159	3 0.0275	6 0.1107
116	YITU-5	4 0.0097	4 0.0096	4 0.0100	3 0.0188	3 0.1111	4 0.0101	4 0.0128	4 0.0163	4 0.0294	4 0.1118

Table 11: **Identification-mode: Effect of N on FNIR at high threshold.** *Values are threshold-based miss rates i.e.* FNIR *at* FPIR *= 0.001 for five enrollment population sizes, N. The left six columns apply for enrollment of a variable number of images per subject. The right six columns apply for enrollment of one image. Missing entries usually apply because another algorithm from the same developer was run instead. Some developers are missing because less accurate algorithms were not run on galleries with N ≥ 3 000 000. Throughout blue superscripts indicate the rank of the algorithm for that column.*

	MISSES NOT AT RANK 1	ENROL LIFETIME						ENROL MOST RECENT					
	FNIR(N, T= 0, R =1)	DATASET: FRVT 2018						DATASET: FRVT 2018					
#	ALGORITHM	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}
1	3DIVI-3	0.0494	0.0645	0.0789	0.0998		$0.0014\ N^{0.367}$	0.0680	0.0887				$0.0043\ N^{0.382}$
2	3DIVI-5	0.0100	0.0133				$0.0002\ N^{0.410}$	0.0163	0.0202	0.0236	0.0279	0.0327	$0.0007\ N^{0.420}$
3	ALCHERA-0	0.0106	0.0121	0.0135	0.0170		$0.0006\ N^{0.307}$	0.0167	0.0186				$0.0035\ N^{0.317}$
4	ALCHERA-3	0.0119	0.0159				$0.0002\ N^{0.412}$	0.0101	0.0127	0.0146	0.0171	0.0204	$0.0004\ N^{0.396}$
5	ANKE-0	0.0077	0.0100				$0.0001\ N^{0.387}$	0.0128	0.0158	0.0181	0.0214	0.0251	$0.0006\ N^{0.391}$
6	AWARE-3	0.0165	0.0209	0.0247	0.0297		$0.0005\ N^{0.393}$	0.0264	0.0332	0.0387	0.0456	0.0532	$0.0011\ N^{0.396}$
7	AWARE-5	0.0163	0.0208				$0.0004\ N^{0.379}$	0.0271	0.0337	0.0392	0.0460	0.0538	$0.0010\ N^{0.392}$
8	AYONIX-0	0.4198	0.4643	0.4969	0.5318		$0.1021\ N^{0.108}$	0.4095	0.4519				$0.0973\ N^{0.108}$
9	AYONIX-2	0.2192	0.2606				$0.0176\ N^{0.289}$	0.2954	0.3432	0.3753	0.4116	0.4480	$0.0449\ N^{0.342}$
10	CAMVI-3	0.0149	0.0368	0.0528	0.1791		$0.0000\ N^{1.156}$	0.0224	0.0544				$0.0000\ N^{0.992}$
11	CAMVI-4	0.0082	0.0326				$0.0000\ N^{1.310}$	0.0140	0.0490	0.0741	0.2382	0.2586	$0.0000\ N^{1.092}$
12	COGENT-0	0.0103	0.0106	0.0109	0.0114	0.0122	$0.0047\ N^{0.059}$	0.0127	0.0131	0.0136	0.0141	0.0151	$0.0058\ N^{0.062}$
13	COGENT-1	0.0103	0.0106				$0.0076\ N^{0.023}$	0.0127	0.0131	0.0136	0.0141	0.0151	$0.0058\ N^{0.062}$
14	COGENT-2	0.0022	0.0027	0.0032	0.0037	0.0043	$0.0001\ N^{0.202}$	0.0054	0.0062	0.0067	0.0075	0.0085	$0.0007\ N^{0.181}$
15	COGENT-3	0.0032	0.0037	0.0042	0.0048	0.0056	$13.4494\ N^{0.467}$	0.0059	0.0064	0.0069	0.0077	0.0087	$0.0009\ N^{0.144}$
16	COGNITEC-0	0.0146	0.0189				$0.0001\ N^{0.366}$	0.0221	0.0278	0.0323	0.0378	0.0443	$0.0010\ N^{0.333}$
17	COGNITEC-1	0.0069	0.0089	0.0106	0.0128	0.0154	$0.0002\ N^{0.278}$	0.0116	0.0143	0.0165	0.0192	0.0225	$0.0006\ N^{0.286}$
18	COGNITEC-2	0.0035	0.0044	0.0052	0.0061	0.0075	$0.0001\ N^{0.284}$	0.0074	0.0083	0.0093	0.0105	0.0121	$0.0008\ N^{0.168}$
19	COGNITEC-3	0.0040	0.0048	0.0055	0.0064	0.0078	$0.0003\ N^{0.180}$	0.0078	0.0088	0.0098	0.0111	0.0126	$0.0009\ N^{0.164}$
20	DAHUA-1	0.0090	0.0093				$0.0002\ N^{0.242}$	0.0074	0.0089	0.0102	0.0115	0.0135	$0.0005\ N^{0.230}$
21	DERMALOG-4	0.0759	0.0961	0.1105	0.1260		$0.0037\ N^{0.229}$	0.1040	0.1274				$0.0054\ N^{0.221}$
22	DERMALOG-5	0.0081	0.0113				$0.0001\ N^{0.350}$	0.0135	0.0171	0.0223	0.0312	0.0470	$0.0004\ N^{0.269}$
23	DERMALOG-6	0.0055	0.0060				$0.0015\ N^{0.095}$	0.0085	0.0102	0.0107	0.0115	0.0125	$0.0027\ N^{0.091}$
24	EVERAI-0	0.0065	0.0166				$0.0000\ N^{1.029}$	0.0102	0.0209	0.0348			$0.0000\ N^{0.795}$
25	EVERAI-1	0.0022	0.0027				$0.0001\ N^{0.222}$	0.0047	0.0056	0.0061			$0.0005\ N^{0.166}$
26	EVERAI-3	0.0020	0.0023	0.0026	0.0028		$0.0004\ N^{0.113}$	0.0041	0.0047	0.0052	0.0059	0.0066	$0.0005\ N^{0.160}$
27	EYEDEA-3	0.0480	0.0613	0.0719	0.0831		$0.0018\ N^{0.240}$	0.0663	0.0824				$0.0028\ N^{0.228}$
28	GLORY-1	0.0818	0.0932	0.1009	0.1091		$0.0147\ N^{0.127}$	0.1154	0.1291				$0.0229\ N^{0.123}$
29	GORILLA-2	0.0102	0.0137				$0.0001\ N^{0.321}$	0.0170	0.0220	0.0261	0.0311	0.0378	$0.0005\ N^{0.289}$
30	HIK-2	0.0155	0.0186	0.0208	0.0240	0.0272	$0.0012\ N^{0.193}$	0.0147	0.0172				$0.0015\ N^{0.179}$
31	HIK-3	0.0085	0.0107				$0.0003\ N^{0.195}$	0.0116	0.0141	0.0164	0.0194	0.0228	$0.0005\ N^{0.235}$
32	HIK-4	0.0083	0.0104	0.0121	0.0146	0.0177	$0.0003\ N^{0.200}$	0.0112	0.0138	0.0159	0.0188	0.0220	$0.0005\ N^{0.231}$
33	HIK-5	0.0026	0.0034	0.0040	0.0049		$0.0002\ N^{0.169}$	0.0057	0.0067	0.0075	0.0087	0.0103	$0.0004\ N^{0.205}$
34	IDEMIA-0	0.0048	0.0063	0.0076	0.0095	0.0116	$0.0001\ N^{0.304}$	0.0095	0.0113	0.0131	0.0153	0.0182	$0.0004\ N^{0.227}$
35	IDEMIA-1	0.0049	0.0065	0.0080	0.0100	0.0124	$0.0001\ N^{0.320}$	0.0096	0.0116	0.0135	0.0162	0.0194	$0.0004\ N^{0.243}$
36	IDEMIA-2	0.0075	0.0099	0.0119	0.0149	0.0183	$0.0001\ N^{0.304}$	0.0105	0.0126				$0.0003\ N^{0.194}$
37	IDEMIA-3	0.0041	0.0054				$0.0001\ N^{0.294}$	0.0080	0.0095	0.0110	0.0127	0.0148	$0.0005\ N^{0.212}$
38	IDEMIA-4	0.0042	0.0052	0.0061	0.0074	0.0089	$0.0001\ N^{0.267}$	0.0080	0.0092	0.0106	0.0124	0.0143	$0.0005\ N^{0.202}$
39	IDEMIA-5	0.0047	0.0062	0.0073	0.0089	0.0107	$0.0001\ N^{0.281}$	0.0090	0.0107	0.0123	0.0144	0.0169	$0.0005\ N^{0.217}$
40	IDEMIA-6	0.0055	0.0071	0.0083	0.0100	0.0119	$0.0001\ N^{0.291}$	0.0102	0.0122	0.0139	0.0161	0.0187	$0.0006\ N^{0.205}$
41	IMAGUS-2	0.1470	0.1833	0.2088	0.2379		$0.0053\ N^{0.215}$	0.7838	0.2223				$0.0115\ N^{0.316}$
42	INCODE-1	0.0098	0.0131	0.0286	0.0466		$0.0000\ N^{0.629}$	0.0151	0.0190				$0.0005\ N^{0.350}$
43	INCODE-3	0.0067	0.0088				$0.0001\ N^{0.316}$	0.0121	0.0153	0.0178	0.0215	0.0258	$0.0004\ N^{0.297}$
44	INNOVATRICS-4	0.0070	0.0081				$0.0008\ N^{0.162}$	0.0120	0.0149	0.0158	0.0158	0.0158	$0.0040\ N^{0.088}$
45	ISYSTEMS-0	0.0074	0.0095	0.0106	0.0105	0.0118	$0.0009\ N^{0.163}$	0.0122	0.0136				$0.0025\ N^{0.115}$
46	ISYSTEMS-1	0.0074	0.0085	0.0094	0.0105	0.0118	$0.0009\ N^{0.158}$	0.0122	0.0126				$0.0025\ N^{0.118}$
47	ISYSTEMS-2	0.0059	0.0046	0.0052			$0.0004\ N^{0.175}$	0.0076	0.0088	0.0096	0.0108	0.0121	$0.0009\ N^{0.156}$
48	ISYSTEMS-3	0.0035	0.0040	0.0044	0.0050	0.0057	$0.0004\ N^{0.166}$	0.0069	0.0075	0.0081	0.0090	0.0100	$0.0012\ N^{0.129}$
49	LOOKMAN-3	0.0186	0.0237	0.0269			$0.0049\ N^{0.042}$	0.0109	0.0114	0.0117	0.0123	0.0131	$0.0049\ N^{0.039}$
50	MEGVII-0	0.0072	0.0095	0.0115	0.0150	0.0182	$0.0001\ N^{0.317}$	0.0095	0.0094	0.0111	0.0134	0.0162	$0.0002\ N^{0.264}$
51	MEGVII-1							0.0124	0.0137	0.0148	0.0163	0.0182	$0.0021\ N^{0.131}$
52	MICROFOCUS-3	0.4791	0.5389	0.5771			$0.0951\ N^{0.121}$	0.5417	0.5953				$0.1370\ N^{0.103}$
53	MICROFOCUS-5	0.3155	0.3701				$0.0307\ N^{0.174}$	0.3716	0.4267	0.4624	0.5013	0.5404	$0.0684\ N^{0.127}$
54	MICROSOFT-0	0.0021	0.0026	0.0031	0.0040	0.0048	$0.0000\ N^{0.280}$	0.0051	0.0058	0.0066	0.0077	0.0090	$0.0003\ N^{0.196}$
55	MICROSOFT-1	0.0020	0.0026	0.0031	0.0038	0.0047	$0.0000\ N^{0.286}$	0.0049	0.0056				$0.0006\ N^{0.158}$
56	MICROSOFT-2	0.0023	0.0029	0.0035	0.0042	0.0051	$0.0001\ N^{0.272}$	0.0052	0.0061				$0.0005\ N^{0.174}$
57	MICROSOFT-3	0.0009	0.0011				$0.0000\ N^{0.285}$	0.0028	0.0032	0.0035	0.0039	0.0045	$0.0003\ N^{0.166}$
58	MICROSOFT-4	0.0008	0.0010	0.0013	0.0015	0.0019	$0.0000\ N^{0.285}$	0.0027	0.0031	0.0034	0.0038	0.0045	$0.0003\ N^{0.174}$
59	MICROSOFT-5	0.0010	0.0013	0.0015	0.0019	0.0025	$0.0000\ N^{0.304}$	0.0028	0.0033	0.0037	0.0044	0.0052	$0.0002\ N^{0.215}$
60	MICROSOFT-6	0.0010	0.0014	0.0018	0.0020	0.0026	$0.0000\ N^{0.317}$	0.0029	0.0033	0.0039	0.0045	0.0053	$0.0002\ N^{0.206}$
61	NEC-0	0.0097	0.0147	0.0154	0.0185	0.0223	$0.0002\ N^{0.284}$	0.0157	0.0196	0.0229	0.0270	0.0320	$0.0006\ N^{0.248}$
62	NEC-1	0.0136	0.0164				$0.0009\ N^{0.202}$	0.0206	0.0235	0.0259	0.0292	0.0329	$0.0024\ N^{0.160}$
63	NEC-2	0.0010	0.0011	0.0012	0.0012	0.0014	$0.0003\ N^{0.086}$	0.0026	0.0028	0.0028	0.0030	0.0031	$0.0012\ N^{0.069}$
64	NEC-3	0.0012	0.0013	0.0014	0.0014	0.0016	$0.0005\ N^{0.061}$	0.0030	0.0031	0.0032	0.0033	0.0035	$0.0016\ N^{0.048}$
65	NEUROTECHNOLOGY-3	0.0161	0.0199				$0.0007\ N^{0.284}$	0.0204	0.0250	0.0288	0.0331	0.0386	$0.0011\ N^{0.256}$
66	NEUROTECHNOLOGY-4	0.0049	0.0058	0.0065	0.0075	0.0087	$0.0004\ N^{0.195}$	0.0072	0.0082	0.0090	0.0100	0.0114	$0.0009\ N^{0.186}$
67	NEUROTECHNOLOGY-5	0.0035	0.0042	0.0043	0.0053	0.0061	$0.0008\ N^{0.102}$	0.0061	0.0068	0.0074	0.0082	0.0094	$0.0008\ N^{0.149}$
68	NEWLAND-2							0.0671	0.0811	0.0913	0.1033		$0.0050\ N^{0.196}$
69	NORUS-2	0.1261	0.1565				$0.0054\ N^{0.236}$	0.1609	0.1916	0.2060	0.2377		$0.0102\ N^{0.216}$
70	NTECHLAB-0	0.0056	0.0077	0.0094	0.0114	0.0139	$0.0001\ N^{0.323}$	0.0092	0.0115	0.0137	0.0164	0.0196	$0.0003\ N^{0.262}$
71	NTECHLAB-1	0.0070	0.0097	0.0119	0.0146	0.0179	$0.0001\ N^{0.317}$	0.0108	0.0139				$0.0003\ N^{0.278}$
72	NTECHLAB-3	0.0037	0.0051				$0.0000\ N^{0.351}$	0.0065	0.0082	0.0096	0.0115	0.0136	$0.0002\ N^{0.291}$

Table 12: **Investigation-mode: Effect of N on FNIR at rank 1** *For five enrollment population sizes, N, with T = 0 and FPIR = 1. The left five columns apply for consolidated enrollment of a variable number of lifetime images from each subject. The right five columns apply for enrollment of one recent image. Missing entries usually apply because another algorithm from the same developer was run instead. Some developers are missing because less accurate algorithms were not run on galleries with N > 1 600 000. Throughout blue superscripts indicate the rank of the algorithm for that column, and yellow highlighting indicates the most accurate value. Caution: The Power-low models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

MISSES NOT AT RANK 1		ENROL LIFETIME						ENROL MOST RECENT					
FNIR(N, T= 0, R=1)		DATASET: FRVT 2018						DATASET: FRVT 2018					
#	ALGORITHM	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}
73	NTECHLAB-6	[36]0.0030	[37]0.0040	[31]0.0049	[35]0.0060	[31]0.0075	[34]0.0000 $N^{0.815\ 93}$	[28]0.0056	[33]0.0068	[29]0.0078	[28]0.0092	[28]0.0107	[30]0.0006 $N^{0.820\ 95}$
74	NTECHLAB-5	[37]0.0028	[31]0.0039				[31]0.0000 $N^{0.840\ 105}$	[24]0.0051	[29]0.0064	[28]0.0076	[31]0.0092	[29]0.0112	[7]0.0001 $N^{0.856\ 104}$
75	NTECHLAB-6	[34]0.0024	[29]0.0034	[32]0.0042	[19]0.0052	[18]0.0066	[3]0.0000 $N^{0.846\ 102}$	[25]0.0047	[24]0.0059	[23]0.0069	[23]0.0081	[24]0.0098	[10]0.0002 $N^{0.850\ 96}$
76	QUANTASOFT-1	[138]0.9857	[140]0.9857				-	[128]0.2198	[126]0.2198	[125]0.2198		[126]0.2198	[125]0.2198 $N^{0.000\ 1}$
77	RANKONE-0	[122]0.0255	[108]0.0319	[93]0.0366	[95]0.0425	[94]0.0486	[88]0.0014 $N^{0.222\ 87}$	[13]0.0375	[122]0.0455	[90]0.0514	[93]0.0564	[93]0.0654	[88]0.0032 $N^{0.192\ 81}$
78	RANKONE-1	[100]0.0182	[97]0.0194	[87]0.0224	[87]0.0260	[88]0.0302	[77]0.0007 $N^{0.123\ 50}$	[11]0.0226	[106]0.0247				[102]0.0062 $N^{0.097\ 16}$
79	RANKONE-2	[96]0.0117	[89]0.0149				[82]0.0003 $N^{0.126\ 65}$	[10]0.0181	[102]0.0221	[88]0.0250	[86]0.0288	[86]0.0330	[72]0.0012 $N^{0.124\ 55}$
80	RANKONE-3	[95]0.0117	[88]0.0149	[89]0.0172	[87]0.0200	[91]0.0236	[71]0.0005 $N^{0.127\ 54}$	[10]0.0181	[101]0.0221	[87]0.0250	[87]0.0288	[84]0.0330	[71]0.0012 $N^{0.124\ 54}$
81	RANKONE-4	[125]0.0246	[107]0.0318				[74]0.0006 $N^{0.232\ 74}$	[15]0.0351	[125]0.0441	[99]0.0508			[77]0.0014 $N^{0.229\ 73}$
82	RANKONE-5	[89]0.0068	[85]0.0092	[86]0.0086	[82]0.0105	[82]0.0122	[67]0.0002 $N^{0.236\ 66}$	[67]0.0102	[80]0.0120	[83]0.0136	[81]0.0158	[81]0.0182	[64]0.0007 $N^{0.231\ 66}$
83	REALNETWORKS-0	[131]0.0397	[126]0.0443	[96]0.0527			[75]0.0007 $N^{0.270\ 81}$	[127]0.0880	[131]0.0426				[96]0.0003 $N^{0.281\ 110}$
84	REALNETWORKS-2	[117]0.0240	[105]0.0320				[64]0.0004 $N^{0.313\ 91}$	[126]0.0823	[126]0.0418	[96]0.0494	[97]0.0587	[96]0.0604	[82]0.0017 $N^{0.222\ 77}$
85	REMAKAI-2	[46]0.0047	[57]0.0062				[23]0.0001 $N^{0.314\ 93}$	[56]0.0085	[58]0.0105	[46]0.0122	[45]0.0145		[36]0.0004 $N^{0.287\ 96}$
86	SENSETIME-0	[15]0.0016	[13]0.0018				-	[37]0.0048	[36]0.0048	[37]0.0050	[36]0.0053	[33]0.0057	[34]0.0018 $N^{0.272\ 9}$
87	SENSETIME-2	[12]0.0016	[12]0.0018				-	[36]0.0046	[37]0.0048	[35]0.0050	[35]0.0053	[34]0.0062	[26]0.0012 $N^{0.295\ 15}$
88	SHAMAN-3	[116]0.0305	[127]0.0969	[62]0.1091			[57]0.0060 $N^{0.198\ 57}$	[120]0.1074	[118]0.1266				[104]0.0097 $N^{0.180\ 30}$
89	SHAMAN-7	[123]0.0290	[106]0.0310				[52]0.0106 $N^{0.098\ 8}$	[120]0.0397	[120]0.0422	[77]0.0442	[76]0.0488	[72]0.0499	[55]0.0139 $N^{0.098\ 10}$
90	SIAT-1	[170]0.2639	[168]0.2639	[62]0.2640			[110]0.2618 $N^{0.021\ 3}$	[31]0.0037	[32]0.0039	[30]0.0041	[9]0.0044	[8]0.0049	[8]0.0010 $N^{0.195\ 37}$
91	SIAT-2	[171]0.2127	[169]0.2128				[109]0.2115 $N^{0.000\ 2}$	[31]0.0037	[31]0.0040	[31]0.0042	[11]0.0045	[5]0.0049	[7]0.0011 $N^{0.202\ 43}$
92	SMILART-4	[132]0.8183	[142]0.9531				[106]0.0894 $N^{0.166\ 22}$	[125]0.9176	[129]0.9643	[88]0.9908			[116]0.4706 $N^{0.050\ 3}$
93	SYNESIS-9	[143]0.1133	[131]0.1350				[100]0.0089 $N^{0.191\ 22}$	[106]0.1478	[107]0.1721	[79]0.1897	[77]0.2108	[77]0.2338	[108]0.0184 $N^{0.198\ 30}$
94	TEVIAN-4	[58]0.0058	[60]0.0080	[47]0.0097			[19]0.0001 $N^{0.361\ 110}$	[71]0.0105	[75]0.0134				[29]0.0003 $N^{0.356\ 102}$
95	TEVIAN-5	[41]0.0040	[44]0.0053				[21]0.0001 $N^{0.349\ 97}$	[44]0.0074	[46]0.0092	[39]0.0104	[42]0.0125	[42]0.0151	[32]0.0003 $N^{0.340\ 92}$
96	TIGER-0	[130]0.0364	[117]0.0450	[88]0.0565	[85]0.0678		[58]0.0009 $N^{0.276\ 71}$	[142]0.0494	[142]0.0638				[91]0.0012 $N^{0.279\ 79}$
97	TIGER-2	[35]0.0034	[28]0.0044				[20]0.0001 $N^{0.295\ 83}$	[39]0.0063	[37]0.0075	[32]0.0088	[32]0.0107	[24]0.0126	[16]0.0003 $N^{0.295\ 88}$
98	TONGYITRANS-1	[81]0.0096	[81]0.0114	[76]0.0127	[74]0.0148		[50]0.0007 $N^{0.252\ 53}$	[37]0.0080	[32]0.0095				[30]0.0006 $N^{0.287\ 84}$
99	TOSHIBA-0	[76]0.0026	[74]0.0035				[17]0.0001 $N^{0.288\ 77}$	[32]0.0058	[32]0.0068	[27]0.0076	[25]0.0085	[46]0.0178	[5]0.0001 $N^{0.325\ 112}$
100	YD-0	[139]0.8583	[143]0.4303	[97]0.4776	[97]0.5281		[115]0.0355 $N^{0.174\ 24}$	[137]0.4073	[135]0.4781				[110]0.0431 $N^{0.193\ 46}$
101	YD-1	[113]0.0184	[102]0.0221				[89]0.0012 $N^{0.210\ 41}$	[128]0.0256	[115]0.0302	[78]0.0341	[79]0.0389	[79]0.0443	[85]0.0021 $N^{0.198\ 53}$
102	VIGILANTSOLUTIONS-3	[133]0.0410	[121]0.0549	[95]0.0654	[94]0.0654		[94]0.0022 $N^{0.219\ 66}$	[143]0.0561	[145]0.0719				[98]0.0015 $N^{0.251\ 107}$
103	VISIONLABS-3	[39]0.0037	[41]0.0050	[40]0.0076	[38]0.0130		[5]0.0000 $N^{0.363\ 107}$	[42]0.0070	[40]0.0089	[38]0.0124	[39]0.0185		[9]0.0000 $N^{0.432\ 113}$
104	VISIONLABS-4	[14]0.0016	[14]0.0020				[34]0.0001 $N^{0.208\ 43}$	[13]0.0057	[12]0.0044	[14]0.0049	[12]0.0062	[12]0.0088	[8]0.0001 $N^{0.282\ 110}$
105	VISIONLABS-5	[11]0.0015	[12]0.0018	[9]0.0020	[10]0.0028	[10]0.0040	[7]0.0000 $N^{0.282\ 100}$	[9]0.0035	[12]0.0041	[12]0.0046	[14]0.0054	[17]0.0068	[15]0.0002 $N^{0.228\ 78}$
106	VISIONLABS-6	[10]0.0013	[9]0.0015				[32]0.0002 $N^{0.142\ 15}$	[7]0.0030	[7]0.0033	[6]0.0037	[7]0.0044	[12]0.0057	[9]0.0002 $N^{0.214\ 73}$
107	VISIONLABS-7	[9]0.0013	[8]0.0014	[7]0.0016	[7]0.0018	[9]0.0022	[33]0.0001 $N^{0.124\ 29}$	[6]0.0030	[5]0.0033	[4]0.0035	[4]0.0039	[7]0.0050	[23]0.0008 $N^{0.165\ 46}$
108	YOOORD-3	[92]0.0055	[95]0.0067	[92]0.0090	[90]0.0095		[42]0.0001 $N^{0.271\ 68}$	[41]0.0070	[43]0.0095				[23]0.0005 $N^{0.234\ 63}$
109	YOOORD-5	[48]0.0040	[46]0.0057				[37]0.0004 $N^{0.192\ 39}$	[34]0.0081	[42]0.0092	[41]0.0104	[39]0.0120	[39]0.0140	[31]0.0006 $N^{0.198\ 53}$
110	YISHING-1	[109]0.0155	[99]0.0208	[84]0.0248	[81]0.0298		[81]0.0003 $N^{0.251\ 91}$	[116]0.0227	[114]0.0290				[82]0.0006 $N^{0.256\ 105}$
111	YITU-0	[40]0.0040	[38]0.0047	[35]0.0053	[34]0.0061	[30]0.0071	[35]0.0002 $N^{0.202\ 60}$	[38]0.0066	[38]0.0074	[31]0.0082	[28]0.0092	[27]0.0103	[29]0.0008 $N^{0.183\ 37}$
112	YITU-1	[26]0.0029	[30]0.0046	[22]0.0051	[21]0.0059	[19]0.0069	[18]0.0002 $N^{0.219\ 35}$	[27]0.0055	[35]0.0072				[30]0.0015 $N^{0.210\ 52}$
113	YITU-2	[7]0.0013	[10]0.0015	[8]0.0017	[7]0.0019	[6]0.0023	[26]0.0001 $N^{0.196\ 38}$	[23]0.0041	[24]0.0044	[18]0.0047	[17]0.0050	[17]0.0055	[28]0.0011 $N^{0.089\ 18}$
114	YITU-3	[17]0.0021	[17]0.0023				[25]0.0006 $N^{0.199\ 35}$	[21]0.0052	[25]0.0054	[19]0.0057	[18]0.0061	[15]0.0065	[33]0.0017 $N^{0.091\ 17}$
115	YITU-4	[5]0.0010	[5]0.0011	[5]0.0012	[5]0.0013	[5]0.0019	[41]0.0002 $N^{0.143\ 19}$	[20]0.0036	[23]0.0037	[17]0.0040	[16]0.0042	[18]0.0072	[17]0.0002 $N^{0.205\ 68}$
116	YITU-5	[16]0.0019	[16]0.0020	[10]0.0021	[9]0.0023	[8]0.0025	[22]0.0005 $N^{0.098\ 10}$	[16]0.0047	[21]0.0048	[16]0.0050	[13]0.0052	[11]0.0055	[35]0.0021 $N^{0.028\ 3}$

Table 13: **Investigation-mode: Effect of N on FNIR at rank 1** *For five enrollment population sizes, N, with T = 0 and FPIR = 1. The left five columns apply for consolidated enrollment of a variable number of lifetime images from each subject. The right five columns apply for enrollment of one recent image. Missing entries usually apply because another algorithm from the same developer was run instead. Some developers are missing because less accurate algorithms were not run on galleries with N > 1 600 000. Throughout blue superscripts indicate the rank of the algorithm for that column, and yellow highlighting indicates the most accurate value. Caution: The Power-law models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

MISSES NOT AT RANK 50		ENROL LIFETIME						ENROL MOST RECENT					
FNIR(N, T= 0, R =50)		DATASET: FRVT 2018						DATASET: FRVT 2018					
#	ALGORITHM	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}
1	3DIVI-3	27 0.0103	28 0.0151	26 0.0192	26 0.0241		30 0.0001 $N^{0.892}$	28 0.0159	28 0.0217				31 0.0002 $N^{0.902}$
2	3DIVI-5	29 0.0130	27 0.0097				32 0.0001 $N^{0.887}$	31 0.0065	32 0.0094	32 0.0083	32 0.0094	30 0.0107	33 0.0007 $N^{0.987}$
3	ALCHERA-0	118 0.0073	99 0.0076	84 0.0079	49 0.0101		96 0.0012 $N^{0.833}$	128 0.0125	121 0.0128				109 0.0079 $N^{0.934}$
4	ALCHERA-3	98 0.0030	90 0.0040				25 0.0000 $N^{1.209}$	97 0.0047	94 0.0052	98 0.0056	49 0.0065	45 0.0070	82 0.0008 $N^{0.936}$
5	ANKE-0	89 0.0024	71 0.0030				43 0.0001 $N^{1.234}$	88 0.0037	83 0.0065	85 0.0072	87 0.0081	84 0.0092	89 0.0006 $N^{1.064}$
6	AWARE-3	32 0.0039	32 0.0050	45 0.0061	42 0.0077		36 0.0001 $N^{1.259}$	126 0.0051	123 0.0101	71 0.0118	89 0.0139	70 0.0170	97 0.0003 $N^{1.018}$
7	AWARE-5	34 0.0041	34 0.0053				39 0.0001 $N^{1.263}$	124 0.0088	118 0.0108	75 0.0127	74 0.0154	86 0.0115	79 0.0017 $N^{1.128}$
8	AYONIX-0	127 0.1723	126 0.2142	97 0.2487	60 0.2990		105 0.0085 $N^{1.235}$	129 0.1967	127 0.2402				111 0.0107 $N^{1.218}$
9	AYONIX-2	108 0.0646	108 0.0873				88 0.0008 $N^{1.227}$	108 0.0974	106 0.1298	90 0.1547	81 0.1850	76 0.2171	95 0.0026 $N^{1.275}$
10	CAMVI-3	116 0.0142	116 0.0867	93 0.0527	81 0.1799		4 0.0000 $N^{1.691}$	114 0.0221	107 0.0541				90 0.0000 $N^{1.591}$
11	CAMVI-4	122 0.0078	124 0.0323				3 0.0000 $N^{1.565}$	132 0.0137	129 0.0486	83 0.0796	87 0.2389	78 0.2888	91 0.0000 $N^{1.724}$
12	COGENT-0	35 0.0021	53 0.0024	51 0.0027	31 0.0031	39 0.0045	48 0.0001 $N^{1.253}$	119 0.0047	94 0.0050	81 0.0054	48 0.0062	64 0.0122	70 0.0001 $N^{1.288}$
13	COGENT-1	54 0.0021	57 0.0024				47 0.0002 $N^{1.183}$	86 0.0047	93 0.0050	82 0.0054	89 0.0062	85 0.0122	93 0.0001 $N^{1.288}$
14	COGENT-2	38 0.0011	37 0.0013	17 0.0014	17 0.0016	19 0.0017	68 0.0002 $N^{1.147}$	97 0.0038	96 0.0041	88 0.0042	96 0.0044	93 0.0047	78 0.0016 $N^{1.066}$
15	COGENT-3	36 0.0014	31 0.0016	19 0.0018	19 0.0020	17 0.0023	121 36.4798 $N^{0.878}$	98 0.0040	90 0.0042	94 0.0044	90 0.0046	90 0.0048	77 0.0017 $N^{1.068}$
16	COGNITEC-0	92 0.0039	87 0.0050				6 0.0000 $N^{1.599}$	100 0.0076	100 0.0092	98 0.0104	98 0.0123	98 0.0148	99 0.0004 $N^{1.018}$
17	COGNITEC-1	87 0.0024	98 0.0028	89 0.0032	89 0.0037	90 0.0044	91 0.0002 $N^{1.200}$	91 0.0056	78 0.0060	87 0.0066	88 0.0072	87 0.0081	88 0.0010 $N^{1.128}$
18	COGNITEC-2	99 0.0040	92 0.0021	86 0.0025	82 0.0025	73 0.0027	74 0.0004 $N^{1.113}$	88 0.0049	89 0.0052	92 0.0054	93 0.0056	92 0.0060	86 0.0021 $N^{1.093}$
19	COGNITEC-3	85 0.0025	84 0.0025	82 0.0026	84 0.0028	82 0.0031	81 0.0007 $N^{1.084}$	78 0.0053	71 0.0056	79 0.0057	72 0.0060	80 0.0063	80 0.0025 $N^{1.087}$
20	DAHUA-1	83 0.0021	88 0.0022				78 0.0005 $N^{1.299}$	85 0.0046	86 0.0049	96 0.0061	96 0.0054	94 0.0058	74 0.0015 $N^{1.088}$
21	DERMALOG-4	120 0.0186	121 0.0272	81 0.0340	86 0.0427		82 0.0001 $N^{1.372}$	116 0.0162	103 0.0346				103 0.0002 $N^{1.393}$
22	DERMALOG-5	112 0.0066	109 0.0092				84 0.0001 $N^{1.362}$	125 0.0113	126 0.0142	78 0.0192	76 0.0275	75 0.0427	87 0.0004 $N^{1.248}$
23	DERMALOG-6	100 0.0046	94 0.0047				100 0.0035 $N^{1.020}$	126 0.0080	92 0.0091	80 0.0083	80 0.0085	82 0.0087	100 0.0053 $N^{1.030}$
24	EVERAI-0	104 0.0050	115 0.0150				5 0.0000 $N^{1.188}$	131 0.0077	128 0.0182	76 0.0317			92 0.0000 $N^{1.918}$
25	EVERAI-1	30 0.0013	29 0.0014				46 0.0004 $N^{0.986}$	28 0.0031	28 0.0033	28 0.0034			30 0.0012 $N^{0.970}$
26	EVERAI-3	28 0.0012	38 0.0013	16 0.0014	15 0.0014		76 0.0006 $N^{1.037}$	21 0.0029	16 0.0030	24 0.0032	15 0.0034	12 0.0035	69 0.0012 $N^{0.965}$
27	EYEDEA-3	120 0.0113	117 0.0160	90 0.0209	85 0.0252		44 0.0001 $N^{1.364}$	110 0.0175	119 0.0236				95 0.0002 $N^{1.236}$
28	GLORY-1	128 0.0415	128 0.0490	88 0.0569	59 0.0600		103 0.0047 $N^{1.164}$	123 0.0604	121 0.0698				108 0.0073 $N^{1.183}$
29	GORILLA-2	97 0.0028	76 0.0029				30 0.0000 $N^{1.287}$	79 0.0050	77 0.0061	58 0.0070	58 0.0084	56 0.0102	59 0.0002 $N^{1.038}$
30	HIK-2	126 0.0084	125 0.0090	99 0.0097	90 0.0106	88 0.0118	56 0.0018 $N^{1.015}$	111 0.0087	104 0.0093				98 0.0035 $N^{1.048}$
31	HIK-3	58 0.0023	54 0.0028				57 0.0001 $N^{1.030}$	48 0.0044	57 0.0051	50 0.0058	51 0.0066	47 0.0076	54 0.0008 $N^{1.069}$
32	HIK-4	52 0.0025	52 0.0028	41 0.0033	36 0.0039	36 0.0048	52 0.0001 $N^{1.046}$	55 0.0045	58 0.0051	51 0.0058	50 0.0066	46 0.0076	53 0.0004 $N^{1.079}$
33	HIK-5	43 0.0009	47 0.0011	42 0.0012	39 0.0014		86 0.0001 $N^{1.040}$	25 0.0029	25 0.0033	25 0.0035	30 0.0038	37 0.0042	62 0.0006 $N^{1.022}$
34	IDEMIA-0	59 0.0016	61 0.0019	56 0.0023	58 0.0026	59 0.0031	61 0.0001 $N^{1.026}$	53 0.0045	56 0.0051	64 0.0055	66 0.0060	65 0.0067	61 0.0008 $N^{1.034}$
35	IDEMIA-1	56 0.0019	51 0.0024	58 0.0029	56 0.0036	58 0.0046	77 0.0000 $N^{1.307}$	57 0.0049	75 0.0058	66 0.0065	56 0.0076	58 0.0089	53 0.0003 $N^{1.101}$
36	IDEMIA-2	62 0.0031	77 0.0040	62 0.0048	62 0.0058	62 0.0074	57 0.0001 $N^{1.240}$	79 0.0061	87 0.0069				80 0.0010 $N^{1.135}$
37	IDEMIA-3	96 0.0019	96 0.0022				94 0.0002 $N^{1.175}$	94 0.0049	98 0.0053	97 0.0057	96 0.0062	93 0.0067	97 0.0011 $N^{1.105}$
38	IDEMIA-4	97 0.0015	96 0.0017	72 0.0020	71 0.0023	71 0.0029	42 0.0001 $N^{1.207}$	46 0.0043	45 0.0046	56 0.0051	57 0.0056	59 0.0062	63 0.0008 $N^{1.121}$
39	IDEMIA-5	98 0.0018	98 0.0023	98 0.0026	94 0.0033	91 0.0042	98 0.0000 $N^{1.289}$	93 0.0048	79 0.0056	74 0.0062	76 0.0070	74 0.0080	91 0.0005 $N^{1.175}$
40	IDEMIA-6	79 0.0022	81 0.0028	80 0.0034	76 0.0043	76 0.0055	97 0.0001 $N^{1.258}$	76 0.0054	74 0.0062	68 0.0072	69 0.0084	67 0.0102	75 0.0003 $N^{1.124}$
41	IMAGUS-2	110 0.0348	110 0.0510	79 0.0641	66 0.0804		87 0.0002 $N^{1.395}$	106 0.0468	106 0.0657				87 0.0003 $N^{1.371}$
42	INCODE-1	73 0.0026	74 0.0033	68 0.0187	56 0.0323		2 0.0000 $N^{1.467}$	73 0.0055	73 0.0063				55 0.0007 $N^{1.015}$
43	INCODE-3	42 0.0017	44 0.0021				26 0.0001 $N^{1.281}$	59 0.0044	61 0.0052	45 0.0057	52 0.0067	49 0.0078	51 0.0003 $N^{1.194}$
44	INNOVATRICS-4	50 0.0020	49 0.0022				70 0.0004 $N^{1.118}$	72 0.0052	72 0.0059	52 0.0061	48 0.0061	39 0.0061	72 0.0026 $N^{1.054}$
45	ISYSTEMS-0	101 0.0043	88 0.0050	46 0.0053	43 0.0056	39 0.0060	94 0.0017 $N^{1.076}$	103 0.0086	102 0.0089				103 0.0043 $N^{1.044}$
46	ISYSTEMS-1	102 0.0043	96 0.0050	48 0.0053	41 0.0056	38 0.0060	95 0.0017 $N^{1.078}$	105 0.0086	101 0.0089				104 0.0049 $N^{1.040}$
47	ISYSTEMS-2	74 0.0026	67 0.0027	59 0.0029			59 0.0012 $N^{1.061}$	77 0.0054	74 0.0056	53 0.0059	45 0.0060	41 0.0063	73 0.0027 $N^{1.051}$
48	ISYSTEMS-3	75 0.0025	60 0.0026	60 0.0027	57 0.0029	51 0.0030	93 0.0012 $N^{1.053}$	74 0.0052	66 0.0054	42 0.0056	39 0.0057	33 0.0059	84 0.0028 $N^{1.048}$
49	LOOKMAN-3	118 0.0075	102 0.0077				103 0.0060 $N^{1.037}$	117 0.0059	112 0.0100	58 0.0101	94 0.0102	88 0.0104	110 0.0079 $N^{1.016}$
50	MEGVII-0	57 0.0012	60 0.0019	50 0.0025	50 0.0032	50 0.0041	7 0.0000 $N^{1.492}$	19 0.0026	22 0.0031	26 0.0034	26 0.0039	26 0.0048	19 0.0002 $N^{1.001}$
51	MEGVII-1							119 0.0091	108 0.0094	64 0.0097	63 0.0101	59 0.0106	102 0.0044 $N^{1.053}$
52	MICROFOCUS-3	129 0.2047	127 0.2825	96 0.3017			104 0.0070 $N^{1.393}$	126 0.2518	124 0.3113				112 0.0114 $N^{1.332}$
53	MICROFOCUS-5	124 0.1040	120 0.1422				86 0.0011 $N^{1.041}$	120 0.1322	120 0.1744	87 0.2086	93 0.2445	90 0.2829	101 0.0042 $N^{1.060}$
54	MICROSOFT-0	8 0.0008	12 0.0010	11 0.0011	11 0.0012	11 0.0014	45 0.0001 $N^{1.074}$	18 0.0028	17 0.0031	22 0.0032	17 0.0035	14 0.0037	58 0.0007 $N^{1.008}$
55	MICROSOFT-1	9 0.0008	10 0.0009	9 0.0011	10 0.0012	10 0.0014	36 0.0001 $N^{1.077}$	15 0.0028	15 0.0030				60 0.0007 $N^{1.058}$
56	MICROSOFT-2	11 0.0008	11 0.0010	9 0.0011	12 0.0012	12 0.0014	38 0.0001 $N^{1.089}$	20 0.0029	21 0.0032				59 0.0007 $N^{1.061}$
57	MICROSOFT-3	6 0.0004	6 0.0004				24 0.0001 $N^{1.057}$	4 0.0018	4 0.0019	4 0.0021	6 0.0022	5 0.0023	52 0.0006 $N^{1.078}$
58	MICROSOFT-4	3 0.0004	4 0.0004	3 0.0005	4 0.0005	4 0.0006	27 0.0001 $N^{1.040}$	8 0.0018	8 0.0019	8 0.0020	8 0.0021	8 0.0022	54 0.0007 $N^{1.070}$
59	MICROSOFT-5	3 0.0004	5 0.0004	5 0.0005	5 0.0005	5 0.0006	34 0.0001 $N^{1.134}$	6 0.0018	3 0.0018	3 0.0019	5 0.0020	6 0.0021	56 0.0007 $N^{1.067}$
60	MICROSOFT-6	4 0.0004	3 0.0004	4 0.0005	3 0.0006	3 0.0006	28 0.0002 $N^{1.085}$	1 0.0018	2 0.0018	2 0.0019	3 0.0021	4 0.0023	41 0.0005 $N^{1.021}$
61	NEC-0	80 0.0025	73 0.0030	42 0.0038	40 0.0047	37 0.0059	35 0.0000 $N^{1.321}$	75 0.0055	82 0.0064	91 0.0074	83 0.0085	69 0.0100	39 0.0003 $N^{1.009}$
62	NEC-1	119 0.0076	119 0.0080				106 0.0038 $N^{1.267}$	121 0.0138	122 0.0138	74 0.0142	91 0.0147	68 0.0154	107 0.0073 $N^{1.046}$
63	NEC-2	40 0.0008	8 0.0009	7 0.0009	7 0.0009	8 0.0009	75 0.0004 $N^{1.046}$	9 0.0022	9 0.0023	9 0.0023	9 0.0024	9 0.0026	75 0.0014 $N^{1.054}$
64	NEC-3	22 0.0011	20 0.0011	12 0.0011	9 0.0011	9 0.0011	82 0.0008 $N^{1.026}$	17 0.0026	12 0.0027	20 0.0028	9 0.0028	10 0.0029	81 0.0019 $N^{1.026}$
65	NEUROTECHNOLOGY-3	40 0.0038	45 0.0051				25 0.0000 $N^{1.526}$	67 0.0068	67 0.0083	62 0.0097	67 0.0116	67 0.0137	23 0.0008 $N^{1.043}$
66	NEUROTECHNOLOGY-4	70 0.0020	50 0.0024	36 0.0027	30 0.0031	27 0.0036	66 0.0002 $N^{1.187}$	60 0.0048	59 0.0061	54 0.0064	41 0.0067	36 0.0080	79 0.0016 $N^{1.081}$
67	NEUROTECHNOLOGY-5	41 0.0017	29 0.0018	31 0.0019	28 0.0021	16 0.0023	72 0.0004 $N^{1.108}$	51 0.0045	49 0.0047	39 0.0048	29 0.0050	29 0.0053	68 0.0021 $N^{1.085}$
68	NEWLSNDI-2							140 0.0285	144 0.0288	93 0.0332	78 0.0391		98 0.0011 $N^{1.029}$
69	NORLI-2	126 0.0366	123 0.0520				66 0.0002 $N^{1.369}$	122 0.0403	120 0.0560	85 0.0682	90 0.0940		94 0.0003 $N^{1.232}$
70	NTECHLAB-0	31 0.0013	33 0.0016	22 0.0021	23 0.0026	29 0.0032	31 0.0000 $N^{1.281}$	29 0.0033	31 0.0039	31 0.0043	36 0.0051	31 0.0058	27 0.0002 $N^{1.195}$
71	NTECHLAB-1	55 0.0018	39 0.0018	29 0.0022	26 0.0029	28 0.0038	5 0.0000 $N^{1.365}$	31 0.0034	26 0.0040				31 0.0003 $N^{1.125}$
72	NTECHLAB-3	21 0.0010	23 0.0012				32 0.0001 $N^{1.217}$	19 0.0028	33 0.0032	33 0.0035	33 0.0039	28 0.0044	36 0.0004 $N^{1.149}$

Table 14: **Investigation-mode: Effect of N on FNIR at rank 50** *For five enrollment population sizes, N, with T = 0 and FPIR = 1. The left five columns apply for consolidated enrollment of a variable number of lifetime images from each subject. The right five columns apply for enrollment of one recent image. Missing entries usually apply because another algorithm from the same developer was run instead. Some developers are missing because less accurate algorithms were not run on galleries with N > 1 600 000. Throughout blue superscripts indicate the rank of the algorithm for that column, and yellow highlighting indicates the most accurate value. Caution: The Power-low models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

MISSES NOT AT RANK 50		ENROL LIFETIME						ENROL MOST RECENT					
FNIR(N, T= 0, R=50)		DATASET: FRVT 2018						DATASET: FRVT 2018					
#	ALGORITHM	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}	N=0.64M	N=1.6M	N=3.0M	N=6.0M	N=12.0M	αN^{β}
73	NTECHLAB-6	[15]0.0009	[15]0.0010	[14]0.0012	[13]0.0014	[13]0.0016	[25]0.0001 $N^{0.200}$ [68]	[14]0.0027	[13]0.0030	[10]0.0032	[8]0.0035	[7]0.0039	[43]0.0005 $N^{0.120}$ [30]
74	NTECHLAB-5	[8]0.0007	[7]0.0008				[14]0.0000 $N^{0.237}$ [66]	[6]0.0021	[5]0.0025	[4]0.0027	[3]0.0031	[1]0.0035	[30]0.0002 $N^{0.185}$ [49]
75	NTECHLAB-8	[7]0.0008	[5]0.0008	[3]0.0009	[2]0.0010	[2]0.0012	[13]0.0000 $N^{0.244}$ [69]	[5]0.0021	[4]0.0023	[2]0.0026	[2]0.0028	[2]0.0032	[36]0.0005 $N^{0.147}$ [38]
76	QUANTASOFT-1	[128]0.9843	[128]0.9843				-	[128]0.1140	[128]0.1140	[83]0.1140			[117]0.1140 $N^{0.001}$ 1
77	RANKONE-6	[11]0.0094	[10]0.0100	[9]0.0120	[12]0.0146	[14]0.0176	[17]0.0001 $N^{0.229}$ [50]	[13]0.0127	[12]0.0159	[7]0.0185	[9]0.0206	[9]0.0252	[48]0.0006 $N^{0.228}$ [59]
78	RANKONE-1	[9]0.0042	[12]0.0055	[8]0.0067	[10]0.0082	[12]0.0100	[30]0.0001 $N^{0.301}$ [92]	[10]0.0079	[9]0.0096				[53]0.0020 $N^{0.215}$ [48]
79	RANKONE-2	[10]0.0037	[11]0.0047				[15]0.0001 $N^{0.218}$ [72]	[9]0.0075	[10]0.0087	[9]0.0098	[10]0.0111	[10]0.0128	[51]0.0006 $N^{0.194}$ [77]
80	RANKONE-3	[9]0.0037	[9]0.0047	[6]0.0055	[9]0.0067	[9]0.0079	[20]0.0001 $N^{0.256}$ [75]	[8]0.0075	[8]0.0087	[6]0.0098	[6]0.0111	[6]0.0128	[41]0.0006 $N^{0.194}$ [76]
81	RANKONE-4	[8]0.0038	[8]0.0079				[7]0.0001 $N^{0.135}$ [9]	[12]0.0099	[12]0.0128	[7]0.0153			[31]0.0002 $N^{0.256}$ [100]
82	RANKONE-5	[6]0.0021	[6]0.0025	[5]0.0029	[5]0.0034	[7]0.0040	[27]0.0001 $N^{0.221}$ [80]	[7]0.0053	[7]0.0058	[5]0.0063	[7]0.0069	[5]0.0077	[14]0.0009 $N^{0.129}$ [56]
83	REALNETWORKS-0	[16]0.0069	[16]0.0083	[14]0.0108			[35]0.0000 $N^{0.393}$ [102]	[16]0.0077	[16]0.0098				[17]0.0002 $N^{0.267}$ [58]
84	REALNETWORKS-2	[5]0.0042	[5]0.0061				[1]0.0000 $N^{0.423}$ [104]	[4]0.0078	[10]0.0098	[7]0.0119	[7]0.0149	[7]0.0185	[21]0.0004 $N^{0.282}$ [87]
85	REMARKAI-2	[4]0.0013	[4]0.0016				[5]0.0001 $N^{0.234}$ [81]	[3]0.0068	[3]0.0042	[3]0.0046	[5]0.0050		[17]0.0007 $N^{0.125}$ [53]
86	SENSETIME-0	[2]0.0012	[2]0.0013				-	[1]0.0041	[2]0.0041	[2]0.0042	[2]0.0043	[3]0.0044	[9]0.0028 $N^{0.028}$ [4]
87	SENSETIME-1	[3]0.0011	[1]0.0012				-	[2]0.0040	[1]0.0041	[1]0.0041	[1]0.0042	[1]0.0048	[3]0.0018 $N^{0.058}$ [25]
88	SHAMAN-3	[15]0.0244	[17]0.0404	[12]0.0452			[28]0.0032 $N^{0.177}$ [29]	[17]0.0449	[16]0.0544				[18]0.0053 $N^{0.169}$ [37]
89	SHAMAN-7	[14]0.0243	[13]0.0248				[10]0.0183 $N^{0.021}$ [8]	[16]0.0334	[14]0.0329	[8]0.0344	[7]0.0352	[7]0.0362	[13]0.0230 $N^{0.028}$ [7]
90	SIAT-1	[13]0.2635	[14]0.2635	[13]0.2636			[11]0.2626 $N^{0.000}$ [2]	[12]0.0029	[12]0.0030	[12]0.0031	[11]0.0032	[9]0.0033	[24]0.0016 $N^{0.048}$ [17]
91	SIAT-2	[12]0.2124	[12]0.2124				[12]0.2116 $N^{0.000}$ [3]	[11]0.0031	[11]0.0032	[11]0.0032	[12]0.0033	[11]0.0034	[34]0.0020 $N^{0.032}$ [11]
92	SMILART-4	[17]0.8160	[17]0.9622				[16]0.0859 $N^{0.169}$ [46]	[15]0.9159	[15]0.9638	[13]0.9906			[16]0.4632 $N^{0.051}$ [19]
93	SYNESIS-3	[10]0.0082	[10]0.0632				[26]0.0174 $N^{0.190}$ [28]	[14]0.0851	[13]0.0891	[11]0.0942	[10]0.1020	[10]0.1126	[44]0.0231 $N^{0.095}$ [42]
94	TEVIAN-4	[4]0.0019	[4]0.0021	[2]0.0025			[18]0.0002 $N^{0.138}$ [61]	[2]0.0041	[2]0.0046				[9]0.0006 $N^{0.142}$ [41]
95	TEVIAN-5	[5]0.0014	[5]0.0017				[19]0.0002 $N^{0.160}$ [91]	[5]0.0034	[5]0.0037	[5]0.0041	[4]0.0044	[4]0.0050	[4]0.0006 $N^{0.134}$ [37]
96	TIGER-0	[16]0.0061	[16]0.0097	[15]0.0125	[15]0.0164		[14]0.0000 $N^{0.444}$ [103]	[17]0.0098	[17]0.0139				[7]0.0001 $N^{0.356}$ [112]
97	TIGER-2	[13]0.0010	[11]0.0012				[23]0.0001 $N^{0.216}$ [67]	[7]0.0028	[6]0.0030	[6]0.0034	[6]0.0038	[5]0.0045	[5]0.0003 $N^{0.161}$ [46]
98	TONGYITRANS-1	[10]0.0057	[9]0.0060	[7]0.0062	[7]0.0067		[9]0.0020 $N^{0.088}$ [27]	[6]0.0049	[6]0.0052				[8]0.0022 $N^{0.091}$ [27]
99	TOSHIBA-0	[3]0.0011	[3]0.0012				[6]0.0002 $N^{0.172}$ [35]	[3]0.0087	[3]0.0089	[3]0.0091	[3]0.0093	[6]0.0127	[1]0.0000 $N^{0.393}$ [107]
100	YD-0	[17]0.1006	[17]0.1421	[16]0.1782	[17]0.2147		[27]0.0001 $N^{0.341}$ [97]	[13]0.1249	[13]0.1692				[37]0.0014 $N^{0.284}$ [126]
101	YD-1	[12]0.0098	[12]0.0105				[21]0.0001 $N^{0.085}$ [24]	[14]0.0145	[14]0.0155	[9]0.0166	[14]0.0179	[12]0.0196	[29]0.0036 $N^{0.103}$ [47]
102	VIGILANTSOLUTIONS-3	[11]0.0092	[11]0.0110	[11]0.0142	[11]0.0143		[24]0.0001 $N^{0.322}$ [86]	[12]0.0118	[11]0.0166				[5]0.0001 $N^{0.392}$ [111]
103	VISIONLABS-3	[6]0.0030	[6]0.0042	[6]0.0066	[6]0.0119		[5]0.0000 $N^{0.612}$ [107]	[9]0.0057	[9]0.0073	[9]0.0106	[13]0.0186		[4]0.0000 $N^{0.481}$ [110]
104	VISIONLABS-4	[2]0.0010	[1]0.0011				[9]0.0002 $N^{0.103}$ [31]	[1]0.0025	[1]0.0027	[1]0.0030	[2]0.0039	[4]0.0059	[8]0.0000 $N^{0.490}$ [113]
105	VISIONLABS-5	[1]0.0009	[1]0.0010	[1]0.0012	[1]0.0016	[1]0.0026	[8]0.0000 $N^{0.548}$ [94]	[10]0.0026	[10]0.0026	[1]0.0029	[1]0.0035	[1]0.0044	[1]0.0002 $N^{0.332}$ [80]
106	VISIONLABS-6	[15]0.0010	[15]0.0010				[22]0.0005 $N^{0.036}$ [15]	[8]0.0023	[8]0.0025	[8]0.0027	[10]0.0031	[15]0.0040	[15]0.0002 $N^{0.197}$ [57]
107	VISIONLABS-7	[14]0.0009	[13]0.0010	[5]0.0010	[5]0.0011	[5]0.0011	[21]0.0004 $N^{0.090}$ [19]	[5]0.0023	[7]0.0024	[6]0.0025	[5]0.0025	[8]0.0032	[46]0.0006 $N^{0.095}$ [44]
108	YOOYOUNG-3	[6]0.0023	[6]0.0025	[4]0.0028	[4]0.0031		[16]0.0004 $N^{0.123}$ [56]	[6]0.0040	[4]0.0042				[28]0.0017 $N^{0.062}$ [29]
109	YOOYOUNG-5	[7]0.0027	[4]0.0029				[26]0.0018 $N^{0.036}$ [34]	[11]0.0051	[11]0.0054	[4]0.0056	[4]0.0060	[4]0.0064	[12]0.0019 $N^{0.094}$ [45]
110	YISHENG-1	[5]0.0035	[5]0.0047	[4]0.0058	[3]0.0072		[12]0.0000 $N^{0.328}$ [88]	[5]0.0069	[5]0.0082				[12]0.0005 $N^{0.191}$ [69]
111	YITU-0	[7]0.0026	[3]0.0027	[2]0.0029	[2]0.0031	[3]0.0034	[14]0.0008 $N^{0.150}$ [26]	[4]0.0048	[3]0.0049	[3]0.0052	[3]0.0054	[2]0.0057	[7]0.0021 $N^{0.040}$ [8]
112	YITU-1	[8]0.0026	[2]0.0027	[3]0.0029	[3]0.0031	[2]0.0034	[13]0.0008 $N^{0.150}$ [27]	[3]0.0048	[2]0.0049				[3]0.0033 $N^{0.028}$ [3]
113	YITU-2	[17]0.0008	[7]0.0009	[7]0.0009	[6]0.0010	[5]0.0010	[11]0.0004 $N^{0.088}$ [18]	[2]0.0034	[3]0.0035	[2]0.0036	[1]0.0036	[1]0.0037	[2]0.0024 $N^{0.022}$ [2]
114	YITU-3	[4]0.0013	[2]0.0018				[3]0.0011 $N^{0.076}$ [10]	[13]0.0045	[4]0.0047	[5]0.0047	[2]0.0048	[2]0.0049	[3]0.0031 $N^{0.029}$ [2]
115	YITU-4	[9]0.0008	[8]0.0008	[1]0.0008	[1]0.0008	[6]0.0011	[2]0.0006 $N^{0.015}$ [4]	[2]0.0032	[1]0.0033	[1]0.0033	[1]0.0033	[3]0.0060	[2]0.0003 $N^{0.062}$ [6]
116	YITU-5	[4]0.0017	[2]0.0017	[2]0.0017	[2]0.0017	[2]0.0018	[2]0.0014 $N^{0.018}$ [5]	[4]0.0044	[4]0.0044	[2]0.0044	[2]0.0044	[2]0.0045	[1]0.0039 $N^{0.018}$ [5]

Table 15: **Investigation-mode: Effect of N on FNIR at rank 50** *For five enrollment population sizes, N, with T = 0 and FPIR = 1. The left five columns apply for consolidated enrollment of a variable number of lifetime images from each subject. The right five columns apply for enrollment of one recent image. Missing entries usually apply because another algorithm from the same developer was run instead. Some developers are missing because less accurate algorithms were not run on galleries with N > 1 600 000. Throughout blue superscripts indicate the rank of the algorithm for that column, and yellow highlighting indicates the most accurate value. Caution: The Power-law models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

MISSES OUTSIDE RANK R FNIR(N, T=0, R)		RESOURCE USAGE TEMPLATE		ENROLL LIFETIME CONSOLIDATED = 1.6M FRVT 2018 MUGSHOTS				ENROL MOST RECENT, N = 1.6M			
#	ALGORITHM	BYTES	MSEC	R=1	R=10	R=50	WORK-10	R=1	R=10	R=50	WORK-10
1	3DIVI-0	184 4096	91 126				289 10.000	118 0.0344	118 0.0644	118 0.0344	117 1.190
2	3DIVI-1	195 4224	94 128				288 10.000	119 0.0375	119 0.0675	119 0.0375	128 1.238
3	3DIVI-2	48 528	97 228				279 10.000	124 0.0404	124 0.0904	124 0.0404	128 1.259
4	3DIVI-3	48 512	100 625	133 0.0645	122 0.0645	122 0.0645	121 1.345	152 0.0857	152 0.0857	152 0.0857	148 1.469
5	3DIVI-4	177 4096	134 628	88 0.0133	88 0.0133	88 0.0133	85 1.069	96 0.0201	96 0.0201	96 0.0201	94 1.115
6	3DIVI-5	182 4096	136 653	96 0.0133	96 0.0133	96 0.0133	91 1.069	97 0.0202	97 0.0202	97 0.0202	96 1.116
7	3DIVI-6	48 528	138 653	96 0.0186	96 0.0186	96 0.0186	101 1.127	110 0.0265	111 0.0265	110 0.0265	116 1.186
8	ALCHERA-0	131 2048	62 263	82 0.0121	86 0.0121	82 0.0121	87 1.085	92 0.0186	92 0.0186	93 0.0186	103 1.138
9	ALCHERA-1	120 2048	8 66	102 0.9824	103 0.9824	103 0.9824	108 0.748	198 0.9869	198 0.9869	198 0.9869	198 9.812
10	ALCHERA-2	131 2048	38 115	124 0.0914	124 0.0914	124 0.0914	125 1.552	153 0.0973	153 0.0973	153 0.0973	152 1.567
11	ALCHERA-3	136 2048	137 548	91 0.0159	91 0.0159	91 0.0159	85 1.085	72 0.0127	72 0.0127	72 0.0127	80 1.074
12	ANRD-0	186 2072	86 431	72 0.0100	72 0.0100	72 0.0100	80 1.055	86 0.0158	86 0.0158	86 0.0158	80 1.096
13	ANRD-1	184 2072	97 433	73 0.0101	73 0.0101	73 0.0101	83 1.055	87 0.0158	87 0.0158	87 0.0158	85 1.096
14	AWARE-0	100 1564	142 653				192 10.000	138 0.0639	138 0.0639	138 0.0639	137 1.439
15	AWARE-1	99 1564	136 651				188 10.000	141 0.0587	141 0.0587	141 0.0587	143 1.382
16	AWARE-2	187 2076	199 912				194 10.000	142 0.0600	142 0.0600	142 0.0600	145 1.416
17	AWARE-3	186 2076	183 716	101 0.0209	101 0.0209	101 0.0209	102 1.110	146 0.0832	145 0.0832	146 0.0832	146 1.186
18	AWARE-4	92	160 712	119 0.0529	119 0.0529	119 0.0529	115 1.275	167 0.0704	167 0.0704	167 0.0704	167 1.378
19	AWARE-5	175 3100	186 817	100 0.0208	100 0.0208	100 0.0208	99 1.110	137 0.0837	137 0.0837	137 0.0837	138 1.191
20	AWARE-6	7 124	175 813	130 0.0538	120 0.0538	120 0.0538	112 1.286	129 0.0722	168 0.0722	167 0.0722	161 1.394
21	AYONIX-0	91 1036	10	144 0.4649	144 0.4649	144 0.4649	144 4.268	188 0.4519	191 0.4519	191 0.4519	190 4.304
22	AYONIX-1	82 1036	12	140 0.3364	140 0.3364	140 0.3364	139 3.073	187 0.3432	187 0.3432	187 0.3432	186 3.244
23	AYONIX-2	77 1036	11	137 0.2606	137 0.2606	137 0.2606	135 2.620	186 0.3432	186 0.3432	186 0.3432	187 3.244
24	CAMVI-1	62 1024	29 177				158 10.000	179 0.2267	178 0.2267	179 0.2267	176 2.419
25	CAMVI-2	71 1024	172 774				179 10.000	189 0.1292	189 0.1292	189 0.1292	189 1.781
26	CAMVI-3	78 1024	188 797	112 0.0368	112 0.0368	112 0.0368	113 1.330	140 0.0544	140 0.0544	140 0.0544	140 1.488
27	CAMVI-4	60 1024	186 718	110 0.0326	110 0.0326	110 0.0326	110 1.291	137 0.0490	137 0.0490	138 0.0490	136 1.438
28	CAMVI-5	66 1024	170 769	116 0.0458	116 0.0458	116 0.0458	128 1.410	149 0.0673	148 0.0673	148 0.0673	153 1.602
29	COGENT-0	45 525	135 551	77 0.0106	77 0.0106	77 0.0106	76 1.062	74 0.0131	74 0.0131	74 0.0131	74 1.111
30	COGENT-1	45 525	135 552	76 0.0106	76 0.0106	76 0.0106	77 1.062	73 0.0131	73 0.0131	75 0.0131	80 1.111
31	COGENT-2	84 1043	205 987	20 0.0027	20 0.0027	20 0.0027	22 1.017	26 0.0062	25 0.0062	26 0.0062	30 1.065
32	COGENT-3	85 1043	205 960	29 0.0037	29 0.0037	29 0.0037	32 1.024	36 0.0064	37 0.0064	35 0.0064	35 1.067
33	COGNITEC-0	154 2052	35 176	96 0.0189	96 0.0189	96 0.0189	105 1.105	212 0.0278	212 0.0278	212 0.0278	211 1.160
34	COGNITEC-1	150 2052	36 202	66 0.0089	66 0.0089	66 0.0089	64 1.048	89 0.0143	89 0.0143	89 0.0143	76 1.086
35	COGNITEC-2	159 2052	36 227	46 0.0044	34 0.0044	34 0.0044	35 1.027	42 0.0083	42 0.0083	42 0.0083	40 1.059
36	COGNITEC-3	128 2052	36 297	39 0.0048	39 0.0048	39 0.0048	40 1.061	48 0.0088	48 0.0088	48 0.0088	52 1.062
37	DAHUA-0	144 2048	78 378	50 0.0070	50 0.0070	50 0.0070	52 1.047	58 0.0115	64 0.0115	64 0.0115	73 1.082
38	DAHUA-1	138 2048	66 371	40 0.0049	40 0.0049	40 0.0049	38 1.030	47 0.0089	47 0.0089	47 0.0089	48 1.058
39	DERMALOG-0	5 128	64 344				170 10.000	161 0.1309	161 0.1309	161 0.1309	158 1.779
40	DERMALOG-1	7 128	28 171				200 10.000	163 0.1563	163 0.1563	163 0.1563	163 1.945
41	DERMALOG-2	23 256	66 344				178 10.000	162 0.1577	162 0.1577	162 0.1577	161 1.817
42	DERMALOG-3	5 128	31 211	128 0.0970	128 0.0970	128 0.0970	127 1.568	188 0.1281	188 0.1281	188 0.1281	187 1.752
43	DERMALOG-4	4 128	25 208	126 0.0961	126 0.0961	126 0.0961	126 1.561	187 0.1274	187 0.1274	187 0.1274	186 1.748
44	DERMALOG-5	5 128	105 532	79 0.0113	79 0.0113	79 0.0113	81 1.089	89 0.0171	89 0.0171	89 0.0171	103 1.137
45	DERMALOG-6	15 256	105 514	48 0.0060	48 0.0060	48 0.0060	53 1.047	59 0.0102	59 0.0102	59 0.0102	72 1.081
46	EVERAI-0	181 2048	94 438	93 0.0166	93 0.0166	99 0.0166	103 1.141	96 0.0209	96 0.0209	96 0.0209	112 1.174
47	EVERAI-1	110 2048	122 590	21 0.0027	21 0.0027	21 0.0027	21 1.017	20 0.0056	20 0.0056	21 0.0056	19 1.038
48	EVERAI-2	128 2048	71 377	23 0.0029	23 0.0029	23 0.0029	26 1.018	22 0.0058	23 0.0058	23 0.0058	21 1.089
49	EVERAI-3	110 2048	186 735	16 0.0023	16 0.0023	16 0.0023	17 1.016	15 0.0047	15 0.0047	15 0.0047	14 1.084
50	EYEDEA-0	194 4152	89 424				234 10.000	214 0.3000	184 0.3000	184 0.3000	184 2.884
51	EYEDEA-1	81 1036	9 311				235 10.000	222 0.1981	222 0.1981	222 0.1981	222 2.226
52	EYEDEA-2	76 1036	59 429				202 10.000	203 0.2000	203 0.2000	203 0.2000	202 2.246
53	EYEDEA-3	79 1036	77 885	122 0.0613	122 0.0613	122 0.0613	121 1.349	151 0.0824	151 0.0824	151 0.0824	149 1.470
54	GLORY-0	45 418	35 160	130 0.1336	130 0.1336	130 0.1336	131 1.965	178 0.1803	166 0.1803	166 0.1803	163 2.318
55	GLORY-1	110 1726	91 405	128 0.0932	128 0.0932	132 0.0932	128 1.876	184 0.1291	187 0.1291	186 0.1291	182 1.925
56	GORILLA-0	816 8300	91 427				275 10.000				250 10.000
57	GORILLA-1	170 2156	70 169	136 0.0414	134 0.0414	134 0.0414	140 1.211	143 0.0627	145 0.0627	143 0.0627	139 1.381
58	GORILLA-2	87 1132	69 341	97 0.0137	97 0.0137	97 0.0137	96 1.087	101 0.0220	100 0.0220	100 0.0220	98 1.116
59	GORILLA-3	168 2156	126 583	103 0.0245	102 0.0245	105 0.0245	97 1.110	122 0.0384	121 0.0384	122 0.0384	114 1.178
60	HISNNO-0	48 520	46 265				168 10.000	163 0.2796	162 0.2746	163 0.2746	163 2.743
61	HIK-0	116 1808	198 876				158 10.000	107 0.0236	107 0.0236	106 0.0236	113 1.176
62	HIK-1	107 1808	178 810				158 10.000	95 0.0173	91 0.0173	91 0.0173	97 1.116
63	HIK-2	106 1808	178 810	94 0.0185	93 0.0185	94 0.0185	100 1.119	90 0.0172	90 0.0172	90 0.0172	93 1.115
64	HIK-3	90 1408	132 633	76 0.0107	75 0.0107	76 0.0107	74 1.057	82 0.0141	82 0.0141	82 0.0141	74 1.082
65	HIK-4	88 1152	101 510	25 0.0104	25 0.0104	25 0.0104	25 1.055	30 0.0138	30 0.0138	30 0.0138	30 1.081
66	HIK-5	89 1408	132 619	26 0.0034	26 0.0034	26 0.0034	26 1.018	29 0.0067	29 0.0067	29 0.0067	27 1.043
67	HIK-6	91 1408	128 610	27 0.0034	27 0.0034	27 0.0034	24 1.018	30 0.0067	30 0.0067	30 0.0067	28 1.043
68	IDEMIA-0	93 564	86 416	52 0.0063	52 0.0063	52 0.0063	57 1.034	81 0.0113	81 0.0113	81 0.0113	81 1.070
69	IDEMIA-1	87 564	89 417	63 0.0065	65 0.0065	66 0.0065	65 1.035	85 0.0116	87 0.0116	86 0.0116	83 1.072
70	IDEMIA-2	9 564	89 417	70 0.0099	70 0.0099	70 0.0099	71 1.056	71 0.0126	71 0.0126	71 0.0126	71 1.081
71	IDEMIA-3	54 528	149 689	46 0.0054	45 0.0054	45 0.0054	43 1.033	54 0.0095	54 0.0095	54 0.0095	57 1.066
72	IDEMIA-4	47 528	147 669	43 0.0052	43 0.0052	42 0.0052	38 1.029	50 0.0092	50 0.0092	50 0.0092	49 1.061

Table 16: **Rank-based accuracy for the FRVT 2018 mugshot sets.** *In columns 3 and 4 are template size and template generation duration. Thereafter values are rank-based FNIR with T = 0 and FPIR = 1. This is appropriate to investigational uses but not those with higher volumes where candidates from all searches would need review. Columns 5 - 9 show FRVT 2018 accuracy for various ranks for galleries unenrolled with all lifetime images. Column 10 is a workload statistic, a small value shows an algorithm front-loads mates into the first 10 candidates. The last four columns gives analogous results for enrollment only of the most recent image - see Figure 8. Throughout, blue superscripts indicate the rank of the algorithm for that column, and the best value is highlighted in yellow.*

MISSES OUTSIDE R RANK R		RESOURCE USAGE		ENROLL LIFETIME CONSOLIDATED, N = 1.6M				ENROLL MOST RECENT, N = 1.6M			
FNIR(N, T=0, R)		TEMPLATE		FRVT 2018 MUGSHOTS							
#	ALGORITHM	BYTES	MSEC	R=1	R=10	R=50	WORK-10	R=1	R=10	R=50	WORK-10
73	IDEMIA-5	25352	20374	91 0.0062	90 0.0062	81 0.0062	46 1.034	87 0.0107	86 0.0107	87 0.0107	58 1.068
74	IDEMIA-6	25352	29373	97 0.0071	97 0.0071	97 0.0071	55 1.039	97 0.0122	97 0.0122	98 0.0122	67 1.078
75	IMAGUS-0	27512	543				138 10.000	138 0.3054	138 0.3054	138 0.3054	138 2.977
76	IMAGUS-2	27512	576	133 0.1833	133 0.1833	133 0.1833	132 2.070	127 0.2223	127 0.2223	127 0.2223	126 2.329
77	IMAGUS-3	27512	757	129 0.3008	129 0.3008	129 0.3008	128 2.951	128 0.3576	128 0.3576	128 0.3576	128 3.360
78	INCODE-0	991024	99150	113 0.0376	113 0.0376	113 0.0376	109 1.201	123 0.0515	123 0.0515	123 0.0515	122 1.285
79	INCODE-1	122048	151690	81 0.0131	84 0.0131	81 0.0131	78 1.066	93 0.0190	93 0.0190	93 0.0190	88 1.106
80	INCODE-2	122048	98291	81 0.0120	81 0.0120	81 0.0120	76 1.060	98 0.0203	98 0.0203	98 0.0203	92 1.113
81	INCODE-3	122048	155704	86 0.0088	86 0.0088	86 0.0088	80 1.044	89 0.0153	85 0.0153	85 0.0153	76 1.086
82	INNOVATRICS-0	84530	100455				139 10.000	127 0.0421	127 0.0421	127 0.0421	126 1.234
83	INNOVATRICS-1	52530	58316				132 10.000	126 0.0421	126 0.0421	126 0.0421	124 1.234
84	INNOVATRICS-2	84550	95255	128 0.0499	128 0.0499	128 0.0499	122 1.354	136 0.0475	136 0.0475	136 0.0475	132 1.343
85	INNOVATRICS-3	84530	90265	104 0.0301	104 0.0301	104 0.0301	104 1.147	113 0.0287	113 0.0287	113 0.0287	108 1.151
86	INNOVATRICS-4	351076	350406	63 0.0081	61 0.0081	61 0.0081	59 1.042	84 0.0149	84 0.0149	84 0.0149	77 1.087
87	ISYSTEMS-0	141024	133222	84 0.0085	84 0.0085	84 0.0085	74 1.059	77 0.0136	77 0.0136	77 0.0136	69 1.098
88	ISYSTEMS-1	241024	92222	83 0.0085	83 0.0085	83 0.0085	73 1.058	76 0.0136	76 0.0136	76 0.0136	68 1.098
89	ISYSTEMS-2	132048	132316	38 0.0046	37 0.0046	39 0.0046	44 1.032	44 0.0088	44 0.0088	44 0.0088	53 1.062
90	ISYSTEMS-3	118048	129866	36 0.0040	32 0.0040	32 0.0040	29 1.029	37 0.0075	37 0.0075	37 0.0075	40 1.057
91	LOOKMAN-3	26292	53842	97 0.0089	97 0.0089	97 0.0099	85 1.074	102 0.0114	102 0.0114	102 0.0114	103 1.096
92	LOOKMAN-4	55548	40325	85 0.0091	88 0.0091	88 0.0091	86 1.074	116 0.0117	116 0.0117	116 0.0117	81 1.096
93	MEGVII-0	188048	134594	71 0.0099	71 0.0099	71 0.0099	62 1.048	92 0.0094	91 0.0094	91 0.0094	92 1.062
94	MEGVII-1	178096	137652				139 10.000	78 0.0137	78 0.0137	78 0.0137	80 1.102
95	MEGVII-2	180096	150656				178 10.000	79 0.0137	79 0.0137	79 0.0137	87 1.102
96	MICROFOCUS-0	15256	100525				168 10.000	158 0.5972	155 0.5972	155 0.5972	158 5.397
97	MICROFOCUS-1	24256	107527				180 10.000	156 0.5972	156 0.5972	156 0.5972	156 5.398
98	MICROFOCUS-2	22256	108529				174 10.000	157 0.6272	157 0.6272	157 0.6272	157 6.039
99	MICROFOCUS-3	21256	96269	146 0.5389	146 0.5389	146 0.5389	146 4.849	154 0.5955	154 0.5955	154 0.5955	154 5.375
100	MICROFOCUS-4	22256	97270	145 0.5191	145 0.5191	145 0.5191	145 4.688	153 0.5775	153 0.5775	153 0.5775	153 5.212
101	MICROFOCUS-5	16256	95269	140 0.3701	140 0.3701	142 0.3701	141 3.437	149 0.4257	150 0.4257	149 0.4257	149 3.877
102	MICROFOCUS-6	17256	94265	142 0.3732	142 0.3732	142 0.3732	142 3.453	150 0.4283	150 0.4283	150 0.4283	150 3.897
103	MICROSOFT-0	56512	58283	19 0.0026	19 0.0026	19 0.0026	16 1.015	23 0.0058	23 0.0058	23 0.0058	20 1.038
104	MICROSOFT-1	261024	60349	19 0.0026	19 0.0026	18 0.0026	15 1.015	21 0.0056	21 0.0056	21 0.0056	19 1.038
105	MICROSOFT-2	531024	120665	23 0.0029	23 0.0029	23 0.0029	20 1.016	25 0.0061	25 0.0061	25 0.0061	25 1.041
106	MICROSOFT-3	781024	130404	4 0.0011	4 0.0011	4 0.0011	3 1.007	4 0.0032	4 0.0032	4 0.0032	3 1.022
107	MICROSOFT-4	1242048	178778	1 0.0010	1 0.0010	1 0.0010	1 1.006	2 0.0031	2 0.0031	2 0.0031	2 1.020
108	MICROSOFT-5	701024	148675	5 0.0013	5 0.0013	5 0.0013	5 1.007	5 0.0033	5 0.0033	5 0.0033	5 1.021
109	MICROSOFT-6	581024	152695	7 0.0014	7 0.0014	7 0.0014	4 1.007	3 0.0033	3 0.0033	3 0.0033	4 1.023
110	NEC-0	122592	10882	32 0.0127	32 0.0127	30 0.0127	29 1.066	31 0.0196	31 0.0196	31 0.0196	29 1.119
111	NEC-1	171592	11888	36 0.0164	32 0.0164	32 0.0164	32 1.101	109 0.0235	108 0.0235	106 0.0235	110 1.158
112	NEC-2	201616	101653	3 0.0011	3 0.0011	3 0.0011	6 1.009	1 0.0028	1 0.0028	1 0.0028	5 1.023
113	NEC-3	201712	180690	8 0.0013	6 0.0013	6 0.0013	9 1.011	7 0.0031	8 0.0031	8 0.0031	9 1.024
114	NEURO TECHNOLOGY-9	1295214	194782				170 10.000	128 0.0497	128 0.0497	128 0.0497	122 1.278
115	NEUROTECHNOLOGY-1	1385214	132661				152 10.000	138 0.0467	138 0.0467	139 0.0467	138 1.250
116	NEUROTECHNOLOGY-2	1385214	134658				158 10.000	138 0.0465	134 0.0465	134 0.0465	128 1.249
117	NEUROTECHNOLOGY-3	1362048	116547	95 0.0199	95 0.0199	93 0.0199	95 1.108	109 0.0250	109 0.0250	109 0.0250	106 1.148
118	NEUROTECHNOLOGY-4	1442048	115643	46 0.0058	46 0.0058	46 0.0058	50 1.037	80 0.0092	80 0.0092	80 0.0092	84 1.058
119	NEUROTECHNOLOGY-5	15256	84412	52 0.0042	53 0.0042	53 0.0042	54 1.026	51 0.0068	51 0.0068	51 0.0068	51 1.050
120	NEUROTECHNOLOGY-6	12256	107746	80 0.0155	80 0.0155	80 0.0155	84 1.070	76 0.0201	76 0.0201	76 0.0201	68 1.102
121	NEWBAND-2	102048	191868				167 10.000	150 0.0811	150 0.0811	150 0.0811	153 1.491
122	NOBLE-1	1412048	91211	132 0.2049	130 0.2049	130 0.2049	135 2.890	165 0.2512	161 0.2512	154 0.2512	165 2.698
123	NOBLE-2	376144	110635	132 0.1565	132 0.1565	132 0.1565	135 1.967	167 0.1816	167 0.1816	167 0.1816	166 2.098
124	NTECHLAB-0	1564442	169730	59 0.0077	59 0.0077	59 0.0077	53 1.038	53 0.0115	53 0.0115	53 0.0115	51 1.064
125	NTECHLAB-1	1581736	82405	97 0.0097	97 0.0097	97 0.0097	93 1.046	91 0.0139	91 0.0139	91 0.0139	91 1.079
126	NTECHLAB-3	1583484	184831	42 0.0051	42 0.0051	42 0.0051	38 1.024	40 0.0082	41 0.0082	41 0.0082	34 1.047
127	NTECHLAB-4	1583484	188829	31 0.0040	31 0.0040	31 0.0040	24 1.019	32 0.0068	32 0.0068	32 0.0068	24 1.041
128	NTECHLAB-5	1581940	124717	32 0.0039	30 0.0039	30 0.0039	31 1.018	28 0.0064	28 0.0064	28 0.0064	17 1.039
129	NTECHLAB-6	1581940	127841	26 0.0034	26 0.0034	26 0.0034	13 1.015	25 0.0059	24 0.0059	24 0.0059	15 1.034
130	QUANTASOFT-1	1252048	70396	199 0.9857	199 0.9857	199 0.9857	198 9.866	176 0.2198	176 0.2198	176 0.2198	182 2.559
131	RANKONE-0	7228	650	116 0.0819	116 0.0819	109 0.0819	118 1.188	133 0.0455	133 0.0455	133 0.0455	131 1.275
132	RANKONE-1	27824	17136	97 0.0194	97 0.0194	97 0.0194	96 1.109	128 0.0247	128 0.0247	128 0.0247	126 1.145
133	RANKONE-2	31133	24115	88 0.0149	88 0.0149	88 0.0149	80 1.086	122 0.0221	122 0.0221	122 0.0221	116 1.135
134	RANKONE-3	5133	16114	88 0.0149	88 0.0149	88 0.0149	85 1.086	115 0.0221	115 0.0221	115 0.0221	102 1.136
135	RANKONE-4	185	436	118 0.0318	118 0.0318	118 0.0318	119 1.171	132 0.0441	132 0.0441	132 0.0441	128 1.249
136	RANKONE-5	1133	1194	76 0.0072	78 0.0072	78 0.0072	76 1.042	98 0.0120	98 0.0120	98 0.0120	91 1.078
137	REALNETWORKS-0	1691100	89244	215 0.0443	215 0.0443	215 0.0443	211 1.223	129 0.0426	131 0.0426	131 0.0426	122 1.222
138	REALNETWORKS-1	1691104	97243	211 0.0329	211 0.0329	211 0.0329	198 1.163	130 0.0426	130 0.0426	130 0.0426	123 1.222
139	REALNETWORKS-2	1691104	92245	210 0.0320	210 0.0320	209 0.0320	205 1.159	124 0.0418	125 0.0418	125 0.0418	121 1.217
140	REMARKAI-0	1352048	158615	54 0.0065	54 0.0065	54 0.0065	48 1.034	89 0.0109	89 0.0109	90 0.0109	90 1.065
141	REMARKAI-2	1682048	98434	51 0.0062	51 0.0062	51 0.0062	43 1.031	88 0.0105	88 0.0105	88 0.0105	46 1.061
142	SENSETIME-0	1684104	122715	13 0.0018	13 0.0018	13 0.0018	14 1.014	18 0.0048	16 0.0048	16 0.0048	23 1.040
143	SENSETIME-1	1674104	142659	11 0.0018	11 0.0018	11 0.0018	13 1.013	17 0.0048	17 0.0048	17 0.0048	22 1.040
144	SHAMAN-0	1594096	113558				182 10.000	168 0.1707	165 0.1707	165 0.1707	165 2.092

Table 17: **Rank-based accuracy for the FRVT 2018 mugshot sets.** *In columns 3 and 4 are template size and template generation duration. Thereafter values are rank-based FNIR with T = 0 and FPIR = 1. This is appropriate to investigational uses but not those with higher volumes where candidates from all searches would need review. Columns 5 - 9 show FRVT 2018 accuracy for various ranks for galleries unenrolled with all lifetime images. Column 10 is a workload statistic, a small value shows an algorithm front-loads mates into the first 10 candidates. The last four columns gives analogous results for enrollment only of the most recent image - see Figure 8. Throughout, blue superscripts indicate the rank of the algorithm for that column, and the best value is highlighted in yellow.*

MISSES OUTSIDE RANK R		RESOURCE USAGE		ENROLL LIFETIME CONSOLIDATED = 1.6M				ENROL MOST RECENT, N = 1.6M			
FNIR(N, T=0, R)		TEMPLATE		FRVT 2018 MUGSHOTS							
#	ALGORITHM	BYTES	MSEC	R=1	R=10	R=50	WORK-10	R=1	R=10	R=50	WORK-10
145	SHAMAN-1	^{152}4096	^{121}557				^{190}10.000	^{166}0.1718	^{166}0.1718	^{166}0.1718	^{166}2.079
146	SHAMAN-2	^{201}8192	^{125}557				^{194}10.000	^{182}0.2620	^{182}0.2620	^{182}0.2620	^{182}2.710
147	SHAMAN-3	^{121}2048	^{128}704	^{127}0.0969	^{127}0.0969	^{127}0.0969	^{123}1.613	^{168}0.1266	^{168}0.1266	^{168}0.1266	^{168}1.811
148	SHAMAN-4	^{120}2048	^{128}642	^{134}0.1867	^{134}0.1867	^{134}0.1867	^{134}2.163	^{178}0.2242	^{178}0.2242	^{178}0.2242	^{177}2.431
149	SHAMAN-6	^{145}2048	^{165}706	^{116}0.0312	^{116}0.0312	^{116}0.0312	^{104}1.249	^{129}0.0424	^{129}0.0424	^{129}0.0424	^{128}1.389
150	SHAMAN-7	^{174}2048	^{165}709	^{115}0.0310	^{115}0.0310	^{115}0.0310	^{103}1.248	^{128}0.0422	^{128}0.0422	^{128}0.0422	^{126}1.387
151	SIAT-0	^{28}1096	^{67}358				^{181}10.000	^{60}0.0101	^{55}0.0101	^{56}0.0101	^{55}1.059
152	SIAT-1	^{149}2052	^{195}842	^{138}0.2689	^{138}0.2689	^{138}0.2689	^{140}3.373	^{36}0.0089	^{30}0.0089	^{30}0.0089	^{31}1.081
153	SIAT-2	^{153}2052	^{195}906	^{130}0.2128	^{130}0.2128	^{130}0.2128	^{130}2.918	^{31}0.0040	^{31}0.0040	^{31}0.0040	^{35}1.032
154	SMILART-0	^{53}1024	^{203}168				^{191}10.000	^{170}0.1981	^{170}0.1981	^{170}0.1981	^{168}2.204
155	SMILART-1	^{53}1024	^{145}562				^{187}10.000	^{173}0.2188	^{175}0.2188	^{175}0.2188	^{175}2.435
156	SMILART-2	^{201}1024	^{128}560				^{193}10.000	^{171}0.1946	^{171}0.1946	^{171}0.1946	^{169}2.195
157	SMILART-4	^{35}512	^{78}167	^{147}0.9581	^{147}0.9581	^{147}0.9581	^{147}9.578	^{188}0.9649	^{188}0.9649	^{188}0.9649	^{188}9.679
158	SMILART-5	^{126}2048	^{103}464				^{178}10.000				^{201}10.000
159	SYNESIS-0	^{35}512	^{39}237				^{186}10.000	^{164}0.1621	^{163}0.1621	^{163}0.1621	^{172}2.380
160	SYNESIS-3	^{181}4096	^{141}102	^{121}0.1350	^{121}0.1350	^{121}0.1350	^{120}1.388	^{167}0.1721	^{167}0.1721	^{167}0.1721	^{167}2.140
161	TEVIAN-0	^{172}2048	^{77}394				^{183}10.000	^{104}0.0225	^{104}0.0225	^{104}0.0225	^{96}1.122
162	TEVIAN-1	^{146}2048	^{77}398				^{205}10.000	^{105}0.0225	^{105}0.0225	^{105}0.0225	^{104}1.122
163	TEVIAN-2	^{173}2048	^{77}397				^{197}10.000	^{103}0.0224	^{103}0.0224	^{103}0.0224	^{98}1.121
164	TEVIAN-3	^{128}2048	^{55}300	^{74}0.0102	^{74}0.0102	^{74}0.0102	^{67}1.052	^{88}0.0169	^{88}0.0169	^{88}0.0169	^{88}1.093
165	TEVIAN-4	^{130}2048	^{55}299	^{60}0.0080	^{60}0.0080	^{60}0.0080	^{60}1.041	^{70}0.0134	^{70}0.0134	^{70}0.0134	^{68}1.076
166	TEVIAN-5	^{132}2048	^{56}416	^{44}0.0053	^{44}0.0053	^{44}0.0053	^{39}1.028	^{48}0.0092	^{48}0.0092	^{48}0.0092	^{45}1.054
167	TIGER-0	^{155}2052	^{93}428	^{117}0.0480	^{117}0.0480	^{117}0.0480	^{112}1.247	^{194}0.0638	^{194}0.0638	^{140}0.0638	^{139}1.334
168	TIGER-1	^{156}2052	^{78}398				^{189}10.000				^{202}10.000
169	TIGER-2	^{152}2052	^{103}464	^{35}0.0044	^{35}0.0044	^{35}0.0044	^{31}1.023	^{39}0.0075	^{39}0.0075	^{39}0.0075	^{32}1.046
170	TIGER-3	^{147}2052	^{101}464				^{192}10.000	^{38}0.0076	^{38}0.0075	^{38}0.0075	^{35}1.046
171	TONGYITRANS-0	^{162}2070	^{27}190	^{49}0.0060	^{49}0.0060	^{49}0.0060	^{50}1.036	^{73}0.0095	^{58}0.0095	^{58}0.0095	^{59}1.062
172	TONGYITRANS-1	^{142}2070	^{27}189	^{90}0.0114	^{90}0.0114	^{82}0.0114	^{84}1.073	^{53}0.0095	^{53}0.0095	^{53}0.0095	^{55}1.062
173	TOSHIBA-0	^{93}1548	^{193}930	^{24}0.0033	^{24}0.0033	^{24}0.0033	^{28}1.018	^{256}0.0068	^{32}0.0068	^{32}0.0068	^{31}1.046
174	TOSHIBA-1	^{159}2060	^{201}931	^{28}0.0035	^{28}0.0035	^{28}0.0035	^{30}1.019	^{34}0.0071	^{34}0.0071	^{34}0.0071	^{36}1.047
175	VD-0	^{78}1028	^{81}337	^{139}0.4303	^{139}0.4303	^{142}0.4303	^{142}3.703	^{193}0.4751	^{192}0.4751	^{184}0.4751	^{191}4.074
176	VD-1	^{151}2052	^{158}695	^{102}0.0221	^{102}0.0221	^{103}0.0221	^{103}1.140	^{115}0.0302	^{115}0.0302	^{115}0.0302	^{125}1.197
177	VIGILANTSOLUTIONS-0	^{93}1544	^{150}823				^{196}10.000	^{154}0.1254	^{154}0.1254	^{154}0.1254	^{154}1.712
178	VIGILANTSOLUTIONS-1	^{158}2056	^{168}739				^{201}10.000	^{174}0.2038	^{174}0.2038	^{174}0.2038	^{170}2.210
179	VIGILANTSOLUTIONS-2	^{93}1544	^{177}820				^{197}10.000	^{180}0.2387	^{180}0.2387	^{180}0.2387	^{179}2.558
180	VIGILANTSOLUTIONS-3	^{99}1544	^{185}882	^{121}0.0549	^{121}0.0549	^{121}0.0549	^{116}1.280	^{148}0.0719	^{148}0.0719	^{148}0.0719	^{148}1.378
181	VIGILANTSOLUTIONS-4	^{93}1544	^{183}830	^{129}0.0993	^{129}0.0993	^{129}0.0993	^{124}1.549	^{156}0.1272	^{156}0.1272	^{156}0.1272	^{153}1.721
182	VIGILANTSOLUTIONS-5	^{94}1544	^{175}778				^{195}10.000	^{89}0.0118	^{89}0.0118	^{89}0.0118	^{89}1.069
183	VIGILANTSOLUTIONS-6	^{96}1544	^{186}884				^{191}10.000	^{76}0.0125	^{76}0.0125	^{76}0.0125	^{84}1.072
184	VISIONLABS-3	^{13}256	^{35}228	^{41}0.0050	^{41}0.0050	^{41}0.0050	^{40}1.041	^{46}0.0089	^{46}0.0089	^{46}0.0089	^{47}1.072
185	VISIONLABS-4	^{25}256	^{57}315	^{14}0.0020	^{14}0.0020	^{13}0.0020	^{12}1.018	^{13}0.0044	^{13}0.0044	^{13}0.0044	^{20}1.081
186	VISIONLABS-5	^{34}512	^{51}300	^{12}0.0018	^{12}0.0018	^{12}0.0018	^{11}1.012	^{12}0.0041	^{12}0.0041	^{12}0.0041	^{9}1.029
187	VISIONLABS-6	^{42}512	^{55}292	^{9}0.0015	^{9}0.0015	^{9}0.0015	^{10}1.011	^{9}0.0033	^{7}0.0033	^{7}0.0033	^{7}1.025
188	VISIONLABS-7	^{36}512	^{51}293	^{8}0.0014	^{8}0.0014	^{8}0.0014	^{8}1.010	^{8}0.0033	^{8}0.0033	^{8}0.0033	^{8}1.025
189	VOCORD-0	^{87}608	^{171}536				^{190}10.000	^{122}0.0403	^{122}0.0403	^{122}0.0403	^{138}1.301
190	VOCORD-1	^{86}608	^{171}536				^{189}10.000	^{123}0.0402	^{122}0.0402	^{122}0.0402	^{134}1.299
191	VOCORD-2	^{131}2048	^{134}635				^{188}10.000	^{120}0.0382	^{120}0.0382	^{120}0.0382	^{128}1.290
192	VOCORD-3	^{87}836	^{191}714	^{65}0.0067	^{59}0.0067	^{59}0.0067	^{54}1.038	^{42}0.0095	^{43}0.0095	^{43}0.0095	^{42}1.054
193	VOCORD-4	^{97}896	^{154}538	^{92}0.0084	^{78}0.0084	^{76}0.0084	^{66}1.051	^{35}0.0102	^{37}0.0102	^{37}0.0102	^{39}1.068
194	VOCORD-5	^{86}768	^{178}822	^{46}0.0057	^{46}0.0057	^{46}0.0057	^{51}1.036	^{42}0.0092	^{47}0.0092	^{42}0.0092	^{44}1.063
195	VOCORD-6	^{201}10240	^{181}825				^{203}10.000	^{202}1.0000	^{202}1.0000	^{202}1.0000	^{203}10.000
196	YISHENG-0	^{193}2108	^{137}615				^{185}10.000	^{111}0.0268	^{111}0.0268	^{111}0.0268	^{116}1.149
197	YISHENG-1	^{175}3704	^{74}387	^{29}0.0208	^{29}0.0208	^{29}0.0208	^{94}1.105	^{124}0.0290	^{124}0.0290	^{124}0.0290	^{126}1.158
198	YITU-0	^{191}4186	^{191}883	^{34}0.0047	^{42}0.0047	^{36}0.0047	^{45}1.031	^{36}0.0074	^{36}0.0074	^{36}0.0074	^{41}1.053
199	YITU-1	^{190}4186	^{200}930	^{30}0.0046	^{30}0.0046	^{30}0.0046	^{41}1.031	^{30}0.0072	^{35}0.0072	^{35}0.0072	^{38}1.052
200	YITU-2	^{185}4138	^{200}870	^{10}0.0015	^{10}0.0015	^{10}0.0015	^{7}1.010	^{14}0.0044	^{14}0.0044	^{14}0.0044	^{14}1.035
201	YITU-3	^{192}4138	^{199}871	^{17}0.0023	^{17}0.0023	^{17}0.0023	^{23}1.018	^{19}0.0054	^{25}0.0054	^{25}0.0054	^{29}1.044
202	YITU-4	^{193}2070	^{196}910	^{5}0.0011	^{5}0.0011	^{5}0.0011	^{5}1.008	^{6}0.0037	^{5}0.0037	^{5}0.0037	^{12}1.031
203	YITU-5	^{191}2070	^{190}861	^{16}0.0020	^{16}0.0020	^{17}0.0020	^{19}1.016	^{18}0.0048	^{18}0.0048	^{18}0.0048	^{26}1.041

Table 18: **Rank-based accuracy for the FRVT 2018 mugshot sets.** *In columns 3 and 4 are template size and template generation duration. Thereafter values are rank-based FNIR with T = 0 and FPIR = 1. This is appropriate to investigational uses but not those with higher volumes where candidates from all searches would need review. Columns 5 - 9 show FRVT 2018 accuracy for various ranks for galleries unenrolled with all lifetime images. Column 10 is a workload statistic, a small value shows an algorithm front-loads mates into the first 10 candidates. The last four columns gives analogous results for enrollment only of the most recent image - see Figure 8. Throughout, blue superscripts indicate the rank of the algorithm for that column, and the best value is highlighted in yellow.*

MISSES BELOW THRESHOLD, T		ENROLMOST RECENT MUGSHOT, N = 1.6M								
FNIR(N, T>0, R>1)		DATASET: FRVT 2018 MUGSHOTS			DATASET: WEBCAM PROBES			DATASET: PROFILE PROBES		
#	ALGORITHM	FPIR=0.001	FPIR=0.01	FPIR=0.1	FPIR=0.001	FPIR=0.01	FPIR=0.1	FPIR=0.001	FPIR=0.01	FPIR=0.1
1	3DIVI-0	128 0.256	128 0.160	135 0.086	135 0.425	119 0.202	115 0.180			
2	3DIVI-1	125 0.256	128 0.160	136 0.087						
3	3DIVI-2	124 0.255	136 0.164	137 0.089						
4	3DIVI-3	145 0.402	151 0.284	152 0.168	131 0.626	133 0.497	138 0.243			
5	3DIVI-4	105 0.171	107 0.096	104 0.047	128 0.345	128 0.237	109 0.138			
6	3DIVI-5	108 0.169	106 0.095	103 0.047	126 0.339	127 0.234	109 0.137	30 0.995	25 0.987	26 0.961
7	3DIVI-6	109 0.170	112 0.098	110 0.051	125 0.342	129 0.238	110 0.142			
8	ALCHERA-0	96 0.140	94 0.073	91 0.035	77 0.216	78 0.146	79 0.087			
9	ALCHERA-1	156 0.999	155 0.999	139 0.995	159 1.000	157 1.000	161 1.000			
10	ALCHERA-2	138 0.480	153 0.304	154 0.184	128 0.591	107 0.442	118 0.295			
11	ALCHERA-3	96 0.169	96 0.072	84 0.030	84 0.239	82 0.152	74 0.081	28 0.999	27 0.993	22 0.921
12	ANKE-0	80 0.120	80 0.065	87 0.033	79 0.220	90 0.151	82 0.088	18 0.991	24 0.985	25 0.972
13	ANKE-1	80 0.122	84 0.065	88 0.033	78 0.220	81 0.151	81 0.088			
14	AWARE-0	154 0.983	128 0.128	138 0.085	145 0.817	111 0.253	116 0.178			
15	AWARE-1	150 0.996	120 0.127	132 0.081						
16	AWARE-2	136 0.997	121 0.120	130 0.078						
17	AWARE-3	95 0.131	100 0.085	108 0.051	89 0.298	100 0.204	102 0.132			
18	AWARE-4	127 0.171	135 0.177	144 0.107	123 0.509	125 0.375	124 0.253			
19	AWARE-5	138 0.378	103 0.088	115 0.050	87 0.253	85 0.163	88 0.099	20 1.000	18 0.999	20 0.998
20	AWARE-6	120 0.278	141 0.178	146 0.109	113 0.398	115 0.285	115 0.188			
21	AYONIX-0	160 0.811	157 0.725	130 0.508	153 0.939	154 0.892	138 0.802			
22	AYONIX-1	133 0.826	122 0.702	128 0.526	148 0.920	150 0.845	151 0.703			
23	AYONIX-2	132 0.825	123 0.702	127 0.526	149 0.920	149 0.845	150 0.702			
24	CAMVI-1	170 0.684	170 0.549	178 0.375	140 0.770	144 0.648	146 0.468			
25	CAMVI-2	161 0.587	165 0.402	141 0.242						
26	CAMVI-3	35 0.074	31 0.050	115 0.055	46 0.182	59 0.108	52 0.094			
27	CAMVI-4	37 0.074	72 0.056	104 0.050	48 0.136	58 0.100	75 0.083	36 0.999	31 0.994	35 0.816
28	CAMVI-5	76 0.102	99 0.078	123 0.069	72 0.179	78 0.132	96 0.110			
29	COGENT-0	46 0.056	54 0.032	81 0.020	52 0.140	68 0.100	71 0.069			
30	COGENT-1	43 0.056	51 0.032	80 0.020	50 0.140	61 0.100	70 0.069			
31	COGENT-2	30 0.047	28 0.020	21 0.010	24 0.098	28 0.063	35 0.036	12 0.997	20 0.993	34 0.983
32	COGENT-3	36 0.051	23 0.018	19 0.009	30 0.095	33 0.061	29 0.037			
33	COGNITEC-0	126 0.163	108 0.098	111 0.053	100 0.303	98 0.200	97 0.115			
34	COGNITEC-1	77 0.105	77 0.055	78 0.027	82 0.230	77 0.135	72 0.071			
35	COGNITEC-2	46 0.056	44 0.027	39 0.014	74 0.178	54 0.101	52 0.060	21 1.000	15 0.947	23 0.936
36	COGNITEC-3	45 0.055	44 0.028	41 0.014	51 0.162	59 0.100	51 0.060			
37	DAHUA-0	67 0.089	71 0.047	83 0.022	47 0.136	47 0.083	45 0.046			
38	DAHUA-1	36 0.075	53 0.039	56 0.018	41 0.122	40 0.075	39 0.042	10 0.953	10 0.862	13 0.679
39	DERMALOG-0	116 0.488	157 0.364	120 0.233	135 0.657	129 0.528	126 0.362			
40	DERMALOG-1	116 0.528	161 0.405	159 0.268						
41	DERMALOG-2	107 0.503	161 0.378	162 0.244						
42	DERMALOG-3	158 0.484	151 0.362	158 0.231	133 0.655	128 0.526	125 0.361			
43	DERMALOG-4	155 0.481	157 0.360	157 0.230	134 0.657	136 0.526	132 0.362			
44	DERMALOG-5	71 0.091	63 0.045	74 0.024	67 0.154	56 0.096	56 0.057			
45	DERMALOG-6	41 0.054	43 0.028	34 0.015	27 0.105	29 0.067	28 0.039	8 0.948	7 0.856	30 0.642
46	EVERAI-0	76 0.092	73 0.047	80 0.028	67 0.170	60 0.100	61 0.060			
47	EVERAI-1	27 0.052	27 0.028	22 0.010	42 0.128	36 0.074	32 0.039			
48	EVERAI-2	26 0.053	26 0.025	27 0.011	40 0.119	31 0.076	30 0.041			
49	EVERAI-3	17 0.038	17 0.018	17 0.008	21 0.096	21 0.060	22 0.034	14 0.979	5 0.535	4 0.247
50	EYEDEA-0	131 0.812	131 0.679	124 0.484	117 0.914	117 0.783	116 0.619			
51	EYEDEA-1	166 0.632	168 0.480	172 0.305						
52	EYEDEA-2	170 0.794	173 0.490	174 0.338						
53	EYEDEA-3	142 0.389	144 0.267	131 0.160	128 0.543	126 0.404	128 0.264			
54	GLORY-0	135 0.349	134 0.297	109 0.238	126 0.547	130 0.470	122 0.290			
55	GLORY-1	138 0.307	147 0.238	103 0.179	124 0.537	128 0.448	127 0.352			
56	GORILLA-0									
57	GORILLA-1	146 0.403	149 0.248	149 0.136	132 0.453	119 0.314	73 0.191			
58	GORILLA-2	105 0.190	114 0.108	106 0.051	92 0.268	80 0.170	68 0.093			
59	GORILLA-3	125 0.326	130 0.160	124 0.078	117 0.434	110 0.447	105 0.181			
60	HENNG-0	175 0.766	182 0.632	183 0.458						
61	HIK-0	92 0.114	93 0.070	95 0.040	58 0.155	56 0.103	62 0.061			
62	HIK-1	87 0.120	91 0.067	90 0.036						
63	HIK-2	88 0.121	90 0.067	89 0.036						
64	HIK-3	76 0.105	82 0.060	85 0.030	50 0.158	67 0.105	63 0.061			
65	HIK-4	74 0.101	83 0.056	82 0.029	56 0.155	63 0.101	58 0.059			
66	HIK-5	26 0.047	23 0.022	29 0.011	31 0.077	31 0.048	15 0.028	27 0.999	29 0.994	32 0.662
67	HIK-6	33 0.050	25 0.022	38 0.011	32 0.086	34 0.052	16 0.029	32 1.000	32 0.997	31 0.645
68	IDEMIA-0	81 0.114	85 0.062	81 0.029	57 0.249	52 0.156	56 0.085			
69	IDEMIA-1	42 0.054	50 0.031	54 0.018						
70	IDEMIA-2	43 0.054	53 0.032	56 0.019						
71	IDEMIA-3	31 0.050	37 0.024	30 0.014	36 0.165	43 0.079	34 0.050			
72	IDEMIA-4	19 0.040	31 0.024	43 0.014	26 0.118	42 0.079	29 0.050	11 0.969	17 0.982	28 0.962

Table 19: **Threshold-based accuracy.** *Values are FNIR(N, T, L) with N = 1.6 million with thresholds set to produce FPIR = 0.001, 0.01, and 0.1 in non-mate searches. Columns 3-5 apply to FRVT-2018 mugshots: Columns 6-8 show the corresponding FNIR values for webcam images searched against the FRVT-2018 mugshot gallery. Finally, the three rightmost columns show FNIR for profile view images searched against the FRVT-2018 frontal gallery. Throughout blue superscripts indicate the rank of the algorithm for that column. Caution: The Power-low models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

MISSES BELOW THRESHOLD, T		ENROL MOST RECENT MUGSHOT, N = 1.6M								
FNIR(N, T> 0, R >1]		DATASET: FRVT 2018 MUGSHOTS			DATASET: WEBCAM PROBES			DATASET: PROFILE PROBES		
#	ALGORITHM	FPIR=0.001	FPIR=0.01	FPIR=0.1	FPIR=0.001	FPIR=0.01	FPIR=0.1	FPIR=0.001	FPIR=0.01	FPIR=0.1
73	IDEMIA-5	26 0.047	26 0.028	47 0.017	52 0.150	56 0.102	69 0.065	132 0.974	119 0.968	97 0.960
74	IDEMIA-6	24 0.046	23 0.028	37 0.018	51 0.226	39 0.161	35 0.108			
75	IMAGUS-0	159 0.784	161 0.608	282 0.453	145 0.872	146 0.779	143 0.635			
76	IMAGUS-2	176 0.751	175 0.566	177 0.377	142 0.818	143 0.645	142 0.460			
77	IMAGUS-3	179 0.808	153 0.670	166 0.512	146 0.909	166 0.809	145 0.667			
78	INCODE-0	134 0.313	134 0.201	143 0.107	118 0.420	103 0.304	117 0.191			
79	INCODE-1	114 0.214	117 0.114	104 0.050	76 0.296	76 0.198	76 0.110			
80	INCODE-2	107 0.186	112 0.102	100 0.046	83 0.269	71 0.176	67 0.100			
81	INCODE-3	103 0.170	101 0.086	96 0.039	80 0.264	69 0.164	61 0.087			
82	INNOVATRICS-0	124 0.255	128 0.165	139 0.089	109 0.361	116 0.253	112 0.159			
83	INNOVATRICS-1	133 0.255	137 0.165	135 0.089						
84	INNOVATRICS-2	130 0.237	132 0.142	131 0.079	102 0.310	104 0.209	101 0.126			
85	INNOVATRICS-3	116 0.224	118 0.134	122 0.068	97 0.297	89 0.203	93 0.116			
86	INNOVATRICS-4	94 0.134	92 0.076	93 0.035	80 0.222	86 0.149	77 0.085	123 0.977	117 0.966	94 0.945
87	IYSTEMS-0	72 0.091	69 0.047	75 0.023	69 0.173	82 0.110	69 0.065			
88	IYSTEMS-1	69 0.090	67 0.047	71 0.023						
89	IYSTEMS-2	52 0.081	53 0.035	45 0.015	43 0.126	45 0.080	45 0.046			
90	IYSTEMS-3	52 0.062	49 0.027	36 0.012	34 0.107	39 0.068	38 0.039	29 1.000	31 0.996	21 0.918
91	IDOKMAN-3	49 0.046	41 0.027	50 0.017	53 0.112	46 0.082	57 0.057			
92	IDOKMAN-4	28 0.047	35 0.027	49 0.016	29 0.105	36 0.075	52 0.052	52 0.980	28 0.978	30 0.977
93	MEGVII-3	91 0.109	91 0.058	96 0.025	85 0.116	81 0.067	28 0.034			
94	MEGVII-1	58 0.078	57 0.039	69 0.022	28 0.097	28 0.061	21 0.033			
95	MEGVII-2	51 0.080	59 0.039	68 0.022	27 0.096	21 0.059	21 0.033	25 0.997	9 0.698	7 0.429
96	MICROFOCUS-0	188 0.933	193 0.867	194 0.749	158 0.985	189 0.950	186 0.877			
97	MICROFOCUS-1	189 0.933	192 0.867	193 0.749						
98	MICROFOCUS-2	191 0.934	194 0.870	196 0.758						
99	MICROFOCUS-3	187 0.931	191 0.866	193 0.748	157 0.992	156 0.948	187 0.876			
100	MICROFOCUS-4	197 0.999	198 0.999	198 0.994	188 0.975	182 0.940	184 0.862			
101	MICROFOCUS-5	184 0.856	189 0.736	189 0.598	151 0.928	152 0.865	181 0.748			
102	MICROFOCUS-6	198 0.978	198 0.963	200 0.641	190 0.923	193 0.958	198 0.739			
103	MICROSOFT-0	21 0.044	12 0.022	21 0.010	24 0.115	30 0.071	31 0.040			
104	MICROSOFT-1	23 0.045	24 0.022	29 0.011						
105	MICROSOFT-2	24 0.050	26 0.025	34 0.012						
106	MICROSOFT-3	16 0.030	16 0.014	15 0.006	17 0.091	10 0.056	14 0.028			
107	MICROSOFT-4	13 0.029	13 0.013	10 0.005	13 0.087	15 0.053	13 0.026			
108	MICROSOFT-5	11 0.028	12 0.012	7 0.005	10 0.070	7 0.041	7 0.021	3 0.838	3 0.838	1 0.123
109	MICROSOFT-6	5 0.014	3 0.008	3 0.004	5 0.037	5 0.024	4 0.016	1 0.203	1 0.148	1 0.109
110	NEC-0	83 0.082	74 0.049	83 0.029	57 0.140	57 0.093	61 0.059			
111	NEC-1	79 0.108	66 0.063	102 0.036	78 0.199	78 0.138	78 0.083			
112	NEC-2	2 0.003	1 0.004	1 0.003	2 0.020	2 0.013	1 0.010			
113	NEC-3	1 0.004	2 0.004	2 0.003	1 0.017	1 0.013	2 0.011	1 0.664	1 0.479	2 0.340
114	NEUROTECHNOLOGY-0	197 0.995	196 0.196	145 0.108	170 0.465	132 0.317	133 0.196			
115	NEUROTECHNOLOGY-1	131 0.299	102 0.195	84 0.105						
116	NEUROTECHNOLOGY-2	122 0.299	101 0.195	84 0.105						
117	NEUROTECHNOLOGY-3	172 0.665	111 0.101	110 0.052	93 0.246	98 0.164	93 0.088			
118	NEUROTECHNOLOGY-4	54 0.066	48 0.030	43 0.014	36 0.117	58 0.073	35 0.040			
119	NEUROTECHNOLOGY-5	48 0.056	37 0.025	33 0.012	44 0.130	52 0.074	41 0.042	21 0.996	25 0.982	25 0.948
120	NEUROTECHNOLOGY-6	123 0.255	103 0.124	100 0.051	113 0.418	101 0.208	88 0.103			
121	NEWLAND-2	150 0.441	153 0.296	140 0.157	120 0.466	122 0.335	122 0.213			
122	NOBLIS-1	199 1.000	197 0.992	189 0.419	199 1.000	199 1.000	192 1.000			
123	NOBLIS-2	196 0.997	173 0.490	169 0.309	196 1.000	196 1.000	162 0.565	93 1.000	34 1.000	34 1.000
124	NTECHLAB-0	84 0.083	72 0.047	69 0.023	84 0.162	82 0.105	62 0.061			
125	NTECHLAB-1	76 0.102	78 0.056	79 0.027						
126	NTECHLAB-3	47 0.056	45 0.030	44 0.015	35 0.118	33 0.075	41 0.043			
127	NTECHLAB-4	31 0.043	29 0.024	30 0.012	30 0.105	9 0.065	27 0.036			
128	NTECHLAB-5	27 0.045	30 0.024	31 0.012	26 0.102	35 0.068	25 0.034			
129	NTECHLAB-6	38 0.039	20 0.021	23 0.010	19 0.091	17 0.059	28 0.032	3 0.566	3 0.443	2 0.317
130	QUANTASOFT-1	190 0.640	178 0.494	174 0.335						
131	RANKONE-0	112 0.219	127 0.129	128 0.076	111 0.391	106 0.291	112 0.196			
132	RANKONE-1	101 0.168	107 0.087	99 0.043						
133	RANKONE-2	38 0.120	87 0.073	90 0.042	89 0.261	86 0.190	100 0.126			
134	RANKONE-3	94 0.120	96 0.073	96 0.042	86 0.255	96 0.187	99 0.122			
135	RANKONE-4	129 0.195	124 0.126	125 0.076	119 0.426	123 0.324	129 0.221			
136	RANKONE-5	90 0.092	90 0.056	64 0.021	90 0.173	76 0.119	75 0.074	89 0.998	21 0.994	32 0.988
137	REALNETWORKS-0	119 0.236	121 0.140	122 0.077	104 0.319	105 0.209	105 0.129			
138	REALNETWORKS-1	118 0.236	120 0.140	121 0.077	103 0.319	104 0.209	103 0.129			
139	REALNETWORKS-2	117 0.234	119 0.139	120 0.077	102 0.318	103 0.209	104 0.129			
140	REMARKAI-0	92 0.130	86 0.062	79 0.025	86 0.203	79 0.123	82 0.064			
141	REMARKAI-2	81 0.126	84 0.061	78 0.024	74 0.196	72 0.122	85 0.063	15 0.980	16 0.958	18 0.878
142	SENSETIME-0	9 0.023	10 0.012	14 0.007	8 0.063	8 0.040	9 0.025	129 1.000	28 0.991	46 0.844
143	SENSETIME-1	11 0.025	11 0.012	15 0.007	9 0.064	10 0.041	11 0.025			
144	SHAMAN-0	152 0.474	140 0.370	164 0.259	140 0.621	128 0.509	138 0.375			

Table 20: **Threshold-based accuracy.** *Values are FNIR(N, T, L) with N = 1.6 million with thresholds set to produce FPIR = 0.001, 0.01, and 0.1 in non-mate searches. Columns 3-5 apply to FRVT-2018 mugshots: Columns 6-8 show the corresponding FNIR values for webcam images searched against the FRVT-2018 mugshot gallery. Finally, the three rightmost columns show FNIR for profile view images searched against the FRVT-2018 frontal gallery. Throughout blue superscripts indicate the rank of the algorithm for that column. Caution: The Power-law models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

MISSES BELOW THRESHOLD, T		ENROL MOST RECENT MUGSHOT, N = 1.6M								
FNIR(N, T> 0, R >1)		DATASET: FRVT 2018 MUGSHOTS			DATASET: WEBCAM PROBES			DATASET: PROFILE PROBES		
#	ALGORITHM	FPIR=0.001	FPIR=0.01	FPIR=0.1	FPIR=0.001	FPIR=0.01	FPIR=0.1	FPIR=0.001	FPIR=0.01	FPIR=0.1
145	SHAMAN-1	[116]0.532	[106]0.406	[102]0.274						
146	SHAMAN-2	[124]0.700	[101]0.582	[75]0.424						
147	SHAMAN-3	[131]0.483	[110]0.348	[100]0.225	[129]0.597	[101]0.472	[107]0.317			
148	SHAMAN-4	[105]0.616	[97]0.490	[78]0.344	[126]0.754	[102]0.639	[103]0.480			
149	SHAMAN-6	[56]0.143	[35]0.095	[33]0.060	[84]0.237	[58]0.168	[52]0.108	[3]0.952	[15]0.935	[21]0.905
150	SHAMAN-7	[57]0.144	[36]0.094	[32]0.060	[58]0.240	[58]0.169	[50]0.107			
151	SIAT-0	[32]0.091	[28]0.067	[26]0.022	[31]0.107	[27]0.064	[22]0.035			
152	SIAT-1	[2]0.020	[2]0.009	[2]0.005	[110]0.365	[123]0.348	[128]0.337			
153	SIAT-2	[10]0.024	[7]0.009	[5]0.005	[127]0.478	[129]0.460	[141]0.481			
154	SMILART-0	[166]0.620	[170]0.486	[170]0.322						
155	SMILART-1	[171]0.641	[177]0.505	[178]0.342						
156	SMILART-2	[169]0.623	[174]0.492	[171]0.325						
157	SMILART-4	[191]0.968	[196]0.965	[197]0.984	[186]0.976	[188]0.973	[189]0.973			
158	SMILART-5									
159	SYNESIS-0	[102]0.554	[102]0.378	[128]0.218	[138]0.734	[141]0.598	[140]0.431			
160	SYNESIS-3	[104]0.583	[108]0.444	[108]0.299	[137]0.696	[128]0.514	[128]0.372			
161	TEVIAN-0	[11]0.208	[17]0.114	[12]0.054	[103]0.381	[106]0.227	[106]0.132			
162	TEVIAN-1	[12]0.208	[18]0.114	[13]0.058						
163	TEVIAN-2	[10]0.202	[16]0.114	[11]0.054						
164	TEVIAN-3	[10]0.180	[12]0.098	[8]0.044	[83]0.298	[57]0.198	[56]0.113			
165	TEVIAN-4	[8]0.120	[8]0.066	[8]0.031	[71]0.176	[72]0.115	[62]0.065			
166	TEVIAN-5	[6]0.090	[5]0.047	[6]0.022	[53]0.144	[40]0.089	[30]0.049	[1]0.910	[1]0.681	[1]0.483
167	TIGER-0	[142]0.392	[132]0.262	[142]0.142	[124]0.500	[124]0.366	[121]0.211			
168	TIGER-1				[127]0.580	[132]0.487	[128]0.356			
169	TIGER-2	[6]0.089	[3]0.042	[3]0.018	[52]0.158	[35]0.095	[40]0.049	[3]0.998	[12]0.927	[7]0.503
170	TIGER-3	[6]0.089	[3]0.042	[3]0.018	[51]0.158	[34]0.095	[40]0.048			
171	TONGYITRANS-0	[9]0.077	[9]0.041	[7]0.019	[38]0.112	[38]0.069	[37]0.038			
172	TONGYITRANS-1	[7]0.069	[4]0.035	[4]0.016	[23]0.101	[26]0.062	[24]0.034			
173	TOSHIBA-0	[5]0.065	[6]0.029	[5]0.013	[39]0.118	[35]0.074	[39]0.041	[17]0.988	[21]0.971	[25]0.899
174	TOSHIBA-1	[3]0.062	[1]0.021	[3]0.010	[18]0.092	[16]0.054	[19]0.032			
175	VD-0	[138]0.917	[130]0.828	[136]0.668	[163]0.946	[163]0.871	[132]0.725			
176	VD-1	[115]0.204	[12]0.118	[17]0.059	[34]0.231	[36]0.188	[39]0.106			
177	VIGILANTSOLUTIONS-0	[161]0.528	[158]0.304	[123]0.247	[127]0.695	[130]0.557	[139]0.389			
178	VIGILANTSOLUTIONS-1	[160]0.637	[176]0.502	[177]0.348						
179	VIGILANTSOLUTIONS-2	[125]0.876	[133]0.731	[105]0.489						
180	VIGILANTSOLUTIONS-3	[149]0.410	[151]0.283	[151]0.163	[136]0.680	[137]0.528	[131]0.356			
181	VIGILANTSOLUTIONS-4	[162]0.550	[157]0.424	[156]0.268	[134]0.817	[145]0.709	[145]0.523			
182	VIGILANTSOLUTIONS-5	[145]0.433	[82]0.045	[81]0.023						
183	VIGILANTSOLUTIONS-6	[146]0.428	[82]0.046	[72]0.023						
184	VISIONLABS-3	[26]0.051	[26]0.026	[26]0.013	[28]0.187	[31]0.091	[28]0.051			
185	VISIONLABS-4	[46]0.060	[25]0.026	[21]0.010	[62]0.159	[57]0.097	[42]0.045			
186	VISIONLABS-5	[28]0.052	[23]0.022	[15]0.008	[54]0.147	[48]0.087	[37]0.041			
187	VISIONLABS-6	[15]0.029	[14]0.012	[11]0.005	[15]0.090	[13]0.051	[12]0.025			
188	VISIONLABS-7	[14]0.025	[13]0.012	[9]0.005	[15]0.090	[12]0.051	[10]0.025	[2]0.461	[3]0.322	[3]0.198
189	VOCORD-0	[144]0.399	[121]0.116	[121]0.062	[85]0.235	[92]0.181	[91]0.108			
190	VOCORD-1	[132]0.299	[125]0.116	[122]0.062						
191	VOCORD-2	[139]0.366	[122]0.107	[116]0.057						
192	VOCORD-3	[50]0.126	[75]0.050	[59]0.020	[39]0.155	[53]0.093	[45]0.048			
193	VOCORD-4	[140]0.378	[76]0.054	[78]0.021	[58]0.175	[51]0.093	[44]0.046			
194	VOCORD-5	[110]0.170	[66]0.046	[65]0.019	[47]0.130	[44]0.080	[42]0.043	[12]0.992	[13]0.929	[24]0.787
195	VOCORD-6	[203]1.000	[203]1.000	[203]1.000	[201]1.000	[201]1.000	[201]1.000			
196	YISHENG-0	[141]0.390	[146]0.209	[134]0.086	[154]0.974	[154]0.275	[152]0.146			
197	YISHENG-1	[128]0.348	[135]0.208	[131]0.090	[151]0.808	[133]0.269	[133]0.144			
198	YITU-0	[58]0.060	[53]0.025	[58]0.012	[14]0.090	[17]0.054	[17]0.030			
199	YITU-1	[29]0.047	[28]0.023	[37]0.011						
200	YITU-2	[7]0.020	[8]0.011	[12]0.006	[6]0.089	[6]0.028	[5]0.016			
201	YITU-3	[8]0.021	[8]0.011	[16]0.007	[9]0.052	[7]0.033	[8]0.021			
202	YITU-4	[8]0.012	[2]0.007	[8]0.004	[3]0.027	[2]0.017	[2]0.011	[8]0.902	[11]0.878	[27]0.845
203	YITU-5	[4]0.013	[4]0.007	[9]0.005	[4]0.032	[4]0.023	[6]0.017			

Table 21: **Threshold-based accuracy.** *Values are FNIR(N, T, L) with N = 1.6 million with thresholds set to produce FPIR = 0.001, 0.01, and 0.1 in non-mate searches. Columns 3-5 apply to FRVT-2018 mugshots: Columns 6-8 show the corresponding FNIR values for webcam images searched against the FRVT-2018 mugshot gallery. Finally, the three rightmost columns show FNIR for profile view images searched against the FRVT-2018 frontal gallery. Throughout blue superscripts indicate the rank of the algorithm for that column. Caution: The Power-low models are mostly intended to draw attention to the kind of behavior, not as a model to be used for prediction.*

		INVESTIGATION MODE				IDENTIFICATION MODE				FAILURE TO EXTRACT			
		RANK ONE MISS RATE, FNIR(N, 0, 1)				HIGH T → FPIR = 0.01, FNIR(N, T, 1)				FEATURES			
#	ALGORITHM	N=1.6M FRVT-18	N=1.6M WEBCAM	N=1.6M PROFILE	N=1.1M WILD	N=1.6M FRVT-18	N=1.6M WEBCAM	N=1.6M PROFILE	N=1.1M WILD*	N=1.6M FRVT-18	N=1.6M WEBCAM	N=1.6M PROFILE	N=1.1M WILD
1	3DIVI-0	118 0.034	111 0.086		59 0.071	124 0.160	119 0.302		66 0.095	0.003	0.007		0.013
2	3DIVI-1	119 0.038			62 0.074	123 0.160			65 0.095	0.002			0.013
3	3DIVI-2	124 0.040			64 0.076	126 0.164			68 0.096	0.003			0.013
4	3DIVI-3	122 0.086	128 0.206		57 0.094	122 0.284	133 0.497		58 0.136	0.002	0.005		0.009
5	3DIVI-4	99 0.020	98 0.062			119 0.096	108 0.237			0.002	0.005		
6	3DIVI-5	97 0.020	96 0.062	23 0.894	32 0.052	116 0.095	107 0.234	29 0.987	42 0.069	0.002	0.005	0.442	0.004
7	3DIVI-6	110 0.027	105 0.074		36 0.060	110 0.098	109 0.238		46 0.072	0.002	0.005		0.004
8	ALCHERA-0	92 0.019	86 0.047		77 0.092	94 0.073	85 0.146		78 0.089	0.006	0.014		0.030
9	ALCHERA-1	159 0.987	160 1.000			158 0.999	160 1.000			0.006	0.018		
10	ALCHERA-2	153 0.097	138 0.166		54 0.098	129 0.304	127 0.442		86 0.135	0.001	0.002		0.012
11	ALCHERA-3	72 0.013	69 0.035	16 0.629	30 0.064	96 0.078	92 0.152	27 0.998	41 0.067	0.001	0.002	0.106	0.012
12	ANKE-0	95 0.016	97 0.038	24 0.897	115 0.289	97 0.065	92 0.151	28 0.985		0.000	0.001	0.080	0.001
13	ANKE-1	97 0.016	96 0.038		111 0.284	88 0.065	91 0.151			0.000	0.001		0.001
14	AWARE-0	125 0.064	122 0.138		125 0.588	128 0.129	131 0.253		122 0.587	0.006	0.054		0.143
15	AWARE-1	121 0.059			124 0.580	125 0.127			121 0.580	0.006			0.143
16	AWARE-2	142 0.060				123 0.130				0.006			0.143
17	AWARE-3	116 0.033	112 0.090		123 0.503	101 0.095	100 0.204		103 0.505	0.004	0.005		0.014
18	AWARE-4	107 0.070	126 0.176			129 0.177	125 0.375			0.003	0.005		
19	AWARE-5	117 0.034	96 0.067	38 0.979	122 0.509	109 0.088	98 0.163	33 0.999	119 0.503	0.001	0.002	0.189	0.002
20	AWARE-6	149 0.072	133 0.128			141 0.178	135 0.283			0.001	0.002		
21	AYONIX-0	131 0.452	139 0.685		120 0.400	137 0.725	134 0.892		124 0.586	0.010	0.051		0.068
22	AYONIX-1	127 0.343	135 0.527		117 0.334	138 0.702	130 0.845		122 0.555	0.010	0.051		0.066
23	AYONIX-2	138 0.343	133 0.527			138 0.702	145 0.845			0.010	0.051		
24	CAMVI-1	129 0.227	143 0.327		50 0.148	128 0.549	144 0.643		95 0.196	0.005	0.009		0.058
25	CAMVI-2	130 0.129			51 0.130	132 0.402			90 0.157	0.005			0.058
26	CAMVI-3	140 0.054	113 0.090		84 0.139	89 0.060	46 0.108		57 0.130	0.006	0.013		0.074
27	CAMVI-4	138 0.049	109 0.077	15 0.640	156 1.000	79 0.058	52 0.100	30 0.994	134 1.000	0.000	0.000	0.000	0.000
28	CAMVI-5	136 0.067	117 0.103		155 1.000	99 0.078	75 0.132		136 1.000	0.000	0.000		0.001
29	COGENT-0	74 0.013	80 0.046		78 0.093	57 0.032	62 0.100		72 0.110	0.000	0.000		0.000
30	COGENT-1	78 0.013	91 0.048			51 0.032	61 0.100			0.000	0.000		
31	COGENT-2	20 0.006	39 0.020	25 0.901	21 0.045	19 0.020	22 0.063	20 0.993	25 0.061	0.000	0.000	0.000	0.000
32	COGENT-3	21 0.006	33 0.021		33 0.053	18 0.018	23 0.061		33 0.063	0.000	0.000		0.000
33	COGNITEC-0	113 0.028	95 0.059			109 0.098	96 0.200			0.003	0.002		
34	COGNITEC-1	93 0.014	67 0.034		61 0.074	77 0.056	71 0.136		40 0.072	0.002	0.002		0.025
35	COGNITEC-2	97 0.008	49 0.025	29 0.941	40 0.065	42 0.027	64 0.101	15 0.947	28 0.061	0.002	0.002	0.924	0.021
36	COGNITEC-3	35 0.009	42 0.025		29 0.051	44 0.028	55 0.100		15 0.049	0.004	0.002		0.012
37	DAHUA-0	64 0.012	59 0.026			70 0.047	47 0.083			0.004	0.008		
38	DAHUA-1	67 0.009	46 0.024	14 0.590	4 0.038	38 0.039	40 0.075	10 0.862	9 0.043	0.002	0.002	0.346	0.001
39	DERMALOG-0	161 0.131	159 0.218		83 0.075	159 0.364	139 0.523		69 0.104	0.003	0.002		0.020
40	DERMALOG-1	163 0.156			78 0.089	158 0.405			80 0.131	0.003			0.020
41	DERMALOG-2	162 0.138			69 0.076	161 0.378			73 0.105	0.003			0.020
42	DERMALOG-3	158 0.128	162 0.217			152 0.362	138 0.526			0.002	0.002		
43	DERMALOG-4	157 0.127	161 0.215		74 0.066	157 0.360	138 0.526		63 0.095	0.001	0.002		0.013
44	DERMALOG-5	89 0.017	71 0.037		52 0.066	74 0.045	56 0.096		39 0.066	0.001	0.002		0.013
45	DERMALOG-6	56 0.010	47 0.024	13 0.517	35 0.056	46 0.028	29 0.067	9 0.856	26 0.064	0.003	0.006	0.181	0.014
46	EVERAI-0	69 0.021	69 0.038			53 0.047	60 0.100			0.000	0.000		
47	EVERAI-1	20 0.006	31 0.020		128 0.918	27 0.023	36 0.074		122 0.927	0.000	0.000		0.000
48	EVERAI-2	23 0.006	36 0.022		113 0.302	34 0.025	41 0.076		109 0.308	0.000	0.000		0.001
49	EVERAI-3	13 0.005	28 0.019	8 0.154	5 0.038	17 0.018	21 0.060	6 0.535	11 0.044	0.000	0.000	0.032	0.001
50	EYEDEA-0	154 0.300	149 0.443		93 0.131	148 0.679	146 0.788		103 0.249	0.001	0.003		0.008
51	EYEDEA-1	170 0.198			90 0.072	139 0.480			87 0.131	0.001			0.008
52	EYEDEA-2	173 0.200			97 0.070	172 0.490			78 0.130	0.000			0.005
53	EYEDEA-3	151 0.082	121 0.148		46 0.064	150 0.267	132 0.404		56 0.091	0.001	0.003		0.008
54	GLORY-0	168 0.180	140 0.320			164 0.297	130 0.470			0.011	0.018		
55	GLORY-1	169 0.129	137 0.267		114 0.315	147 0.238	128 0.448		110 0.355	0.011	0.013		0.114
56	GORILLA-0				153 0.994				153 0.994	0.001			0.008
57	GORILLA-1	160 0.063	115 0.095		72 0.057	148 0.248	119 0.314		49 0.076	0.001	0.001		0.007
58	GORILLA-2	100 0.022	77 0.044		19 0.045	114 0.108	90 0.170		42 0.049	0.001	0.001		0.006
59	GORILLA-3	121 0.038	103 0.070		86 0.069	133 0.160	120 0.247		83 0.080	0.001	0.001		0.007
60	HBINNO-0	153 0.275			126 0.335	152 0.632			91 0.411	0.007			0.151
61	HIK-0	107 0.024	80 0.033		97 0.158	93 0.070	66 0.108		89 0.155	0.010	0.004		0.027
62	HIK-1	91 0.017			103 0.162	79 0.067			96 0.166	0.002			0.013
63	HIK-2	90 0.017			53 0.094	92 0.067			66 0.103	0.001			0.008
64	HIK-3	92 0.014	83 0.027			82 0.060	67 0.105			0.000	0.000		
65	HIK-4	90 0.014	80 0.027		82 0.064	80 0.056	48 0.101		45 0.075	0.000	0.000		0.008
66	HIK-5	29 0.007	45 0.017	10 0.371		28 0.022	31 0.048	29 0.994		0.000	0.000	0.000	0.001
67	HIK-6	30 0.007	15 0.017	11 0.371	135 1.000	26 0.022	14 0.052	32 0.997	133 1.000	0.000	0.000	0.000	0.001
68	IDEMIA-0	81 0.011	66 0.034		104 0.166	86 0.062	53 0.156		119 0.288	0.003	0.000		0.002
69	IDEMIA-1	55 0.012			79 0.157	30 0.031			59 0.205	0.003			0.002
70	IDEMIA-2	71 0.013			107 0.198	83 0.032			100 0.242	0.005			0.031
71	IDEMIA-3	54 0.010	61 0.024			32 0.024	42 0.079			0.000	0.000		
72	IDEMIA-4	80 0.009	59 0.022	27 0.934	27 0.051	41 0.024	42 0.079	17 0.962	35 0.064	0.000	0.000	0.041	0.002

Table 22: **Miss rates by dataset:** *At left, rank 1 miss rates relevant to investigations; at right, with threshold set to target FPIR = 0.01 for higher volume, low prior, uses. *For the WILD set, FPIR = 0.1 Yellow indicates most accurate algorithm. Throughout blue superscripts indicate the rank of the algorithm for that column.*

		INVESTIGATION MODE RANK ONE MISS RATE, FNIR(N, 0, 1)				IDENTIFICATION MODE HIGHT → FPIR = 0.01, FNIR(N, T, 1)				FAILURE TO EXTRACT FEATURES			
		N=1.6M	N=1.6M	N=1.6M	N=1.1M	N=1.6M	N=1.6M	N=1.6M	N=1.1M	N=1.6M	N=1.6M	N=1.6M	N=1.1M
#	ALGORITHM	FRVT-18	WEBCAM	PROFILE	WILD	FRVT-18	WEBCAM	PROFILE	WILD*	FRVT-18	WEBCAM	PROFILE	WILD
73	IDEMIA-5	[29]0.011	[12]0.029	[8]0.943	[18]0.044	[48]0.028	[60]0.102	[7]0.968	[27]0.055	0.000	0.000	0.041	0.000
74	IDEMIA-6	[29]0.012	[10]0.022		[7]0.052	[43]0.028	[34]0.161		[29]0.067	0.000	0.000		0.000
75	IMAGUS-0	[120]0.305	[140]0.482		[120]0.222	[116]0.609	[142]0.779		[139]0.311	0.009	0.013		0.049
76	IMAGUS-2	[127]0.222	[139]0.301		[98]0.154	[127]0.566	[143]0.645		[115]0.252	0.004	0.008		0.023
77	IMAGUS-3	[138]0.358	[150]0.513			[133]0.670	[146]0.809			0.004	0.008		
78	INCODE-0	[129]0.051	[116]0.100			[144]0.201	[132]0.304			0.001	0.004		
79	INCODE-1	[93]0.019	[83]0.046		[81]0.052	[115]0.114	[96]0.198		[81]0.062	0.001	0.004		0.009
80	INCODE-2	[88]0.020	[84]0.048		[9]0.059	[117]0.102	[91]0.176		[12]0.045	0.000	0.001		0.001
81	INCODE-3	[89]0.015	[78]0.040		[10]0.039	[105]0.096	[88]0.164		[12]0.044	0.000	0.001		0.001
82	INNOVATRICS-0	[127]0.042	[115]0.076		[115]0.188	[138]0.165	[115]0.258		[115]0.245	0.002	0.008		0.008
83	INNOVATRICS-1	[126]0.042			[100]0.193	[127]0.165			[77]0.221	0.002			0.008
84	INNOVATRICS-2	[130]0.048	[116]0.074			[136]0.142	[105]0.209			0.000	0.001		
85	INNOVATRICS-3	[112]0.023	[86]0.055		[82]0.071	[128]0.134	[95]0.203		[82]0.081	0.000	0.001		0.007
86	INNOVATRICS-4	[94]0.019	[79]0.040	[6]0.940	[85]0.067	[78]0.076	[79]0.149	[5]0.966	[43]0.071	0.000	0.001	0.046	0.013
87	ISYSTEMS-0	[77]0.014	[71]0.038		[103]0.183	[83]0.047	[70]0.110		[96]0.169	0.003	0.015		0.005
88	ISYSTEMS-1	[76]0.014			[105]0.162	[87]0.047			[72]0.139	0.003			0.005
89	ISYSTEMS-2	[94]0.009	[62]0.026		[24]0.049	[56]0.035	[42]0.080		[22]0.051	0.002	0.002		0.009
90	ISYSTEMS-3	[87]0.007	[42]0.023	[3]0.718	[12]0.043	[42]0.027	[34]0.068	[2]0.995	[12]0.044	0.002	0.002	0.142	0.003
91	LOOKMAN-3	[92]0.011	[76]0.038		[121]1.000	[91]0.027	[62]0.082			0.000	0.000		0.000
92	LOOKMAN-4	[86]0.012	[82]0.039	[38]0.932	[108]1.000	[89]0.027	[80]0.095	[21]0.973		0.000	0.000	0.000	0.000
93	MEGVII-0	[51]0.009	[12]0.017		[61]0.061	[81]0.058	[30]0.067		[20]0.094	0.000	0.000		0.005
94	MEGVII-1	[78]0.014	[15]0.017			[97]0.039	[22]0.061			0.002	0.000		
95	MEGVII-2	[97]0.014	[32]0.017	[7]0.295		[128]0.039	[22]0.059	[8]0.698		0.002	0.000	0.033	
96	MICROFOCUS-0	[135]0.397	[161]0.782		[116]0.316	[132]0.867	[138]0.950		[115]0.434	0.005	0.030		0.005
97	MICROFOCUS-1	[136]0.597			[116]0.316	[129]0.867			[115]0.434	0.005			0.005
98	MICROFOCUS-2	[137]0.627			[114]0.342	[134]0.870			[117]0.447	0.005			0.005
99	MICROFOCUS-3	[134]0.595	[160]0.781		[110]0.279	[130]0.866	[136]0.948		[113]0.412	0.001	0.005		0.014
100	MICROFOCUS-4	[133]0.577	[123]0.759			[127]0.999	[135]0.940			0.001	0.005		
101	MICROFOCUS-5	[127]0.426	[126]0.601		[100]0.153	[126]0.736	[130]0.865		[106]0.261	0.001	0.005		0.011
102	MICROFOCUS-6	[120]0.418	[128]0.583		[98]0.146	[125]0.963	[131]0.898		[102]0.246	0.001	0.005		0.011
103	MICROSOFT-0	[23]0.006	[31]0.021		[20]0.065	[23]0.022	[35]0.071		[37]0.065	0.000	0.001		0.019
104	MICROSOFT-1	[21]0.006			[24]0.062	[26]0.022			[29]0.061	0.000			0.019
105	MICROSOFT-2	[28]0.006			[35]0.063	[30]0.026			[36]0.065	0.000			0.019
106	MICROSOFT-3	[4]0.003	[5]0.012			[16]0.014	[18]0.056			0.000	0.001		
107	MICROSOFT-4	[2]0.005	[7]0.012		[5]0.039	[15]0.013	[15]0.053		[9]0.043	0.000	0.001		0.004
108	MICROSOFT-5	[3]0.003	[5]0.011	[1]0.097	[3]0.035	[12]0.012	[7]0.041	[3]0.188	[4]0.041	0.000	0.001	0.049	0.000
109	MICROSOFT-6	[3]0.003	[3]0.011	[2]0.092		[5]0.008	[5]0.024	[1]0.148		0.000	0.001	0.049	
110	NEC-0	[34]0.020	[47]0.041		[124]0.999	[76]0.049	[52]0.093		[122]0.999	0.001	0.002		0.004
111	NEC-1	[106]0.024	[89]0.056			[87]0.063	[76]0.133			0.005	0.003		
112	NEC-2	[1]0.003	[2]0.009		[80]0.093	[3]0.004	[3]0.023		[72]0.102	0.000	0.001		0.025
113	NEC-3	[3]0.003	[3]0.010	[4]0.272	[74]0.088	[2]0.004	[1]0.013	[3]0.479	[58]0.092	0.000	0.001	0.041	0.025
114	NEURO TECHNOLOGY-0	[138]0.050	[133]0.104		[127]1.000	[145]0.196	[120]0.317		[125]1.000	0.004	0.022		0.091
115	NEURO TECHNOLOGY-1	[132]0.047			[120]0.954	[142]0.195			[122]0.953	0.001			0.028
116	NEURO TECHNOLOGY-2	[134]0.047			[121]0.983	[140]0.195			[123]0.983	0.001			0.029
117	NEURO TECHNOLOGY-3	[129]0.025	[98]0.042			[119]0.101	[88]0.164			0.000	0.001		
118	NEURO TECHNOLOGY-4	[91]0.009	[59]0.020		[76]0.090	[89]0.050	[58]0.079		[74]0.122	0.000	0.001		0.007
119	NEURO TECHNOLOGY-5	[31]0.007	[46]0.024	[2]0.884	[16]0.408	[35]0.025	[9]0.074	[2]0.982	[14]0.415	0.000	0.000	0.020	0.000
120	NEURO TECHNOLOGY-6	[35]0.020	[52]0.045		[28]0.050	[123]0.124	[101]0.206		[26]0.065	0.000	0.000		0.001
121	NEWLAND-2	[130]0.081	[119]0.117			[139]0.296	[122]0.535			0.007	0.012		
122	NOBLIS-1	[81]0.251	[43]0.522		[122]0.734	[137]0.992	[102]1.000		[124]0.744	0.000	0.000		0.000
123	NOBLIS-2	[124]0.182	[144]0.292	[3]0.971		[123]0.490	[102]1.000	[4]1.000		0.000	0.000	0.000	
124	NTECHLAB-0	[52]0.012	[28]0.031		[12]0.041	[72]0.047	[66]0.105		[7]0.048	0.000	0.001		0.005
125	NTECHLAB-1	[8]0.014			[30]0.045	[78]0.076			[21]0.049	0.000			0.005
126	NTECHLAB-3	[21]0.008	[29]0.028			[97]0.030	[85]0.075			0.000	0.000		
127	NTECHLAB-4	[50]0.007	[42]0.019		[14]0.043	[32]0.024	[25]0.065		[15]0.048	0.000	0.000		0.003
128	NTECHLAB-5	[23]0.006	[24]0.018		[7]0.038	[20]0.024	[26]0.063		[8]0.042	0.000	0.000		0.000
129	NTECHLAB-6	[24]0.006	[17]0.017	[3]0.208	[9]0.038	[23]0.021	[16]0.059	[4]0.443	[9]0.042	0.000	0.000	0.040	0.000
130	QUANTASOFT-1	[126]0.220	[132]0.727		[126]0.620	[118]0.484			[123]0.780	0.000	0.000		0.000
131	RANKONE-0	[133]0.045	[120]0.117		[29]0.114	[122]0.129	[118]0.291		[75]0.161	0.000	0.000		0.000
132	RANKONE-1	[103]0.025			[88]0.097	[103]0.087			[67]0.102	0.000			0.000
133	RANKONE-2	[103]0.022	[102]0.091			[97]0.093	[95]0.150			0.000	0.000		
134	RANKONE-3	[13]0.022	[9]0.068		[49]0.078	[98]0.073	[93]0.157		[57]0.095	0.000	0.000		0.000
135	RANKONE-4	[132]0.044	[123]0.141		[82]0.094	[128]0.126	[123]0.324		[75]0.126	0.000	0.000		0.000
136	RANKONE-5	[28]0.012	[35]0.041	[36]0.981	[29]0.061	[96]0.036	[72]0.119	[26]0.994	[4]0.068	0.000	0.000	0.489	0.000
137	REALNETWORKS-0	[131]0.042	[110]0.078		[62]0.076	[131]0.140	[104]0.209		[58]0.084	0.001	0.000		0.004
138	REALNETWORKS-1	[131]0.043	[110]0.078			[131]0.140	[105]0.209			0.001	0.000		
139	REALNETWORKS-2	[125]0.042	[110]0.078		[132]0.992	[129]0.139	[110]0.209		[130]0.992	0.001	0.000		0.000
140	REMARKAI-0	[91]0.011	[37]0.030			[88]0.062	[78]0.128			0.000	0.001		
141	REMARKAI-2	[36]0.010	[25]0.029	[13]0.708	[59]0.046	[84]0.061	[75]0.122	[30]0.958	[35]0.062	0.000	0.001	0.017	0.000
142	SENSETIME-0	[10]0.005	[13]0.016	[12]0.446		[11]0.012	[9]0.040	[30]0.971		0.004	0.000	0.042	0.000
143	SENSETIME-1	[7]0.005	[12]0.016		[3]0.038	[11]0.012	[10]0.041		[1]0.796	0.004	0.000		0.000
144	SHAMAN-0	[108]0.171	[136]0.262		[30]0.115	[101]0.270	[134]0.507		[98]0.146	0.020	0.011		0.043

Table 23: **Miss rates by dataset:** *At left, rank 1 miss rates relevant to investigations; at right, with threshold set to target FPIR = 0.01 for higher volume, low prior, uses. * For the WILD set, FPIR = 0.1 Yellow indicates most accurate algorithm. Throughout blue superscripts indicate the rank of the algorithm for that column.*

		INVESTIGATION MODE				IDENTIFICATION MODE				FAILURE TO ENROLL			
		RANK ONE MISS RATE, FNIR(N, 0, 1)				HIGHT → FPIR = 0.01, FNIR(N, T, 1)				FEATURES			
		N=1.6M	N=1.6M	N=1.6M	N=1.1M	N=1.6M	N=1.6M	N=1.6M	N=1.1M	N=1.6M	N=1.6M	N=1.6M	N=1.1M
#	ALGORITHM	FRVT-18	WEBCAM	PROFILE	WILD	FRVT-18	WEBCAM	PROFILE	WILD*	FRVT-18	WEBCAM	PROFILE	WILD
145	SHAMAN-1	166 0.172			88 0.113	168 0.406			88 0.158	0.020			0.043
146	SHAMAN-2	182 0.262			93 0.132	180 0.582			90 0.201	0.020			0.043
147	SHAMAN-3	185 0.127	167 0.172		88 0.109	156 0.348	131 0.472		82 0.132	0.020	0.011		0.043
148	SHAMAN-4	178 0.224	169 0.219			174 0.490	142 0.620			0.020	0.011		
149	SHAMAN-6	125 0.042	71 0.058	36 0.910		125 0.095	98 0.168	34 0.935		0.020	0.011	0.869	
150	SHAMAN-7	133 0.042	80 0.057		50 0.078	104 0.094	85 0.169		80 0.079	0.020	0.010		0.029
151	SIAT-0	46 0.010	32 0.021		71 0.078	88 0.047	75 0.064		124 0.250	0.000	0.000		0.008
152	SIAT-1	31 0.004	142 0.333		31 0.040	6 0.009	123 0.248		3 0.041	0.000	0.000		0.003
153	SIAT-2	71 0.004	148 0.446			7 0.009	126 0.460			0.000	0.000		
154	SMILART-0	170 0.198	147 0.325		105 1.000	160 0.485			120 1.000	0.008			0.121
155	SMILART-1	155 0.219			134 1.000	157 0.505			103 1.000	0.021			0.006
156	SMILART-2	171 0.196			145 1.000	154 0.492			144 1.000	0.000			0.048
157	SMILART-4	176 0.965	165 0.974		135 0.834	170 0.965	158 0.973		126 0.833	0.011	0.013		0.089
158	SMILART-5									0.011	0.013		
159	SYNESIS-0	164 0.162	138 0.361			152 0.378	141 0.598			0.002	0.009		0.081
160	SYNESIS-3	160 0.172	134 0.235			148 0.444	135 0.524			0.006	0.015		0.042
161	TEVIAN-0	101 0.022	57 0.066		51 0.054	117 0.114	108 0.227		94 0.074	0.002	0.005		0.007
162	TEV(AI)-1	105 0.022			48 0.062	118 0.114			25 0.078	0.002			0.007
163	TEVIAN-2	106 0.022			58 0.093	116 0.114			72 0.118	0.002			0.008
164	TEVIAN-3	85 0.017	59 0.052			105 0.098	97 0.198			0.001	0.002		
165	TEVIAN-4	77 0.013	48 0.038		28 0.050	80 0.066	71 0.115		53 0.063	0.001	0.002		0.005
166	TEVIAN-5	45 0.009	54 0.028	9 0.329		76 0.047	45 0.089	7 0.661		0.001	0.002	0.116	
167	TIGER-0	142 0.064	115 0.095		128 1.000	149 0.263	128 0.366		125 1.000	0.000	0.000		0.005
168	TIGER-1		144 0.351				132 0.487			0.000	0.000		
169	TIGER-2	36 0.008	61 0.023	7 0.356		61 0.042	65 0.095	13 0.027		0.000	0.000	0.056	
170	TIGER-3	42 0.008	40 0.023			72 0.042	54 0.095			0.000	0.000		
171	TONGYITRANS-0	56 0.010	38 0.022			81 0.041	32 0.069			0.003	0.001		
172	TONGYITRANS-1	50 0.010	37 0.022		39 0.112	74 0.036	28 0.062		83 0.134	0.003	0.001		0.009
173	TOSHIBA-0	30 0.007	36 0.012	15 0.689		49 0.029	35 0.074	21 0.971		0.000	0.000	0.070	0.002
174	TOSHIBA-1	38 0.007	35 0.022			23 0.021	16 0.054			0.000	0.000		
175	VD-0	109 0.475	154 0.551		108 0.217	130 0.828	153 0.871		111 0.364	0.011	0.018		0.026
176	VD-1	135 0.030	87 0.055			121 0.118	95 0.185			0.005	0.001		0.017
177	VIGILANTSOLUTIONS-0	154 0.125	130 0.212		67 0.075	153 0.304	140 0.587		87 0.152	0.000	0.001		0.003
178	VIGILANTSOLUTIONS-1	174 0.204			86 0.108	150 0.502			89 0.209	0.000			0.008
179	VIGILANTSOLUTIONS-2	124 0.239			47 0.064	135 0.751			76 0.129	0.000			0.003
180	VIGILANTSOLUTIONS-3	148 0.072	128 0.151		91 0.065	121 0.283	137 0.526		79 0.131	0.000	0.001		0.003
181	VIGILANTSOLUTIONS-4	126 0.127	118 0.244			127 0.424	145 0.709			0.000	0.001		
182	VIGILANTSOLUTIONS-5	67 0.012				58 0.045				0.000	0.001		
183	VIGILANTSOLUTIONS-6	79 0.013				53 0.046				0.000	0.001		
184	VISIONLABS-3	48 0.009	60 0.030		28 0.051	27 0.026	50 0.091		18 0.046	0.002	0.003		0.014
185	VISIONLABS-4	18 0.004	35 0.020			38 0.026	57 0.097			0.001	0.001		
186	VISIONLABS-5	13 0.004	23 0.019		18 0.043	25 0.022	46 0.087		16 0.046	0.001	0.001		0.006
187	VISIONLABS-6	7 0.003	15 0.015			14 0.012	15 0.051			0.001	0.001		
188	VISIONLABS-7	5 0.003	10 0.015	5 0.130	1 0.033	18 0.012	12 0.051	5 0.322	1 0.035	0.001	0.001	0.051	0.001
189	VOCORD-0	138 0.040	100 0.068			120 0.118	92 0.191			0.015	0.025		0.019
190	VOCORD-1	132 0.040				119 0.116				0.015			0.018
191	VOCORD-2	130 0.038				113 0.107				0.015			0.015
192	VOCORD-3	42 0.008	48 0.024		38 0.057	35 0.050	83 0.093		81 0.062	0.001	0.011		0.006
193	VOCORD-4	58 0.010	51 0.021			76 0.054	84 0.093			0.000	0.000		
194	VOCORD-5	45 0.009	43 0.023	20 0.739	37 0.044	66 0.046	44 0.080	13 0.929	44 0.045	0.001	0.009	0.554	0.003
195	VOCORD-6	208 1.000	202 1.000			208 1.000	201 1.000			0.001	0.009		
196	YISHENG-0	131 0.027	93 0.060		84 0.067	146 0.209	125 0.275		98 0.100	0.002	0.005		0.014
197	YISHENG-1	134 0.029	84 0.060		82 0.061	145 0.208	119 0.269		81 0.087	0.002	0.005		0.014
198	YITU-0	28 0.007	29 0.020		72 0.036	32 0.025	70 0.054		86 0.094	0.003	0.001		0.026
199	YITU-1	35 0.007			72 0.036	28 0.023			87 0.092	0.003			0.026
200	YITU-2	24 0.004	4 0.010		22 0.046	8 0.011	8 0.028		24 0.051	0.000	0.000		0.000
201	YITU-3	26 0.005	16 0.016			9 0.011	9 0.033			0.003	0.001		
202	YITU-4	9 0.004	3 0.008	20 0.831	16 0.044	5 0.007	5 0.017	11 0.875	19 0.047	0.000	0.000	0.900	0.006
203	YITU-5	16 0.005	9 0.014			4 0.007	4 0.023			0.003	0.001		

Table 24: **Miss rates by dataset:** *At left, rank 1 miss rates relevant to investigations; at right, with threshold set to target FPIR = 0.01 for higher volume, low prior, uses. *For the WILD set, FPIR = 0.1 Yellow indicates most accurate algorithm. Throughout blue superscripts indicate the rank of the algorithm for that column.*

MISSES OUTSIDE RANK R		MUGSHOT SEARCHES, N = 1.6M IDENTITIES							
FNIR(N, T, R)		INVESTIGATION MODE, T = 0				IDENTIFICATION MODE, T > 0 FOR FPIR = 0.001			
		PROPORTION MATED SEARCHES				PROPORTION MATED SEARCHES			
		WITHOUT THE MATE AT RANK 1		WITH NO MATE AT RANK 1	WITH K-TH MATE NOT IN TOP K	WITH THE MATE BELOW THRESHOLD		WITHOUT ANY MATE ABOVE THRESH	WITHOUT ALL MATES ABOVE THRESH
DALLERY		RECENT	CONSOLIDATED	UNCONSOLIDATED		RECENT	CONSOLIDATED	UNCONSOLIDATED	
1	3DIVI-5	29 0.0202	43 0.0133	46 0.0133	48 0.0449	35 0.1691	30 0.1339	28 0.1339	30 0.3186
2	3DIVI-6	37 0.0245	46 0.0186	51 0.0172	47 0.0410	31 0.1705	27 0.1345	27 0.1350	28 0.3160
3	ALCHERA-2	54 0.0978	57 0.0914	65 0.0734	53 0.1876	64 0.4899	59 0.3756	53 0.4418	60 0.6820
4	ANKE-0	42 0.0158	39 0.0100	48 0.0100	43 0.0338	41 0.1199	38 0.0989	32 0.0999	42 0.2558
5	ANKE-1	36 0.0158	40 0.0101	44 0.0101	34 0.0337	42 0.1219	40 0.1001	41 0.1001	40 0.2581
6	AWARE-5	14 0.0337	47 0.0208	58 0.0230	36 0.0740	60 0.3729	59 0.2984	61 0.3777	61 0.6534
7	AWARE-6	28 0.0722	36 0.0538	60 0.0538	31 0.1551	58 0.2779	57 0.2419	58 0.2465	58 0.5140
8	AYONIX-1	70 0.3432	62 0.3364	67 0.2841	58 0.4764	68 0.8249	61 0.8583	60 0.7936	66 0.9037
9	AYONIX-2	69 0.3432	67 0.2606	66 0.2841	57 0.4763	67 0.8246	69 0.8068	66 0.7933	65 0.9036
10	CAMVI-4	30 0.0490	54 0.0326	68 0.0469	45 0.0475	31 0.0741	37 0.0505	34 0.0661	34 0.1105
11	CAMVI-5	21 0.0673	55 0.0458	62 0.0633	56 0.0638	42 0.1020	36 0.0727	40 0.0922	37 0.1513
12	COGENT-2	14 0.0062	12 0.0027	32 0.0027	12 0.0086	39 0.0476	42 0.0299	39 0.0391	20 0.1275
13	COGENT-3	15 0.0064	18 0.0037	35 0.0029	14 0.0091	21 0.0515	17 0.0341	27 0.0450	28 0.1448
14	COGNITEC-2	25 0.0083	22 0.0044	26 0.0043	24 0.0145	28 0.0560	23 0.0401	25 0.0400	25 0.1342
15	COGNITEC-3	26 0.0088	24 0.0048	28 0.0048	25 0.0148	29 0.0555	21 0.0397	23 0.0397	26 0.1322
16	DAHUA-0	38 0.0115	32 0.0070	38 0.0072	34 0.0204	36 0.0891	32 0.0624	35 0.0691	30 0.1967
17	DAHUA-1	27 0.0089	25 0.0049	27 0.0052	27 0.0173	38 0.0755	28 0.0521	30 0.0577	32 0.1738
18	DERMALOG-5	39 0.0171	41 0.0113	48 0.0139	41 0.0254	39 0.0909	39 0.0649	38 0.0767	39 0.2072
19	DERMALOG-6	30 0.0102	28 0.0060	30 0.0061	15 0.0119	22 0.0542	20 0.0383	31 0.0416	21 0.1280
20	EVERAI-2	12 0.0058	13 0.0029	28 0.0032	13 0.0099	27 0.0526	19 0.0370	22 0.0410	22 0.1312
21	EVERAI-3	8 0.0047	13 0.0023	21 0.0024	11 0.0073	31 0.0397	11 0.0256	21 0.0285	11 0.0978
22	GORILLA-2	31 0.0220	42 0.0157	50 0.0158	55 0.0570	52 0.1902	49 0.1379	57 0.1537	57 0.3589
23	GORILLA-3	53 0.0384	49 0.0245	58 0.0283	50 0.1032	58 0.3260	58 0.2730	59 0.3043	59 0.5786
24	HIK-5	17 0.0087		36 0.0037	33 0.0140	17 0.0467		26 0.0564	18 0.1238
25	HIK-6	18 0.0087	16 0.0054	37 0.0037	32 0.0140	20 0.0500	16 0.0324	24 0.0392	23 0.1310
26	IDEMIA-5	35 0.0107	29 0.0062	39 0.0064	39 0.0192	26 0.0465	18 0.0319	17 0.0348	17 0.1125
27	IDEMIA-6	37 0.0122	30 0.0071	36 0.0076	37 0.0188	24 0.0458	14 0.0316	28 0.0342	12 0.1082
28	INCODE-2	41 0.0203	34 0.0120	40 0.0137	30 0.0480	32 0.1861	35 0.1360	36 0.1507	31 0.3500
29	INCODE-3	34 0.0153	26 0.0088	45 0.0103	46 0.0368	30 0.1703	33 0.1227	33 0.1388	35 0.3290
30	INNOVATRICS-4	33 0.0149	23 0.0081	42 0.0081	43 0.0293	40 0.1340	31 0.0928	42 0.0927	41 0.2479
31	ISYSTEMS-5	23 0.0075	20 0.0040	22 0.0041	16 0.0106	25 0.0620	23 0.0402	29 0.0500	27 0.1519
32	LOOKMAN-5	24 0.0114	27 0.0089	31 0.0067	17 0.0109	35 0.0463	26 0.0425	37 0.0338	15 0.1015
33	LOOKMAN-6	26 0.0117	38 0.0091	35 0.0072	21 0.0134	19 0.0472	24 0.0417	35 0.0346	14 0.1036
34	MEGVII-1	41 0.0137		41 0.0096	38 0.0231	33 0.0746		31 0.0577	33 0.1688
35	MEGVII-2	32 0.0137		43 0.0097	35 0.0236	34 0.0756		33 0.0623	36 0.1810
36	MICROFOCUS-5	71 0.4257	63 0.3701	69 0.3701	69 0.5522	69 0.8361	67 0.9835	67 0.8139	67 0.9189
37	MICROFOCUS-6	72 0.4283	64 0.3732	70 0.3732	70 0.5566	71 0.8780	68 0.8195	68 0.8196	68 0.9215
38	MICROSOFT-5	3 0.0033	3 0.0015	5 0.0015	10 0.0062	5 0.0229	7 0.0171	7 0.0193	7 0.0755
39	MICROSOFT-6	6 0.0033	5 0.0014	7 0.0015	9 0.0060	8 0.0141	8 0.0080	10 0.0213	10 0.0772
40	NEC-2	1 0.0028	2 0.0011	1 0.0008	1 0.0019	2 0.0047	2 0.0024	2 0.0021	2 0.0060
41	NEC-3	5 0.0031	4 0.0013	2 0.0010	5 0.0019	1 0.0044	1 0.0021	3 0.0022	1 0.0060
42	NEUROTECHNOLOGY-5	19 0.0068	21 0.0042	29 0.0032	14 0.0094	26 0.0561	29 0.0527	26 0.0438	26 0.1364
43	NEUROTECHNOLOGY-6	28 0.0201	48 0.0153	47 0.0142	52 0.0554	37 0.2555	51 0.2695	52 0.2125	37 0.4458
44	NEWLAND-2	52 0.0811		61 0.0559	54 0.1562	62 0.4105		56 0.3750	62 0.6252
45	NOBLIS-1	48 0.2512	61 0.2049	64 0.2032	53 0.3681	63 0.9996	63 0.9998	72 0.9994	70 0.9997
46	NOBLIS-2	50 0.1816	59 0.1565	66 0.2517	60 0.3944	72 0.9974	65 0.9969	71 0.9967	71 0.9987
47	NTECHLAB-5	16 0.0064	19 0.0039	31 0.0039	31 0.0179	15 0.0448	18 0.0347	16 0.0347	19 0.1235
48	NTECHLAB-6	13 0.0059	15 0.0034	17 0.0034	26 0.0154	13 0.0891	13 0.0301	15 0.0301	13 0.1088
49	QUANTASOFT-1	67 0.2198	66 0.9857	71 0.9426	71 0.9502	66 0.8399	66 0.9915	69 0.9640	69 0.9801
50	RANKONE-4	49 0.0441	52 0.0318	59 0.0318	59 0.0945	54 0.1951	50 0.1595	53 0.1545	56 0.3690
51	RANKONE-5	45 0.0120	44 0.0072	49 0.0072	49 0.0237	45 0.0617	46 0.0447	49 0.0447	49 0.1404
52	REALNETWORKS-2	46 0.0418	53 0.0320	54 0.0268	48 0.0908	46 0.2341	48 0.2049	46 0.1775	48 0.3949
53	REMARKAI-0	43 0.0109	31 0.0065	33 0.0065	40 0.0238	44 0.1201	41 0.1020	43 0.1020	42 0.2691
54	REMARKAI-2	51 0.0105	50 0.0062	34 0.0062	37 0.0235	43 0.1264	40 0.0991	45 0.0991	45 0.2615
55	SENSETIME-0	9 0.0048	9 0.0018	9 0.0018	4 0.0037	9 0.0234	9 0.0165	9 0.0168	9 0.0609
56	SENSETIME-1	10 0.0048	8 0.0018	8 0.0018	7 0.0041	7 0.0245	8 0.0175	8 0.0177	8 0.0628
57	SHAMAN-6	58 0.0424	51 0.0312	57 0.0312	52 0.0542	48 0.1432	43 0.1109	50 0.1109	46 0.2629
58	SHAMAN-7	57 0.0422	50 0.0310	56 0.0310	51 0.0529	47 0.1426	44 0.1112	47 0.1112	45 0.2624
59	SMILART-4	73 0.9649	68 0.9531	72 0.9722	72 0.9738	70 0.9683	70 0.9569	70 0.9740	67 0.9781
60	SYNESIS-5	65 0.1721	58 0.1350	64 0.1350	64 0.2571	65 0.5832	66 0.5296	64 0.5295	64 0.7489
61	TEVIAN-5	20 0.0092	10 0.0055	23 0.0055	22 0.0218	56 0.0599	54 0.0667	48 0.0770	44 0.2079
62	TIGER-2	24 0.0075	23 0.0044	25 0.0044	28 0.0177	50 0.0888	55 0.0898	54 0.0898	54 0.2016
63	TIGER-3	22 0.0075		24 0.0044	29 0.0177	55 0.0888		55 0.0898	53 0.2015
64	TOSHIBA-0	31 0.0088	14 0.0033	55 0.0033	18 0.0110	53 0.0642	31 0.0529	52 0.0529	52 0.1599
65	TOSHIBA-1	44 0.0091	17 0.0035	52 0.0035	20 0.0120	52 0.0618	52 0.0596	51 0.0595	52 0.1819
66	YD-1	55 0.0502	45 0.0221	53 0.0221	44 0.0560	57 0.2036	53 0.1464	44 0.1458	43 0.3667
67	VIGILANTSOLUTIONS-5	40 0.0118				67 0.4827			
68	VIGILANTSOLUTIONS-6	47 0.0125		63 0.0077	42 0.0258	61 0.4260		62 0.4165	63 0.6577
69	VISIONLABS-5	4 0.0033	7 0.0015	4 0.0015	8 0.0040	10 0.0289	10 0.0185	5 0.0201	5 0.0737
70	VISIONLABS-7	4 0.0033	6 0.0014	4 0.0014	8 0.0040	8 0.0289	8 0.0185	8 0.0201	6 0.0737
71	VOCORD-5	22 0.0092	35 0.0057	39 0.0054	27 0.0182	49 0.1697	45 0.1076	45 0.1717	50 0.3775
72	VTEU-4	7 0.0037	11 0.0011	3 0.0012	3 0.0035	3 0.0123	3 0.0074	3 0.0090	3 0.0237
73	VTEU-5	11 0.0048	10 0.0020	10 0.0020	10 0.0041	4 0.0128	4 0.0076	4 0.0098	4 0.0260

Table 25: **Comparing enrollment styles for the FRVT 2018 mugshot sets.** *Consolidated refers to enrollment of all lifetime images in one template Unconsolidated refers to enrollment of those images separately under different identifiers. Columns 3 - 6 values are FNIR at rank 1 and with T = 0. Columns 7 - 10 values are high threshold FNIR. Throughout, blue superscripts indicate the rank of the algorithm for that column, and the best three values are highlighted in yellow and green.*

2019/09/11 17:24:52	FNIR(N, R, T) = FPIR(N, T) =	False neg. identification rate False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification

Figure 19: **[Mugshot Dataset] Error rate reductions in 2018.** *For each* FRVT*2018 participant, the plot shows accuracy gains between Phase 1 (Feb 2018), Phase 2 (Jun 2018) and Phase 3 (Nov 2018) according to two metrics: rank one miss rate, FNIR(N, 1, 0), and high threshold, FNIR(N, L, T), with T set to achieve FPIR = 0.003. The text "Red=" gives the best reduction multiplier for the given metric on the recent enrollment strategy - a smaller value is better.*

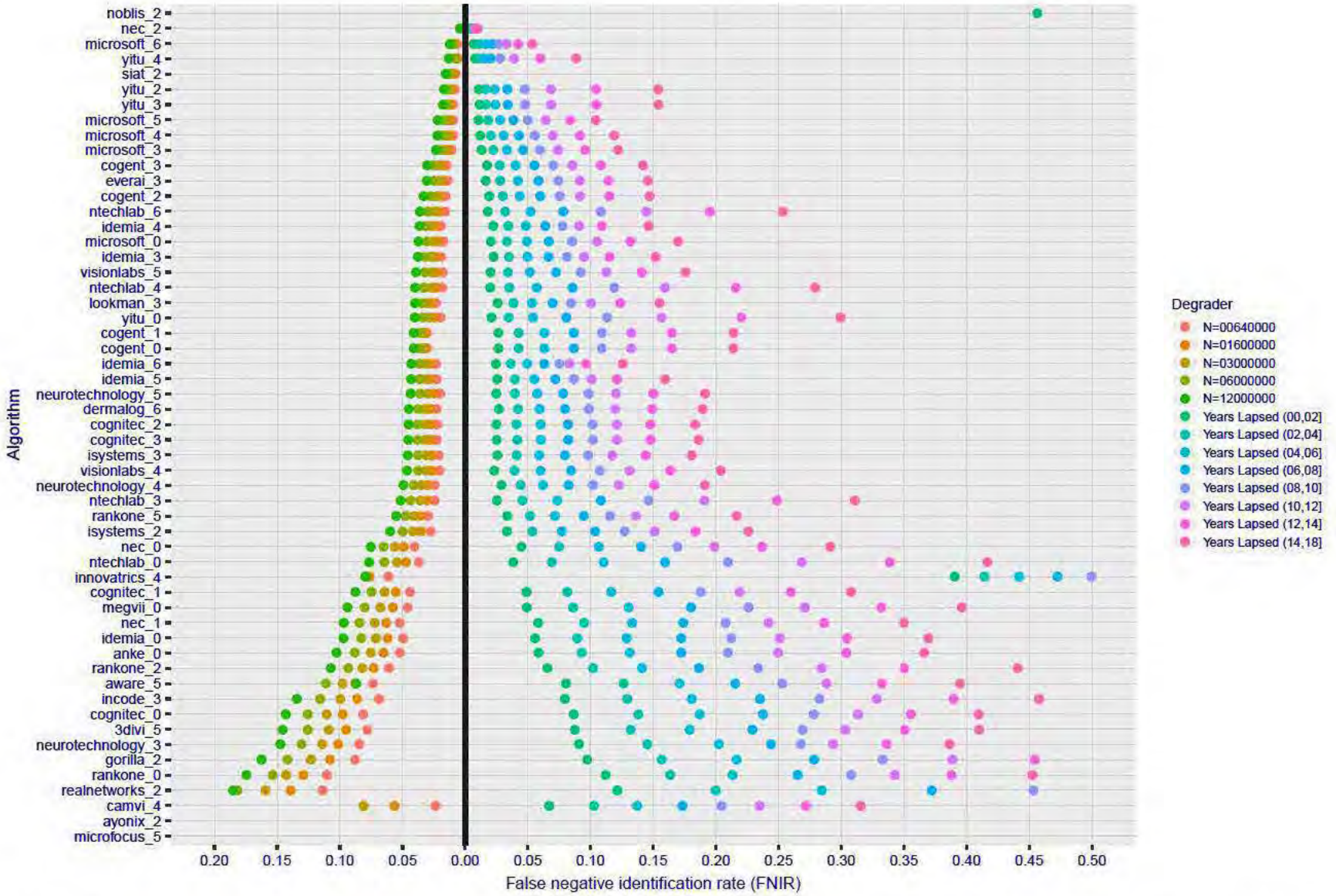


Figure 20: **[FRVT-2018 Mugshot Ageing Dataset] Contrast of ageing and population size dependency.** *The Figure shows, at left, the dependence FNIR(N) for the FRVT-2018, as tabulated in Table 12. At right, is FNIR(N = 3 000 000, ΔT) from Figure 62. Ageing miss rates are computed over all searches binned by number of years between search and initial enrollment. In all cases, FPIR = 0.01.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

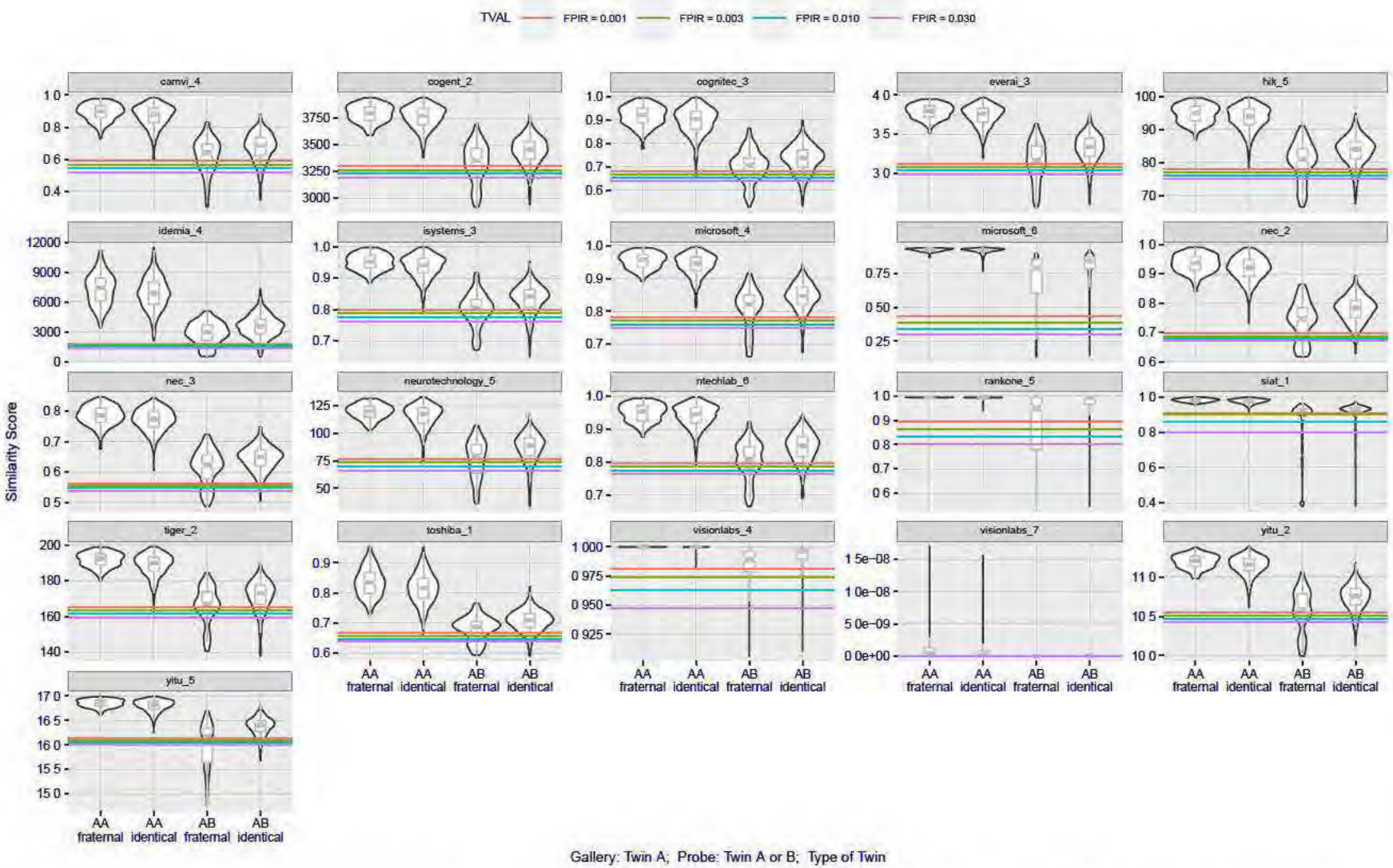


Figure 21: **[Twins Dataset] High scores from twins.** *The Figure shows native similarity scores from searches into a dataset of N = 640 000 background mugshot images plus 104 portrait images, one from each of one of a pair of twins. Two distributions of scores are plotted for each of monozygotic (identical) and dizygotic (fraternal) twins. The first distribution ("AA") shows the mate score from Twin A against their own enrollment. The second ("AB") shows scores from searches of Twin B against the Twin A enrollment: As these are non-mate scores they should be below the various thresholds shown as horizontal lines. That they usually are not is an indication that twins produce very high non-mate scores. Note in theory half of dizygotic (fraternal) twins are different sex. In the sample used here some fraternal twins are correctly rejected.*

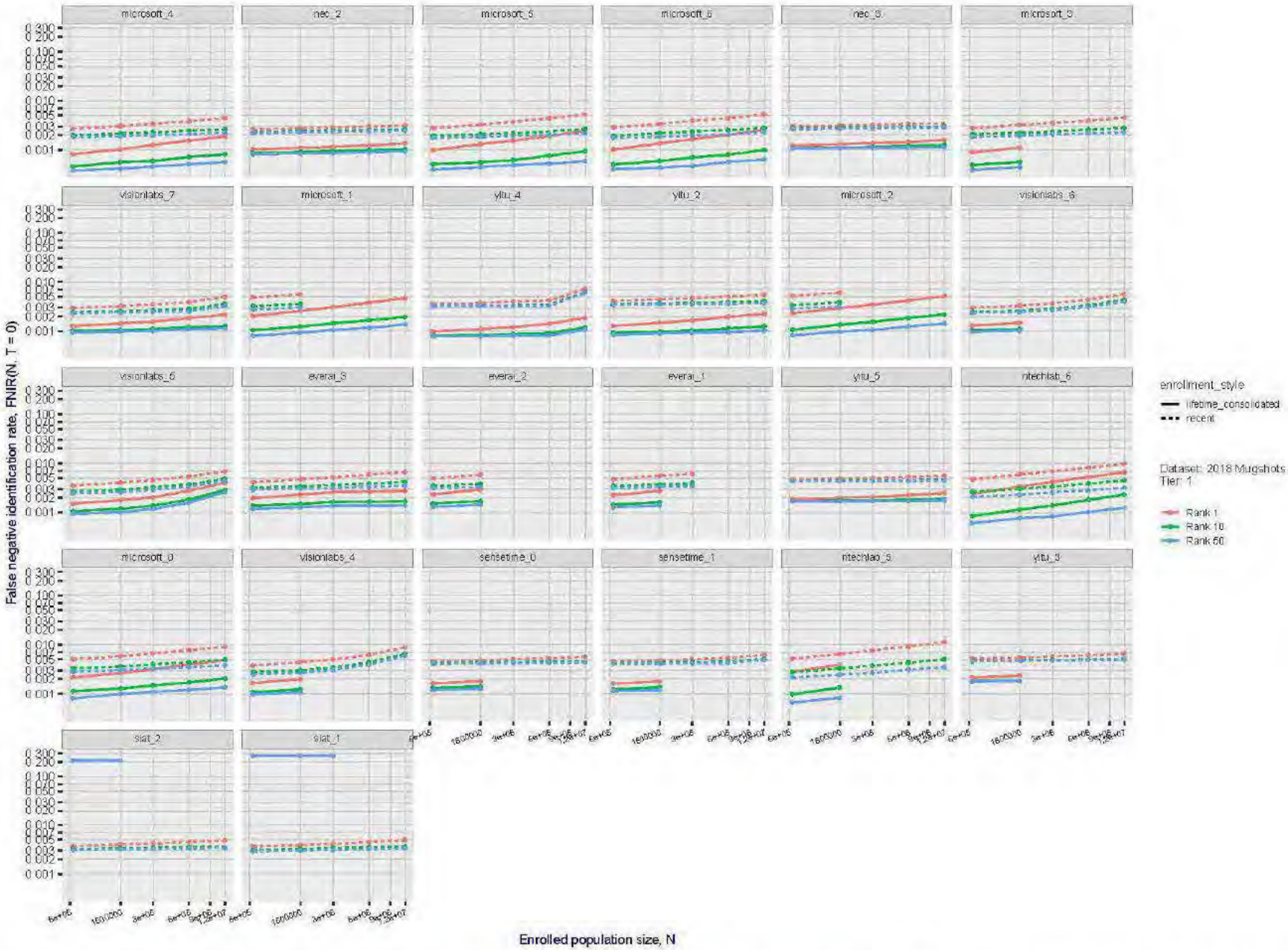
2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

Appendices

Appendix A Accuracy on large-population FRVT 2018 mugshots

2019/09/11 17:24:52	FNIR(N, R, T) = FPIR(N, T) =	False neg. identification rate False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification



2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

Figure 22: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

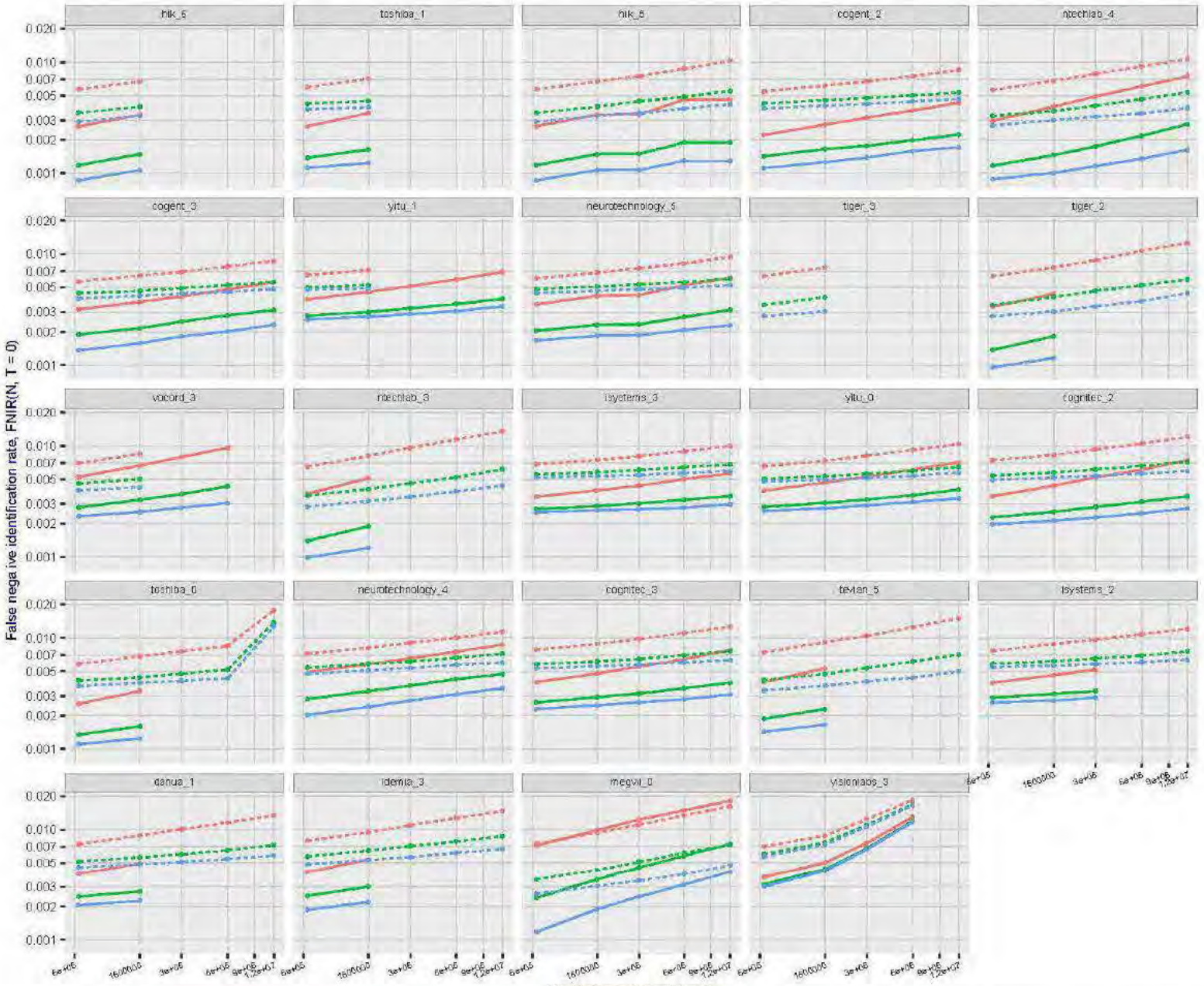


Figure 23: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

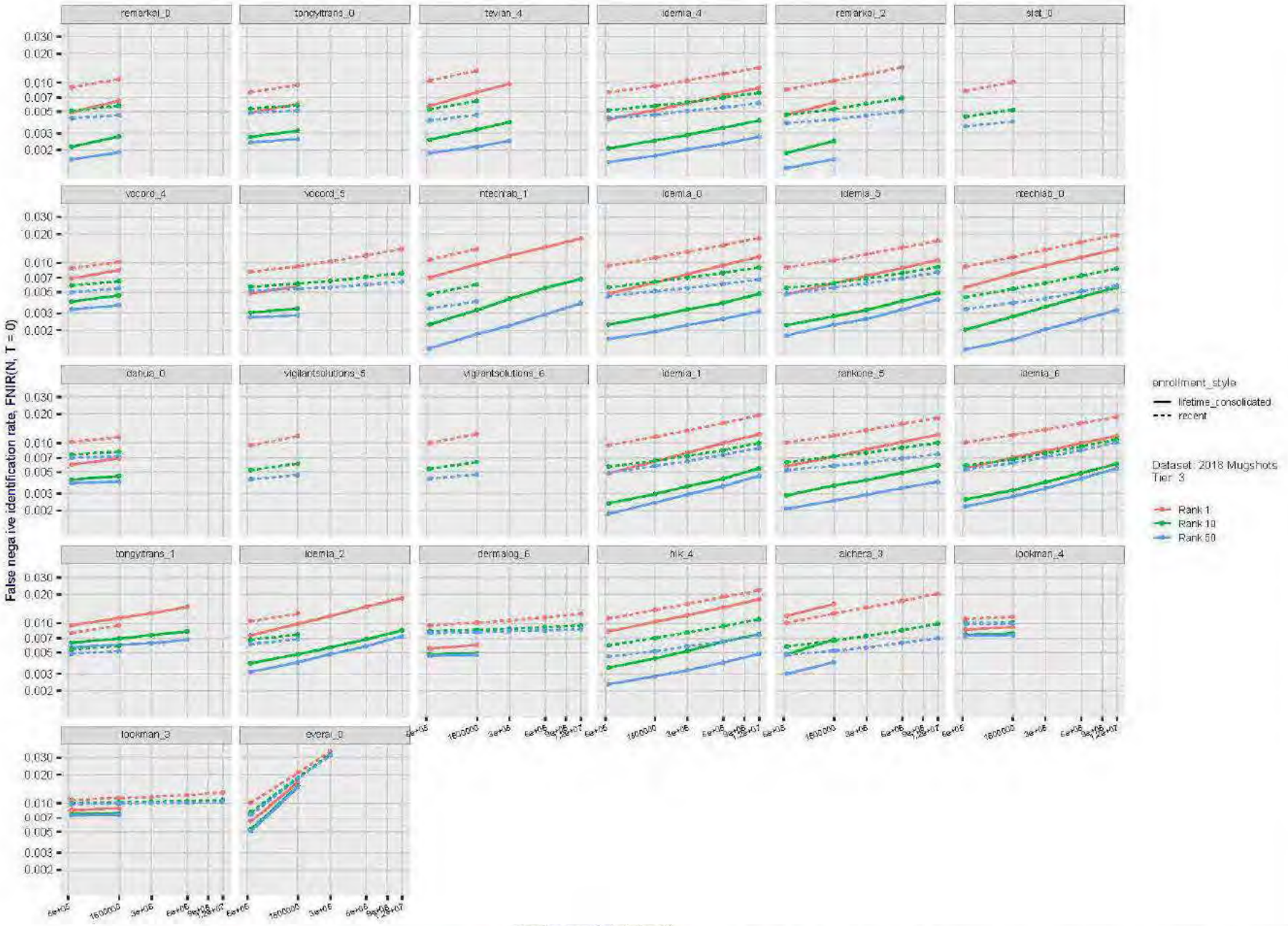


Figure 24: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

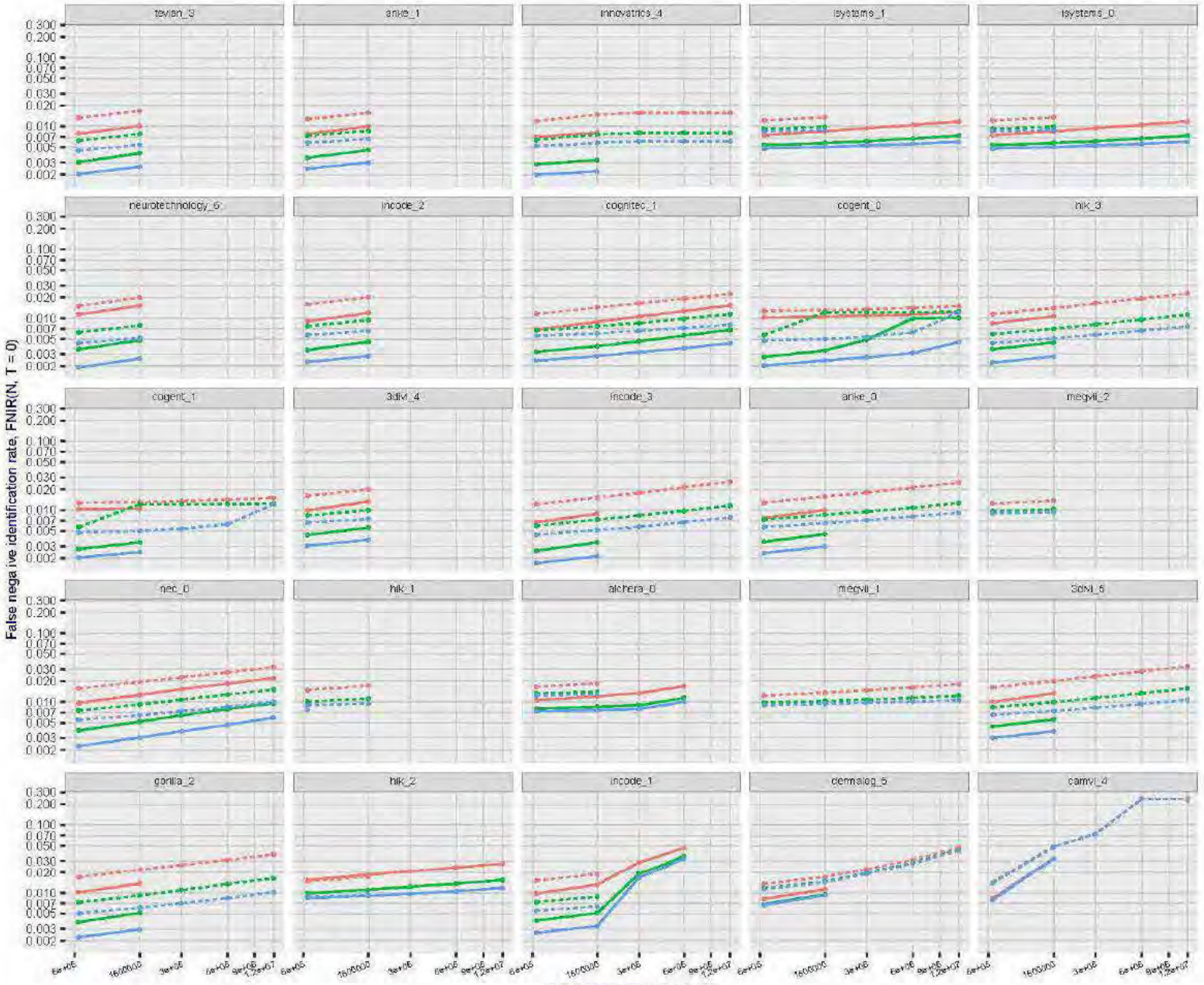


Figure 25: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

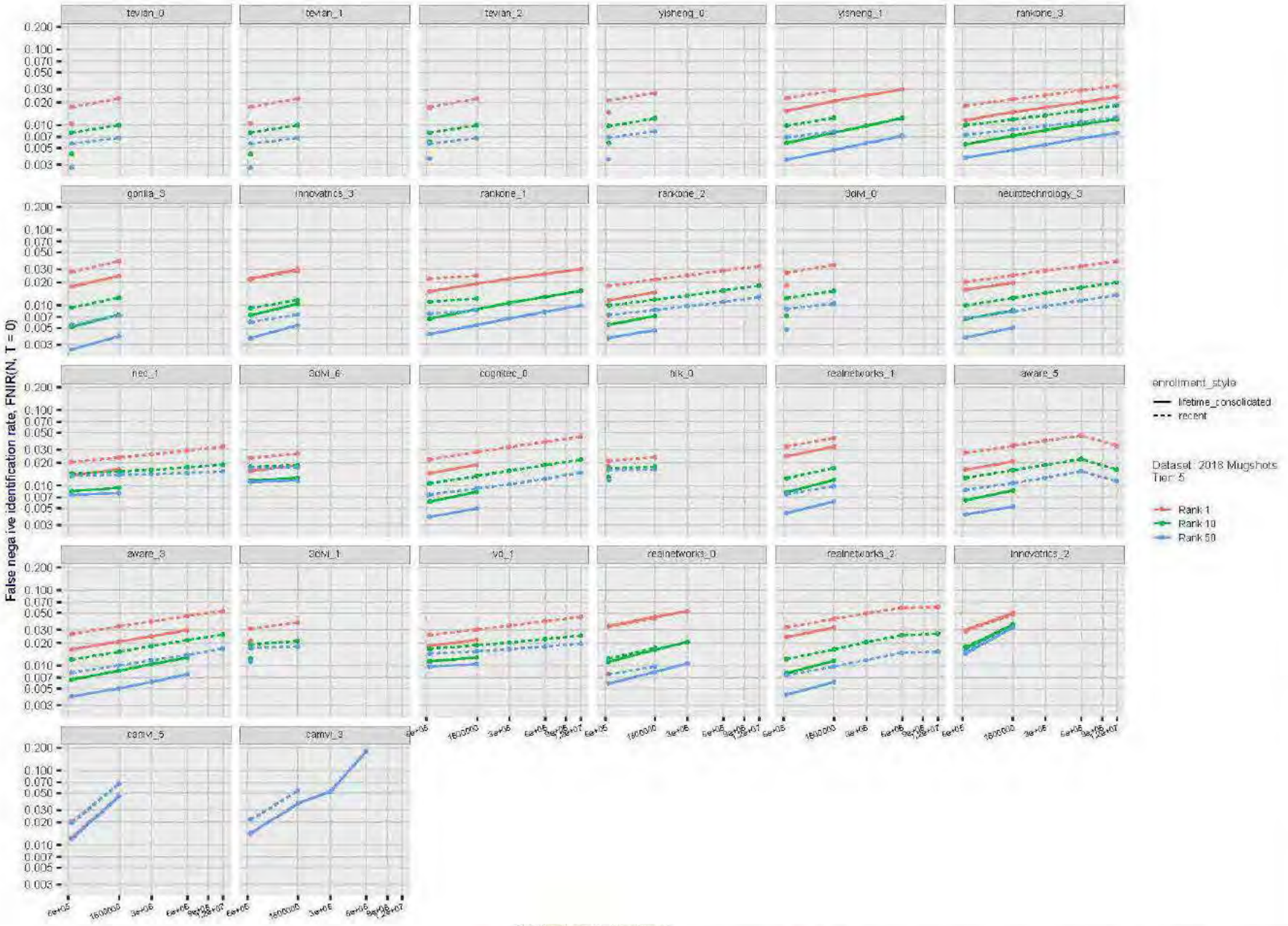
2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification



FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

2019/09/11
17:24:52

Figure 26: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

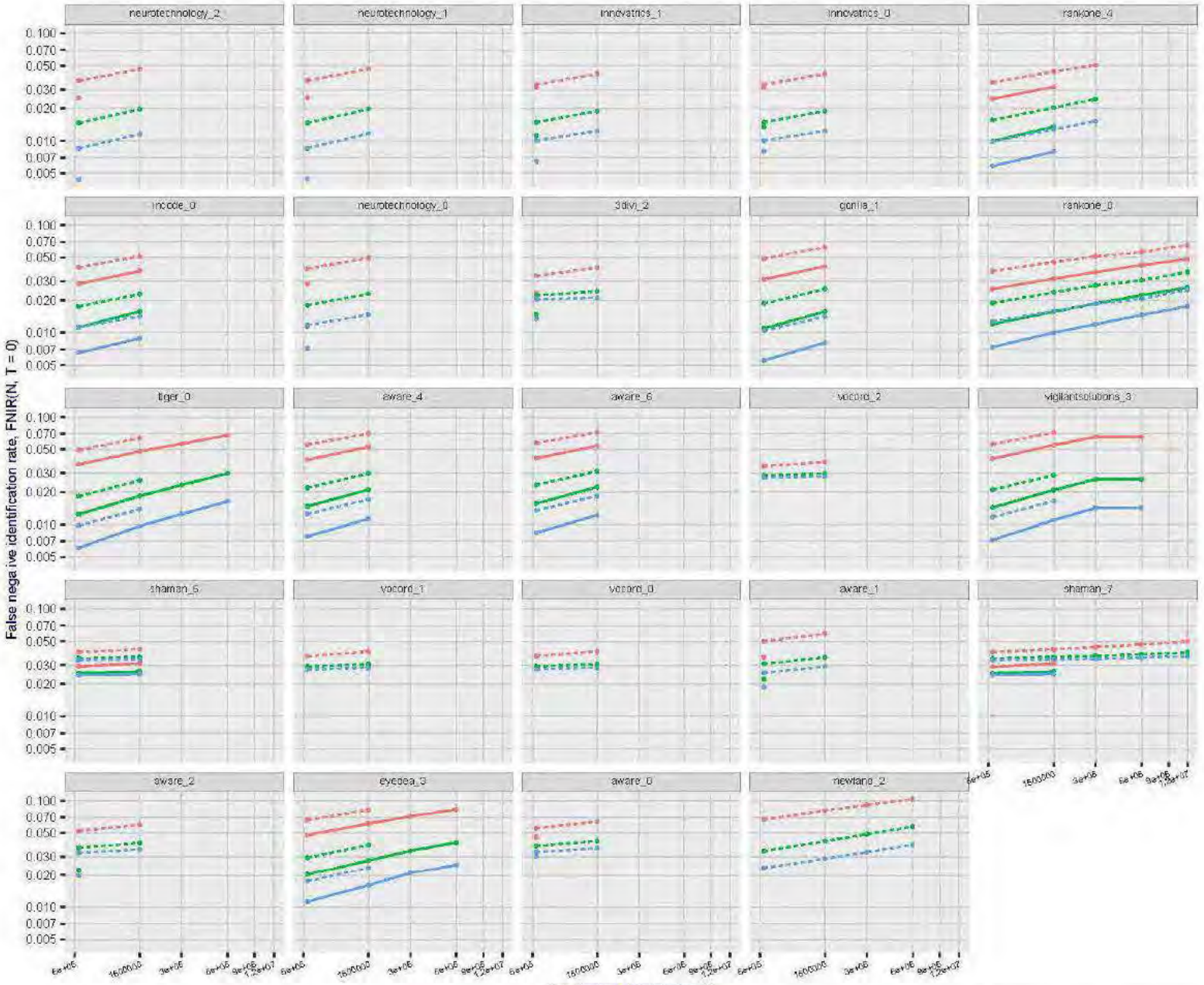


Figure 27: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640000.*

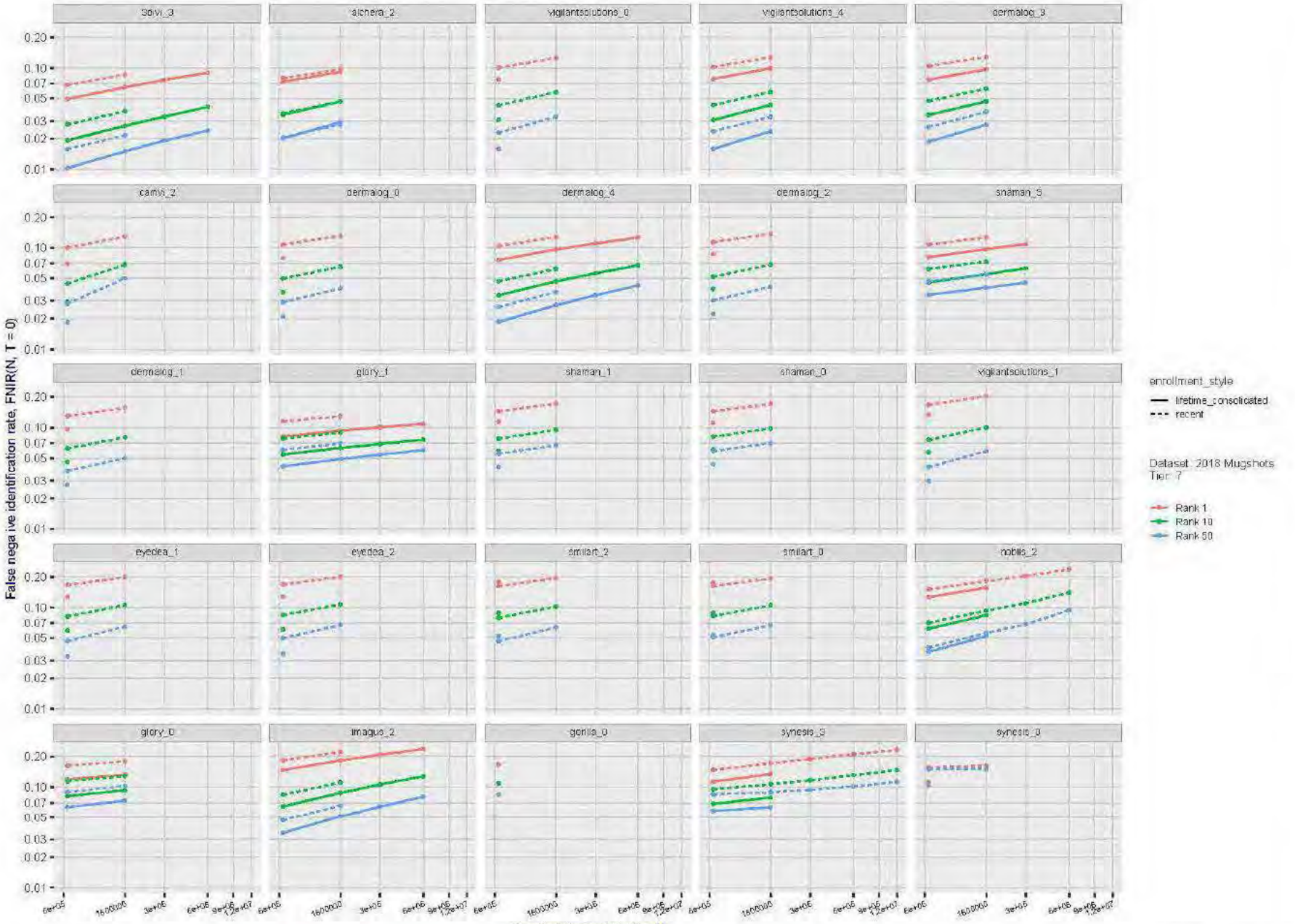


Figure 28: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

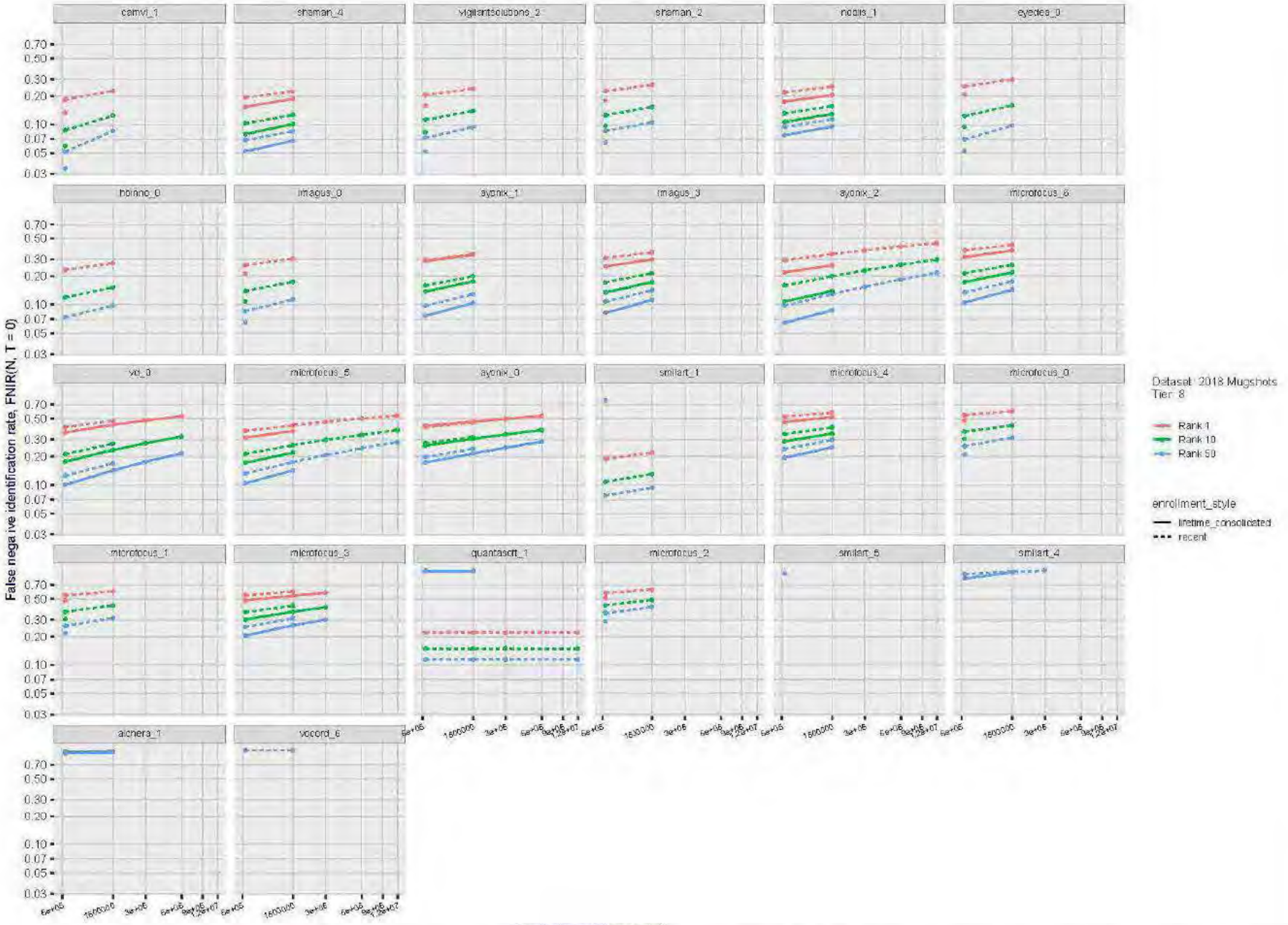


Figure 29: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. number of enrolled subjects.** *The figure shows false negative identification rates, FNIR(N, R), across various gallery sizes and ranks 1, 10 and 50. The threshold is set to zero, so this metric rewards even weak scoring rank 1 mates. This also means FPIR = 1, so any search without an enrolled mate will return non-mated candidates. For clarity, results are sorted and reported into tiers spanning multiple pages, the tiering criteria being rank 1 hit rate on a gallery size of 640 000.*

2019/09/11 17:24:52	FNIR(N, R, T) = False neg. identification rate FPIR(N, T) = False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification	

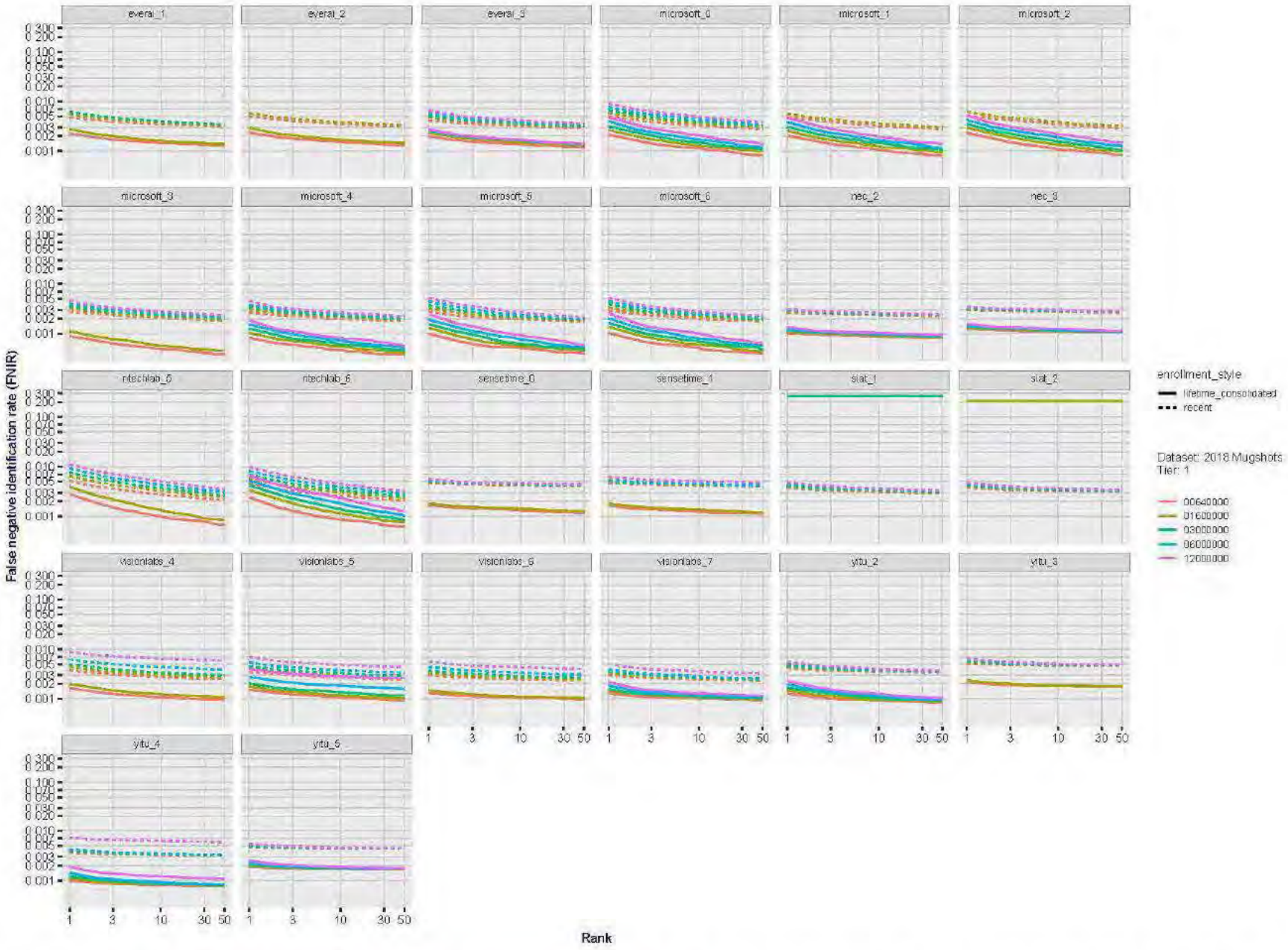


Figure 30: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

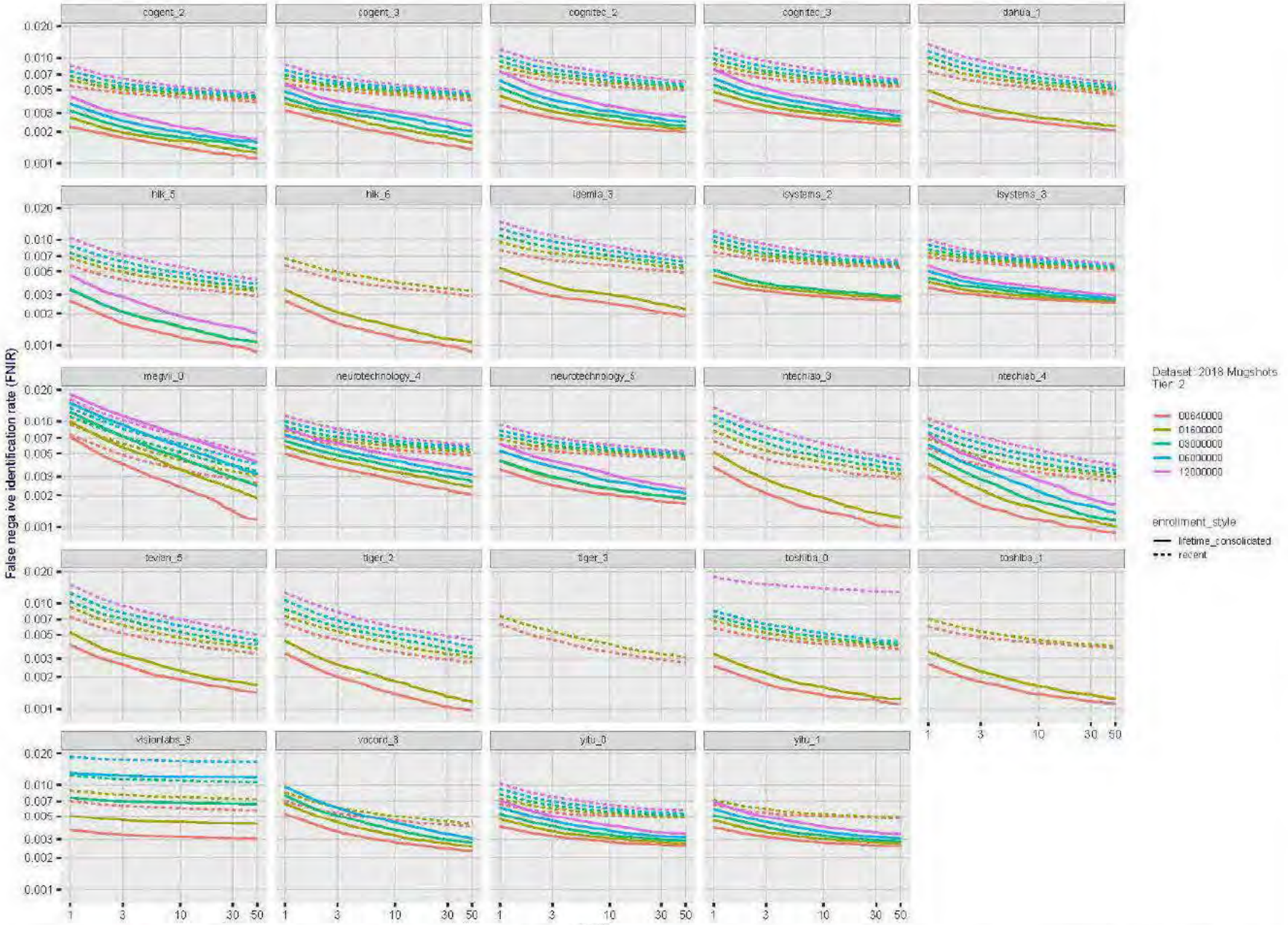


Figure 31: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

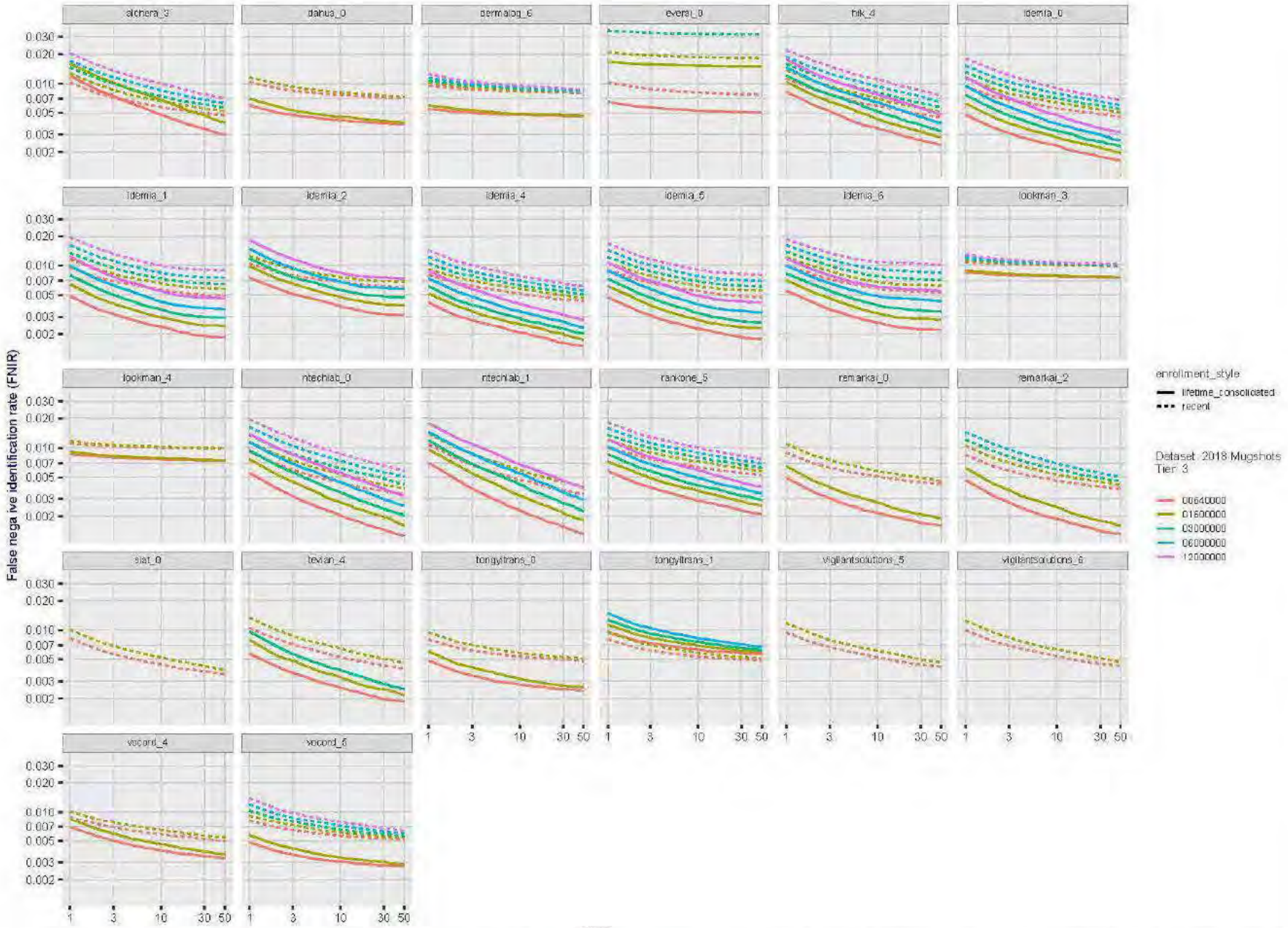


Figure 32: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

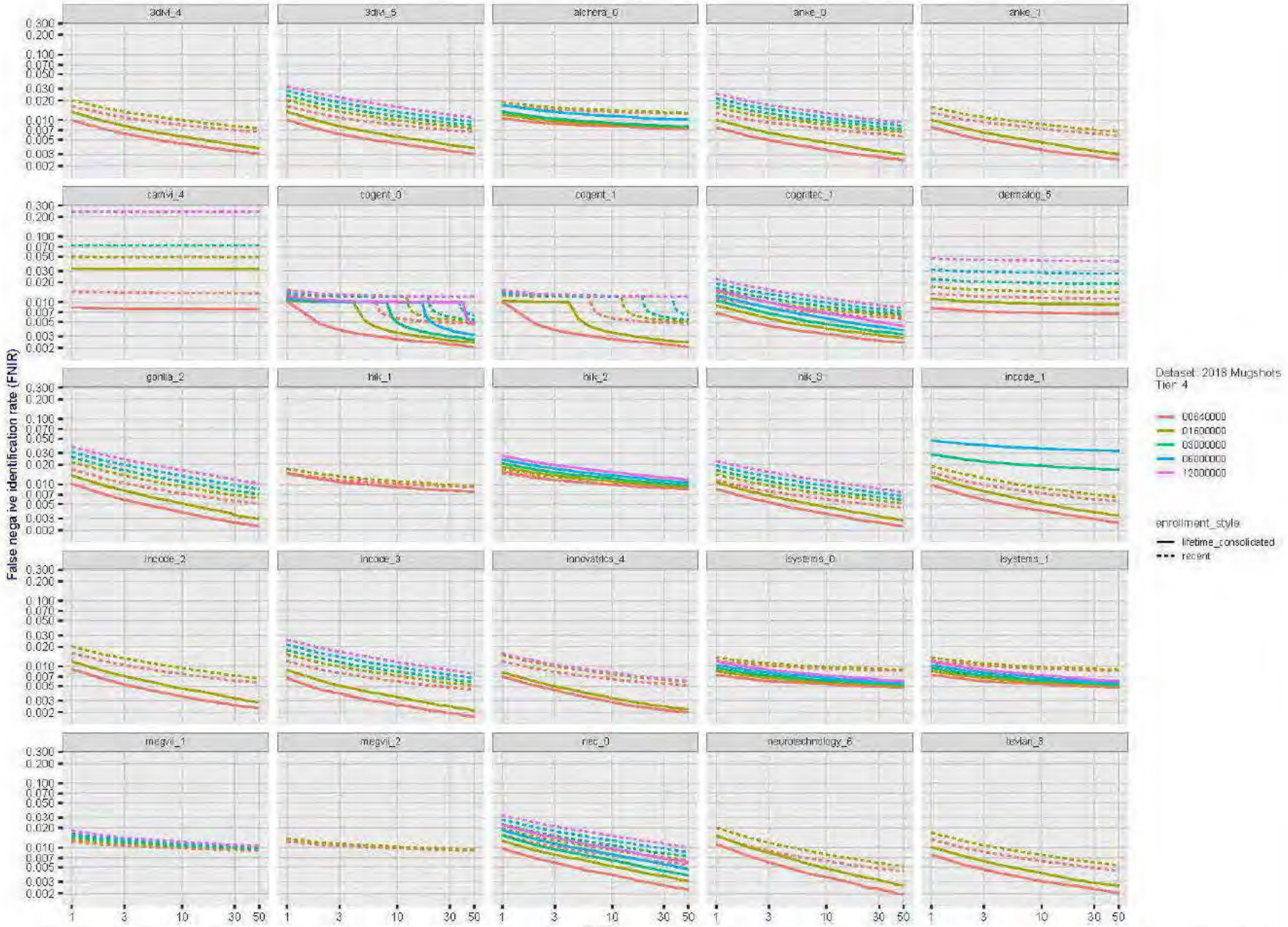


Figure 33: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

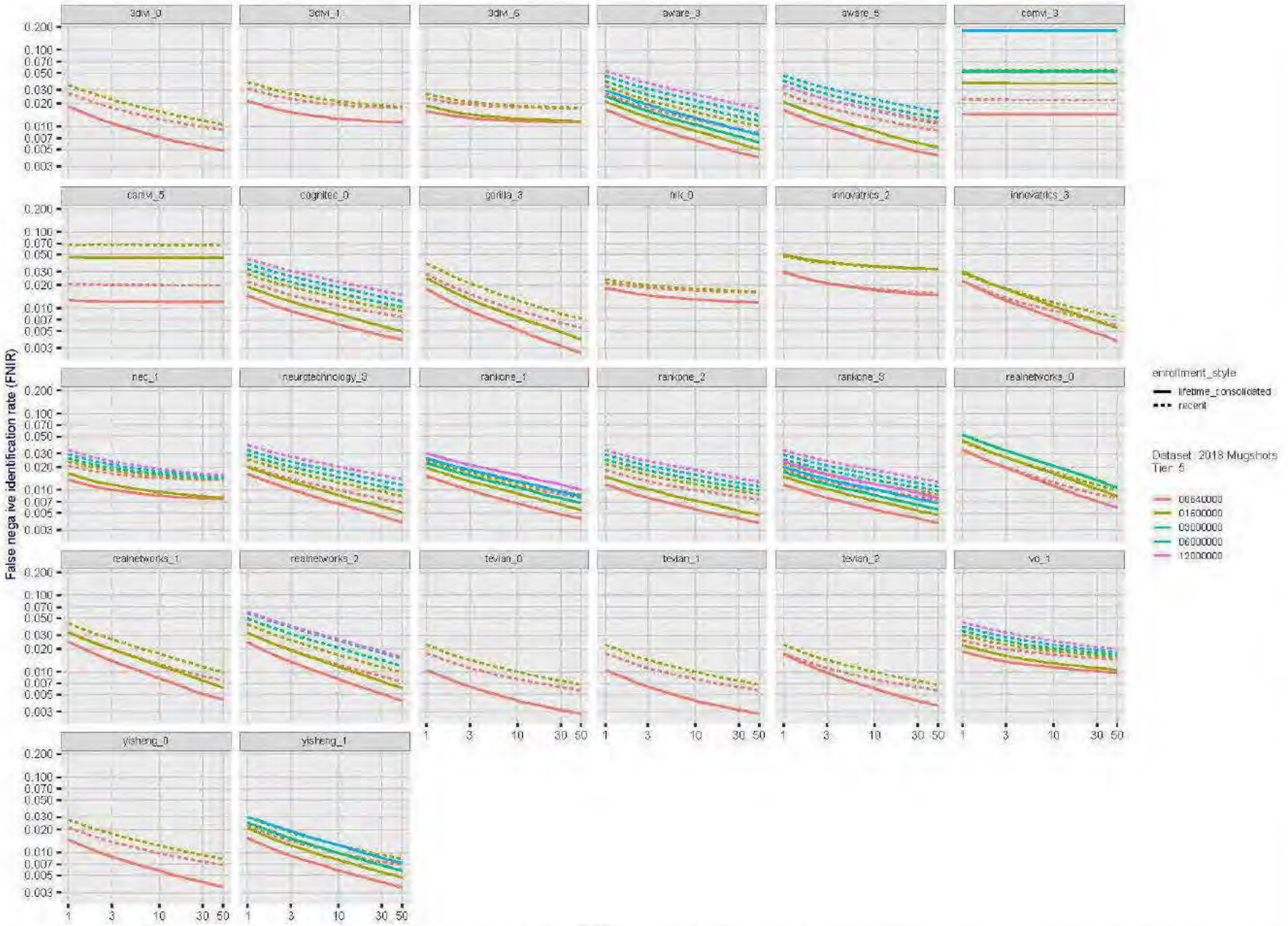


Figure 34: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

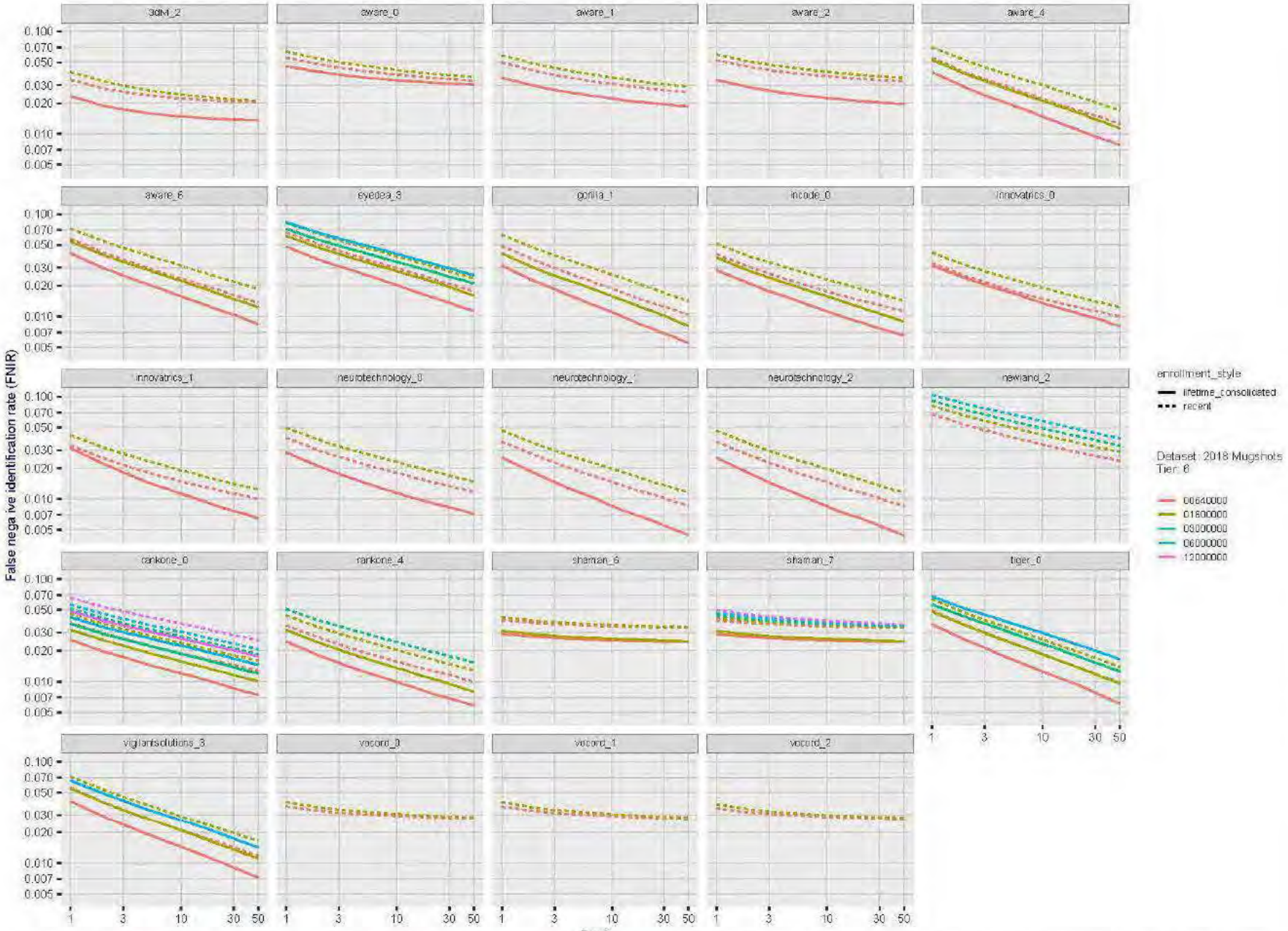


Figure 35: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

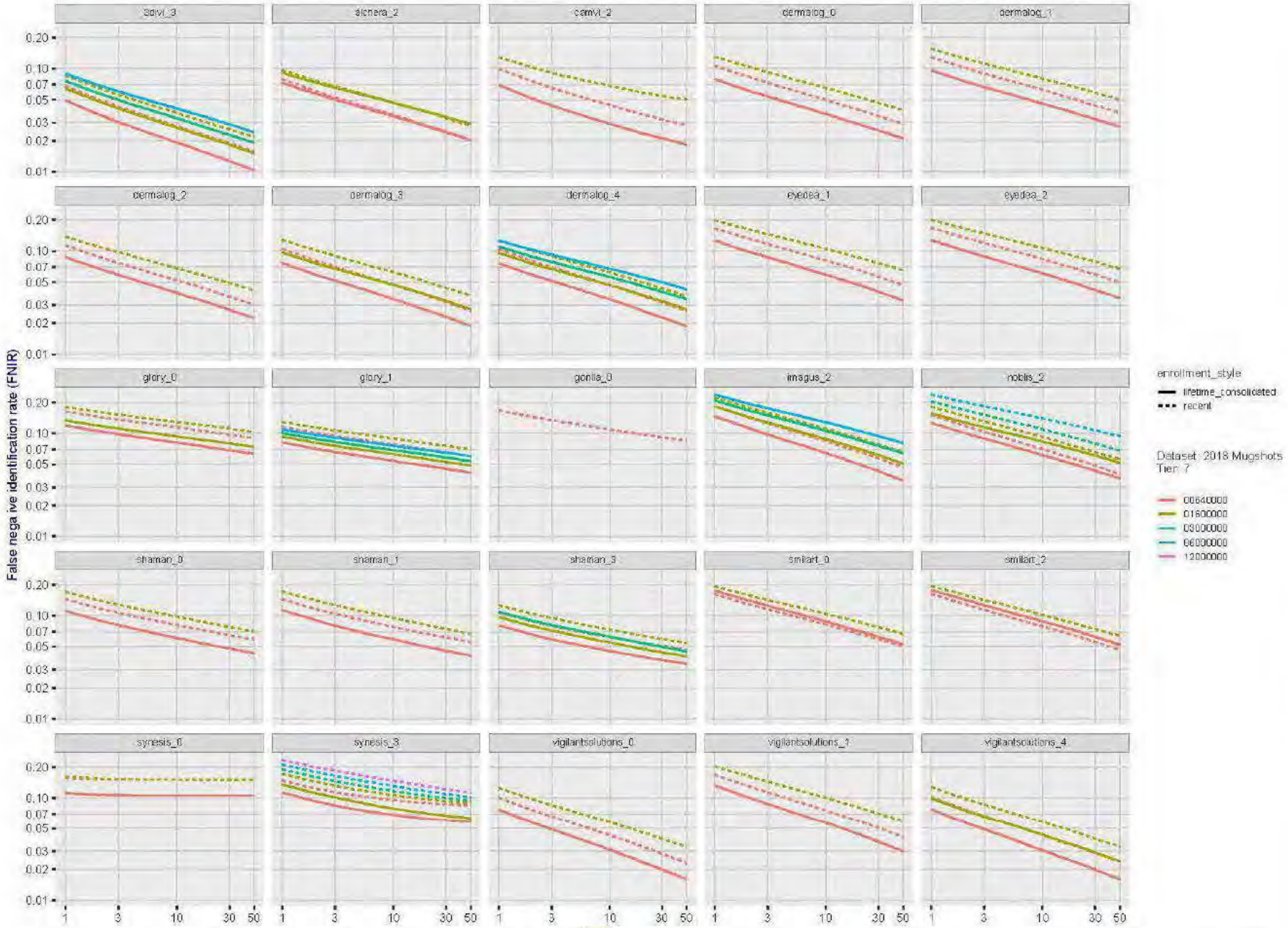


Figure 36: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

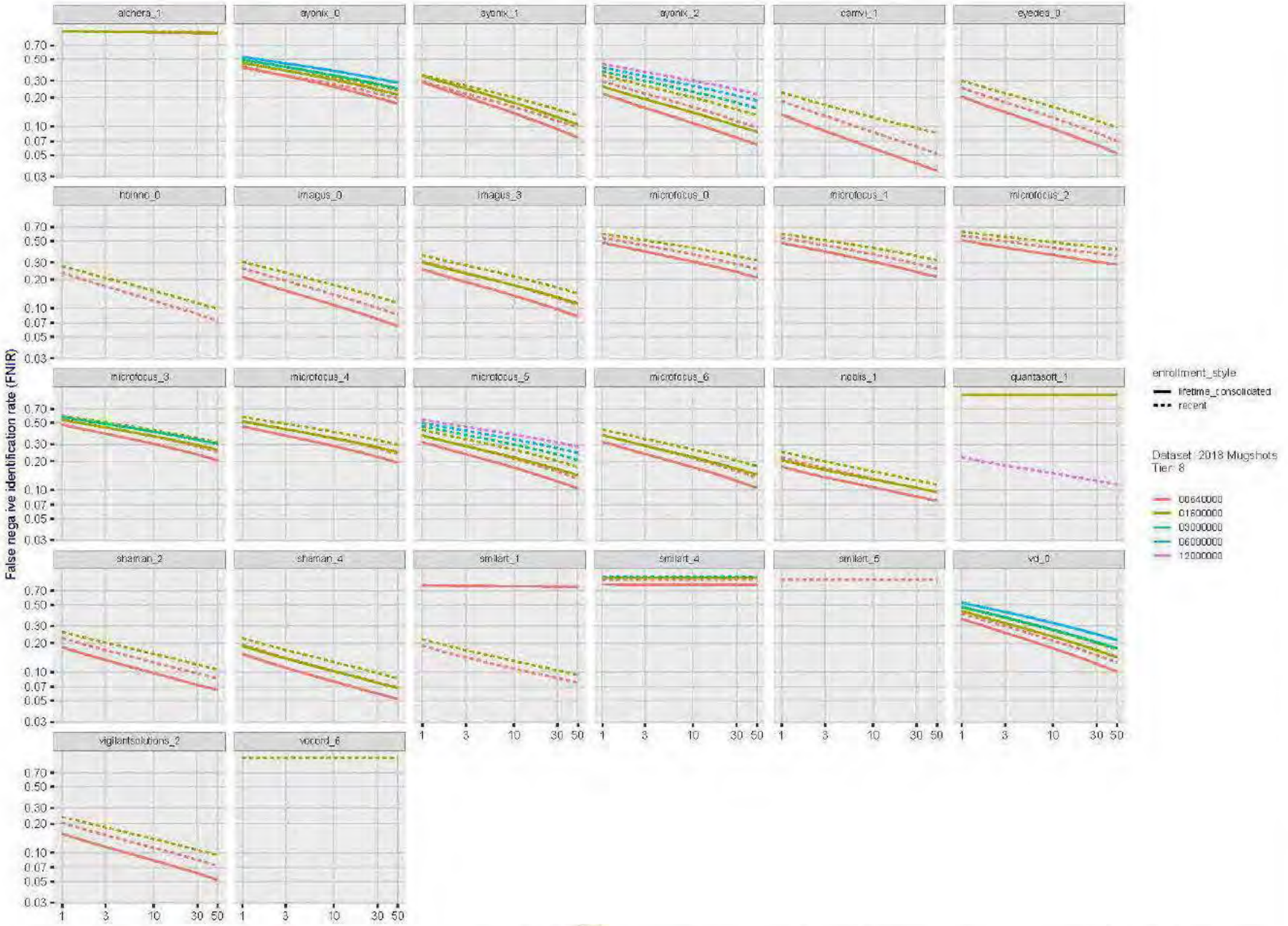


Figure 37: **[FRVT-2018 Mugshot Dataset] Rank-based identification miss rates vs. rank.** *The figure shows false negative identification rates (FNIR) for ranks up to 50. This metric is appropriate to investigational applications where human reviewers will adjudicate sorted candidate lists. Note that with threshold set to zero, FPIR = 1, i.e. any search without an enrolled mate will return non-mated candidates. Results are sorted and reported into tiers for clarity, with the tiering criteria being rank 1 hit rate on a gallery size of N = 640 000 subjects.*

2019/09/11 17:24:52	FNIR(N, R, T) = FPIR(N, T) =	False neg. identification rate False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification

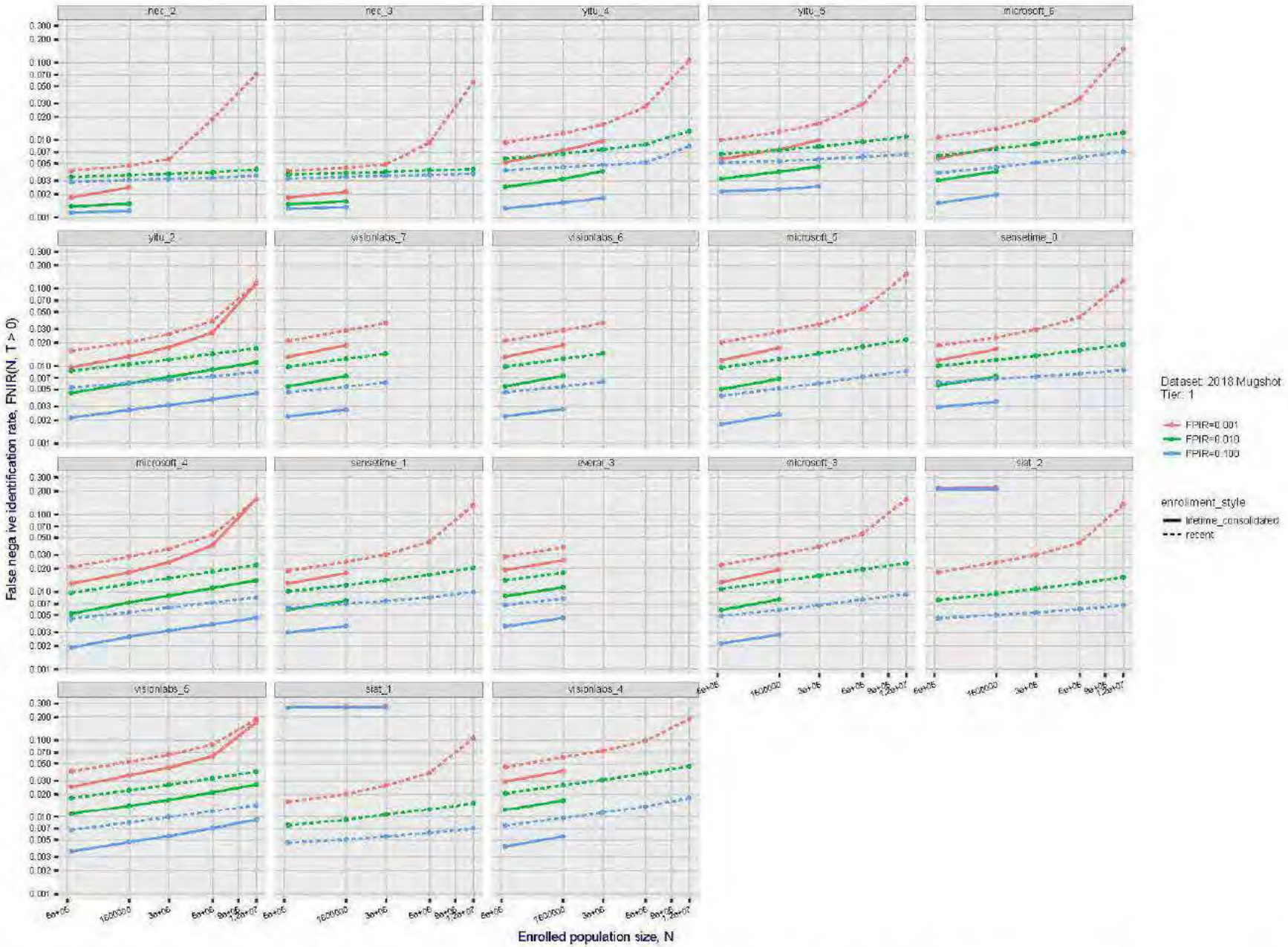


Figure 38: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

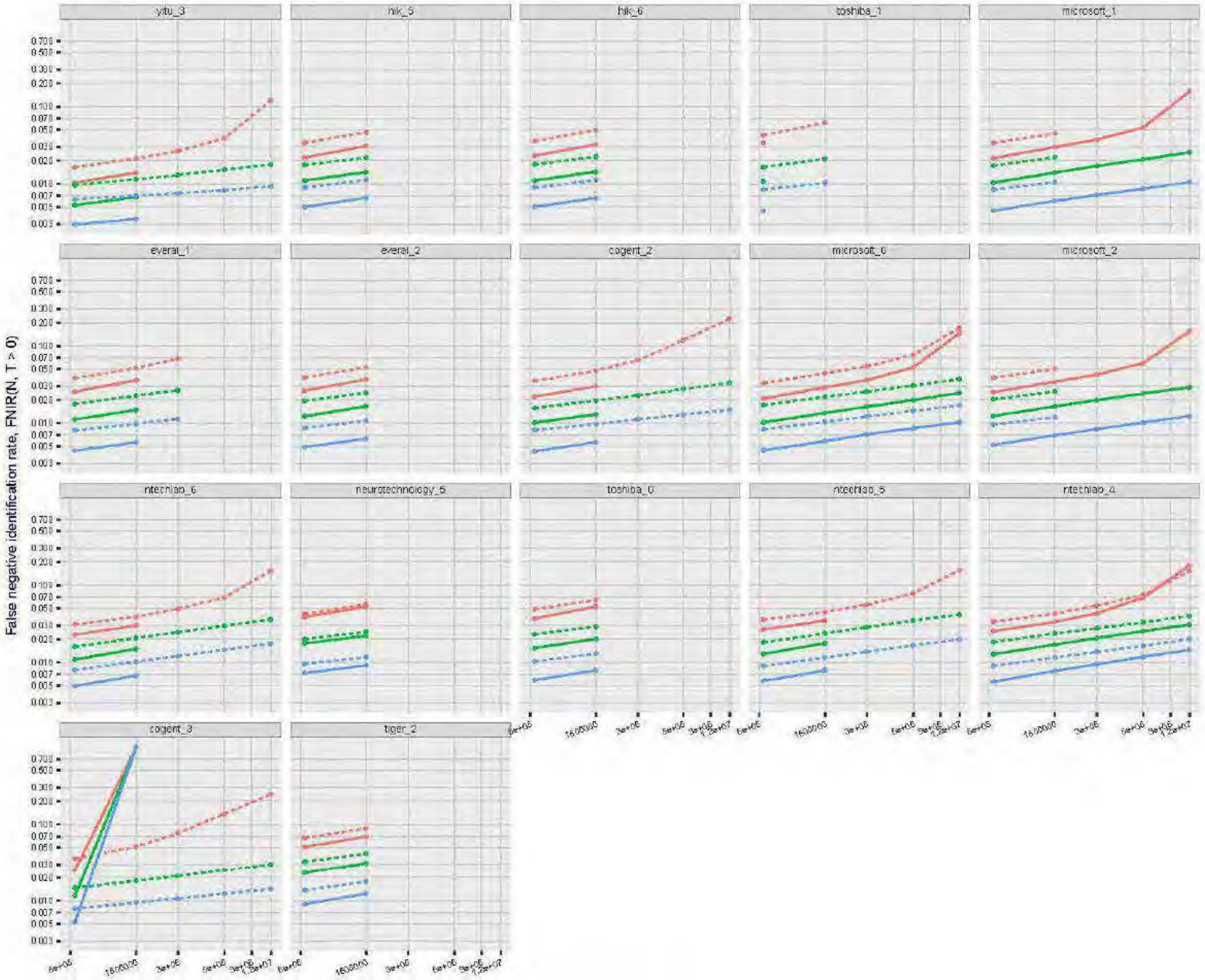


Figure 39: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

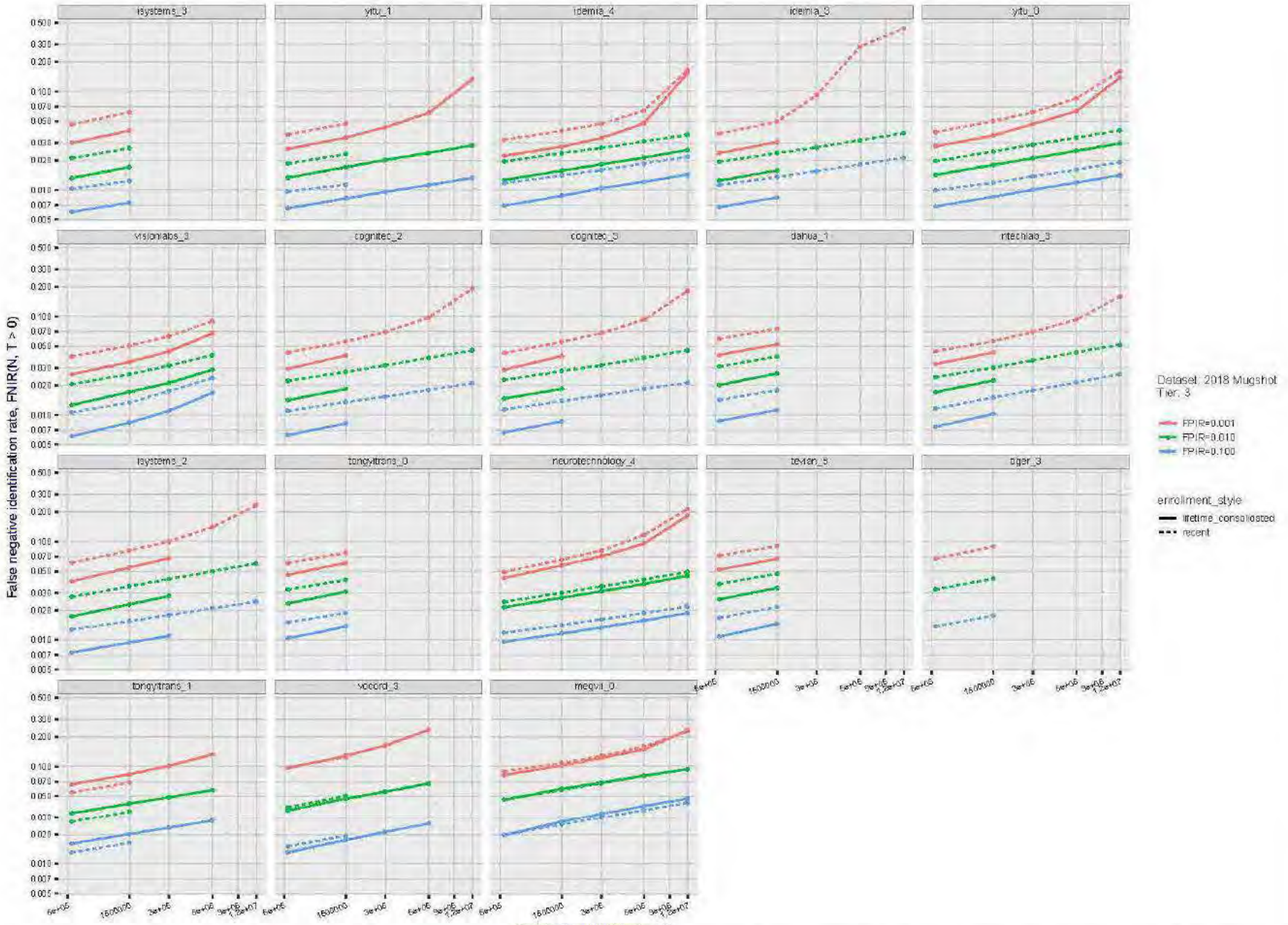


Figure 40: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

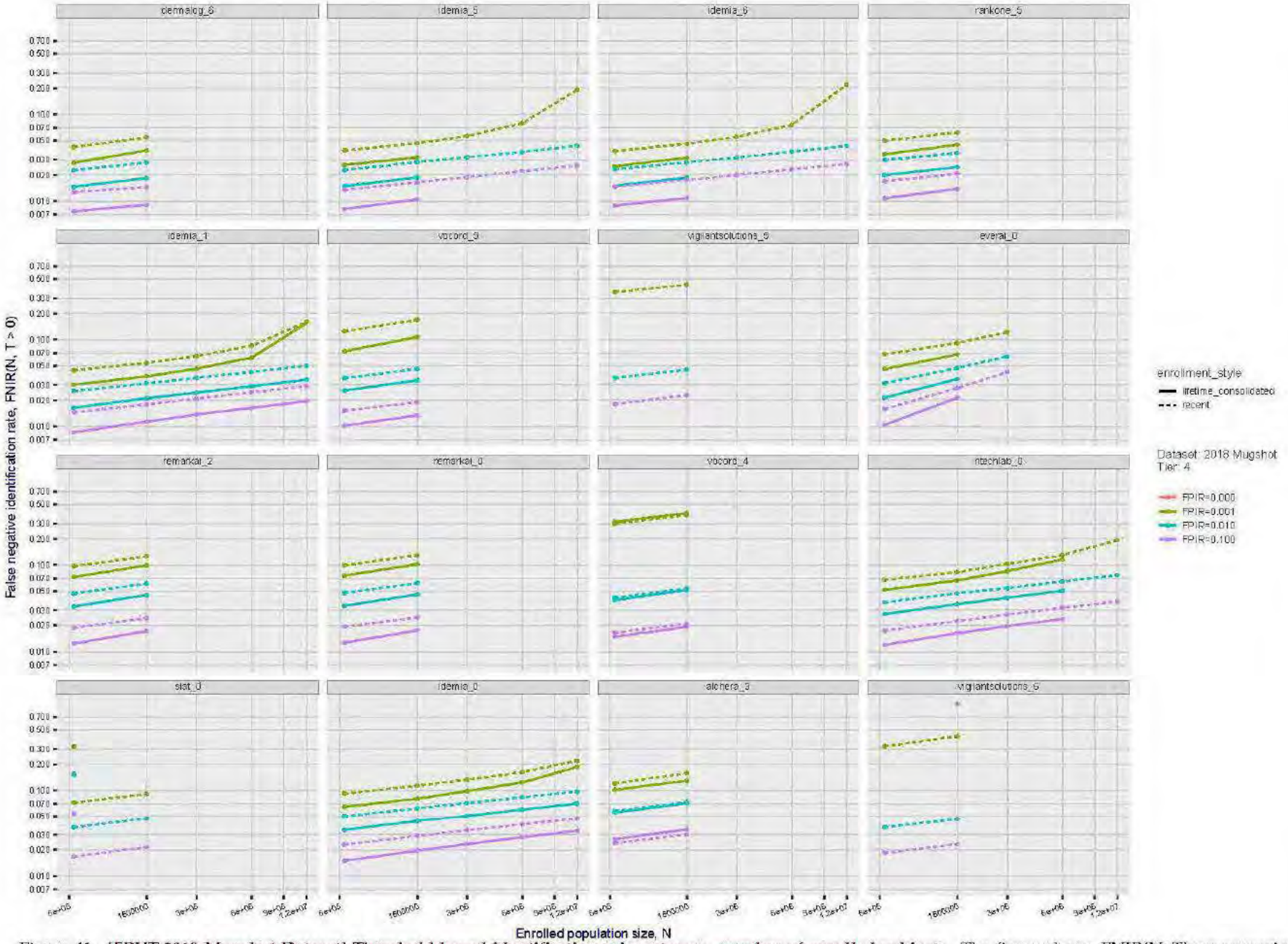


Figure 41: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

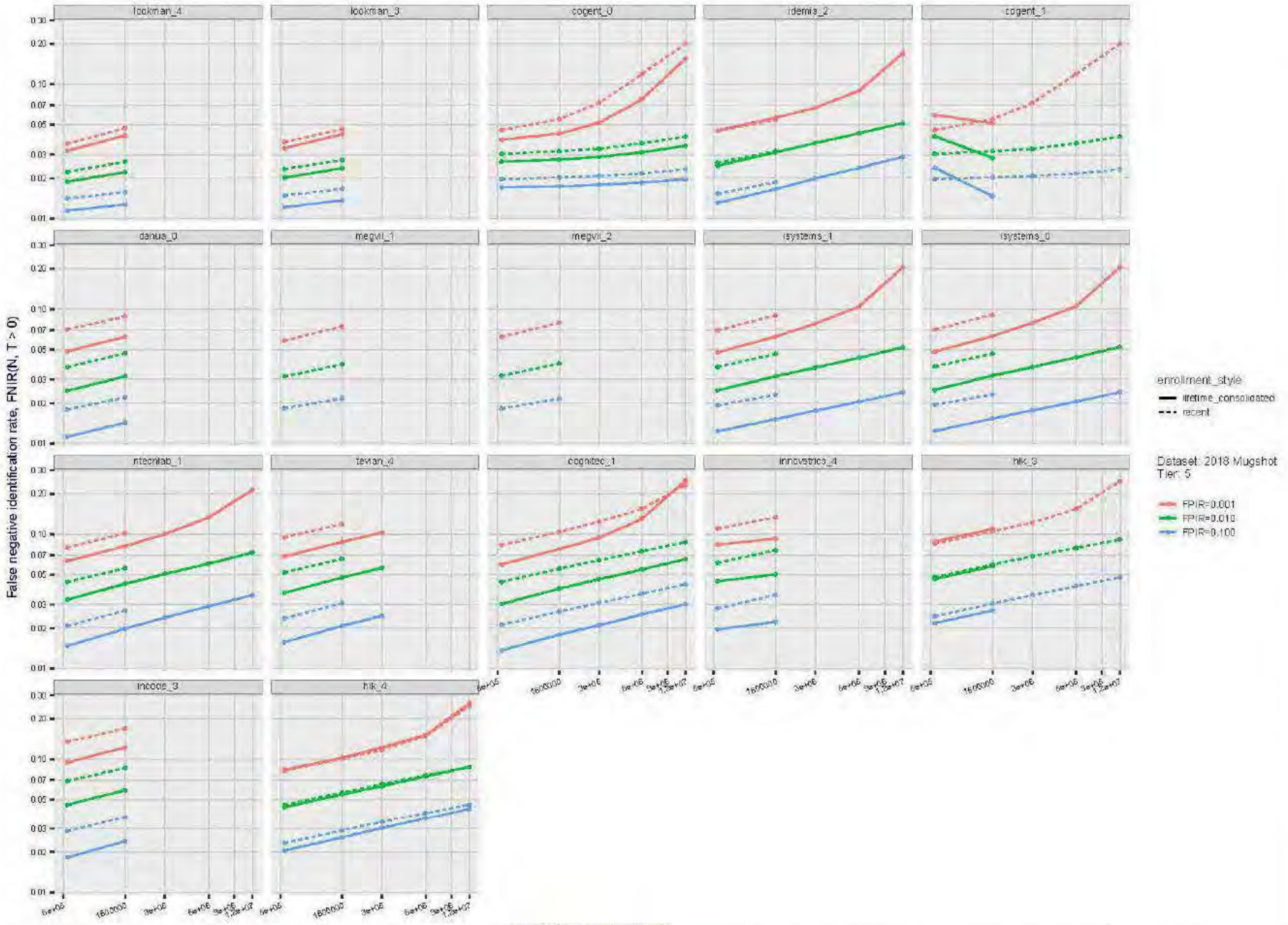


Figure 42: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

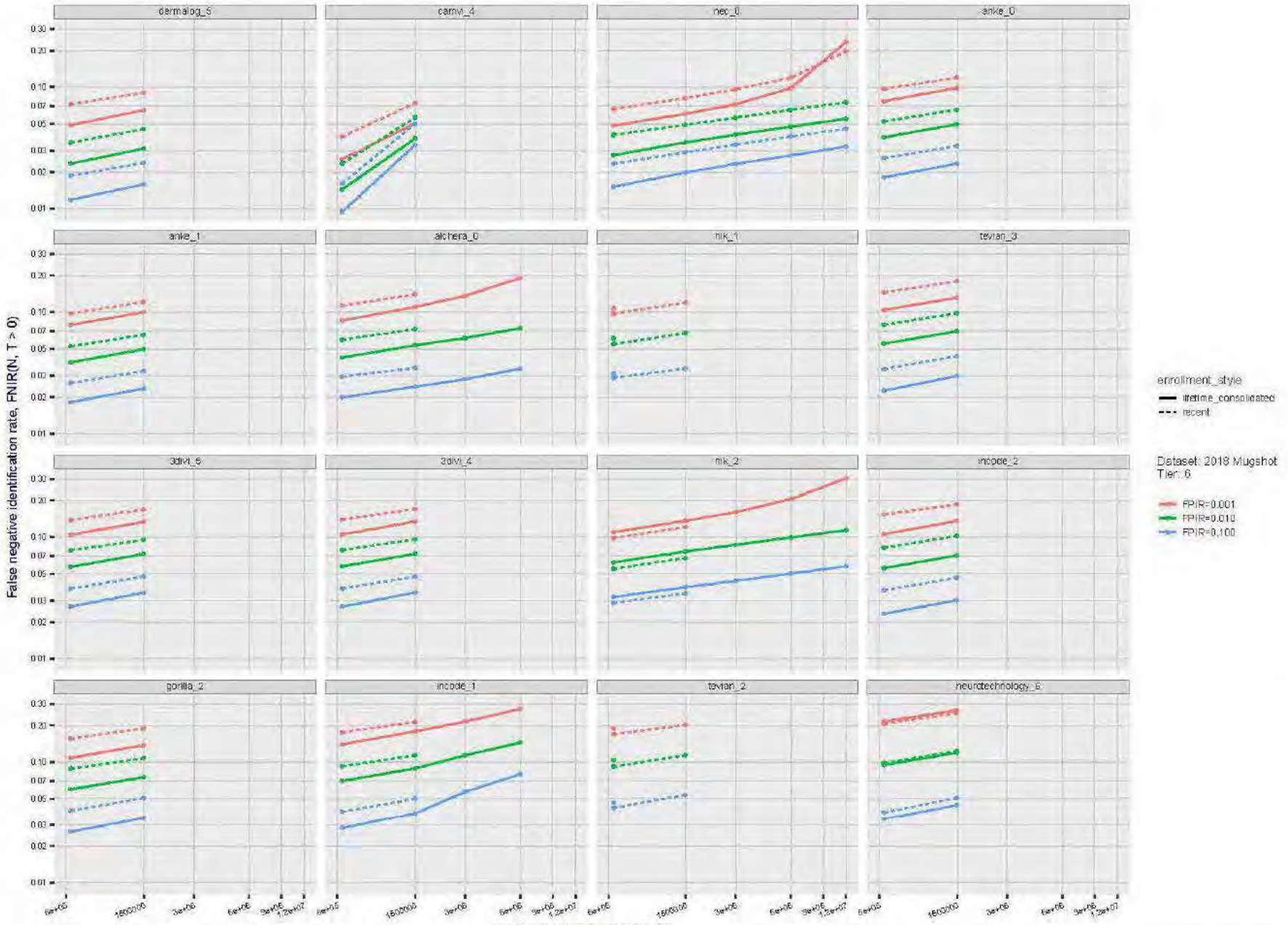


Figure 43: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

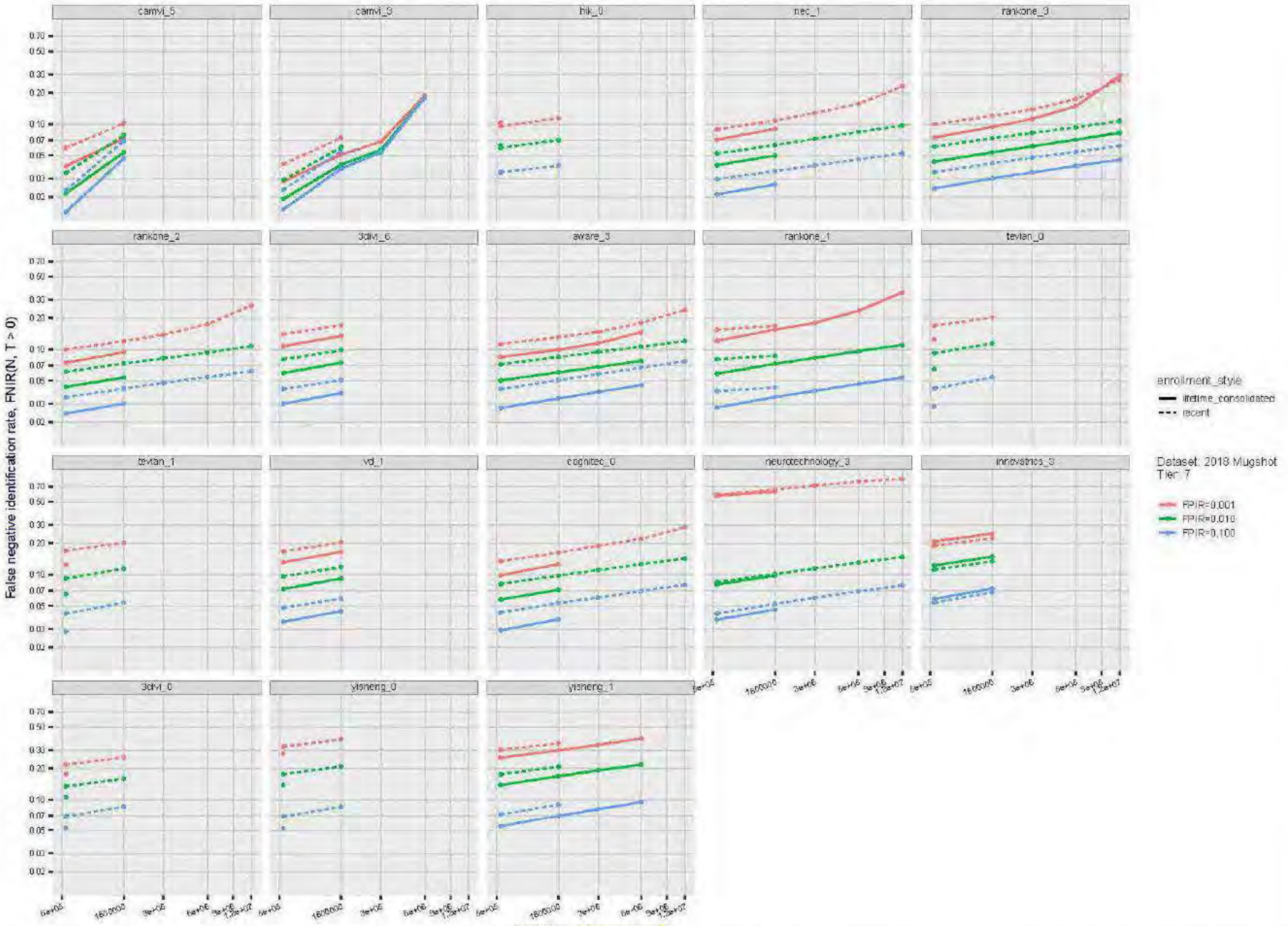


Figure 44: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

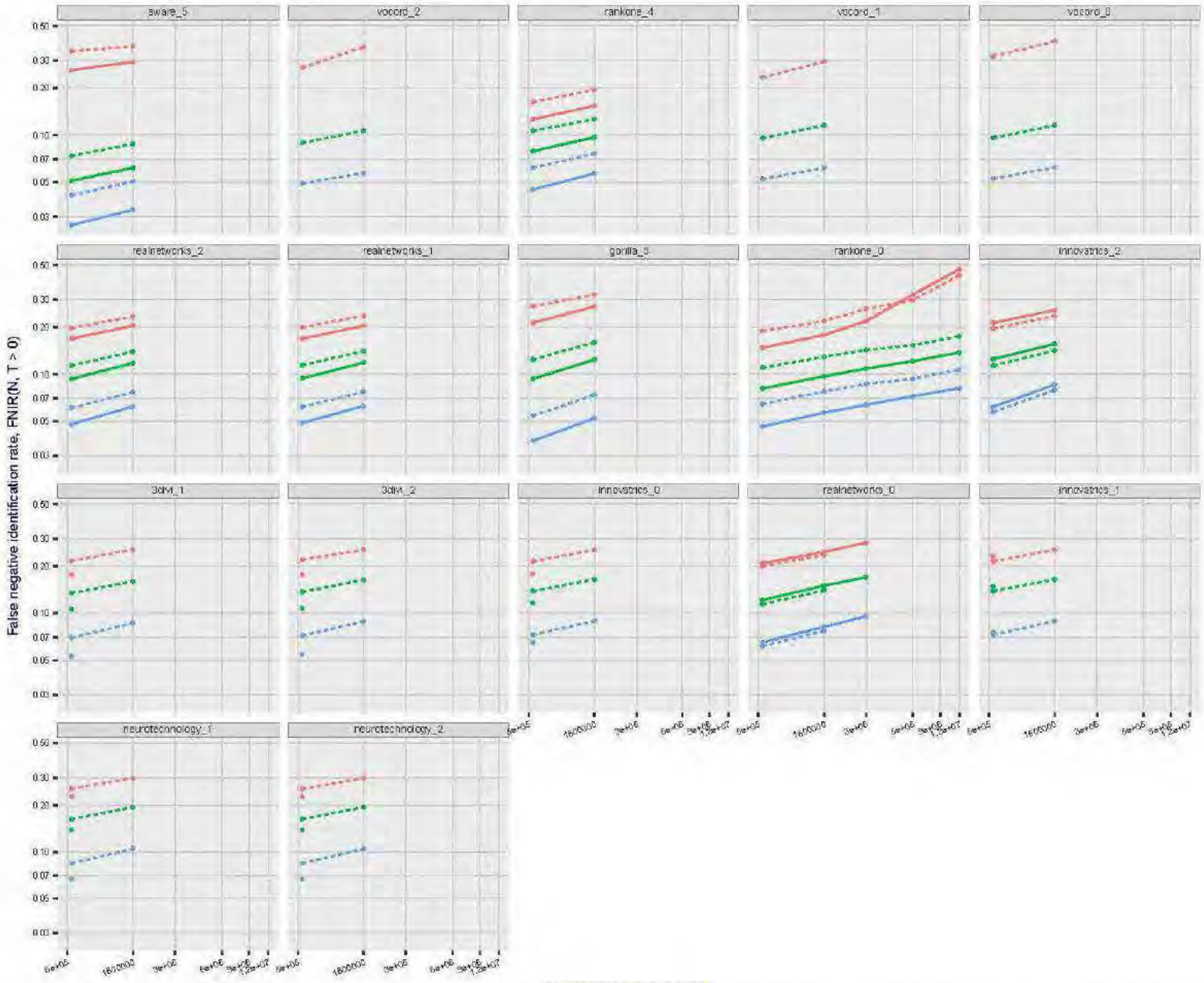


Figure 45: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

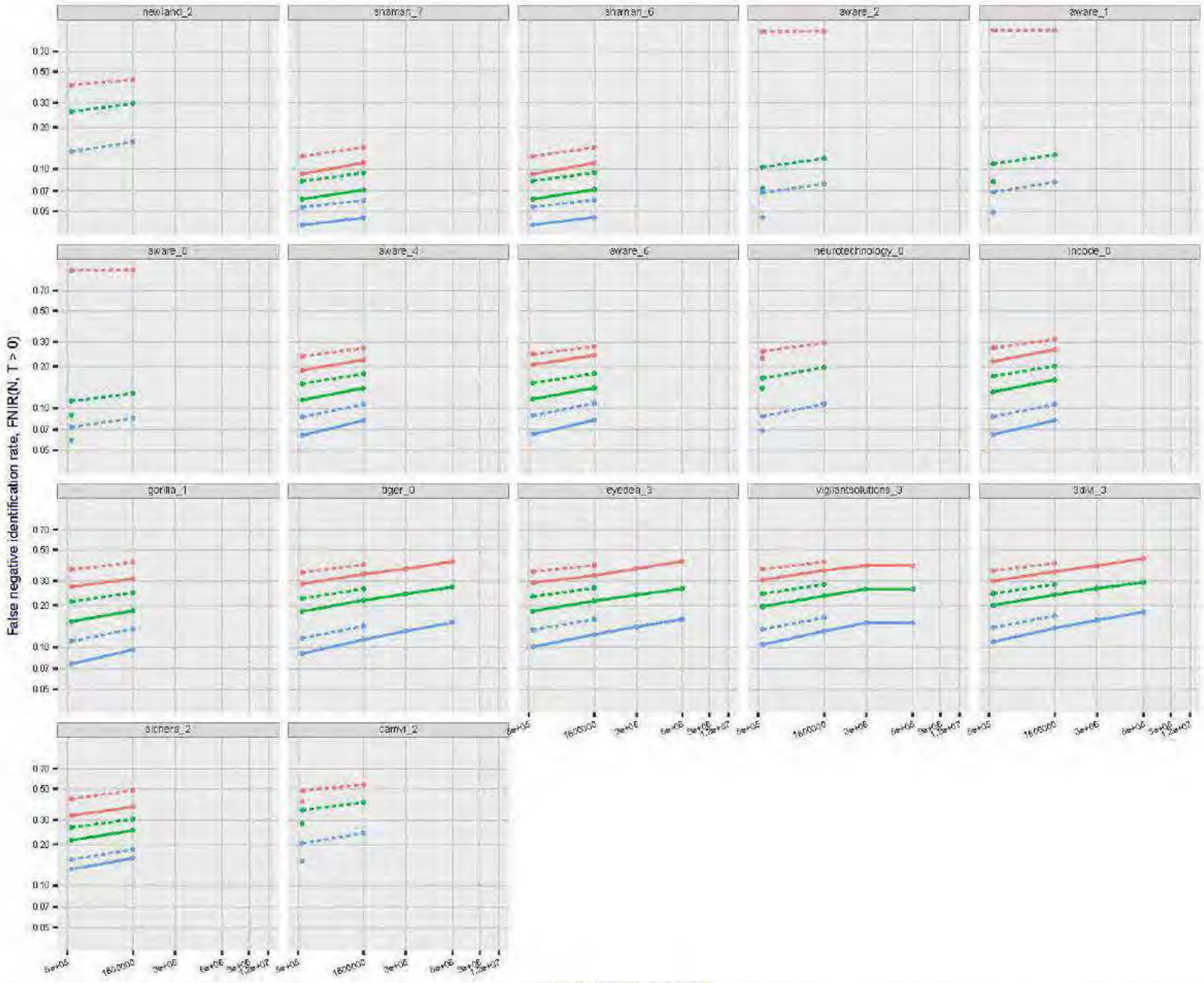


Figure 46: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

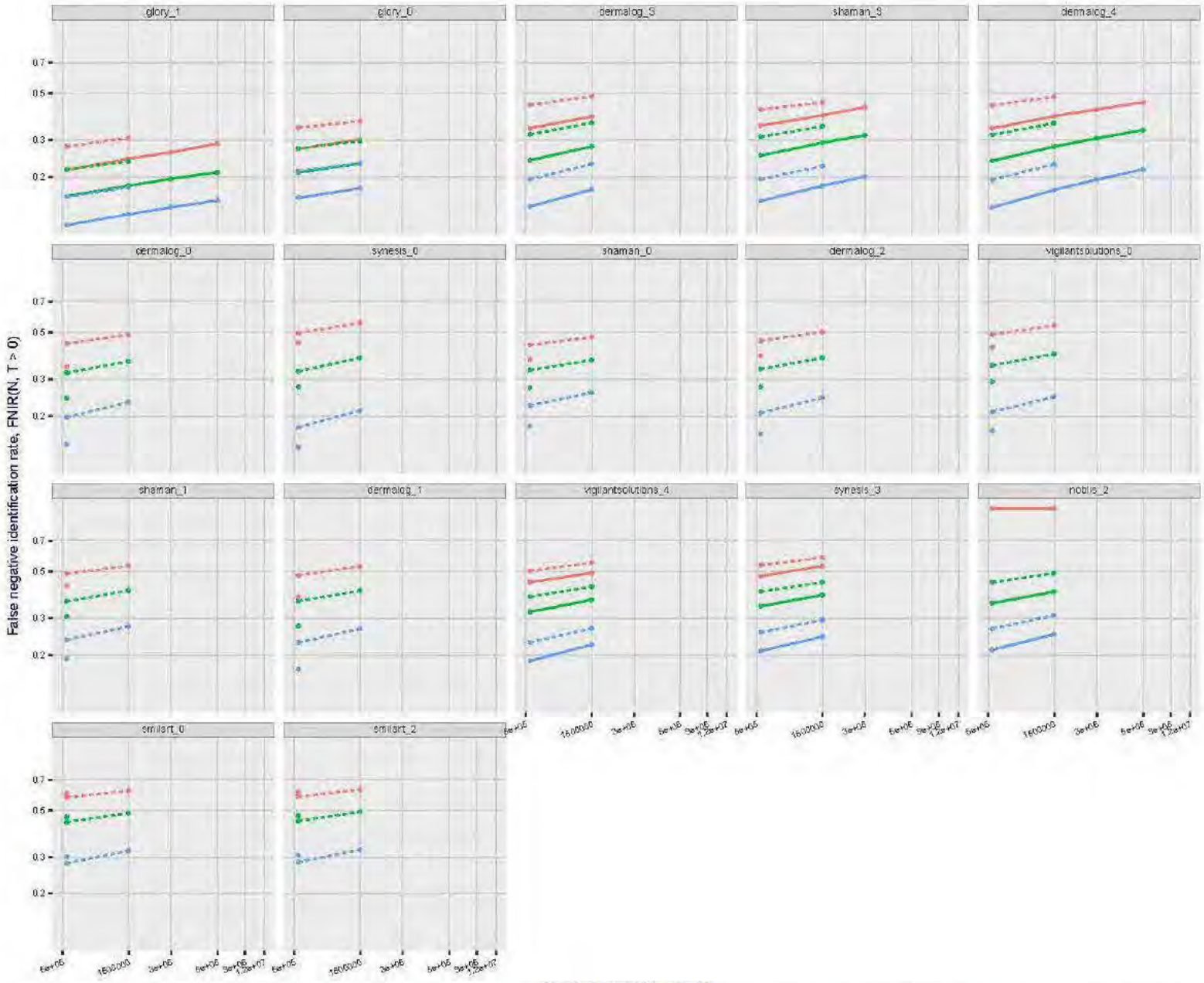


Figure 47: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

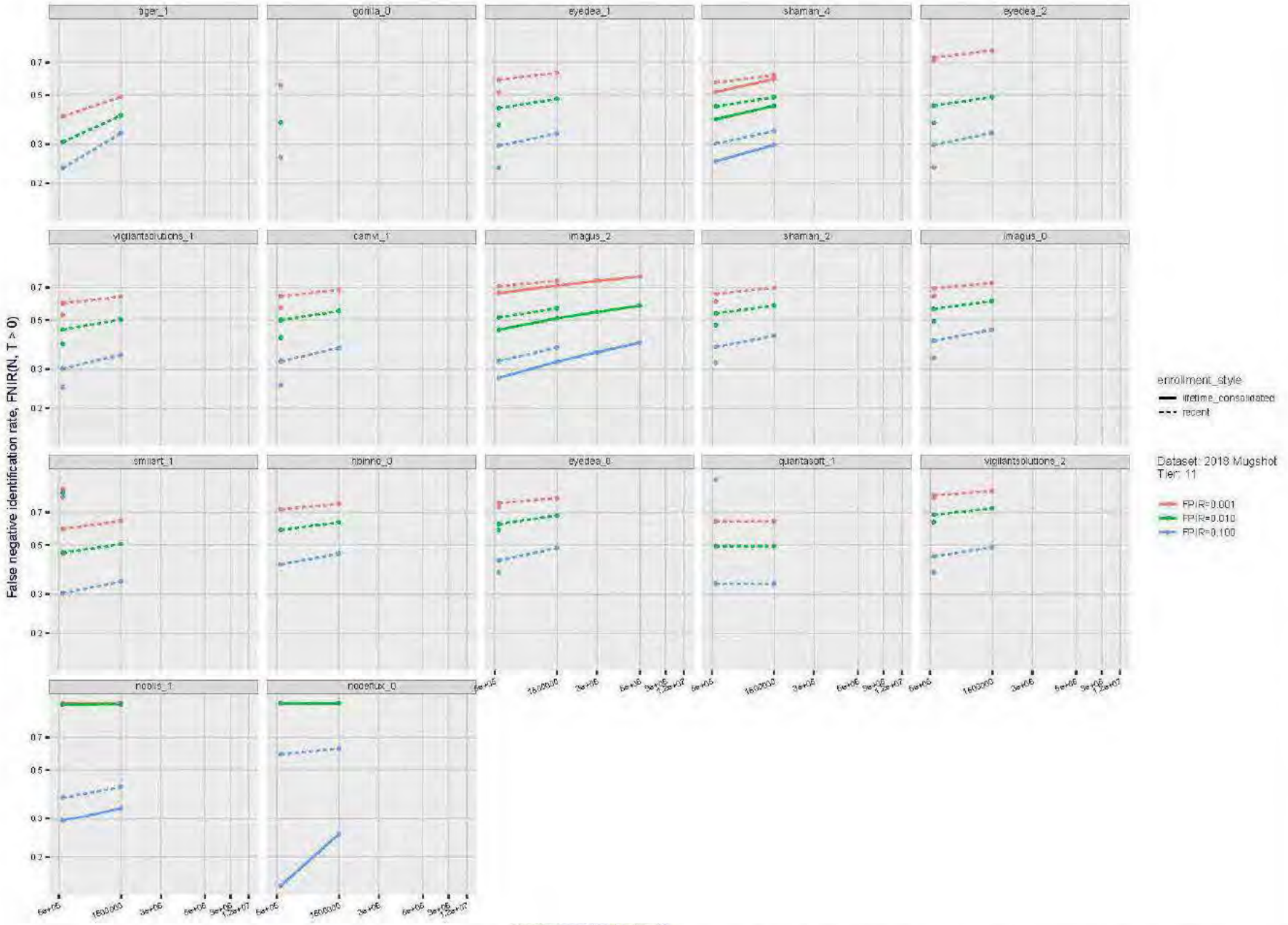


Figure 48: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

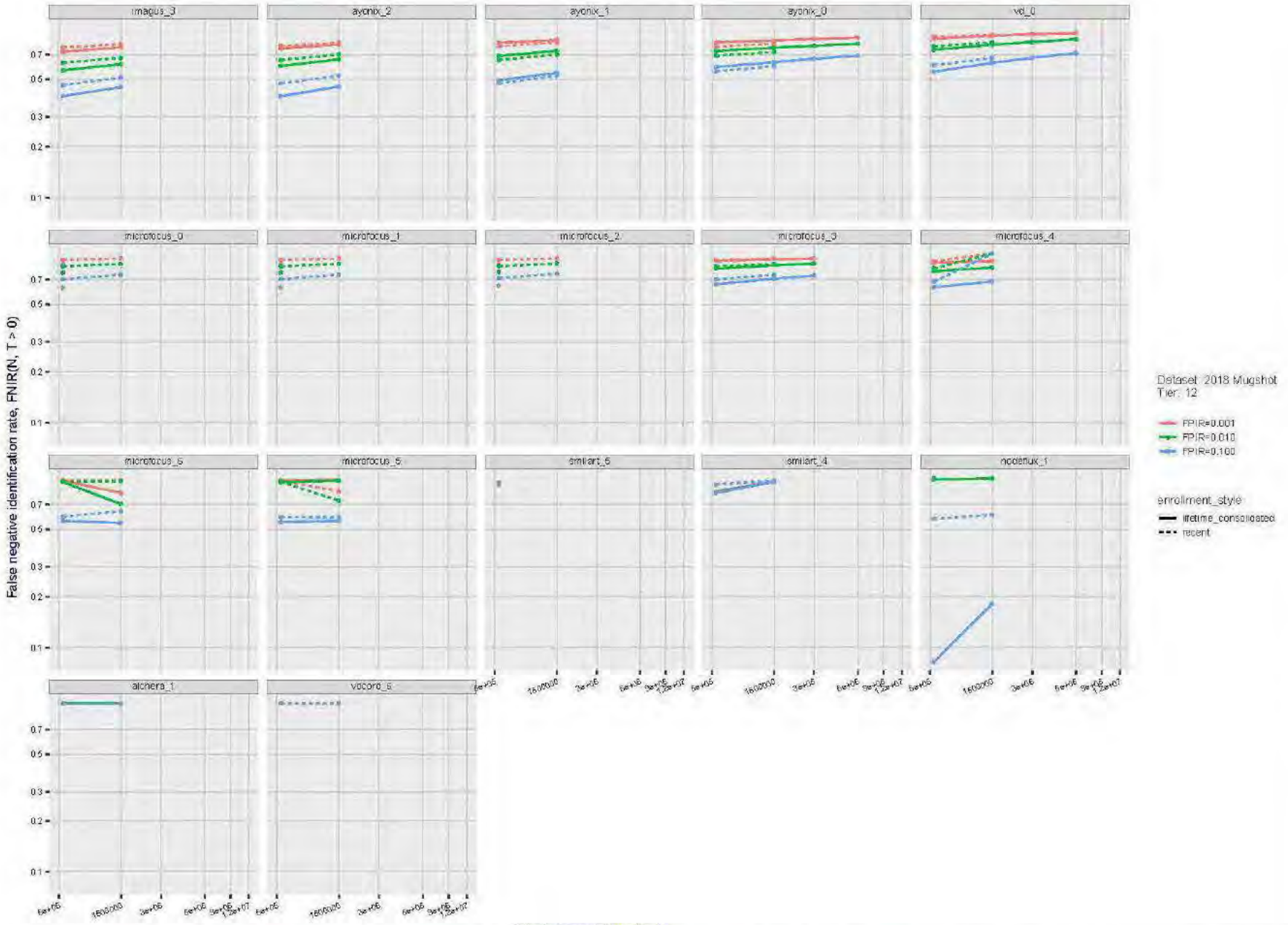


Figure 49: **[FRVT-2018 Mugshot Dataset] Threshold-based identification miss rates vs. number of enrolled subjects.** *The figure shows FNIR(N, T) across various gallery sizes when the threshold is set to achieve the given FPIRs. The rank criterion is irrelevant at high thresholds as mates are always at rank 1. The results are computed from the trials listed in rows 1-10 of Table 5. Less accurate algorithms were not run on large N, so results are missing. For clarity, results are sorted and reported into tiers spanning multiple pages. The tiering criteria is complicated: First paging by FNIR(N_b, 1, 0), then sorting by median FNIR(N_b, T), N_b = 640 000.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

2019/09/11 17:24:52	FNIR(N, R, T) = False neg. identification rate FPIR(N, T) = False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification	

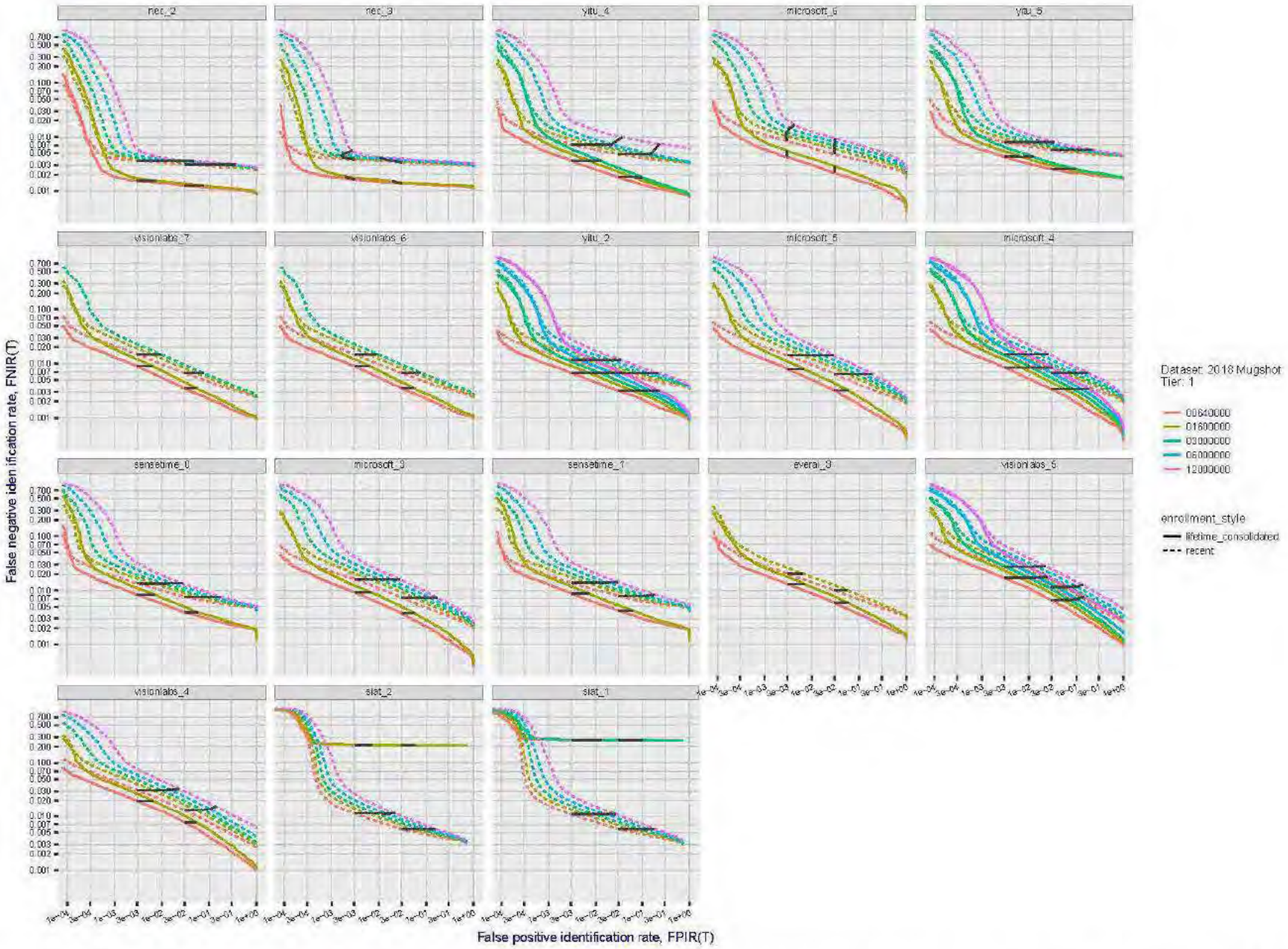


Figure 50: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

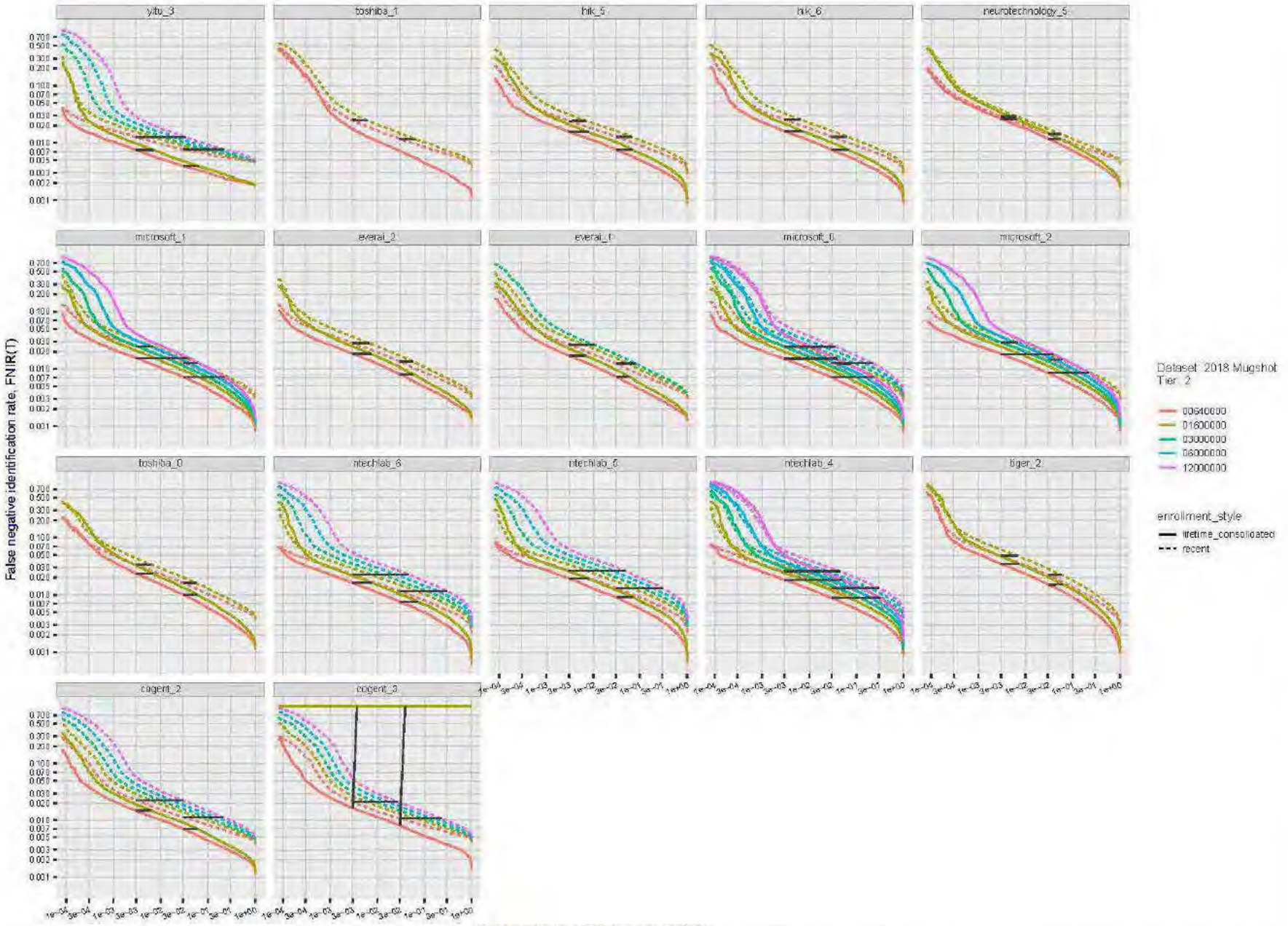


Figure 51: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

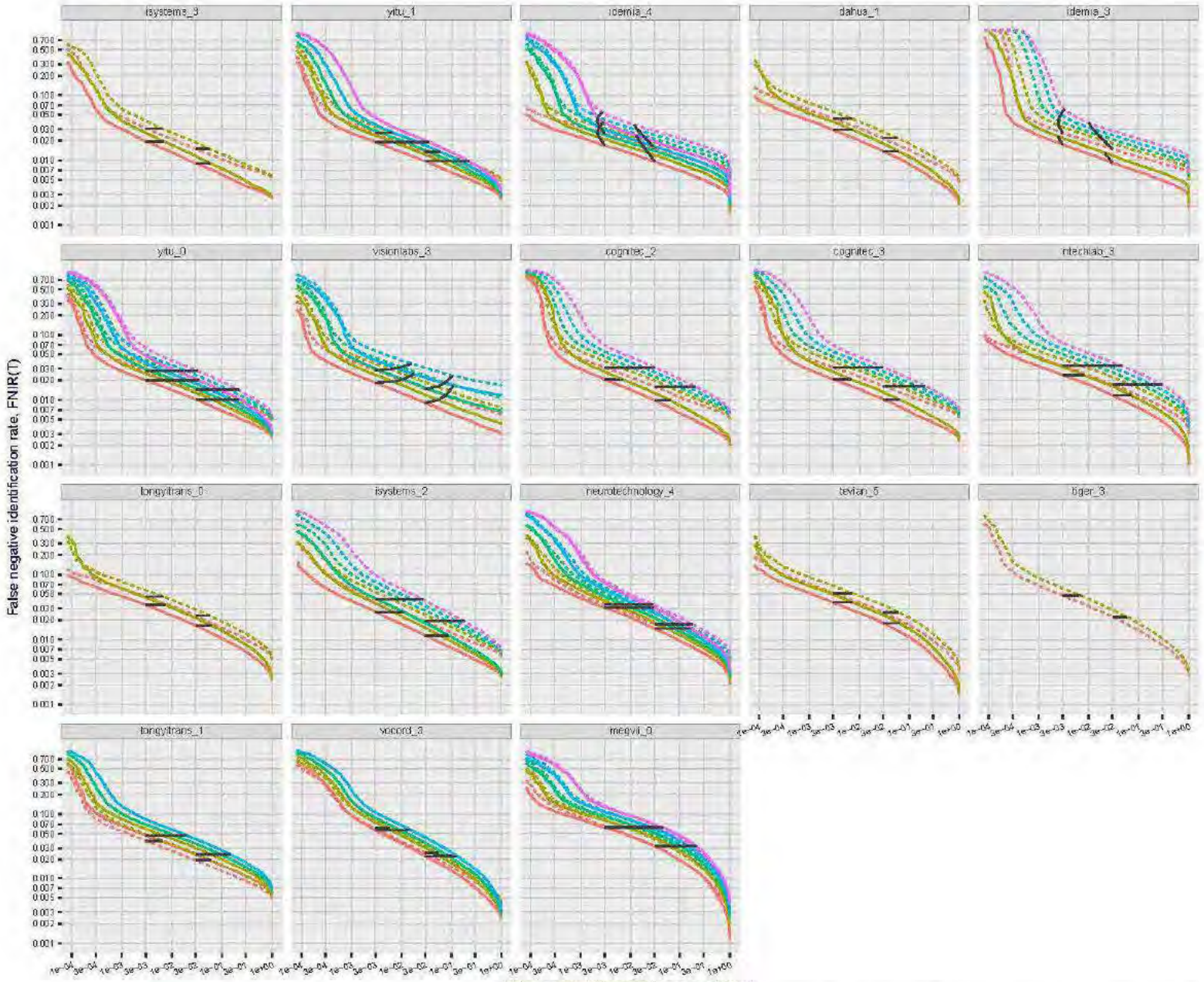
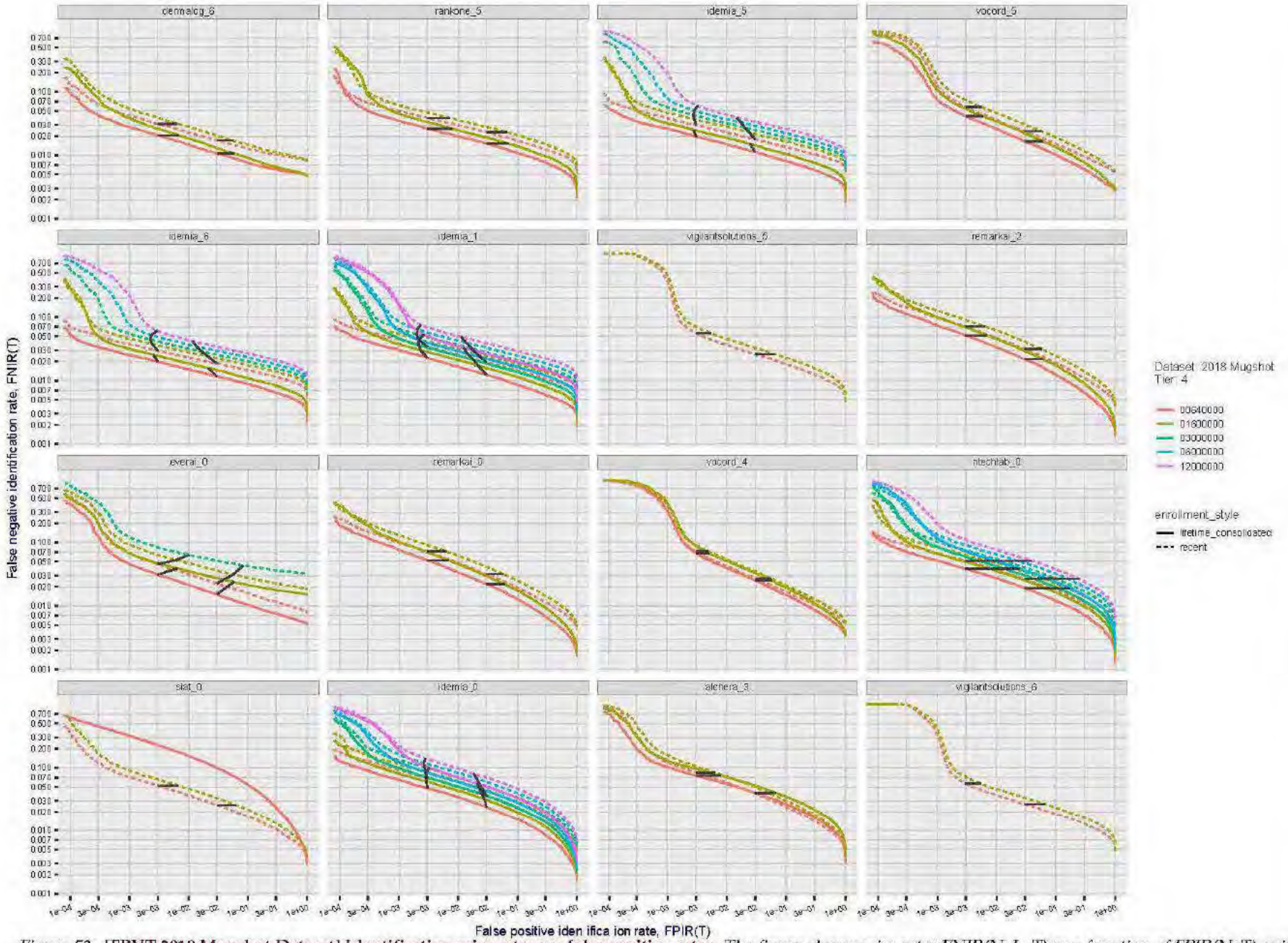


Figure 52: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification



2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

Figure 53: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

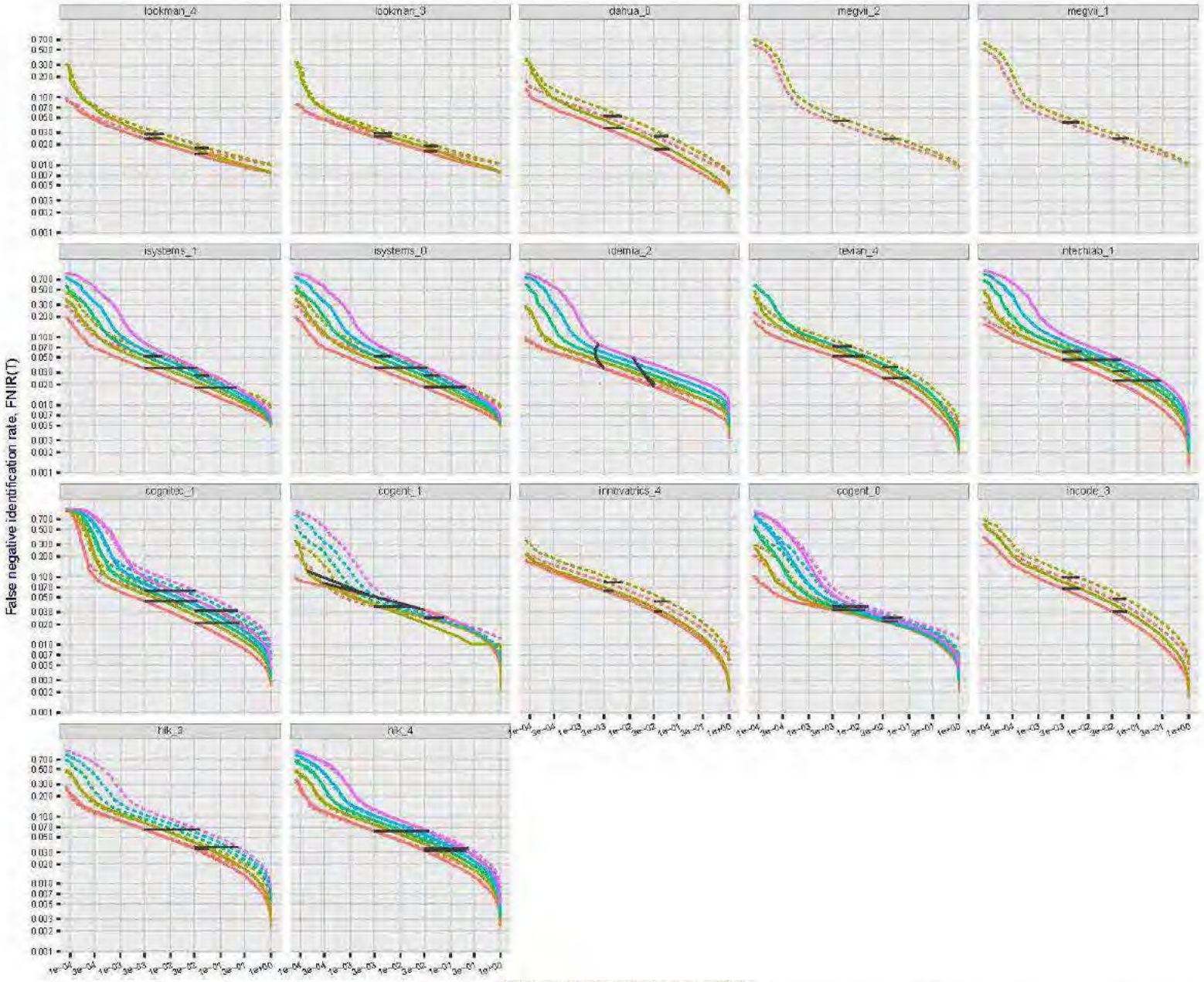


Figure 54: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

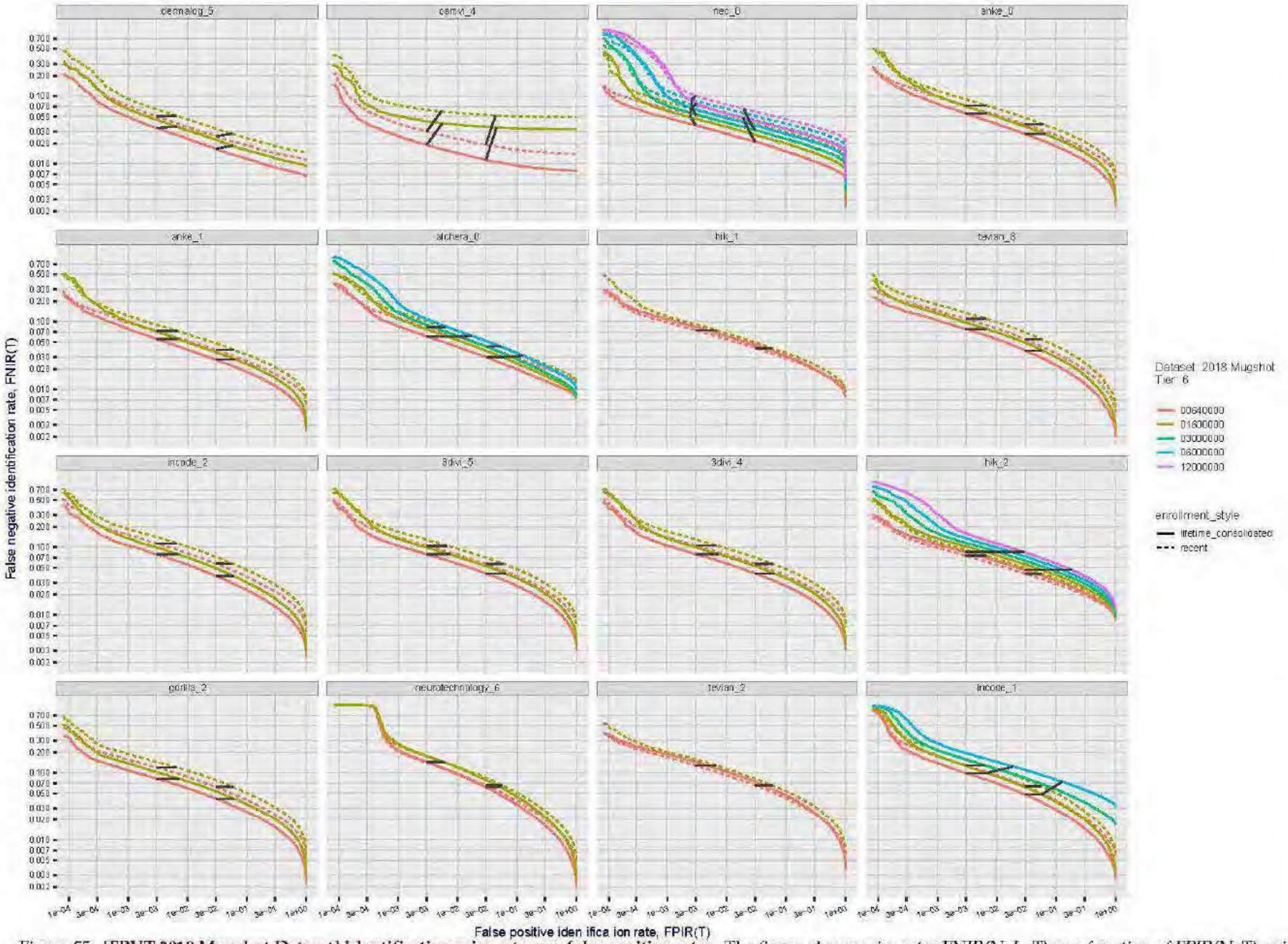


Figure 55: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

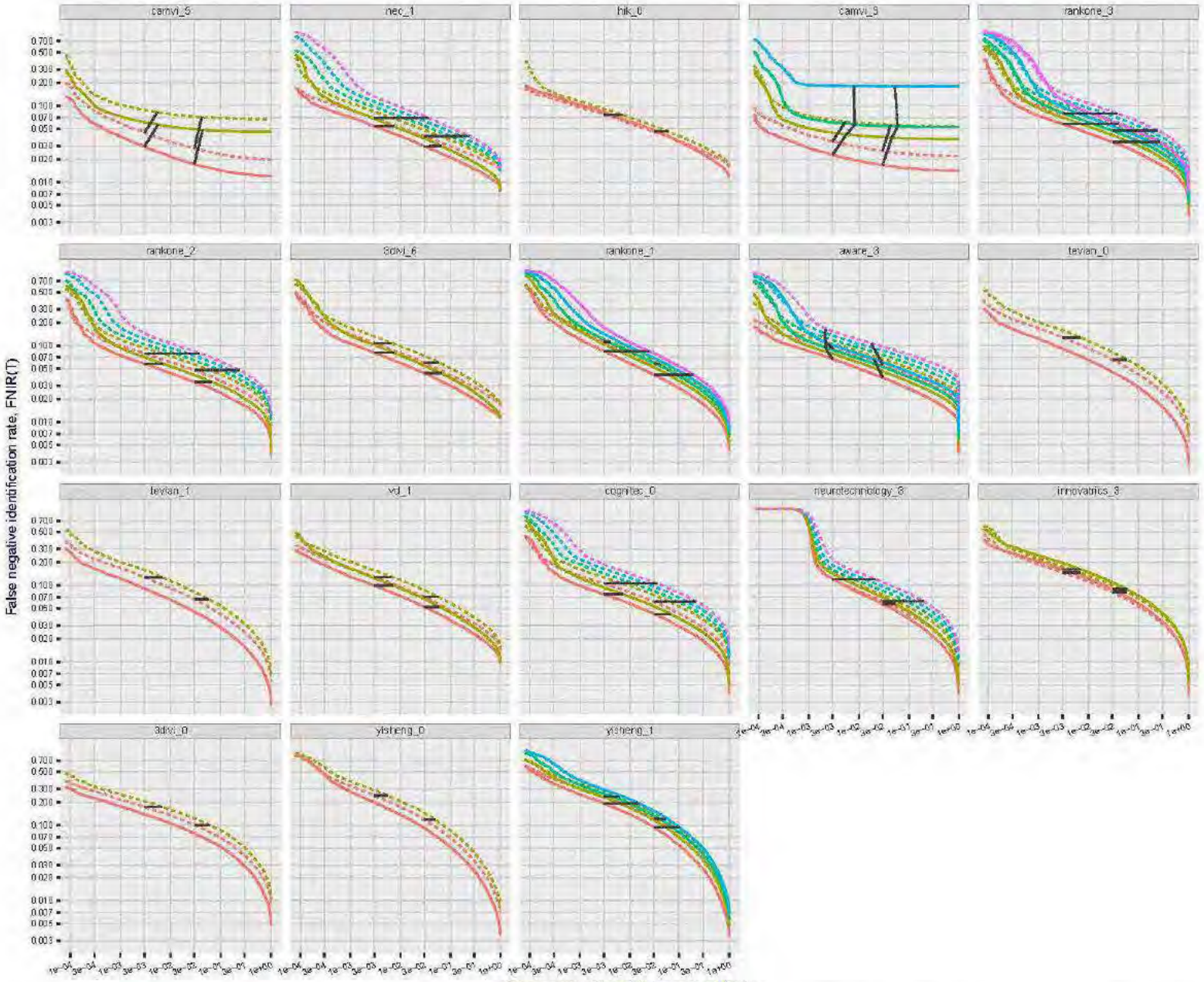


Figure 56: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

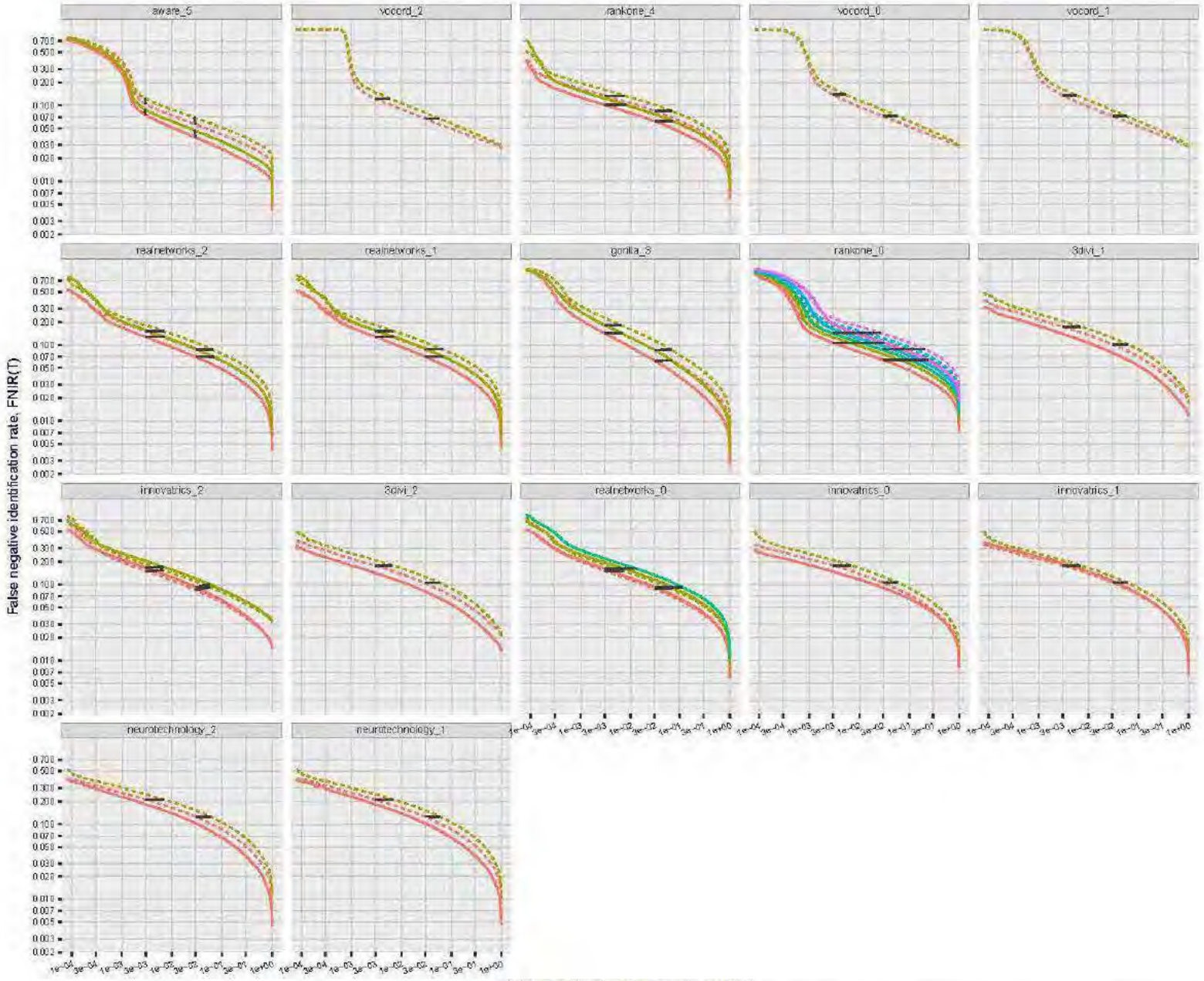


Figure 57: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

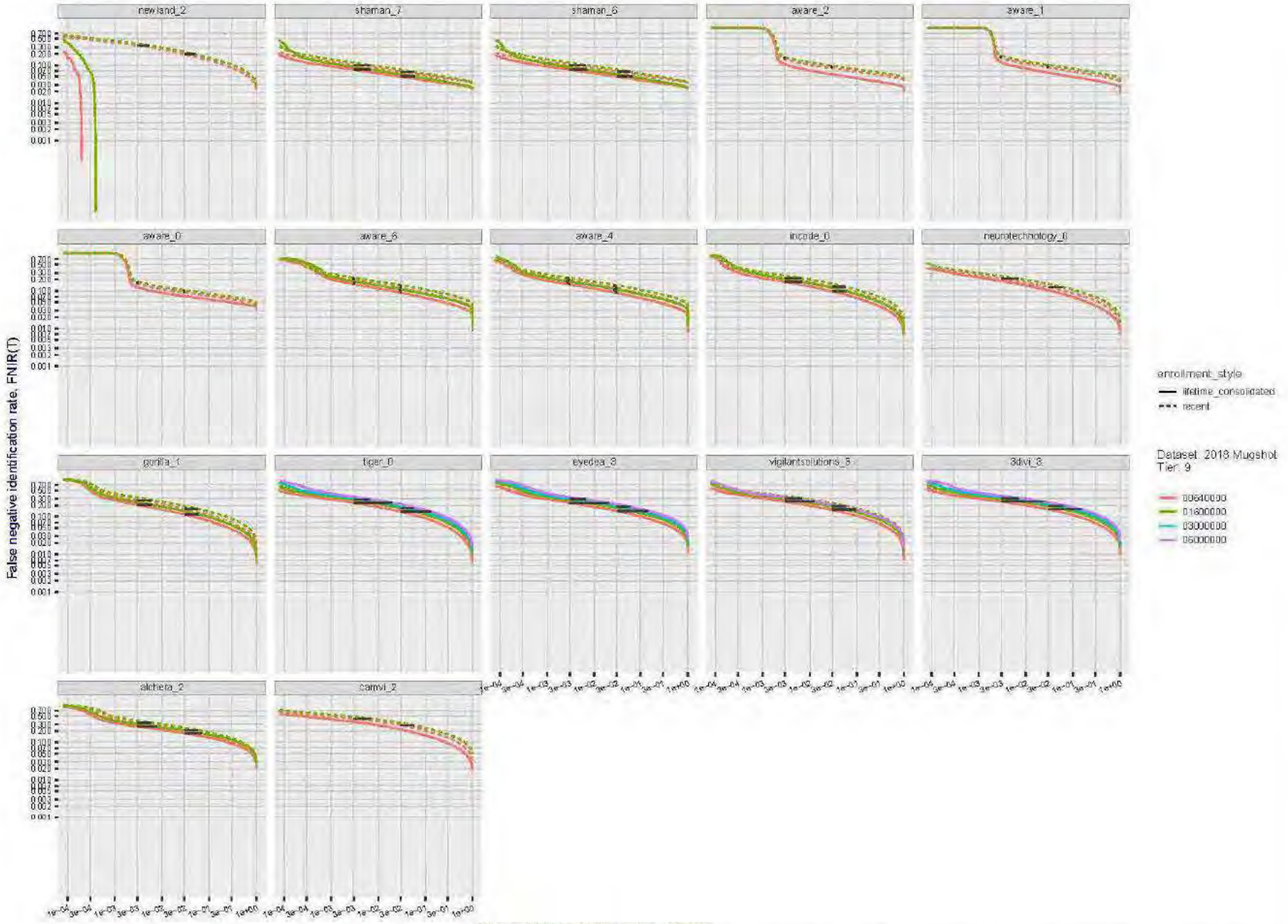


Figure 58: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

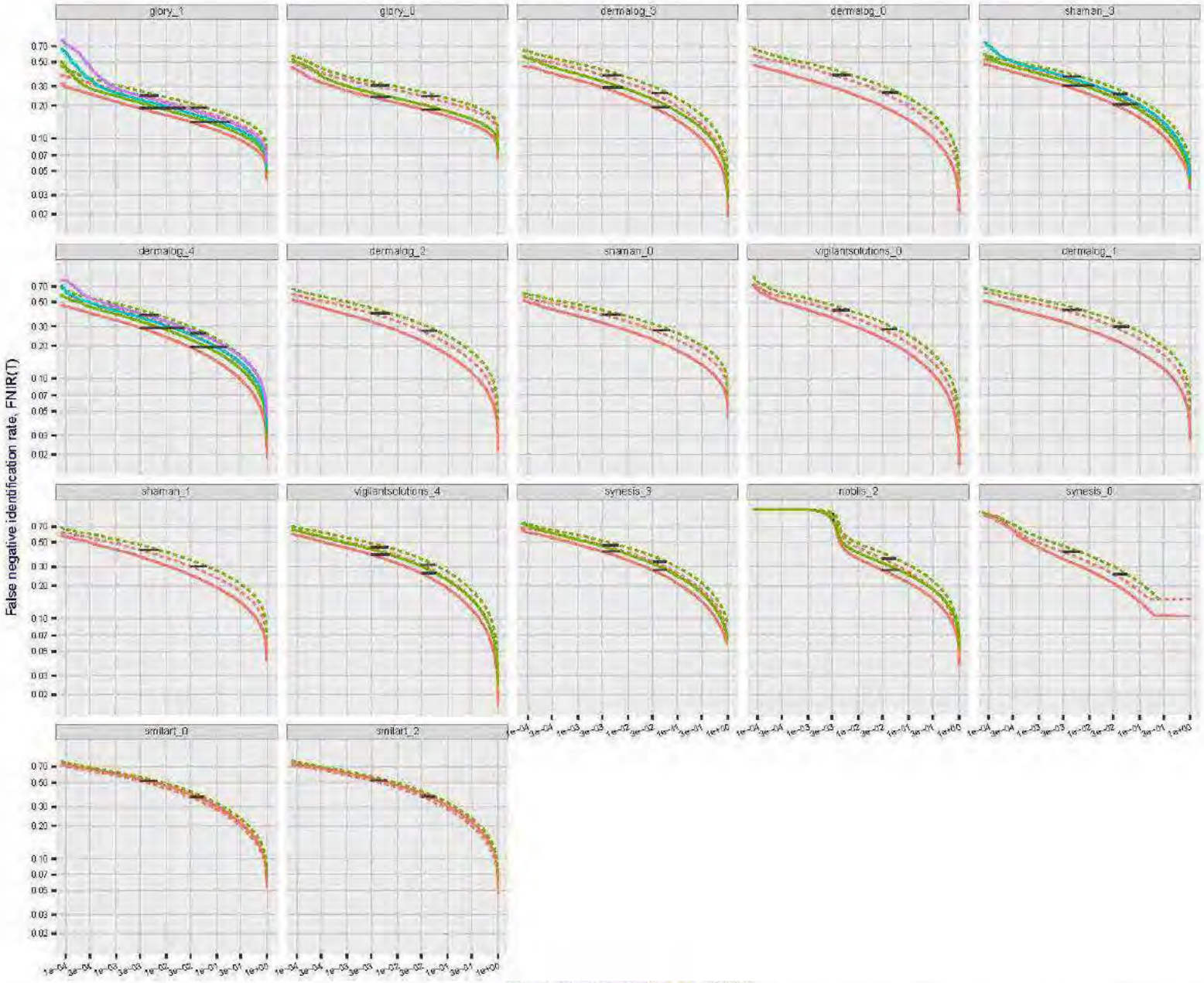


Figure 59: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

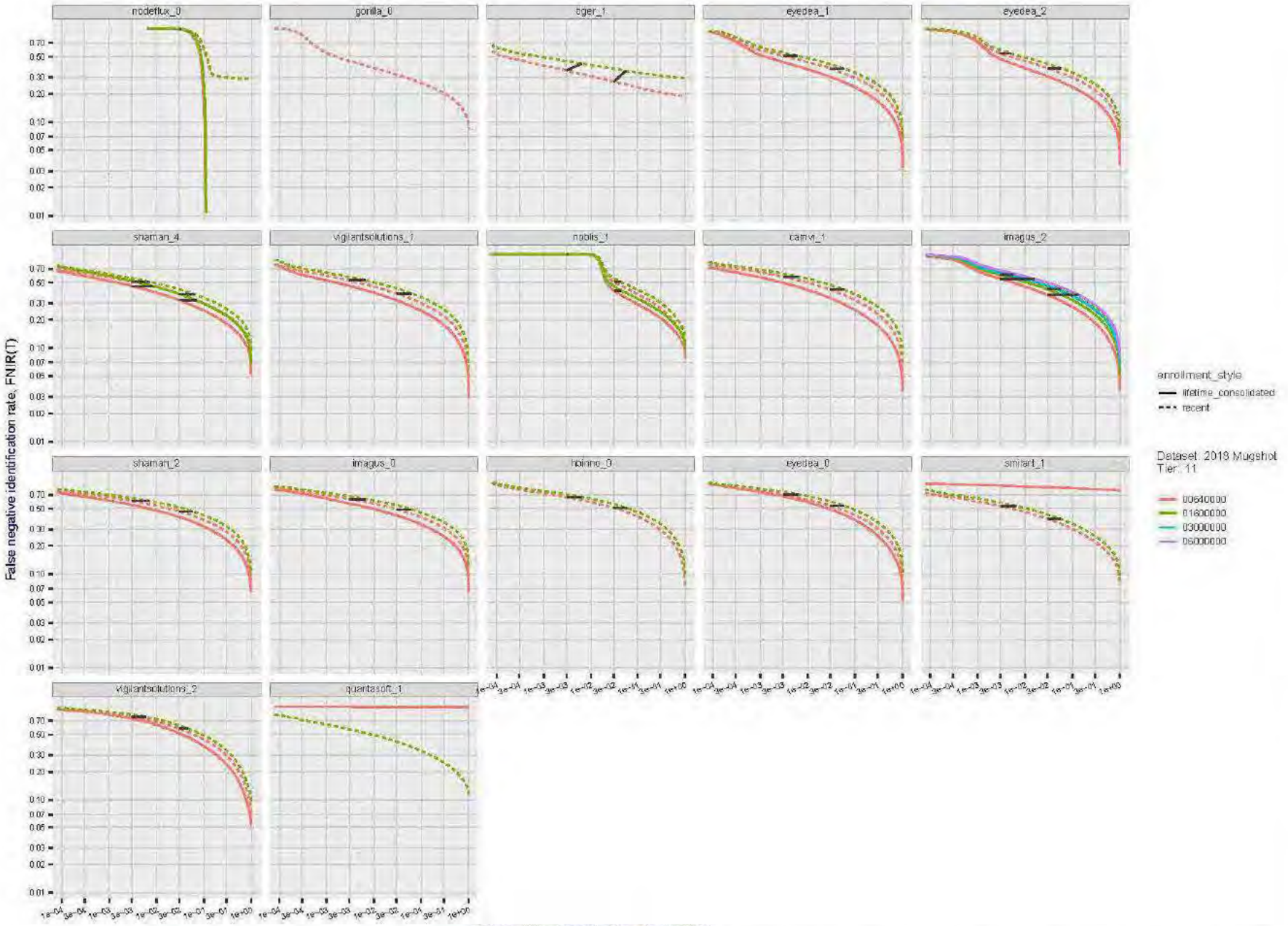


Figure 60: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

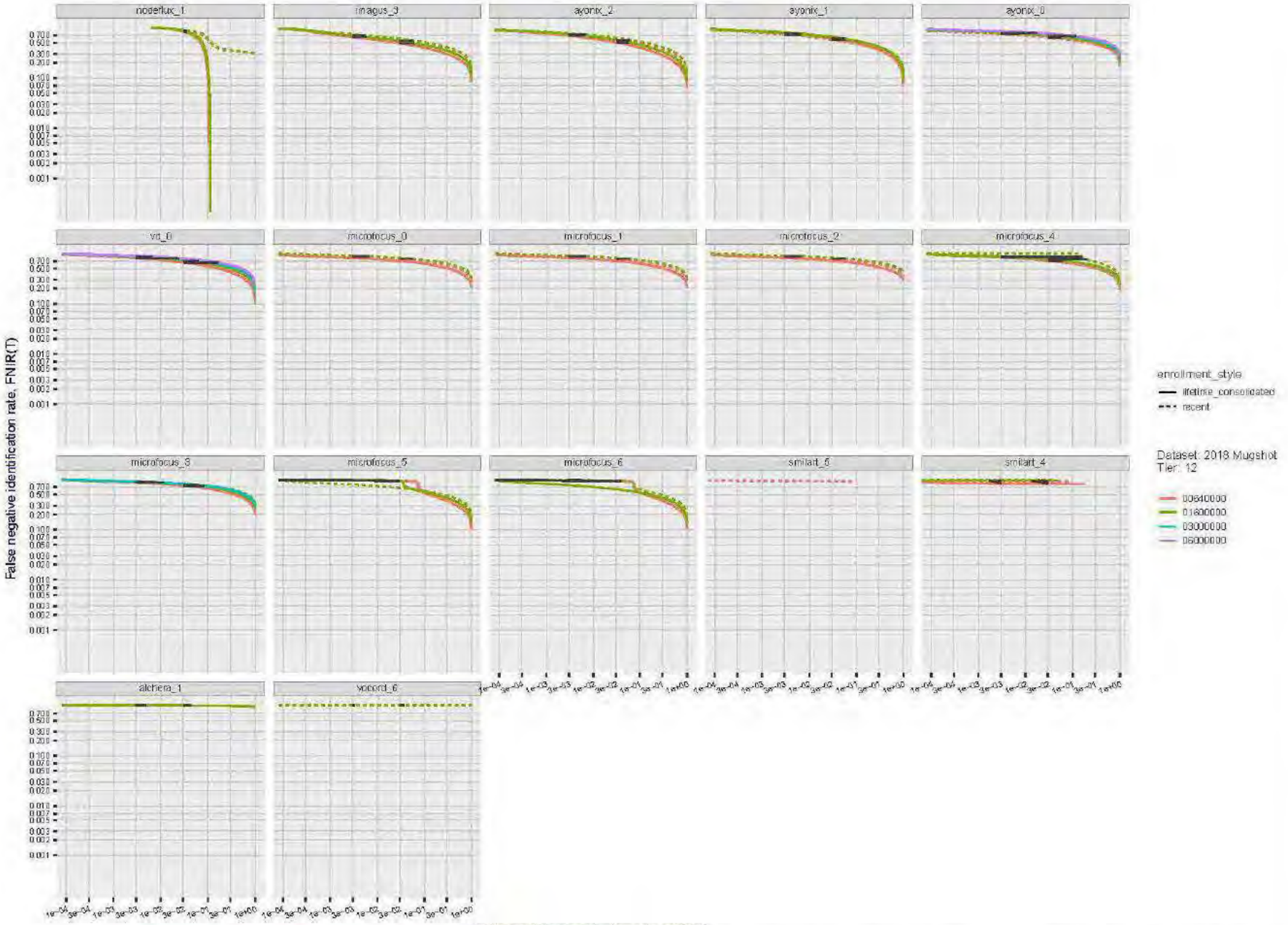


Figure 61: **[FRVT-2018 Mugshot Dataset] Identification miss rates vs. false positive rates.** *The figure shows miss rates FNIR(N, L, T) as a function of FPIR(N, T), with N ranging from 640 000 to 12 000 000 as noted in rows 1-10 of Table 5. These error tradeoff characteristics are useful for applications where a threshold must be elevated to limit false positives, such as when human reviewer labor is not matched to the volume of searches. Dark lines join points of equal threshold: If horizontal, FPIR(T) rises with N, and mate scores are independent of N. Other algorithms adjust scores in an attempt to make FPIR independent of N.*

Appendix B Effect of time-lapse: Accuracy after face ageing

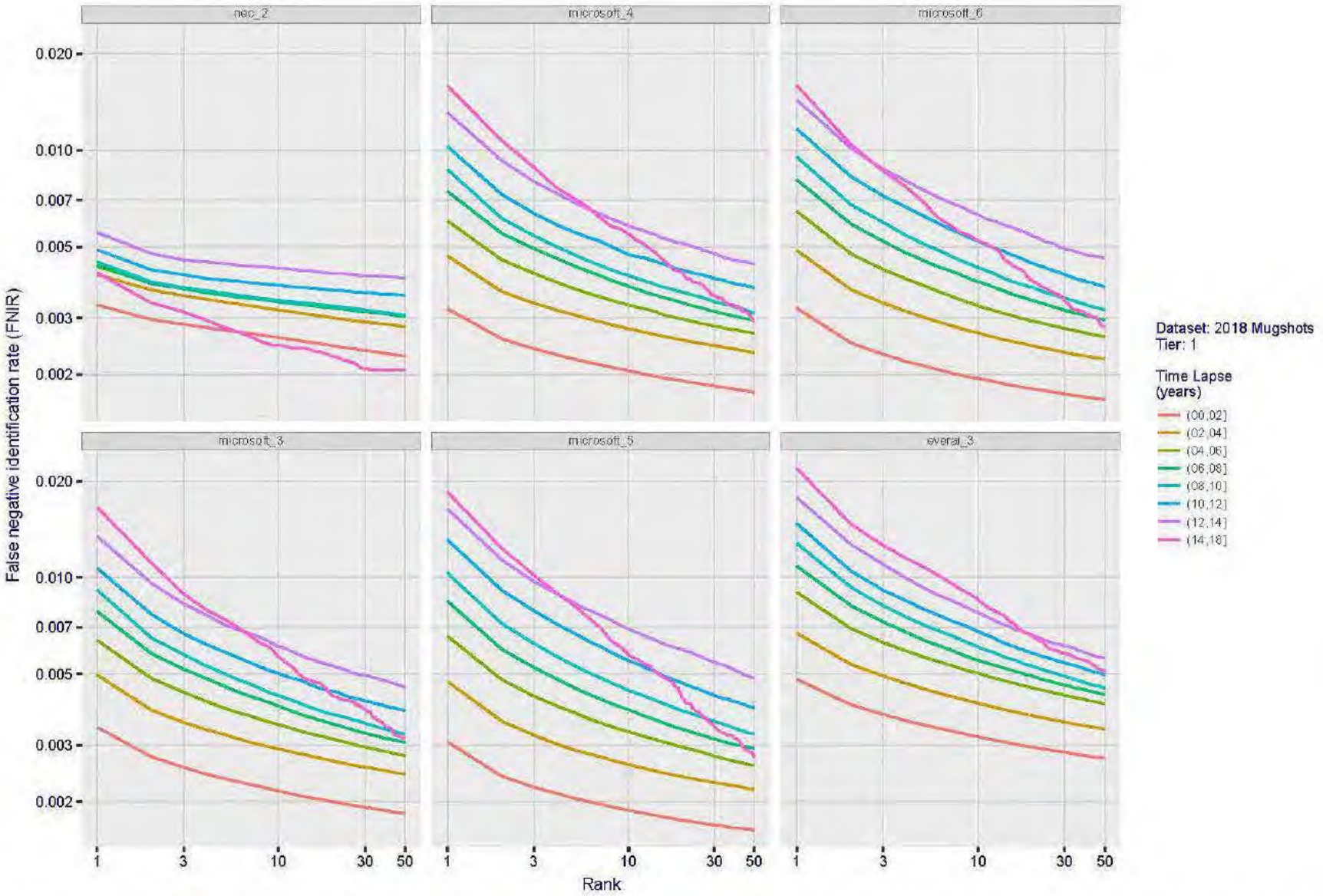


Figure 62: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

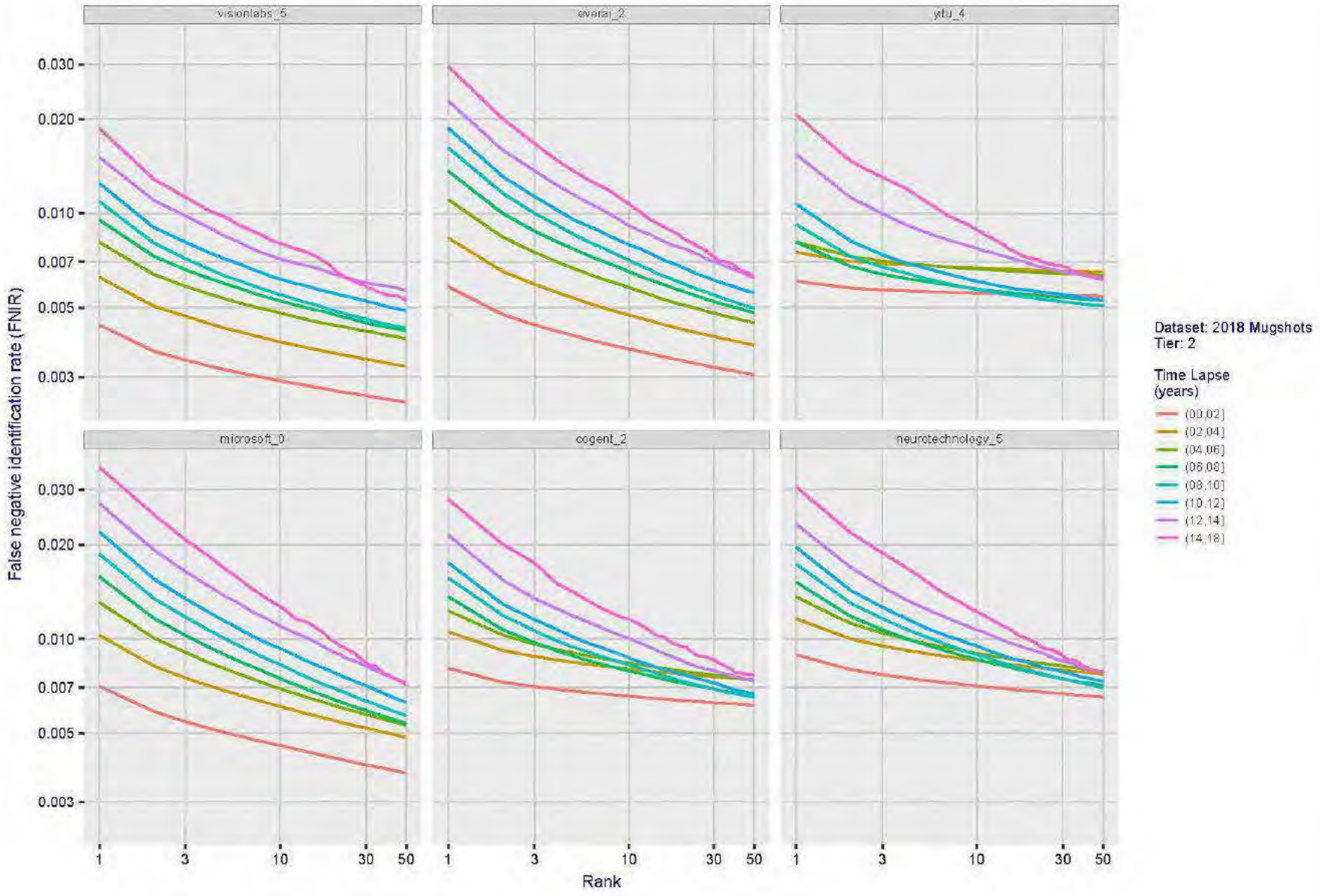


Figure 63: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

2019/09/11
17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

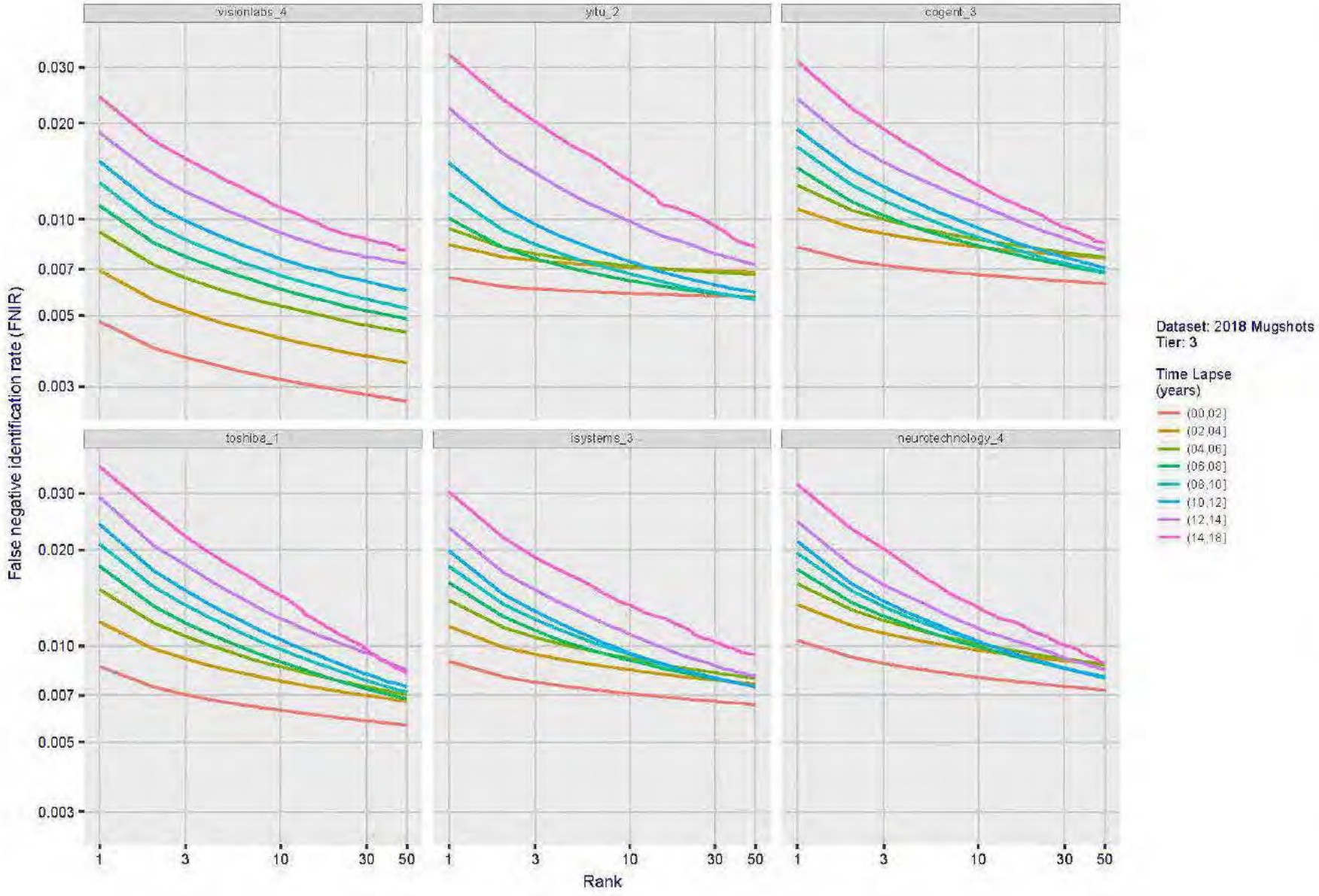


Figure 64: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

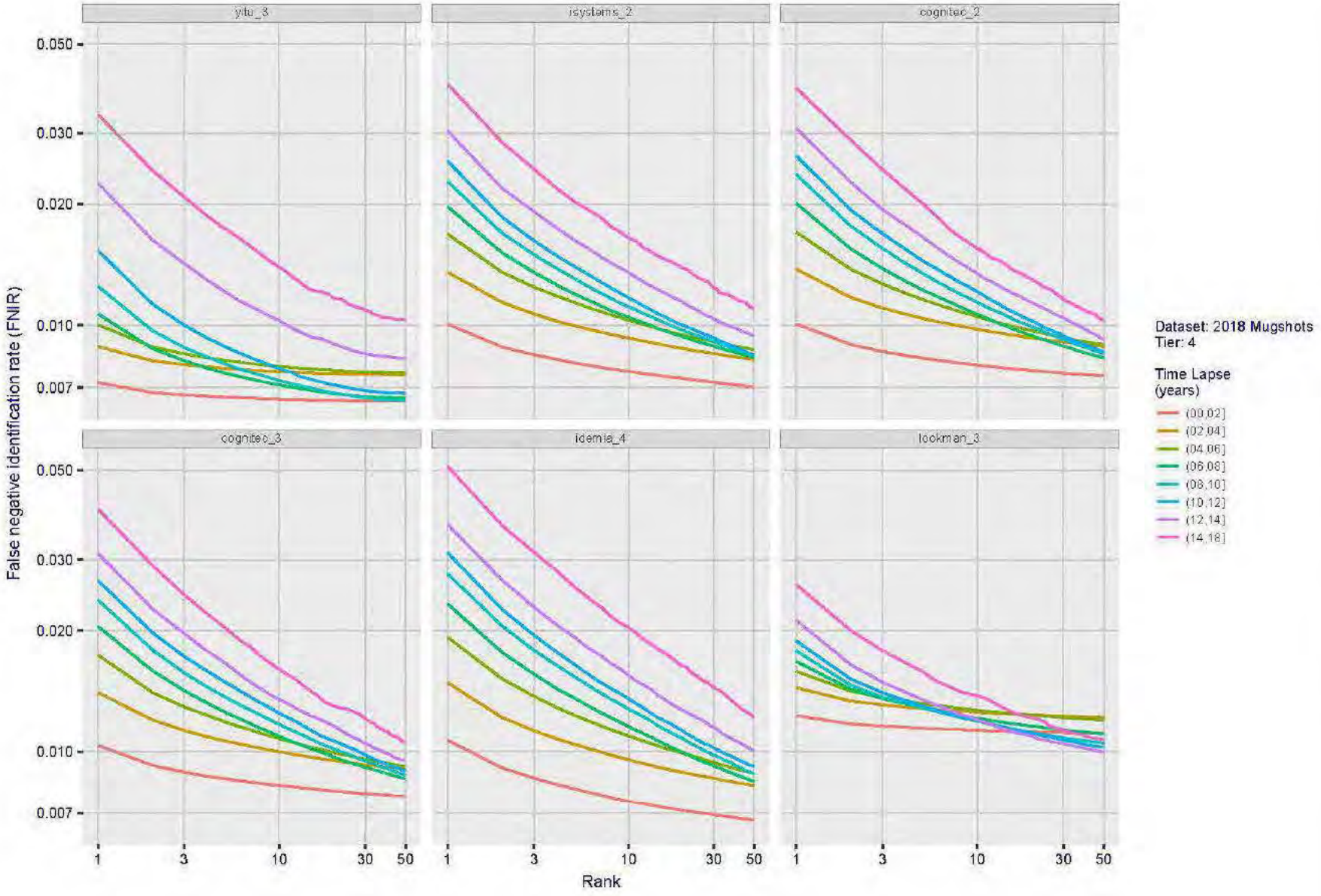


Figure 65: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

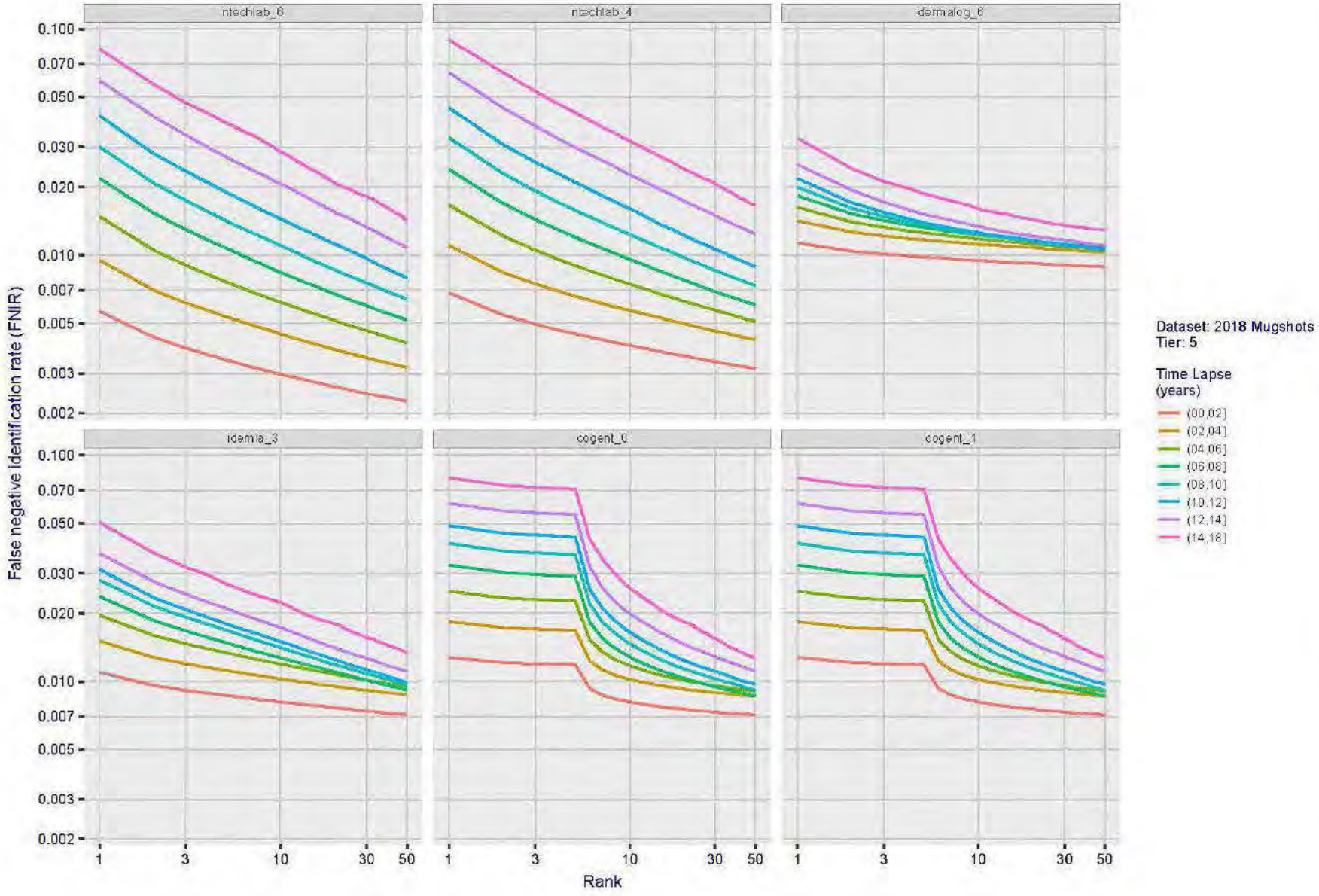


Figure 66: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

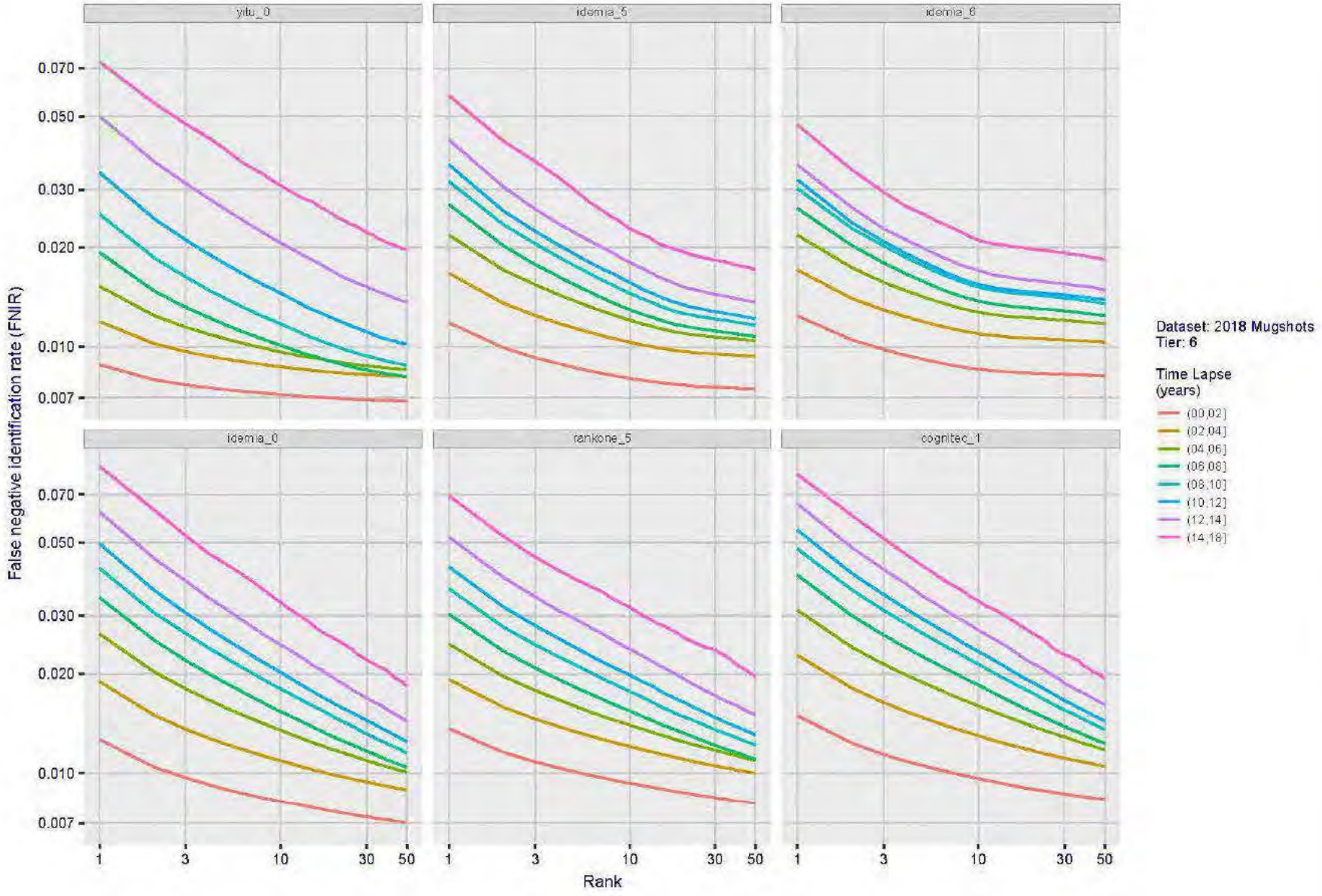


Figure 67: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

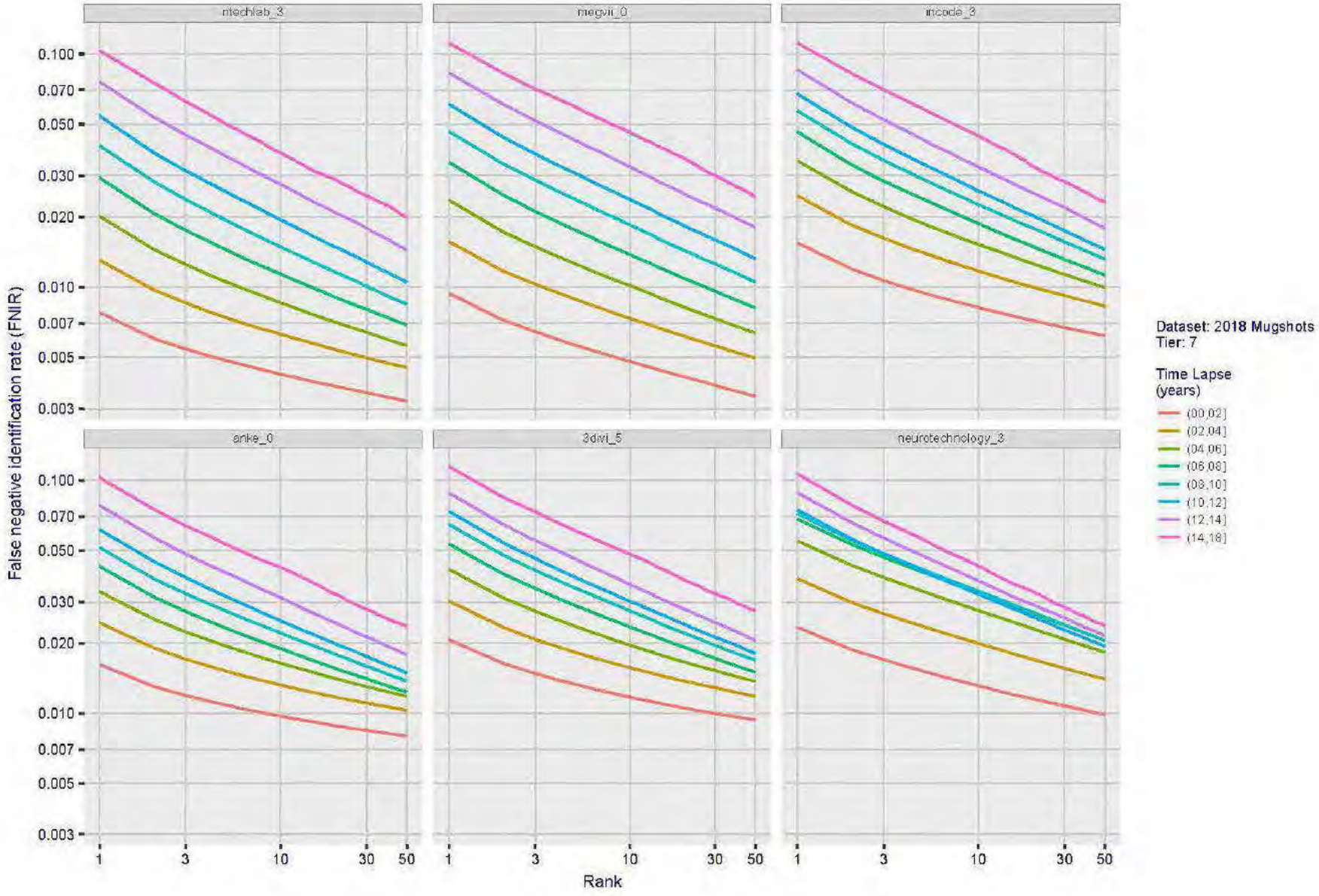


Figure 68: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

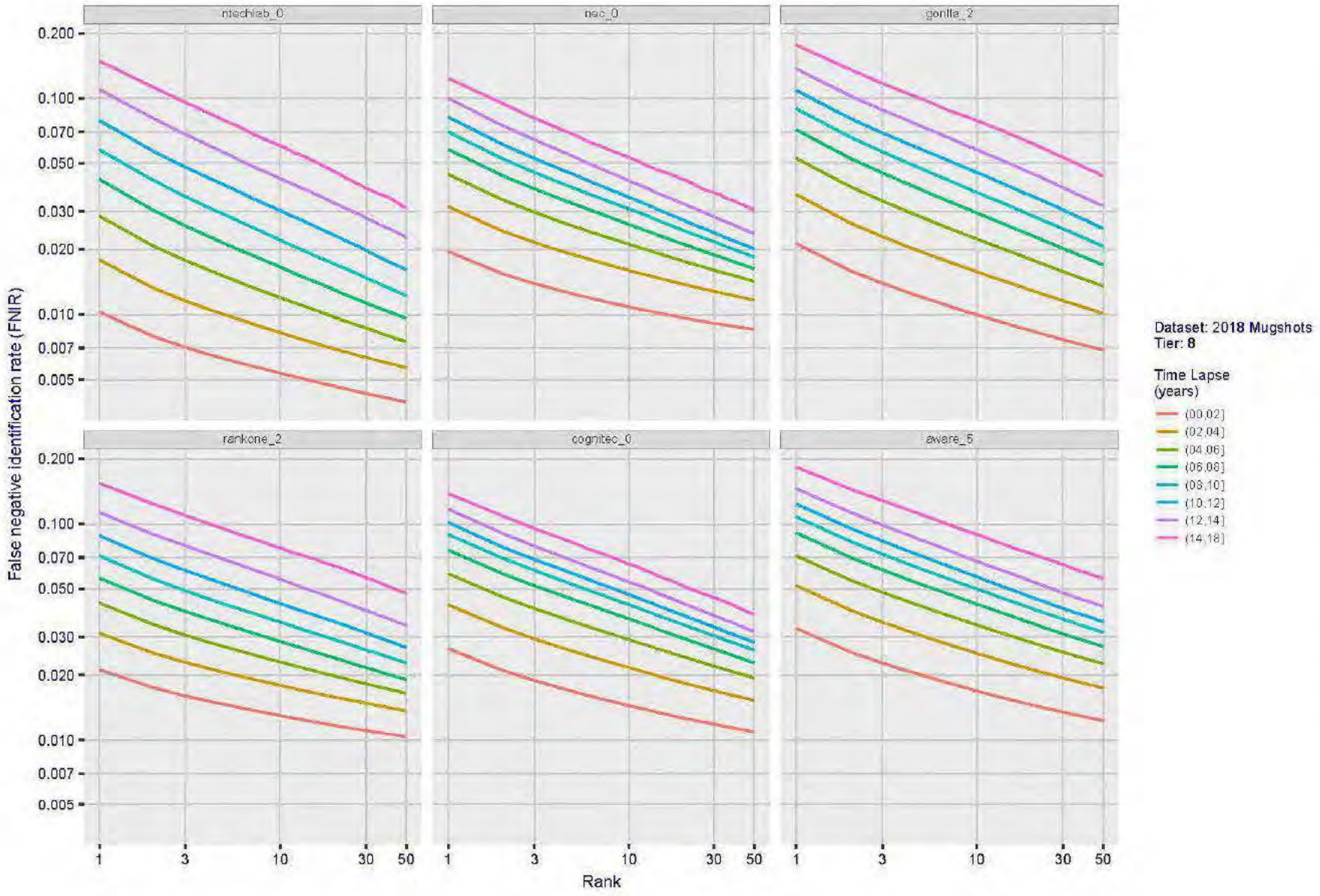


Figure 69: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

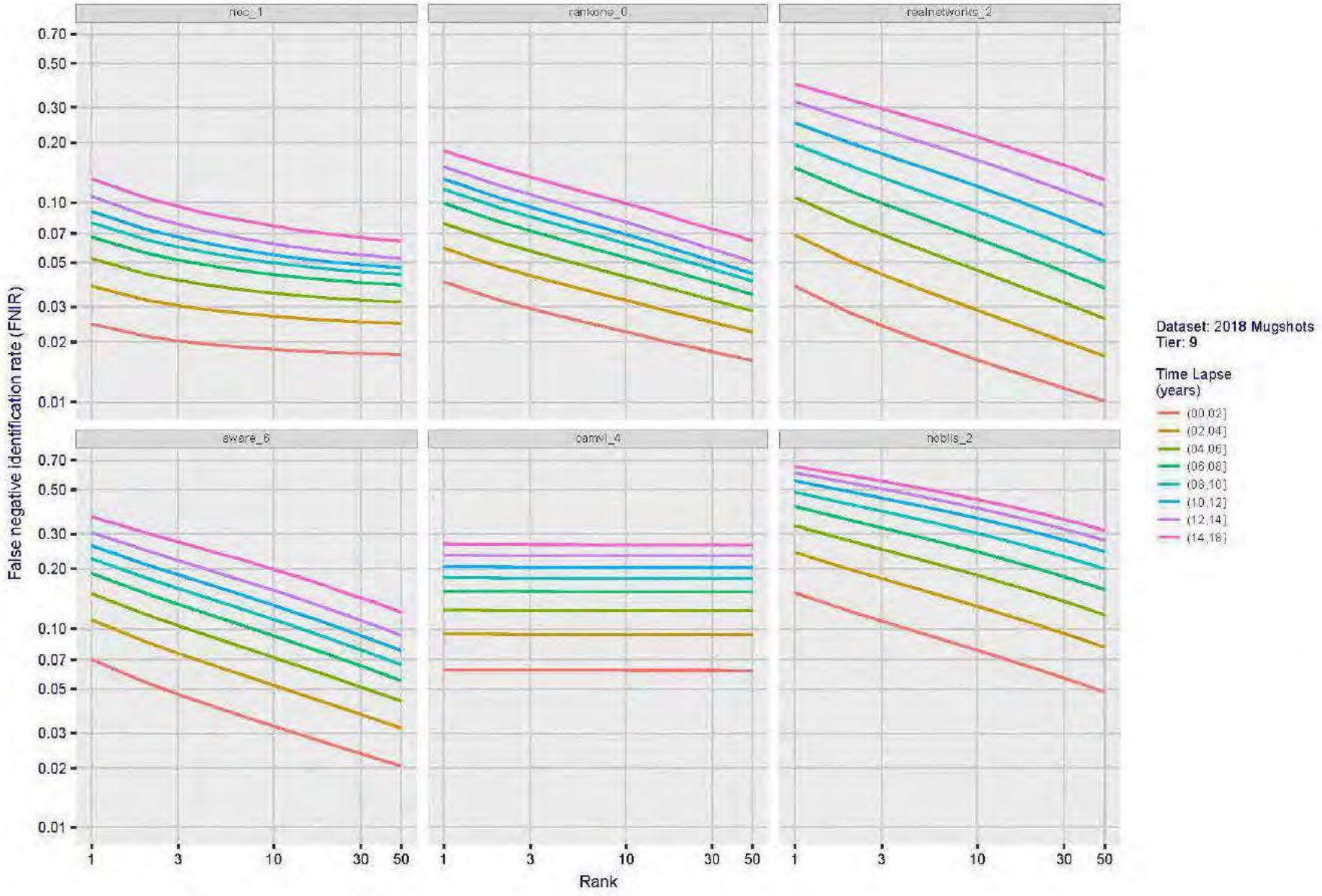


Figure 70: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

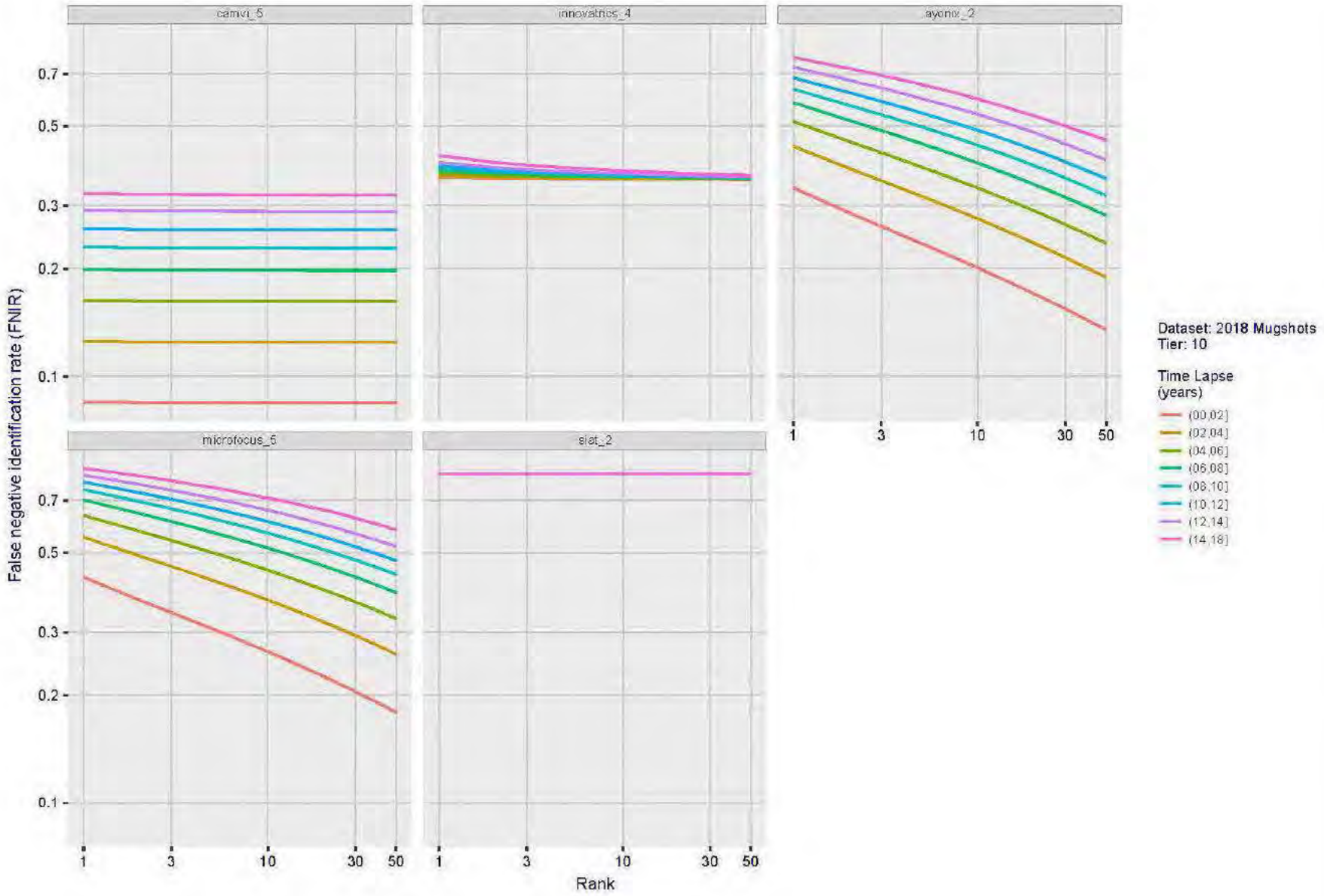


Figure 71: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. rank by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

2019/09/11 17:24:52	FNIR(N, R, T) = False neg. identification rate FPIR(N, T) = False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification	

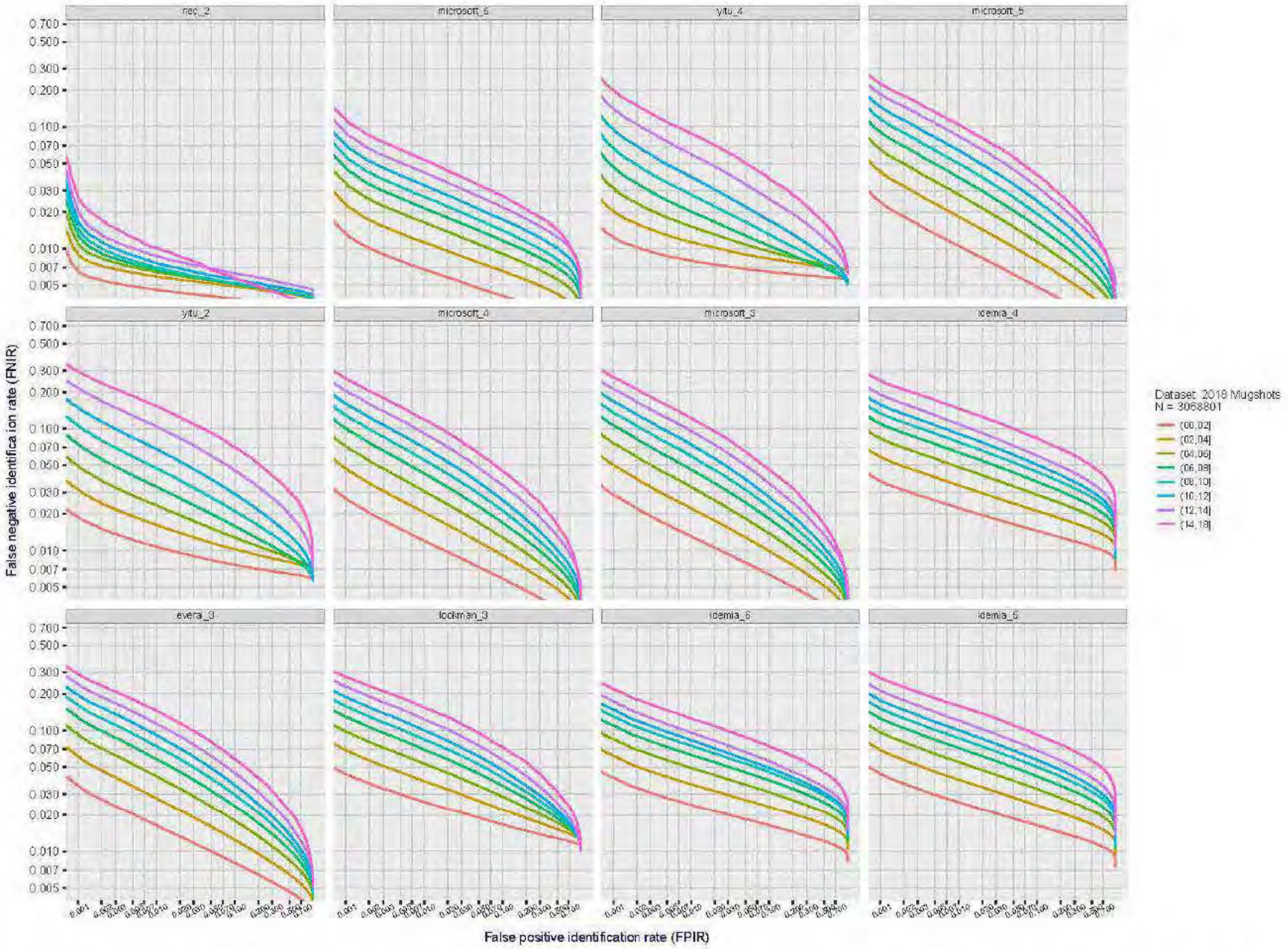


Figure 72: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. FPIR by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment. FPIR is computed from the same FRVT 2018 non-mates noted in row 3 of Table 5 with N = 3 000 000.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

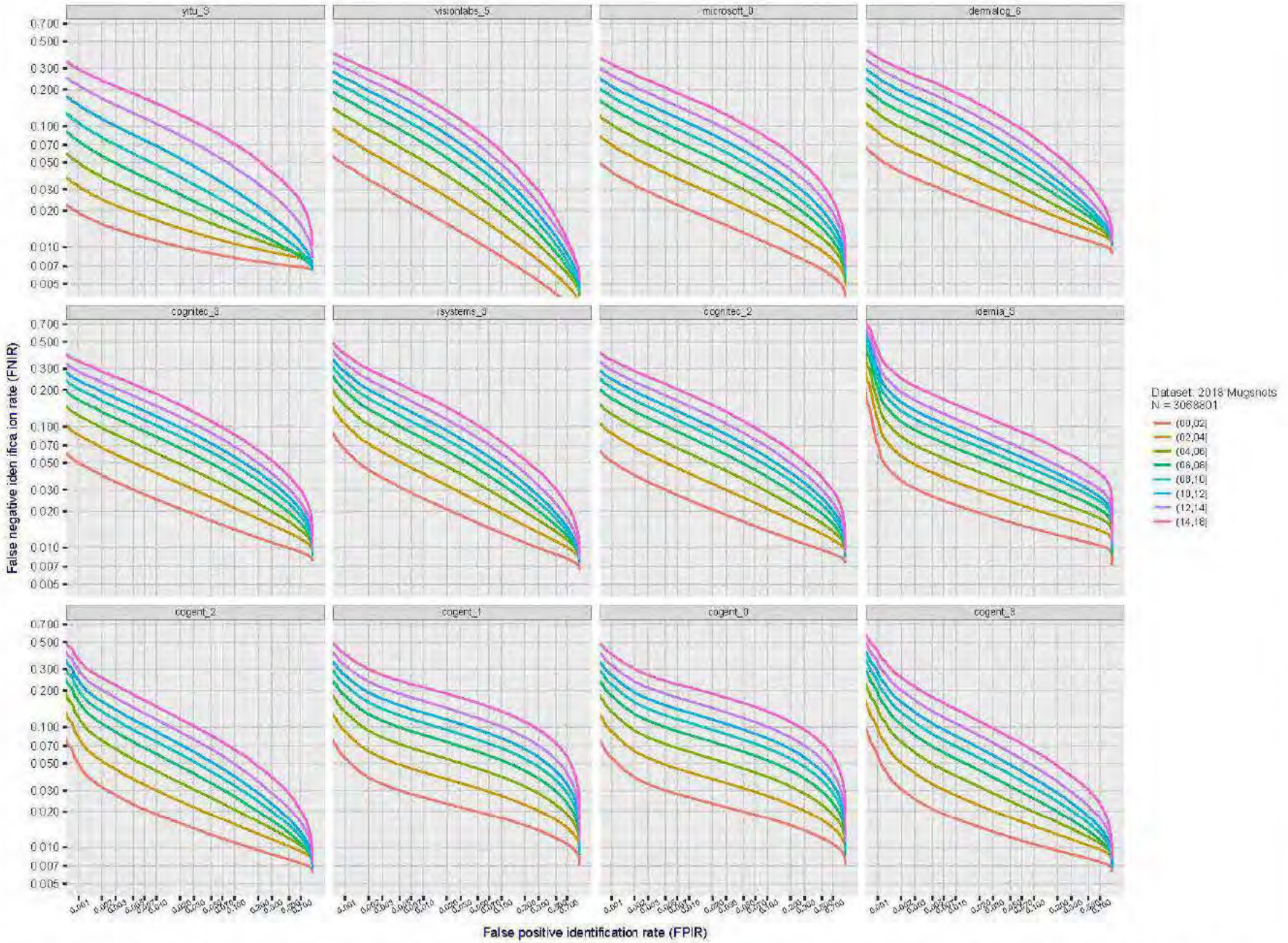


Figure 73: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. FPIR by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment. FPIR is computed from the same FRVT 2018 non-mates noted in row 3 of Table 5 with N = 3 000 000.*

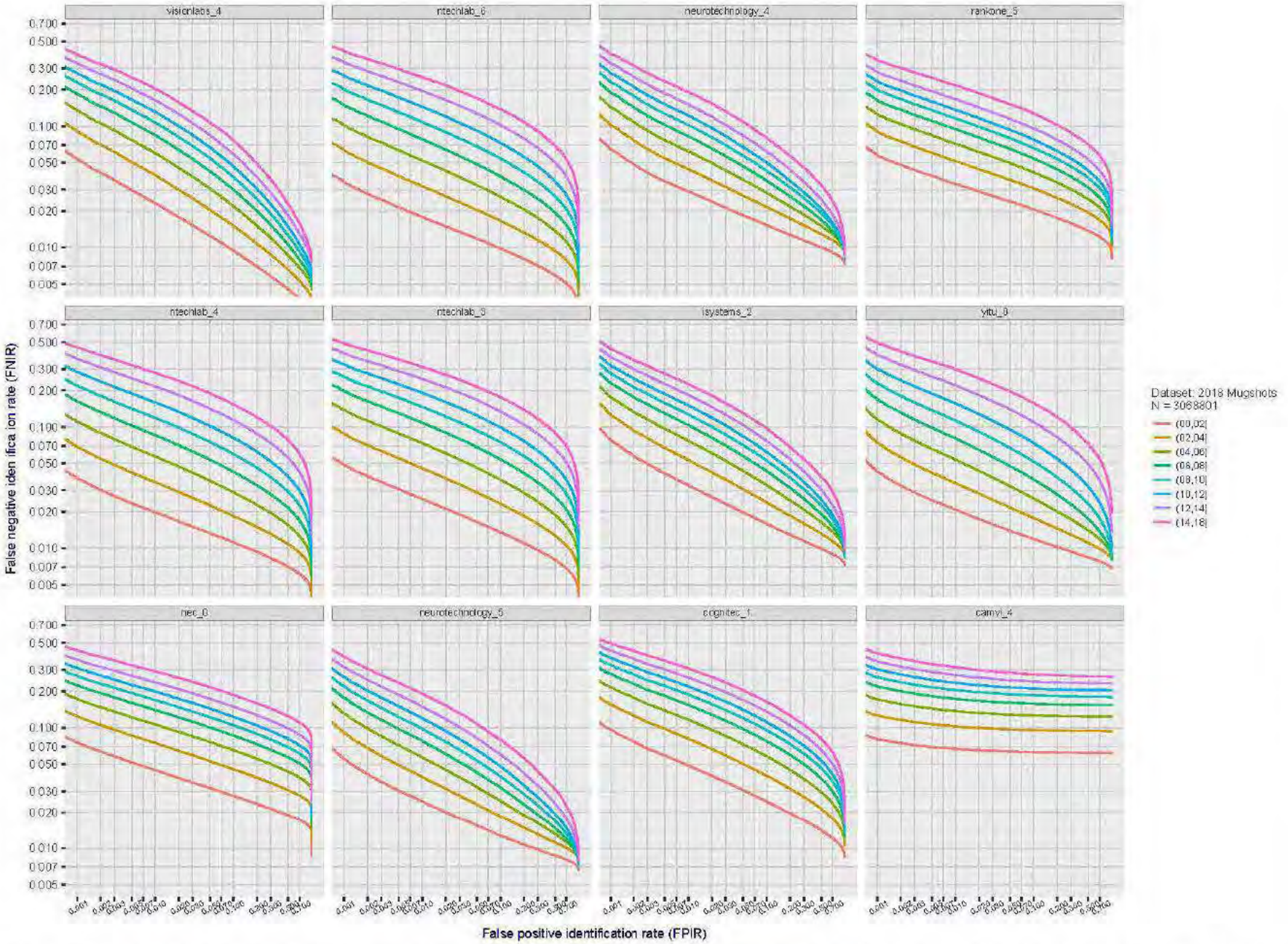


Figure 74: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. FPIR by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment. FPIR is computed from the same FRVT 2018 non-mates noted in row 3 of Table 5 with N = 3 000 000.*

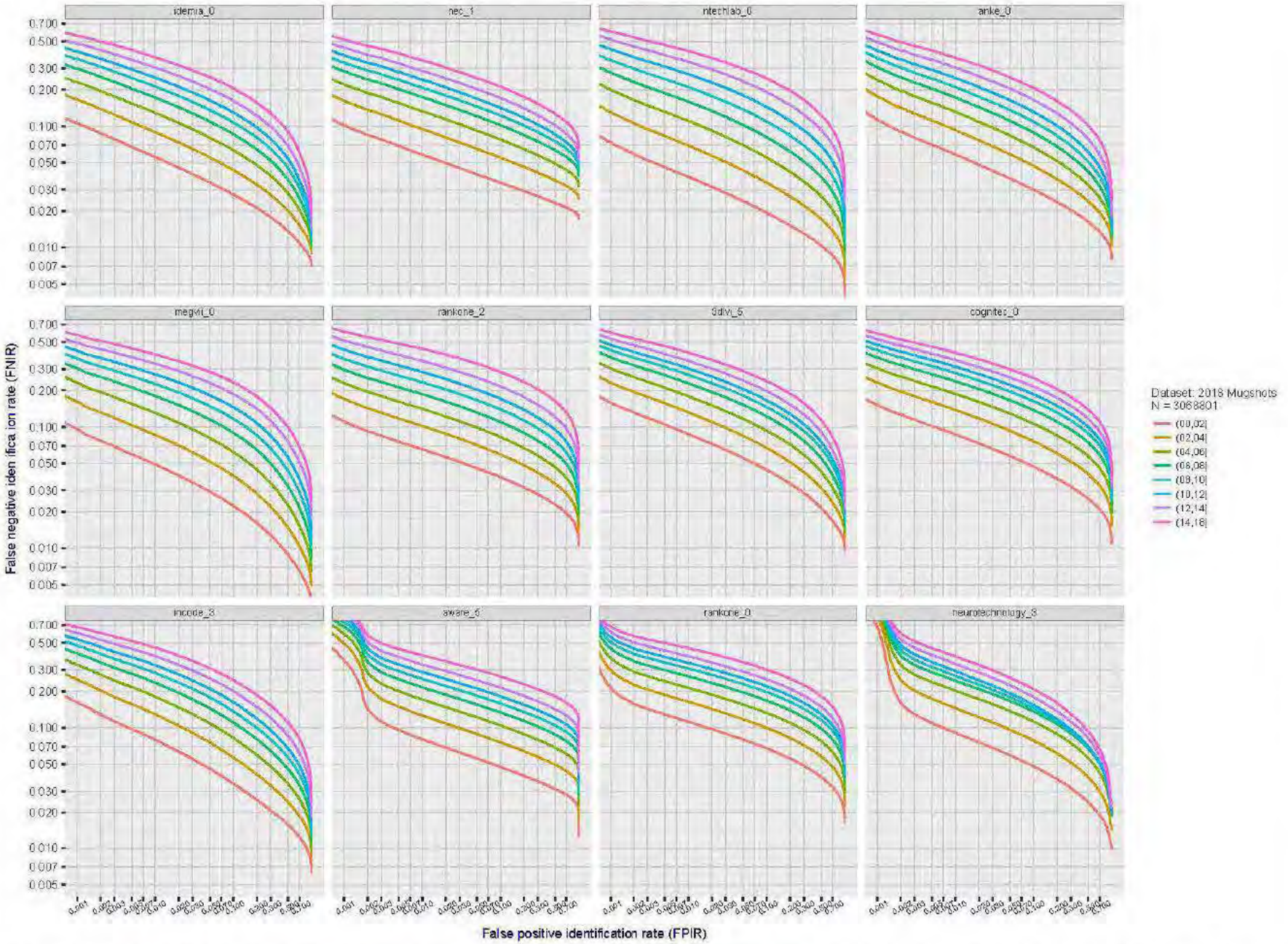


Figure 75: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. FPIR by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment. FPIR is computed from the same FRVT 2018 non-mates noted in row 3 of Table 5 with N = 3 000 000.*

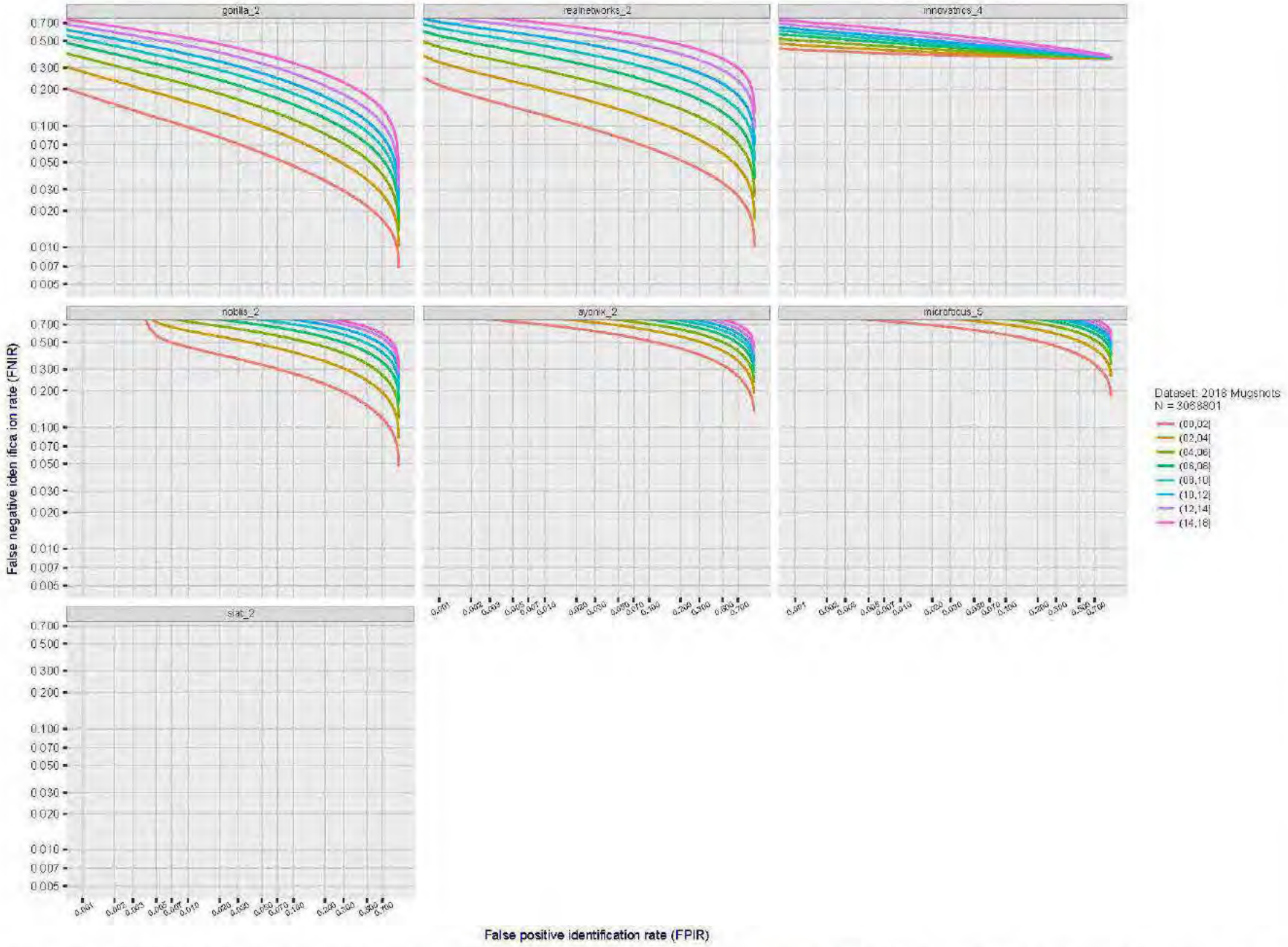


Figure 76: **[FRVT-2018 Mugshot Ageing Dataset] Identification miss rates vs. FPIR by time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Miss rates are computed over all searches noted in row 17 of Table 5 and binned by number of years between search and initial enrollment. FPIR is computed from the same FRVT 2018 non-mates noted in row 3 of Table 5 with N = 3 000 000.*

