

NISTIR 8331

**Ongoing Face Recognition
Vendor Test (FRVT)
Part 6B: Face recognition accuracy with
face masks using post-COVID-19
algorithms**

Mei Ngan
Patrick Grother
Kayee Hanaoka

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.IR.8331>

NISTIR 8331

**Ongoing Face Recognition
Vendor Test (FRVT)
Part 6B: Face recognition accuracy with
face masks using post-COVID-19
algorithms**

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

This publication is available free of charge from:
<https://doi.org/10.6028/NIST.IR.8331>

November 2020



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

Certain commercial entities, equipment, or materials may be identified in this document in order to describe an experimental procedure or concept adequately. Such identification is not intended to imply recommendation or endorsement by the National Institute of Standards and Technology, nor is it intended to imply that the entities, materials, or equipment are necessarily the best available for the purpose.

**National Institute of Standards and Technology Interagency or Internal Report 8331
Natl. Inst. Stand. Technol. Interag. Intern. Rep. 8331, 83 pages (November 2020)**

**This publication is available free of charge from:
<https://doi.org/10.6028/NIST.IR.8331>**

Executive Summary

OVERVIEW

This is the second of a series of reports on the performance of face recognition algorithms on faces occluded by protective face masks [2] commonly worn to reduce inhalation and exhalation of viruses. Inspired by the COVID-19 pandemic response, this is a continuous study being run under the Ongoing Face Recognition Vendor Test (FRVT) executed by the National Institute of Standards and Technology (NIST). In our first report [8], we tested “pre-pandemic” algorithms that were already submitted to FRVT 1:1 prior to mid-March 2020. This report augments its predecessor with results for more recent algorithms provided to NIST after mid-March 2020. While we do not have information on whether or not a particular algorithm was designed with face coverings in mind, the results show evidence that a number of developers have adapted their algorithms to support face recognition on subjects potentially wearing face masks. The algorithms tested were one-to-one algorithms submitted to the FRVT 1:1 Verification track. Future editions of this document will also report accuracy of one-to-many algorithms.

WHAT’S NEW

This report includes

- ▷ Results from evaluating 65 face recognition algorithms provided to NIST since mid-March 2020
- ▷ Assessment of when both the enrollment and verification images are masked (in addition to when only the verification image is masked)
- ▷ Results for red and white colored masks (in addition to light-blue and black)
- ▷ Cumulative results for 152 algorithms evaluated to date (submitted both prior to and after mid-March 2020)

MOTIVATION

Traditionally, face recognition systems (in cooperative settings) are presented with mostly non-occluded faces, which include primary facial features such as the eyes, nose, and mouth. However, there are a number of circumstances in which faces are occluded by masks such as in pandemics, medical settings, excessive pollution, or laboratories. Inspired by the COVID-19 pandemic response, the widespread requirement that people wear protective face masks in public places has driven a need to understand how cooperative face recognition technology deals with occluded faces, often with just the periocular area and above visible. An increasing number of research publications have surfaced on the topic of face recognition on people wearing masks along with face-masked research datasets [10]. A number of commercial providers have announced the availability of face recognition algorithms capable of handling face masks, and this report documents performance results for 65 algorithms submitted to NIST after mid-March 2020. This report includes results for all algorithms evaluated to date. At the time of this writing, we are not aware of any large-scale, independent, and publicly reported evaluation on the effects of face mask occlusion on face recognition.

WHAT WE DID

The NIST Information Technology Laboratory (ITL) quantified the accuracy of face recognition algorithms on faces occluded by masks applied digitally to a large set of photos that has been used in an FRVT verification benchmark since 2018. These algorithms were submitted to FRVT 1:1 and includes 65 new algorithms provided to NIST since mid-March 2020. While we do not have information on whether or not a particular algorithm was designed with face coverings in mind, the algorithms were submitted with the expectation that NIST would execute them on masked face images. Using the original unmasked images to form a baseline for accuracy, we measured the impact of occlusion by digitally applying a mask to the face and varying mask shape, mask color, and nose coverage.

We ran these algorithms over a large set of photographs collected in U.S. governmental applications that are currently in operation: **application photographs** from a global population of applicants for immigration benefits and **border crossing photographs** of travelers entering the United States. Both datasets were collected for authorized travel or immigration processes.

**WHAT WE DID
(CONTINUED)**

The application photos (used as reference images) have good compliance with image capture standards. The border crossing photos (used as probe images) are not in good compliance with image capture standards given possible constraints on capture equipment, duration, facilities, and environment. We evaluated the case where the application photos were left unmasked, and synthetic masks were applied to the border crossing photos. This mimics an operational scenario where a person wearing a mask attempts to authenticate against a prior visa or passport photo. We also evaluated when both the application photos and border crossing photos were masked. This mimics, for example, a seamless travel scenario through an airport where a masked face image captured at check-in is enrolled and used during subsequent authentication attempts of the passenger still wearing a face mask. Together these datasets allowed us to process a total of 6.2 million images through a cumulative total of 152 algorithms.

Our use of software to apply masks to face images has the following advantages: it allows very rapid characterization of the effect of masks on face recognition; it allows controlled exploration of factors such as mask size, shape, and color; it affords repeatability, which is key to the fair comparison of algorithms; it scales to very large datasets - in our study, some 6.2 million photographs - which allows fine-grained characterization of false positive rates in addition to false negative rates. Conversely, our use of digital masks presents a number of limitations - please see the *Limitations* section of this executive summary for a more detailed discussion on the limitations of this study.