2019/09/11 17:24:52	FNIR(N, R, T) = False neg. identification rate FPIR(N, T) = False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification	

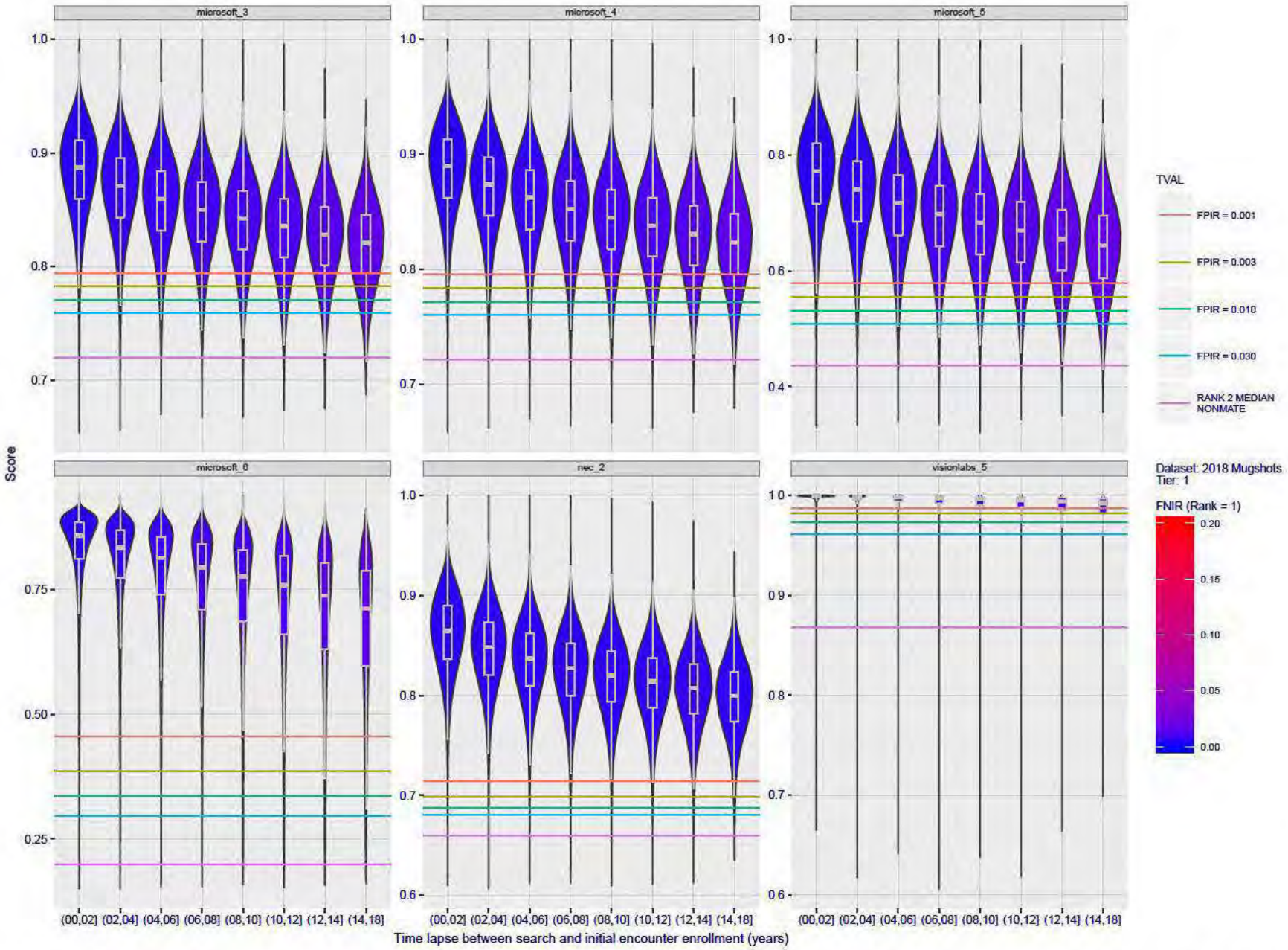


Figure 77: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

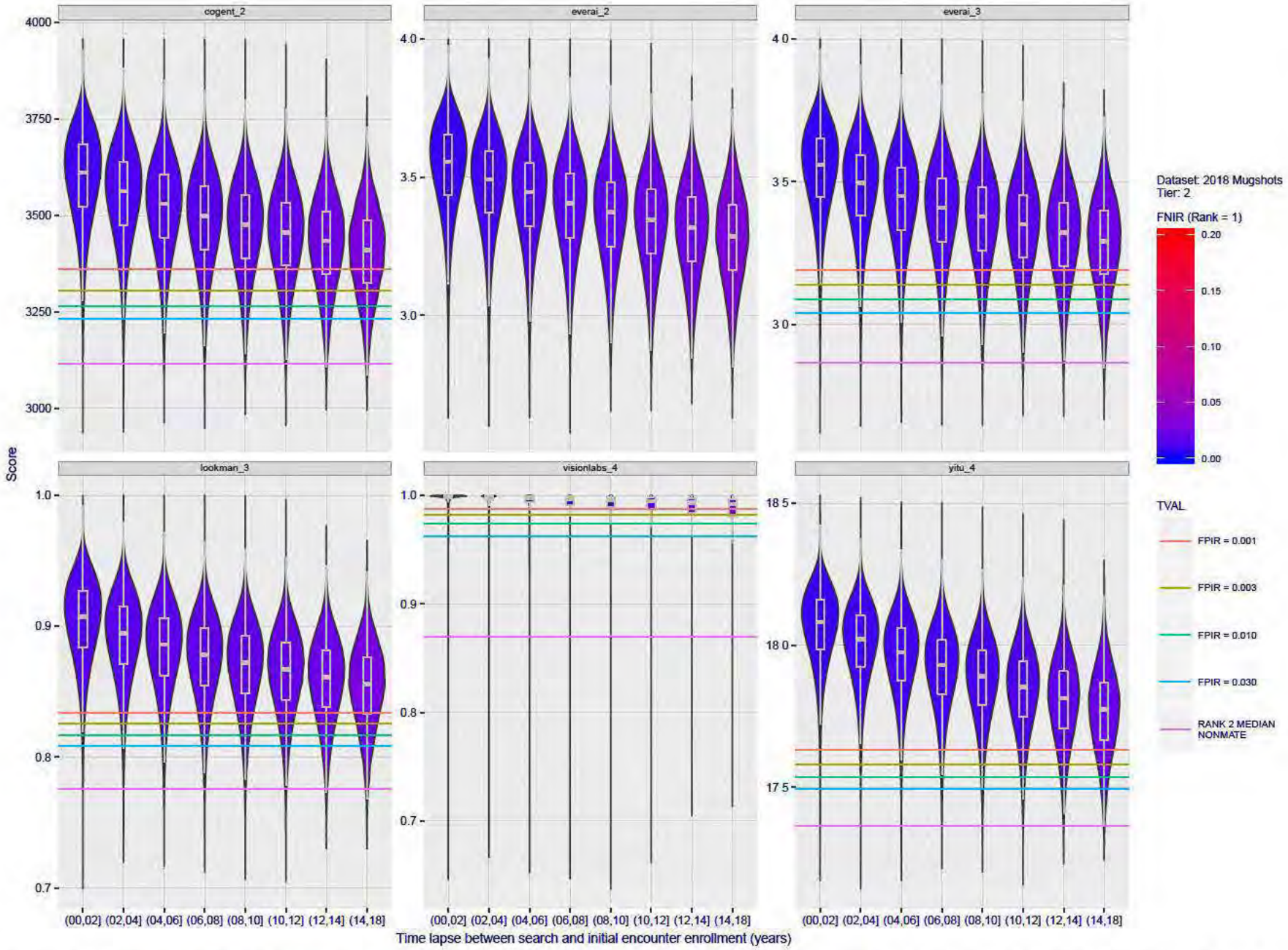


Figure 78: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

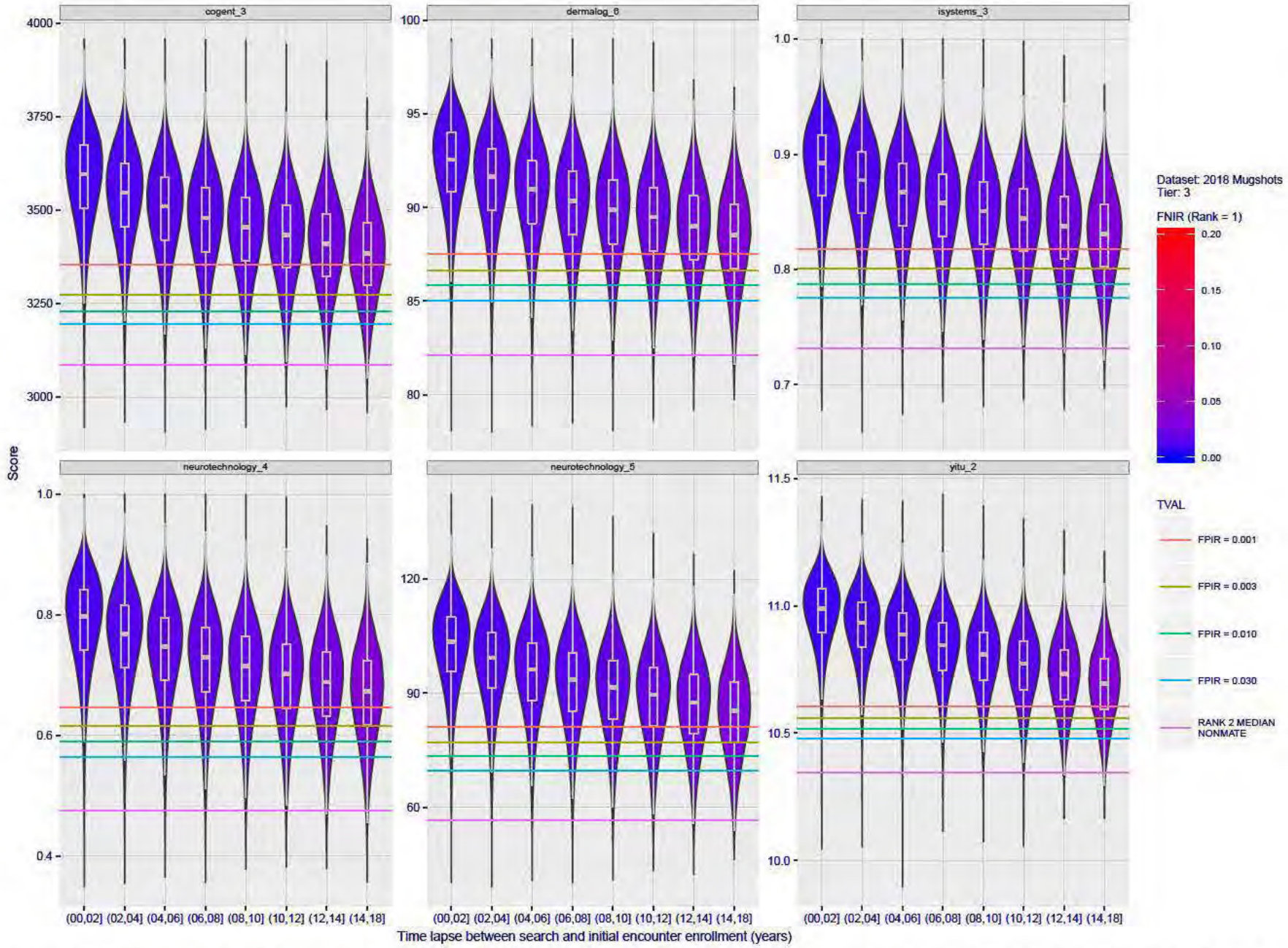


Figure 79: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

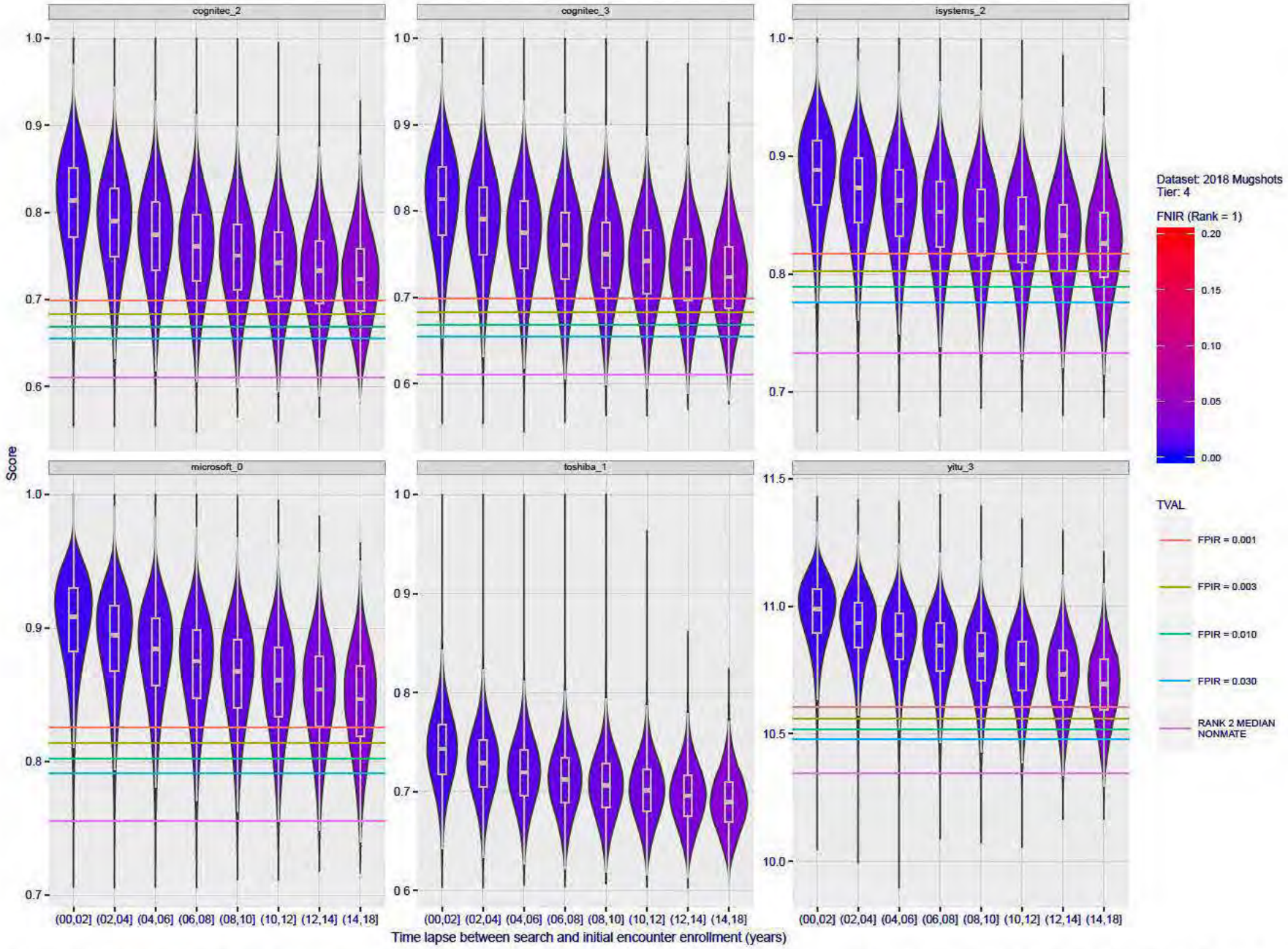


Figure 80: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

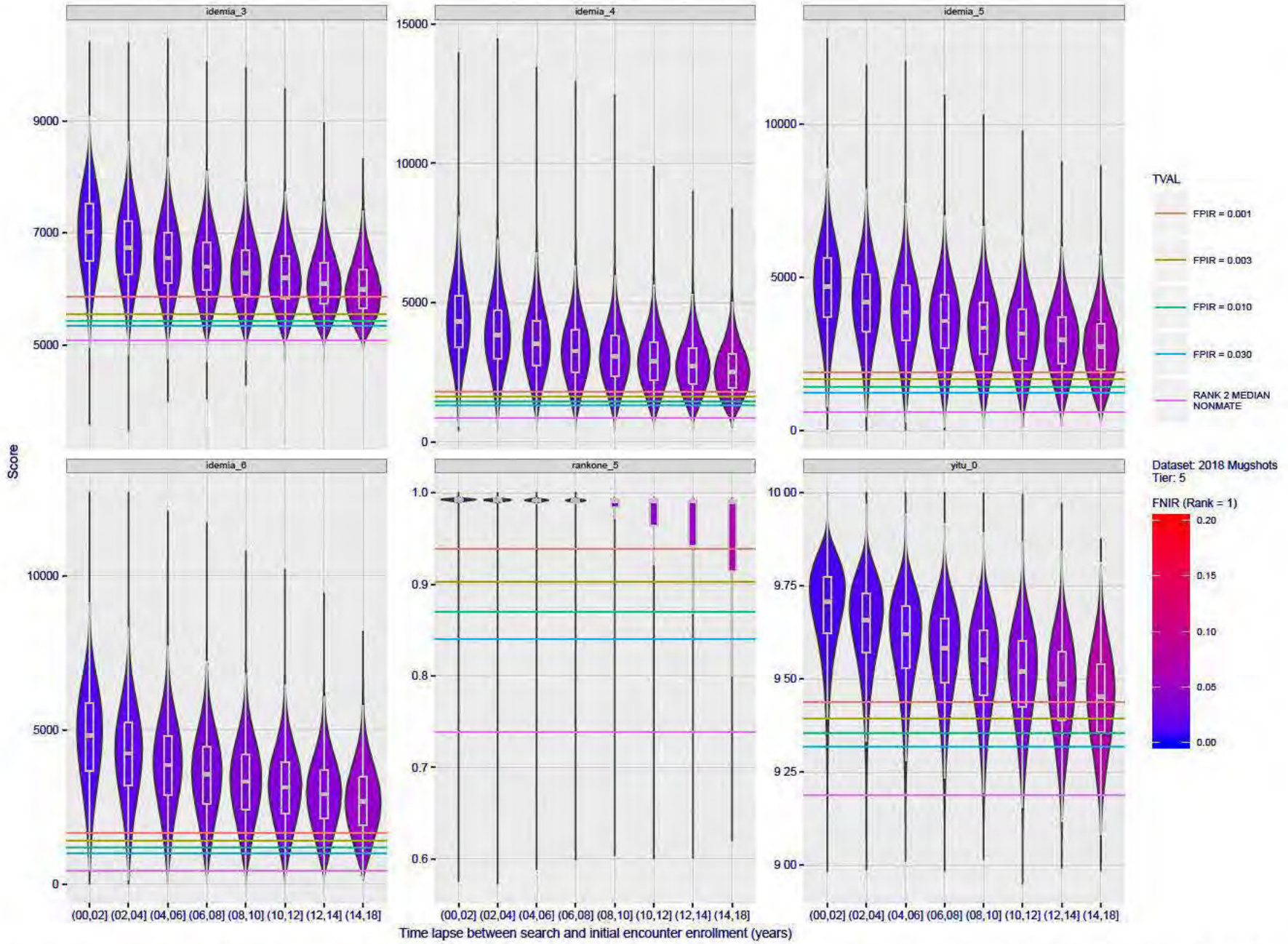


Figure 81: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

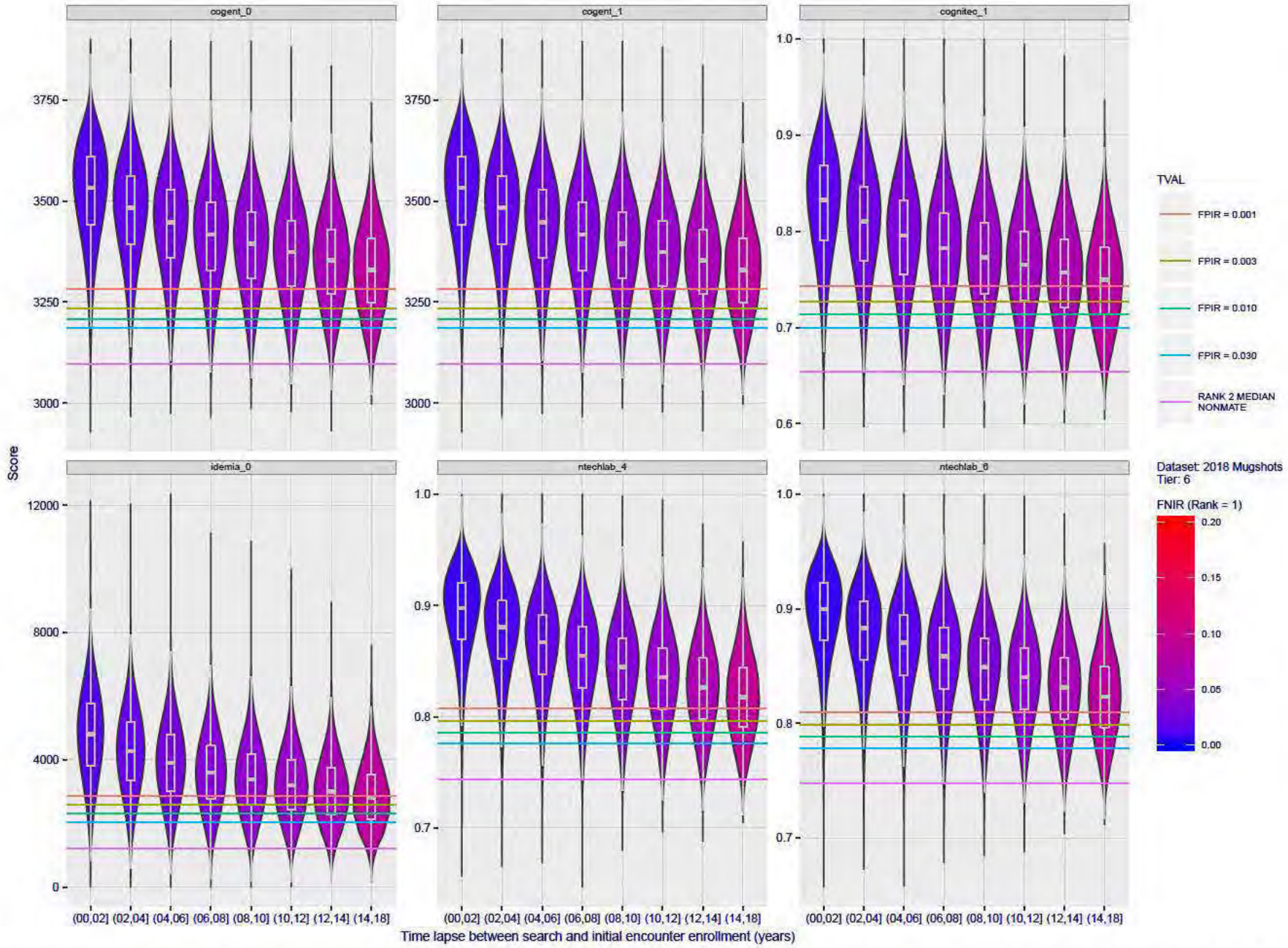


Figure 82: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

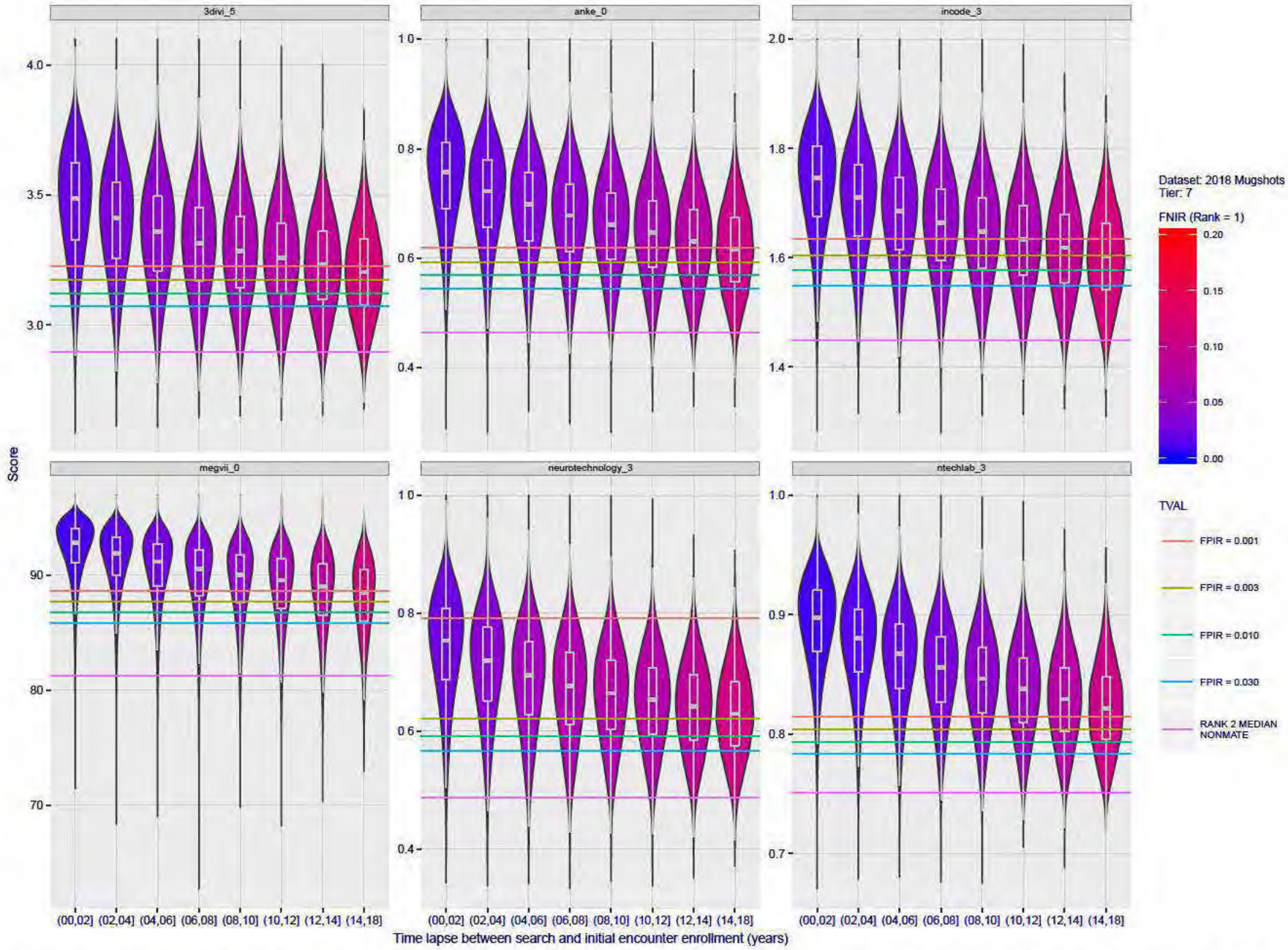


Figure 83: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

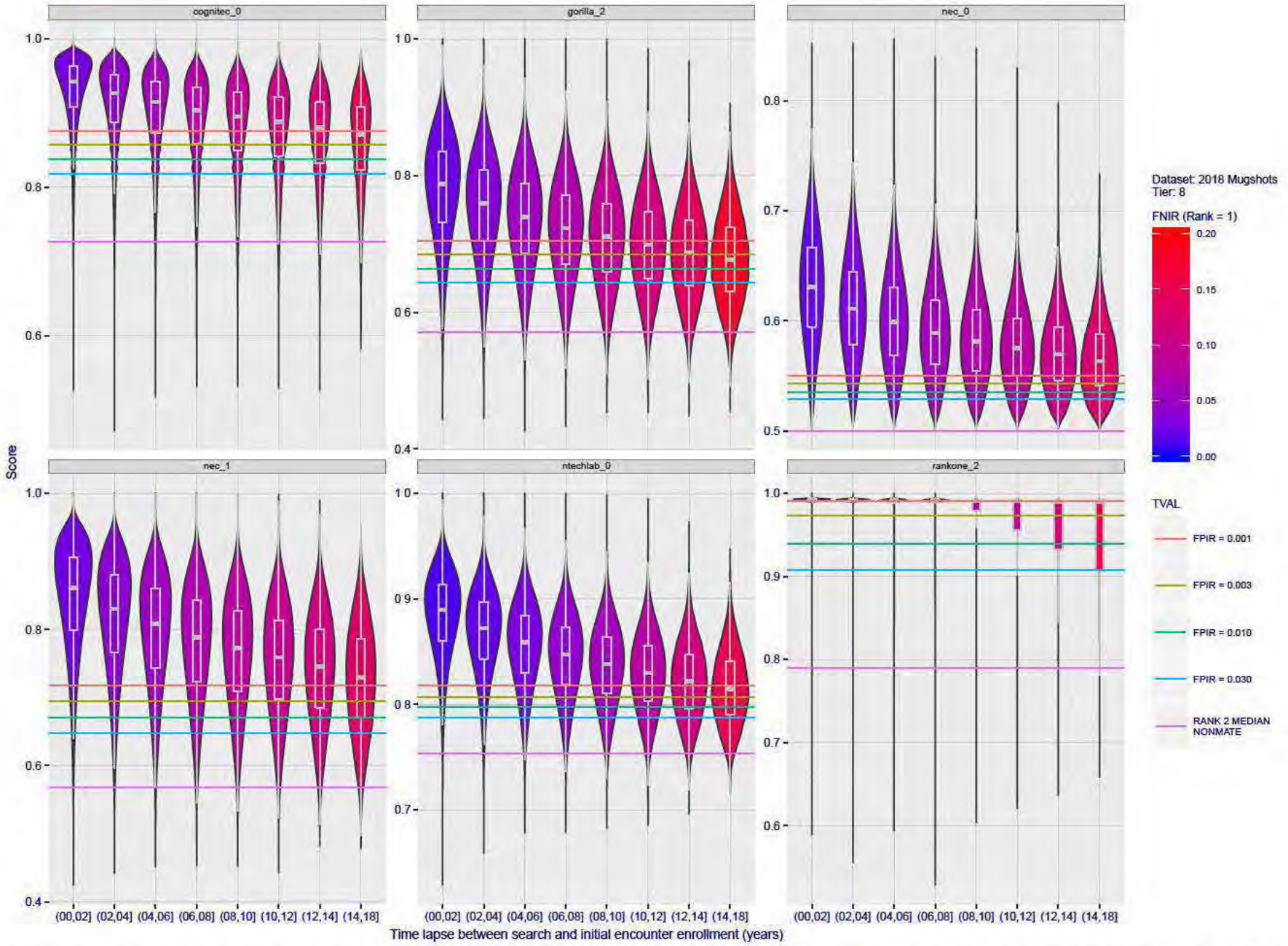


Figure 84: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

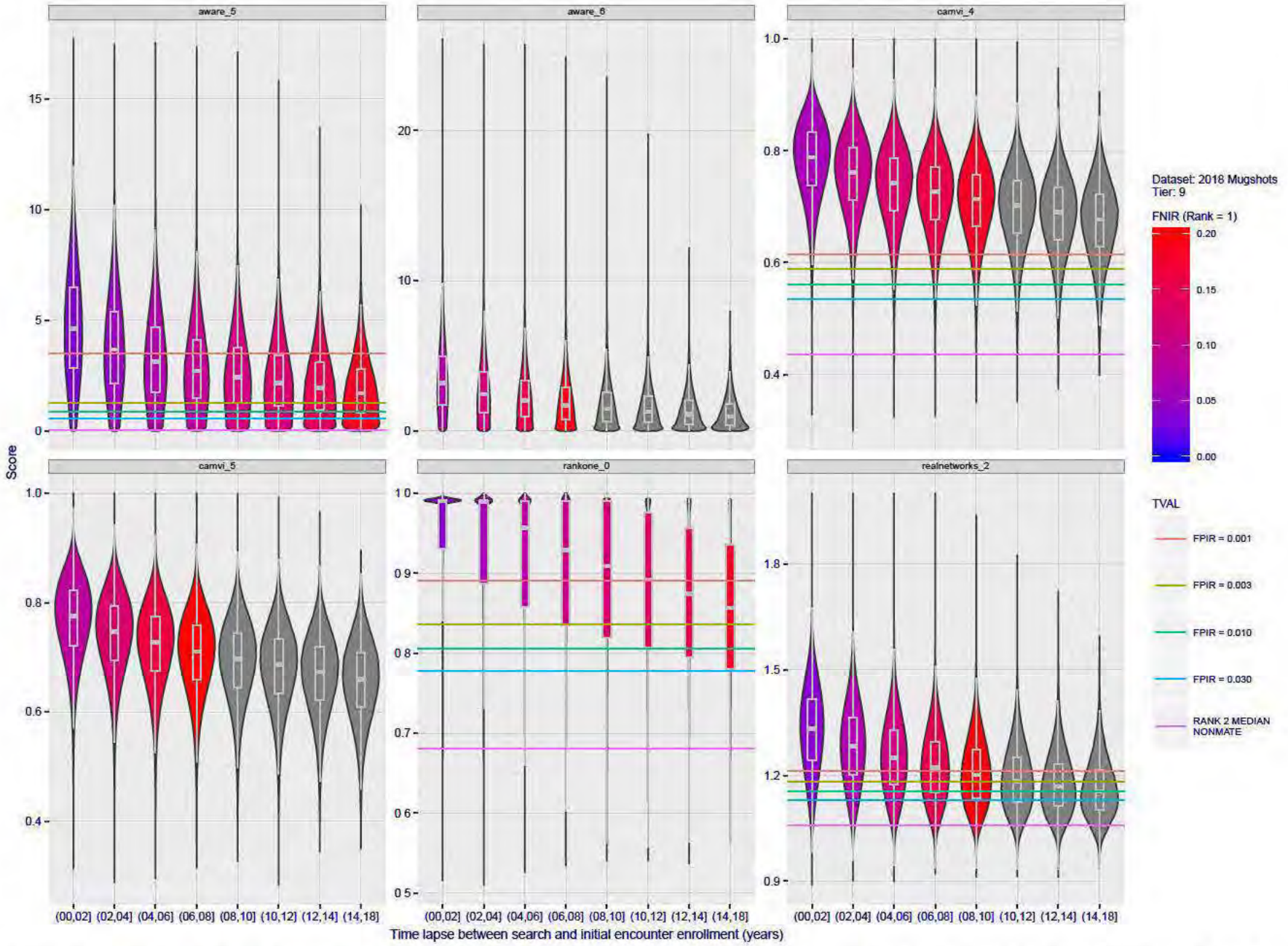


Figure 85: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

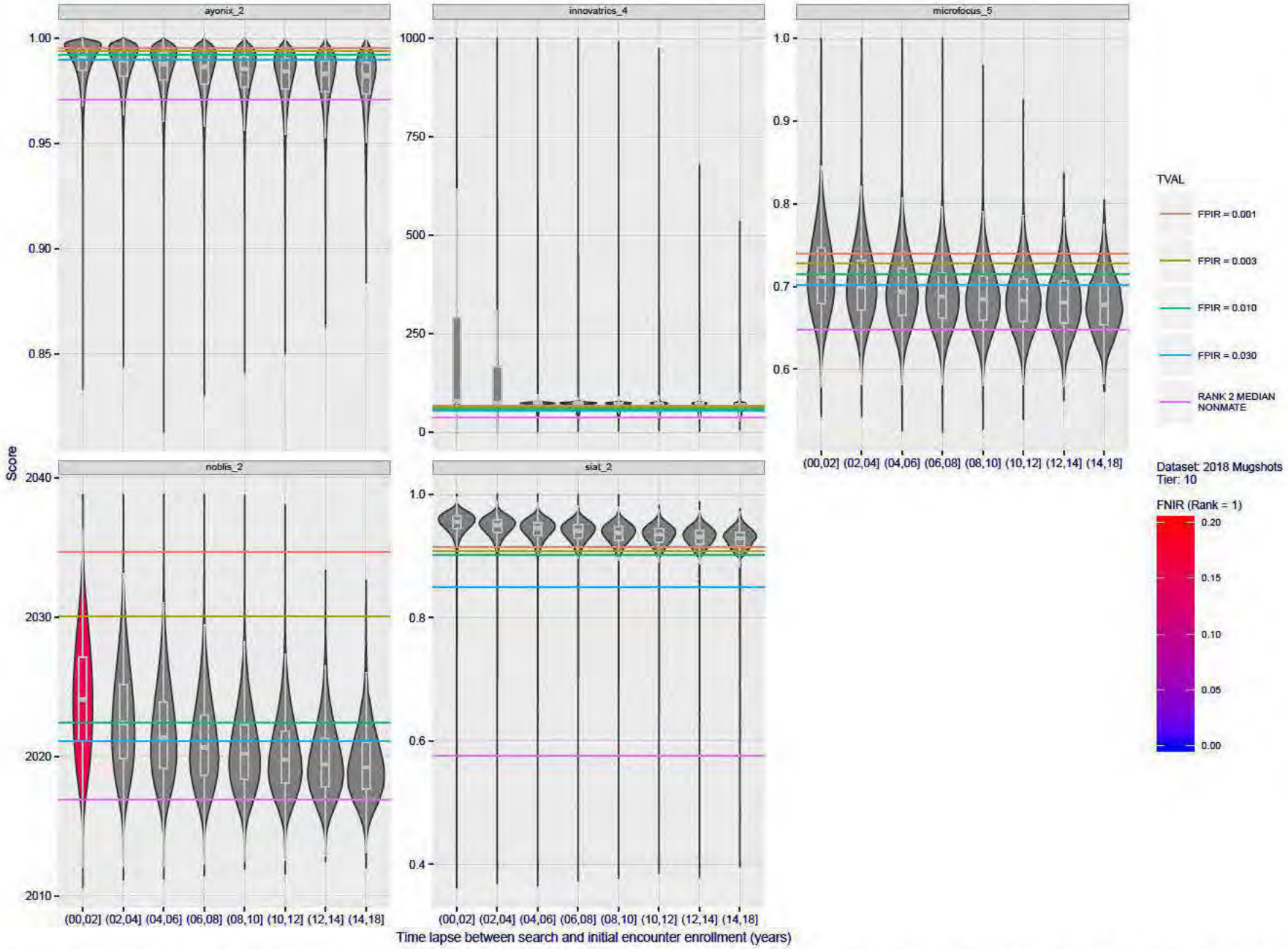


Figure 86: **[FRVT-2018 Mugshot Ageing Dataset] Native mate scores vs. time-elapsed.** *The oldest image of each individual is enrolled. Thereafter, all more recent images are searched. Mated score distributions are computed over all searches noted in row 17 of Table 5 binned by number of years between search and initial enrollment.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

Appendix C Effect of enrolling multiple images

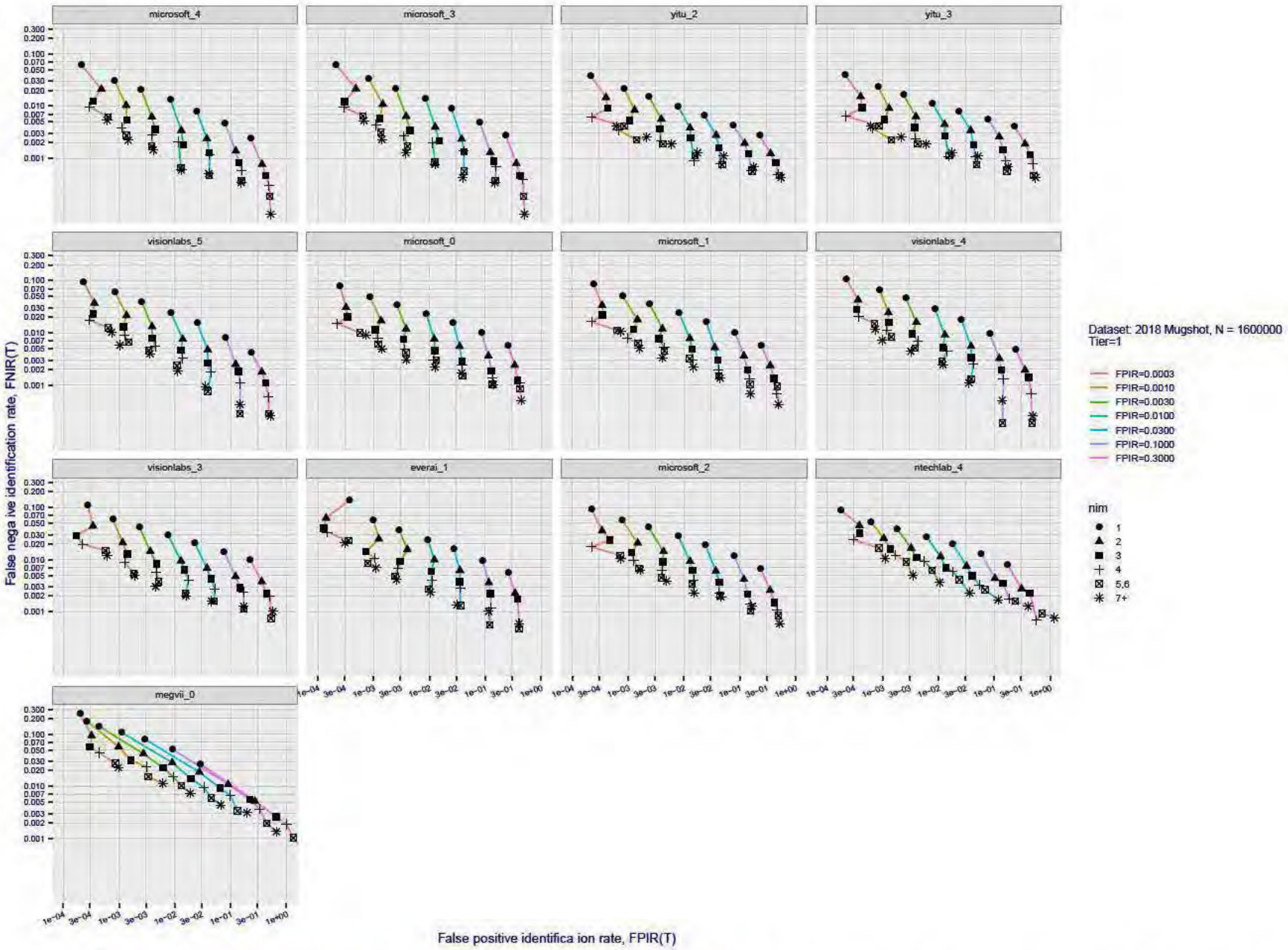


Figure 87: **[FRVT-2018 Mugshot Dataset] Effect of enrolling multiple images for each identity.** *The plot shows an identification miss rates vs. false positive rates, at seven operating thresholds. The enrolled population size is fixed. The images are enrolled with lifetime-consolidation - see section 2.3.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

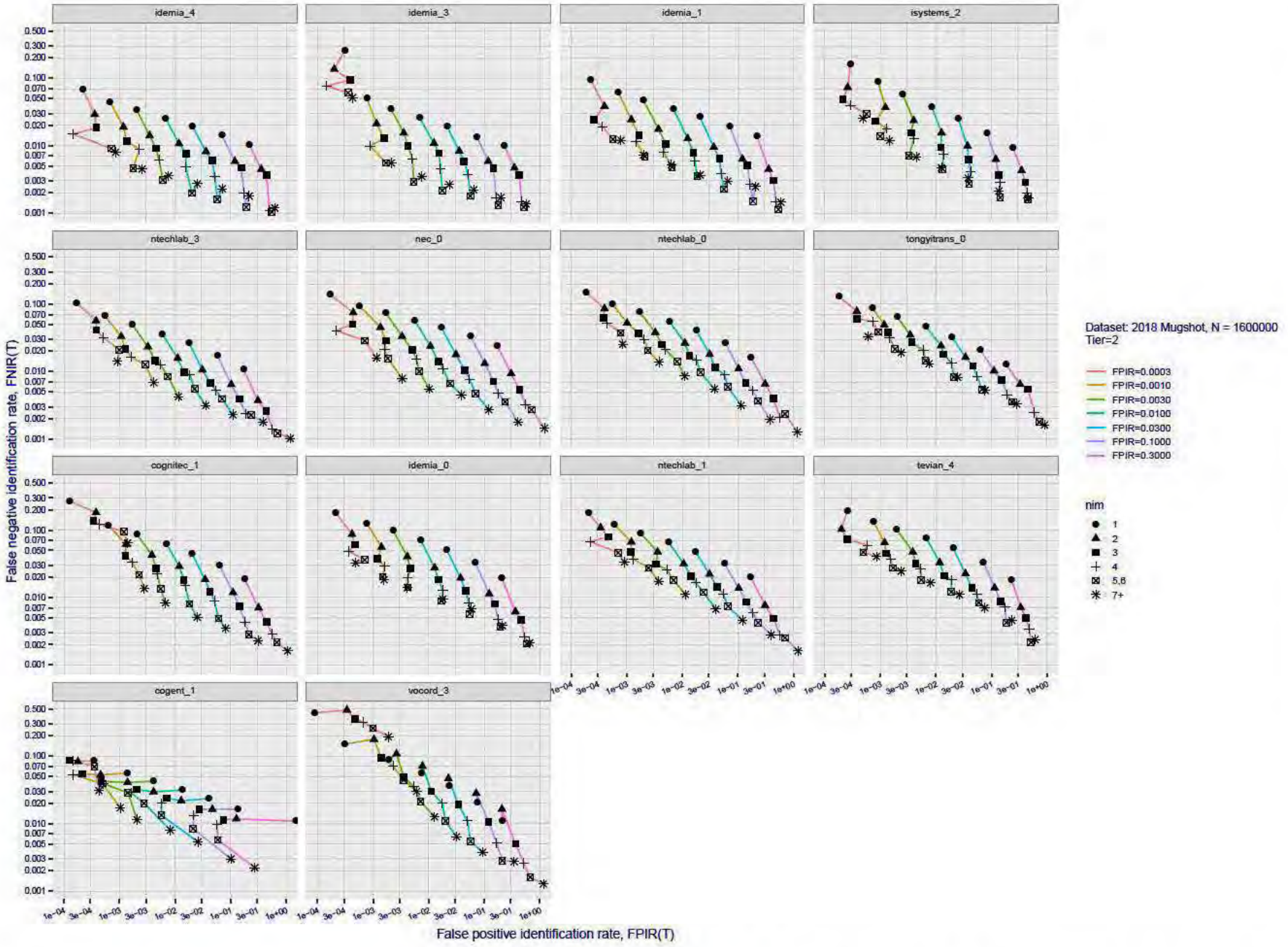


Figure 88: **[FRVT-2018 Mugshot Dataset] Effect of enrolling multiple images for each identity.** *The plot shows an identification miss rates vs. false positive rates, at seven operating thresholds. The enrolled population size is fixed. The images are enrolled with lifetime-consolidation - see section 2.3.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

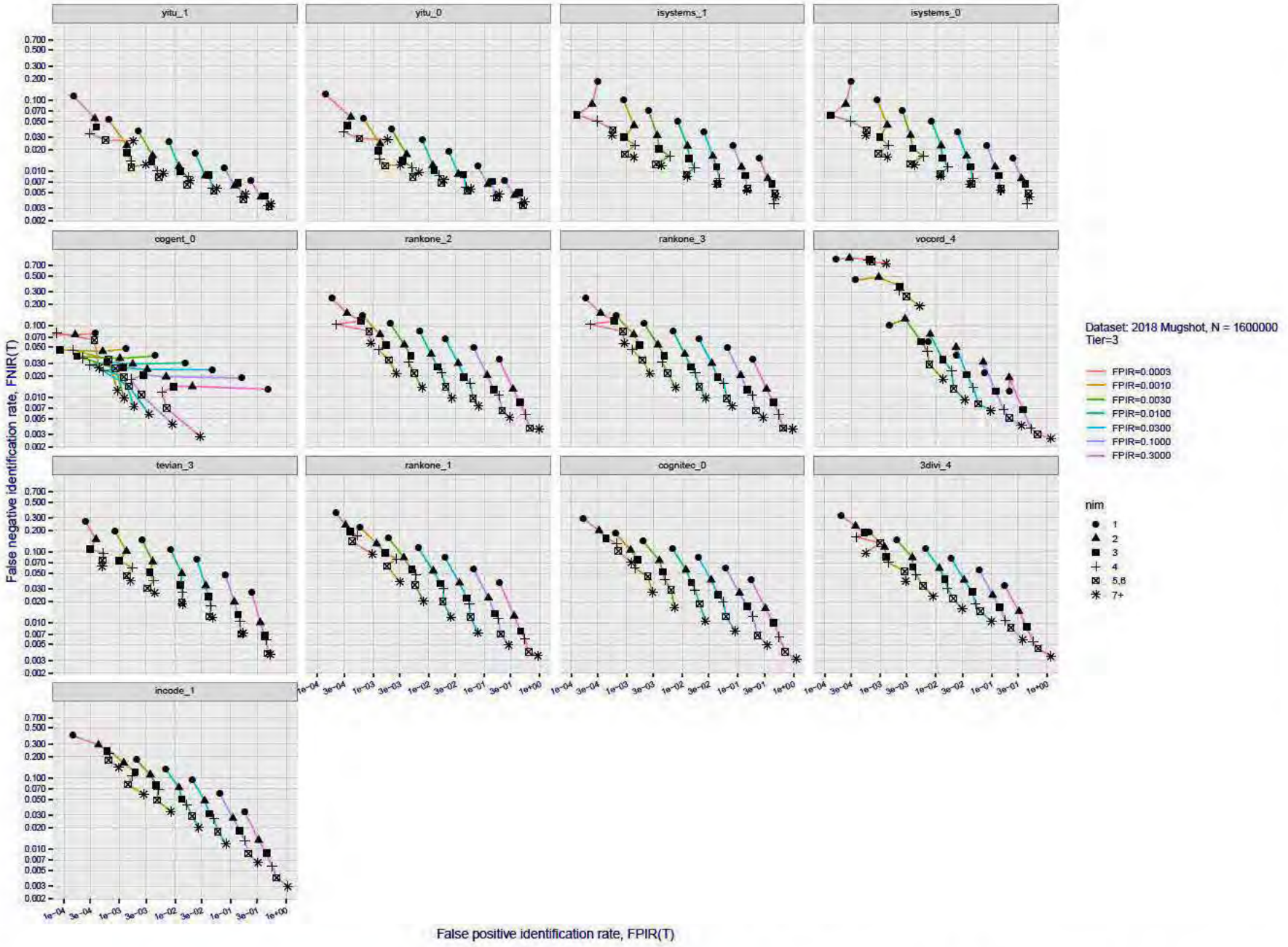


Figure 89: **[FRVT-2018 Mugshot Dataset] Effect of enrolling multiple images for each identity.** *The plot shows an identification miss rates vs. false positive rates, at seven operating thresholds. The enrolled population size is fixed. The images are enrolled with lifetime-consolidation - see section 2.3.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

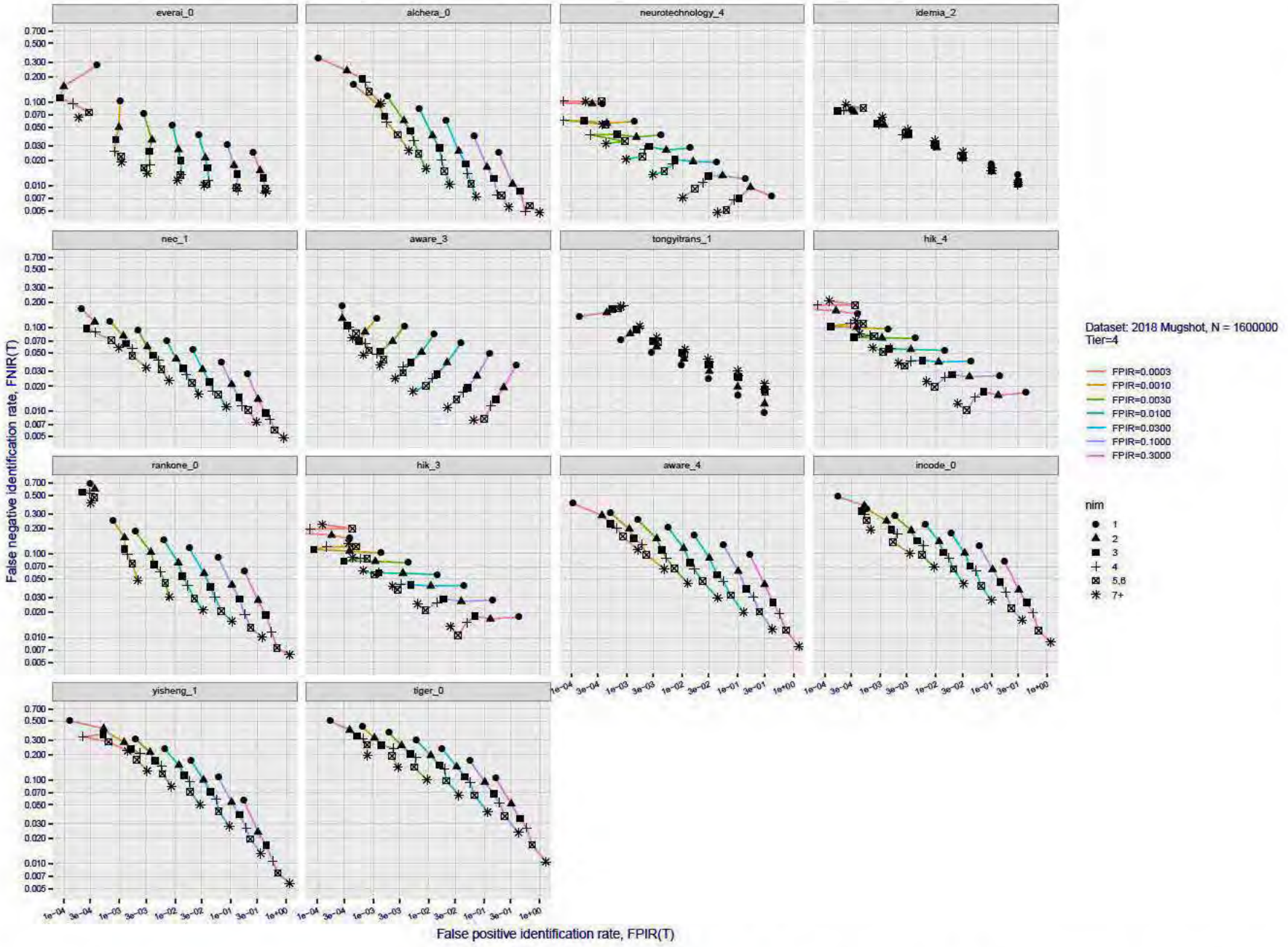


Figure 90: **[FRVT-2018 Mugshot Dataset] Effect of enrolling multiple images for each identity.** *The plot shows an identification miss rates vs. false positive rates, at seven operating thresholds. The enrolled population size is fixed. The images are enrolled with lifetime-consolidation - see section 2.3.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

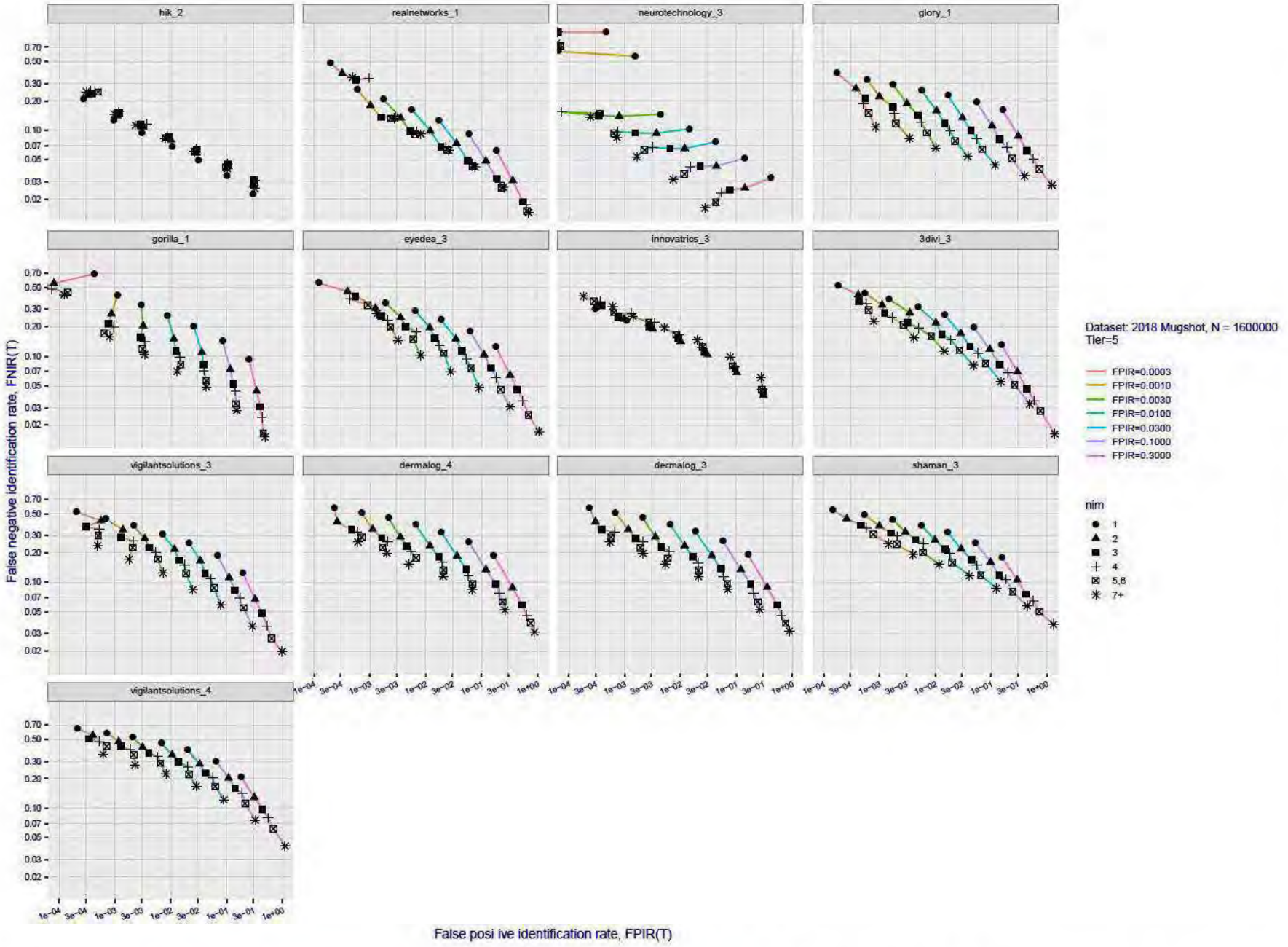


Figure 91: **[FRVT-2018 Mugshot Dataset] Effect of enrolling multiple images for each identity.** *The plot shows an identification miss rates vs. false positive rates, at seven operating thresholds. The enrolled population size is fixed. The images are enrolled with lifetime-consolidation - see section 2.3.*

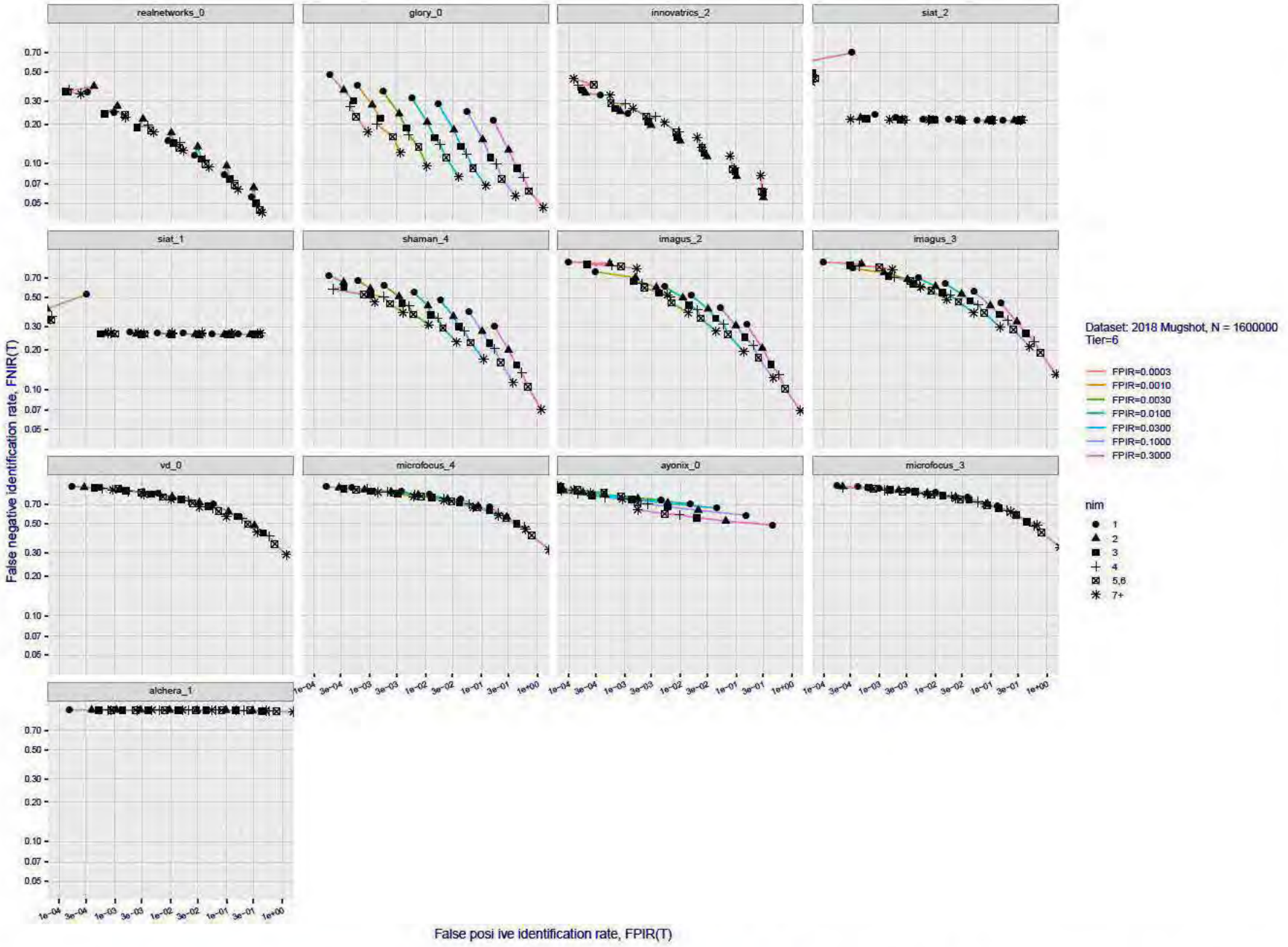


Figure 92: **[FRVT-2018 Mugshot Dataset] Effect of enrolling multiple images for each identity.** *The plot shows an identification miss rates vs. false positive rates, at seven operating thresholds. The enrolled population size is fixed. The images are enrolled with lifetime-consolidation - see section 2.3.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

Appendix D Accuracy with poor quality webcam images

2019/09/11 17:24:52	FNIR(N, R, T) = FPIR(N, T) =	False neg. identification rate False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification

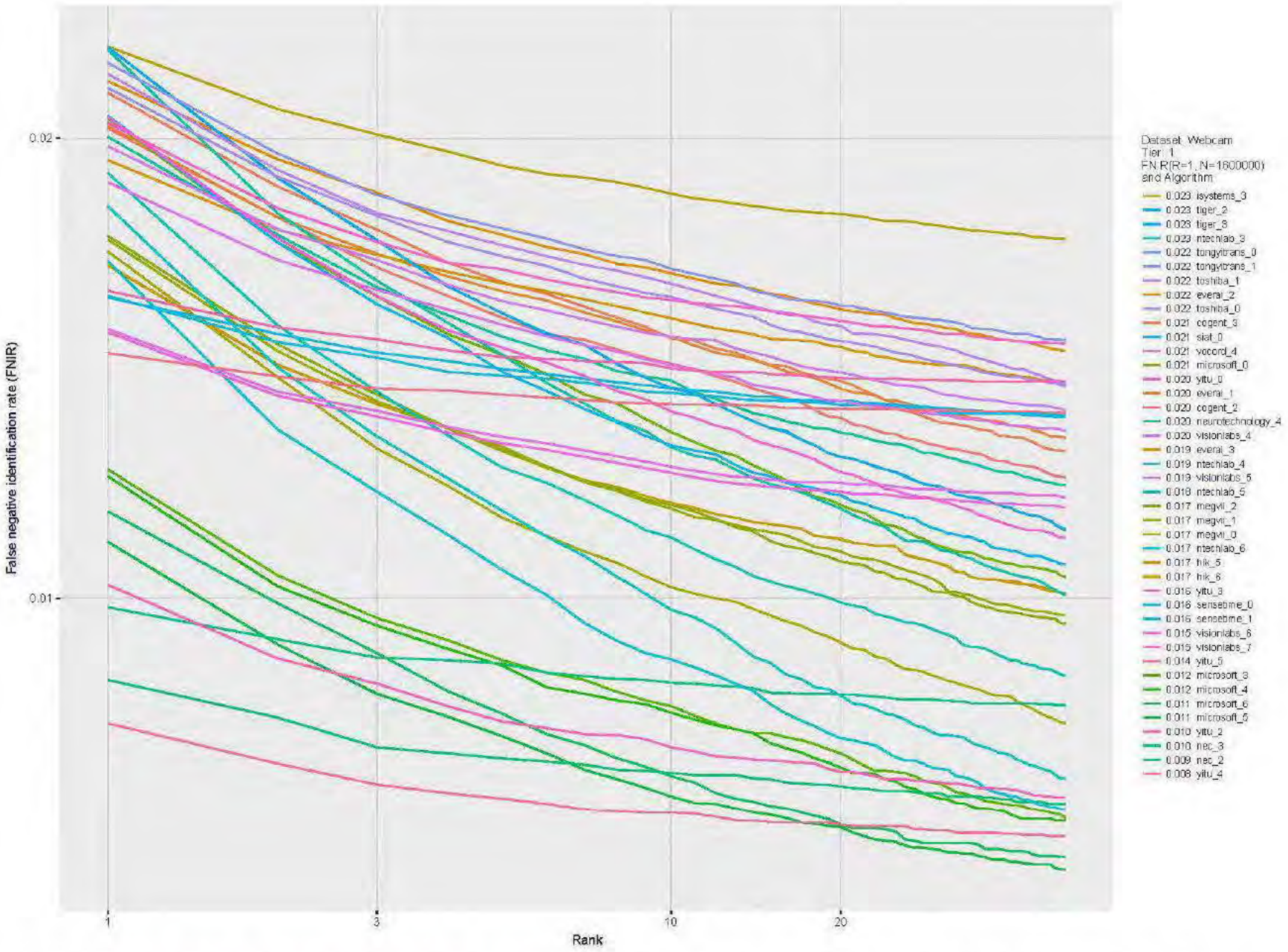


Figure 93: **[Webcam Dataset] Identification miss rates vs. rank.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

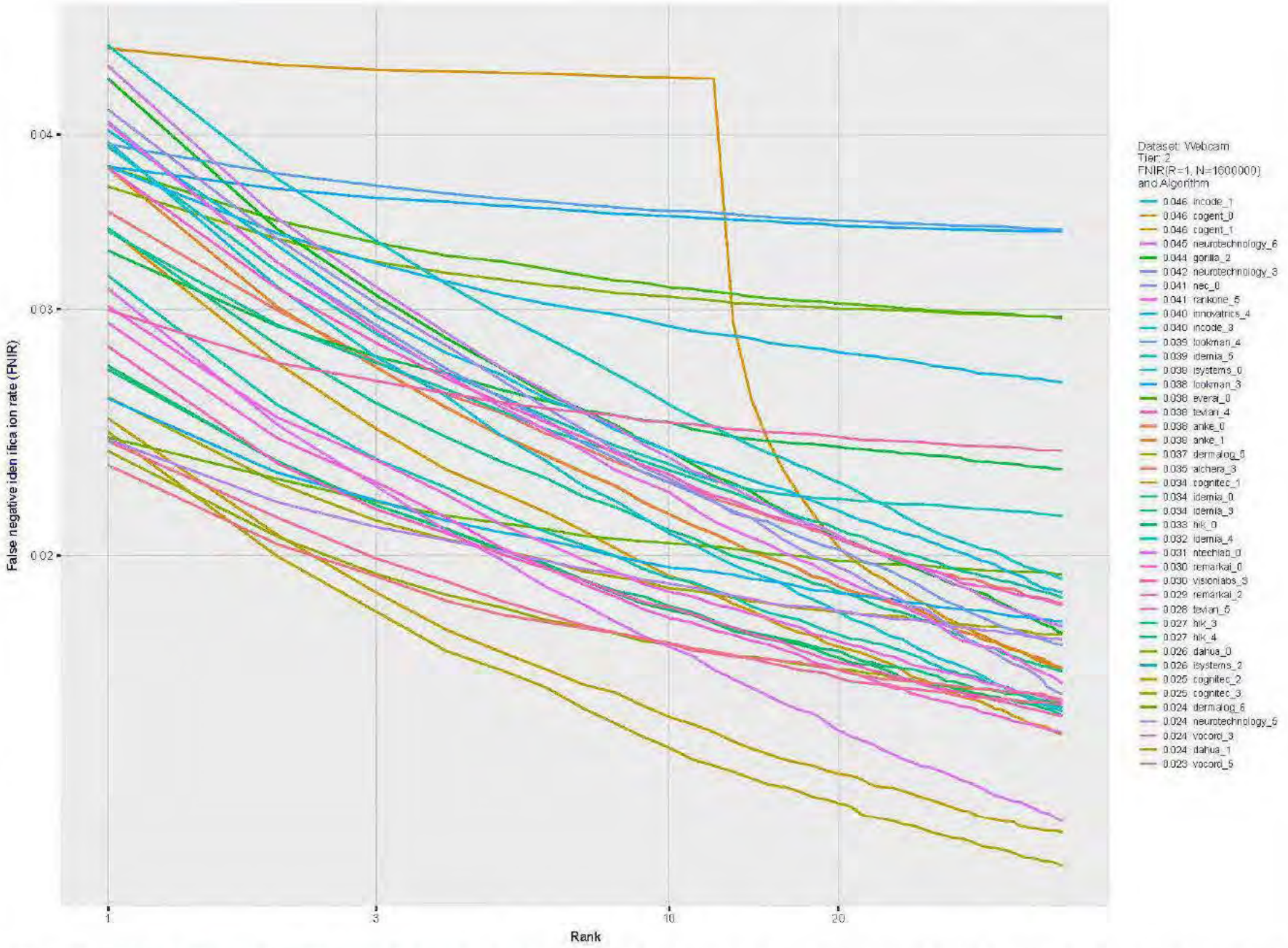


Figure 94: **[Webcam Dataset] Identification miss rates vs. rank.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11
17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

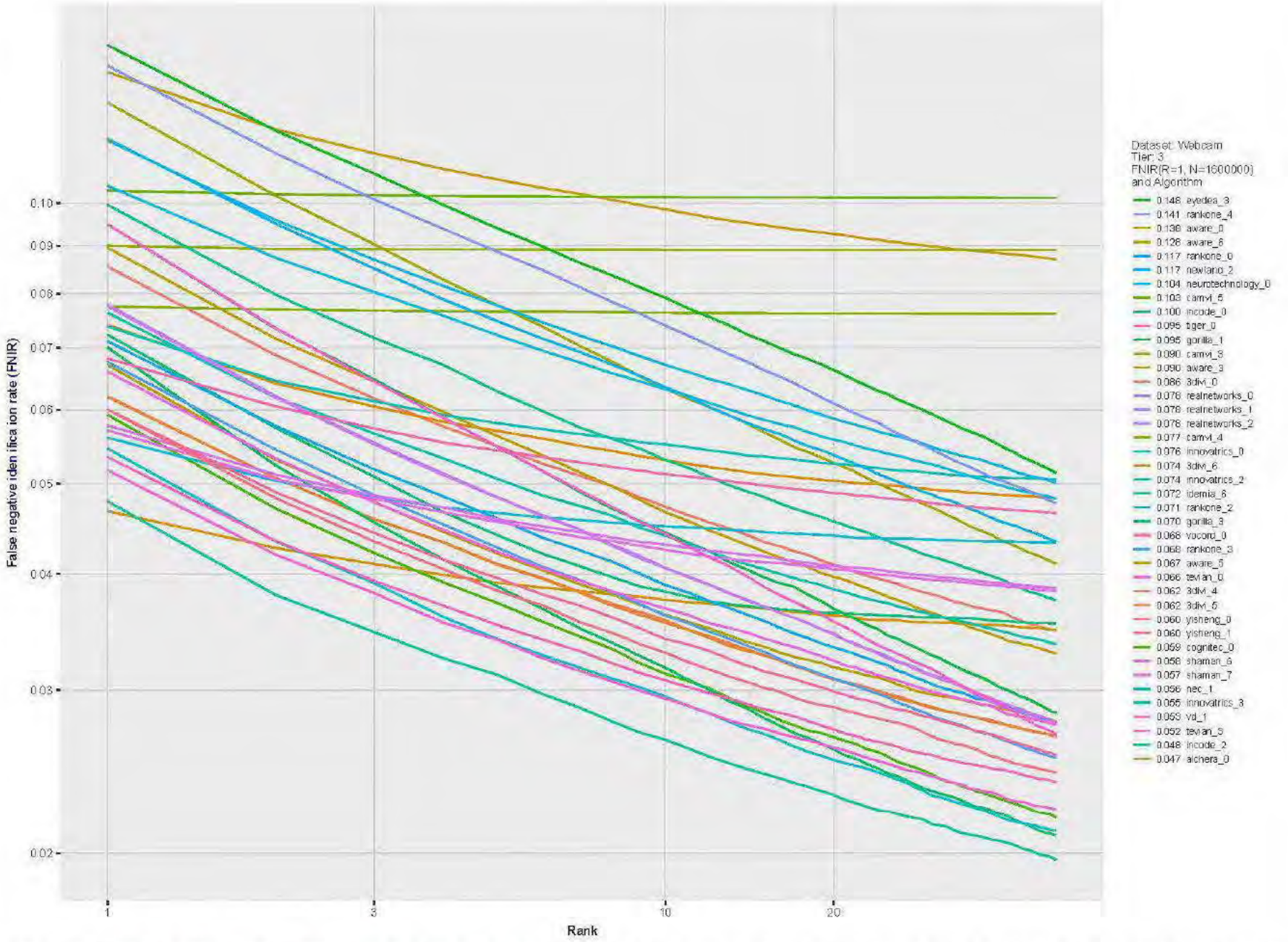


Figure 95: **[Webcam Dataset] Identification miss rates vs. rank.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11
17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

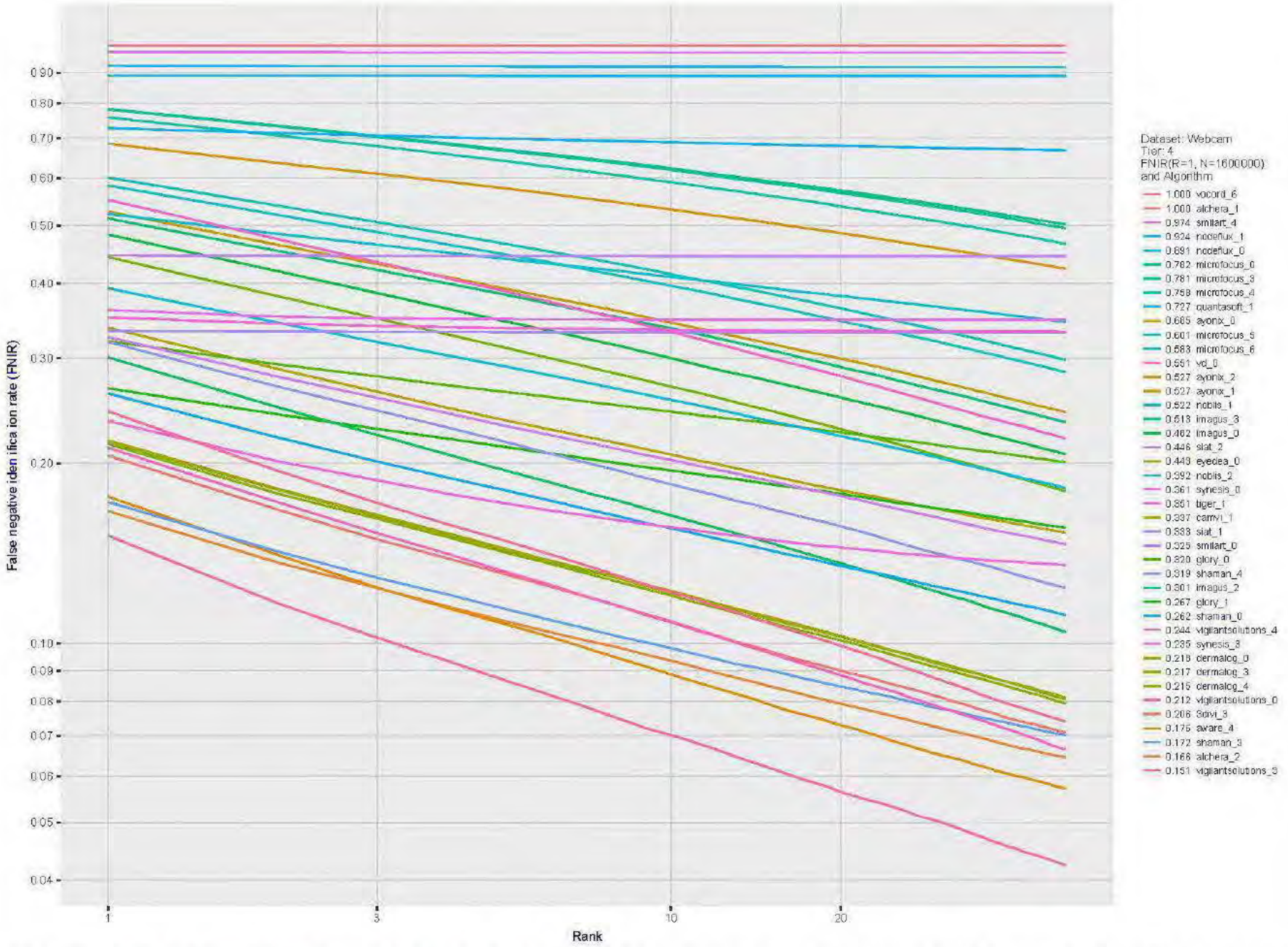


Figure 96: **[Webcam Dataset] Identification miss rates vs. rank.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11 17:24:52
FNIR(N, R, T) = FPIR(N, T) =
False neg. identification rate
False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

2019/09/11 17:24:52	FNIR(N, R, T) = FPIR(N, T) =	False neg. identification rate False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification

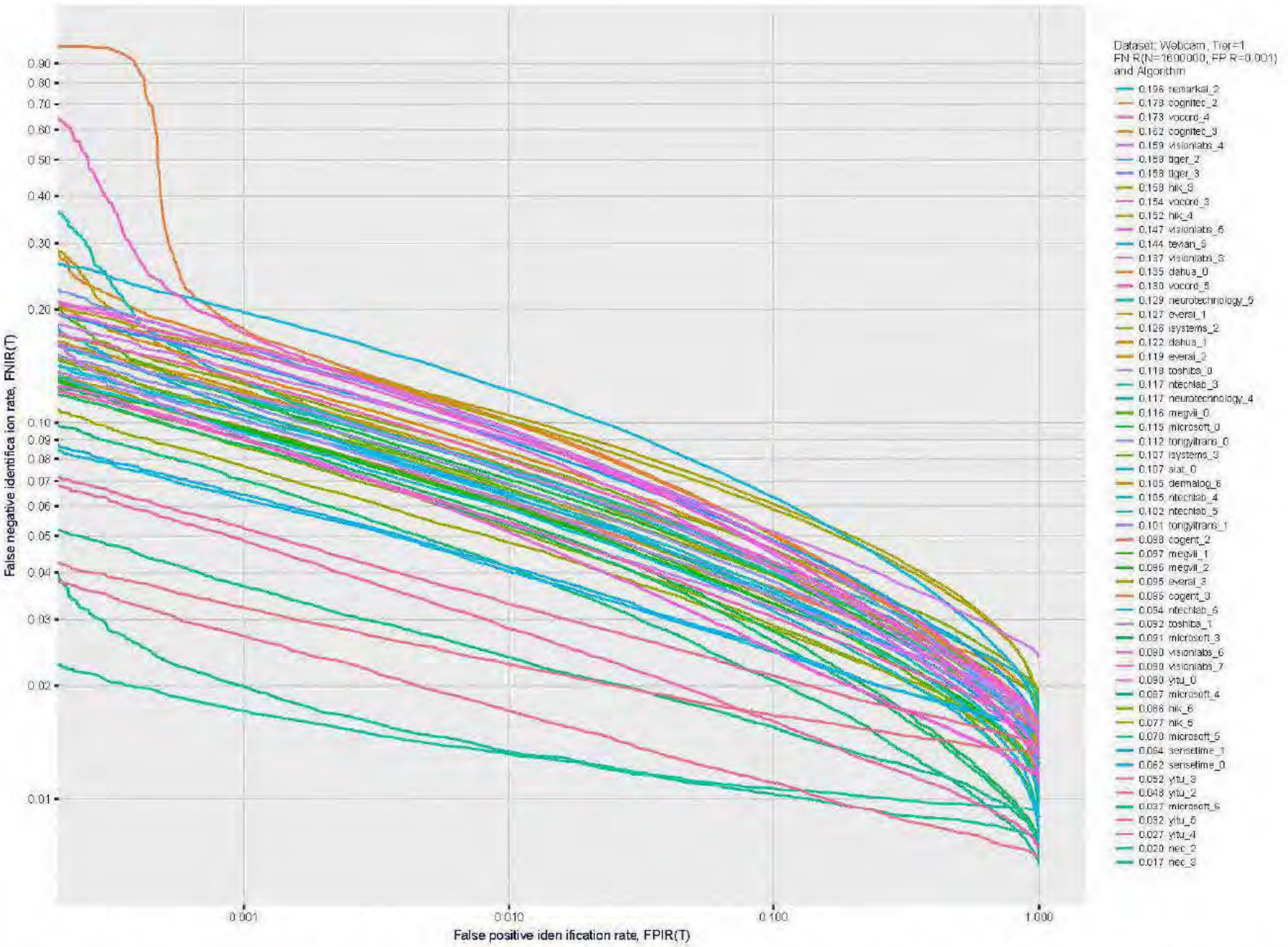


Figure 97: **[Webcam Dataset] Identification miss rates vs. false positive rates.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11 17:24:52
FNIR(N, R, T) = FPIR(N, T) =
False neg. identification rate
False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

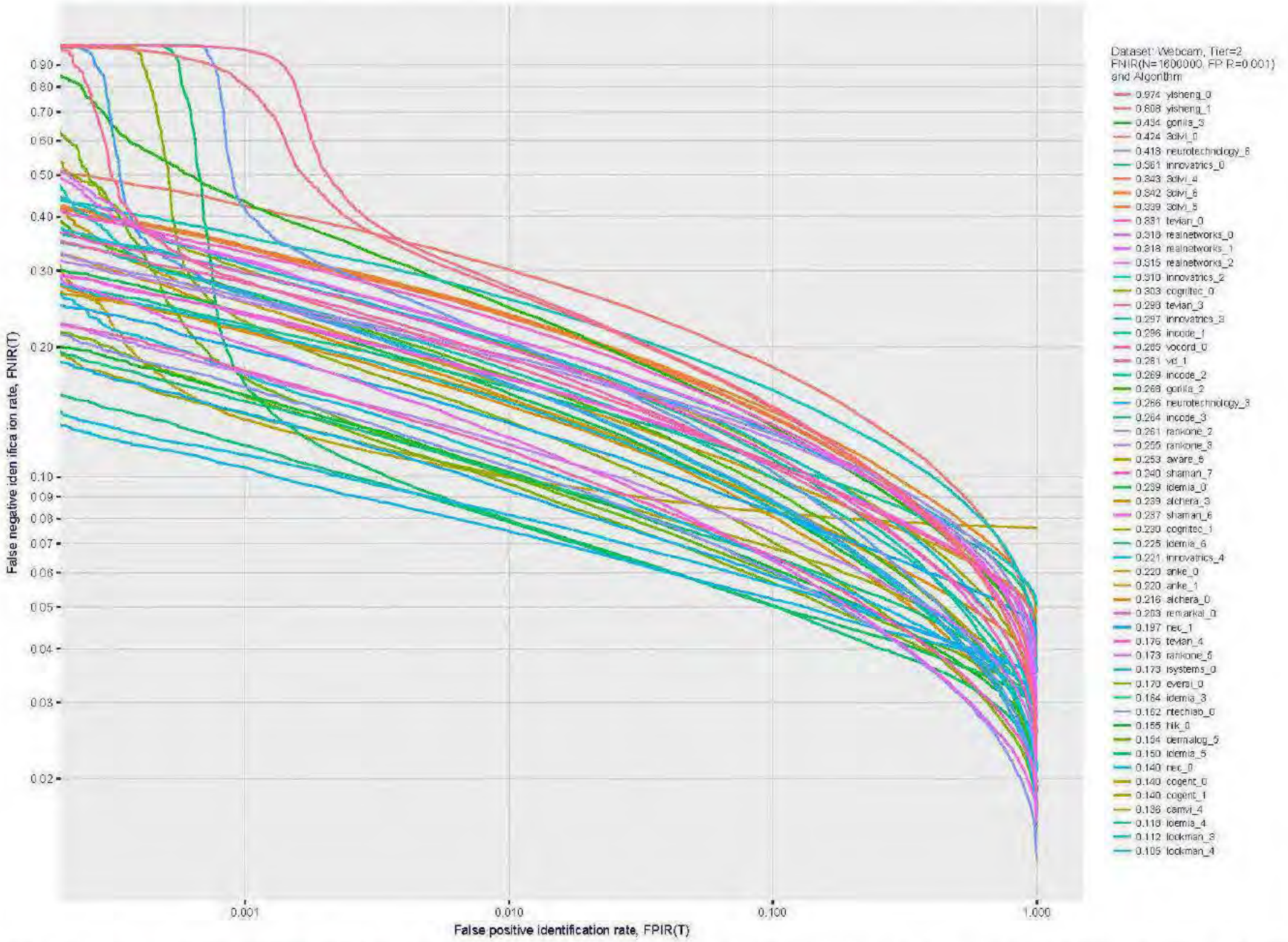


Figure 98: **[Webcam Dataset] Identification miss rates vs. false positive rates.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11 17:24:52
FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

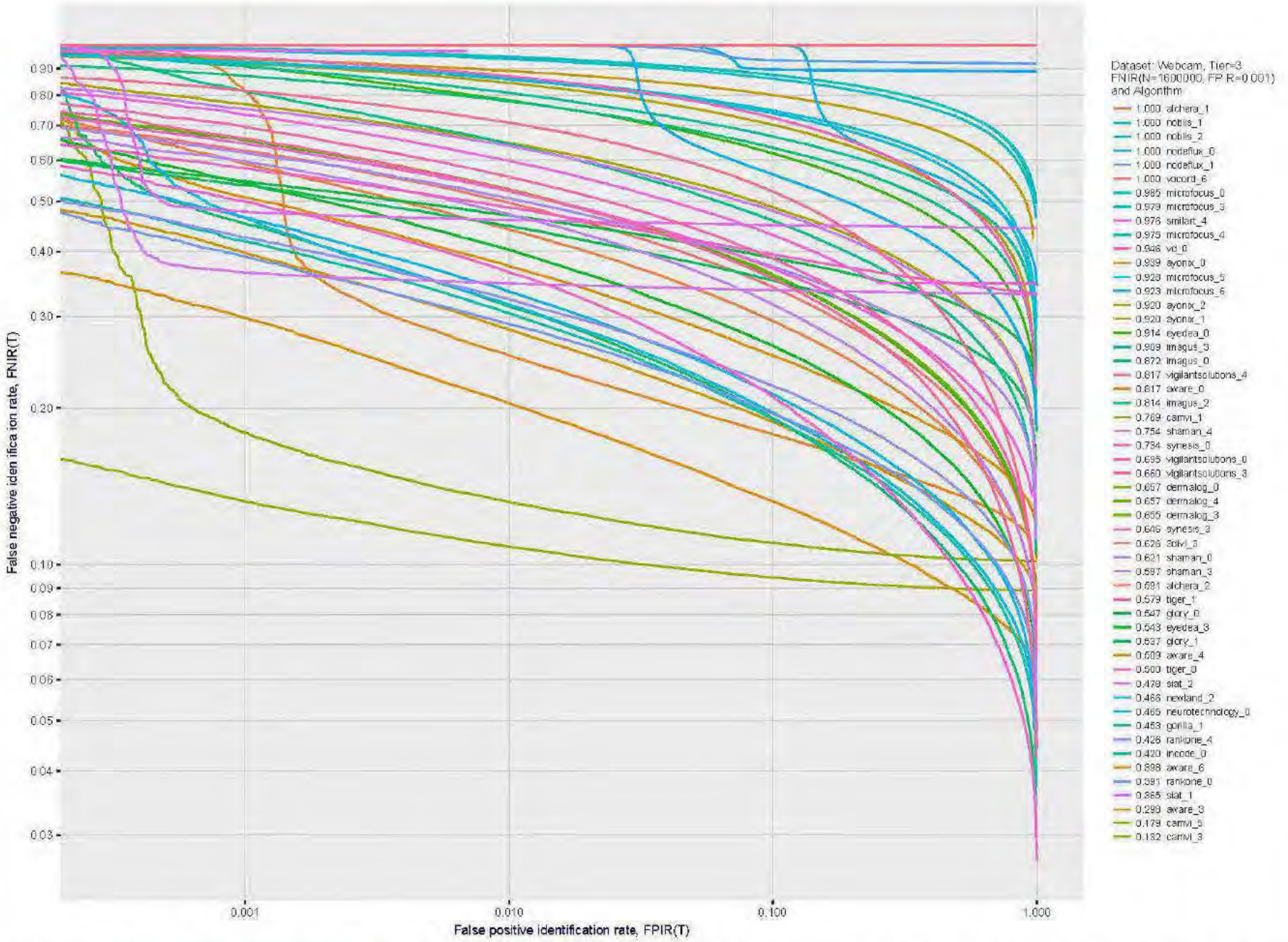


Figure 99: **[Webcam Dataset] Identification miss rates vs. false positive rates.** *The results apply to cross-domain recognition in which webcams are searched against enrolled mugshots. The FNIR values are higher than those for mugshot-mugshot identification due to low image resolution, lighting and less constrained subject pose in webcam images - see Figure 4.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

Appendix E Accuracy for profile-view to frontal recognition

Figures 100 - 102 gives accuracy results for searching 100 000 mated and 100 000 non-mated profile-view images against the same FRVT 2018 frontal enrollment dataset, N = 1 600 000, used in the main mugshot trials. This experiment corresponds to row-13 of Table 5. An example of profile-view image is given in Figure 5.

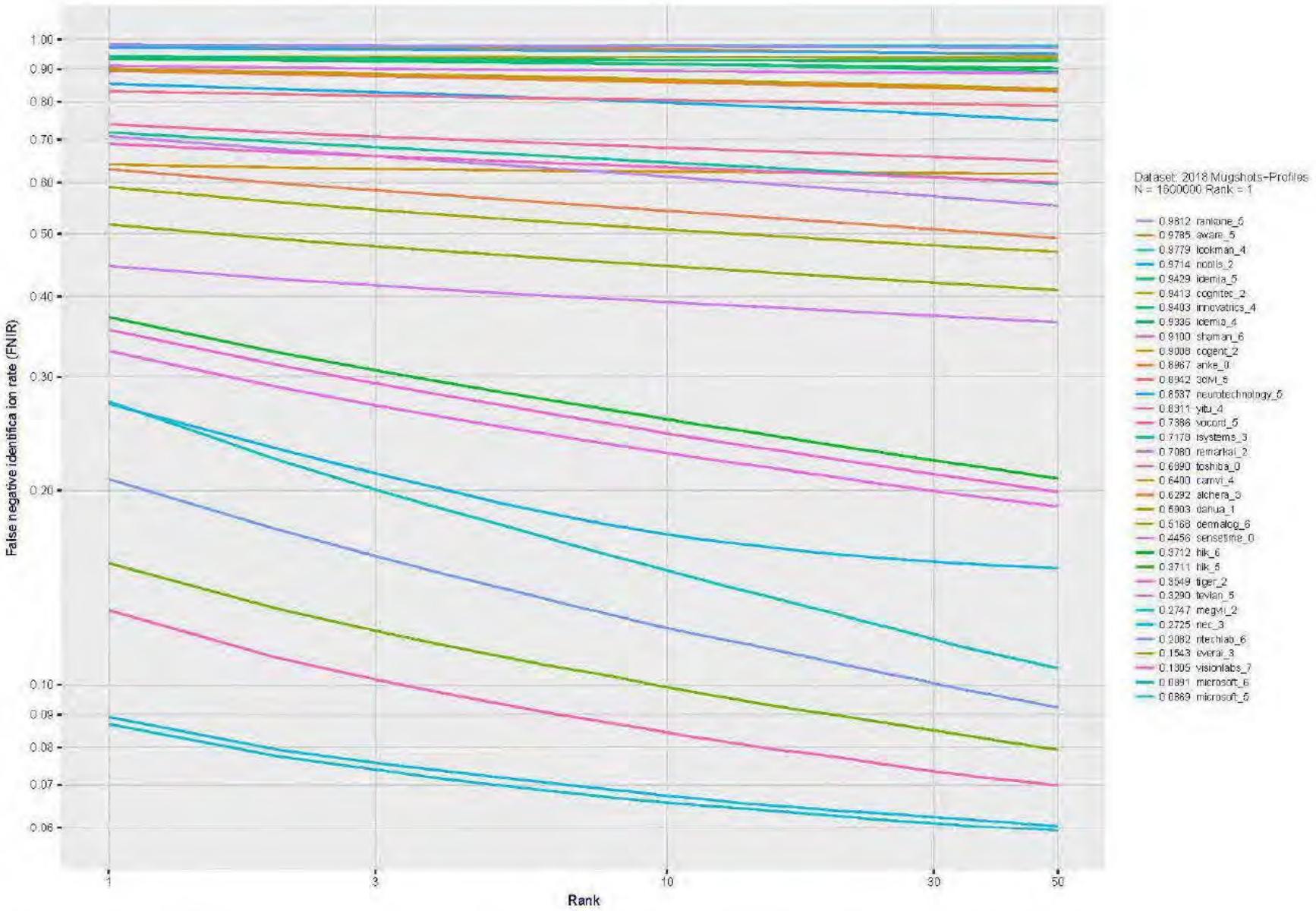


Figure 100: **[Mugshot and profile-view dataset] Rank-based accuracy.** *For some of the more accurate Phase 3 algorithms the figure plots error tradeoff characteristics for frontal and profile-view searches into an enrolled set of N = 1 600 000 frontal images. Note that some algorithms fail on profile-view images with FNIR → 1 - this evaluation did not ask developers to provide profile-view capability. Some algorithms, on the other hand, give FNIR approaching that for frontal-view searches using c. 2010 algorithms. The best result is that 91% of profile-view searches yield the correct mate at rank 1, and better than 94% in the top-50 candidates.*

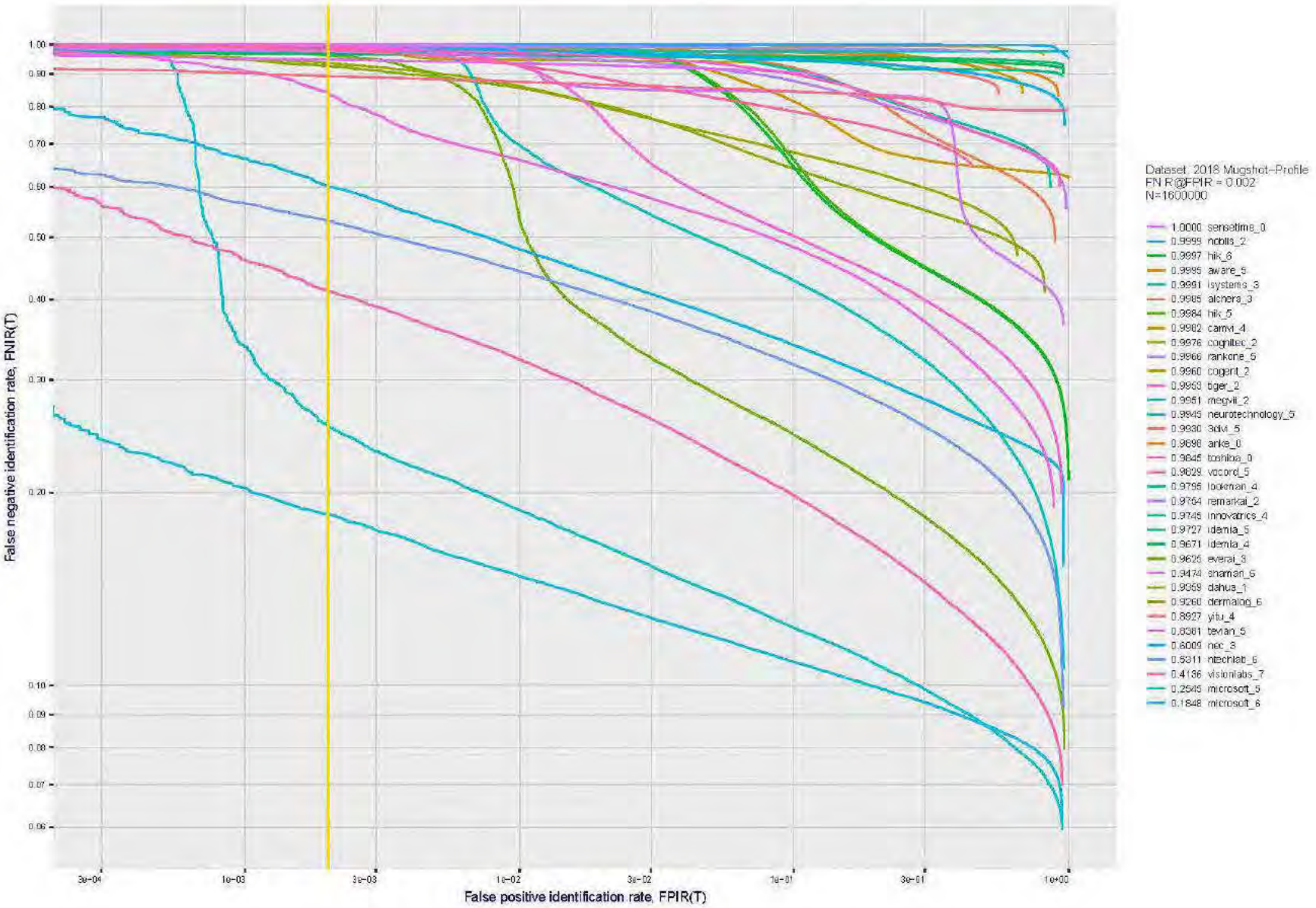


Figure 101: **[Mugshot and profile-view dataset] Threshold-based accuracy.** *For some of the more accurate Phase 3 algorithms the figure plots error tradeoff characteristics for frontal and profile-view searches into an enrolled set of N = 1 600 000 frontal images. Note that some algorithms fail on profile-view images with FNIR → 1 - this evaluation did not ask developers to provide profile-view capability. Some algorithms, on the other hand, give FNIR approaching that for frontal-view searches using c. 2010 algorithms.*

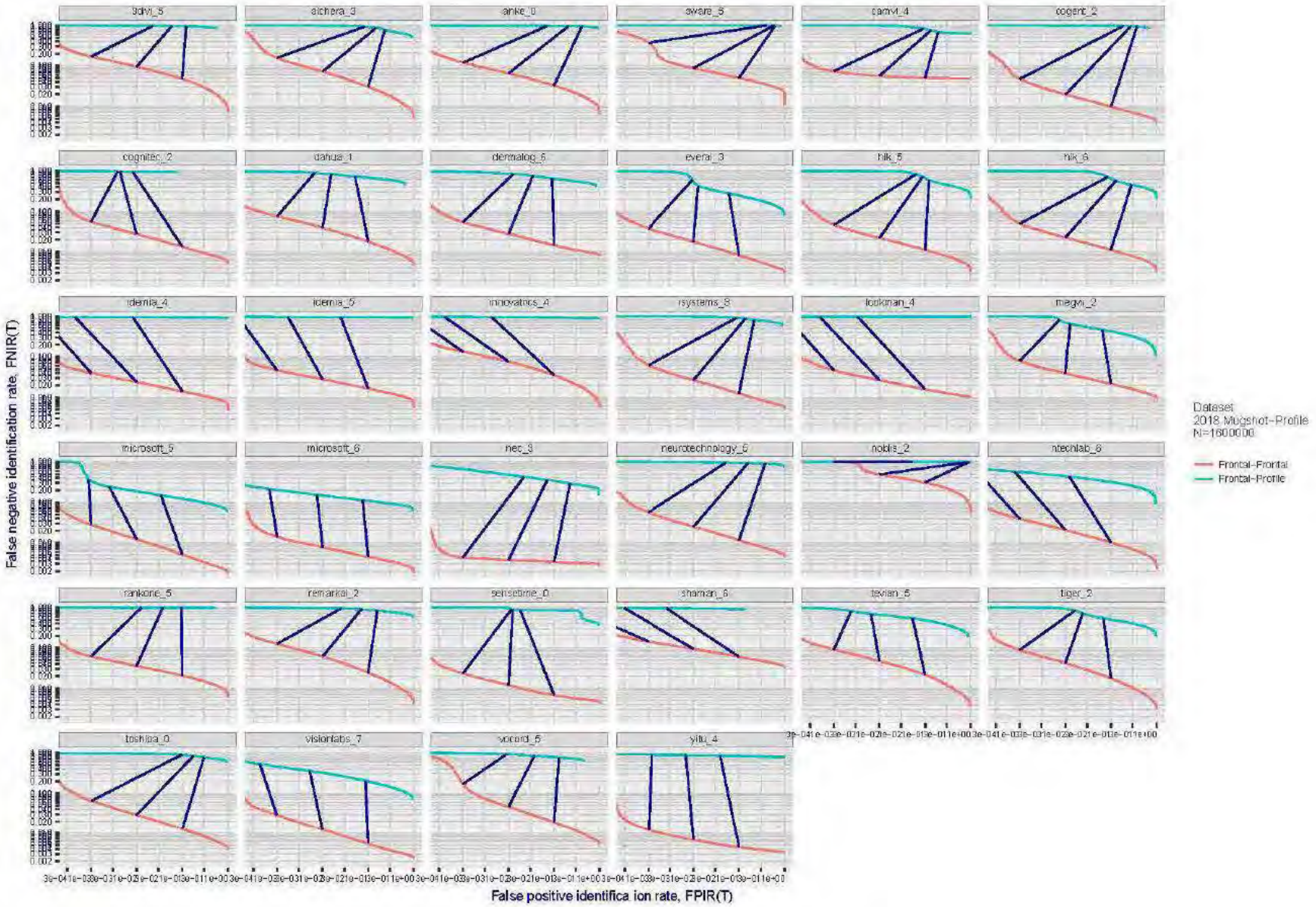


Figure 102: **[Mugshot and profile-view dataset] Speed-accuracy tradeoff.** *For some of the more accurate Phase 3 algorithms the figure plots error tradeoff characteristics for frontal and profile-view searches into an enrolled set of N = 1 600 000 frontal images. Some algorithms fail on profile-view images with FNIR → 1 - this evaluation did not ask developers to provide profile-view capability. Some algorithms, on the other hand, give FNIR approaching that for frontal-view searches using c. 2010 algorithms. Blue lines connect points of equal threshold from which it is evident that some algorithms would give markedly higher false positive outcomes if profile-view images were searched in a system configured for frontal searches. This would be a vulnerability in an access control system.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

Appendix F Accuracy when identifying wild images

2019/09/11 17:24:52	FNIR(N, R, T) = False neg. identification rate FPIR(N, T) = False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification	

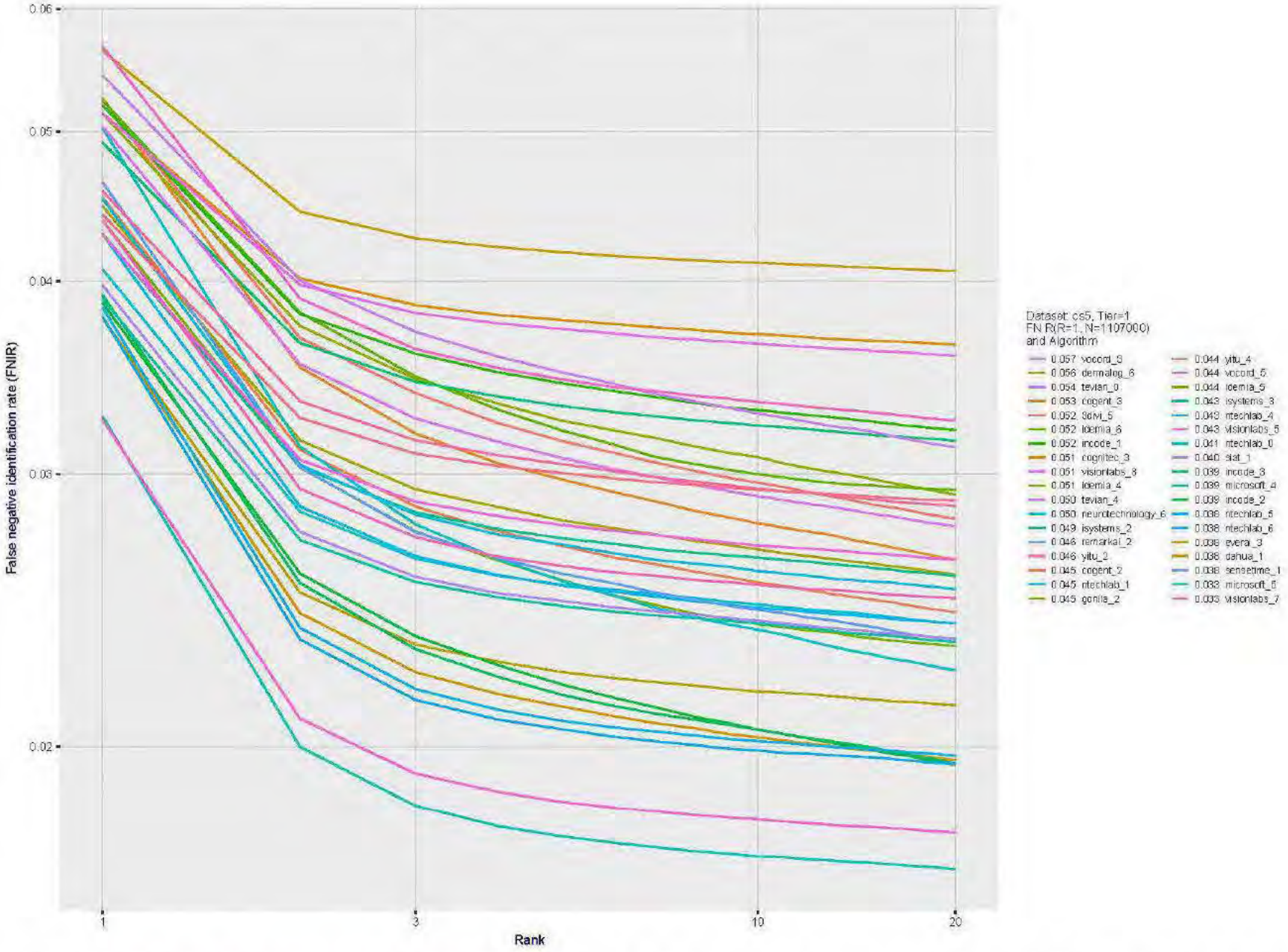


Figure 103: **[Wild Dataset] Identification miss rates vs. rank.** *For the wild dataset, the figure shows false negative identification rates (FNIR) vs. rank when the threshold is set to zero. This metric is relevant to human reviewers who will traverse candidate lists checking whether any of the returned identities match to the search imagery. Specifically, wild images were searched against 1.1 million individuals enrolled with wild images as well.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

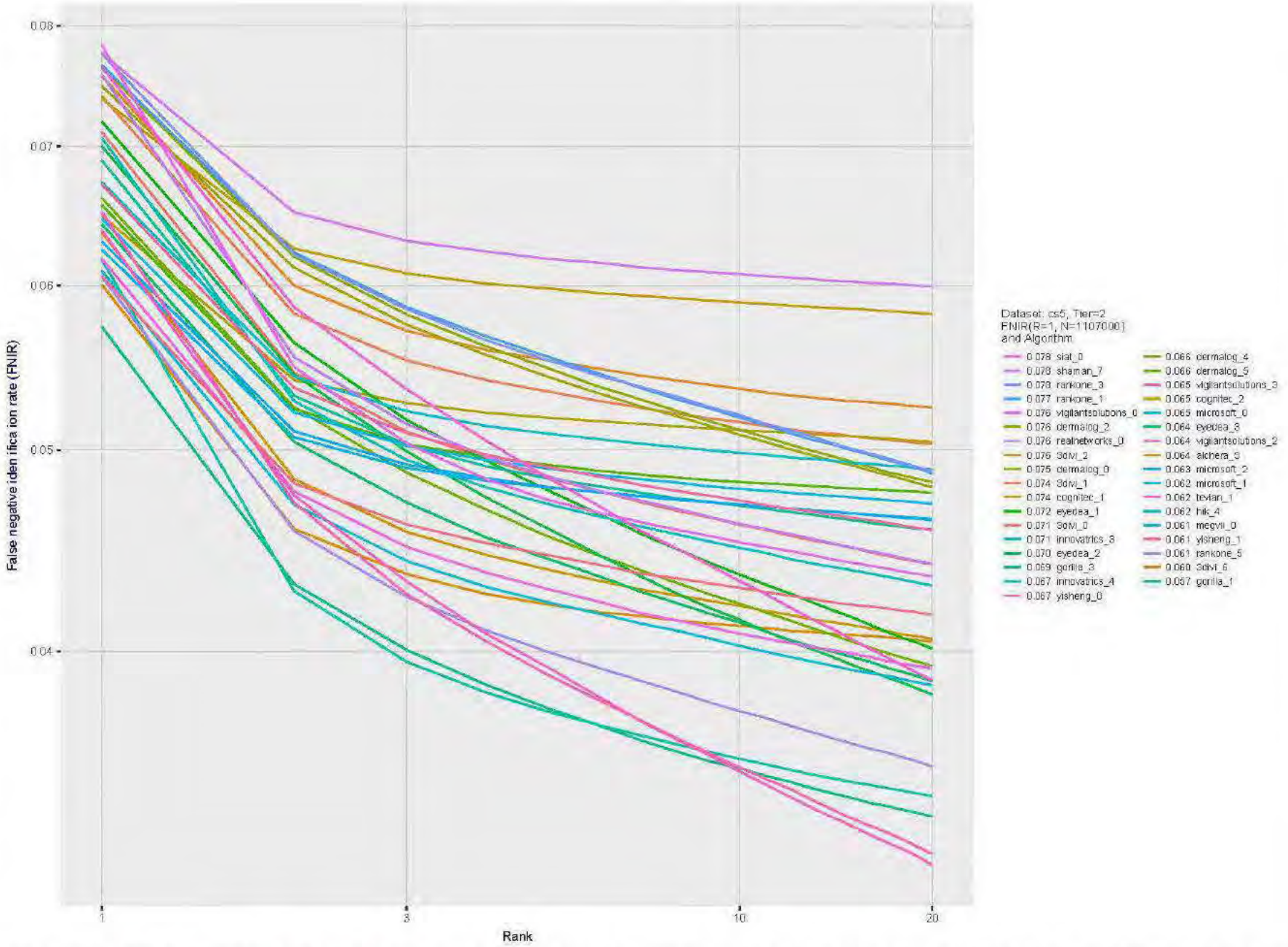


Figure 104: **[Wild Dataset] Identification miss rates vs. rank.** *For the wild dataset, the figure shows false negative identification rates (FNIR) vs. rank when the threshold is set to zero. This metric is relevant to human reviewers who will traverse candidate lists checking whether any of the returned identities match to the search imagery. Specifically, wild images were searched against 1.1 million individuals enrolled with wild images as well.*

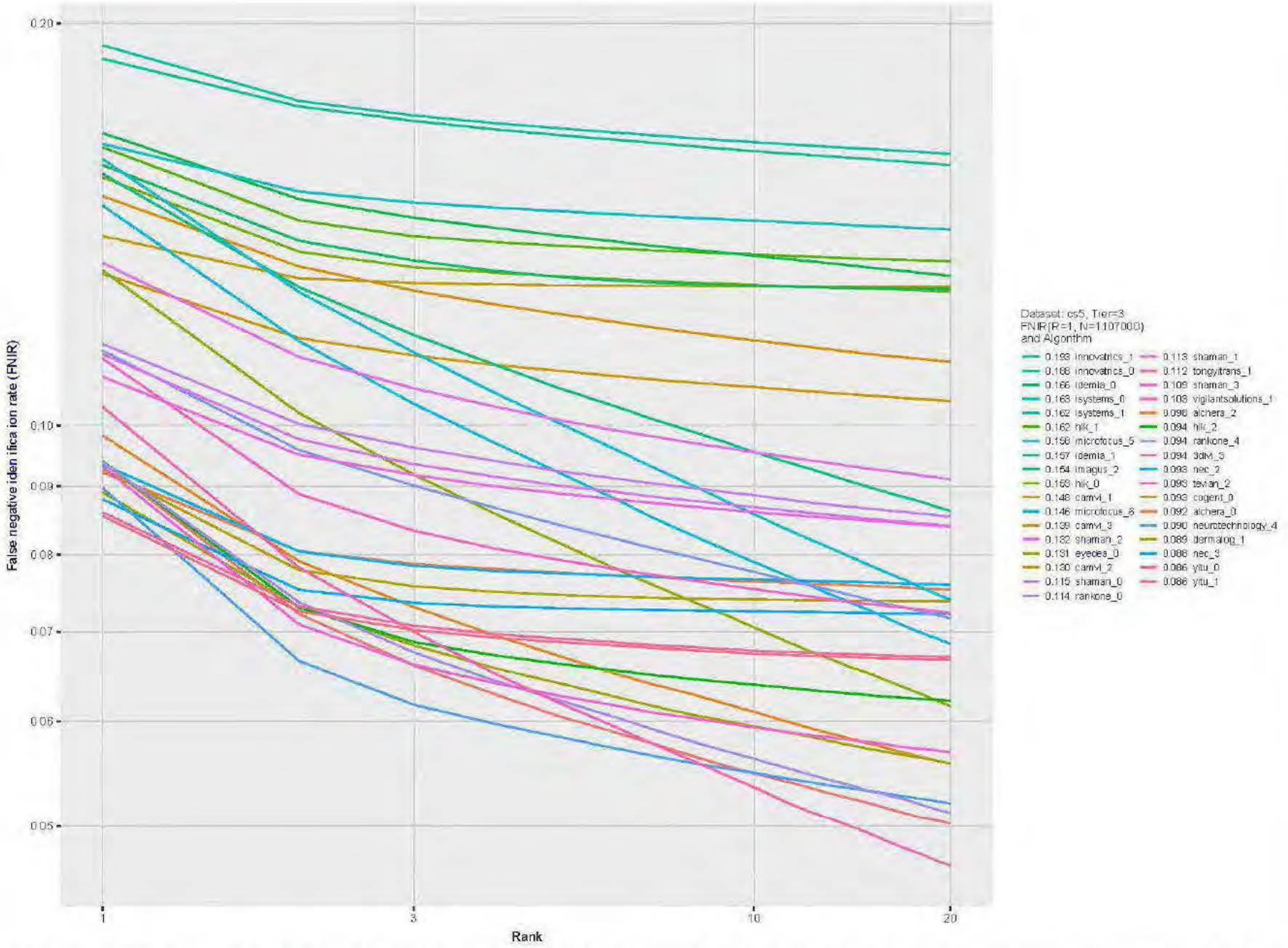


Figure 105: **[Wild Dataset] Identification miss rates vs. rank.** *For the wild dataset, the figure shows false negative identification rates (FNIR) vs. rank when the threshold is set to zero. This metric is relevant to human reviewers who will traverse candidate lists checking whether any of the returned identities match to the search imagery. Specifically, wild images were searched against 1.1 million individuals enrolled with wild images as well.*

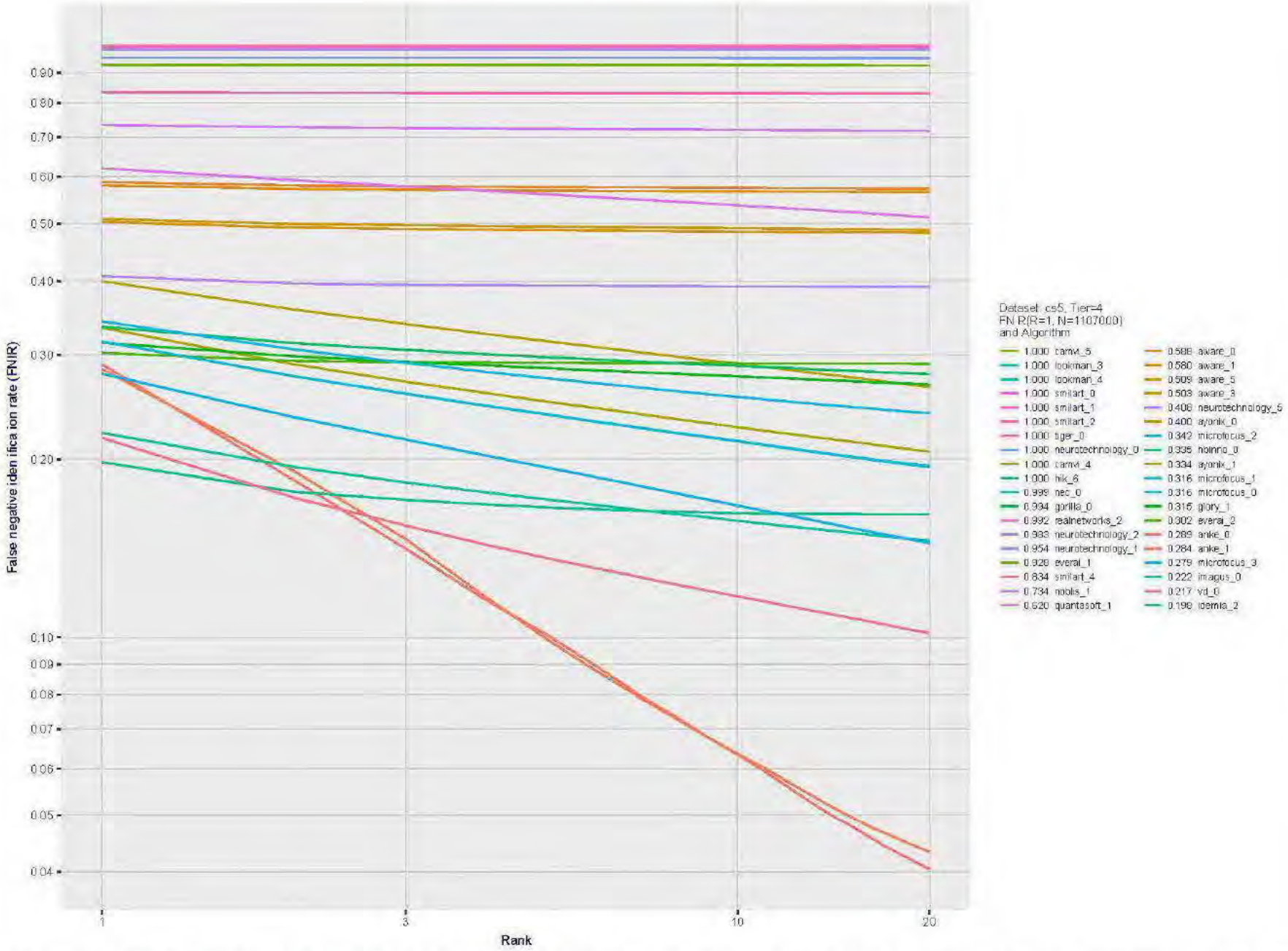
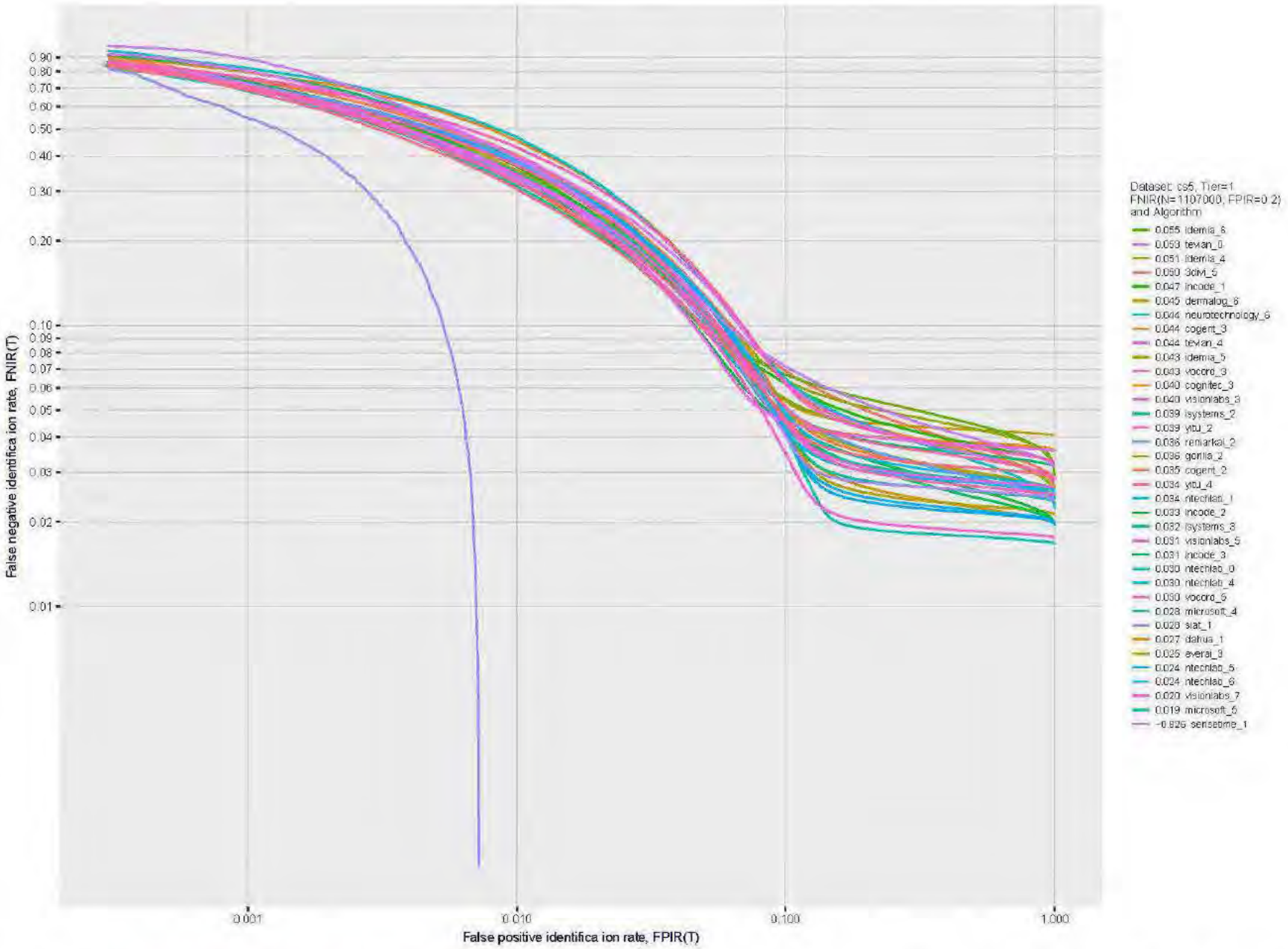


Figure 106: **[Wild Dataset] Identification miss rates vs. rank.** *For the wild dataset, the figure shows false negative identification rates (FNIR) vs. rank when the threshold is set to zero. This metric is relevant to human reviewers who will traverse candidate lists checking whether any of the returned identities match to the search imagery. Specifically, wild images were searched against 1.1 million individuals enrolled with wild images as well.*

2019/09/11 17:24:52	FNIR(N, R, T) = FPIR(N, T) =	False neg. identification rate False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification



2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

Figure 107: **[Wild Dataset] Identification miss rates vs. false positive rates.** *The figure shows accuracy of algorithms on wild images searched against wild images of 1.1 million individuals enrolled into a gallery. On the vertical axis is miss rate FNIR(N, T, L) with N = 1 107 000, as a function of false positive identification FPIR(N, T). The rapid increase in FNIR below FPIR = 0.1 suggests that some background identities in the gallery are actually present in the non-mated search sets. This issue will be addressed in the 2019 revision of this report.*

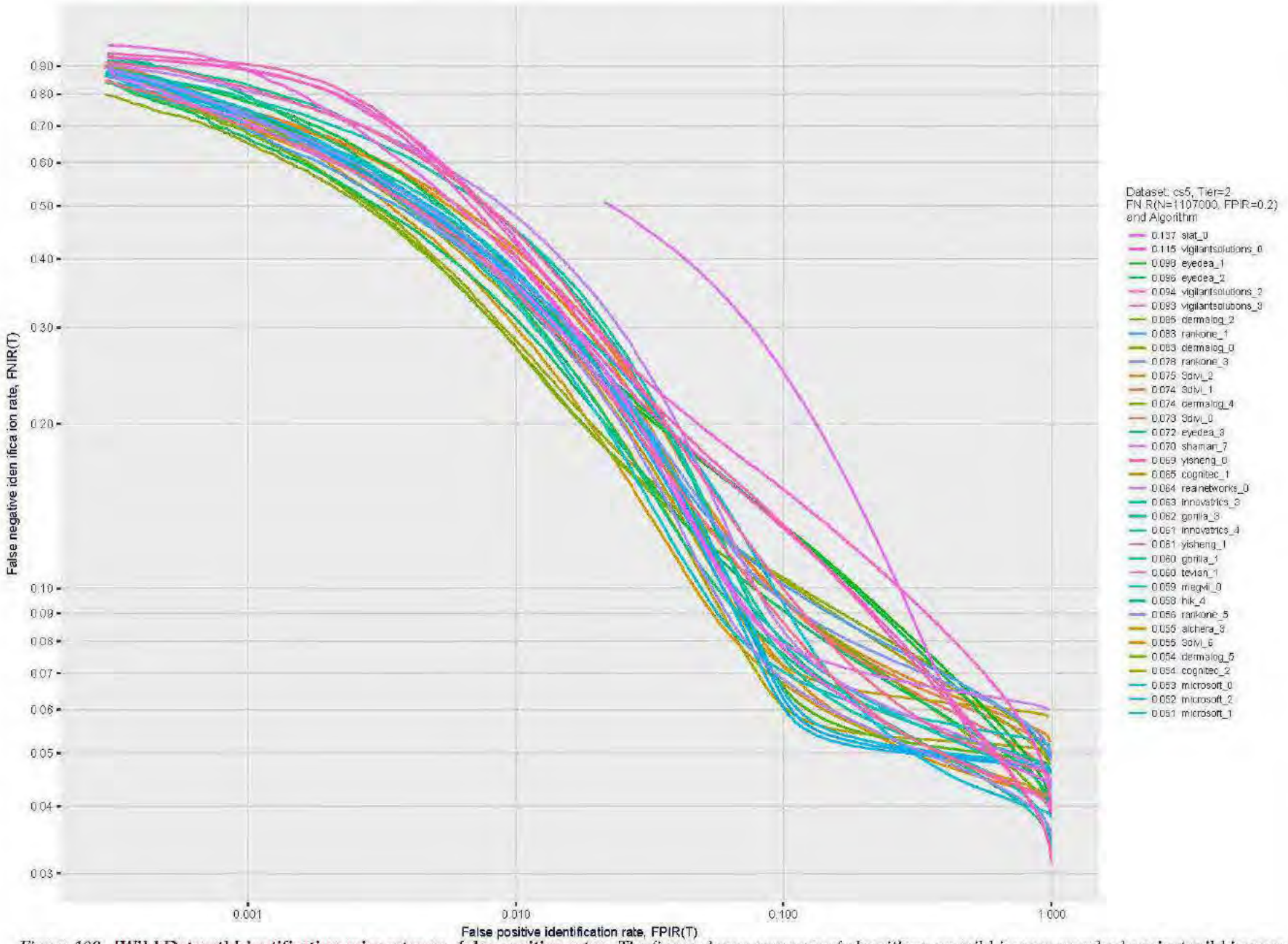


Figure 108: **[Wild Dataset] Identification miss rates vs. false positive rates.** *The figure shows accuracy of algorithms on wild images searched against wild images of 1.1 million individuals enrolled into a gallery. On the vertical axis is miss rate FNIR(N, T, L) with N = 1 107 000, as a function of false positive identification FPIR(N, T). The rapid increase in FNIR below FPIR = 0.1 suggests that some background identities in the gallery are actually present in the non-mated search sets. This issue will be addressed in the 2019 revision of this report.*

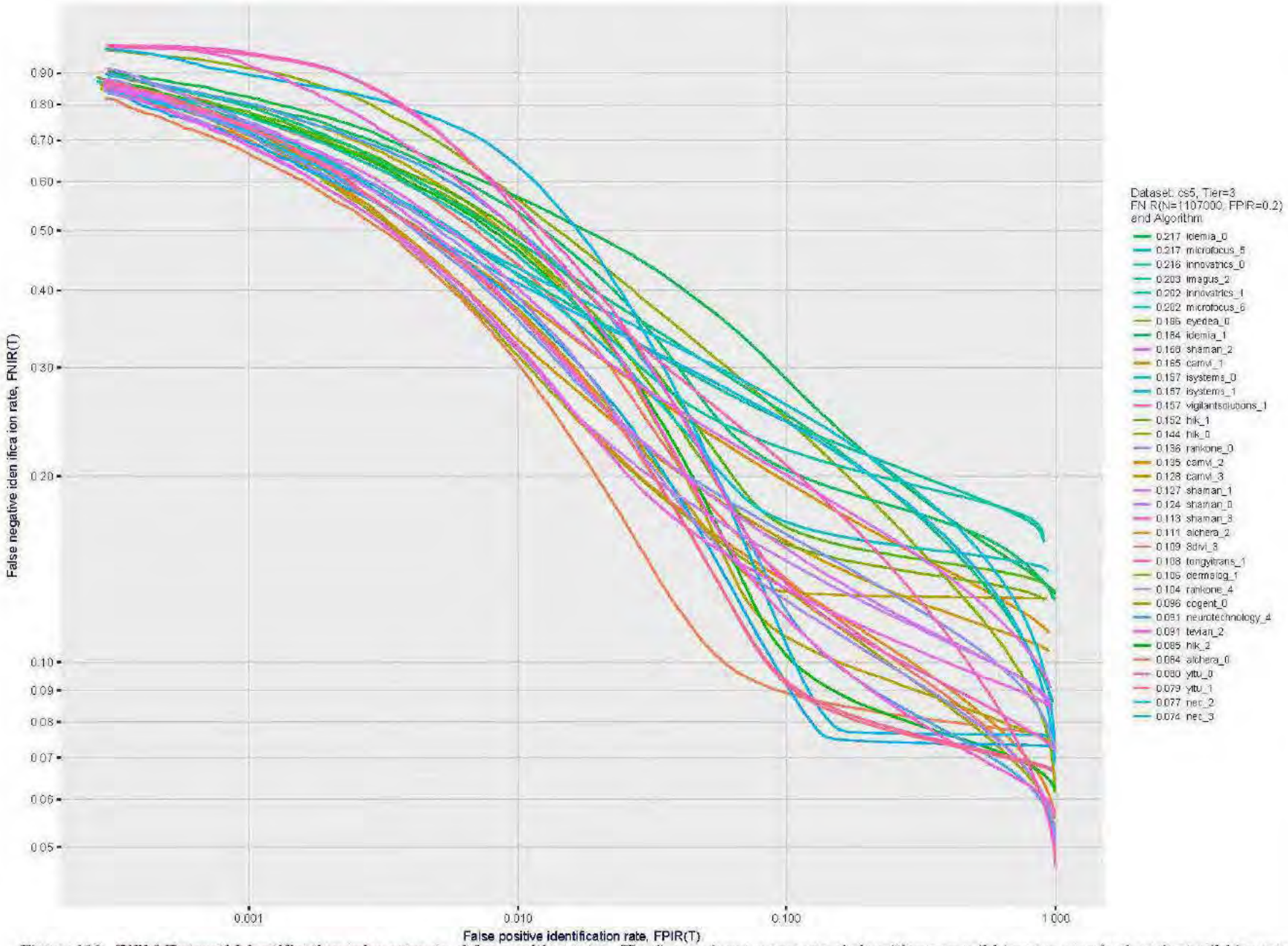


Figure 109: **[Wild Dataset] Identification miss rates vs. false positive rates.** *The figure shows accuracy of algorithms on wild images searched against wild images of 1.1 million individuals enrolled into a gallery. On the vertical axis is miss rate FNIR(N, T, L) with N = 1 107 000, as a function of false positive identification FPIR(N, T). The rapid increase in FNIR below FPIR = 0.1 suggests that some background identities in the gallery are actually present in the non-mated search sets. This issue will be addressed in the 2019 revision of this report.*

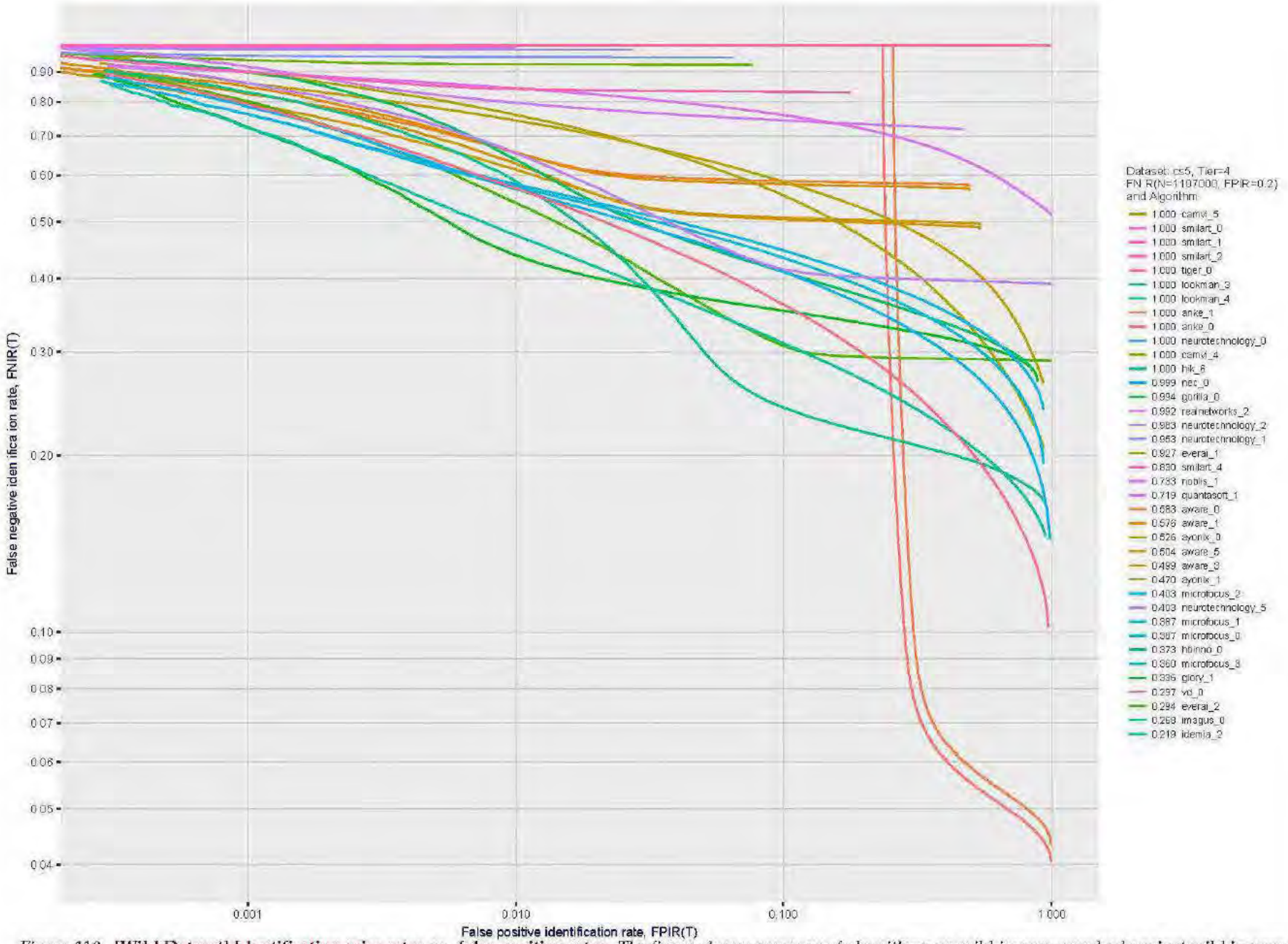


Figure 110: **[Wild Dataset] Identification miss rates vs. false positive rates.** *The figure shows accuracy of algorithms on wild images searched against wild images of 1.1 million individuals enrolled into a gallery. On the vertical axis is miss rate FNIR(N, T, L) with N = 1 107 000, as a function of false positive identification FPIR(N, T). The rapid increase in FNIR below FPIR = 0.1 suggests that some background identities in the gallery are actually present in the non-mated search sets. This issue will be addressed in the 2019 revision of this report.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

Appendix G Search duration

As in and prior tests, this section documents search speeds spanning three orders of magnitude. In applications where search volumes are high enough, this will have implications for hardware requirements especially for large N or when search duration is appreciably larger than the time it takes to prepare a template from the search image(s). Further, given very large (and growing) operational databases, the scalability of algorithms is important. It has been reported previously [8] that search duration can scale sublinearly with enrolled population size N. Further there has been considerable recent research on indexing, exact [13] and approximate nearest neighbor search [1,13] and fast-search [14,16].

Figure 111 charts the search duration measurements presented earlier in Tables 6 - 9.

- Most algorithms scale linearly. For those in that category, there is a wide range in speed with search durations ranging from 82 milliseconds for a 12 million gallery (for NEC-3) to more than 40 seconds (for Yitu-3, Toshiba-2) and even higher for less accurate algorithms.
- Some developers (Camvi, Dermalog, EverAI, Innovatrics, and Visionlabs) provide algorithms whose template search durations grow logarithmically i.e. approximately $T(N) = a \log N$ with the constant a varying between implementations. In the figure this model is fit using the point $T(1) = 0$, and $T(640\,000)$. This very sublinear behaviour affords extremely fast search times in very large galleries. One caveat for the sublinear algorithms is that the fast-search data structures require considerable computation time - on the order of hours - for N in the millions, and this scales mildy super-linearly, i.e. $O(N^b), b > 1$. There are exceptions: the Camvi algorithms take minutes; and Innvovatrics' scale sublinearly.

2019/09/11 17:24:52	FNIR(N, R, T) = False neg. identification rate FPIR(N, T) = False pos. identification rate	N = Num. enrolled subjects R = Num. candidates examined	T = Threshold	T = 0 → Investigation T > 0 → Identification	

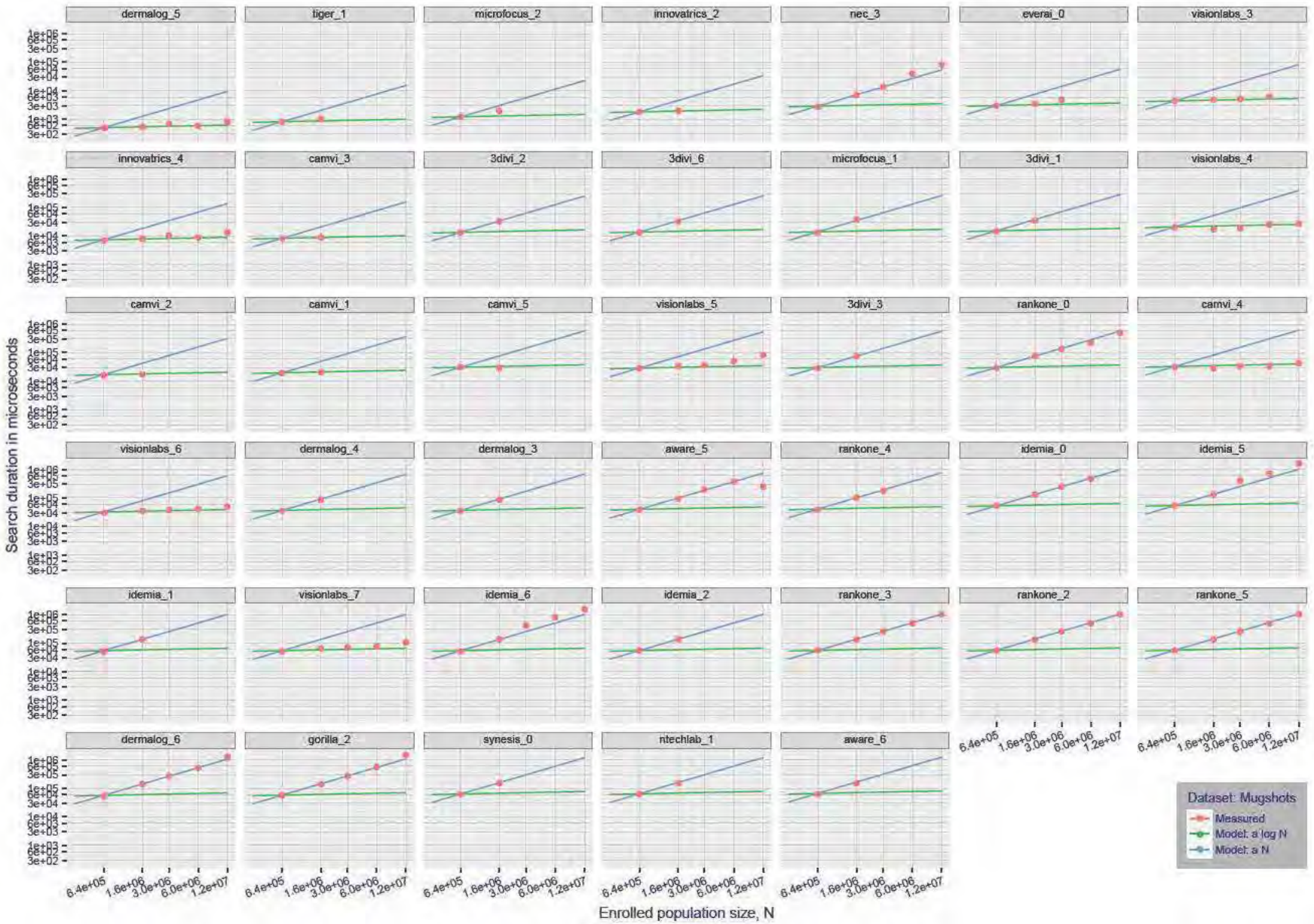


Figure 111: **[Mugshot Dataset] Search duration vs. enrolled population size.** *In red are the actual point durations measured on a single c. 2016 core. The blue shows linear growth from $N = 640\,000$. The green line shows logathmic growth from that point to $N = 1\,600\,000$. Note the sublinear growth from algorithms from Camvi, Dermalog, EverAI, Innovatrics, and Visionlabs. The tiger_1 algorithm is also sublinear, but inaccurate and inoperable at $N \geq 3000000$. This capability sometimes comes at the additional expense of converting a linear gallery data structure into whatever fast-search data structure is used. Note that search times are sometimes dominated by the template generation times shown in Table 16.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate
N = Num. enrolled subjects
R = Num. candidates examined
T = Threshold
T = 0 → Investigation
T > 0 → Identification

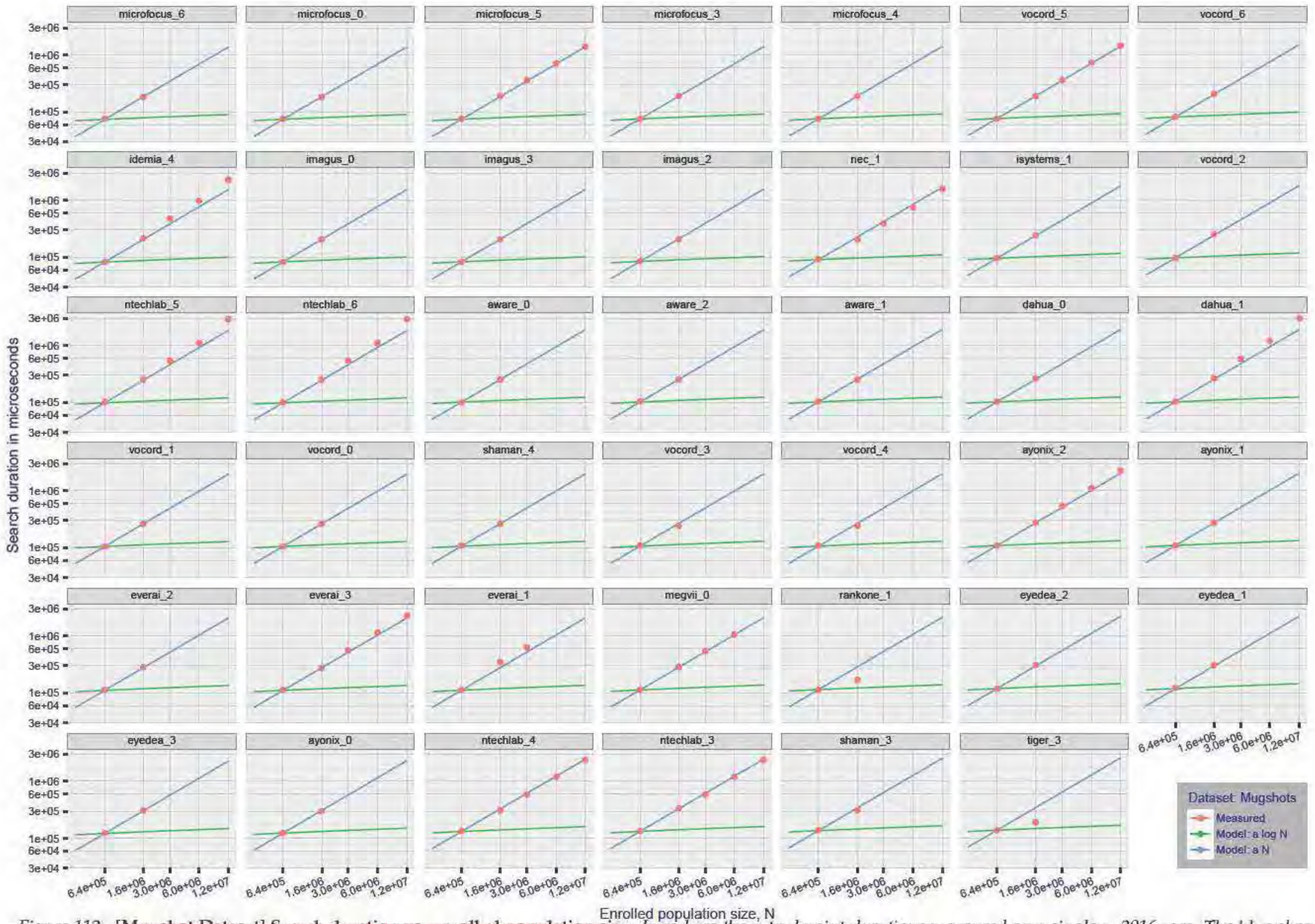


Figure 112: **[Mugshot Dataset] Search duration vs. enrolled population size.** *In red are the actual point durations measured on a single c. 2016 core. The blue shows linear growth from $N = 640\,000$. The green line shows logathmic growth from that point to $N = 1\,600\,000$. Note the sublinear growth from algorithms from Camvi, Dermalog, EverAI, Innovatrics, and Visionlabs. The tiger_1 algorithm is also sublinear, but inaccurate and inoperable at $N \geq 3000000$. This capability sometimes comes at the additional expense of converting a linear gallery data structure into whatever fast-search data structure is used. Note that search times are sometimes dominated by the template generation times shown in Table 16.*

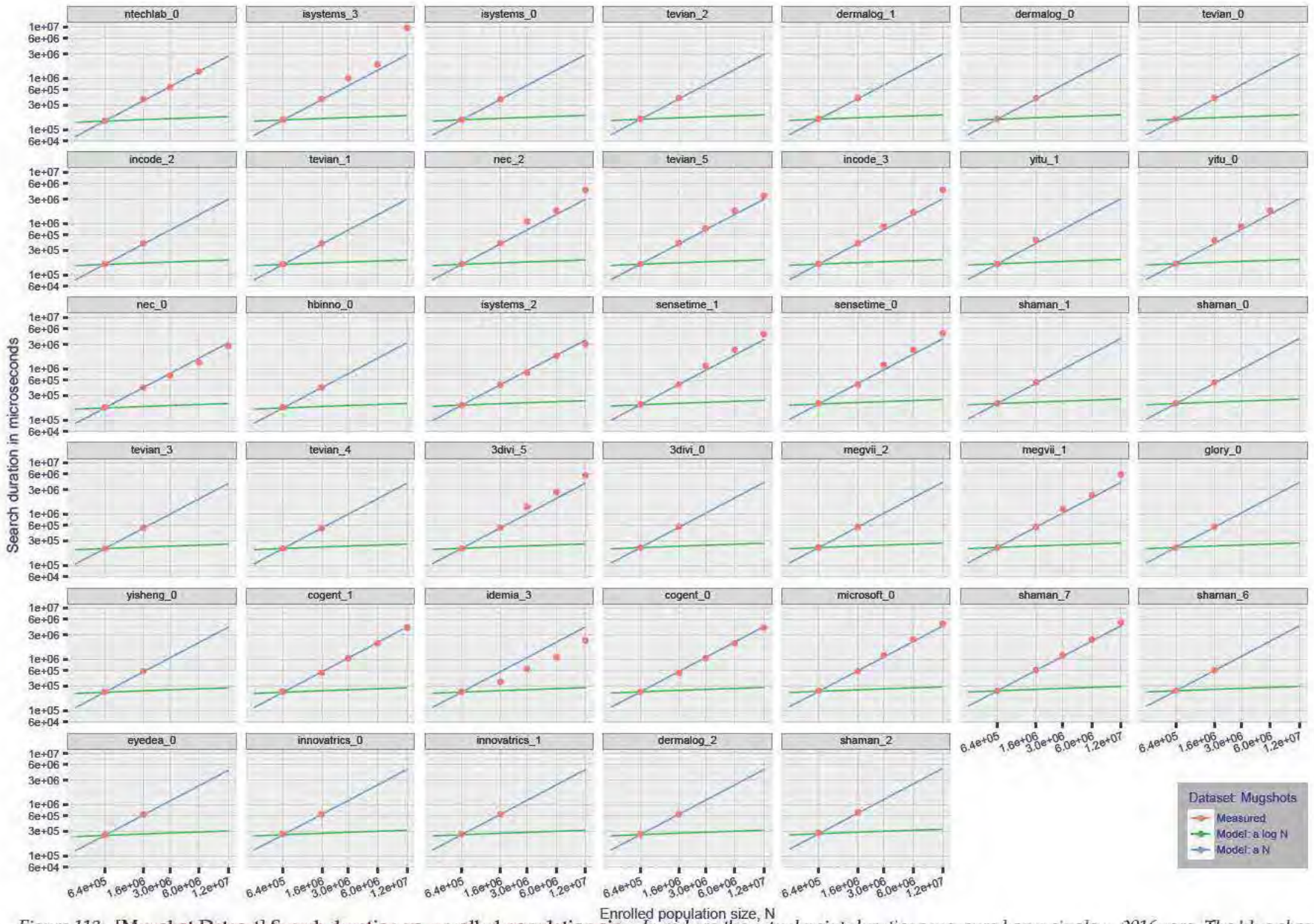


Figure 113: **[Mugshot Dataset] Search duration vs. enrolled population size.** *In red are the actual point durations measured on a single c. 2016 core. The blue shows linear growth from $N = 640\,000$. The green line shows logathmic growth from that point to $N = 1\,600\,000$. Note the sublinear growth from algorithms from Camvi, Dermalog, EverAI, Innovatrics, and Visionlabs. The tiger_1 algorithm is also sublinear, but inaccurate and inoperable at $N \geq 3000000$. This capability sometimes comes at the additional expense of converting a linear gallery data structure into whatever fast-search data structure is used. Note that search times are sometimes dominated by the template generation times shown in Table 16.*

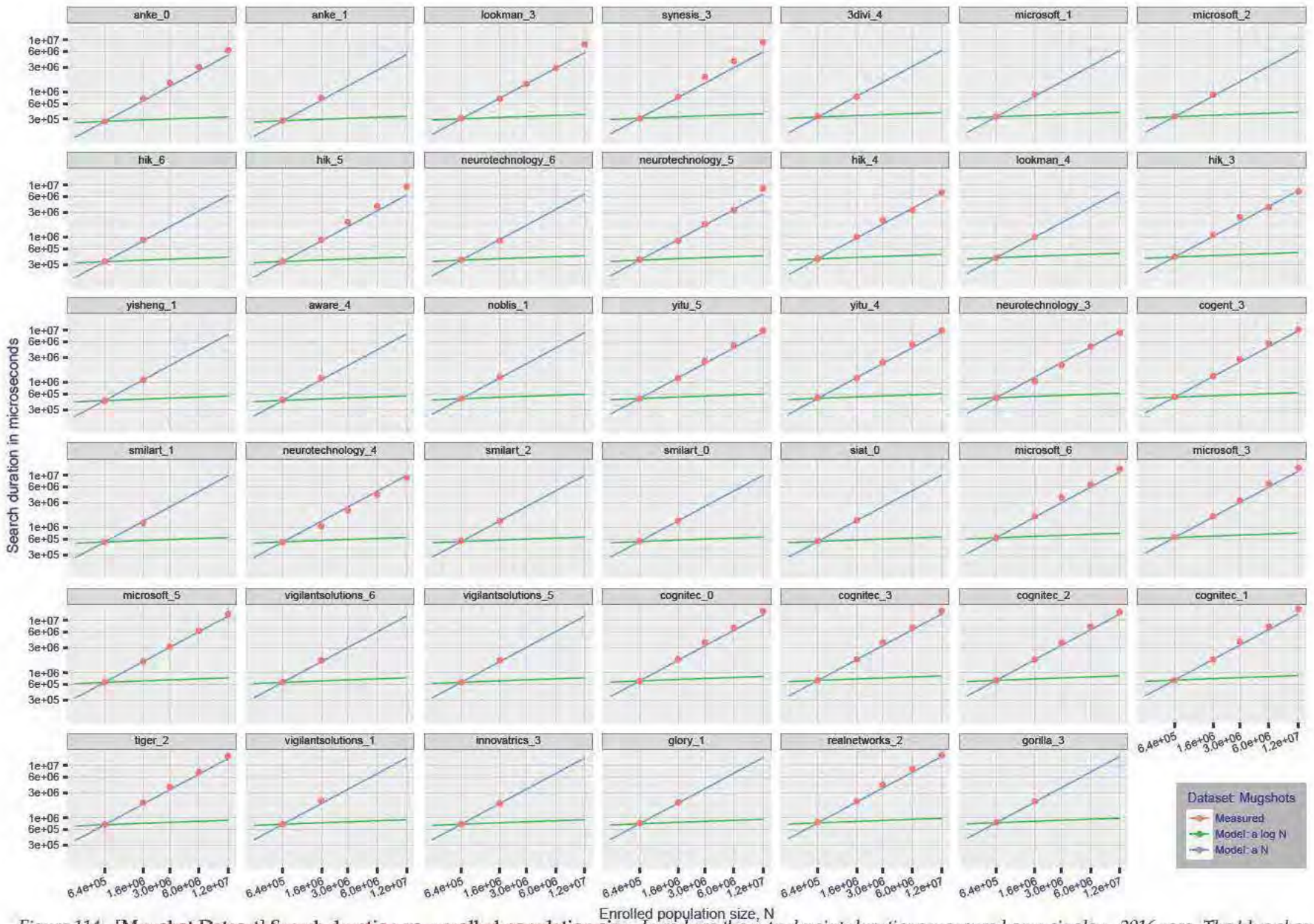


Figure 114: **[Mugshot Dataset] Search duration vs. enrolled population size.** *In red are the actual point durations measured on a single c. 2016 core. The blue shows linear growth from $N = 640\,000$. The green line shows logathmic growth from that point to $N = 1\,600\,000$. Note the sublinear growth from algorithms from Camvi, Dermalog, EverAI, Innovatrics, and Visionlabs. The tiger_1 algorithm is also sublinear, but inaccurate and inoperable at $N \geq 3000000$. This capability sometimes comes at the additional expense of converting a linear gallery data structure into whatever fast-search data structure is used. Note that search times are sometimes dominated by the template generation times shown in Table 16.*

2019/09/11 17:24:52

FNIR(N, R, T) = False neg. identification rate
FPIR(N, T) = False pos. identification rate

N = Num. enrolled subjects
R = Num. candidates examined

T = Threshold

T = 0 → Investigation
T > 0 → Identification

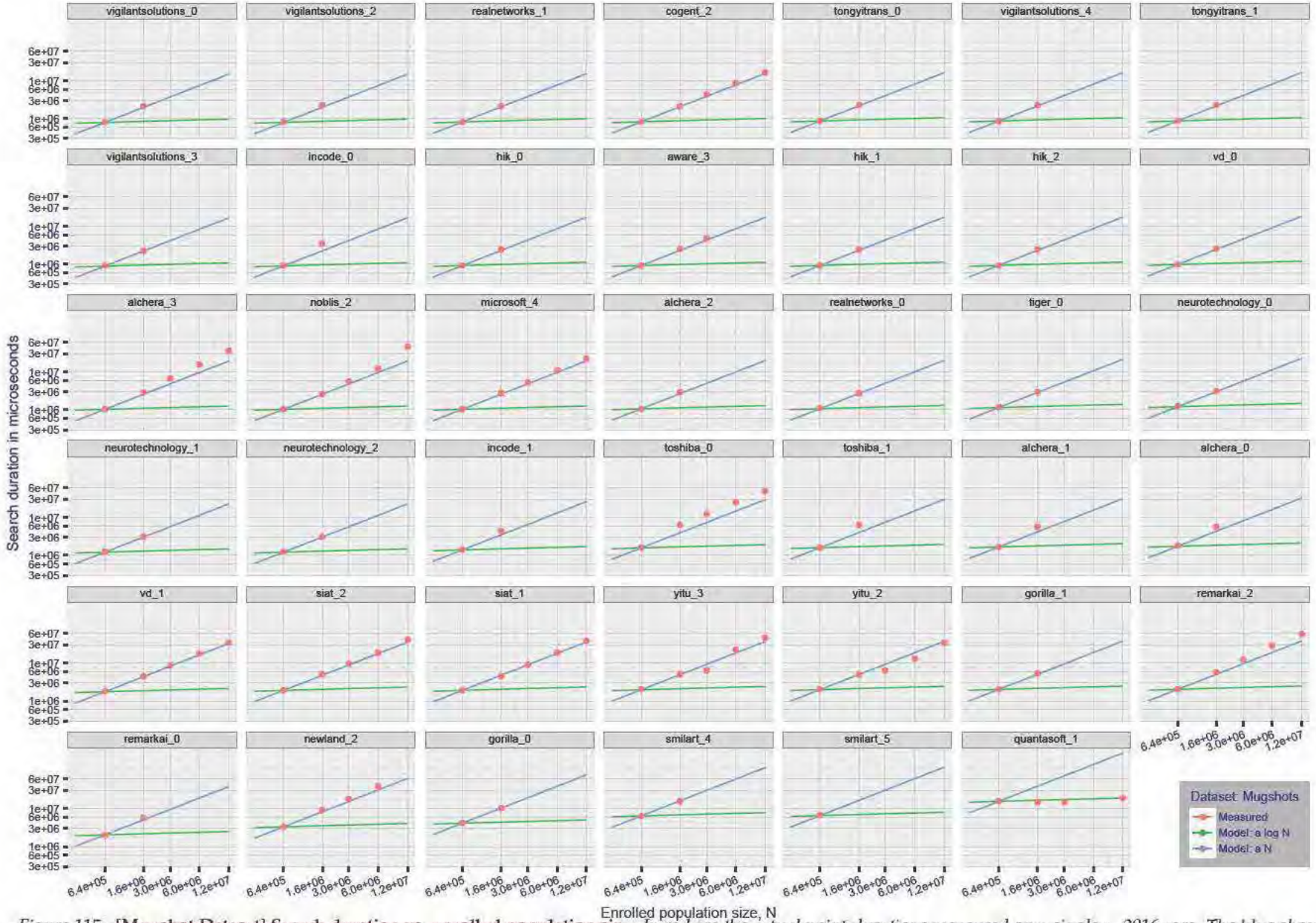


Figure 115: **[Mugshot Dataset] Search duration vs. enrolled population size.** *In red are the actual point durations measured on a single c. 2016 core. The blue shows linear growth from $N = 640\,000$. The green line shows logathmic growth from that point to $N = 1\,600\,000$. Note the sublinear growth from algorithms from Camvi, Dermalog, EverAI, Innovatrics, and Visionlabs. The tiger_1 algorithm is also sublinear, but inaccurate and inoperable at $N \geq 3000000$. This capability sometimes comes at the additional expense of converting a linear gallery data structure into whatever fast-search data structure is used. Note that search times are sometimes dominated by the template generation times shown in Table 16.*

Appendix H Gallery Insertion Timing

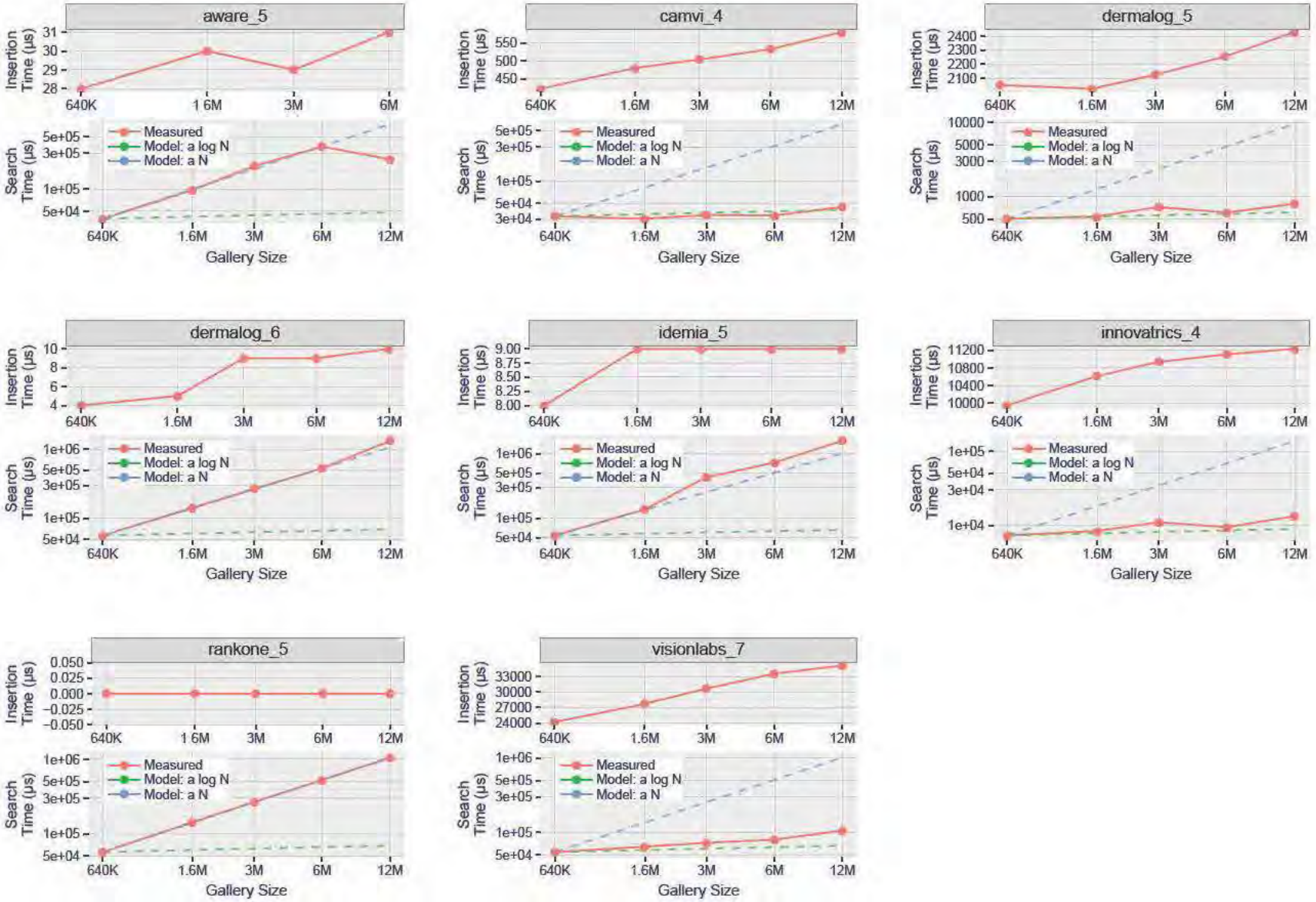


Figure 116: **[Mugshot Dataset] Gallery insertion duration vs. enrolled population size.** *This chart plots the time it takes to insert a single template into a finalized gallery, illustrated over increasing gallery sizes. For reference, search times on finalized galleries of corresponding sizes are plotted right underneath. Gallery insertion time plots were generated on algorithms that 1) successfully implemented gallery insertion with no errors and 2) that were run on galleries with N up to 12 000 000. Generally, only the more accurate algorithms were run on galleries with N up to 12 000 000.*

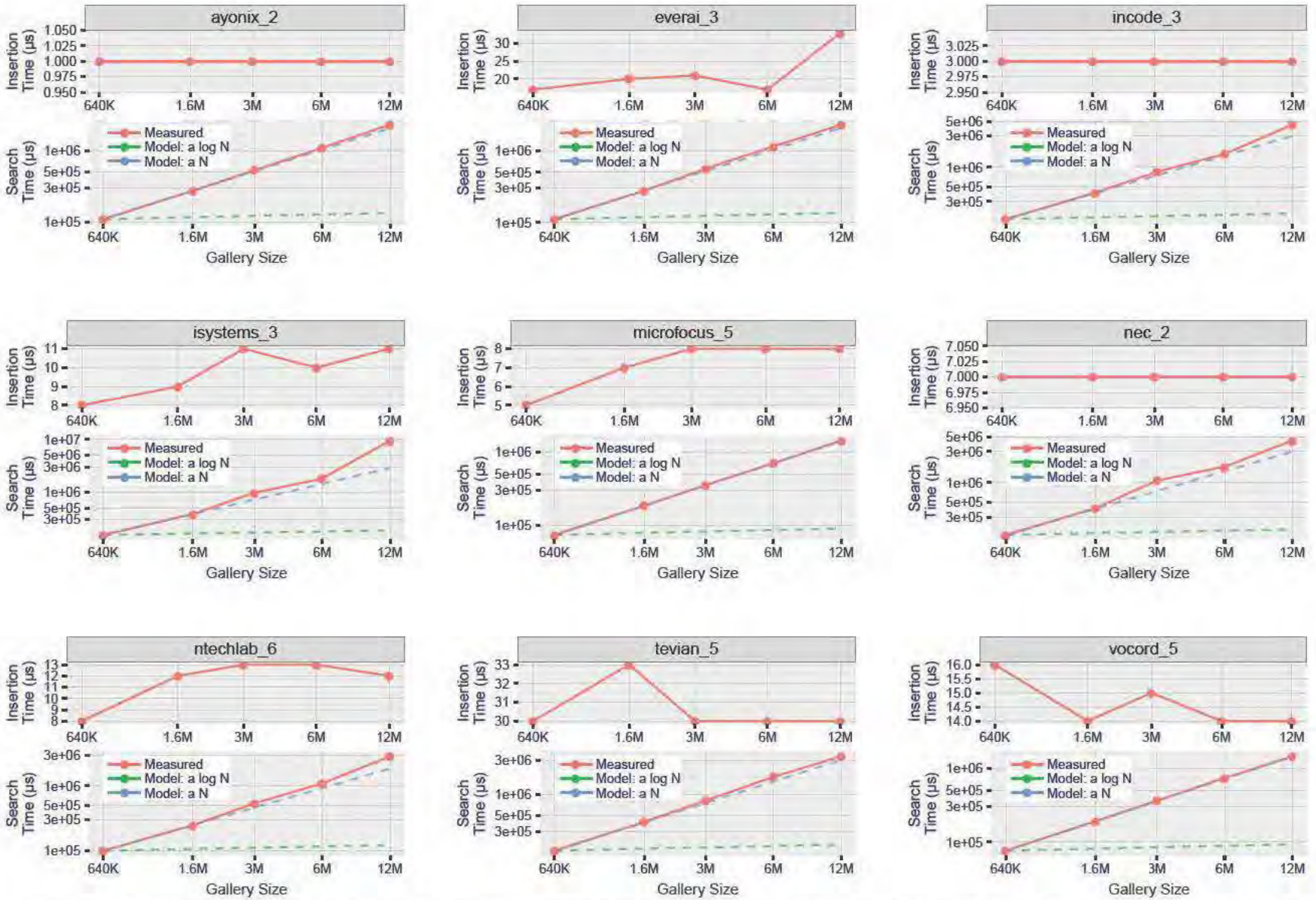


Figure 117: **[Mugshot Dataset] Gallery insertion duration vs. enrolled population size.** *This chart plots the time it takes to insert a single template into a finalized gallery, illustrated over increasing gallery sizes. For reference, search times on finalized galleries of corresponding sizes are plotted right underneath. Gallery insertion time plots were generated on algorithms that 1) successfully implemented gallery insertion with no errors and 2) that were run on galleries with N up to 12 000 000. Generally, only the more accurate algorithms were run on galleries with N up to 12 000 000.*

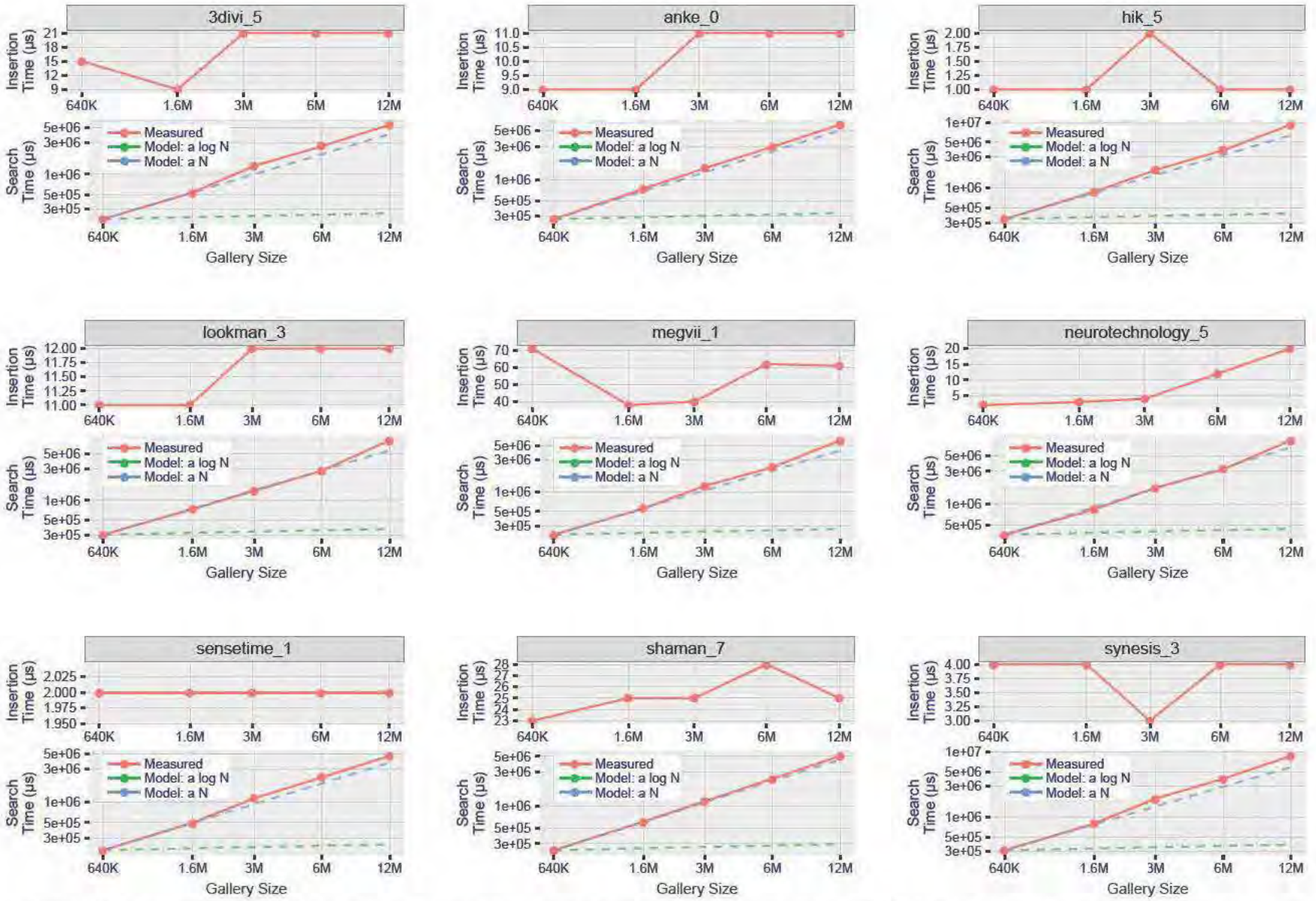


Figure 118: **[Mugshot Dataset] Gallery insertion duration vs. enrolled population size.** *This chart plots the time it takes to insert a single template into a finalized gallery, illustrated over increasing gallery sizes. For reference, search times on finalized galleries of corresponding sizes are plotted right underneath. Gallery insertion time plots were generated on algorithms that 1) successfully implemented gallery insertion with no errors and 2) that were run on galleries with N up to 12 000 000. Generally, only the more accurate algorithms were run on galleries with N up to 12 000 000.*

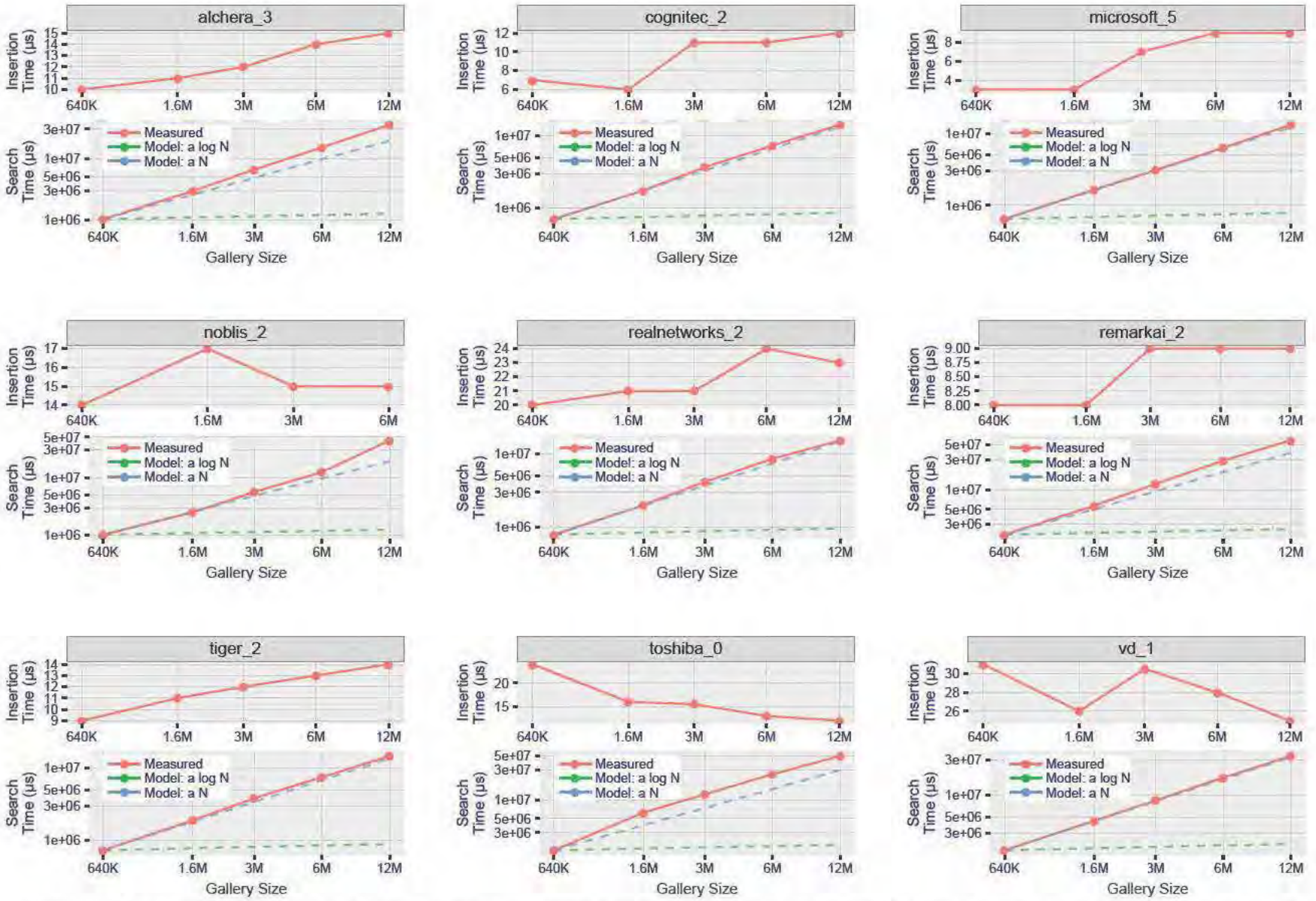


Figure 119: **[Mugshot Dataset] Gallery insertion duration vs. enrolled population size.** *This chart plots the time it takes to insert a single template into a finalized gallery, illustrated over increasing gallery sizes. For reference, search times on finalized galleries of corresponding sizes are plotted right underneath. Gallery insertion time plots were generated on algorithms that 1) successfully implemented gallery insertion with no errors and 2) that were run on galleries with N up to 12 000 000. Generally, only the more accurate algorithms were run on galleries with N up to 12 000 000.*

References

[1] Artem Babenko and Victor Lempitsky. Efficient indexing of billion-scale datasets of deep descriptors. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[2] L. Best-Rowden and A. K. Jain. Longitudinal study of automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(1):148–162, Jan 2018.

[3] Blumstein, Cohen, Roth, and Visher, editors. *Random parameter stochastic models of criminal careers*. National Academy of Sciences Press, 1986.

[4] Thomas P. Bonczar and Lauren E. Glaze. Probation and parole in the united statesm 2007, statistical tables. Technical report, Bureau of Justice Statistics, December 2008.

[5] White D., Kemp R. I., Jenkins R., Matheson M, and Burton A. M. Passport officers errors in face matching. *PLoS ONE*, 9(8), 2014. e103510. doi:10.1371/journal. pone.0103510.

[6] P. Grother, G. W. Quinn, and P. J. Phillips. Evaluation of 2d still-image face recognition algorithms. NIST Interagency Report 7709, National Institute of Standards and Technology, 8 2010. http://face.nist.gov/mbe as MBE2010 FRVT2010.

[7] P. J. Grother, R. J. Micheals, and P. J. Phillips. Performance metrics for the frvt 2002 evaluation. In *Proceedings of Audio and Video Based Person Authentication Conference (AVBPA)*, June 2003.

[8] Patrick Grother and Mei Ngan. Interagency report 8009, performance of face identification algorithms. *Face Recognition Vendor Test (FRVT)*, May 2014.

[9] Patrick Grother, George Quinn, and Mei Ngan. Face in video evaluation (five) face recognition of non-cooperative subjects. Interagency Report 8173, National Institute of Standards and Technology, March 2017. https://doi.org/10.6028/NIST.IR.8173.

[10] Patrick Grother, George W. Quinn, and Mei Ngan. Face recognition vendor test - still face image and video concept, evaluation plan and api. Technical report, National Institute of Standards and Technology, 7 2013. http://biometrics.nist.gov/cs_links/face/frvt/frvt2012/NIST_FRVT2012_api_Aug15.pdf.

[11] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 770–778, June 2016.

[12] Gary B. Huang, Manu Ramesh, Tamara Berg, and Erik Learned-Miller. Labeled faces in the wild: A database for studying face recognition in unconstrained environments. Technical Report 07-49, University of Massachusetts, Amherst, October 2007.

[13] Masato Ishii, Hitoshi Imaoka, and Atsushi Sato. Fast k-nearest neighbor search for face identification using bounds of residual score. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pages 194–199, Los Alamitos, CA, USA, May 2017. IEEE Computer Society.

[14] Jeff Johnson, Matthijs Douze, and Hervé Jégou. Billion-scale similarity search with gpus. *CoRR*, abs/1702.08734, 2017.

[15] Ira Kemelmacher-Shlizerman, Steven M. Seitz, Daniel Miller, and Evan Brossard. The megaface benchmark: 1 million faces for recognition at scale. *CoRR*, abs/1512.00596, 2015.

[16] Yury A. Malkov and D. A. Yashunin. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *CoRR*, abs/1603.09320, 2016.

[17] Joyce A. Martin, Brady E. Hamilton, Michelle J.K. Osterman, Anne K. Driscoll, , and Patrick Drake. National vital statistics reports. Technical Report 8, Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System, Division of Vital Statistics, November 2018.

[18] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition. In *British Machine Vision Conference*, 2015.

[19] P. Jonathon Phillips, Amy N. Yates, Ying Hu, Carina A. Hahn, Eilidh Noyes, Kelsey Jackson, Jacqueline G. Cavazos, Géraldine Jeckeln, Rajeev Ranjan, Swami Sankaranarayanan, Jun-Cheng Chen, Carlos D. Castillo, Rama Chellappa, David White, and Alice J. O'Toole. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences*, 115(24):6171–6176, 2018.

[20] Florian Schroff, Dmitry Kalenichenko, and James Philbin. Facenet: A unified embedding for face recognition and clustering. *CoRR*, abs/1503.03832, 2015.

[21] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. *CoRR*, abs/1409.1556, 2014.

[22] Jeroen Smits and Christiaan Monden. Twinning across the developing world. *PLOS ONE*, 6(9):1–5, 09 2011.

[23] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott E. Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. *CoRR*, abs/1409.4842, 2014.

[24] Yaniv Taigman, Ming Yang, Marc'Aurelio Ranzato, and Lior Wolf. Deepface: Closing the gap to human-level performance in face verification. In *Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, CVPR '14, pages 1701–1708, Washington, DC, USA, 2014. IEEE Computer Society.

[25] A. Towler, R. I. Kemp, and D White. *Unfamiliar face matching systems in applied settings*. Nova Science, 2017.

[26] Working Group 3. Ed. M. Werner. *ISO/IEC 19794-5 Information Technology - Biometric Data Interchange Formats - Part 5: Face image data*. JTC1 :: SC37, 2 edition, 2011. http://webstore.ansi.org.

[27] David White, James D. Dunn, Alexandra C. Schmid, and Richard I. Kemp. Error rates in users of automatic face recognition software. *PLoS ONE*, 10:1–14, October 2015.

[28] Bradford Wing and R. Michael McCabe. Special publication 500-271: American national standard for information systems data format for the interchange of fingerprint, facial, and other biometric information part 1. Technical report, NIST, September 2015. ANSI/NIST ITL 1-2015.

[29] Andreas Wolf. Portrait quality - (reference facial images for mrtd). Technical report, ICAO, April 2018.

[30] D. Yadav, N. Kohli, P. Pandey, R. Singh, M. Vatsa, and A. Noore. Effect of illicit drug abuse on face recognition. In *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 1–7, Los Alamitos, CA, USA, mar 2016. IEEE Computer Society.

NISTIR 8280

Face Recognition Vendor Test (FRVT)
Part 3: Demographic Effects

Patrick Grother
Mei Ngan
Kayee Hanaoka

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8280

NIST
National Institute of
Standards and Technology
U.S. Department of Commerce

NISTIR 8280

Face Recognition Vendor Test (FRVT)
Part 3: Demographic Effects

Patrick Grother
Mei Ngan
Kayee Hanaoka
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8280

December 2019

U.S. Department of Commerce
Wilbur L. Ross, Jr., Secretary

National Institute of Standards and Technology
Walter Copan, NIST Director and Undersecretary of Commerce for Standards and Technology

National Institute of Standards and Technology Interagency or Internal Report 8280
Natl. Inst. Stand. Technol. Interag. Intern. Rep. 8280, 81 pages (December 2019)

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8280

EXECUTIVE SUMMARY

OVERVIEW This is the third in a series of reports on ongoing face recognition vendor tests (FRVT) executed by the National Institute of Standards and Technology (NIST). The first two reports cover, respectively, the performance of one-to-one face recognition algorithms used for verification of asserted identities, and performance of one-to-many face recognition algorithms used for identification of individuals in photo data bases. This document extends those evaluations to document accuracy variations across demographic groups.

MOTIVATION The recent expansion in the availability, capability, and use of face recognition has been accompanied by assertions that demographic dependencies could lead to accuracy variations and potential bias. A report from Georgetown University [14] work noted that prior studies [22], articulated sources of bias, described the potential impacts particularly in a policing context, and discussed policy and regulatory implications. Additionally, this work is motivated by studies of demographic effects in more recent face recognition [9, 16, 23] and gender estimation algorithms [5, 36].

AIMS AND SCOPE NIST has conducted tests to quantify demographic differences in contemporary face recognition algorithms. This report provides details about the recognition process, notes where demographic effects could occur, details specific performance metrics and analyses, gives empirical results, and recommends research into the mitigation of performance deficiencies.

NIST intends this report to inform discussion and decisions about the accuracy, utility, and limitations of face recognition technologies. Its intended audience includes policy makers, face recognition algorithm developers, systems integrators, and managers of face recognition systems concerned with mitigation of risks implied by demographic differentials.

WHAT WE DID The NIST Information Technology Laboratory (ITL) quantified the accuracy of face recognition algorithms for demographic groups defined by sex, age, and race or country of birth.

We used both one-to-one verification algorithms and one-to-many identification search algorithms. These were submitted to the FRVT by corporate research and development laboratories and a few universities. As prototypes, these algorithms were not necessarily available as mature integrable products. Their performance is detailed in FRVT reports [16, 17].

We used these algorithms with four large datasets of photographs collected in U.S. governmental applications that are currently in operation:

- ▷ **Domestic mugshots** collected in the United States.
- ▷ **Application photographs** from a global population of applicants for immigration benefits.
- ▷ **Visa photographs** submitted in support of visa applicants.
- ▷ **Border crossing photographs** of travelers entering the United States.

All four datasets were collected for authorized travel, immigration or law enforcement processes. The first three sets have good compliance with image capture standards. The last set does not, given constraints on capture duration and environment. Together these datasets allowed us to process a total of 18.27 million images of 8.49 million people through 189 mostly commercial algorithms from 99 developers.

The datasets were accompanied by sex and age metadata for the photographed individuals. The mugshots have metadata for race, but the other sets only have country-of-birth information. We restrict the analysis to 24 countries in 7 distinct global regions that have seen lower levels of long-distance immigration. While country-of-birth information may be a reasonable proxy for race in these countries, it stands as a meaningful factor in its own right particularly for travel-related applications of face recognition.

The tests aimed to determine whether, and to what degree, face recognition algorithms differed when they processed photographs of individuals from various demographics. We assessed accuracy by demographic group and report on false negative and false positive effects. False negatives are the failure to associate one person in two images; they occur when the similarity between two photos is low, reflecting either some change in the person's appearance or in the image properties. False positives are the erroneous association of samples of two persons; they occur when the digitized faces of two people are similar.

In background material that follows we give examples of how algorithms are used, and we elaborate on the consequences of errors noting that the impacts of demographic differentials can be advantageous or disadvantageous depending on the application.

WHAT WE FOUND

The accuracy of algorithms used in this report has been documented in recent FRVT evaluation reports [16, 17]. These show a wide range in accuracy across developers, with the most accurate algorithms producing many fewer errors. These algorithms can therefore be expected to have smaller demographic differentials.

Contemporary face recognition algorithms exhibit demographic differentials of various magnitudes. Our main result is that false positive differentials are much larger than those related to false negatives and exist broadly, across many, but not all, algorithms tested. Across demographics, false positives rates often vary by factors of 10 to beyond 100 times. False negatives tend to be more algorithm-specific, and vary often by factors below 3.

- **False positives:** Using the higher quality Application photos, false positive rates are highest in West and East African and East Asian people, and lowest in Eastern European individuals. This effect is generally large, with a factor of 100 more false positives between countries. However, with a number of algorithms developed in China this effect is reversed, with low false positive rates on East Asian faces. With domestic law enforcement images, the highest false positives are in American Indians, with elevated rates in African American and Asian populations; the relative ordering depends on sex and varies with algorithm.

  We found false positives to be higher in women than men, and this is consistent across algorithms and datasets. This effect is smaller than that due to race.

  We found elevated false positives in the elderly and in children; the effects were larger in the oldest and youngest, and smallest in middle-aged adults.

▷ **False negatives:** With domestic mugshots, false negatives are higher in Asian and American Indian individuals, with error rates above those in white and African American faces (which yield the lowest false negative rates). However, with lower-quality border crossing images, false negatives are generally higher in people born in Africa and the Caribbean, the effect being stronger in older individuals. These differing results relate to image quality: The mugshots were collected with a photographic setup specifically standardized to produce high-quality images across races; the border crossing images deviate from face image quality standards.

In cooperative access control applications, false negatives can be remedied by users making second attempts.

The presence of an enrollment database affords one-to-many identification algorithms a resource for mitigation of demographic effects that purely one-to-one verification systems do not have. Nevertheless, demographic differentials present in one-to-one verification algorithms are usually, but not always, present in one-to-many search algorithms. One important exception is that some developers supplied highly accurate identification algorithms for which false positive differentials are undetectable.

More detailed results are introduced in the Technical Summary.

IMPLICATIONS OF THESE TESTS

Operational implementations usually employ a single face recognition algorithm. Given algorithm-specific variation, it is incumbent upon the system owner to know their algorithm. While publicly available test data from NIST and elsewhere can inform owners, it will usually be informative to specifically measure accuracy of the operational algorithm on the operational image data, perhaps employing a biometrics testing laboratory to assist.

Since different algorithms perform better or worse in processing images of individuals in various demographics, policy makers, face recognition system developers, and end users should be aware of these differences and use them to make decisions and to improve future performance. We supplement this report with more than 1200 pages of charts contained in seventeen annexes that include exhaustive reporting of results for each algorithm. These are intended to show the breadth of the effects, and to inform the algorithm developers.

There are a variety of techniques that might mitigate performance limitations of face recognition systems performance issues overall and specifically those that relate to demographics. This report includes recommendations for research in developing and evaluating the value, costs, and benefits of potential mitigation techniques - see sections 8 and 9.

Reporting of demographic effects often has been incomplete in academic papers and in media coverage. In particular, accuracy is discussed without stating the quantity of interest be it false negatives, false positives or failure to enroll. As most systems are configured with a fixed threshold, it is necessary to report both false negative and false positive rates for each demographic group at that threshold. This is rarely done - most reports are concerned only with false negatives. We make suggestions for augmenting reporting with respect to demographic difference and effects.

BACKGROUND: ALGORITHMS, ERRORS, IMPACTS

FACE ANALYSIS: CLASSIFICATION, ESTIMATION, RECOGNITION

Before presenting results in the Technical Summary we describe what face recognition is, contrasting it with other applications that analyze faces, and then detail the errors that are possible in face verification and identification and their impacts.

Much of the discussion of face recognition bias in recent years cites two studies [5, 36] showing poor accuracy of face gender classification algorithms on black women. Those studies did not evaluate face recognition algorithms, yet the results have been widely cited to indict their accuracy. Our work was undertaken to quantify analogous effects in face recognition algorithms. We strongly recommend that reporting of bias should include information about the class of algorithm evaluated. We use the term **face analysis** as an umbrella for any algorithm that consumes face images and produces some output. Within that are **estimation** algorithms that output some continuous quantity (e.g., age or degree of fatigue). There are **classification** algorithms that aim to determine some categorical quantity such as the sex of a person or their emotional state. Face classification algorithms are built with inherent knowledge of the classes they aim to produce (e.g., happy, sad). Face **recognition** algorithms, however, have no built-in notion of a particular person. They are not built to identify particular people; instead they include a face detector followed by a feature extraction algorithm that converts one or more images of a person into a vector of values that relate to the identity of the person. The extractor typically consists of a neural network that has been trained on ID-labeled images available to the developer. In operations, they act as generic extractors of identity-related information from photos of persons they have usually never seen before. Recognition proceeds as a differential operator: Algorithms compare two feature vectors and emit a similarity score. This is a vendor-defined numeric value expressing how similar the parent faces are. It is compared to a threshold value to decide whether *two* samples are from, or represent, the same person or not. Thus, recognition is mediated by persistent identity information stored in a feature vector (or "template"). Classification and estimation, on the other hand, are single-shot operations from *one* sample alone, employing machinery that is different from that used for face recognition.