**WHAT WE
FOUND**

The following results represent observations on algorithms provided to NIST both before and after the COVID-19 pandemic to date. We do not have information on whether or not a particular algorithm was designed with face coverings in mind. The results documented capture a snapshot of algorithms submitted to the FRVT 1:1 in face recognition on subjects potentially wearing face masks.

- ▷ **False rejection performance:** All algorithms submitted after the pandemic continue to give increased false non-match rates (FNMR) when the probes are masked. While a few pre-pandemic algorithms still remain within the most accurate on masked photos, some developers have submitted algorithms after the pandemic showing significantly improved accuracy and are now among the most accurate in our test. Using border crossing images, without masks, the most accurate algorithms will fail to authenticate about 0.3% of persons while falsely accepting no more than 1 in 100000 impostors (i.e. FNMR= 0.003 at FMR= 0.00001). With the highest coverage mask we tested and the most accurate algorithms, this failure rate rises to about 5% (FNMR = 0.05) among all algorithms. This is noteworthy given that around 70% of the face area is occluded by the mask. However, many algorithms submitted since mid-March 2020 remain much less tolerant: some algorithms that are quite competitive with unmasked faces (FNMR < 0.01) still fail to authenticate between 10% to 40% of masked images (FNMR → 0.4).

See Figures 6, 7 and Table 5

For the case where both the enrollment and verification images are masked, interestingly, many algorithms show a reduction in false non-match rates compared to when only the verification image is masked, at a fixed threshold. While the reduction in FNMR is favorable, we observe much larger false match rates when both images are masked. These findings are discussed in subsequent sections of this executive summary. *See Figure 37*

In cooperative access control applications, false rejections can traditionally be remedied by users making second attempts. This is effective when users correct pose, expression, or illumination aspects of their presentation. With masked faces, however, a second attempt may not be effective if the failure is a systematic property of the algorithm.

WHAT WE
FOUND
(CONTINUED)

- ▷ **Evolution of algorithms on face masks:** We observe that a number of algorithms submitted since mid-March 2020 show notable reductions in error rates with face masks over their pre-pandemic predecessors. When comparing error rates for unmasked versus masked faces, the median FNMR across algorithms submitted since mid-March 2020 has been reduced by around 25% from the median pre-pandemic results. The figure below presents examples of developer evolution on both masked and unmasked datasets. For some developers, false rejection rates in their algorithms submitted since mid-March 2020 decreased by as much as a factor of 10 over their pre-pandemic algorithms, which is evidence that some providers are adapting their algorithms to handle face masks. However, in the best cases, when comparing results for unmasked images to masked images, false rejection rates have increased from 0.3%-0.5% (unmasked) to 2.4%-5% (masked). The current performance of face recognition with face masks is comparable to the state-of-the-art on unmasked images in mid-2017 [4].

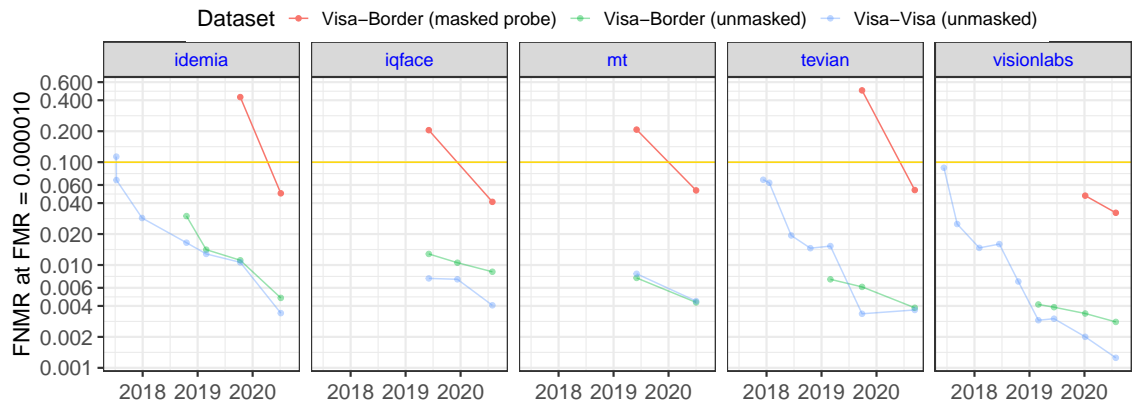


Figure 1: Examples of developer evolution of accuracy on masked and unmasked datasets.

See Figures 6, 5, and 7

- ▷ **False acceptance performance:** As most systems are configured with a fixed threshold, it is necessary to report both false negative and false positive rates for each group at that threshold. When comparing a masked probe to an unmasked enrollment photo, in most cases, false match rates (FMR) are reduced by masks. The effect is generally modest with reductions in FMR usually being smaller than a factor of two. This property is valuable in that masked probes do not impart adverse false match security consequences for verification.

However, when both the enrollment and verification images are masked, most algorithms give elevated false match rates, with FMR ranging from 10 to 100 times higher than when only the probe is masked or both images are unmasked, at the same threshold. This behavior applies to most algorithms tested, with the exception of particular algorithms submitted by Idemia (idemia-006) and Pensees (pensees-001). These two algorithms generate approximately constant FMR across any masked/unmasked combination, and similarly, yield approximately constant FNMR across masked-probe and masked-enrollment-and-probe combinations. *See Figure 37*

- ▷ **Mask-agnostic face recognition:** All 1:1 verification algorithms submitted to the FRVT test since the start of the pandemic are evaluated on both masked and unmasked datasets. The test is designed this way to mimic operational reality: some images will have masks, some will not (especially enrollment samples from a database or ID card). And to the extent that the use of protective masks will exist for some time, our test will continue to evaluate algorithmic capability on verifying all combinations of masked and unmasked faces.