VERIFICATION

Errors: A comparison of images from the same person yields a genuine or "mate" score. A comparison of images from different people yields an imposter or "nonmate" score. Ideally, nonmate scores should be low and mate scores should be high. In practice, some imposter scores are above a numeric threshold giving false positives, and some genuine comparisons yield scores below threshold giving false negatives.

Applications: One-to-one verification is used in applications including logical access to a phone or physical access through a security check point. It also supports non-repudiation e.g. to authorize the dispensing of a prescription drug. Two photos are involved: one in the database that is compared with one taken of the person seeking access to answer the question: "Is this the same person or not?"

Impact of errors: Errors have different implications for the system owner and for the individual whose photograph is being used, depending upon the application. In verification applications, false negatives cause inconvenience for the user. For example, an individual may not be able to get into their phone or they are delayed entering a facility or crossing a border. These errors can usually be remediated with a second attempt. False positives, on the other hand, present a security concern to the system owner, as they allow access to imposters.

IDENTIFICATION

Identification algorithms, referred to commonly as one-to-many or "1-to-N" search algorithms, notionally compare features extracted from a search "probe" image with all feature vectors previously enrolled from "gallery" images. The algorithms return either a fixed number of the most similar candidates, or only those that are above a preset threshold. A candidate is an index and a similarity score. Some algorithms execute an exhaustive search of all N enrollments and a sort operation to yield the most similar. Other algorithms implement "fast-search" techniques [2,19,21,26] that avoid many of the N comparisons and are therefore highly economical [17].

Identification applications: There are two broad uses of identification algorithms. First, they can be used to facilitate positive access like in one-to-one verification but without presentation of an identity claim. For example, a subject is given access to a building solely on the basis of presentation a photograph that matches *any* enrolled identity with a score above threshold. Second, they can be used for so-called negative identification where the system operator claims implicitly that searched individuals are not enrolled - for example, checking databases of gamblers previously banned from a casino.

Impacts: As with verification, the impact of a demographic differential will depend on the application. In one-to-many searches, false positives primarily occur when a search of a subject who is not present in the database yields a candidate identity for human review. This type of "one to many" search is often employed to check for a person who might be applying for a visa or driver's license under a name different than their own. False positives may also occur when a search of someone who is enrolled produces the wrong identity with, or instead of, the correct identity. Identification algorithms produce such outcomes when the search yields a comparison score above a chosen threshold.

In identification applications such as visa or passport fraud detection, or surveillance, a false positive match to another individual could lead to a false accusation, detention or deportation. Higher false negatives would be an advantage to an enrollee in such a system, as their fraud would go undetected, and a disadvantage to the system owner whose security goals will be undermined.

Investigation: This is a special-case application of identification algorithms where the threshold is set to zero so that all searches will produce a fixed number of candidates. In such cases, the false positive identification rate is 100% because any search of someone not in the database will still yield candidates. Algorithms used in this way are part of a hybrid *machine-human system*: The algorithm offers up candidates for human adjudication, for which labor must be available. In such cases, false positive differentials from the algorithm are immaterial - the machine returns say 50 candidates regardless. What matters then is the human response, and the evidence there is for both poor [10, 42] and varied human capability, even without time constraints [34], and sex and race performance differentials, particularly an interaction between the reviewer's demographics with those of the photographs under review [7]. The interaction of machine and human is beyond the scope of this report, as is human efficacy.

TECHNICAL SUMMARY

This section summarizes the results of the study. This is preceded by an introduction to terminology and discussion of a vital aspect in reporting demographic effects, namely that it is necessary to report both false negative and false positive error rates.

ACCURACY DIFFERENTIALS

When similarity scores are computed over a collection of images from demographic A (say elderly Asian men) they may be higher than from demographic B (say young Asian women). We adopt terminology from a Department of Homeland Security Science and Technology Directorate article [20] and define **differential performance** as a "difference in the genuine or imposter [score] distributions". Such differentials are inconsequential unless they prompt a **differential outcome.** An outcome occurs when a score is compared with an operator-defined threshold. A genuine score below threshold yields a false negative outcome, and an imposter score at or above threshold, a false positive outcome. The subject of this report is to quantify differential outcomes between demographics. The term demographic differential is inherited from an ISO technical report [6] now under development.

FIXED THRESHOLD OPERATION

A crucial point in reasoning about differentials is that the vast majority of biometric systems are configured with a fixed threshold against which all comparisons are made (i.e., the threshold is not tailored to cameras, environmental conditions or, particularly, demographics). Most academic studies ignore this point (even in demographics e.g., [13]) by reporting false negative rates at fixed false positive rates rather than at fixed thresholds, thereby hiding excursions in false positive rates and misstating false negative rates. This report includes documentation of demographic differentials about typical operating thresholds.

We report false positive and false negative rates separately because the consequences of each type of error are of importance to different communities. For example, in a one-to-one access control, false negatives inconvenience legitimate users; false positives undermine a system owners security goals. On the other hand, in a one-to-many deportee detection application, a false negative would present a security problem, and a false positive would flag legitimate visitors. The prior probability of imposters in each case is important. For example, in some access control cases, imposters almost never attempt access and the only germane error rate is the false negative rate.

RESULTS OVERVIEW

We found empirical evidence for the existence of demographic differentials in the majority of contemporary face recognition algorithms that we evaluated. The false positive differentials are much larger than those related to false negatives. False positive rates often vary by one or two orders of magnitude (i.e., 10x, 100x). False negative effects vary by factors usually much less than 3. The false positive differentials exist broadly, across many, but not all, algorithms. The false negatives tend to be more algorithm-specific. Research toward mitigation of differentials is discussed in sections 9 and 8.

The accuracy of algorithms used in this report has been documented in recent FRVT evaluation reports [16, 17]. These show a wide range in accuracy across algorithm developers, with the most accurate algorithms producing many fewer errors than lower-performing variants. More accurate algorithms produce fewer errors, and will be expected therefore to have smaller demographic differentials.

FALSE NEGATIVES

With regard to false negative demographic differentials we make the observations below. Note that in real-time cooperative applications, false negatives can often be remedied by making second attempts.

- ▷ False negative error rates vary strongly by algorithm, from below 0.5% to above 10%. For the more accurate algorithms, false negative rates are usually low with average demographic differentials being, necessarily, smaller still. This is an important result: use of inaccurate algorithms will increase the magnitude of false negative differentials. See Figure 22 and Annex 12 .
- ▷ In domestic mugshots, false negatives are higher in Asian and American Indian individuals, with error rates above those in white and black faces. The lowest false negative rates occur in black faces. This result might not be related to race - it could arise due to differences in the time elapsed between photographs because ageing is highly influential on face recognition false negatives. We will report on that analysis going forward. See Figure 17.
- ▷ False negative error rates are often higher in women and in younger individuals, particularly in the mugshot images. There are many exceptions to this, so universal statements pertaining to algorithms false negative rates across sex and age are not supported.
- ▷ When comparing high-quality application photos, error rates are very low and measurement of false negative differentials across demographics is difficult. This implies that better image quality reduces false negative rates and differentials. See Figure 22.
- ▷ When comparing high-quality application images with lower-quality border crossing images, false negative rates are higher than when comparing the application photos. False negative rates are often higher in recognition of women, but the differentials are smaller and not consistent. See Figure 21.
- ▷ In the border crossing images, false negatives are generally higher in individuals born in Africa and the Caribbean, the effect being stronger in older individuals. See Figure 18.

FALSE POSITIVES

Verification Algorithms: With regard to false positive demographic differentials we make the following observations.

- ▷ We found false positives to be between 2 and 5 times higher in women than men, the multiple varying with algorithm, country of origin and age. This increase is present for most algorithms and datasets. See Figure 6.
- ▷ With respect to race, false positive rates are highest in West and East African and East Asian people (but with exceptions noted next). False positive rates are also elevated but slightly less so in South Asian and Central American people. The lowest false positive rates generally occur with East European individuals. See Figure 5.
- ▷ A number of algorithms developed in China give low false positive rates on East Asian faces, and sometimes these are lower than those with Caucasian faces. This observation - that the location of the developer as a proxy for the race demographics of the data they used in training - matters was noted in 2011 [33], and is potentially important to the reduction of demographic differentials due to race and national origin.

- We found elevated false positives in the elderly and in children; the effects were larger in the oldest adults and youngest children, and smallest in middle aged adults. The effects are consistent across country-of-birth, datasets and algorithms but vary in magnitude. See Figure 14 and Figure 15.
- With mugshot images, the highest false positives are in American Indians, with elevated rates in African American and Asian populations; the relative ordering depends on sex and varies with algorithm. See Figure 12 and Figure 13.

Identification Algorithms: The presence of an enrollment database affords one-to-many algorithms a resource for mitigation of demographic effects that purely one-to-one verification systems do not have. We note that demographic differentials present in one-to-one verification algorithms are usually, but not always, present in one-to-many search algorithms. See Section 7.

One important exception is that some developers supplied identification algorithms for which false positive differentials are undetectable. Among those is Idemia, who publicly described how this was achieved [15]. A further algorithm, NEC-3, is on many measures, the most accurate we have evaluated. Other developers producing algorithms with stable false positive rates are Aware, Toshiba, Tevian and Real Networks. These algorithms also give false positive identification rates that are approximately independent of the size of enrollment database. See Figure 27.

PRIOR WORK This report is the first to describe demographic differentials for identification algorithms. There are, however, recent prior tests of verification algorithms whose results comport with ours regarding demographic differentials between races.

- Using four verification algorithms applied to domestic mugshots, the Florida Institute of Technology and its collaborators showed [23] simultaneously elevated false positives and reduced false negatives in African Americans vs. Caucasians.
- Cavazos et al. [8] applied four verification algorithms to GBU challenge images [32] to show order-of-magnitude higher false positives in Asians vs. Caucasians. The paper articulates five lessons related to measurement of demographic effects.
- In addition, a recent Department of Homeland Security (DHS) Science and Technology / SAIC study [20] using a leading commercial algorithm showed that pairing of imposters by age, sex and race gives false positive rates that are two orders of magnitude higher than by pairing individuals randomly.
- On an approximately monthly schedule starting in 2017, NIST has reported [16] on demographic effects in one-to-one verification algorithms submitted to the FRVT process. Those tests employed smaller sets of mugshot and visa photographs than are used here.

WHAT WE DID NOT DO

This report establishes context, gives results and impacts, and discusses additional research that can support mitigation of observed deficiencies. It does not address the following:

- **Training of algorithms:** We did not train algorithms. The prototype algorithms submitted to NIST are fixed and were not refined or adapted. This reflects the usual operational situation in which face recognition systems are not adapted on customers local data. We did not attempt, or invite developers to attempt, mitigation of demographic differentials by retraining the algorithms on image sets maintained at NIST. We simply ran the tests using algorithms as submitted.

- **Analyze cause and effect:** We did not make efforts to explain the technical reasons for the observed results, nor to build an inferential model of them. Specifically, we have not tried to relate recognition errors to skin tone or any other phenotypes evident in faces in our image sets. We think it likely that modeling will need richer sets of covariates than are available. In particular, efforts to estimate skin tone and other phenotypes will involve an algorithm that itself may exhibit demographic differentials.

  We did not yet pursue regression approaches due to the volume of data, the number of algorithms tested, and the need to model each recognition algorithms separately, as they are built and trained independently. Due to their ability to handle imbalanced data, we note, however, the utiltity of mixed effects models [3, 4, 9] previously developed for explaining recognition failure. Such approaches can use subject-specific variables (age, sex, race, etc.) and image-specific variables (contrast, brightness, blur, uniformity, etc.). Models are often useful, even though it is inevitable that germane quantities will be unavailable to the analysis.

- **Consider the effect of cameras:** The possible role of the camera, and the subject-camera interaction, has been detailed recently [9]. This is particularly important when standards-compliant photography is not possible, or not intended, for example, in high throughput access control. Without access to human-camera interaction data, we do not report on quantities like satisfaction, difficulty of use, and failure to enroll. Along these lines, it has been suggested [41] that NISTs tests using standards-compliant images "don't translate to everyday scenarios".

  In fact, we note demographic effects *even* in high-quality images, notably elevated false positives. Additionally, we quantify false negatives on a border crossing dataset which is collected at a different point in the trade space between quality and speed than are our other three mostly high-quality portrait datasets.

  Finally, some governmental organizations dedicated resources to advancing standards so that the "real-world" images in their applications are high-quality portraits. For example, the main criminal justice application is supported by the FBI and others being proactive in the 1990s in establishing portrait capture standards, and then promulgating them.

- **Use wild images:** We did not use image data from the Internet nor from video surveillance. This report does not capture demographic differentials that may occur in such photographs.

RESEARCH RECOMMENDATIONS

We now discuss research germane to the quantification, handling and mitigation of demographic differentials.

Testing: Since 2017 NIST has provided demographic differential data to developers of one-to-one verification algorithms. Our goal has been to encourage developers to remediate the effects. While that may have happened in some cases, a prime incentive for a developer when participating in NIST evaluations is to reduce false negatives rates globally. Going forward, we plan to start reporting accuracy that pushes developers to produce approximately equal false positive rates across all demographics.

Mitigation of false positive differentials: With adequate research and development, the following may prove effective at mitigating demographic differentials with respect to false positives: Threshold elevation, refined training, more diverse training data, discovery of features with greater discriminative power - particularly techniques capable of distinguishing between twins - and use of face and iris as a combined modality. These are discussed in section 9. We also discuss, and discount, the idea of user-specific thresholds.

Mitigation of false negative differentials: False negative error rates, and demographic differentials therein, are reduced in standards-compliant images. This motivates the suggestions of further research into image quality analysis, face-aware cameras and improved standards-compliance discussed in section 8.

Policy research: The degree to which demographic differentials could be tolerated has never been formally specified in any biometric application. Any standard directed toward limiting allowable differentials in the automated processing of digitized biological characteristics might weigh the actual consequences of differentials which are strongly application dependent.

REPORTING OF DEMOGRAPHIC EFFECTS

Reporting of demographic effects has been incomplete, in both academic papers and in media coverage. In particular, accuracy is discussed without specifying, particularly, false positives or false negatives. We therefore suggest that reports covering demographic differentials should describe:

- The purpose of the system - initial enrollment of individuals into a system, identity verification or identification;
- The stage at which the differential occurred - at the camera, during quality assessment, in the detection and feature extraction phase, or during recognition;
- The relevant metric: false positive or false negative occurrences during recognition, failures to enroll, failed detections by the camera, for example;
- Any differentials in duration of processes or difficulty in using the system;
- Any known information on recognition threshold value, whether the threshold is fixed, and what the target false positive rate is;
- Which demographic group has the elevated failure rates - for example by age, sex, race, height, or in some intersection thereof; and
- Consequences of any error, if known, and procedures for error remediation.

ACKNOWLEDGMENTS

The authors are grateful to Yevgeniy Sirotin and John Howard of SAIC at the Maryland Test Facility, Arun Vemury of DHS S&T, Michael King of Florida Institute of Technology, and John Campbell of Bion Biometrics for detailed discussions of their work in this area.

The authors are grateful to staff in the NIST Biometrics Research Laboratory for infrastructure supporting rapid evaluation of algorithms.

DISCLAIMER

Specific hardware and software products identified in this report were used in order to perform the evaluations described in this document. In no case does identification of any commercial product, trade name, or vendor, imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose. Developers participating in FRVT grant NIST permission to publish evaluation results.

ANNEXES

We supplement this report with more than 1200 pages of charts contained in 17 Annexes which include exhaustive reporting of results for each algorithm. These are intended to show the breadth of the effects and to inform the algorithms' developers. We do not take averages over algorithms, for example the average increase of false match rate in women, because averages of samples from different distributions are seldom meaningful (by analogy, taking the average of temperatures in Montreal and Miami). Applications typically employ just one algorithm, so averages and indeed any statements purporting to summarize the entirety of face recognition will not always be correct.

The annexes to this report are listed in Table 1. The first four detail the datasets used in this report. The remaining annexes contain more than 1200 pages of automatically generated graphs, usually one for each algorithm evaluated. These are intended to show the breadth of the effects, and to inform the algorithms' developers.

#	CATEGORY		DATASET	CONTENT
Annex 1	Datasets		Mugshot	Description and examples of images and metadata: Mugshots
Annex 2	Datasets		Application	Description and examples of images and metadata: Application portraits
Annex 3	Datasets		Visa	Description and examples of images and metadata: Visa portraits
Annex 4	Datasets		Border crossing	Description and examples of images and metadata: Border crossing photos
Annex 5	Results	1:1	Application	False match rates for demographically matched impostors
Annex 6	Results	1:1	Mugshot	Cross-race and sex false match rates in United States mugshot images
Annex 7	Results	1:1	Application	Cross-race and sex false match rates in worldwide application images
Annex 8	Results	1:1	Application	False match rates with matched demographics using application images
Annex 9	Results	1:1	Application	Cross-age false match rates with application photos
Annex 10	Results	1:1	Visa	Cross age false match rates with visa photos
Annex 11	Results	1:1	Mugshot	Cross age and country with application photos
Annex 12	Results	1:1	Mugshot	Error tradeoff characteristics with United States mugshots
Annex 13	Results	1:1	Mugshot	False negative rates in United States mugshot images by sex and race
Annex 14	Results	1:1	Mugshot	False negative rates by country for global application and border crossing photos
Annex 15	Results	1:1	Mugshot	Genuine and impostor score distributions for United States mugshots
Annex 16	Results	1:N	Mugshot	Identification error characteristics by race and sex
Annex 17	Results	1:N	Mugshot	Candidate list score magnitudes by sex and race

Table 1: **Annexes and their content.**

TERMS AND DEFINITIONS

The following table defines common terms appearing in this document. A more complete, consistent biometrics vocabulary is available as ISO/IEC 2382 Part 37.

DATA TYPES	Feature vector	A vector of real numbers that encodes the identity of a person
	Sample	One or more images of the face of a person
	Similarity score	Degree of similarity of two faces in two samples, as rendered by a recognition algorithm
	Template	Data produced by face recognition algorithm that includes a feature vector
	Threshold	Any real number, against which similarity scores are compared to produce a verification decision
ALGORITHM COMPONENTS	Face detector	Component that finds faces in an image
	Comparator	Component that compares two templates and produces a similarity score
	Searcher	Component that searches a database of templates to produce a list of candidates
	Template generator	Component of a face recognition algorithm that converts a sample into a template; this component implicitly embeds a face detector
ONE-TO-ONE VERIFICATION	Imposter comparison	Comparison of samples from different persons
	Genuine comparison	Comparison of samples from the same person
	False match	Incorrect association of two samples from different persons, declared because similarity score is at or above a threshold
	False match rate	Proportion of imposter comparisons producing false matches
	False non-match	Failure to associate two samples from one person, declared because similarity score is below a threshold
	False non-match rate	Proportion of genuine comparisons producing false non-matches
	Verification	The process of comparing two samples to determine if they belong to the same person or not
ONE-TO-MANY IDENTIFICATION	Gallery	A set of templates, each tagged with an identity label
	Consolidated gallery	A gallery for which all samples of a person are enrolled under one identifier, whence $N = N_G$
	Unconsolidated gallery	A gallery for which samples of a person are enrolled under different identifiers, when $N < N_G$
	Identity label	Some index or pointer to an identifier for an individual
	Identification	The process of searching a probe into gallery
	Identification decision	The assignment either of an identity label or a declaration that a person is not in the gallery
SYMBOLS	FMR	Verification false match rate (measured over comparison of samples)
	FNMR	Verification false non-match rate (measured over comparison of samples)
	FPIR	Identification false match rate (measured over comparison of samples)
	FNIR	Identification false non-match rate (measured over comparison of samples)
	N	The number of subjects whose faces are enrolled into a gallery
	N_G	The number of samples enrolled into a gallery, $N_G \geq N$.
	N_{NM}	The number of non-mated searches conducted
	N_M	The number of mated searches conducted

Contents

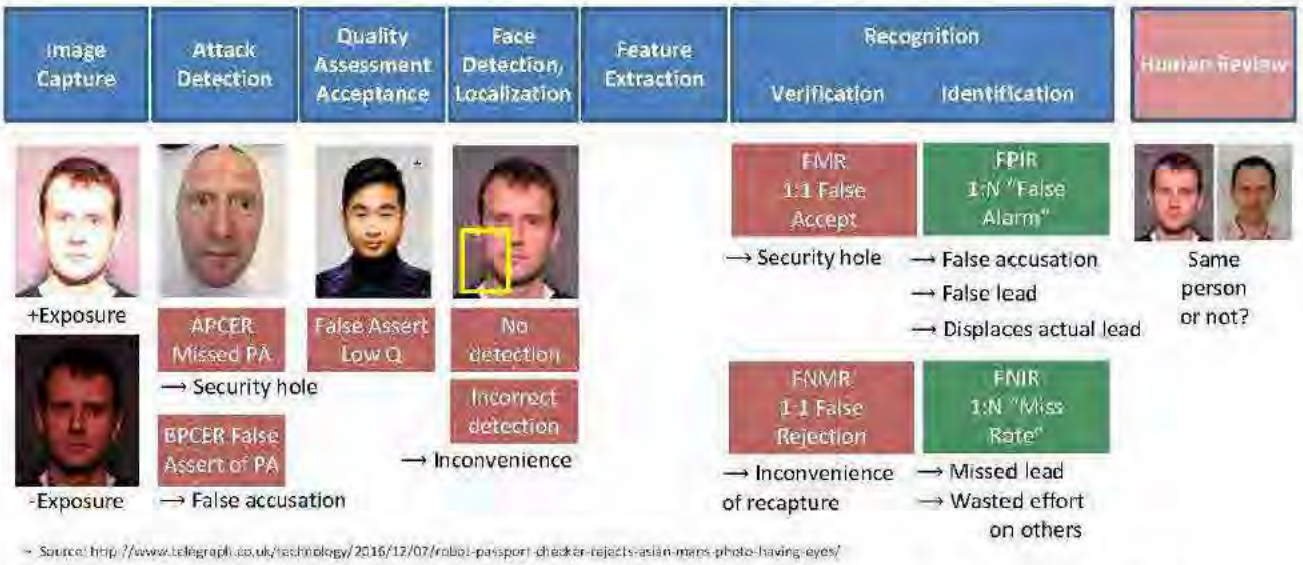


Figure 1: The figure is intended to show possible stages in a face recognition pipeline at which demographic differentials could, in principle, arise. Note that none of these stages necessarily includes algorithms that may be labelled artificial intelligence, though typically the detection and feature extraction modules are AI-based now.

1 Introduction

Over the last two years there has been expanded coverage of face recognition in the popular press. In some part this is due to the expanded capability of the algorithms, a larger number of applications, lowered barriers to algorithm development[1], and, not least, reports that the technology is somehow biased. This latter aspect is based on Georgetown [14] and two reports by MIT [5, 36]. The Georgetown work noted prior studies [22] articulated sources of bias, and described the potential impacts particularly in a policing context, and discussed policy and regulatory implications. The MIT work did not study face recognition, instead it looked at how well publicly accessible cloud-based *estimation* algorithms can determine gender from a single image. The studies have widely cited as evidence that face *recognition* is biased.

This stems from a confusion in terminology: Face classification algorithms, of the kind MIT reported on, accept one face image sample and produce an estimate of age, or sex, or some other property of the subject. Face recognition algorithms, on the other hand, operate as differential operators: They compare identity information in features vectors extract from two face image samples and produce a measure of similarity between the two, which can be used to answer the "question same person or not?". Face algorithms, both one-to-one identity verification and one-to-many search algorithms, are built on this differential comparison. The salient point, in the demographic context, is that one or two people are involved in a comparison and, as we will see, the age,

[1]Gains in face recognition performance stem from well-capitalized AI research in industry and academic leading to the development of convolutional neural networks, and open-source implementations thereof (Caffe, Tensorflow etc.). For face recognition the availability of large numbers of identity-labeled images (from the web, and in the form of web-curated datasets [VGG2, IJB-C]), and the availability of ever more powerful GPUs has supported training those networks.

sex, race and other demographic properties of both will be material to the recognition outcome.

The MIT reports nevertheless serve as a cautionary tale in two respects. First, that demographic group membership can have a sizeable effect on algorithms that process face photographs; second, that algorithm capability varies considerably by developer.

1.1 Potential sources of bias in face recognition systems

Lost in the discussion of bias is specificity on exactly what component of the process is at fault. Accordingly, we introduce Figure 1 to show that a face recognition system is composed of several parts. The figure shows a notional face recognition pipeline consisting of a capture subsystem, primarily a camera, followed by a presentation attack detection (PAD) module intended to detect impersonation attempts, a quality acceptance (QA) step aimed at checking portrait standard compliance, then the recognition components of feature extraction and 1:1 or 1:N comparison, the output of which may prompt human involvement. The order of the components may be different in some systems, for example the QA component may be coupled to the capture process and would precede PAD. Some components may not exist in some systems, particularly the QA and PAD functions may not be necessary.

The Figure shows performance metrics, any of which could notionally have a demographic differential. Errors at one stage will generally have downstream consequences. In a system where subjects make cooperative presentation to the camera, a person could be rejected in the early stages before recognition itself. For example, a camera equipped with optics that have too narrow a field of view could produce an image of a tall individual in which in which the top part of the head was cropped. This could cause rejection at almost any stage and a system owner would need to determine the origin of errors.

1.2 The role of image quality

Recent research [9] has shown that cameras can have an effect on a generic downstream recognition engine. A poor image can undermine detection or recognition, and it is possible that certain demographics yield photographs ill-suited to face recognition e.g. young children [28], or very tall individuals. As pointed out above there is potential for demographic differentials to appear at the capture stage, that is when only a single image is being collected before any comparison with other images. Demographic differentials that occur during collection could arise from (at least) inadequacies of the camera, from the environment or "stage", and from client-side detection or quality assessment algorithms. Note that manifestly poor (and unrecognizable) images can be collected from mis-configured cameras, without any algorithmic or AI culpability. Indeed, after publication of the MIT studies [5, 36] on bias in gender-estimation algorithms, suspicion fell upon the presence of poor

photographs, due to under-exposure of dark-skinned individuals in that dataset. An IBM gender estimation algorithm had been faulted in the MIT study; in response, and previously, IBM has been active in addressing AI bias. Relevant here is that it produced a better algorithm[2], and examined whether skin tone itself drove gender classification accuracy [30, 31] - in short, "skin type *by itself* has a minimal effect on the classification decision".

False negatives occur in biometric systems when samples from one individual yield a comparison score below a threshold. This will occur when the features extracted from two input photographs are insufficiently similar. Recall that face recognition is implemented as a differential operator: two samples are analyzed and compared. So a false negative occurs when two from the same face appear different to the algorithm.

It is very common to attribute false negatives to factors such as pose, illumination and expression so much so that dedicated databases have been built up to support development of algorithms with immunity to such[3]. Invariance to such "nuisance" factors has been the focus of the bulk of face recognition research for more two decades. Indeed over the last five years there have been great advances in this respect due to the adoption of deep convolutional neural networks which demonstrate remarkable tolerance to very sub-standard photographs i.e. those that deviate from formal portrait standards most prominently ISO/IEC 39794-5 and its law-enforcement equivalent ANSI/NIST ITL 1-2017.

However, here we need to distinguish between factors that are expected to affect one photo in a mated pair - due to poor photography (e.g. mis-focus), poor illumination (e.g. too dark), and poor presentation (e.g. head down) - and those that would affect both photographs over time, potentially including properties related to demographics.

1.3 Photographic Standards

In the late 1990s the FBI asked NIST to establish photographic best-practices for mugshot collection[4]. This was done to guide primarily state and local police departments in the capture of photographs that would support forensic (i.e. human) review. It occurred more than a decade before the FBI deployed automated face recognition. That standardization work was conducted in anticipation of digital cameras[5] being available to replace film cameras that had been used for almost a century. The standardization work included consideration of cameras, lights and geometry[6]. There was explicit consideration of the need to capture images of both dark and light skinned individuals, it being understood that it is relatively easy to produce photographs for which

[2] See Mitigating Bias in AI Models.
[3] The famous PIE databases, for example.
[4] Early documents, such as Best Practice Recommendation for the Capture of Mugshots, 1999, seeded later formal standardization of ISO/IEC 19794-5.
[5] See NIST Interagency Report 6322, 1999.
[6] See this overview.

large areas of dark or bright pixels can render detection of anatomical features impossible.

Face recognition proceeds as a differential operation on features extracted from two photographs. Accuracy can be undermined by poor photography/illumination/presentation and by differences in those i.e. any change in the digital facial appearance. Of course an egregiously underexposed photograph will have insufficient information content, but two photographs taken with even moderately poor exposure can match, and leading contemporary algorithms are highly tolerant of quality degradations.

1.4 Age and ageing

Ageing will change appearance over decades and will ultimately undermine automated face recognition[7]. In the current study, we don't consider ageing to be a demographic factor because it is a slow, more-or-less graceful, process that happens to all of us. However, there is at least one demographic group that ages more quickly than others - children - who are disadvantaged in many automated border control systems either by being excluded by policy, or by encountering higher false negatives. Age itself is a demographic factor as accuracy in the elderly and the young differ for face recognition (usually) and also for fingerprint authentication. This applies even without significant time lapse between two photographs.

Clearly injury or disease can change appearance on very short timescales, so such factors should be excluded, when possible, from studies dedicated to detection of broad demographic effects. Development of equipment and algorithms, and studies thereof, that are dedicated to the inclusive use of biometrics are valuable of course - for example recognition of photosensitive subjects wearing sunglasses, or finger amputees presenting fingerprints.

[7]See recent results for verification algorithms in the FRVT reports, and for identification algorithms in NIST Interagency Report 8271 [17]. For a formal longitudinal analysis of ageing, using mixed-effects models, see Best-Rowden [3].

#	SOURCE	IMAGE	NUMBER OF SUBJECTS	NUMBER OF IMAGES	DISCUSSION
1	Cavazos et al. [8] at UT Dallas	Notre Dame GBU [32] portraits	389	<1085	The study showed order-of-magnitude elevations in false positive rates in university volunteer Asian vs. Caucasian faces. The study reported FMR(T). As the study showed neither FNMR(T) nor linked error tradeoff characteristics the false negative differential is not apparent. It discusses the effect of "yoking" i.e the pairing of imposters by sex and race. It deprecates area-under-the-curve (AUC). The study used two related algorithms from the University of Maryland, one open-source algorithm [38], and one older inaccurate pre-DCNN algorithm.
2	Krishnapriya et al. [23] at Florida Inst. Tech	Operational mugshots: Morph db [37]	10350 African Am. + 2769 Caucasians	42620 African Am. + 10611 Caucasians	The study reported: order-of-magnitude elevated false positives in African Americans vs. Caucasians; lower false negative rates in African Americans; and reduced differentials in higher quality images [23, 24]. That study used three open-source algorithms, and one commercial algorithm. Two of the open-source algorithms are quite inaccurate and not representative of commercial deployment. Importantly, the study also noted the inadequacies of error tradeoff characteristics for documenting fixed-threshold demographic differentials.
3	Howard et al. [20] at SAIC/MdTF with DHS S&T	Lab collected, adult volunteers [9]	363	-	The study establish useful definitions for "differential performance" and "differential outcome" and for broad and narrow heterogeneity of imposter distributions. It showed order-of-magnitude variation in false positive rates with age, sex and race, establishing an information gain approach to formally ordering their effect. The study employed images from 11 capture devices, and applied one leading commercial verification algorithm.

Table 2: Prior studies.

2 Prior work

All prior work relates to one-to-one verification algorithms. This report, in contrast, includes results for many recent, mostly commercial, algorithms implementing both verification and identification.

Except as detailed below, this report is the first to properly report and distinguish between false positive and false negative effects, something that is often missing in other reports.

The broad effects given in this report concerning age and sex have been known as far back as 2003 [35]. Since 2017, our ongoing FRVT report [16] has reported large false positive differential across sex, age and race.

Tables 2 and 3 summarize recent work in demographic effects in automated face recognition.

#	SOURCE	IMAGE	NUMBER OF SUBJECTS	NUMBER OF IMAGES	DISCUSSION
4	Cook et al. [9] at SAIC/MdTF with DHS S&T	Lab collected, adult volunteers	525		The study deployed mixed-effects regression models to examine dependence of genuine similarity scores on sex, age, height, eyewear, skin reflectance and on capture device. The report displayed markedly different images of the same people from different capture devices, showing potential for the camera to induce demographic differential performance. The study found lower similarity scores in those identifying as Black or African American, comporting with [22] but contrary to the best ageing study [3]. The study also showed that comparison of samples collected on the same day have different demographic differentials than those collected up to four years apart, in particular that women give lower genuine scores than men with time separation. Same-day biometrics are useful for short-term recognition applications like transit through an airport.
5	El Khiyari et al. [13]	Operational mugshots: Morph db [37]	724 adult, balanced on race + sex	2896 = 1448 each African Am. + Cau-casians, balanced on sex	The paper used a subset of the MORPH database with two algorithms([38], modified and one COTS) to show better verification error rates in the men, the elderly, and in whites. The study should be discounted for two reasons: First the algorithms give high error rates at very modest false match rates: the best FNMR = 0.06 at FMR = 0.01. Second the paper reports FNMR at fixed FMR, not at fixed thresholds thereby burying FMR differentials. Moreover, the paper does not disclose how imposters were paired e.g. randomly or, say, with same age, race, and sex.

Table 3: Prior studies (continued).

3 Performance metrics

Both verification and identification systems generally commit two kinds of errors, the so-called Type I error where an individual is incorrectly associated with another, and Type II where the individual is incorrectly not associated with themselves.

The ISO/IEC 19795-1 performance testing and reporting standard requires different metrics to be reported for identification and verification implementations. Accordingly the following subsections define the formal metrics used throughout this document.

3.1 Verification metrics

Verification accuracy is estimated by forming two sets of scores: Genuine scores are produced from mated pairs; imposter scores are produced from non-mated pairs. These comparisons should be done in random order so that the algorithm under test cannot infer that a comparison is mated or not.

From a vector of N genuine scores, u, the false non-match rate (FNMR) is computed as the proportion below some threshold, T:

$$\text{FNMR}(T) = 1 - \frac{1}{N}\sum_{i=1}^{N} H(u_i - T) \tag{1}$$

where $H(x)$ is the unit step function, and $H(0)$ taken to be 1.

Similarly, given a vector of M imposter scores, v, the false match rate (FMR) is computed as the proportion above T:

$$\text{FMR}(T) = \frac{1}{M}\sum_{i=1}^{M} H(v_i - T) \tag{2}$$

The threshold, T, can take on any value. We typically generate a set of thresholds from quantiles of the observed imposter scores, v, as follows. Given some interesting false match rate range, $[\text{FMR}_L, \text{FMR}_U]$, we form a vector of K thresholds corresponding to FMR measurements evenly spaced on a logarithmic scale. This supports plotting of FMR on a logarithmic axis. This is done because typical operations target false match rates spanning several decades 10^{-6} to as high as 10^{-2}.

$$T_k = Q_v(1 - \text{FMR}_k) \tag{3}$$

where Q_v is the quantile function, and FMR_k comes from

$$\log_{10}\text{FMR}_k = \log_{10}\text{FMR}_L + \frac{k}{K}\left[\log_{10}\text{FMR}_U - \log_{10}\text{FMR}_L\right] \tag{4}$$

Error tradeoff characteristics are plots of FNMR(T) vs. FMR(T). These are plotted with $\text{FMR}_U \rightarrow 1$ and FMR_L as low as is sustained by the number of imposter comparisons, M. This should be somewhat higher than the "rule of three" limit $3/N$ because samples are generally not independent due to the use of the same image in multiple comparisons.

3.2 Identification metrics

Identification accuracy is estimated from two sets of candidate lists: First, a set of candidate lists obtained from mated-searches; second, a set from non-mated searches. These searches should not be conducted by randomly ordering mated and non-mated searches so that the algorithm under test cannot infer that a search has a mate or not. Tests of open-set biometric identification algorithms must quantify frequency of two error conditions:

- **False positives**: Type I errors occur when search data from a person who has never been seen before is incorrectly associated with one or more enrollees' data.
- **Misses**: Type II errors arise when a search of an enrolled person's biometric does not return the correct identity.

Many practitioners prefer to talk about "hit rates" instead of "miss rates" - the first is simply one minus the other as detailed below. Sections 3.2.1 and 3.2.2 define metrics for the Type I and Type II performance variables.

Additionally, because recognition algorithms sometimes fail to produce a template from an image, or fail to execute a one-to-many search, the occurrence of such events must be recorded. Further because algorithms might elect to not produce a template from, for example, a poor quality image, these failure rates must be combined with the recognition error rates to support algorithm comparison. This is addressed in section 3.4.

3.2.1 Quantifying false positives

It is typical for a search to be conducted into an enrolled population of N identities, and for the algorithm to be configured to return the closest L candidate identities. These candidates are ranked by their score, in descending order, with all scores required to be greater than or equal to zero. A human analyst might examine either all L candidates, or just the top $R \leq L$ identities, or only those with score greater than threshold, T.

From the candidate lists, we compute **false positive identification rate** as the proportion of non-mate searches that erroneously return candidates:

$$\text{FPIR}(N, T) = \frac{\text{Num. non-mate searches with one or more candidates returned with score at or above threshold}}{\text{Num. non-mate searches attempted.}} \tag{5}$$

Under this definition, FPIR can be computed from the highest non-mate candidate produced in a search - it is not necessary to consider candidates at rank 2 and above. An alternative quantity, selectivity, accounts for multiple candidates above threshold - see [17].

3.2.2 Quantifying hits and misses

If L candidates are returned in a search, a shorter candidate list can be prepared by taking the top $R \leq L$ candidates for which the score is above some threshold, $T \geq 0$. This reduction of the candidate list is done because thresholds may be applied, and only short lists might be reviewed (according to policy or labor availability, for example). It is useful then to state accuracy in terms of R and T, so we define a "miss rate" with the general name **false negative identification rate** (FNIR), as follows:

$$\text{FNIR}(N, R, T) = \frac{\text{Num. mate searches with enrolled mate found outside top R ranks or score below threshold}}{\text{Num. mate searches attempted.}} \tag{6}$$

This formulation is simple for evaluation in that it does not distinguish between causes of misses. Thus a mate that is not reported on a candidate list is treated the same as a miss arising from face finding failure, algorithm intolerance of poor quality, or software crashes. Thus if the algorithm fails to produce a candidate list, either because the search failed, or because a search template was not made, the result is regarded as a miss, adding to FNIR.

Hit rates, and true positive identification rates: While FNIR states the "miss rate" as how often the correct candidate is either not above threshold or not at good rank, many communities prefer to talk of "hit rates". This is simply the **true positive identification rate**(TPIR) which is the complement of FNIR giving a positive statement of how often mated searches are successful:

$$\text{TPIR}(N, R, T) = 1 - \text{FNIR}(N, R, T) \tag{7}$$

This report does not report true positive "hit" rates, preferring false negative miss rates for two reasons. First, costs rise linearly with error rates. For example, if we double FNIR in an access control system, then we double user inconvenience and delay. If we express that as decrease of TPIR from, say 98.5% to 97%, then we mentally have to invert the scale to see a doubling in costs. More subtly, readers don't perceive differences in numbers near 100% well, becoming inured to the "high nineties" effect where numbers close to 100 are perceived indifferently.

Reliability is a corresponding term, typically being identical to TPIR, and often cited in automated (fingerprint) identification system (AFIS) evaluations.

An important special case is the **cumulative match characteristic**(CMC) which summarizes accuracy of mated-searches only. It ignores similarity scores by relaxing the threshold requirement, and just reports the fraction of mated searches returning the mate at rank R or better.

$$\mathrm{CMC}(N, R) = 1 - \mathrm{FNIR}(N, R, 0) \tag{8}$$

We primarily cite the complement of this quantity, $\mathrm{FNIR}(N, R, 0)$, the fraction of mates *not* in the top R ranks.

The **rank one hit rate** is the fraction of mated searches yielding the correct candidate at best rank, i.e. CMC(N, 1). While this quantity is the most common summary indicator of an algorithm's efficacy, it is not dependent on similarity scores, so it does not distinguish between strong (high scoring) and weak hits. It also ignores that an adjudicating reviewer is often willing to look at many candidates.

3.3 DET interpretation

In biometrics, a false negative occurs when an algorithm fails to match two samples of one person – a Type II error. Correspondingly, a false positive occurs when samples from two persons are improperly associated – a Type I error.

Matches are declared by a biometric system when the native comparison score from the recognition algorithm meets some threshold. Comparison scores can be either similarity scores, in which case higher values indicate that the samples are more likely to come from the same person, or dissimilarity scores, in which case higher values indicate different people. Similarity scores are traditionally computed by fingerprint and face recognition algorithms, while dissimilarities are used in iris recognition. In some cases, the dissimilarity score is a distance possessing metric properties. In any case, scores can be either mate scores, coming from a comparison of one person's samples, or nonmate scores, coming from comparison of different persons samples.

The words "genuine" or "authentic" are synonyms for mate, and the word "imposters"" is used as a synonym for nonmate. The words "mate" and "nonmate" are traditionally used in identification applications (such as law enforcement search, or background checks) while genuine and imposter are used in verification applications (such as access control).

An error tradeoff characteristic represents the tradeoff between Type II and Type I classification errors. For identification this plots false negative vs. false positive identification rates i.e. FNIR vs. FPIR parametrically with T. Such plots are often called detection error tradeoff (DET) characteristics or receiver operating characteristic (ROC). These serve the same function – to show error tradeoff – but differ, for example, in plotting the complement of an error rate (e.g. TPIR = 1 – FNIR) and in transforming the axes, most commonly using logarithms, to show multiple decades of FPIR.

3.4 Failure to extract features

During enrollment some algorithms fail to convert a face image to a template. The proportion of failures is the failure-to-enroll rate, denoted by FTE. Similarly, some search images are not converted to templates. The corresponding proportion is termed failure-to-extract, denoted by FTX. We do not report FTX because we assume that the same underlying algorithm is used for template generation for enrollment and search.

In verification, we do not need to explicitly include failure to extract rates into the FNMR and FMR accuracy statements, because we regard any comparison that involves an image for which a failure-to-extract occurred as producing a zero similarity score. This increases FNMR and decreases FMR. Gaming opportunities that theoretically arise from this treatment of FMR are generally not of concern because the algorithm under test does not know whether any given image will be used in genuine comparisons, imposter comparisons or both.

For identification, we similarly incorporate failure-to-extract events into FNIR and FPIR measurements as follows.

- **Enrollment templates**: Any failed enrollment is regarded as producing a zero length template. Algorithms are required by the API [18] to transparently process zero length templates. The effect of template generation failure on search accuracy depends on whether subsequent searches are mated, or non-mated: Mated searches will fail giving elevated FNIR; non-mated searches will not produce false positives so, to first order, FPIR will be reduced by a factor of $1-\text{FTE}$.
- **Search templates and 1:N search**: In cases where the algorithm fails to produce a search template from input imagery, the result is taken to be a candidate list whose entries have no hypothesized identities and zero score. The effect of template generation failure on search accuracy depends on whether searches are mated, or non-mated: Mated searches will fail giving elevated FNIR; Non-mated searches will not produce false positives, so FPIR will be reduced.

This approach is the correct treatment for positive-identification applications such as access control where cooperative users are enrolled and make attempts at recognition. This approach is not appropriate to negative identification applications, such as visa fraud detection, in which hostile individuals may attempt to evade detection by submitting poor quality samples. In those cases, template generation failures should be investigated as though a false alarm had occurred.

	Developer	Verification algorithms	Identification algorithms
1	3Divi	3divi-003 3divi-004	3divi-0 3divi-3
2	Adera Global PTE Ltd	adera-001	
3	Alchera Inc	alchera-000 alchera-001	alchera-0
4	Alivia / Innovation Sys	isystems-001 isystems-002	isystems-0 isystems-3
5	AllGoVision	allgovision-000	allgovision-000
6	AlphaSSTG	alphaface-001	
7	Amplified Group	amplifiedgroup-001	
8	Anke Investments	anke-004	anke-0 anke-002
9	AnyVision	anyvision-002 anyvision-004	
10	Aware	aware-003 aware-004	aware-0 aware-3
11	Awidit Systems	awiros-001	
12	Ayonix	ayonix-000	ayonix-0
13	Beijing Vion Technology Inc	vion-000	
14	Bitmain	bm-001	
15	CSA IntelliCloud Technology	intellicloudai-001	
16	CTBC Bank Co Ltd	ctbcbank-000	
17	Camvi Technologies	camvi-002 camvi-004	camvi-1 camvi-3 camvi-4
18	China Electronics Import-Export Corp	ceiec-001 ceiec-002	
19	China University of Petroleum	upc-001	
20	Chunghwa Telecom Co. Ltd	chtface-001	
21	Cognitec Systems GmbH	cognitec-000 cognitec-001	cognitec-0 cognitec-2
22	Cyberextruder	cyberextruder-001 cyberextruder-002	
23	Cyberlink Corp	cyberlink-002 cyberlink-003	
24	DSK	dsk-000	
25	Dahua Technology Co Ltd	dahua-002 dahua-003	dahua-0 dahua-1 dahua-002
26	Deepglint	deepglint-001	deepglint-001
27	Dermalog	dermalog-005 dermalog-006	dermalog-0 dermalog-5 dermalog-6
28	DiDi ChuXing Technology Co	didiglobalface-001	
29	Digital Barriers	digitalbarriers-002	
30	Eyedea Recognition		eyedea-0 eyedea-3
31	FaceSoft Ltd	facesoft-000	
32	FarBar Inc	f8-001	f8-001
33	Gemalto Cogent	cogent-003 cogent-004	
34	Glory Ltd	glory-001	glory-0
35	Gorilla Technology	gorilla-003	gorilla-0
36	Guangzhou Pixel Solutions Co Ltd	pixelall-002	pixelall-002
37	Hengrui AI Technology Ltd	hr-001 hr-002	
38	Hikvision Research Institute	hik-001	hik-0 hik-5
39	ID3 Technology	id3-003 id3-004	
40	ITMO University	itmo-005 itmo-006	
41	Idemia	idemia-004 idemia-005	idemia-0 idemia-4 idemia-5
42	Imagus Technology Pty Ltd	imagus-000	imagus-0
43	Imperial College London	imperial-000 imperial-002	imperial-000
44	Incode Technologies Inc	incode-004	incode-0 incode-004

Table 4: Algorithms evaluated in this report.

	Developer	Verification algorithms	Identification algorithms
45	Innovatrics	innovatrics-004 innovatrics-006	innovatrics-0
46	Institute of Information Technologies	iit-001	
47	Intel Research Group	intelresearch-000	
48	Intellivision	intellivision-001 intellivision-002	
49	Is It You	isityou-000	
50	Kakao Corp	kakao-001 kakao-002	
51	Kedacom International Pte	kedacom-000	kedacom-001
52	Kneron Inc	kneron-003	
53	Lomonosov Moscow State University	intsysmsu-000	intsysmsu-000
54	Lookman Electroplast Industries	lookman-002 lookman-004	
55	Megvii/Face++	megvii-001 megvii-002	megvii-0 megvii-1
56	MicroFocus	microfocus-002 microfocus-001	microfocus-0
57	Microsoft		microsoft-0 microsoft-5
58	Momentum Digital Co Ltd	sertis-000	
59	Moontime Smart Technology	mt-000	
60	N-Tech Lab	ntechlab-006 ntechlab-007	ntechlab-0 ntechlab-6 ntechlab-007
61	NEC		nec-2 nec-3
62	Neurotechnology	neurotechnology-005 neurotechnology-006	neurotechnology-0 neurotechnology-5 neurotechnology-007
63	Nodeflux	nodeflux-001 nodeflux-002	
64	NotionTag Technologies Private Limited	notiontag-000	
65	Panasonic R+D Center Singapore	psl-002 psl-003	
66	Paravision (EverAI)	everai-paravision-003 paravision-004	everai-0 everai-3 everai-paravision-004
67	Rank One Computing	rankone-007	rankone-0 rankone-5 rankone-006 rankone-007
68	Realnetworks Inc	realnetworks-002 realnetworks-003	realnetworks-0 realnetworks-2 realnetworks-003
69	Remark Holdings	remarkai-001	remarkai-0 remarkai-000
70	Rokid Corporation Ltd	rokid-000	
71	Saffe Ltd	saffe-001 saffe-002	
72	Sensetime Group Ltd	sensetime-002	sensetime-0 sensetime-1 sensetime-002
73	Shaman Software	shaman-000 shaman-001	shaman-0
74	Shanghai Jiao Tong University	sjtu-001	
75	Shanghai Ulucu Electronics Technology Co. Ltd	uluface-002	
76	Shanghai Universiy - Shanghai Film Academy	shu-001	
77	Shanghai Yitu Technology	yitu-003	yitu-0 yitu-4 yitu-5
78	Shenzhen EI Networks Limited	einetworks-000	
79	Shenzhen Inst Adv Integrated Tech CAS	siat-004 siat-002	siat-0
80	Shenzhen Intellifusion Technologies Co Ltd	intellifusion-001	
81	Smilart	smilart-002 smilart-003	smilart-0
82	Star Hybrid Limited	starhybrid-001	
83	Synesis	synesis-005	synesis-0
84	Tech5 SA	tech5-002 tech5-003	tech5-001
85	Tencent Deepsea Lab	deepsea-001	deepsea-001
86	Tevian	tevian-004 tevian-005	tevian-0 tevian-4
87	Thales		cogent-0 cogent-3
88	TigerIT Americas LLC	tiger-002 tiger-003	tiger-0

Table 5: Algorithms evaluated in this report.

	Developer	Verification algorithms	Identification algorithms
89	Tong'Yi Transportation Technology	tongyi-005	
90	Toshiba	toshiba-002 toshiba-003	toshiba-0 toshiba-1
91	Trueface.ai	trueface-000	
92	ULSee Inc	ulsee-001	
93	Veridas Digital Authentication Solutions S.L.	veridas-002	
94	Via Technologies Inc.	via-000	
95	Videonetics Technology Pvt Ltd	videonetics-001	
96	Vigilant Solutions	vigilantsolutions-006 vigilantsolutions-007	vigilantsolutions-0
97	Visidon	vd-001	vd-0
98	Vision-Box	visionbox-000 visionbox-001	
99	VisionLabs	visionlabs-006 visionlabs-007	visionlabs-7 visionlabs-008
100	Vocord	vocord-006 vocord-007	vocord-0 vocord-3
101	Winsense Co Ltd	winsense-000	
102	X-Laboratory	x-laboratory-000	
103	Xiamen Meiya Pico Information Co. Ltd	meiya-001	
104	Zhuhai Yisheng Electronics Technology	yisheng-004	yisheng-0
105	iQIYI Inc	iqface-000	
106	iSAP Solution Corporation	isap-001	

Table 6: Algorithms evaluated in this report.

4 False positive differentials in verification

False positives occur in biometric systems when samples from two individuals yield a comparison score at or above a set threshold. Most systems are configured with a threshold that is fixed for all users. False positives present a security hazard to one-to-one verification applications. They have similarly serious consequences in one-to-many identification applications. For example, in applications where subjects apply for some benefit more than once under different biographic identities e.g. visa-shopping, driving license issuance, benefits fraud, an otherwise undetected false positive might lead to various downstream consequences such a financial loss. In a surveillance application a false positive may lead to a false accusation.

This section gives empirical quantification of the variation in verification false match rates across demographics. We present results for one-to-many identification later in section 7.

We conduct several experiments with images drawn from both domestic United States and worldwide populations.

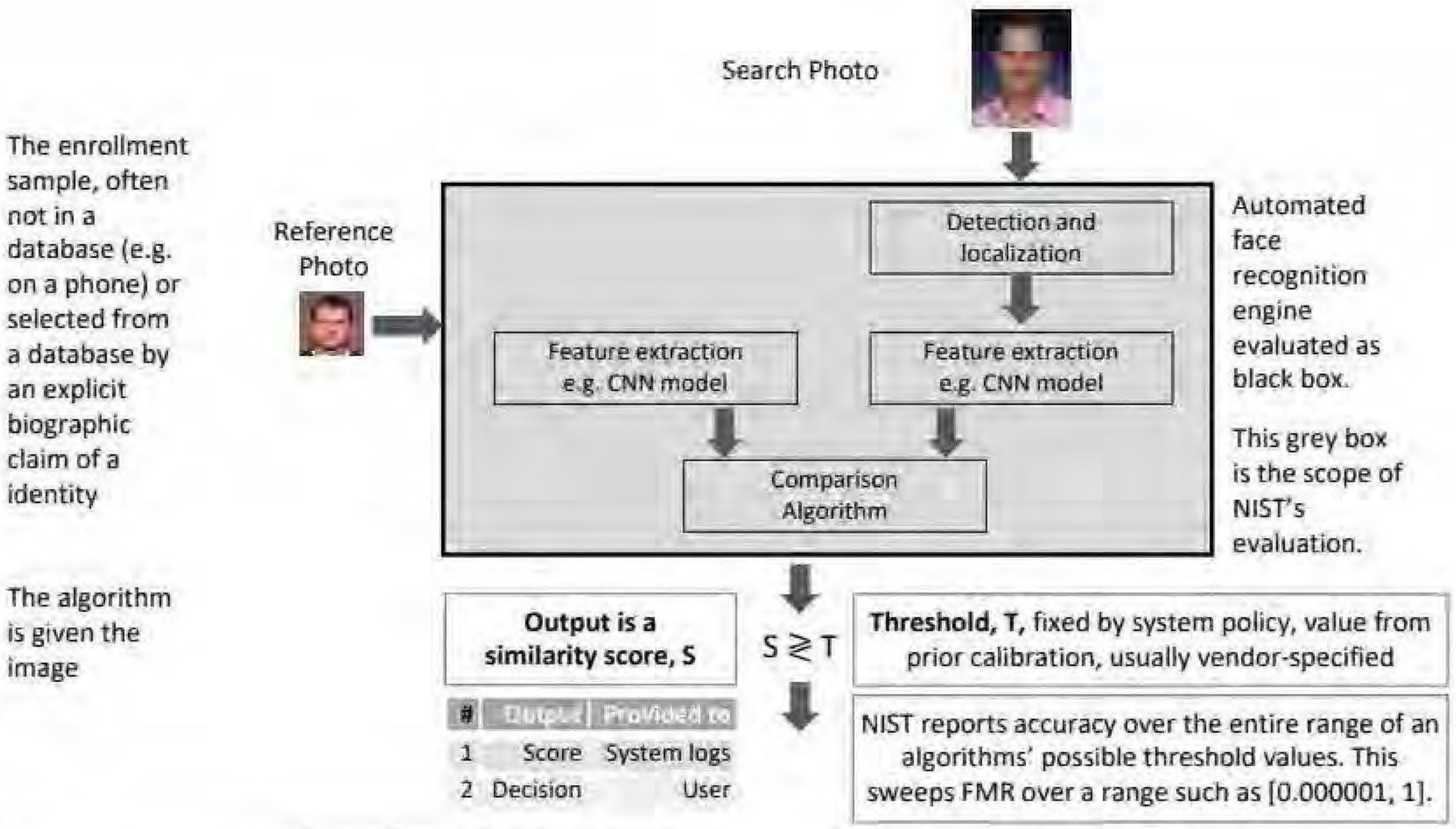
1. One-to-one application photo cross comparison, by age, sex, country-of-birth.
2. One-to-one mugshot cross comparison by age, sex, and race.
3. One-to-one visa photo cross comparison by age.

4.1 Metrics

The metrics appropriate to verification have been detailed in section 3.1. These are related to particular applications in Figure 2. The discussion in subsequent sections centers on false match rates at particular thresholds, i.e. FMR(T).

4.2 False match rates under demographic pairing

It is necessary in many biometric tests to estimate false match rates. This is done by executing imposter comparisons, and measuring false positive outcomes at some threshold(s). Historically biometric evaluations generated imposter comparisons by randomly pairing individuals, or by exhaustively comparing all individuals. As we will show in this section, this practice is inappropriate for evaluation of face recognition algorithms as it underestimates false match rates that would occur in practice. The random pairing of imposters is sometimes referred to as zero-effort pairing, mean that no effort is expended by an imposter to look like the target of the recognition attempt.



	Access Control	Non-repudiation
Role	Afford access of a person to a physical or logical resource.	Record the presence of a specific individual
Example	Door unlock. Phone unlock.	Refutation of a claim by a pharmacist that they did not dispense a particular drug, or an employer that an employee did not arrive for work.
Claim of identity	Explicit claim with an identity token such as a phone, passport or ID card.	Claim with a prior login to a system.
Threshold	High, to limit false positives	Moderate, to prevent confederates using system
Result	Acceptance decision Y/N.	Logged verification decision
Human role	Adjudicate failed rejections, to determine a false rejection, or detect an actual impostor attempt	Retrieve records to resolve a dispute
Intended human involvement frequency	Rare – approx. the false rejection rate identification rate plus prior probability of an actual mate	Rare – approx. the fraud rate multiplied by the false positive rate
Performance metric of interest	**FNMR at low FMR. See sec. 3.1, 3.2 and Tables 10, 19**	**FNMR at moderate FMR.**

Figure 2: Verification applications and relevant metrics.

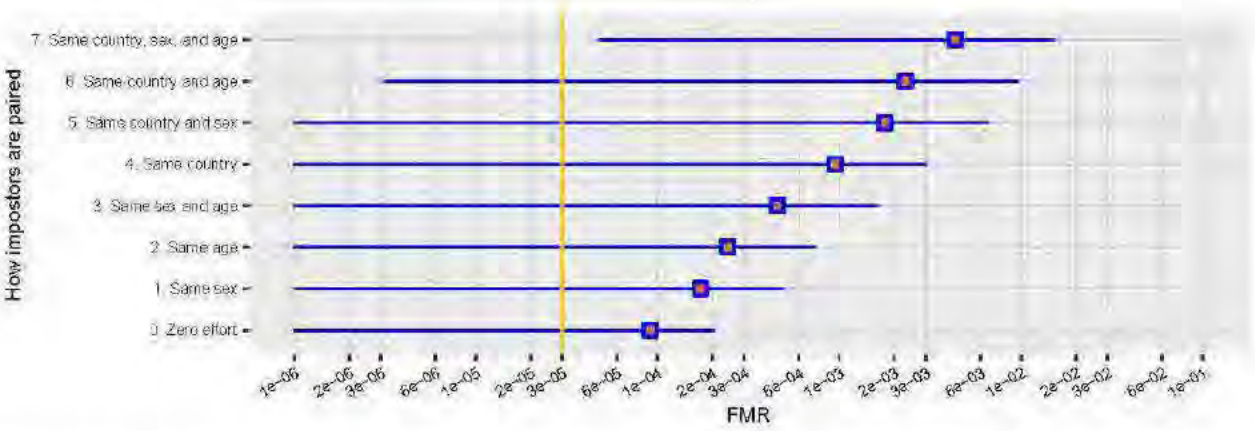


Figure 3: For application photos, the figure shows growth in one-to-one verification false match rates as the imposter demographic pairings are made more similar. At each level the point shows the mean FMR over all countries, age groups, and sexes. For example, in the second row "6. Same country and age" the mean is taken over 24 within-country times 5 within age-group times 4 within and cross sex FMR estimates, i.e. 480 FMR values. The blue line spans the 5-th to 95-th percentiles of the FMR estimates. The vertical line shows a nominal FMR value of 0.00003 obtained by setting the threshold on randomly associated i.e. zero-effort pairs of mugshot images.

Method: We used each verification algorithm to compare 442 019 application images with a disjoint set of 441 517 other application images. The two sets are subject-disjoint. The subjects were born in 24 countries. This produced 195 billion imposter scores. The images are described in Annex 2 .

The red point in the plot shows the mean of false match rates over particular sets of demographic groups.

- **Row 7**: The uppermost point corresponds to the mean over 240 FMR estimates, namely those comparing each of two sexes with each other, in each of five age-groups, and within each of 24 countries (2 x 5 x 24 = 240).
- **Row 6**: As row 7, but the average is over 480 FMR estimates that now includes different sex FMR estimates also.
- **Row 5**: As row 7, but now the average is over 1200 FMR estimates that additionally includes all cross-age group imposter scores.
- **Row 4**: As row 7, but now the average is over 2400 FMR estimates that additionally includes all cross-age and cross-sex imposter scores.
- **Row 3**: The average is over 5760 FMR estimates that includes 24^2 cross-country comparisons within each sex and age group.
- **Row 2**: The average is over 11520 FMR estimates now including different sex FMR estimates also.

- **Row 1**: The average is over 28880 FMR estimates now including five different within-age FMR estimates also.
- **Row 0**: The average is over 57600 FMR estimates reflecting within- and between-group estimates for 24 countries, 5 age groups and 2 sexes ($24^2.5^2.2^2$).

The ordering of these rows is hand-crafted. Evaluators at DHS' Maryland Test Facility developed [20] a formal approach to showing the most influential pairing factor by quantifying information gained about FMR by having knowledge of the demographic factors, age, sex and race.

The figure shows how false match rates increase when imposters are drawn from increasingly similar demographics. This shows that fully zero-effort imposter pairings understate false match rates relative to the situation of a slightly more active imposters who would chose to present (stolen) credentials from subjects of the same sex, age and ethnicity. The practice of using zero-effort imposter pairings in tests, we think, stems from tests of fingerprint algorithm that use where friction ridge structure, particularly minutiae point arrangements, that are thought to be a developmental trait without clear genetic influence[8]

Note that our analysis has not so far documented whether particular demographic groups give higher false match rates. To address this question we introduce Figure 4 which shows results similar to those above but now for each specific country of birth.

We make the following observations:

- **Restricted pairing increases FMR**: Within each country, there is a more than order of magnitude increase in FMR between the zero-effort pair anyone-with-anyone setting, and the same-age, same-sex, same-country pairing. This re-iterates the results of the previous section, and shows it applies globally.
- **Country-of-birth matters**: For many of the different levels of demographic pairing there is between one and two orders of magnitude between the 24 countries represented in this dataset. For example when imposters are from the same sex and country but of any age, the algorithm gives FMR of 0.000046 on Polish faces and 0.0024 on Vietnamese, a fifty fold increase.
- **Regions with highest and lowest FMR**: Across algorithms often the lowest FMR is observed in Eastern European populations and the highest in East Asian populations. However there are important exceptions: Some algorithms developed in East Asia tend to give lower FMR in photos of subjects born in East Asian countries[9]. This observation and the topic of demographic differentials associated with na-

[8] Genetic influence on friction ridge structure is known: The absence of the SMARCAD1 gene leads to absence of fingerprints at birth. Further, the distance between friction ridges is smaller, on average in women than in men, and this may well be under genetic influence. The distance itself is likely not used as a biometric feature, at least not explicitly. Fingerprint pattern classes (arch, whorl etc.), however, have been shown to have regional (geographic) variations, and these were, at least historically, used in one-to-many multi-finger search strategies.

[9] See, for example, the figure in Annex 8 for algorithms from HIK, Dahua, Yitu, Alphaface, Deepsea Tencent, Toshiba.

tional origin are covered more completely in the next section which includes results for comparison of individuals within and across national boundaries.

Discussion: The results above show that false match rates for imposter pairings in likely real-world scenarios are much higher than those from measured when imposters are paired with zero-effort. For this reason NIST has been reporting "matched-covariate" accuracy results in its FRVT evaluation of face verification algorithms [16]. Along similar lines the Australian Department of Foreign Affairs and Trade in tests it sponsors only uses same-sex imposter pairings. The effect of this is to raise thresholds, and thereby raise false non-march rates also. Thresholds increase because they are determined from non-mate scores, s, via the quantile function Q, as that value, T, which gives a proportion, FMR, at or above threshold:

$$T = Q(s, 1 - \text{FMR}) \tag{9}$$

and the set of demographically matched scores is smaller than if all possible comparisons is used.

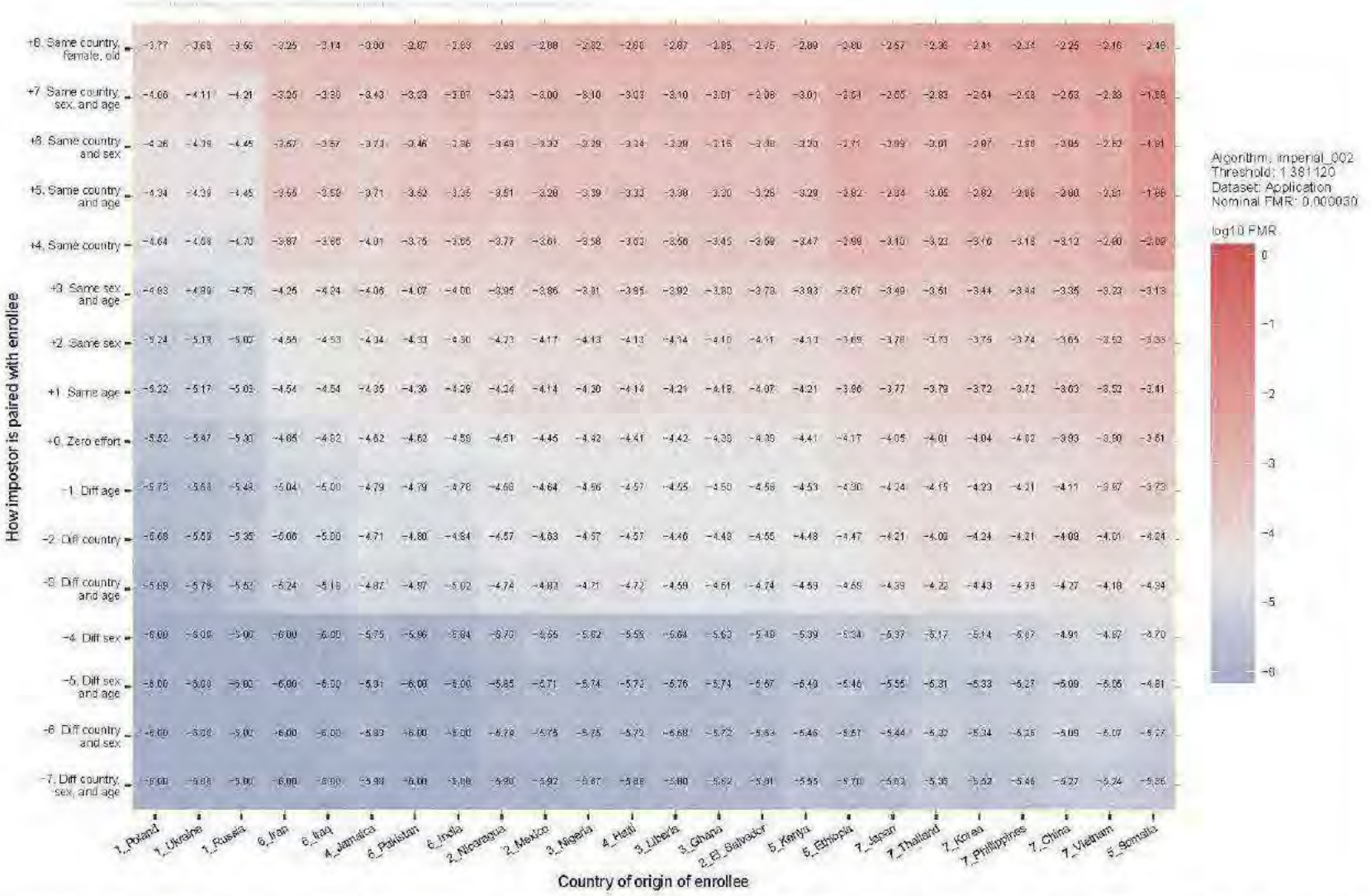


Figure 4: The heatmap shows FMR for each country-of-birth, when the imposter comparisons are drawn from increasingly demographically-matched individuals. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(\text{FMR})$ with large negative values encoding superior false match rates. The center row ("0. Zero effort") row compares individuals without regard to demographics. Rows above that pair imposters more closely until, in the second row, the imposters are of the same sex, age and country of origin. The top row corresponds to one particular demographic often associated with the highest FMR values. The rows below center pair for increasingly unlikely imposter pairings. For example "-5. Diff sex and age" shows FMR for imposters of different sex and age group. The countries appear in order of increasing mean FMR. Values below -6 are pinned to -6. Annex 8 contains the corresponding figure for all algorithms.

4.3 False match rates within and across countries

Method: Using high quality application portraits drawn from the corpus described in Annex 2, we compared 442 019 images from 24 countries with 441 517 images of different individuals from the same countries, yielding 195.2 billion imposter comparisons. We executed this set of comparisons with 126 verification algorithms submitted to the FRVT Verification track. These are listed in Table 4-6. We compared scores with a set of 10 thresholds to produce FMR estimates at each of those thresholds. The thresholds were computed over a set of 93 070 400 imposter comparisons made using a different set of images, namely the law enforcement mugshots detailed in Annex 1. Each threshold was selected as the lowest value that gave FMR at or below a target FMR. The target FMR value was 0.00003.

Each photograph was assigned to the age groups defined by the intervals $(00-20]$, $(20-35]$, $(35-50]$, $(50-65]$, and $(65-99]$.

We excluded small numbers of photographs for which country of birth was not available, or for which sex was not listed as male or female.

Each comparison is accompanied by sex, country of birth and age group metadata for the two individuals represented in the photographs. Given many comparisons with the same demographic pairing, we can produce a measurement of FMR when comparing individuals from two demographic groups, for example Polish men over the age of 65 with Mexican women between 20 and 35.

Analysis: To address the issue addressed in the title of this section we produced figures depicting cross-country false match rates. Figure 5 is an example. We restricted the demographics to just men in the largest age group, $(35-50]$, and then repeated that for women. We remove sex and age from the discussion for two reasons: First, to isolate the country-of-origin effect, and, second, to reflect what real-world imposters would do: procure identity credentials from persons of the same age and sex.

Figure 5 shows cross-country FMR for one of the more accurate algorithms. Annex 7 contains corresponding figures for all algorithms, for both men and women. The annex therefore extends to more than 250 pages. We could repeat this visualization for other age groups - the results are similar. We discuss the effect of age itself later. Likewise, we could repeat the visualization for other recognition thresholds. The one adopted corresponds to a FMR = 0.00003. The trends are very similar at any threshold.

The Figure shows FMR as a heatmap. It uses a logarithmic scale, so that a FMR of 0.0001 is represented by a color and a text value of -4, i.e. log to the base 10. Low FMR values are shown in blue. High FMR values are shown in red. A grey color connotes the target FMR value ($\log_{10} 0.00003 = -4.5$). High FMR values present a security concern in verification applications.

Discussion: From the Figure and those in the annexes, we make a number of observations. First by assigning

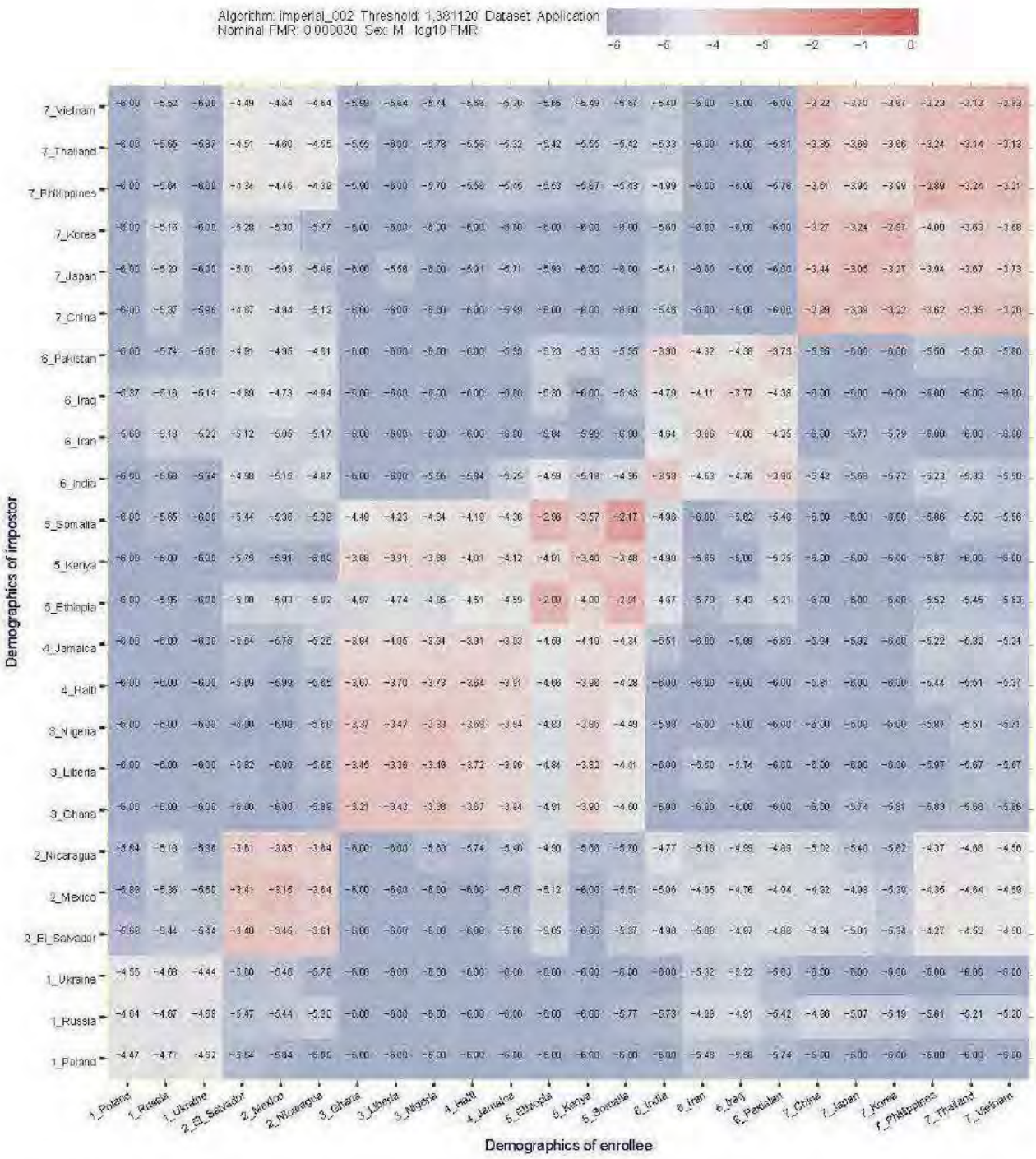


Figure 5: For 24 countries in seven regions the figure shows false positive rates when the reference algorithm is used to compare single photos of mid-aged male subjects from the countries identified in the respective columns. The threshold is to a preset fixed value everywhere. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(\text{FMR})$ *with large negative values encoding superior false match rates. Annex 7 contains the corresponding figure for all algorithms.*

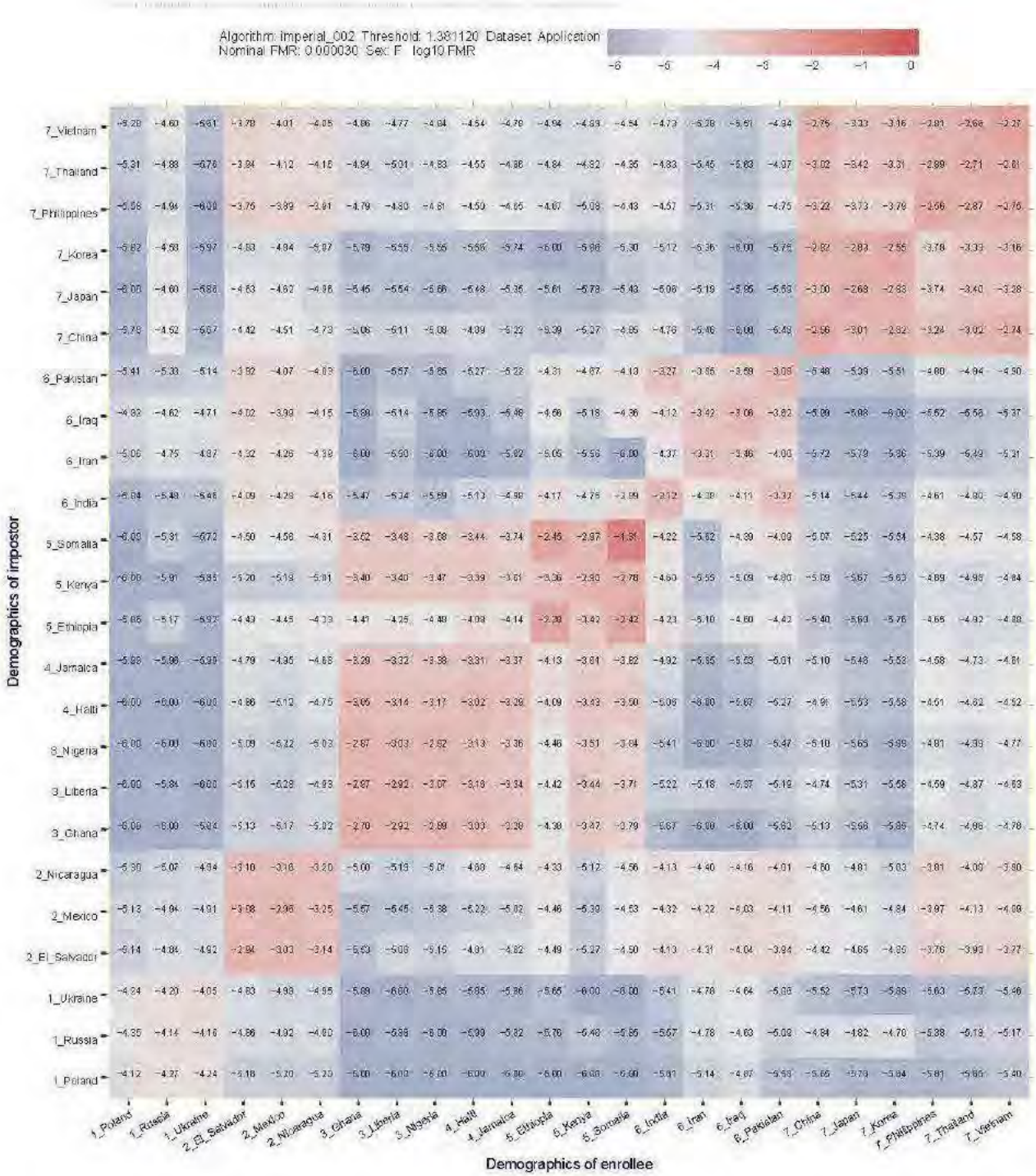


Figure 6: For 24 countries in seven regions the figure shows false positive rates when the reference algorithm is used to compare single photos of mid-aged female subjects from the countries identified in the respective columns. The threshold is to a preset fixed value everywhere. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(\text{FMR})$ with large negative values encoding superior false match rates. Annex 7 contains the corresponding figure for all algorithms.

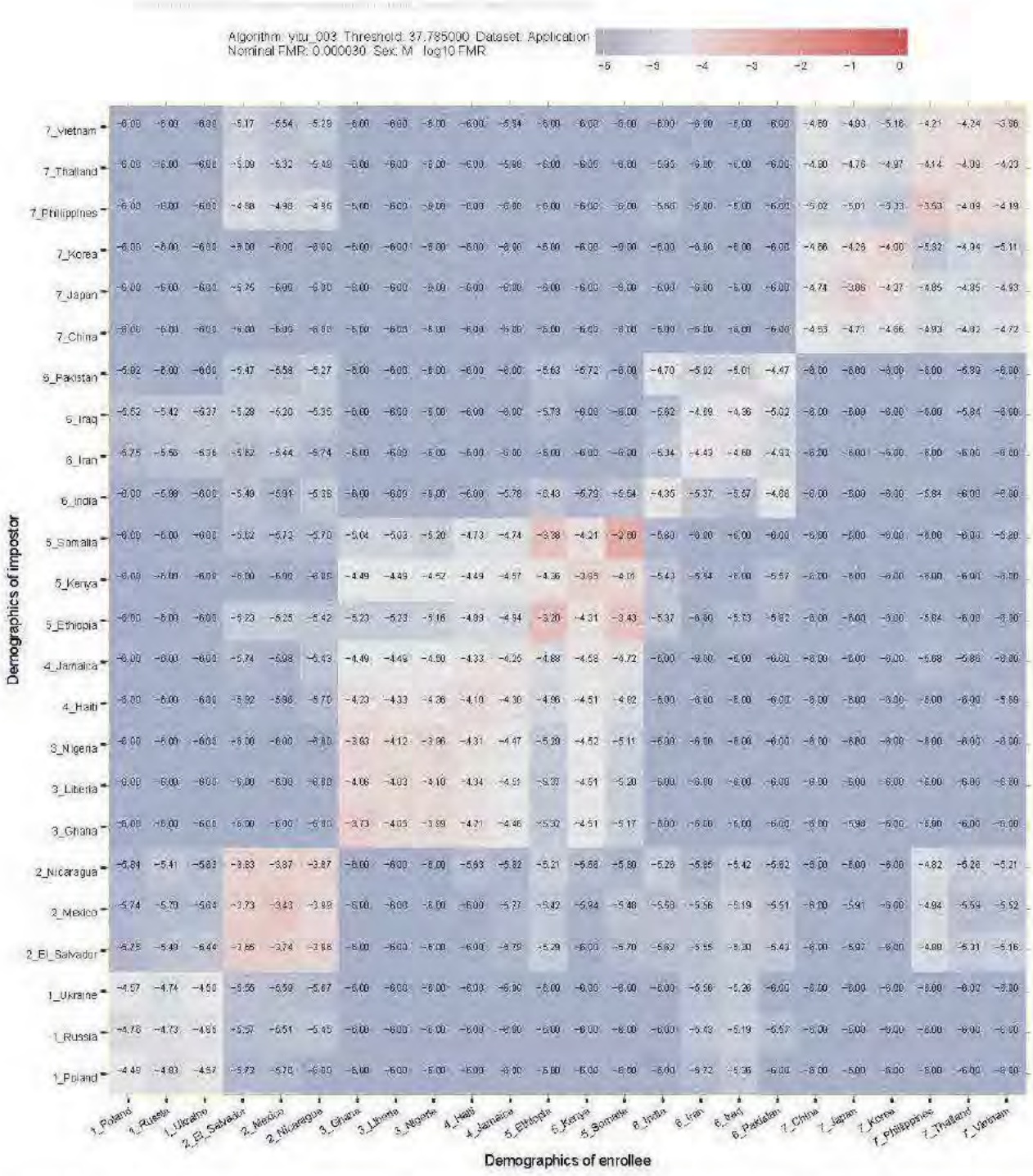


Figure 7: For 24 countries in seven regions the figure shows false positive rates when the **Chinese-developed** *algorithm is used to compare single photos of mid-aged male subjects from the countries identified in the respective columns. The threshold is to a preset fixed value everywhere. Each cell depicts FMR on a logarithmic scale. The text value is* $\log_{10}(\text{FMR})$ *with large negative values encoding superior false match rates. Annex 7 contains the corresponding figure for all algorithms.*

countries into the following regions

- 1: Eastern Europe - Russia, Poland and Ukraine
- 2: Central America - Mexico, Honduras, El Salvador, Nicaragua
- 3: West Africa - Ghana, Liberia, Nigeria
- 4: The Caribbean - Haiti, Jamaica
- 5: East Africa - Ethiopia, Kenya, Somalia
- 6: South Asia - India, Iran, Iraq, Pakistan
- 7: East Asia - China, Japan, Korea, Philippines, Thailand, Vietnam

we see a block structure, in particular a block-diagonal structure indicative of strongly correlated false match rates within region. For example it is true that when comparing photos of individuals from East Africa with those from Eastern Europe, most algorithms give very low FMR. The more interesting results are within-region, around the diagonal, and between regions along the diagonal. We now note the following common trends, and then some notable exceptions. We then conclude with some comments on what the ideal situation would be, and on the meaning. Each Annex includes a "contact sheet" which shows all heatmaps on a single page as thumbnails. The idea is to show macroscopic behavior across all algorithms. When viewed on a computer the figure has very high resolution and zooming in reveals full detail; when printed it will likely just show coarse trends.

- Nominal FMR in Eastern Europe: For many algorithms, FMR within Eastern Europe is close to the nominal target false match rate i.e. a grey color, $-5 \leq \log_{10} \text{FMR} \leq -4$. There are few exceptions to this, even for algorithms developed in China, Western Europe and the USA.
- Higher FMR in East Africa: For almost all algorithms the highest FMR is for comparison of Somali faces. We suspected this could be due to mislabeled data or statistical (in)significance but rejected those possibilities[10]. Further the FMR is high within Ethiopia and between Ethiopia and Somalia. Similarly Kenya-Kenya comparisons give high FMR, although somewhat reduced. In a substantial majority of photos of Somalian women, the subject is wearing full head dress that typically covers the hair and ears leaving only the face exposed. While this might produce false positives, headwear is almost always absent in photographs of men. Further work is needed to explain the observation in more detail.

[10] We discount that this result is anomalous as follows: 1. The sample size may be small for this study, but not absolutely small: The Somalia-Somalia FMR measurement is obtained from 1 733 116 comparisons involving 2632 images of 1974 males. 2. The effect persists when comparing Somalian and Ethiopian faces, and we'd suspect that ground-truth labelling errors - instances of one person being present two IDs - would not persist across national boundaries. 3. In addition to high FMR, which is a count of high imposter scores, the mean similarity score is also very high, an observation that again applies to all algorithms.

- ▷ Higher FMR in West Africa too: The countries with the second highest FMR tend to be in West Africa, i.e. Ghana, Liberia and Nigeria. These countries do not share any borders. The high FMR values occur almost equally within and between countries.

- ▷ Higher FMR between West Africa and the Caribbean: Elevated FMR occurs when comparing faces of individuals from countries in West Africa with those in the Caribbean.

- ▷ Higher FMR between West and East Africa: Elevated FMR occurs when comparing faces of individuals from countries in West Africa and Kenya. The effect is often lower than within either region alone. However, the high FMR does not extend to comparisons of West African and Ethiopian or Somali faces.

- ▷ Higher FMR in East Asia: It is very common for algorithms to give high FMR within East Asian countries and between them. For the algorithm shown, Vietnamese faces strongly match other Vietnamese, and with all the other countries in the region. The East Asian block often divides into northern and southern blocks with reduced, but still high, FMR when individuals are compared between those blocks (e.g. Korea and Vietnam).

- ▷ Some Chinese algorithms give nominal FMR when comparing Chinese: As shown in Annex 7 some algorithms developed in China exhibit much reduced FMR on the East Asian population - for example, see Figure 7. These algorithms are from Megvii, Meiya, Hik Vision, Dahua, X-Laboratory, Yitu and SHU (Shanghai University Film Academy). For Deepsea Tencent the same applies, but less prominently in South East Asia. In some cases the effect is only apparent for comparisons involving images of Chinese, e.g. Star Hybrid. Other Chinese algorithms, however, exhibit the more common trend of producing elevated FMR across East Asia. These include developers of more accurate algorithm such as Alphaface, Deepglint and Sensetime. Thus it is not sufficient for an algorithm to be developed in China for it to mitigate the FMR increase on images from the local population.

- ▷ One of the most accurate algorithms produces more uniform FMR: The corresponding Figure for the Yitu-003 algorithm - Figure 7) - shows that the demographic differentials in FMR are attenuated. As noted the FMR values for comparisons within East Asia are near the nominal value. Notably, however, this applies to West Africa also. This appears to be an important result, as it is a proof that some algorithms do not exhibit higher FMR in those populations. Yitu reported in a meeting in London in October 2017 that its training data included on order of 10^9 photographs of an unspecified (lower) number of Chinese nationals. Whether that is the entirety of their training data is not known.

- ▷ Developer dependency does not apply to South Asia: Neither Lookman nor Tiger IT's algorithm produce nominal FMR on the S. Asian imposter comparisons.

- Magnitudes are large: The East African FMR values are often two orders of magnitude higher than the nominal value and those recorded within Eastern Europe. That is, the $\log_{10}$ FMR values are +2 higher corresponding to FMR that is of order 100 times larger than the de-facto baseline. From a security perspective this is analogous to using a two-digit PIN instead of the common four digits. For West Africa, the FMR values are between one and two orders of magnitude above baseline. A shift of 1.4 on the logarithmic scale corresponds to a factor of 25 increase, for example.

- Anomalies in the figures: The cross-country heatmaps for the SIAT-004, Panasonic PSL-001, and Sensetime-002 algorithms are mostly red, indicating high false match rates for all comparisons. This may arise because the threshold used was computed over comparisons of a different kind of images - mugshots not application portraits. The algorithms are told what kind of image they are being given at the time features are extracted from the image. The consequence is that the imposter distribution for mugshots looks different to that for the application images, and thus thresholds are not portable. This would present an operational issue to any end-user not informed to set the threshold accordingly. In any case, while the heatmaps are mostly red, they still exhibit the same kind of FMR variations seen for many other algorithms.

Discussion: The heatmap figures of Annex 7 show a widespread elevation of false match rates in African faces relative to those in Eastern Europe. The reasons for these shifts are unknown. We did not make any attempts to explain the effects. To summarize the effect we include the scatter plots of Figures 10 - 9. Each point corresponds to one algorithm. Its coordinates show false match rates within West Africa against those within Eastern Europe. The degree to which the point is above the diagonal line shows the extent that FMR in the African countries exceeds that in the Eastern European ones.

We note several outcomes of this visualization.

- **Worst case** In the scatter plot for African women Figure 9 there is a cluster of algorithms located near $x = 0.00012$ and $y = 0.003$. Compared to the target FMR value of 0.00003 (the vertical line) there is a near four-fold increase in FMR of women over men. Much more significantly there is a more than 100-fold vertical excursion from white men to African women.

- **Dispersion** Some algorithms, most notably those from Sensetime give FMR much different to the target value. The threshold was set using Annex 1 mugshots but the Figure reflects FMR measured over comparison of Annex 2 application photos. Both sets of photos are well illuminated portraits, so this instability across datasets would be unwelcome, especially if an algorithm were to be fielded on imagery qualitatively different. Many algorithms do give the expected FMR for white men FMR = 0.00003 as seen in Figure 8.

Figures 10 and 11 repeat the scatterplot summaries for the East Asian demographic too. The picture there is more interesting. While the same pattern is present, it is clear that some algorithms developed in China do not give elevated false match rates relative to Eastern Europeans. The absence of the effect is important in that it implies high FMR in that population is not inevitable. We did not see a corresponding improvement for South Asian faces for the few algorithms we understand were submitted by developers there (in India and Bangladesh).

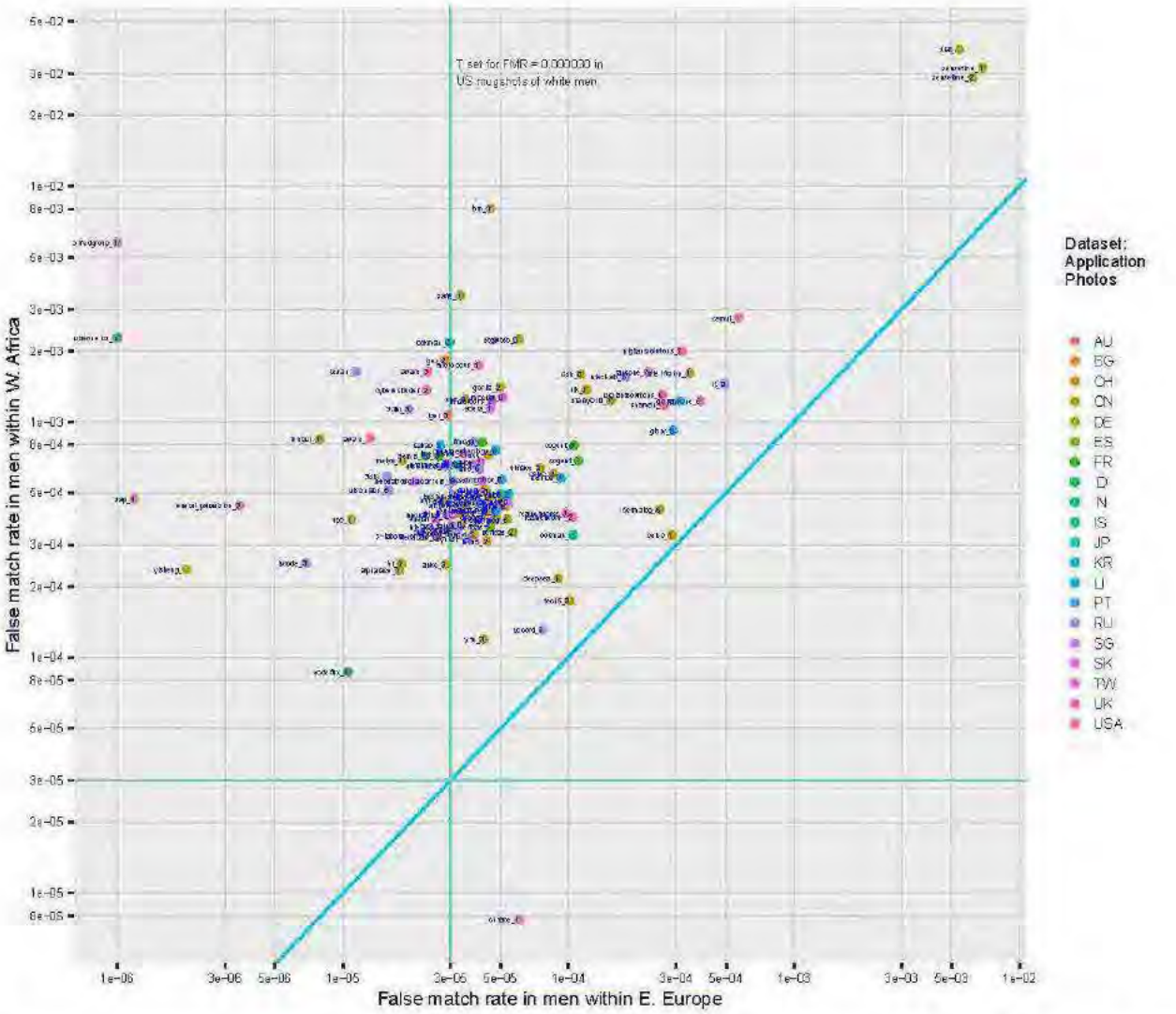


Figure 8: The scatter plot shows FMR when comparing same-age **men** *within and across three Eastern European countries (Russia, Ukraine, Poland), against FMR obtained comparing men within and across three West African countries (Ghana, Liberia, Nigeria). The threshold is fixed for each algorithm to give the FMR noted in the annotation over white men in the U.S. mugshot database. This is indicated by the vertical and horizontal green lines. The blue diagonal line $y = x$ is included to show "over/under". The color code identifies the domicile of the developer - some multinationals conduct research elsewhere. Training data likewise may originate elsewhere.*

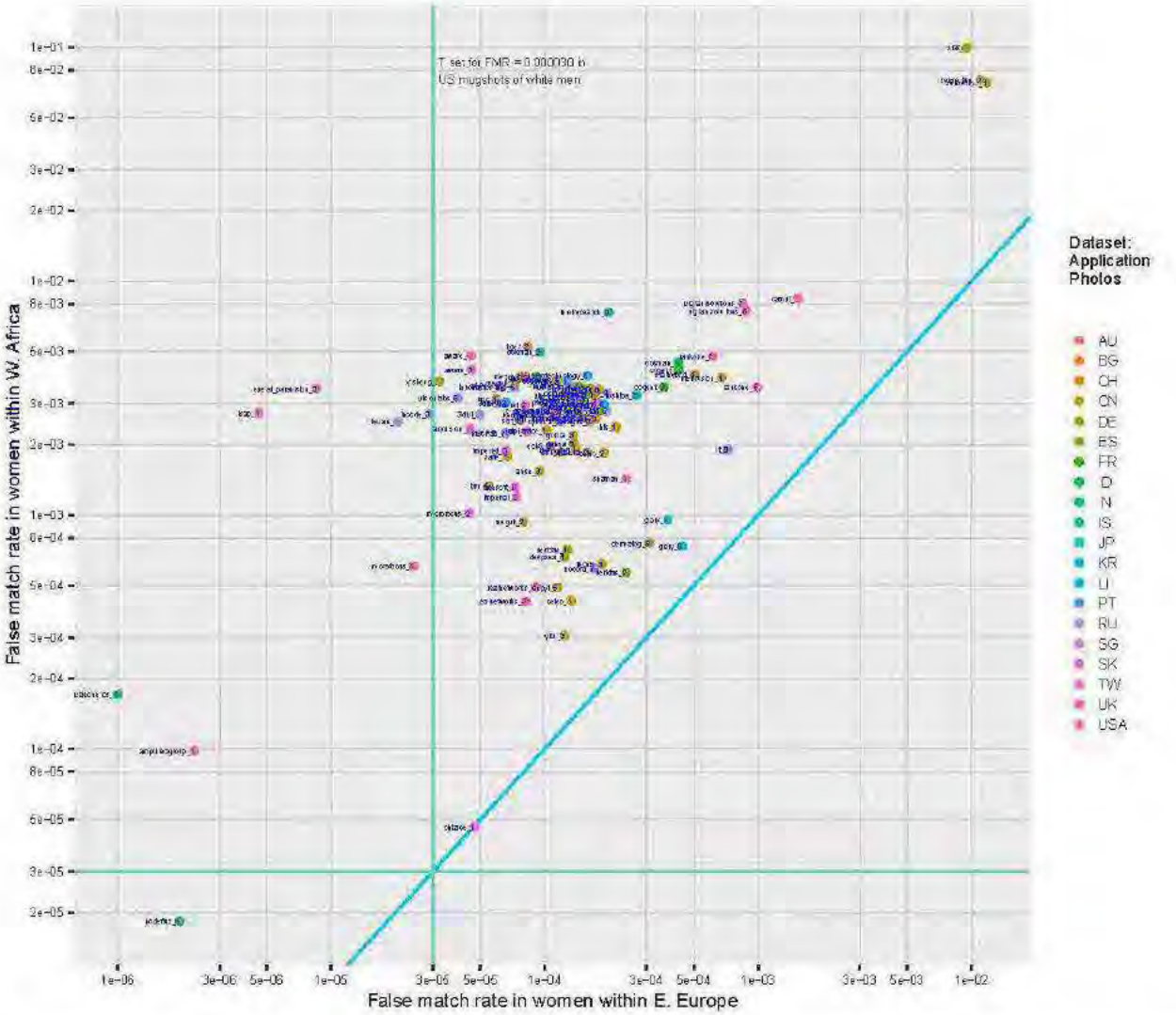


Figure 9: The scatter plot shows FMR when comparing same-age **women** *within and across three Eastern European countries (Russia, Ukraine, Poland), against FMR obtained comparing women within and across three West African countries (Ghana, Liberia, Nigeria). The threshold is fixed for each algorithm to give the FMR noted in the annotation over white men in the U.S. mugshot database. This is indicated by the vertical and horizontal green lines. The blue diagonal line $y = x$ is included to show "over/under". The color code identifies the domicile of the developer - some multinationals conduct research elsewhere. Training data likewise may originate elsewhere.*

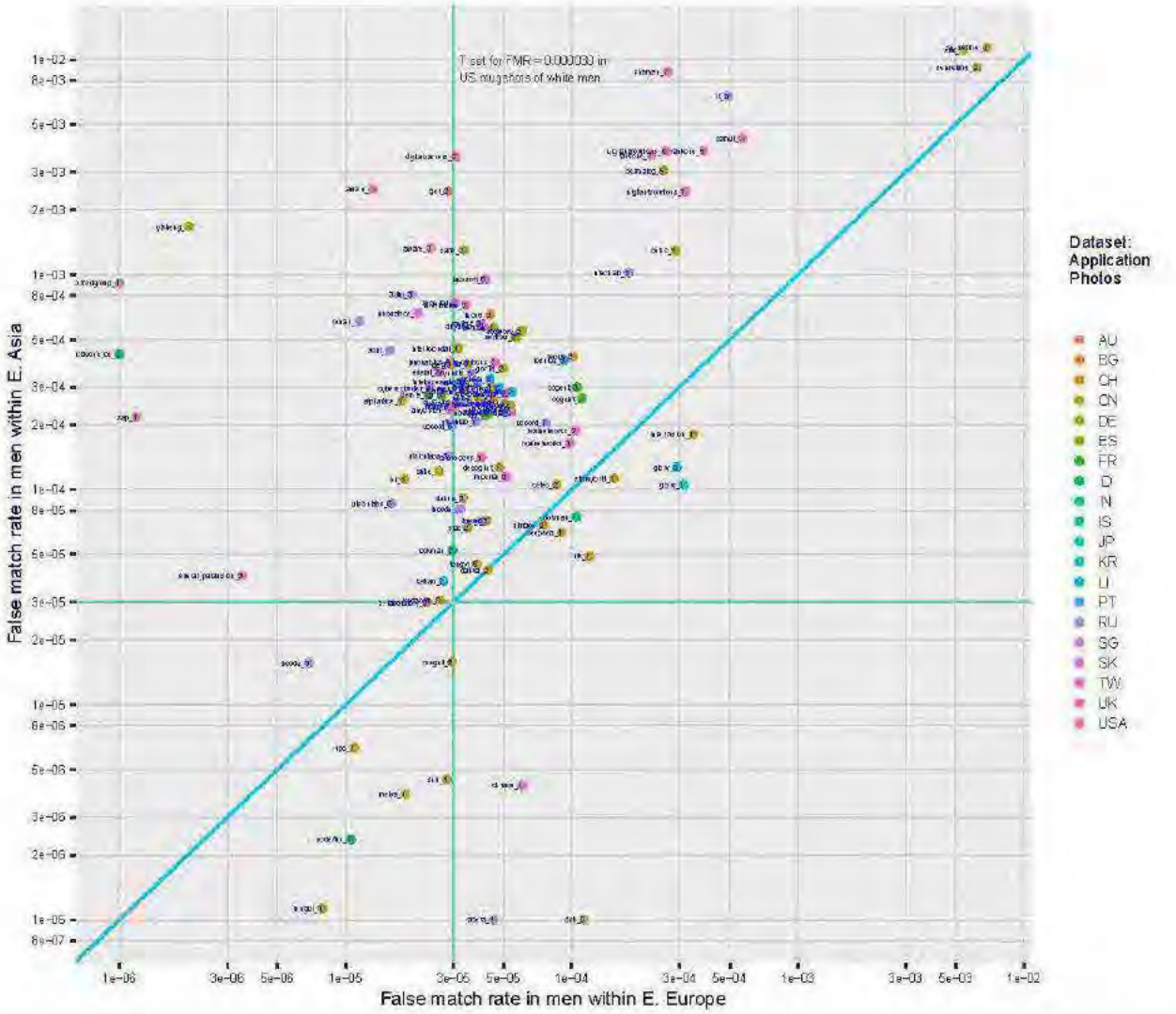


Figure 10: The scatter plot shows FMR when comparing same-age **men** *within and across three Eastern European countries (Poland, Russia, Ukraine), against FMR obtained comparing men within and across six East Asian countries (China, Japan, Korea, Philippines, Thailand and Vietnam). The threshold is fixed for each algorithm to give the FMR noted in the annotation over white men in the U.S. mugshot database. This is indicated by the vertical and horizontal green lines. The blue diagonal line $y = x$ is included to show "over/under". The color code identifies the domicile of the developer - some multinationals conduct research elsewhere. Training data likewise may originate elsewhere.*

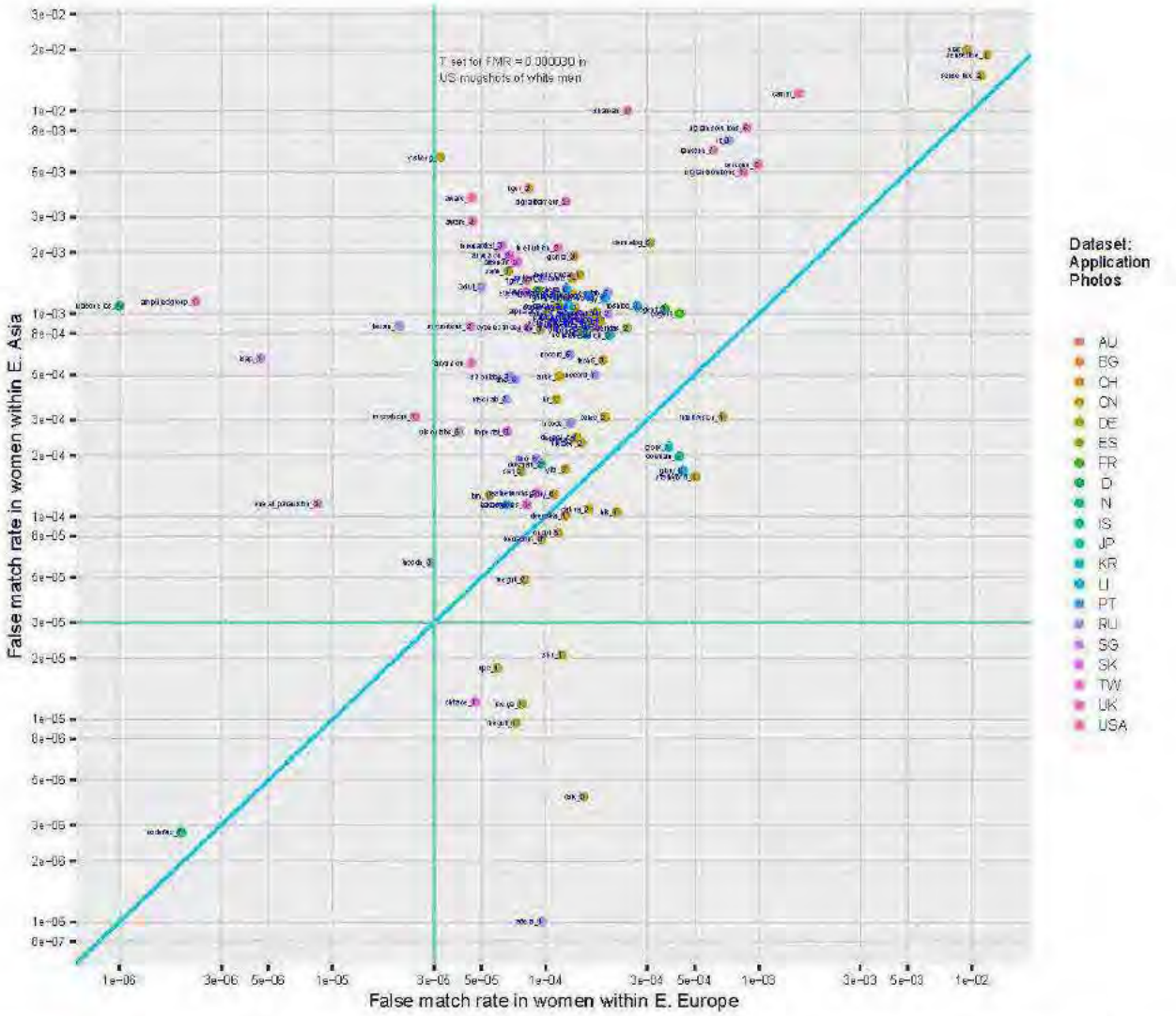


Figure 11: The scatter plot shows FMR when comparing same-age **women** *within and across three Eastern European countries (Poland, Russia, Ukraine), against FMR obtained comparing women within and across six East Asian countries (China, Japan, Korea, Philippines, Thailand and Vietnam). The threshold is fixed for each algorithm to give the FMR noted in the annotation over white men in the U.S. mugshot database. This is indicated by the vertical and horizontal green lines. The blue diagonal line $y = x$ is included to show "over/under". The color code identifies the domicile of the developer - some multinationals conduct research elsewhere. Training data likewise may originate elsewhere.*

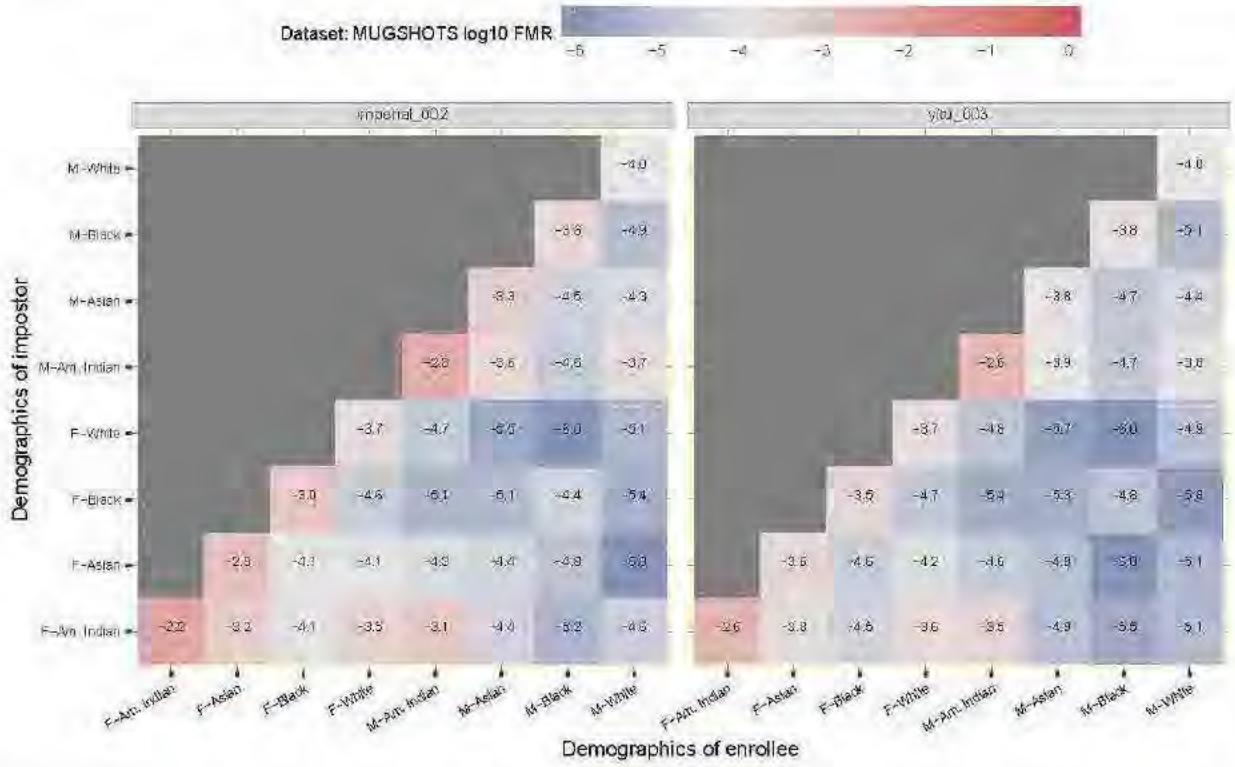


Figure 12: For mugshot photos tagged with one of four race labels and a sex label, the heatmaps show false positive rates for comparison of randomly selected photos from the groups identified in the respective rows and columns. Two algorithms are used, one in each panel, and the threshold for each is set to a fixed value everywhere. The value is the smallest threshold that gives FMR $\leq$ 0.0001 on the white male imposters. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(\text{FMR})$ *with large negative values encoding superior false match rates. Annex 6 contains the corresponding figure for all algorithms.*

4.4 Dependence of FMR on race in United States mugshots

Method: Using high quality mugshot portraits from the mugshot images detailed Annex 1 , we apply each verification algorithm to conduct 3 million comparisons for each of the eight demographics defined by two sexes and four races. The origin and meaning of these labels is described in the Annex. We executed this set of comparisons with 126 verification algorithms submitted to the FRVT Verification track. These are listed in Tables 4-6. We compared scores with a threshold to produce FMR estimates for each demographic pairing. Each threshold was selected as the lowest value that gave FMR at or below a target FMR. The target FMR value was 0.0001. The threshold was computed over the set of 3 000 000 mugshot imposter comparisons made for white males. Thus, by design, the FMR for that demographic is exactly 0.0001.

We excluded photographs for which race or sex was unavailable or unknown. We did not report comparisons by age-group.

Analysis: As with the international set of application photos, we use the heatmap to show cross-demographic

false match rates, including cross-sex. Heatmaps for two algorithms are shown in Figure 12. The Figure shows FMR as a heatmap. It uses a logarithmic scale, so that a FMR of 0.0001 is represented by a color and a text value of -4, i.e. log to the base 10. Low FMR values are shown in blue. High FMR values are shown in red. A grey color connotes the target FMR value ($\log_{10} 0.0001 = -4$). High FMR values present a security concern in verification applications. Corresponding figures for all algorithms appear in Annex 6

Figure 13 extracts the within-sex and within-race diagonal elements of those figures and summarizes the results for all algorithms, ordering the result by worst-case FMR elevation.

Discussion: From the figure, and those in the annex, we make a number of observations.

- Higher FMR in women: As with application photos, most algorithms give systematically higher false match rates in women than in men. The magnitude of this difference is lower with mugshots than with application photos.
- Highest FMR in American Indians: First, the highest FMR occurs in images of American Indians[11]. For the Imperial-002 algorithm featured in Figure 12 the FMR for American Indian women is 0.0068, i.e. a 68 fold increase over the FMR of 0.0001 in white males. In men, the multiple is 47. Why such large increases occur is not known. One component of the increase may stem from database identity labelling errors[12]. We discount this possibility because the database has otherwise excellent ground-truth integrity, supported by fingerprint enrollments.
- Higher FMR in Asian and Black women: There are order-of-magnitude increases in FMR in mugshots of Asian and Black women. Some algorithms developed in China reduce this differential, for example Yitu-003 in the right panel of Figure 12.

[11] The data supplied to NIST tags this group with letter "I" per the EBTS standard which describes this group as "American Indian, Eskimo, Alaskan native, or a person having origins in any of the 48 contiguous states of the United States or Alaska who maintains cultural identification through tribal affiliation or community recognition". In the figures we replace the letter "I" with "American Indian" to distinguish from subjects from India in the international datasets.

[12] Specifically instances of "one person under two IDs" can cause apparent false positives, that are actually true positives.

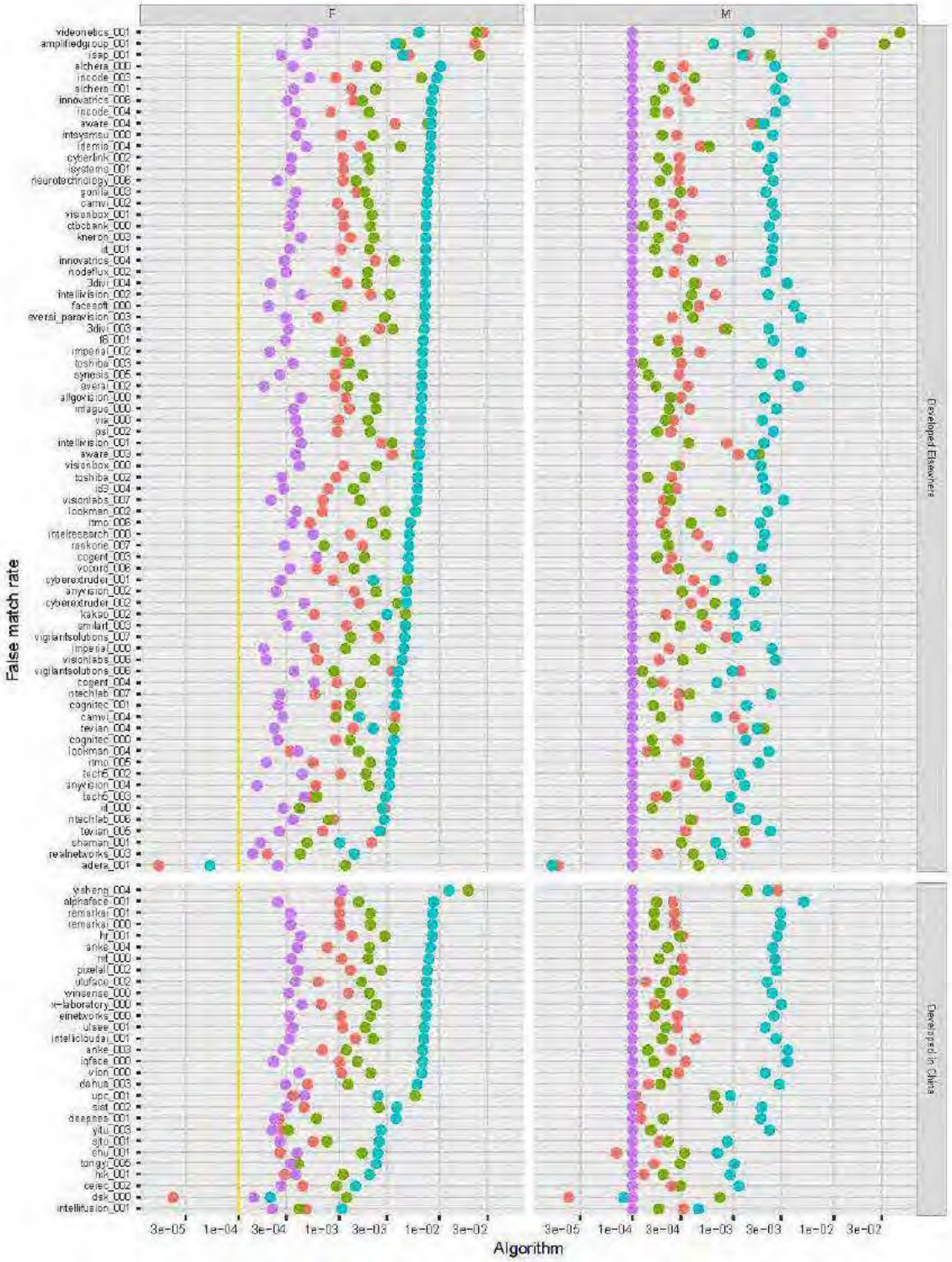


Figure 13: For each verification algorithm, the dots give the false match rates for same-sex and same-race imposter comparisons. The threshold is set for each algorithm to give FMR = 0.0001 on white males (the purple dots in the right hand panel). The algorithms are sorted in order of worst case FMR, usually for American Indian women. Algorithms developed in China appear in the lower panel.

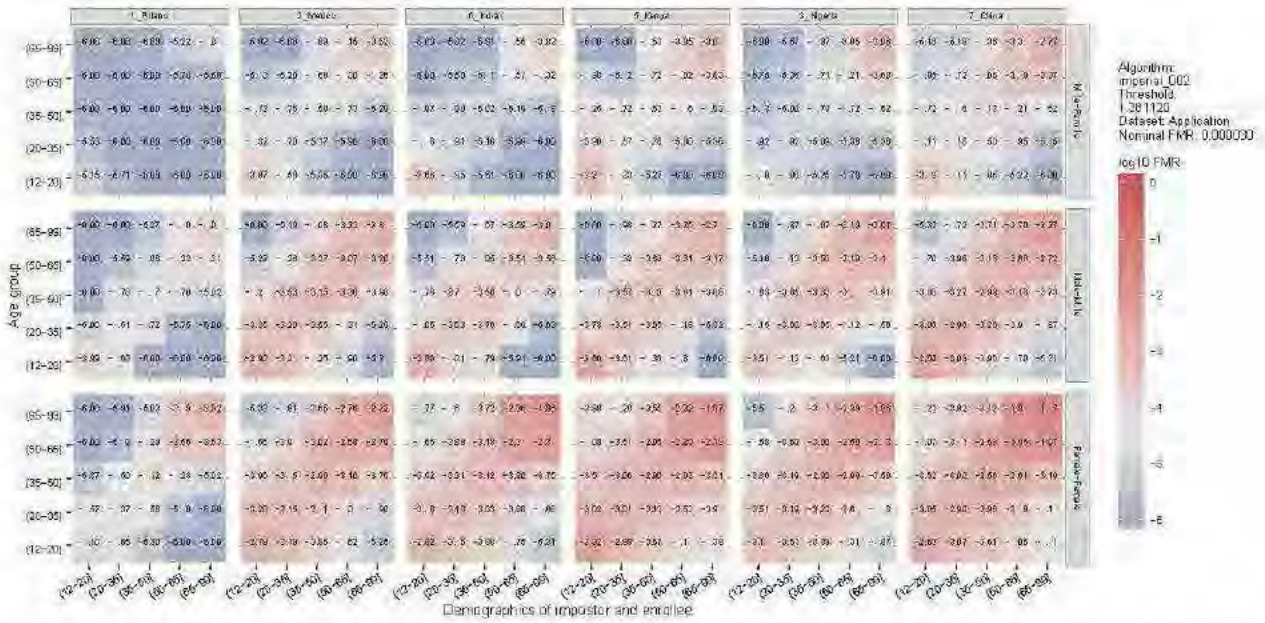


Figure 14: For six countries selected for the high number of images in the dataset and from distinct regions the heatmaps show cross-age false match rates for imposters of the same sex from the age groups given on the respective axes. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(\text{FMR})$ with large negative values encoding superior false match rates. Annex 9 contains the corresponding figure for all algorithms.

4.5 Do some or all algorithms yield more false positives on certain age groups

Method: Using high quality application portraits drawn from the corpus described in Annex 2, we compared 442 019 images from 24 countries with 441 517 images of different individuals within and across age groups $(00-20]$, $(20-35]$, $(35-50]$, $(50-65]$, and $(65-99]$.

We executed this set of comparisons with 126 verification algorithms submitted to the FRVT Verification track. These are listed in Tables 4-6. Each comparison yield a score. When many scores are compared with a fixed threshold, we obtain an estimate of the false match rate. The threshold was computed over a set of 93 070 400 imposter comparisons made using a different set of images, namely the mugshots detailed in Annex 1. The threshold is the smallest value that for which the FMR is less than or equal to 0.00003. This was repeated for other thresholds giving FMR $\{0.000001, 0.000003, 0.00001, 0.00003, 0.0001, 0.0003, 0.001, 0.003, 0.01, 0.03\}$.

Each comparison is accompanied by sex, country of birth and age group metadata for the two individuals represented in the photographs. We excluded small numbers of photographs for which age information was unavailable or for which sex was not listed as male or female.

Given many comparisons with the same demographic pairing, we can produce a measurement of FMR when comparing individuals from two age groups, for example Polish men over the age of 65 with Polish men under 20.

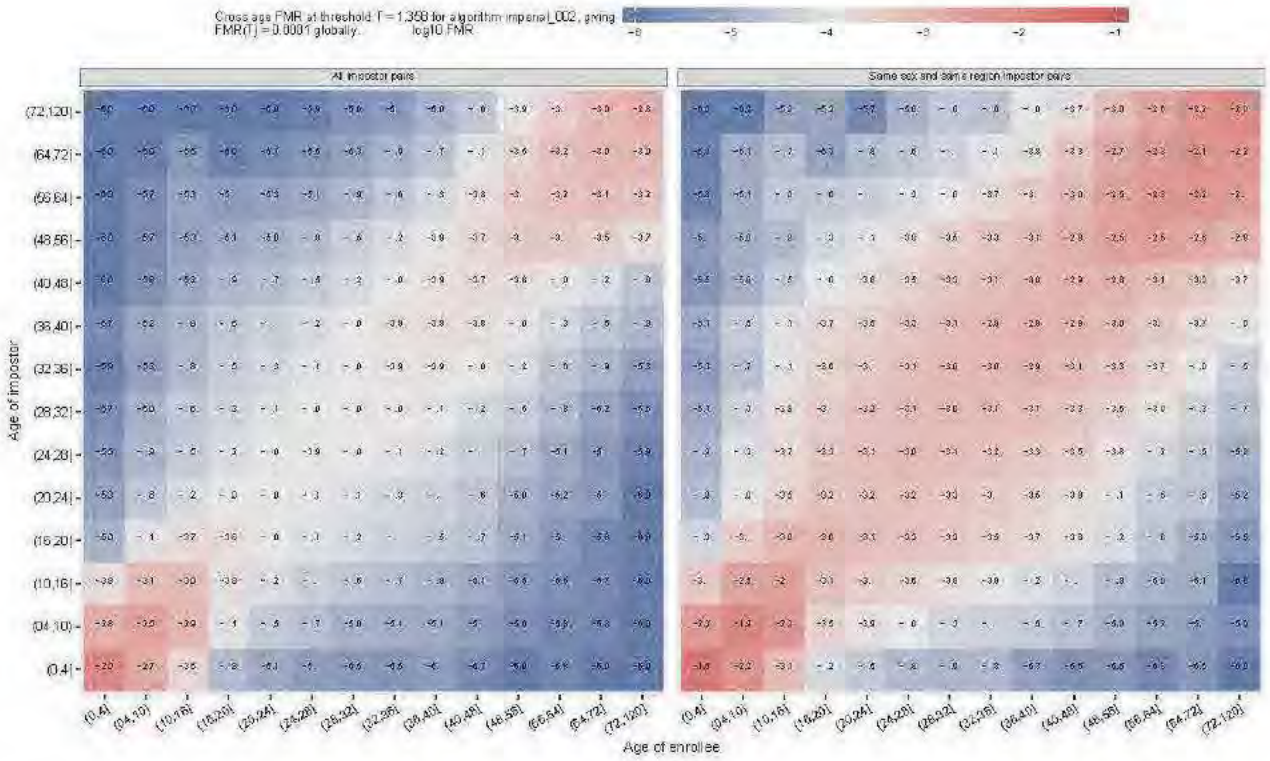


Figure 15: For visa photos from all countries, the heatmap shows for one algorithm cross-age false match rates for imposters of the same sex. Each cell depicts FMR on a logarithmic scale. The text value is $\log_{10}(\text{FMR})$ with large negative values encoding superior false match rates. The threshold is fixed to the value that gives a FMR of 0.0001 over all zero-effort imposter pairs. Annex 10 contains the corresponding figure for all algorithms.

Analysis: To address the issue of age we produced figures depicting cross-age false match rates. We do this within-country only, as cross-country effects have been covered in section 4.3. We include male-male, female-female, and also male-female comparisons (although they are of less interest operationally). Figure 14 is an example, showing results for one of the more accurate algorithms. The Figure includes results for six countries, one per region. We dropped one region (the Caribbean) and 18 of the 24 countries because the effects are similar everywhere.

Figure 14 shows cross-age group FMR for one of the more accurate algorithms. Annex 9 contains corresponding Figures for all algorithms, and therefore extends to more than 130 pages.

Discussion: From Figure 14 and those in the annex, we make these observations.

- Lower FMR for persons in different groups: In almost all cases - for all algorithms, countries of origin and both sexes, comparison of images of persons in different age groups yields lower (better) false match rates than for persons in the same age group. This, obviously, is an aggregate result; it will generally be possible to find some individuals from different age groups who produce high imposter scores but this will be increasingly difficult as the age difference increases.

- Highest FMR in the oldest age group: For women from all most countries, comparison of images of individuals in the 65-and-over age group produce the highest false match rates. For men this is often true also.

- High FMR in the youngest age group: For both sexes, but men in particular, comparison of images of persons in the $12-20$ age group produce high false match rates. The dataset does not include any subjects below 12. Below that age we consider a smaller dataset of visa photographs (see Annex 3) that includes individuals in age groups $(0, 4]$ and $(4, 10]$. The results are included in the heatmap of Figure 15. Note that each FMR estimate is formed from comparisons from all countries, not just one, so they hide the geographic idiosyncrasies of the algorithms.

  These results are similar to those reported by Michalski et al. [28] for false positives in children using one commercial algorithm. The report also shows false negative ageing effects broken out by age at enrolment, and time lapse.

- Lower FMR across sex: Comparison of images of persons of different sex usually produces very low FMR. However, within the youngest and oldest age groups, FMR is again higher and substantially above the nominal FMR.

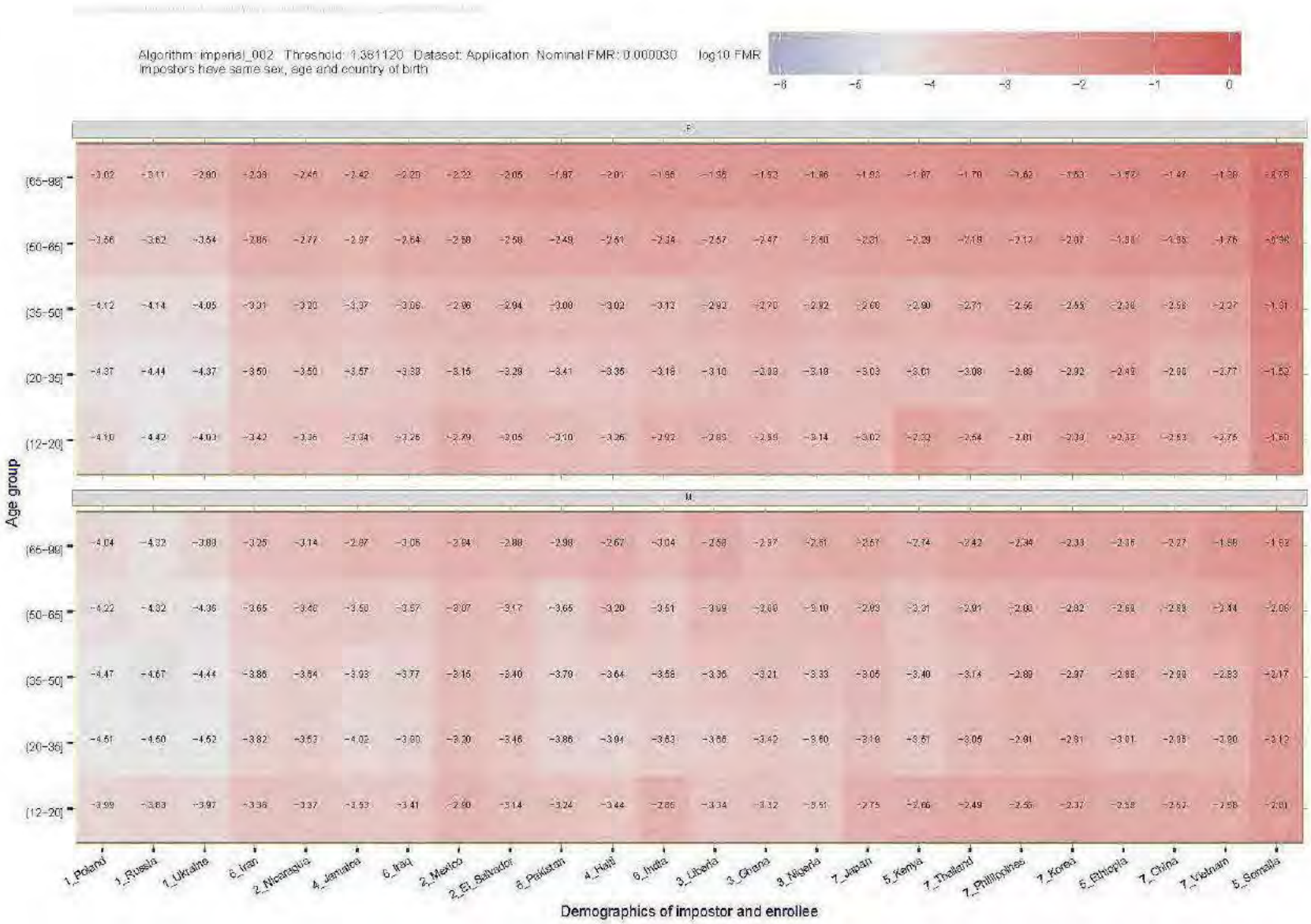


Figure 16: For application photos, the The heatmap shows one-to-one false match rates for same-sex, same-age and same-country of birth imposters, broken out by age and country. The text value is $\log_{10}(\text{FMR})$ *with large negative values encoding superior false match rates. Each cell depicts FMR on a logarithmic scale. The text value is* $\log_{10}(\text{FMR})$ *with large negative values encoding superior false match rates. Annex 11 contains the corresponding figure for all algorithms.*

5 False negative differentials in verification

5.1 Introduction

False negatives occur in biometric systems when samples from one individual yield a comparison score below a threshold. This will occur when the features extracted from two input photographs are insufficiently similar. Recall that face recognition is implemented as a differential operator: two samples are analyzed and compared. So a false negative occurs when two from the same face appear different to the algorithm.

5.2 Tests

This section gives empirical quantification of the variation in false negative rates across demographics. We base this on recognition results from three one-to-one verification tests:

- **Mugshot - Mugshot:** In the first test we look for demographic effects in the groups defined by the sex and race labels provided with these United States images - see Annex 1 .
- **Application - Application photo:** We consider also a high quality dataset collected from subjects hailing from twenty four countries in seven global regions.
- **Application - Border crossing photo:** As discussed in Annex 4 , the border crossing photos are collected under time constraints, in high volume immigration environments. The photos there present classic pose and illumination challenges to algorithms.

5.3 Metrics

The metrics appropriate to verification have been detailed in section 3.1. These are related to particular applications in Figure 2. The dicussion in subsequent sections centers on false non-match rates at particular thresholds, i.e. FNMR(T).

5.4 Results

Figure 17 summarizes the false non-match rates for the 52 most accurate algorithms comparing mugshot photos. It does this for each of four race categories and two sexes[13]. Figure 18 takes the same approach but for 20 countries of birth and two age groups (over/under 45). It summarizes comparison of high quality immigration

[13]See Annex 1 for descriptions of the images and metadata.

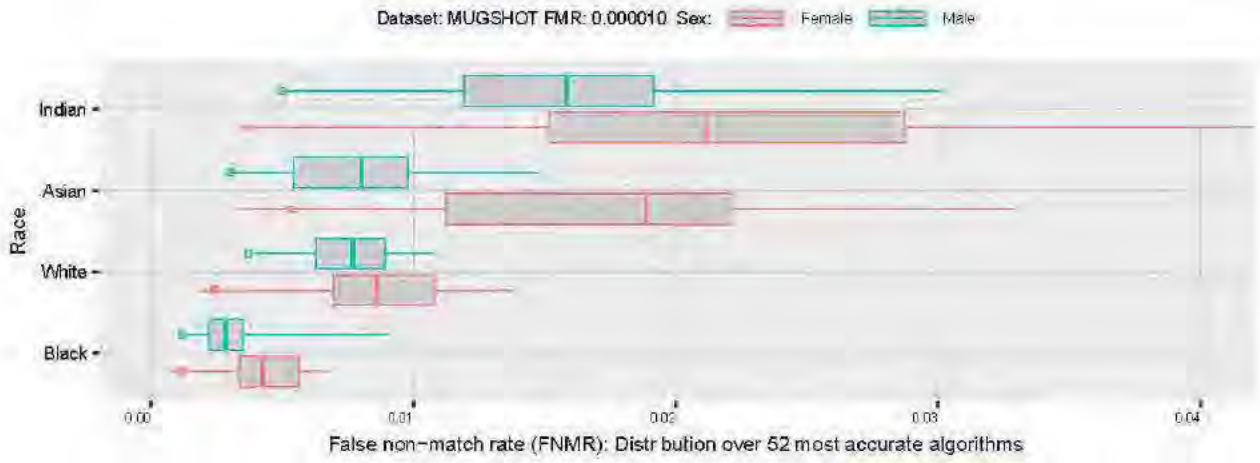


Figure 17: For mugshot comparisons, the figure shows the distribution of FNMR values over the 52 most accurate verification algorithms, by sex and race. The threshold was set for each algorithm to achieve FMR = 0.00001 over all imposter comparisons. The line within each box is the median over those algorithms; the box itself spans the interquartile range (26 algorithms) and the lines here extend to minimum and maximum values. The small box on the left side indicates the accuracy for best algorithm overall, on this dataset alphaface-001.

application photos with lower quality border crossing photos. These are described in Annex 2 and Annex 4 respectively.

We make the following observations.

- FNMR is absolutely low: In one-to-one verification of mugshots, the best algorithms give FNMR below 0.5% at the reasonably stringent FMR criterion of 0.00001. FNMR is generally below 1% with exceptions discussed below. For the more difficult application-border crossing comparisons, the best algorithm almost always gives FNMR below 1%. These error rates are far better than the gender-classification error rates that spawned widespread coverage of bias in face recognition. In that study [5], two algorithms assigned the wrong gender to black females almost 35% of the time. The recognition error rates here, even from middling algorithms, are an order of magnitude lower. Thus, to the extent there are demographic differentials, they are much smaller than those that (correctly) motivated criticisms of the 2017-era gender classification algorithms.

- FNMR in African and African American subjects: In domestic mugshots, the lowest FNMR in images of subjects whose race is listed as black. However, when comparing high-quality appliction photos with border-crossing images, FNMR is often highest in African born subjects. We don't formally measure contrast or brightness in order to determine why this occurs, but inspection of the border quality images shows underexposure of dark skinned individuals often due to bright background lighting in the border crossing environment. In mugshots this does not occur. In neither case is the camera at fault.

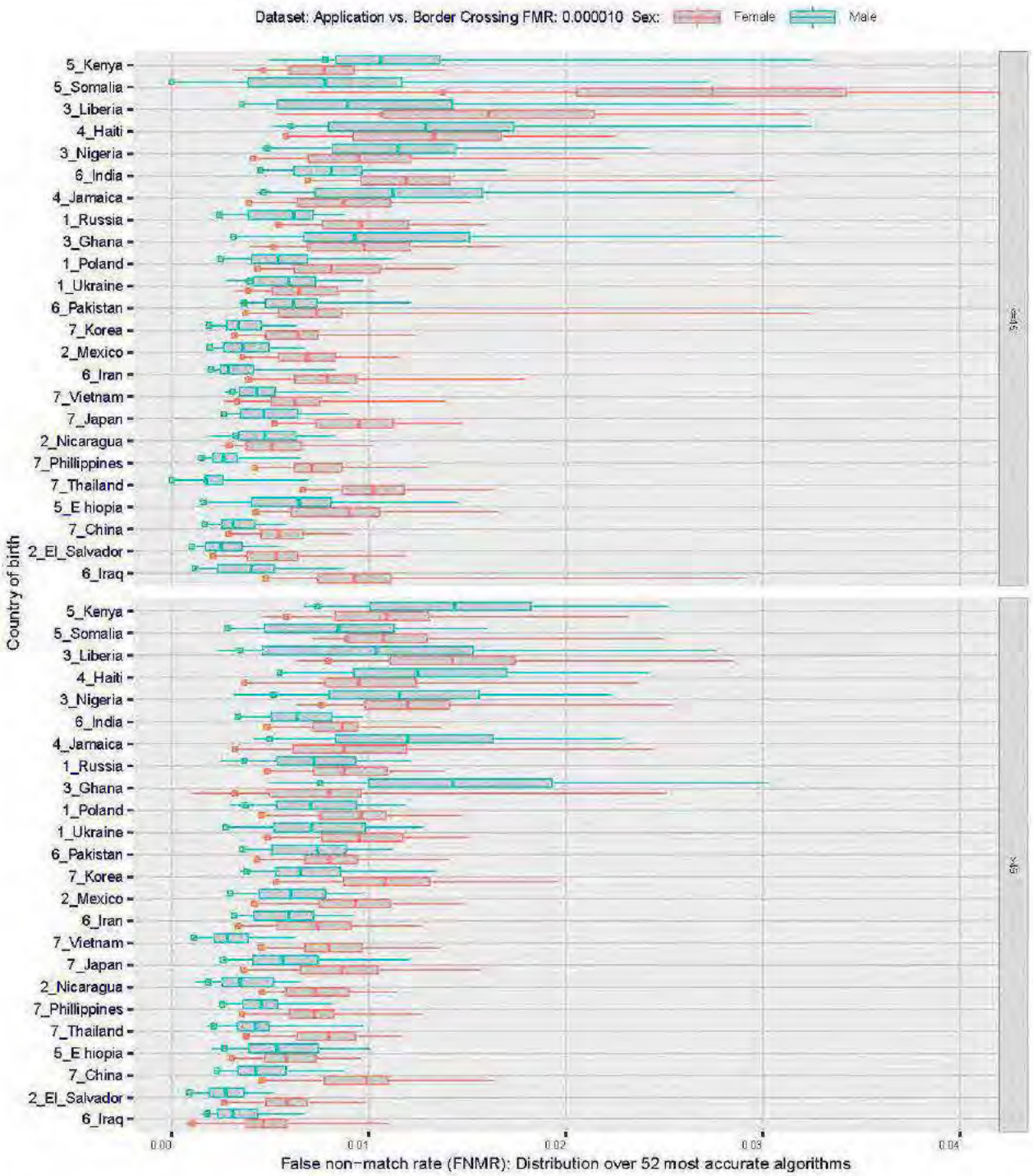


Figure 18: For the application - border crossing photo comparisons, the boxplots show the distribution of FNMR values over the 52 most accurate algorithms, by sex, country of birth, and age group. The threshold was set for each algorithm to achieve FMR = 0.00001 over all imposter comparisons. The line within each box is the median over those algorithms; the box itself spans the interquartile range (26 algorithms) and the lines here extend to minimum and maximum values. The small box on the left side indicates the accuracy for best algorithm overall, on this dataset visionlabs-007.

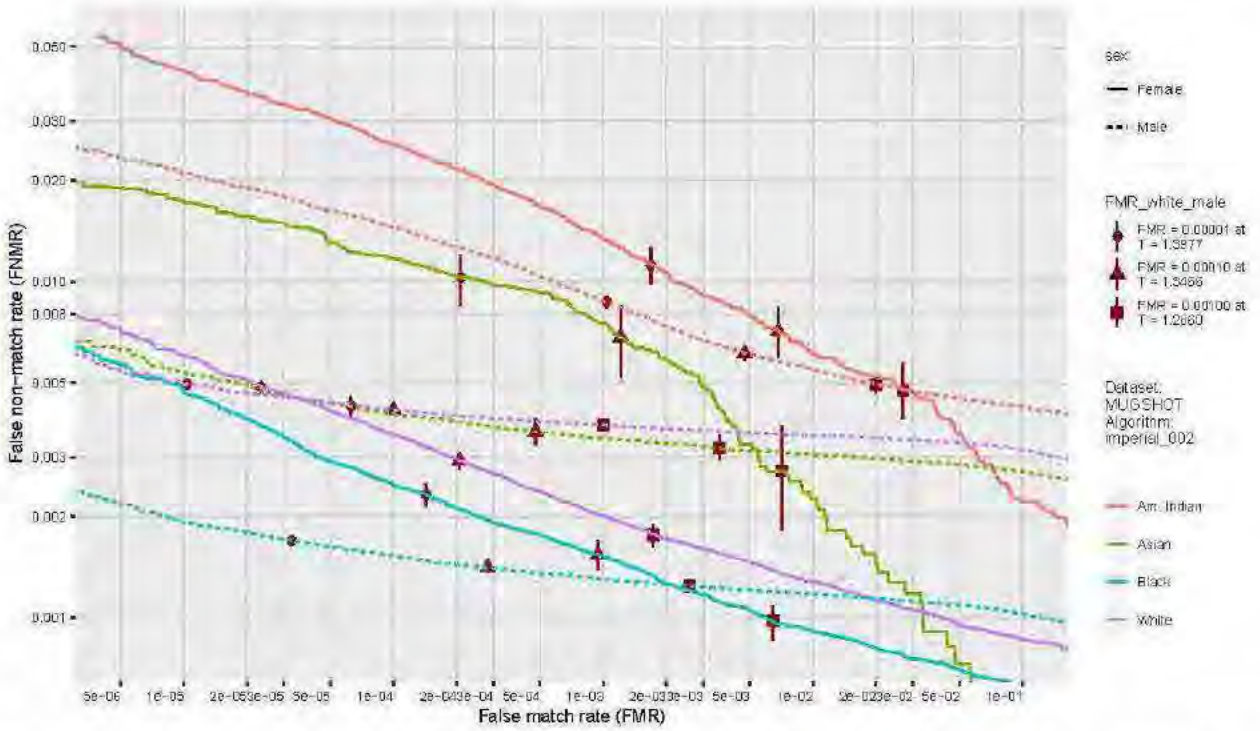


Figure 19: For one algorithm verifying mugshot images, the error tradeoff characteristics show false non-match vs. false match rates. The FMR estimates are computed for same-sex and same-race imposter comparisons. Each symbol (circle, triangle, square) corresponds to a fixed threshold - their vertical and horizontal displacements reveal, respectively, differences in FNMR and FMR between demographic groups. The vertical line through each symbol indicates uncertainty related to sample size - it spans 95% of bootstrap samples of the genuine scores. Annex 12 contains the corresponding figure for all algorithms.

- Women give higher FNMR: In most cases, algorithms give higher false non-match rates in women than men. Note that this is a marginal effect - perhaps 98% of women are still correctly verified - so the effect is confined to fewer than 2% of comparisons where algorithms fail to verify. It is possible that the error differences are due to relative prevalence some unknown covariate. There are some exceptions, however: In Kenya, Nigeria, Jamaica men give higher FNMR. This applies in Haiti and Ghana also but only for people aged 45 or over.

These aggregations of results over a large number of algorithms is intended to expose coarse differences between demographic groups. In so doing it hides that certain algorithms may differ from the trends evident in the Figure. Full error tradeoff characteristics appear in Annex 12 .

The false negative results for law enforcement images apply to high quality mugshots, collected with deliberate consideration of standards. When image quality degrades, false negatives are expected to increase. We next consider results for the comparison of high quality Annex 2 application reference photos with Annex 4 border crossing images collected in a less controlled environment under some (implicit) time constraint. We report

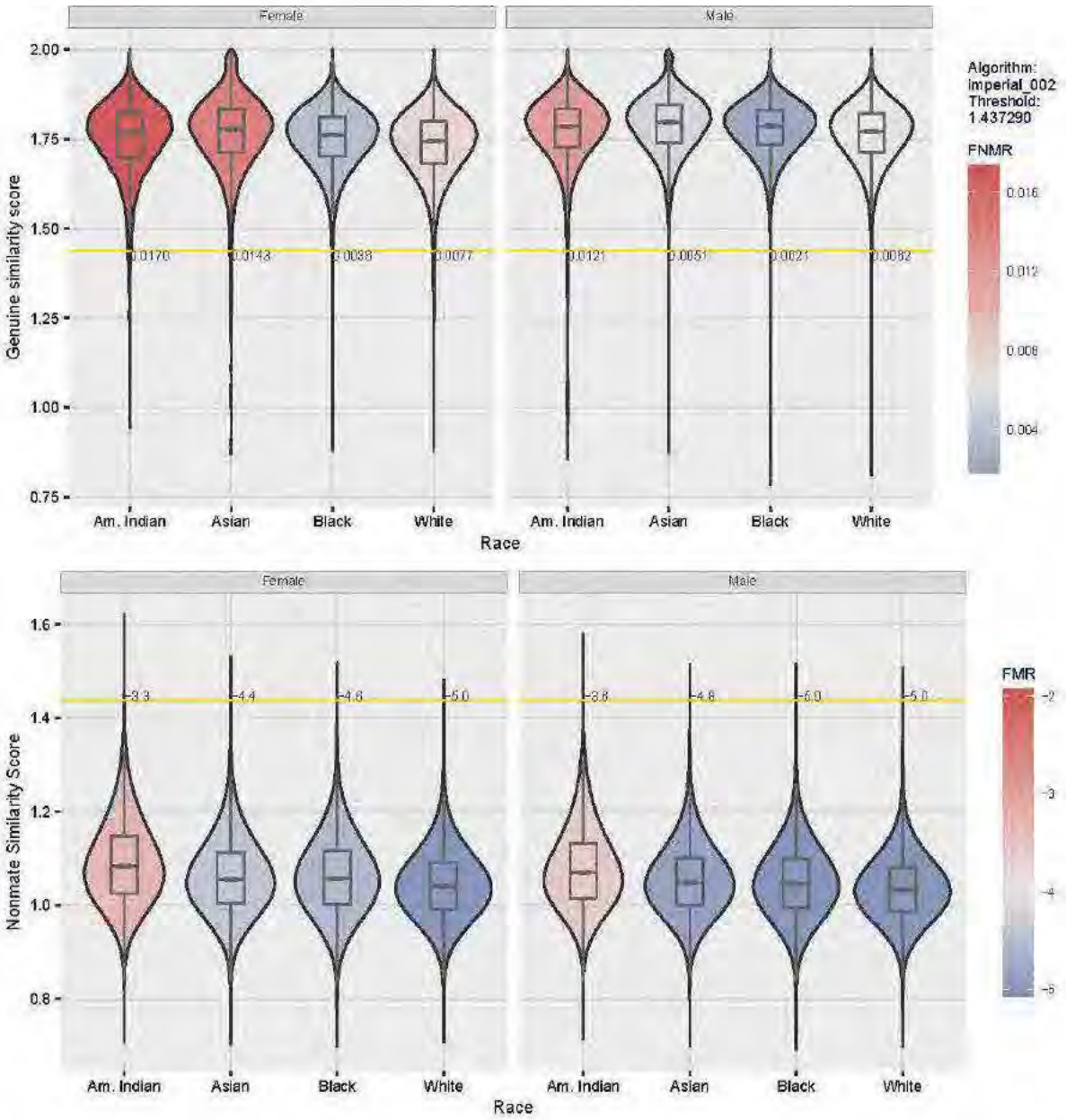


Figure 20: For one algorithm verifying mugshots, the violin plots show native similarity score distributions. The horizontal line shows the threshold that gives FMR = 0.0001 over all the imposter pairs. The imposters have the same sex and race. The upper figure shows genuine scores and the color indicates FNMR at the given threshold on a linear scale. The lower figure shows imposter scores with color indicating FMR on a logarithmic scale. FMR values below 10^{-5} are pinned to that value. Annex 15 contains the corresponding figure for all algorithms.

results in two ways:

- ▷ Per algorithm: Annex 14 shows FNMR by country of birth for two sexes and two age groups (above and below age 45).
- ▷ As Figure 22 heatmap showing results for all algorithms and all countries of birth. Each FNMR is the arithmetic mean of the four FNMR estimates for male and female and age over and under 45. The rows of the figure are sorted in order of mean FNMR, the mean being taken over all twenty four countries. The columns of the figure are sorted in order of mean FNMR from the 50 most accurate algorithms - this statistic was chosen so that high FNMR estimates from poor algorithms did not skew the results.

From these figure we note the following:

- **Wide variation across algorithms:** False non-match rates range from near 0.1% up to above 10%. This two-orders-of-magnitude range shows that some algorithms are intolerant of the quality problems inherent in the image the border crossing images. These problems are: low contrast, non-centered and cropped faces, non-frontal pose, and poor resolution, in part due to poor compression.
- **The most accurate algorithms give low FNMR:** The most accurate algorithms given FNMR below 1% for almost all countries and demographic groups. For example, the Visionlabs-007 algorithm has outliers only for Liberian and Somali women under the age of 45, for whom FNMR is below 1.4%.
- **Lower variation across countries:** For the more accurate algorithms, false non-match rates generally range by a factor of two or three from the left side of Figure 22 to the right i.e. FNMR in El Salvador is almost always lower than that in Somalia.
- **No clear patterns by age and sex:** By considering the Figures of Annex 14 , the differences between the over- and under-45s is often small, varies by country and by algorithm. However, broad statements do not mean that certain algorithms do not exhibit demographic differentials.
- **Higher FNMR in subjects from Africa and the Caribbean:** The heatmap is constructed with countries appearing in order of the mean FNMR over the fifty most accurate algorithms. This reveals higher FNMR in Africa and the Caribbean. After those two regions, the next highest FNMR is in the Eastern Europe countries.

The low error rates stem from efforts over the last decade to train algorithms that are invariant to nuisance variables such as non-frontal pose and poor contrast. The absolute magnitude of FNMR drives inconvenience. In many applications, any subject experiencing a false rejection could make a second attempt at recognition.

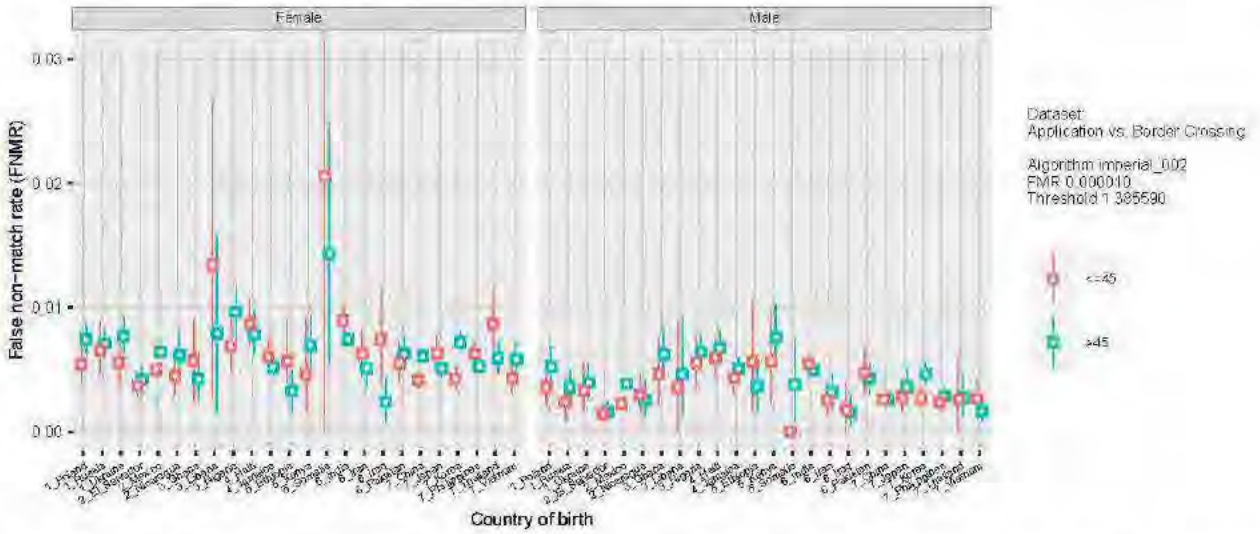


Figure 21: For 24 countries the figure shows false negative rates when the reference algorithm is used to compare two photos of subjects from the countries identified in the respective columns. The square box gives the median false non-match rate computed over 2000 bootstrap resamples of the genuine scores. The ends of the line span 95% of those re-samples, thereby giving a measure of uncertainty in the FNMR estimate. The threshold is set to a fixed value everywhere; it is the lowest value that gives FMR $\leq$ 0.00001. Annex 14 contains the corresponding figure for all algorithms.

Why these effects occur would require some multivariate analysis of image- and subject-specific properties. We suggest that analysis might start with measurement of image related quantities from the digital images to include such as contrast, intensity, areas of over and under exposure, presence of cropping, and head orientation. For tools, mixed-effects regression models could be an initial starting point [4] but such work would need to address correlation between quantities such race and contrast. We have not yet initiated such work and it is possible that such analysis would be incomplete due to influential but unknown covariates. In particular, given the border crossing images were collected with cameras mounted at fixed height and are steered by the immigration officer toward the face it is possible that subject height influences genuine matching scores. For example very tall subjects might be subject be underexposed because strong ceiling lights in the background might cause underexposure. Inspection of failure cases invariably leads to insight in such cases. We have not yet conducted that work.

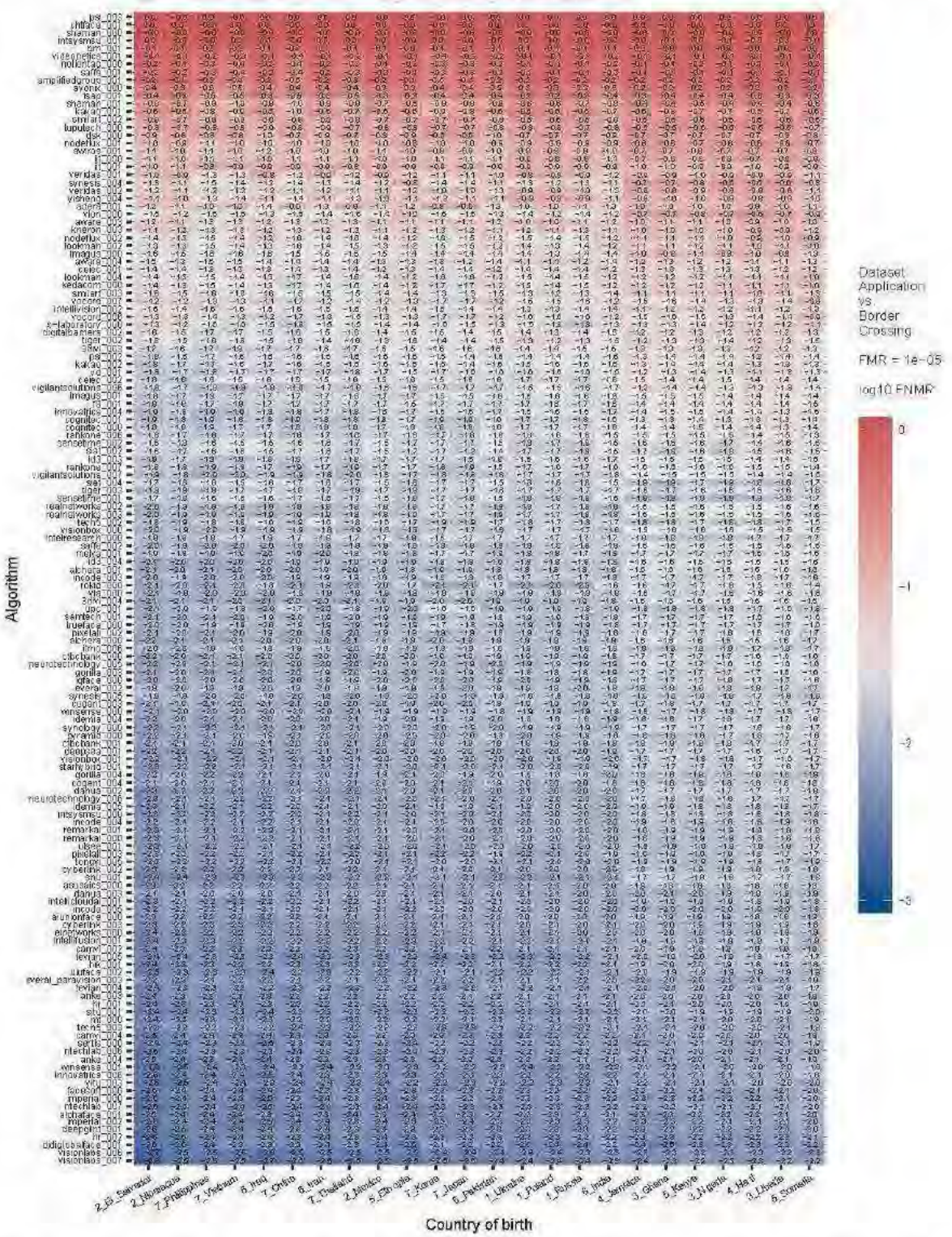


Figure 22: For 24 countries in seven regions the figure shows verification false non-match rates when the reference algorithm is used to compare two photos of subjects from the countries identified in the respective columns. The FNMR value is the mean over men/women and over/under age 45, so represents FNMR in situations where those four populations were balanced. The threshold is set to a fixed value everywhere; is is the lowest value that gives FMR $\leq$ 0.00001. Each cell depicts FNMR on a logarithmic scale. The text value is $\log_{10}(FNMR)$ with large negative values encoding superior false match rates.

6 False negative differentials in identification

The three identification trials all use just mugshot photographs. They were conceived of to isolate specific demographic factors as follows.

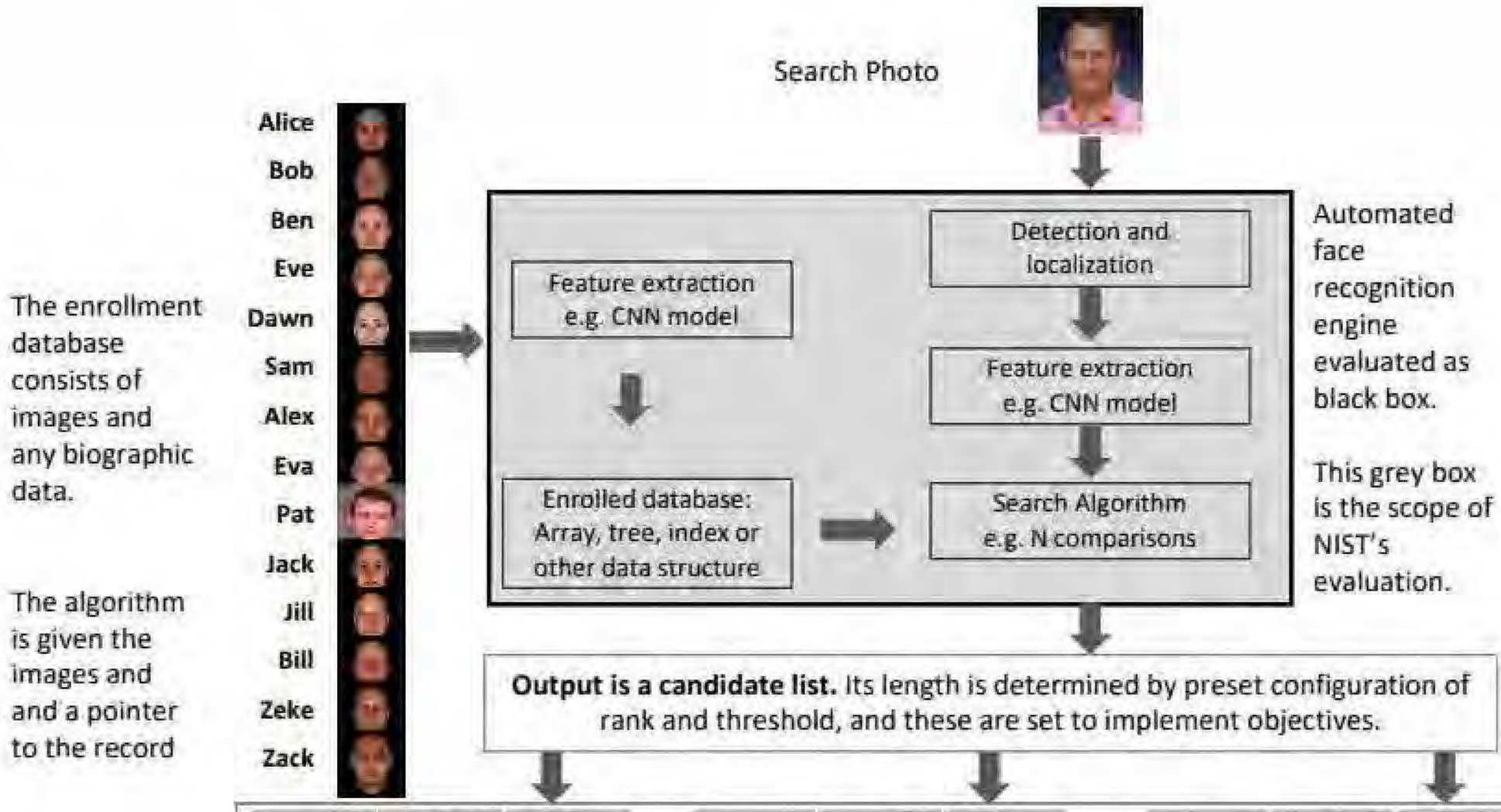
- **Sex**: We construct a gallery containing 800 000 white men, and 800 000 white women, aged 20 - 40. We search that with mated probes taken in a different calendar year to the enrolled photo but no longer than 5 years after. We search with balanced sets of non-mate probes, also aged 20-40.
- **Sex**: We construct a gallery containing 500 000 black men, and 500 000 black women, aged 20 - 40. We search that with mated probes taken in a different calendar year to the enrolled photo but no longer than 5 years after. We search with balanced sets of non-mate probes, also aged 20-40.
- **Race**: We construct a gallery containing 800 000 black men, and 800 000 white men, aged 20 - 40. We search that with mated probes taken in a different calendar year to the enrolled photo but no longer than 5 years after. We search with balanced sets of non-mate probes, also aged 20-40.

More detail appears in Annex 16 . In each case the mated probes are used to measure false negative identification rate, and the nonmated probes are used to measure false positive identification rate. These tests all employ domestic mugshots, and only younger adults. Further work will extend analysis to a global population with more range in age.

6.1 Metrics

The metrics appropriate to identification have been detailed in section 3.2. These are related to particular applications in Figure 23 reflecting two modes of operation. The general metric FNIR(N, R, T) covers both as follows:

- **Investigation:** For investigators willing to traverse long candidate lists in pursuit of a lead, the metric FNIR($N, R, 0$) is the proportion of missed mates when searching an N-enrollee gallery and considering the R most similar candidates without applying a threshold (T = 0). The utility of longer lists is shown by plotting FNIR vs. R.
- **Identification:** For those applications where a non-zero threshold is used to only return results when a search has a likely enrolled mate, the metric is FNMR(N, R, T). The use of thresholds $T > 0$ will suppress many false positives, but will also elevate false negatives, the tradeoff being shown as a plot of FNIR(T) vs. FPIR(T).



	Positive identification	identification	Post-event investigation
Example	Access to a gym or cruise ship	Watchlist e.g. Detection of deportee or duplicate drivers license applications	Crime scene photos, or of detainee without ID documents.
Claim of identity	Implicit claim to be enrolled	Implicit claim to not be enrolled	No claim: Inquiry
Threshold	High, to implement security objective	High, to limit false positives	Zero
Num. candidates	1	0	L, set by request to algorithm
Human role	Review candidate to assist user in resolution of false negatives, or to detect impostor	Review candidate to determine false positive or correct hit	Review multiple candidates, refer possible hits to examiners see [26]
Intended human involvement frequency	Rare – approx. the false negative identification rate plus prior probability of impostor	Rare – approx. the false positive identification rate plus prior probability of an actual mate	Always
Performance metric of interest	**FNIR at low FPIR. See sec. 3.1, 3.2 and Tables 10, 19**	**FNIR at low FPIR. See sec. 3.1, 3.2 and Tables 10, 19**	**FNIR at ranks 1... 50, say. FPIR = 1. See sec. 3.2 and Table 12, 14, 16**

Figure 23: Identification applications and relevant metrics.

6.2 Results

Figures 24 and 25 show identification error rates for two algorithms. Plots for all algorithms are included in Annex 16 . In each case, the upper panels show FNIR vs R. The lower panels show FNIR vs. FPIR. We make the following observations

- Differentials by race in men: From the left-side panels, black men invariably give lower false negative identification rates than white men. This applies in the investigate and identification modes, and particularly for the more accurate algorithms. The differentials are often small, well below a factor of two. The are some exceptions including algorithms from 3DiVi, Aware, Eyedea, Idemia, Kedacom, Tevian and Vocord.

- Differentials between the sexes: Women invariably give higher false negative rates than men. This applies within both racial groups. There are exceptions, notably that searches of white women are more likely to produce the correct result in the top ranks than are search of men. This is less true for black women. A possible mechanism for this is available from section 4 verification results, namely that black women tend to produce high one-to-one false match rates. High non-mate scores may be displacing the correct black women from rank 1 position.

- Low FPIR is not attainable: The error tradeoff characteristics show a rapid increase FNIR as the threshold is increased to reduce FPIR. For example, in FNIR Figure 24, FNIR reaches 50% when FPIR is reduced to 0.0001. This is due to the presence of high scoring non-mates in the imposter searches. They can occur for several reasons. First, ground truth identity labeling errors in which photos of a person are in the database under multiple IDs. These cause apparent false positives. We discount this because the mugshot ground truth integrity is excellent, and underpinned by ten-print fingerprint matching. A second reason is the presence of twins in the population. Given the population represented by the dataset, we estimate a few percent of the United States adult population is present in the dataset. Given well documented twinning rates[14] [27], we expect twins to be in the data, both identical and, more commonly, fraternal. Siblings will be expected to give elevated similarities along the same lines.

- Higher false positive identification rates in black women: The lines connecting points of fixed threshold are often long and slanted in the error tradeoff plots in the center column of the bottom row - see Figure 24, for example. This is a common occurence revealing an order-of-magnitude increase in FPIR, with magnitudes varying by algorithm. Notably some algorithms do not exhibit this excursion. For example, the algorithm featured in Figure 25 gives much smaller excursions in FPIR.

[14] See the CDC's National Vital Statistics Report for 2017: https://www.cdc.gov/nchs/data/nvsr/nvsr67/nvsr67_08-508.pdf

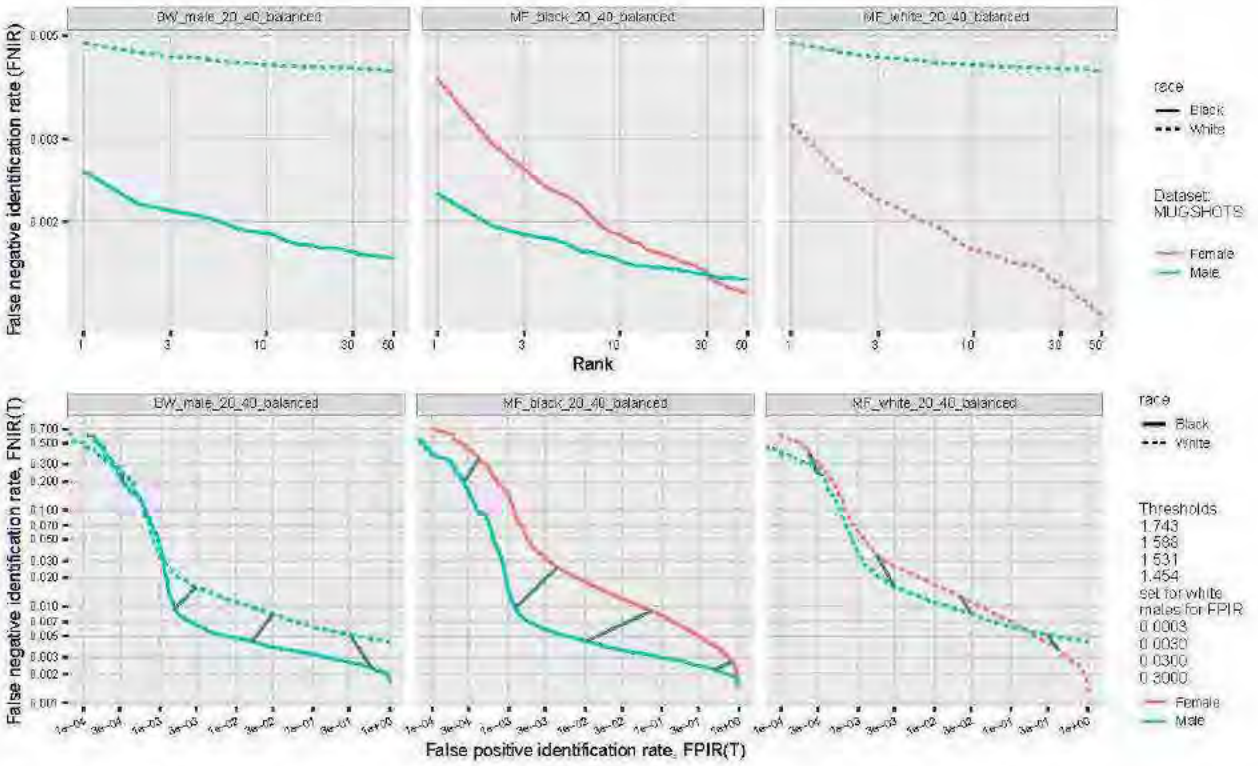


Figure 24: For mugshot identification, the top row shows false negative identification "miss" rates as a function of rank, a metric appropriate to investigators traversing candidate lists for lead generation. The bottom row shows miss rates as a function of false positive identification rate, where a threshold is swept from a low value on the right to high values on the left. This metric is appropriate to organizations for which the volume of searches is high enough that they cannot afford labor to review results from every search. The left panels show the effect of race in young men. The center and right panels show difference between men and women, in black then white subjects respectively. The grey lines join points of equal threshold. The four thresholds are chosen to give FPIR of $\{0.0003, 0.003, 0.03, 0.3\}$ respectively for one baseline demographic, here white males. The figure applies to one algorithm, provided to NIST in August 2019. The corresponding figures for all identification algorithms appear in Annex 16 .

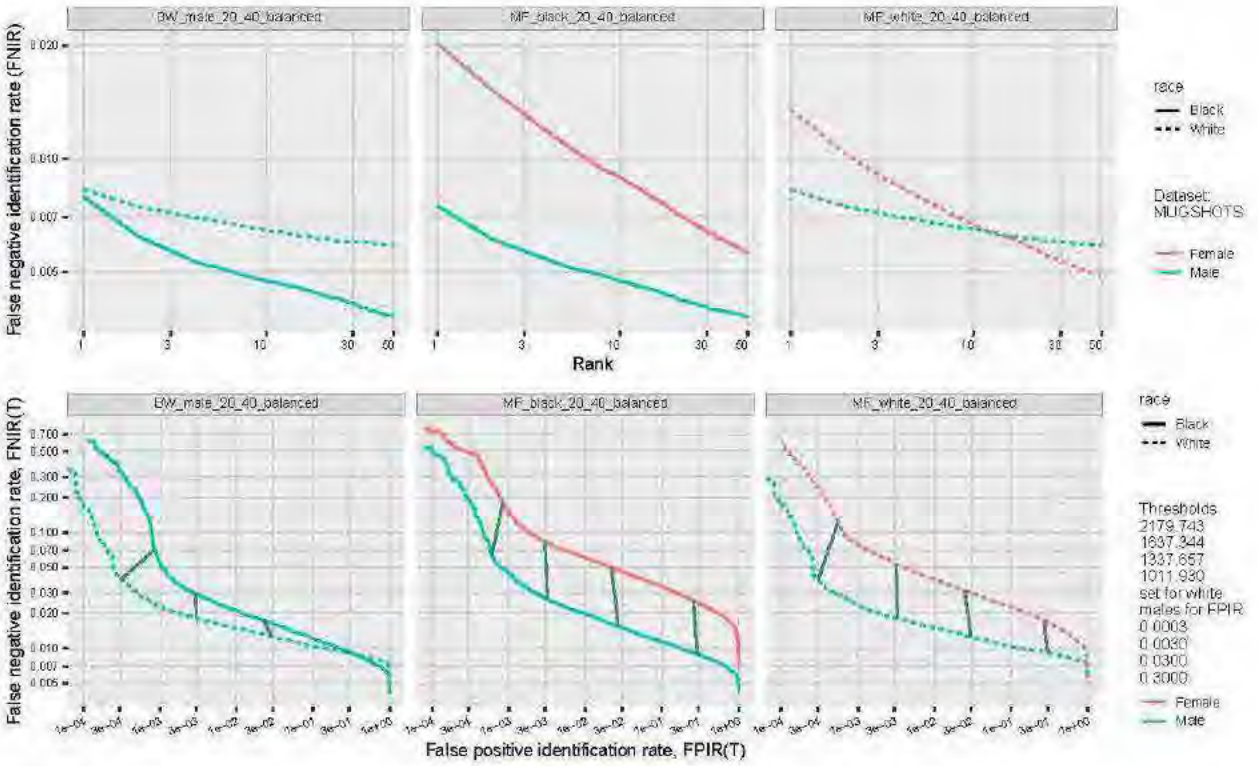


Figure 25: For mugshot identification, the top row shows false negative identification "miss" rates as a function of rank, a metric appropriate to investigators traversing candidate lists for lead generation. The bottom row shows miss rates as a function of false positive identification rate, where a threshold is swept from a low value on the right to high values on the left. This metric is appropriate to organizations for which the volume of searches is high enough that they cannot afford labor to review results from every search. The left panels show the effect of race in young men. The center and right panels show difference between men and women, in black then white subjects respectively. The grey lines join points of equal threshold. The four thresholds are chosen to give FPIR of $\{0.0003, 0.003, 0.03, 0.3\}$ respectively for one baseline demographic, here white males. The figure applies to one algorithm, provided to NIST in June 2018. The corresponding figures for all identification algorithms appear in Annex 16 .

7 False positive differentials in identification

The section addresses whether identification algorithms exhibit similar false positive differentials to verification algorithms. We first note that large-scale one-to-many identification deployments typically operate at false match rates much lower than those targeted in verification applications. It is typical in verification access control to target false match rates (FMR) between 0.00001 and 0.001, i.e. between one per hundred thousand and one per thousand. Identification applications, howveer, often enroll very large numbers of individuals numbering into the 10s or 100s of millions. If such systems are configured with thresholds aimed at producing false positive outcomes say one in 100 times, i.e. FPIR = 0.01, then the implied likelihood that a comparison will yield a false match is given by this formula

$$\text{FMR} = \frac{FPIR}{N} \tag{10}$$

where N is the size of the enrolled population. With FPIR= 0.01, and N= 10^6s this formula implies FMR= 10^{-8}. The formula gives a first order equivalence of identification with verification: the former needs low false positive rates in large galleries. Metrics are discussed in section 3.

Some one-to-many search algorithms implement a 1:N search of a probe image as N 1:1 comparisons of the probe with the N enrolled items. This is followed by a sort operation which yields N candidates sorted in decreasing order of similarity. The result of that is returned in either of two ways: The system will return an operator-specified number of candidates, or it will return however many candidates are above an operator-specified threshold[15]. In the case where a threshold is used, the number of candidates returned will be a random-variable that is dependent on the image data itself.

Other algorithms do not implement 1:N search as N 1:1 comparisons. Instead they might employ a set of fast-search algorithms aimed at expediting search [2, 19, 21, 26]. These include various techniques to partition the enrollment data so that far fewer than N comparisons are actually executed. However, this does not mean that false positive occurences will be reduced because the algorithms are still tasked with finding the most similar enrollments.

For the three experiments listed in section 6, Figure 26 shows median scores returned by one identification algorithm when non-mated searches are conducted. It is clear that if a threshold is applied there will be demographic differences in the number of candidates returned, and in the score values. Such behavior applies to many algorithms - see Annex 17 .

This effect disappears in the algorithm featured in Figure 27. This is an important result because it implies much more equitable likelihoods of false positives. This is especially important result in negative identification

[15] The "operator-specified" parameters might sometimes be set by-policy, or by the manufacturer of the system.

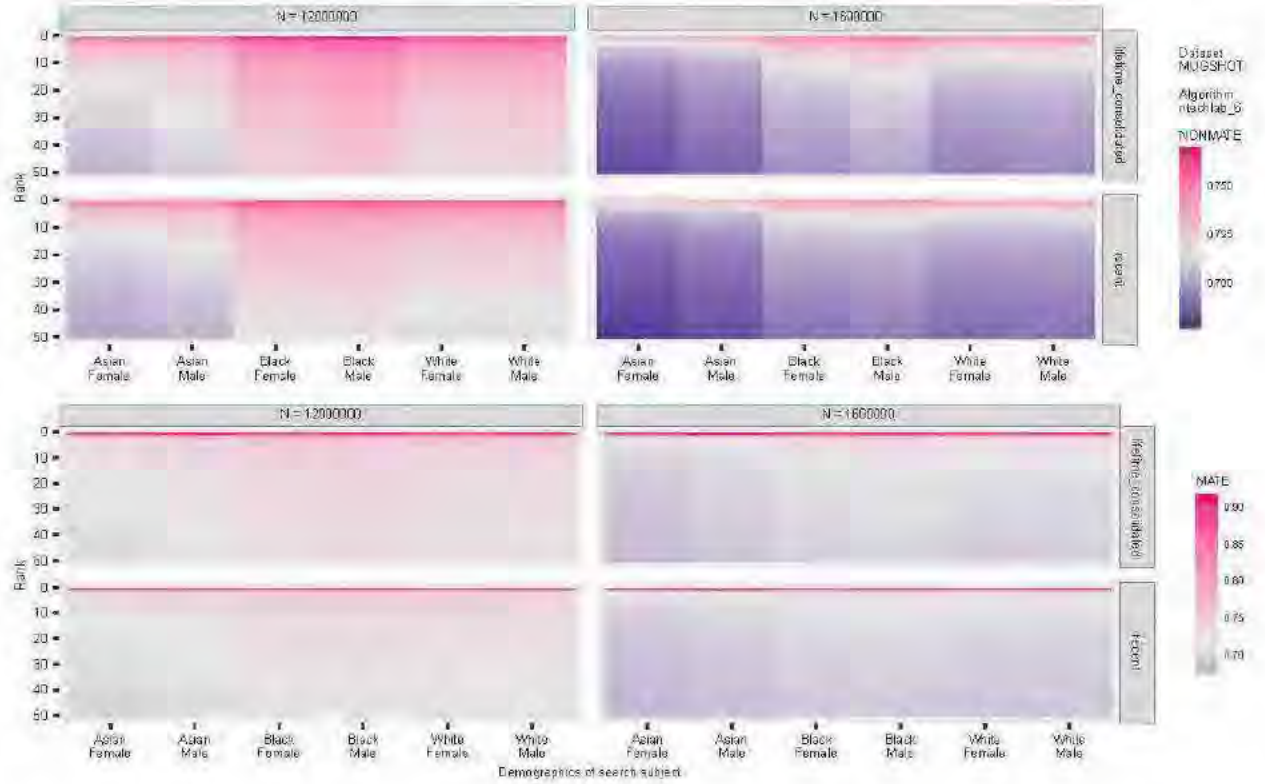


Figure 26: For searches of Asian, black, white men and women's faces into mixed galleries of mugshot photos the heatmaps show median similarity scores for candidates placed at rank 1 to 50. The upper four panels are produced in nonmated searches; the lower four from mated searches. The left-side panels are produced from searches into galleries with 12 000 000 people enrolled. The right-side uses galleries with N = 1 600 000 enrolled. The "lifetime consolidated" and "recent" labels refer to inclusion of multiple images per person, or just one - see [17].
Contrast the behavior here with that in Figure 27 and the corresponding figures for developers Aware, Idemia, NEC, Tevian, and Toshiba that are included in Annex 17 .

applications where the prior probability of a searched person actually being is in the database is low, e.g. card-sharp surveillance in a casino, or soccer hooligans at a sports game[16]. The lack of an effect on false positive identification rates is evident in Figure 25 where the grey lines join points of equal threshold. From left-to-right, the FPIR values for black and white males, black men and women, and white men and women are closely similar. The more normal behavior (see Figure 24 and Annex 16) is for larger shifts in false positive rates.

We now consider the implications for **investigative "lead generation" applications**. In such cases, algorithms return a fixed number of candidates and human reviewers compare the probe photo alongside each candidate gallery photo to determine if the photos are a match. In mugshot-mugshot searches the reviewer will very often look no further than rank 1 per the very high accuracy results documented in NIST Interagency Report

[16] For example, a recent news article noted the use of automated face recognition to search around 21 000 spectators at soccer games against a watch-list of about 50 people.

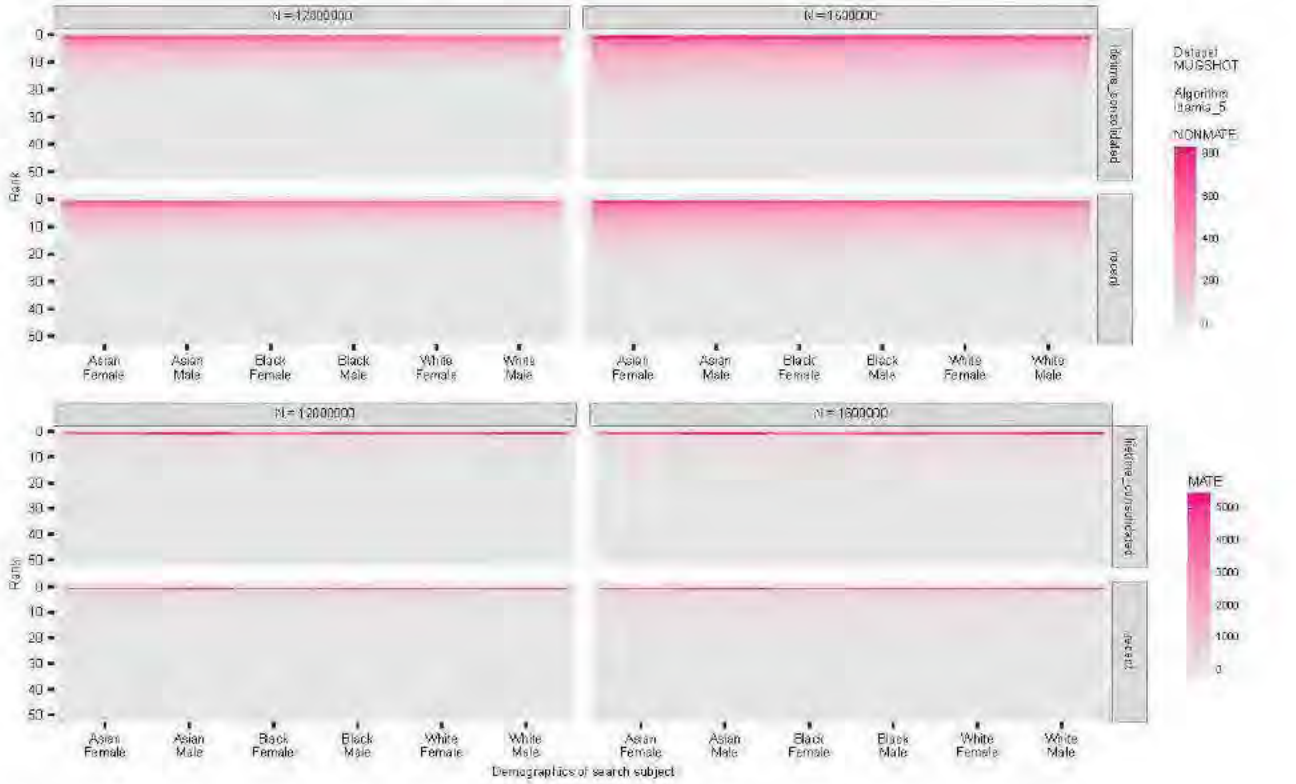


Figure 27: For searches of Asian, black, white men and women's faces into mixed galleries of mugshot photos the heatmaps show median similarity scores for candidates placed at rank 1 to 50. The upper four panels are produced in nonmated searches; the lower four from mated searches. The left-side panels are produced from searches into galleries with 12 000 000 people enrolled. The right-side uses galleries with N = 1 600 000 enrolled. The "lifetime consolidated" and "recent" labels refer to inclusion of multiple images per person, or just one - see [17]. The uniformity of the scores across demographic groups is in contrast to that evident in Figure 26 and many others in the Annex 17 compendium.

8271. That report also includes a workload measure summarizing the expected number of candidates that will have to be reviewed before a mate is located. A very important parameter in such applications, however, is the prior probability that a mate is actually present. In boarding a cruise ship for example, almost everyone attempting to board would be present in the gallery. In a casino application aimed at detecting "high rollers" the likelihood a patron of the casino is in that set is much lower. In such cases a human reviewer, if so employed, would in most searches review all say 50 candidates on the list. That's laborious and may not be tenable from an operations research perspective due to fatigue and reward factors in humans.

But in whatever circumstances human reviewers are tasked with reviewing candidate lists, how are demographic differentials such as those in Figure 26 expected to influence the human? The human will see fifty candidates regardless. However, if those candidates are accompanied by scores, presented as text in a GUI for example, the reviewer will see higher scores in the black female population and potentially elsewhere. Over time this may influence the human, though one earlier study [12] looked at cognitive bias issues in the human review of fingerprint search results, without demographic effects, and found scant evidence that scores influence the reviewer. That study did, however, find that just the order in which candidates are presented to reviewers affects both false positives and false negatives. For example, reviewers are more likely to miss (i.e. a false negative) a mated candidate that appears far down the candidate list. The issues involved in human review are beyond the scope of this document, but full consideration of *systems* comprised of automated face search algorithms and human reviewers is an experimental psychology, human factors and operations research issue.

8 Research toward mitigation of false negatives

False negative error rates, and demographic differentials therein, are reduced in standards-compliant images. This motivates the following two research and development possibilities.

- **Improved standards compliance:** The ISO/IEC 19794-5 standard includes requirements regulating geometry and exposure. Recent research [24] noted that higher quality images, as determined by an automated quality assessment algorithm, yields a reduced false negative differential. While commercial packages exist for the automated assessment of quality, and NIST has an ongoing assessment of the underlying algorithms, rejection of single images on quality grounds can itself have demographic problems [1]. The ISO/IEC SC 37 biometrics subcommittee has recently initiated work on quality (ISO/IEC 29794-5 and 24357).

- **Face-aware cameras:** The same ISO/IEC committee has recently initiated work on specifications for capture subsystems that may require real-time face detection, pose estimation, and exposure measurement. Analogous "auto-capture" quality control mechanisms exist in iris and fingerprint scanners. That standard, ISO/IEC 24358, will be developed through 2020 with completion expected in 2021. Participation is open via national standardization groups.

  Along similar lines further research into automated image quality assessment, and particularly specifications for closed-loop face-aware capture would prove valuable in averting low-contrast and over- and under-exposed images. Many enrollment operations still rely on documentary photography standards with cameras that are not detecting and metering off faces.

This work would be supported by research into two further topics:

Analysis: There is a need for improved models of demographic effects, particularly to how subject-specific properties including phenotypes, imaging artefacts and algorithms interact. Such models would extend work [9] in separating the relative contributions of at least, sex, age, race and height. Efforts to automatically estimate phenotypic information from images will involve algorithms that may themselves exhibit demographic differentials. Such work will need to address this possibility.

Information theoretic analysis: Given the potential for poorly illuminated photographs to produce false negatives, via under- or over-exposure of dark or light skin, an information theoretic approach to characterize algorithmic response to poor lighting would be useful for future standardization. In particular, the ISO/IEC 19794-5 standard has, since 2004, required portrait photos to have at least 7 bits of content in each color channel. Such work should quantify both false negative and false positive dependence.

9 Research toward mitigation of false positives

9.1 Summary

The threshold manipulation strategies described above would be irrelevant if the algorithm developer provided software with homogeneous false match rates. That will prove impossible as there will always be some distribution around a mean - the goal should be much more homogeneous false match rates than is currently the case.

9.2 Algorithm training

A longer-term mitigation is prompted by our observation that many algorithms developed in China do not give the elevated false positive rates on Chinese faces that algorithms developed elsewhere do. This affirms a prior finding of an "other-race effect" for algorithms [33] though that paper did not separate false positive from false negative shifts. This suggests that training data, or perhaps some other factor intrinsic to the development, can be effective at reducing particular false positive differentials. Thus, the longer-term mitigation would be for developers to investigate the utility of more diverse, globally derived, training data. Absent such data, developers might consider whether their cost functions can be altered to reduce differentials. One developer advanced such a concept in November 2018 [15].

9.3 Greater discriminative power

Face recognition algorithms measure similarity between face images. Facial appearance is partially determined by genes, the phenotypic expression of which determines skin tone and a large set of characteristics related to shape of the face. In NIST recognition tests [17], identical twins invariably cause false positives at all practical operational thresholds. Twins are characterized by very similar features given identical genes. Similarities in faces in fraternal twins [17] are expected to extend also to siblings (which also share half of the genes), and then to more distant relatives. In 2004, an algorithm was patented that can correctly distinguish twins [US Patent: US7369685B2]; it operates by extracting features from skin texture (adjacent to the nose, and above the eyebrows). This algorithm requires high resolution and, moreover, knowledge that any given image has that resolution. However, contemporary deployments of face recognition are very often based on processing of images at or below VGA spatial sampling rates (i.e., 480 x 640 pixel images), and this is often insufficient for skin texture to be viable. The human reviewer community has long specified much higher resolution for forensic purposes (see ANSI/NIST Face Acquisition Profiles).

9.4 Collection and use of face and iris

The texture of the human iris is known to have a structure that when imaged and processed by published feature extraction algorithms [11, 29] will correctly discriminate between identical twins [40] - something that contemporary marketplace face algorithms do not [17]. The reason for this appears to be that the iris features detected by automated algorithms are not genetically determined. However genetics research [25] does show iris textures have some genetic linkage, so a better characterisation of the tails of the impostor distribution is needed, at least for large scale one-to-many identification. Nevertheless, a 2019 DHS Science and Technology study noted that false positives are no higher within individuals of the same sex, age and race as they are across those groups [39]. As shown in Figure 4 and Annex 8 that is not the case for face recognition. NIST has near-term plans to investigate the impostor distribution in twins more fully.

Given the marketplace presence of multiple cameras that collect face and iris essentially simultaneously, one approach to consider for mitigation of false positive differentials in face recognition would be for face records to include adjunct iris images. The standards infrastructure is in place for this (ANSI/NIST Type 17, ISO/IEC 39794-6, and ICAO 9303 Data Group 4). This would afford very low false positive rates.

The apparent lack of genetic influence, and demonstrated low false match likelihoods, has been the primary property in establishing the use of the iris for the identification of individuals in large populations - most notably in the Indian National ID program Aadhaar. The iris recognition industry has multiple camera developers, multiple algorithm suppliers, and image interchange and quality standards that support interoperable recognition across cameras.

These aspects afford solutions to higher and heterogeneous false positive rates in face recognition. The first is simply to replace face with iris. There would be advantages and disadvantages to this - detailing and weighing those is beyond our scope here. However a second solution would be to augment face with iris, to produce a compound biometric "face-and-iris"[17]. This is made possible by the marketplace availability for at least a decade now of cameras that collect iris and face images essentially simultaneously. Recognition of the combined biometric would involve a particular kind of biometric fusion that in which both the face and iris must match (against respective thresholds) so as to limit false positives. This differs from some convenience-driven implementations that authenticate a person with either face or iris alone.

Use of iris in some applications, for example surveillance, is limited by the difficulty and expense of imaging the iris at long distances.

We don't mention fingerprints in this context because even though genetic influence is considered to be absent

[17]Such a compound biometric would conventionally still require collection of two images: First an iris image with near infrared illumination and the face image either entirely in ambient light, or ambient light with a near infrared component. The recognition of irises in purely visible-light images is highly problematic in brown-eyed people as melanin in the iris absorbs incident light at visible wavelengths.

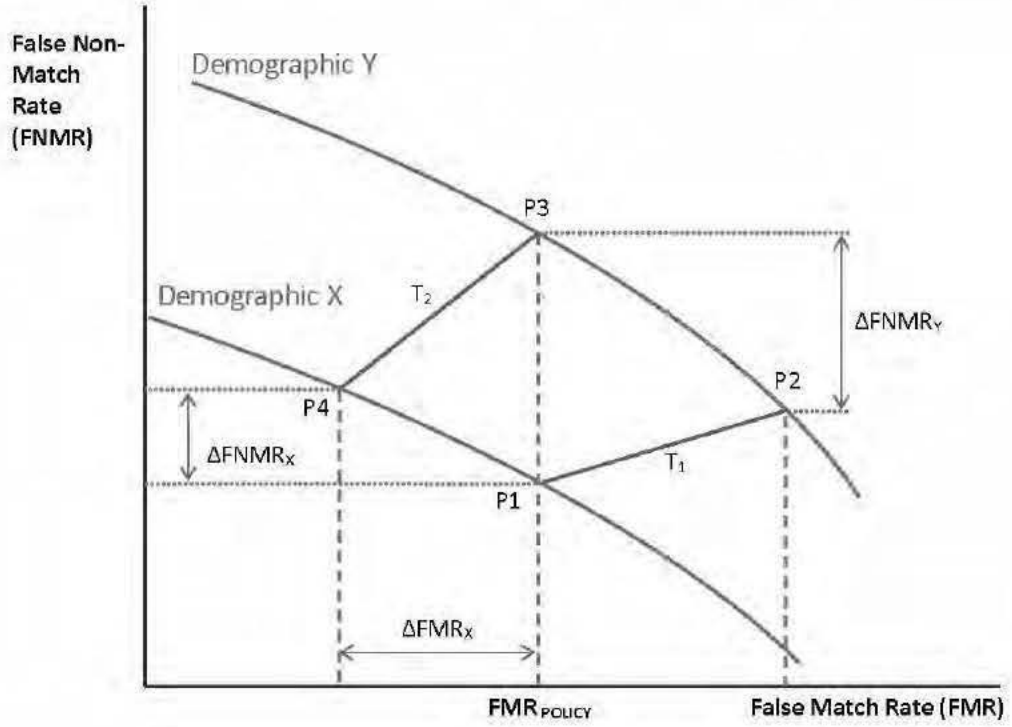


Figure 28: The figure shows the increases in FNMR implied by increasing the operating threshold to achieve the target FMR on the high-FMR demographic, Y.

or minimal, the collection of both fingerprint and face is not simultaneous.

9.5 Threshold elevation

We detail one mitigation of heterogeneous variable false match rates, and its consequences, as follows. The explanation uses a graphical construct based on the error tradeoff characteristics shown in Figure 28.

- We start with a target false match rate $\mathrm{FMR}_{\mathrm{POLICY}}$ that has been set to implement some security objective. This value, in a verification application might reasonably be set to say 1 in 5000 (i.e. 0.0002). This is implemented by setting a threshold T_1. Suppose that this threshold was perfectly calibrated for Demographic X i.e. $\mathrm{FMR}(T_1) = \mathrm{FMR}_{\mathrm{POLICY}}$. This corresponds to the point P1.
- Now suppose that we later discover, perhaps as a result of some biometric performance test or audit that, for some new group Demographic Y, that the observed false match rate at the fixed threshold T_1 is much higher, a factor of five say (0.001). This point P2 therefore represents therefore a failure to meet the original security objective for that group.
- To bring the overall system into policy compliance, the system owner consults the error tradeoff charac-

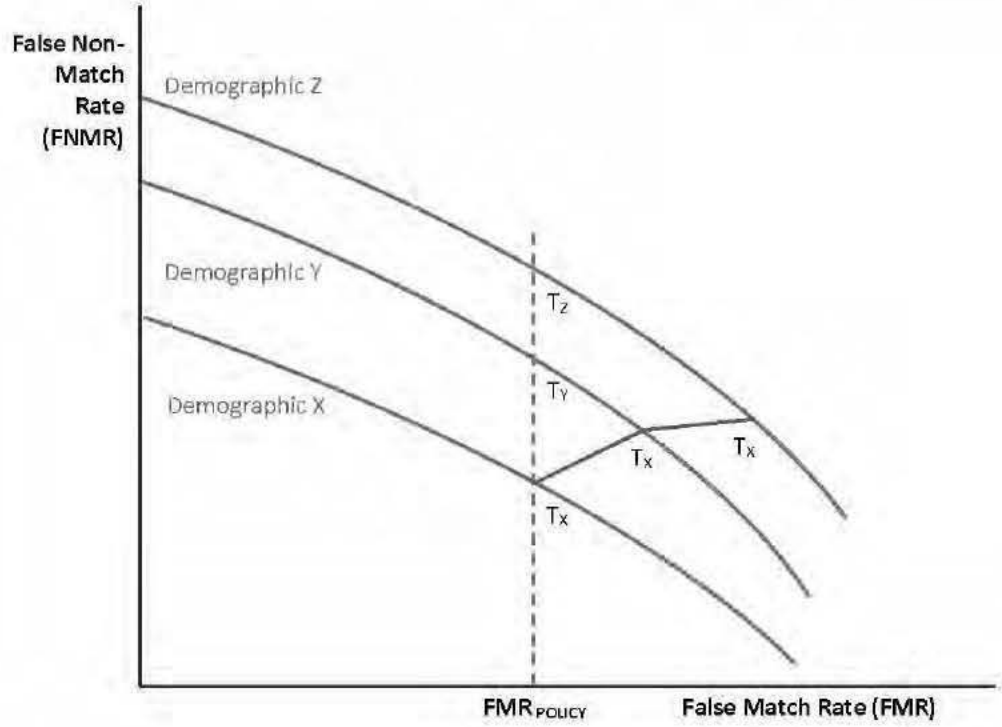


Figure 29: The figure shows the effect of setting thresholds to achieve the target FMR on demographics X and Y.

teristic for Demographic Y and notes that by elevating the threshold to T_2, the false match rate would be returned to policy compliance, at point P3.

- The effect of this however is that FNMR is necessarily elevated both demographic groups. This is because the new threshold T_2 is higher than T_1, and applies to all transactions from all demographics. These increases are shown as $\Delta FNMR_X$ and $\Delta FNMR_Y$ would have a magnitude that depends on the gradients of the error-tradeoff characteristics (which may differ). The only gain is a reduction in FMR for Demographic X, to a value which beats the original target policy.

Using this kind of construct, we see the benefit in having a biometric algorithm for which false match rates are homogenous i.e. do not vary (much) over any demographics.

The above argument assumes that the original high $FMR_Y(T_1)$ is indeed problematic. It may be tolerable in cases where individuals in that Demographic are rare, e.g. elderly persons entering a gym or nightclub. Any decision to not elevate the threshold to T_2 should be deliberated in the security context defined by threat, risk and cost.

9.6 Explicit thresholds for each demographic

In this section we discuss the suggestion [23] to address heterogeneous false match rates by assigning a threshold to each demographic. The proposal is for a verification system to set the threshold each time a subject executes a verification transaction tailoring it on the basis of who is using the system. Referencing Figure 29, this would correspond to adopting thresholds T_1 and T_2 (i.e. points P1 and P3) on-the-fly. How to do this presents a problem. Naively one could encode in an identity document (e.g. a passport) some indication of the demographic group (e.g. female, middle aged, south Asian) and the system would read this information, consult a lookup table, and set T accordingly. This would be effective for genuine legitimate users of the system. The security consequences of this are, however, more complicated. Consider what an imposter would do given knowledge that thresholds are variable.

- If the imposter were from a demographic for which the threshold is low, he would procure / steal a credential from somebody of the same age, sex and ethnicity. This would be typical behavior for any imposter. However, if particular countries passports were known to be used with low-thresholds, we'd expect genesis of a black-market for stolen credentials in those places.
- If the imposter were from a demographic for which the threshold is high he might procure / steal a credential from somebody in one of the low-threshold demographics, matching age and sex minimally the same sex. To better induce a false match the imposter would still need to have the same age, sex and ethnicity. This would be typical behavior anyway.

  Note that societal construction will often naturally afford opportunities for imposters to have access to identity credentials from other persons who, naturally, have the same ethnicity, sex and age group.

Another aspect to this approach is that it shifts responsibility for threshold management to the system owner rather than the developer. That may sound fully appropriate but imposes two responsibilities on the operator: First, figuring out what the thresholds should be via some appropriate testing, and secondly to implement the strategy with capture of demographic information and use of that in software.

References

[1] December 2016. https://www.telegraph.co.uk/technology/2016/12/07/robot-passport-checker-rejects-asian-mans-photo-having-eyes/.

[2] Artem Babenko and Victor Lempitsky. Efficient indexing of billion-scale datasets of deep descriptors. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2016.

[3] L. Best-Rowden and A. K. Jain. Longitudinal study of automatic face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40(1):148–162, Jan 2018.

[4] J. Ross Beveridge, Geof H. Givens, P. Jonathon Phillips, and Bruce A. Draper. Factors that influence algorithm performance in the face recognition grand challenge. *Computer Vision and Image Understanding*, 113(6):750–762, 2009.

[5] Joy Buolamwini. Gender shades: Intersectional phenotypic and demographic evaluation of face datasets and gender classifiers. Technical report, MIT Media Lab, 01 2017.

[6] J. Campbell and M. Savastano. Iso/iec 22116 identifying and mitigating the differential impact of demographic factors in biometric systems. Technical report, ISO/IEC JTC 1, SC 37, Working Group 6, http://iso.org/standard/72604.html, 11 2018.

[7] Jacqueline G. Cavazos, Eilidh Noyes, and Alice J. O'Toole. Learning context and the other-race effect: Strategies for improving face recognition. *Vision Research*, 157:169 – 183, 2019. Face perception: Experience, models and neural mechanisms.

[8] Jacqueline G. Cavazos, P. Jonathon Phillips, Carlos D. Castillo, and Alice J. O'Toole. Accuracy comparison across face recognition algorithms: Where are we on measuring race bias? In *https://arxiv.org/abs/1912.07398*, 12 2019.

[9] Cynthia Cook, John Howard, Yevgeniy Sirotin, Jerry Tipton, and Arun Vemury. Demographic effects in facial recognition and their dependence on image acquisition: An evaluation of eleven commercial systems. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, PP:1–1, 02 2019.

[10] White D., Kemp R. I., Jenkins R., Matheson M, and Burton A. M. Passport officers errors in face matching. *PLoS ONE*, 9(8), 2014. e103510. doi:10.1371/journal. pone.0103510.

[11] J. Daugman. How iris recognition works. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1):21–30, Jan 2004.

[12] Itiel Dror and Kasey Wertheim. Quantified assessment of afis contextual information on accuracy and reliability of subsequent examiner conclusions. Technical Report 235288, National Institute of Justice, July 2011.

[13] H El Khiyari and Wechsler H. Face verification subject to varying (age, ethnicity, and gender) demographics using deep learning. *Journal of Biometrics and Biostatistics*, 7:323, 11 2016. doi:10.4172/2155-6180.1000323.

[14] C. Garvie, A. Bedoya, and J. Frankle. The perpetual line-up: Unregulated police face recognition in america. Technical report, Georgetown University Law School, Washington, DC, 10 2018.

[15] Stéphane Gentric. Face recognition evaluation @ idemia. In *Proc. International Face Performance Conference, National Institute of Standards and Technology NIST, Gaithersburg, MD*, November 2018.

[16] Patrick Grother, Mei Ngan, and Kayee Hanaoka. Face recognition vendor test (frvt) part 1: Verification. Interagency Report DRAFT, National Institute of Standards and Technology, October 2019. https://nist.gov/programs-projects/frvt-11-verification.

[17] Patrick Grother, Mei Ngan, and Kayee Hanaoka. Face recognition vendor test (frvt) part 2: Identification. Interagency Report 8271, National Institute of Standards and Technology, September 2019. https://doi.org/10.6028/NIST.IR.8271.

[18] Patrick Grother, George W. Quinn, and Mei Ngan. Face recognition vendor test - still face image and video concept, evaluation plan and api. Technical report, National Institute of Standards and Technology, 7 2013. http://biometrics.nist.gov/cs_links/face/frvt/frvt2012/NIST_FRVT2012_api_Aug15.pdf.

[19] Feng Hao, John Daugman, and Piotr Zielinski. A fast search algorithm for a large fuzzy database. *IEEE Transactions on Information Forensics and Security*, 3(2):203–212, 2008.

[20] John J. Howard, Yevgeniy Sirotin, and Arun Vermury. The effect of broad and specific demographic homogeneity on the imposter distributions and false match rates in face recognition algorithm performance. In *Proc. 10-th IEEE International Conference on Biometrics Theory, Applications and Systems, BTAS 2019, Tampa Florida,USA*, September 2019.

[21] Masato Ishii, Hitoshi Imaoka, and Atsushi Sato. Fast k-nearest neighbor search for face identification using bounds of residual score. In *2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, pages 194–199, Los Alamitos, CA, USA, May 2017. IEEE Computer Society.

[22] B. F. Klare, Burge M. J., Klontz J. C., Vorder Bruegge R. W., and Jain A. K. Face recognition performance: Role of demographic information. *IEEE Trans. on Information Forensics and Security*, 7(6):1789–1801, 9 2012.

[23] K. S. Krishnapriya, Kushal Vangara, Michael C. King, Vitor Albiero, and Kevin Bowyer. Characterizing the variability in face recognition accuracy relative to race. *CoRR*, abs/1904.07325, 2019. http://arxiv.org/abs/1904.07325.

[24] K. S. Krishnapriya, Kushal Vangara, Michael C. King, Vitor Albiero, and Kevin Bowyer. Us study: better image quality could cut face system bias. *Biometric Technology Today*, 2019(5):11 – 12, 2019.

[25] Mats Larsson, David L. Duffy, Gu Zhu, Jimmy Z. Liu, Stuart Macgregor, Allan F. McRae, Margaret J. Wright, Richard A. Sturm, David A. Mackey, Grant W. Montgomery, Nicholas G. Martin, and Sarah E. Medland. GWAS findings for human iris patterns: Associations with variants in genes that influence normal neuronal pattern development. *American Journal of Human Genetics*, 89(2):334–343, August 2011.

[26] Yury A. Malkov and D. A. Yashunin. Efficient and robust approximate nearest neighbor search using hierarchical navigable small world graphs. *CoRR*, abs/1603.09320, 2016.

[27] Joyce A. Martin, Brady E. Hamilton, Michelle J.K. Osterman, Anne K. Driscoll, , and Patrick Drake. National vital statistics reports. Technical Report 8, Centers for Disease Control and Prevention, National Center for Health Statistics, National Vital Statistics System, Division of Vital Statistics, November 2018.

[28] Dana Michalski, Sau Yee Yiu, and Chris Malec. The impact of age and threshold variation on facial recognition algorithm performance using images of children. In *International Conference on Biometrics*, February 2018.

[29] D. M. Monro, S. Rakshit, and D. Zhang. Dct-based iris recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 29(4):586–595, April 2007.

[30] Vidya Muthukumar. Color-theoretic experiments to understand unequal gender classification accuracy from face images. In *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.

[31] Vidya Muthukumar, Tejaswini Pedapati, Nalini Ratha, Prasanna Sattigeri, Chai-Wah Wu, Brian Kingsbury, Abhishek Kumar, Samuel Thomas, Aleksandra Mojsilovic, and Kush R. Varshney. Understanding unequal gender classification accuracy from face images. *CoRR*, abs/1812.00099, November 2018.

[32] P. Jonathon Phillips, J. Beveridge, Bruce Draper, Geof Givens, Alice O'Toole, David Bolme, Joseph Dunlop, Yui Lui, Hassan Sahibzada, and Samuel Weimer. The good, the bad, and the ugly face challenge problem. *Image and Vision Computing*, 30:177185, 03 2012.

[33] P. Jonathon Phillips, Fang Jiang, Abhijit Narvekar, Julianne Ayyad, and Alice J. O'Toole. An other-race effect for face recognition algorithms. *ACM Trans. Appl. Percept.*, 8(2):14:1–14:11, February 2011.

[34] P. Jonathon Phillips, Amy N. Yates, Ying Hu, Carina A. Hahn, Eilidh Noyes, Kelsey Jackson, Jacqueline G. Cavazos, Géraldine Jeckeln, Rajeev Ranjan, Swami Sankaranarayanan, Jun-Cheng Chen, Carlos D. Castillo, Rama Chellappa, David White, and Alice J. O'Toole. Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms. *Proceedings of the National Academy of Sciences*, 115(24):6171–6176, 2018.

[35] P.J. Phillips, P. Grother, R. J. Micheals, D. M. Blackburn, E. Tabassi, and M. Bone. Face recognition vendor test 2002. Evaluation Report IR 6965, National Institute of Standards and Technology, www.itl.nist.gov/iad/894.03/face/face.html or www.frvt.org, March 2003.

[36] Inioluwa Raji and Joy Buolamwini. Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. In *Conference on AI, Ethics and Society*, pages 429–435, 01 2019.

[37] K. Ricanek and T. Tesafaye. Morph: a longitudinal image database of normal adult age-progression. In *7th International Conference on Automatic Face and Gesture Recognition (FGR06)*, pages 341–345, April 2006.

[38] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition. In *Proc. International Conference on Learning Representations*, volume https://arxiv.org/abs/1409.1556v6, 2015.

[39] Yevgeniy Sirotin. A comparison of demographic effects in face and iris recognition. Technical report, Iris Experts Group, Gaithersburg, MD, 6 2019.

[40] Zhenan Sun, Alessandra Paulino, Jianjiang Feng, Zhenhua Chai, Tieniu Tan, and Anil Jain. A study of multibiometric traits in identical twins. *Proc. of the International Society of Optical Engineering*, 7667, 04 2010.

[41] Darrell M. West. 10 actions that will protect people from facial recognition software. Technical report, Brookings Institution, Artificial Intelligence and Emerging Technology Initiative, Washington, DC, 10 2019.

[42] David White, James D. Dunn, Alexandra C. Schmid, and Richard I. Kemp. Error rates in users of automatic face recognition software. *PLoS ONE*, October 2015.

NISTIR 8381

Face Recognition Vendor Test (FRVT)
Part 7: Identification for Paperless Travel and Immigration

Patrick Grother
Austin Hom
Mei Ngan
Kayee Hanaoka
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8381

National Institute of
Standards and Technology
U.S. Department of Commerce

NISTIR 8381

Face Recognition Vendor Test (FRVT)
Part 7: Identification for Paperless Travel and Immigration

Patrick Grother
Austin Hom
Mei Ngan
Kayee Hanaoka
Information Access Division
Information Technology Laboratory

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8381

July 2021

U.S. Department of Commerce
Gina M. Raimondo, Secretary

National Institute of Standards and Technology
James K. Olthoff, Performing the Non-Exclusive Functions and Duties of the Under Secretary of Commerce for Standards and Technology & Director, National Institute of Standards and Technology

DISCLAIMER

Specific hardware and software products identified in this report were used in order to perform the evaluations described in this document. In no case does identification of any commercial product, trade name, or vendor, imply recommendation or endorsement by the National Institute of Standards and Technology, nor does it imply that the products and equipment identified are necessarily the best available for the purpose.

INSTITUTIONAL REVIEW BOARD

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 Department of Commerce Regulations, 15 CFR 27, also known as the Common Rule for the Protection of Human Subjects (45 CFR 46, Subpart A).

ACKNOWLEDGMENTS

The authors are grateful for the support and collaboration of the U.S. Customs and Border Protection (CBP) component of the Department of Homeland Security.

Additionally we are indebted to staff at DHS' Science & Technology Directorate (S&T) and Office of Biometric Identity Management (OBIM) for discussions and image data that supports this work.

RELEASE NOTES

This report will be updated periodically with results for new algorithms, new analyses, and new datasets as they become available.

The report is open for comment - correspondence should be directed to frvt@nist.gov.

Executive Summary

We investigate the use of one-to-many facial recognition in airport transit settings in which travelers' faces are matched against galleries of individuals expected to be present. We primarily consider the case where face recognition serves double-duty for access control (to an aircraft) and facilitation (of recording a visa-holder's departure from a country). This is done in a paperless mode in which a boarding pass (something you have) is replaced with presentation of a biometric (something you are) to a camera, representing an implicit claim to be entitled to board. We describe how such systems can fail, discussing errors during gallery creation, photo capture at boarding, attack detection, and face matching. We discuss how errors might be estimated, citing relevant standards, and their consequences.

We quantify face matching errors by simulating departing flights, populating galleries with an airport ENTRY photo of 420 travelers, then measuring accuracy by running searches of EXIT photos. We repeat this with galleries populated with multiple photos per person, and with galleries as large as 42000, modelling the same concept of operations but at a centralized airport checkpoint. We report that accuracy varies greatly across algorithms, that use of multiple images per person reduces errors considerably, and that error rates when searching 42000-person galleries are often three times higher than in 420-person galleries, but still sometimes below 1%. We consider demographics, and note that for the more accurate algorithms, error rates are so low that accuracy variations across sex and race are insignificant. We include additionally a discussion of how our accuracy estimates might differ from those measured operationally due to by factors that we could not control, such as camera type and imaging environment.

Technical Summary

Background: One-to-many biometric search systems are discussed in their role of positive and negative identification - the former refers to the expectation that person in a probe sample is present in the database (as in access to an office) while the latter presumes the person is not (as in compulsive gamblers entering a casino). The distinction is useful because the applications differ in their tolerance for false negatives and false positives. This report addresses the positive use of one-to-many facial recognition in airport transit settings in which travelers' faces are matched against galleries of individuals expected to be present. We primarily consider the case where face recognition serves double-duty for access control (to an aircraft) and facilitation (of recording a visa-holder's exit).

In late 2018 the United States commenced a pilot of face-based confirmation of departure system in which passengers boarding an aircraft make cooperative presentations to a camera and the captured photos are immediately searched against a gallery comprised of photos of persons expected on the flight. This process is intended to biometrically bind the traveler to the departure. A positive biometric match in used two ways: First, by the airline, to grant access to the aircraft in lieu of a boarding-pass presentation; second, by passport control authorities to record the departure from the United States of in-scope passengers (e. g. visa holders), notionally replacing the long-standing airline manifest-based biographic process.

Overview: This report summarizes three NIST activities: First, to describe the biometric aspects of the traveler departure application and factors that are expected to affect its performance; second, to document results from running offline simulations in which recent accurate face recognition algorithms are applied to actual ENTRY and EXIT images with the goals of establishing a methodology, estimating accuracy, and exposing some factors that will affect those estimates; third to consider the use of face recognition at other airport touchpoints where higher populations are expected.

EXIT simulations: We simulate traveler EXIT by preparing 567 galleries each containing exactly 420 individuals representing the population expected on a flight. The individuals are not selected by age or sex. They are selected to have the same region of travel document (for example, South America or East Asia). We search each gallery with a fixed set of actual 132 931 EXIT photos using recent commercial one-to-many face recognition search engines. Notably we do not have camera, location and timestamp metadata so we cannot "replay" biometric boarding of actual flights. We

include this and other caveats in section 5.

Algorithms: Our EXIT simulations make use of one-to-many search algorithms submitted to NIST's ongoing Face Recognition Vendor Test between mid 2018 and April 2021. These algorithms are prototypes from the R&D laboratories of commercial developers of face recognition. These include two variants from the incumbent provider to the face matching facility used in the U.S., including the NEC-3 algorithm that was broadly the most accurate algorithm evaluated in 2018 as reported in NIST Interagency Report 8271 [1].

Images: This report makes use of images provided by DHS Office of Biometric Identity Management in May 2019. That collection is comprised of images and limited metadata indicating in which operation the data was collected e.g. airport-entry, pedestrian land entry, or exit. Some images were accompanied by metadata including date of capture, year of birth, sex, and country-of-birth[1]. From that database, this report uses 132 931 EXIT images of 128 384 individuals to search 567 air-ENTRY galleries each represented a departing flight[2]. Those galleries hold images drawn from the 825 976 airport ENTRY images of the 122 387 EXIT individuals who have a prior ENTRY image. The EXIT images were collected in 2018 and the first four months of 2019.

Other content: Section 1 discusses more general error sources and metrics relevant to EXIT and departure, putting matching results into the broader context of aircraft boarding. Section 2 guides readers toward different testing methodologies appropriate to answering a broader range of questions. Section 3 details our simulations and results. Section 4 considers use of one-to-many traveler verification systems (TVS) with a much larger population of N = 42 000 enrollees for use at other airport touchpoints. Importantly, section 5 discusses various reasons that would render the accuracy estimates in this report too high or too low.

Biometric results: We show that as many as 428 of 567 simulated flights each carrying 420 passengers can be boarded using one-to-many face recognition without any false negative errors - see Table 1 column 5. Stated in terms of error rates, this corresponds to at least 99.5% of travelers being able to board with a single presentation to a camera. This is attainable by enrolling a single prior ENTRY image in the galleries and using any of seven 2020-2021 face recognition algorithms - see Table 2 column 5.

For many travelers, multiple prior images can be enrolled in a gallery. Here, if we enroll an average of six prior air-ENTRY images, then the most accurate algorithm will now board 545 of 567 flights without any errors - see Table 1 column 4. Large gains are realized by all algorithms: Now at least 18 developers' algorithms are effective at boarding greater than 99.5% of travelers - see Table 2 column 3.

In 2007, U.S. legislation[3] specified that 97% of travelers' exits should be verified. That requirement can be met with almost all of the algorithms tested here[4]. Note that there are various systematic reasons why such accuracy may not be achieved in practice - see section 5.

In test of late 2018 algorithms [1], the most accurate algorithms on large population mugshot searches were NEC-2 and NEC-3. They remain in the top five on that benchmark today. However, when matching lower quality EXIT to ENTRY images the algorithms are less accurate than a 2018 Microsoft algorithm and many other more recent algorithms. By taking 100 minus the miss percentages in Table 2, NEC-3 correctly identifies 98.7% of individuals enrolled with a single image and 99.0% of those enrolled with images from multiple prior encounters. For the most accurate algorithm, Visionlabs-10, these values are 99.9% and 100% respectively, corresponding to about a factor of 10 fewer errors than NEC-3. Note that NEC-3 is now more than two years old and we may assume NEC has since improved its capability.

[1] This metadata was vital to our 2019 quantification of demographic effects in NIST Interagency Report 8280. [2]
[2] The terms ENTRY and EXIT refer respectively to inbound and outbound border crossings to, in this case, the United States.
[3] See 8 U.S.C. 1187(c).
[4] We chose to run only recent and high-performing algorithms and also some widely used prior-generation algorithms. Many more algorithms have been entered into the 1:N search track of FRVT.

ALGORITHM			NUM ZERO FALSE NEGATIVE SIMULATIONS		
			$N = 420$	$N = 420$	$N = 42000$
#	NAME	DATE	$k \geq 1$	$k = 1$	$k = 1$
1	VISIONLABS-010	2021-02-05	[1]545	[1]428	[3]177
2	IDEMIA-008	2021-03-15	[2]536	[2]422	[2]215
3	VISIONLABS-009	2020-08-04	[3]533	[3]406	[5]125
4	CLOUDWALK-HR-000	2021-02-10	[4]528	[4]393	[1]265
5	DEEPGLINT-001	2020-07-23	[5]519	[5]336	[4]153
6	CANON-CIB-000	2020-10-19	[6]518	[7]307	[13]19
7	XFORWARDAI-001	2021-01-21	[7]513	[6]309	[7]113
8	PARAVISION-007	2021-02-01	[8]490	[8]237	[6]124
9	TRUEFACE-000	2021-01-27	[9]476	[15]154	[15]4
10	NEUROTECHNOLOGY-008	2021-03-26	[10]470	[12]169	[17]2
11	COGENT-004	2021-02-10	[11]454	[11]182	[14]10
12	PARAVISION-005	2019-12-11	[12]453	[13]156	[10]72
13	NTECHLAB-008	2020-01-06	[13]451	[17]125	[20]1
14	PIXELALL-004	2020-07-02	[14]435	[16]146	[23]0
15	TECH5-002	2021-04-07	[15]416	[18]110	[18]2
16	DERMALOG-008	2021-01-25	[16]382	[20]71	[27]0
17	IDEMIA-007	2020-01-17	[17]374	[21]66	[21]0
18	MICROSOFT-006	2018-10-29	[18]361	[14]155	[16]3
19	SENSETIME-005	2020-12-17	[19]319	[9]233	[8]99
20	SENSETIME-004	2020-08-10	[20]316	[10]208	[9]96
21	RANKONE-010	2020-11-05	[21]300	[19]76	[22]0
22	RANKONE-009	2020-06-26	[22]203	[24]38	[19]1
23	COGNITEC-004	2021-03-08	[23]201	[26]11	[28]0
24	NEC-003	2018-10-30	[24]111	[22]66	[12]30
25	NEC-002	2018-10-30	[25]111	[23]65	[11]32
26	NEUROTECHNOLOGY-007	2019-10-03	[26]90	[25]21	[25]0
27	DERMALOG-007	2020-02-12	[27]30	[27]3	
28	IDEMIA-004	2018-06-30	[28]3	[29]0	[24]0
29	NEC-000	2018-06-21	[29]0	[28]0	[26]0

Table 1: **Number of simulations (out of 567) completed without errors.** *The second row N values give the number of individuals enrolled in each gallery. The 420 person galleries represent aircraft boarding; the 42000 case represents a airport security line where many more people are expected. The third row k values give the number of images of each enrollee in each gallery.*
The second and third columns identify the algorithm and the date it was submitted to NIST. The remaining columns give the number of simulations, out of 567, for which all 420 travelers boarded the flight (cols. 4, 5), or passed the checkpoint (column 6), without experiencing a false negative. Higher values are better, and the table is sorted on the first results column. The threshold is set so that only a fraction, 0.0003, of non-mated searches would return any match. The shaded cells indicate the three most accurate algorithms for that trial.

Our demonstration of considerably higher accuracy from newer algorithms is an existence proof that EXIT accuracy on operational images can be improved. Given the pace of developments associated with the industrial migration to various convolutional neural networks, it is incumbent on end-users to establish contractual provisions for technology refreshment, factoring in such quantities as speed, scalability, stability, and cost.

The accuracy values noted above correspond to correct identification of an individual – here "correct" requires the algorithm to report the correct identity with a score above a set threshold. The threshold is set to limit false positives – this is necessary to prevent illicit boarding of an aircraft in an access-control context, and to limit visa-holder's status indicator mistakes in an EXIT facilitation context. The false positive identification rate (FPIR) in this report is usually set to 1 in 3 333, i.e. the proportion of searches of people not entitled to board an aircraft who succeed in doing so. A false positive occurs when a photo from such a traveler matches any (random) gallery photo. The consequences of such events, and a more detailed discussion of security, appears in section 1.5. We also include figures showing the tradeoff of false negative and positive identification rates, noting that some algorithms can afford lower FPIR without greatly degrading accuracy. An FPIR of 1 in 3 333 would imply that a mismatch would occur once during the boarding of about eight flights (3 333/420) – whether that is too frequent or to scarce is essentially policy issue informed by the error tradeoff characteristics of section 3.2.3 and the demographic dependencies given in sections 3.2.4 and 3.2.5.

Discussion: The report documents accuracy or small-gallery identification simulations showing a strong algorithm effect - accuracy is much improved with some algorithms versus others. This dominates two other main effects - first

ALGORITHM			PERCENT TRAVELERS NOT MATCHED		
			$N = 420$	$N = 420$	$N = 42000$
#	NAME	DATE	$k \geq 1$	$k = 1$	$k = 1$
1	VISIONLABS-010	2021-02-05	[1]0.02	[1]0.13	[3]0.61
2	IDEMIA-008	2021-03-15	[2]0.02	[2]0.15	[2]0.49
3	VISIONLABS-009	2020-08-04	[3]0.03	[3]0.16	[5]0.74
4	CLOUDWALK-HR-000	2021-02-10	[4]0.03	[4]0.18	[1]0.43
5	DEEPGLINT-001	2020-07-23	[5]0.04	[5]0.24	[6]0.80
6	CANON-CIB-000	2020-10-19	[6]0.04	[7]0.30	[13]1.79
7	XFORWARDAI-001	2021-01-21	[7]0.05	[6]0.28	[7]0.81
8	PARAVISION-007	2021-02-01	[8]0.07	[8]0.41	[4]0.72
9	TRUEFACE-000	2021-01-27	[9]0.08	[14]0.66	[16]3.72
10	NEUROTECHNOLOGY-008	2021-03-26	[10]0.08	[11]0.59	[18]4.12
11	PARAVISION-005	2019-12-11	[11]0.10	[13]0.62	[10]1.04
12	NTECHLAB-008	2020-01-06	[12]0.11	[17]0.81	[19]4.52
13	COGENT-004	2021-02-10	[13]0.11	[12]0.59	[14]2.34
14	PIXELALL-004	2020-07-02	[14]0.12	[15]0.69	[17]3.88
15	TECH5-002	2021-04-07	[15]0.14	[18]0.86	[21]5.21
16	DERMALOG-008	2021-01-25	[16]0.19	[19]1.04	[23]6.39
17	IDEMIA-007	2020-01-17	[17]0.19	[21]1.12	[20]5.19
18	MICROSOFT-006	2018-10-29	[18]0.23	[16]0.71	[15]3.21
19	SENSETIME-005	2020-12-17	[19]0.28	[9]0.45	[8]0.85
20	SENSETIME-004	2020-08-10	[20]0.29	[10]0.50	[9]0.89
21	RANKONE-010	2020-11-05	[21]0.31	[20]1.06	[22]5.71
22	COGNITEC-004	2021-03-08	[22]0.49	[26]2.18	[25]9.20
23	RANKONE-009	2020-06-26	[23]0.52	[24]1.52	[24]7.85
24	NEC-002	2018-10-30	[24]0.99	[22]1.29	[11]1.61
25	NEC-003	2018-10-30	[25]0.99	[23]1.29	[12]1.78
26	NEUROTECHNOLOGY-007	2019-10-03	[26]1.02	[25]2.02	[27]31.93
27	DERMALOG-007	2020-02-12	[27]1.97	[27]3.66	
28	IDEMIA-004	2018-06-30	[28]4.96	[28]8.13	[26]17.81
29	NEC-000	2018-06-21	[29]15.41	[29]18.85	[28]91.97

Table 2: **False negative rates by gallery size and number of enrolled images per person.** *The second row N values give the number of individuals enrolled in each gallery. The 420 person galleries represent aircraft boarding; the 42000 case represents a airport security line where many more people are expected. The third row k values give the number of images of each enrollee in each gallery.*
The second and third columns identify the algorithm and the date it was submitted to NIST. The remaining columns give false negative identification "miss" rates i.e. the proportion of travelers not matched to their gallery photo(s), expressed as a percentage. Lower values are better, and the table is sorted on the first results column. The superscripts give the rank of the algorithm for that column. The threshold is set so that only a fraction, 0.0003, of non-mated searches would return any match. The shaded cells indicate the three most accurate algorithms for that trial.

that more prior enrollment images for each enrollee improves accuracy and, second, that even a 100-fold population size increase degrades accuracy only modestly.

The report gives some information on demographic dependencies. Many algorithms give somewhat higher false negative rates on women compared to men. This is not true for, or has reduced magnitude, for the more accurate algorithms. With high accuracy, and with opportunities in real operations to make second identification attempts, these differentials are either small or can be remediated. The report also notes demographic dependence on false positive rates, particularly that women and people of certain nationalities, often East Asia, tend to give higher false positive identification rates. Again some algorithms are considerably superior to others in this respect. Note that security context matters: In particular that passive non-mate, and active attack, presentations will be very small percentages of all attempts.

The accuracy estimates in this report are just that, estimates. Section 5 notes several factors that would drive accuracy higher or lower. Primary among those is that we can't be sure how well the images we possess represent the actual paired ENTRY galleries and their EXIT photos. A passport control authority has two complementary options for improving on our estimates: First is to run exhaustive clipboard style operational tests; second is to provide NIST or some other laboratory with a) actual images, and b) the operational algorithm. This latter option had been planned in 2019 but was derailed for several reasons, including the COVID pandemic.

1 Errors and Their Consequences in Biometric Exit

The following subsection describe mechanisms by which an EXIT system, as comprised, makes errors. We distinguish biometric errors (from cameras and algorithms) from operational issues deriving from business processes.

1.1 Failure to Enroll

Nature: In the context of TVS' manifest-driven gallery construction, some individuals who are legitimately booked on an aircraft will not be enrolled in the face recognition gallery. This number will usually be zero but could be non-zero for several reasons, among them:

1. Absence of historical photo. For various policy-related issues a PCA may not have a prior photo - these could include first-time visitors, foreign passport holders born in the country, and bilateral trade-related visa exemptions. In such cases a PCA might legitimately have no ENTRY record. This circumstance might be termed an *operational failure to enroll*.

   Measurement: A PCA can estimate the prevalence of missing enrollments by cross-referencing airline manifests and the lack of prior reference photos. This estimate will include instances of 2 below.

   Consequences: Failures to enroll will manifest as false negatives (see section 1.4 below). Airline staff can resolve by biographic and human visual biometric inspection.

2. Biographic errors. It is possible that the manifest provided to the PCA by the air carriers includes biographic errors from well understood sources such as recent marriage and change of name, and typographical errors.

   Measurement: A PCA can estimate the prevalence of missing enrollments by cross-referencing airline manifests and the lack of prior reference photos. This estimate will include instances of 1 above.

   Consequences: Failures to enroll will manifest as false negatives (see section 1.4 below). Airline staff can resolve by biographic and human visual biometric inspection.

3. Poor image quality. It is possible the photographs that a PCA has on an individual are of poor enough quality that the TVS feature extraction software fails to produce a template from the photograph. This could occur because the face detector fails to find the face, or because the software deems the photo to be of low utility to their downstream recognition engine so, electively, does not produce a template. Such outcomes would constitute *biometric failures to enroll*.

   Measurement: A PCA can estimate algorithm enrollment failures by direct analysis of TVS logs.

   Consequences: Failures to enroll will manifest as false negatives (see section 1.4 below). Airline staff can resolve by biographic and human visual biometric inspection.

1.2 Failure to Capture

Nature: During aircraft boarding TVS never receives photos of some travelers for at least two reasons:

1. Camera failure: Some cameras might fail to trigger and take a photograph. This can occur due to failed face detection (e.g. due to sunglasses, or subject not being in the field-of-view), or because an on-board quality algorithm deemed the captured photograph of insufficient utility, or due to some system fault of the kind remedied

by rebooting the system. During observations at various airports in June 2019, some cameras would not trigger; others would trigger only after the subject disengaged by moving away, and then re-engaged.

Measurement: We can put an upper bound on the frequency of such events by subtracting the number of people verified from the number of people on the manifest. This quantity will include outright recognition failures too. This estimate will include people who never appeared before the camera (e.g. because the airline allowed traditional paper-based boarding).

2. Airline operations: An operational source of "failure to capture" can be that airline staff might redirect the traveler to some human-adjudicated boarding process such as the traditional passport or boarding-pass based biographic confirmation. This could occur a) because the staff perceive the traveler has had difficulty, or b) that they will have difficulty (e.g. because they're too tall or short), or c) simply because the airline staff are trying to expedite boarding by using the biometric process and the biographic process.

   Measurement: Such events can only be documented by observation, most readily human observation, but also via some automated supervisor or logging system.

Consequences: For an in-scope traveler the consequence will be that EXIT will only be recorded biographically according to the information used in forming the passenger manifest – this is essentially the legacy biographic process. An immediate operational consequence is that the passenger will have to be processed manually (by airline) staff. Downstream, this may cause the PCA to perform overstay inquiries.

1.3 Failure to Extract Features

Nature: It is possible the photographs that the PCA has on an individual are of poor enough quality that the TVS feature extraction software fails to produce a template from the photograph. This can occur during gallery construction or during EXIT operations.

Measurement: Such events can be measured from algorithm logs such as those produced in FRVT, and likely by operational systems.

Consequences: If TVS fails to extract features during EXIT, the traveler's boarding attempt will be rejected, possibly silently. He or she may make a second attempt, perhaps after being prompted. In June 2019 observation of boarding, the author noticed airline staff directing passengers to the gate-agent biographic process. This would likely lead to the PCA having to revert to its reliance on biographic recording of EXIT.

1.4 False Negative During Identification

Nature: In a positive identification application like EXIT, the one-to-many search algorithm generally grants access if the rank-one (i.e. highest-scoring) candidate has a score above threshold. The identity of the person in the live photo is taken to be that returned by the system even if it is incorrect. From a testing perspective, an error occurs if the rank-1 candidate is of the wrong identity or has score below threshold. This gives us the following performance metric, the false negative identification rate (FNIR):

$$\text{FNIR}(N,T) = \frac{\text{Num. searches where top-scoring candidate has wrong ID or score below threshold}}{\text{Number of searches conducted}} \quad (1)$$

This definition automatically incorporates "failure to extract" feature events as they won't return high-scoring candidates. The dependence on gallery size, N, and threshold, T, are present as they are design choices affecting FNIR. For

an audience who likes to think in terms of accuracy or "hit rates", we can convert the "miss rate" or Eq. 1 to True Positive Identification Rate using TPIR = 1 – FNIR, so a 3% FNIR becomes 97% TPIR. However, that definition is naïve in that it assumes every traveler was photographed. It ignores instances of failure-to-capture, and also cases where travelers are photographed, not matched, and then make further attempts. Then an operational definition of false negative identification rate is

$$\text{FNIR}(N,T) = \frac{\text{Num. travelers who are not matched to the correct ID in one or more presentations to the camera}}{\text{Number of travelers}} \quad (2)$$

The two measures would be equivalent if each traveler executes just once search. To the extent that is true, our Equation 1 estimates in this report will approximate Equation 2. We use the 1 thoughout this report. We discuss in section 5, factors that can make our estimates too high or too low.

Measurement: In this report, we don't have insight into the transactional nature of aircraft boarding, with failed captures or failed searches. Instead all we see are images that can be used in simulations of boarding. For measuring duration of boarding, and quantities such as the number of travelers who need to make further presentations, an operational observational test is most appropriate. Many aspects may be measurable in scenario tests in which passengers and airline staff model the actual target boarding process.

Consequences: False negatives will usually be resolved by biographic and human visual biometric inspection by airline staff. For an in-scope traveler the consequence will be that EXIT will only be recorded biographically according to the information used in forming the passenger manifest – this is essentially the legacy biographic process. Downstream, this may cause the PCA to perform overstay inquiries. The PCA would possess an aircraft boarding photo, but one that is not bound to an identity – such images are provided to PCA staff monitoring a flight departure.

1.5 False Positive During Identification

Nature: False positives occur when images of two people are erroneously associated. In biometric EXIT there are three kinds of false positive:

- First is the **in-gallery false positive** in which a legitimately enrolled traveler matches the wrong identity. Such a possibility necessarily implies that the correct identity would be displaced from the rank-1 position on the candidate list, usually to rank 2. That list is a data structure internal to the particular TVS and is not typically presented to airline staff or anyone else. Depending on how the system is built, an in-gallery false positive may result in a false negative for the correct passenger if he or she boards later in the process. Such errors were observed by the author in June 2019 during visits to observe the boarding process in five different airports.

  Measurement: The in-gallery false positive rate is not currently defined in performance testing standards as it is approximately the proportion of mated searches yielding the mate at rank 2 or higher. Such an outcome would most often occur because the search imagery is of poor quality, but could occur if the enrolled imagery was poor. Formal measurement can be achieved by careful online observation of the boarding process. Error rates can be estimated approximately from recognition logs by counting instances of a passenger apparently boarding the plane twice – once legitimately as themselves and secondly when another traveler incorrectly matched their identity.

  Consequence: Such errors will likely be resolved by airline staff, who may become familiar with such an event.

- Second is the **incorrect acceptance** of people who are not in the gallery and not expected on the departing flight.

This population includes travelers who mistakenly arrive at the wrong gate[5] without subversive intent. The frequency of occurrence is usually stated by the False Positive Identification Rate (FPIR). FPIR is the primary security-related parameter in a one-to-many access control system. Its value is chosen by a system owner to target security objectives and is implemented by setting the system threshold according to some calibration[6].

Measurement: Such errors were observed by the author in June 2019 when airline staff in the gate area were accidentally captured by the camera and incorrectly matched to an actual passenger. While this kind of error could be measured by making in-person attempts, this approach does not scale. An offline approach in which images are matched after-the-fact affords more precise FPIR estimates – this report takes just this approach.

Consequence: The consequence for the airline is potentially a stowaway. However, airlines usually count passenger totals and may thereby be able to detect such events. While there is little consequence for the PCA's EXIT processing, these events, if undetected, could cause erroneous updates to the PCA's systems, undermining integrity.

- Third is a false positive from someone who is illicitly trying to gain access. This category would include stowaways and potentially visa overstayers.

> **Passive vs. active attack**
>
> False match rates usually express the likelihood that a face recognition algorithm will compare two photographs and return a high score from two individuals who are selected entirely randomly, or perhaps with the restriction that they have the same demographics such as age, sex, and race.
>
> However, if someone makes more deliberate efforts to impersonate an identity e.g. via cosmetics or wearing a face mask, then additional algorithms must be employed to detect the presentation attack (PA). To succeed an attacker must defeat the PAD subsystem, if installed and enabled, AND match the intended identity – see section 1.6

  - **Casual attack:** If someone is making a low-effort attack – for example as a stowaway – they might rely on matching any identity essentially fortuitously, and then hoping the airline staff does not notice nor take steps to resolve the match. A second intent here would be to fake someone's departure from a country. This possibility - to overstay a visa by sending a confederate to verify a particular identity - is notable in that it would be difficult for an overstayer to select a confederate who would match the *particular* identity in a biometric search – In this respect a one-to-many system where there is no claim to an identity is more secure to passive attack. However, the security context is that such a system is prone to circumvention attack: a confederate failing to match an enrolled identity might appeal to airline staff who would make biographic or visual biometric efforts to verify the person, with the likely outcome that passenger would be allowed to board.
  - **Active attack:** An overstay attempt would be much more successful if the confederate actively impersonates the visa-holder. This could be achieved using a presentation attack instrument such as a face mask.

    Measurement: Vulnerability to active attack could be demonstrated via "red-team" presentations to the operational system. More formal quantification of the vulnerabilities is best conducted in laboratory trials using identical equipment to that used in the operation. Each approach will require controlled, defined and

[5] This can occur because of a gate change, or because someone goes to the wrong gate. The author, for example, has accidentally tried paper-based boarding at the adjacent gate on several occasions.

[6] Threshold calibration is an imprecise process because FPIR often depends on demographics and image quality related properties. A threshold is set starting with vendor recommendation and refined using offline tests (such as FRVT) or empirical instrumentation and tests or logging of the operational system.

repeatable production of presentation artefacts (masks, cosmetics etc.). The metrics relevant to this kind of attack are standardized – see section 2.2.

- **Comparison with existing paper-based boarding:** Attacks on non-biometric paper-based departure systems are possible also: A stowaway could find, or steal, a boarding pass. A confederate seeking to depart for a visa-overstayer would only have to present a boarding pass and possibly a cursory inspection by the airline staff of the passport. In these cases, a biometric system, if used and not circumvented, will improve security compared over the legacy process.
- **Consequences:** For an IA, a successful impersonation attack would likely produce an undetected overstay. The attack assumes the confederate either does not need or want to return to the United States or could do so using other documents. There are no consequences for the airline.

1.6 Presentation Attack Detection Metrics

The ISO/IEC 30107-3 standard establishes the metric Impostor Attack Presentation Match rate (IAPMR) which expresses the proportion of attackers who both defeat the PA detection software AND match the correct identity. That metric is appropriate to access, say, to a mobile phone. In one-to-many processing such as paper-less EXIT, a traveler would have to defeat the PAD and match the specific intended enrollment.

1.7 Demographic Differentials

Biometrics generally give different error rates for different populations. For example, fingerprints are known to give higher false negative rates in the very young and the elderly[7]. NIST Interagency Report 8280 [2] documented error rate differentials for face recognition examining the effect of sex, age and race on accuracy of many commercial algorithms. That report made an important distinction between differentials in false negative and false positive error rates, the former affecting how well a single individual is not matched as him or herself, the latter affecting how often two individuals are erroneously associated. The consequences of such errors, and differentials in their rate of occurrence, are very different. We include visualizations of false negative differentials in section 3.2.1 and false positives demographic differentials in sections 3.2.4 and 3.2.5.

[7] In the young, typical contact sensors have inadequate resolution to resolve the fine friction ridge structure. In the elderly the factors include inelasticity of the skin and inability to present flat impressions e.g. due to arthritis.

2 Operational Questions

2.1 Context

This report gives extensive documentation of biometric identification performance. However larger questions exist, and core biometric performance statements only inform answers to those questions. For example,

- An airline might ask "which camera and boarding solution should we procure?" – this report is silent on that because we would at least need to know what cameras were used for collecting the data, and this is not information we have. Dedicated laboratory tests of camera equipment[8] are appropriate to such tests.
- An airline might ask "what is the proportion of passengers being referred to gate agents" – such a quantity could be approximately estimated from TVS logs, but is more precisely answered only by observation of the operational system.
- A security analyst might ask "what is the chance on an active impersonation attack succeeding" – this question can be addressed potentially by laboratory trials if the fielded system can be copied and if access access to the TVS recognition engine is granted. It may be easier to conduct operational "red team" trials with an appropriately motivated staff. Active attacks (e.g. using face masks) are not the fault of the recognition algorithm per-se, but are enabled by lack of (or use of poor) presentation attack detection algorithms[9].
- A policy maker might ask "is biometrics better than biographic matching for overstay detection?" – we can't address that without biographic data and extant biographic matching algorithms.

2.2 Standardized Tests

Since 2003, there have been significant worldwide investments in supporting development of biometrics performance testing and reporting standards in the ISO/IEC JTC 1 Subcommittee 37. That body develops very well vetted consensus standards in working groups (WGs) dedicated to vocabulary (WG1), interfaces (WG2), data interchange and image quality (WG3), application aspects including face-aware capture devices (WG4), performance testing and reporting (WG5) and societal issues (WG6). Table 3 lists standards that may be valuable in the measurement of performance in a PCA's ENTRY-EXIT processes.

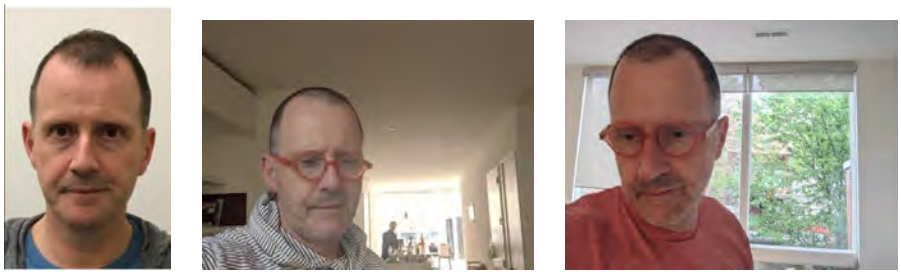
There are a number of other testing standards supporting other domains of use.

[8] See the scenario tests conducted at the Maryland Test Facility, for example.

[9] PAD approaches have advanced in recent years, both in software and hardware. However, their use will often increases false negatives because they sometimes erroneously flag a bona-fide presentation. Their use may be more appropriate on inbound arrival processing (ENTRY).

Table 3: Testing standards supporting performance measurement in ENTRY-EXIT

Number	Title	Relevance
ISO/IEC 19795-1	Principles and Framework	This foundational document establishes requirements on all biometric tests regarding design of tests of enrollment, verification and identification, and how to put uncertainty estimates on measured error rates.
ISO/IEC 19795-2	Technology and Scenario Testing	Regulates two kinds of "in-vitro" test: "Technology" tests which are most often offline sample comparison and search tests such as those documented herein, and "scenario tests" that are usually human-in-the-loop laboratory tests intended to mimic operational systems.
ISO/IEC 19795-3	Environmental Aspects	A technical report guiding testing and reporting in the presence of environmental variations such as humidity and illumination
ISO/IEC 19794-4	Interoperability Testing	Relevant to tests where components of a system, possibly from different manufacturers must produce and consume standardized data, for example cameras must produce images that will be consumed by remote recognition algorithms.
ISO/IEC 19795-6	Operational Testing	Establishes requirements on "in-situ" tests, where identity ground truth is not necessarily known, and where the act of measuring accuracy or duration can potentially disturb the estimates. This kind of test is advantaged by considering the actual system on its native population in its native environment. These aspects are often material and difficult to approximate in lab tests.
ISO/IEC 30107-3	Presentation Attack Detection	This standard regulates tests of PAD components and PAD-enabled systems and gives detailed guidance on measuring and naming of error rates that are available for various levels of logging and instrumentation.
ISO/IEC 19795-10	Demographic dependence	This standard (2020-11) is in the early stages of development. It will establish requirements on various kinds of tests intended to measure demographic differentials in biometric devices, algorithm and systems.



a) Visa or Passport | b) Immigration Lane ENTRY | Departure Point EXIT

Figure 1: Image (a) is representative of passport-like data that would ordinarily be available to a PCA's TVS from all in-scope travelers and citizens. However, such images were not available for the trials conducted here. The remaining images have size 240x240 pixels and are representative of some poorer quality ENTRY images: Image (b) is typical of ENTRY photos in that it has has non-frontal pose, and strong background illumination reducing contrast on the subject's face. Image (c) is typical of EXIT photos in that it exhibits some close-range distortion, mild non-frontal pose arising from "don't wait for frontal presentation" fast-capture and adverse background lighting. contrast.

3 Simulation and Results

3.1 Air-Exit Simulations

We simulate biometric EXIT by running simulations using archived images as follows.

1. We form a departing flight by placing ENTRY images from N = 420 individuals into a gallery. We use 420 because that number is reasonable for a large commercial twin-aisle jet such as the Boeing 777 or the Airbus A380[10]. The exact gallery size is not that important because accuracy is an insensitive function of N. We later increase the population size to 42 000 to simulate an airport security checkpoint, for example.

2. We populate an EXIT gallery in two ways.

   (a) First, with one ENTRY image per person.

   (b) Second with a multiple such images, the average is about 6, with some variance per individual.

   While it is common practice to populate the gallery with images from all prior encounters of a person[11], we include the one-image case to show "worst-case" accuracy i.e. that expected when only one prior encounter is available. We include results for single- and multiple-image enrollment in sections 3.2.1 and 3.2.2 respectively[12].

3. We populate a gallery with individuals from the same region of the world. We do this for two reasons: As discussed in section 5, we list various factors that will push our error rate estimates up, and down. that flights departing the U.S. tend to have some racial homogeneity – flights departing for Japan have more individuals from East Asian countries than do flights departing to Nigeria, and more than would be expected by random selection. Another reason is that face recognition accuracy will be worse for homogenous galleries because false positives will be more common. Our practice of building homogenous galleries biases the test toward higher error.

4. The 12 regions are: Europe, W. Africa, E. Africa, N. Africa, Middle East, S. Asia, E. Asia, Oceania, N. America, C. America, Caribbean and S. America. We assign individuals to a region based on the issuers of their travel document. Occasionally some travelers will travel on a different country's travel document; in such cases we assume their region to be that of the gallery ENTRY image.

5. We form 567 galleries, with one image per person. We form another set of 567 galleries with variable numbers of images per person. The number of galleries we can form per region varies because we have more images from some regions than others.

6. We search each gallery using a single probe-set containing 127 258 EXIT images of 123 075 people. By visual inspection it is evident that the images are collected using different cameras in different locations. For a given gallery only small proportion of the searches will have an enrolled mate in the departure gallery, at most 420 of 123 075 people. These mated pairs afford estimates of false negative identification rate. The remaining images, from persons of all regions of the world, form a non-mated search set used for estimating false positive identification rates.

[10] Aircraft configuration makes a difference, so that while the A380 is capable of carrying 560 economy class passengers it is atypical for that to occur for aircraft departing the United States.

[11] Not all, as the U.S. PCA stated, their TVS "does not enroll recent crossing images of U.S. travelers into the gallery, but does enroll recent crossing images of foreign nationals into the gallery."

[12] The single-image enrollment will be more pertinent to processing of citizens of a country for whom, often, only one photo exists in the gallery. The multiple-image enrollments yield better accuracy, and are pertinent to foreign travelers.

7. We run multiple algorithms, in some cases more than one from each developer. These were submitted to the one-to-many identification track of the FRVT between May 2018 and the present. The list of algorithms includes the NEC-3 algorithm that was broadly the most accurate through November 2018 as reported in NIST Interagency Report 8271 [1], but which has been eclipsed in accuracy by newer algorithms submitted since.

8. We compute 10 thresholds for each algorithm corresponding respectively to the 10 false positive identification error rates: 0.00003, 0.0001, 0.0003, 0.001, 0.003, 0.01, 0.03, 0.1, 0.3, 1. We get the threshold value by looking at the highest non-mate score produced when running all non-mate searches against all galleries. Given, say, 126838 non-mate searches into each of 567 galleries, the threshold for FPIR = 0.0003 is taken to be the 126838 x 567 x 0.0003 = 21575-th highest observed rank-one comparison score.

3.2 Results

3.2.1 Attainable accuracy with single entry image

Figure 2 shows accuracy for two algorithms submitted to NIST 28 months apart. These are the NEC-3 algorithm submitted to NIST in November 2018, and the Visionlabs-10 algorithm from February 2021. The gallery size is N = 420 subjects, each person enrolled with exactly one ENTRY image. The vertical axis is a count of the individuals who are not biometrically authenticated during boarding. The horizontal axis shows the region of the enrolled population. The dots correspond of one departing flight. The dots are jittered horizontally around the region label, and vertically around the integer value, to avoid over-plotting and show the distribution.

The notable observations from the graphs are:

1. The number of false negative recognition errors is spread between zero and 16, with the most common value being 6. These errors would need to be resolved via a second attempt at biometrics, or via an airline-defined biographic process.

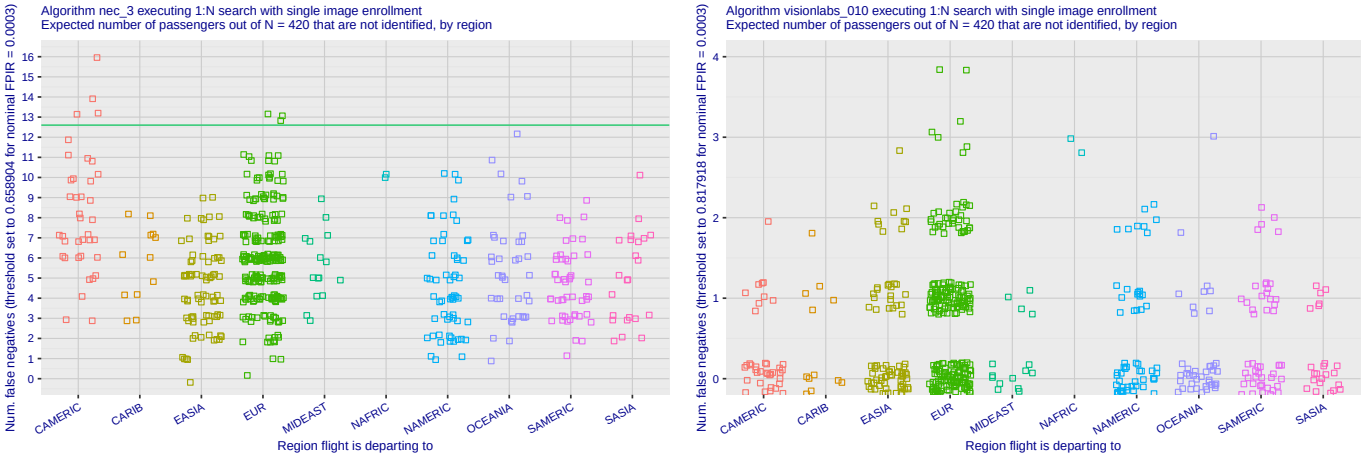
2. The distributions across regions are similar. The Central American flights give modestly higher FNIR, but this may simply be the result of chance. To the extent that some of the regions here are proxies for race, the results comport with those published in NIST Interagency Report 8280 [2] showing little dependence of false negative rates on race. Any false negative demographic differentials should be corrected for:

   (a) Ageing: It is possible that different travelers from certain regions travel less frequently such that the gallery photos are older – time lapse affects appearance and accuracy.

   (b) Age: It is possible that absolute age affects accuracy. For example, although not the subject of the simulation here, flights into Orlando are disproportionately populated with children[13] whose lower height can affect head pitch angle and accuracy.

3. We report the rate of false negatives (FNIR) in a subsequent figure – but note here that a count of 13 (i.e. 0.03 * 420) corresponds to a 3% failure rate. On that basis, the overall error rate is below 3% corresponding to better than the 97% verification rate required in 2007 legislation.

[13] In visits to observe EXIT boarding processes June 2019, the author observed children, without instruction, standing on tip-toes in order to present their face to the camera mounted above five feet. This was sometimes effective.



(a) NEC-3 (b) VISIONLABS-10

Figure 2: Count of false negatives on simulated flights by region using the NEC-3 algorithm from Nov. 2018 and the Visionlabs-10 version from February 2021. The gallery is populated with one ENTRY image from each of N = 420 individuals. The threshold is set to target a false positive identification rate of 0.0003 corresponding to 1 false positive in 3 333 impostor search attempts. The false negative identification rate for a flight can be stated by dividing the number of false negatives by the number of passengers, 420.

Figure 2(b) shows accuracy for a recent algorithm that is among the most accurate submissions to the one-to-many track of FRVT. The number of errors now is much lower, ranging from 0 to 4, with most common value being 0. A false negative count of zero corresponds to correct recognition of all passengers.

Note that the most accurate algorithms have been submitted to NIST recently, in early 2021, showing accuracy gains are still being realized by developer innovation. Several algorithms, including the VisionLabs-10 algorithm used in Figure 2(b), are more accurate than leading algorithms submitted to NIST in 2018 - see the ongoing FRVT webpage for names, dates, and more general accuracy results. The implication is that a PCA will realize accuracy gains if its technology refresh process is active and frequent.

Figure 3 shows the same figure for the most accurate algorithms tabulated appearing in Table 2. We note the following:

1. Figure 3 includes, in blue text, values for FNIR, the estimated proportion of passengers who will not be able to board with a single probe capture. The values are well near 0.1% for the most accurate algorithms, and often above 1% for the less accurate ones.

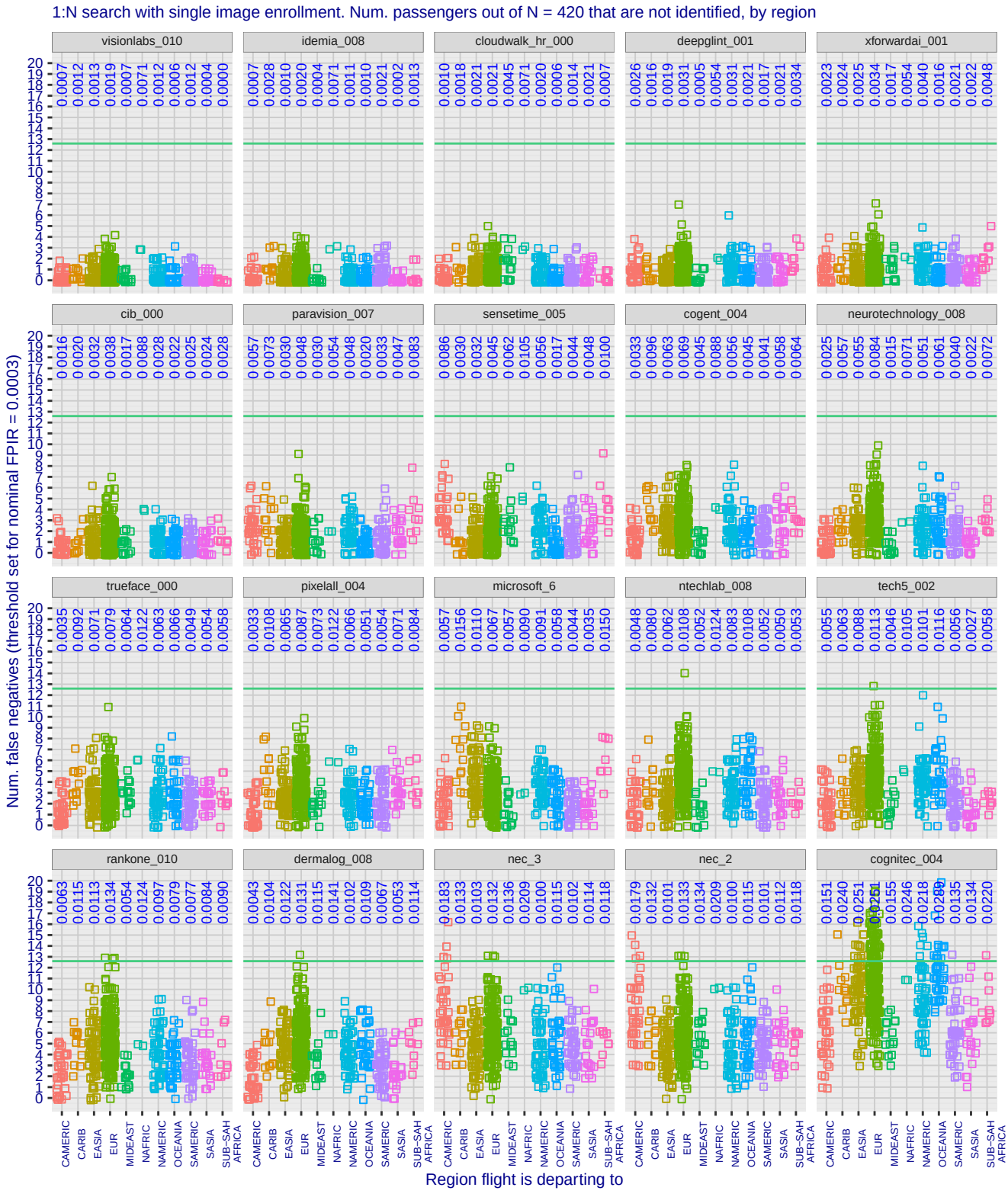


Figure 3: Count of false negatives on simulated flights by region. Each point corresponds to one flight with a gallery populated with one ENTRY image from each of N = 420 individuals. The threshold is set to target a false positive identification rate of 0.0003 corresponding to 1 false positive in 3 333 impostor search attempts. The blue text gives FNIR. The panels are arranged left-to-right, top-to-bottom in order of mean false negative count. The horizontal green line corresponds to the 3% false negative goal implied by legislation in the U.S.

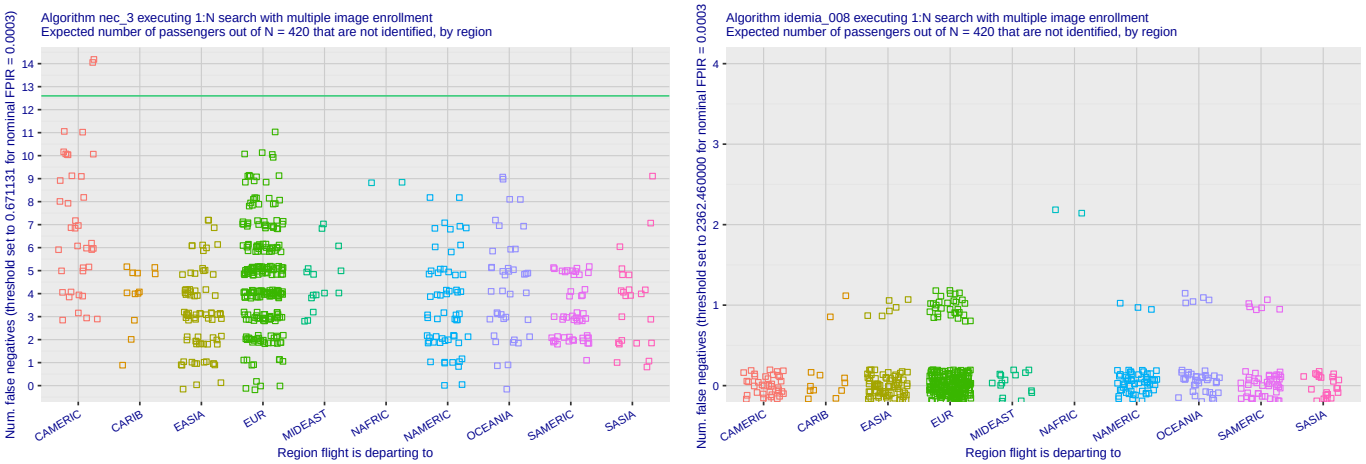
PCA = PASSPORT CONTROL AGENCY
TVS = TRAVELER VERIFICATION SERVICE
FNIR(N, R, T) = FALSE NEG. ID RATE
FPIR(N, T) = FALSE POS. ID RATE
N = NUM. ENROLLED SUBJECTS
T = THRESHOLD
$T = 0 \rightarrow$ Investigation
$T > 0 \rightarrow$ Identification

3.2.2 Attainable accuracy with multiple entry images

Figure 4 shows the accuracy results for two algorithms for a gallery of size N = 420 subjects each now enrolled with *multiple* ENTRY images. The algorithms were submitted 29 months apart, in November 2018 (NEC-3) and March 2021 (Idemia-8). From the two figures we note the following:

1. The use of multiple enrollment images reduces the number of false negative recognition errors modestly for NEC-3 (2018). It produces around 4 errors on average instead of 6 with a single image. The worst case count is reduced from 16 to 14.

2. With Idemia-8 (2021) the effect of enrolling more images is a more substantial reduction in false negative outcomes such that a large majority of flights will see all passengers board without any errors. The worst case count of error is reduced from 4 to 2.



(a) NEC-3

(b) IDEMIA-8

Figure 4: Count of false negatives on simulated flights by region using the November 2018 NEC-3 and March 2019 Idemia-8 algorithms. Each point corresponds to one flight the gallery for which is populated with multiple ENTRY image from each of N = 420 individuals. The threshold is set to target a false positive identification rate of 0.0003 corresponding to 1 false positive in 3 333 impostor search attempts. The horizontal green line corresponds to the 3% false negative goal implied by legislation in the U.S.

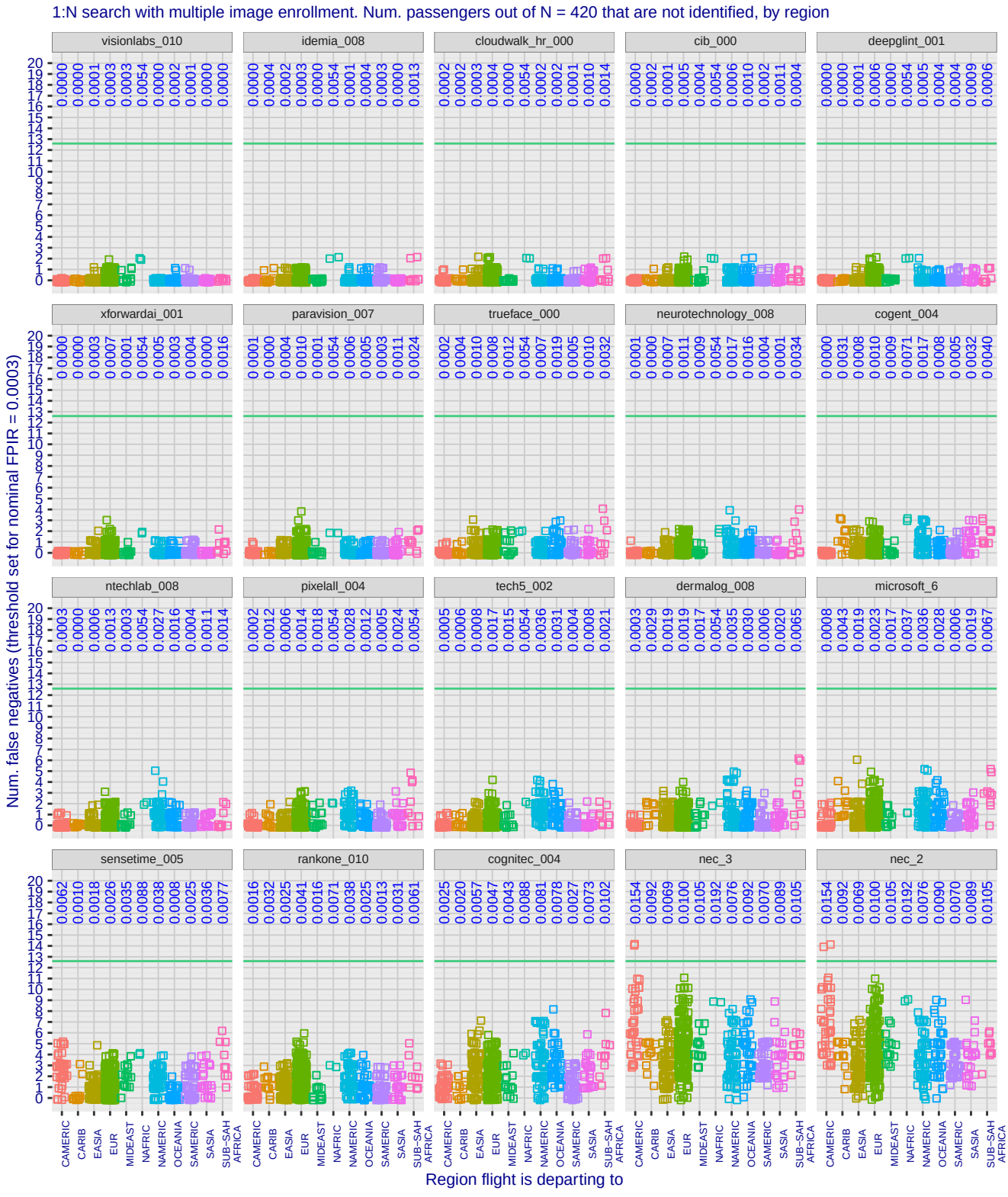


Figure 5: Comparison of false negative identification rates between number of images enrolled (one per person, vs. several) and between algorithms. The algorithms were submitted to between June 2018 and April 2021. Each point corresponds to one flight to the identified region the gallery for which is populated with ENTRY images from each of N = 420 individuals. The blue text is a false negative identification rate (FNIR), often below 1%. The orange text is the number of simulated flights, out of 567, for which the number of false negative errors is zero. The threshold is set to target a false positive identification rate of 0.0003 corresponding to 1 false positive in 3 333 impostor search attempts.

The failure of NEC-3 to exploit multiple images may stem from how we provided images to the gallery. In FRVT we typically provide all images of an individual to the algorithm in one call to the template generation function – the algorithm consumes multiple images and has the opportunity to select or fuse images as it sees fit. However, that is atypical operationally: images are provided to the algorithm serially such that multiple images of the same person result in separate enrolled templates - that is the model followed here[14] even though it denies the algorithms an explicit fusion opportunity. Figure 5 shows analogous results for nine of the more accurate algorithms evaluated in FRVT through November 2011. The panels are included in order of mean overall number of false negatives. Notably:

1. For the majority of flights, the most accurate algorithms correctly identify every passenger, and only ever fail to on up to 2 out 420 people.
2. On this metric there are multiple algorithms affording lower false negative identification error rates than does NEC-3. This is an existence proof of better accuracy that suggests an PCA will benefit from monitoring of test results and regular technology refresh. A PCA would need to factor other variables into procurement from a new developer including performance aspects (speed, scalability to large galleries, and demographic equitability) and contractual factors like capital and transaction costs, including those of integration.

3.2.3 False negative vs. false positive tradeoff

The results for each algorithm thus far have been stated at a single threshold. If we had set this threshold to a higher value the false negative rates would also have been higher, but with the advantage of lower false positive rates. Conversely, if the threshold had been low, false negatives would be better and false positives could occur more easily. The threshold is conventionally set to achieve a low enough probability that an impostor could match an enrolled identity thereby meeting some planned security objective. In one-to-many applications, an impostor only needs to match any enrolled identity to gain access – he has N opportunities. This generally necessitates higher thresholds than one-to-one verification where the impostor claims one particular identity.

[14] See Figure 8 and section 3.2 in the FRVT 1:N report, NIST Interagency Report 8271 [1] for details on multi-image enrollment and metrics. See the FRVT page for newer algorithm results.

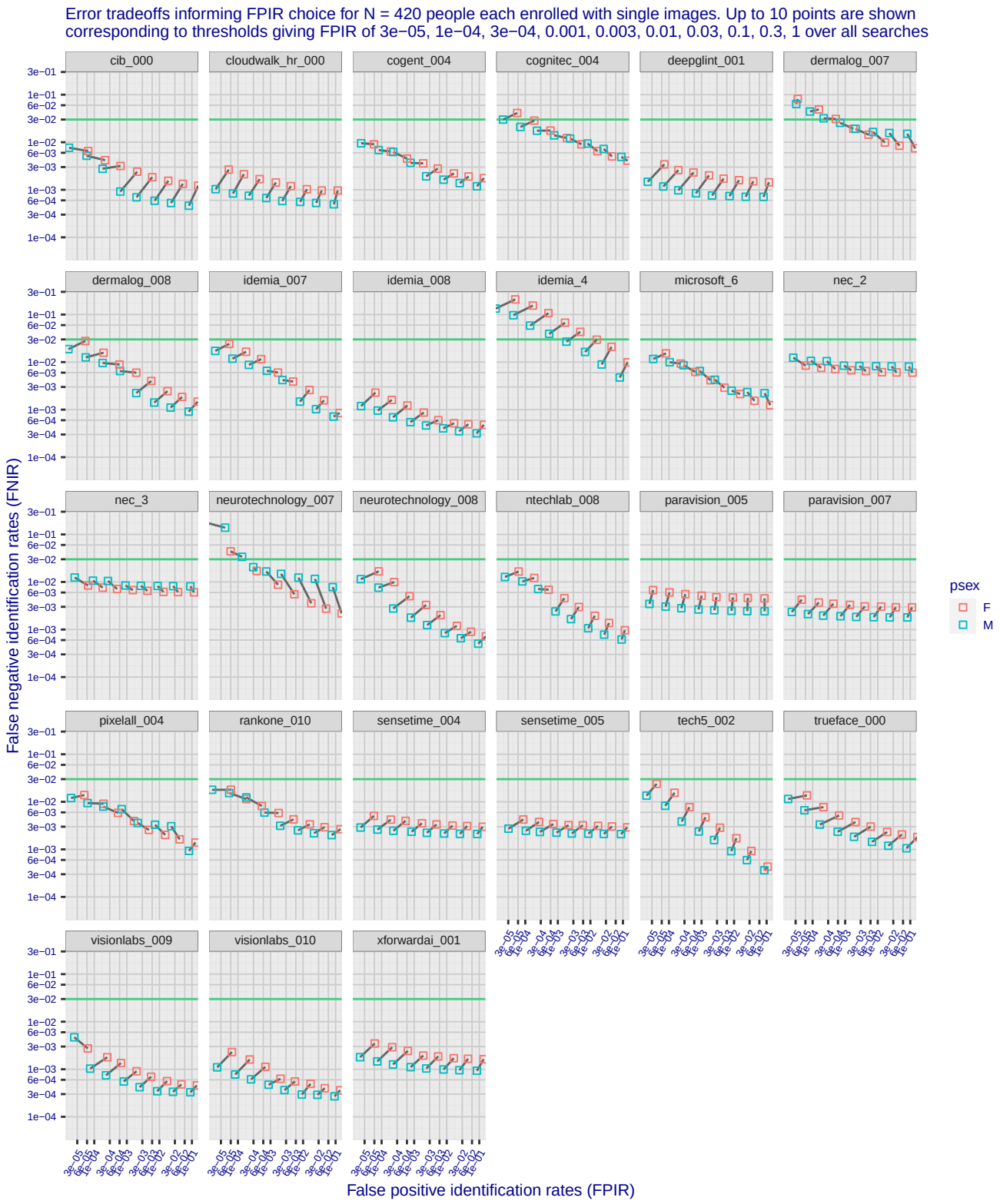


Figure 6: Error tradeoff for 26 algorithms executing 1:N searches with N = 420 people enrolled with a single image. The 10 point-pairs correspond to 10 possible thresholds for each algorithm. The red and blue boxes correspond to female and male travelers; their relative displacement indicates generally higher false positives and false negative rates in women. Smaller displacement indicates smaller (better) demographic differential (see NEC and Paravision-5, for example). The horizontal green line corresponds to 3% false negative goal implied by legislation in the U.S.

What should the false positive identification rate be?

This question is about policy. As discussed in in the introduction a TVS can serve double duty as an aircraft access control system and as a visa-holder EXIT status facilitation system. The discussion centers on what the system is trying to prevent – stowaways, evasion of traveler-to-bag matching, or faking someone's departure.

One factor is the prior probability that someone would try to board a flight at all. It's likely quite common and not necessarily nefarious - the author has accidentally gone to the wrong gate on more than occasion. For pure facilitation a low threshold could be used, but in its access control role that would allow any traveler to board, and potentially get free passage. A conventional value for access control is for false positive rate of 1 in 10 000. Lower values can used but impostors will switch to active attack techniques to achieve a false positive. One factor is variability in false positive rates with demographics: many algorithms can give 100 times more false positives on elderly, female people from certain countries.

Figure 6 shows the error rate tradeoff by plotting false negative identification rates against false positive identification rates at ten operating thresholds spread over four decades of FPIR, from 1 in 33 333 to nearly 1 in 3. Instead of showing the full curves, the ten-point pairs expose the increase in FNIR at low FPIR but also show the difference in error rates for men and women.

We note the following points:

1. Some algorithms give generally lower FNIR across the range of FPIR. This is simply a re-iteration that accuracy varies markedly.
2. Some algorithms give a flat error tradeoff characteristic. This is most evident for the idemia-8, deepglint-1, nec-3, paravision-5 and sensetime-4 algorithms. This is an attractive property of any biometric system because it allows very low false positive identification rates to be attained without intolerable increases in false negative identification rates. This becomes important later when we increase the enrolled population size by a factor of 100.
3. Most algorithms give FNIR below 0.03 (the green line in the plots) for a wide range of FPIR, thereby meeting the legislative mandate to be able to verify the EXIT of 97% of (in-scope) travelers.

Comparing Figure 7 with Figure 6 shows that across the four-decade range of FPIR, the FNIR values are reduced by using multiple enrollment images. The single-image enrollment represents "worst-case" of having just one prior encounter. The multi-image case is more typical.

Note that this analysis doesn't answer the technical question of whether enrolling multiple images per subject increases FPIR versus using just single image. The reason is that the thresholds for multiple enrollments are generally higher than for singles. There are exceptions – Idemia-7 for example. The question is important in situations where some travelers might have dozens of enrollment images and the algorithm response could be to attract false matches i.e. to make such enrollees lambs[15].

[15]The term lamb, a category defined in "The Biometric Zoo", refers to an enrollee who attracts more than average number of false matches.

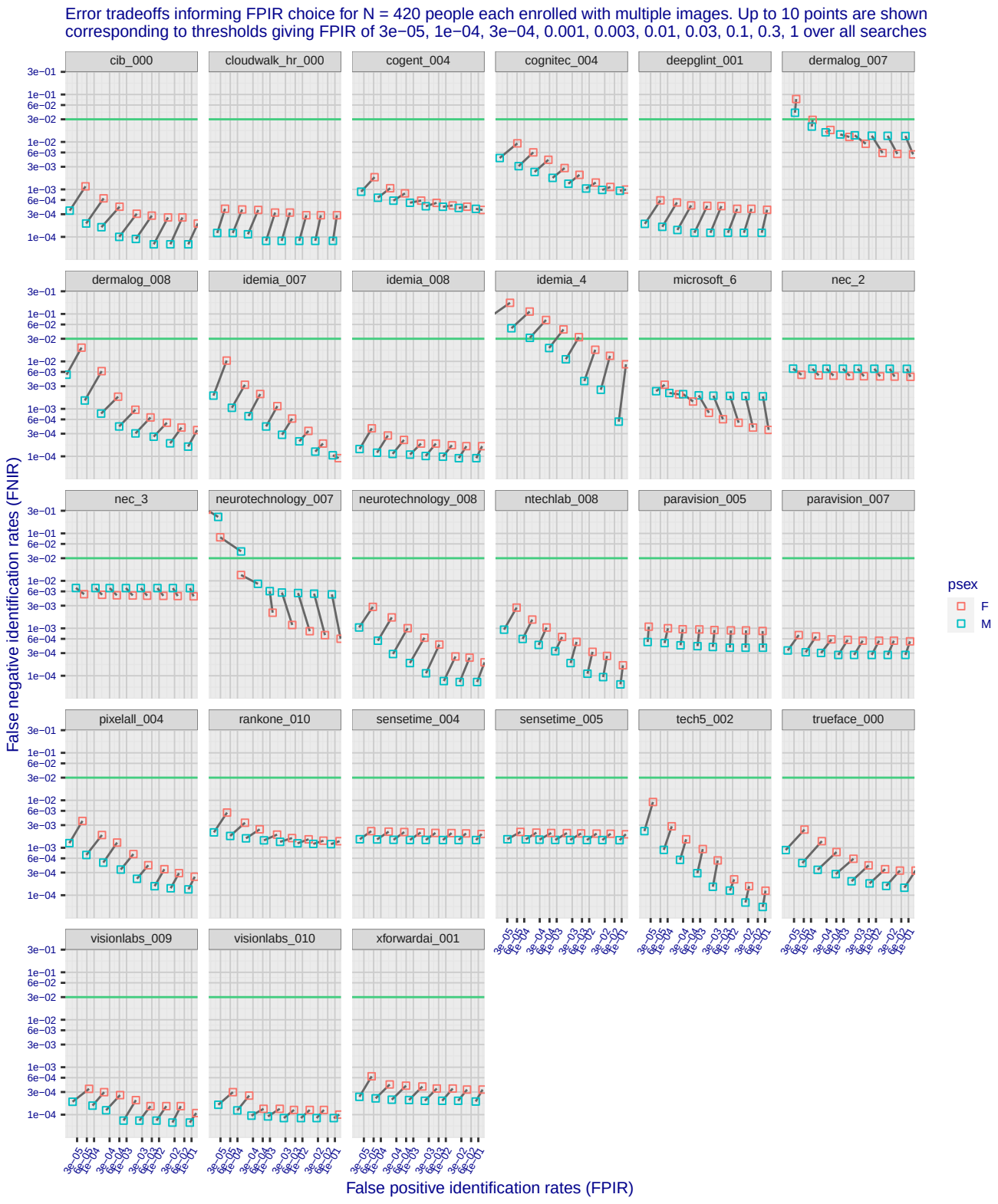


Figure 7: Error tradeoff for 26 algorithms executing 1:N searches with N = 420 people each enrolled with multiple images. The 10 point-pairs correspond to 10 possible thresholds for each algorithm. The red and blue boxes correspond to female and male travelers; their relative displacement indicates generally higher false positives and false negative rates in women. Smaller displacement indicates smaller (better) demographic differential (see NEC and Paravision-5, for example). The horizontal green line corresponds to 3% false negative goal implied by legislation in the U.S.

3.2.4 Demographics: Differentials by sex

Vertical displacement of point pairs in Figure 6 and Figure 7 reveal broadly higher False Negative Identification Rates in women than in men. This is consistent with NIST IR 8280 using other kinds of images. The cause of this is not known. Note some algorithms, including NEC-3, Microsoft-6 and Neurotechnology-7, give the opposite behavior or fairly equitable rates.

The horizontal displacement in the figures show that all algorithms give a factor of 2 or 3 times higher false positive identification rates in women; this means that women will be mismatched against a wrong identity somewhat more often than men. This will be rare but over enough flights it will disadvantage more women than men. Algorithms from Microsoft, NEC, and Cognitec give notably smaller differentials.

Figure 8 summarizes false negative rates by sex: the difference often amounts to 1 additional false negative in women than men. Note that there are many flights with zero false negative flights for both sexes.

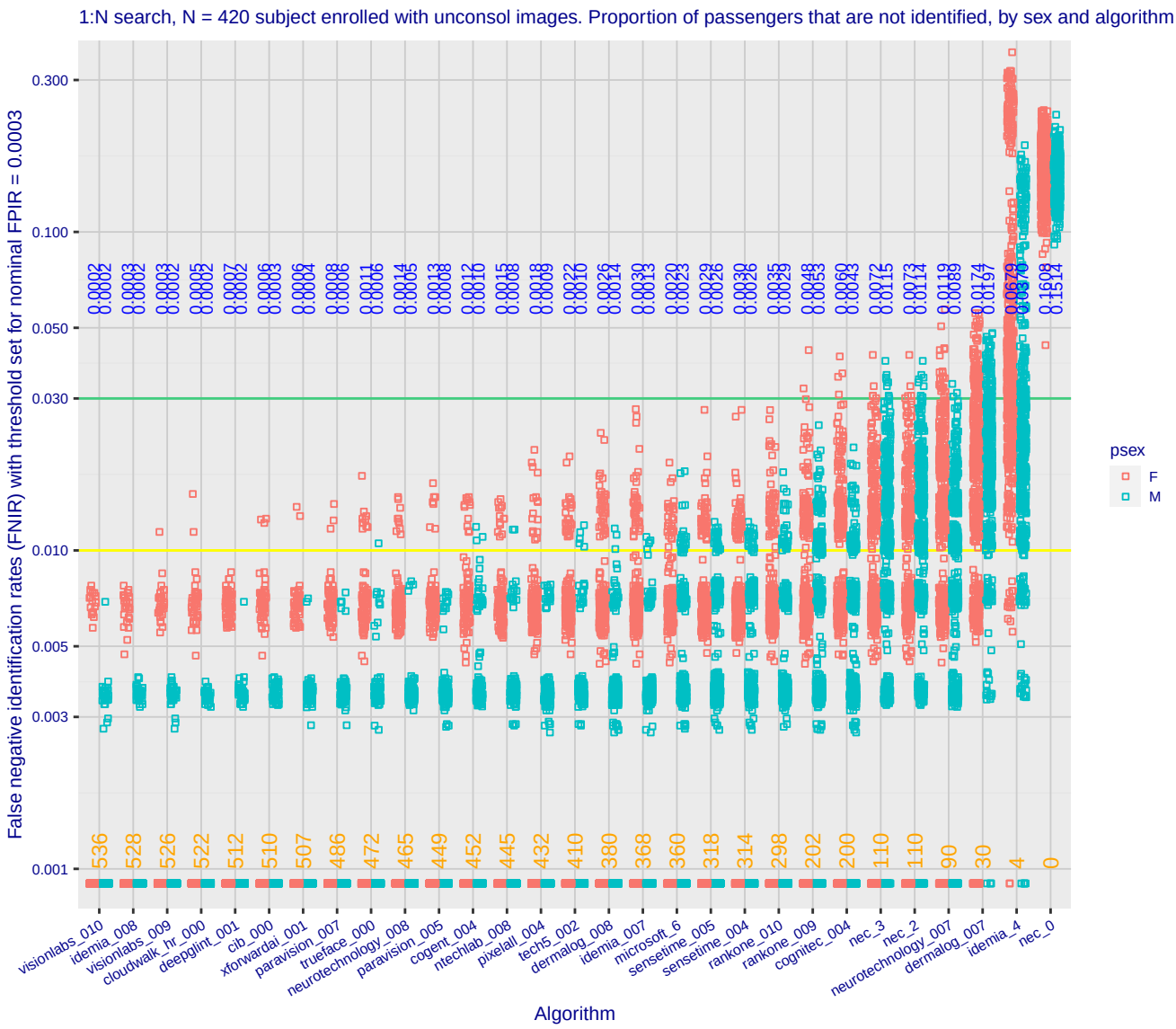


Figure 8: False negative identification rates by algorithm and sex. Each point corresponds to the boarding of N = 420 people on to one flight where each is enrolled with multiple images. The blue text is a FNIR value for that algorithm on that sex. The green line connotes a 3% FNIR (reflecting a legislative mandate). The yellow line is at 1% FNIR. The cluster of points at 0.0009 corresponds to zero errors (adjusted to plot on a log scale) - the orange text gives the number of simulated flights, out of 567, for which there are no false negative errors. The next cluster near 0.004 corresponds to 1 error out of around 210 males.

3.2.5 Demographics: False positve differentials by region

Figure 9 shows false positive identification rates by region and by sex for two algorithms from NEC and Canon. Appendix B gives analagous figures for all algorithms. We make the following comments:

1. **Magnitude:** The NEC-3 algorithms shows FPIR is quite insensitive to geography and sex with false positive identification rates estimates mostly clustered between $2x10^{-4}$ and $7x10^{-4}$ In constrast Canon's cib-000 algorithm gives FPIR estimates between $7x10^{-5}$ and $2x10^{-3}$. As noted in NIST Interagency Report 8280 the NEC-3 algorithm is taking steps to normalize false positive rates in one-to-many searches.

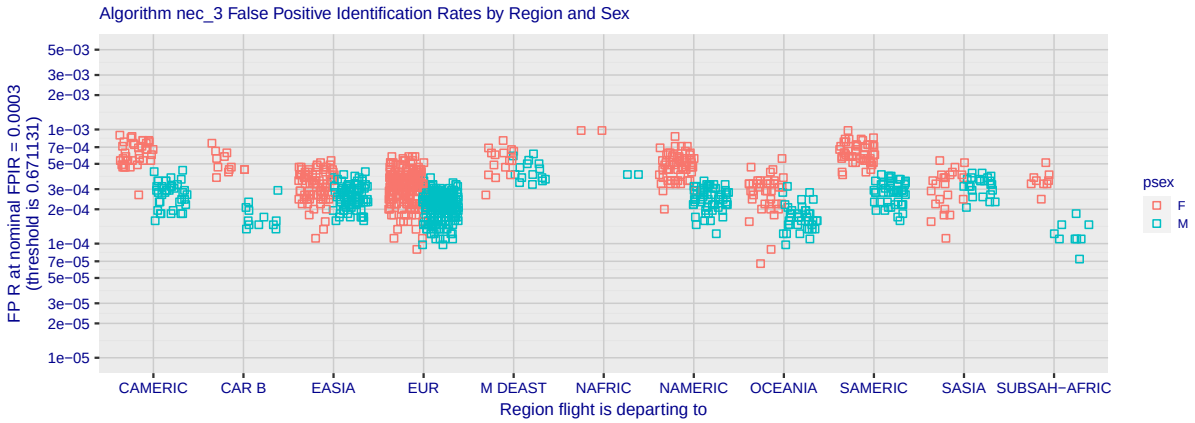
2. **Sex:** It is very common across algorithms for women to give higher FPIR than men. The NEC-3 algorithm gives

broadly the smallest differential in FPIR value – cf. the y-axis in the next Figure.

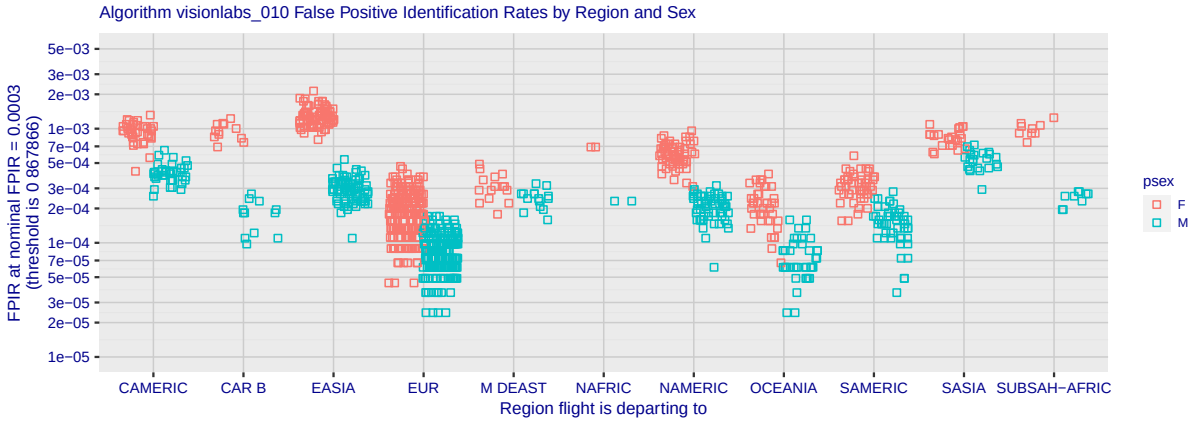
3. **Region:** False positive identification rates are commonly an order of magnitude higher in Asian women than in European men. For Canon's cib-0 algorithms there is a factor of 30 variation.

As always, with the observation of a demographic differential, the question is "what is the impact"? The overall target FPIR was 0.0003, achieved by setting an algorithm specific target. The worst upside departure from that is Canon's cib-1 algorithm (see Appendix B) which gives FPIR for Asian women near 0.003. This FPIR, 1 in 333, is still low but implies between one and two false positives per flight boarding – these would likely manifest as an in-gallery false match described in section 1.5. This may be an acceptable cost, but does constitute a disadvantage for Asian women attempting to record their departure from the United States.

The error tradeoff characteristics of figure 11 are, for some algorithms, quite flat implying that even lower false positive identification rates could be targetted (by increasing the threshold) without great adverse implications for false negative identification rates.



(a) NEC-3



(b) VisionLabs-10

Figure 9: For two algorithms, each point shows a false positive identification rate estimated by running c. 120 000 searches against that flight's gallery. Red and blue connote male and female enrollees. Analogous figures for all algorithms appear in Appendix B.

4 Scaling One-To-Many Authentication to Larger Populations

4.1 Motivation

Thus far we have discussed the use of 1:N face recognition for recording the exit of travelers while boarding an aircraft. A PCA can appropriately limit enrollment in FR galleries to just the population expected on the flight. This data minimization reduces mis-match possibilities. However, the travel industry has articulated a vision for paperless travel. In its simplest form, this starts when a traveler authenticates to an authoritative travel document (passport) using a one-to-one biometric verification of a live photo and then proceeds through an airport's "touchpoints" such as the TSA line and airline lounges and aircraft boarding, without presenting a boarding pass. Instead, the traveler engages a camera which submits a photo as a query into a database of individuals expected and authorized to proceed. For example, such a system could be fielded at a security checkpoint. In such cases many more people would need to be enrolled into the face recognition engine than at a departure gate – for example, all people expected in the airport during a time window extending from a few hours before their respective flights to the time of expected or actual departure. The number of individuals could readily extend into the tens of thousands, and more if airside locations would additionally recognize inbound passengers (e.g. buying in duty-free shops).

4.2 Background

The dependence of recognition accuracy on enrolled population size is well known. Qualitatively, as enrolled population grows any given search has a greater possibility of a false match. Such outcomes can occur for two types of traveler.

1. An illegitimate traveler – someone who is not expected in the airport – makes a presentation to the camera in attempt to pass the checkpoint. This succeeds if the traveler matches any enrolled identity with a comparison score above threshold.

2. A legitimate traveler – someone who is expected at the touchpoint – presents to the camera but matches an identity other than self. This may be inconsequential at a TSA line, but would be consequential in a hypothetical duty-free store application of this approach should the biometric result allow purchase without further authentication.

The rate at which false positives occur is the false positive identification rate (FPIR). In a biometric test, FPIR is estimated by conducting non-mated searches into an enrolled population. FPIR is stated as the number of searches resulting in a false positive divided by the number of non-mated searches. How does FPIR scale with the number of enrolled identities? There are two classes of face search algorithms: Class A is those that implement a 1:N search as N 1:1 comparisons followed by a sort operation, and Class B is comprised of everything else including those that implement some more complex search strategy.

- **Class A** algorithms are expected to give a FPIR increases with the number of enrolled identities. It may increase further also if those identities are enrolled with several images each. Given a system in which N people are enrolled, with one image each, a standard binomial model gives

$$\text{FPIR}(N, T) = 1 - (1 - \text{FMR}(T))^N \tag{3}$$

where the system owner sets the threshold T, and has an estimate of FMR(T), the false match rate in purely one-to-one comparisons. For small FMR, this approximates to

$$\text{FPIR}(N, T) = N\text{FMR}(T) \tag{4}$$

implying that the one-to-many false positive hazard grows linearly with N.

- **Class B** one-to-many algorithms are those that do not implement 1:N search using N 1:1 comparisons - these can include fast-search algorithms (using trees, indexes etc) and those that normalize scores across some or all of the gallery entries. These algorithms may not exhibit the (near) linear dependence of equations 3 and 4. This can occur for other reasons also. Some algorithms adjust comparison scores to the database size such that FPIR becomes approximately independent of N. The NEC-3 and Idemia algorithms exhibit such behavior (see FRVT Part 2 and its report cards). This relieves the system owner of the need to configure thresholds for the given population size. A system owner might consult vendor documentation, or consult NIST's FRVT Part 2 report which documents the dependence of FPIR on N and T.

4.3 Simulation of Large-N Accuracy

4.3.1 Experimental design

We repeated the EXIT simulation given previously but instead of enrolling N = 420 individuals with one ENTRY image, we mixed in a further 41580 such images from a disjoint population selected without regard to demographics. The result is a set of 567 galleries, each with N = 42 000 individuals. This population size is somewhat larger relative to the number of international passengers appearing daily in large U.S. airports in 2019.

4.3.2 Results

Figure 10 shows the number of false negatives expected when using the algorithm named on the horizontal axis to search three different kinds of galleries. The first enrolls 420 people with a single PCA ENTRY image; the second enrolls those same people with all prior PCA ENTRY encounters; the last enrolls 42 000 people with a single image. The kind of gallery is encoded by the shape. The vertical position of each point is the mean (over 567 regional galleries) of the number of false negatives when 420 test subjects are searched. The color of each point encodes the fraction of all 567 trials that give three or fewer false negatives.

We make the following observations:

1. The number of false negatives is higher with N = 42 000 than with N = 420, as expected.
2. Some algorithms nevertheless give only modest increases in the number of false negatives. In the most accurate case, the mean number of passengers being rejected would be below 1% (4/420), and more than 75% of flights (trials) would have three or fewer false negatives out of the 420 people making attempts.
3. Other algorithms give substantially higher false negative rates – the graph shows FNIR approaching about 8% (34/420) for legacy algorithms.
4. Note that this analysis does not consider variance around the point estimates, nor sex or regional differences.

4.3.3 Discussion

Large enrolled populations require algorithms to be configured to operate at lower false match rates – following Equation 4 an increase in N from 420 to 42 000 will necessitate a 100-fold FMR reduction to maintain constant FPIR. This puts a premium on algorithms that maintain relatively low FNIR at lower FPIR.

By inspecting Figure 11, for N = 42 000, all algorithms except some from Cloudwalk, Deepglint, Idemia, NEC, Paravision, Sensetime, Visionlabs and X-ForwardAI cannot maintain FNIR below 0.03 so, depending on FPIR, would not be meeting a legislative mandate for FNIR < 0.03 and TPIR > 0.97.

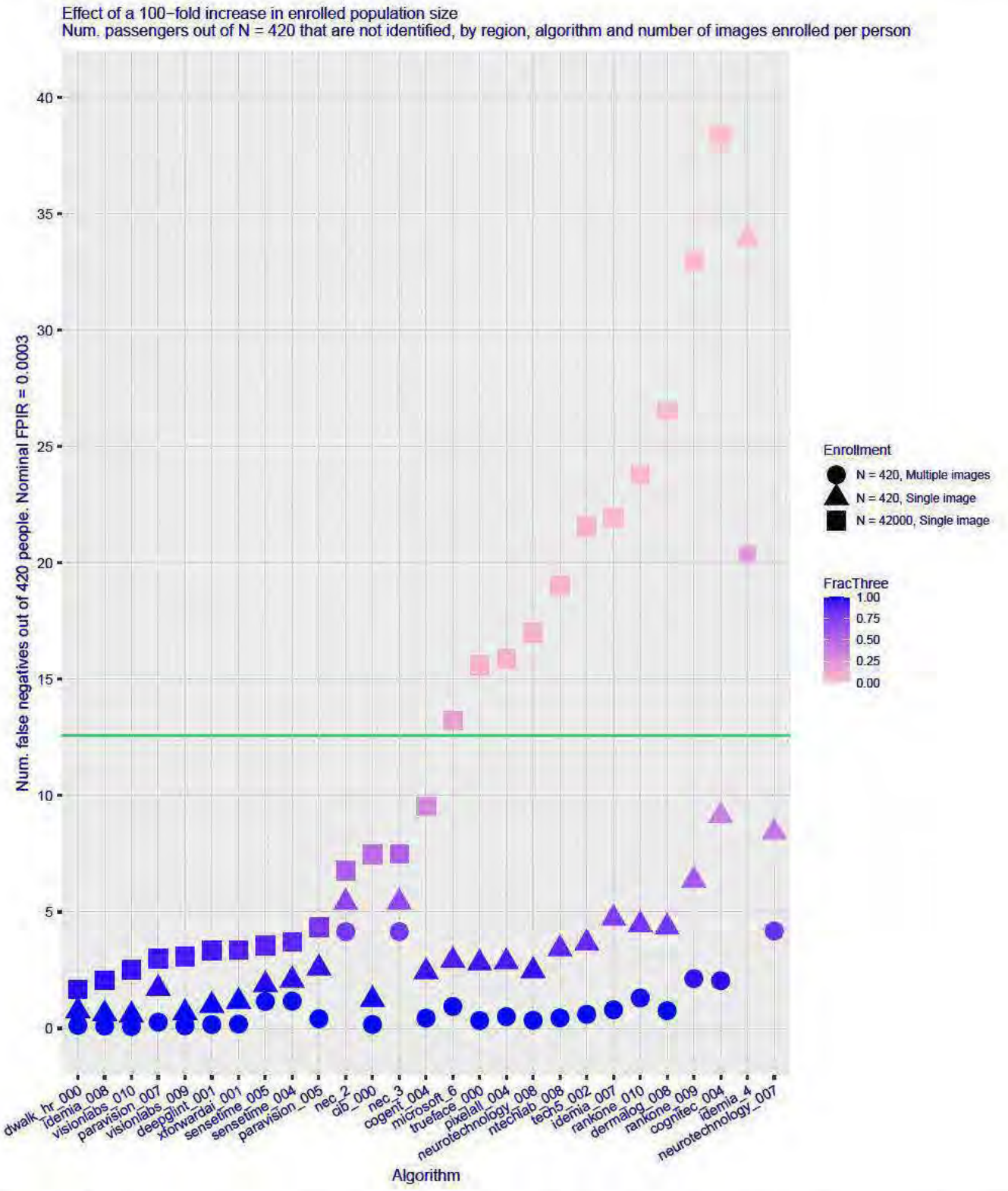


Figure 10: Comparing the number of false negatives expected when N = 420 people are searched against galleries containing imagery from enrolled populations of size of N = 420 and N = 42 000. The threshold is fixed in all cases to produce a false positive identification error rate of 1 in 3 333 (FPIR = 0.0003). The threshold value for each algorithm will usually be higher for the larger gallery to maintain the same false positive likelihood. The horizontal green line corresponds to a 3% false negative rate.

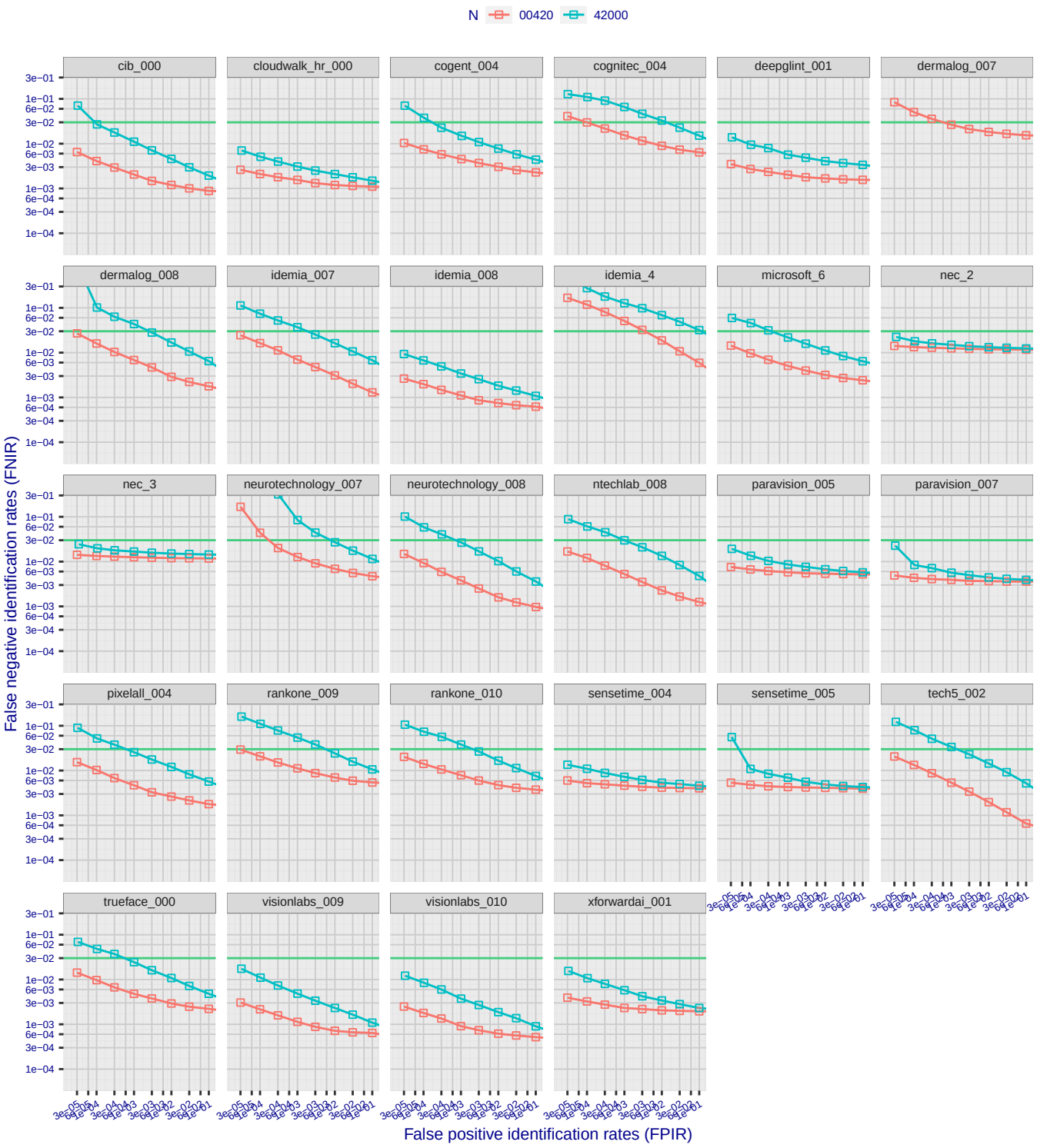


Figure 11: Error tradeoff characteristics for twelve algorithms conducting identical sets of searches into galleries of size N = 420 and N = 42 000. The horizontal line corresponds to a 3% false negative identification rate. The left side of each panel is relevant to the more "lights-out" use of FR in positive access control and EXIT facilitation; the right side of each panel corresponds to high false positive identification rates for investigative uses of FR where humans review candidate lists. A flat profile confers the advantage of being able to run at lower FPIR without much elevation in FNIR.

5 Factors That Render Accuracy Estimates Approximate

The result in this report do not constitute an answer to the questions "how well does a particular TVS work", "does TVS satisfy a 97% legislative verification mandate" or "what is the accuracy of a PCA's EXIT solution". Why? Because the questions are different and because the tests we have reported here, while extensive, depart from the intended and desirable tests as follows. For each factor discussed, we note in blue the expected effect on accuracy.

1. **Airline re-direction of passengers.** During the EXIT pilot, airlines diverted some customers toward the legacy paper-based boarding process. This was particularly true when boarding was proceeding slowly or when cameras of the network to TVS were malfunctioning. This is not expected to bias accuracy in an offline test either way, but would lead to complication in using TVS logs to measure accuracy.

   The population so diverted was sometimes not random – very tall or short travelers, and those with children, would be directed from the FR line. To the extent that this occurs, and to the extent that the NIST EXIT collection is not itself affected by this, our estimates of accuracy may be too high.

2. **Algorithms:** NIST does not have access to the actual algorithms deployed in TVS systems. Instead this study uses prototypes submitted to the one-to-many search track of the FRVT. These prototypes are identified using a name and a number. For example, "NEC-3" is from the NEC Corporation, and the three is simply a sequence number of algorithms sent to NIST. NIST is unable to confirm whether any prototype in FRVT has ever been deployed. Indeed a developer may make decisions on whether to productize a prototype on the basis of FRVT-derived technical information. In any case, a developer will maintain their own versioning designations. NIST is not provided with copies of operational algorithms.

3. **Active development**: Given persistent improvements in accuracy, as documented in FRVT, it is incumbent on end-users to instantiate a "technology-refresh" procedure so as to realize accuracy gains. Note that results in this report for 2018-era algorithms are likely out-of-date. Thus, for any given developer, it is likely that higher accuracy is available than is estimated here.

4. **Algorithm post-processing:** Accuracy will change if any software is used to post-process candidate lists produced by the algorithm. Conventionally the face recognition algorithm issues a candidate identity and a similarity score, which is compared to a system-wide threshold. If post-processing is used to re-score or re-rank then its effect on both false positive and false negative identification rates should be measured by comparing with that available from the raw candidate lists alone.

5. **Image data:** International travel has long been predicated on presentation of a passport. With e-Passports it is common for passport images to be retrieved and used for 1:1 verification of the the traveler. If on ENTRY those images are retained by the PCA , they can be used in downstream EXIT face recognition processes. The same applies to visa portraits collected as part of a visa application.

   (a) We did not have passport or passport-equivalant images for use in this study. These include visa images of various travelers and passport photos of the citizens persons. Instead we used airport arrivals hall photos with reduced quality. To the extent that a TVS makes use of high-quality passport and visa images, the accuracy values reported here are likely to be worse than for a system for which such images are available.

   (b) In this study we used an extract of a much larger corpus provided to NIST in May 2019. These images were anonymized and accompanied by limited metadata. The set included images labelled *exit*, and *ENTRY*. The former were collected in airport departures. A few of the exit images appear to have been collected from

persons in a vehicle, as could occur at a land-border. This factor will tend cause our accuracy estimates to differ from those of an operational TVS.

(c) Our exit images were collected in 2018 and the first four months of 2019. We assume that subsequent cameras, and their refined deployment by airlines, will yield improved images today compared to those used in this study, so we would expect improved accuracy over that noted here.

(d) Moreover, our exit images are not accompanied by camera make and model information, nor flight manifest information. It was therefore not possible for NIST to exactly reconstruct "a flight" – instead we pooled all exit images as probes searched against each gallery. Our search set therefore pools exit images from quite different cameras and locations (airports). We are therefore unable to compare cameras and collection sites. From observations made during site visits, we note markedly different approaches to the quality-speed tradeoff. We expect therefore that our accuracy estimates have reduced variance compared to that from a TVS.

(e) If exit images are retained, even for a short period, they may be useful in offline "after-hours" accuracy estimation. For example, images from one flight could be used to make non-mated searches into a gallery of another flight, so as to estimate FPIR.

(f) Our galleries were constructed to hold people from one travel region as inferred from the nation that issued their travel document. This means that the galleries in this document will contain people who never flew together.

6. **Homogenous galleries.** Our practice of making galleries from people holding travel documents from countries in the same region of the world probably means that false positive rates are higher than if the galleries had been composed of a more mixed population. This practice would tend to depress our accuracy relative to those in TVS.

7. **Presence of images active attack.** It is possible that some of the captured images are from a presentation attack that went undetected, for example using a face mask. The occurrence of this is considered to be very small. Note that since we didn't have passport images, we do not expect the dataset to contain morphed images. While this is increasing possibility operationally, it can be averted by live, trusted capture of images as in a primary passport control lane.

Appendices

Appendix A Figures summarizing false negatives for each algorithm

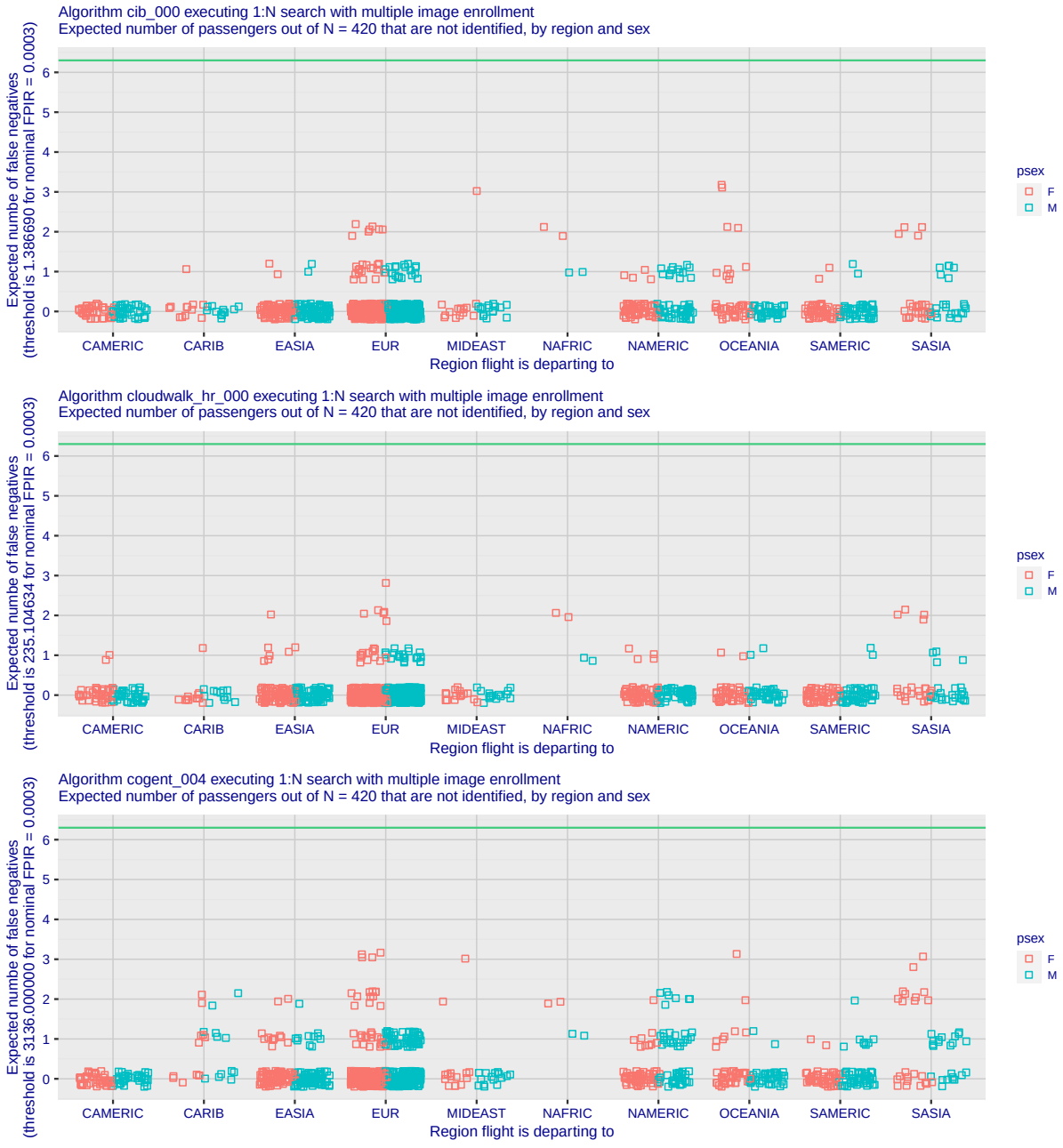


Figure 12: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

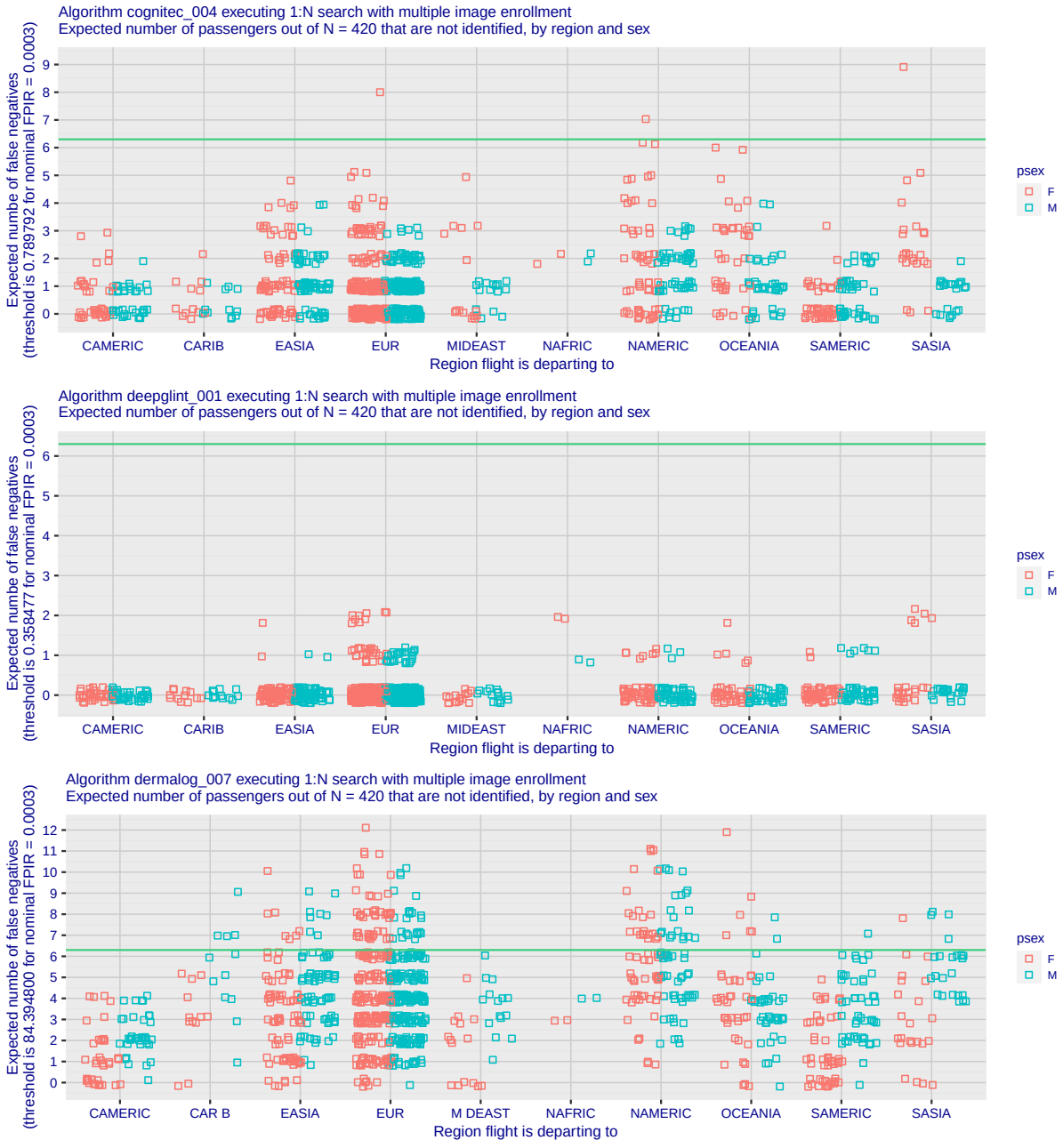


Figure 13: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

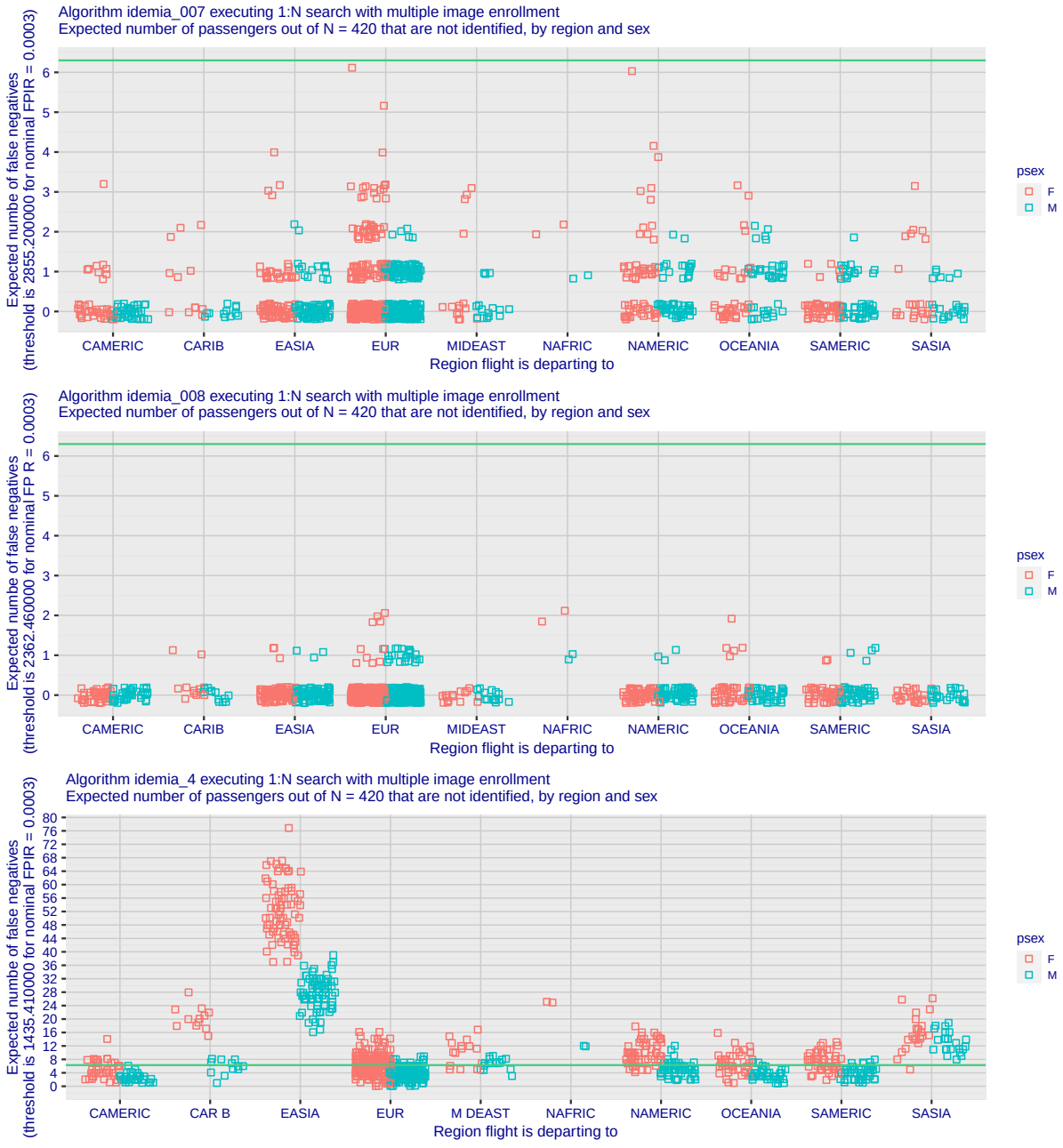


Figure 14: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

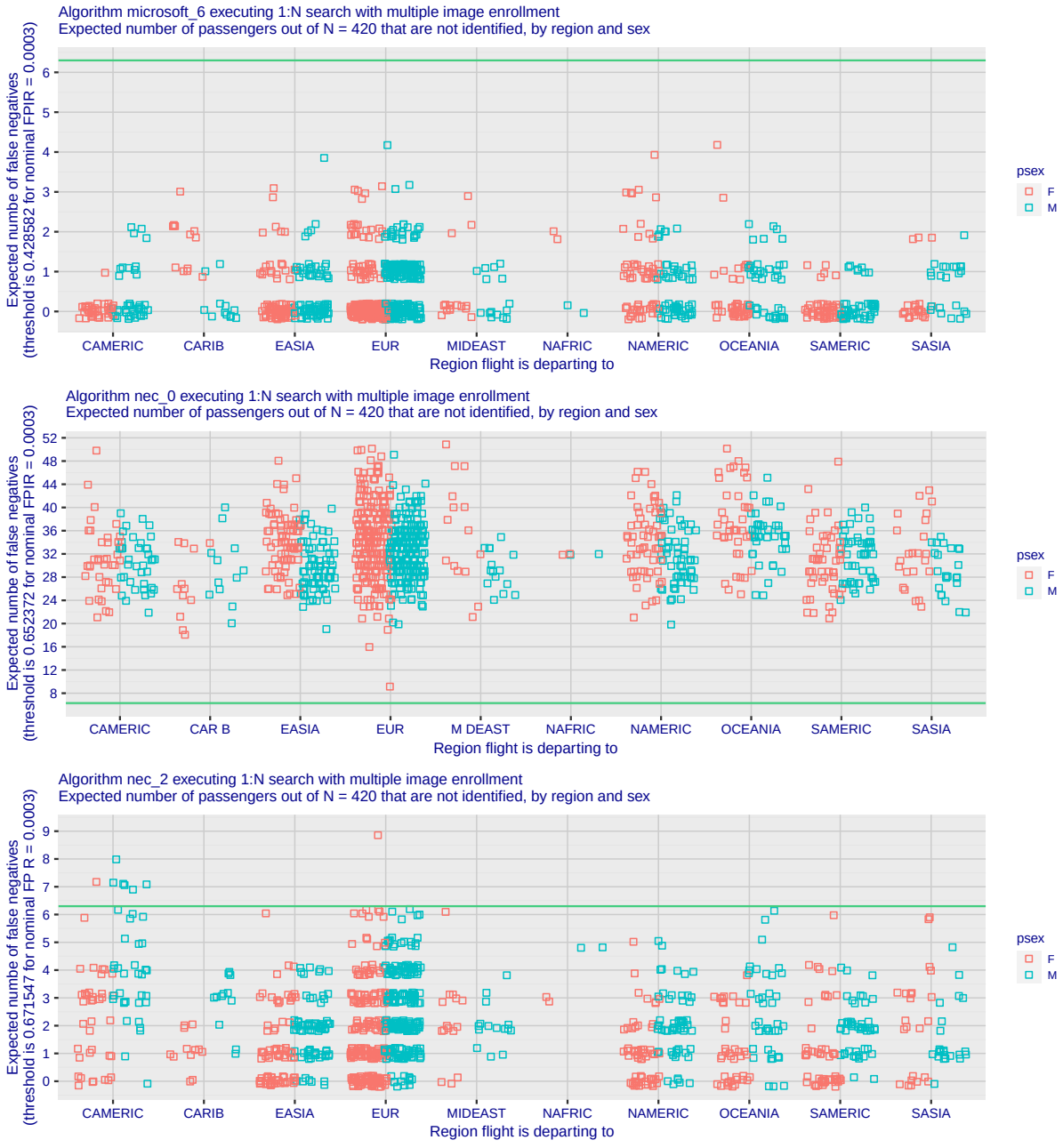


Figure 15: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

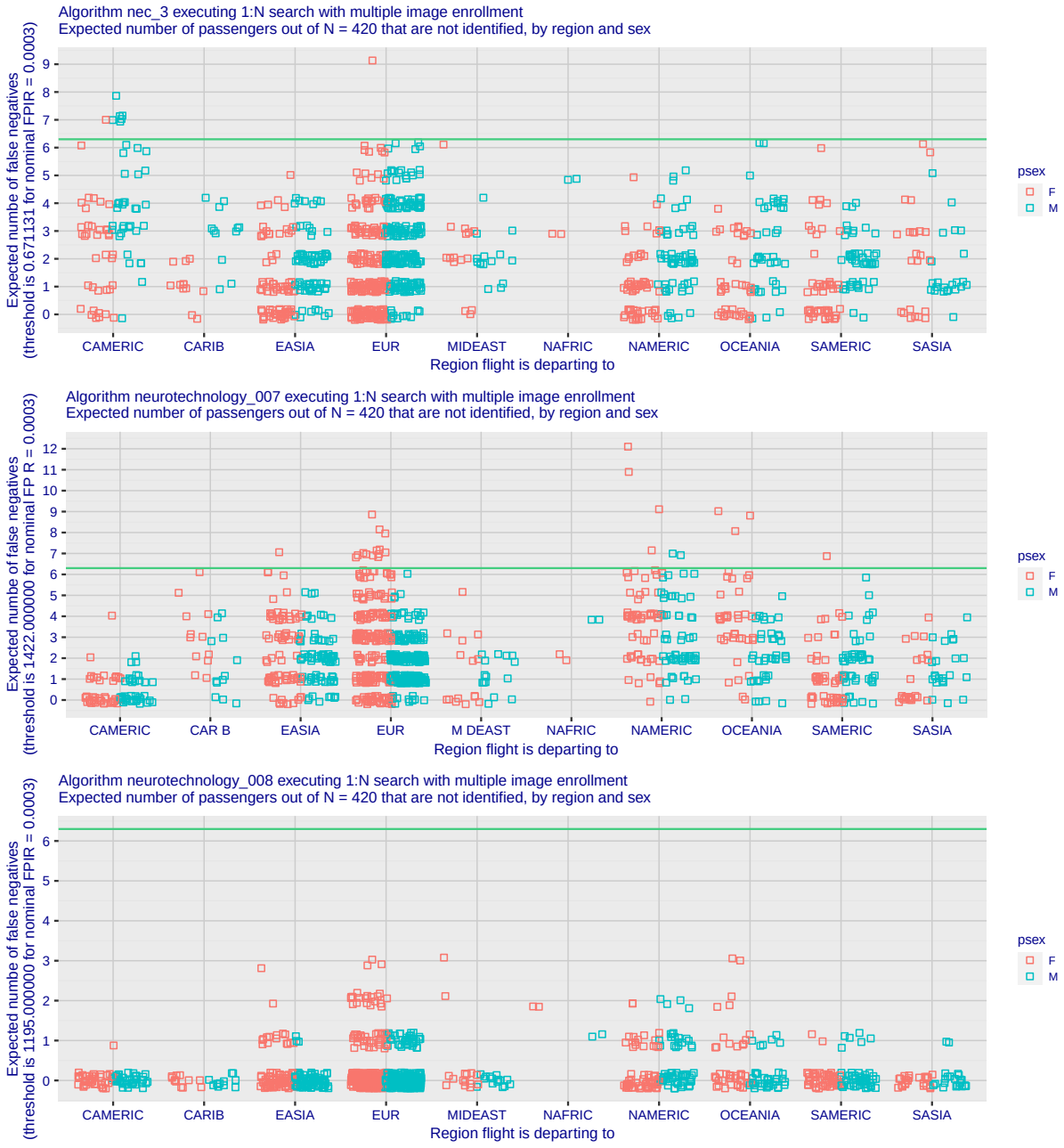


Figure 16: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

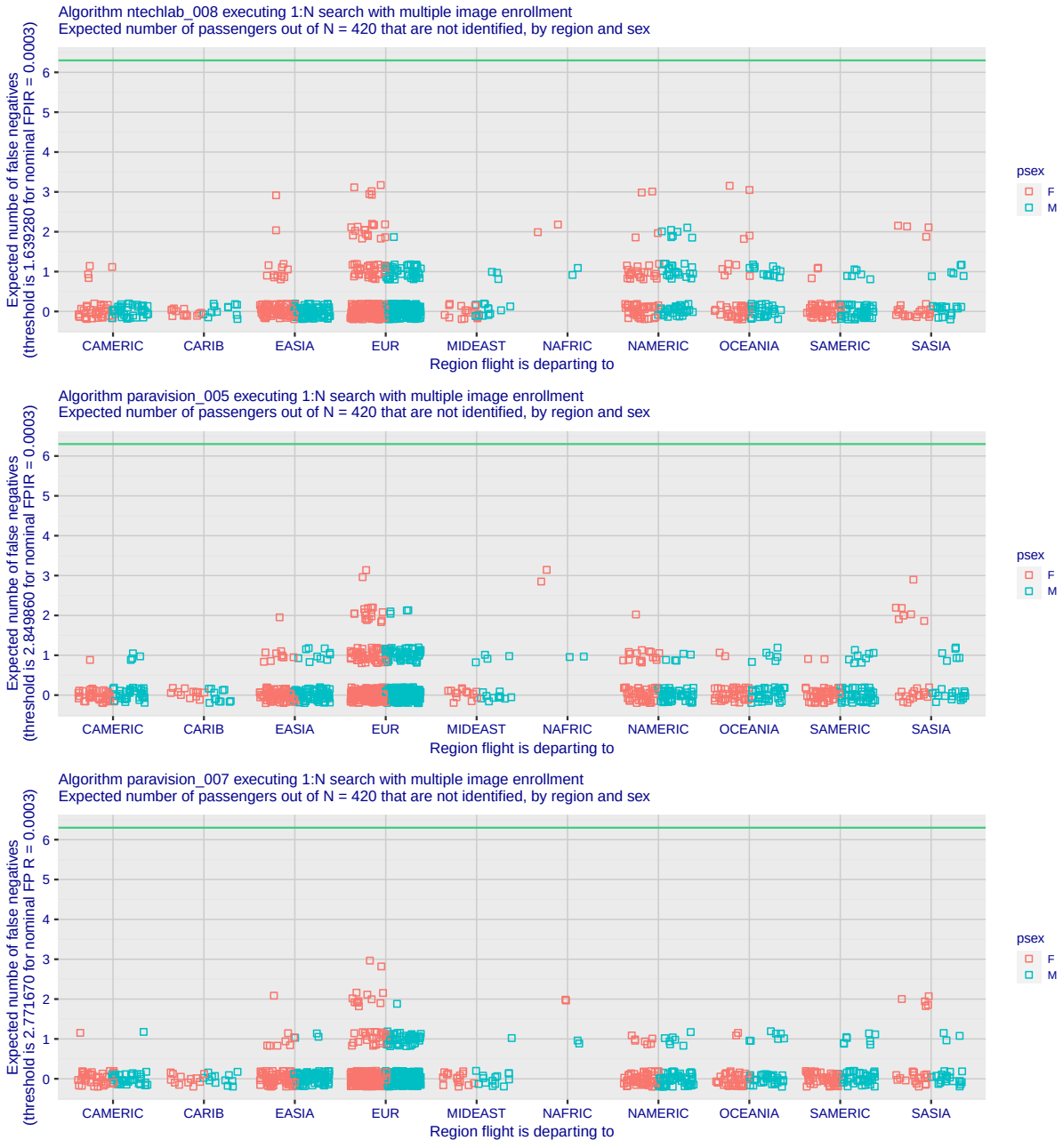


Figure 17: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

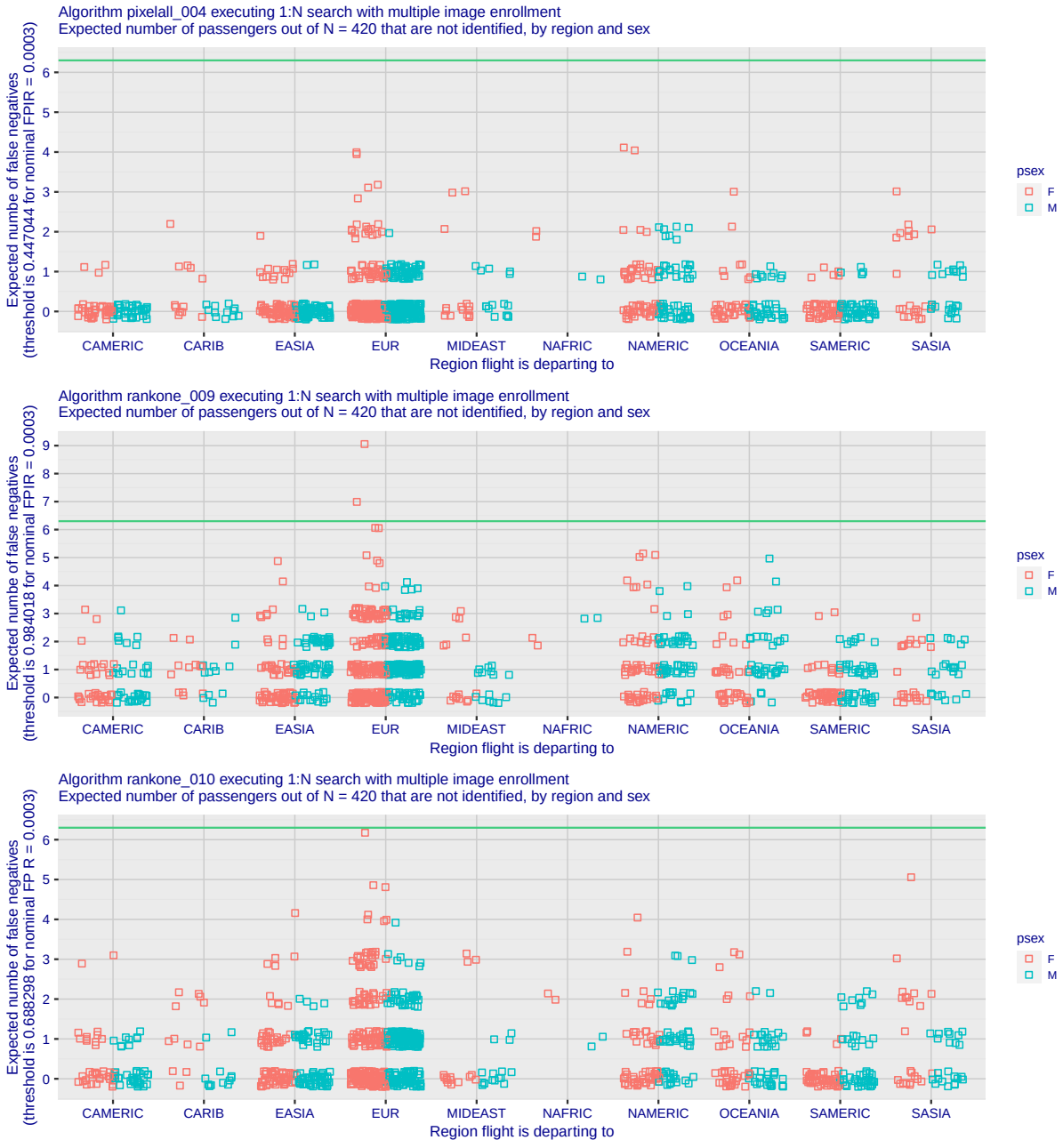


Figure 18: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

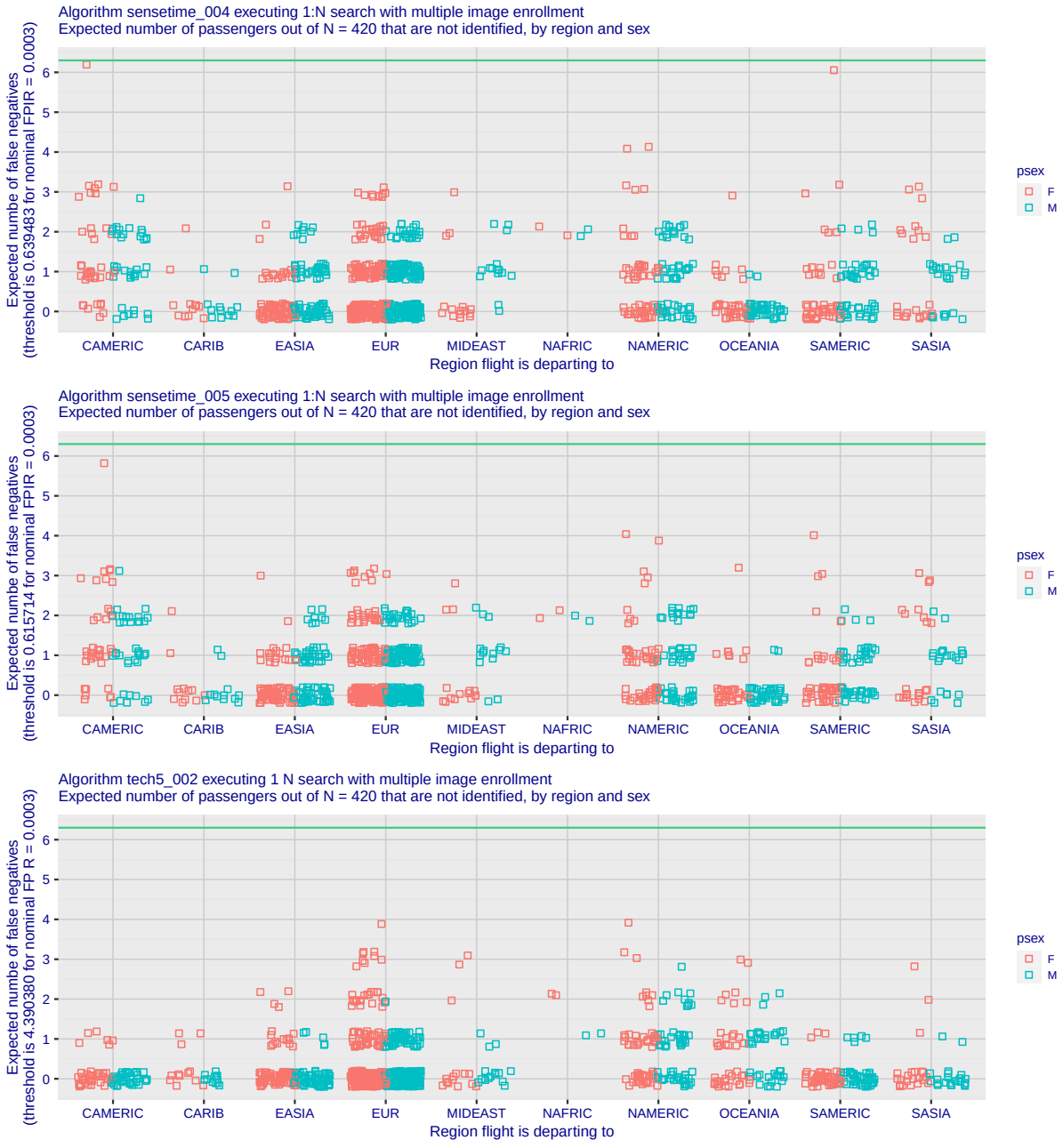


Figure 19: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

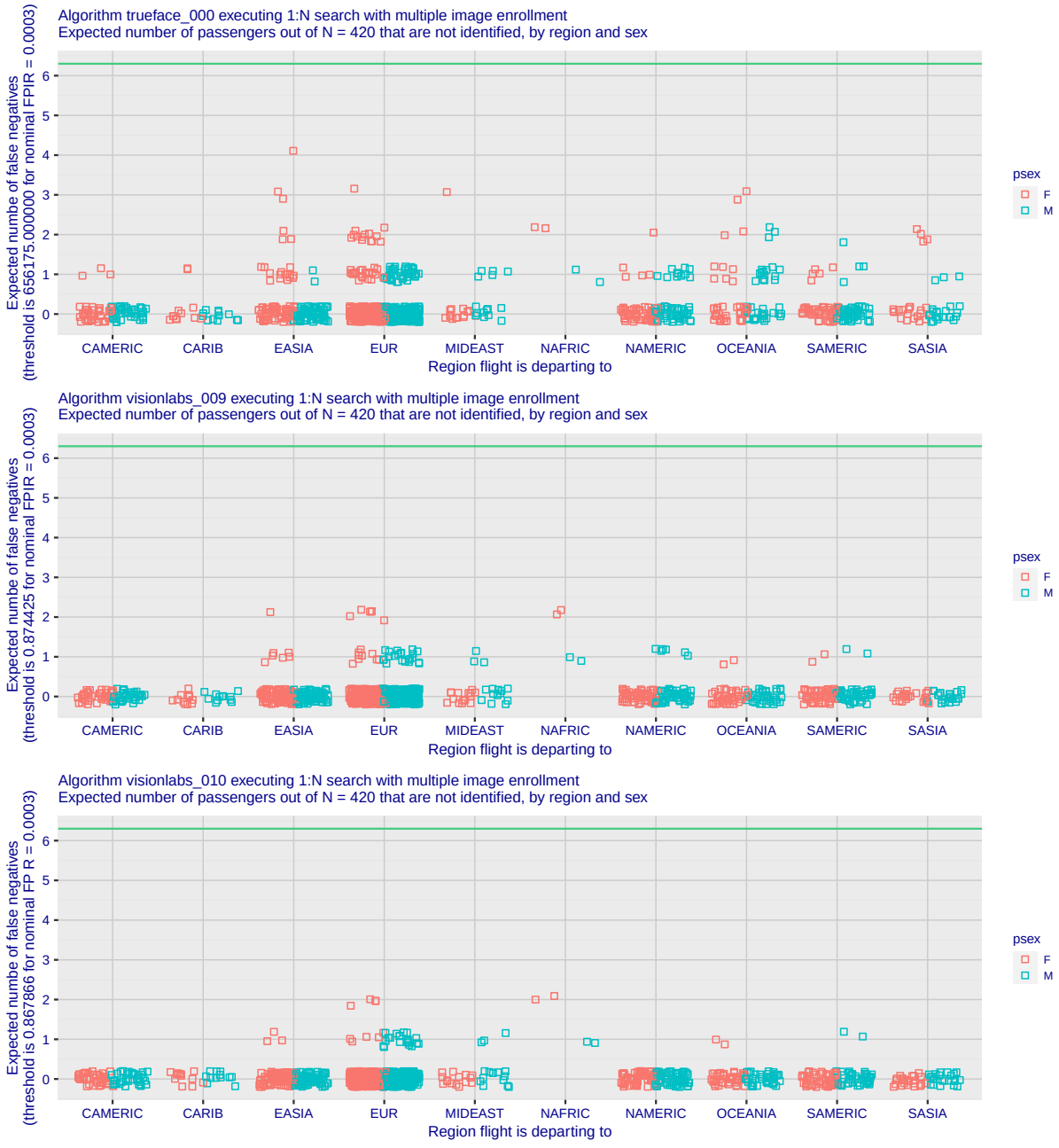


Figure 20: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

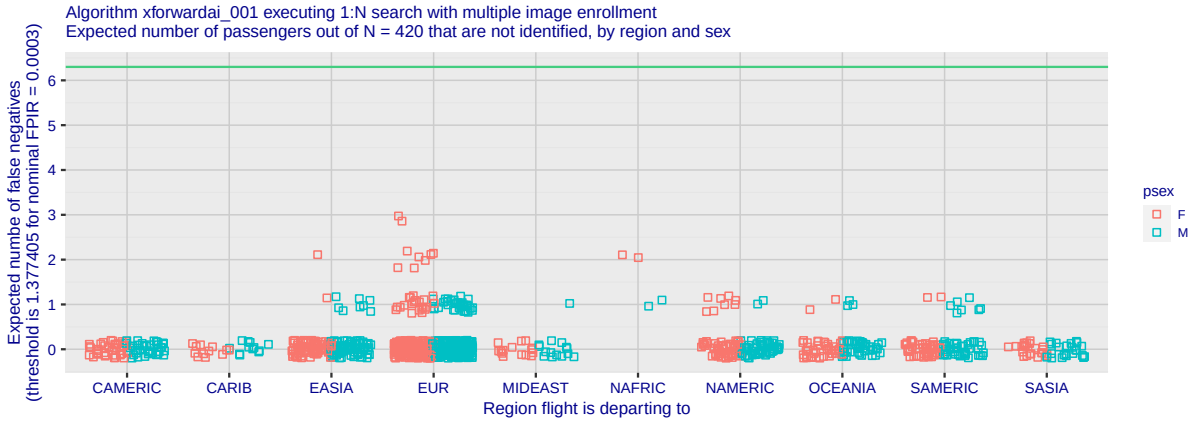


Figure 21: For the eleven regions and two sexes, each point give the expected number of false negatives for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The numbers are stated by scaling measured numbers of false negatives to 210 per sex. The points' positions are jittered horizontally and vertically to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have. The number of individuals in the gallery is exactly 420.

Appendix B Figures summarizing false positive identification rate for each algorithm

Figure 22: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

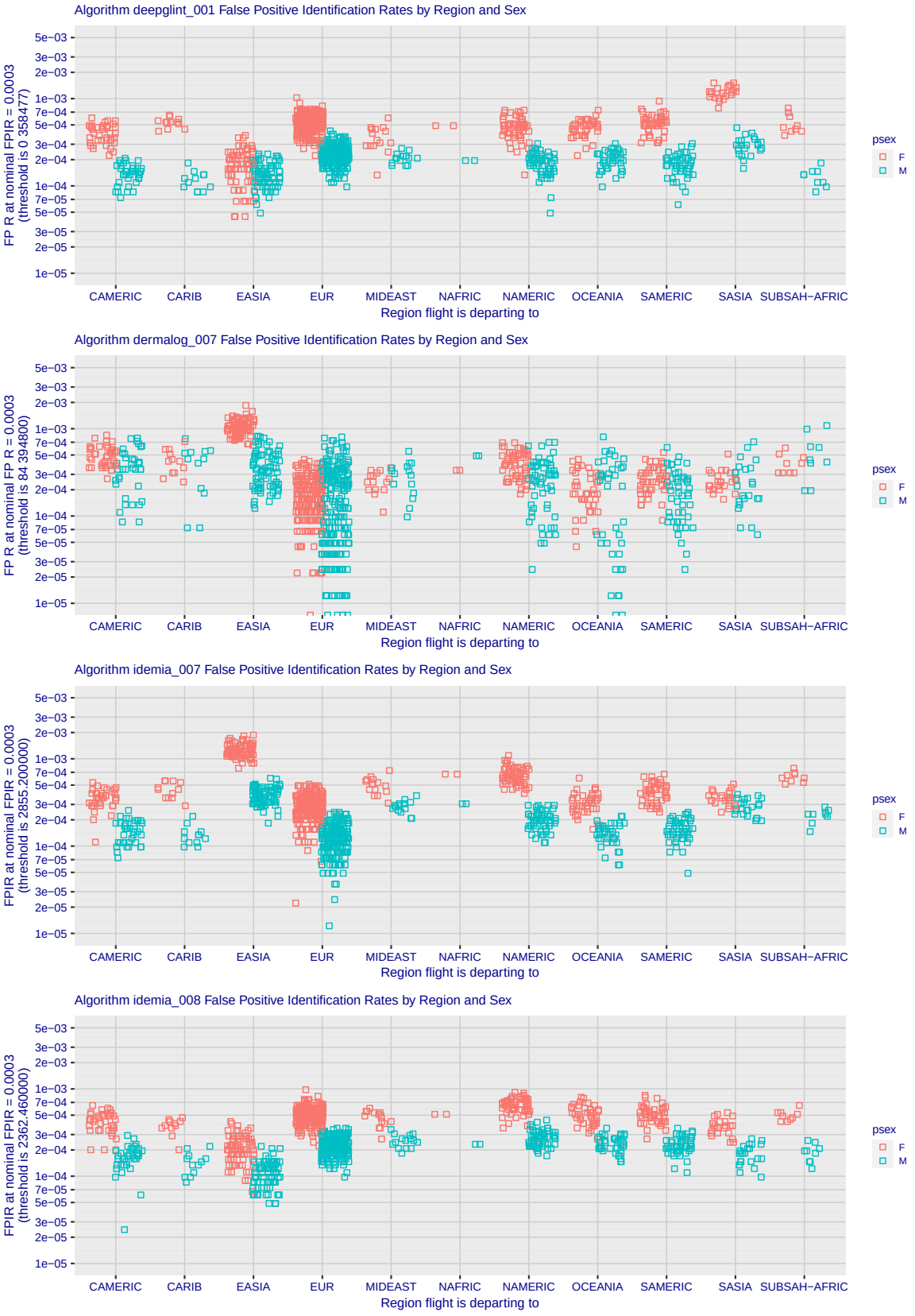


Figure 23: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

Figure 24: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

Figure 25: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

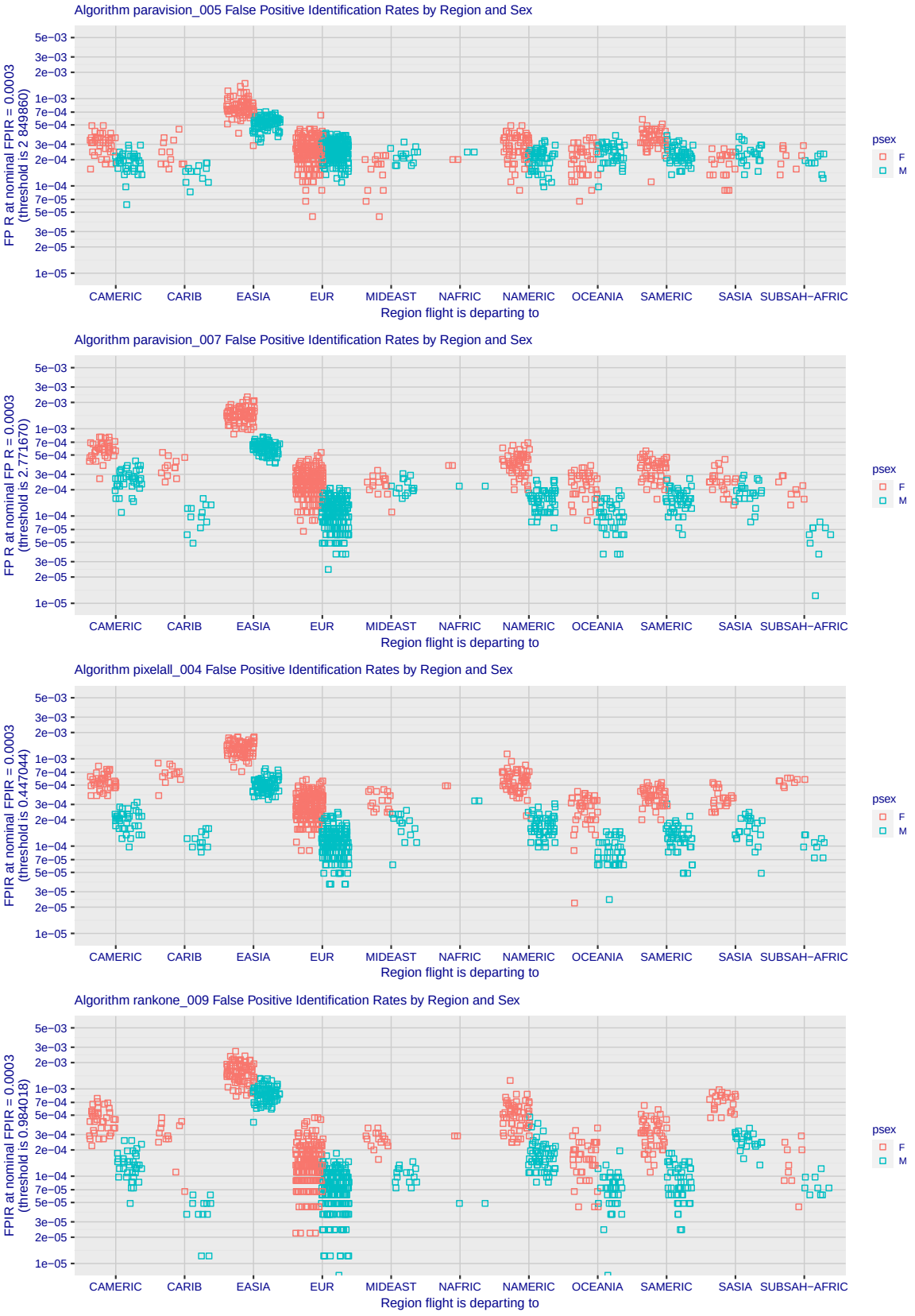


Figure 26: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

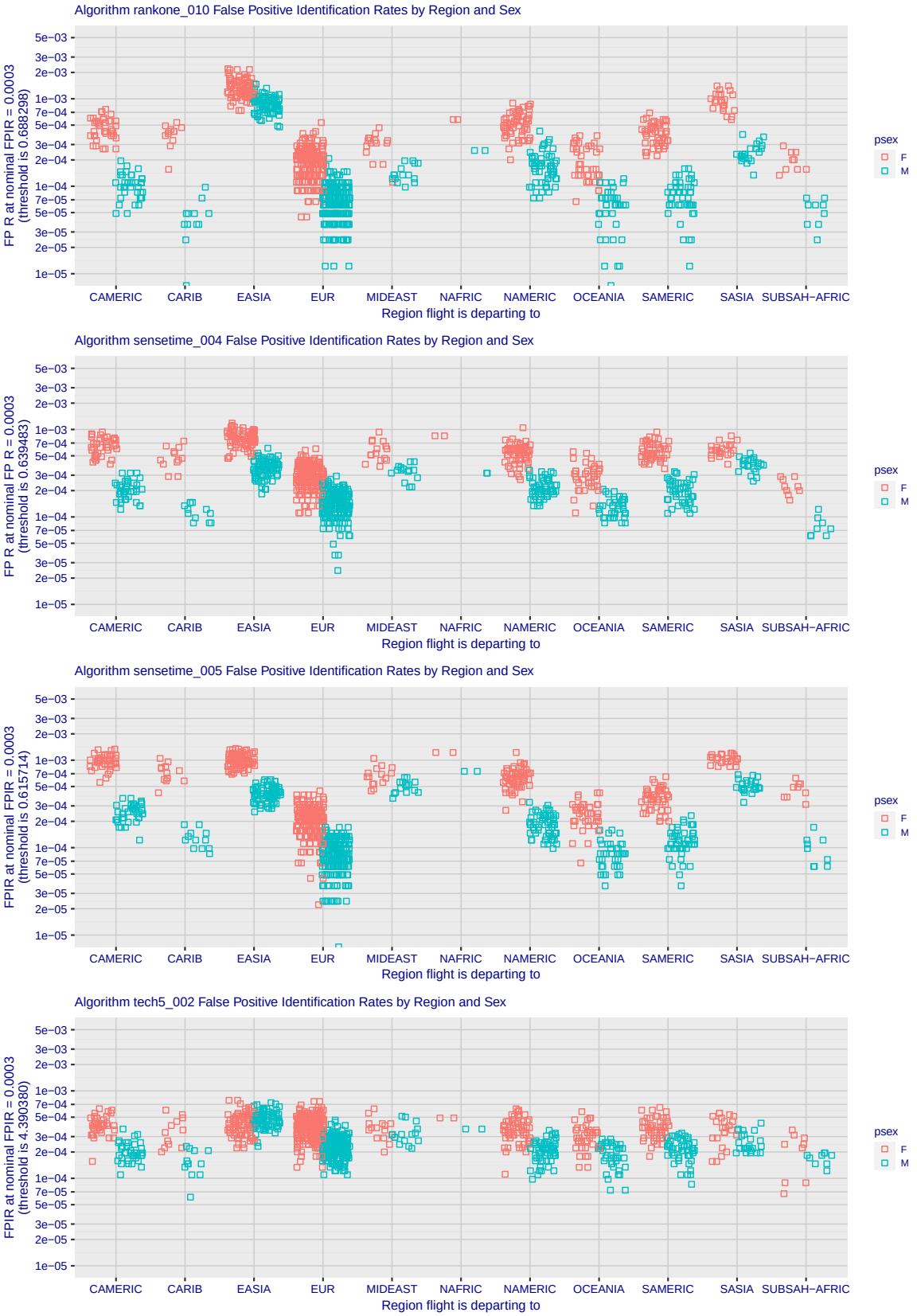


Figure 27: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

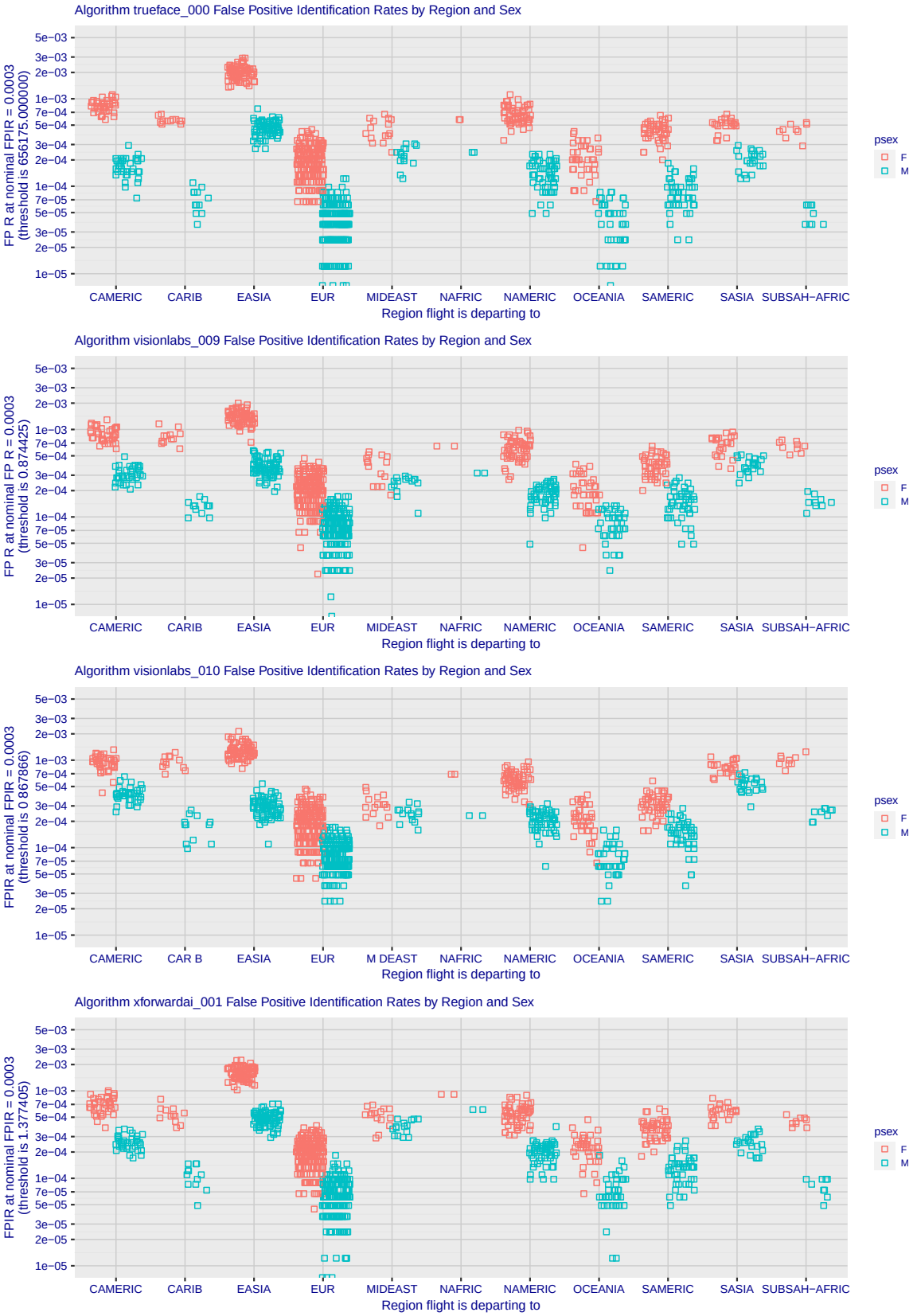


Figure 28: For the eleven regions and two sexes, each point give the false positive identification rarte for a simuliated flight in which 420 passengers, 210 men and 210 women, attempt boarding after being enrolled with multiple images each. The points' positions are jittered horizontally to mitigate over-plotting invisibility. There are many more flights to Europe, particularly, and East Asia simply because of their representation in the EXIT image corpus we have The number of individuals in the gallery is exactly 420.

References

[1] Patrick Grother, Mei Ngan, and Kayee Hanaoka. Face recognition vendor test (frvt) part 2: Identification. Interagency Report 8271, National Institute of Standards and Technology, Home: https://pages.nist.gov/frvt/html/frvt1N.html, September 2019. https://doi.org/10.6028/NIST.IR.8271.

[2] Patrick Grother, Mei Ngan, and Kayee Hanaoka. Face recognition vendor test (frvt) part 3: Demographic effects. Interagency Report 8280, National Institute of Standards and Technology, Home: https://pages.nist.gov/frvt/html/frvt11.html, December 2019. https://doi.org/10.6028/NIST.IR.8280.

CBP One

Modernizing the Travel Experience

U.S. Customs and Border Protection
October 2020

Executive Summary

CBP One is a mobile application that is intended to act as an intuitive single point of entry for travelers and stakeholders to access CBP services and mobile application such as ROAM, I-94 Entry/Exit and Appointment feature. Through a series of intuitive questions, it will guide each type of user to the appropriate services based on their needs.

CBP ONE ALLOWS USERS TO:

- ✓ Access appropriate CBP services for your specific needs
- ✓ Report your Arrival via Air or Sea
- ✓ Schedule inspections for perishable cargo
- ✓ Request inspection of Hand-Carry Permit Items*
- ✓ Enable user to apply for and quickly reference I-94 applications for US Entry
- ✓ Biometrically validate US Exit and close out I-94 forms
- ✓ Pay I-94 fees or User Fee Decals
- ✓ Apply for the Global Entry Program
- ✓ Apply for and manage Cruising Licenses
- ✓ Chat or Video chat with CBPOs (when applicable)
- ✓ Receive push notifications for statuses on requests/submissions

BENEFITS

- Eliminates need for multiple CBP applications
- Provides live updates on the status of inspections/appointments
- Saves profile and traveler data to reduce data entry
- Scan passports to directly upload document data
- Offers quick access to recently visited features
- Eliminates long lines and phones calls (where applicable)

History of CBP One

CBP leadership identified challenges from having multiple detached services and applications available to various user types. CBP One was developed with emerging technology to modernize and streamline access to CBP services and features. CBP One offers faster processing times for CBP Officers, due to advanced submissions and real-time feedback from users.

PROBLEM

- Stakeholders are increasingly demanding mobile capabilities for CBP services
- CBP applications and services were scattered and decentralized
- No consistent infrastructure or user interface for public-facing apps
- No mobile compatible format for most stakeholders to interact with CBP

SOLUTION

- Centralized platform for all CBP applications and services
- Modernized user experience
- Improved traveler compliance
- Mobile friendly solution for Trade/Commercial community
- Available at all Field Offices and Ports of Entry

INTENDED BENEFITS FOR CBP

- Establishes CBP One as the comprehensive public facing application for OFO services
- Reduces duplicative efforts and procurements for app development within OFO.
- Utilizes modern technologies to streamline and automate CBP operations
- Allows officers to shift focus from administrative tasks to enforcement responsibilities

Vision

Create a modern and simplified traveler and stakeholder experience, while making it easier for Officers to review and process users through its efficient data intake method.

Security

Secure sign-in through login.gov with access to previously saved information. Documents and biometric data are validated using cutting edge technology.

Intuitive Interface

Users are notified of the required information, forms, and visas they will need to provide through an easy and intuitive process based on who they are and what they need to do.

Pre-Vetting/Notifications In New Environments

CBP One enables Officers to prioritize targeting and vetting efforts in advance of traveler arrival at the border.

Integrated, Public-Facing Mobility

CBP One will consolidate existing, public-facing CBP mobile apps/services and make them accessible in one mobile application.

Capability Roadmap

Services in the CBP One app will be rolled out in phases. The first phase will include inspection requests for perishable cargo; the ability to apply, and pay for an I-94; biometric confirmation of third country nationals' exit from the U.S.; CBP ROAM's integration, with the ability to apply for updated cruising licenses. Additional services will be available in 2021

*Cargo Crew Processing is limited based on prior EAC approval
**Pending Policy Review and Approval

Available by November '20 | Available by February '21 | Available in 2021

CBP One Core Team and Stakeholders

The core CBP One team consists three product owners, each with a unique role and position. However, the broader CBP One team spans across the CBP enterprise and includes a large range of stakeholders

Program Manager

(b)(6), (b)(7)(C)

Product Owner(s)

(b)(6), (b)(7)(C)

OFO Stakeholders

APP
- Bus Processing
- Trusted Traveler
- Notice of Arrival
- Landing Rights Requests

APTL
- Scheduling Perishable Cargo
- Advance Notice Hand-Carry Permit Items

PPAE
- CBPBOAM – Report Arrivals for Pleasure boats
- Provisional I-94 Entry
- Third Country National Exit Reporting

DTOPS
- Fee payments

CCS
- Cruising License
- Vessel Arrivals
- Scheduling Cargo Inspections

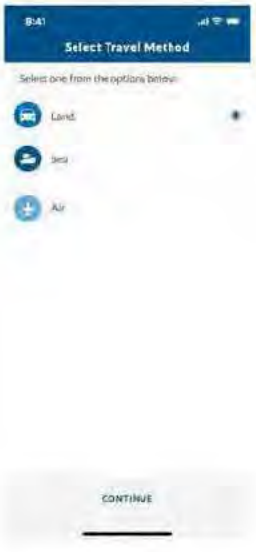
Technical Field SMEs

(b)(6), (b)(7)(C)

OIT Development Team

Close-up of CBP One

User Type: Traveler



User Type: Broker, Carrier, Forwarder

Appointment Dashboard (Officer Facing)

(b)(6), (b)(7)(C), (b)(7)(E)

DEPARTMENT OF HOMELAND SECURITY
U.S. Customs and Border Protection

OFO INTER-OFFICE ROUTING SHEET

Date: 8/14/2018

FROM

OFO Directorate/Division: PPAE

Writer's Full Name: (b)(6), (b)(7)(C)

Telephone No.: (b)(6), (b)(7)(C)

Room No.: (b)(6), (b)(7)(C)

OFO Tracking No.: (b)(7)(E)

Subject: CBP One Application (App) Vision and Next Steps

Priority

- ☐ S1/S2 — Date & Time Due
- ☐ C1/C2 _____ _____

Type

- ☒ Needs Approval
- ☐ For Situational Awareness

Comments

- ☒ Return to writer for edits
- ☐ Let's Discuss

TO	Destination	Initials	Date
	OFO Front Office		
☐	Executive Assistant Commissioner	(b)(6), (b)(7)(C)	8/27/18
☒	Deputy Executive Assistant Commissic		8/17/18
☒	Chief of Staff		8/16/2018
☒	Deputy Chief of Staff		
☐			
	Admissibility & Passenger Programs		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Agriculture Programs & Trade Liaison		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Cargo & Conveyance Security		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Mission Support		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	National Targeting Center		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Operations		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Planning Program Analysis & Evaluat	(b)(6), (b)(7)(C)	
☒	Executive Director		8/15/18
☒	Deputy Executive Director		8/15/18
☒	Director		8/14/18
☐			

This routing sheet is strictly for the use of in office documents only. Please do not use to distribute documents that need to be signed by other offices within CBP.

CBP HQ OFO Form 123 (12/16)

Office of Field Operations
Planning, Program Analysis and Evaluation Directorate (PPAE)
August 14, 2018

Action: Information Only

Issue: CBP One Application (App) Vision and Next Steps

Executive Summary:

- PPAE Innovative Program Acquisitions (IPA) is formally launching the development of the CBP One app to provide travelers a consistent experience when interacting with CBP mobile products.
- CBP One will be an intuitive application suite of CBP traveler facing mobile apps, for existing, and upcoming efforts.
- CBP One will guide travelers to the appropriate application based upon a series of questions the traveler will answer about their trip and travel needs.

Background:

- In recent years, CBP has launched multiple mobile apps to streamline the ways in which travelers interact with CBP, and additional apps are planned for the near future.
- While digital engagement has increased, the growing number of apps risks confusing travelers and impeding the overall user experience.
- The idea for CBP One was developed as a singular, intuitive mobile application suite of applications for travelers to interact with CBP on their smart devices and avoid having to self-navigate between different CBP mobile app options (e.g., CBP ROAM, I-94 application, I-94 exit, etc.).
- (b)(5)

Current Status:

- IPA has identified relevant stakeholders for the development of the CBP One app, including product owners, sponsors, and technical Field subject matter experts.
- These stakeholders met in-person in Washington DC during the week of July 30th (b)(5) (b)(5)
- The ultimate vision for CBP One is that it will serve as a platform to access present and future CBP mobile apps.
- Moving forward, PPAE will lead the coordination across OFO and OIT to complete the design and development of CBP One.
- OFO has funded CBP One and will work with OIT to develop a schedule for application release.

Approved/Date: _____ **Disapproved/Date:** _____

Needs Discussion/Date: (b)(6), (b)(7)(C) 8/17 **Modify/Date:** _____

CBP One Mobile Application

U.S. Customs and Border Protection

CBP One mobile application will be designed with an user centric interface to provide users with an intuitive and guided border entry/exit experience regardless of geographic location, mode of transportation or citizenship. CBP One is intended to compliment other traveler facing apps (e.g., ROAM, I-94 Entry, I-94 Exit), and be conversational in nature to direct users to the appropriate functionality given their travel scenario.

CBP One prepares all travelers for U.S. entry/exit by allowing users to:

- Integrate with login.gov and pay.gov
- Create and save traveler profiles, including those of their travel group
- Scan passports to directly import information into CBP One
- Declare port of entry and geographic location (when applicable)
- Select mode of entry
- Access the appropriate CBP applications

Notional CBP One Mockup

(b)(5)

Potential Traveler Scenarios:

- Allow users to remotely report U.S. Entry by small vessel via the ROAM App
- Enable user to enter, submit and quickly reference I-94 applications for US Entry
- Biometrically validate US Exit and close out I-94 forms
- Use Ready Lanes when traveling by vehicle via the I'm Ready App
- Leverage CBP One as a bus carrier to enter and pre-submit all traveler data before US Entry

*Note: Prior to the first screen shown CBP One will show users a welcome screen, the onboarding steps, and prompt users to agree to Terms & Conditions

U.S. Customs and Border Protection

CBP One™ : Allows **International Organizations** to submit advance information for individuals

01

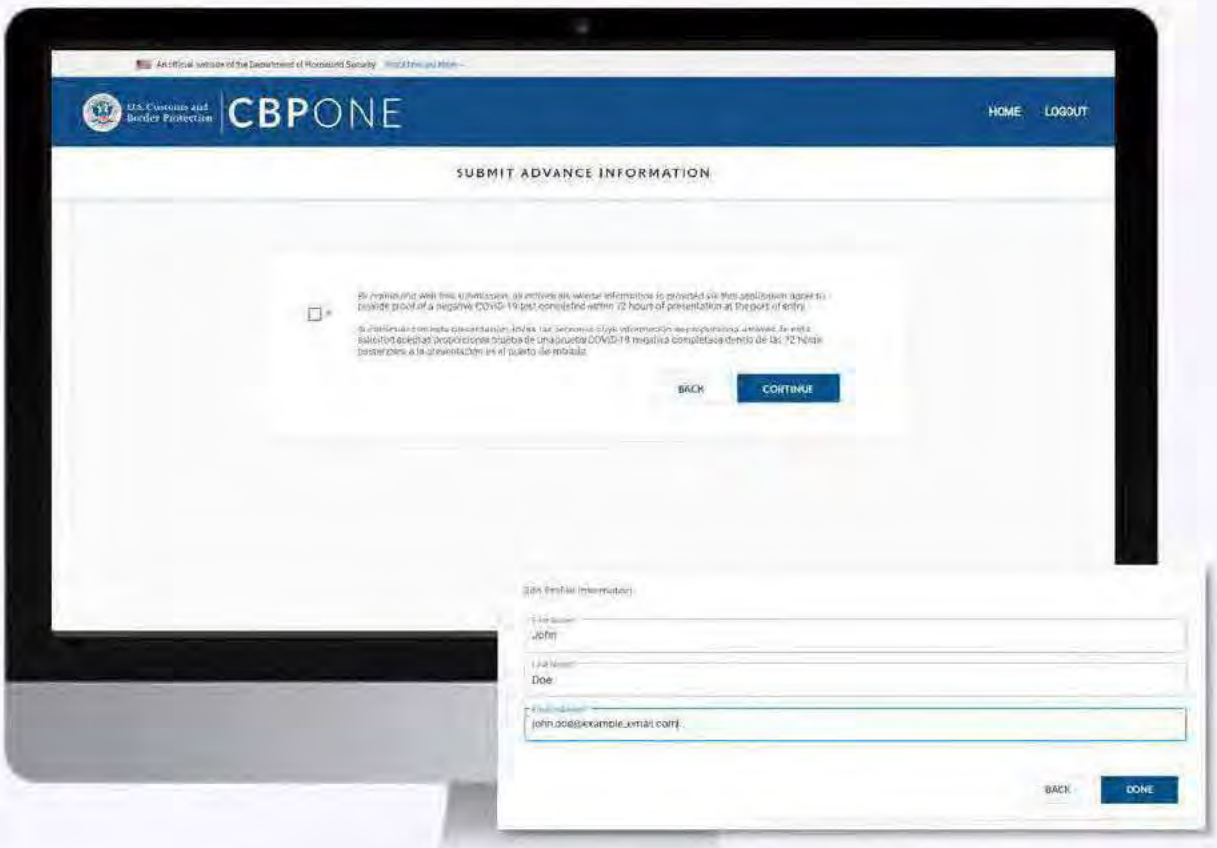
To get started, go to https://cbpone.cbp.dhs.gov/

Tap on "International Organization" to begin and select "Submit Advance Information".

02

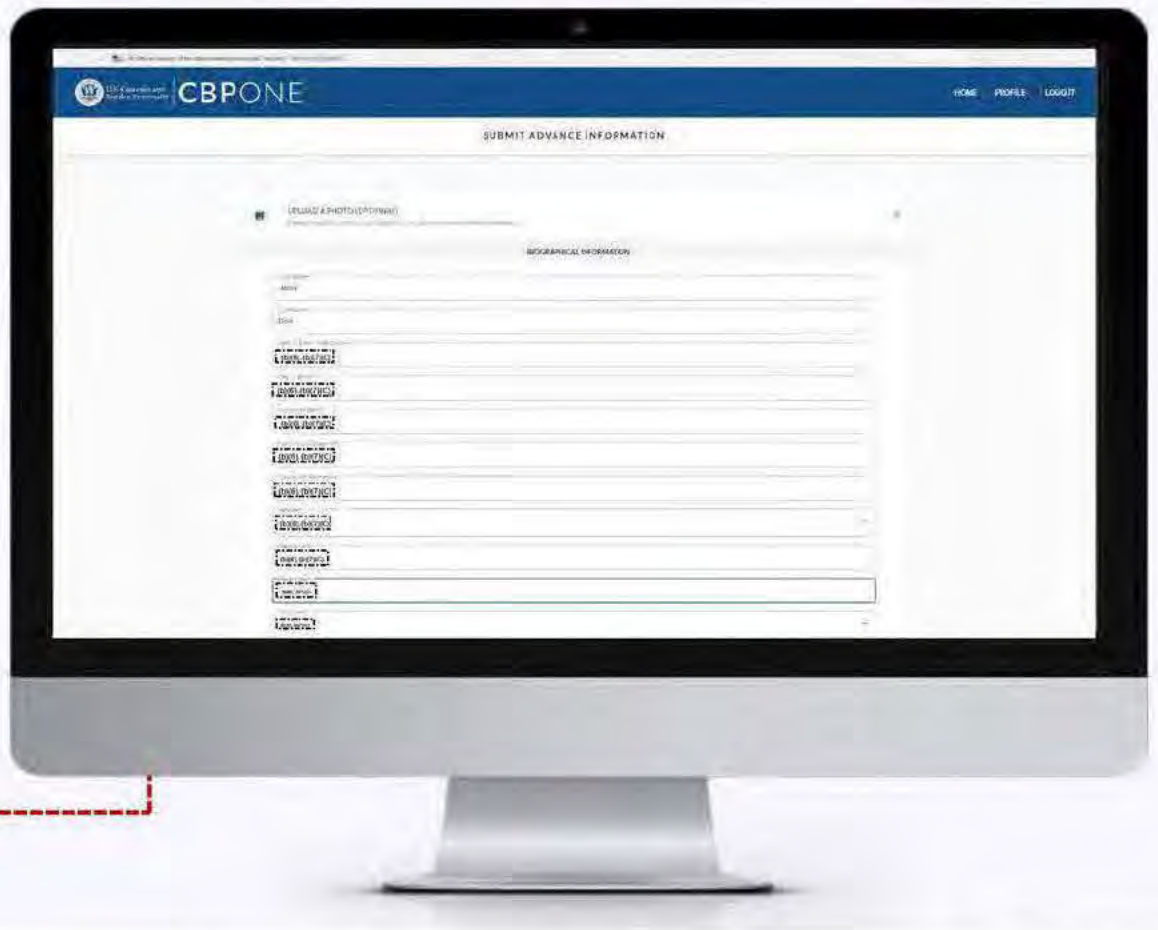
Acknowledge the COVID-19 checkbox to proceed. First time users will be prompted to create a profile.

Note: All individuals whose information is provided via this application are required to provide proof of a negative COVID-19 test completed within 72 hours of presentation at the port of entry.



03

Select "Add Individual" and fill out the required fields for all individuals traveling together on this trip.

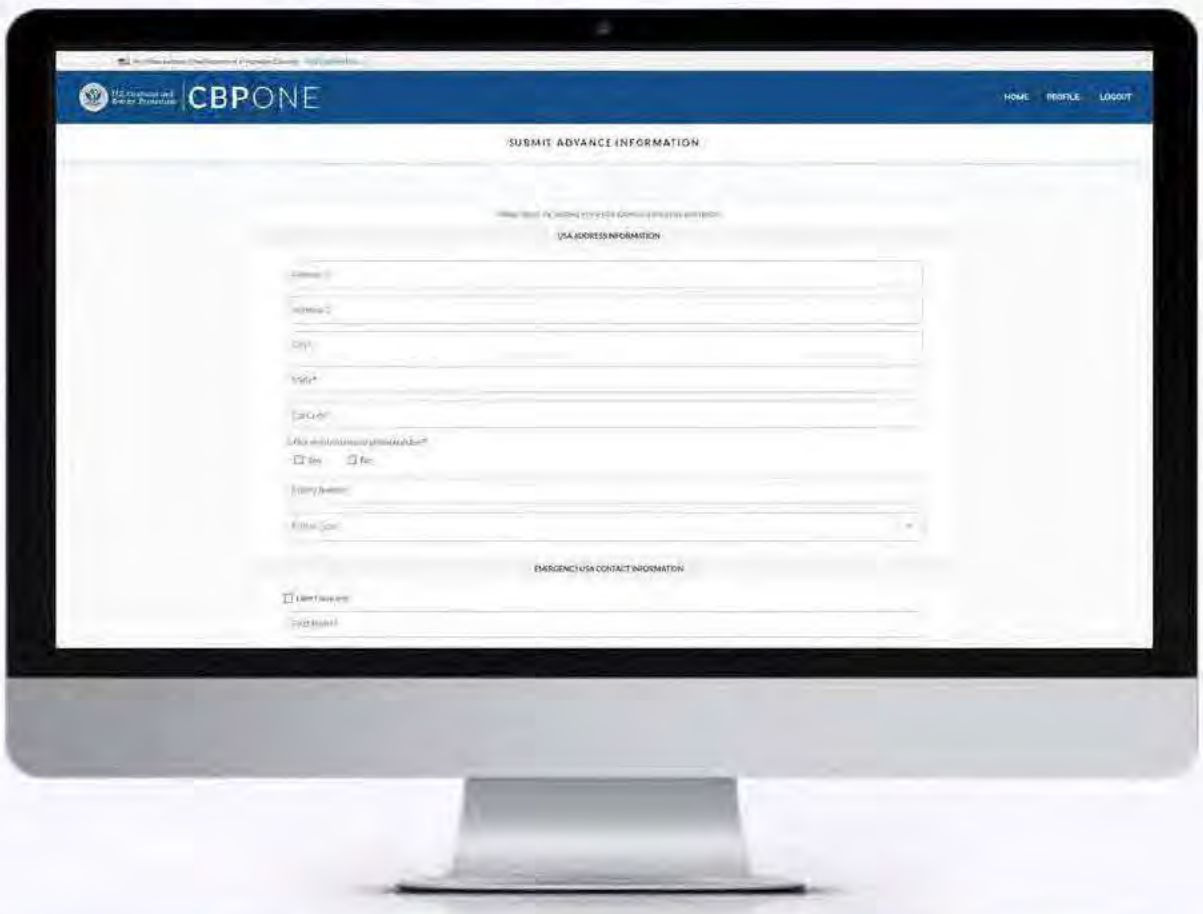


U.S. Customs and Border Protection

CBP One™ : Allows **International Organizations** to submit advance information for individuals

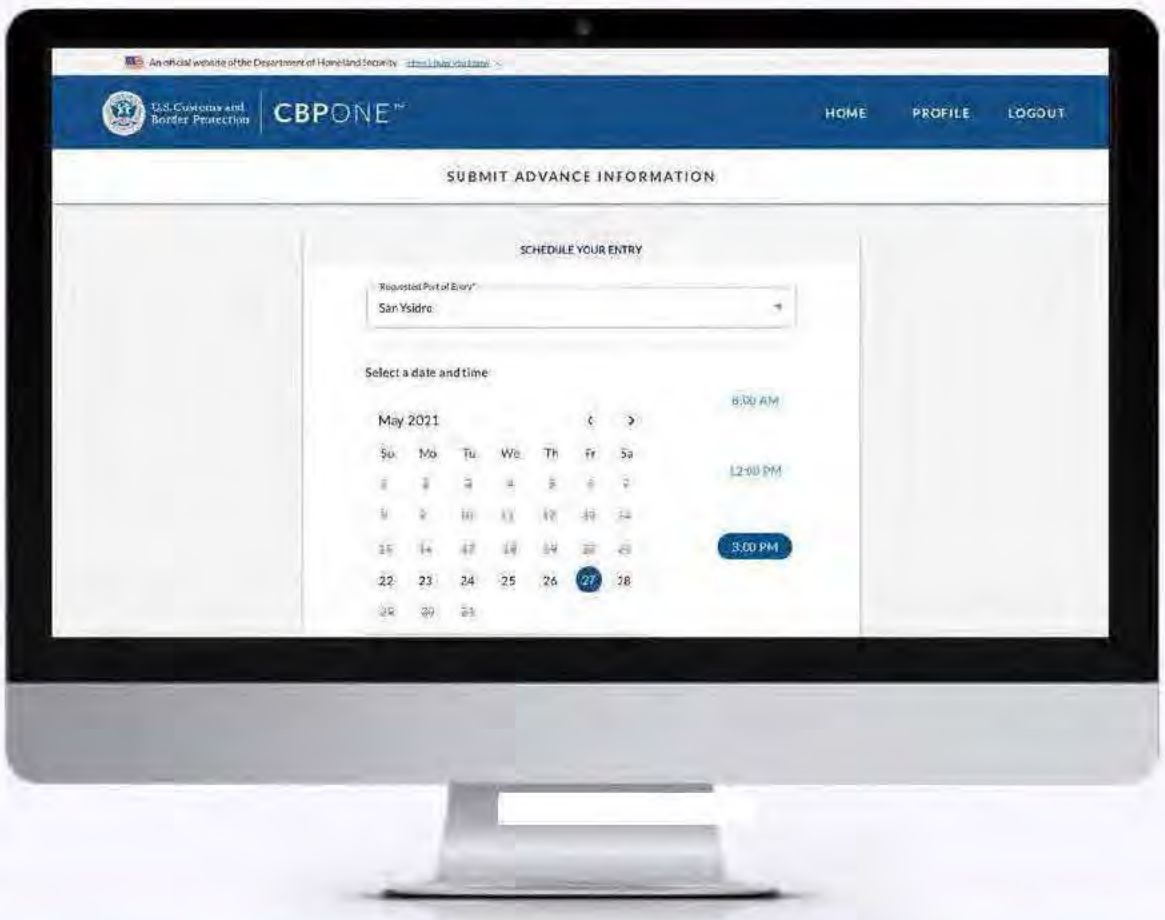
04

Fill out the address in the USA where the individual(s) will arrive and reside, along with an emergency USA point of contact. Then, fill out the address for where they lived before coming to the USA and Title 42 remarks.



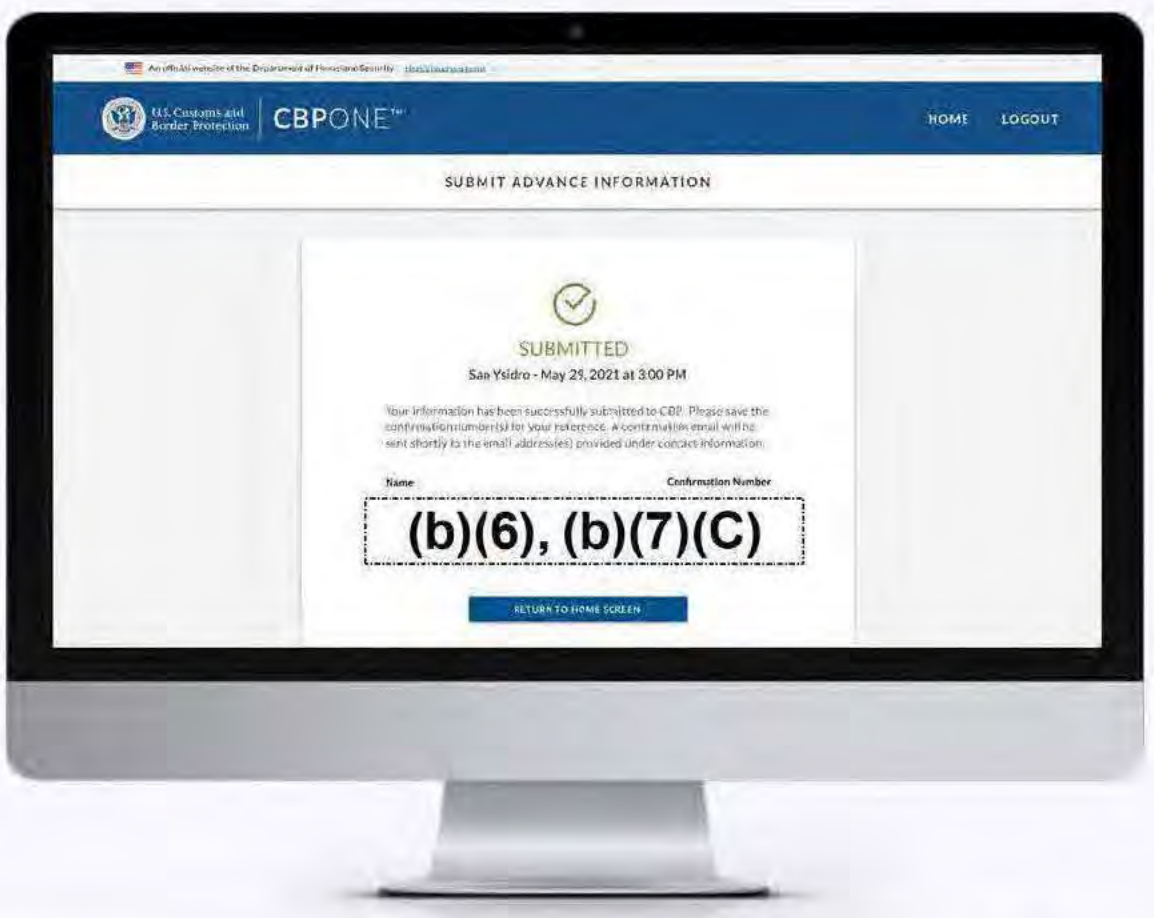
05

Lastly, select the requested port of entry and date of entry and choose an available time slot.