WHAT WE
FOUND
(CONTINUED)

At least two developers - Idemia and Pensees - have developed algorithms that work with any combination of masked and unmasked images, favorably generating approximately constant FMR across any masked/unmasked combination.

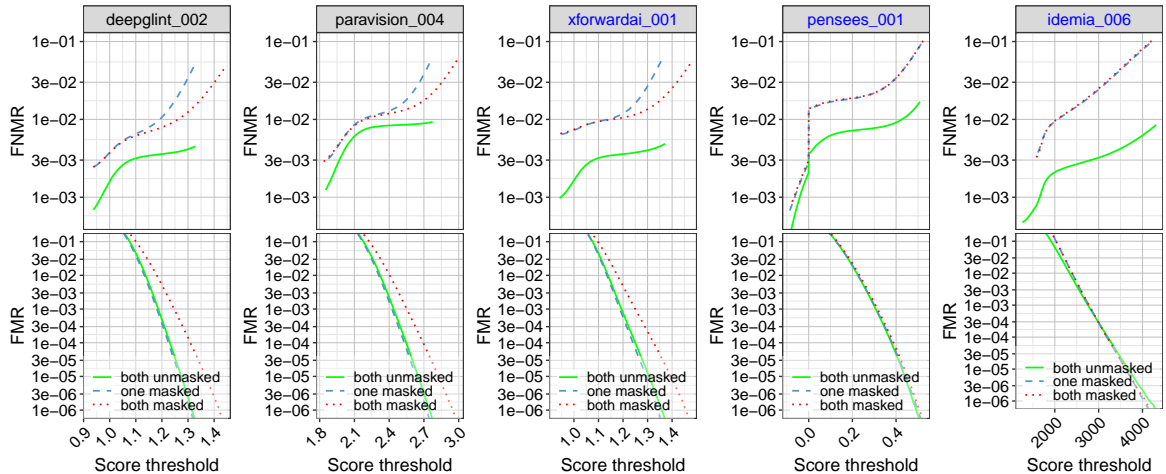


Figure 2: FNMR and FMR calibration curves on masked/unmasked combinations.

An example of an integrated approach might be: 1) inclusion of a mask detector 2) on an unmasked image, extract features from both the full face and the periocular region 3) on a detected masked image, extract features from the periocular region 4) at match time, compare full-face templates when both images are unmasked, and periocular templates otherwise.

- ▷ **Coverage of the masks:** Masks that occlude more of the face give larger false non-match rates. We surveyed over the extent to which the mask covers the nose, from not at all (“low”) to typical (“medium”) to near the eyes (“high”). We baselined those with unmasked faces with the result that FNMR increases by factors of around 6, 20, and 30 respectively for the median algorithm submitted since mid-March 2020. However, as noted, algorithms vary considerably in their tolerance of coverage. Readers should consult tabulated values for specific algorithms.

See Table 5 and Figures 9, 11

We included the “low” option not because it is a common position for a mask but as an option for authentication applications where it would be tenable to ask the user to pull the mask down to just below the nose for the duration of the authentication attempt.

- ▷ **Color of the masks:** We considered white, light-blue, red, and black masks. Many algorithms have higher error rates in black and red masks than light-blue and white masks. The reason for observed accuracy differences between mask color is unknown but is a point for consideration by impacted developers. Mask color also affects the rate at which some algorithms fail to produce a template from an image.

See Figure 24 and Table 9

- ▷ **Shape of the masks:** The shape of the masks matters. Full-face-width masks generally cover more of the face than rounder N95 type masks. Post-pandemic algorithm results show that wide-width masks generally give false negative rates about a factor of 1.5 higher than do rounder type masks.

See Figure 8

- ▷ **Failure to detect and template:** The false negative rates in this report include the effects of both face detection and localization errors, and low-similarity matching errors. We separately include tables detailing how often an algorithm does not make a template from an input image. While many algorithms give low failure-to-template rates, some give high values ranging close to 100%. Conversely, the successful creation of a template does not guarantee proper facial localization. Such localization failures will not be captured as a failure to detect and template event but will impact accuracy rates nonetheless.

See Table 9 and Figure 10

LIMITATIONS

As a simulation, this study likely doesn't fully capture the effects of masks on face recognition. Particularly the following points should be weighed by readers in the near term. Some of these will be addressed in subsequent work at NIST.

- ▷ **Train algorithms:** As with all NIST evaluations, we regard the software as a black box whose parameters (models) remain fixed for the entirety of its use without learning from the test data. We do not train or fine-tune algorithms.
- ▷ **Evaluate one-to-many algorithms:** We have only run one-to-one verification algorithms with masks. This elicits data on the effect of masks on the underlying feature extraction and discrimination of algorithms and can therefore be expected to give first-order indications of the effect on one-to-many identification algorithms.
- ▷ **Consider the effect of eye occlusion:** We did not address the effect of eye-glasses or eye-protection. While our dataset includes examples of people wearing glasses, we didn't collect such data nor simulate it with digital addition.
- ▷ **Test with images of real masks:** Given time and resource constraints, we didn't collect photos of subjects wearing masks. The possible downsides of this are several. First, our digital masks are tailored to faces; while a few don't fit realistically, mass-produced real masks may not fit all actual persons correctly either. We were not able to pursue an exhaustive simulation of the endless variations in color, design, shape, texture, bands, and ways masks can be worn. Second, because many cameras run with exposure-control, it is possible that a dark mask will cause less light to be reflecting and the camera to increase gain on the sensor causing overexposure of the periocular region. Likewise a white mask could lead to underexposure problems. Third, it is possible that some cameras that include a face detector, may fail to focus or acquire a masked face correctly.
- ▷ **Use textured masks:** All masks synthesized by NIST in this study have a uniform color. The consequences of this are that we do not capture the impact of mask texture or pattern on face recognition. The possibility exists for patterned masks to induce higher facial localization errors, which is not captured in our current study. We received a suggestion that such information may serve as a soft biometric, in that a subject that always wears the same textured mask will be more identifiable. We don't intend to encourage algorithm development along this line, because as mass-produced high-efficacy masks become more common, mask diversity may actually drop.
- ▷ **Study demographic effects on masked images:** This report estimates overall performance of existing algorithms on recognition of faces occluded by masks. We deferred tabulating accuracy for different demographic groups until more capable mask-enabled algorithms have been submitted to FRVT.
- ▷ **Evaluate algorithms on non-cooperative, unconstrained imagery:** This report documents results for matching masked webcam images to unmasked portrait-style photos. While the properties of the two sets of images differ, subjects are operating in cooperative mode and are for the most part, looking at the camera.
- ▷ **Consider effects of human examination:** This report does not consider the various ways humans are involved in face recognition systems. For example, analysts can correct face detection or localization errors induced by masks, prior to automated recognition. Likewise, humans are often tasked with adjudication of images following a rejection or other exception from an automated system. Analysis of human capability and role is pertinent to those operations, but is beyond the scope of this study.

**IMPLICATIONS
AND FUTURE
WORK**

Know Your Algorithm: Operational implementations usually employ a single face recognition algorithm. Given algorithm-specific sensitivities to masks and other image or subject properties, 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 collected with actual masks.

NIST plans on releasing a series of reports, iteratively assessing different aspects and use cases of face masking on recognition performance.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

ACKNOWLEDGMENTS

This work was conducted in collaboration with the Department of Homeland Security's Science & Technology Directorate (S&T), Office of Biometric Identity Management (OBIM), and Customs and Border Protection (CBP). Additionally, 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.

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).

Contents

EXECUTIVE SUMMARY	I
ACKNOWLEDGMENTS	VII
DISCLAIMER	VII
INSTITUTIONAL REVIEW BOARD	VII
1 FACE MASK EFFECTS	1
1.1 STATUS	1
1.2 INTRODUCTION	1
2 IMAGE DATASETS	2
2.1 APPLICATION IMAGES	2
2.2 WEBCAM IMAGES	2
2.3 SYNTHETICALLY MASKED IMAGES	3
3 METRICS	5
3.1 MATCHING ACCURACY	5
3.2 FAILURE TO ENROLL	5
4 ALGORITHMS	5
5 RESULTS	9
EVOLUTION OF ALGORITHM PERFORMANCE WITH FACE MASKS	9
FNMR - SUMMARY	15
MASK VS. NO MASK	22
MASK SHAPE	22
MASK NOSE COVERAGE	22
ROLE OF FTE	24
IMPACT OF MASK NOSE COVERAGE AND SHAPE	24
IMPACT OF MASK COLOR	38
FNMR AND FMR VS. THRESHOLD	52
APPENDIX A DLIB MASKING METHODOLOGY	71

List of Tables

1 DATASET SUMMARY	2
2 ALGORITHM SUMMARY	6
3 ALGORITHM SUMMARY	7
4 ALGORITHM SUMMARY	8
5 FALSE NON-MATCH RATE	12
6 FALSE NON-MATCH RATE	13
7 FALSE NON-MATCH RATE	14
8 FALSE NON-MATCH RATE	15
9 FAILURE TO ENROL RATES	19
10 FAILURE TO ENROL RATES	20
11 FAILURE TO ENROL RATES	21
12 FAILURE TO ENROL RATES	22

List of Figures

1	EXAMPLES OF DEVELOPER EVOLUTION OF ACCURACY ON MASKED AND UNMASKED DATASETS.	iii
2	FNMR AND FMR CALIBRATION CURVES ON MASKED/UNMASKED COMBINATIONS.	iv
3	ENROLLMENT IMAGE EXAMPLES	2
4	SYNTHETIC FACE MASK EXAMPLES	4
5	ACCURACY GAINS	10
6	PRE VS. POST-COVID ALGORITHMS	11
7	FNMR GAIN	16
8	FNMR GAIN	17
9	FNMR GAIN	18
10	ROLE OF FTE	23
11	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	25
12	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	26
13	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	27
14	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	28
15	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	29
16	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	30
17	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	31
18	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	32
19	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	33
20	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	34
21	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	35
22	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	36
23	DET UNMASKED VERSUS MASKED - SHAPE AND COVERAGE	37
24	DET UNMASKED VERSUS MASKED - COLOR	39
25	DET UNMASKED VERSUS MASKED - COLOR	40
26	DET UNMASKED VERSUS MASKED - COLOR	41
27	DET UNMASKED VERSUS MASKED - COLOR	42
28	DET UNMASKED VERSUS MASKED - COLOR	43
29	DET UNMASKED VERSUS MASKED - COLOR	44
30	DET UNMASKED VERSUS MASKED - COLOR	45
31	DET UNMASKED VERSUS MASKED - COLOR	46
32	DET UNMASKED VERSUS MASKED - COLOR	47
33	DET UNMASKED VERSUS MASKED - COLOR	48
34	DET UNMASKED VERSUS MASKED - COLOR	49
35	DET UNMASKED VERSUS MASKED - COLOR	50
36	DET UNMASKED VERSUS MASKED - COLOR	51
37	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	53
38	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	54
39	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	55
40	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	56
41	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	57
42	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	58
43	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	59
44	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	60
45	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	61
46	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	62
47	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	63
48	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	64
49	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	65
50	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	66
51	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	67
52	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	68
53	MASKED-TO-UNMASKED VS. MASKED-TO-MASKED	69
54	DLIB MASKING METHODOLOGY	71