06

Review all information and tap on submit. A confirmation screen will display the confirmation number(s) for each individual. A confirmation email will be sent to the email address(es) provided under contact information.



From: (b)(6), (b)(7)(C)
To:
Cc:
Subject: RE: NIST and S&T Reports
Date: Wednesday, May 4, 2022 11:13:58 AM

Hi (b)(6), (b)(7)(C)

I am not aware of any other reports from NIST or DHS S&T that are specific to TVS.

Thank You,

(b)(6), (b)(7)(C)
US Customs and Border Protection
Office of Information and Technology
Targeting and Analysis Systems Program Directorate (TASPD)
(C) (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, May 4, 2022 11:09 AM
To: (b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
Subject: NIST and S&T Reports

Hi (b)(6), (b)(7)(C)—

Apologies if I have already asked the following question in the past, I just want to trust but verify. We received the attached FOIA in September 2021. The American Immigration Council requests, among other items, final reports "received by CBP from the Department of Homeland Security's ("DHS") Science and Technology Directorate, any other component of DHS, or the National Institute of Standards and Technology, relating to the efficacy of facial-recognition technology."

I know we've received the 2 attached reports from S&T, as well as the public NIST reports, but just wanted to double check if OIT has received any other reports from either NIST or S&T that would be responsive to the FOIA request. Please advise.

I am available to discuss.

Thanks,

(b)(6), (b)(7)(C)
Management & Program Analyst
Biometric Entry-Exit Strategic Transformation (BEST)

Admissibility and Passenger Programs
Office of Field Operations
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

❖ **Deployments This Month:**

6/18 – NGOs can Schedule Undocumented Noncitizens (UN)

6/30 – Ability for UN to Submit Directly

U.S. Customs and Border Protection

June CBP One Status Report

What is Available Now

Oct 2020

- ✓ Air Perishable Cargo Scheduling
  - Miami, LAX, Dallas, Houston Boston
- ✓ Apply for I-94
- ✓ MPP for Intl Organizations and TSA

February 2020

- ✓ IOs Verify Eligibility for MPP Winddown

March 2020

- ✓ TSA Verify Identity at Checkpoints
- ✓ NGOs Submit Advance Info on Non-MPP

What is Coming Next

July 2021

- Transition ROAM, Bus APIS, Cargo Crew (Detroit)
- Apply for and Update Cruising Licenses.
- Air Traveler Notice Biological/Ag Inspection

September 2021

- I-94 Exit

August or Later

- Pilot Landing Rights and Diversion Notices
- Schedule Commercial Vessel Arrivals
- Schedule Non-Perishable Cargo Sea/Land

Current Statistics

- Downloads
  - 6,500
- Scheduling Requests
  - 25,000 inspection/500,000 items
- MPP IOs/TSA Queries
  - 25,00 queries
- TSA Queries
  - 18,000
- Non-MPP
  - 1,300
- Average App Store Rating
  - Android: 5 Stars
  - iOS: 4.4 Stars

Challenges or Roadblocks

- (b)(5)

U.S. Customs and Border Protection (CBP)
Office of Field Operations (OFO)
Planning, Program Analysis and Evaluation Directorate (PPAE)
October 1, 2020

Action: Approval

Issue: Official Launch of CBP One

Executive Summary:

- In August 2018, the Executive Assistant Commissioner (EAC), OFO approved the development of CBP One™ mobile application that is intended to act as an intuitive single portal for travelers and stakeholders to access CBP mobile applications and services such as CBP ROAM, I-94 Entry/Exit and the Appointment Request feature (Attachment A).
- On October 28, 2020, CBP One will be available on the IOS and Android Play Stores under "CBP One™".
- The rollout of CBP One will be a phased approach with the initial capabilities limited to the I-94 Apply and a Scheduling/Appointment feature for brokers to schedule perishable exams.
- Subsequent rollouts will include incorporate I-94 Exit, CBP ROAM along with the ability to apply for and update cruising licenses, the ability to apply for and check Global Entry status, and the upcoming CBP version of Mobile Passport Control.

Background:

- In 2018, the EAC approved the development of CBP One as a single portal for CBP stakeholder applications.
- The CBP One app was developed to consolidate the public facing apps to reduce confusion and the need for individual stakeholders to use multiple CBP apps to provide information and/or request services.
- CBP One guides stakeholders through a series of intuitive questions to direct each type of user to the appropriate services based on their needs.
- CBP individual applications are now "capabilities" accessible through one app that provides seamless navigation by the public (Attachment A).

Current Status:

- The OFO Innovation Center developed the CBP One application in collaboration with the Office of Information Technology and has been piloting the first iteration of the scheduling feature in the Miami Field Office since July 2020.
- The I-94 Apply is a mobile version of the current I-94 Apply website offering users the ability to apply, pay for, and view their I-94 form and I-94 history on their phone.
- PPAE is working with Admissiblity and Passenger Programs (APP) on options for current Visa Waiver applicants to answer questions through the app as the current process requires them to pay Electronic System Travel Authorization fees to use the advance pay option.

- The OFO Innovation Center has coordinated with the following OFO program office and directorates: APP, Land Border Integration and Biometrics (LBIB), Cargo Conveyance Security and Agriculture Programs and Trade Liaison to incorporate individual app capabilities (Attachment B – Capability Roadmap).
- CBP One allows individual programs within OFO to integrate their capabilities into CBP One while maintaining full discretion and responsibility for privacy, internal messaging, external marketing and operations and maintenance of their capabilities. PPAE will not assume nor interfere with these responsibilities.
- The Executive Directors will be advised of the CBP One app by email with a request to coordinate any future public-facing apps with the Innovation Center. This coordination will leverage the desired app's existing capabilities and ensure it is conducive to integrating with CBP One from the onset.

Recommendation:

- Distribute the draft memo from the Executive Assistant Commissioner announcing the launch of CBP One.

Approved/Date: (b)(6), (b)(7)(C) 10/20/20 **Disapproved/Date:** _____

Needs Discussion/Date: _____ **Modify/Date:** _____

Submitted by: (b)(6), (b)(7)(C) Director Strategic Transformation
Office: Planning, Program Analysis and Evaluation
Date: September 28, 2020

CBP One™ | August 2021 Status Report

U.S. Customs and Border Protection

Deployments for Next Month:

- Stakeholder Testing for APTL Advance Air Traveler Info
- Stakeholder Testing for CBP ROAM
- Perishable Cargo Launches for San Juan and New York

Available Now

- ✓ Scheduling for inspections of perishable cargo (air)
  - ✓ Miami, LAX, Dallas, Houston, Boston
- ✓ Apply for I-94
- ✓ MPP for International Organizations and TSA

Upcoming Releases

September 2021

- CBP ROAM Cruising License Feature
- Bus APIS
- Air Traveler Notice Biological/Ag Inspection
- Apply for and update Cruising Licenses

October 2021

- I-94 Exit

November 2021 and onwards

- Pilot Landing Rights and Diversion Notices
- Schedule Commercial Vessel Arrivals
- Schedule Non-Perishable Cargo (Sea and Land)

Challenges & Roadblocks

- (b)(5)

Current Statistics

Downloads
Android: 18,398
iOS: 29,375

MPP IOs & TSA Queries
85,000 queries have been performed through CBP One™

Scheduling Request
37,600 inspections of perishable cargo have been processed through CBP One™

Provisional I-94
299 (will increase with resumption of travel)

- **Deployments This Month:**
- **Stakeholder Testing for APTL Advance Air Traveler Info**
- **Stakeholder Testing APP Bus APIS**
- **Testing/Training for Cruising licenses in *ROAM***

U.S. Customs and Border Protection

August CBP One Status Report

What is Available Now

Oct 2020

- ✓ Air Perishable Cargo Scheduling
  - Miami, LAX, Dallas, Houston Boston
- ✓ Apply for I-94
- ✓ MPP for Intl Organizations and TSA

February 2020

- ✓ IOs Verify Eligibility for MPP Winddown

March 2020

- ✓ TSA Verify Identity at Checkpoints
- ✓ NGOs Submit Advance Info on Non-MPP

What is Coming Next

August 2021

- **Direct Transmission of Non-MPP Advance Info – On Hold**

September 2021

- Bus APIS
- Air Traveler Notice Biological/Ag Inspection
- Apply for and Update Cruising Licenses.

October 2021

- I-94 Exit

November or Later

- Pilot Landing Rights and Diversion Notices
- Schedule Commercial Vessel Arrivals
- Schedule Non-Perishable Cargo Sea/Land

Current Statistics

- Downloads
  - 6,500
- Scheduling Requests
  - 25,000 inspection/500,000 items
- MPP IOs/TSA Queries
  - 25,00 queries
- TSA Queries
  - 18,000
- Non-MPP
  - 1,300
- Average App Store Rating
  - Android: 5 Stars
  - iOS: 4.4 Stars

Challenges or Roadblocks

- (b)(5)

U.S. Customs and Border Protection

CBP One™

Allows **International Organizations** to check the case status of individuals

Download CBP One™

Download on the App Store

GET IT ON Google Play

To get started, download CBP One™ from the Apple App Store or Google Play Store.

Sign In Using Login.gov

The app will redirect to **login.gov** where you can either create or login to your existing account.

Questions?
Contact us at: CBPOne@cbp.dhs.gov

1. Who Are You

Tap on "International Organization" to begin. First time users will be prompted to create a profile.

2. Check Case Status & Photo Options

Select "Check Case Status", then select one of the following options: take photo, upload photo or decline photo. "Decline Photo" will take the user to step 4.

(b)(6), (b)(7)(C)

3. Photo Results

Green indicates the individual is approved to go to primary, red indicates the individual is not found, yellow indicates a record was found but it is no longer active, and blue (see step 6) indicates the record has been closed and is not eligible for processing.

4. Search by A-Number

The individual will be prompted for their A-Number if there is no match on the photo. The search will return one of the results outlined in step 3 or 6.

5. Search by Biographical Data

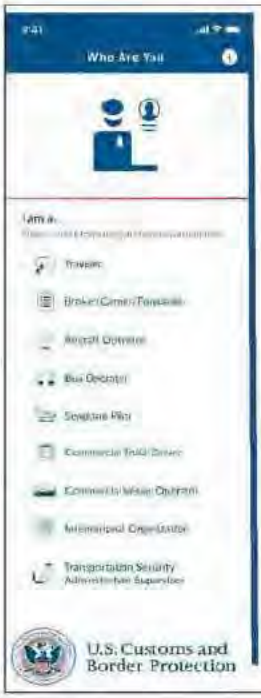
The individual can do a third query by first/last name, DOB, and country of citizenship if and only if there is no match on the photo and A-number.

6. A-Number and Biographical Data Results

The search will return one of the results outlined below or in step 3. Tap on "Check Next Case Status" to go to the next individual, or "Done" to return to the IO home screen.

GAO-20-568 updates on public facing website:

Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues | U.S. GAO



March CBP One Status Report

What is Available Now

- ✓ Air Perishable Cargo Scheduling
  - Miami, LAX, Dallas, Houston Boston
- ✓ Apply for I-94
- ✓ MPP for Intl Organizations and TSA

What is Coming Next

May 2021

- Report Arrivals for Pleasure Boaters, Bus Travelers, Cargo Crew (Detroit)
- Apply for and Update Cruising Licenses.

July 2021

- Pilot Landing Rights and Diversion Notices
- Advance Notice of Permit Items – Air

August or Later

- Schedule Commercial Vessel Arrivals
- Schedule Non-Perishable Cargo Sea/Land

Current Statistics

- Downloads
  - Android: 1,070
  - iOS: 2,560
- Scheduling Requests
  - 15,500 inspection
- MPP IOs/TSA Queries
  - 20,000 queries
- Average App Store Rating
  - Android: 5 Stars
  - iOS: 4.4 Stars

Challenges or Roadblocks

- (b)(5)

- ❖ **Deployments for Next Month:**
- ❖ **Stakeholder Testing for APTL Advance Air Traveler Info**
- ❖ **Stakeholder Testing APP Bus APIS**
- ❖ **Testing/Training for Cruising licenses in *ROAM***

U.S. Customs and Border Protection

July CBP One Status Report

What is Available Now

Oct 2020

- ✓ Air Perishable Cargo Scheduling
  - Miami, LAX, Dallas, Houston Boston
- ✓ Apply for I-94
- ✓ MPP for Intl Organizations and TSA

February 2020

- ✓ IOs Verify Eligibility for MPP Winddown

March 2020

- ✓ TSA Verify Identity at Checkpoints
- ✓ NGOs Submit Advance Info on Non-MPP

What is Coming Next

August 2021

- **Direct Transmission of Non-MPP Advance Info – On Hold**

September 2021

- Bus APIS
- Air Traveler Notice Biological/Ag Inspection
- Apply for and Update Cruising Licenses.

October 2021

- I-94 Exit

November or Later

- Pilot Landing Rights and Diversion Notices
- Schedule Commercial Vessel Arrivals
- Schedule Non-Perishable Cargo Sea/Land

Current Statistics

- Downloads
  - 6,500
- Scheduling Requests
  - 25,000 inspection/500,000 items
- MPP IOs/TSA Queries
  - 25,00 queries
- TSA Queries
  - 18,000
- Non-MPP
  - 1,300
- Average App Store Rating
  - Android: 5 Stars
  - IOS: 4.4 Stars

Challenges or Roadblocks

- (b)(5)

CBP ONE – CAPABILITY ROADMAP

DRAFT

U.S. Customs and Border Protection

(b)(5), (b)(7)(E)

APPENDIX

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

CBP One Advance Information Collection for Undocumented Population
Generally Resolved Discussion Points
July 15, 2021

(b)(5)

(b)(5)

(b)(5)

(b)(5)

(b)(5)

U.S. Customs and Border Protection

April 14, 2021

MEMORANDUM FOR: Directors, Field Operations
Executive Directors
Office of Field Operations

FROM: (b)(6), (b)(7)(C) (b)(6), (b)(7)(C)
Acting Executive Director
Planning, Program Analysis and Evaluation
Office of Field Operations

SUBJECT: New CBP One™ Pleasure Boat Capabilities/Training

On October 20th, 2020, the Office of Field Operations (OFO) formally launched the CBP One™ mobile app for travelers and stakeholders to access CBP mobile apps and services. Since its launch, CBP One™ has released capabilities to allow brokers to schedule perishable cargo exams and land border travelers to apply and pay for provisional I-94s.

In May 2021, the existing CBP ROAM app will transition all its services into CBP One™ and the CBP ROAM app will no longer be available for use. This transition will also introduce new features which will allow pleasure boaters to apply for cruising licenses, report their cruising port of call, and confirm their identity via facial biometric verification technology. These new capabilities are expected to reduce the number of boaters that must travel to CBP offices for services.

The transition of CBP ROAM capabilities into CBP One™ will have minimal impact on both the application users and CBP Officers. Existing CBP ROAM users will only be required to perform a one-time re-entry of traveler profile information and all ROAM services will continue to function as they did in CBP ROAM. However, with the application's significant enhancements, pleasure boaters will now be able to opt into facial biometric verification, apply for cruising licenses and report arrivals at cruising ports of call. There will be several outreach and marketing efforts at a national level to inform the traveling public of these changes and materials will also be provided to facilitate local messaging.

(b)(5)
(b)(5) Attached is the current list of CBP ROAM points of contact (POCs). Please confirm the POC for your area of responsibility or provide a new POC to (b)(6), (b)(7)(C) at (b)(6), (b)(7)(C) by April 20, 2021.

If you have any questions, please feel free to contact me or have a member of your staff contact (b)(6), (b)(7)(C)

Attachment

COMPUTATIONAL BIOMEDICINE LABORATORY
Department of Computer Science
University of Houston

BOA number:
Task Order number: (b)(7)(E)
Project Description: CBP I94/Mobile Exit App Technical Vulnerability Evaluation and Hackathon
PI: (b)(6), (b)(7)(C)

Report on Hackathon Results
Report date: December 11, 2020

Contractor's name and address:
University of Houston System
(b)(6), (b)(7)(C)

Name of person submitting report:
(b)(6), (b)(7)(C)

(b)(7)(E)

Hackathon Results

Objectives

The specific objectives were the following:

1. Evaluate the liveness measures of human photos taken from a mobile device app (smartphone);
2. Establish a hack challenge for geolocation coordinates presented from that mobile device app (spoofing).

Technical review results

(b)(7)(E)

(b)(7)(E)

Appendix

Security Assessment

(b)(7)(E)

(b)(7)(E)

Appendix B

(b)(7)(E) The following individuals were curated to participate in this Challenge from (b)(7)(E) community:

(b)(6), (b)(7)(C), (b)(7)(E)

Appendix C

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

(b)(7)(E)

DEPARTMENT OF HOMELAND SECURITY
U.S. Customs and Border Protection

OFO INTER-OFFICE ROUTING SHEET

Date: 07/11/2019

FROM

OFO Directorate/Division: PPAE / STO

Writer's Full Name: (b)(6), (b)(7)(C)

Telephone No.: (b)(6), (b)(7)(C)

Room No.: (b)(6), (b)(7)(C)

OFO Tracking No.: (b)(7)(E)

Subject: Stakeholder Scheduling and Cruising License Application

Priority

- ☐ S1/S2 Date & Time Due
- ☐ C1/C2 _____ _____

Type

- ☒ Needs Approval
- ☐ For Situational Awareness

Comments

- ☐ Return to writer for edits
- ☐ Let's Discuss

App Concurrence provided
19 CFR 4.94 Requirements
are included

TO	Destination	Initials	Date
	OFO Front Office		
☐	Assistant Commissioner		
☒	Deputy Assistant Commissioner	(b)(6), (b)(7)(C)	7/31
☒	Chief of Staff		
☒	Deputy Chief of Staff		7/22/19
☐	OBI Director		
☐	CMO Director		
☐	Senior Policy Advisor		
☐			
	Admissibility & Passenger Programs		
☒	Executive Director	(b)(6), (b)(7)(C)	7/25/19
☐	Deputy Executive Director (b)(6), (b)(7)(C)		
☐	Director		
☐			
	Agriculture Programs & Trade Liaison		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Cargo & Conveyance Security		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Mission Support		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	National Targeting Center		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Operations		
☐	Executive Director		
☐	Deputy Executive Director		
☐	Director		
☐			
	Planning Program Analysis & Evaluat		
☒	Executive Director	(b)(6), (b)(7)(C)	7/29/19
☒	Deputy Executive Director		7/29/19
☒	Director		7/11/2019
☒	Director Davis		7/18/2019

This routing sheet is strictly for the use of in office documents only. Please do not use to distribute documents that need to be signed by other offices within CBP.

CBP HQ OFO Form 123 (04/14)

Office of Field Operations
Planning, Program Analysis and Evaluation Directorate
July 9, 2019

Action: Approval

Issue: Development of Stakeholder Scheduling and Cruising License Applications in Support of Miami Field Office Innovation Efforts

Executive Summary:

The Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations (OFO) Innovation Center (IC) will support the Miami Field Office (MFO) and the Emerging Technologies Team at the Office of Information Technology (OIT) to develop, pilot and implement two stakeholder applications. One application will facilitate stakeholders scheduling CBP services such as perishable exams and vessel or private aircraft arrivals while the second will facilitate the application and issuance of cruising licenses for foreign pleasure boats. The applications will be developed utilizing current platforms with consistent user interfaces as other stakeholder applications.

Background:

App for Stakeholders to Schedule CBP Services

- In August 2018, the MFO held and the PPAE OFO IC participated in a Shark Tank Event for local employees to pitch innovative ideas. One of the ideas was the development of an online scheduling tool for perishable exams. Currently, (b)(7)(E) Supervisory Agriculture Specialists at the Miami International Airport currently manage the scheduling of perishable exams 24 hours a day 7days a week through an exchange of phone calls with assignments on a first-come first-serve basis.
- The idea was selected for development and a Project Zone request was approved by the Executive Director of PPAE in September 2018.
- The PPAE OFO IC, Agriculture Programs and Trade Liaison (APTL) division, MFO and OIT had several conference calls to discuss the current process for scheduling perishable cargo exams and create story boards for an online application.
- In March 2019, the PPAE OFO IC held a workshop in Long Beach, California, to develop requirements for the online scheduling application for perishable cargo along with the expansion to include additional opportunities to schedule appointments (e.g. NII exams at large seaports).
- The application will significantly reduce the workload and data input by Supervisory Agriculture Specialists, increase the effectiveness of assigning and reassigning CBP Agriculture Specialists and enhance the stakeholder experience by scheduling, updating and receiving messaging on appointments through on online application.

App for the Application and Issuance of Cruising Licenses for Pleasure Boats

- A cruising license is authorized by 19 CFR 4.94, *a license to Cruise in the Waters of the United States.* A cruising license allows a foreign vessel not imported for resale to travel

from port to port within the U.S. or to a foreign port make reentry to the U.S. without filing a formal entrance and clearance and paying associated fees.

- The current process includes:
  - Pleasure boaters traveling to a CBP Office, at times up to two hours, to request a cruising license;
  - CBP determining whether to approve the license;
  - Notating the approval in the Pleasure Boat Reporting System; and
  - Issuing a paper letter to the traveler as evidence of the cruising license.
- The pleasure boater is then required to call each subsequent Port of Entry (POE) to report their arrival and advise they have a cruising license.
- The PPAE OFO IC and the MFO have discussed the app with the Admissibility and Passenger Programs (APP) Division and they concur with the project.
- The application will enable a consistent nationwide approach to issuing cruising license, facilitate the query of cruising licenses by subsequent POEs, streamline the reporting of domestic arrivals by foreign pleasure boats by leveraging current reporting platforms and automate the enforcement of cruising license periods of validity.

Current Status:

- The PPAE OFO IC will work with all applicable directorates/programs and the field on development of requirements. (b)(5)
  (b)(5)
- A Privacy Threshold Assessment will be submitted to the Privacy Office. (b)(5)
  (b)(5)
- (b)(5)
- The development of the application does not require additional funding.
- The initial plans for piloting will be in the Miami and Los Angeles Field Offices.
- The application will be on an existing CBP consolidated application platform.

Recommendation:

1. Approve the development, pilot and implementation of a stakeholder application to schedule CBP services on an existing CBP consolidated platform for stakeholder apps.

(b)(6), (b)(7)(C) Approve _____ Disapprove _____ Let's Discuss

2. Approve the development of an application for the application for and issuance of cruising licenses on an existing CBP consolidated platform for stakeholder apps.

_____ Approve _____ Disapprove _____ Let's Discuss

Submitted by: (b)(6), (b)(7)(C)
Director, Strategic Transformation Office
Planning, Program Analysis and Evaluation

U.S. Customs and Border Protection

CBP One™

Exiting the United States by land? Report your departure on your mobile device

Download CBP One™

Download on the App Store

GET IT ON Google Play

To get started, download CBP One™ from the Apple App Store or Google Play Store.

Sign In Using Login.gov

The app will redirect to **login.gov** where you can either create or login to your existing account.

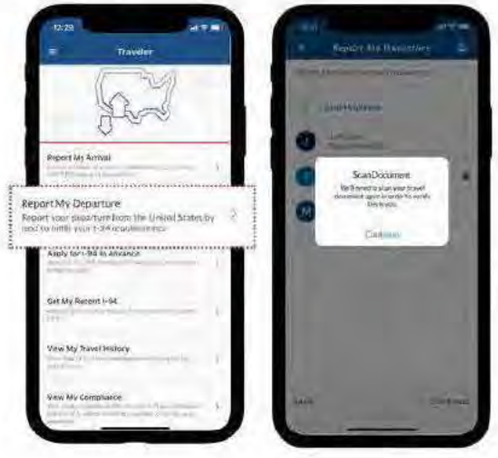
Questions?
Contact us at: CBPOne@cbp.dhs.gov

1. Who Are You

Tap 'Traveler' then "Land" to begin. You must be outside the U.S. to report your departure.

2. Report My Departure

Select "Report My Departure". Scan your passport or BCC for multi-factor authentication.



3. Departure Details

Please report where you left the U.S. and review your information.

4. Liveness Detection

In order to complete the exit, CBP One will need to take a quick video selfie.

5. Successfully Submitted

Once submitted, you should be able to view your exit information in your travel history (please note this can take up to 48 hours).

6. Liveness Detection Failed

If your liveness detection fails the first two times, you may try again. If it fails after a third try, you can try again or cancel.

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C)

U.S. Customs and Border Protection (CBP)
Office of Field Operations (OFO)
Planning, Program Analysis and Evaluation Directorate (PPAE)
May 21, 2021

Action: Approval

Issue: Streamlining I-94 Issuance at Land Borders

Executive Summary:

- U.S Customs and Border Protection (CBP), Office of Field Operation's (OFO) Planning, Program Analysis and Evaluation (PPAE) office, in coordination with Admissibility and Passenger Programs (APP), is streamlining the I-94 issuance process at land borders.
- Effective immediately CBP Officers will no longer be required to print paper I-94s.
- In addition, a robust communications campaign is being developed to promote the traveling public to apply and pay for I-94s ahead of arrival at the land border.

Background:

- On March 27, 2013, DHS published an Interim Final Rule (IFR) in the Federal Register (FR) and later a Final Rule that allowed CBP to automate Form I-94 at air and seaports of entry. The IFR expanded the definition of an I-94 to include electronic formats.
- In April 2013, OFO implemented automation of Form I-94 in the air and sea environment and has successfully issued electronic forms since that date.
- To provide the traveling public access to their automated I-94 records, an I-94 retrieval website was launched at the same time.
- Since the launch of the original I-94 website, additional functionality and capabilities have been added to include, view travel history, apply and pay for a land border I-94, submit a large group application for land borders and view travel compliance and overstay information.
- While many efforts to streamline and enhance the I-94 issuance process have been implemented, the current land border I-94 process, to include the I-94W, is still somewhat labor intensive for the CBP Officer. The process involves data entry, biometric collection, fee collection, a CBP interview and printing/stamping of the I-94.
- Analysis shows that the largest impact to operations is realized when the traveler applies and pays for an I-94 in advance of arrival at the land border (approximately 3 minutes per person time savings).
- To further encourage this process, CBP added the I-94 website functionality to CBP One™ to provide a mobile mechanism to apply and pay for I-94s and to access I-94 information even while at the Port of Entry (POE).

Current Status:

- Effectively immediately upon issuance of the attached memorandum, CBP will no longer require the printing of I-94s at the land border.
- OFO has deployed a paperless process to meet the demands of the Migrant Protection Protocols at the land border. Returning asylees are provided with tear sheets to advise how to retrieve the electronic I-94 from the CBP I-94 website or CBP One™ mobile app.

- The Electronic System for Travel Authorization (ESTA) Program Management Office has a pending regulation change that will require ESTA for Visa Waiver Program travelers on the land border, further streamlining the admission process.
- PPAE continues to deploy SA-PED and SA I-94 at ports of entry on both the Northern and Southern land border.
- In locations where these systems are deployed, ports will have the option to admit provisional I-94 applicants on primary.
- CBP will still need to refer travelers in vehicle lanes to capture biometrics; however, they will still benefit from the elimination of printing and no cash collections.

Next Steps:

- PPAE will issue the memo and muster to the field.
- The attached memorandum will be issued to the Chief, U.S. Border Patrol.
- PPAE will work with APP to provide outreach to relevant stakeholder groups and affected agencies on the change to electronic I-94 issuance.
- PPAE will coordinate with the Communications Management Office (CMO) to deploy a robust communication plan to promote CBP One™ and the CBP I-94 Website to encourage travelers to apply in advance of arrival.
    - The communications plan will include:
        - Press Release, Public Affairs Guidance and social media materials that can be tailored to each port's operation.
        - Tear sheets and signage for ports of entry.
        - IDS Slides and internal communications.
- As new facial biometric comparison deployments roll-out, new paperless and CBP One™ messaging will be incorporated.

Recommendation:

Approve the plan and sign memorandum for issuance to U.S. Border Patrol. The PPAE Executive Director will issue the memo to the field.

(b)(6), (b)(7)(C)

Approved/Date: (b)(6), (b)(7)(C) 05/26/2021 **Disapproved/Date:** _____

Needs Discussion/Date: _____ **Modify/Date:** _____

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5)

U.S. Customs and Border Protection

CBP One™ : Allows **International Organizations** to submit advance information for individuals

01

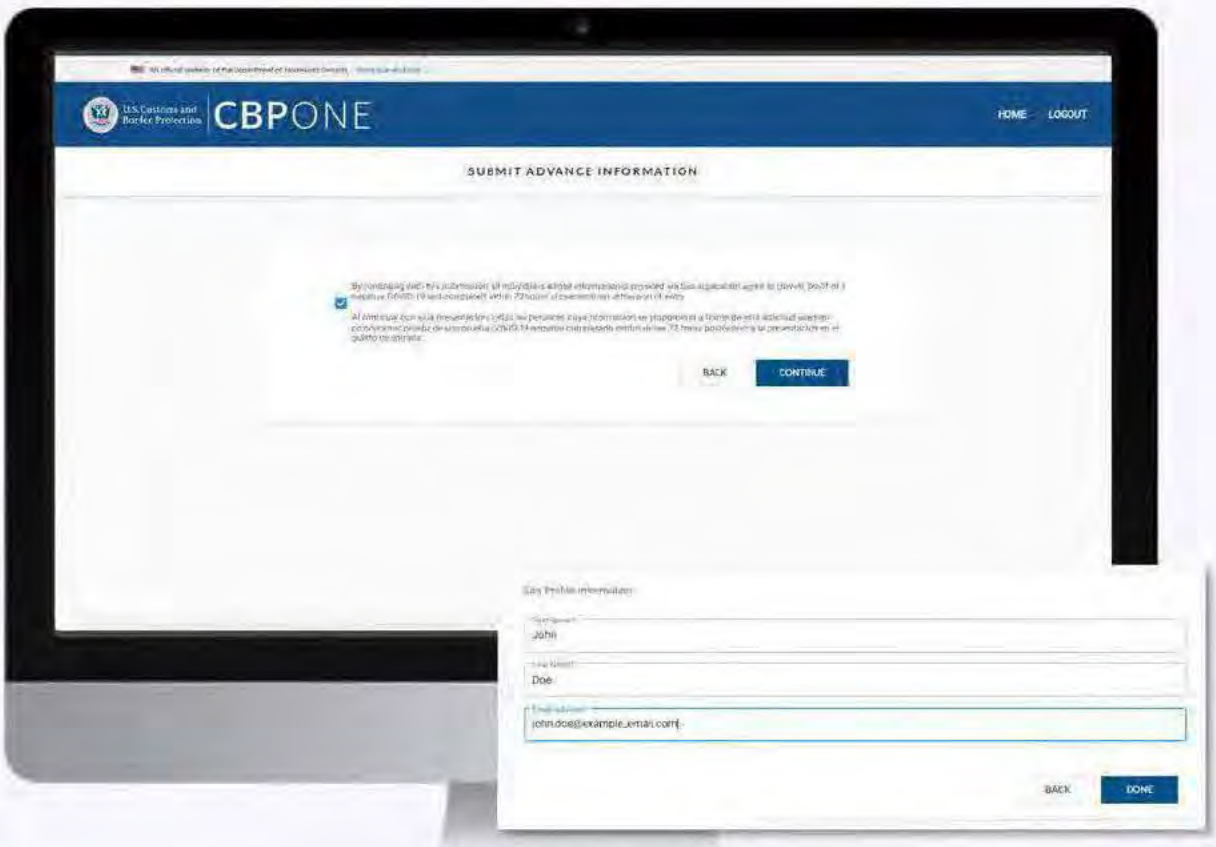
To get started, go to https://cbpone.cbp.dhs.gov/

Tap on "International Organization" to begin and select "Submit Advance Information".

02

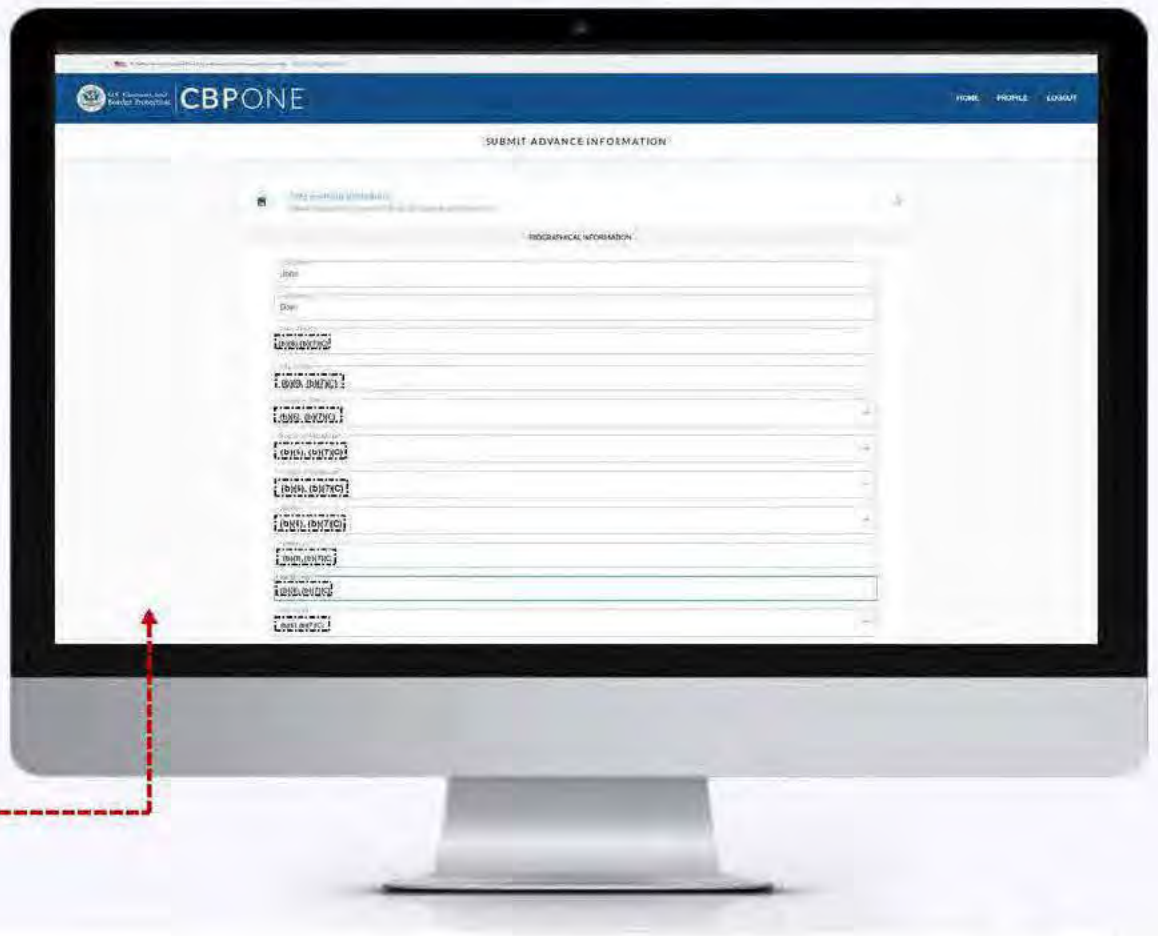
Acknowledge the COVID-19 checkbox to proceed. First time users will be prompted to create a profile.

Note: All individuals whose information is provided via this application are required to provide proof of a negative COVID-19 test completed within 72 hours of presentation at the port of entry.



03

Select "Add Individual" and fill out the required fields for all individuals traveling together on this trip.

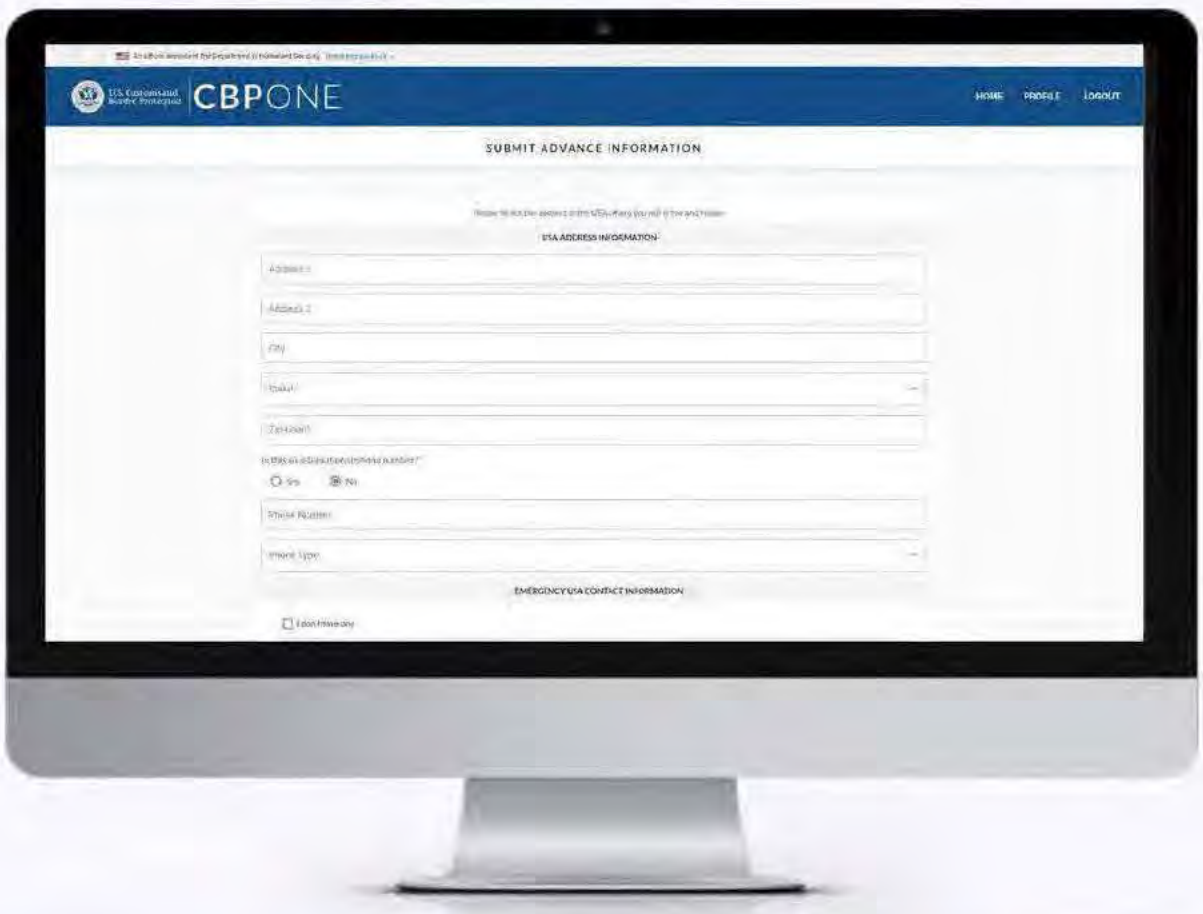


U.S. Customs and Border Protection

CBP One™ : Allows **International Organizations** to submit advance information for individuals

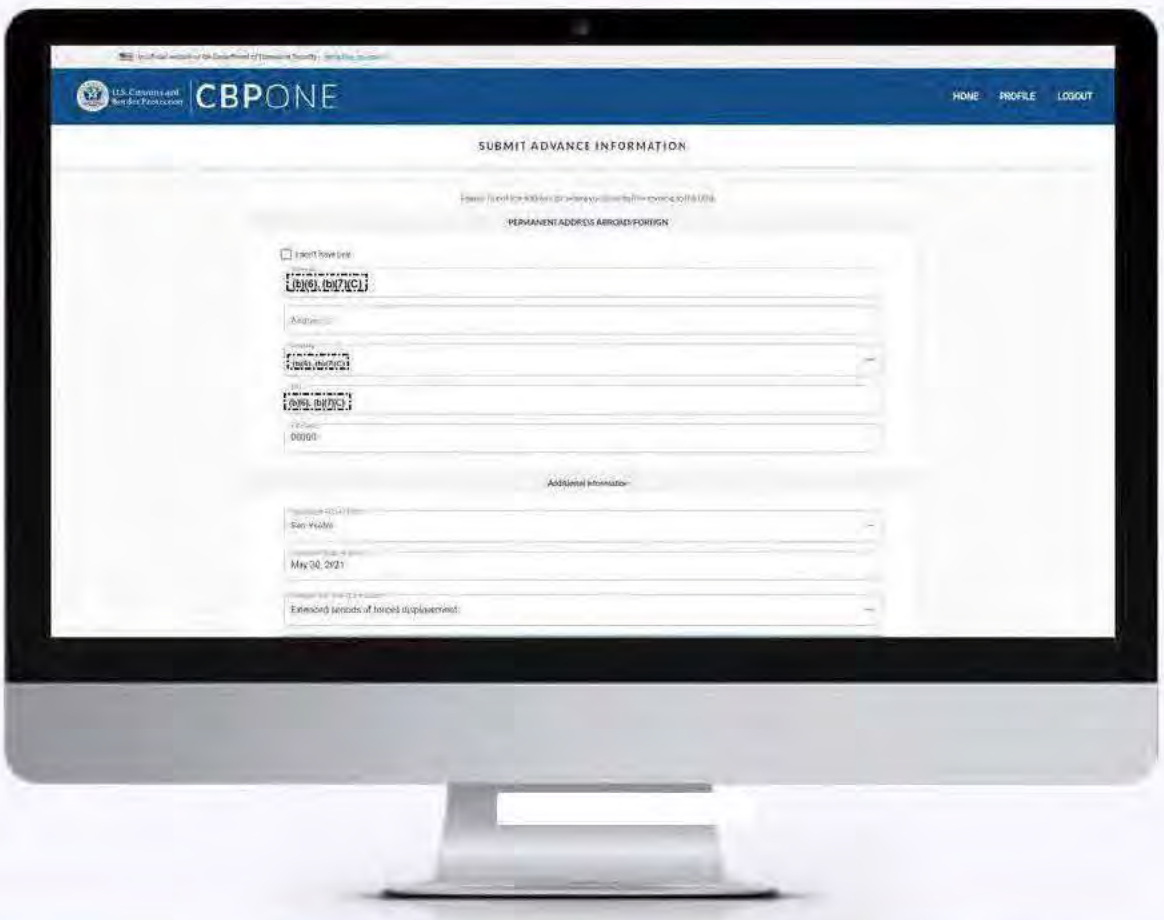
04

Fill out the address in the USA where the individual(s) will arrive and reside, along with an emergency USA point of contact.



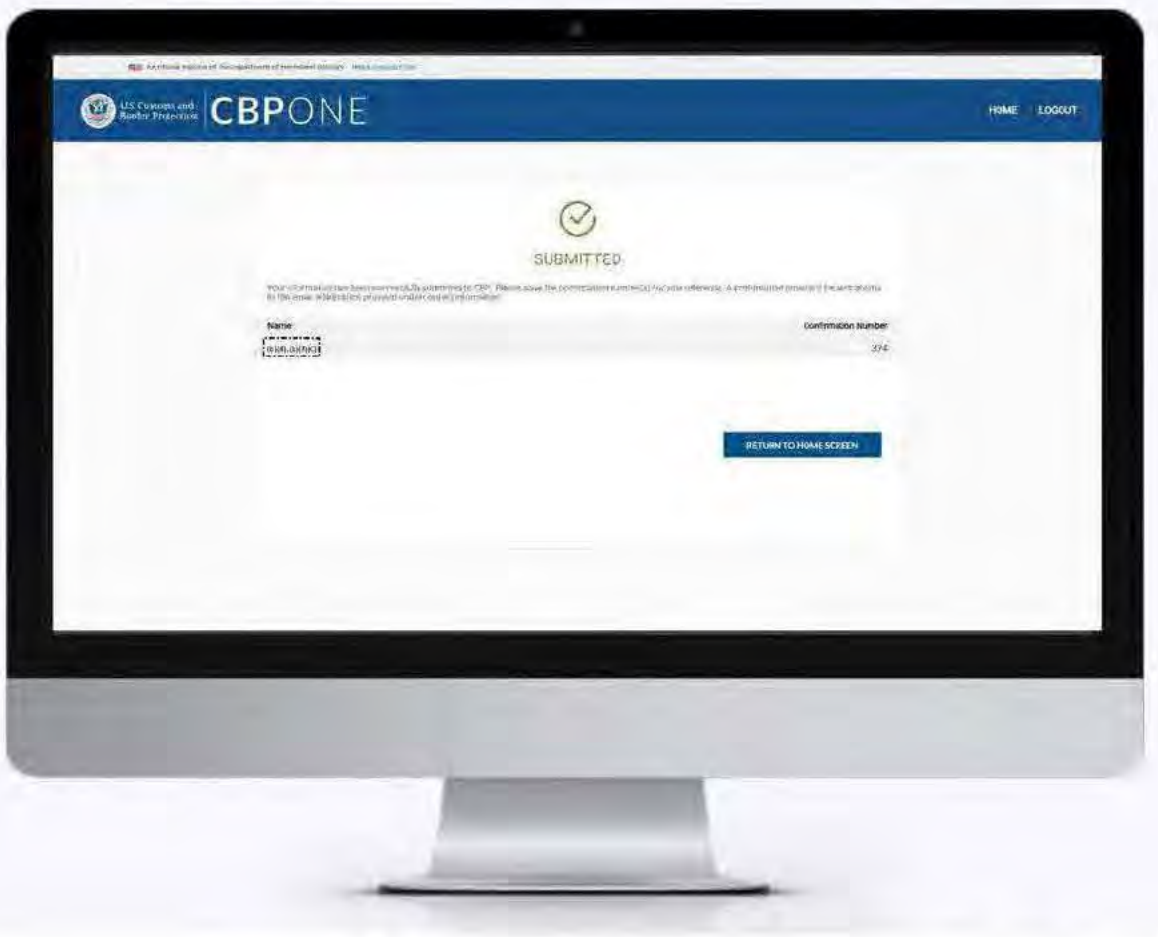
05

Lastly, fill out the address for where they lived before coming to the USA along with their requested port and date of entry, and Title 42 remarks.



06

Review all information and tap on submit. A confirmation screen will display the confirmation number(s) for each individual. A confirmation email will be sent to the email address(es) provided under contact information.



(b)(5)

(b)(5)

(b)(5)

U.S. Customs and Border Protection

December 8, 2020

MEMORANDUM FOR: (b)(6), (b)(7)(C)
Executive Assistant Commissioner
Enterprise Services

FROM: (b)(6), (b)(7)(C)
Executive Assistant Commissioner
Office of Field Operations

(b)(6), (b)(7)(C)

SUBJECT: 180-Day Update for U.S. Government Accountability Office Final Report: *FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (GAO-20-568)

Pursuant to the requirements of 31 U.S.C. Section 720, the Office of Field Operations (OFO) is submitting this written statement on actions taken, on-going, or planned regarding the U.S. Government Accountability Office (GAO) recommendations contained in its report, *CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (GAO-20-568).

The report contained five recommendations for U.S. Customs and Border Protection (CBP), OFO.

Recommendation 1: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's (BEEP) privacy notices contain complete and current information, including all of the locations where facial recognition is used and how travelers can request to opt out as appropriate.

Original Response: Concur. CBP's OFO will collaborate with the CBP Office of Public Affairs to publish: 1) Biometric Entry-Exit privacy information; 2) locations where facial recognition is used; and 3) traveler opt-out procedures on CBP's public-facing website, as well as to review and update that information on a monthly basis. CBP's OFO will also ensure that information provided in response to inquiries via the CBP Call Center is also reviewed and updated monthly.

Estimated Completion Date (ECD): December 31, 2020

Update to Recommendation: CBP launched its updated biometrics website on September 1, 2020 (www.biometrics.cbp.gov). The purpose of the site is to deliver information to the public and other stakeholder groups. The site provides a user-friendly communication channel for promoting facial comparison technology and biometrics information in a dynamic and inviting manner. As a testament to CBP's commitment to privacy protections, outlined in the DHS Fair

Information Practice Principles, the website includes the current locations using facial comparison technology as well as information on how to request alternative screening and copies of CBP's privacy signage on display. The information provided, including a link to CBP's Traveler Verification Service (TVS) Privacy Impact Assessment (PIA), is yet another tool in CBP's arsenal to ensure technology sustains and does not erode privacy protections.

CBP OFO met with the CBP Call Center Teams and discussed the Centers' existing and needed biometric resources. On October 1, 2020, OFO conducted three TVS 101 briefings for both the Traveler Call Center and Information Center staff. OFO provided the Call Center with the Public Affairs Guidance, to include Frequently Asked Questions (FAQs). In addition, OFO provided the Information Center with a few FAQs to post on www.help.cbp.gov.

CBP will continue to ensure that content is up to date on the CBP website, as required, and when updates are made we will provide new details to both the Info Center and Call Center.

Recommendation 2: The Commissioner of CBP should ensure that the BEEP's privacy signage is consistently available at all locations where CBP is using facial recognition.

Original Response: Concur. It is important to note that, unlike Federal Inspection Services areas, the airport departure areas are not managed by CBP personnel. However, CBP OFO will continue to work with its airlines/airport partners to ensure that privacy signage is available, on display, and reflective of current privacy messaging for travelers. For example, CBP provides notice to individuals regarding the collection, use, dissemination, and maintenance of personally identifiable information as part of efforts to promote transparency. While CBP acknowledges that operational constraints may affect the placement of signs or the timely posting of updated signage, the overall public is informed that stakeholders are taking photos in coordination with CBP. Further, CBP's OFO regularly conducts periodic signage audits that include local CBP personnel to ensure signs are accurate and placed appropriately.

In addition, CBP notifies travelers at these ports using verbal announcements, signs, and message boards, as appropriate, that CBP takes these photos for identity verification purposes. Travelers are also informed of their ability to request alternative identity verification procedures. Also publicly stated are notifications that, should a traveler decide to request alternative identity verification procedures, the airline would conduct manual identity verification using his/her travel document, and may notify CBP to collect biometrics, such as fingerprints, if applicable. CBP's OFO will also continue to work with airline and airport partners to identify other methods to communicate the use of facial recognition and travelers' privacy rights.

ECD: June 30, 2021

Update to Recommendation: OFO is working with CBP's Enterprise Services, Office of Facilities and Asset Management, Administrative Services Program Management Office, Printing, Graphics and Distribution Branch, to ship out additional signs with the

most recent/updated language, to the ports of entry (POE). OFO is also drafting policy guidance to the Field Offices/POEs requesting a signage auditor point of contact who would be responsible for ensuring signage (post deployment) is correct and on display.

Every Simplified Arrival deployment includes a Signage Team to ensure current signs are correctly displayed and visible to the traveling public. The CBP Signage Team also conducts an exit sign audit to ensure current language is displayed.

Recommendation 3: The Commissioner of CBP should direct the BEEP to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information.

Original Response: Concur. In the air exit environment, CBP OFO will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, as detailed in the BEEP audit plan, provided to GAO in April 2020. This audit plan enables a comprehensive review of compliance with security and privacy requirements on the part of CBP and CBP's partners. As mentioned in the draft report, CBP completed one partner audit thus far.

Although, CBP planned additional audits for 2020, due to the COVID-19 global health pandemic and subsequent travel restrictions, CBP paused the planned audit activities. Once pandemic travel restrictions are lifted, CBP's OFO and Office of Information and Technology (OIT) will resume conducting audits. (b)(5)

(b)(5)

ECD: June 30, 2021

Update to Recommendation: Due to the COVID-19 global health pandemic and subsequent travel restrictions, CBP paused the planned audit activities; (b)(5)

(b)(5)

(b)(5)

Recommendation 4: The Commissioner of CBP should develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.

Original Response: Concur. The CBP's BEEP's Air Exit Segment was granted Acquisition Decision Event 3 in December 2019. One of the action items from this decision was to complete an update to the Operational Requirements Document (ORD). (b)(5)

(b)(5)

ECD: June 30, 2021

Update to Recommendation: (b)(5)
(b)(5)

Recommendation 5: The Commissioner of CBP should develop a process by which BEEP officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Original Response: Concur. CBP's OFO has a suite of tools that allows for system and operational performance management and OFO generates three types of performance reports that are automatically generated and distributed weekly within CBP and to external stakeholders. These reports include:

1) Saturation Report: Notes the percentage of flights biometrically processed out of the total number of possible international departures segmented by airport.

2) Biometric Air Exit Overview Report: Includes a daily synopsis of operational performance data including numbers of biometrically processed flights and travelers together with biometric match rates.

3) Stakeholder Raw Data Reports: Provides Air Exit stakeholders with operational performance data by flight number, passenger counts, and biometric match rates.

The OFO's Biometric Entry-Exit Air team monitors the reports for performance issues and addresses any anomalies with stakeholders as they arise. The reports are also used to promote/increase usage by stakeholders.

CBP's OFO also conducts random sampling to determine the technical match rates and identify any system or equipment issues. The random sampling is conducted on a weekly basis and includes two flights per airport per week. CBP's OFO receives alert notifications if TVS experiences an outage, and has a Gallery Assembly System monitor that provides notifications when a flight gallery is not created. Depending on the severity and impact to end users, OFO generates stakeholder notifications, as appropriate.

CBP requests that GAO consider this recommendation resolved and closed, as implemented.

ECD: Pending Complete

Update to Recommendation: Awaiting GAO's confirmation that this recommendation is closed, as implemented.

If we can be of further assistance, please contact me, at (b)(6), (b)(7)(C) or have a member of your staff contact (b)(6), (b)(7)(C) Executive Director, Planning, Program Analysis, and Evaluation, at (b)(6), (b)(7)(C)

Attachment

Transportation Security Administration and U.S. Customs and Border Protection: Deployment of Biometric Technologies

Report to Congress

Homeland Security

U.S. Department of Homeland Security

Assistant Secretary for Legislative Affairs
U.S. Department of Homeland Security
Washington, DC 20528

Homeland Security

August 30, 2019

Message from the Assistant Secretary for Legislative Affairs

I am pleased to present the following report, "Transportation Security Administration and U.S. Customs and Border Protection: Deployment of Biometric Technologies," which has been prepared by the Transportation Security Administration (TSA) and the U.S. Customs and Border Protection (CBP). This report is required by Section 1919 of the *FAA Reauthorization Act of 2018* (P.L. 115-254), signed into law on October 5, 2018.

The report describes CBP and TSA's development and implementation of biometric technology pilots. It includes assessments on the operational and security impact of biometric technology; potential effects on privacy with the expanded use of biometric technologies methods to mitigate privacy risks; methods to analyze and address matching performance errors; and special assessments on the biometric entry-exit program.

This report is being provided to the following Members of Congress:

The Honorable Roger Wicker
Chairman, Senate Committee on Commerce, Science, & Transportation

The Honorable Maria Cantwell
Ranking Member, Senate Committee on Commerce, Science, & Transportation

The Honorable Ron Johnson
Chairman, Senate Committee on Homeland Security and Governmental Affairs

The Honorable Gary C. Peters
Ranking Member, Senate Committee on Homeland Security and Governmental Affairs

The Honorable Bennie G. Thompson
Chairman, House Committee on Homeland Security

The Honorable Mike Rogers
Ranking Member, House Committee on Homeland Security

Please do not hesitate to contact us at (202) 447- 5890 if we may be of further assistance.

Respectfully,

Christine M. Ciccone

CHRISTINE M. CICCONE
Assistant Secretary for Legislative Affairs

Executive Summary

TSA, CBP, travelers, and travel industry partners recognize that identity and vetting are critically important elements in the air environment. Travelers are repeatedly asked to prove their identity within the travel continuum. Governments and industry partners must repeatedly verify travelers' asserted identity at check-in, bag-check, security checkpoint, and at departure. Projected increases in air travel volume, combined with current infrastructure and operational constraints, underscore the need to automate current processes. Facial biometric technology has potential to modernize and streamline the process without sacrificing safety and security by reducing the reliance on manual identity verification processes.

At the direction of Congress, CBP developed a pilot biometric entry-exit program to aid in the identity verification of travelers upon entry into and exit from the United States. CBP and the Department of Homeland Security (DHS) invested in developing an identity as a service solution (IDaaS) that uses facial comparison to automate manual identity verification. This solution is called the Traveler Verification Service (TVS). The biometric entry-exit program is carried out through a privacy-by-design model and firmly situated within the DHS Fair Information Practice Principles.[1]

TVS offers a secure system that works quickly and reliably. It uses existing traveler data to build small galleries of faces associated with each departing flight and enables CBP and its partners such as TSA, select air carriers and airport authorities to simply take and submit a traveler's photo for identity verification. Live photos are compared against the correlating flight gallery[2] and TVS returns verification results in seconds. For travelers at the gate, this means the traveler's facial biometric can serve as a boarding pass. For industry partners, it can mean a convenient, efficient, and safe travel experience redefined by biometrics.

CBP established a rigorous process to review data associated with matching performance of biometric facial comparison. Although TVS true match rates can vary, CBP's analysis found a negligible effect in regards to biometric matching attributed to demographic variables. Further, because data privacy, protection, and mitigation of algorithmic or operational bias are prime concerns, CBP actively makes improvements while seeking to ensure there are no signs of bias,[3]

---

[1] See DHS Privacy Policy Directive 140-06, available at: https://www.dhs.gov/publication/privacy-policy-guidance-memorandum-2008-01-fair-information-practice-principles.

[2] A pre-positioned "gallery" of traveler face templates is created using the biographic data from the airline manifest to retrieve the photos from government holdings, such as passports, visas, and previous entries.

[3] CBP measures and evaluates true match and non-match rates, as well as false match and non-match rates to provide a comprehensive understanding of system effectiveness in alignment with its mission. CBP analyzes for

and engages in a robust public dialogue on appropriate standards. CBP also engages in outreach with privacy advocates, the National Institute of Standards and Technology (NIST), and U.S. Department of Homeland Security's (DHS) Science and Technology Directorate (S&T) to monitor performance and progress.

While TSA is not evaluating the use of facial comparison for law enforcement purposes, it is assessing its use for traveler identity verification as part of its mission to protect the Nation's transportation systems to ensure freedom of movement for people and commerce. TSA is using CBP's TVS for international travelers in this assessment process. In October 2018, TSA published the *TSA Biometrics Roadmap for Aviation Security and the Passenger Experience.*[4] The Biometrics Roadmap defines clear pathways to improve security, safeguard the Nation's transportation system, and accelerate the speed of action through smart investments and collaborative partnerships. In pursuing these goals, TSA seeks to use innovative collaboration concepts and solutions to enhance security effectiveness, improve operational efficiency, and yield a consistent, streamlined traveler experience. As it works to test the use of opt-in facial image collection and matching processes for additional populations, including TSA Pre✓® travelers and the general flying public, TSA is grounding its solutions in rigorous scientific study and analysis. TSA is committed to protecting traveler privacy as part of its biometrics effort, and as such, incorporates privacy considerations into each phase of biometric solution development.

Beginning in March 2017, CBP and TSA began evaluating the use of facial comparison at the security checkpoint through a series of multi-phased pilots. Early success on initial proof of concept testing in October 2018 encouraged TSA and CBP to explore the viability of expanded use of TVS at the checkpoint through data integration between TVS and TSA Secure Flight systems. Both agencies will continue to build on their efforts to evaluate the ways in which biometrics technology can improve the traveler experience. TSA and CBP are committed to enhancing security consistent with their homeland security missions and biometrics efforts, including facial comparison.

---

demographic biases in its biometric exit systems. No bias based on demographics has been statistically identified in its approach. However, operational and environmental conditions, such as lighting, show much greater correlation.

[4] https://www.tsa.gov/sites/default/files/tsa_biometrics_roadmap.pdf

U.S. Transportation Security Administration and U.S. Customs and Border Protection: Deployment of Biometric Technologies

Table of Contents

I. Legislative Language

This Report to Congress was compiled pursuant to Section 1919(c) of the *FAA Reauthorization Act of 2018* (P.L. 115-254), signed into law on October 5, 2018, which states in part:

> (c) REPORT REQUIRED.—Not later than 270 days after the date of enactment of this Act, the Secretary shall submit to the appropriate committees of Congress, and to any Member of Congress upon the request of that Member, a report that includes specific assessments from the Administrator and the Commissioner of U.S. Customs and Border Protection with respect to the following:
>
> (1) The operational and security impact of using biometric technology to identify travelers.
> (2) The potential effects on privacy of the expansion of the use of biometric technology under paragraph (1), including methods proposed or implemented to mitigate any risks to privacy identified by the Administrator or the Commissioner related to the active or passive collection of biometric data.
> (3) Methods to analyze and address any matching performance errors related to race, gender, or age identified by the Administrator with respect to the use of biometric technology, including the deployment of facial comparison technology;
> (4) With respect to the biometric entry-exit program, the following:
>
> (A) Assessments of— (i) the error rates, including the rates of false positives and false negatives, and accuracy of biometric technologies; (ii) the effects of biometric technologies, to ensure that such technologies do not unduly burden categories of travelers, such as a certain race, gender, or nationality; (iii) the extent to which and how biometric technologies could address instances of travelers to the United States overstaying their visas, including— (I) an estimate of how often biometric matches are contained in an existing database; (II) an estimate of the rate at which travelers using fraudulent credentials identifications are accurately rejected; and (III) an assessment of what percentage of the detection of fraudulent identifications could have been accomplished using conventional methods; (iv) the effects on privacy of the use of biometric technologies, including methods to mitigate any risks to privacy identified by the Administrator or the Commissioner of U.S. Customs and Border Protection related to the active or passive collection of biometric data; and (v) the number of individuals who stay in the United States after the expiration of their visas each year.
>
> (B) A description of— (i) all audits performed to assess— (I) error rates in the use of biometric technologies; or (II) whether the use of biometric technologies and error rates in the use of such technologies disproportionately affect a certain race, gender, or nationality; and (ii) the results of the audits described in clause (i).

(C) A description of the process by which domestic travelers are able to opt-out of scanning using biometric technologies.

(D) A description of— (i) what traveler data is collected through scanning using biometric technologies, what agencies have access to such data, and how long the agencies possess such data; (ii) specific actions that the Department and other relevant Federal departments and agencies take to safeguard such data; and (iii) a short-term goal for the prompt deletion of the data of individual United States citizens after such data is used to verify traveler identities.

II. Background

Biometrics are recognized as unique physical characteristics that can be used to identify a person. Physiological traits such as fingerprints, facial images, iris patterns, hand geometry, speech, and gait, are all examples of biometric indicators. Today, biometrics are commonly used to accurately identify a person or authenticate an individual's identity. The U.S. Department of Homeland Security (DHS) uses biometric information for a variety of mission purposes. For example, U.S. Customs and Border Protection (CBP) uses biometrics as part of its border security mission and under its mandate to establish and implement a biometric entry-exit system. As part of its mission to protect the Nation's transportation systems and to ensure freedom of movement for people and commerce, the Transportation Security Administration (TSA) is exploring the use of biometrics for identity verification for both traveler screening, and to enable access to airport sterile areas by airport workers.

Over the past decade, significant developments and improvements in biometrics technology have occurred. At the same time, the use of biometrics technology has also prompted concerns about accuracy, privacy, and security, among other issues. While CBP and TSA explore the use of biometrics consistent with their respective missions, they are mindful of those considerations as well as the need to build to and utilize enterprise biometric services offered through DHS's Office of Biometric Identity Management (OBIM).

A. CBP's Progress Toward a Biometric Exit System

CBP has used biometrics to verify the identities of foreign nationals entering the United States at air ports of entry since the mid-2000s. In recent years, it has also made significant progress towards achieving a biometric entry and exit solution mandated by federal statute and executive orders. Under existing laws[5] and Executive Order 13780,[6] CBP is required to implement measures to verify identities of travelers upon entry to and exit from the United States. After receiving the biometric entry-exit mission in 2013 and through the authorization of fee funds,[7] CBP accelerated the implementation of a capability to biometrically verify the identities of travelers arriving and departing the United States by air while facilitating travel processes.

In 2017, after several successful biometric pilots, CBP began vetting the Traveler Verification Service (TVS), a facial image matching service that uses biographic data to retrieve all associated traveler facial images from DHS holdings and segment them into smaller, more manageable data sets,[8] for use in the live environment. TVS uses the product of a fusion of

---

[5] *See, e.g., Intelligence Reform and Terrorism Prevention Act of 2004* (Pub. L. No. 108-458, 118 Stat 3638 (2004)) and the *Implementing Recommendations of the 9/11 Commission Act of 2007* (Pub. L. No. 110-53, 121 Sat. 266 (2007)).

[6] https://www.federalregister.gov/documents/2017/03/09/2017-04837/protecting-the-nation-from-foreignterrorist-entry-into-the-united-states

[7] The *FY 2016 Consolidated Appropriations Act* (P.L. 114-113) funded the Biometric Entry-Exit Program through the authorization of up to $1B in fee collections on H-1B and L-1visa applications through FY 2025.

[8] For example, by flight, by cruise, or by frequent border crossers.

biometric and biographic information, enabling the biometric data to be the key to verify the traveler identity with the advance biographic data. The matching service compares the traveler's live photo to source images such as the travel document, enabling CBP to confirm the entry and departure of in-scope,[9] aliens. TVS was initially demonstrated at airports across the United States, as well as in the sea environment in 2017. CBP began piloting the capability at land ports of entry in the pedestrian environment in August 2018.

CBP's facial matching service is being leveraged to support biometric entry and exit processing for sea and land operations. Each travel mode offers unique challenges that require integrated solutions to mitigate any potential negative impacts to travel and trade. Biometric solutions must be thoroughly designed and tested to ensure that they are effective; compatible with expediting travel; can be integrated into existing infrastructure, systems, and processes; are not cost prohibitive, and do not put individuals' privacy at undue risk.

Figure 1 Biometric Entry Exit Statistics (as of June 2019)

Air Entry and Exit

CBP envisions the facial matching service will significantly reduce the need to manually check paper travel documents by providing an automated process which can replace manual checks of travel document across the travel continuum. In 2017, CBP demonstrated TVS at eight international airports at boarding gates using CBP officers to process each traveler. CBP also partnered with JetBlue Airways, Delta Air Lines, British Airways, and Los Angeles International Airport (LAX) to evaluate biometric exit boarding integrated with stakeholder departure control systems. In Fiscal Year (FY) 2018, CBP's transformed entry process using facial comparison was reengineered and deployed in the air entry environment at 15 airports including four preclearance locations, with plans to expand further in 2019.

[9] An "in-scope" traveler is any person who may be required by law to provide biometrics upon entry into the United States pursuant to 8 CFR 235.1(f)(ii), or upon exit from the United States pursuant to 8 CFR 215.8. "In-scope" travelers include any alien other than those specifically exempt as outlined in the CFR. Exempt aliens include: Canadian citizens under Section 101(a)(15)(B) of the Immigration and Nationality Act who are not otherwise required to present a visa or be issued a form I-94 or Form I-95; aliens younger than 14 or older than 79 on the data of admission; aliens admitted A-1, A-2, C-3 (except for attendants, servants, or personal employees of accredited officials), G-1, G-2, G-3, G-4, NATO-1, NATO-2, NATO-3, NATO-4, NATO-5, or NATO-6 visas, and certain Taiwan officials who hold E-1 visas and members of their immediate families who hold E-1 visas unless the Secretary of State and the Secretary of Homeland Security jointly determine that a class of such aliens should be subject to the requirements of paragraph (d)(1)(ii); classes of aliens to whom the Secretary of Homeland Security and the Secretary of State jointly determine it shall not apply; or an individual alien to whom the Secretary of Homeland Security, the Secretary of State, or the Director of Central Intelligence determines it shall not apply.

Prior to departure, the TVS creates a pre-positioned "gallery" of traveler face templates using the biographic data from the airline manifest to retrieve the photos from government holdings, such as passports, visas, and previous entries. During boarding, the stakeholder system takes a picture of the traveler. The TVS compares the picture against the gallery and provides a biometric match result.

Due to the success of CBP's stakeholder engagement strategy to date, CBP has received letters of commitment from 26 airports and airlines to begin implementation of biometric exit using CBP's matching service. CBP is actively working with each committed partner to implement biometric exit solutions. In FY2018, biometric air exit started at nine airports and ended at 16 airports. Total in-scope travelers exiting the country processed started at 40,000 monthly and ended FY2018 with 157,000 monthly. These numbers continued to grow steadily during FY2019, growing 54% since the beginning of the calendar year, with 548,000 being processed in the month of April 2019.

By 2022, CBP's goal is to deploy biometric exit to the top 20 airports, which account for more than 97 percent of departing commercial air travelers from the United States. CBP is actively working to expand stakeholder partnerships and adoption, prioritizing the highest volume of international airports and carriers to achieve the biometric air exit implementation goal. CBP continues efforts to consider innovative ways to utilize TVS with mobiles phones, tablets and watches. CBP will look to expand partnerships with international airports and governments and to further expand capabilities in preclearance locations to continually improve security and facilitation of traveler processes.

Sea Environment

Leveraging the investment in TVS for the air environment, CBP is partnering with the cruise industry to modernize traveler and crew inspections by implementing facial matching technology in the sea environment. Preparations are underway to apply the use of facial comparison technology in the debarkation (arrival) and embarkation (departure) points at seaports. These improvements will enable increased security and enforcement as well as facilitate traveler inspections.

Today, five major cruise lines are engaged with CBP to develop facial biometric processing supported by the TVS for closed-loop cruises.[10] Going forward, a focus on expanding integration with cruise partners will be implemented, focused initially on closed-loop cruises for debarkation. Through FY2020, CBP will seek to expand across closed-loop embarkation. Beyond FY2020, capabilities will be expanded to open-loop cruise routes.

Land Environment

[10] A closed-loop cruise is a term that refers to a cruise itinerary which begins and ends at the same U.S. location. An open-loop cruise is one that begins and ends in different ports, either departing from or arriving in the United States.

The Land Biometric Exit strategy focuses on implementing an interim exit capability while simultaneously investigating innovative technologies to reach the long-term goal of a comprehensive exit solution. CBP is actively piloting capabilities at the land border in both the pedestrian and vehicle environments to determine the best long-term approach for a comprehensive biometric entry-exit capability. Since September 2018, 139 impostors were identified on entry using the TVS capability in a land pedestrian environment. Details on the challenges of implementing biometrics in the land border are detailed in section VI, and CBP's strategy to mitigate those challenges are in section IV.

In late 2017, CBP began the initial implementation of an interim land exit approach to provide a capability for CBP to report the final departure from the United States of third-country nationals at land ports of entry.[11] The third country nationals' capability is a short-term solution that leverages the biometric exit mobile platform from the air environment and allows compliant in-scope travelers a means to biometrically report departure. Since January 2018, more than 180 mobile devices have been deployed to 74 land border ports of entry to support this initiative. CBP personnel have deployed to more than 50 locations to provide training courses for the mobile app to support these deployments.

CBP will continue to evaluate concepts of operation and technologies in the land environment to determine the final approach. Solutions being evaluated leverage the underlying TVS architecture in both the pedestrian and vehicle environments.

B. CBP and TSA Partnership to Evaluate Biometrics at the Checkpoint

In March 2017, CBP and TSA began evaluating the use of facial comparison at the TSA checkpoint for identity verification. In April 2018, the TSA Administrator and CBP Commissioner signed a policy memorandum promoting a collaborative approach to the continued development and use of biometric technology at airports.

The goal of the partnership is to enhance security and promote effective use of resources. CBP and TSA established multi-phased pilots involving volunteer international travelers. The first phase at John F. Kennedy International Airport (JFK) began in October 2017 to collect data and validate the technology. In the second phase at LAX in August 2018 and Hartsfield-Jackson Atlanta International Airport (ATL) in November 2018, TSA used CBP's TVS to test biometrics for identity verification in an operational environment. In the third phase, CBP and TSA will explore data-sharing and integration between biometric and traveler vetting systems. The goal will be to create a consolidated traveler identity verification that meets the operational needs of both agencies. In 2019, CBP and TSA plan to continue working on the necessary technical integration and pilot planning activities. The results of the pilot will help inform the rollout plans at TSA checkpoints.

---

[11] [11] DHS/CBP/PIA-026(a), *Biometric Exit Mobile Program* (June 29, 2018), https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp026a-bemobile-june2018.pdf.

C. TSA's Exploration of Biometrics

TSA protects the Nation's transportation systems to ensure freedom of movement for people and commerce. The TSA Administrator's 2018-2026 Strategy[12] details the three strategic priorities that will guide the agency as it seeks to further enhance transportation security.

- **Improve security and safeguard the transportation system:** TSA will lead by example by strengthening operations through powerful and adaptable detection capabilities, intelligence-driven operations, and enhanced vetting.
- **Accelerate action:** TSA will build a culture of innovation that anticipates and rapidly counters the changing threats across the transportation system. TSA will develop its ability to make timely, data-driven decisions, and rapidly field innovative solutions.
- **Commit to our people:** TSA will foster a diverse, inclusive, and transparent work environment, establishing itself as a choice federal employer. TSA will use available tools and authorities to cultivate a skilled workforce equipped to meet the challenges of tomorrow.

Identity verification and traveler vetting are integral to TSA's multi-layered security processes and core security mission. The identity verification process ensures that the person seeking access to the airport sterile environment is the person who was vetted by TSA's Secure Flight against intelligence-driven watch lists, and receives the appropriate level of physical screening. Currently, TSA relies on a manual identity verification process through which Transportation Security Officers (TSOs) and, for checked baggage, airline employees, verify a traveler's identity by manually reviewing their boarding pass and a valid form of identification (ID). For photo ID documents, TSOs must visually confirm the photo on the document matches the traveler. Once a TSO confirms a traveler's identity, he/she direct the traveler to proceed to security screening based on their Secure Flight vetting status as it appears on the boarding pass. Automated facial recognition capabilities can play an important role, in increasing the effectiveness of this travel document checker (TDC) position at the checkpoint.

TSA is deploying Credential Authentication Technology (CAT) to increase security at checkpoints. CAT addresses ID fraud vulnerabilities by verifying the security features on a traveler's ID and boarding pass. CAT also provides automated access to real-time Secure Flight traveler vetting information at the checkpoint. In the future, biometrics will complement the capabilities CAT offers by enabling TSA to match the person's facial image against the facial image on file or on their ID.

In 2013, TSA established the TSA Pre✓® Application Program. Under this trusted traveler program, TSA conducts significant additional vetting of applicants; those individuals that TSA has determined are low risk are then eligible for expedited screening at participating U.S. airports. Members of the traveling public voluntarily pay a fee and provide their biographic and information and fingerprints to conduct the enrollment and vetting to check an applicant's criminal history, potential ties to terrorism, enrollment eligibility, and citizenship. As of September 2018, TSA has transitioned from single-factor biometric enrollment (fingerprints) to

[12] Available at: https://www.tsa.gov/sites/default/files/tsa_strategy.pdf.

multi-modal biometric enrollment (fingerprints and face), so that facial images can be used for identity verification.

In June 2017, TSA assessed, as a proof of concept, the use of biometric authentication technology to verify the identity of TSA Pre✓® travelers. As part of this proof of concept, TSA compared a fingerprint scanned using this technology with the fingerprint provided at the time of TSA Pre✓® enrollment. This proof of concept demonstrated the potential for biometrics to enhance security through increased assurance of traveler identity. It also underscored the need for additional work to explore other biometric technologies, such as facial images, and integrate those biometrics into airport checkpoint operations.

Additionally, in 2018, TSA conducted a three-week proof of concept at LAX using facial comparison to provide automated verification of identities at the TDC. This proof of concept was available to e-Passport[13] holders who volunteered to test the technology. Travelers scanned their e-Passports to verify the name on the e-Passport matched the name on the traveler's boarding pass. If it matched, the system extracted the traveler's digital photo from the e-Passport chip. The traveler was then prompted to complete a photo capture with a facial comparison camera. Facial comparison technology compared the e-Passport photo to the real-time photo and prompted the e-Gate to open if they matched. After the e-Gate opened, the travelers proceeded to the TDC; those who did not match were directed to the TDC officer. All passengers were required to complete the standard TDC process for manual identity and travel document verification, regardless of the e-Gate biometric matching results.

Recognizing the need for TSA to take a more comprehensive approach to biometrics, Administrator David Pekoske championed the development of the Biometrics Roadmap,[14] published in October 2018. The roadmap provides the following:

- Defines clear pathways to improve security, safeguard the Nation's transportation system, and accelerate the speed of action through smart investments and collaborative partnerships;
- Incorporates feedback gathered during more than 40 engagements with aviation security leaders from airlines, airports, and solution providers; and
- Includes feedback gathered from key government stakeholders, including TSA internal offices, DHS headquarters, and operational components.

[13] E-Passports contain an electronic chip that holds the same information that is printed on the passport's data page including a digital photograph of the holder. See https://www.dhs.gov/e-passports.
[14] Available at: https://www.tsa.gov/sites/default/files/tsa_biometrics_roadmap.pdf.

It also outlines four goals to achieve TSA's vision for biometrics.

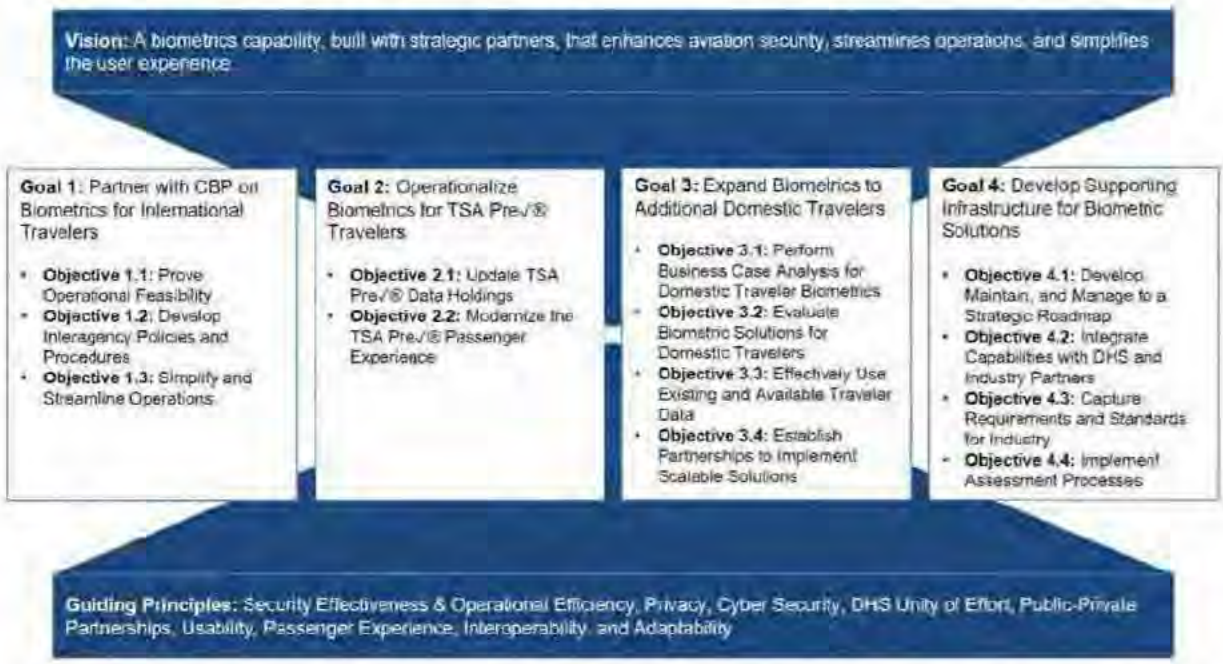


Figure 2 TSA's Vision, Goals, Objectives, and Principles for Checkpoint Biometrics

Goal 1: In partnering with CBP on biometric technology pilots, TSA is exploring the feasibility of applying biometric solutions at the TSA checkpoint. While CBP and TSA mission requirements differ in some regards, CBP's biometric air exit program offers the opportunity to conduct joint operational pilot projects, collect data, refine solutions, and exchange data. TSA's partnership with CBP will also enable TSA to identify and examine technical, legal, and regulatory issues before broader deployment.

Goal 2: To further implement biometrics for TSA Pre✓®, TSA continues enhancing the trusted traveler experience for TSA Pre✓® travelers. As of September 2018, TSA is capturing photos for those who renew in person or who are enrolling for the first time in the TSA Pre✓® Application Program.

Goal 3: TSA will explore opt-in biometric solutions for additional travelers beyond international outbound and trusted travelers. An assessment of the appropriate authorities, privacy issues, and potential risks and benefits as it explores ways to improve the screening experience for standard (non-TSA Pre✓®) domestic travelers will be conducted. As TSA explores biometric solutions for additional travelers, it will conduct pilot projects and seek input from a diverse group of stakeholders. Additionally, TSA will continue to partner with DHS and interagency partners, including DHS Science and Technology (S&T) Directorate, and OBIM, as well as CBP, and the DHS Office of Privacy and DHS Office of Civil Rights and Civil Liberties, to evaluate biometric solutions for domestic travelers.

Goal 4: TSA will develop supporting infrastructure for biometric solutions that align with legal and policy authorities. TSA's biometrics efforts will also align with the DHS-wide transition to enterprise biometric services offered through OBIM's Homeland Advanced Recognition Technology (HART) system. Common standards will also allow TSA to establish assessment processes, making it possible to quickly evaluate security procedure changes, assess cybersecurity posture, develop qualified product and service lists, and implement audits and controls to ensure operations adhere to applicable laws, policies, and compliance authorities.

III. Operational and Security Impacts of Using Biometric Technology

Recognizing the important role that biometric technology can play in enhancing security and improving operations, CBP and TSA are methodically studying the impact of these technologies through a number of pilots and demonstrations. Though the operational and security factors that are driving the use of biometric technologies are distinct for both agencies, CBP and TSA's assessments are helping to refine biometric solutions and biometrics efforts throughout DHS.

On an average day, CBP processes more than one million travelers arriving at air, sea, and land ports of entry. Innovative technologies are being used to enhance a wide range of its operational capabilities. The use of biometrics, specifically facial comparison technology, assists CBP in confirming the departure of non-U.S. citizens and facilitates future processing at entry and exit. Through CBP's development of biometrics at entry-exit, it has found that biometrics are an effective tool in combatting the use of stolen and fraudulent travel and identity documents. The goal is to ultimately enhance identity verification while facilitating a more secure travel experience.

A. CBP Operational and Security Impacts

In addition to the responsibilities referenced in Section II B, CBP has the ongoing mission to inspect all incoming and departing travelers and conveyances to determine admissibility to the United States and enforce and administer U.S. immigration laws.

A key aspect of effective enforcement is the ability to discern individuals who are lawfully present in the United States from those who have violated their terms of admission. An effective immigration system requires an end-to-end process that collects exit data and matches that to entry data. Without exit data, there is no meaningful way to determine whether foreign nationals have overstayed their periods of admission.

Biometric data, when used with biographic data, allows CBP to confirm with greater assurance a traveler's true identity, ensuring the traveler matches the biographic identity that has been vetted through DHS databases. As biometric technology has evolved, the ability to use individual characteristics to confirm identity for all travelers, including U.S. citizens, is now a reality for all modes of transportation.

To implement a biometric entry-exit solution that is both operationally feasible and realistic, CBP established key parameters based on existing operational constraints and infrastructure limitations.

CBP's Key Strategic Parameters Table 1

Key Strategic Parameter	Description
Do not add another processing layer to known travel processes	Avoid a stove piped, independent approach by integrating biometrics into already existing travel processes.
Utilize existing infrastructure	The solution will work in existing port infrastructure for entry and exit processing.
Utilize existing business models	Leverage existing stakeholder (airline, cruise line) systems, processes, and business models.
Leverage current traveler behavior	Leverage traveler behaviors and expectations that require minimal new or unexpected steps for travelers.
Leverage existing data and IT infrastructure	Leverage existing traveler data, such as passport and visa information, and leverage existing government IT infrastructure as much as possible.
Utilize existing DHS enterprise biometric services, capabilities, and investments	Leverage and integrate with DHS Enterprise Services for shared biometric matching capabilities.

For the initial implementation of biometric exit solutions in the air environment, CBP is working in partnership with the air travel industry to lead the transformation of air travel using biometrics as the key to enhancing security and unlocking benefits, which will dramatically improve the entire traveler experience. The strategic benefits are described in the following table:

CBP Strategic Benefits Table 2

Strategic Benefit	Description
Improved business process	An enhanced entry-exit business process that integrates within existing government and stakeholder business models.
Stronger relationships	An environment that allows CBP and stakeholders to work together and that allows for further airline modernization.
A positive impact on inbound security and throughput	Enhanced inbound security and more efficient throughput.
Improved traveler experience	An overall enhanced traveler experience.
Improved data integrity	Utilize DHS enterprise biometric repositories provided to ensure accurate biometric identity records.
Enhanced visa overstay enforcement	Support the ID and tracking of visa overstays by closing information gaps associated with current exit reporting capabilities allowing for improved enforcement action.

CBP is transforming the way the agency identifies travelers by shifting the key to unlocking a traveler's record from biographic identifiers to biometric ones – primarily a traveler's face.

Pre-staging the existing traveler data upstream in the travel process enables all stakeholders to transform from manual and redundant processes to safer and automated traveler movement. CBP will continue to increase security by using a live facial biometric to match the traveler to advance traveler information, while also checking any existing fingerprints on file against the biometric watch list, which decreases dependency on less reliable paper travel documents, such as passports and visas. New facial comparison processes will enhance CBP's biometric capabilities alongside of the existing fingerprint processes.

CBP is partnering with the air travel industry and TSA to deploy a biometric air entry-exit solution that improves and streamlines the overall traveler experience. The four primary goals of this large-scale transformation is to make air travel more:

- **Secure** - Providing increased certainty as to the identity of travelers at multiple points in the travel continuum;
- **Simple** - Eliminating the need for physical document and boarding pass checks, as well as the collection of fingerprints;
- **Facilitative** - Establishing a clear and easily understood process that will reduce the potential for major "bottlenecks" within the air travel process; and
- **Compliant** - Employing a high integrity biometric entry and exit system that not only increases CBP's certainty as to the identity of travelers, but also more ably holds accountable those violating terms of admittance.

B. TSA Operational and Security Impacts

For TSA, biometrics can provide important benefits in air travel. TSA experienced a milestone year in 2018, screening a record setting 813.8 million travelers.[15] This amounts to more than 2 million travelers per day. TSA is already on track to exceed this in 2019. Like TSA, airlines, airports, and security regulators around the globe are faced with an ever-rising volume of air travelers to screen. In light of rising air travel volume and operational constraints, TSA must look to innovative technologies, like biometrics, to enhance security and efficiency while improving the traveler experience.

[15] https://www.tsa.gov/blog/2019/02/07/tsa-year-review-record-setting-2018

TSA evaluates potential changes to its aviation security programs and technology solutions through the lens of the Risk Mitigation Trade Space Framework.[16] The framework contains the following elements:

Figure 4 TSA's Risk Mitigation Trade Space Framework

- **Operational Efficiency** – What is the effect of a new security technology or procedure on operational footprint, wait times, and TSA's workforce staffing?
- **Security Effectiveness** – What is the effect of a new security technology or procedure on TSA's ability to detect, deter, or otherwise mitigate threats? How may adversaries shift their tactics in response to such changes?
- **Traveler Satisfaction** – What does the new technology or procedure do to improve the traveler experience?
- **Industry Vitality** – What, if any, is the economic impact of implementation? Is there an industrial base capable of supporting implementation or production of new systems?
- **Fiscal/Policy Issues** – What are the relevant issues at play and how will TSA address them?

Biometrics could potentially improve the traveler experience and open the door to innovative models of public-private cooperation between TSA and aviation industry stakeholders. That said, biometric solutions raise unique issues about privacy and accuracy that are addressed later in this report.

Operational Impacts – From an operational perspective, the introduction of biometrics to the TSA checkpoint will most directly affect the TDC position. This position is staffed by a TSO who gathers boarding passes and identity credentials from each traveler in the queue to quickly perform a series of screening steps (see *Figure 5*).

The planned use of CAT will help automate *steps 1, 3, and 5*. The automation of these tasks will increase TSA's confidence in the validity of credentials used to travel and the accuracy of the biographic data used to conduct Secure Flight vetting. CAT will also mitigate the threat of altered and counterfeit IDs, reduce the need for boarding passes at the checkpoint for many travelers (eliminate *step 4*), and automatically look up a traveler's vetting status in near-real time from Secure Flight's vetting engine.

The use of biometrics (for example, facial comparison) will also largely automate *step 2* by increasing assurance of identity beyond what is currently possible in a manual, human-based

[16] Strategic Five-Year Technology Investment Plan for Aviation Security: 2015 Report to Congress.

operation.[17] Specifically, biometrics will help mitigate threats posed by impostors using valid credentials for fraudulent purposes at the checkpoint (see subsection on *security impacts* for more detail).

For *step 6*, further integration of access control solutions with credential authentication and biometric technologies will help more fully automate the TDC process.

The development of this biometrically enabled solution will allow TSA to better secure access to the airport sterile environment and evaluate how to potentially reinvest valuable officer resources to other screening tasks. The automation of TDC functions will create a need for a 'TSO resolution' *step 7* in the event of system issues (for example, biometric match error, and alarm resolution).[18] In the future, TSOs will oversee biometric operations at the TDC to help travelers use the technology and address issues as they happen. TSOs will continue to provide important security safeguards, including directing travelers to the correct screening lane based on the travelers vetting status.

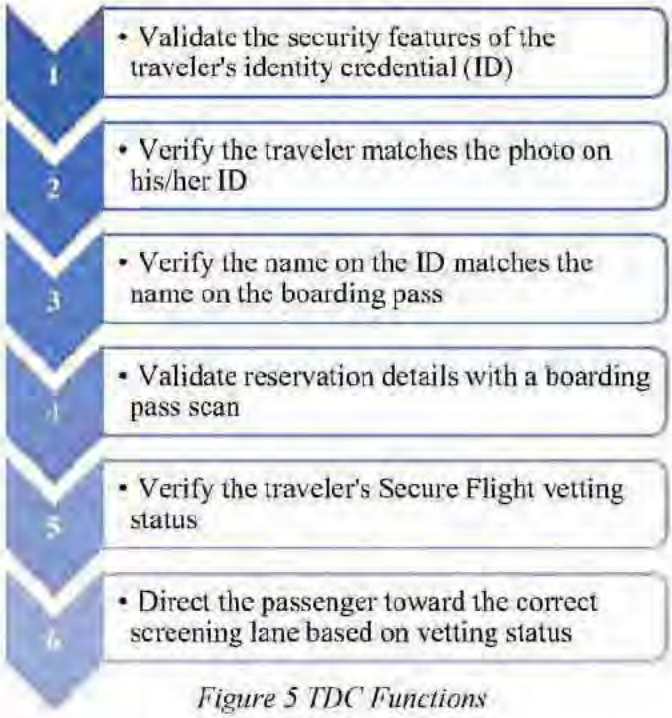


Figure 5 TDC Functions

Given the diversity of airports across the United States and their unique layouts, the operational placement and use of a fully integrated biometric solution will vary from facility to facility. For example, the use of an automated, biometric solution at a relatively small checkpoint may result in faster TDC processing times. However, the throughput of the checkpoint may be largely unaffected because a faster TDC process would merely shift traveler volume from the queue into the screening lane itself. A screening lane can only operate as fast as its slowest piece of transportation security equipment. This result underscores the need for continued investment across the entire checkpoint security enterprise.

On the other hand, at larger checkpoints with more lanes the operational efficiencies of an automated, biometric TDC may be greater. This would especially be true if the ratio of biometrically enabled TDCs to screening lanes was higher than the ratio of manual or CAT TDCs to screening lanes, thus freeing up TSO resources that could be used elsewhere. TSA will continue to explore this area as it tests checkpoint biometric solutions.

[17] Except for a relatively small number of "super-recognizers," human beings are generally outperformed by facial comparison technologies, especially when presented with the faces of persons not familiar to them such as the thousands of travelers a TSO greets and processes each day. See: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0150036

[18] Per initial modeling conducted by the Homeland Security Systems Engineering & Development Institute (HSSEDI), keeping match error rates low through the use of reliable and accurate biometric systems and ensuring the use of swift error resolution procedures will be key to maintaining and improving checkpoint throughput.

In summary, the operational efficiencies TSA could gain from integrated biometric solutions may be different depending on airport facility layouts, sizes, checkpoint lane counts, and traveler volumes. New procedures and robust workforce training will be required to maximize the operational benefits of biometric solutions.

Security Impacts – TSA uses a multi-layered, risk-based approach to securing the Nation's transportation systems. Today, during the airline reservation process, the traveler provides their first name, last name, date of birth, gender, and, if applicable, known traveler number, or DHS redress number. The airline transmits this information to TSA's Secure Flight system for vetting against intelligence-driven watch lists. The result of this vetting process, known as the Boarding Pass Print Result, is sent to the airline and encoded on the traveler's printed or mobile boarding pass.

When the traveler arrives at the checkpoint, the TSO must quickly perform a series of complex tasks (see *Figure 6*) using a variety of tools. TSOs assess whether the presented ID credential is authentic, determine whether the traveler matches the picture on their ID credential, decide whether the name on the boarding pass matches the name on their ID credential, distinguish between various forms of ID (state driver's licenses, passports, and government IDs, among others), validate the boarding pass, and direct the traveler to the appropriate level of screening based on their Boarding Pass Print Result.

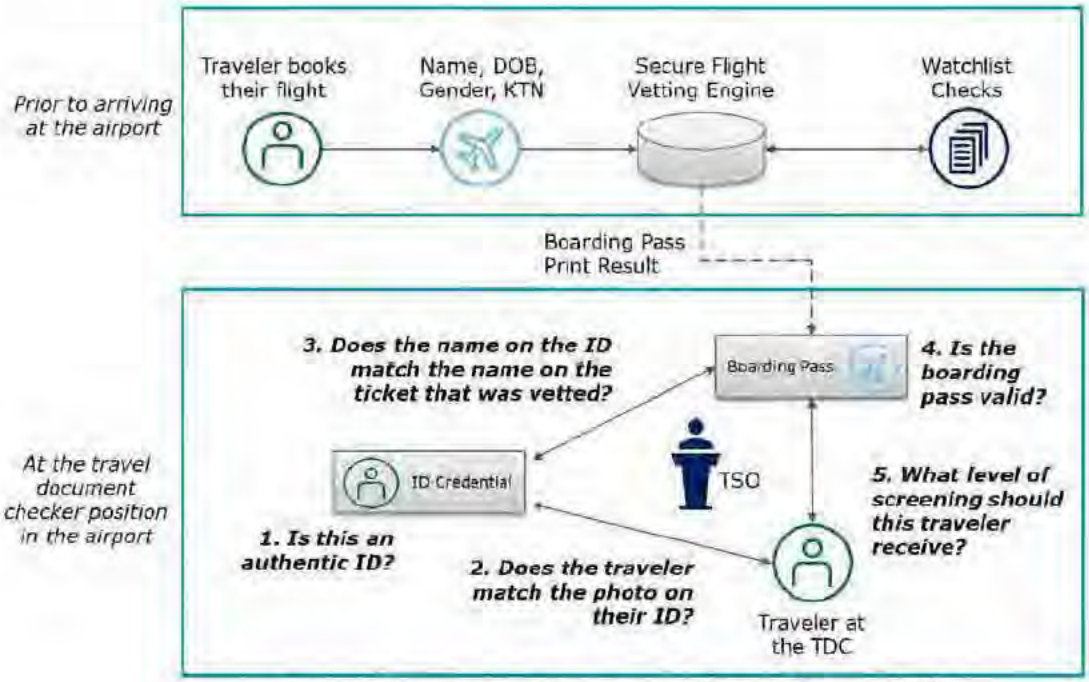


Figure 6 Systems and Operational View of Current TDC

Using an integrated, biometric TDC solution (see *Figure 7*), TSA can automate certain repetitive tasks and enable the system to verify the traveler's identity using the facial image and biographic information encoded on the ID or through the use of previously enrolled biometric and biographic data (for example, Trusted Traveler information). This technology will help

eliminate human errors and biases in face matching, lower TSA's reliance on the boarding pass, and enable a near-real time connection to TSA vetting systems for up-to-date results.

This model shifts the burden of the security decision onto the system while reducing TSO burden of repetitive, manual face comparisons and name matching between travel documents. Automating this process will enable TSOs to focus on the operation of the systems and intervene as needed to resolve problems or process travelers who cannot or do not wish to use the biometric system.[19]

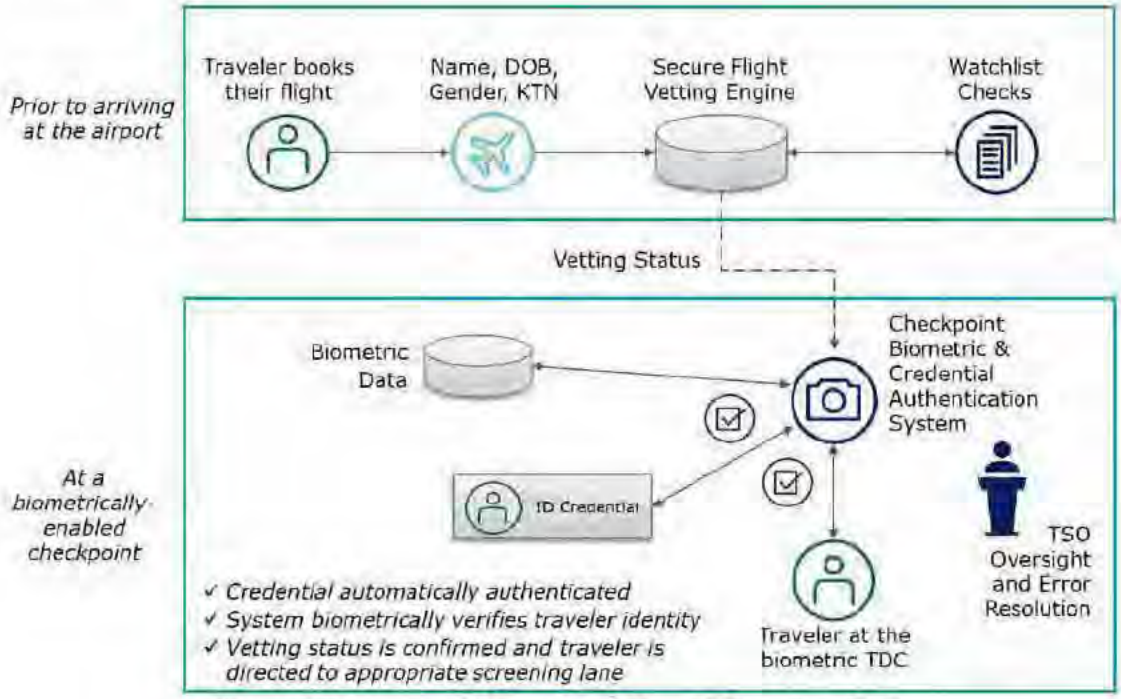


Figure 7 Systems and Operational View of Biometric TDC

Applying a biometric TDC to TSA Pre✓® and standard lanes would measurably increase security effectiveness and deter adversaries, or force a shift in their tactics. For example, individuals hoping to avoid detection using a fake ID or impostors using an authentic, stolen ID would be prevented from gaining access to the sterile area of the airport. In addition, integrated biometric solutions will help ensure individuals receive the correct level of screening based on their vetting status; making it more difficult for adversaries to avoid higher levels of screening by falsifying their identity.

While the rate of adversaries attempting to gain access to the checkpoint is difficult to determine, TSA can look to intelligence estimates and the experience of other organizations that use similar biometric solutions. CBP, for example, has used biometric facial comparison technology to identify more than 130 impostors trying to gain entry through air and pedestrian environments. Integrating biometrics into the checkpoint will enable TSA to further strengthen its security

[19] For example, minors under age 16 without state-issued driver's licenses would still be processed using traditional boarding pass scans. Travelers who opt out to a biometric experience will also require TSO assistance to proceed into the screening lane.

baseline, more effectively deter and detect bad actors, and better measure performance of security measures against adversaries trying to gain access to the airport sterile environment.

IV. Potential Effects on Privacy and Mitigation Methods

As they evaluate biometric technologies, CBP and TSA are committed to protecting travelers' information and privacy. In accordance with Office of Management and Budget (OMB) Directives 03-22, *OMB Guidance for Implementing the Privacy Provisions of the E-Government Act of 2002*,[20] any use of personally identifiable information (PII), including use of facial comparison technology, requires a thorough analysis of its privacy impact through a Privacy Impact Assessment (PIA). Both CBP and TSA have submitted and published a number of PIAs on related pilots and programs to the DHS Privacy Office for adjudication and publication. DHS PIAs use the Fair Information Practice Principles (FIPPs) to assess and mitigate any impact on an individual's privacy. These principles are rooted in the *Privacy Act of 1974* and govern the use of PII.[21] The FIPPs help guide CBP and TSA as they seek to protect privacy and improve the traveler experience while gaining the operational and security benefits of biometrics technology.

TSA and CBP collaborate regularly with their respective Privacy Offices and DHS's Privacy Office. On September 11, 2017, the DHS Privacy Office commissioned the DHS Data Privacy and Integrity Advisory Committee (DPIAC) to advise the Department on best practices for the use of facial comparison technology. CBP briefed the DPIAC in September 2017, May 2018, and July 2018, when CBP provided a tour of biometric entry and exit operations at Orlando International Airport, and again in December 2018. The DPIAC published its report 2019-01 of the *DHS DPIAC: Privacy Recommendations in Connection with the Use of Facial Recognition Technology*,[22] on February 26, 2019. CBP has implemented, and is working to implement many of the DPIAC recommendations. CBP also met with privacy and civil liberties advocates twice since 2017 to discuss the biometric entry-exit program, including technical demonstrations, the future biometric vision, privacy and security protections, notice to the public, retention policies, and alternative screening procedures. Each meeting included a lengthy question and answer session. Similarly, in August 2019, TSA held a privacy roundtable with privacy and civil liberties groups to discuss its exploration of biometrics technology.

It also noted that "it is critical for the success of the Biometric Exit Program and/or other biometric programs that data intended to be used only for screening purposes is not further transferred, shared, or used for other purposes, including without limitation private-sector purposes (e.g. marketing) or other government purposes (e.g. law enforcement or intelligence purposes)." The DPIAC's detailed recommendations will be particularly helpful as TSA and CBP consider the privacy impacts of biometrics technology.[23] For instance, TSA and CBP consider issues such as timely and transparent notice; alternative screening processes; data

---

[20] https://www.whitehouse.gov/sites/whitehouse.gov/files/omb/memoranda/2003/m03_22.pdf

[21] https://www.dhs.gov/sites/default/files/publications/privacy_policyguide_2008-01_0.pdf ; see: DHS Privacy Policy Directive 140-06, The Fair Information Practice Principles: Framework for Privacy Policy at the Department of Homeland Security, available at: https://www.dhs.gov/publication/privacy-policy-guidance-memorandum-2008-01-fair-information-practice-principles.

[22] https://www.dhs.gov/sites/default/files/publications/Report%202019-01_Use%20of%20Facial%20Recognition%20Technology_02%2026%202019.pdf

[23] https://www.dhs.gov/publication/dpiac-recommendations-report-2019-01

minimization; reliability testing; data quality and integrity; accuracy; and accountability and auditability of facial comparison technology. Both agencies will address the FIPPs in their biometric technology efforts and associated privacy compliance documentation, to ensure the protection of personal information at all stages of the information lifecycle.

A. CBP Approach to Mitigating Privacy Impacts

CBP is fully committed to protecting privacy and ensuring the integrity of its facial comparison matching service. In developing and expanding the use of the TVS, CBP is implementing a privacy by design[24] approach to ensure that privacy protections are embedded into its use of facial comparison technology. CBP employs four primary safeguards to secure the data, including secure storage, brief retention periods, irreversible biometric templates, and strong encryption during data storage and transfer.

CBP complies with the requirements of the *Privacy Act of 1974*, as amended, the *E-Government Act of 2002*, and Departmental and government-wide policies governing the collection, use, and maintenance of PII. As with other biometric collections, facial comparison poses privacy risks that are mostly mitigated. CBP's phased deployment has illustrated the success of the use of facial comparison technology in a variety of operational scenarios, meeting CBP's business requirements while requiring minimal infrastructure investments and space redesign as well as minimal impacts upon travelers. Additionally, the approach has allowed CBP to ensure that biometrics are collected, maintained, and used consistent with privacy law and best practices. CBP analyzes the privacy impact of its collection, use, dissemination, storage, and sharing of PII through the lens of the DHS FIPPs as described above.[25] The eight FIPPs principles, rooted in the tenets of the *Privacy Act*, have served as the framework for privacy policy at DHS for more than a decade.

When a traveler presents himself or herself for entry, exit, or at a TSA security checkpoint, the traveler will encounter a camera connected to the biometric cloud matching service via a secure, encrypted connection. The biometric matching service converts the live photos into secure templates and matches them against templates of gallery images, which travelers have already provided to the U.S. Government for travel purposes. The templates cannot be reverse engineered to reconstruct the photo. Finally, CBP does not share any photos with travel stakeholders, but rather provides the travelers and partner airlines with the results of the biometric match (match or no-match) through a response message data value. In implementing biometric matching through the TVS, CBP is simply replacing the existing document checks with a biometric facial comparison process, which will greatly reduce the need for travelers to continually present identity documentation at multiple stops along their journey.

---

[24] *See* DHS/CBP/PIA-056 Traveler Verification Service (November 14, 2018), *available at* www.dhs.gov/privacy.

[25] Additionally, the DHS Privacy Office conducted a Privacy Compliance Review of CBP's Southwest Border Pedestrian Exit Field Test that resulted in 10 recommendations to improve the privacy of individuals' biometric information, including facial and iris images. Available at: https://www.dhs.gov/sites/default/files/publications/SW%20Border%20PCR%20report%20FINAL%2020161230.pdf.

CBP provides transparency and general notification to the public through program information, such as frequently asked questions, available on the CBP website at www.cbp.gov/biometrics, and the TVS Privacy Impact Assessments (PIAs) and System of Records Notices (SORNs) published at www.dhs.gov/privacy.[26] The PIAs and SORNs for the TVS and its predecessor projects explain all aspects of CBP's biometric entry-exit programs, including policies and procedures for the collection, storage, analysis, use, dissemination, retention, and deletion of data. These PIAs and SORNs describe in detail CBP's approach to ensuring both the processes and systems integrate controls to mitigate privacy risks.

Following the DHS FIPP of transparency, CBP works closely with airline, airport, and cruise line partners to incorporate notifications and processes into their current business models, such as gate announcements or visible signage that explain the facial matching process and alternative inspection procedures. If processes or procedures change, CBP will update these channels to ensure all outreach material is current and clear for the traveling public. Because facial comparison can be performed quickly with minimal instruction and with a high degree of accuracy, the approach implemented represents the best operational means of verifying the identity of the traveler, and the data is collected in a manner perceived as less invasive to the traveler. Facial comparison requires no actual physical contact to collect the biometric data, and there is less risk of the loss of traveler documents that contain the date of birth and other sensitive PII.

Prior to admission into the United States, CBP must ensure that each traveler is a U.S. citizen, lawful permanent resident, or is otherwise an alien eligible for admission, and that the traveler is not attempting to import any merchandise in violation of U.S. laws. Similarly, CBP officers may inspect travelers departing the United States in order to create exit records and as required for law enforcement operations. The website www.cbp.gov/biometrics, along with signage, verbal announcements, tear sheets, and the TVS PIA contain details on the current biometric entry-exit process, including alternative procedures. In accordance with the FIPP of individual participation, a U.S. citizen and otherwise exempt aliens[27] may notify either the CBP officer or the airline boarding agent that he or she would like to opt out at the time of boarding and, instead, present credentials for a manual identity verification using their travel document. In adherence to the FIPP of purpose specification, CBP stipulates that PII collected through the biometric entry-exit program be used primarily to verify that the traveler attempting to board the flight or cross the border is, in fact, the rightful bearer of the travel document he or she is presenting.

Throughout its history, CBP has maintained productive partnerships with the travel industry, where the flow of PII between entities is well-defined in law and regulations. In line with the FIPPs, data minimization and use limitation, CBP has taken noteworthy steps to protect privacy,

---

[26] *See* DHS/CBP/PIA-056 Traveler Verification Service (November 14, 2018), *available at* www.dhs.gov/privacy. www.dhs.gov/privacy. The SORNs associated with CBP's Traveler Verification Service are: DHS/CBP-007 Border Crossing Information, DHS/CBP-021 Arrival and Departure Information System, DHS/CBP-006 Automated Targeting System, DHS/CBP-011 U.S. Customs and Border Protection TECS.

[27] Certain aliens are exempt from any requirement to provide biometrics upon entry into the United States pursuant to 8 CFR 235.1(f)(ii), or upon exit from the United States pursuant to 8 CFR 215.8.

such as its commitment to prohibit the sharing the photos captured and matched through the TVS with CBP industry partners. Only the results of the "match/no-match" determination are shared. In fact, CBP's business requirements for partner airline and technology vendors do not permit the retention of photos for commercial purposes, following transmittal to CBP for matching. In addition, TVS only utilizes the irreversible biometric templates of source and newly-captured photos for matching and uses a unique identifier[28] to disassociate the biographic information associated with the new facial images.

While CBP does not retain U.S. citizens' images submitted as part of the traveler verification process,[29] photos of foreign nationals (and those dual national U.S. citizens traveling on foreign documentation) are retained for up to 14 days in secure systems to confirm traveler's identities, evaluate the technologies, and to assure continued accuracy of the algorithms. In addition, CBP transmits facial images for in-scope travelers to the DHS Automated Biometric Identification System (IDENT) for retention as the traveler's biometric encounter with CBP. For U.S. citizens, only a confirmation of the border crossing and the associated biographic information is retained.

In line with the FIPP of accountability and auditing, the CBP Privacy Office will conduct a CBP Privacy Evaluation by the end of calendar year 2019 to ensure that all parties, including airlines, airport authorities, and cloud providers, are in compliance with the privacy protections described in the TVS PIA. The results of the evaluation will be shared with the DHS Privacy Office.

B. TSA Approach to Mitigating Privacy Impacts

TSA is committed to protecting traveler privacy and ensuring the traveling public's trust as it modernizes identity verification through its exploration of biometric technology. TSA will comply with DHS privacy policy throughout each phase of TSA's biometric solution development – from initial design to implementation. Solutions will be designed to secure data as it is collected, stored, and transmitted between systems to protect both travelers and system integrity.

TSA recognizes that biometric technologies, particularly facial comparison, pose unique privacy concerns with respect to privacy and passengers' civil rights and civil liberties. There is significant risk to individuals should the facial images be compromised or used for purposes beyond those specified for its collection. There is also a risk to both individuals and transportation security in the event that the biometric technology is not sufficiently accurate. To mitigate these risks, TSA will evaluate issues such as:

- Robust notice of facial comparison deployment for traveler screening;
- Meaningful choice of screening choices for the traveler;
- Robust cyber-security measures to protect traveler data from collection through transmission to receipt; and

[28] The unique identifier is generated by either the travel agent, travel website hosting service, or the airline at the time of the reservation. It is comprised of a sequential number (which is only valid for the particular airline and the specific flight), plus the record locator, a six-digit code used to access additional information about the traveler.
[29] Photos of U.S. citizens are held in secure CBP systems for no more than 12 hours after identity verification, in case of an extended system outage.

- Limitation on the use of the facial images to those necessary for transportation security, consistent with the Privacy Act.

TSA will integrate privacy protections as it continues to partner with CBP on biometrics for international travelers, implement new biometric capabilities for TSA Pre✓® travelers, and explore the expansion of biometric collections, such as use of facial images, to additional domestic travelers. TSA will also adhere to DHS privacy policy in its adoption of new biometric-based vetting solutions for non- traveler groups such as aviation workers, law enforcement officers, and crew members.

Privacy and Facial Comparison for International Travelers

Since beginning to explore the use of facial comparison technology for traveler identity verification, TSA has taken steps to provide notice to the public about its efforts, assess privacy risks, and establish strategies to protect traveler privacy. The challenge of traveler identity verification through facial comparison for TSA is significant for international and domestic travelers for whom established, government-owned facial image databases do not exist. In comparison of this challenge, TSA engaged in several pilots involving international travelers. For instance, in January 2018, a PIA was published for a three-week proof of concept at LAX using passports.[30] The proof of concept was to validate the use of facial comparison technology to automate identity verification during the TDC process.

TSA compared the facial images of aviation passengers with e-Passports on outward-bound international flights and who voluntarily entered the screening checkpoint through automated electronic security gates or "e-Gate." The e-Gate device captured an image of the passenger's face and compared it to the biometric image in the passenger's e-Passport. The e-Gate attempted to replicate the function of the TDC and authenticated the passenger's e-Passport and boarding pass.

Additionally, privacy protections have been embedded in TSA's partnership with CBP on facial recognition pilots. These pilots took place in international terminals at a select number of airports to limit biometric collection to travelers on international flights. They enabled both agencies to collect data, refine solutions, and exchange information on the operational performance of facial comparison technology. Privacy compliance documents for each of these pilots have analyzed the potential effects on privacy and identified methods to lessen privacy risks.

In the first phase of the partnership, which took place in October 2017, TSA and CBP conducted an operational pilot at JFK to test the ability of CBP's TVS to match traveler identities against galleries of pre-staged photos at the TSA checkpoint. The second phase consisted of a pilot at LAX, which tested the TVS with a larger gallery and enhanced automation, from August to October 2018. Additionally, in November 2018, TSA, CBP, and Delta Air Lines began testing

[30] https://www.dhs.gov/sites/default/files/publications/privacy-pia-tsa-046-tdcautomationusingfacialrecognition-january2018.pdf

biometrics for identity verification at Terminal F at ATL. CBP published PIAs on each phase of the TVS pilot.[31]

Privacy and Biometric Solutions for TSA Pre✓® Travelers and Additional Populations

TSA is using biometrics to modernize the trusted traveler experience for TSA Pre✓® travelers. For first-time enrollees or for individuals renewing their membership in that program in person, TSA started capturing facial images to help verify identity. At this time, a facial image is not required for individuals who renew their TSA Pre✓® Application Program membership online.

TSA will also evaluate the possibility of allowing additional trusted travelers to access the TSA Pre✓® lanes (for example, members of the Department of Defense), as well as the general flying public to opt in to biometric screening and verification. However, before making biometric solutions available to these travelers, TSA will work with OBIM and DHS oversight offices, including the DHS Privacy Office and the Office of Civil Rights and Civil Liberties to evaluate options, conduct pilots, and to ensure compliance with privacy law and policy and civil rights and civil liberties requirements.

In any biometric technology solutions involving the collection, maintenance, use, or dissemination of PII, TSA will be transparent by notifying the public and explaining the steps the agency is taking to safeguard individuals' information. In its development of biometric technologies for additional populations, TSA will comply with Section 208 of the *E-Government Act of 2002*, Section 222 of the *Homeland Security Act of 2002*, and DHS' privacy compliance process. As such, TSA will conduct appropriate privacy threshold analyses, PIAs, and system of records notices when considering the use of biometric solutions with potential privacy impacts. TSA will also comply with applicable TSA, DHS, and Office of Management and Budget policies and authorities governing the handling of PII.

TSA will comply with law and DHS privacy policy related to the use of facial comparison technology for identity verification such as notice to travelers, opt-in policies, consent protocols, specific use limitations, and alternative screening procedures for travelers that do not wish to provide their facial image for identity verification purposes. Consistent with information technology security policies and authorities, TSA will also develop biometric solutions that meet cybersecurity protocols so that data is protected at all stages of the information lifecycle. Additionally, public education and outreach will be conducted to provide awareness of the agency's future biometrics efforts.

Stakeholder Engagement on Privacy

As part of its commitment to protecting traveler privacy in the use of biometrics technology, TSA will continue to:

- Engage with non-governmental stakeholders to obtain input on best practices for protecting privacy;

[31] https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp030-tvs-november2018_2.pdf

- Coordinate with internal TSA offices, DHS Headquarters, oversight entities and interagency partners to track compliance with privacy authorities and requirements, develop privacy-protective policies, and appropriately manage identified privacy risks;
- Seek information and feedback from industry, privacy groups, academic institutions, and other privacy professionals and research organizations as it considers the expansion of biometrics solutions to increase to increase security and streamline the passenger experience; and
- Share information with key stakeholders on its development of biometrics technology capabilities.

V. TSA Methods to Analyze and Address Matching Performance Errors

While TSA has been using fingerprints since 2004 to conduct security threat assessments—including checks on an applicant's criminal history, potential ties to terrorism, and citizenship—the use of biometrics to verify traveler identity has begun only recently. As of September 2018, the TSA Pre✓® Application Program has transitioned from single factor enrollment (fingerprints) to multi-modal biometric (fingerprints and facial image) enrollment. See Section II.C for an overview of TSA's biometric testing efforts to date.

TSA's exploration of the use of biometric data, namely facial images, as a means of facilitating secure travel is coming at an ideal time in the biometric industry. According to the most recent National Institute of Standards and Technology (NIST) Face Comparison Vendor Test, facial verification algorithms have become significantly more accurate over the 2013-2018 period. The NIST Interagency Report 8238 states:

> While the industry gains are broad—at least 28 developers' algorithms now outperform the most accurate algorithm from late 2013—there remains a wide range of capabilities. With good quality portrait photos, the most accurate algorithms will find matching entries, when present, in galleries containing 12 million individuals, with error rates below 0.2 percent. The remaining errors are in large part attributable to long-run ageing and injury.[32]

According to NIST, these gains have been largely facilitated by a revolution in algorithm development, fueled by new machine learning approaches. Whereas algorithms of five years ago may have struggled to match images that differed in pose, illumination, and facial expression, today's algorithms are increasingly tolerant of such variations in image quality. Indeed, improvements to the technology are being made in months rather than years.

Despite these gains, however, facial comparison systems are shadowed by reports of variable performance across demographic characteristics; namely race, age, and gender.[33] Much of the discussion has focused on the ability of various facial comparison algorithms to accurately process younger subjects, female subjects, and subjects with darker skin. Indeed, a 2019 article published by DHS S&T and informed by testing conducted in 2018 at S&Ts Maryland Test Facility (MdTF) found evidence of some variation in facial comparison performance along

---

[32] *See* NIST Interagency Report 8238, available at https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8238.pdf

[33] TSA does not collect data on traveler race, ethnicity, or skin color for the purposes of making security and screening decisions; however, TSA may collect such data – in accordance with standard test protocols – during operational testing to ensure systems perform accurately under operational conditions.

demographic lines.[34] Interestingly, however, this performance variation was not solely based on the face comparison algorithm, but resulted from an interaction between the matching algorithm and image acquisition hardware.[35]

S&T tested 11 commercially available facial image acquisition systems using a demographically diverse population of 363 volunteer subjects.[36] The live images ("probes") gathered by each system were matched against historical and same-day enrollment images using a leading commercial algorithm for facial comparison. The variation in facial matching performance across different image acquisition systems versus when images are matched against a single, industry-leading algorithm suggests the hardware used to capture the probe image significantly affects matching accuracy.

As a result, using a superior biometric acquisition system capable of capturing higher quality facial images may significantly reduce or even eliminate performance differences along demographic lines. Logically, it follows that a lower quality acquisition system can increase the likelihood of performance variation along demographic characteristics. This key finding will influence TSA's testing, development, and potential procurement of checkpoint facial comparison capabilities.

S&T's recent round of testing, which took place in May and June of 2019, examined the performance of an additional 10 commercial facial acquisition systems against eight commercial facial comparison algorithms. When completed, the findings of this research may give more insight into the best mix of hardware and software assets needed to ensure the accuracy of checkpoint biometric systems for the diverse traveling public. Additionally, TSA will join interagency efforts to ensure DHS biometric systems (for example, CBP TVS, OBIM IDENT/HART) are designed to enhance performance across missions, use cases, and demographics.

Other variables encountered in the airport environment can affect system performance as well. Inconsistent lighting (for example, sun glare through large windows), changes in a traveler's facial structure relative to previous encounter images, and eyewear or other face/head wear can affect system performance. This underscores the need for TSA to continue to invest time and energy into ensuring its checkpoint biometric solutions, as well as other transportation security equipment, are designed with the human-system interface in mind. Intuitive, highly usable solutions combined with the right TSO procedures, biometric acquisition hardware, and matching software will help ensure TSA's mission requirements are met while also ensuring a streamlined security experience for air travelers.

---

[34] Note: S&T found "relative skin reflectance" to be a better indicator of system performance than U.S. Census categories (e.g. "White", "Black", and "Other").

[35] *See* Demographic Effects in Facial Recognition and Their Dependence on Image Acquisition: An Evaluation of Eleven Commercial Systems, available at https://ieeexplore.ieee.org/document/8636231

[36] For more information on the S&T test facility, protocols, and results *see* https://mdtf.org/rally/.

Given the wide diversity of the millions of travelers moving through airport checkpoints daily, accuracy in biometric solutions is a key issue. Therefore, TSA is grounding its exploration of biometric solutions in rigorous scientific study and analysis to ensure the full benefits of biometrics technology are realized. Efforts will continue to ensure biometric checkpoint solutions are designed to mitigate performance variations based on demographic characteristics.

VI. Performance Assessments and Audits of the Biometric Entry-Exit Program

CBP has a robust process for performing operational assessments of CBP's biometric system performance, including evaluating the performance of biometric transactions performed during arrival and departure operations in the air environment, as well as continual performance assessments of technical demonstrations to determine the best concept of operations in other operational environments such as land and sea. Third parties such as S&T and NIST are also engaged to both evaluate CBP operational data and make recommendations for performance enhancements that include biometric capture and matching.

A. Performance Assessments

Biometric Performance Analysis of CBP Systems

CBP has a rigorous process in place to review data and metrics associated with biometric exit facial comparison matching performance. Biometric Air Exit Key Performance Parameters (KPPs) mandate that the system's True Acceptance Rate (TAR) [37] must equal or exceed 97 percent of all in-scope travelers and that the system's False Acceptance Rate (FAR) [38] must not exceed 0.1 percent of all in-scope travelers.

To establish whether or not TVS is fulfilling these KPPs, CBP is systematically analyzing actual flight data for the airlines using Biometric Air Exit. The evaluation team periodically prepares summary reports that present the actual performance of TVS against its KPPs in production.

On a weekly basis, operational performance analysis of CBP biometric operations are conducted, including Air Entry, Air Exit, Preclearance, and Pedestrian Entry (currently in technical demonstration). CBP's performance analysis is focused on the ability to match travelers captured by airports and airlines against the gallery created using the Advanced Passenger Information System (APIS) manifest. Beginning in November 2018, CBP moved to a sampling method to assess the technical match rate for biometric exit and aspects of the CBP-TSA pilot. The technical match rate is a measure of how well the matching algorithm is performing. It includes U.S. citizens who choose not to opt out and individuals who are in-scope (pursuant to 8 CFR 215 and 235) that had a photo in the CBP gallery from existing DHS sources and were

[37] The **TAR** is the number of valid matches divided by the sum of the valid matches and the invalid non-matches. Note that this sum (valid matches plus invalid non-matches) equals the number of matches that should have occurred, and includes all the travelers with a valid encounter photo and at least one valid gallery photo. This definition of the TAR is generally equivalent to the Technical Match Rate (TMR), as defined by CBP's Office of Field Operations.

[38] The **FAR** is the number of invalid matches divided by the sum of the invalid matches and the valid non-matches. Note that this sum (invalid matches plus valid non-matches) equals the number of matches that should NOT have occurred, and includes all the travelers with a valid encounter photo for whom there is no valid gallery photo.

successfully captured by the camera. The following table shows recent match results for each production mode of operation, as a per day average[39].

Modality	Number of Locations	Flight Count	Number of Travelers	Technical Match Rate
Air Entry	11	446	34,716	99.2%
Air Exit	16	92	11,545	97.6%
Air Preclearance	4	45	6,559	99.4%
Pedestrian Entry	4		12,591	97.7%

The estimated false positive rate based on the internal CBP analysis is .0103 percent, which is within the established KPP target of less than .1 percent. As a comparison, a 2014 study "*Passport Officers' Errors in Face Matching*[40]," found that even individuals with specialist experience and training in the task, passport-issuing officers had a 14 percent false positive rate when conducting a person-to-photo comparison test.

Ensuring Biometric Technologies Do Not Unduly Burden Travelers

CBP continuously monitors the biometric matching service and conducts a variety of statistical tests to bolster performance thresholds and minimize any possible bias impact on travelers of certain race, gender, or nationality.

CBP requires that all airlines submit traveler information to the Advanced Passenger Information System (APIS). Among the data submitted is gender, date of birth, travel document type, number, and nationality. Using a subset of this data, CBP conducted extensive statistical analysis including *chi squared*[41] independence tests to determine whether traveler demographics (age, gender, and nationality) affect facial comparison match rates. As CBP does not collect race/ethnicity nor is this information included in the APIS manifest, citizenship is used as a proxy to conduct its analysis.

CBP's analysis found a negligible effect in regards to biometric matching based on citizenship[42], age, or gender while achieving a technical match rate (TMR) in the high 90 percentile.[43] As of December 2018, TMR continues to be at a steady state, above 98 percent. Significant improvements to the algorithm and exit operations continue to be made, which has led to a substantial reduction in the initial gaps in matching for ages and genders. On average, U.S.

---

[39] Data shown indicates the averages per day for the period March 20, 2019 to April 2, 2019.
[40] https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0103510
[41] A *chi-squared* ($\chi 2$) independence test is a statistical test applied to sets of categorical data to evaluate how likely it is that any observed difference between the sets arose by chance. The tests can be used to determine whether there is a significant difference between the expected frequencies and the observed frequencies in one or more categories.
[42] While CBP uses citizenship as a general proxy because it does not collect race/ethnicity data, it takes into account in its analysis that this is clearly a more effective proxy when looking at homogenous countries than diverse ones.
[43] Based on June 2017 – November 2018 CBP Air Exit data from biometric exit locations: JFK, MIA, IAH, HOU, ORD, SEA, SFO, LAS, DTW, LAX, IAD, MCO, ATL, BOS, and FLL.

citizens typically match at a lower rate as they have fewer and older photos which decreases matching rates. Travelers between ages 26 and 65 match slightly better than "young" (ages 14 to 25) travelers (0.3 percent) and "old" (ages 66 to 79) travelers (0.1 percent), compared to 2.8 percent and 8 percent, respectively, during the initial pilot period. Similarly, women match slightly better than men (0.2 percent), compared to matching worse initially (1.7 percent) during the pilot period. Much of the bias seen in the initial period also relates to much lower flight volume during that timeframe.

As NIST concluded during its 2018 Face Comparison Vendor Test[44], there have been massive improvements in the accuracy of face comparison algorithms in the last five years (2013-2018). The performance of CBP's TVS continues to improve over time due to technical, operational, and procedural advancements including threshold adjustments and testing multiple vendors. CBP has enhanced the photo selection process used to build the galleries, which reduces the number of travelers with no photos and improves the accuracy of the system.[45] Additionally, CBP has enhanced the manner in which the galleries are populated, ensuring that the information included in the flight manifest is used to its maximum potential to include more higher-quality photographs.[46] CBP has also issued various update to the matching algorithms, which increase the algorithm's ability to create biometric templates from non-frontal images taken during the U.S. entry or exit process.

There have also been software changes to the cameras to allow travelers posing for the photos to receive visual feedback. Furthermore, as CBP continues and expands its usage of TVS, personnel using the technology become more aware of the optimal camera positions to ensure better images and increase the traveler throughput. Some cameras are also now equipped with multiple lenses to capture images for various angles, which may increase photo quality depending on the height of the traveler.

Biometric Technology Impact on Travelers Overstaying Their Lawful Period of Admission

CBP has the ability to accurately report overstay numbers in the air and sea environments today. In FY2018, DHS calculated a total overstay rate of 1.22 percent, or 666,582, overstay events. In other words, 98.78 percent of in-scope nonimmigrant entries in FY2018 departed the United States on time and in accordance with the terms of their admission. Annual statistics on visa overstays are provided by DHS to Congress in the Annual Entry Exit Overstay Report.[47]

Adding biometric verification to an already robust biographic exit capability enables CBP to better detect travelers seeking to depart the country under a false identity, including aliens

---

[44] *See* NIST Interagency Report 8238, available at https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8238.pdf.
[45] A 2010 NIST evaluation of face comparison showed that considerable accuracy benefits accrue with retention and use of all historical images. See https://www.nist.gov/publications/report-evaluation-2d-still-image-face-recognition-algorithms.
[46] Additional information about CBP's gallery building process can be found in the DHS/CBP/PIA-056, Privacy Impact Assessment for the Traveler Verification Service, issued Nov. 14, 2018, available at https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp030-tvs-november2018_2.pdf.
[47] See FY 2018 Entry-Exit Overstay Report at: https://www.dhs.gov/sites/default/files/publications/19_0417_fy18-entry-and-exit-overstay-report.pdf

seeking to fraudulently use validly issued U.S. travel documents. The addition of biometrics has assisted CBP officers in detecting impostors attempting to gain entry to the United States. As part of the continued expansion of biometric exit capabilities, CBP will measure and report on the number of impostors detected by the biometric exit program.

Utilizing biometric technology, CBP has been able to biometrically confirm more than 14,000 travelers that overstayed their lawful period of admission on exit. As of April 2019, 130 impostors have been positively identified using the TVS system across air entry and pedestrian entry environments. All biometric encounters of in-scope foreign nationals are recorded in the enterprise biometrics system IDENT.

B. Audits Performed

DHS Office of Inspector General

The DHS Office of Inspector General (OIG) audit (OIG-18-80), *Review of CBP Biometric Exit Capability*[48], evaluated CBP's efforts to develop and implement a biometric exit capability and assess whether biometric data collected has improved DHS's ability to verify foreign visitor departures at U.S. airports. The final report was issued on September 24, 2018, and included four recommendations:

1) Develop an internal plan to institute enforcement mechanisms or back-up procedures to prevent airlines from bypassing biometric processing prior to flight boarding;
2) Take steps to coordinate with airport and airline stakeholders to increase bandwidth to meet the operational demands of biometric processing at the Nation's top airports;
3) Continue to refine the TVS algorithm to ensure the highest possible traveler match rate, with allowances for photo age and quality; and
4) Develop internal contingency plans for funding and staffing the program, in the event that airlines do not agree to partner with CBP in implementing the biometric capability nationwide.

The OIG conducted fieldwork from September 2017 to January 2018 and reviewed data from the earliest start of the technology demonstrations, which were never intended to be a final implementation model. However, regarding recommendation three and as addressed previously in this report, CBP continues to monitor and improve algorithm performance through incremental updates and improvements with system development and image quality requirements. CBP data analytic teams are evaluating any anomalies and providing feedback to development teams to improve entity resolution and refine matching performance

DHS Science and Technology (S&T) Directorate

In order to continually improve upon the quality of the images, DHS S&T is assisting CBP by testing the efficiency, effectiveness, user satisfaction, and equitability of biometric systems. This

[48] Available at: https://www.oig.dhs.gov/sites/default/files/assets/2018-09/OIG-18-80-Sep18.pdf.

includes performing independent scenario testing of state of the art commercial biometric systems at the MdTF as well as performing analyses using a sample of operational TVS images.

Starting in 2018, DHS S&T has performed independent biometric analyses using a sample of operational TVS probe and gallery facial images[49][50]. These analyses focused on answering specific questions regarding biometric performance. DHS S&T found that the algorithm used in TVS was superior in performance to all other algorithms tested.

Calculating standard biometric performance metrics in operational systems is challenging. DHS S&T developed a method for estimating the false positive identification rate (FPIR) using operational TVS system data. DHS S&T presented the new method, termed "Virtual Red Team" analysis, to CBP. DHS S&T used this method to estimate FPIR. DHS S&T concluded that FPIR for TVS varies by flight, such that some flight routes could have FPIR values 6-fold higher than others.

Based on these analyses, DHS S&T made specific recommendations to CBP including:

1. To ensure that only ticketed travelers are allowed to use TVS for boarding OR to increase match thresholds used for biometric exit; and
2. To carry out an exhaustive "Virtual Red Team" analysis to calculate the risk of false matches based on the demographics (age, country of origin, gender) of travelers on individual flights.

National Institute of Standards and Technology

CBP is also collaborating with NIST to perform an independent and comprehensive scientific analysis of CBP's operational face matching performance, including impacts due to traveler demographics and image quality. This independent study will help verify results and provide a more in-depth analysis on various factors. Upon analyzing a comprehensive set of data, NIST will provide objective recommendations regarding matching algorithms, optimal thresholds, and gallery creation, optimizing face matching performance for large-scale traveler ID at air, land, and sea entry and exit ports of entry. CBP will continue to actively monitor and refine the performance of this process and associated algorithms in order to make incremental improvements and minimize signs of bias, and ensure the high accuracy of facial matching for all travelers.

---

[49] DHS S&T Port of Entry- People Screening. February, 2018. Analysis of Data and Algorithms Related to the Traveler Verification System: Estimating Effects of Gallery Size and Traveler Demographics on False Positive Identification Rates.

[50] DHS S&T Biometric and Identity Technology Center. January, 2019. Analysis of Data and Algorithms Related to the Traveler Verification System: Estimating False Match Rate and False Positive Identification Rate.

VII. Conclusion

Biometric technologies have the potential to greatly enhance operational efficiencies and security for both CBP and TSA. CBP has made significant progress in implementing biometric solutions across air, land, and sea since receiving the biometric entry-exit mission in 2013. Following publication of the joint policy memorandum on CBP and TSA's partnership on the development and implementation of biometric technologies, particularly facial comparison, both agencies have worked together on a number of operational pilots. These volunteer-based pilots have allowed both agencies to test, evaluate, and continue to refine biometric technology solutions, while working to achieve a more streamlined traveler experience. CBP and TSA's efforts have been grounded in transparency and a commitment to traveler privacy. CBP and TSA will continue to work together and seek input from their stakeholders as they examine the impact of biometric technology and work to align with DHS initiatives, strategies, and capabilities on biometrics.

VIII. Appendices

Appendix A. DHS Fair Information Practice Principles

Transparency	DHS should be transparent and provide notice to the individual regarding its collection, use, dissemination, and maintenance of personally identifiable information (PII).
Individual Participation	DHS should involve the individual in the process of using PII, and to the extent practicable, seek individual consent for the collection, use, dissemination, and maintenance of PII.
Purpose Specification	DHS should specifically articulate the authority that permits the collection of PII and specifically articulate the purpose(s) for which the PII is intended to be used.
Data Minimization	DHS should only collect PII that is directly relevant and necessary to accomplish the specified purpose(s) and only retain PII for as long as is necessary to fulfill the specified purpose(s).
Use Limitation	DHS should use PII solely for the purpose(s) specified in the notice. Sharing PII outside the Department should be for a purpose compatible with the purpose for which the PII was collected.
Data Quality and Integrity	DHS should to the extent practicable, ensure that PII is accurate, relevant, timely, and complete.
Security	DHS should protect PII (in all media) through appropriate security safeguards against risks such as loss, unauthorized access or use, destruction, modification, or unintended or inappropriate disclosure.
Accountability and Auditing	DHS should be accountable for complying with these principles, providing training to all employees and contractors who use PII, and auditing the actual use of PII to demonstrate compliance with these principles and all applicable privacy protection requirements.

Appendix B. Acronyms

Acronym	Definition
APIS	Advance Passenger Information System
ATL	Hartsfield-Jackson Atlanta International Airport
CAT	Credential Authentication Technology
CBP	U.S. Customs and Border Protection
DCA	Ronald Reagan Washington National Airport
DHS	U.S. Department of Homeland Security
DPIAC	Data Privacy and Integrity Advisory Committee
FAR	False Acceptance Rate
FIPP	Fair Information Practice Principles
FOUO	For Official Use Only
FPIR	False Positive Identification Rate
FY	Fiscal Year
HART	Homeland Advanced Recognition Technology
HSSEDI	Homeland Security Systems Engineering & Development Institute
ID	Identification
IDENT	DHS Automated Biometric Identification System
JFK	John F. Kennedy International Airport
KTN	Known Traveler Number
KPP	Key Performance Parameters
LAX	Los Angeles International Airport
MdTF	S&Ts Maryland Test Facility
NIST	National Institute of Standards and Technology
OBIM	Office of Biometric Identity Management
OIG	DHS Office of Inspector General
OMB	Office of Management and Budget
PIA	Privacy Impact Assessment
PII	personally identifiable information
S&T	DHS Science and Technology Directorate
SORN	System of Records Notices
TAR	True Acceptance Rate
TDC	Travel Document Checker
TMR	Technical Match Rate
TSA	Transportation Security Administration
TSO	Transportation Security Officer
TVS	Traveler Verification Service

Office of Field Operations
180-Day Update to Final Report Recommendation for
U.S. Government Accountability Office (GAO) Final Report Entitled: *FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (GAO-20-568)

Recommendation	180-Day Update	Due Date
Recommendation 1: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's (BEEP) privacy notices contain complete and current information, including all of the locations where facial recognition is used and how travelers can request to opt out as appropriate. **Response: Concur.** CBP's OFO will collaborate with the CBP Office of Public Affairs to publish: 1) Biometric Entry-Exit privacy information; 2) locations where facial recognition is used; and 3) traveler opt-out procedures on CBP's public-facing website, as well as to review and update that information on a monthly basis. CBP's OFO will also ensure that information provided in response to inquiries via the CBP Call Center is also reviewed and updated monthly. **Estimated Completion Date (ECD):** December 31, 2020	CBP launched its updated biometrics website on September 1, 2020 (www.biometrics.cbp.gov). The purpose of the site is to deliver information to the public and other stakeholder groups. The site provides a user-friendly communication channel for promoting facial comparison technology and biometrics information in a dynamic and inviting manner. As a testament to CBP's commitment to privacy protections, outlined in the DHS Fair Information Practice Principles, the website includes the current locations using facial comparison technology as well as information on how to request alternative screening and copies of CBP's privacy signage on display. The information provided, including a link to CBP's Traveler Verification Service (TVS) Privacy Impact Assessment (PIA), is yet another tool in CBP's arsenal to ensure technology sustains and does not erode privacy protections. CBP OFO met with the CBP Call Center Teams and discussed the Centers' existing and needed biometric resources. On October 1, 2020, OFO conducted three TVS 101 briefings for both the Traveler Call Center and Information Center staff. OFO provided the Call Center with the Public Affairs Guidance, to include Frequently Asked Questions (FAQs). In addition, OFO provided the Information Center with a few FAQs to post on www.help.cbp.gov. CBP will continue to ensure that content is up to date on the CBP website, as required, and when updates are made we will provide new details to both the Information Center	December 31, 2020

Recommendation	180-Day Update	Due Date
	and Call Center. **Points of Contact / Telephone Numbers / E-mail Addresses:** (b)(6), (b)(7)(C) Acting Director Entry/Exit Policy and Planning Division Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations U.S. Customs and Border Protection (b)(6), (b)(7)(C) (b)(6), (b)(7)(C) Director, Biometric Exit-Entry Strategic Transformation Division PPAE Office of Field Operations U.S. Customs and Border Protection O: (b)(6), (b)(7)(C) M: (b)(6), (b)(7)(C)	
Recommendation 2: The Commissioner of CBP should ensure that the BEEP's privacy signage is consistently available at all locations where CBP is using facial recognition. **Response:** Concur. It is important to note that, unlike Federal Inspection Services areas, the airport departure areas are not managed by CBP personnel. However, CBP OFO will continue to work with its airlines/airport partners to ensure that privacy signage is available, on display, and reflective of current privacy messaging for travelers. For example, CBP provides	OFO is working with CBP's Enterprise Services, Office of Facilities and Asset Management, Administrative Services Program Management Office, Printing, Graphics and Distribution Branch, to ship out additional signs with the most recent/updated language, to the ports of entry (POE). OFO is also drafting policy guidance to the Field Offices/POEs requesting a signage auditor point of contact who would be responsible for ensuring signage (post deployment) is correct and on display. Every Simplified Arrival deployment includes a Signage Team to ensure current signs are correctly displayed and visible to the traveling public. The CBP Signage Team also conducts an exit sign audit to	June 30, 2021

Recommendation	180-Day Update	Due Date
notice to individuals regarding the collection, use, dissemination, and maintenance of personally identifiable information as part of efforts to promote transparency. While CBP acknowledges that operational constraints may affect the placement of signs or the timely posting of updated signage, the overall public is informed that stakeholders are taking photos in coordination with CBP. Further, CBP's OFO regularly conducts periodic signage audits that include local CBP personnel to ensure signs are accurate and placed appropriately. In addition, CBP notifies travelers at these ports using verbal announcements, signs, and message boards, as appropriate, that CBP takes these photos for identity verification purposes. Travelers are also informed of their ability to request alternative identity verification procedures. Also publicly stated are notifications that, should a traveler decide to request alternative identity verification procedures, the airline would conduct manual identity verification using his/her travel document, and may notify CBP to collect biometrics, such as fingerprints, if applicable. CBP's OFO will also continue to work with airline and airport partners to identify other methods to communicate the use of facial recognition and travelers' privacy rights.	ensure current language is displayed. **Points of Contact / Telephone Numbers / E-mail Addresses:** (b)(6), (b)(7)(C) Acting Director Entry/Exit Policy and Planning Division Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations U.S. Customs and Border Protection **(b)(6), (b)(7)(C)** (b)(6), (b)(7)(C) Director, Biometric Exit-Entry Strategic Transformation Division PPAE Office of Field Operations U.S. Customs and Border Protection O: **(b)(6), (b)(7)(C)** M: **(b)(6), (b)(7)(C)**	

Recommendation	180-Day Update	Due Date
ECD: June 30, 2021		
Recommendation 3: The Commissioner of CBP should direct the BEEP to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information. **Response: Concur.** In the air exit environment, CBP OFO will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, as detailed in the BEEP audit plan, provided to GAO in April 2020. This audit plan enables a comprehensive review of compliance with security and privacy requirements on the part of CBP and CBP's partners. As mentioned in the draft report, CBP completed one partner audit thus far. Although, CBP planned additional audits for 2020, due to the COVID-19 global health pandemic and subsequent travel restrictions, CBP paused the planned audit activities. Once pandemic travel restrictions are lifted, CBP's OFO and Office of Information and Technology (OIT) will resume conducting audits. **(b)(5)**	Due to the COVID-19 global health pandemic and subsequent travel restrictions, CBP paused the planned audit activities; **(b)(5)** to **(b)(5)** **Point of Contact / Telephone Number / E-mail Address:** (b)(6), (b)(7)(C) Acting Director Entry/Exit Policy and Planning Division Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations U.S. Customs and Border Protection **(b)(6), (b)(7)(C)** (b)(6), (b)(7)(C) Director, Biometric Exit-Entry Strategic Transformation Division PPAE Office of Field Operations U.S. Customs and Border Protection O: M: (b)(6), (b)(7)(C)	June 30, 2021

Recommendation	180-Day Update	Due Date
(b)(5) **ECD:** June 30, 2021	(b)(6), (b)(7)(C)	
Recommendation 4: Develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement. **Response: Concur.** The CBP's BEEP's Air Exit Segment was granted Acquisition Decision Event 3 in December 2019. One of the action items from this decision was to complete an update to the Operational Requirements Document (ORD). (b)(5) **ECD:** June 30, 2021	CBP has drafted the updated ORD and is currently circulating the document for review. **Points of Contact / Telephone Number / E-mail Addresses:** (b)(6), (b)(7)(C) Acting Director Entry/Exit Policy and Planning Division Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations (b)(6), (b)(7)(C) (b)(6), (b)(7)(C) Acquisition Program Manager – BEEP Portfolio Acquisition Executive Division Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations U.S. Customs and Border Protection (b)(6), (b)(7)(C)	June 30, 2021
Recommendation 5: Develop a process by which BEEP officials are alerted when the performance of air exit facial recognition falls below established thresholds. **Original Response:** Concur. CBP's OFO has a suite of tools that allows for system and operational performance	Awaiting GAO's confirmation that this recommendation is closed, as implemented. **Points of Contact / Telephone Numbers / E-mail Addresses:** (b)(6), (b)(7)(C) Acting Director Entry/Exit Policy and Planning Division	Pending Complete

Recommendation	180-Day Update	Due Date
management and OFO generates three types of performance reports that are automatically generated and distributed weekly within CBP and to external stakeholders. These reports include: 1) *Saturation Report*: Notes the percentage of flights biometrically processed out of the total number of possible international departures segmented by airport. 2) *Biometric Air Exit Overview Report*: Includes a daily synopsis of operational performance data including numbers of biometrically processed flights and travelers together with biometric match rates. 3) *Stakeholder Raw Data Reports:* Provides Air Exit stakeholders with operational performance data by flight number, passenger counts, and biometric match rates. The OFO's Biometric Entry-Exit Air team monitors the reports for performance issues and addresses any anomalies with stakeholders as they arise. The reports are also used to promote/increase usage by stakeholders. CBP's OFO also conducts random sampling to determine the technical match rates and identify any system or equipment issues. The random sampling is conducted on a weekly basis and includes two flights per	Planning, Program Analysis and Evaluation (PPAE) Office of Field Operations U.S. Customs and Border Protection **(b)(6), (b)(7)(C)** **(b)(6), (b)(7)(C)** Director, Biometric Exit-Entry Strategic Transformation Division PPAE Office of Field Operations U.S. Customs and Border Protection O: **(b)(6), (b)(7)(C)** M: **(b)(6), (b)(7)(C)**	

Recommendation	180-Day Update	Due Date
airport per week. Finally, CBP's OFO receives alert notifications if TVS experiences an outage, and has a Gallery Assembly System monitor that provides notifications when a flight gallery is not created. Depending on the severity and impact to end users, OFO generates stakeholder notifications, as appropriate. CBP requests that GAO consider this recommendation resolved and closed, as implemented.		

Message

From: (b)(6), (b)(7)(C)

Sent: 11/6/2020 4:38:58 PM

To:

(b)(6)

(b)(6)

CC:

(b)(6), (b)(7)(C)

Subject: CBP ONE MOBILE APPLICATION - INTRODUCING A NEW WAY TO REQUEST INSPECTION APPOINTMENTS
Attachments: CBP One - Scheduling QRG_vF.pdf; CBP One Mobile Application Demo.ics

Good Afternoon,

On October 28, 2020, CBP launched the CBP OneT mobile application with its Inspection Appointment feature, introducing a new way of requesting inspection appointments for perishable cargo entering the U.S. by air. Inspection

requests are available via the mobile application in the Miami region. We are excited that after 1 week we already have 18 brokers using the app with more starting each day.

Demo Sessions:
Over the last week, CBP has held multiple demo sessions for brokers, carriers, and forwarders to familiarize them with CBP One and the Inspection Appointment request feature.

We will be hosting a third demo session on **November 10, 2020** for those who have not yet been able to join one. Attached is the calendar invite for this Webex meeting. Please accept the meeting invite if you plan to attend.

How to download CBP OneT
CBP OneT is now live and available for free on the Apple App Store and the Google Play Store. Additional POEs will be available in the future.

Benefits of the mobile application include:

- Secure access to the CBP OneT app via Login.gov
- Convenient submissions for cargo inspections from your mobile device
- Quickly access all outstanding/completed inspection requests
    - Easily review appointment details
- Live stats updates on your appointments to include ***Enroute*** and ***Complete***

For more information, please visit our CBP OneT website. You can also view our CBP OneT: Appointment Feature video on YouTube.

Please contact us at CBPOne@cbp.dhs.gov for questions or additional information.

Sincerely,

CBP OneT Team

U.S. Customs and Border Protection

CBP One™

Use the Appointments Feature Designed for Brokers, Carriers and Forwarders

Download CBP One™ Today!

Streamlines inspection requests and appointment updates

Reduces/Eliminates unnecessary wait time for runners

Enhances communications through email status updates to your group inbox.

Download on the App Store

GET IT ON Google Play

To get started, download CBP One™ on the Apple App Store or Google Play Store.

Sign In Using Login.gov

The app will redirect to **login.gov**, where you can either create or login to your existing account.

Desktop version coming soon!

Questions? Contact us at: cbpone@cbp.dhs.gov

1. Who Are You

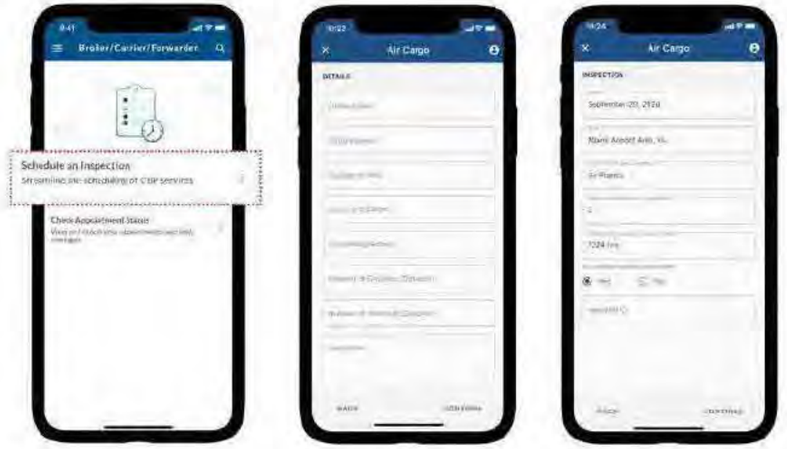
Tap on 'Broker/Carrier/Forwarder' in order to begin.

2. Create Profile

Add and save all necessary information. Your profile can be edited at any time in the future by tapping on the profile icon in the top right corner.

3. Request Inspection

Tap on 'Schedule an Inspection', select appointment type and cargo type, and fill in the required fields.

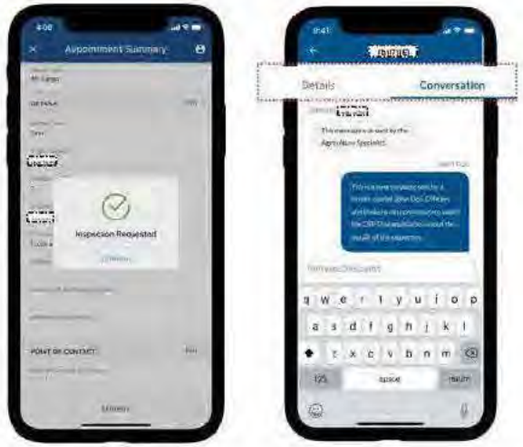


4. Review Information

Review all information and tap on submit. You may edit information on this page by tapping "edit".

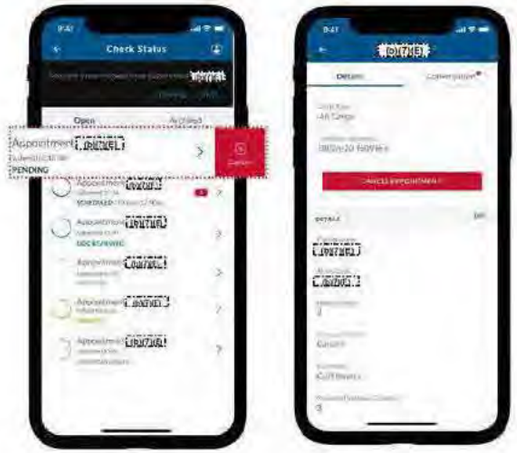
5. Successfully Submitted!

You will receive in app and push notifications, along with emails on the status of your appointments. A CBPAS may initiate a chat, which you can respond to under the 'Conversation' tab.



6. Cancel/Edit an Inspection

View/edit the details of your inspection under "check status". If you need to cancel an inspection, simply swipe to the left or click on the "Cancel Appointment" button in the details screen. Completed/Cancelled appointments will be archived.*



*Only pending inspections can be edited, while pending, acknowledged, scheduled, doc reviewed, and assigned inspections can be cancelled.

Organizer: CBPONE : cbpone@cbp.dhs.gov
Subject: CBP One Mobile Application Demo
Location: Webex
Start Time: 2020-11-10T10:30:00-05:00
End Time: 2020-11-10T11:00:00-05:00
Attendees: CBPONE : cbpone@cbp.dhs.gov

-- Do not delete or change any of the following text. --

When it's time, join your Webex meeting here.

Meeting number (access code): (b)(6), (b)(7)(C)

Meeting password: (b)(6), (b)(7)(C) from phones)

Join meeting

Tap to join from a mobile device (attendees only)

(b)(6), (b)(7)(C)

Some mobile devices may ask attendees to enter a numeric meeting password.

Join by phone

(b)(6), (b)(7)(C)

Global call-in numbers

Join from a video system or application

Dial (b)(6), (b)(7)(C)

Join using Microsoft Lync or Microsoft Skype for Business

Dial (b)(6), (b)(7)(C)

If you are a host, click here to view host information.

U.S. Customs and Border Protection

CBP One™

Allows **International Organizations** to submit advance information for individuals

Download CBP One™

Download on the App Store

GET IT ON Google Play

To get started, download CBP One™ from the Apple App Store or Google Play Store.

Sign In Using Login.gov

The app will redirect to **login.gov** where you can either create or login to your existing account.

Questions?

Contact us at: CBPOne@cbp.dhs.gov

1. Who Are You

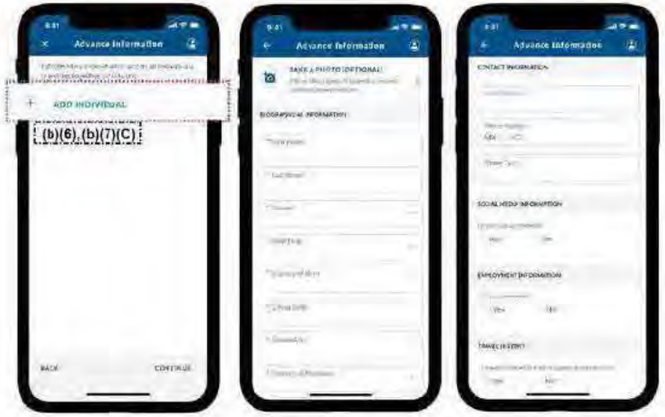
Tap on "International Organization" to begin. First time users will be prompted to create a profile.

2. Submit Advance Information

Select "Submit Advance Information", then acknowledge the Covid-19 checkbox to proceed.*

3. Add Individual

Select "Add Individual" and fill out the required fields for all individuals traveling together on this trip.



4. Address Information

Fill out the address in the USA where they will arrive and reside, along with an emergency USA point of contact. Then, fill out the address for where they lived before coming to the USA and Title 42 remarks.

5. Schedule an Appointment Time

Lastly, select the requested port of entry and date of entry and choose an available time slot.

6. Review and Submit

Review all information and tap on submit. A confirmation screen will display the confirmation number(s) for each individual. A confirmation email will be sent to the email address(es) provided under contact information.

*All individuals whose information is provided via this application are required to provide proof of a negative COVID-19 test completed within 72 hours of presentation at the port of entry

Message

From: (b)(6), (b)(7)(C)

Sent: 6/13/2022 8:26:38 PM

To: (b)(6), (b)(7)(C)

Subject: FW: REQUEST CLOSURE REC 1: GAO-20-568SU: CBP and TSA's Use of Facial Recognition Technology (Job Code GAO-103508)

(b)(6), (b)(7)(C)

I found this email. I think OFO sent the document and copied me. I have to see what I have from (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)

Sent: Thursday, April 22, 2021 9:29 AM

To:

Cc: (b)(6), (b)(7)(C)

Subject: RE: REQUEST CLOSURE REC 1: GAO-20-568SU: CBP and TSA's Use of Facial Recognition Technology (Job Code GAO-103508)

No problem and thank you (b)(6)

(b)(6), (b)(7)(C)
Audit Program Manager
Quality Assurance Enterprise Division
Planning, Program Analysis & Evaluation
Office of Field Operations
U.S. Customs & Border Protection
1300 Pennsylvania Avenue
Room (b)(6), (b)(7)(C)
Washington, DC 20229
Office: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)

From: (b)(6)

Sent: Thursday, April 22, 2021 9:12 AM

To:

Cc: (b)(6), (b)(7)(C)

Subject: RE: REQUEST CLOSURE REC 1: GAO-20-568SU: CBP and TSA's Use of Facial Recognition Technology (Job Code GAO-103508)

Thank you (b)(6), (b)(7)(C)

The files opened just fine, and I'll be reviewing the documents later today and let you know if we have any questions. We really appreciate your help.

V/r

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Thursday, April 22, 2021 8:28 AM
To:
Cc: (b)(6), (b)(7)(C)
Subject: REQUEST CLOSURE REC 1: GAO-20-568SU: CBP and TSA's Use of Facial Recognition Technology (Job Code GAO-103508)

Good morning:

Per my conversation with the Component Audit Liaison, (b)(6), (b)(7)(C) the zipped supporting documents file, that we sent for the closure of the Recommendation 1 audit: GAO-20-568SU: *CBP and TSA's Use of Facial Recognition Technology*, was too large.

I've attached a new and smaller zipped file containing our supporting documentation. It is now 5 MB. Please let me know if there are any issues with this email or if you will need anything further.

Please confirm receipt of this email.

Thank you,

(b)(6), (b)(7)(C)
Audit Program Manager
Quality Assurance Enterprise Division
Planning, Program Analysis & Evaluation
Office of Field Operations
U.S. Customs & Border Protection
1300 Pennsylvania Avenue
Room (b)(6), (b)(7)(C)
Washington, DC 20229
Office: (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)

~~This document and any attachment(s) may contain restricted, sensitive, and/or law enforcement sensitive information belonging to the U.S. Government. It is not for release, review, retransmission, dissemination, or use by anyone other than the intended recipient.~~

Message

From: (b)(6), (b)(7)(C)

Sent: 7/7/2021 6:36:27 PM

To:

CC:

(b)(6), (b)(7)(C)

Subject: Recommendation 3 update for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good Afternoon (b)(6)

CBP changed the ECD to June 30, 2022, for recommendation #3 in audit report GAO-20-568SU entitled *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs*. This change was made based on the below feedback CBP received in response to our request for closure submitted June 23, 2021, for recommendation 3.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone:

From: (b)(6)

Sent: Tuesday, June 29, 2021 11:30 AM

To: (b)(6), (b)(7)(C)

(b)(6)

Cc: (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)

Subject: RE: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Hello (b)(6), (b)(7)(C)

Thank you for the updates, and for the information on recommendation #3. We are happy to see that CBP has made progress addressing the recommendation. In order to close this recommendation as implemented, we suggest waiting

until 2022. Waiting until 2022 will give CBP an opportunity to complete additional assessments, since CBP has only completed two as of June 2021. Providing CBP with additional time may also allow CBP to resolve any issues (i.e. travel restrictions, staff availability) caused by the pandemic. In addition, it would be helpful if CBP could provide us with a plan that includes, 1) the locations and/or names of the partners it plans to assess/audit, and 2) the proposed dates for each of these assessments/audits. Providing that additional information, and level of detail would provide our leadership, and our clients with additional reassurance that CBP has a plan to conduct these audits.

Thank you for the updates on recommendations # 2 and #4. We will follow up with you and your team in 2022 on those recommendations. Please let us know if you and your team are amenable to this plan, or if you need anything further from us. We appreciate your continued assistance.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6), (b)(7)(C)
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, June 23, 2021 5:03 PM
To: (b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
Subject: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good Afternoon (b)(6)

Please see the below updates for recommendations 2, 3 and 4 for the GAO-20-568SU audit report entitled: *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (103508). CBP is requesting closure of recommendation 3 based on the update provided and attached supporting documentation. Also, recommendations 2 and 4 have new estimated completion dates.

Recommendation 2: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's privacy signage is consistently available at all locations where CBP is using facial recognition. ECD: June 30, 2021 **New ECD: December 31, 2021**

June Update: CBP continues to monitor biometric exit signage and has engaged in an extensive signage deployment for entry signage in the air environment. In addition to CBP's ongoing monitoring of signage and continued deployment of signage to new locations, CBP has developed a plan to ensure privacy signage is consistently available at all locations. As part of that plan, CBP is reviewing the signage language and updating it to be more understandable to include making it clearer that travelers can request alternative screening procedures. Additionally, in the upcoming weeks, CBP will be conducting a signage survey of all Field locations. The survey questions have been recently updated to request information on sign version, number of signs, visibility of signs, and operational/infrastructure restrictions, etc. The results of the survey may inform signage distribution, placement, etc. CBP needs additional time to complete this recommendation and has a new estimated completion date of December 31, 2021.

Recommendation 3: The Commissioner of CBP should direct the Biometric Entry-Exit Program to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information. **ECD:** June 30, 2021

June 2021: By April 2021, CBP completed the Metropolitan Washington Airports Authority (MWAA) assessment, and drafted a report with recommended remediations based on the results of the assessment. CBP also conducted an assessment of the Port Authority of New York and New Jersey (PANYNJ) in Newark. As of June 2021, CBP completed two partner assessments. (b)(5)

(b)(5)

See the attached documents that CBP provides to stakeholders prior to conducting the assessments. CBP will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, using the attached documents. Using these documents, CBP can conduct a comprehensive review of partners' compliance with security and privacy requirements.

- Updated Rules of Engagement
- Requested Artifacts
- Penetration Testing Questionnaire

CBP is requesting closure of this recommendation based on CBP's assessment plan and progress made with conducting assessments.

Recommendation 4: The Commissioner of CBP should develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.
ECD: June 30, 2021 **New ECD: March 31, 2022.**

June 2021: (b)(5)

(b)(5)

(b)(5) Currently the CONOPS is being finalized and the estimated approval is August 2021. The ORD has an estimated approval by March 2022. Both of these documents are reviewed by DHS Joint Requirements Council, Science and Technology, Systems Engineering, Chief Technology Office, Chief Information Office, and other DHS entities. CBP needs additional time to complete this recommendation and has a new estimated completion date of March 31, 2022.

Thank you,
(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Message

From: (b)(6), (b)(7)(C)

Sent: 4/14/2021 4:12:43 PM

To:

CC: (b)(6), (b)(7)(C)

Subject: Request for Closure: Recommendation 1 GAO-20-568, "FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues"

Attachments: Biometrics_TCCCannedResponse (002).docx; InfoCenter Briefing.pdf; TCC Biometric Brief 1.ics; TCC Biometric Brief 2.ics

Hi (b)(6)

I am sending this information again. For some reason, (b)(6) has not been receiving my emails with the attachments.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone:

From: (b)(6), (b)(7)(C)

Sent: Monday, March 29, 2021 11:00 AM

To: (b)(6)

Subject: FW: Request for Closure: Recommendation 1 GAO-20-568, "FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues"

Hi (b)(6)

Here is the request for rec 1 with the documents.

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone:

From: (b)(6), (b)(7)(C)

Sent: Monday, January 4, 2021 10:26 AM

To:

Cc: (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)

Subject: Request for Closure: Recommendation 1 GAO-20-568, "FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues"

Good Morning,

Please find below U.S. Customs & Border Protection's (CBP) request to close recommendation 1 from report GAO-20-568, "FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues". For reference, the recommendation and original response/corrective action were:

Recommendation 1: Ensure that the Biometric Entry-Exit Program's privacy notices contain complete and current information, including all of the locations where facial recognition is used and how travelers can request to opt out as appropriate.
Response: Concur. CBP's OFO will collaborate with the CBP Office of Public Affairs to publish: 1) Biometric Entry-Exit privacy information; 2) locations where facial recognition is used; and 3) traveler opt-out procedures on CBP's public-facing website, as well as to review and update that information on a monthly basis. CBP's OFO will also ensure that information provided in response to inquiries via the CBP Call Center is also reviewed and updated monthly. Estimated Completion Date (ECD): December 31, 2020.

December 2020: On September 1, 2020, CBP launched its updated biometrics website (www.biometrics.cbp.gov). The purpose of the site is to deliver information to the public and other stakeholder groups. The site provides a user-friendly communication channel for promoting facial comparison technology and biometrics information in a dynamic and inviting manner. As a testament to CBP's commitment to privacy protections, outlined in the DHS Fair Information Practice Principles, the website includes the current locations using facial comparison technology as well as information on how to request alternative screening and copies of CBP's privacy signage on display. The information provided, including a link to CBP's Traveler Verification Service (TVS) Privacy Impact Assessment (PIA), is yet another tool in CBP's arsenal to ensure technology sustains and does not erode privacy protections. CBP's Office of Field Operations (OFO) met with the CBP Call Center Teams and discussed the Centers' existing and needed biometric resources. On October 1, 2020, OFO conducted three TVS 101 briefings for both the Traveler Call Center and Information Center staff. OFO provided the Call Center with the Public Affairs Guidance, to include Frequently Asked Questions (FAQs). In addition, OFO provided the Information Center with a few FAQs to post on www.help.cbp.gov. CBP will continue to ensure that content is up to date on the CBP website, as required, and when updates are made OFO will provide new details to both the Information Center and Call Center.

CBP respectfully requests that GAO consider this recommendation resolved and closed as implemented.

Supporting documents to show that OFO collaborated with the CBP Office of Public Affairs to publish:

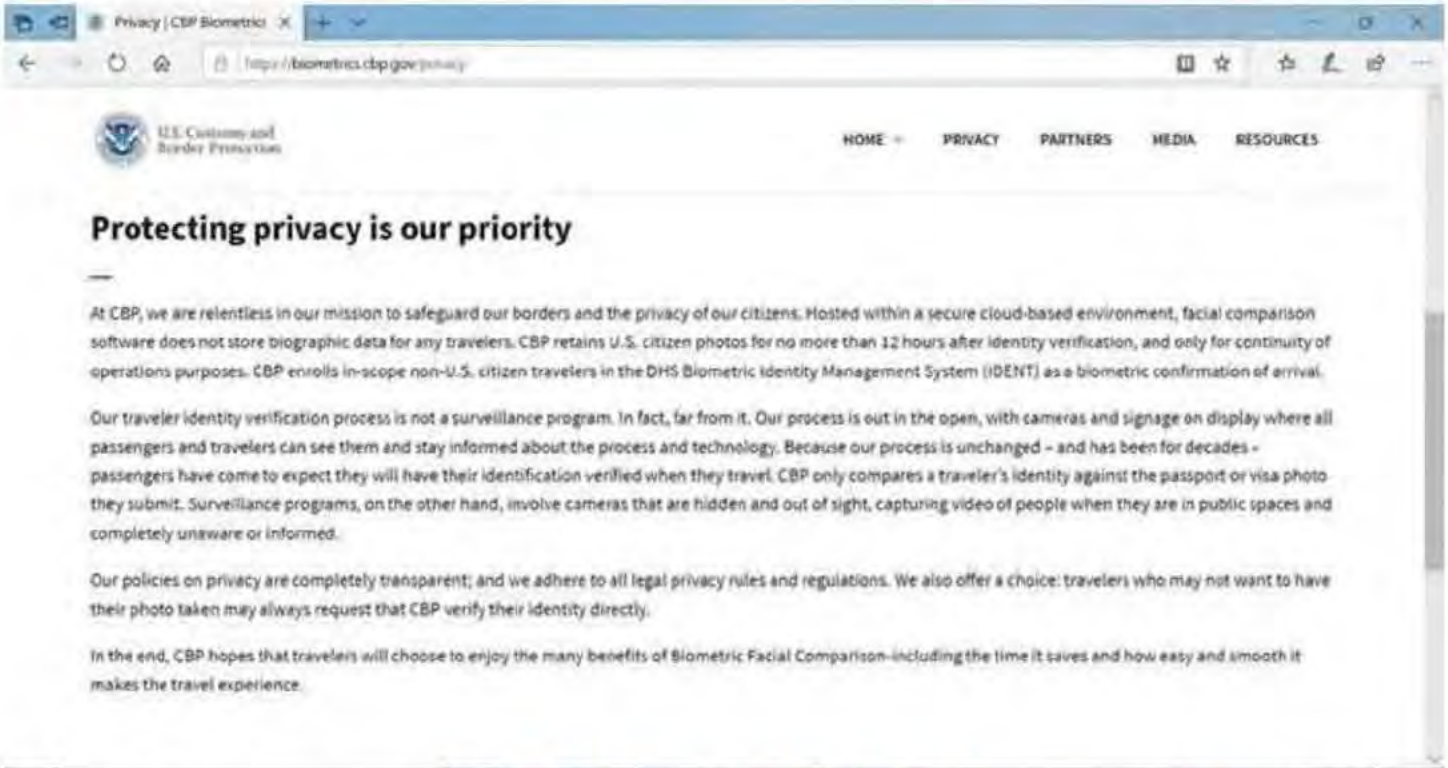
1) Biometric Entry-Exit privacy information;
2) Locations where facial recognition is used;
3) Traveler opt-out procedures on CBP's public-facing website, as well as to review and update that information on a monthly basis; and
4) OFO ensures that information provided in response to inquiries via the CBP Call Center is also reviewed and updated monthly.

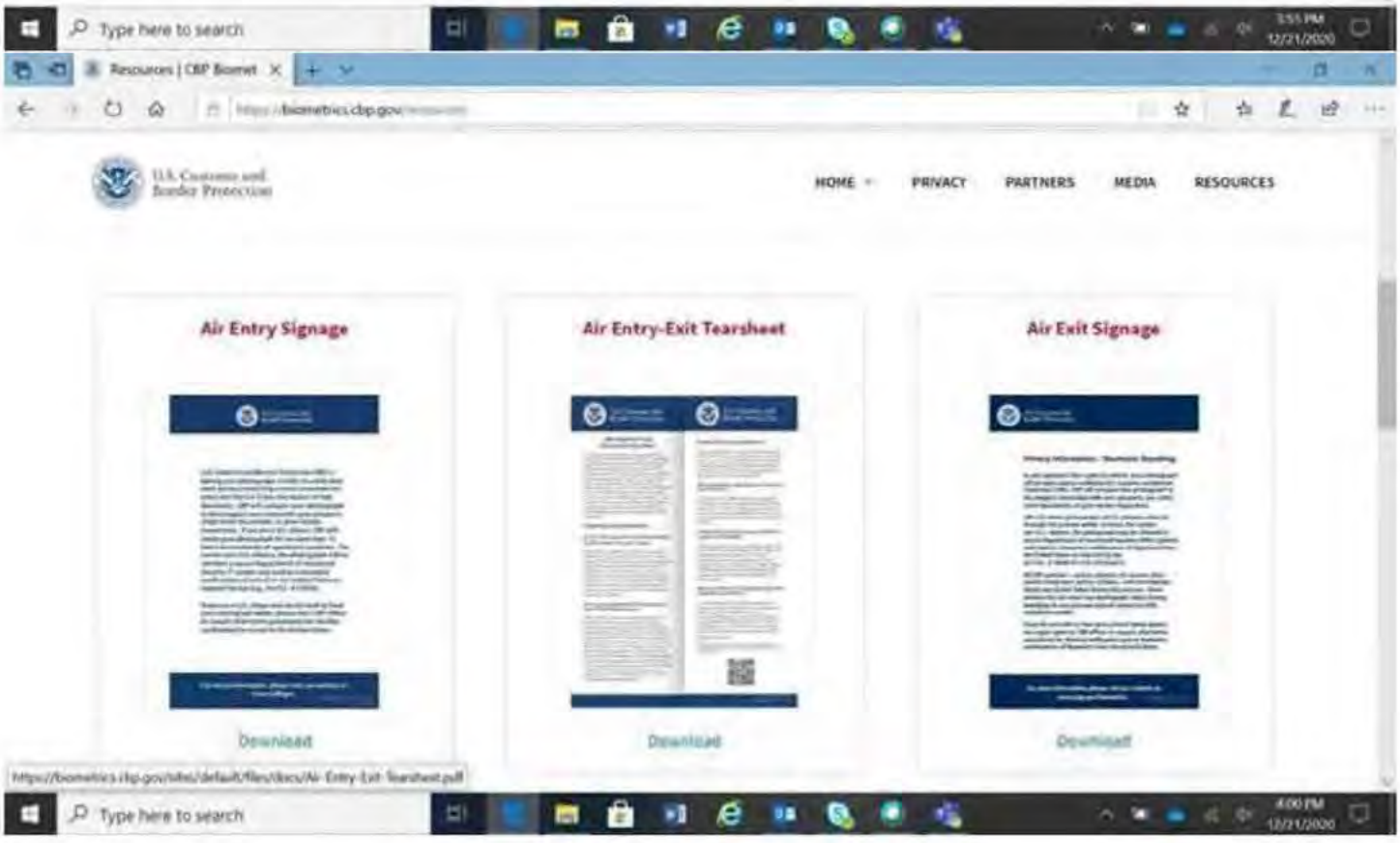
Regarding the website (item 3 above), the website has a Privacy page, Resource page (which includes copies of our actual privacy signage), and a Where page that lists all of the locations with facial comparison technology. Screen shots included below.

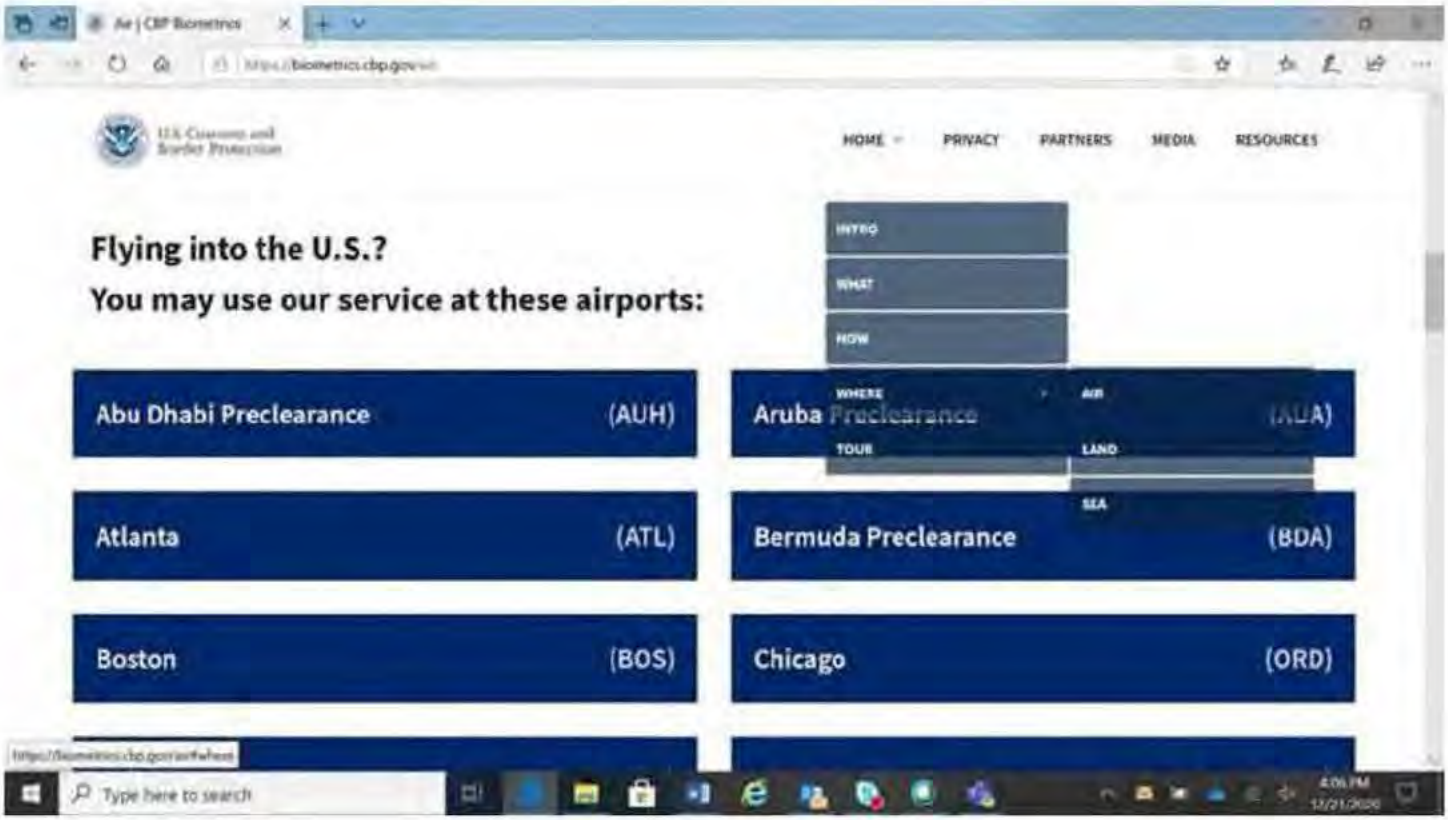
Regarding OFO's collaboration with the Call Center (item 4 above), the following is attached:

- Two Meeting Invites for Biometrics 101 Briefs for Call Center Staffers (one in the morning and one in the afternoon of October 1, 2020 to ensure all staffers were briefed);

- Meeting invite for OFO to join weekly Info Center staff meeting; and
- Responses for Call Center staff.







Should you have any questions, please let me know.

Thank you,
(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Thank you for contacting the CBP Traveler Communications Center.

Regarding U.S. Customs and Border Protection's Biometric Technology, we have included some information below that addresses frequently asked questions on this program. For additional information, please see www.biometrics.cbp.gov

General Information about Biometric Facial Comparison Technology
U.S. Customs and Border Protection (CBP) Office of Field Operations (OFO) implementation of Biometric Facial Comparison Technology (BFCT) for entry and exit in the air environment is refining the arrival and departure process. Biometric entry, also known as Simplified Arrival, is a coordinated effort to design and implement a modernized approach to the air arrival process. In the exit environment, CBP is pursuing public-private partnerships to achieve its mandate to biometrically confirm travelers departing the United States. This BFCT confirms the identity of the traveler, creates an arrival record, and biometrically confirms the entry or departure of eligible, in-scope travelers.

Why Facial Comparison?
BFCT streamlines and modernizes the CBP arrival inspection process without unduly burdening the traveler. Facial comparison transforms the current process to achieve safe, efficient, frictionless air travel while enabling CBP to advance travel facilitation and law enforcement priorities. Keep in mind, all travelers need to carry a passport for international travel regardless of whether not they use BFCT.

How it Works: Biometrics on Entry
The Traveler Verification Service (TVS) is the backend system that performs the facial comparison function.

- CBP anticipates arrival of travelers and builds galleries using traveler photos. Travelers will pause for a photo at the CBP primary inspection area.
- In a matter of seconds, CBP's BFCT will automatically match the new photo of the traveler to high-quality images that the traveler has already provided to the government, such as passport, visa photos, and photos captured on prior encounters.
- An officer performs an admissibility interview and determines whether the traveler should be referred or admitted.

Biometrics on Exit
Prior to exiting the United States, CBP generates galleries of the historical images of travelers for a given flight.

- Each traveler approaches the departure gate during boarding to stand for a photo, which is matched by a camera either operated by CBP or approved partners such as airlines, airport authorities.
- Images are compared to the photos in the flight specific photo gallery.
- If a match is found, the service returns the results to CBP systems, and the traveler boards; if not, a CBP Officer or partner representative will verify the traveler's identity via alternative methods. CBP confirms the traveler's departure from the U.S. in its existing systems.

Privacy
CBP is committed to its privacy obligations and has taken steps to safeguard the privacy of all travelers. CBP complies with the requirement of the Privacy Act of 1974, as amended, the E-Government Act of 2002, the Homeland Security Act of 2002, and Departmental policies that govern the collection, use, and maintenance of personally identifiable information. DHS has published more than ten Privacy Impact

Assessments on all aspects of CBP's Biometric Entry-Exit program, which include policies and procedures for the gathering, storage, analysis, use, dissemination, retention, and deletion of data. CBP discards new photos of U.S. citizens within 12 hours of the identity verification process. CBP temporarily retains the facial images of non-U.S. citizen travelers for up to 14 days in secure CBP systems to support system audits, to evaluate performance, and to ensure accuracy of the BFCT. CBP enrolls eligible non-U.S. citizen travelers in the DHS Biometric Identity Management System (IDENT) as a biometric confirmation of entry or departure and retains the photos for up to 75 years, consistent with the Border Crossing Information (BCI) System of Records Notice (SORN).

Biometrics Beyond Air Environments

Land Environment - CBP has a multi-pronged strategy to integrate biometrics on the land border. CBP will capture images from travelers in vehicles, use BE-Mobile for surge operations, and exchange biographic data with foreign partners.

Sea Environment - CBP is partnering with 7 major cruise lines, testing facial biometric processing for arriving passengers on closed-loop cruises. The goal is to enhance data sharing between CBP and cruise industry and biometrically confirm the identities of passengers and crew.

CBP, in partnership with the Transportation Security Administration (TSA), launched a pilot testing CBP's facial matching service for identity verification at TSA checkpoints. Testing is under way where BCFT is used at baggage drop. The ultimate goal would be BCFT will be used throughout the entire travel continuum.

Thank you again for contacting the Traveler Communications Center.

CBP Biometrics Website: www.biometrics.cbp.gov.

From:
To:
Cc:

(b)(6), (b)(7)(C)

Subject: FW: CBP Information Center All Staff Meeting

-----Original Appointment-----
From: (b)(6), (b)(7)(C)
Sent: Thursday, October 1, 2020 9:27 AM
To: (b)(6), (b)(7)(C)
Cc:

(b)(6), (b)(7)(C)

Subject: CBP Information Center All Staff Meeting
When: Monday, October 5, 2020 9:00 AM-9:30 AM (UTC-05:00) Eastern Time (US & Canada).
Where:

-----Original Appointment-----
From: (b)(6), (b)(7)(C)
Sent: Friday, September 25, 2020 3:52 PM
To: (b)(6), (b)(7)(C)
Cc:

(b)(6), (b)(7)(C)

Subject: CBP Information Center All Staff Meeting
When: Monday, October 5, 2020 9:00 AM-9:30 AM (UTC-05:00) Eastern Time (US & Canada).
Where:

(b)(7)(E)

This Teams Meeting is hosted on a U.S. Government information system and is provided for U.S. Government-authorized use only. Unauthorized or improper use or access of this system may result in disciplinary action as well as civil and criminal penalties.

Organizer: (b)(6), (b)(7)(C)
Subject: Biometrics 101 Briefing
Location: Microsoft Teams Meeting
Start Time: 2020-10-01T14:30:00-04:00
End Time: 2020-10-01T15:30:00-04:00
Attendees:

(b)(6), (b)(7)(C)

Biometrics CBP 101.pptx

The Biometric Entry-Exit Strategic Transformation Team (BEST) will provide call center and CIC staff with a Biometrics 101 Briefing in order to familiarize the team with CBP's biometric programs.

Please forward invite as appropriate.

Presentation Attached.

Join Microsoft Teams Meeting

Learn more about Teams | Meeting options

This Teams Meeting is hosted on a U.S. Government information system and is provided for U.S. Government-authorized use only. Unauthorized or improper use or access of this system may result in disciplinary action as well as civil and criminal penalties.

U.S. Customs and Border Protection

Biometrics 101

September 2020

ENTRY-EXIT | HOW DID WE GET HERE?

U.S. Customs and Border Protection

TIMELINE: HOW CBP HAS CHANGED THE FACE OF TRAVEL

1996: Illegal Immigration Reform and Immigrant Responsibility Act
Congress directs the implementation of a Biographic Entry and Exit System

2000

2002: Border Security Act
Act requires all ports of entry to have equipment that allows biometric comparison and travel document authentication

2004: U.S. Visit Begins to Collect Biometrics
In January 2004, U.S. Visit begins to capture a photograph and fingerprints of foreign travelers

2004: Intelligence Reform and Terrorism Prevention Act
The 9/11 Committee advises Congress to direct DHS to accelerate the implementation of an automated biometric entry/exit data system

2010

2013: Congress Transfers the Biometric Exit Mission to CBP

2015: Biometric Exit (BE) Mobile
Facial biometric pilot at JFK, IAD, ATL, ORD, LAX verifies gaps in biographic data quality

2016: Departure Information System Test
First facial comparison exit pilot at ATL

2016: Pedestrian Biometric Field Test
Pilot at Otay Mesa, California demonstrates that travelers preferred ease of use of facial technology

2017: Executive Order (E.O.) 13780
EO calls for expedited completion of the Biometric Entry-Exit Tracking System

2017: Using the Cloud
Set up commercial cloud matching environment for Sprint operations

2017: Vehicle Test at Anzalduas
Test a front-end system to identify a vendor to conduct future biometric capture capabilities

2019: More to Come!

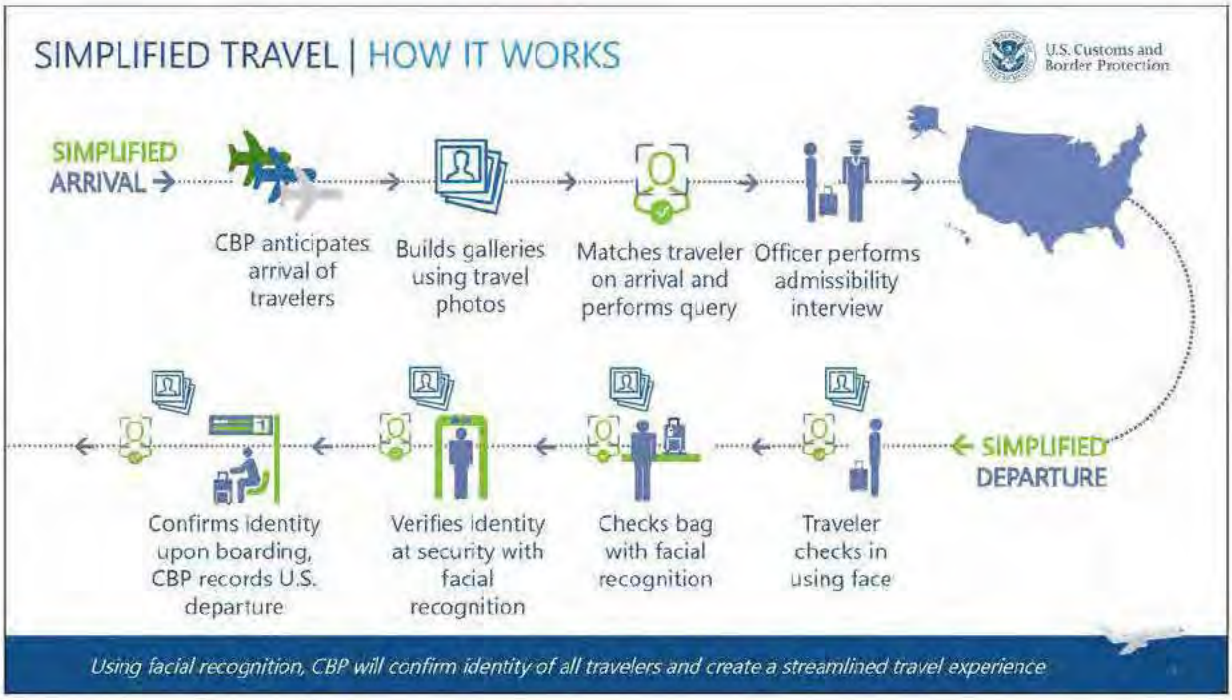
DEFINING THE VISION | BIOMETRIC ENTRY-EXIT CONCEPT

U.S. Customs and Border Protection

Through a biometric matching service and the use of biometric data CBP will transform travel processing by:

- Retrieving all associated traveler images from DHS holdings and segregating them into smaller, more manageable data sets (i.e., by flight, by cruise)
- Fusing biometric and biographic information enabling the biometric to be the key to verifying traveler identity

The biometric matching service is a device-agnostic, secure, cloud-based technical infrastructure to support advanced identity verification.


SIMPLIFIED TRAVEL | HOW IT WORKS
U.S. Customs and Border Protection
SIMPLIFIED ARRIVAL
CBP anticipates arrival of travelers
Builds galleries using travel photos
Matches traveler on arrival and performs query
Officer performs admissibility interview
SIMPLIFIED DEPARTURE
Traveler checks in using face
Checks bag with facial recognition
Verifies identity at security with facial recognition
Confirms identity upon boarding, CBP records U.S. departure
Using facial recognition, CBP will confirm identity of all travelers and create a streamlined travel experience

BIOMETRICS | WHY FACIAL COMPARISON?

U.S. Customs and Border Protection

1. CBP is using **already existing traveler biometric (facial) and biographic information** including visa photos, Primary encounters, certain enforcement data, U.S. Passports, and LPR card photos.

2. Facial comparison **eliminates the need to take fingerprints** each time a foreign national returns to the United States, only requiring on first encounters and expired prints. The system checks the fingerprints already on file against watchlists as part of the vetting process.

3. Facial comparison on Entry allows officers to **focus on the traveler interview and threat detection** that technology cannot identify. **The technology identifies the *knowns*, and the officer determines the *unknowns*.** On Exit, facial comparison provides a more secure means of identity verification and allows airlines to onboard travelers in their departure control systems with just a photograph.

4. Facial comparison technology is **very accurate** and has **drastically increased CBP's ability to detect imposters** when compared to human detection.

5. The use of facial comparison technology has led to **positive tangible results**. To date, 290+ imposters have been caught.

Facial recognition is a tool that enables officers to focus on enforcement and threat detection

TRAVELER VERIFICATION SERVICE | WHY IT WORKS

U.S. Customs and Border Protection

- Uses existing traveler biometrics
- No new data requirements
- Matches one to few utilizing cloud infrastructure
- Token-less processing
- Integrates into existing airport infrastructure
- Extends to land and sea environments
- Trusted source for identity verification

(b)(6), (b)(7)(C), (b)(7)(E)

CBP's matching service enables travel partners to improve traveler experience while meeting the biometric exit mandate

BIOMETRIC PROCESS | HOW IT WORKS

U.S. Customs and Border Protection

BIOMETRIC ENTRY
HOW IT WORKS

- The Officer will initiate traveler processing by taking a photograph rather than a document scan.
- The Officer will review the live and system-matched (gallery) photograph and traveler biographic information.
- If a match is determined and there is no adverse information discovered during primary, the traveler may be granted entry.

BIOMETRIC EXIT
HOW IT WORKS

- During the boarding process, travelers stand in front of the camera and a photo is taken.
- The picture taken is then matched against the photo provided with a traveler's passport, visa, or previous entries.
- Once the traveler is matched to the document with facial comparison technology, then they can proceed to board their flight.

PRECLEARANCE
HOW IT WORKS

- Similar to Biometric Entry, the traveler is processed biometrically in a designated Preclearance facility at the airport before arriving in the U.S.
- Upon arrival, travelers proceed to onward connections or baggage claim.

BIOMETRIC PROCESS | PHOTO CAPTURE AND STORAGE

U.S. Customs and Border Protection

GUIDELINES FOR PHOTO CAPTURE

Entry

	Can I photograph the traveler in primary?	Notes
In-Scope Non-U.S. Citizen	✓ Always	Photos are stored by DHS for 75 years for future identity verification.
Out-of-Scope Non-U.S. Citizen	✓ Can opt out	This group includes travelers under the age of 14 and over the age of 79, diplomats, Canadians, and otherwise exempt aliens. Photos are deleted within 14 days.
U.S. Citizen	✓ Can opt out	Photos of U.S. Citizens are deleted immediately from the system upon matching. U.S. Citizens can opt out if requested. If the camera has automatically captured a photo, you must delete it.

Exit

	Facial Comparison Opt Out	Biometric Capture	Notes
In-Scope Non U.S. Citizen	✓ Can opt out	✓ Required	In-scope travelers follow U.S. Visit rules – Non-immigrants (minus exceptions) and LPRs 14-79 years of age.
Out-of-Scope Non-U.S. Citizen	✓ Can opt out	⊗ Exempt	This group includes minors, diplomats, and Canadians. Manually verify identity through passport/document examination.
U.S. Citizen	✓ Can opt out	⊗ Exempt	Manually verify identity through passport/document examination.

BIOMETRIC PROCESS | *OPERATIONAL PERFORMANCE

U.S. Customs and Border Protection

	Sea	Air Exit	Air Entry	Air Preclearance
	749 voyages processed	100,872 flights processed	282,201 flights processed	32,584 flights processed
	2,132,000+ participating passengers	13,029,000+ participating passengers	21,486,000+ participating passengers	4,378,000+ participating passengers
	96.9% Biometric match rate	97.8% Biometric match rate	98.8% Biometric match rate	99.0% Biometric match rate

*As of September 2020

CBP's TVS has proved reliable in sea and air environments

TANGIBLE BENEFITS | IMPOSTOR APPREHENSION

U.S. Customs and Border Protection

On October 31, 2018, a female subject and her daughter presented B1/B2 BCCs. The biometric algorithm helped the officers determine the facial "mismatch."