1 Face Mask Effects

1.1 Status

NIST has conducted the second out of a series of tests aimed at quantifying face recognition accuracy for people wearing masks. Our initial approach has been to apply masks to faces digitally (i.e., using software to apply a synthetic mask). This allowed us to leverage large datasets that we already have. This report documents results for 1:1 verification algorithms. In our first report [8], we tested “pre-pandemic” algorithms that were already submitted to FRVT 1:1 prior to mid-March 2020. This report augments its predecessor with results for more recent algorithms provided to NIST **after the COVID-19 pandemic**. While we do not have information on whether or not a particular algorithm was designed with face coverings in mind, the algorithms were submitted with the expectation that NIST would execute them on masked face images. In addition to reporting results for when only the verification image is masked, we also document the effects for the case when both enrollment and verification images are masked. This report quantifies the effect of masks on both false negative and false positives match rates and tracks developer evolution of face recognition accuracy with face masks.

The FRVT evaluation is an ongoing test that remains open to new participation. Comments and suggestions should be directed to frvt@nist.gov.

1.2 Introduction

The majority of face recognition systems have been deployed in applications where subjects make cooperative presentations to a camera, for example as part of an application for a benefit or ID credential, or as during access control. With very few exceptions such images would not include face masks or other occlusions. However, with the COVID-19 pandemic, we can anticipate a demand to authenticate persons wearing masks, for example in immigration settings, without the need to the subjects to remove those masks. This presents a problem for face recognition, because regions of the face occluded by masks - the mouth and nose - include information useful for both recognition and, potentially, the detection stage that precedes it.

Previous work on face recognition of occluded faces has been directed at situations such as crime scenes where subjects were actively un-cooperative i.e. acting to evade face detection and recognition. Those applications are often characterized by image properties (low resolution, video compression, uncontrolled head orientation) that are known [5] to degrade recognition accuracy.

2 Image Datasets

2.1 Application Images

The images are collected in an attended interview setting using dedicated capture equipment and lighting. The images are of size 300x300 pixels. The images are all high-quality frontal portraits collected in immigration offices and with a white background. As such, potential quality related drivers of high false match rates (such as blur) can be expected to be absent. The images are encoded as ISO/IEC 10918 i.e. JPEG. Over a random sample of 1000 images, the images have compressed file sizes (mean: 42KB, median: 58KB, 25-th percentile: 15KB, and 75-th percentile: 66KB). The implied bit-rates are mostly benign and superior to many e-Passports. When these images are provided as input into the algorithm, they are labeled with the type "ISO".



Figure 3: Examples of images with properties similar to the enrollment application photos used in this study. The subjects in the photos are all NIST employees.

2.2 Webcam Images

These images are taken with a camera oriented by an attendant toward a cooperating subject. This is done under time constraints, so there are roll, pitch, and yaw angle variations. Also, background illumination is sometimes bright, so the face is under exposed. Sometimes, there is perspective distortion due to close range images. The images are generally in poor conformance with the ISO/IEC 19794-5 Full Frontal image type. The images have mean interocular distance of 38 pixels. The images are all live capture. When these images are provided as input into the algorithm, they are labeled with the type "WILD". Examples of such images are included in Figure 4 and Figure 4 in NIST Interagency Report 8271. Results for verification of these images (unmasked) appear in FRVT Part 1 - Verification both compared against images of the same type, and with those described in section 2.1.

Description	#
Total images	6 244 865
Application (enrollment) images	1 019 232
Subjects in application images	1 019 232
Webcam (verification) images	5 225 633
Subjects in webcam (verification) images	2 535 329
Mated comparisons	3 225 633
Impostor comparisons	200 000 000
Subjects in mated comparisons	535 329
Subjects in impostor comparisons	3 019 232

Table 1: Summary quantities of the dataset used in this evaluation.

2.3 Synthetically Masked Images

In this test, synthetically-generated masks were overlaid on top of 1) just the probe image (webcam images described in Section 2.2) or 2) both the enrollment (application photos described in Section 2.1) and probe images. The Dlib [7] C++ toolkit version 19.19 was used to detect and establish key facial points on the face, and with the facial points, solid masks of different shape, height, and color were drawn on the face. The exact Dlib facial points and details used to generate the masks are documented in Appendix A. In the event that Dlib was unable to detect a face in the image, eye coordinates were used to generate a mask leveraging standardized token frontal geometry [1].

Examples of synthetically-masked probe images are presented in Figures 4.