Captured Photo	Travel Document Photo	Captured Photo	Travel Document Photo
(b)(6), (b)(7)(C)	(b)(6), (b)(7)(C)	(b)(6), (b)(7)(C)	(b)(6), (b)(7)(C)

Imposter at Nogales, AZ

SIMPLIFIED LAND TRAVEL | PROGRESS AND PLANS

U.S. Customs and Border Protection

Vehicle At Speed
- Capture of private vehicle occupants' faces at both entry and exit

Third Country National Plan
- BE-Mobile for Pulse and Surge Operations and Mobile App for TCN self-report

Pedestrian Technical Demonstration:
- Test Facial Recognition technology at entry and exit

Engagement with Foreign Partners
- Biographic data exchange

CBP has a multi-pronged strategy to integrate biometrics on the land border

CBP is making progress on its biometric exit land border strategy with initial focus on TCNs. National deployment of BE-Mobile for outbound pulse and surge by end of 2018. CBP will initial a vehicle at speed test in Anzalduas, TX in August 2018. CBP will test pedestrian exit using facial matching starting in June 2018.

CBP initiated a data sharing of biographic entry/exit records pilot with Mexico.

CBP is making progress on its biometric exit land border strategy with initial focus on TCNs. National deployment of BE-Mobile for outbound pulse and surge by end of 2018. CBP will initial a vehicle at speed test in Anzalduas, TX in August 2018. CBP will test pedestrian exit using facial matching starting in June 2018.

CBP initiated a data sharing of biographic entry/exit records pilot with Mexico.

CBP is making progress on its biometric exit land border strategy with initial focus on TCNs. National deployment of BE-Mobile for outbound pulse and surge by end of 2018. CBP will initial a vehicle at speed test in Anzalduas, TX in August 2018. CBP will test pedestrian exit using facial matching starting in June 2018.

CBP initiated a data sharing of biographic entry/exit records pilot with Mexico.

SIMPLFIED TRAVEL | COMMITMENT TO PRIVACY

U.S. Customs and Border Protection

- Transparency Efforts
  - Briefing sessions with privacy advocates and stakeholders
- Notification to the Public
  - Privacy Impact Assessments completed for CBP and airline-led projects
  - Online content at CBP.GOV
    - Fact sheets
    - Frequently Asked Questions
  - Signage at inspection sites and gate announcements
  - U.S. citizens not wishing to have a photo taken can request an alternative ID verification process
- Retention Period for Facial Images
  - No more than two weeks for out-of-scope non-U.S. travelers, for confirmation of travelers' identities, evaluation of the technology, assurance of accuracy of the algorithms, and system audits
  - U.S. citizen photos deleted immediately

CBP commitment to transparency builds public trust while enhancing security and facilitate travel

SIMPLIFIED TRAVEL | THANK YOU

U.S. Customs and Border Protection

(b)(6), (b)(7)(C) Branch Chief
Biometric Exit-Entry Strategic Transformation

Planning, Program Analysis and Evaluation
Office of Field Operations
U.S. Customs and Border Protection

(b)(6), (b)(7)(C)

Organizer: (b)(6), (b)(7)(C)

Subject: Call Center Biometrics 101 Briefing

Location: Teams Meeting

Start Time: 2020-10-01T07:30:00-04:00

End Time: 2020-10-01T08:30:00-04:00

Attendees: (b)(6), (b)(7)(C)

Biometrics CBP 101.pptx

The Biometric Entry-Exit Strategic Transformation Team (BEST) will provide call center staff with a Biometrics 101 Briefing in order to familiarize the team with CBP's biometric programs.

Please forward invite as appropriate to midnight shift.

Presentation attached.

Join Microsoft Teams Meeting

Learn more about Teams | Meeting options

U.S. Customs and Border Protection

Biometrics 101

September 2020

ENTRY-EXIT | HOW DID WE GET HERE?

U.S. Customs and Border Protection

TIMELINE: HOW CBP HAS CHANGED THE FACE OF TRAVEL

1996: Illegal Immigration Reform and Immigrant Responsibility Act
Congress directs the implementation of a Biographic Entry and Exit System

2000

2002: Border Security Act
Act requires all ports of entry to have equipment that allows biometric comparison and travel document authentication

2004: U.S. Visit Begins to Collect Biometrics
In January 2004, U.S. Visit begins to capture a photograph and fingerprints of foreign travelers

2004: Intelligence Reform and Terrorism Prevention Act
The 9/11 Committee advises Congress to direct DHS to accelerate the implementation of an automated biometric entry/exit data system

2010

2013: Congress Transfers the Biometric Exit Mission to CBP

2015: Biometric Exit (BE) Mobile
Facial biometric pilot at JFK, IAD, ATL, ORD, LAX verifies gaps in biographic data quality

2016: Departure Information System Test
First facial comparison exit pilot at ATL

2016: Pedestrian Biometric Field Test
Pilot at Otay Mesa, California demonstrates that travelers preferred ease of use of facial technology

2017: Executive Order (E.O.) 13780
EO calls for expedited completion of the Biometric Entry-Exit Tracking System

2017: Using the Cloud
Set up commercial cloud matching environment for Sprint operations

2017: Vehicle Test at Anzalduas
Test a front-end system to identify a vendor to conduct future biometric capture capabilities

2019: More to Come!

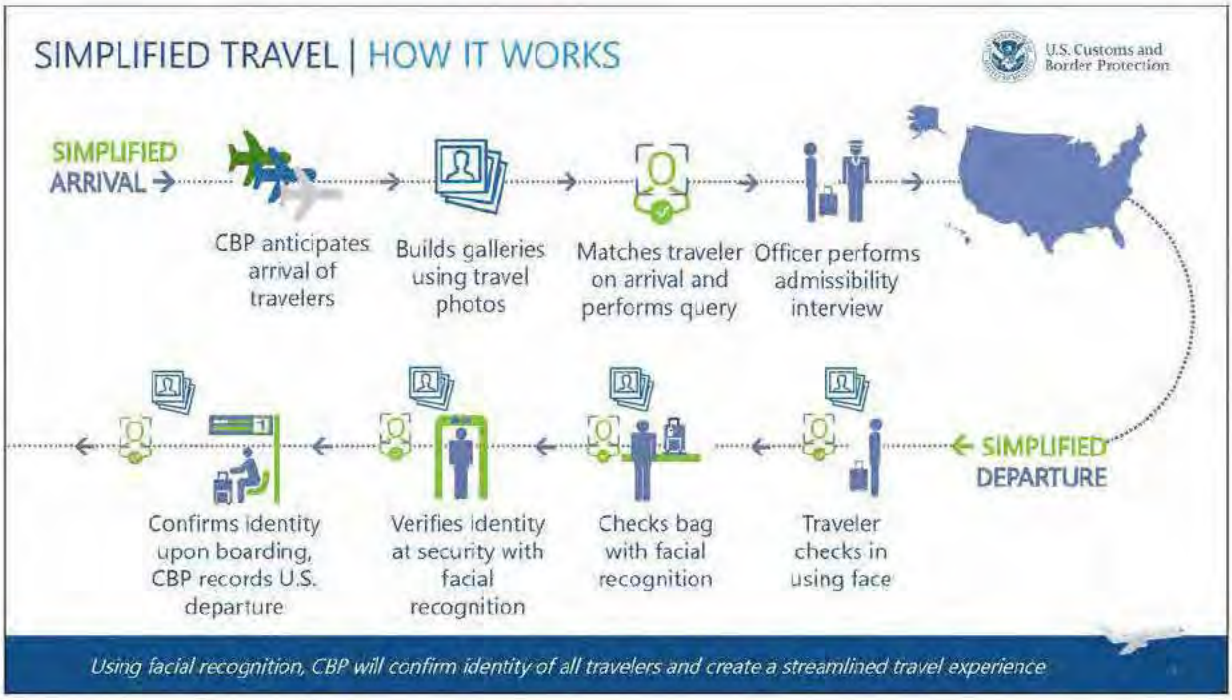
DEFINING THE VISION | BIOMETRIC ENTRY-EXIT CONCEPT

U.S. Customs and Border Protection

Through a biometric matching service and the use of biometric data CBP will transform travel processing by:

- Retrieving all associated traveler images from DHS holdings and segregating them into smaller, more manageable data sets (i.e., by flight, by cruise)
- Fusing biometric and biographic information enabling the biometric to be the key to verifying traveler identity

The biometric matching service is a device-agnostic, secure, cloud-based technical infrastructure to support advanced identity verification.


SIMPLIFIED TRAVEL | HOW IT WORKS
U.S. Customs and Border Protection
SIMPLIFIED ARRIVAL
CBP anticipates arrival of travelers
Builds galleries using travel photos
Matches traveler on arrival and performs query
Officer performs admissibility interview
SIMPLIFIED DEPARTURE
Traveler checks in using face
Checks bag with facial recognition
Verifies identity at security with facial recognition
Confirms identity upon boarding, CBP records U.S. departure
Using facial recognition, CBP will confirm identity of all travelers and create a streamlined travel experience

BIOMETRICS | WHY FACIAL COMPARISON?

U.S. Customs and Border Protection

1. CBP is using **already existing traveler biometric (facial) and biographic information** including visa photos, Primary encounters, certain enforcement data, U.S. Passports, and LPR card photos.

2. Facial comparison **eliminates the need to take fingerprints** each time a foreign national returns to the United States, only requiring on first encounters and expired prints. The system checks the fingerprints already on file against watchlists as part of the vetting process.

3. Facial comparison on Entry allows officers to **focus on the traveler interview and threat detection** that technology cannot identify. **The technology identifies the *knowns*, and the officer determines the *unknowns*.** On Exit, facial comparison provides a more secure means of identity verification and allows airlines to onboard travelers in their departure control systems with just a photograph.

4. Facial comparison technology is **very accurate** and has **drastically increased CBP's ability to detect imposters** when compared to human detection.

5. The use of facial comparison technology has led to **positive tangible results**. To date, 290+ imposters have been caught.

Facial recognition is a tool that enables officers to focus on enforcement and threat detection

TRAVELER VERIFICATION SERVICE | WHY IT WORKS

U.S. Customs and Border Protection

- Uses existing traveler biometrics
- No new data requirements
- Matches one to few utilizing cloud infrastructure
- Token-less processing
- Integrates into existing airport infrastructure
- Extends to land and sea environments
- Trusted source for identity verification

(b)(6), (b)(7)(C), (b)(7)(E)

CBP's matching service enables travel partners to improve traveler experience while meeting the biometric exit mandate

BIOMETRIC PROCESS | HOW IT WORKS

U.S. Customs and Border Protection

BIOMETRIC ENTRY
HOW IT WORKS

- The Officer will initiate traveler processing by taking a photograph rather than a document scan.
- The Officer will review the live and system-matched (gallery) photograph and traveler biographic information.
- If a match is determined and there is no adverse information discovered during primary, the traveler may be granted entry.

BIOMETRIC EXIT
HOW IT WORKS

- During the boarding process, travelers stand in front of the camera and a photo is taken.
- The picture taken is then matched against the photo provided with a traveler's passport, visa, or previous entries.
- Once the traveler is matched to the document with facial comparison technology, then they can proceed to board their flight.

PRECLEARANCE
HOW IT WORKS

- Similar to Biometric Entry, the traveler is processed biometrically in a designated Preclearance facility at the airport before arriving in the U.S.
- Upon arrival, travelers proceed to onward connections or baggage claim.

BIOMETRIC PROCESS | PHOTO CAPTURE AND STORAGE

U.S. Customs and Border Protection

GUIDELINES FOR PHOTO CAPTURE

Entry

	Can I photograph the traveler in primary?	Notes
In-Scope Non-U.S. Citizen	✓ Always	Photos are stored by DHS for 75 years for future identity verification.
Out-of-Scope Non-U.S. Citizen	✓ Can opt out	This group includes travelers under the age of 14 and over the age of 79, diplomats, Canadians, and otherwise exempt aliens. Photos are deleted within 14 days.
U.S. Citizen	✓ Can opt out	Photos of U.S. Citizens are deleted immediately from the system upon matching. U.S. Citizens can opt out if requested. If the camera has automatically captured a photo, you must delete it.

Exit

	Facial Comparison Opt Out	Biometric Capture	Notes
In-Scope Non U.S. Citizen	✓ Can opt out	✓ Required	In-scope travelers follow U.S. Visit rules – Non-immigrants (minus exceptions) and LPRs 14-79 years of age.
Out-of-Scope Non-U.S. Citizen	✓ Can opt out	⊗ Exempt	This group includes minors, diplomats, and Canadians. Manually verify identity through passport/document examination.
U.S. Citizen	✓ Can opt out	⊗ Exempt	Manually verify identity through passport/document examination.

BIOMETRIC PROCESS | *OPERATIONAL PERFORMANCE

U.S. Customs and Border Protection

	Sea	Air Exit	Air Entry	Air Preclearance
	749 voyages processed	100,872 flights processed	282,201 flights processed	32,584 flights processed
	2,132,000+ participating passengers	13,029,000+ participating passengers	21,486,000+ participating passengers	4,378,000+ participating passengers
	96.9% Biometric match rate	97.8% Biometric match rate	98.8% Biometric match rate	99.0% Biometric match rate

*As of September 2020

CBP's TVS has proved reliable in sea and air environments

TANGIBLE BENEFITS | IMPOSTOR APPREHENSION

U.S. Customs and Border Protection

On October 31, 2018, a female subject and her daughter presented B1/B2 BCCs. The biometric algorithm helped the officers determine the facial "mismatch."

Captured Photo	Travel Document Photo	Captured Photo	Travel Document Photo
(b)(6), (b)(7)(C)	(b)(6), (b)(7)(C)	(b)(6), (b)(7)(C)	(b)(6), (b)(7)(C)

Imposter at Nogales, AZ

SIMPLIFIED LAND TRAVEL | PROGRESS AND PLANS

U.S. Customs and Border Protection

Vehicle At Speed
- Capture of private vehicle occupants' faces at both entry and exit

Third Country National Plan
- BE-Mobile for Pulse and Surge Operations and Mobile App for TCN self-report

Pedestrian Technical Demonstration:
- Test Facial Recognition technology at entry and exit

Engagement with Foreign Partners
- Biographic data exchange

CBP has a multi-pronged strategy to integrate biometrics on the land border

CBP is making progress on its biometric exit land border strategy with initial focus on TCNs. National deployment of BE-Mobile for outbound pulse and surge by end of 2018. CBP will initial a vehicle at speed test in Anzalduas, TX in August 2018. CBP will test pedestrian exit using facial matching starting in June 2018.

CBP initiated a data sharing of biographic entry/exit records pilot with Mexico.

CBP is making progress on its biometric exit land border strategy with initial focus on TCNs. National deployment of BE-Mobile for outbound pulse and surge by end of 2018. CBP will initial a vehicle at speed test in Anzalduas, TX in August 2018. CBP will test pedestrian exit using facial matching starting in June 2018.

CBP initiated a data sharing of biographic entry/exit records pilot with Mexico.

CBP is making progress on its biometric exit land border strategy with initial focus on TCNs. National deployment of BE-Mobile for outbound pulse and surge by end of 2018. CBP will initial a vehicle at speed test in Anzalduas, TX in August 2018. CBP will test pedestrian exit using facial matching starting in June 2018.

CBP initiated a data sharing of biographic entry/exit records pilot with Mexico.

SIMPLFIED TRAVEL | COMMITMENT TO PRIVACY

U.S. Customs and Border Protection

- Transparency Efforts
  - Briefing sessions with privacy advocates and stakeholders
- Notification to the Public
  - Privacy Impact Assessments completed for CBP and airline-led projects
  - Online content at CBP.GOV
    - Fact sheets
    - Frequently Asked Questions
  - Signage at inspection sites and gate announcements
  - U.S. citizens not wishing to have a photo taken can request an alternative ID verification process
- Retention Period for Facial Images
  - No more than two weeks for out-of-scope non-U.S. travelers, for confirmation of travelers' identities, evaluation of the technology, assurance of accuracy of the algorithms, and system audits
  - U.S. citizen photos deleted immediately

CBP commitment to transparency builds public trust while enhancing security and facilitate travel

SIMPLIFIED TRAVEL | THANK YOU

U.S. Customs and Border Protection

(b)(6), (b)(7)(C) Branch Chief
Biometric Exit-Entry Strategic Transformation

Planning, Program Analysis and Evaluation
Office of Field Operations
U.S. Customs and Border Protection

(b)(6), (b)(7)(C)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/1/2022 3:55:04 PM

To: (b)(6), (b)(7)(C)

Subject: FW: Recommendation 3 update for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

FYI

From: (b)(6)

Sent: Wednesday, July 7, 2021 2:42 PM

To: (b)(6), (b)(7)(C)

Cc: (b)(6), (b)(7)(C)

Subject: RE: Recommendation 3 update for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good afternoon (b)(6), (b)(7)(C)

We appreciate the update and will follow up with you and the CBP team in 2022.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

From: (b)(6), (b)(7)(C)

Sent: Wednesday, July 7, 2021 2:36 PM

To: (b)(6), (b)(7)(C)

Cc: (b)(6), (b)(7)(C)

Subject: Recommendation 3 update for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good Afternoon (b)(6)

CBP changed the ECD to June 30, 2022, for recommendation #3 in audit report GAO-20-568SU entitled *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs*. This change was made based on the below feedback CBP received in response to our request for closure submitted June 23, 2021, for recommendation 3.

Thank you,
(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Tuesday, June 29, 2021 11:30 AM
To: (b)(6), (b)(7)(C)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6)
Subject: RE: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Hello (b)(6), (b)(7)(C)

Thank you for the updates, and for the information on recommendation #3. We are happy to see that CBP has made progress addressing the recommendation. In order to close this recommendation as implemented, we suggest waiting until 2022. Waiting until 2022 will give CBP an opportunity to complete additional assessments, since CBP has only completed two as of June 2021. Providing CBP with additional time may also allow CBP to resolve any issues (i.e. travel restrictions, staff availability) caused by the pandemic. In addition, it would be helpful if CBP could provide us with a plan that includes, 1) the locations and/or names of the partners it plans to assess/audit, and 2) the proposed dates for each of these assessments/audits. Providing that additional information, and level of detail would provide our leadership, and our clients with additional reassurance that CBP has a plan to conduct these audits.

Thank you for the updates on recommendations # 2 and #4. We will follow up with you and your team in 2022 on those recommendations. Please let us know if you and your team are amenable to this plan, or if you need anything further from us. We appreciate your continued assistance.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, June 23, 2021 5:03 PM
To: (b)(6), (b)(7)(C)
Cc:
(b)(6)

Subject: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good Afternoon (b)(6)

Please see the below updates for recommendations 2, 3 and 4 for the GAO-20-568SU audit report entitled: *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (103508). CBP is requesting closure of recommendation 3 based on the update provided and attached supporting documentation. Also, recommendations 2 and 4 have new estimated completion dates.

Recommendation 2: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's privacy signage is consistently available at all locations where CBP is using facial recognition. ECD: June 30, 2021 **New ECD: December 31, 2021**

June Update: CBP continues to monitor biometric exit signage and has engaged in an extensive signage deployment for entry signage in the air environment. In addition to CBP's ongoing monitoring of signage and continued deployment of signage to new locations, CBP has developed a plan to ensure privacy signage is consistently available at all locations. As part of that plan, CBP is reviewing the signage language and updating it to be more understandable to include making it clearer that travelers can request alternative screening procedures. Additionally, in the upcoming weeks, CBP will be conducting a signage survey of all Field locations. The survey questions have been recently updated to request information on sign version, number of signs, visibility of signs, and operational/infrastructure restrictions, etc. The results of the survey may inform signage distribution, placement, etc. CBP needs additional time to complete this recommendation and has a new estimated completion date of December 31, 2021.

Recommendation 3: The Commissioner of CBP should direct the Biometric Entry-Exit Program to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information. **ECD:** June 30, 2021

June 2021: By April 2021, CBP completed the Metropolitan Washington Airports Authority (MWAA) assessment, and drafted a report with recommended remediations based on the results of the assessment. CBP also conducted an assessment of the Port Authority of New York and New Jersey (PANYNJ) in Newark. As of June 2021, CBP completed two partner assessments. (b)(5)

(b)(5)

See the attached documents that CBP provides to stakeholders prior to conducting the assessments. CBP will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, using the attached documents. Using these documents, CBP can conduct a comprehensive review of partners' compliance with security and privacy requirements.

- Updated Rules of Engagement
- Requested Artifacts
- Penetration Testing Questionnaire

CBP is requesting closure of this recommendation based on CBP's assessment plan and progress made with conducting assessments.

Recommendation 4: The Commissioner of CBP should develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.

ECD: June 30, 2021 **New ECD: March 31, 2022.**

June 2021: (b)(5)

(b)(5)

(b)(5) Both of these documents are reviewed by DHS Joint Requirements Council, Science and Technology, Systems Engineering, Chief Technology Office, Chief Information Office, and other DHS entities. CBP needs additional time to complete this recommendation and has a new estimated completion date of March 31, 2022.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Message

From: (b)(6), (b)(7)(C)

Sent: 11/19/2020 8:40:41 PM
To:
CC: (b)(6), (b)(7)(C)

Subject: Status of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Afternoon (b)(6)

I am touching base to see if GAO made a decision on this request for closure.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Thursday, September 17, 2020 8:56 AM
To:
Cc: (b)(6), (b)(7)(C)
(b)(6)
Subject: RE: Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good morning (b)(6), (b)(7)(C)

Thank you for reaching out to us and double checking on that recommendation. Based on the timing of when CBP sent us the documentation associated with this recommendation, we were unable to fully assess the documentation and CBP's actions to determine whether the actions CBP took fully addressed our recommendation before we issued our report. As we mention towards the end of our report, in the "Agency Comments and Our Evaluation" section on pg. 73, once we have an opportunity to fully review CBP's documentation, we will determine the extent to which CBP's actions fully address the recommendation, and then determine if we can close the recommendation. We expect to finish our assessment, and provide CBP with an update within 60 days of the issuance of our final report. Expect to hear back from us on around 11/2/20. Please let us know if you have any other questions.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, September 16, 2020 4:37 PM
To:
Cc: (b)(6), (b)(7)(C)
Subject: Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Afternoon (b)(6)

CBP requested closure of this recommendation in the attached management response letter, but this recommendation remained in the final report. Can you provide more details regarding the additional actions or information needed for GAO to considered the below recommendation resolved and closed?

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/1/2022 3:54:29 PM

To: (b)(6)

Subject: FW: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

FYI

From: (b)(6)

Sent: Tuesday, June 29, 2021 11:30 AM

To: (b)(6), (b)(7)(C)

(b)(6)

Cc: (b)(6), (b)(7)(C)

(b)(6)

Subject: RE: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Hello (b)(6), (b)(7)(C)

Thank you for the updates, and for the information on recommendation #3. We are happy to see that CBP has made progress addressing the recommendation. In order to close this recommendation as implemented, we suggest waiting until 2022. Waiting until 2022 will give CBP an opportunity to complete additional assessments, since CBP has only completed two as of June 2021. Providing CBP with additional time may also allow CBP to resolve any issues (i.e. travel restrictions, staff availability) caused by the pandemic. In addition, it would be helpful if CBP could provide us with a plan that includes, 1) the locations and/or names of the partners it plans to assess/audit, and 2) the proposed dates for each of these assessments/audits. Providing that additional information, and level of detail would provide our leadership, and our clients with additional reassurance that CBP has a plan to conduct these audits.

Thank you for the updates on recommendations # 2 and #4. We will follow up with you and your team in 2022 on those recommendations. Please let us know if you and your team are amenable to this plan, or if you need anything further from us. We appreciate your continued assistance.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

From: (b)(6), (b)(7)(C)

Sent: Wednesday, June 23, 2021 5:03 PM

To:

Cc: (b)(6), (b)(7)(C)

(b)(6)

Subject: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good Afternoon (b)(6)

Please see the below updates for recommendations 2, 3 and 4 for the GAO-20-568SU audit report entitled: *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (103508). CBP is requesting closure of recommendation 3 based on the update provided and attached supporting documentation. Also, recommendations 2 and 4 have new estimated completion dates.

Recommendation 2: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's privacy signage is consistently available at all locations where CBP is using facial recognition. ECD: June 30, 2021 **New ECD: December 31, 2021**

June Update: CBP continues to monitor biometric exit signage and has engaged in an extensive signage deployment for entry signage in the air environment. In addition to CBP's ongoing monitoring of signage and continued deployment of signage to new locations, (b)(5)

(b)(5)

(b)(5) CBP needs additional time to complete this recommendation and has a new estimated completion date of December 31, 2021.

Recommendation 3: The Commissioner of CBP should direct the Biometric Entry-Exit Program to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information. **ECD:** June 30, 2021

June 2021: By April 2021, CBP completed the Metropolitan Washington Airports Authority (MWAA) assessment, and drafted a report with recommended remediations based on the results of the assessment. CBP also conducted an assessment of the Port Authority of New York and New Jersey (PANYNJ) in Newark. As of June 2021, CBP completed two partner assessments. (b)(5)

(b)(5)

See the attached documents that CBP provides to stakeholders prior to conducting the assessments. CBP will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, using the attached documents. Using these documents, CBP can conduct a comprehensive review of partners' compliance with security and privacy requirements.

- Updated Rules of Engagement
- Requested Artifacts
- Penetration Testing Questionnaire

CBP is requesting closure of this recommendation based on CBP's assessment plan and progress made with conducting assessments.

Recommendation 4: The Commissioner of CBP should develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.
ECD: June 30, 2021 **New ECD: March 31, 2022.**

June 2021: (b)(5)

(b)(5) re

(b)(5) Both of these documents are reviewed by DHS Joint Requirements Council, Science and Technology, Systems Engineering, Chief Technology Office, Chief Information Office, and other DHS entities. CBP needs additional time to complete this recommendation and has a new estimated completion date of March 31, 2022.

Thank you,
(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/1/2022 3:54:00 PM

To: (b)(6), (b)(7)(C)

Subject: FW: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

FYI

From: (b)(6)

Sent: Thursday, June 24, 2021 10:03 AM

To: (b)(6), (b)(7)(C)

Cc: (b)(6)

Subject: RE: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good morning (b)(6), (b)(7)(C)

Thank you for this update. We'll review this information and get back to you at some point next week. We appreciate your help on addressing these outstanding recommendations.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

From: (b)(6), (b)(7)(C)

Sent: Wednesday, June 23, 2021 5:03 PM

To: (b)(6), (b)(7)(C)

Cc: (b)(6)

Subject: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Good Afternoon (b)(6)

Please see the below updates for recommendations 2, 3 and 4 for the GAO-20-568SU audit report entitled: *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (103508). CBP is requesting closure of recommendation 3 based on the update provided and attached supporting documentation. Also, recommendations 2 and 4 have new estimated completion dates.

Recommendation 2: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's privacy signage is consistently available at all locations where CBP is using facial recognition. ECD: June 30, 2021
New ECD: December 31, 2021

June Update: CBP continues to monitor biometric exit signage and has engaged in an extensive signage deployment for entry signage in the air environment. In addition to CBP's ongoing monitoring of signage and continued deployment of signage to new locations, (b)(5)

(b)(5) CBP needs additional time to complete this recommendation and has a new estimated completion date of December 31, 2021.

Recommendation 3: The Commissioner of CBP should direct the Biometric Entry-Exit Program to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information. **ECD:** June 30, 2021

June 2021: By April 2021, CBP completed the Metropolitan Washington Airports Authority (MWAA) assessment, and drafted a report with recommended remediations based on the results of the assessment. CBP also conducted an assessment of the Port Authority of New York and New Jersey (PANYNJ) in Newark. As of June 2021, CBP completed two partner assessments. (b)(5)

(b)(5)

See the attached documents that CBP provides to stakeholders prior to conducting the assessments. CBP will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, using the attached documents. Using these documents, CBP can conduct a comprehensive review of partners' compliance with security and privacy requirements.

- Updated Rules of Engagement
- Requested Artifacts
- Penetration Testing Questionnaire

CBP is requesting closure of this recommendation based on CBP's assessment plan and progress made with conducting assessments.

Recommendation 4: The Commissioner of CBP should develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.
ECD: June 30, 2021 **New ECD: March 31, 2022.**

June 2021 (b)(5)

(b)(5)

(b)(5)

(b)(5) Both of these documents are reviewed by DHS Joint Requirements Council, Science and Technology, Systems Engineering, Chief Technology Office, Chief Information Office, and other DHS entities. CBP needs additional time to complete this recommendation and has a new estimated completion date of March 31, 2022.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/23/2021 9:03:03 PM

To:

CC: (b)(6), (b)(7)(C)

Subject: Recommendation 2, 3 and 4 updates for the GAO-20-568SU audit report entitled: Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues (103508).R

Attachments: Updated Rules of Engagement - 06.16.21.docx; Stakeholder Assessment Document Request 02.01.21.docx; Penetration Testing Pre-Assessment Questionnaire Target List.xlsx

Good Afternoon (b)(6)

Please see the below updates for recommendations 2, 3 and 4 for the GAO-20-568SU audit report entitled: *Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues* (103508). CBP is requesting closure of recommendation 3 based on the update provided and attached supporting documentation. Also, recommendations 2 and 4 have new estimated completion dates.

Recommendation 2: The Commissioner of CBP should ensure that the Biometric Entry-Exit Program's privacy signage is consistently available at all locations where CBP is using facial recognition. ECD: June 30, 2021 **New ECD: December 31, 2021**

June Update: CBP continues to monitor biometric exit signage and has engaged in an extensive signage deployment for entry signage in the air environment. In addition to CBP's ongoing monitoring of signage and continued deployment of signage to new locations, CBP has developed a plan to ensure privacy signage is consistently available at all locations. As part of that plan, (b)(5)

(b)(5)

(b)(5) CBP needs additional time to complete this recommendation and has a new estimated completion date of December 31, 2021.

Recommendation 3: The Commissioner of CBP should direct the Biometric Entry-Exit Program to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information. **ECD:** June 30, 2021

June 2021: By April 2021, CBP completed the Metropolitan Washington Airports Authority (MWAA) assessment, and drafted a report with recommended remediations based on the results of the assessment. CBP also conducted an assessment of the Port Authority of New York and New Jersey (PANYNJ) in Newark. As of June 2021, CBP completed two partner assessments. (b)(5)

(b)(5)

(b)(5)

See the attached documents that CBP provides to stakeholders prior to conducting the assessments. CBP will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, using the attached documents. Using these documents, CBP can conduct a comprehensive review of partners' compliance with security and privacy requirements.

- Updated Rules of Engagement
- Requested Artifacts
- Penetration Testing Questionnaire

CBP is requesting closure of this recommendation based on CBP's assessment plan and progress made with conducting assessments.

Recommendation 4: The Commissioner of CBP should develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.
ECD: June 30, 2021 **New ECD: March 31, 2022.**

June 2021: (b)(5)

(b)(5)

(b)(5) Both of these documents are reviewed by DHS Joint Requirements Council, Science and Technology, Systems Engineering, Chief Technology Office, Chief Information Office, and other DHS entities. CBP needs additional time to complete this recommendation and has a new estimated completion date of March 31, 2022.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email:
Phone: (b)(6), (b)(7)(C)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/1/2022 3:51:48 PM

To: (b)(6), (b)(7)(C)

Subject: FW: GAO Report: GAO-20-568.

FYI

From: (b)(6)
Sent: Tuesday, June 15, 2021 10:40 AM
To: (b)(6), (b)(7)(C)
Subject: RE: GAO Report: GAO-20-568.

Hi (b)(6), (b)(7)(C)

I'm so sorry, I thought I sent you an update on these two recommendations. We have closed both recommendations (#1 and #5) for GAO report number GAO-20-568, as closed and implemented. If you need anything else on this please let me know. Thank you for your help during this process.

V/R

(b)(6), (b)(7)(C)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 15, 2021 10:31 AM
To: (b)(6)
Subject: GAO Report: GAO-20-568.

Hi (b)(6)

Do you have an update on these closure requests? Per the below you was going to give me an update in May.

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone:

From: (b)(6)
Sent: Friday, April 30, 2021 10:33 AM
To: (b)(6), (b)(7)(C)
Subject: RE:

Good Morning (b)(6), (b)(7)(C)

Just a quick update, that we're processing these two recommendations, and I expect to have an answer for you next week on whether we can close Rec #1 and Rec #5 related to GAO Report: GAO-20-568.

(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, April 21, 2021 7:16 AM
To: (b)(6)
Subject:

Good Morning,

Here is the password to access the documents for recommendation 5: (b)(7)(E)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone:

From: (b)(6)
Sent: Tuesday, April 20, 2021 5:37 PM
To: (b)(6), (b)(7)(C)
Subject: RE: Hi (b)(6), (b)(7)(C)

We appreciate your help

(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, April 20, 2021 4:36 PM
To (b)(6)
Subject: RE: Hi (b)(6), (b)(7)(C)

program office and will get back with you.

From: (b)(6)
Sent: Tuesday, April 20, 2021 2:20 PM
To: (b)(6), (b)(7)(C)
Subject: Hi (b) (6), (b) (7)(C)

Is there a pass word associated with the information you sent us today? thank you

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)							
Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved
								File fields should be arrival time.	

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)								
Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)							
Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)								
Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)								
Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)									
Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)							
Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)							
Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)							
Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary		(b)(7)(E)		Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(7)(E)

Issue Type	Key	Summary	(b)(7)(E)	Status	Detailed Acceptance Criteria	Description	Resolved

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Office of Field Operations
Planning, Program Analysis, and Evaluation
Land Border Integration and Biometric Program
April 28, 2020

Action Required: Informational

Issue: Pilot Test of Third Country Nationals (TCN) Reporting Departures at Land Borders Using the I-94/Self Reporting Mobile Exit (SRME) Mobile Application

Executive Summary:

- Update to the original Issue Paper dated, January 19, 2020 with replacing Laredo Port of Entry to Blaine Port of Entry, see Issue Paper from the Seattle Field Office titled, "*Facial Recognition Technology for the Land Border,*" dated April 6, 2020.
- Other updates include timeline adjustments in response to border closures due to COVID-19.
- CBP has developed a public facing I-94/SRME mobile application to be used by the public to submit their face biometrics after exiting the U.S. The first test of this public facing mobile application was done using the CBP Mobility Subject Matter Experts (SMEs) using the test flight environment offered by GooglePlay and iTunes app stores. The test was conducted in San Diego at the San Ysidro port of entry in July 2018.
- In addition, CBP contracted with the University of Houston to conduct a white hat hackathon on testing the vulnerabilities of the geolocation and liveness detection functions within the app.
- Data was gathered from these tests and based on the feedback and technical recommendations, the public facing application was modified to improve the mobile apps user experience and close the vulnerability gaps exposed.
- OFO plans on testing the updated version of this public facing mobile application with the public exiting the U.S. from the Blaine and Champlain ports of entry, and conducting a second white hat hackathon to test the revised apps' technical vulnerabilities. These results will be combined into a final report.

Pilot Test of I-94/SRME Public-Facing Mobile Application:

- CBP I-94/SRME mobile app is designed to provide travelers a capability to use their mobile device to biometrically record their exit from the U.S. by submitting a live photo and use of location services on the phone to verify the report is being made from outside the U.S.
- To use the exit features of the mobile app, the traveler provides biographic travel document information, port of departure, and a submission of his or hers facial biometric. The application performs liveness verification of the photo, geolocation verification, retrieves the active I-94 number, and timestamps the submission with the date and time.
- The data submitted is verified in the backend and if verified, an encounter record is created and sent through TDED to be displayed as a "reported" departure encounter record in ADIS.

Limited Test by the Public – Blaine/Champlain CONOPS:

- The limited test will be conducted with actual travelers. The Seattle and Buffalo Field Offices, specifically Blaine and Champlain Ports of Entry, are planned to be the pilot sites for testing the I-94/SRME mobile app. The targeted population are travelers in need of an I-94 with an admit until date (AUD) that expires during the testing period.

- A robust public outreach campaign will be used to publicize the availability of the mobile app for use on entry and benefit of using to report one's exit in meeting compliance requirements of I-94.
- The test will start in August 2020 for up to 180 days. The duration may be adjusted to align with I-94 traveler departure patterns. It is estimated that up to 5000 travelers will be participants in this pilot.
- (b)(5)
- Data gathered from the participants will also be used to make recommendations on improving the mobile app's user experience and to address any identified vulnerabilities.
- Any subsequent possible publishing of the CBP Exit Application to iTunes and Google Play stores for full public use would not be done until the application's vulnerability risk level is deemed acceptable by CBP.

Planned Testing and Reporting Schedule:

- (b)(5)
- Adjustments will be made to the schedule based on impacts of COVID-19 directives.

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

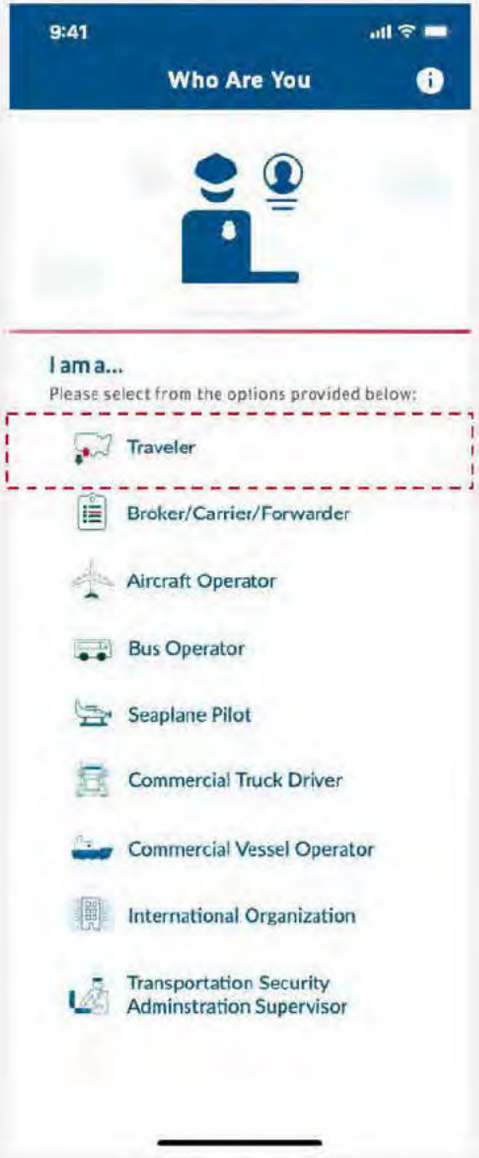
Withheld in full pursuant to (b)(5), (b)(7)(E)

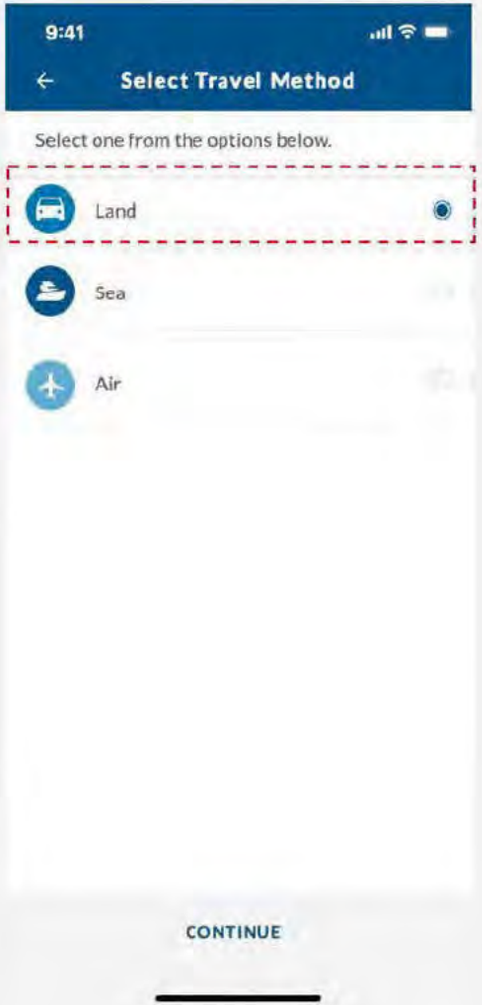
CBP One™

September 29, 2021

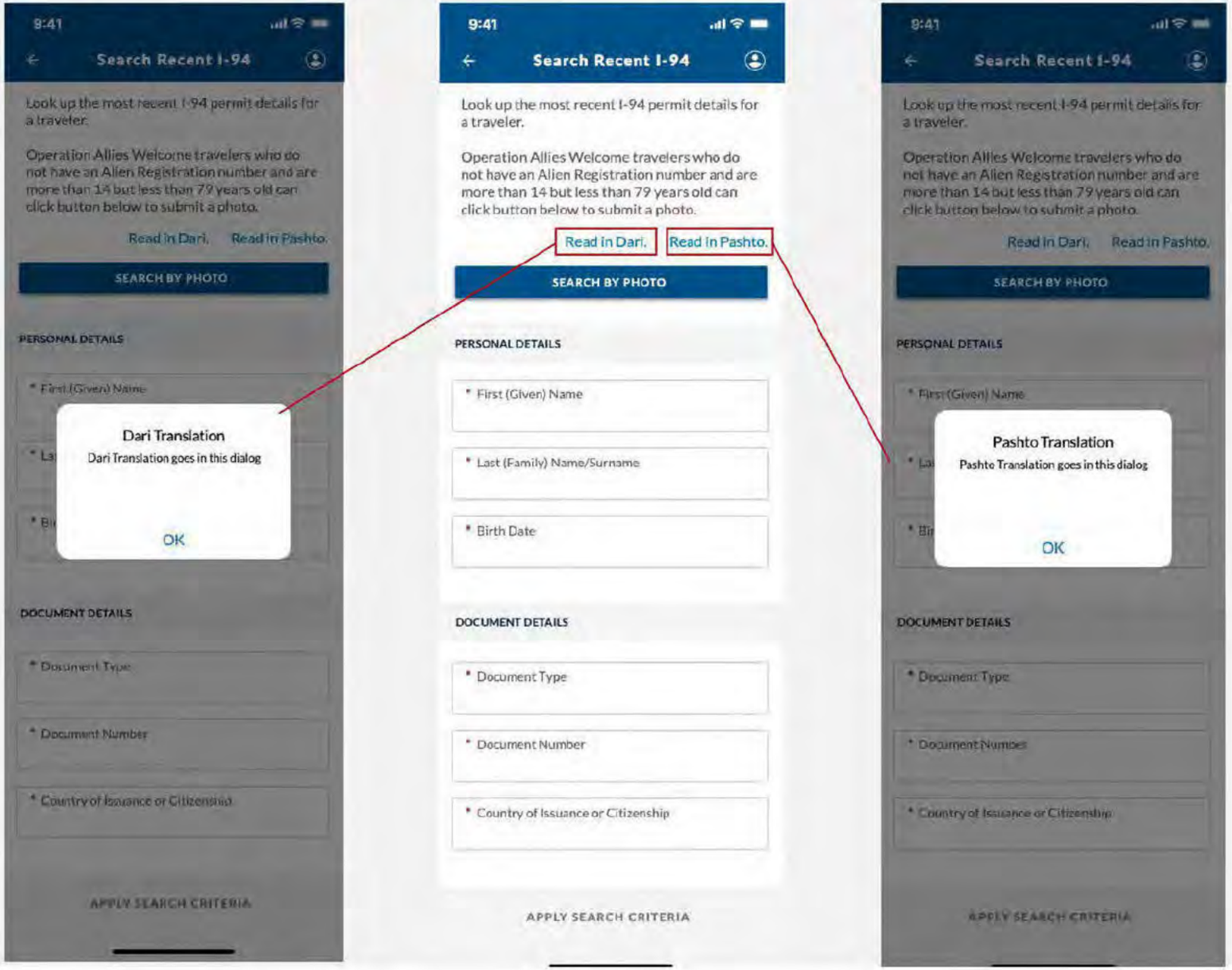
U.S. Customs and Border Protection

CBP One – **Traveler: Land > Get My Recent I-94**

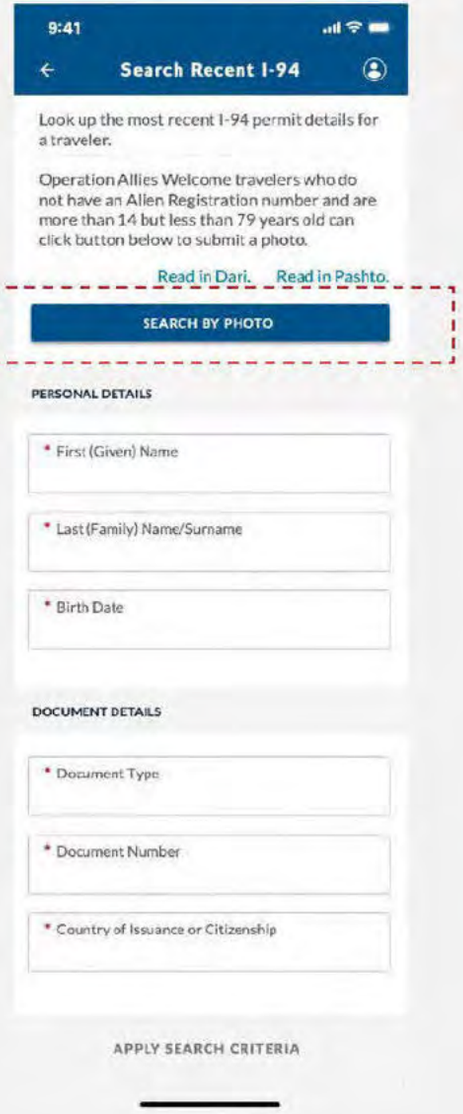


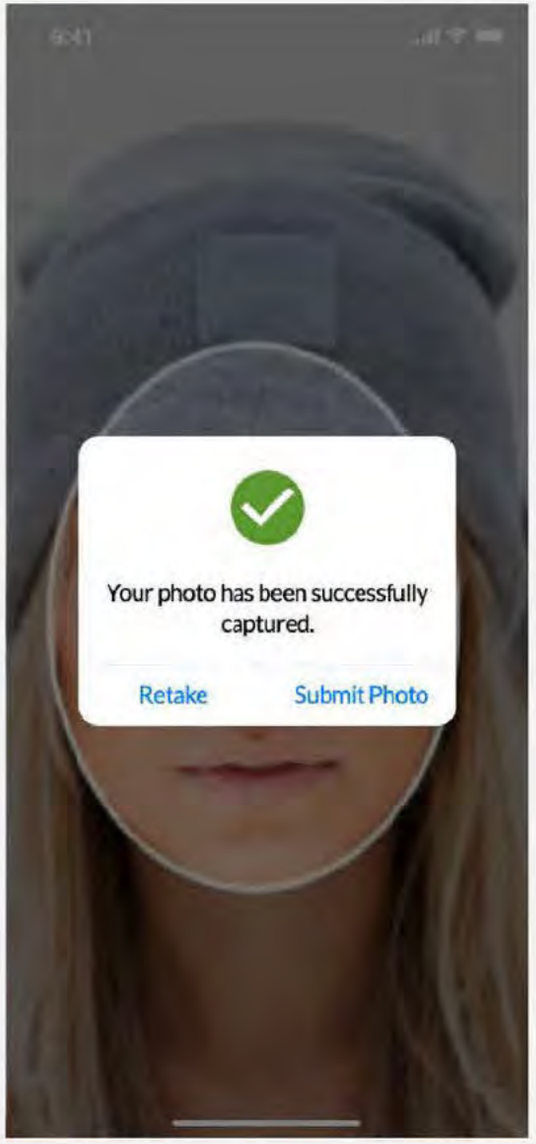


CBP One – **Traveler: Land Get My Recent I-94 > Translations**

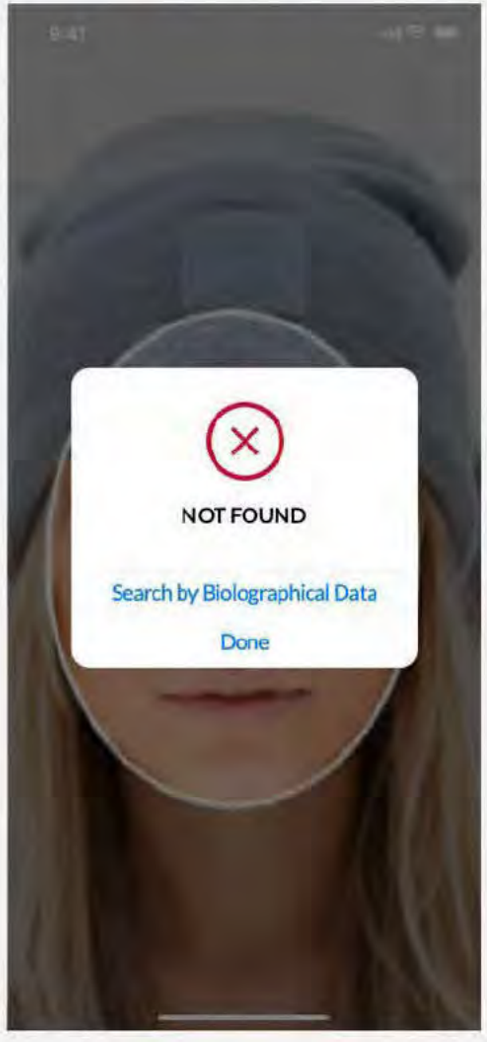


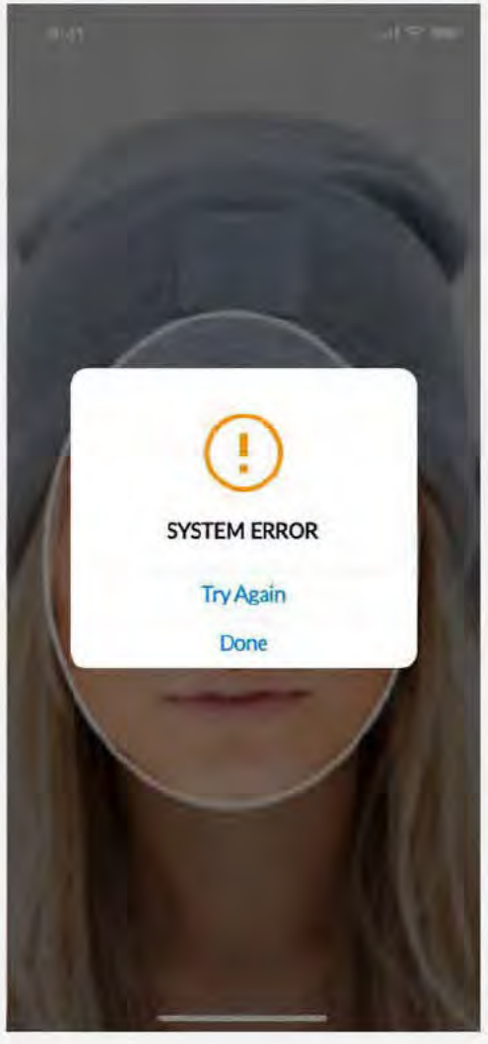
CBP One – **Traveler: Land Get My Recent I-94 > Take Photo**



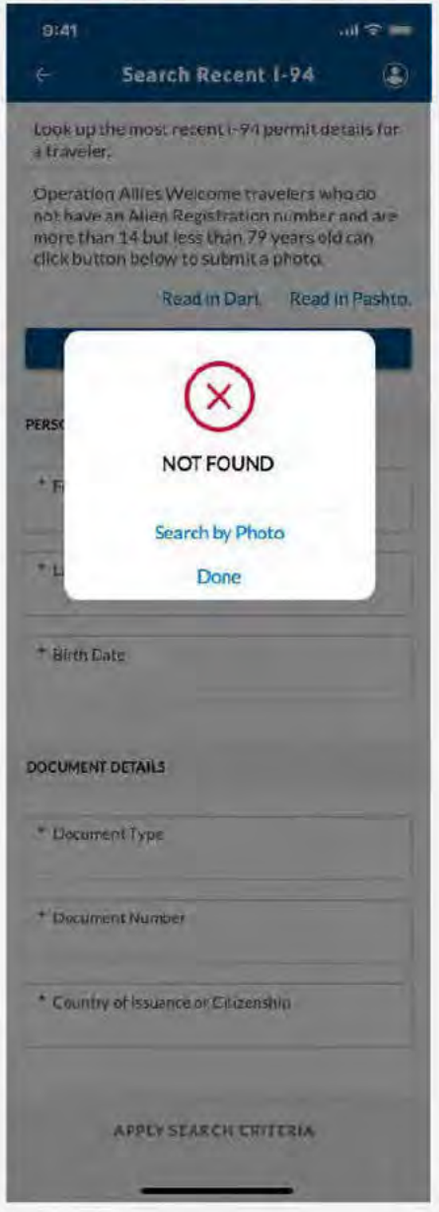


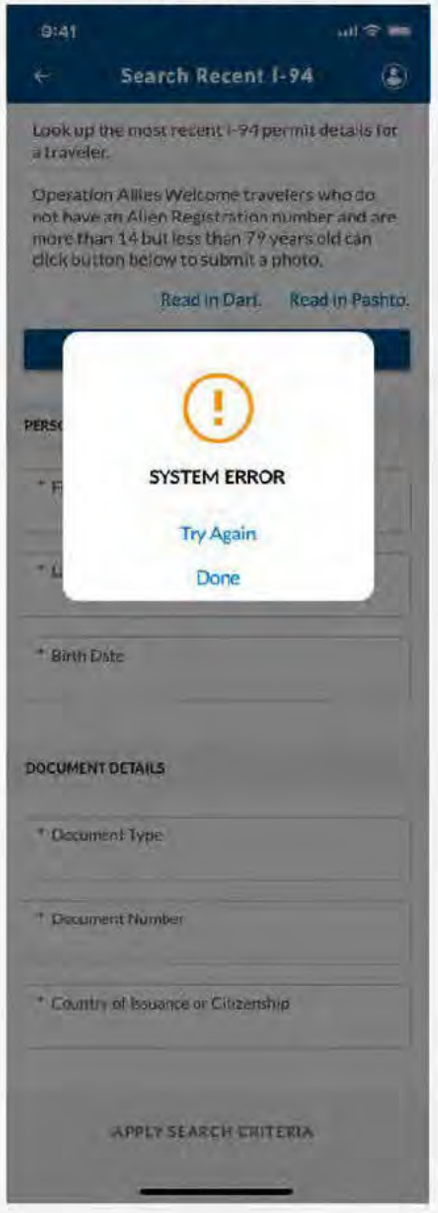
CBP One – **Traveler: Land > Get My Recent I-94 > Query Results after Photo**





CBP One – **Traveler: Land - Get My Recent I-94 > Search By Biographical Data**





Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

U.S. Customs and Border Protection: Evaluating Possible Bias

Executive Summary:
U.S Customs and Border Protection (CBP) mitigates potential algorithmic bias in the Traveler Verification Service (TVS)[1] in a number of ways.

First, CBP relies on test, research, and evaluation activities performed by National Institute of Standards and Technology (NIST) and U.S. Department of Homeland Security Science and Technology Directorate to inform the procurement process for algorithms. Additionally, CBP evaluates the algorithms used in TVS using diverse training sets.[2] CBP also continuously monitors the biometric matching service and conducts a variety of statistical tests to enhance the effectiveness of the algorithms and minimize any possible bias impact.

Additionally, as described in more detail below, CBP is further collaborating with NIST to perform an independent and comprehensive scientific analysis of CBP's operational face-matching performance, including impacts of traveler demographics and image quality.

Methodology for CBP's Continued Monitoring and Bias Assessment:
CBP requires that all carriers submit Advanced Passenger Information System (APIS) data for flights to and from the United States. Amongst the data submitted is gender, date of birth, citizenship, and country of residence. Using this data, CBP has conducted extensive statistical analysis (chi squared independence tests) to determine whether traveler demographics (age, gender, and nationality) affect facial recognition match rates. CBP does not collect race/ethnicity information, nor is this information included in the APIS manifest. As a result, CBP uses citizenship as a proxy for this data.

Performance Results for CBP's Continued Monitoring Bias Assessment and Continued Improvements:
CBP's own analysis found a negligible effect in regard to its biometric matching based on citizenship, age, and/or gender while achieving a technical match rate (TMR) in the high 90 percentile.[3] TMR defines how well the algorithm performs for each traveler who could be identified. As of December 2018, TMR continues to be at a steady state, above 98%. However, U.S. citizens tend to have fewer photos and older photos in government databases, which may affect the TMR. CBP continues to make significant improvements to the algorithm and has increased the number of the exit operations, which has led more data, and thus to a substantial reduction in the initial gaps in matching for age and gender. Following these improvements, travelers between ages 26 and 65 match only slightly better than "young" (ages 14 to 25) travelers (by 0.3%) and "old" (ages 66 to 79) travelers (by 0.1%), compared to during initial

---

[1] Additional information about CBP's TVS can be found in the DHS/CBP/PIA-056, Privacy Impact Assessment for the Traveler Verification Service, issued Nov. 14, 2018, available at https://www.dhs.gov/sites/default/files/publications/privacy-pia-cbp030-tvs-november2018_2.pdf

[2] Across most flights processed, CBP was able to encounter diverse age and gender. CBP also worked to ensure diversity of citizenship across flights tested. For example, CBP selects flights to airports in various regions of the world to increase the likelihood for diversity of citizenship. This ensured that the matching algorithm tested on a diverse training set.

[3] Based on June 2017 – November 2018 CBP Air Exit data from biometric exit locations: JFK, MIA, IAH, HOU, ORD, SEA, SFO, LAS, DTW, LAX, IAD, MCO, ATL, BOS, and FLL.

tests, when they matched better by 2.8% and 8%, respectively. Similarly, women currently match slightly better than men (by 0.2%), compared to initial tests, where men matched better (by 1.7%). Much of the bias seen in earlier flights also relates to much lower flight volume in the pilot period. It should be noted that volume of flights was much lower during the initial pilot period.

The performance of CBP's TVS continues to improve over time due to technical, operational, and procedural advancements, including threshold adjustments and testing multiple vendors. NIST concluded during its 2018 Face Recognition Vendor Test[4] that there have been massive improvements in the accuracy of face recognition algorithms over the last five years (2013-2018) and CBP continues to test and employ new, more refined, algorithms. CBP has issued various updates to the matching algorithms, which increase the algorithm's ability to create biometric templates from non-frontal images taken during the U.S. entry or exit process.

CBP has also enhanced the photo selection process to ensure the most recent photos are selected. CBP also has enhanced the manner in which the galleries are populated, by utilizing biographic information to build the galleries, which reduces the number of travelers with no photos and improves the accuracy of the system.[5] Furthermore, as CBP continues and expands its usage of TVS, personnel using the technology become more aware of the optimal camera positions to ensure better images and increase the traveler throughput. Some cameras are also now equipped with multiple lenses to capture images for various angles, which may increase photo quality depending on the height of the traveler.

Future Assessments:
CBP is further collaborating with NIST to perform an independent and comprehensive scientific analysis of CBP's operational face matching performance. During these tests, CBP is sharing facial images of certain in-scope travelers[6] captured during technical demonstrations, which will enable NIST to conduct a scientific analysis CBP's performance. By analyzing the image, NIST will be able to identify impacts due to image quality and traveler characteristics. This will help independently verify results and provide a more in-depth study controlling for various factors.

Upon analyzing a comprehensive data set, NIST will provide objective recommendations regarding matching algorithms, optimal thresholds, and gallery creation, optimizing face matching performance for large-scale traveler identification at air, land and sea ports of entry. CBP will continue to actively monitor and refine the performance of this process and associated algorithms in order to make improvements, minimize potential bias impact, and ensure the high accuracy of facial matching for all travelers.

---

[4] *See* NIST Interagency Report 8238, available at https://nvlpubs.nist.gov/nistpubs/ir/2018/NIST.IR.8238.pdf.
[5] A 2010 NIST evaluation of face recognition showed that "[w]hen all prior images of a person are enrolled under one identity, accuracy improvements in both verification and identification trials are realized. The value of multiple images increases with the number of images." See NIST Interagency Report 7709, at 40, available at https://www.nist.gov/publications/report-evaluation-2d-still-image-face-recognition-algorithms.
[6] "In-scope" aliens are any aliens that are not exempt from a requirement to provide biometric identifiers to confirm their admissibility pursuant to 8 CFR 235.1(f)(ii) or, at specified airports, their departure pursuant to 8 CFR 215.8(a)(1). CBP may share these photos only in accordance with applicable law and consistent with the relevant Privacy Act System of Records Notice (SORN). More information is available at www.dhs.gov/privacy-impact-assessments.

Vehicle Biometric – Anzalduas Phase I Post Technology Demonstration

Evaluation

March 2019

Vehicle Biometric Anzalduas Technology Demonstration Goals

Can we successfully capture facial biometric quality photos of POV travelers?

Goal of Field Test

Determine the feasibility of cameras in vehicle lanes to capture images for biometric matching.

Success of the technology was defined as rate of image capture for all occupants in a vehicle.

Successful image capture was measured as template was produced from the captured image for use in biometric matching by TVS.

About the Technology Demonstration

Comparison of two vendors for 3+ months

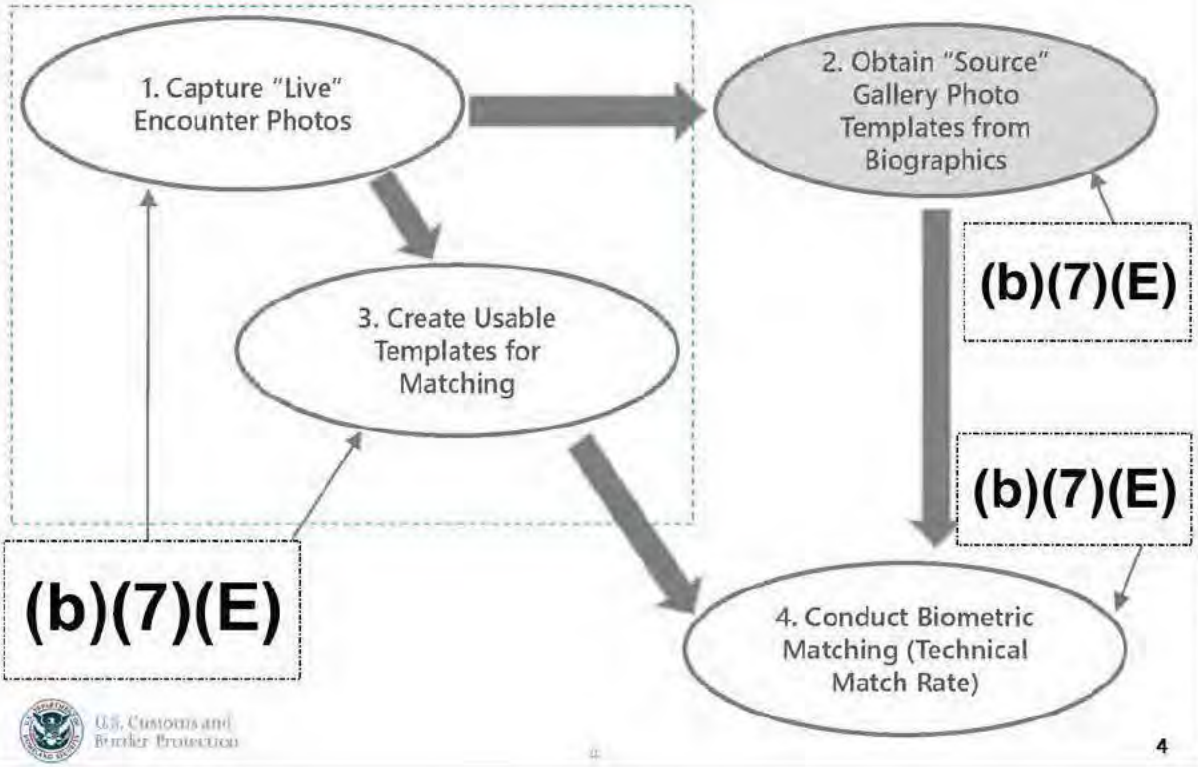
- Anzalduas, TX
  - (b)(4) cameras in 1- IB lane / 1- OB lane
  - (b)(4) cameras in 1- IB lane / 1- OB lane
- Duration:
  - (b)(7)(E) September 1 – December 16, 2018[1]
  - (b)(7)(E) December 17, 2018 – February 28, 2019[2]

(b)(7)(E)

(b)(4), (b)(7)(E)

About Biometric Matching and Vendor Comparison

Encounters, Galleries, and Matching



U.S. Customs and Border Protection

Bottom-line Up Front

When the driver is the only traveler, template creation rate increases to 90%

Single Traveler Template Rate
90%

Vendor camera technology is capable of producing "match-quality" templates for 78% of all travelers in a vehicle.

Template Rate
78%

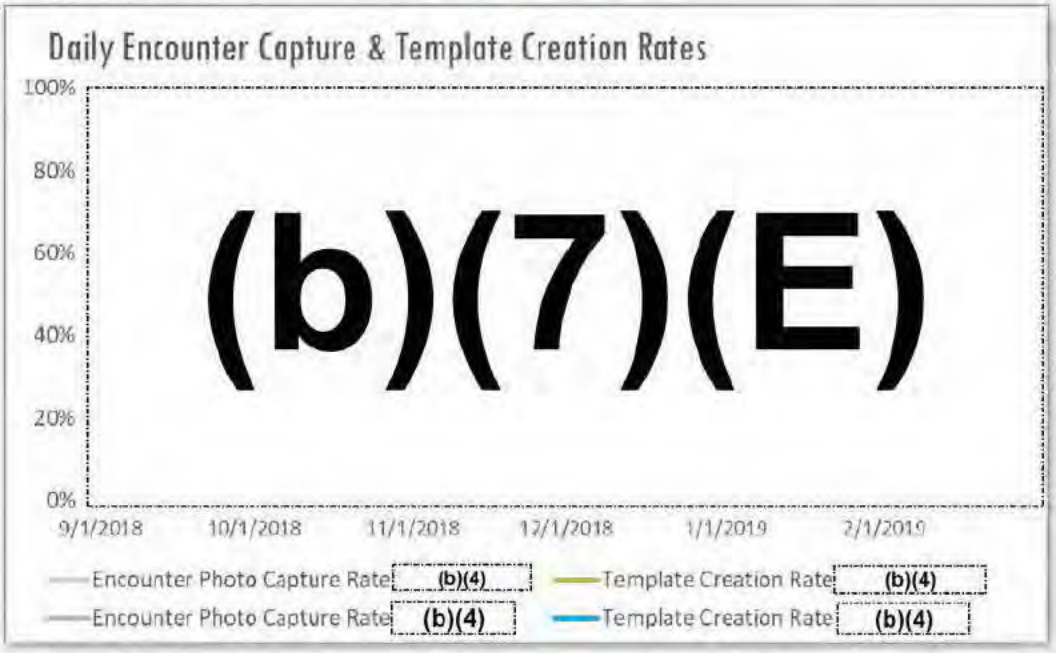
A or B (b)(4) outperformed (b)(4) by a significant margin as measured by template creation rate

(b)(4)	
78%	56%

Vendor camera technology work better than anticipated; One vendor distinguished themselves.

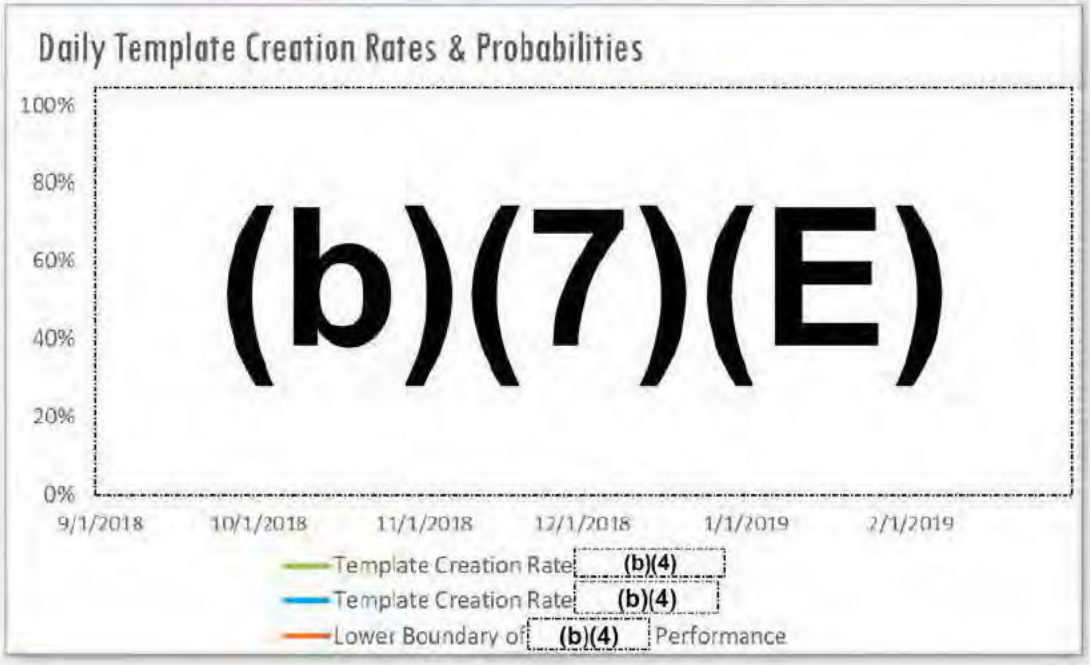
Capturing Traveler Photos & Creating Templates Rates by Vendor

(b)(4) struggles to meet capture requirements



Statistical Look at Template Rate Comparison between Vendors

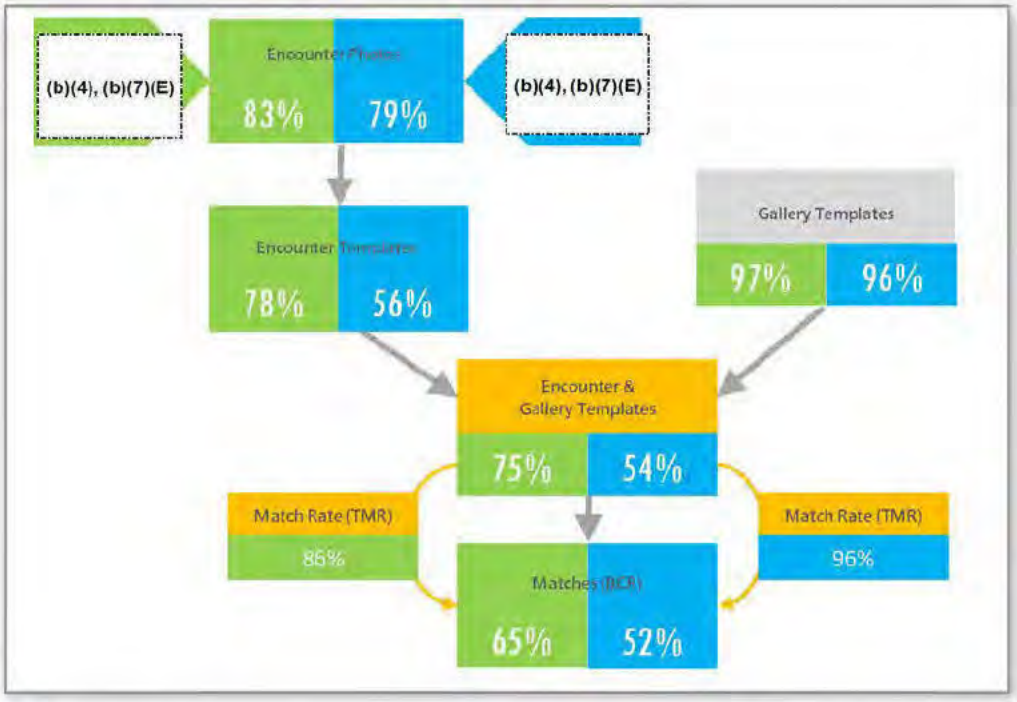
(b)(4) best day creating templates is below (b)(4) (b)(7)(E) percentile line



U.S. Customs and Border Protection

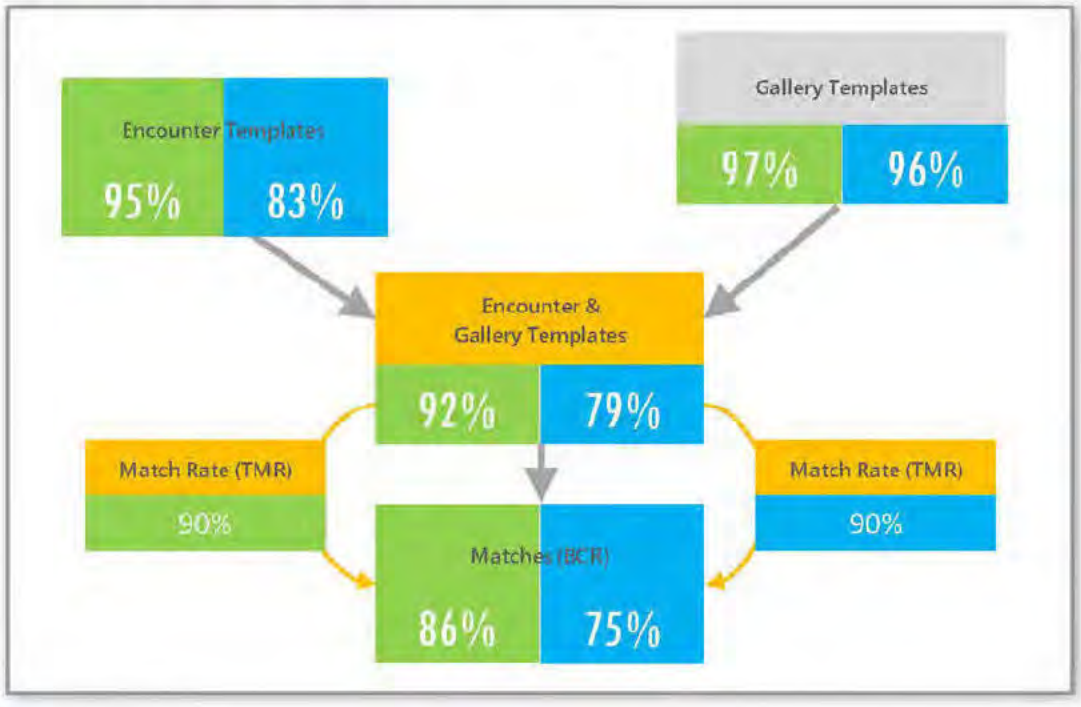
Results at a Glance – All Travelers

Summary Encounters, Galleries, and Matching



What are the results in the front seat?

Dramatically improved Template Creation Rate



Technical Issues Encountered

(b)4), (b)(7)(E)

Other Observations

(b)(4), (b)(7)(E)

Examples of Images Captured by Vendor

(b)(4)

(b)(6), (b)(7)(C)

(b)(4)

(b)(6), (b)(7)(C)

Challenges Moving Forward

- Integration with TVS to increase the gallery photos available for matching; search other photo sources
- VPC redesign and integration with TVS.
- Similar to PED, begin analyzing gallery options for the vehicle environment
- Further improve the automatic photo capture rates to minimize need for officer's to capture photos of travelers at the booth.
- Share non-USC "frequent crosser" encounter photos with IDENT, for inclusion in future matching galleries
- (b)(7)(E)

Some challenges can be addressed without equipment in the field

Next Steps

Ending Phase I Pilot Activities and Phase II Activities

(b)(7)(E)

CBP One™

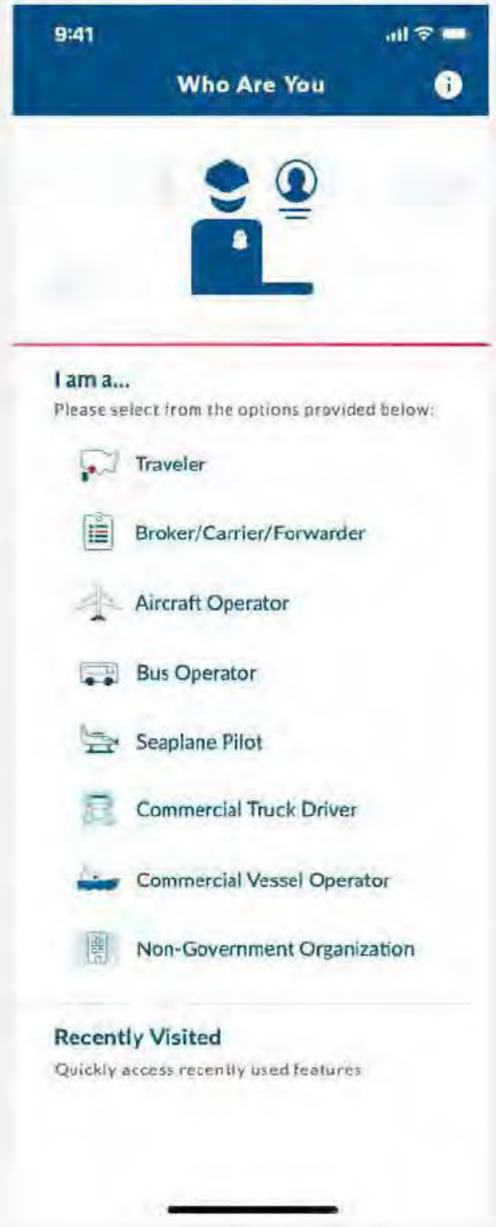
Submit Advance Information

U.S. Customs and Border Protection

CBP One Background

U.S. Customs and Border Protection

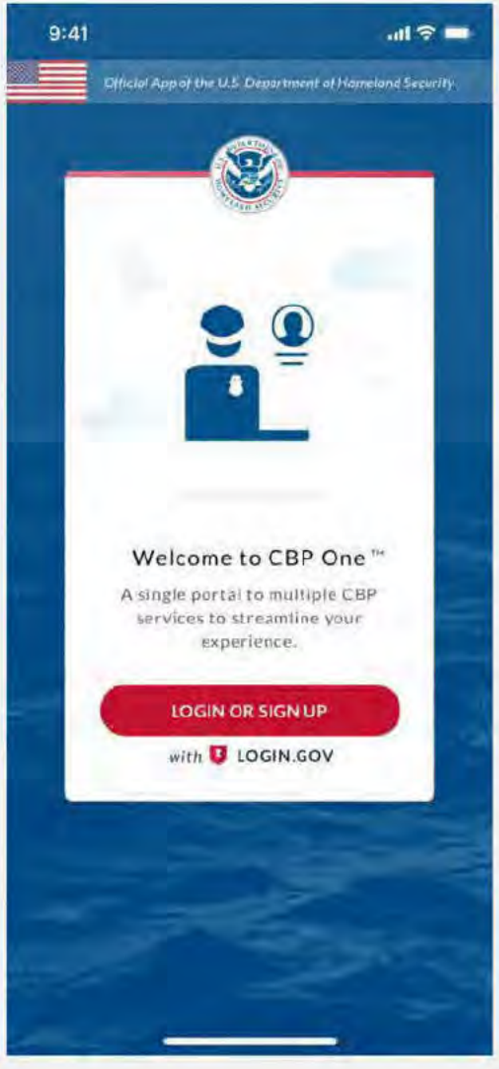
- U.S. Customs and Border Protection built a mobile application to serve as the single point of entry for travelers and stakeholders to access CBP mobile applications and services.
- Through a series of intuitive questions, the app will guide each type of user to the appropriate services based on their particular needs.
- CBP One is currently available on the Apple App Store and Google Play Store with limited functionality.

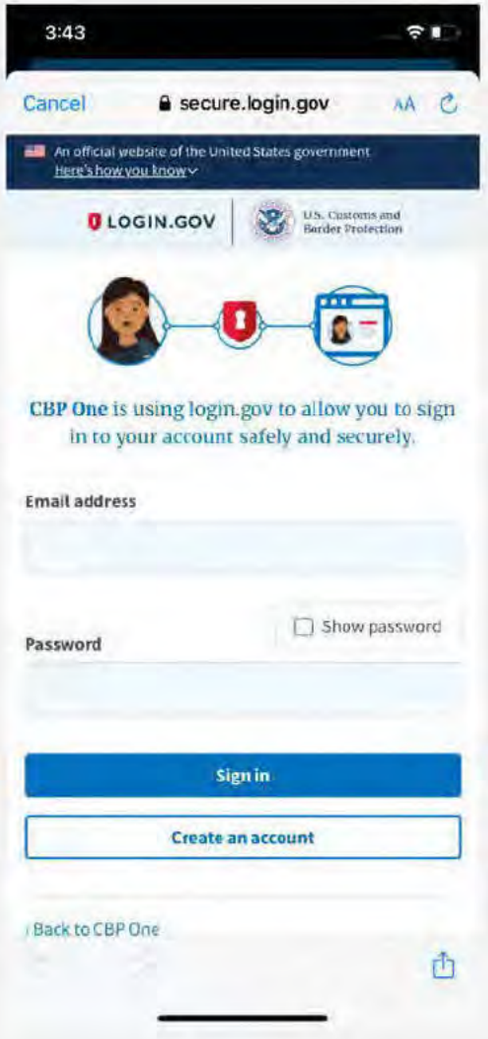


To access CBP One, users are required to Sign up or Login through Login.gov

U.S. Customs and Border Protection

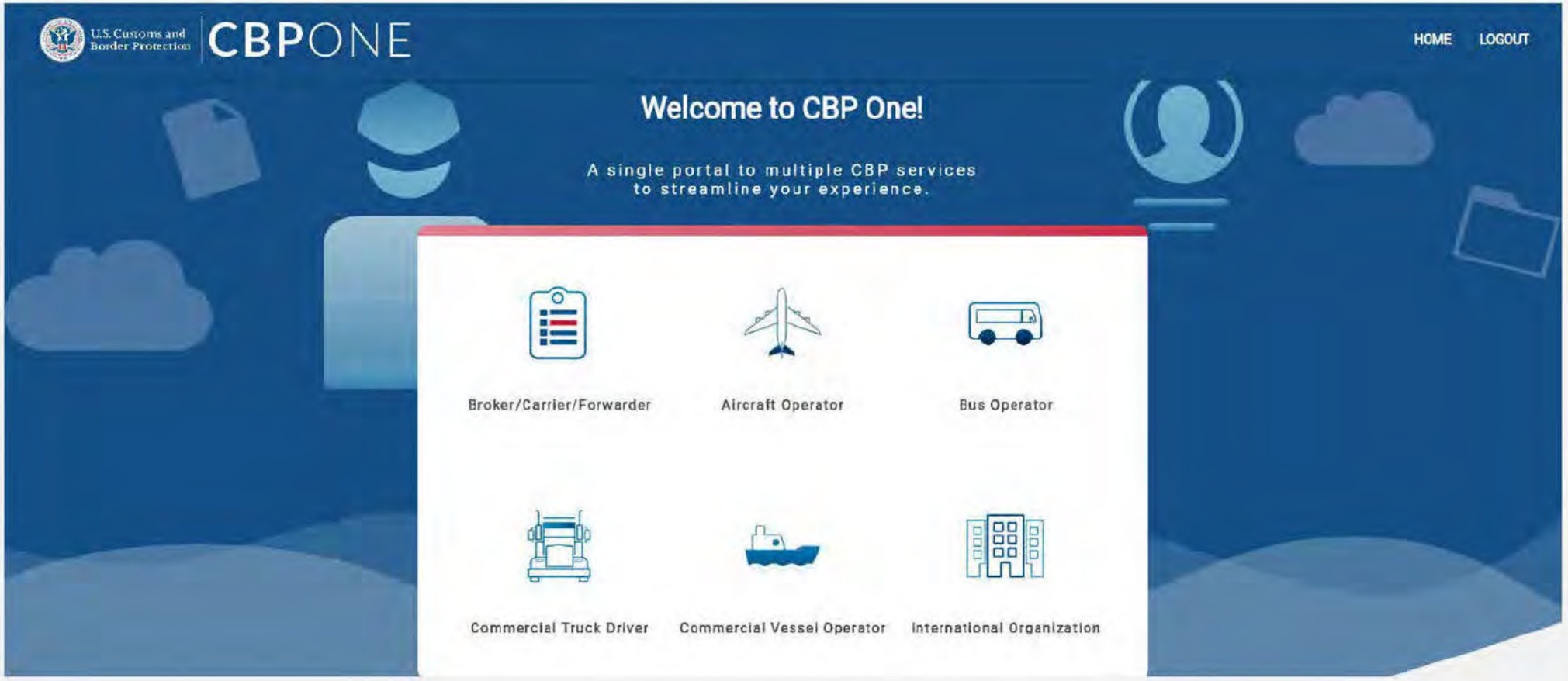
1. NGOs will use their **organization email** to sign up or login through Login.gov.
    - A personal email **will not** work
2. Enter a password.
3. Login requires a two-step authentication. NGOs will select one or more authentication method such as:
    a. **More secure**
        - ✓ Security Key
        - ✓ Authentication application
    b. **Less secure**
        - ✓ SMS/Text messages
        - ✓ Backup codes





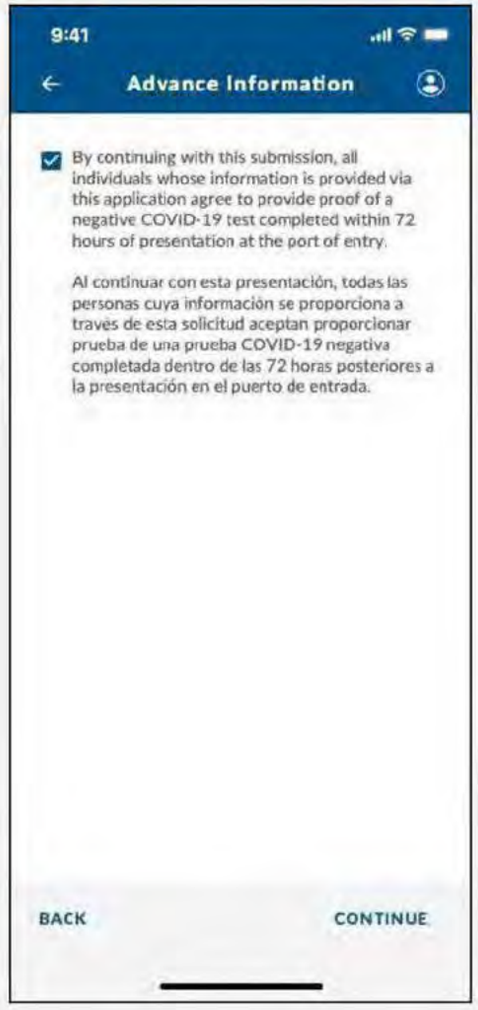
U.S. Customs and Border Protection

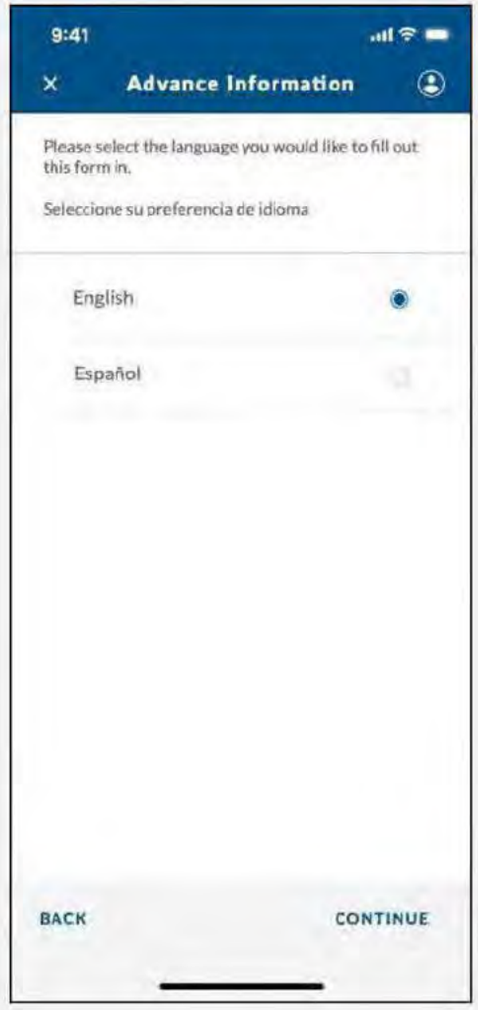
Submit Advance Information Mobile



Submit Advance Information Mobile

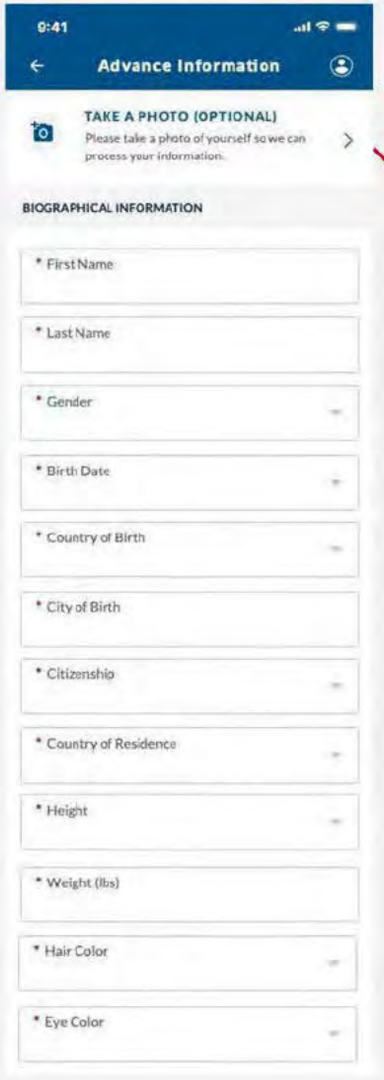
U.S. Customs and Border Protection

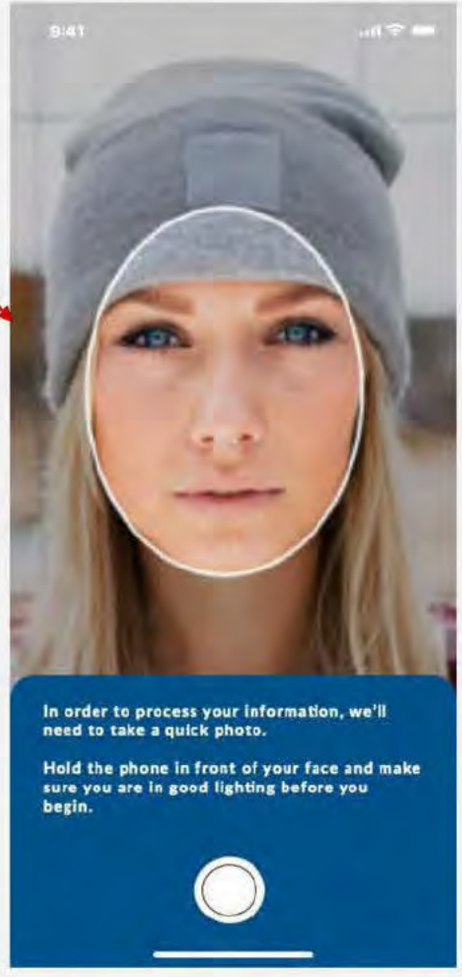


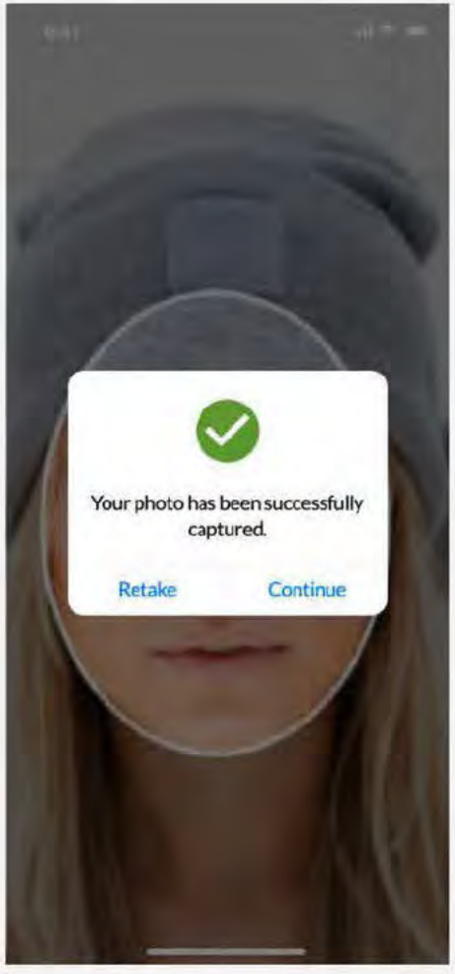


Individual Information

U.S. Customs and Border Protection

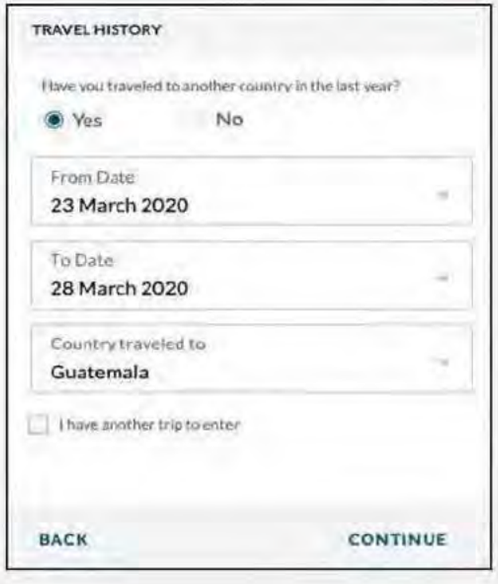






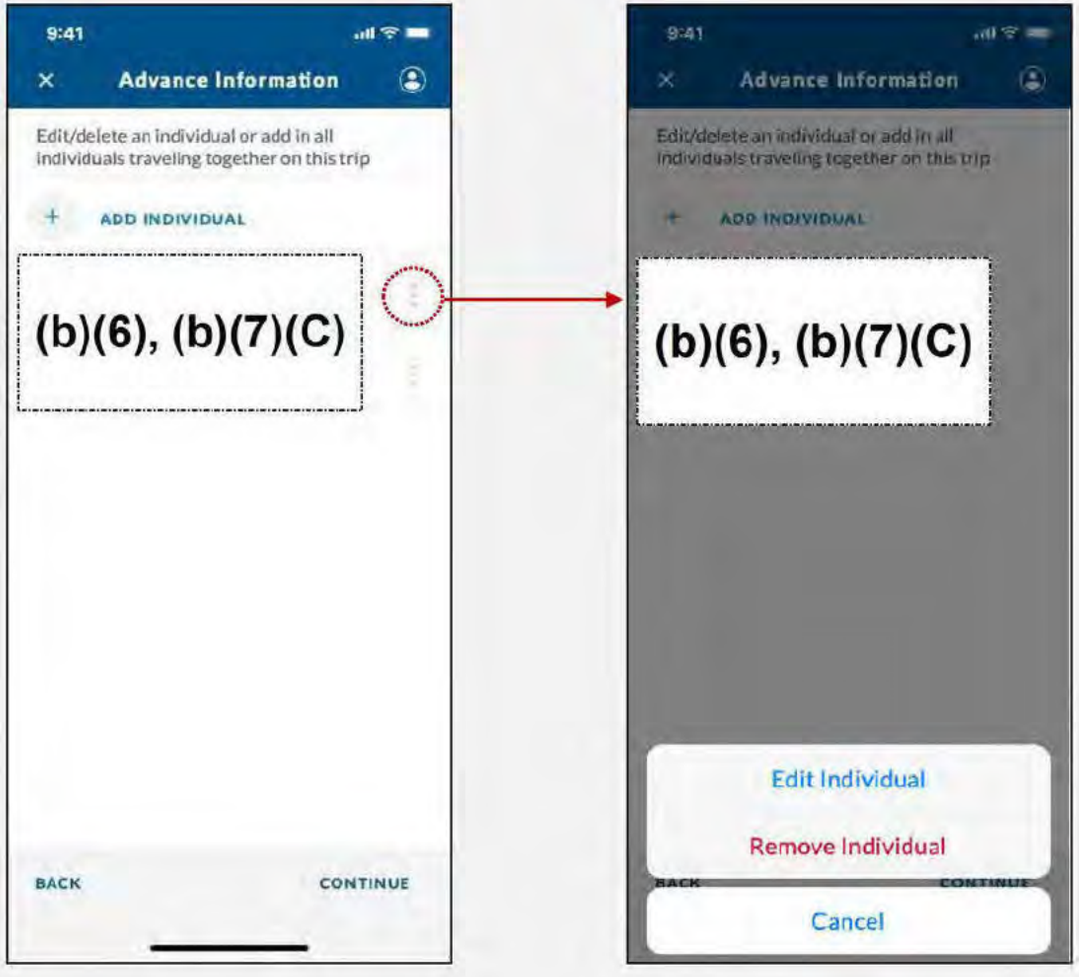
Individual Information

U.S. Customs and Border Protection



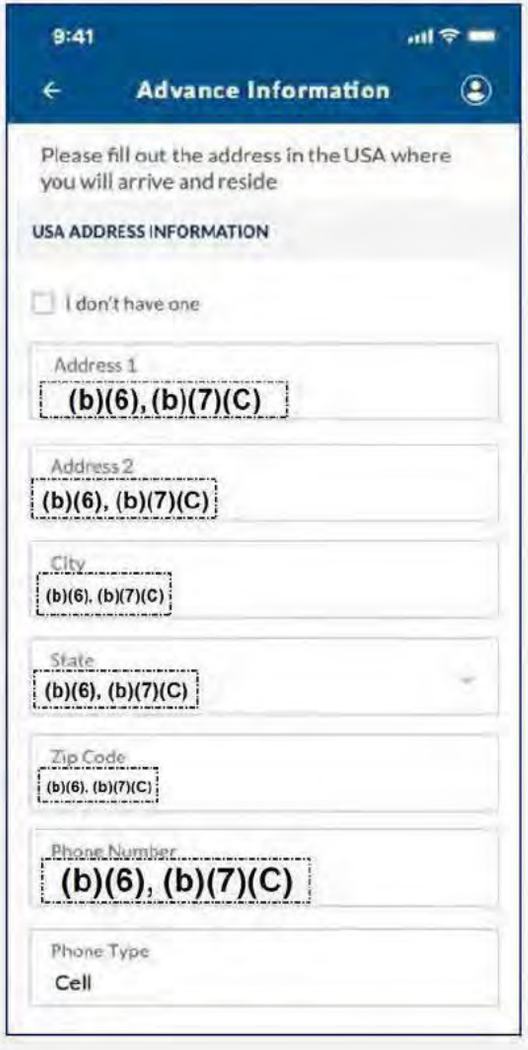
Add Additional Individuals Who Share Common Addresses

U.S. Customs and Border Protection



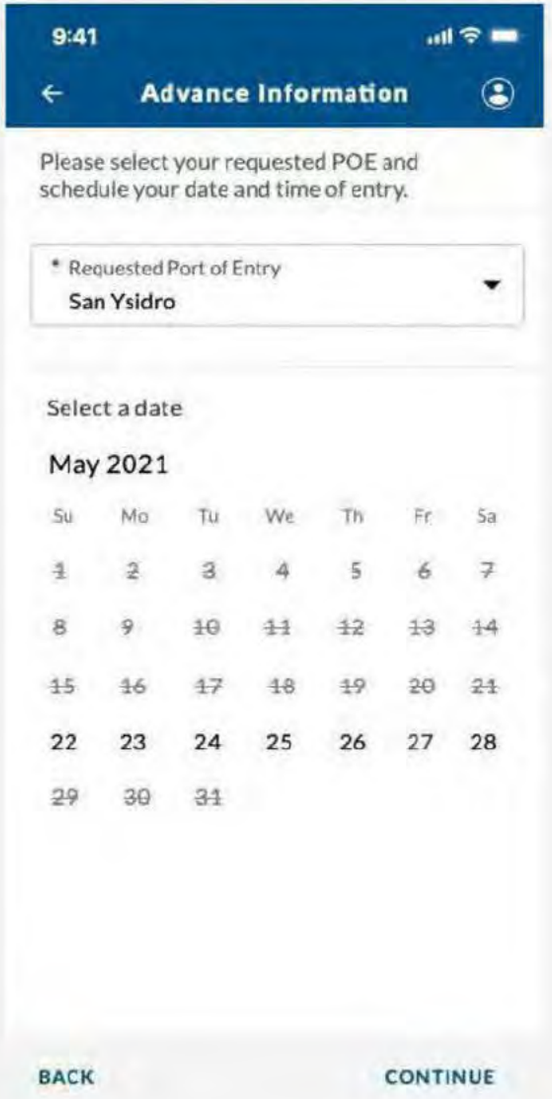
Contact Information

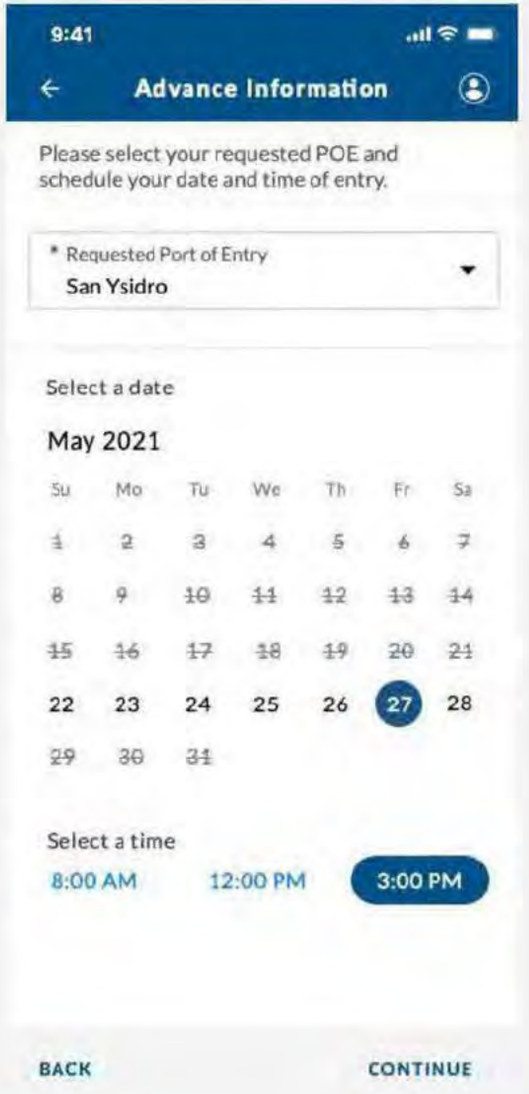
U.S. Customs and Border Protection



Requesting Day POE and Day/Time for Presentation

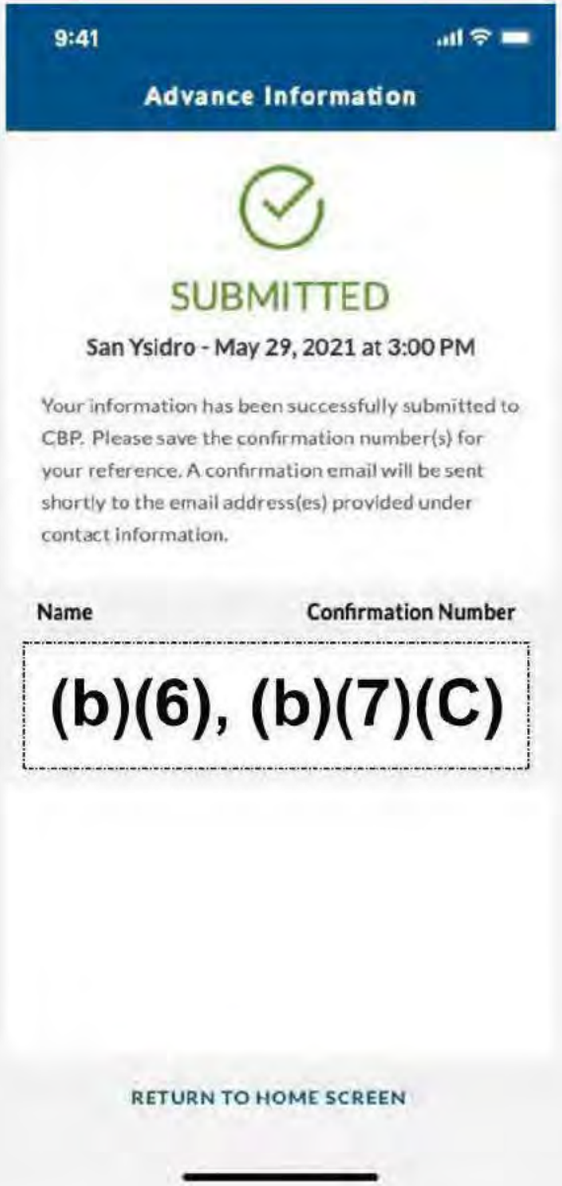
U.S. Customs and Border Protection





Confirmation Page with Unique Record ID - POE Day and Time

U.S. Customs and Border Protection



U.S. Customs and Border Protection

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One™

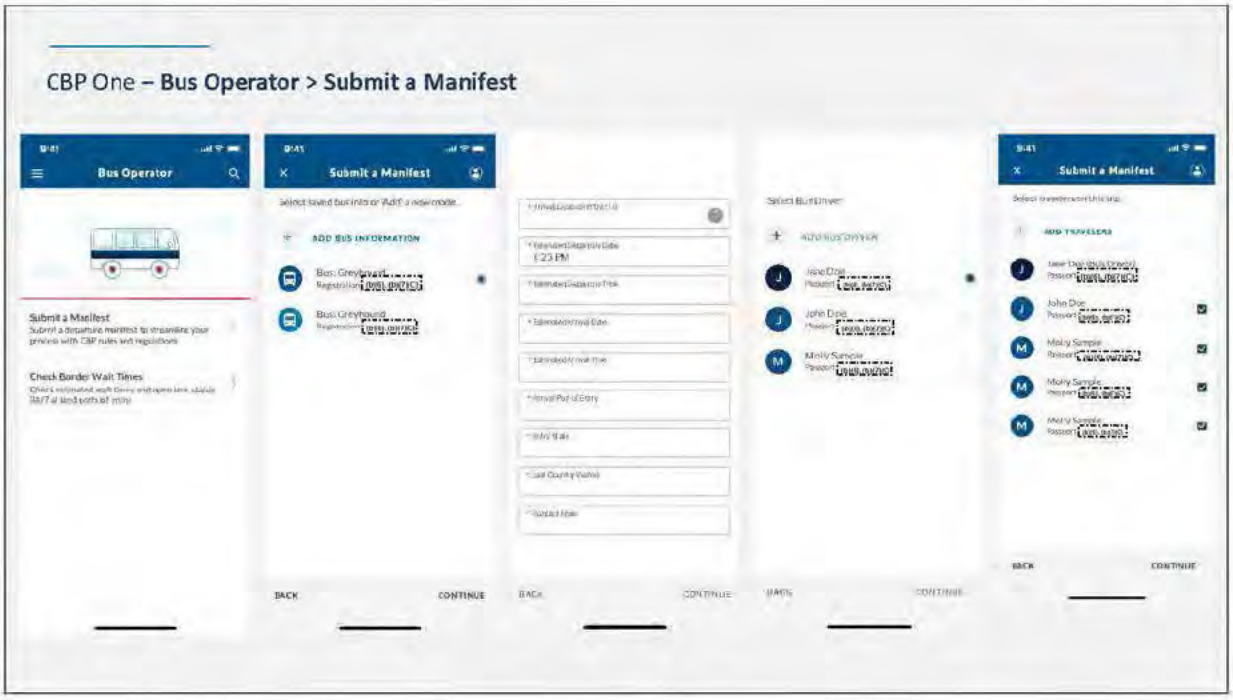
ROAM Workflows

May 24, 2021

U.S. Customs and Border Protection

Bus Operator ROAM Flow (Submit a Manifest)

CBP One – **Bus Operator > Submit a Manifest**



CBP One – **Bus Operator >**
Submit a Manifest Continued

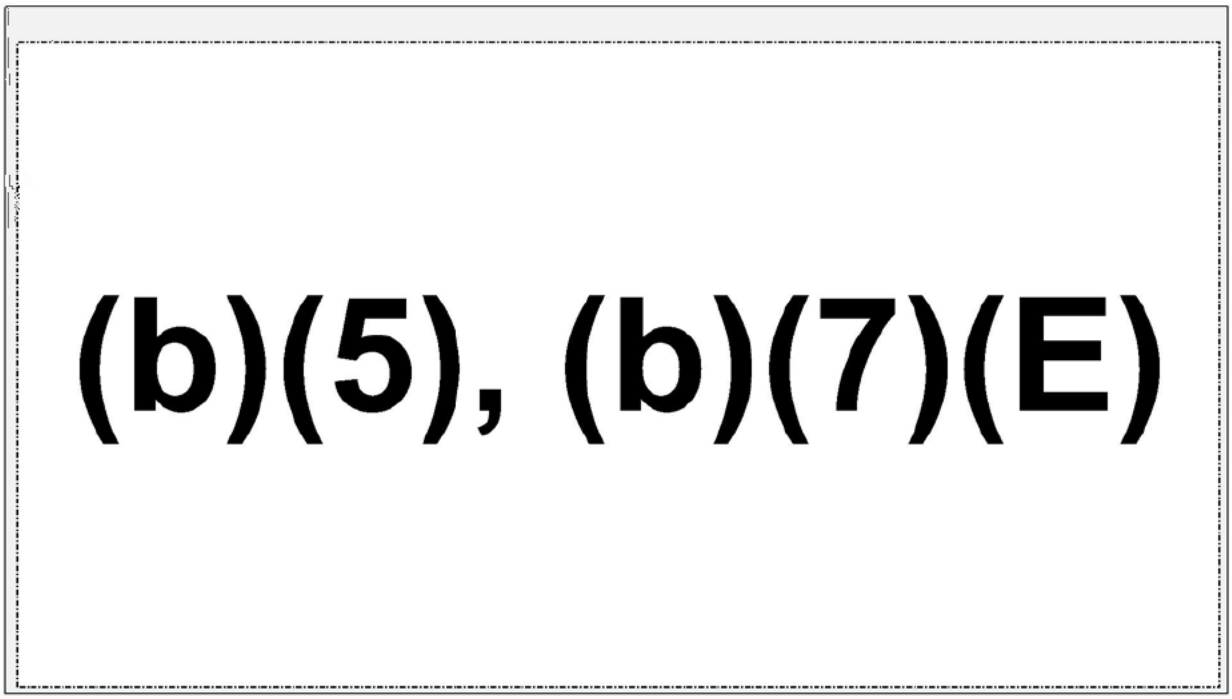
CBP One™

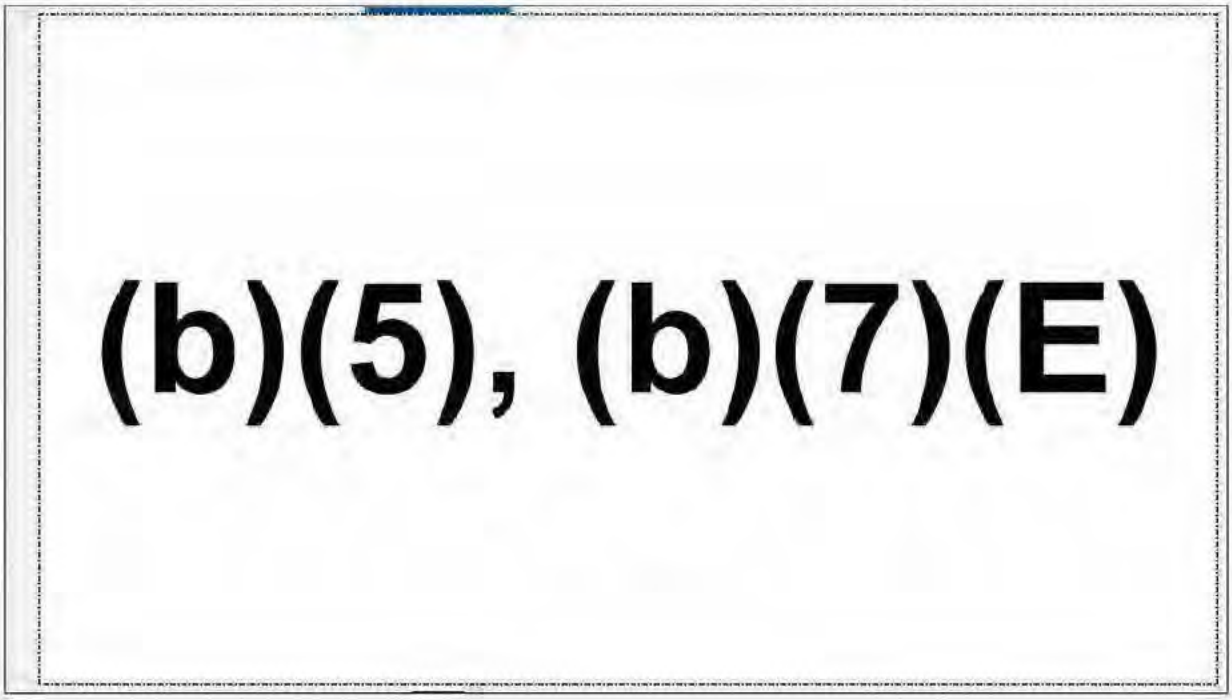
(b)(5)

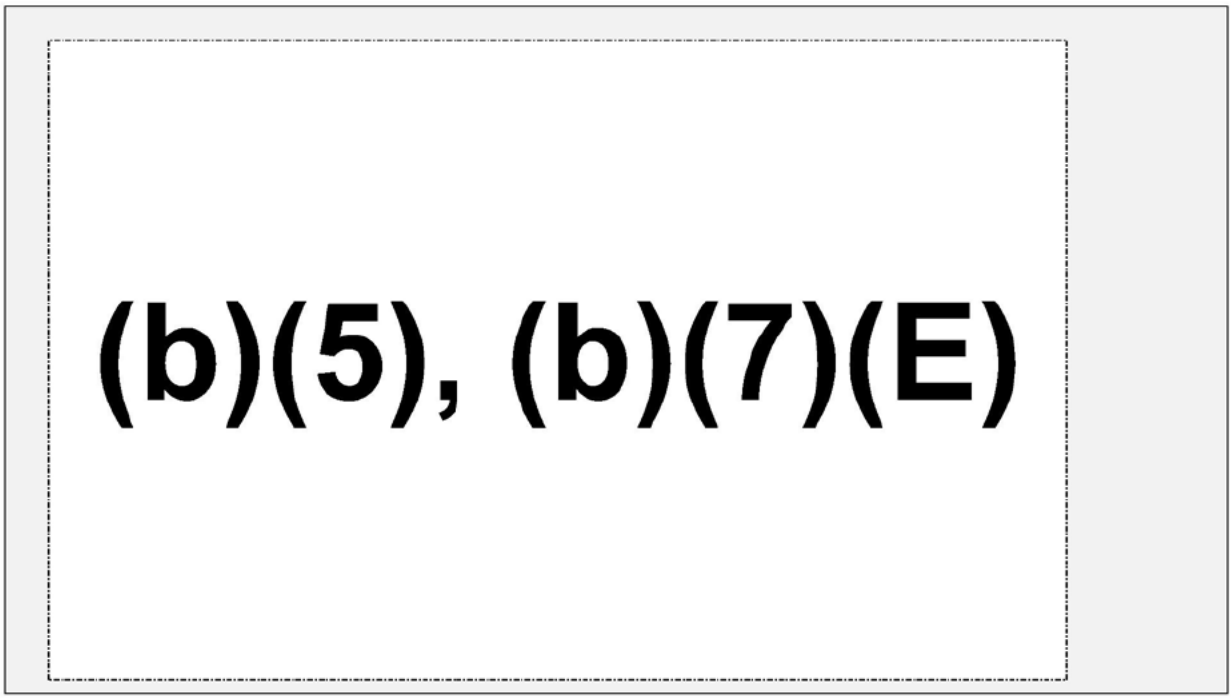
October 20, 2021

U.S. Customs and Border Protection

(b)(5), (b)(7)(E)

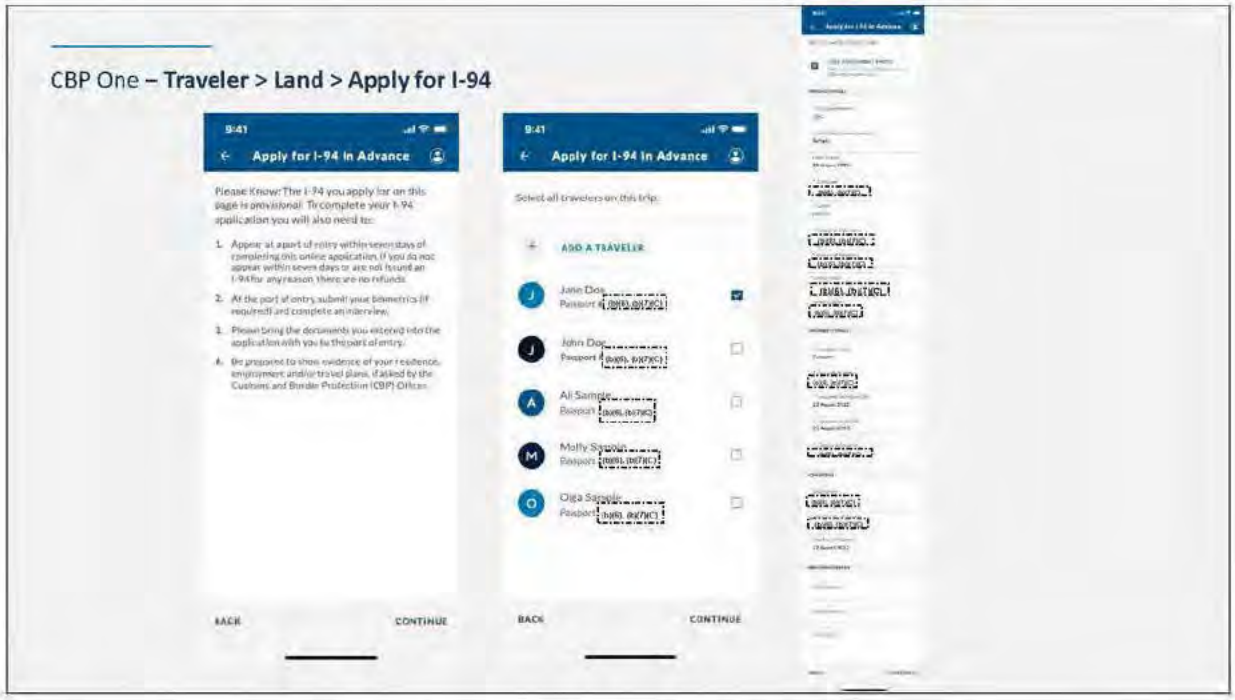

(b)(5), (b)(7)(E)


(b)(5), (b)(7)(E)

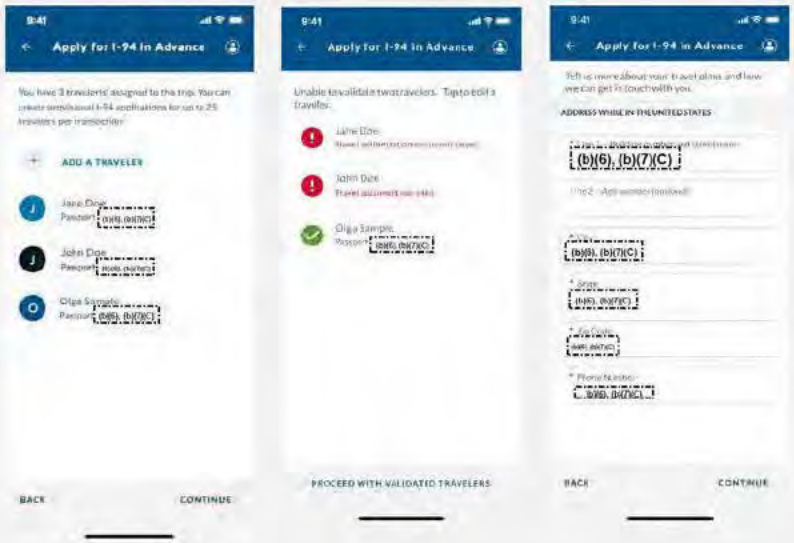

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

CBP One – **Traveler > Land > Apply for I-94**



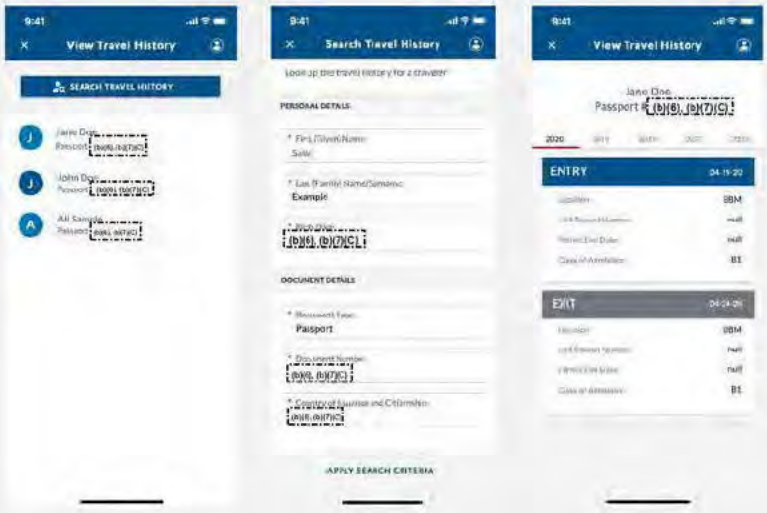
CBP One – **Traveler > Land > Apply for I-94**



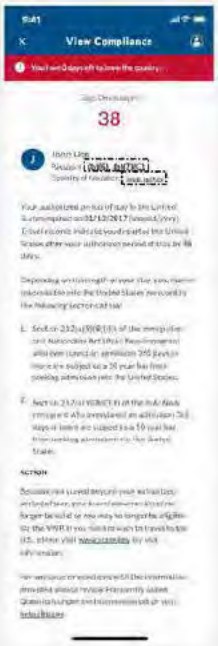
CBP One – **Traveler > Land > Apply for I-94**

CBP One – **Traveler > Land > Get Recent I-94**

CBP One – **Traveler > Land > View Travel History**



CBP One – **Traveler > Land > View Compliance**



Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Office of Field Operations
Planning, Program Analysis, and Evaluation
Land Border Integration and Biometric Program
May 11, 2020

Action Required: Information Only

Issue: Reporting Departure of Third Country Nationals (TCN) at Land Borders using Self Reporting Mobile Exit (SRME) Mobile Application and Other Proposed Technical Demonstrations

Executive Summary:

- Beginning in December 2017 and concluding September 2018, CBP deployed the capability to biometrically capture the departure of Third Country Nationals using BE-Mobile when departing at the land borders to all land Field Offices including 74 land border ports of entry.
- CBP also developed a public facing mobile application for use by the public to submit their face biometrics after exiting the U.S. The first test of this public facing mobile application was done using the CBP Mobility Subject Matter Experts (SMEs) using the test flight environment offered by GooglePlay and iTunes app stores. The test was conducted in San Diego at the San Ysidro port of entry in July 2018.
- In addition, CBP contracted with the University of Houston to conduct a white hat hackathon on testing the vulnerabilities of the geolocation and liveness detection functions within the app.
- OFO plans on testing the updated version of this public facing mobile application with nonimmigrant valid I-94 holders exiting the U.S. from the Blane, WA port of entry, and again conducting a white hat hackathon to test the revised apps' technical vulnerabilities.
- Additionally, CBP plans to demonstrate the capability of taking photos of departing travelers using CBP equipment and comparing those photos to documents on file within government holdings.

Current Status and Highlights:

BE-Mobile Pulse and Surge

- With the rollout of BE-Mobile completed, OFO can increase the usage of the BE-Mobile Land application running on CBP mobile devices (**(b)(7)(E)**) as well as desktop/tablet (**(b)(7)(E)**) meet any pulse and surge operations requested by OFO management.
- **(b)(5), (b)(7)(E)**
- **(b)(7)(E)**
  - **(b)(7)(E)** using this solution could be a short-term option if expanded collection of biometric exits is needed.
- In 2018, OFO surveyed the land ports to determine outbound reporting for TCNs.
  - **(b)(7)(E)**

~~For Official Use Only~~

- (b)(5)

Pilot Test of Self Reporting Mobile Exit (SRME) Public-Facing Mobile Application

- CBP Self Reporting Mobile Exit (SRME) is designed to provide travelers a capability to use their mobile device to biometrically record their exit from the U.S. by submitting a live photo and use of location services on the phone to verify the report is being made from outside the U.S.
- To use the exit features of the mobile app, the traveler provides biographic travel document information, port of departure, and a submission of his or hers facial biometric. The application performs liveness verification of the traveler, geolocation verification, retrieves the active I-94 number, and timestamps the submission with the date and time.
- The limited test will be conducted with actual travelers. The Seattle and Buffalo Field Offices, specifically Blaine and Champlain Ports of Entry, are planned to be the pilot sites for testing the I-94/SRME mobile app. The targeted population are travelers in need of an I-94 with an admit until date (AUD) that expires during the testing period.
- A robust public outreach campaign will be used to publicize the availability of the mobile app for use on entry and benefit of using to report one's exit in meeting compliance requirements of I-94.
- The test will start in August 2020 for up to 180 days. The duration may be adjusted to align with I-94 traveler departure patterns. It is estimated that up to 5000 travelers will be participants in this pilot.
- A final report on test results will include analysis of participant demographics, biometric match results, geographic reporting data (e.g., distance from the port of entry when report is submitted), and a detailed technical analysis from both the pilot test and the white hat hackathon of geo-location data and liveness detection. CBSA departure records will be used as part of this analysis.

Pedestrian Exit Technology Demonstration

- CBP is planning to evaluate acquiring photos of pedestrians departing from the U.S. and compare those photos to the photos associated with the documents on file in government holdings, without requiring the traveler to present a document. CBP will deploy a camera system with software capable of taking an image of an approaching traveler, and will leverage technology under consideration or currently deployed to the land entry environment.

(b)(5), (b)(7)(E)

Next Steps and Challenges:

- Increase BE-Mobile pulse and surge operations at select locations – as requested.
- Complete development of the SRME Mobile Application (Android and iOS) and test use of SRME Mobile Application at the Blaine, WA and Champlain, NY POEs.
- Conduct White Hat Hackathon of SRME Mobile Application.

~~For Official Use Only~~

- Conduct pedestrian exit technology demonstration, based on currently deployed and expanding pedestrian entry biometric solution – by December 2020.
- As with other programs, the Covid-19 Pandemic delayed planned deployment work in March – May 2020.
- (b)(5) Discussions are on-going as to schedule impact of the pandemic.

Submitted By: (b)(6), (b)(7)(C) Land Border Integration and Biometric Programs

U.S. Customs and Border Protection

CBP One™
Use the Appointments Feature Designed for Brokers, Carriers and Forwarders

Download CBP One™ Today!

Streamlines inspection requests and appointment updates

Reduces/Eliminates unnecessary wait time for runners

Enhances communications through email status updates to your group inbox.

Download on the App Store

GET IT ON Google Play

To get started, download CBP One™ on the Apple App Store or Google Play Store.

Sign In Using Login.gov

The app will redirect to **login.gov**, where you can either create or login to your existing account.

Desktop version coming soon!

Questions? Contact us at: cbpone@cbp.dhs.gov

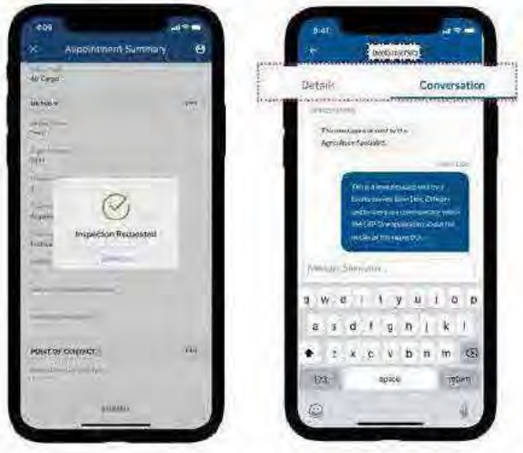
1. Who Are You
Tap on 'Broker/Carrier/Forwarder' in order to begin.

2. Create Profile
Add and save all necessary information. Your profile can be edited at any time in the future by tapping on the profile icon in the top right corner.

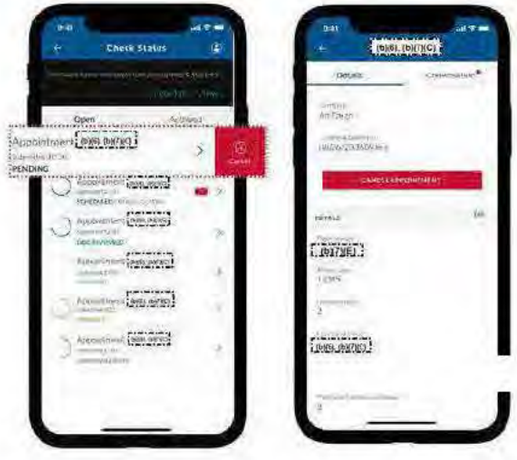
3. Request Inspection
Tap on 'Schedule an Inspection', select appointment type and cargo type, and fill in the required fields.

4. Review Information
Review all information and tap on submit. You may edit information on this page by tapping "edit".

5. Successfully Submitted!
You will receive in app and push notifications, along with emails on the status of your appointments. A CBPAS may initiate a chat, which you can respond to under the 'Conversation' tab.



6. Cancel/Edit an Inspection
View/edit the details of your inspection under "check status". If you need to cancel an inspection, simply swipe to the left or click on the "Cancel Appointment" button in the details screen. Completed/Cancelled appointments will be archived.*



*Only pending inspections can be edited, while pending, acknowledged, scheduled, doc reviewed, and assigned inspections can be cancelled.

CBP One ™

Convenient, Faster Inspection Appointments for Brokers, Carriers and Forwarders

Download CBP One™ on the Apple App Store or Google Play Store.

Benefits

- Real-time appointment status updates
- View inspection request history
- Reduced wait time for runners
- Interactive chat feature with CBP

U.S. Customs and Border Protection

Contact your local Port of Entry (POE) to see if inspection appointment requests can be submitted via CBP One ™

Questions? Contact CBPOne@cbp.dhs.gov

CBP One ™

Quickly Request Cargo Inspection Appointments

To get started, download CBP One™ from the Apple App Store or Google Play Store.

ANDROID APP ON Google play | Download on the App Store

Use the Appointment feature designed for Brokers, Carriers and Forwarders

- Real-time appointment status updates
- View inspection request history
- Reduced wait time for runners
- Interactive chat feature with CBP

U.S. Customs and Border Protection

Contact your local Port of Entry (POE) to see if inspection appointment requests can be submitted via CBP One ™

Questions? Contact CBPOne@cbp.dhs.gov

CBP One™ | Allows **TSA Supervisors** to retrieve traveler information

U.S. Customs and Border Protection

Download CBP One™

Download on the App Store

GET IT ON Google Play

To get started, download CBP One™ from the Apple App Store or Google Play Store.

Sign In Using Login.gov

The app will redirect to **login.gov** where you can either create or login to your existing account.

Questions?
Contact us at: CBPOne@cbp.dhs.gov

1. Who Are You

(b)(7)(E)

First time users will be prompted to create a profile.

(b)(7)(E)

2. Retrieve Traveler Information

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

3. Photo Results

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

4. (b)(7)(E) Results

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

5. Biographical Data

(b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

6.Biographical Data Results

(b)(6), (b)(7)(C), (b)(7)(E)

(b)(6), (b)(7)(C), (b)(7)(E)

Message

From: (b)(6), (b)(7)(C)

Sent: 9/21/2020 2:42:34 PM

To: (b)(6), (b)(7)(C)

Subject: Trusted Traveler Programs work flow for CBP One applicatoion

Attachments: TTP Mobile App WorkFlow4.pptx

(b)(6), (b)(7)(C)

Like we discussed last week everything related to Trusted Traveler Programs will start when the user selects the traveler option. How we envision it is once the user selects traveler and they choose either Land, Sea, or Air one of the options on the next screen will be check TTP Status with a brief description of TTP. All the functionality of TTP Programs will be available after the option is selected. Please see the attached power point for reference. The power just shows a workflow from "Air" but our PSPD team has built it for all the travel methods.
The first functionality will be a status check and then submitting the application. If you have any additional questions please let me know.

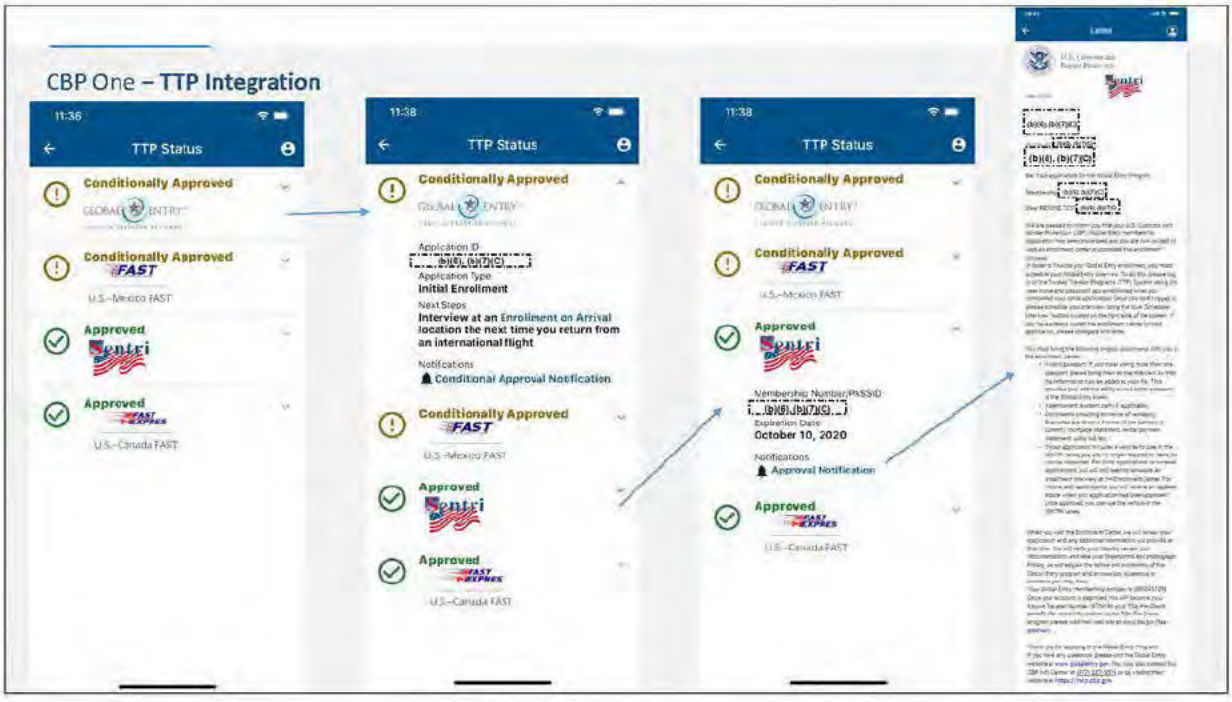
Thank you

(b)(6), (b)(7)(C)
Supervisory CBPO (Program Manager)
U. S. Customs and Border Protection
OFO/ Trusted Traveler Programs
Washington D.C.
(b)(6), (b)(7)(C) (Office)

CBP One **Onboarding**

Log In to access TTP Mobile

TTP Mobile App – Plug In


CBP One – TTP Integration
11:38
TTP Status
Conditionally Approved
GLOBAL ENTRY
Conditionally Approved
FAST
U.S.–Mexico FAST
Approved
Sentri
Approved
U.S.–Canada FAST
Application ID
Application Type
Initial Enrollment
Next Steps
Interview at an Enrollment on Arrival location the next time you return from an international flight
Notifications
Conditional Approval Notification
Membership Number/PASSID
Expiration Date
October 10, 2020
Approval Notification

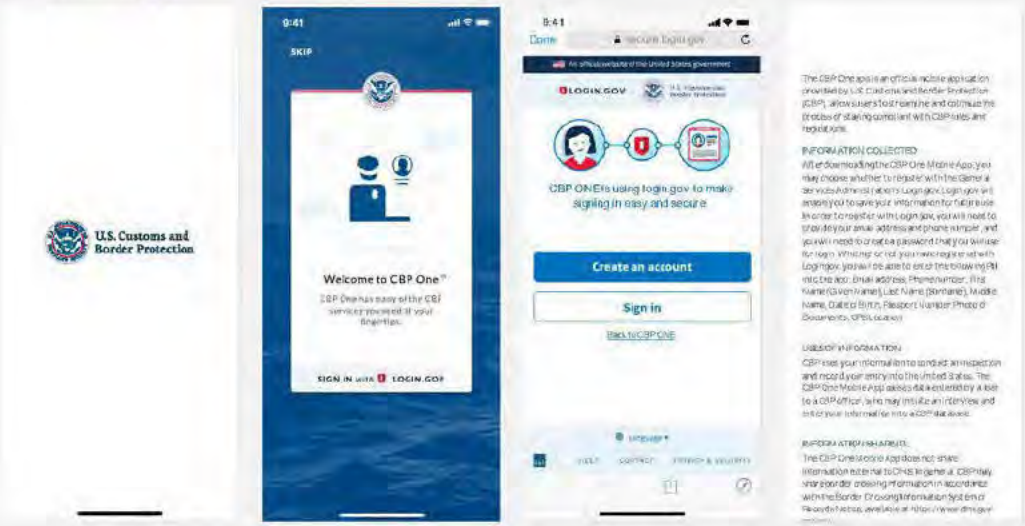
CBP One – **TTP Profile**

CBP One™

MOCK UPS

4.0.0.v – 7/27/20

CBP One **Onboarding**

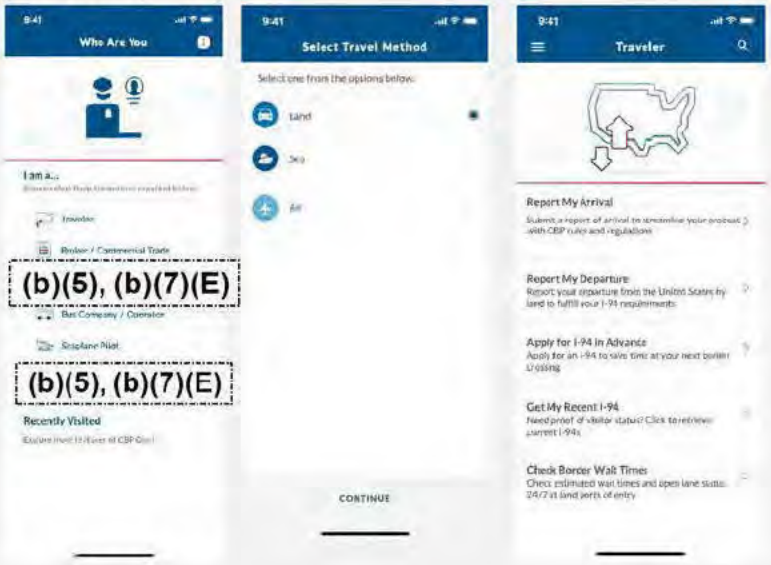


CBP One **Home**

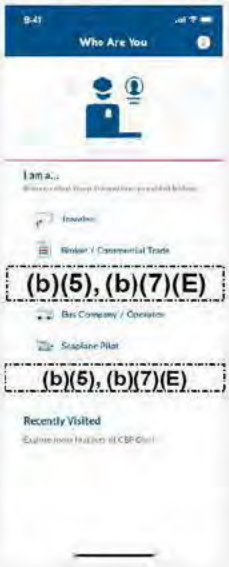

9:41
Who Are You
I am...
Traveler
Broker / Commercial Trade
(b)(5), (b)(7)(E)
Bus Company / Operator
Seaplane Pilot
(b)(5), (b)(7)(E)
Recently Visited
9:41
Who Are You
I am...
Traveler
Broker / Commercial Trade
(b)(5), (b)(7)(E)
Bus Company / Operato
Seaplane Pilot
(b)(5), (b)(7)(E)
Recently Visited

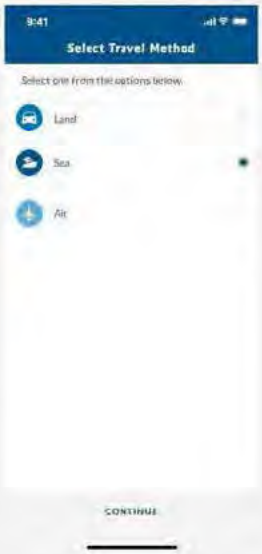
CBP One **I Am... > Traveler > Method of Travel**

CBP One – **Traveler Home (Land)**

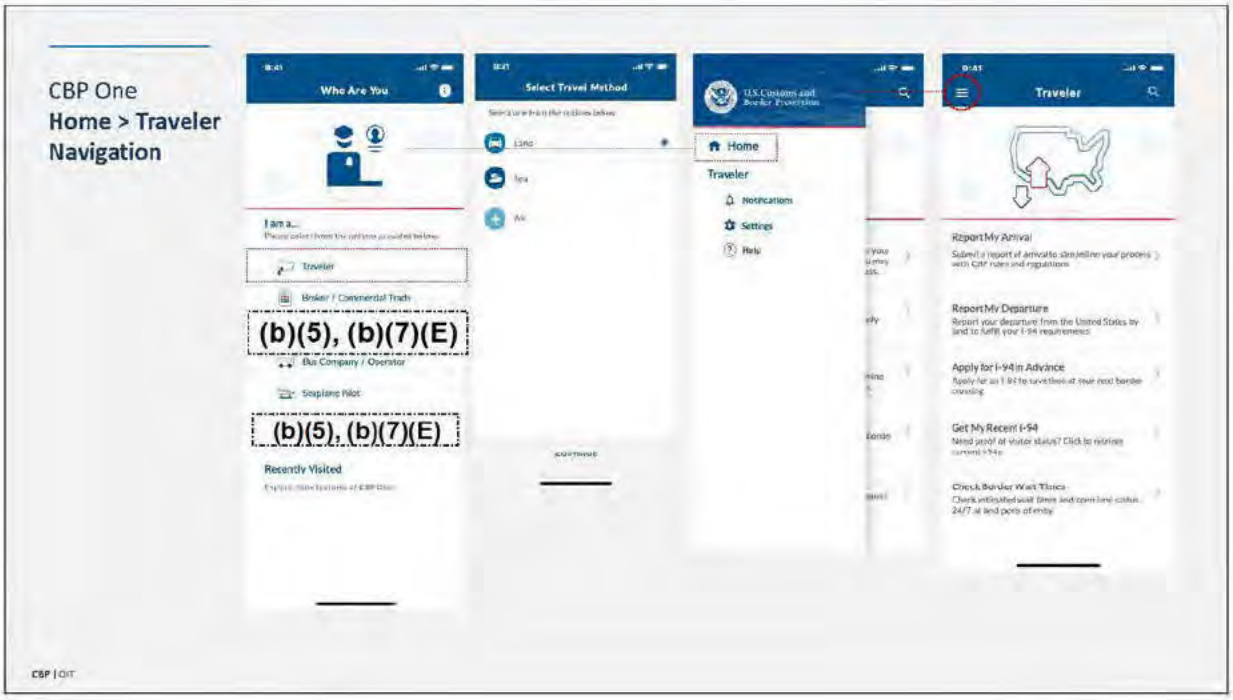


CBP One – **Traveler Home (Sea)**





CBP One – **Traveler Home (Air)**


CBP One
Home > Traveler
Navigation
Who Are You
I am a...
Traveler
Broker / Commercial Trade
(b)(5), (b)(7)(E)
Bus Company / Operator
Seaplane Pilot
(b)(5), (b)(7)(E)
Recently Visited
Select Travel Method
Land
Sea
Air
U.S. Customs and Border Protection
Home
Traveler
Notifications
Settings
Help
Traveler
Report My Arrival
Report My Departure
Apply for I-94 in Advance
Get My Recent I-94
Check Border Wait Times
CBP | OIT

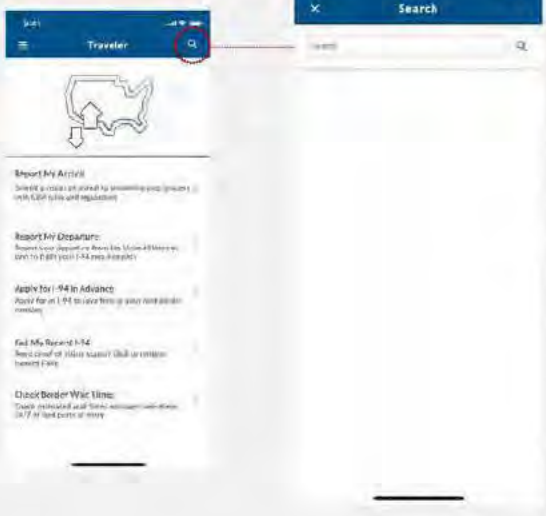
(b)(7)(E)

CBP One **Home > Traveler Navigation > Traveler Profile**

CBP One **Notifications**

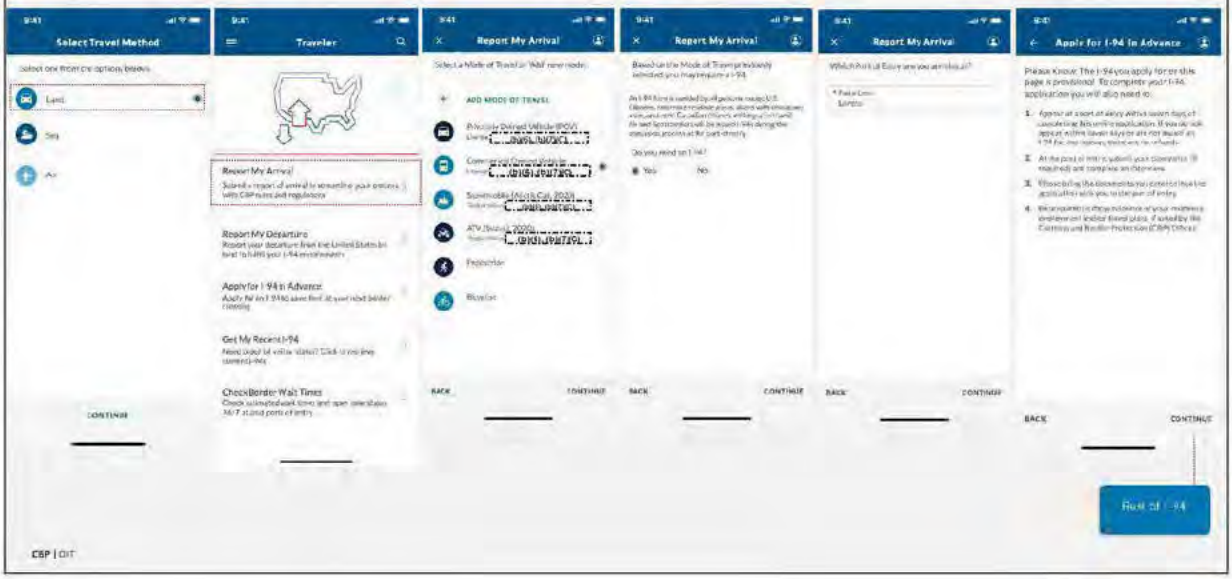
(b)(5), (b)(7)(E)

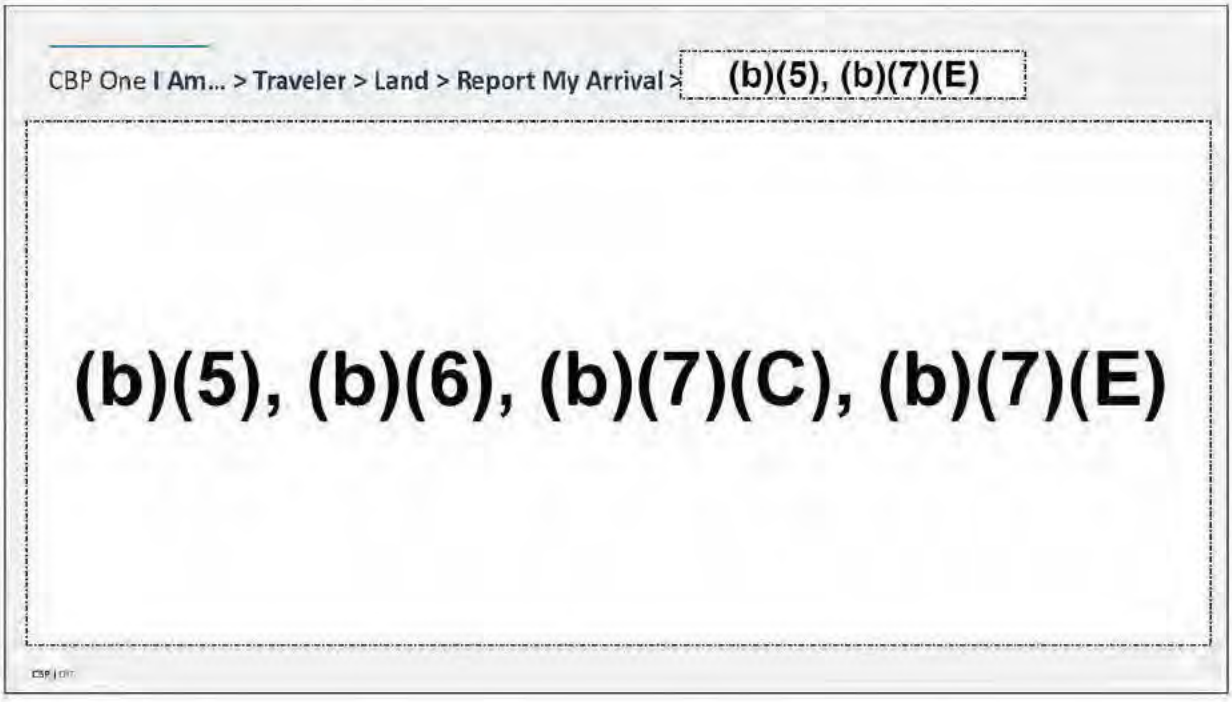
CBP One **Search**



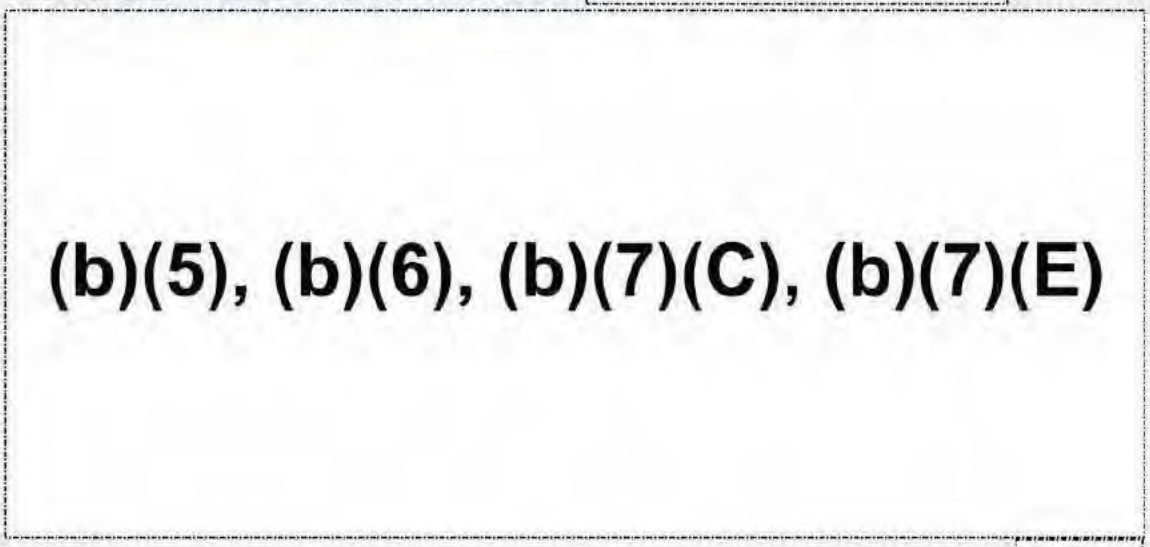
(b)(5), (b)(7)(E)

CBP One **I Am... > Traveler > Land > Report My Arrival > Need an I-94?**



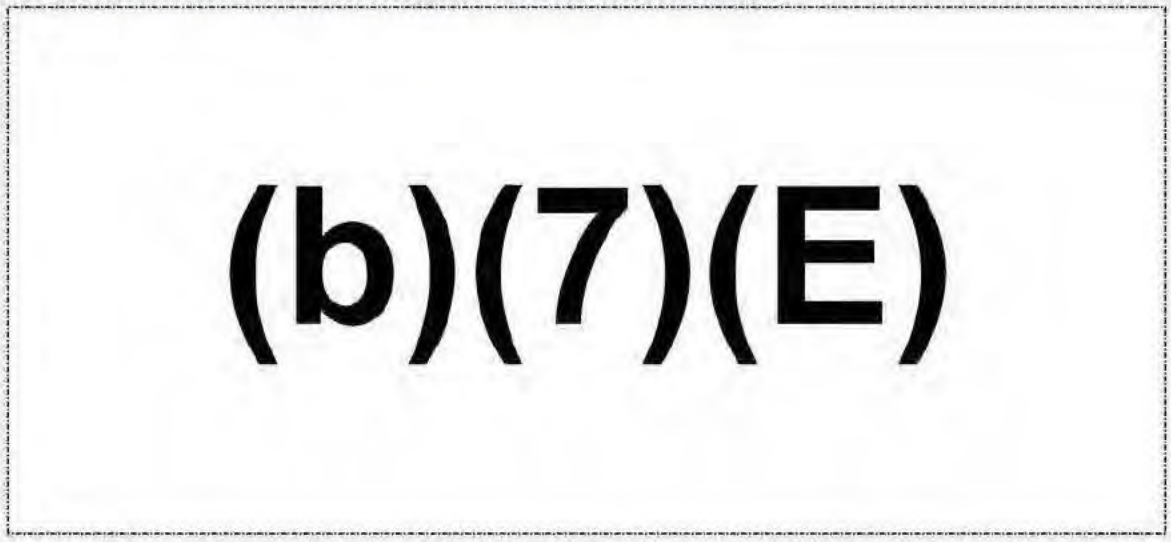

CBP One I Am... > Traveler > Land > Report My Arrival
(b)(5), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One I Am... > Traveler > Land > Report My Arrival (b)(5), (b)(7)(E)



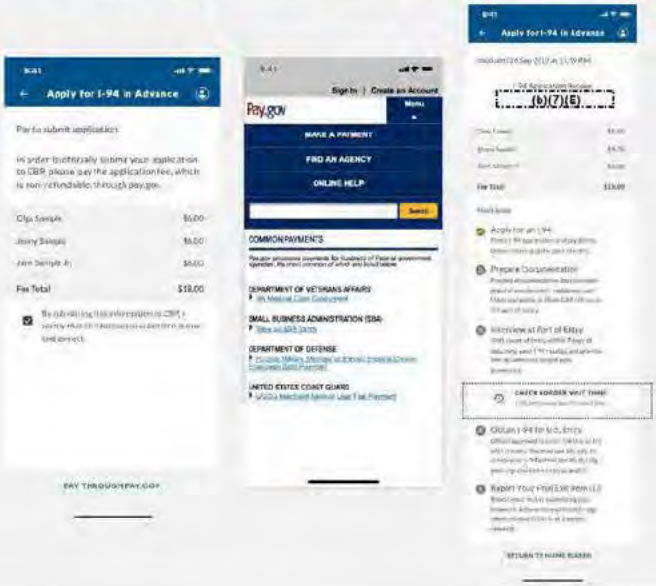
(b)(5), (b)(7)(E)

CBP One **I Am... > Traveler > Land > Report My Departure> I-94 Exit**

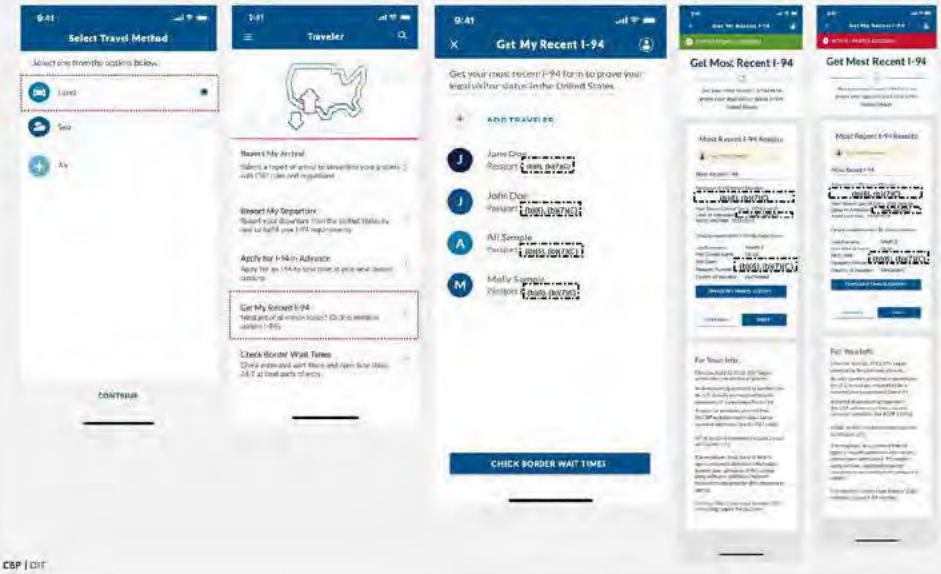


CBP One I Am... > Traveler > Land > Apply for I-94 in Advance

CBP One **I Am... > Traveler > Land > Apply for I-94 in Advance > Check For Border Wait Times After Application**



CBP One I Am... > Traveler > Get Existing I-94 > Get My Recent I-94 Workflow



CBP One I Am... > Traveler > (b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(5), (b)(7)(E)

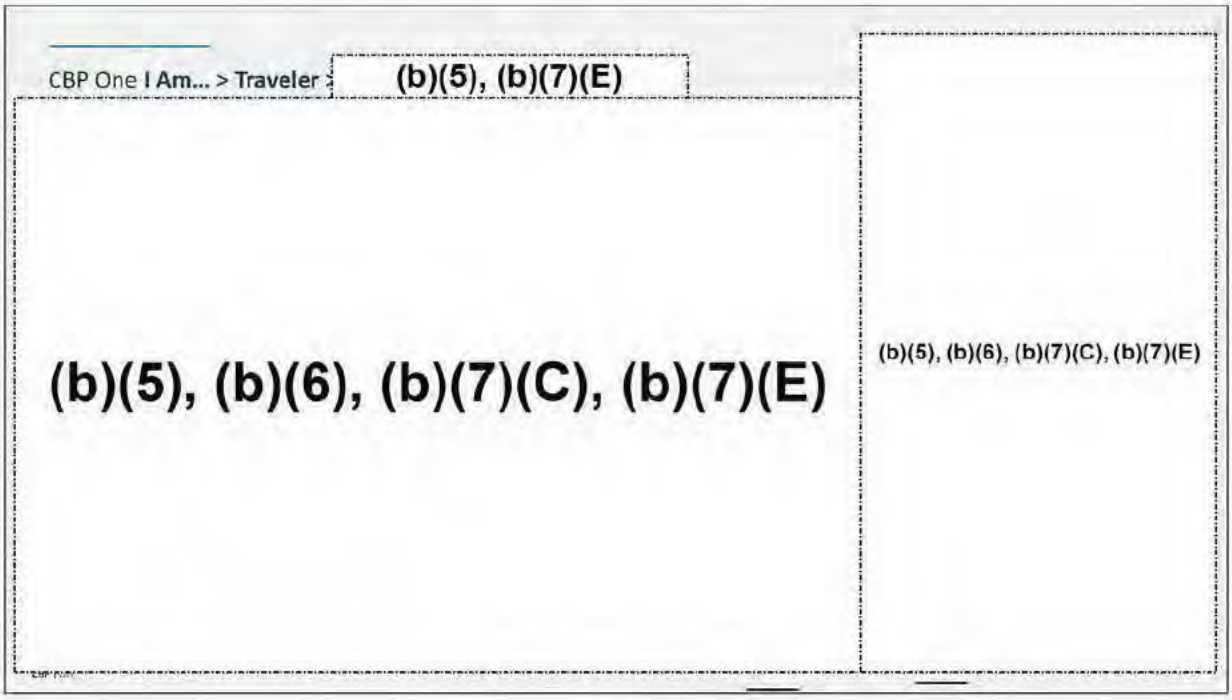
CBP One I Am... > Traveler (b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One I Am... > Traveler > (b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

CBP | DHS


CBP One I Am... > Traveler
(b)(5), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One I Am... > Traveler (b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One **I Am... > Traveler > Air > Request Inspection/Check Status**

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One **I Am... > Traveler > Air > Request Inspection/Check Status Continued**

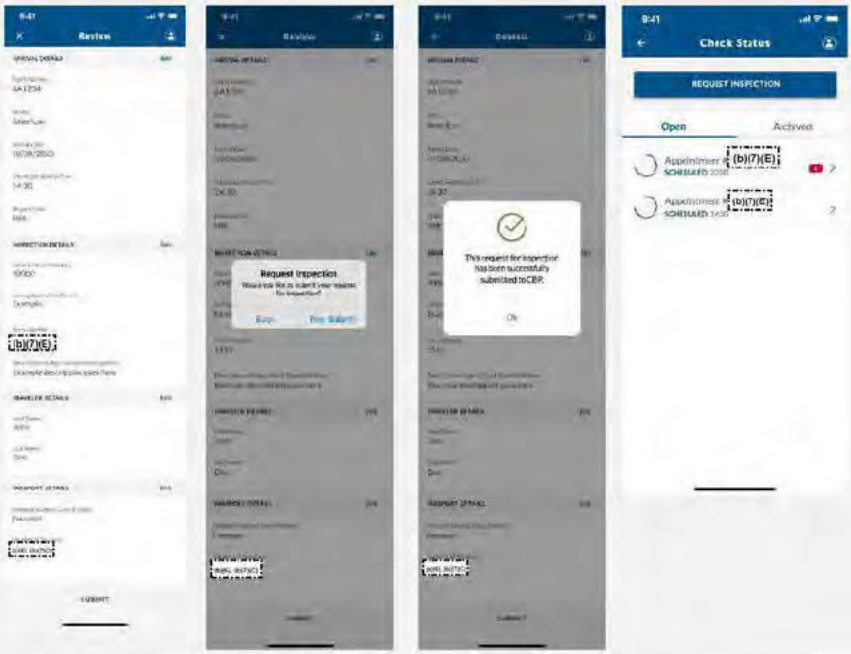
CBP One **I Am... > Traveler > Air > Request Inspection/Check Status Continued**

CBP One **Request Inspection/Check Status Continued**

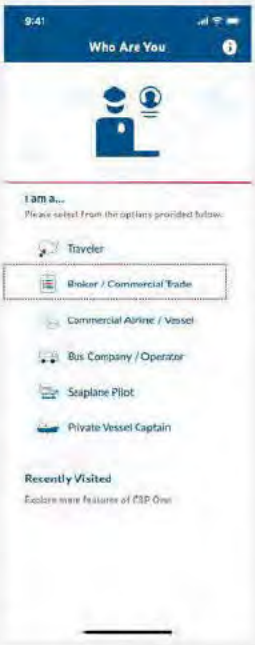
First time users would input all their information here

Single Select: Returning users would see this screen (their information would have been saved from a previous visit)

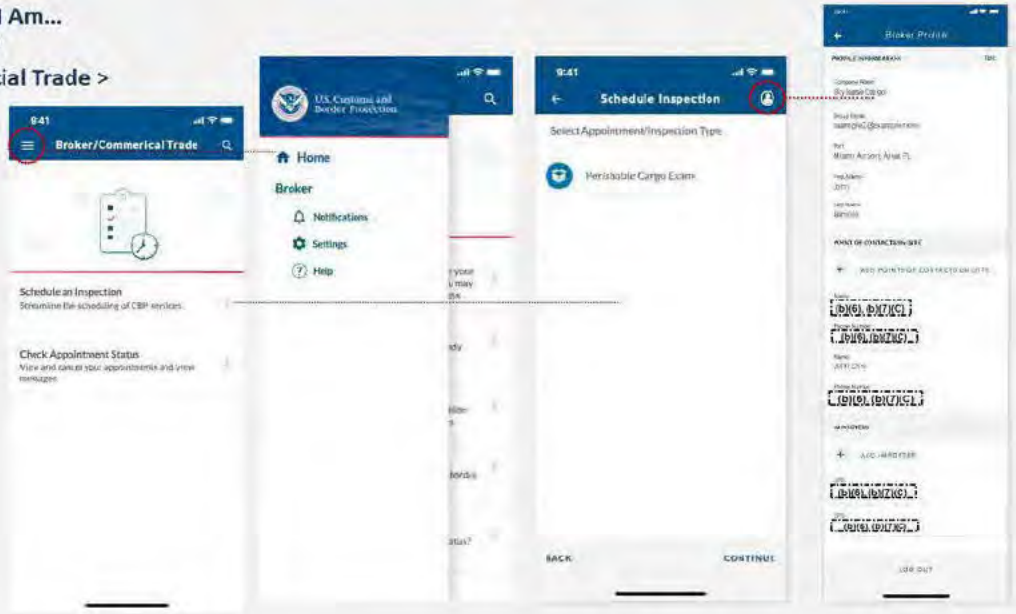
CBP One **Request Inspection/Check Status Continued**



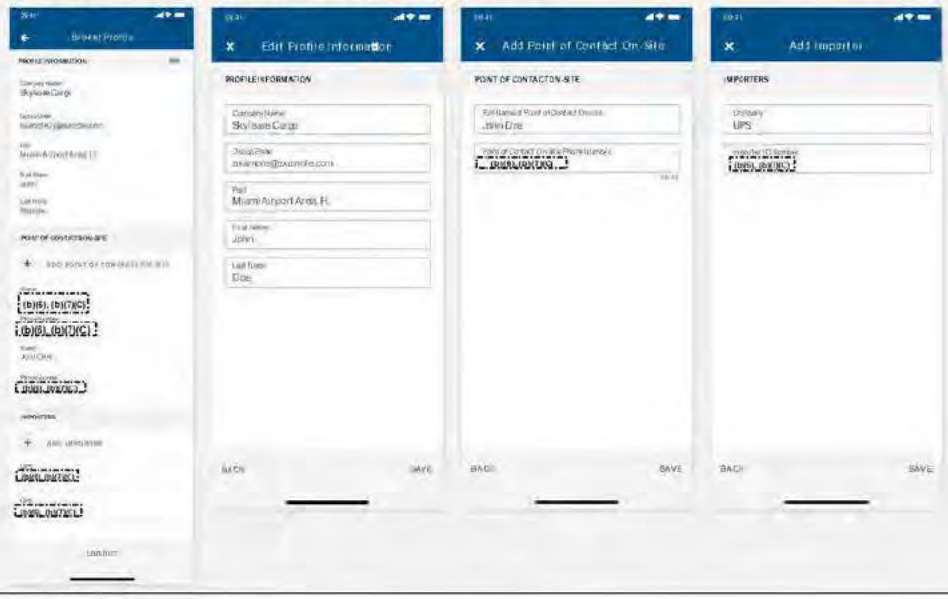
CBP One **I Am...**
> Broker/
Commercial Trade



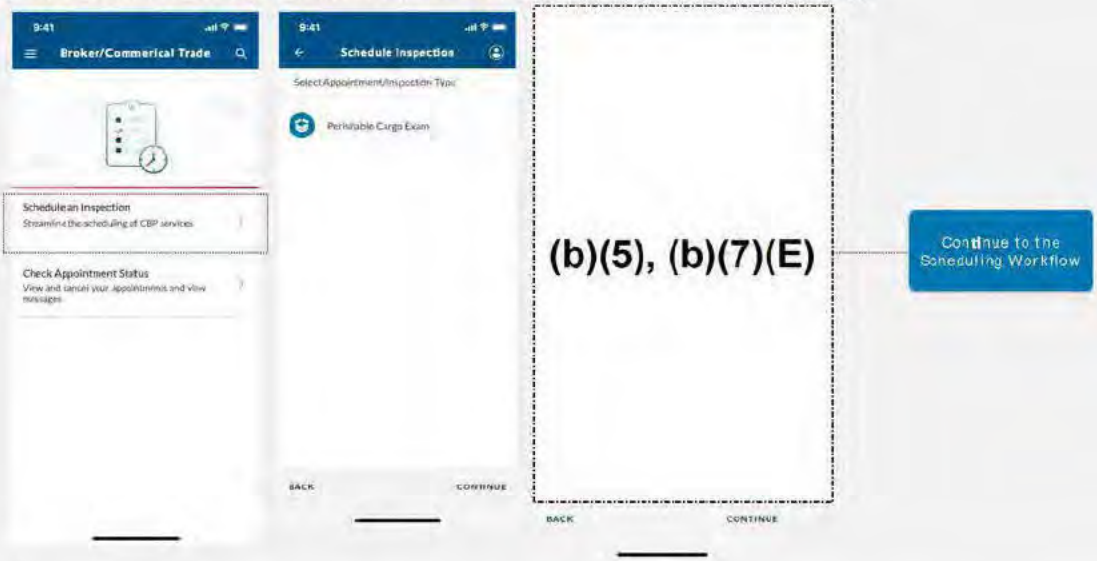
CBP One **I Am...**
> Broker/
Commercial Trade >
Profile



CBP One **I Am... > Broker/Commercial Trade > Profile Continued**

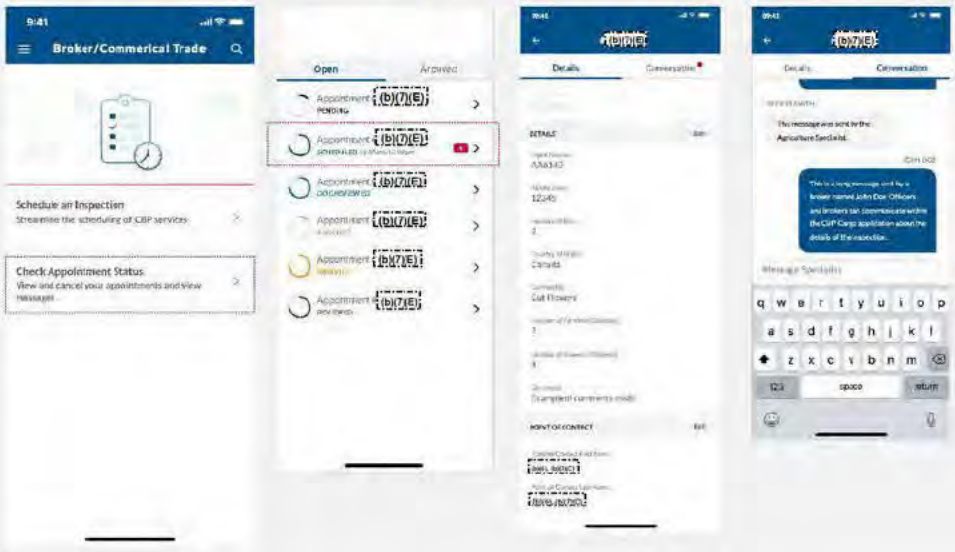


CBP One **I Am... > Broker/Commercial Trade > Schedule an Inspection > Scheduling**

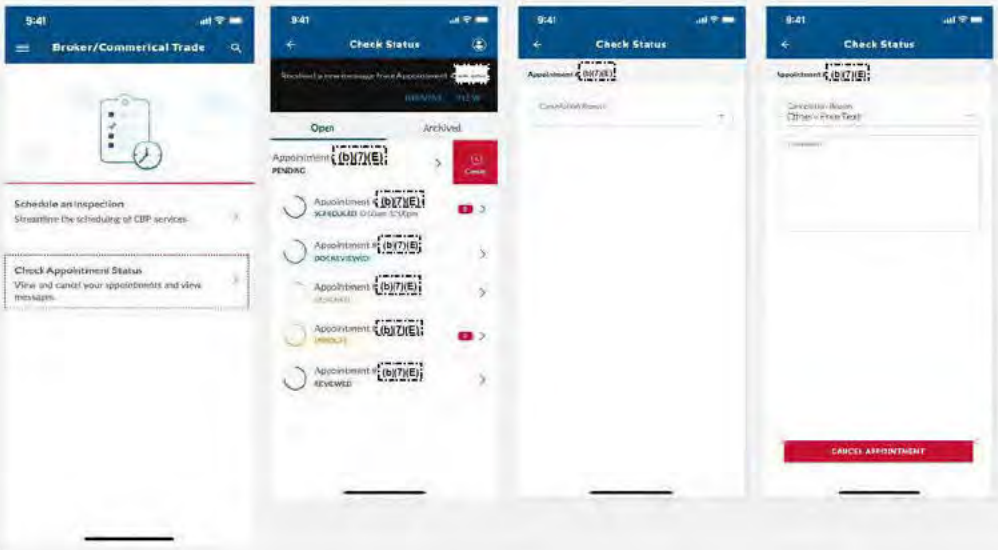


CBP One **I Am... > Broker/Commercial Trade > Check Appointment Status > Open/Archived**

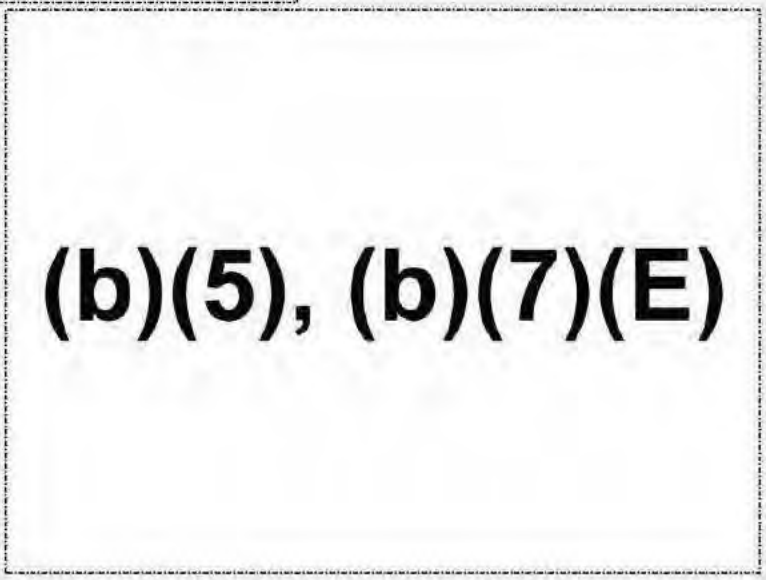
CBP One **I Am... > Broker/Commercial Trade > Check Appointment Status > Chat**

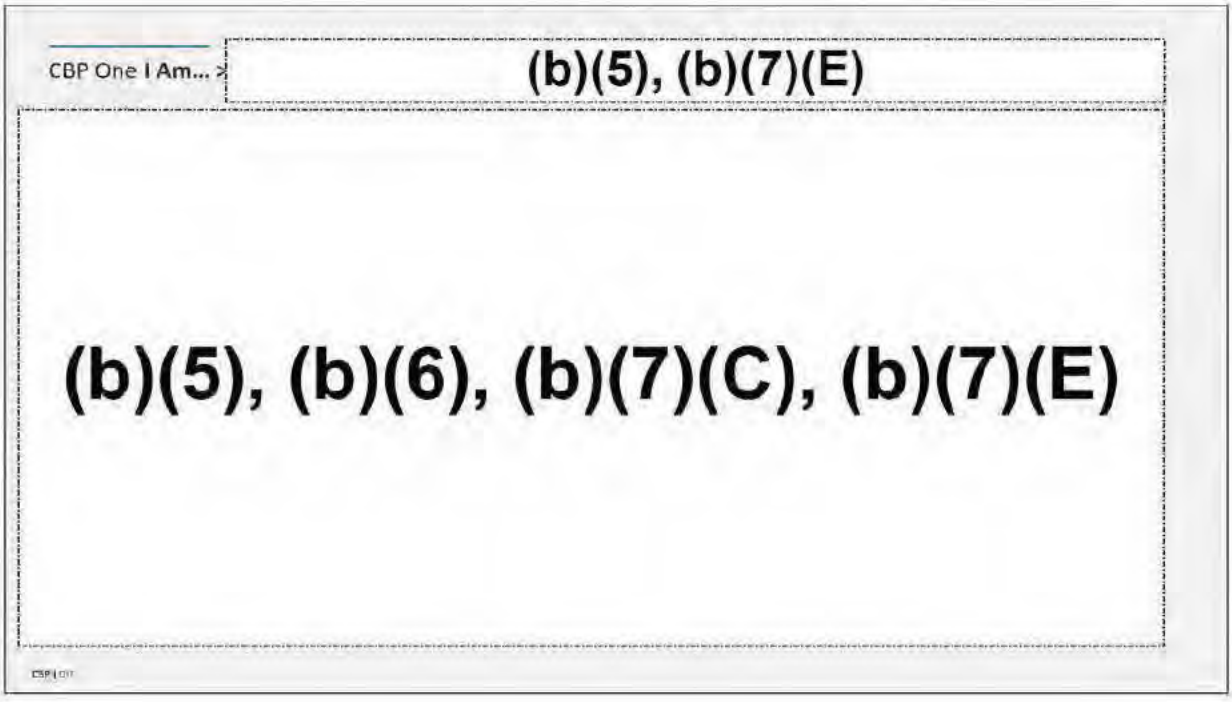


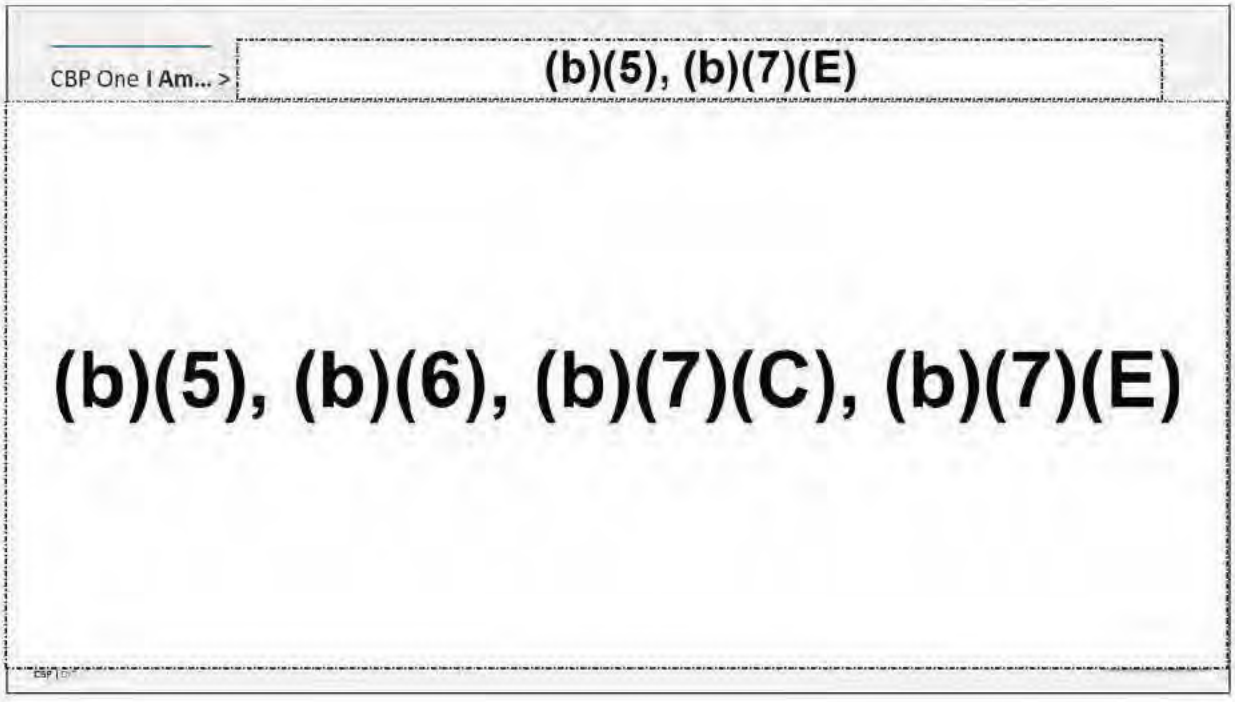
CBP One **I Am... > Broker/Commercial Trade > Check Appointment Status > Cancel**



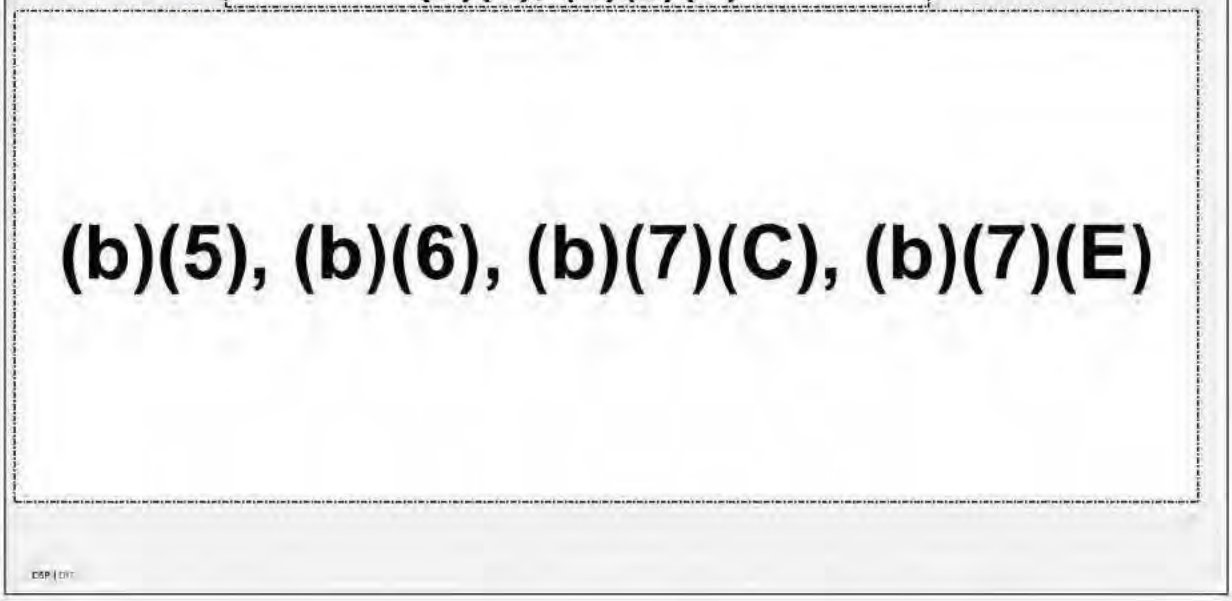
CBP One I Am... > (b)(5), (b)(7)(E)




CBP One I Am... >
(b)(5), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)


CBP One I Am... >
(b)(5), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

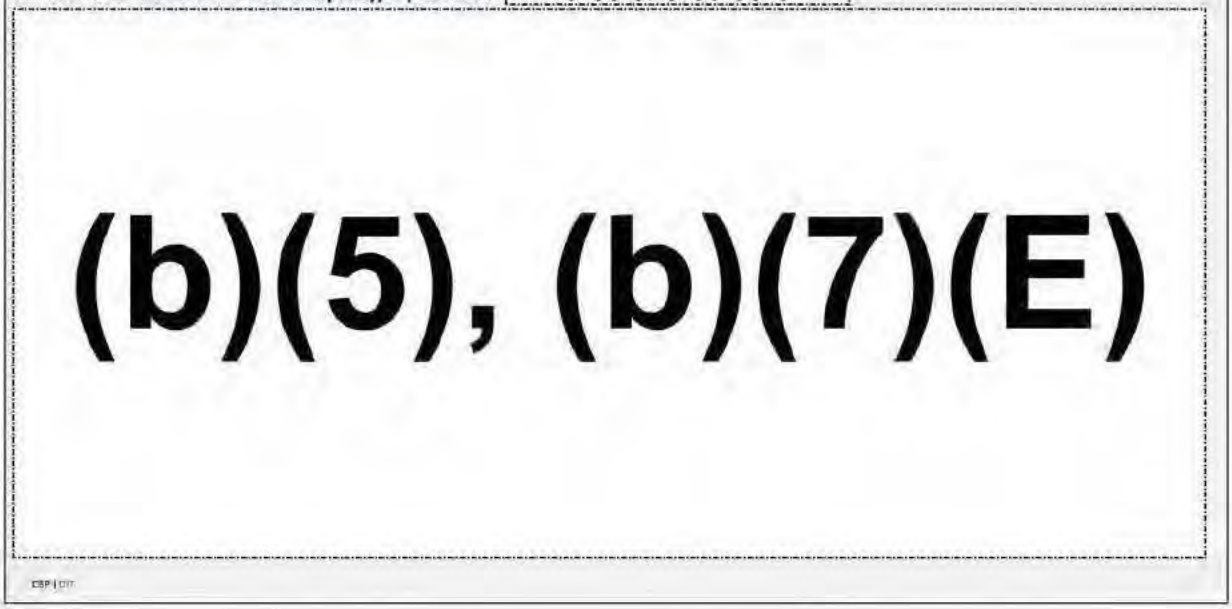
CBP One I Am... > (b)(5), (b)(7)(E)



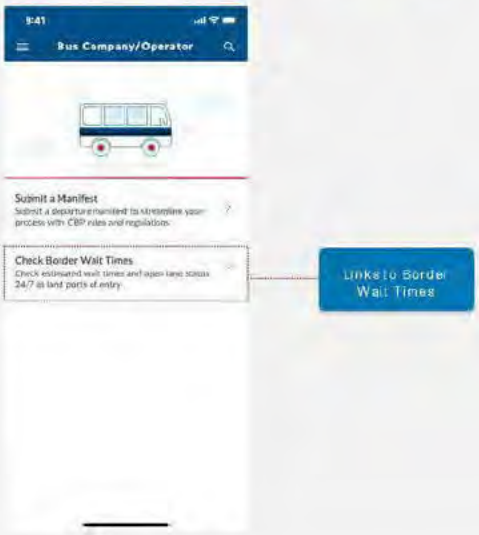
CBP | DHS

CBP One **I Am...**
> Bus Company/Operator

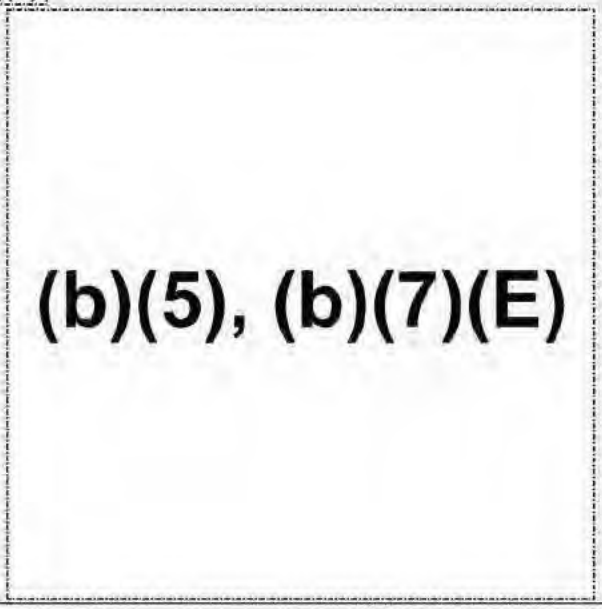
CBP One I Am... > Bus Company/Operator > (b)(5), (b)(7)(E)



CBP One **I Am... > Bus Company/Operator > Check Border Wait Times**



CBP One I Am... (b)(5), (b)(7)(E)



CBP One I Am... > (b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One **I Am...**

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

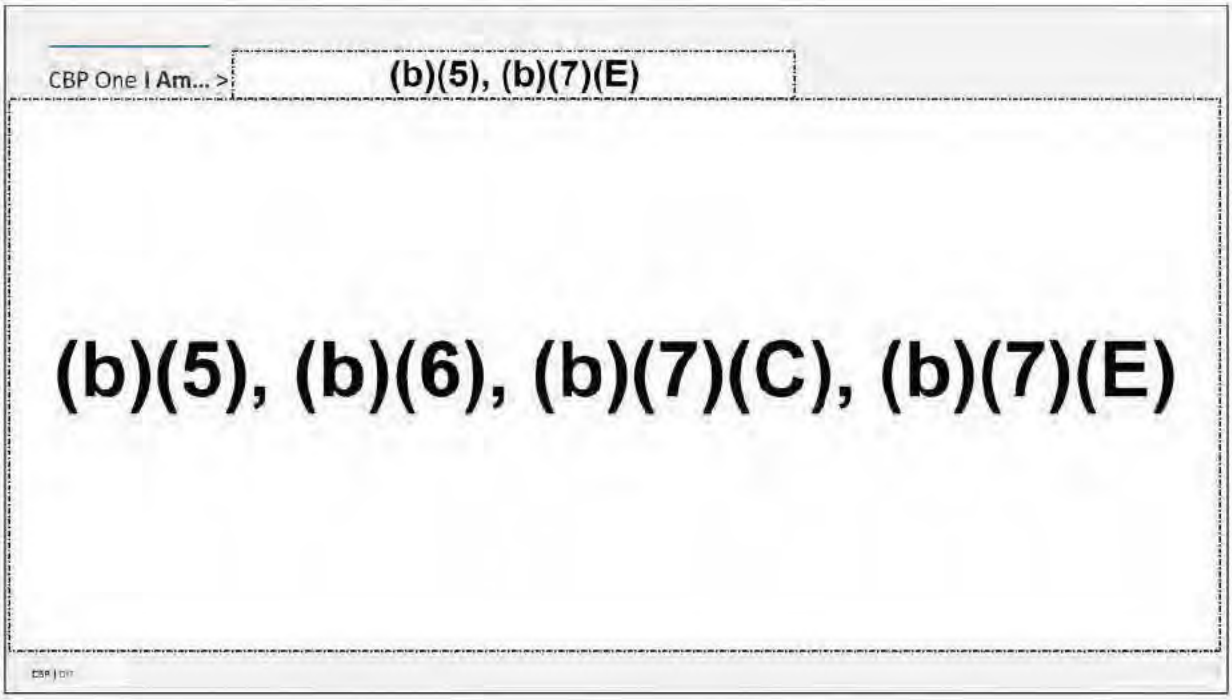
CBP | DHS

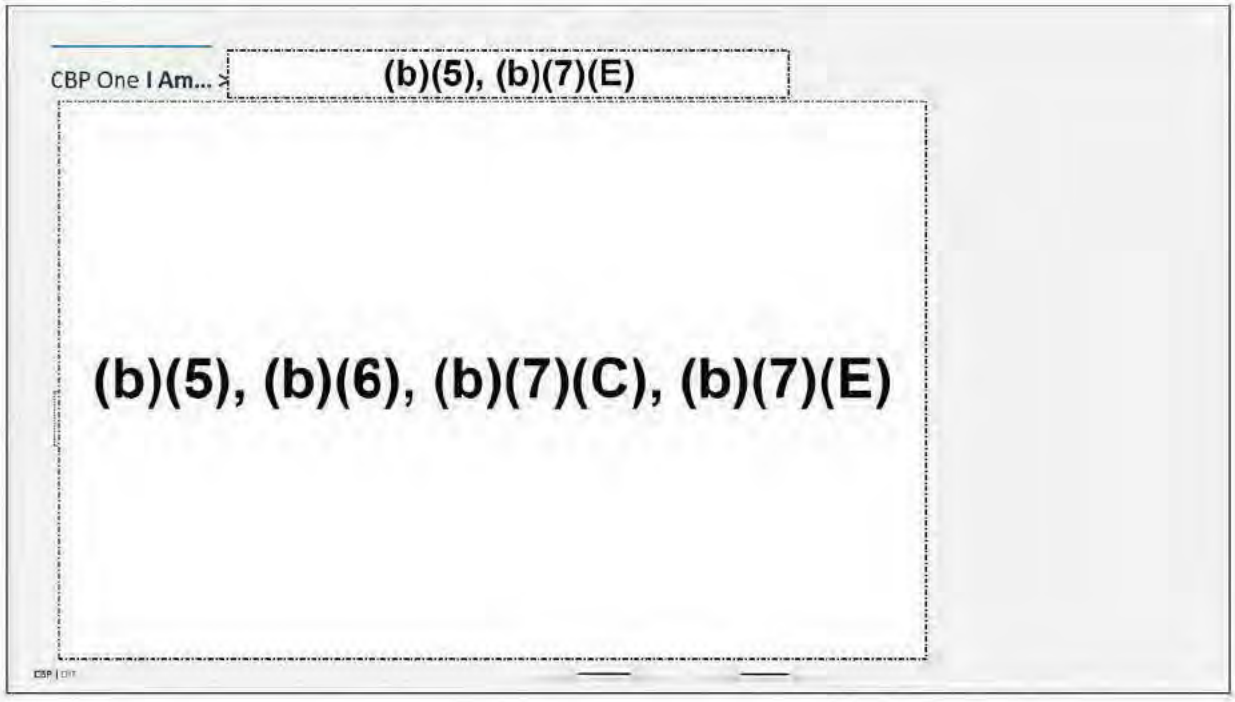
CBP One I Am.. (b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

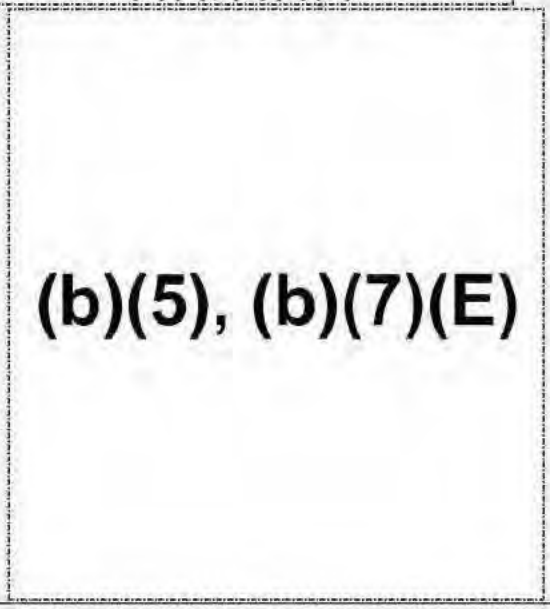
CBP One I Am... > (b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)


CBP One I Am... >
(b)(5), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)


CBP One I Am...
(b)(5), (b)(7)(E)
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

CBP One I Am... (b)(5), (b)(7)(E)



Message

From: (b)(6), (b)(7)(C)

Sent: 11/2/2021 9:16:52 PM

To: (b)(6), (b)(7)(C)

Subject: RE: (b)(5) for CBP One

Hi (b)(6), (b)(7)(C)

Did this get approved? I did not see a final email on this.