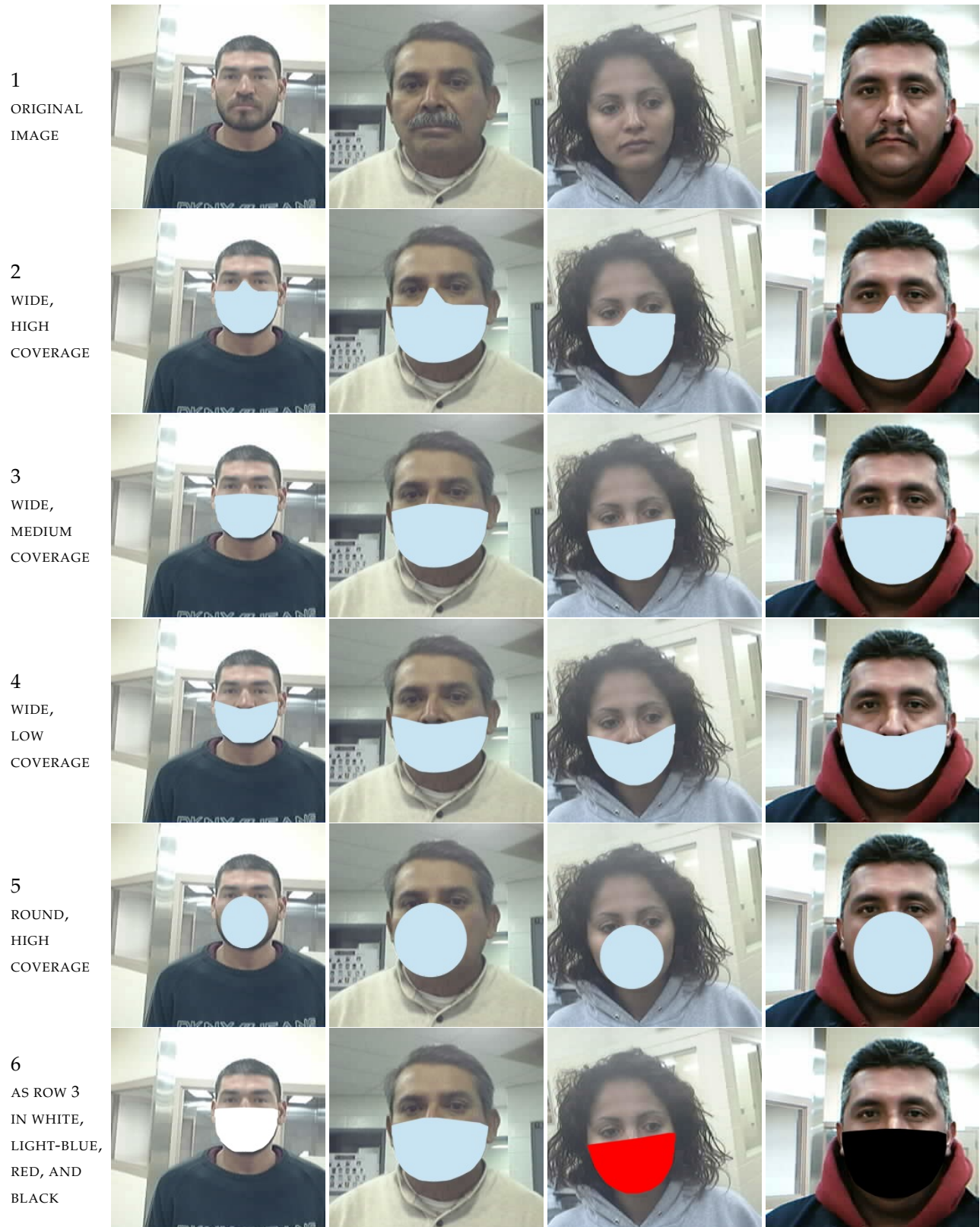


Figure 4: Examples of synthetically-generated face masks used in this study. The original images are from the NIST MEDS-II Dataset [3]. They were collected in operational settings using the same camera and procedure as is used for the border images that form the mainstay of the experiments in this report.

3 Metrics

3.1 Matching accuracy

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

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

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

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

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

The threshold, T , can take on any value. We typically generate a set of thresholds from quantiles of the observed impostor 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

$$T_k = Q(1 - \text{FMR}_k) \quad (3)$$

where Q is the quantile function, and FMR_k comes from

$$\log_{10} \text{FMR}_k = \log_{10} \text{FMR}_L + \frac{k}{K} [\log_{10} \text{FMR}_U - \log_{10} \text{FMR}_L] \quad (4)$$

Detection error tradeoff (DET) characteristics are plots of $\text{FNMR}(T)$ vs. $\text{FMR}(T)$. These are plotted with $\text{FMR}_U \rightarrow 1$ and FMR_L as low as is sustained by the number of impostor comparisons, N . This is somewhat higher than the “rule of three” limit $3/N$ [6] because samples are not independent, due to re-use of images.

3.2 Failure to Enroll

Failure to enroll (FTE) is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image. This would typically occur if a face is not detected. FTE is measured as the number of function calls that give EITHER a non-zero error code OR that give a “small” template. This is defined as one whose size is less than 60 bytes. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails yet do return a valid default data structure.

The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template to produce a low similarity score. Thus higher FTE results in higher FNMR and lower FMR.

4 Algorithms

The FRVT activity is open to participation worldwide, and the test will evaluate submissions on an ongoing basis. There is no charge to participate. The process and format of algorithm submissions to NIST are described in the FRVT 1:1 Verification Application Programming Interface (API) [9] document. Participants provide their submissions in the form of libraries compiled on a specific Linux kernel, which are linked against NIST’s test harness to produce executables. NIST provides a validation package to participants to ensure that NIST’s execution of submitted libraries produces the expected output on NIST’s test machines.

This report documents the results of algorithms submitted to FRVT 1:1 for testing both before (prior to mid-March 2020) and after the COVID-19 pandemic. Table 4 lists the algorithms that were tested. Note that algorithms that are expired or retired are not included in this report.

	Developer	Algorithm	Submission Date
1	3Divi	3divi-004	2019-07-22
2	3Divi	3divi-005	2020-08-28
3	ACI Software	acisw-003	2020-08-03
4	ADVANCE.AI	advance-002	2019-12-19
5	ASUSTek Computer Inc	asusaics-000	2019-10-24
6	AYF Technology	ayfttech-001	2020-07-06
7	Acer Incorporated	acer-000	2020-01-08
8	Acer Incorporated	acer-001	2020-06-30
9	Ai First	aifirst-001	2019-11-21
10	AiUnion Technology	aiunionface-000	2019-10-22
11	Aigen	aigen-001	2020-10-06
12	Akurat Satu Indonesia	ptakuratsatu-000	2020-09-11
13	AlphaSSTG	alphaface-002	2020-02-20
14	Anke Investments	anke-005	2019-11-21
15	Antheus Technologia	antheus-000	2019-12-05
16	Antheus Technologia	antheus-001	2020-06-25
17	Aware	aware-005	2020-02-27
18	Awidit Systems	awiros-001	2019-09-23
19	Awidit Systems	awiros-002	2020-10-28
20	Beihang University-ERCACAT	ercacat-001	2020-07-06
21	Beijing Alleyes Technology	alleyes-000	2020-03-09
22	BioID Technologies SA	bioidtechswiss-000	2019-11-15
23	BioID Technologies SA	bioidtechswiss-001	2020-08-28
24	Bresee Technology	bresee-000	2020-08-07
25	CSA IntelliCloud Technology	intelicloudai-001	2019-08-13
26	CTBC Bank	ctbcbank-000	2019-06-28
27	Camvi Technologies	camvi-004	2019-07-12
28	Canon Information Technology (Beijing)	cib-000	2019-12-11
29	Canon Information Technology (Beijing)	cib-001	2020-08-05
30	China University of Petroleum	upc-001	2019-06-05
31	Chinese Univeristy of Hong Kong	cuhkee-001	2020-03-18
32	Chosun University	chosun-000	2020-02-12
33	Chosun University	chosun-001	2020-07-01
34	Chunghwa Telecom	chtface-002	2019-12-07
35	Chunghwa Telecom	chtface-003	2020-06-24
36	Cortica	cor-001	2020-09-24
37	Cybercore	cybercore-000	2020-08-26
38	Cyberlink Corp	cyberlink-004	2020-02-27
39	Cyberlink Corp	cyberlink-005	2020-07-31
40	DSK	dsk-000	2019-06-28
41	Dahua Technology	dahua-004	2019-12-18
42	Dahua Technology	dahua-005	2020-08-13
43	Decatur Industries Inc	decatur-000	2020-08-18
44	Deepglint	deepglint-002	2019-11-15
45	DiDi ChuXing Technology	didiglobalface-001	2019-10-23
46	Expasoft LLC	expasoft-000	2020-01-06
47	Expasoft LLC	expasoft-001	2020-09-03
48	FaceSoft	facesoft-000	2019-07-10
49	Fiberhome Telecommunication Technologies	fiberhome-nanjing-002	2020-08-10
50	Fujitsu Research and Development Center	fujitsulab-000	2020-02-04
51	Fujitsu Research and Development Center	fujitsulab-001	2020-09-30
52	GeoVision Inc	geo-000	2020-06-29
53	Glory	glory-002	2019-11-12
54	Gorilla Technology	gorilla-005	2020-03-11
55	Gorilla Technology	gorilla-006	2020-07-31
56	Guangzhou Pixel Solutions	pixelall-003	2019-10-15
57	Guangzhou Pixel Solutions	pixelall-004	2020-07-02

Table 2: List of algorithms included in this report. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

	Developer	Algorithm	Submission Date
58	Hengrui AI Technology	hr-003	2020-09-25
59	ID3 Technology	id3-005	2020-08-04
60	ITMO University	itmo-007	2020-01-06
61	Idemia	idemia-005	2019-10-11
62	Idemia	idemia-006	2020-07-06
63	Imageware Systems	iws-000	2020-08-12
64	Imagus Technology Pty	imagus-001	2019-10-22
65	Imperial College London	imperial-002	2019-08-28
66	Incode Technologies Inc	incode-006	2020-02-20
67	Incode Technologies Inc	incode-007	2020-08-25
68	Innef Labs	innefulabs-000	2020-09-04
69	Innovative Technology	innovativetechnologyltd-002	2020-02-26
70	Innovatrics	innovatrics-006	2019-08-13
71	Institute of Information Technologies	iit-002	2019-12-04
72	Intel Research Group	intelresearch-001	2020-01-14
73	Intel Research Group	intelresearch-002	2020-07-24
74	Intellivision	intellivision-002	2019-08-23
75	Kakao Enterprise	kakao-003	2020-02-26
76	Kedacom International Pte	kedacom-000	2019-06-03
77	Kneron Inc	kneron-005	2020-02-21
78	Kookmin University	kookmin-001	2020-09-28
79	Lomonosov Moscow State University	intsymsu-002	2020-03-12
80	Lookman Electroplast Industries	lookman-004	2019-06-03
81	Luxand Inc	luxand-000	2019-11-07
82	MVision	mvision-001	2019-11-12
83	Momentum Digital	sertis-000	2019-10-07
84	Momentum Digital	sertis-001	2020-07-30
85	Moontime Smart Technology	mt-000	2019-06-03
86	Moontime Smart Technology	mt-002	2020-07-02
87	N-Tech Lab	ntech-008	2020-01-06
88	NSENSE Corp	nsensecorp-001	2020-10-20
89	Netbridge Technology Incoporation	netbridgetech-001	2020-01-08
90	Netbridge Technology Incoporation	netbridgetech-002	2020-08-11
91	Neurotechnology	neurotechnology-008	2020-01-08
92	Neurotechnology	neurotechnology-009	2020-07-07
93	Nodeflux	nodeflux-002	2019-08-13
94	NotionTag Technologies Private Limited	notiontag-000	2019-06-12
95	Oz Forensics LLC	oz-001	2020-07-29
96	PXL Vision AG	pxl-001	2020-06-30
97	Panasonic R+D Center Singapore	psl-005	2020-07-06
98	Paravision (EverAI)	paravision-004	2019-12-11
99	Pensees Pte	pensees-001	2020-08-17
100	Rank One Computing	rankone-009	2020-06-26
101	Rank One Computing	rankone-010	2020-11-05
102	Remark Holdings	remarkai-002	2019-11-21
103	Rokid Corporation	rokid-000	2019-08-01
104	Samsung S1 Corp	s1-001	2019-12-06
105	Satellite Innovation/Eocortex	eocortex-000	2020-08-26
106	Scanovate	scanovate-001	2019-11-12
107	Scanovate	scanovate-002	2020-06-26
108	Securif AI	securifai-001	2020-10-06
109	Sensetime Group	sensetime-003	2019-06-04
110	Shanghai Jiao Tong University	sjtu-002	2020-02-12
111	Shanghai Jiao Tong University	sjtu-003	2020-11-02
112	Shanghai Ulucu Electronics Technology	uluface-002	2019-07-10
113	Shanghai University - Shanghai Film Academy	shu-002	2019-12-10
114	Shanghai University - Shanghai Film Academy	shu-003	2020-06-24