Thank you,

(b)(6), (b)(7)(C)
Acting Director, Strategic Transformation Office
Planning, Program Analysis and Evaluation
Office of Field Operations
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, July 28, 2021 9:04 AM
To: (b)(6)
(b)(6)
Cc: (b)(6), (b)(7)(C), (b)(7)(E)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: FW: (b)(5) for CBP One

Good morning. I reached out to our team here are CBP and have a response for OMB.

(b)(7)(E)

1. (b)(7)(E)
(b)(7)(E) Individuals can upload a photo and it does not need to be a live photo or a "passport" quality photo meeting specific requirements.

2. The photo is the most efficient source of identification to ensure the person presenting themselves at a limit line, with a paper copy of a confirmation email, is the person for whom the CBP On submission was made. (b)(7)(E)
(b)(7)(E)

3. CBP One is a voluntary program. It may not be feasible for all individuals. However the NGOs we have breifed acknowledge and seem ready and willing to support individuals. This is not NGOs submitting on their behalf, but NGOs providing support and access to the tools needed to utilize the application on a mobile device or desktop. If someone can not provide a photo, they can still present themselves to the POE directly.

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade

U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Monday, July 19, 2021 5:32 PM
To: (b)(6), (b)(7)(C)
(b)(6)
Cc: (b)(6), (b)(7)(C), (b)(7)(E)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(b)(6)
Subject: RE: (b)(5) for CBP One

(b)(6), (b)(7)(C)

OMB has one initial comment. It is related to the mandate of a photograph. Could CBP elaborate on what is the process if someone does not have access or the ability to provide a photograph? Does CBP feel that this requirement can be fulfilled by any respondent to the collection? If someone can not provide a photograph, what is the process? Overall, OMB is looking to understand the requirement for the photograph and if this requirement could potential be an issue if someone can not provide a photograph due to limitation of resources.

Thanks.

(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Friday, June 25, 2021 11:10 AM
To: (b)(6)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: (b)(5) for CBP One

Good morning. (b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection

(b)(6), (b)(7)(C)

Cell: (b)(6), (b)(7)(C)

U.S. Department of Homeland Security

U.S. Customs and Border Protection

Entry/Exit Transformation

Houston – George Bush Intercontinental Airport

United Air Lines Post Deployment Site Visit

October 31 – November 2, 2017

Introduction

U.S. Customs and Border Protection (CBP) intends to demonstrate the initial implementation of the Traveler Verification Service (TVS) through the expansion of air exit capabilities at select airports. The limited expansion will demonstrate to airlines and airports how biometrics can be integrated into current boarding processes, provide real-time, centralized biometric matching capabilities, and record biometrically verified outbound departures in CBP systems. Specifically, live photos of passengers will be compared against the photos stored in CBP systems utilizing the flight departure manifest. While completing analysis of existing biometric exit experiments, CBP began the implementation of a biometric air exit field trial 2016 at an airport in partnership with a large air carrier.

Purpose

Beginning in June 2017, United Airlines (UA) collaborated with CBP and NEC Corporation[1] to test facial recognition (FR) as part of ongoing trials to implement biometrics at air exit. On behalf of CBP Headquarters Office of Field Operations, (b)(7)(E) was contracted to perform a time and motion study and record observations of UA flight departing from Houston George Bush Intercontinental Airport (IAH) to Tokyo, Japan on October 31 – November 2, 2017. The purpose of the study at IAH was to determine the total boarding time of all passengers on the selected UA flights and the individual passenger process time at the NEC NeoFace® Express facial recognition station. Information obtained from the study will be compared to metrics derived during the baseline study performed at IAH on June 1-3, 2017 to include the total flight boarding time and process time for passengers to use the self-boarding gates to scan the boarding pass prior to departure.

Approach

For the post deployment time and motion study, a team of (b)(7)(E) analysts, escorted by CBP, were stationed at Gate Terminal E to observe the following flights (Table 1):

Table 1. Flights Observed During Post Deployment Time and Motion Study

Date	Flight #	Destination	Gate	Number of Passengers	Scheduled Departure Time
October 31, 2017	(b)(7)(E)	Tokyo, Japan	(b)(7)(E)	(b)(7)(E)	11:20AM[2]
November 1, 2017		Tokyo, Japan			11:20AM
November 2, 2017		Tokyo, Japan			11:20AM

[1] https://www.necam.com/docs/?id=6c812b4d-2a12-40ed-9fea-fae81550c7aa
[2] Flight was originally scheduled to depart at 10:30AM but changed due to off season travel.

While on site, (b)(7)(E)

(b)(7)(E)

The metrics captured include:

(b)(7)(E)

Qualitative Analysis

Figure 1 below is a diagram of the departure gates at IAH Terminal E. International UA flight (b)(7)(E) departed from Gate (b)(7)(E)

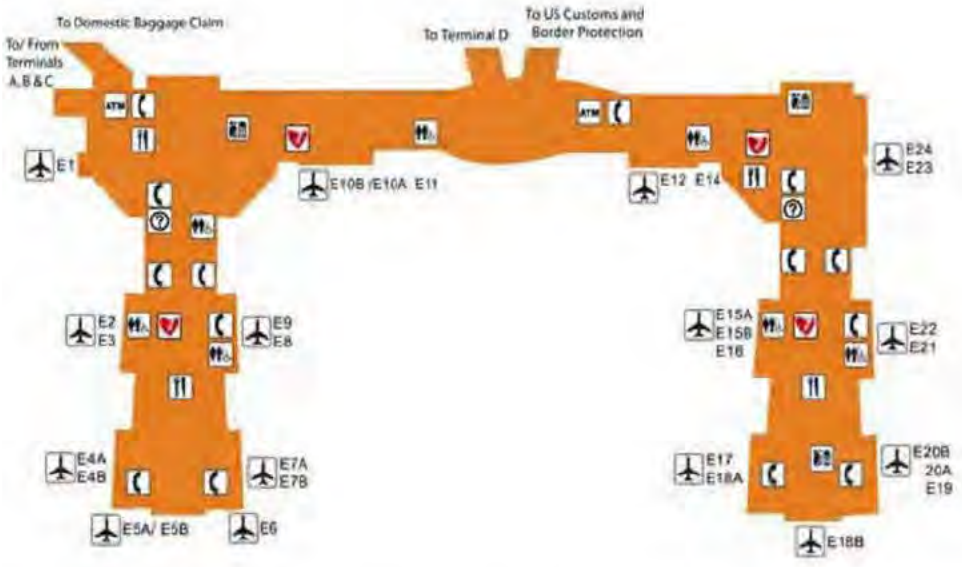


Figure 1. Layout of departure gates at IAH Terminal E

Two adjoining customer service counters were housed in the gate area, and each counter was equipped with two workstations (Figure 2). The far-right workstation at the customer service counter contained a stationary boarding scanner to process crew, eligible pre-boarding passengers, and late arrivals.

Figure 2. Boarding gate E for UA flight at IAH Terminal E

Four self-boarding gates with bar code scanners were positioned adjacent to the customer service counter at Gate . Digital signs were exhibited above each self-boarding gate and were activated by the UA agents to display the proper boarding group, during the boarding process. Since the baseline site visit in June 2017, three NeoFace® Express facial recognition stations were installed in front of the self-boarding gates, approximately 5 feet from the head of queue (Figure 3).

Figure 3. NeoFace® Express facial recognition stations at gate IAH Terminal E

While on-site, (b)(7)(E)

(b)(7)(E)

Five free-standing queue signs, separated by stanchions for Groups 1 through 5, were positioned approximately ten feet in front of the self-boarding gates (Figure 4). (b)(7)(E)

(b)(7)(E)

[3] A post deployment study was also conducted at William P. Hobby Airport on October 31 – November 2, 2017. A separate report and stats will be provided to CBP Headquarters.

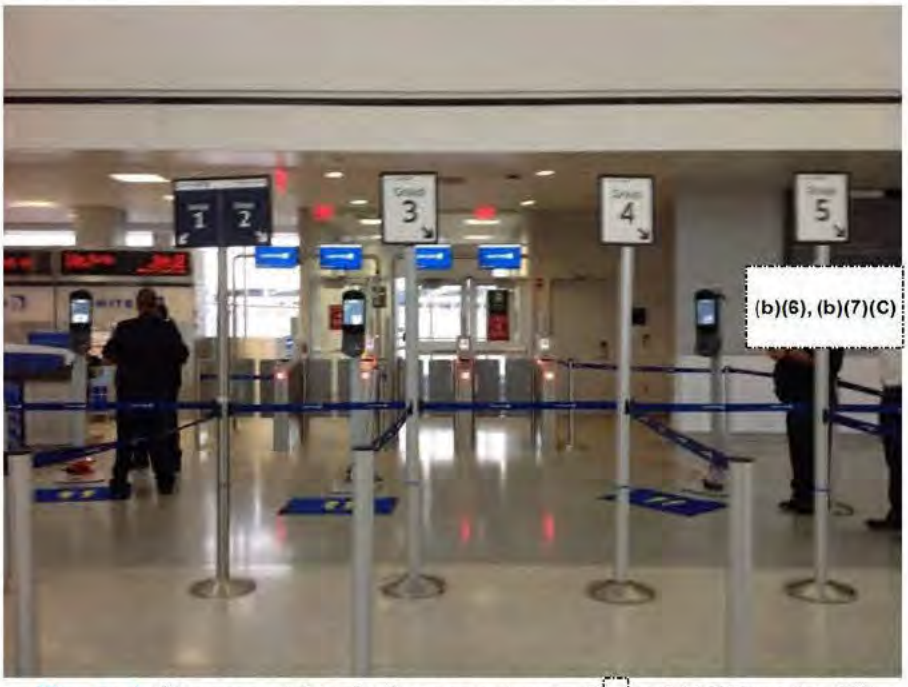


Figure 4. Signage at head of queue at gate E at IAH Terminal E

A video animation, demonstrating the use of the self-boarding gates, was observed at Gate (Figure 5),

(b)(7)(E)

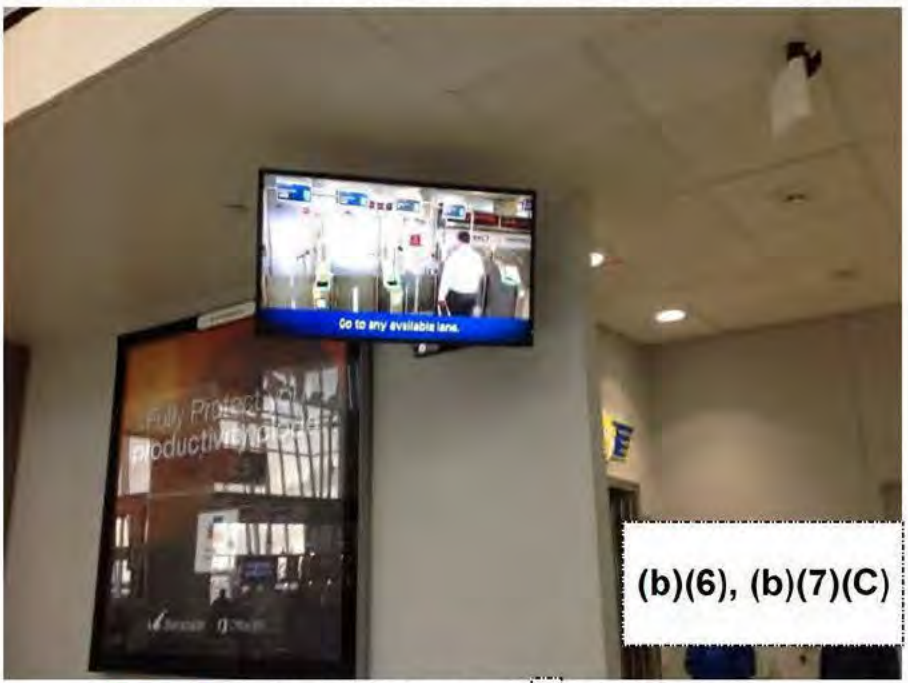


Figure 5. Animated video at gate at IAH Terminal E

(b)(7)(E)

Figure 6. Left: LED screens in the gate area; Right: UA digital screens above self-boarding gates

A UA agent in the gate area announced to all waiting passengers the time boarding would begin, as well as informed passengers that CBP would be "collecting biometrics". (b)(7)(E)

(b)(7)(E)

(b)(7)(E)

The order of UA flight boarding was conducted as follows:

1. Pre-boarding for military, passengers that required special assistance, including passengers in wheelchairs, and families with children under the age of two.
2. Group 1: First Class
3. Group 2: Premier
4. Groups 3 and 4: Economy

A UA agent was positioned at the head of queue and removed the stanchion in front of the appropriate group that was authorized to board. The UA agent reviewed each passenger's passport and directed one passenger to each of the three cameras. (b)(7)(E)

(b)(7)(E)

As shown in Figure 7, the NEC NeoFace® Express camera stood an estimated 5.5 feet high, and the system used a dual-camera design to capture the facial image. A floor mat was placed in front of the camera to stage the passenger to capture a photo to compare to the gallery for facial recognition. The bar code scanner was situated below the camera for passengers to scan boarding passes, which then subsequently activated the photo capture. While on site, all passengers in wheelchairs were required to approach a facial recognition station for processing. On occasion, CBP would tilt the equipment for passengers in wheelchairs to capture the photo.

At William P. Hobby Airport and Atlanta International Airport, (b)(7)(E)

(b)(7)(E)

Figure 7. Right: NeoFace® Express facial recognition; Left: Passenger processing

The facial recognition screen depicted an animation illustrating the procedure to scan the boarding pass (Figure 8). Passengers were expected to independently scan their own boarding passes, (b)(7)(E)

(b)(7)(E)

(b)(7)(E)

Figure 8. Boarding pass illustration on the facial recognition station

After the photo capture was complete, the system responded with three possible outcomes:

1) Green Screen - indicated a positive match between the current photo and a matching photo in the CBP photo gallery. The message "Thank you – enjoy your flight" was displayed.
2) Yellow Screen – indicated a quality issue with the passenger photo compared to the gallery photo. (b)(7)(E)

(b)(7)(E)

3) Blue Screen – There was no match with pictures in the CBP gallery.

(b)(7)(E)

After boarding authorization, passengers traversed through the gate and entered a sterile corridor, approximately 15-20 yards long, and turned right leading to the jet bridge (Figure 9). (b)(7)(E)

(b)(7)(E)

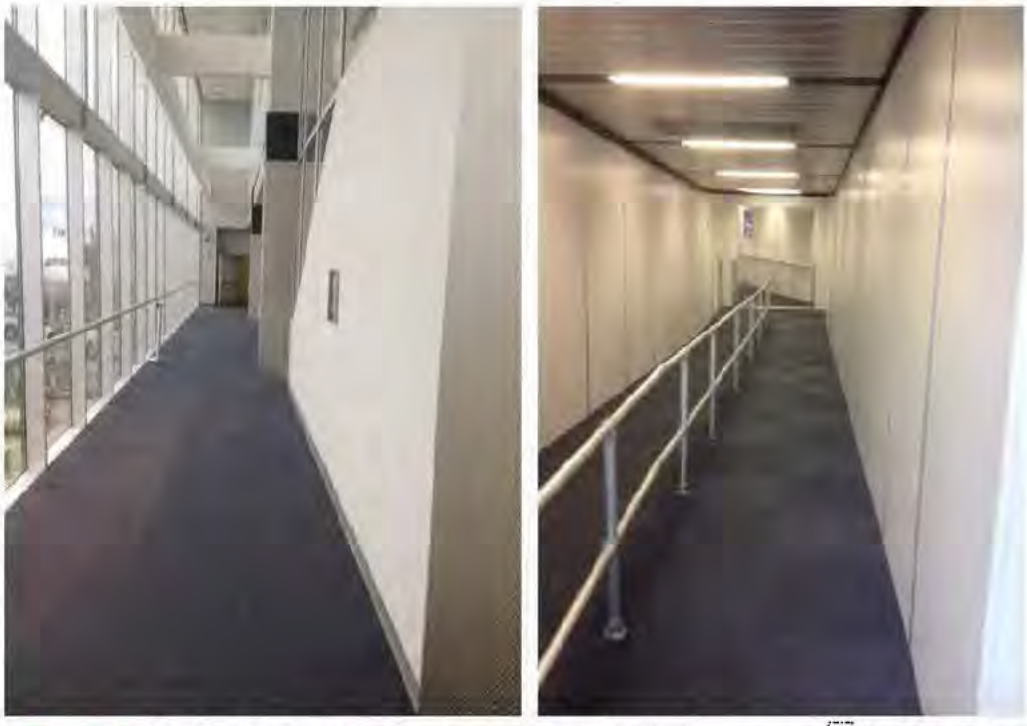


Figure 9. Right: Sterile corridor: Left: Path to jet bridge at gate (b)(7)(E) Terminal E

Quantitative Analysis

Facial Recognition Station Metrics

A total of [redacted] passengers from three outbound flights to Tokyo, Japan were observed using the NEC NeoFace® Express camera during the post deployment site visit. Table 2 lists the number of passengers observed and recorded using the facial recognition station for each of the three flights.

Table 2. United Airlines International Boarding Metrics

Date	Flight #	Number of Passengers Observed at Facial Recognition	Total Number of Passengers on Flight
October 31, 2017	(b)(7)(E)		
November 1, 2017			
November 2, 2017			

Table 3 provides key metrics derived from the post deployment time and motion study. (b)(7)(E) (b)(7)(E) (b)(7)(E) Out of (b)(7)(E) passengers observed, (b)(7)(E) passengers' photos successfully matched the gallery on the first attempt, resulting in a green light response at the facial recognition station and an average process time of (b)(7)(E) seconds.

Table 3. Facial Recognition Metrics (time in seconds)

Metric Name	
Passenger Facial Recognition Walk-Up Time – All Passengers	(b)(7)(E)
Passenger Facial Recognition Process Time –All Passengers	
Passenger Facial Recognition Process Time –Green Light Response Only	
Passenger Facial Recognition Process Time – Blue/Yellow Light Response Only	

Figure 10 presents the breakout of the walk-up time per passenger and Figure 11, the facial recognition process time per passenger across all (b)(7)(E) passengers observed.

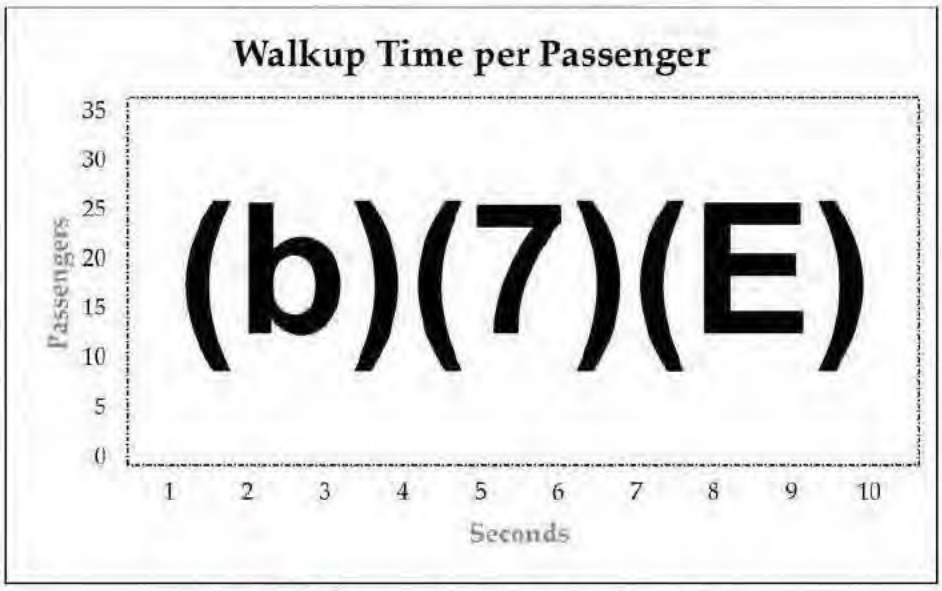


Figure 10. Average passenger walk-up time

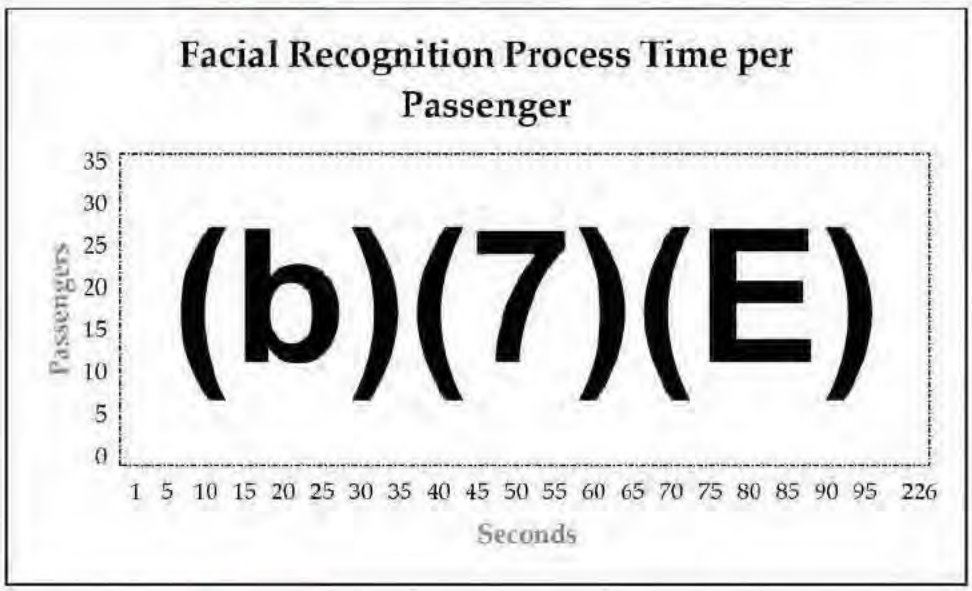
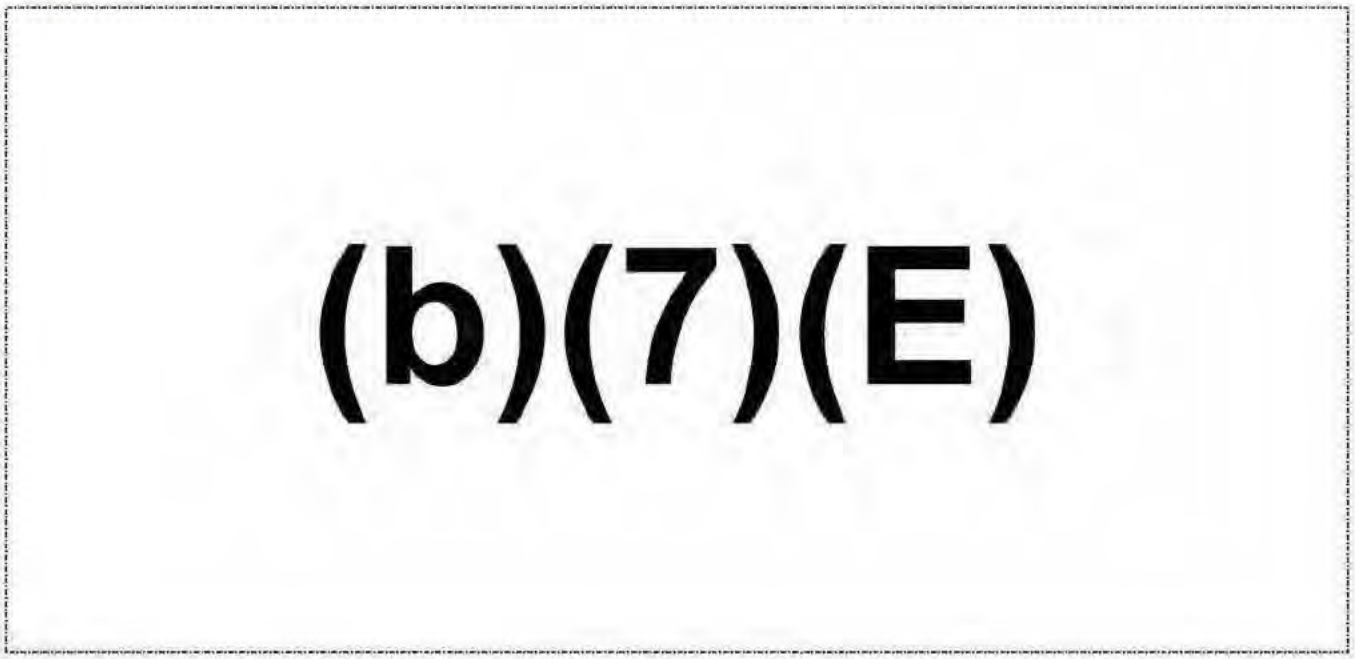


Figure 11. Average passenger process time

The average passenger process time of (b)(7)(E) seconds was derived by recording the following steps at the facial recognition station:

- Step 1 – Boarding pass scanned: Time passenger arrived at the facial recognition station and scanned the boarding pass until the photo capture process begins.
- Step 2 – Facial Recognition Complete: Time the photo capture and facial recognition was complete.
- Step 3 – Departure: Time the photo capture and facial recognition was complete until the passenger steps away from the facial recognition station.

Figure 12 presents a breakout, by percentage, of the average time spent by the passenger on each step of the process. Based on observations, total processing time was impacted for the following reasons:



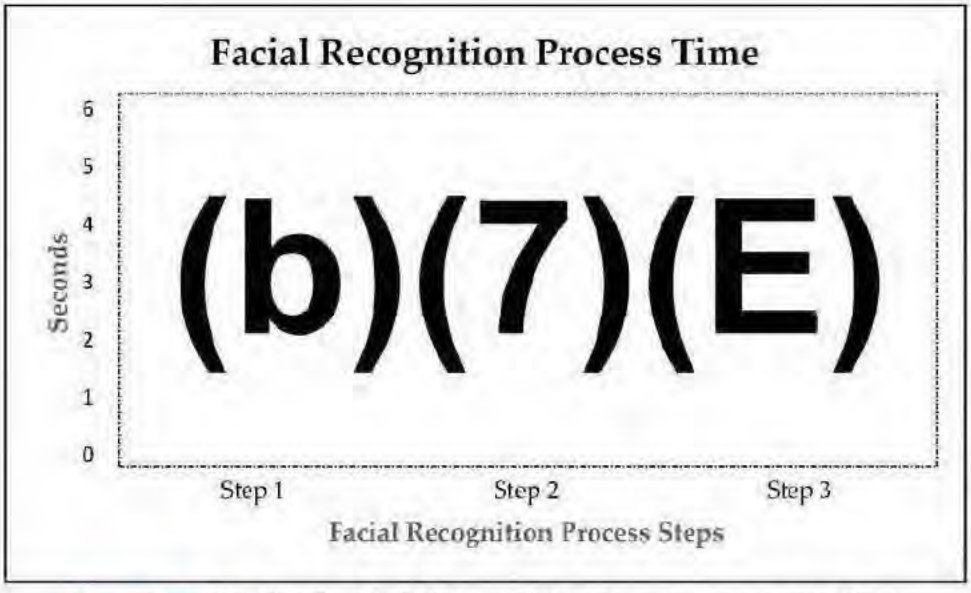


Figure 12 Breakout of facial recognition passenger process time

The sum of the average passenger walk-up time and the facial recognition boarding process time equals the average passenger cycle time. A calculated passenger throughput rate (average passengers processed per minute) is derived from the passenger cycle time. Given the average walk-up time of (b)(7)(E) seconds and average process time of (b)(7)(E) seconds, the cycle time at the facial recognition station is

(b)(7)(E) seconds per passenger. The result of the calculated throughput rate is (b)(7)(E) passengers per minute per facial recognition station.

By comparison, during the baseline study performed in June 2017, the observed average walk-up time was (b)(7)(E) seconds, plus (b)(7)(E) seconds passenger process time per self-boarding gate, resulting in a cycle time of (b)(7)(E) seconds or (b)(7)(E) passengers per minute per gate. Three self-boarding gates were operating in June for a total throughput of (b)(7)(E) passengers per minute. Table 4 compares the calculated cycle time and throughput from the baseline time and motion study at the self-boarding gates and post deployment time and motion study at the facial recognition stations.

Table 4. Derived Throughput: Self-Boarding Gates vs. FR Station

Metric Name and Definition	Baseline – Self Boarding Gates	Post Deployment – Facial Recognition Station
Average Passenger Cycle Time (seconds)	(b)(7)(E)	
Calculated Throughput (passengers per minute)		

(b)(7)(E)

(b)(7)(E) Table 5 lists the total boarding times for all general boarding passengers, including pre-boarding and first class. (b)(7)(E)

(b)(7)(E)

Table 5. General Passenger Boarding Times by Flight

Departure Date	Total Passengers Boarded	Total Flight Process Time (minutes)	Total General Passengers Boarded[4]	Total General Boarding Time (minutes)	Number of Passengers Boarded Per minute
June 1, 2017	(b)(7)(E)				
June 2, 2017					
June 3, 2017					
October 31, 2017					
November 1, 2017					
November 2, 2017					

While on site, there were four self-boarding gates in operation (b)(7)(E) (b)(7)(E) (Figure 13); (b)(7)(E) (b)(7)(E)

Figure 13. View of available self-boarding gates during (b)(7)(E) boarding

[4] Includes pre-boarding, first class and all general boarding. Passengers that arrived late are excluded

(b)(7)(E)

Table 6 lists the initial and follow up responses for the (b)(7)(E) passengers observed at the facial recognition station. Initial green light responses were received for (b)(7)(E) passengers (b)(7)(E) and initial yellow or blue light responses were received for (b)(7)(E) passengers (b)(7)(E)

(b)(7)(E)

Table 6. Passenger outcomes at Facial Recognition Station

FR Scan Response	Count	Percentage
Green	(b)(7)(E)	
Yellow / Green		
Yellow / Yellow *		
Yellow / Yellow / Green		
Yellow / Blue *		
Blue / Green		
Blue / Yellow		
Blue / Blue *		
Total Observed		
(b)(7)(E)		

Table 7 lists the percentage of passengers, with selected attributes, observed at the facial recognition station. (b)(7)(E)

(b)(7)(E)

Table 7. Passenger Attributes

Attribute	Count	Percentage of Passengers Observed
(b)(7)(E)		

Flight Boarding Metrics

(b)(7)(E)

(b)(7)(E)

The metrics presented in Table 8 were calculated as follows:

- *Pre-boarding time* is defined as the time the first passenger eligible for pre-boarding is processed by an agent at the customer service counter or the facial recognition station until the time all pre-boarding passengers have been processed.
- *First class and General boarding time* is divided into four groups, Groups 1-4.
  - Groups 1 and 2 are first-class and UA premier passengers.
  - Groups 3 and 4 are economy passengers. (b)(7)(E)

(b)(7)(E)

Table 8. Total Boarding Time by Flight

Departure Date	Flight #	Pre/ Priority Boarding Time (mm:ss)	General Boarding Time (mm:ss)	General PAX Count	Late Arrival Boarding Time (mm:ss)	Late Arrival PAX Count	Total Flight Boarding Time (mm:ss)
October 31, 2017	(b)(7)(E)						
November 1, 2017							
November 2, 2017							

(b)(7)(E)

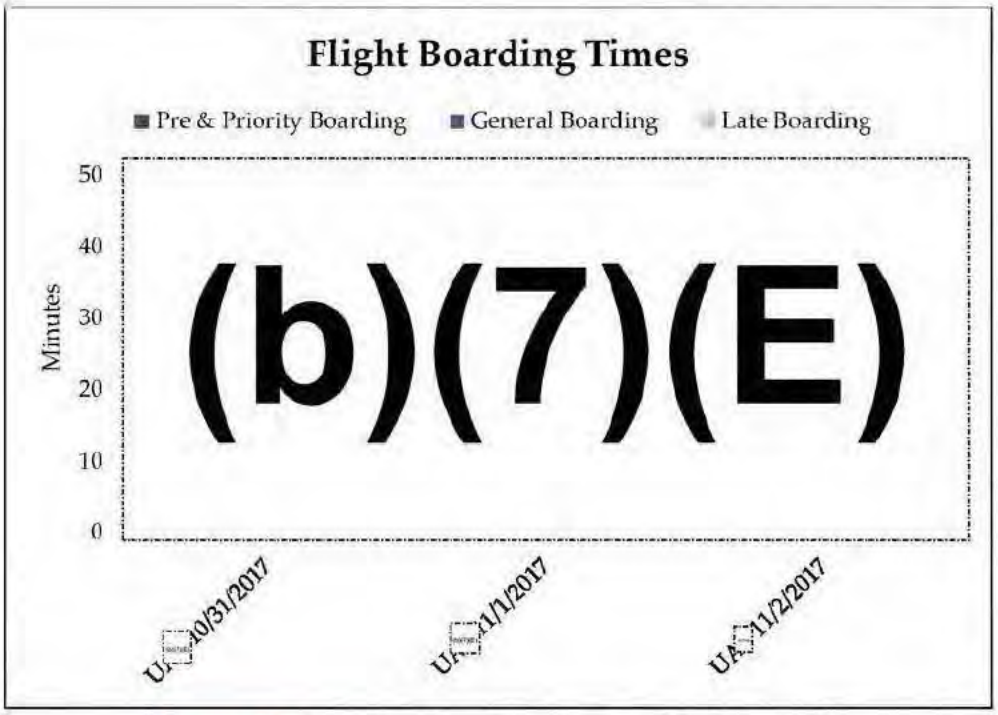


Figure 14. Boarding time by groups and late arrivals

(b)(7)(E)

Table 9. Key Flight Boarding Metrics

Departure Date	First Passenger Processed for Boarding	Time Last Passenger Processed at FR Station	Time Last Passenger Processed for Boarding	Last Passenger Enters Aircraft	Time Door to Jet Bridge Closed	Time Aircraft Pushed Back
October 31, 2017	(b)(7)(E)					
November 1, 2017						
November 2, 2017						

Conclusion

(b)(7)(E)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/21/2021 4:48:16 PM

To: (b)(6), (b)(7)(C)

CC: (b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
DURST, CASEY OWEN
DAVIES, MATTHEW S

Subject: RE: CBP One (b)(5)

Good afternoon all,

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

Please let us know if you would like to discuss via call (b)(5), (b)(7)(E) (b)(5), (b)(7)(E) If you need more formal or senior level concurrence with this path forward, also let me know. Thank you.

Best,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C) | U.S. Customs and Border Protection (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Thursday, June 10, 2021 9:45 AM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: CBP One (b)(5)

Good morning,

(b)(6), (b)(7)(C)

Could we have a check-in on this on Monday or Tuesday? (b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Thank you!

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
CBP Privacy Officer
Privacy and Diversity Office (PDO), Office of the Commissioner
U.S. Customs and Border Protection
1300 Pennsylvania Avenue NW, Room (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 8, 2021 1:28 PM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: CBP One (b)(5)

(b)(5), (b)(7)(E)

(b)(6), (b)(7)(C) | U.S. Customs and Border Protection | (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 8, 2021 11:00 AM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)

Subject: RE: CBP One (b)(5)

(b)(5), (b)(7)(E)

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 8, 2021 10:41 AM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)

Subject: RE: CBP One (b)(5)

Good morning!

We are moving along with an update to the CBP One Appendix, (b)(5)
(b)(5)

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

Thanks!

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
CBP Privacy Officer
Privacy and Diversity Office (PDO), Office of the Commissioner
U.S. Customs and Border Protection
1300 Pennsylvania Avenue NW, Room (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (m)

(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Thursday, May 20, 2021 1:29 PM
To: (b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: CBP One (b)(5)

Good afternoon. I'm writing to fill everyone in on our approach regarding this information collection. Short answer:

(b)(5), (b)(7)(E)

Please let me know if you have any questions.

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

U.S. Department of Homeland Security

U.S. Customs and Border Protection

Entry/Exit Transformation

John F. Kennedy International Airport

Delta Air Lines Post Deployment Site Visit

August 1 – August 2, 2017

Introduction

U.S. Customs and Border Protection (CBP) intends to demonstrate the initial implementation of the Traveler Verification Service (TVS) through the expansion of air exit capabilities at select airports. The limited expansion will demonstrate to airlines and airports how biometrics can be integrated into current boarding processes, provide real-time, centralized biometric matching capabilities, and record biometrically verified outbound departures in CBP systems. Specifically, live photos of passengers will be compared against the photos stored in CBP systems utilizing the flight departure manifest. While completing analysis of existing biometric exit experiments, CBP began the implementation of a biometric air exit field trial 2016 at an airport in partnership with a large air carrier.

Purpose

Delta Air Lines (DL) recently collaborated with CBP and Vision-Box[1] to test facial recognition as part of ongoing trials to implement biometrics at air exit.

On behalf of CBP Headquarters Office of Field Operations, (b)(7)(E) was contracted to perform a time and motion study and record observations of three international DL departures from John F. Kennedy International Airport (JFK) to Amsterdam on August 1 – 2, 2017. The purpose of the study at JFK was to determine the total boarding time of all passengers on the selected DL flights and the individual passenger process time at the facial recognition station. Information obtained from the post deployment study will be compared to metrics derived during the baseline study performed at JFK on March 21-23, 2017 to include the total flight process time and individual passenger process time for the agent to review the passport and scan the boarding pass.

Approach

For the post deployment time and motion study, a team of (b)(7)(E) analysts, escorted by CBP, were stationed in the gate (b)(7)(E) area of Terminal 4 to observe three Amsterdam flights (Table 1):

Table 1. Flights Observed During Post Deployment Time and Motion Study

Date	Flight #	Destination	Gate	Number of Passengers	Scheduled Departure Time
August 1, 2017	(b)(7)(E)	Amsterdam	(b)(7)(E)	(b)(7)(E)	08:30PM
August 2, 2017		Amsterdam			04:21PM
August 2, 2017		Amsterdam			10:15PM

[1] www.vision-box.com/solutions/bordercontrol/

While on site, (b)(7)(E)

(b)(7)(E)

Key metrics from the post deployment time and motion study will be compared to metrics derived from the March 2017 baseline site visit, in which ten flights were observed, including two DL departures to Amsterdam (Table 2). Comparison of findings are addressed in the Quantitative section beginning on page 14.

Table 2. Flights Observed During Baseline Time and Motion Study

Date	Flight #	Destination	Gate	Number of Passengers	Scheduled Departure Time
March 21, 2017	(b)(7)(E)	**Amsterdam Netherlands**	(b)(7)(E)		**05:21PM**
March 21, 2017		Frankfurt Germany			06:52PM
March 22, 2017		**Amsterdam Netherlands**			**05:21PM**
March 22, 2017		Paris France			07:00PM
March 22, 2017		Sau Paulo Brazil			08:10PM
March 22, 2017		London Heathrow United Kingdom			09:30PM
March 22, 2017		Tel Aviv Israel			11:37PM
March 23, 2017		Accra Ghana			05:02PM
March 23, 2017		Frankfurt Germany			07:30PM
March 23, 2017		Reykjavik Iceland			08:45PM

The metrics captured include:

(b)(7)(E)

(b)(7)(E)

Qualitative Analysis

Figure 1 below is a diagram of the departure gates at JFK Terminal 4. While on site, facial recognition was only utilized as a pilot for the daily outbound flights to Amsterdam at gate [redacted].

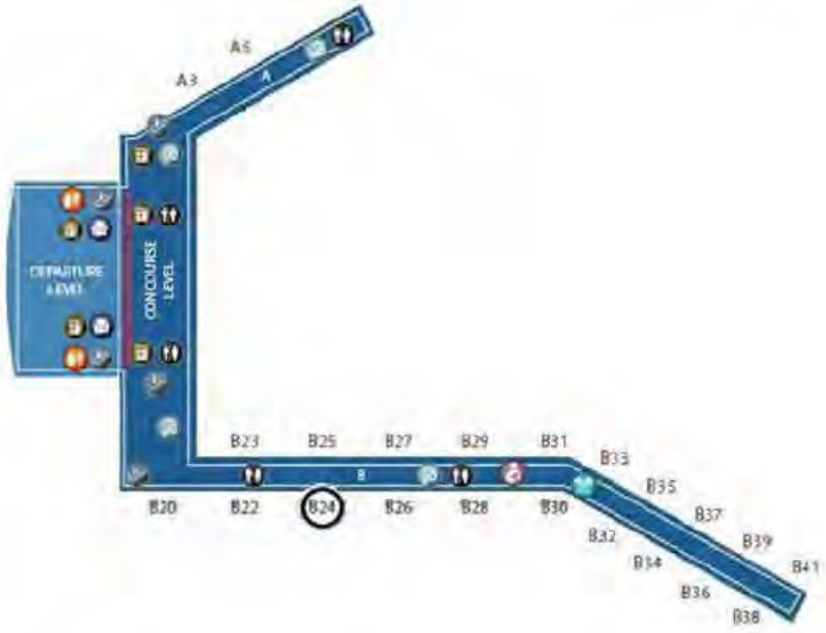


Figure 1. Layout of departure gates at JFK Terminal 4

Two "vb i-match™ eGates" were installed, and the facial recognition stations replaced one of the two stand-alone podiums adjacent to the customer service counter. Figure 2 represents the gate area before and after the installation of the facial recognition stations.

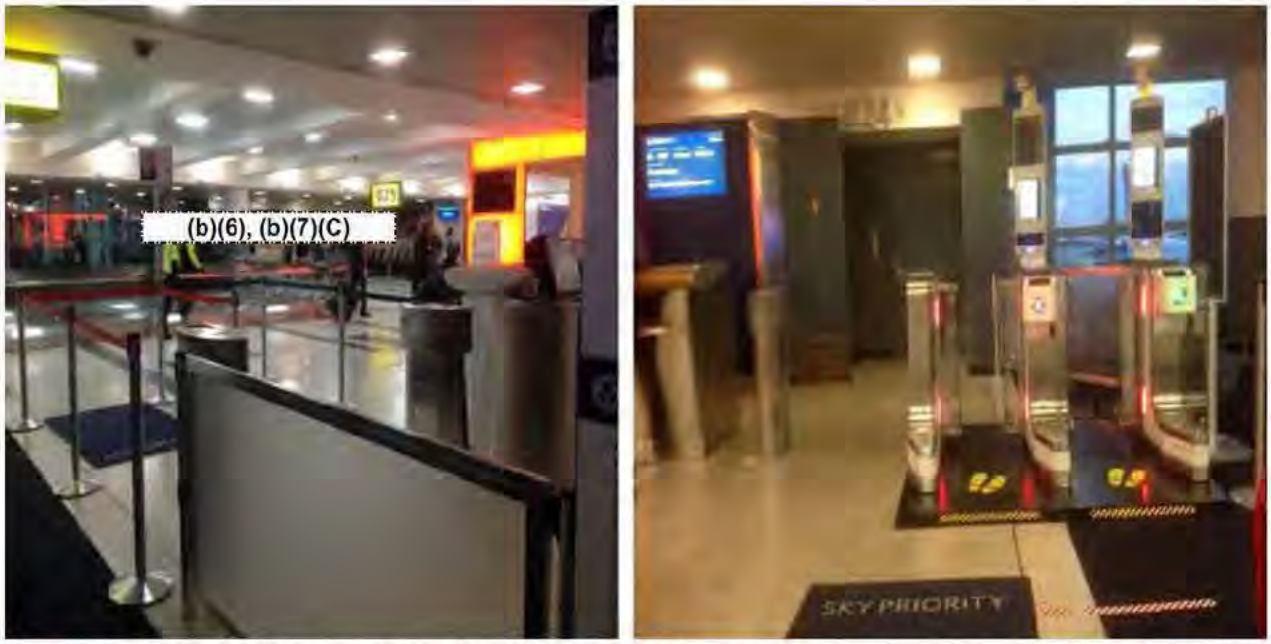


Figure 2. Left: View of two podiums baseline study; Right: Configuration of facial recognition station

According to a Vision-Box press release: "The vb i-match™ eGate ensures accurate ISO-compliant photo capture, assessed by built-in facial recognition analytical tools and assisted by dynamic height and light adjustments. The automated height feature adjusts to passengers of all heights, while the dynamic lighting system enables a high-quality image capture regardless ambient lighting within the terminal. The advanced telescopic doors ensure a smooth, secure and sequenced automated boarding process by managing a consistent passenger flow. The integration with Delta Air Line Inc.'s Departure Control Systems unleashes a cohesive boarding flow compliant with aviation industry standards, in order to optimize the on-boarding timing and allow the boarding agents to focus on the traveler experience[2]." Each facial recognition station contained a bar code scanner, adjustable camera, and eGate to board passengers (Figure 3).

[2]http://www.vision-box.com/news/vision-box-implements-a-facial-recognition-pilot-program-at-new-york-jfk-airport/

Figure 3. Facial recognition station at gate (b)(7)(E) at JFK Terminal 4

Presumably, the configuration of the podium and two facial recognition stations enabled three passengers to board simultaneously, rather than the two queues as observed during the baseline study.

(b)(7)(E)

For DL to operate the facial recognition stations, a CBP Officer was present at the gate while boarding outbound flights to Amsterdam. (b)(7)(E)

(b)(7)(E)

A CBP Officer was also stationed about 90 feet in the hallway of the sterile corridor to conduct and mitigate outbound operations (Figure 4). (b)(7)(E)

(b)(7)(E)

Figure 4. Approximate location of CBP Officer in the sterile corridor at gate (b)(7)(E) Terminal 4

Three to four agents were staffed at the customer service counter at least one hour prior to the Amsterdam flight departure to perform administrative tasks in preparation for boarding. A DL Supervisor was also on hand to operate the facial recognition stations during boarding.

A DL agent announced to all passengers waiting in the gate area the time boarding would begin. (b)(7)(E)

(b)(7)(E)

To board, one agent was staffed at the podium, adjacent to the customer service counter, to process passengers with special assistance needs and general boarding. A DL Supervisor was also stationed between the two facial recognition stations. The order of Delta Air Lines boarding was conducted in the following manner.

1) Pre-boarding for passengers that required special assistance, including passengers in wheelchairs
2) Premium Zone
3) Sky Priority Zone
4) Groups 1, 2, and 3, respectively
5) All remaining passengers

Prior to official boarding, one of the agents walked to the head of the queue and requested premium and sky priority passengers to form a single line to the right and general boarding to the left of the 8′ free-standing queue sign (Figure 5). (b)(7)(E)

(b)(7)(E)

Figure 5. Passenger queue at gate (b)(7)(E)

(b)(7)(E)

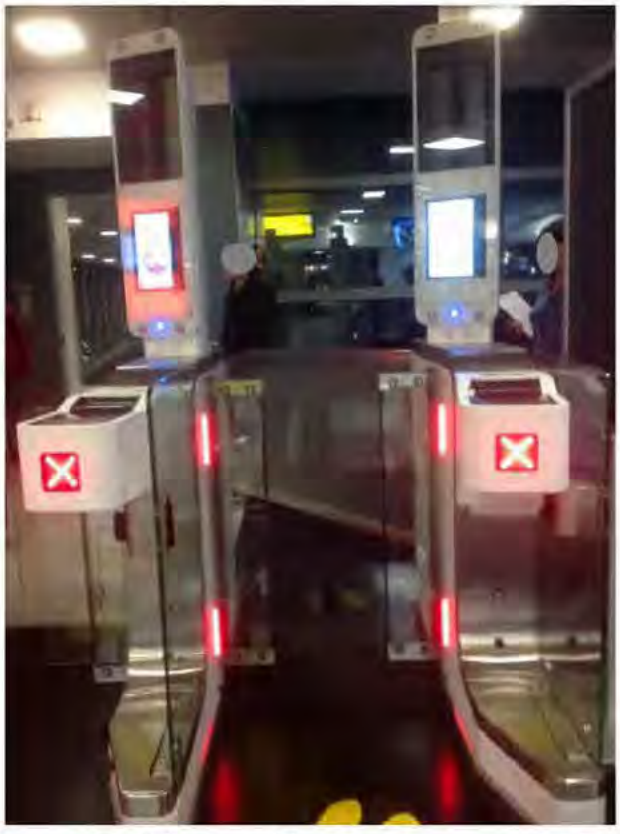


Figure 6. Facial recognition stations not operational

The DL agent posted at the facial recognition station motioned to the passenger at the head of queue to move forward and requested the passenger to place their feet on the symbols on the floor mat and position luggage outside the area of the facial recognition station (Figure 7).

(b)(7)(E)

(b)(6), (b)(7)(C)

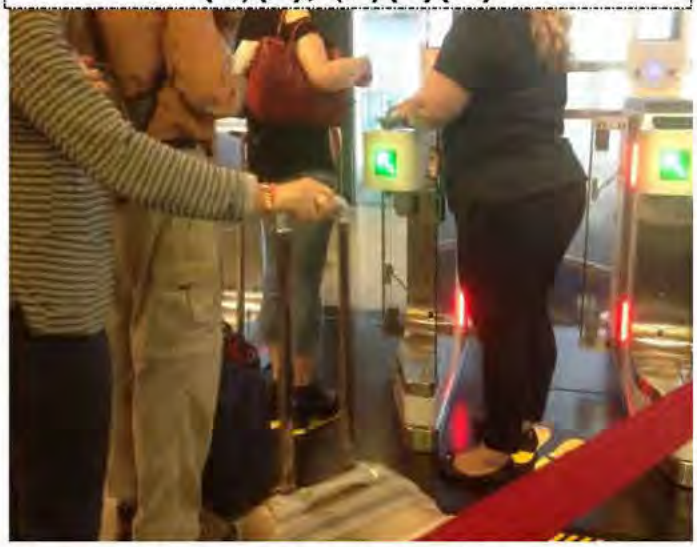


Figure 7. Left: Floor mat at facial recognition station; Right: DL agent processing passenger

(b)(7)(E)

Upon boarding, the facial recognition station contained a screen above the scanner that displayed an animation informing the passenger to position the boarding pass face down (Figure 8). The system then generated a green circle indicating the boarding pass was scanning.

Figure 8. Left: Display screen to illustrate boarding procedure; Right: Boarding scanner

As soon as the boarding pass was successfully read, the affixed camera automatically adjusted up and down to compensate for the passenger's height to capture the passenger's photo and compare it to CBP's gallery. However, the passenger's photo was captured whether the passenger was looking directly at the camera or not. Once the photo capture was complete, the gate automatically opened for boarding regardless of a match (Figure 9). After the eGate opened, passengers walked approximately 6' to access the sterile corridor and traversed down three ramps, approximately 270 ', before approaching the jet bridge.

Figure 9. Facial recognition stations not operational

(b)(7)(E)

(b)(7)(E)

(b)(7)(E) When a referral was generated, a red X was displayed at the facial recognition station (Figure 10).

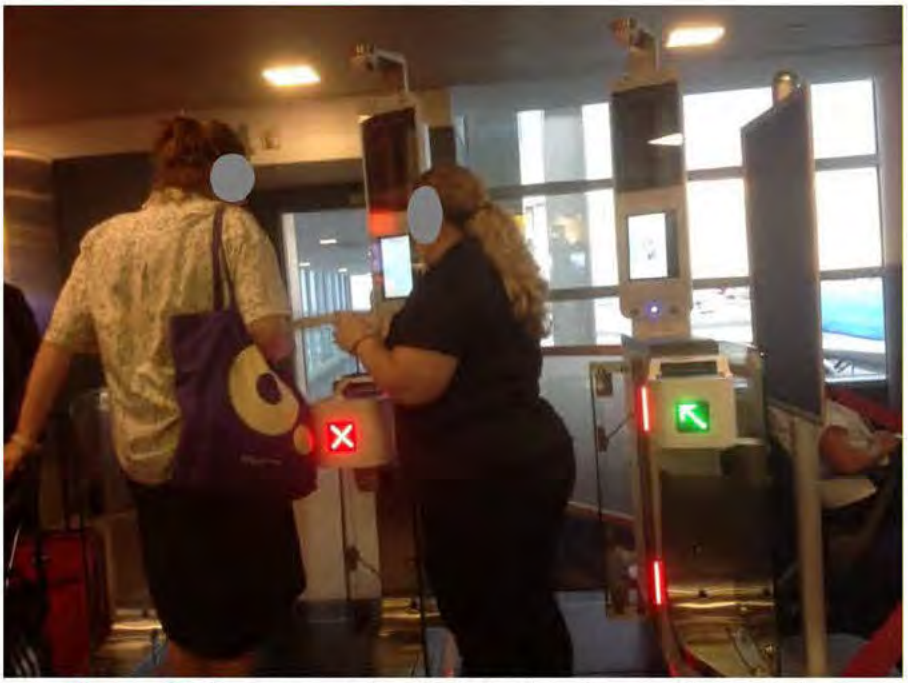


Figure 10. Passenger referral at the facial recognition station

Quantitative Analysis

Comparison of Baseline and Post Deployment Metrics

Data analysts were initially on-site March 21-23, 2017 to perform a baseline time and motion study for ten selected DL departures (Table 3). In this section, metrics from the baseline time and motion study are compared against metrics collected on-site August 1-2, 2017 during the post deployment facial recognition time and motion study of the three Amsterdam flights observed.

Table 3. Flights Observed During Baseline and Post Deployment Time and Motion Studies

Date	Flight #	Destination	Gate	Number of Passengers	Scheduled Departure Time
March 21, 2017	(b)(7)(E)	**Amsterdam**	(b)(7)(E)		05:21PM
March 21, 2017[3]		Frankfurt			06:52PM
March 22, 2017		**Amsterdam**			05:21PM
March 22, 2017		Paris			07:00PM
March 22, 2017		Sau Paulo			08:10PM
March 22, 2017		London			09:30PM
March 22, 2017		Tel Aviv			11:37PM
March 23, 2017		Accra			05:02PM
March 23, 2017		Frankfurt			07:30PM
March 23, 2017		Iceland			08:45PM
August 1, 2017		Amsterdam			08:30PM[4]
August 2, 2017		Amsterdam			04:21PM
August 2, 2017		Amsterdam			10:15PM

(b)(7)(E)

[3] Site team was unable to capture individual passenger process times due to procedural issue with Delta Air Lines.
[4] Baggage was loading at 08:30PM but site team did not remain on-site for actual departure, given available CBP resources.

A total of (b)(7)(E) passengers from three Amsterdam outbound flights were observed boarding using the facial recognition station during the two-day post deployment site visit. Of the (b)(7)(E) passengers observed, (b)(7)(E) passengers did not complete the process at the facial recognition station and were referred to a DL agent at the podium. The remaining (b)(7)(E) passengers were processed by a DL agent at the podium or customer service counter. *The passengers not processed at the facial recognition station are excluded from the individual post deployment metrics presented.* Table 4 lists the number of passengers observed that used the facial recognition station for each of the three flights.

Table 4. Delta International Boarding Metrics

Date	Flight #	Number of Passengers Processed at Facial Recognition[5]	Number of Passengers Processed at Podium or Counter	Total Number of Passengers
August 1, 2017	(b)(7)(E)			
August 2, 2017				
August 2, 2017				

The average passenger walk-up time to the facial recognition station was (b)(7)(E) seconds, which was similar to the average passenger walk-up time observed during the baseline study. The post deployment average passenger process time was (b)(7)(E) seconds, (b)(7)(E) seconds longer than the baseline average passenger process time of (b)(7)(E) seconds observed at the podium across all nine flights and (b)(7)(E) seconds higher than the two Amsterdam flights observed during the baseline study.

Table 5 compares passenger walk-up and boarding process times derived from the baseline and post deployment time and motion studies.

[5]The site team was unable to obtain the precise number of passengers that used the facial recognition stations but attempted to collect as many transactions as possible.

Table 5. Delta International Boarding Metrics

Metric Name	(b)(7)(E)
Passenger Boarding Walk Up Time – Baseline All Flights (seconds)	
Passenger Boarding Walk Up Time – Baseline Amsterdam Only (seconds)	
Passenger Boarding Walk Up Time – Post Deployment (seconds)	
Passenger Boarding Process Time – Baseline All Flights (seconds)	
Passenger Boarding Process Time – Baseline Amsterdam Only (seconds)	
Passenger Boarding Process Time - Post Deployment (seconds)	

The post deployment passenger process time of (b)(7)(E) seconds was derived by recording the following steps at the facial recognition station:

- Step 1 – Read Boarding Pass: Time passenger arrived at the facial recognition station until the boarding pass was scanned.
- Step 2 – Facial Recognition Complete: Time the photo capture and facial recognition was complete.
- Step 3 – Departure: Time the passenger departed after the gate was automatically opened to board.

Figure 11 presents a breakout, by percentage, of the average time spent by the passenger on each step of the boarding process. Two issues contributing to the amount of time to scan the boarding pass and successfully read the document were:

(b)(7)(E)

[6] Family of three required additional assistance at the podium after all general boarding passengers were processed.

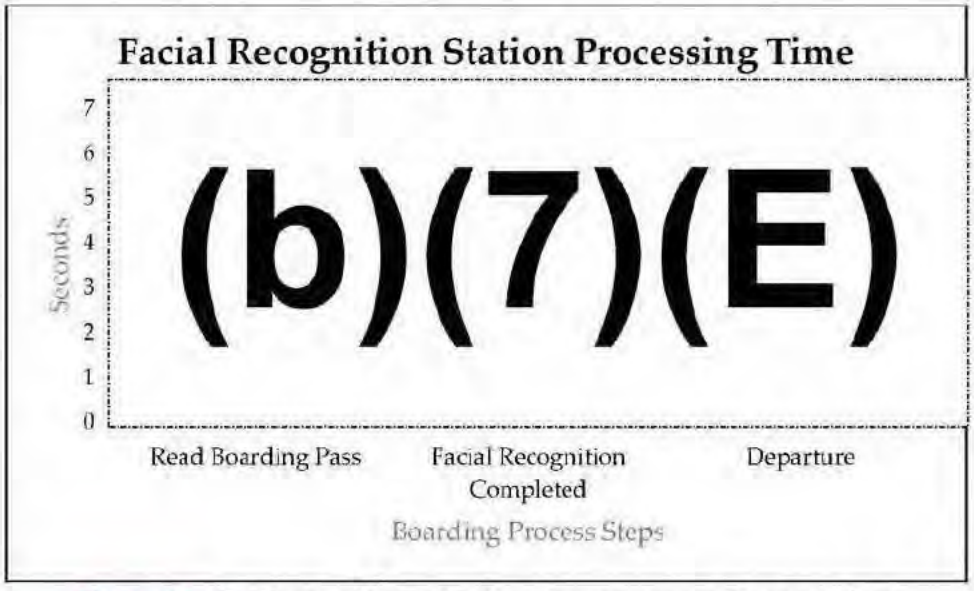


Figure 11. Breakout of post deployment passenger process time

(b)(7)(E)

(b)(7)(E) Table 6 compares the average cycle time and calculated throughput rate between the baseline and post deployment flights observed.

Table 6. Passenger Cycle Time and Throughput Rate

Metric Name and Definition	Baseline – All Flights	Baseline – Amsterdam Flights Only	Post Deployment – Amsterdam Flights Only
Average Passenger Boarding Cycle Time (seconds)	(b)(7)(E)		
Calculated Throughput (passengers per minute)			

The post deployment average passenger boarding cycle time at the facial recognition station was (b)(7)(E) seconds higher than the baseline average passenger process cycle time of (b)(7)(E) seconds for the two Amsterdam flights observed. Figure 12 is a comparison of the Amsterdam baseline and post deployment cycle time per passenger.

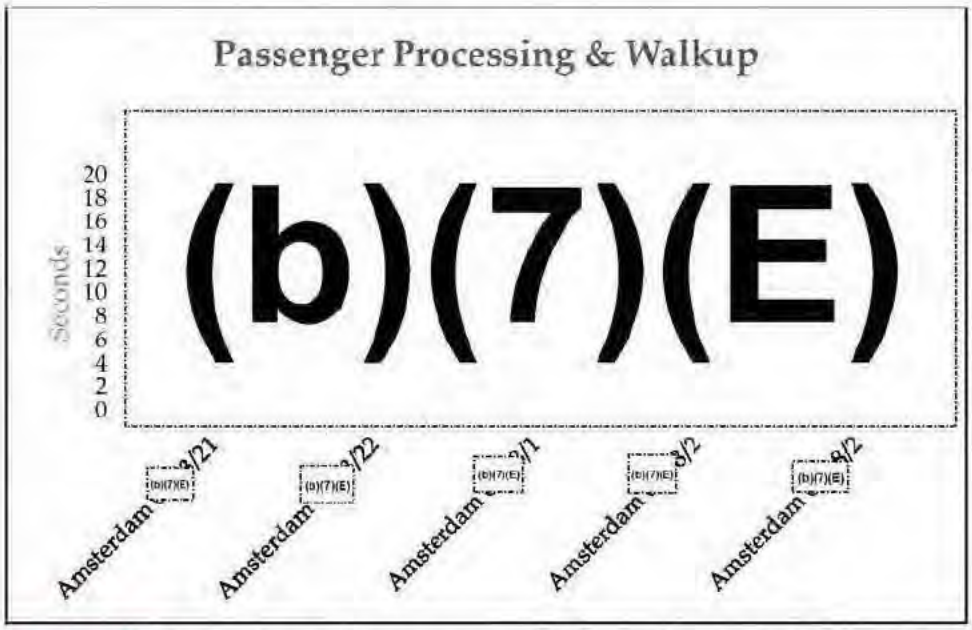


Figure 12. Comparison of passenger cycle time – Amsterdam departures

(b)(7)(E)

Figure 13 and Figure 14 illustrates the breakout of the walk-up time and boarding process time per passenger, based on the percentage of all passengers observed:

- (b)(7)(E) passengers boarded across the nine flights during the baseline site visit.
- (b)(7)(E) passengers boarded across the two Amsterdam flights during the baseline site visit.
- (b)(7)(E) passengers boarded across the three Amsterdam flights via the facial recognition station during the post deployment site visit.

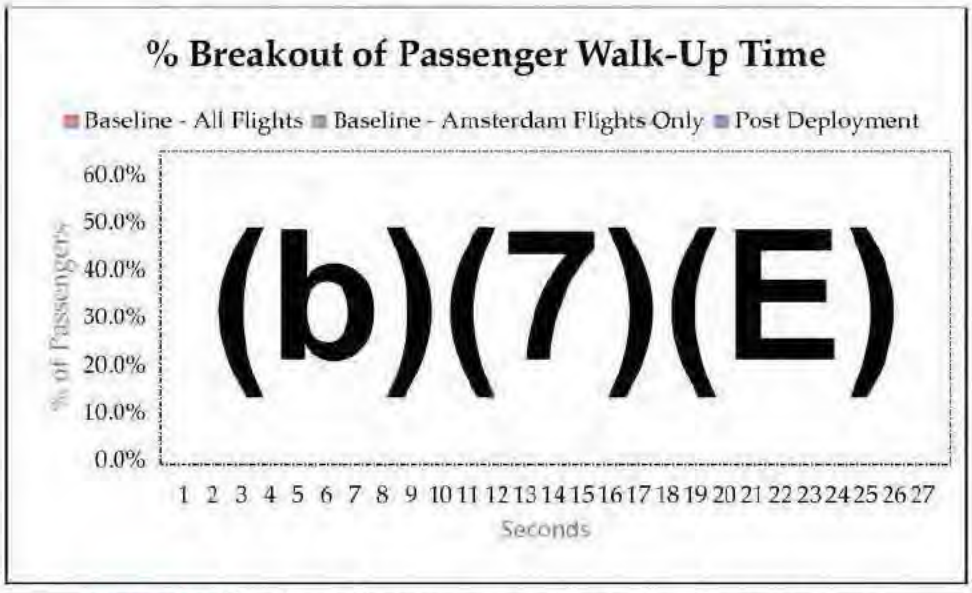


Figure 13. Breakout of passenger walk up time, by percentage

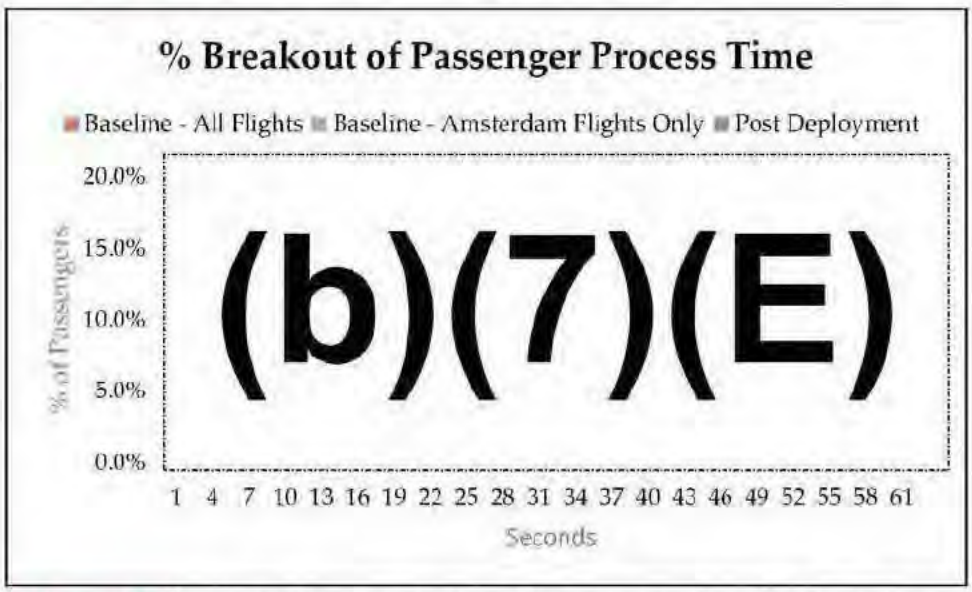


Figure 14. Breakout of passenger boarding process time, by percentage

Within the total flight boarding time, the site team also recorded the total amount of elapsed time between boarding groups, which included pre-boarding and general boarding for all groups and the time waiting for passengers that arrived late after general boarding (Table 7). The number of passengers confirmed on each flight was provided by a DL agent. Metrics were calculated as follows:

- Pre-boarding time is defined as the time the first passenger is processed for special assistance until the time all passengers eligible to pre-board have been processed at the customer service counter, agent podium (baseline) or facial recognition station (post deployment).

- General boarding time is defined as the time of the official DL announcement for the first general boarding Group 1 until the time all remaining passengers in the gate area through Group 3 have been processed at the customer service counter, agent podium, or facial recognition station.
- Late arrival boarding time is defined as time the last passenger waiting in the gate area in the general boarding group has been processed until the end time of the last passenger arriving late is processed.
- Total flight boarding time is defined as the time the first passenger was processed by the airline agent or at the facial recognition station until the last passenger was processed for each flight observed.

Table 7. Total Boarding Time by Flight

Date	Flight #	Destination	Priority/ General Boarding Time (mm:ss)	Priority/ General Boarding Passenger Count	Late Arrival Boarding Time (mm:ss)	Late Arrival Boarding Passenger Count	Total Flight Boarding Time (mm:ss)
March 21, 2017	(b)(7)(E)	**Amsterdam**	(b)(7)(E)				
March 21, 2017[7]		Frankfurt					
March 22, 2017		**Amsterdam**					
March 22, 2017		Paris					
March 22, 2017		Sau Paulo					
March 22, 2017		London					
March 22, 2017		Tel Aviv					
March 23, 2017		Accra					
March 23, 2017		Frankfurt					
March 23, 2017		Iceland					
August 1, 2017		Amsterdam					
August 2, 2017		Amsterdam					
August 2, 2017		Amsterdam					

[7] Site team was unable to capture individual passenger process times due to procedural issue with Delta Air Lines.

The recorded total flight boarding time of (b)(7)(E) for the Amsterdam flight observed March 21, 2017 during the baseline study is relatively the same as the total flight boarding time of (b)(7)(E) for the Amsterdam flight observed August 1, 2017. (b)(7)(E)

(b)(7)(E)

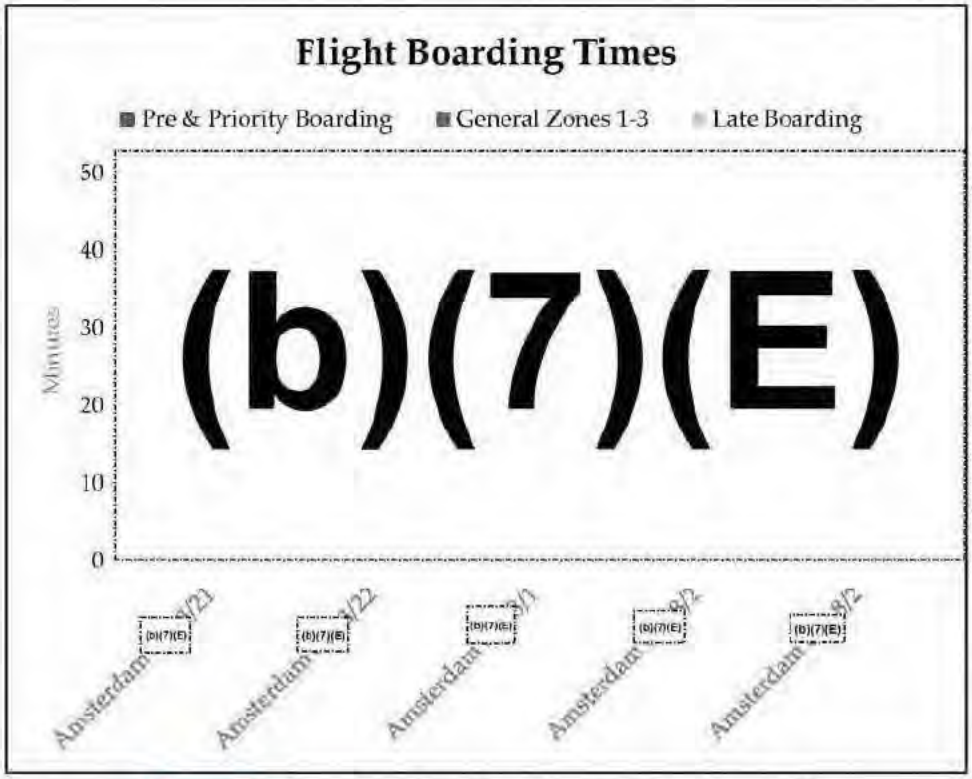


Figure 15. Boarding time by groups and late arrivals

In addition, the site team recorded the time the first and last passenger was processed, the time the jet bridge doorway was closed, and the time the last passenger stepped onto to the aircraft for boarding (Table 8). (b)(7)(E)

(b)(7)(E)

Table 8. Key Flight Boarding Metrics

Departure Date	Flight #	First Passenger Processed	Time Last Passenger Processed	Last Passenger Enters Aircraft	Time Door to Jet Bridge Closed	Time Aircraft Pushed Back
March 21, 2017	(b)(7)(E)					
March 22, 2017						
March 22, 2017						
March 22, 2017						
March 22, 2017						
March 23, 2017						
March 23, 2017						
March 23, 2017						
August 1, 2017						
August 2, 2017						
August 2, 2017						

Table 9 lists the percentage of passengers, with selected attributes, observed during the baseline and post deployment time and motion studies. Passengers in wheelchairs were processed at the podium.

At the facial recognition station, the system captured the passenger's photo regardless if they were not looking directly at the camera. (b)(7)(E)

(b)(7)(E)

Table 9. Passenger Attributes

Attribute	Baseline – All Flights	Baseline – Amsterdam Flights Only	Post Deployment – Amsterdam Flights Only
(b)(7)(E)			

(b)(7)(E)

Post Deployment Metrics Only

Metrics by Flight Number

Figure 16 displays the individual walk-up time and processing time per passenger, exclusively at the facial recognition station, for each of the three DL Amsterdam flight departures observed during the post deployment site visit. (b)(7)(E)

(b)(7)(E)

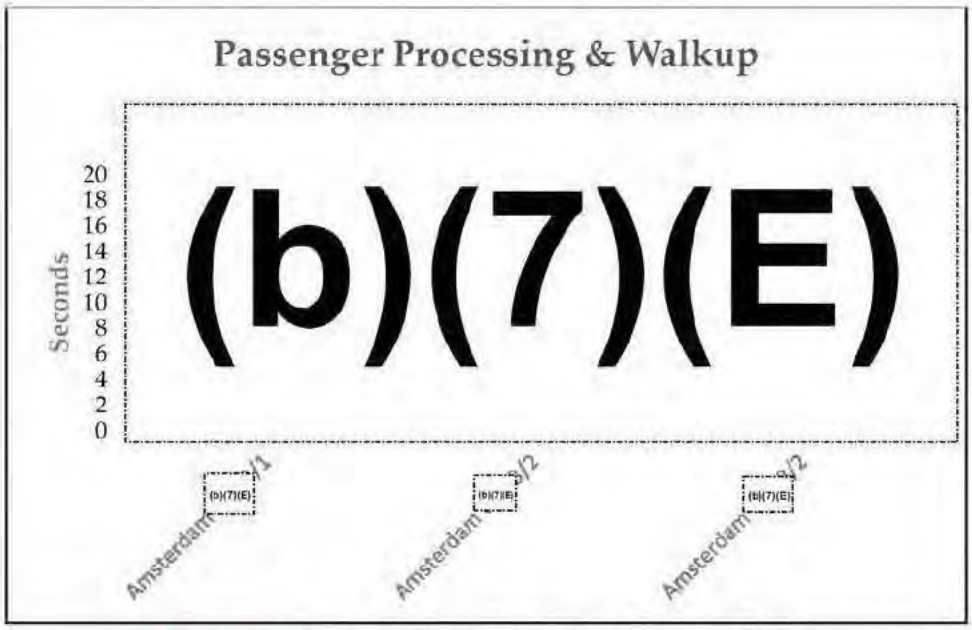


Figure 16. Average passenger cycle time by flight

(b)(7)(E)

Table 10. Total Flight Process Time by Boarding Group[8]

Date	Flight #	Destination	Pre-Boarding Time (minutes)	General Boarding Time (minutes)	General Passenger Count	Late Arrival Boarding Time (minutes)	Late Arrival Passenger Count
August 1, 2017	(b)(7)(E)	Amsterdam	(b)(7)(E)				
August 2, 2017		Amsterdam					
August 2, 2017		Amsterdam					

[8] See page 20 for metrics for total flight process time by boarding group.

Based on observations, total flight process time (Figure 17) (b)(7)(E)

(b)(7)(E)

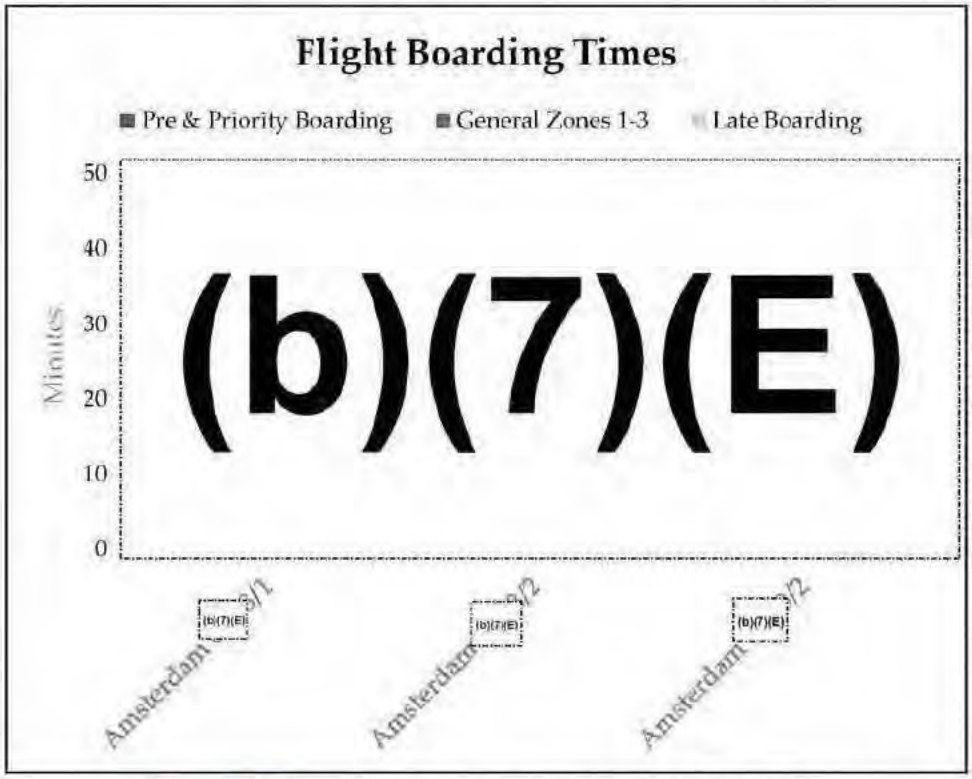


Figure 17. Boarding time by groups and late arrivals

Message

From: (b)(6), (b)(7)(C)

Sent: 6/10/2021 2:05:34 PM

To: (b)(6), (b)(7)(C)

Subject: RE: CBP One (b)(5)

(b)(5), (b)(7)(E)

Respectfully,

(b)(6), (b)(7)(C)
U.S. Customs and Border Protection Officer
(A) Branch Chief, Enforcement Programs Division
Office of Field Operations
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)

Sent: Thursday, June 10, 2021 9:53 AM

To (b)(6), (b)(7)(C)

Subject: FW: CBP One (b)(5)

(b)(5), (b)(7)(E)

Thank you.

(b)(6), (b)(7)(C)
Program Manager, Strategic Transformation Office
Planning, Program Analysis, and Evaluation
Office of Field Operations
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Thursday, June 10, 2021 9:45 AM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: CBP One (b)(5)

Good morning,

(b)(6), (b)(7)(C)

Could we have a check-in on this on Monday or Tuesday? (b)(5), (b)(7)(E)
(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Thank you!

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
CBP Privacy Officer
Privacy and Diversity Office (PDO), Office of the Commissioner
U.S. Customs and Border Protection
1300 Pennsylvania Avenue NW, Room (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (m)

(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 8, 2021 1:28 PM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: CBP One (b)(5)

(b)(5), (b)(7)(E)

(b)(6), (b)(7)(C) | U.S. Customs and Border Protection (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 8, 2021 11:00 AM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: CBP One (b)(5)

(b)(5), (b)(7)(E)

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)

Cell: (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, June 8, 2021 10:41 AM
To: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: CBP One (b)(5)

Good morning!

We are moving along with an update to the CBP One Appendix, and a standalone PIA for the Advanced Arrival of Undocumented Individuals process.

(b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

Thanks!

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
CBP Privacy Officer
Privacy and Diversity Office (PDO), Office of the Commissioner
U.S. Customs and Border Protection
1300 Pennsylvania Avenue NW, Room (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (m)

(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Thursday, May 20, 2021 1:29 PM
To: (b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C) DAVIES, MATTHEW S (b)(6), (b)(7)(C)
DURST, CASEY OWEN (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C) (OCC)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)

Subject: CBP One (b)(5)

Good afternoon. I'm writing to fill everyone in on our approach regarding this information collection. Short answer:

(b)(5), (b)(7)(E)

Please let me know if you have any questions.

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

CBP One™

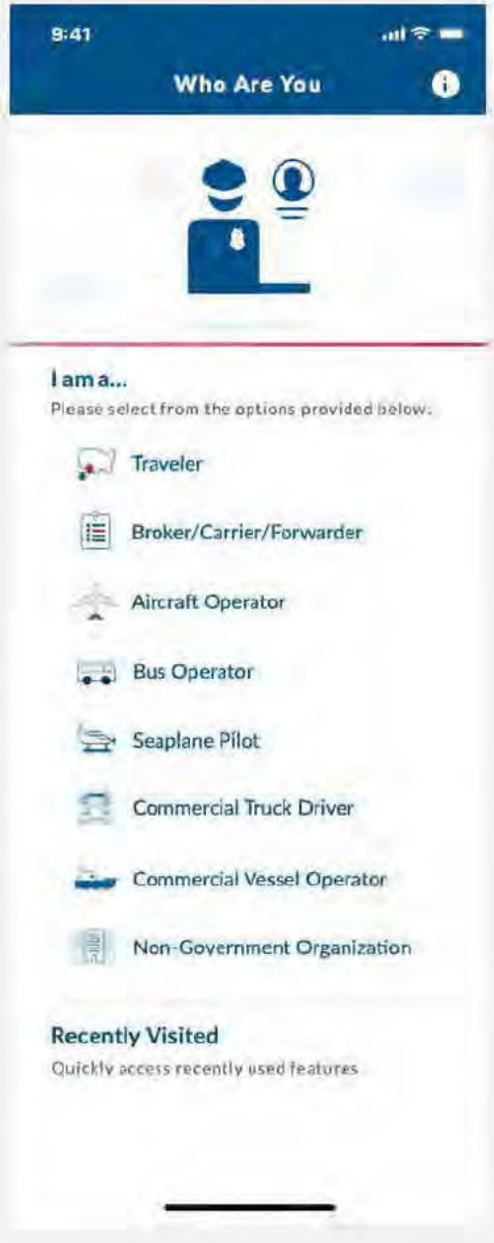
MPP Workflow – Check Case Status

February 9, 2021

U.S. Customs and Border Protection

Background

- U.S. Customs and Border Protection built a mobile application to serve as the single point of entry for travelers and stakeholders to access CBP mobile applications and services.
- Through a series of intuitive questions, the app will guide each type of user to the appropriate services based on their particular needs.
- CBP is currently available on the Apple App Store and Google Play Store with limited functionality

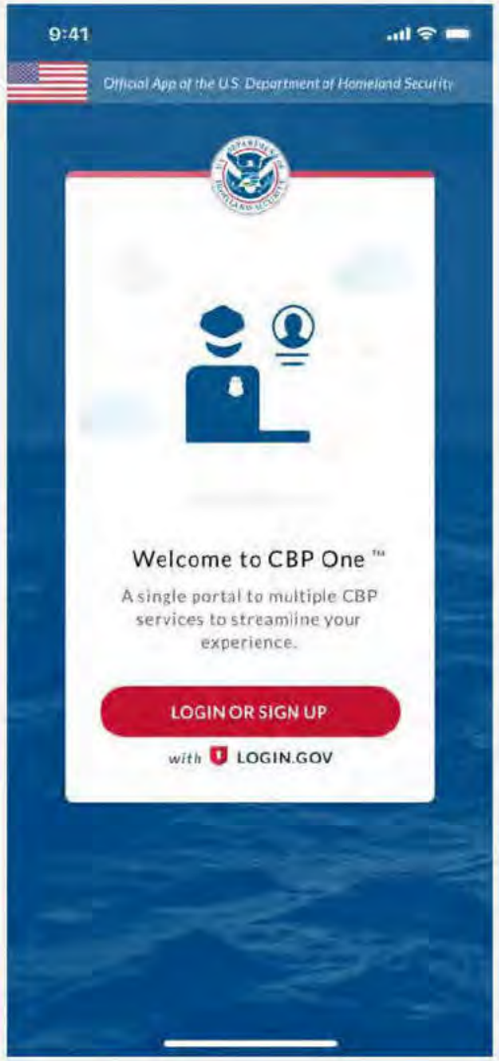


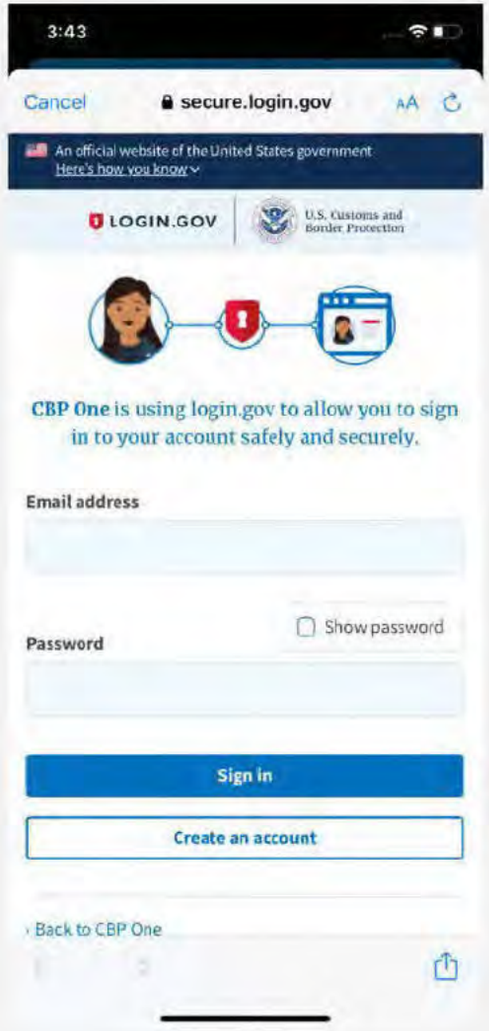
CBP One Screens – Login.gov

1. **Use your organization email.**
    - A personal email will not work for MPP

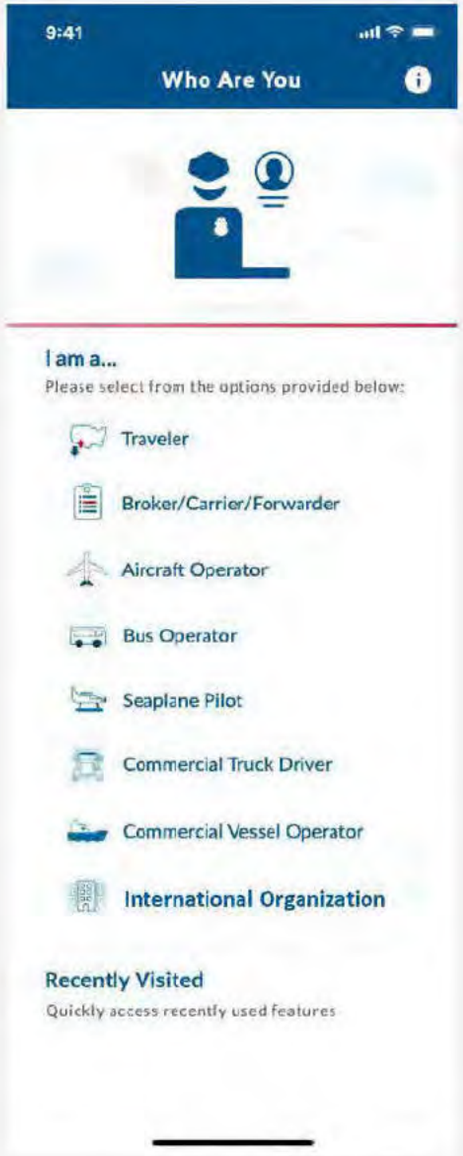
2. **Enter a password**

3. **Select one or more authentication methods such as:**
    - **More secure**
        - Security Key
        - Authentication application
    - **Less secure**
        - SMS/Text messages
        - Backup codes





CBP One - **Home Screen**



CBP One – **Check Case Status > Take Photo > Photo Capture – Best Practices**

(b)(7)(E)

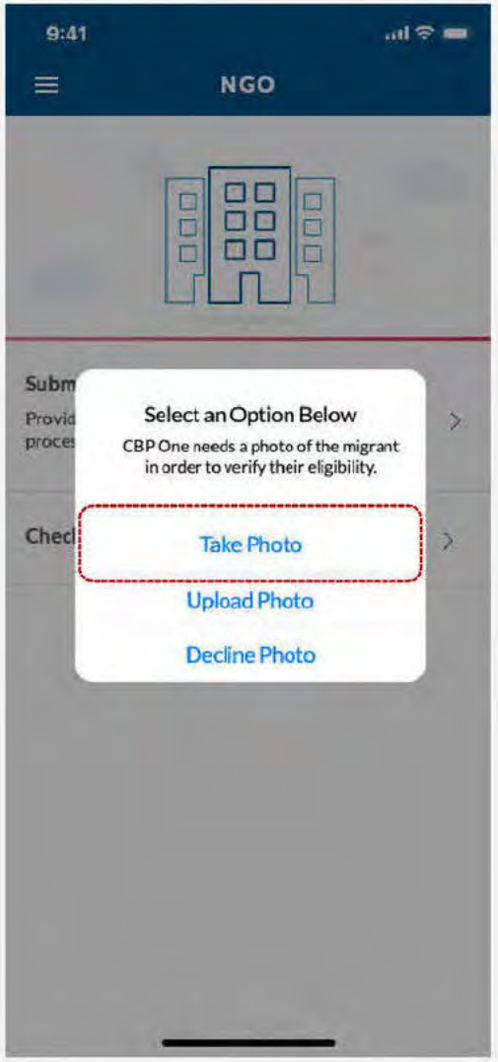
(b)(6), (b)(7)(C)

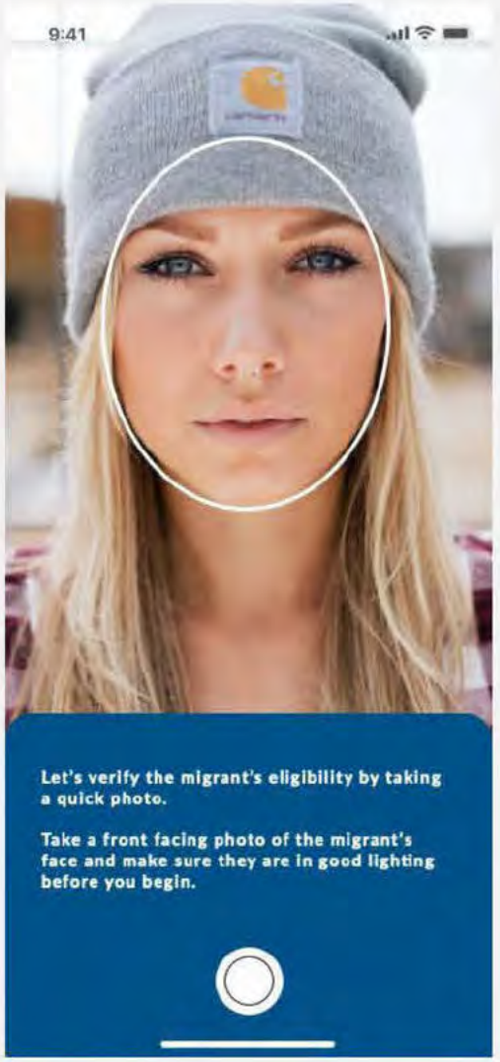
(b)(6), (b)(7)(C)

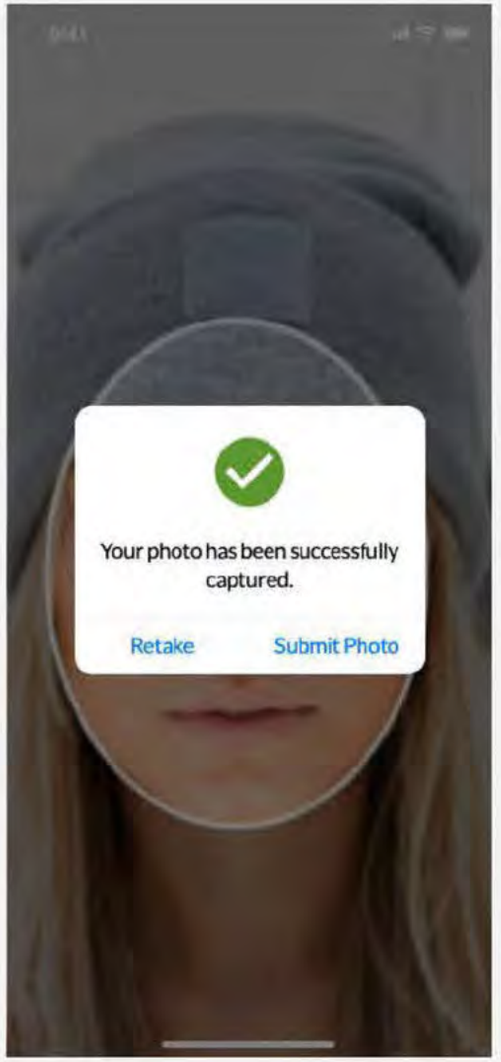
Ask travelers to **remove hats or glasses** when possible.

COVID Protections - If safe to do so, with sufficient safeguards and distancing (**use camera zoom**), you may ask the traveler to pull the face mask away to capture the photo and then reapply the covering.

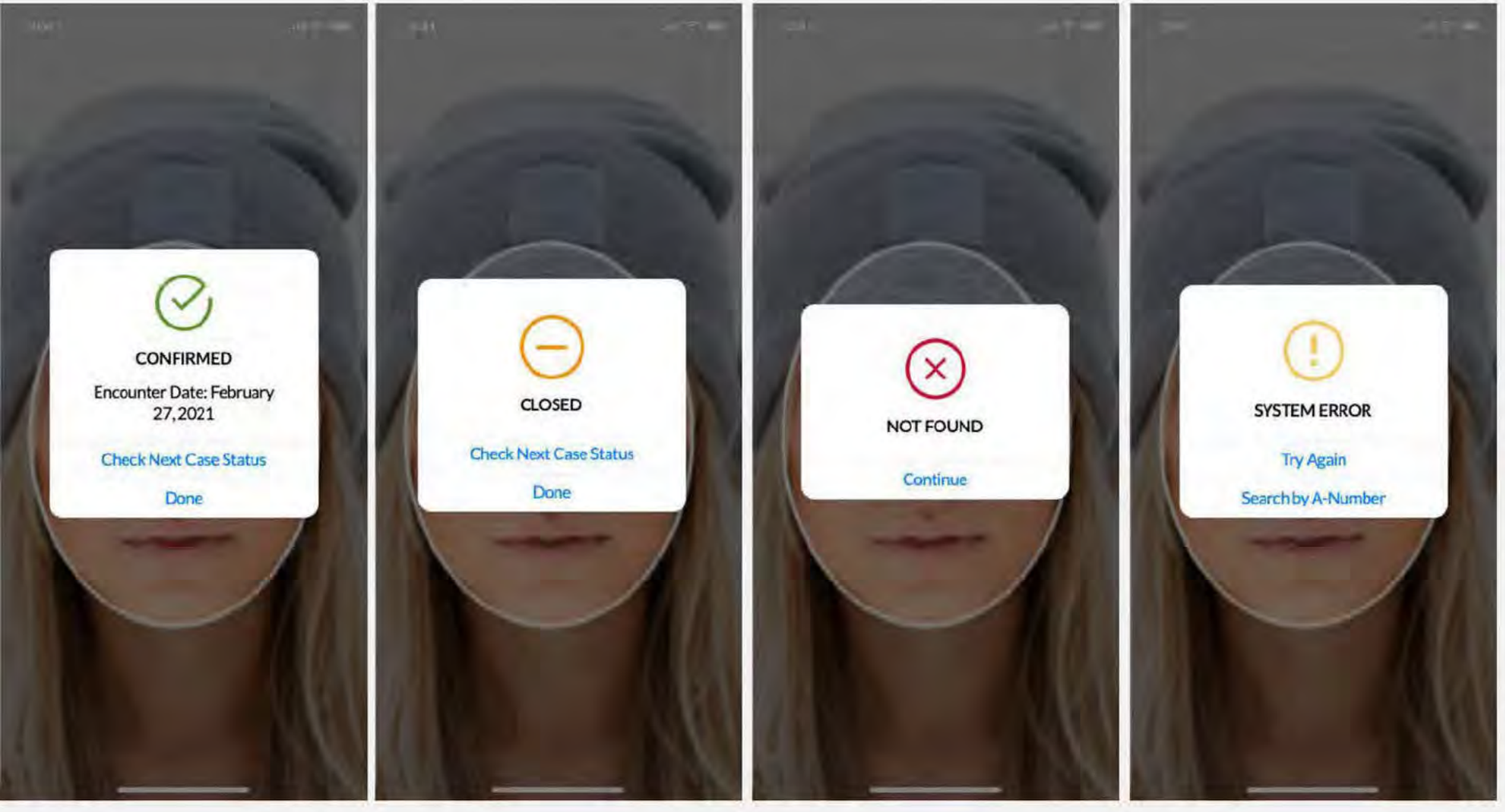
CBP One – **Check Case Status > Take Photo**







CBP One – **IO > Check Case Status > Photo Results**



If photo match <u>is found</u>:

- **Green light** message with **Encounter Date** will display;
- Select "done" and exit back to the IO home screen; or
- Select "Check Next Case Status" to go to the next case.

If photo match <u>is found but the case if closed on not found</u>:

- **Yellow Light** will display; and
- Select "Check Next Case or "Done"

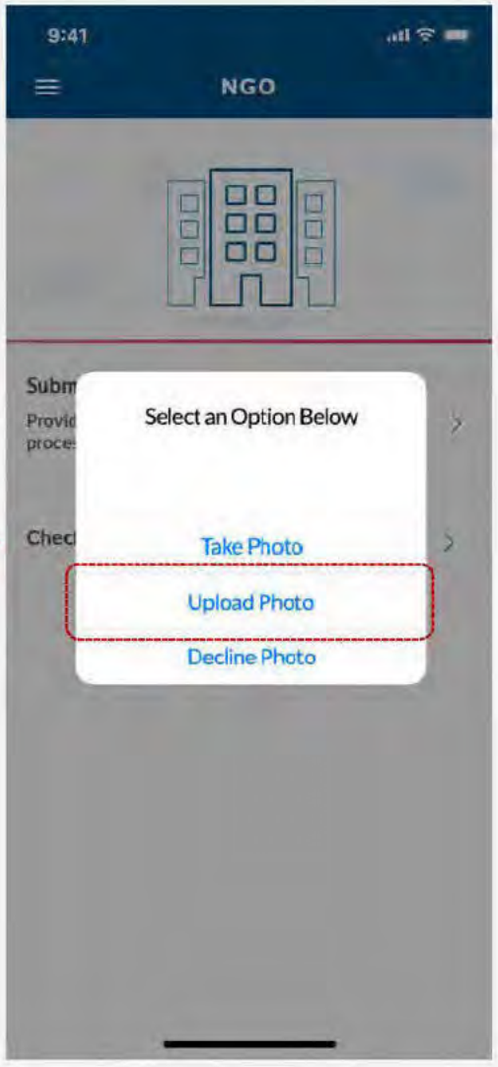
If photo match is <u>not found</u>:

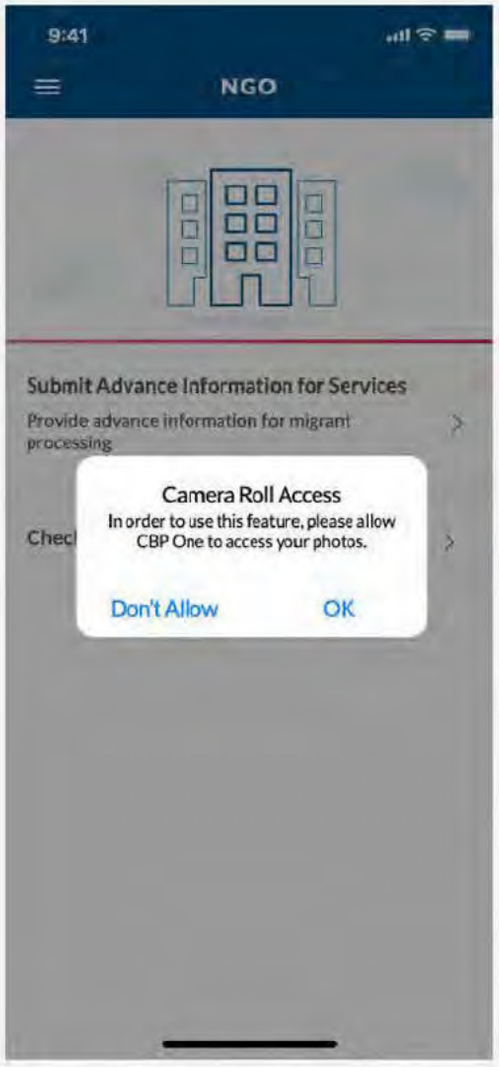
- **Red-light** message will display; and
- Select "Continue" to enter an A-Number.

An <u>error message</u> will pop up if there is a system error:

- User can try again; or
- Search by A-number.

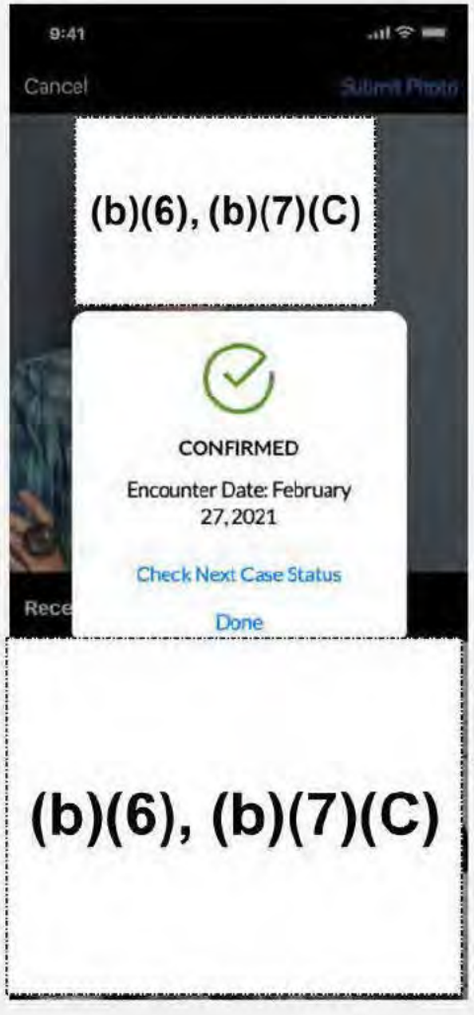
CBP One – **Check Case Status > Upload Photo**



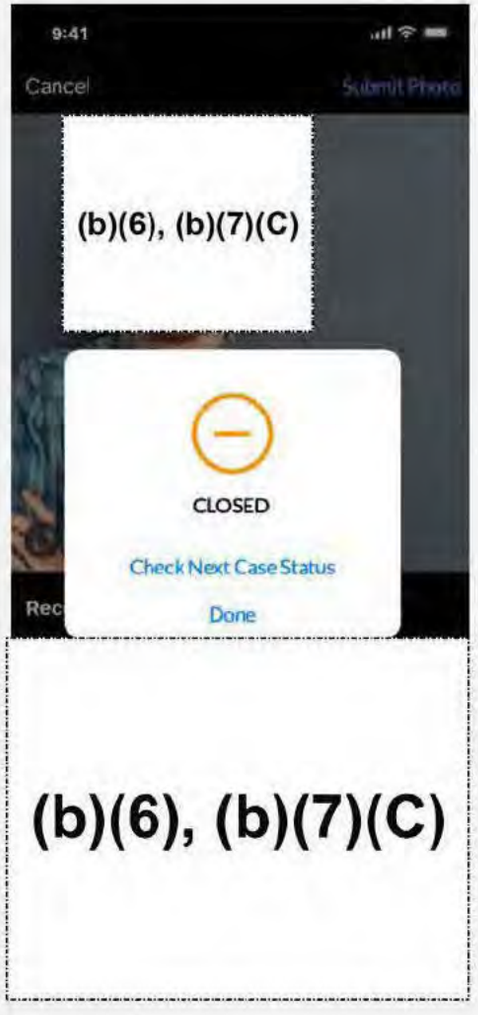


First time user would have to allow CBP One to access their camera roll

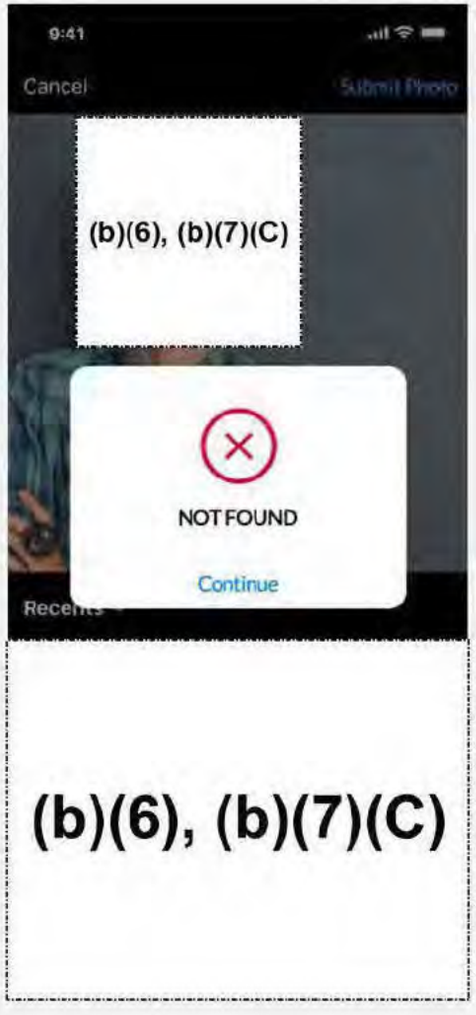
CBP One – **Check Case Status > Upload Photo**



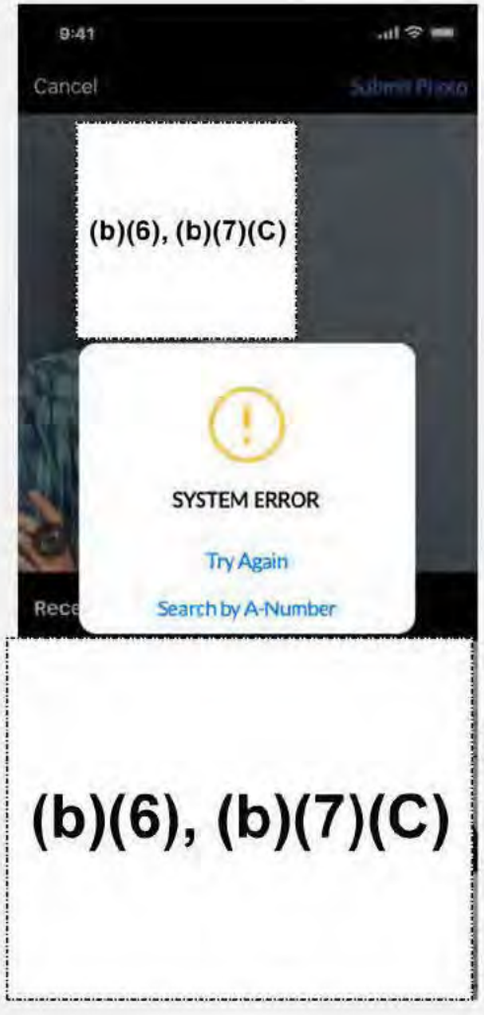
(b)(6), (b)(7)(C)



(b)(6), (b)(7)(C)

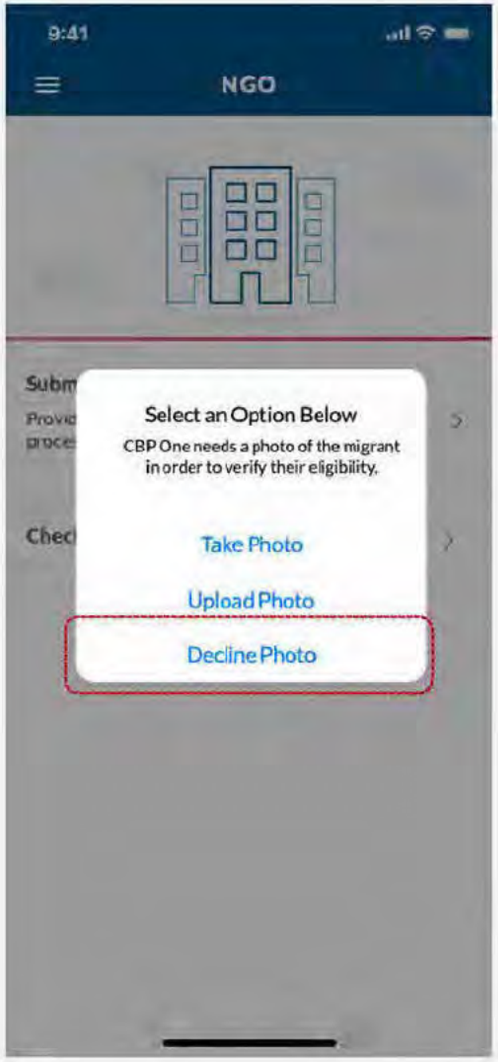


(b)(6), (b)(7)(C)



(b)(6), (b)(7)(C)

CBP One – **Check Case Status > Decline Photo**



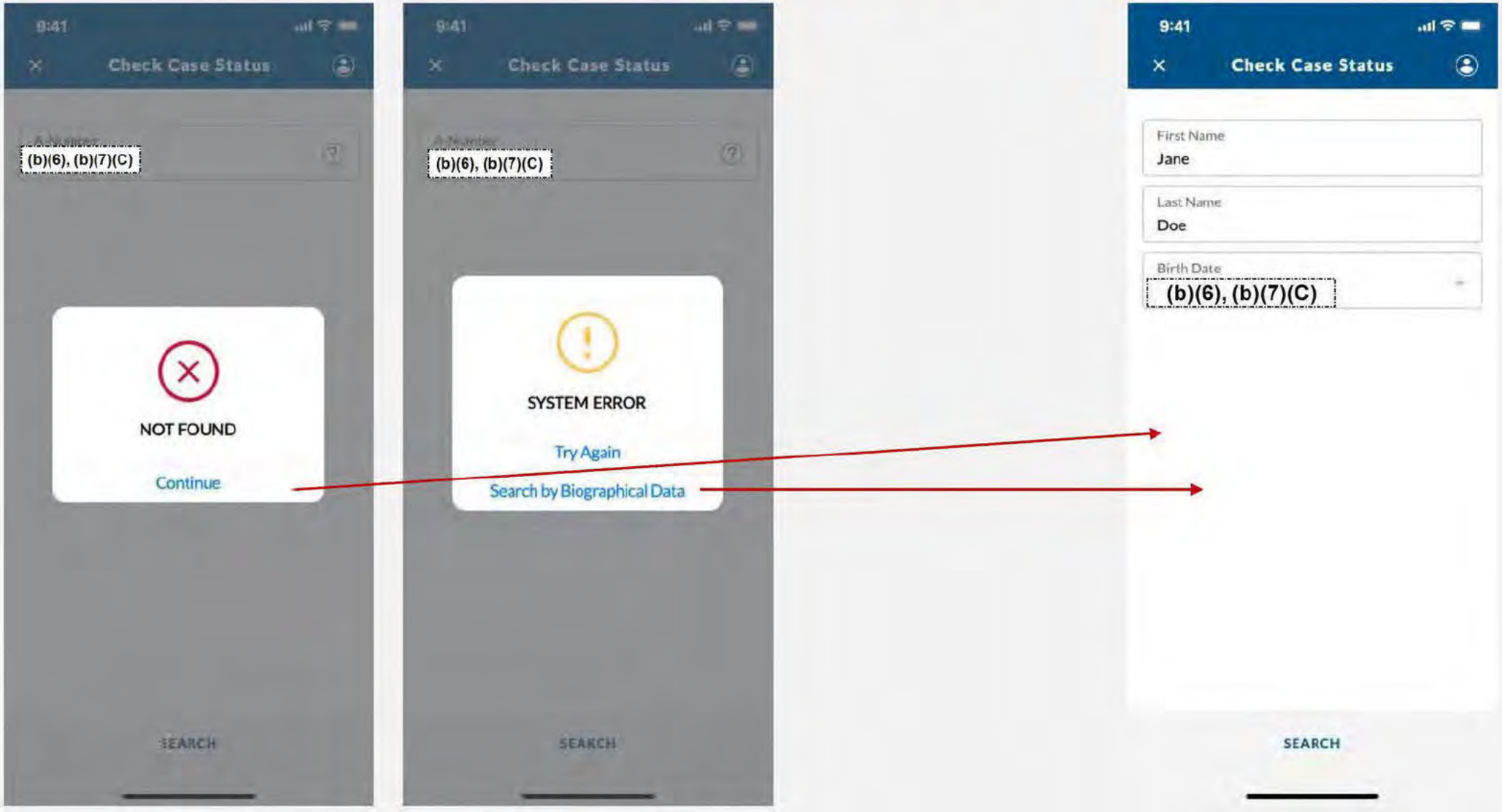
If user selects "**Decline Photo**", then the app will take them directly **to search by their A-number.**

CBP One – **NGO > Check Case Status > Query Results after A-Number**

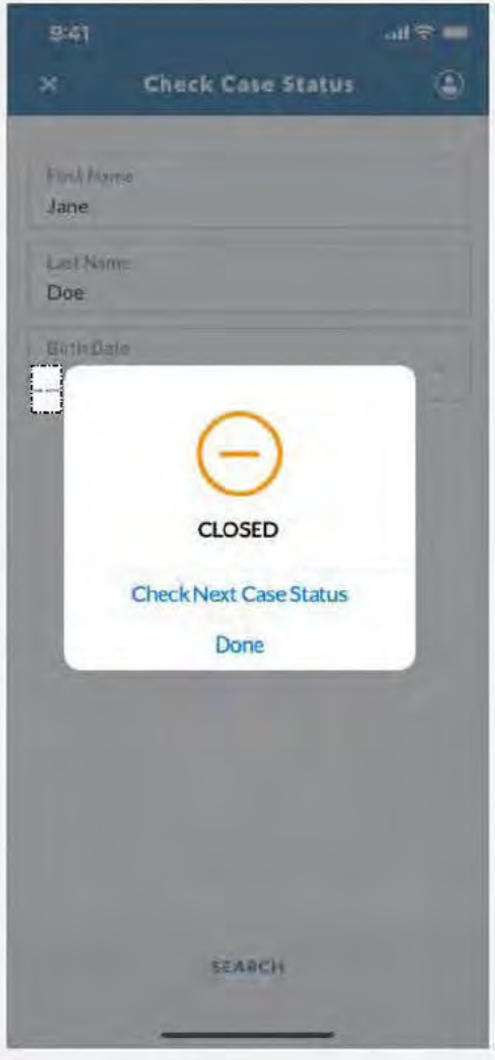
Potential Enhancement

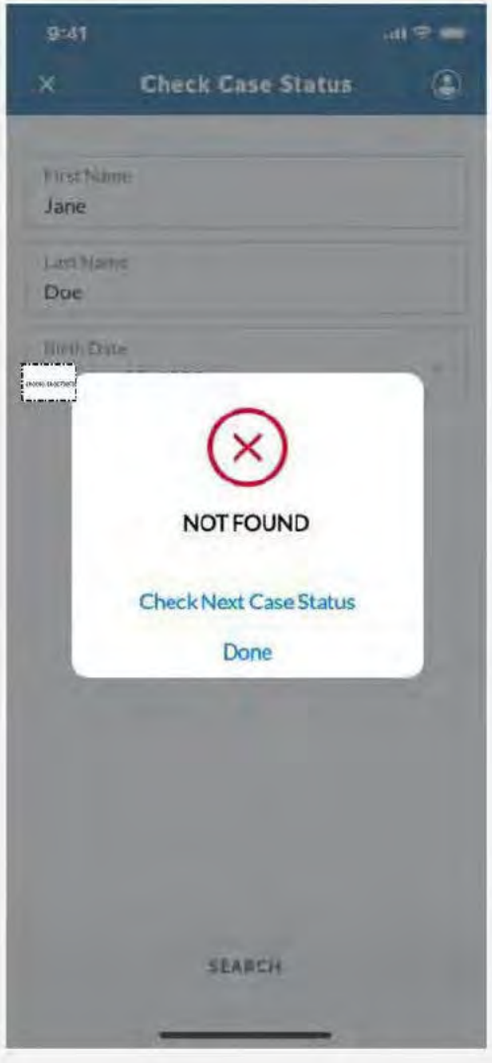
If photo capture and A-number do not pull a result,
user can do a third search using their first name, last name, and date of birth.

CBP One – **IO > Check Case Status > Search by Biographical Data**



CBP One – **NGO > Check Case Status > Query Results after Biographical Data**





October 20, 2020

U.S. Customs and Border Protection

MEMORANDUM FOR: Directors, Field Operations
Executive Directors
Office of Field Operations

FROM: William F. Ferrara (b)(6), (b)(7)(C)
Executive Assistant Commissioner
Office of Field Operations

SUBJECT: Implementation of CBP One

The Office of Field Operations (OFO) Innovation Center (IC) has formally launched the development of **CBP One**™, a mobile app intended to act as an intuitive single portal for travelers and stakeholders to access CBP mobile apps and services such as CBP ROAM, I-94 and Appointment features. Through a series of guided questions, the user is directed to the appropriate services based on their needs.

The CBP One app will be rolled out in phases, starting by offering stakeholders the ability to schedule agriculture inspections for perishable cargo and to apply and pay for an I-94 at land ports of entry (POEs). The second phase will allow (b)(5) (b)(5) with the new capability to apply for and update cruising licenses. (b)(5) (b)(5) Additional features such as landing rights applications and diversion notifications are planned for implementation in 2021.

The first phased will launch on October 28, 2020. It will offer be a mobile version of the current I-94 website (via CBP One) that will allow users to apply and pay for a provisional I-94. There will be no change to the current operating procedures at the POEs for issuance of I-94s. If the traveler has paid through the app or the current site, a pop-up will appear that indicates a provisional I-94 is on file. Once the inspection is completed and the traveler is admitted into the proper class of admission an electronic I-94 will be issued and will be immediately available in the TECS I-94 database and on I-94 website.

The mobile ability to request an inspection of perishable cargo will also be available. This capability was developed based on an idea from a Shark Tank innovation event hosted by the Miami Field Office. The CBP One feature allows stakeholders to schedule inspections during port designated hours. POE personnel manage these requests through a dashboard with the

ability to review details and assign Agriculture Specialists while offering interactive messaging and live status updates to users. This feature has been piloted in Miami since August 2020. A demo of the feature can be accessed here: CBP One - Appointment App

If you are interested in offering stakeholders the ability to coordinate perishable cargo examinations or have other scheduled examinations that you feel could benefit from this app, please provide a specific point of contact to (b)(6), (b)(7)(C) Program Manager at (b)(6), (b)(7)(C) by October 30, 2020.

If you would like to learn more or have any questions about the new CBP One app, please reach out to Strategic Transformation Office Director (b)(6), (b)(7)(C) for more information.

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Estimated Time Savings for CBP One								
	CBP			Travelers/Stakeholder				
	Hours	Dollars	(b)(5)	Hours	Dollars	(b)(5)		

(b)(5), (b)(7)(E)

U.S. Customs and Border Protection
Office Field Operations
February 12, 2021
10:00 – 10:40 EST
MPP CBP IO One Application Overview

Hosts:
Office of Field Operations (OFO) Admissibility and Passenger Programs (APP) and Systems Enforcement Analysis & Review (SEAR)

Meeting Notes:

- APP and SEAR hosted a joint overview of the CBP One IO App showcasing the modified technology to the existing CBP One application which will allow the International Organizations (IOs) to interface with CBP systems regarding case status of those migrants enrolled in the Migrant Protection Protocol (MPP) process.

- CBP utilizes the facial recognition technology as the interactive platform for the IO's to work with the active MPP population. This streamlined process ensures data and identity of the MPP enrollee are validated.

- Validation on the front end allows for future increased application capacity.
  - (b)(7)(E) photo from gallery is used from a smartphone device, not stored on CBP's device and the IO can check the MPP enrollee's status.
  - Alien registration number (A#) can also be utilized, if known, instead of photo.

- CBP One IO app return indicators:
  - Green - MPP enrollee has an active case and cleared for processing at a POE.
  - Red - MPP enrollee's case may be closed or false negative.
    - If red X is returned, the IO or MPP enrollee will need to provide supporting documentation, call the Executive Office of Immigration Review (EOIR) hotline, or check the EOIR system for additional information.
  - Yellow - there is a system error. Begin verification process again or contact EOIR.

Comments and Issues addressed:

- The CBP One IO capability should go live today or no later than Monday February 15, 2021.
- Processing will be used with A# and biographical information if needed.
- (b)(5)
  - (b)(5)
- CBP's technology division continues to pursue opportunities for data batching and CBP One IO application will continue to transform and expand capabilities.
  - The expansion will take a while longer as CBP One was not initially designed for this particular process, but OFO is confident it can be accomplished.
- APP and SEAR will distribute the CBP One IO App Powerpoint presentation and one page write up.

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Message

From: (b)(6), (b)(7)(C)
Sent: 8/31/2022 5:50:35 PM
To: (b)(6), (b)(7)(C)
Subject: FW: (b)(5) for CBP One
Attachments: 60FRN 1651-0140 Collection of Advance Information from Certain Undocumented Individuals on the Land Border.pdf; (b)(5), (b)(7)(E)
(b)(5), (b)(7)(E)

(b)(6), (b)(7)(C)
Chief Economist
Office of Trade: Regulations & Rulings
U.S. Customs & Border Protection
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, November 3, 2021 12:27 PM
To:
Cc: (b)(6), (b)(7)(C)
Subject: RE: (b)(5) for CBP One

Hi, (b)(6), (b)(7)(C) No, OMB has not approved that (b)(5)

(b)(5), (b)(7)(E)

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, November 2, 2021 5:17 PM
To: (b)(6), (b)(7)(C)
Subject: RE: (b)(5) for CBP One

Hi (b)(6), (b)(7)(C)

Did this get approved? I did not see a final email on this.

Thank you,

(b)(6), (b)(7)(C)
Acting Director, Strategic Transformation Office
Planning, Program Analysis and Evaluation
Office of Field Operations
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, July 28, 2021 9:04 AM
To: (b)(6)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: FW: (b)(5) for CBP One

Good morning. I reached out to our team here are CBP and have a response for OMB.

(b)(7)(E)

1. (b)(7)(E)
(b)(7)(E) Individuals can upload a photo and it does not need to be a live photo or a "passport" quality photo meeting specific requirements.

2. The photo is the most efficient source of identification to ensure the person presenting themselves at a limit line, with a paper copy of a confirmation email, is the person for whom the CBP On submission was made. (b)(7)(E)
(b)(7)(E)

3. CBP One is a voluntary program. It may not be feasible for all individuals. However the NGOs we have breifed acknowledge and seem ready and willing to support individuals. This is not NGOs submitting on their behalf, but NGOs providing support and access to the tools needed to utilize the application on a mobile device or desktop. If someone can not provide a photo, they can still present themselves to the POE directly.

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Monday, July 19, 2021 5:32 PM
To: (b)(6), (b)(7)(C)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(b)(6)
Subject: RE: (b)(5) for CBP One

(b)(6), (b)(7)(C)

OMB has one initial comment. It is related to the mandate of a photograph. Could CBP elaborate on what is the process if someone does not have access or the ability to provide a photograph? Does CBP feel that this requirement can be

fulfilled by any respondent to the collection? If someone can not provide a photograph, what is the process? Overall, OMB is looking to understand the requirement for the photograph and if this requirement could potential be an issue if someone can not provide a photograph due to limitation of resources.

Thanks.

(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Friday, June 25, 2021 11:10 AM
To: (b)(6)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: (b)(5) for CBP One

Good morning. (b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

Dated: September 22, 2021.

Victoria E. Townsend,

Program Analyst, Office of Federal Advisory Committee Policy.

[FR Doc. 2021–21003 Filed 9–27–21; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HEALTH AND HUMAN SERVICES

National Institutes of Health

National Institute of Biomedical Imaging and Bioengineering; Notice of Closed Meeting

Pursuant to section 10(d) of the Federal Advisory Committee Act, as amended, notice is hereby given of a meeting of the National Institute of Biomedical Imaging and Bioengineering Special Emphasis Panel.

The meetings will be closed to the public in accordance with the provisions set forth in sections 552b(c)(4) and 552b(c)(6), Title 5 U.S.C., as amended. The grant applications and the discussions could disclose confidential trade secrets or commercial property such as patentable material, and personal information concerning individuals associated with the grant applications, the disclosure of which would constitute a clearly unwarranted invasion of personal privacy.

Name of Committee: National Institute of Biomedical Imaging and Bioengineering Special Emphasis Panel; Career Development (Ks) and Conference support (R13).

Date: November 5, 2021.

Time: 10:00 a.m. to 5:00 p.m.

Agenda: To review and evaluate grant applications.

Place: National Institutes of Health, Democracy II, 6707 Democracy Blvd., Bethesda, MD 20892 (Virtual Meeting).

Contact Person: John P. Holden, Ph.D., Scientific Review Officer, National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, 6707 Democracy Blvd., Suite 920, Bethesda, MD 20892, (301) 496–8775, *john.holden@nih.gov.*

(Catalogue of Federal Domestic Assistance Program Nos. 93.866, National Institute of Biomedical Imaging and Bioengineering, National Institutes of Health, HHS)

Dated: September 22, 2021.

Victoria E. Townsend,

Program Analyst, Office of Federal Advisory Committee Policy.

[FR Doc. 2021–21001 Filed 9–27–21; 8:45 am]

BILLING CODE 4140–01–P

DEPARTMENT OF HOMELAND SECURITY

U.S. Customs and Border Protection

[1651–0140]

Collection of Advance Information From Certain Undocumented Individuals on the Land Border

AGENCY: U.S. Customs and Border Protection (CBP), Department of Homeland Security.

ACTION: 60-Day notice and request for comments; revision of an existing collection of information.

SUMMARY: The Department of Homeland Security, U.S. Customs and Border Protection will be submitting the following information collection request to the Office of Management and Budget (OMB) for review and approval in accordance with the Paperwork Reduction Act of 1995 (PRA). The information collection is published in the **Federal Register** to obtain comments from the public and affected agencies.

DATES: Comments are encouraged and must be submitted (no later than November 29, 2021) to be assured of consideration.

ADDRESSES: Written comments and/or suggestions regarding the item(s) contained in this notice must include the OMB Control Number 1651–0140 in the subject line and the agency name. Please use the following method to submit comments:

Email. Submit comments to: *CBP_PRA@cbp.dhs.gov.*

Due to COVID–19-related restrictions, CBP has temporarily suspended its ability to receive public comments by mail.

FOR FURTHER INFORMATION CONTACT: Requests for additional PRA information should be directed to Seth Renkema, Chief, Economic Impact Analysis Branch, U.S. Customs and Border Protection, Office of Trade, Regulations and Rulings, 90 K Street NE, 10th Floor, Washington, DC 20229–1177, Telephone number 202–325–0056 or via email *CBP_PRA@cbp.dhs.gov.* Please note that the contact information provided here is solely for questions regarding this notice. Individuals seeking information about other CBP programs should contact the CBP National Customer Service Center at 877–227–5511, (TTY) 1–800–877–8339, or CBP website at *https://www.cbp.gov/.*

SUPPLEMENTARY INFORMATION: CBP invites the general public and other Federal agencies to comment on the proposed and/or continuing information collections pursuant to the Paperwork Reduction Act of 1995 (44 U.S.C. 3501 *et seq.*). This process is conducted in accordance with 5 CFR 1320.8. Written comments and suggestions from the public and affected agencies should address one or more of the following four points: (1) Whether the proposed collection of information is necessary for the proper performance of the functions of the agency, including whether the information will have practical utility; (2) the accuracy of the agency's estimate of the burden of the proposed collection of information, including the validity of the methodology and assumptions used; (3) suggestions to enhance the quality, utility, and clarity of the information to be collected; and (4) suggestions to minimize the burden of the collection of information on those who are to respond, including through the use of appropriate automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, *e.g.,* permitting electronic submission of responses. The comments that are submitted will be summarized and included in the request for approval. All comments will become a matter of public record.

Overview of This Information Collection

Title: Collection of Advance Information from Certain Undocumented Individuals on the Land Border.

OMB Number: 1651–0140.

Form Number: N/A.

Current Actions: Revision.

Type of Review: Revision.

Affected Public: Individuals.

Abstract: The Department of Homeland Security (DHS), in consultation with U.S. Customs and Border Protection (CBP), has established a process to streamline the processing of undocumented noncitizens under Title 8 of the United States Code at certain ports of entry (POEs), as these individuals require secondary processing upon their arrival, which takes longer than when individuals arrive with sufficient travel documentation.

CBP is proposing extending and amending this data collection, which was established on an emergency basis on May 3, 2021. This data collection expands on the previous collection process for persons who may warrant an exception to the CDC's Order *Suspending the Right To Introduce Certain Persons from Countries Where a Quarantinable Communicable Disease Exists* ("CDC Order") (85 FR 65806), to include undocumented noncitizens who

will be processed under Title 8 at the time they arrive at the POE after the CDC Order is rescinded, in whole or in part. The purpose is to continue to achieve efficiencies to process undocumented noncitizens under Title 8 upon their arrival at the POE, consistent with public health protocols, space limitations, and other restrictions.

CBP collects certain biographic and biometric information from undocumented noncitizens prior to their arrival at a POE, to streamline their processing at the POE. The requested information is that which CBP would otherwise collect from these individuals during primary and/or secondary processing. This information is voluntarily provided by undocumented noncitizens, directly or through non-governmental organizations (NGOs) and international organizations (IOs). Providing this information is not a prerequisite for processing under Title 8, but reduces the amount of data entered by CBP Officers (CBPOs) and the length of time an undocumented noncitizen remains in CBP custody.

The biographic and biometric information being collected in advance, that would otherwise be collected during primary and/or secondary processing at the POEs includes, but is not limited to, descriptive information such as: Name, Data of birth, Country of Birth, City of Birth, Country of Residence, Contact Information, Addresses, Nationality, Employment history (optional), Travel history, Emergency Contact (optional), U.S. and foreign addresses, Familial Information (optional), Marital Status (optional), Identity Document (not a WHTI compliant document) (optional), Gender, Preferred Language, Height, Weight, Eye color and Photograph.

This information is submitted to CBP by undocumented noncitizens on a voluntary basis, for the purpose of facilitating and implementing CBP's mission. This collection is consistent with DHS' and CBP's authorities, including under 6 U.S.C. 202 and 211(c). Pursuant to these sections, DHS and CBP are generally charged with "[s]ecuring the borders, territorial waters, ports, terminals, waterways, and air, land, and sea transportation systems of the United States," and "implement[ing] screening and targeting capabilities, including the screening, reviewing, identifying, and prioritizing of passengers and cargo across all international modes of transportation, both inbound and outbound."

Proposed Changes: This information collection is being changed to require the submission of the photograph—previously optional—for all who choose to provide advance information. The submission of a photograph in advance will provide CBPOs with a mechanism to match a noncitizen who arrives at the POE with the photograph submitted in advance, therefore identifying those individuals, and verifying their identity. The photograph is particularly important for identity verification once NGOs/IOs are no longer facilitating the presentation of all individuals for CBP processing (NGOs/IOs will be able to continue assisting for some individuals but others will be able to participate on their own).

CBP will also allow individuals to request to present themselves for processing at a specific POE on a specific day and time, although such a request does not guarantee that an individual will be processed at a given time. Individuals will have the opportunity to modify their requests within the CBP One™ application to an alternate day or time. In all cases, CBP will inspect, and process individuals based on available capacity at the POE. This new functionality does not require the collection of new Personal Identifiable Information (PII) data elements.

Type of Information Collection: Advance Information on Undocumented Travelers.

Estimated Number of Respondents: 91,250.

Estimated Number of Annual Responses per Respondent: 1.

Estimated Number of Total Annual Responses: 91,250.

Estimated Time per Response: 16 minutes.

Estimated Total Annual Burden Hours: 24,333.

Dated: September 23, 2021.

Seth D. Renkema,

Branch Chief, Economic Impact Analysis Branch, U.S. Customs and Border Protection.

[FR Doc. 2021–20988 Filed 9–27–21; 8:45 am]

BILLING CODE P

DEPARTMENT OF HOMELAND SECURITY

Federal Emergency Management Agency

[Docket ID FEMA–2021–0002]

Final Flood Hazard Determinations

AGENCY: Federal Emergency Management Agency, Department of Homeland Security.

ACTION: Notice.

SUMMARY: Flood hazard determinations, which may include additions or modifications of Base Flood Elevations (BFEs), base flood depths, Special Flood Hazard Area (SFHA) boundaries or zone designations, or regulatory floodways on the Flood Insurance Rate Maps (FIRMs) and where applicable, in the supporting Flood Insurance Study (FIS) reports have been made final for the communities listed in the table below.

The FIRM and FIS report are the basis of the floodplain management measures that a community is required either to adopt or to show evidence of having in effect in order to qualify or remain qualified for participation in the Federal Emergency Management Agency's (FEMA's) National Flood Insurance Program (NFIP). In addition, the FIRM and FIS report are used by insurance agents and others to calculate appropriate flood insurance premium rates for buildings and the contents of those buildings.

DATES: The date of January 28, 2022 has been established for the FIRM and, where applicable, the supporting FIS report showing the new or modified flood hazard information for each community.

ADDRESSES: The FIRM, and if applicable, the FIS report containing the final flood hazard information for each community is available for inspection at the respective Community Map Repository address listed in the tables below and will be available online through the FEMA Map Service Center at *https://msc.fema.gov* by the date indicated above.

FOR FURTHER INFORMATION CONTACT: Rick Sacbibit, Chief, Engineering Services Branch, Federal Insurance and Mitigation Administration, FEMA, 400 C Street SW, Washington, DC 20472, (202) 646–7659, or (email) *patrick.sacbibit@fema.dhs.gov;* or visit the FEMA Mapping and Insurance eXchange (FMIX) online at *https://www.floodmaps.fema.gov/fhm/fmx_main.html.*

SUPPLEMENTARY INFORMATION: The Federal Emergency Management Agency (FEMA) makes the final determinations listed below for the new or modified flood hazard information for each community listed. Notification of these changes has been published in newspapers of local circulation and 90 days have elapsed since that publication. The Deputy Associate Administrator for Insurance and Mitigation has resolved any appeals resulting from this notification.

This final notice is issued in accordance with section 110 of the Flood Disaster Protection Act of 1973, 42 U.S.C. 4104, and 44 CFR part 67.

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Message

From: (b)(6), (b)(7)(C)
Sent: 8/31/2022 5:50:40 PM
To: (b)(6), (b)(7)(C)
Subject: FW: (b)(7)(E) CBP One
Attachments: (b)(5), (b)(7)(E)

(b)(6), (b)(7)(C)
Chief Economist
Office of Trade: Regulations & Rulings
U.S. Customs & Border Protection
(b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Friday, June 25, 2021 11:10 AM
To: (b)(6)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) (OCC) (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: (b)(5) for CBP One

Good morning. (b)(5), (b)(7)(E)

(b)(5), (b)(7)(E)

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

(b)(6), (b)(7)(C)
Branch Chief, Economic Impact Analysis Branch
Regulations & Rulings, Office of Trade
U.S. Customs and Border Protection
(b)(6), (b)(7)(C)
Cell: (b)(6), (b)(7)(C)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	1	ECD Count:	1
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	1	ECD Count:	1
Action Taken Date	Resource Name	Action Taken			
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)					

Supporting Statement
Collection of Advance Information from Certain Undocumented Individuals on the Land Border
1651-NEW

A. Justification

1. **Explain the circumstances that make the collection of information necessary. Identify any legal or administrative requirements that necessitate the collection. Attach a copy of the appropriate section of each statute and regulation mandating or authorizing the collection of information.**

The Department of Homeland Security (DHS), in consultation with U.S. Customs and Border Protection (CBP), is working to establish a process to streamline the processing of undocumented noncitizens under Title 8 at certain ports of entry (POEs), as these individuals require secondary processing upon their arrival, which takes longer than when individuals arrive with sufficient travel documentation.

CBP is proposing this new data collection, which expands upon the existing collection process for persons who may warrant an exception to the CDC Order, to include undocumented noncitizens who will be processed under Title 8 at the time they arrive at the port of entry after the CDC order is rescinded, in whole or in part. The purpose is to continue to achieve efficiencies to process undocumented noncitizens under Title 8 upon their arrival at the POE, consistent with public health protocols, space limitations, and other restrictions.

CBP plans to collect certain biographic and biometric information from undocumented noncitizens prior to their arrival at a POE, to streamline their processing at the POE. The requested information is that which CBP would otherwise collect from these individuals during primary and/or secondary processing. This information will be voluntarily provided by undocumented noncitizens, directly or through non-governmental organizations (NGOs) and international organizations (IOs). Providing this information will not be a prerequisite for processing under Title 8, but will reduce the amount of data entered by CBP Officers (CBPOs) and the length of time an undocumented noncitizen remains in CBP custody.

The biographic and biometric information being collected in advance, that would otherwise be collected during primary and/or secondary processing at the POEs includes, but is not limited to, descriptive information such as: Name, Data of birth, Country of Birth, City of Birth, Country of Residence, Contact Information, Addresses, Nationality, Employment history (optional), Travel history, Emergency Contact (optional), U.S. and foreign addresses, Familial Information (optional), Marital Status (optional), Identity Document (not a WHTI compliant document) (optional), Gender, Preferred Language, Height, Weight, Eye color and Photograph.

This information collection is being changed to require the submission of the photograph – previously it was optional – for all who choose to provide advance information. The submission of a photograph in advance will provide CBPOs with a mechanism to match a noncitizen who arrives at the port with the photograph submitted in advance, therefore identifying those individuals, and verifying their identity. The photograph is particularly important for identity verification once NGOs/IOs are no longer facilitating the presentation of all individuals for CBP processing (NGOs/IOs will be able to continue assisting for some individuals but others will be able to participate on their own).

CBP will also allow individuals to request to present themselves for processing at a specific POE on a specific day and time, although such a request does not guarantee that an individual will be processed at a given time. Individuals will have the opportunity to modify their requests within the CBP One™ application to an alternate day or time. In all cases, CBP will inspect and process individuals based on available capacity at the POE. This new functionality does not require the collection of new PII data elements.

This information will be submitted to CBP by undocumented noncitizens on a voluntary basis, for the purpose of facilitating and implementing CBP's mission. This collection is consistent with DHS' and CBP's authorities, including under 6 U.S.C. §§ 202 and 211(c). Pursuant to these sections, DHS and CBP are generally charged with "securing the borders, territorial waters, ports, terminals, waterways, and air, land, and sea transportation systems of the United States," and "implement[ing] screening and targeting capabilities, including the screening, reviewing, identifying, and prioritizing of passengers and cargo across all international modes of transportation, both inbound and outbound."

2. **Indicate how, by whom, and for what purpose the information is to be used. Except for a new collection, indicate the actual use the agency has made of the information received from the current collection.**

Individuals directly, or through IOs/NGOs, will use the CBP One mobile or desktop application to voluntarily submit biographic information, as well as a photograph, prior to their arrival at a CBP POE for processing. Collecting this information in advance will significantly streamline processing of undocumented noncitizens upon their arrival at the POE. Typically, once an undocumented noncitizen arrives at the POE, CBPOs spend significant time collecting and verifying basic biographic data about the noncitizen during the inspection process. One at a time, the CBPOs interview and collect information from these individuals during secondary inspection. The CBPOs manually enter the information into the Unified Secondary System (USEC). To facilitate processing upon arrival and reduce the amount of manual

data entry into secondary processing systems, CBP One data will be used to populate the fields in secondary processing systems, which can then be verified by the CBPO.

Undocumented noncitizens, or IOs/NGOs on their behalf, will submit the biographic information and a photograph to CBP via the CBP One Application prior to the individual's arrival at the POE. While no information is stored locally in the CBP One Application or on a user's device, this data is stored in a segregated backend database within the Automated Targeting System (ATS). The information will be tagged as coming from CBP One. CBP will store a templatized copy of the photograph in a standalone Traveler Verification Service (TVS) gallery to be matched against a photograph taken by a CBPO once the individual arrives at the POE using Simplified Arrival. The TVS gallery will be populated by the new backend dataset ingesting into ATS specifically for the non-MPP population. When photographs are submitted to ATS from CBP One, the new TVS gallery will stage those photographs until the individual arrives at the POE.

Using Simplified Arrival, once an undocumented noncitizen arrives at the POE for processing, CBP will take a new photograph to search against the new gallery within TVS. If no match is made, CBPOs will manually query ATS based on biographic data to populate Simplified Arrival for processing in primary or query by CBP One confirmation numbers, which are provided to the individual after they submit their advance information through CBP One. As with any undocumented noncitizen who arrives at the POE, the CBPO will use Simplified Arrival to create a referral to secondary for further processing, which will include the confirmation number received from CBP One. Once referred to secondary, CBPOs may import the information captured through the CBP One application into USEC, the secondary processing system. This will reduce the time spent by CBPOs manually entering data in secondary. In secondary, the officers will review the advanced data collected for accuracy, edit the data, and save the information in USEC.

The overall goal of the advance information collection is to achieve efficiencies to process undocumented noncitizens under Title 8, consistent with public health protocols, space limitations, and other restrictions. When data is collected in advance, it helps expedite secondary processing because it will reduce manual data entry into USEC. Such processing will significantly reduce the time these individuals spend in congregate settings, which may contribute to the spread of communicable diseases such as SARS-CoV-2, the virus that causes COVID-19.

3. **Describe whether, and to what extent, the collection of information involves the**

use of automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses, and the basis for the decision for adopting this means of collection. Also describe any consideration of using information technology to reduce burden.

CBP will begin collecting this information through a mobile or computer application. CBP will collect this information electronically, directly from individuals or from IOs/NGOs on behalf of these individuals, via the CBP One application. The CBP One application is currently available as a mobile app on both Google and Apple play stores, as well as a website (https://cbpone.cbp.dhs.gov/#/home) accessible from any browser.

4. **Describe efforts to identify duplication. Show specifically why any similar information already available cannot be used or modified for use for the purposes described in Item 2 above.**

This information is not duplicated for this population in any other place or any other form.

5. **If the collection of information impacts small businesses or other small entities describe any methods used to minimize burden.**

This information collection does not have an impact on small businesses or other small entities.

6. **Describe consequences to Federal program or policy activities if the collection is not conducted or is conducted less frequently.**

Not collecting information in advance would lead to longer processing times for undocumented individuals at POEs, as well as increase the time these individuals will remain in a congregate setting, increasing the risk of transmission of communicable diseases such as COVID-19 among these individuals and CBP employees.

7. **Explain any special circumstances.**

This information is collected in a manner consistent with the guidelines of 5 CFR 1320.5(d)(2).

8. **If applicable, provide a copy and identify the date and page number of publication in the Federal Register of the agency's notice, required by 5 CFR 1320.8(d), soliciting comments on the information collection prior to submission to OMB. Summarize public comments received in response to that notice and describe actions taken by the agency in response to these comments. Specifically address comments received on cost and hour burden.**

(b)(5)

9. Explain any decision to provide any payment or gift to respondents, other than remuneration of contractors or grantees.

There is no offer of a monetary or material value for this information collection.

10. Describe any assurance of confidentiality provided to respondents and the basis for the assurance in statute, regulation, or agency policy.

CBP is publishing a new Privacy Impact Assessment (PIA) for this information collection entitled "Advance Information Collection from Undocumented Individuals." The collection is generally covered by the PIA for the DHS/CBP/PIA-068 CBP One™ Mobile Application (originally published February 2021); b) the PIA for the DHS/CBP/PIA-067 U.S. Customs and Border Protection Unified Secondary (originally published December 2020); and c) and the PIA for the DHS/CBP/PIA-056 Traveler Verification Service (originally published November 2018).

The Systems of Records Notices (SORNs) that will be included in this ICR include the ATS SORN (DHS/CBP-006 Automated Targeting System, May 22, 2012, 77 FR 30297), which pertains to the collection of information in advance of travel. All information collected at the time of inspection and processing is covered by the DHS/CBP-016 Nonimmigrant Information System (March 13, 2015, 80 FR 13398) and DHS/CBP-011 U.S. Customs and Border Protection TECS (December 19, 2008, 73 FR 77778) SORNs.

There are no assurances of confidentiality provided to the respondents of this information collection.

11. Provide additional justification for any questions of a sensitive nature, such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private. This justification should include the reasons why the agency considers the questions necessary, the specific uses to be made of the information, the explanation to be given to persons from whom the information is requested, and any steps to be taken to obtain their consent.

There are no questions of a sensitive nature.

12. Provide estimates of the hour burden of the collection of information.

INFORMATION	TOTAL ANNUAL	NO. OF RESPONDENTS	NO. OF RESPONSES	TOTAL	TIME PER

COLLECTION	BURDEN HOURS		PER RESPONDENT	RESPONSES	RESPONSE
Advance Information on Undocumented Travelers	24,333	91,250	1	91,250	16 minutes

Public Cost

The estimated cost to the respondents is $496,393. This is based on the estimated burden hours (24,333) multiplied by ($20.40). CBP used the U.S. Department of Transportation's guidance on value of travel time for value of time estimates ($20.40)[1] for travel by land.

13. Provide an estimate of the total annual cost burden to respondents or record keepers resulting from the collection of information.

There are no record keeping, capital, start-up or maintenance costs associated with this information collection. Use of the CBP One app is free of charge. CBP assumes that basic internet access is a customary cost of doing business and will not additionally burden any NGO/IO assisting individuals in submitting this form.

14. Provide estimates of annualized cost to the Federal Government. Also provide a description of the method used to estimate cost, which should include quantification of hours, operational expenses (such as equipment overhead, printing, and support staff), and any other expense that would not have been incurred without this collection of information.

The estimated annual cost to the Federal Government associated with the review of these records is $298,935. This is based on the number of responses (91,250) multiplied by the time to review and process each response (3 minutes) = 4,562.50 hours multiplied by the average hourly rate ($65.52) = $298,935. The previous review time of 4 hours was adjusted to reflect the time spent on reviewing the CBP One data in primary and secondary processing systems only, 3 minutes. The 4-hour estimate included the entire case processing time.

15. Explain the reasons for any program changes or adjustments reported in Items 12 or 13.

This information collection is being extended beyond persons who may warrant an exception to the CDC order to undocumented noncitizens who will be processed

1 2016 Revised Value of Travel Time Guidance.pdf (transportation.gov)

under Title 8 at the time of arrival to a POE. Collection of this information will reduce the amount of data manually entered by CBPOs, which is expected to expedite secondary process and thus reduce the length of time an undocumented noncitizen remains in CBP custody. It is also being changed to incorporate a scheduling component. CBP is implementing the ability for individuals, directly or through NGOs/IOs, to request to present at a specific POE on a specific date and time. This will automate the manual process that is currently being utilized for those individuals who may warrant an exception to the CDC Order, which requires the exchange of numerous phone calls and emails. This will reduce the amount of time CBP, individuals, and NGOs/IOs spend on this activity. Providing undocumented noncitizens a prescribed process to request processing at a specific POE and day/time may reduce the number of individuals attempting to enter between the POEs. Finally, the collection is being changed to require those individuals who choose to submit advance information to submit photographs, rather than leaving them as optional. This will provide CBPOs with a mechanism to match a noncitizen who arrives at the port with the photograph submitted in advance, thereby facilitating identify verification and matching to data previously submitted.

16. For collection of information whose results will be published, outline plans for tabulation, and publication.

This information collection will not be published for statistical purposes.

17. If seeking approval to not display the expiration date, explain the reasons that displaying the expiration date would be inappropriate.

CBP will display the expiration date for OMB approval of this information collection.

18. "Certification for Paperwork Reduction Act Submissions."

CBP does not request an exception to the certification of this information collection.

B. Collection of Information Employing Statistical Methods

No statistical methods were employed.

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	2	ECD Count:	3
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	2	ECD Count:	3
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b) (7)(E)	Recommendation #:	2	ECD Count:	3
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b) (7)(E)	Recommendation #:	2	ECD Count:	3
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	4	ECD Count:	5
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	4	ECD Count:	5
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	4	ECD Count:	5
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	4	ECD Count:	5
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	4	ECD Count:	5
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	4	ECD Count:	5
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

U.S. Customs and Border Protection

PLANNING, PROGRAM ANALYSIS, AND EVALUATION

Biometric Exit: Evaluating Bias and Performance Metrics

Biometric Bias | Problem Statement and Overview

Problem statement

- Various biometric facial recognition algorithms have been known to be biased against Race/Ethnicity, Age and/or Gender. CBP run Biometrics facial recognition needs to be evaluated for potential bias.

Results

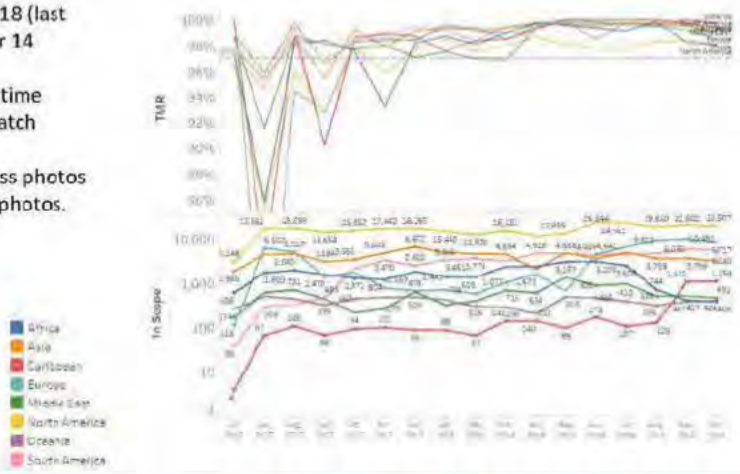
- Our analysis suggests that currently there is no detectable bias (or effects are negligible) in regards to biometric matching based on Race/Ethnicity, Age or Gender.
- 3rd party review would help independently verify results, and provide a more in depth study controlling for various factors.
- CBP is participating in a joint DHS, NIST initiative to evaluate facial recognition performance using operationally relevant face image sets.
- Over 80 million DHS face images including 60 million from CBP are being provided to NIST for independent algorithm and performance bias evaluations.

Biometric Bias | Citizenships by Region Over time

U.S. Customs and Border Protection

Matching improves over time

- Updated through October 2018 (last full month of data. November 14 moved to sampling)
- Matching has improved over time
- Regions are converging on match rates.
- USC matches worse due to less photos on average and having older photos.

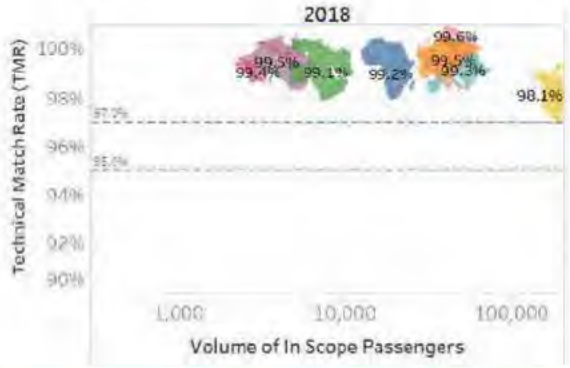


Match rates are improving over time, all regions are now above 97% match.

Biometric Bias | Citizenships by Region

U.S. Customs and Border Protection

- Updated through first half of November 2018
- Match Rates by Citizenship (clustered by region) were reviewed
- No major differences by region
- USCs tend to match worse (Less photos and older photos on average)



Match rates continue to improve. Since 2018, all regions are preforming relatively equally.

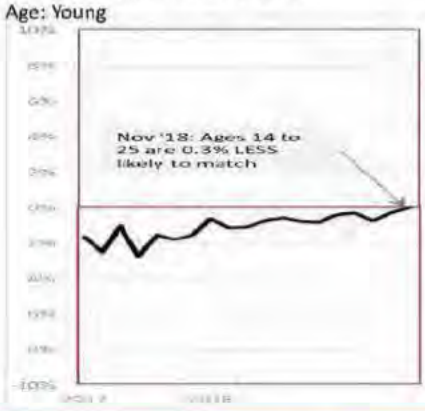
Biometric Bias by month | Age

U.S. Customs and Border Protection

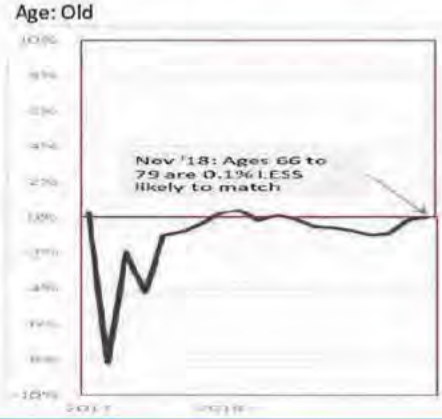
Age bias- initially there was a large difference 2.8% for young individuals and 8% of old individuals relative to middle aged people, travelers between 26 and 65 match slightly better than "young" travelers (0.3%) and "old" travelers (0.1%).

- Updated through November 2018

Age: Young



Age: Old



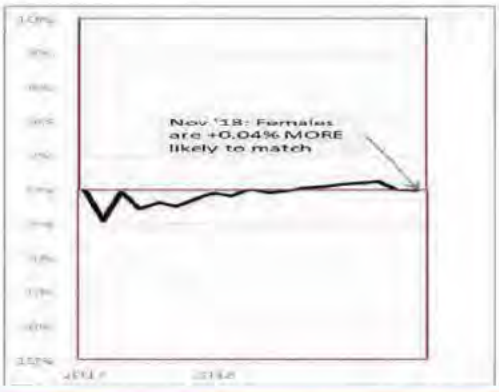
Age bias has decreased over time and now is negligible.

Biometric Bias | Gender

U.S. Customs and Border Protection

Gender Bias Initial gaps (bias) in matching for genders and ages have reduced substantially. In the latest month (Nov '18), women match slightly better than men (0.2%)

- Updated through November 2018



Gender bias has reduced over time and is now negligible.

Technical match rate (TMR) is a ratio of in-scope passengers that matched divided by all in-scope passengers with both gallery and encounter photo. TMR only applies to Air Exit, Entry, Preclearance, and Land. TSA and Sea are ratio of in-scope passengers that matched divided by all passengers/crew, which is considered a biometric confirmation rate (BCR).

TSA BCR is lower due to the inclusion of domestic travelers and TSO daily test photos at each camera (increases the TMR denominator). The duplicates and non-travelers are not being removed/ground-truth because we don't know their identify. Additionally, TSA is not making travelers remove hats, sunglasses etc.

Sea BCR is lower SEA is greatly impacted by the fact that travelers do not need WHTI compliant documents. A driver's license is fine. Getting gallery photos is less successful for a population which may have a drivers license but not a passport.

Biometric Performance | Matching Enhancements

U.S. Customs and Border Protection

- Statistical tests performed to optimize matching threshold
- NEC (matching algorithm) version updated
- OIT improved gallery photo retrieval which increased number of traveler photos for matching and gallery quality
- Camera enhancements such as "selfie-mode" improves the quality of the photo

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	3	ECD Count:	2
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	3	ECD Count:	2
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	3	ECD Count:	2
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	3	ECD Count:	2
Action Taken Date	Resource Name	Action Taken			

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Supporting Statement
Collection of Advance Information from Certain Undocumented Individuals on the Land Border
1651-NEW

A. Justification

1. **Explain the circumstances that make the collection of information necessary. Identify any legal or administrative requirements that necessitate the collection. Attach a copy of the appropriate section of each statute and regulation mandating or authorizing the collection of information.**

The Department of Homeland Security (DHS), in consultation with U.S. Customs and Border Protection (CBP), is working to establish a process to streamline the processing of undocumented noncitizens under Title 8 at certain ports of entry (POEs), as these individuals require secondary processing upon their arrival, which takes longer than when individuals arrive with sufficient travel documentation.

CBP is proposing this new data collection, which expands upon the existing collection process for persons who may warrant an exception to the CDC Order, to include undocumented noncitizens who will be processed under Title 8 at the time they arrive at the port of entry after the CDC order is rescinded, in whole or in part. The purpose is to continue to achieve efficiencies to process undocumented noncitizens under Title 8 upon their arrival at the POE, consistent with public health protocols, space limitations, and other restrictions.

CBP plans to collect certain biographic and biometric information from undocumented noncitizens prior to their arrival at a POE, to streamline their processing at the POE. The requested information is that which CBP would otherwise collect from these individuals during primary and/or secondary processing. This information will be voluntarily provided by undocumented noncitizens, directly or through non-governmental organizations (NGOs) and international organizations (IOs). Providing this information will not be a prerequisite for processing under Title 8, but will reduce the amount of data entered by CBP Officers (CBPOs) and the length of time an undocumented noncitizen remains in CBP custody.

The biographic and biometric information being collected in advance, that would otherwise be collected during primary and/or secondary processing at the POEs includes, but is not limited to, descriptive information such as: Name, Data of birth, Country of Birth, City of Birth, Country of Residence, Contact Information, Addresses, Nationality, Employment history (optional), Travel history, Emergency Contact (optional), U.S. and foreign addresses, Familial Information (optional), Marital Status (optional), Identity Document (not a WHTI compliant document) (optional), Gender, Preferred Language, Height, Weight, Eye color and Photograph.

This information collection is being changed to require the submission of the photograph – previously it was optional – for all who choose to provide advance information. The submission of a photograph in advance will provide CBPOs with a mechanism to match a noncitizen who arrives at the port with the photograph submitted in advance, therefore identifying those individuals, and verifying their identity. The photograph is particularly important for identity verification once NGOs/IOs are no longer facilitating the presentation of all individuals for CBP processing (NGOs/IOs will be able to continue assisting for some individuals but others will be able to participate on their own).

CBP will also allow individuals to request to present themselves for processing at a specific POE on a specific day and time, although such a request does not guarantee that an individual will be processed at a given time. Individuals will have the opportunity to modify their requests within the CBP One™ application to an alternate day or time. In all cases, CBP will inspect and process individuals based on available capacity at the POE. This new functionality does not require the collection of new PII data elements.

This information will be submitted to CBP by undocumented noncitizens on a voluntary basis, for the purpose of facilitating and implementing CBP's mission. This collection is consistent with DHS' and CBP's authorities, including under 6 U.S.C. §§ 202 and 211(c). Pursuant to these sections, DHS and CBP are generally charged with "securing the borders, territorial waters, ports, terminals, waterways, and air, land, and sea transportation systems of the United States," and "implement[ing] screening and targeting capabilities, including the screening, reviewing, identifying, and prioritizing of passengers and cargo across all international modes of transportation, both inbound and outbound."

2. **Indicate how, by whom, and for what purpose the information is to be used. Except for a new collection, indicate the actual use the agency has made of the information received from the current collection.**

Individuals directly, or through IOs/NGOs, will use the CBP One mobile or desktop application to voluntarily submit biographic information, as well as a photograph, prior to their arrival at a CBP POE for processing. Collecting this information in advance will significantly streamline processing of undocumented noncitizens upon their arrival at the POE. Typically, once an undocumented noncitizen arrives at the POE, CBPOs spend significant time collecting and verifying basic biographic data about the noncitizen during the inspection process. One at a time, the CBPOs interview and collect information from these individuals during secondary inspection. The CBPOs manually enter the information into the Unified Secondary System (USEC). To facilitate processing upon arrival and reduce the amount of manual

data entry into secondary processing systems, CBP One data will be used to populate the fields in secondary processing systems, which can then be verified by the CBPO.

Undocumented noncitizens, or IOs/NGOs on their behalf, will submit the biographic information and a photograph to CBP via the CBP One Application prior to the individual's arrival at the POE. While no information is stored locally in the CBP One Application or on a user's device, this data is stored in a segregated backend database within the Automated Targeting System (ATS). The information will be tagged as coming from CBP One. CBP will store a templatized copy of the photograph in a standalone Traveler Verification Service (TVS) gallery to be matched against a photograph taken by a CBPO once the individual arrives at the POE using Simplified Arrival. The TVS gallery will be populated by the new backend dataset ingesting into ATS specifically for the non-MPP population. When photographs are submitted to ATS from CBP One, the new TVS gallery will stage those photographs until the individual arrives at the POE.

Using Simplified Arrival, once an undocumented noncitizen arrives at the POE for processing, CBP will take a new photograph to search against the new gallery within TVS. If no match is made, CBPOs will manually query ATS based on biographic data to populate Simplified Arrival for processing in primary or query by CBP One confirmation numbers, which are provided to the individual after they submit their advance information through CBP One. As with any undocumented noncitizen who arrives at the POE, the CBPO will use Simplified Arrival to create a referral to secondary for further processing, which will include the confirmation number received from CBP One. Once referred to secondary, CBPOs may import the information captured through the CBP One application into USEC, the secondary processing system. This will reduce the time spent by CBPOs manually entering data in secondary. In secondary, the officers will review the advanced data collected for accuracy, edit the data, and save the information in USEC.

The overall goal of the advance information collection is to achieve efficiencies to process undocumented noncitizens under Title 8, consistent with public health protocols, space limitations, and other restrictions. When data is collected in advance, it helps expedite secondary processing because it will reduce manual data entry into USEC. Such processing will significantly reduce the time these individuals spend in congregate settings, which may contribute to the spread of communicable diseases such as SARS-CoV-2, the virus that causes COVID-19.

3. **Describe whether, and to what extent, the collection of information involves the**

use of automated, electronic, mechanical, or other technological collection techniques or other forms of information technology, e.g., permitting electronic submission of responses, and the basis for the decision for adopting this means of collection. Also describe any consideration of using information technology to reduce burden.

CBP will begin collecting this information through a mobile or computer application. CBP will collect this information electronically, directly from individuals or from IOs/NGOs on behalf of these individuals, via the CBP One application. The CBP One application is currently available as a mobile app on both Google and Apple play stores, as well as a website (https://cbpone.cbp.dhs.gov/#/home) accessible from any browser.

4. **Describe efforts to identify duplication. Show specifically why any similar information already available cannot be used or modified for use for the purposes described in Item 2 above.**

This information is not duplicated for this population in any other place or any other form.

5. **If the collection of information impacts small businesses or other small entities describe any methods used to minimize burden.**

This information collection does not have an impact on small businesses or other small entities.

6. **Describe consequences to Federal program or policy activities if the collection is not conducted or is conducted less frequently.**

Not collecting information in advance would lead to longer processing times for undocumented individuals at POEs, as well as increase the time these individuals will remain in a congregate setting, increasing the risk of transmission of communicable diseases such as COVID-19 among these individuals and CBP employees.

7. **Explain any special circumstances.**

This information is collected in a manner consistent with the guidelines of 5 CFR 1320.5(d)(2).

8. **If applicable, provide a copy and identify the date and page number of publication in the Federal Register of the agency's notice, required by 5 CFR 1320.8(d), soliciting comments on the information collection prior to submission to OMB. Summarize public comments received in response to that notice and describe actions taken by the agency in response to these comments. Specifically address comments received on cost and hour burden.**

(b)(5)

9. **Explain any decision to provide any payment or gift to respondents, other than remuneration of contractors or grantees.**

There is no offer of a monetary or material value for this information collection.

10. **Describe any assurance of confidentiality provided to respondents and the basis for the assurance in statute, regulation, or agency policy.**

CBP is publishing a new Privacy Impact Assessment (PIA) for this information collection entitled "Advance Information Collection from Undocumented Individuals." The collection is generally covered by the PIA for the DHS/CBP/PIA-068 CBP One™ Mobile Application (originally published February 2021); b) the PIA for the DHS/CBP/PIA-067 U.S. Customs and Border Protection Unified Secondary (originally published December 2020); and c) and the PIA for the DHS/CBP/PIA-056 Traveler Verification Service (originally published November 2018).

The Systems of Records Notices (SORNs) that will be included in this ICR include the ATS SORN (DHS/CBP-006 Automated Targeting System, May 22, 2012, 77 FR 30297), which pertains to the collection of information in advance of travel. All information collected at the time of inspection and processing is covered by the DHS/CBP-016 Nonimmigrant Information System (March 13, 2015, 80 FR 13398) and DHS/CBP-011 U.S. Customs and Border Protection TECS (December 19, 2008, 73 FR 77778) SORNs.

There are no assurances of confidentiality provided to the respondents of this information collection.

11. **Provide additional justification for any questions of a sensitive nature, such as sexual behavior and attitudes, religious beliefs, and other matters that are commonly considered private. This justification should include the reasons why the agency considers the questions necessary, the specific uses to be made of the information, the explanation to be given to persons from whom the information is requested, and any steps to be taken to obtain their consent.**

There are no questions of a sensitive nature.

12. **Provide estimates of the hour burden of the collection of information.**

INFORMATION	TOTAL ANNUAL	NO. OF RESPONDENTS	NO. OF RESPONSES	TOTAL	TIME PER

COLLECTION	BURDEN HOURS		PER RESPONDENT	RESPONSES	RESPONSE
Advance Information on Undocumented Travelers	24,333	91,250	1	91,250	16 minutes

Public Cost

The estimated cost to the respondents is $496,393. This is based on the estimated burden hours (24,333) multiplied by ($20.40). CBP used the U.S. Department of Transportation's guidance on value of travel time for value of time estimates ($20.40)[1] for travel by land.

13. Provide an estimate of the total annual cost burden to respondents or record keepers resulting from the collection of information.

There are no record keeping, capital, start-up or maintenance costs associated with this information collection. Use of the CBP One app is free of charge. CBP assumes that basic internet access is a customary cost of doing business and will not additionally burden any NGO/IO assisting individuals in submitting this form.

14. Provide estimates of annualized cost to the Federal Government. Also provide a description of the method used to estimate cost, which should include quantification of hours, operational expenses (such as equipment overhead, printing, and support staff), and any other expense that would not have been incurred without this collection of information.

The estimated annual cost to the Federal Government associated with the review of these records is $389,409. This is based on the number of responses (91,250) multiplied by the time to review and process each response (3 minutes) = 4,563 hours multiplied by the average hourly rate ($85.35) = $298,935. The previous review time of 4 hours was adjusted to reflect the time spent on reviewing the CBP One data in primary and secondary processing systems only, 3 minutes. The 4-hour estimate included the entire case processing time.

15. Explain the reasons for any program changes or adjustments reported in Items 12 or 13.

This information collection is being extended beyond persons who may warrant an exception to the CDC order to undocumented noncitizens who will be processed

1 2016 Revised Value of Travel Time Guidance.pdf (transportation.gov)

under Title 8 at the time of arrival to a POE. Collection of this information will reduce the amount of data manually entered by CBPOs, which is expected to expedite secondary process and thus reduce the length of time an undocumented noncitizen remains in CBP custody. It is also being changed to incorporate a scheduling component. CBP is implementing the ability for individuals, directly or through NGOs/IOs, to request to present at a specific POE on a specific date and time. This will automate the manual process that is currently being utilized for those individuals who may warrant an exception to the CDC Order, which requires the exchange of numerous phone calls and emails. This will reduce the amount of time CBP, individuals, and NGOs/IOs spend on this activity. Providing undocumented noncitizens a prescribed process to request processing at a specific POE and day/time may reduce the number of individuals attempting to enter between the POEs. Finally, the collection is being changed to require those individuals who choose to submit advance information to submit photographs, rather than leaving them as optional. This will provide CBPOs with a mechanism to match a noncitizen who arrives at the port with the photograph submitted in advance, thereby facilitating identify verification and matching to data previously submitted.

16. For collection of information whose results will be published, outline plans for tabulation, and publication.

This information collection will not be published for statistical purposes.

17. If seeking approval to not display the expiration date, explain the reasons that displaying the expiration date would be inappropriate.

CBP will display the expiration date for OMB approval of this information collection.

18. "Certification for Paperwork Reduction Act Submissions."

CBP does not request an exception to the certification of this information collection.

B. Collection of Information Employing Statistical Methods

No statistical methods were employed.

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	5	ECD Count:	2

Action Taken Date	Resource Name	Action Taken

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Action Taken Quick View					
Report Number:	(b)(7)(E)	Recommendation #:	5	ECD Count:	2

Action Taken Date	Resource Name	Action Taken

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Message

From: (b)(6), (b)(7)(C)

Sent: 6/1/2022 4:13:08 PM

To: (b)(6)

Subject: FW: Additional DHS Technical Comments to GAO-20-568 "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508)

Attachments: DHS Tech Comments GAO 103508 TSA additional Submission 8-17-2020.docx

From: (b)(6)

Sent: Wednesday, August 19, 2020 12:32 PM

To: (b)(6) (b)(6)

Cc: OBIM Audit Liaison (b)(6) (b)(6) MGMTCAL (b)(6), (b)(7)(C) (b)(6), (b)(7)(C) (b)(6), (b)(7)(C) PARMExecSec (b)(6) (b)(6) PLCYAudits (b)(6) S&T GAO Liaison (b)(6) TSA_ALDauditsmail (b)(6) (b)(6) GAO-OIG Liaison (b)(6) (b)(6)

Subject: Additional DHS Technical Comments to GAO-20-568 "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508)

Good afternoon (b)(6) and all,

I am reaching out to you with the attached additional technical comments to draft report GAO-20-568 "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508).

We regret any inconvenience this may cause, and appreciate your flexibility in accepting these comments late in GAO's final report process.

Recognizing that this may impact GAO's planned date to issue the final report, please let us know if you have an update to your estimated final report date.

As always, please include our organizational mailbox, (b)(6) on the distribution of the final report to ensure proper and timely dissemination within DHS.

Thank you,

(b)(6)

Assistant Director
Departmental GAO-OIG Liaison Office
U.S. Department of Homeland Security
(b)(6) (cell)

From: (b)(6)
Sent: Friday, August 14, 2020 12:39 PM
To: (b)(6)
Subject: RE: DHS Management response to GAO-20-568 "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508)

Just as a quick follow up, we plan to issue the report on September 2, 2020. Thanks again for your help.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)

From: (b)(6)
Sent: Thursday, August 13, 2020 1:33 PM
To: (b)(6)
(b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6) MGMTCAL (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) PARMExecSec (b)(6)
(b)(6)
PLCYAudits (b)(6) S&T GAO Liaison (b)(6) TSA_ALDauditsmail
(b)(6) GAO-OIG Liaison (b)(6)
(b)(6)
Subject: DHS Management response to GAO-20-568 "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508)

Good afternoon,

Thank you for the opportunity to review and comment on draft report GAO-20-568 "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508). DHS's formal management response to the draft report is attached. Technical comments were previously provided under separate cover. Please include our organizational mailbox, (b)(6) on the distribution of the final report to ensure proper and timely dissemination within DHS.

Our records reflect an estimated final report date on/about September 4, 2020. Please let me know if there is a more accurate estimate.

Respectfully,

(b)(6)

Assistant Director
Departmental GAO-OIG Liaison Office
U.S. Department of Homeland Security
(b)(6) (cell)

From: (b)(6)
Sent: Friday, July 31, 2020 12:41 PM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)
(b)(6) MGMTCAL (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
(b)(6), (b)(7)(C) PARMExecSec (b)(6)
(b)(6)
PLCYAudits (b)(6) S&T GAO Liaison (b)(6) TSA_ALDauditsmail
(b)(6) GAO-OIG Liaison (b)(6)
(b)(6)
Subject: DHS Technical Comments to GAO-20-568SU "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508)

Good afternoon,

Thank you for the opportunity to review and comment on draft report GAO-20-568SU "Facial Recognition: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues" (103508). Please see the attached technical comments. DHS's formal management response letter will be sent under separate cover.

Please include our organizational mailbox, (b)(6) on the distribution of the final report.

Once GAO has an opportunity to review these technical comments, please reach back to the Department to confirm whether the final report will contain restrictive markings, as well as other changes that GAO may make to the final report in response to the Technical Comments, as this information will impact the Management Response Letter.

Thank you,

Respectfully,

(b)(6)

Assistant Director
Departmental GAO-OIG Liaison Office
U.S. Department of Homeland Security
(b)(6) (cell)

Technical Comments for ("X" or highlight one):

	GAO Statement of Facts		OIG Discussion Draft/NFR
X	GAO Draft Report		OIG Draft Report

Engagement # / Report #:	(b)(7)(E)
Engagement Title:	FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues
Date:	August 17, 2020

Report Page	Line or Bullet	Comments	Component /POC	Type
Use page number from the report rather than the document .		*Provide specific remarks, including suggested revised report language, as appropriate.*	*Identify the name, title, office, phone #, email address, and Component of the program official/SME submitting each comment.*	*Choose one or more options to characterize each comment:* 1. Accuracy 2. Sensitivity 3. Context and Perspective 4. Editorial

(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Report Page	Line or Bullet	Comments	Component /POC	Type
(b)(5), (b)(6), (b)(7)(C), (b)(7)(E)				

AFFIRM: This work product has been reviewed for sensitivity concerns which DO/DO NOT exist (highlight one).

Component/POC (name, title, email address, and phone number)	(b)(6), (b)(7)(C), (b)(7)(E)

Report Page	Line or Bullet	Comments	Component /POC	Type
		(b)(6), (b)(7)(C), (b)(7)(E)		

Message

From: (b)(6), (b)(7)(C)

Sent: 4/20/2021 4:24:16 PM

To: (b)(6)

CC: (b)(6), (b)(7)(C)

Subject: RE: Follow-up to Request for Closure of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Attachments: GAO Document List_Rec 5.xlsx; GAO Recommendation 5 Documents.zip

Good Afternoon (b)(6) and (b)(6)

CBP's Office of Field Operations provided the attached in response to the requested additional information. I am waiting for them to provide the attachments for recommendation on a zip drive.

1. A list of all of the alerts CBP receives when there is a system problem/failure. Having a comprehensive list of the types of alerts CBP receives would allow us to further explain the system's current reporting capabilities. Examples of the system alerts CBP mentioned during the call were alerts it receives when the system is down due to a power outage, gallery failure, or a system availability issue. If you can provide examples/copies/screen shots of these reports/alerts, that would also help us further explain some of the processes in place to alert CBP during system failures. Note: Please don't limit these alerts to the what the Biometric Entry-Exit Program receives. We are also interested in reviewing what alerts OIT may receive at its Network Operations Center.

2. At least three examples that illustrate actions taken by CBP to correct/address an issue with the performance of the program identified from reviewing the weekly performance reports, such as when the biometric confirmation rate or technical match rate is significantly below expected performance for a given flight, terminal, airline, or airport. Having these examples (i.e. an email to a port director or an airline) would help us further explain how CBP actively uses the weekly performance reports to improve the performance of the program. Please provide any documentation of these examples, if available.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)

Sent: Wednesday, April 14, 2021 12:16 PM

To: (b)(6)

Cc: (b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)

Subject: Follow-up to Request for Closure of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Hi (b)(6) and (b)(6),

Per (b)(6) request, I am sending this information to you as well. (b)(6) has not been receiving my emails with the attachments.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Wednesday, March 24, 2021 10:06 AM
To: (b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
Subject: RE: Follow-up to Request for Closure of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Thank you (b)(6),(b)(7)(C) Let me gather the information and I'll provide you a response by the end of the week.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, March 24, 2021 9:23 AM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)
Subject: Follow-up to Request for Closure of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Morning (b)(6)

I am following up to get the status of CBP's request for closure for recommendation 5. During the meeting on January 21, 2021, CBP/OFO explained in detail why the agency is confident that the weekly review would be more appropriate, as it is the best way to identify a daily pattern of a particular issue. GAO is suggesting daily review and CBP explained that would make it more difficult to identify the pattern. Please let me know when CBP can expect a decision from GAO.

Background

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds

Actions Taken: CBP's OFO has a suite of tools for system and operational performance management, and OFO creates three types of performance reports that are automatically generated and distributed on a weekly basis within CBP and to external stakeholders. These reports include: 1. Saturation Report: Notes the percentage of flights biometrically processed out of the total number of possible international departures segmented by airport. 2. Biometric Air Exit Overview Report: Includes a daily synopsis of operational performance data including numbers of biometrically processed flights and travelers together with biometric match rates. 6 3. Stakeholder Raw Data Reports: Provides Air Exit stakeholders with operational performance data by flight number, passenger counts, and biometric match rates. The OFO's Biometric Entry-Exit Air team monitors these reports for performance issues and addresses any anomalies with stakeholders as they arise. These reports are also used to promote/increase usage by stakeholders. CBP's OFO also conducts random sampling to determine the technical match rates and identify any system or equipment issues. The random sampling is conducted on a weekly basis and includes two flights per airport per week. Finally, CBP's OFO receives alert notifications if TVS experiences an outage, and has a Gallery Assembly System monitor that provides notifications when a flight gallery is not created. Depending on the severity and impact to end users, OFO generates stakeholder notifications, as appropriate.

GAO's response to request for closure: **4 C.F.R. § 81.6(j) (analogous to U.S.C. § 552 (b)(5))**

4 C.F.R. § 81.6(j) (analogous to U.S.C. § 552 (b)(5))

Thank you,
(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6), (b)(7)(C)
Sent: Tuesday, January 12, 2021 7:54 AM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)

Subject: Follow-up to Request for Closure of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Morning (b)(6)

While GAO is now recommending a daily review, CBP/OFO is confident that the weekly review would be more appropriate as it is the best way to identify a daily pattern of a particular issue. A daily review would make it more difficult to identify the pattern. CBP/OFO would like to meet with GAO to provide additional clarification. CBP/OFO is available next Thursday, January 21 any time from 9:00am-12:00pm? Does that date/time work for GAO? If so, I will arrange the meeting and send a meeting invite.

BACKGROUND

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Actions Taken: CBP's OFO has a suite of tools for system and operational performance management, and OFO creates three types of performance reports that are automatically generated and distributed on a weekly basis within CBP and to external stakeholders. These reports include: 1. Saturation Report: Notes the percentage of flights biometrically processed out of the total number of possible international departures segmented by airport. 2. Biometric Air Exit Overview Report: Includes a daily synopsis of operational performance data including numbers of biometrically processed flights and travelers together with biometric match rates. 6 3. Stakeholder Raw Data Reports: Provides Air Exit stakeholders with operational performance data by flight number, passenger counts, and biometric match rates. The OFO's Biometric Entry-Exit Air team monitors these reports for performance issues and addresses any anomalies with stakeholders as they arise. These reports are also used to promote/increase usage by stakeholders. CBP's OFO also conducts random sampling to determine the technical match rates and identify any system or equipment issues. The random sampling is conducted on a weekly basis and includes two flights per airport per week. Finally, CBP's OFO receives alert notifications if TVS experiences an outage, and has a Gallery Assembly System monitor that provides notifications when a flight gallery is not created. Depending on the severity and impact to end users, OFO generates stakeholder notifications, as appropriate.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Monday, December 7, 2020 9:32 AM
To: (b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
Subject: RE: Status of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Hello (b)(6), (b)(7)(C)

4 C.F.R. § 81.6(j) (analogous to U.S.C. § 552 (b)(5))

V/r

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6), (b)(7)(C)
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Thursday, November 19, 2020 3:41 PM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)
Subject: Status of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Afternoon (b)(6)

I am touching base to see if GAO made a decision on this request for closure.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Thursday, September 17, 2020 8:56 AM
To: (b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
(b)(6), (b)(7)(C)
Subject: RE: Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good morning (b)(6), (b)(7)(C)

Thank you for reaching out to us and double checking on that recommendation. Based on the timing of when CBP sent us the documentation associated with this recommendation, we were unable to fully assess the documentation and CBP's actions to determine whether the actions CBP took fully addressed our recommendation before we issued our report. As we mention towards the end of our report, in the "Agency Comments and Our Evaluation" section on pg. 73, once we have an opportunity to fully review CBP's documentation, we will determine the extent to which CBP's actions fully address the recommendation, and then determine if we can close the recommendation. We expect to finish our assessment, and provide CBP with an update within 60 days of the issuance of our final report. Expect to hear back from us on around 11/2/20. Please let us know if you have any other questions.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, September 16, 2020 4:37 PM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)
Subject: Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Afternoon (b)(6)

CBP requested closure of this recommendation in the attached management response letter, but this recommendation remained in the final report. Can you provide more details regarding the additional actions or information needed for GAO to considered the below recommendation resolved and closed?

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

Document Title	Description	CBP Response/Action	Internal Notes
(b)(5), (b)(7)(E)			

Document Title	Description	CBP Response/Action	Internal Notes
(b)(5), (b)(7)(E)			

Message

From: (b)(6), (b)(7)(C)

Sent: 1/12/2021 12:54:11 PM

To: (b)(6)

CC: (b)(6), (b)(7)(C)

Subject: Follow-up to Request for Closure of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Morning (b)(6)

While GAO is now recommending a daily review, CBP/OFO is confident that the weekly review would be more appropriate as it is the best way to identify a daily pattern of a particular issue. A daily review would make it more difficult to identify the pattern. CBP/OFO would like to meet with GAO to provide additional clarification. CBP/OFO is available next Thursday, January 21 any time from 9:00am-12:00pm? Does that date/time work for GAO? If so, I will arrange the meeting and send a meeting invite.

BACKGROUND

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Actions Taken: CBP's OFO has a suite of tools for system and operational performance management, and OFO creates three types of performance reports that are automatically generated and distributed on a weekly basis within CBP and to external stakeholders. These reports include: 1. Saturation Report: Notes the percentage of flights biometrically processed out of the total number of possible international departures segmented by airport. 2. Biometric Air Exit Overview Report: Includes a daily synopsis of operational performance data including numbers of biometrically processed flights and travelers together with biometric match rates. 6 3. Stakeholder Raw Data Reports: Provides Air Exit stakeholders with operational performance data by flight number, passenger counts, and biometric match rates. The OFO's Biometric Entry-Exit Air team monitors these reports for performance issues and addresses any anomalies with stakeholders as they arise. These reports are also used to promote/increase usage by stakeholders. CBP's OFO also conducts random sampling to determine the technical match rates and identify any system or equipment issues. The random sampling is conducted on a weekly basis and includes two flights per airport per week. Finally, CBP's OFO receives alert notifications if TVS experiences an outage, and has a Gallery Assembly System monitor that provides notifications when a flight gallery is not created. Depending on the severity and impact to end users, OFO generates stakeholder notifications, as appropriate.

Thank you,

(b)(6), (b)(7)(C)

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6)

Sent: Monday, December 7, 2020 9:32 AM

To: (b)(6), (b)(7)(C)

Cc: (b)(6), (b)(7)(C)

Subject: RE: Status of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Hello (b)(6), (b)(7)(C)

4 C.F.R. § 81.6(j) (analogous to U.S.C. § 552 (b)(5))

V/r

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Thursday, November 19, 2020 3:41 PM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)
Subject: Status of Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Afternoon (b)(6)

I am touching base to see if GAO made a decision on this request for closure.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison

Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

From: (b)(6)
Sent: Thursday, September 17, 2020 8:56 AM
To: (b)(6), (b)(7)(C)
Cc: (b)(6), (b)(7)(C)
(b)(6)
Subject: RE: Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good morning (b)(6), (b)(7)(C)

Thank you for reaching out to us and double checking on that recommendation. Based on the timing of when CBP sent us the documentation associated with this recommendation, we were unable to fully assess the documentation and CBP's actions to determine whether the actions CBP took fully addressed our recommendation before we issued our report. As we mention towards the end of our report, in the "Agency Comments and Our Evaluation" section on pg. 73, once we have an opportunity to fully review CBP's documentation, we will determine the extent to which CBP's actions fully address the recommendation, and then determine if we can close the recommendation. We expect to finish our assessment, and provide CBP with an update within 60 days of the issuance of our final report. Expect to hear back from us on around 11/2/20. Please let us know if you have any other questions.

V/R

(b)(6)
Senior Analyst, Homeland Security & Justice
U.S. Government Accountability Office
(b)(6)
(b)(6)

From: (b)(6), (b)(7)(C)
Sent: Wednesday, September 16, 2020 4:37 PM
To: (b)(6)
Cc: (b)(6), (b)(7)(C)
Subject: Recommendation 5 - GAO-20-568, FACIAL RECOGNITION: CBP and TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy and System Performance Issues

Good Afternoon (b)(6)

CBP requested closure of this recommendation in the attached management response letter, but this recommendation remained in the final report. Can you provide more details regarding the additional actions or information needed for GAO to considered the below recommendation resolved and closed?

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Thank you,

(b)(6), (b)(7)(C)
Component Audit Liaison
Management Inspection Division
Office of Accountability
U.S. Customs and Border Protection
Email: (b)(6), (b)(7)(C)
Phone: (b)(6), (b)(7)(C)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(6), (b)(7)(C), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

Withheld in full pursuant to (b)(5), (b)(7)(E)

U.S. Department of Homeland Security
Washington, DC 20528

Homeland Security

August 13, 2020

Rebecca Gambler
Director, Homeland Security and Justice
U.S. Government Accountability Office
441 G Street, NW
Washington, DC 20548

Re: Management Response to Draft Report GAO-20-568, "FACIAL RECOGNITION: CBP & TSA are Taking Steps to Implement Programs, but CBP Should Address Privacy & System Performance Issues"

Dear Ms. Gambler,

Thank you for the opportunity to comment on this draft report. The U.S. Department of Homeland Security (DHS or the Department) appreciates the U.S. Government Accountability Office's (GAO) work in planning and conducting its review and issuing this report.

The Department is pleased to note GAO's recognition of progress made by U.S. Customs and Border Protection (CBP) with testing and deploying facial recognition technology (FRT) at ports of entry to create entry-exit records for foreign nationals, as part of its Biometric Entry-Exit Program, and on the Transportation Security Administration's (TSA) pilot tests to assess the feasibility of using FRT. DHS remains committed to facilitating legitimate travel and securing U.S. borders through expanded use of facial recognition matching that sustains privacy protections, while also maintaining high standards of transparency and accountability.

For example, in 2017, CBP developed and implemented the Traveler Verification Service (TVS) as its facial recognition matching service as part of the congressional mandate to implement a biometric entry-exit system. TVS is an efficient, accurate, and secure manner to verify identity. As acknowledged in the draft report, CBP, in partnership with airlines, deployed FRT to 27 airports to biometrically confirm travelers' identities when they depart the United States (air exit) and was in the early stages of assessing FRT at sea and land ports of entry during field work for this report. CBP is currently working to

implement TVS for biometric air exit for 97 percent of in-scope[1] departing commercial air travelers from the United States, by the end of Fiscal Year 2021.

To monitor the progress towards meeting the 97 percent goal, CBP's Office of Field Operations (OFO) generates weekly-automated reports that track usage and biometric confirmation rates, as well as weekly performance reports for each stakeholder to encourage increased TVS usage. As this effort involves usage by airlines, airport authorities, sea cruise lines, and seaport authorities, and other stakeholders, CBP's increase of availability of TVS at all air, land, and sea ports of entry will encourage an increase in the operational use of TVS.

In addition to ensuring the accuracy of TVS, CBP remains committed to ensuring that the use of technology sustains, and does not erode, privacy protections. CBP takes privacy very seriously, and is dedicated to protecting the privacy of all travelers. For example, CBP provides notice to individuals regarding the collection, use, dissemination, and maintenance of personally identifiable information as part of efforts to promote transparency.

To ensure that CBP and its stakeholders take appropriate measures to mitigate privacy and security risks associated with biometric data collection, CBP developed a comprehensive audit plan, which was provided to the GAO in April 2020. The audit plan includes security interviews with partner information technology departments, security scans of biometric processing systems, and penetration tests of those systems. CBP uses the totality of this information to determine whether a system is secure and subsequently, if the information exchanged in the system is equally safeguarded.

Further, TSA's participation in this audit is consistent with its commitment to transparency and accountability regarding the use of biometric technology for identity verification at the TSA checkpoint. As acknowledged in the draft report, it is too early to fully assess TSA's compliance with privacy protection principles. We reiterate, however, our commitment that DHS' Fair Information Practice Principle will continue to guide TSA as it seeks to protect passenger privacy while achieving the operational and security benefits of biometric technology and improving the passenger experience.

It is important to note, however, that DHS believes that many of the performance rates for face recognition algorithms outlined in the National Institute of Standards and Technology (NIST) December 2019 report are not acceptable for use in CBP operations. For example, CBP uses an algorithm evaluated by NIST and confirmed to be high-performing, ranking first or second in most categories evaluated, including match performance in galleries that are much bigger than those used by CBP, calling the

[1] An "in-scope" traveler is any person who is required by law to provide biometrics upon exit from the United States pursuant to 8 CFR 235.1(f)(ii).

demographic differential for "undetectable."[2] The performance metrics described by NIST are consistent with CBP operational performance metrics for entry-exit, and CBP's operational data continues to show there is no measurable differential performance in matching based on demographic factors. Moreover, the NIST FRVT report shows a wide range in accuracy across algorithm developers, with the most accurate algorithms producing many "fewer errors" and "undetectable false positive" differentials.

Since many of the performance rates specified in the NIST report are not acceptable for use in CBP operations, CBP does not use them. CBP believes the only relevant parts of the report, for the purposes of GAO's draft report, are the specific sections on algorithm performance for algorithms that CBP actually uses.

The draft report contained five recommendations, with which the Department concurs. Attached find our detailed response to each recommendation. DHS previously submitted technical comments under a separate cover for GAO's consideration.

Again, thank you for the opportunity to review and comment on this draft report. Please feel free to contact me if you have any questions. We look forward to working with you again in the future.

Sincerely,

JIM H CRUMPACKER

Digitally signed by JIM H CRUMPACKER
Date: 2020.08.13 08:56:03 -04'00'

JIM H. CRUMPACKER, CIA, CFE
Director
Departmental GAO-OIG Liaison Office

Attachment

[2] Face Recognition Vendor Test (FRVT), Part 3: Demographic Effects, National Institute of Standards and Technology, U.S. Department of Commerce (December 2019), p.8.

Attachment: Management Response to Recommendations Contained in GAO-20-568

GAO recommended that the Commissioner of CBP:

Recommendation 1: Ensure that the Biometric Entry-Exit Program's privacy notices contain complete and current information, including all of the locations where facial recognition is used and how travelers can request to opt out as appropriate.

Response: Concur. CBP's OFO will collaborate with the CBP Office of Public Affairs to publish: 1) Biometric Entry-Exit privacy information; 2) locations where facial recognition is used; and 3) traveler opt-out procedures on CBP's public-facing website, as well as to review and update that information on a monthly basis. CBP's OFO will also ensure that information provided in response to inquiries via the CBP Call Center is also reviewed and updated monthly. Estimated Completion Date (ECD): December 31, 2020.

Recommendation 2: Ensure that the Biometric Entry-Exit Program's privacy signage is consistently available at all locations where CBP is using facial recognition.

Response: Concur. It is important to note that, unlike Federal Inspection Services areas, the airport departure areas are not managed by CBP personnel. However, CBP OFO will continue to work with its airlines/airport partners to ensure that privacy signage is available, on display, and reflective of current privacy messaging for travelers. For example, CBP provides notice to individuals regarding the collection, use, dissemination, and maintenance of personally identifiable information as part of efforts to promote transparency. While CBP acknowledges that operational constraints may affect the placement of signs or the timely posting of updated signage, the overall public is informed that stakeholders are taking photos in coordination with CBP. Further, CBP's OFO regularly conducts periodic signage audits that include local CBP personnel to ensure signs are accurate and placed appropriately.

In addition, CBP notifies travelers at these ports using verbal announcements, signs, and message boards, as appropriate, that CBP takes these photos for identity verification purposes. Travelers are also informed of their ability to request alternative identity verification procedures. Also publicly stated are notifications that, should a traveler decide to request alternative identity verification procedures, the airline would conduct manual identity verification using his/her travel document, and may notify CBP to collect biometrics, such as fingerprints, if applicable. CBP's OFO will also continue to work with airline and airport partners to identify other methods to communicate the use of facial recognition and travelers' privacy rights. ECD: June 30, 2021.

Recommendation 3: Direct the Biometric Entry-Exit Program to develop and implement a plan to conduct privacy audits of its commercial partners', contractors', and vendors' use of personally identifiable information.

Response: Concur. In the air exit environment, CBP OFO will continue to conduct security reviews on partner biometric capture equipment and all interfaces with CBP's TVS, as detailed in the Biometric Entry-Exit Program audit plan, provided to GAO in April 2020. This audit plan enables a comprehensive review of compliance with security and privacy requirements on the part of CBP and CBP's partners. As mentioned in the draft report, CBP completed one partner audit thus far. Although, CBP planned additional audits for 2020, due to the COVID-19 global health pandemic and subsequent travel restrictions, CBP paused the planned audit activities. Once pandemic travel restrictions are lifted, CBP's OFO and Office of Information Technology (OIT) will resume conducting audits. Further, CBP's Privacy and Diversity Office is finalizing its CBP Privacy Evaluation of TVS, which evaluates TVS program protections identified in previously issued compliance documentation, such as Privacy Impact Assessments.

CBP's OFO and OIT plan to conduct four to six reviews per year that will begin after COVID-19 travel restrictions are lifted. ECD: June 30, 2021.

Recommendation 4: Develop and implement a plan to ensure that the biometric air exit capability meets its established photo capture requirement.

Response: Concur. The CBP's Biometric Entry-Exit Program's Air Exit Segment was granted Acquisition Decision Event 3 in December 2019. One of the action items from this decision was to complete an update to the Operational Requirements Document (ORD). CBP's OFO will update the ORD by removing the photo capture requirement, as this requirement is not applicable to current air exit operations. ECD: June 30, 2021.

Recommendation 5: Develop a process by which Biometric Entry-Exit program officials are alerted when the performance of air exit facial recognition falls below established thresholds.

Response: Concur. CBP's OFO has a suite of tools for system and operational performance management, and OFO creates three types of performance reports that are automatically generated and distributed on a weekly basis within CBP and to external stakeholders. These reports include:

1. Saturation Report: Notes the percentage of flights biometrically processed out of the total number of possible international departures segmented by airport.
2. Biometric Air Exit Overview Report: Includes a daily synopsis of operational performance data including numbers of biometrically processed flights and travelers together with biometric match rates.

3. Stakeholder Raw Data Reports: Provides Air Exit stakeholders with operational performance data by flight number, passenger counts, and biometric match rates.

The OFO's Biometric Entry-Exit Air team monitors these reports for performance issues and addresses any anomalies with stakeholders as they arise. These reports are also used to promote/increase usage by stakeholders.

CBP's OFO also conducts random sampling to determine the technical match rates and identify any system or equipment issues. The random sampling is conducted on a weekly basis and includes two flights per airport per week.

Finally, CBP's OFO receives alert notifications if TVS experiences an outage, and has a Gallery Assembly System monitor that provides notifications when a flight gallery is not created. Depending on the severity and impact to end users, OFO generates stakeholder notifications, as appropriate.

We request that GAO consider this recommendation resolved and closed, as implemented.