Table 3: List of algorithms included in this report. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

	Developer	Algorithm	Submission Date
115	Shenzhen AiMall Tech	aimall-002	2020-03-12
116	Shenzhen AiMall Tech	aimall-003	2020-08-12
117	Shenzhen Intellifusion Technologies	intellifusion-002	2020-03-18
118	Staqu Technologies	staqu-000	2020-07-15
119	Star Hybrid Limited	starhybrid-001	2019-06-19
120	Su Zhou NaZhiTianDi intelligent technology	nazhai-000	2020-06-25
121	Synology Inc	synology-000	2019-10-23
122	TUPU Technology	tuputech-000	2019-10-11
123	Taiwan AI Labs	ailabs-001	2019-12-18
124	Tech5 SA	tech5-004	2020-03-09
125	Tech5 SA	tech5-005	2020-07-24
126	Tencent Deepsea Lab	deepsea-001	2019-06-03
127	Tevian	tevia-005	2019-09-21
128	Tevian	tevia-006	2020-09-11
129	Trueface.ai	trueface-000	2019-10-08
130	Trueface.ai	trueface-001	2020-07-20
131	Universidade de Coimbra	visteam-000	2020-01-14
132	Veridas Digital Authentication Solutions S.L.	veridas-004	2020-07-21
133	Via Technologies Inc	via-001	2020-01-08
134	Videmo Intelligente Videoanalyse	videmo-000	2019-12-19
135	Videonetics Technology Pvt	videonetics-002	2019-11-21
136	Viettel Group	vts-000	2020-11-04
137	Vigilant Solutions	vigilantsolutions-007	2019-06-27
138	Vigilant Solutions	vigilantsolutions-008	2020-08-03
139	VinAI Research VietNam	vinai-000	2020-09-24
140	VisionLabs	visionlabs-008	2020-01-06
141	VisionLabs	visionlabs-009	2020-07-27
142	Vocord	vocord-008	2020-01-031
143	Winsense	winsense-001	2019-10-16
144	X-Laboratory	x-laboratory-001	2020-01-21
145	Xforward AI Technology	xforwardai-000	2020-02-06
146	Xforward AI Technology	xforwardai-001	2020-09-25
147	Xiamen University	xm-000	2020-10-19
148	YooniK	yooniK-000	2020-06-24
149	Yuan High-Tech Development	yuan-000	2020-06-30
150	iQIYI Inc	iqface-000	2019-06-04
151	iQIYI Inc	iqface-002	2020-07-30
152	iSAP Solution Corporation	isap-002	2020-09-01

Table 4: List of algorithms included in this report. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

5 Results

This section includes accuracy results for the 152 one-to-one verification algorithms listed in Section 4, of which 65 were submitted to FRVT after mid-March 2020 and are labeled in blue in figures and tables throughout this report. We do not include speed and computational resource requirements - they are given in Table 1 in the FRVT 1:1 [report](#). Missing entries generally mean the algorithm was not run on that particular mask variation due to time and resource constraints. The results, which span many pages, are comprised of:

- ▷ **Evolution of algorithm performance with face masks:** Figure 5 shows the evolution of performance with face masks for developers that have submitted algorithms since mid-March 2020.
- ▷ **FNMR - summary:** Figure 6 gives a summary of false non-matches rates between pre and post-COVID algorithms with a common type of mask. FNMR values are stated at a fixed threshold calibrated to give FMR = 0.00001 on unmasked images.
- ▷ **FNMR - detailed:** Table 5 tabulates false non-match rates by color, shape, and nose coverage. It includes also FNMR without any mask. FNMR values are stated at a fixed threshold calibrated to give FMR = 0.00001 on unmasked images.
- ▷ **Mask vs. no mask:** The scatter plot in Figure 7 shows variation across all algorithms of FNMR without masks against FNMR with a common type of mask, broken out by pre and post-COVID algorithms.
- ▷ **Mask shape:** The scatter plot in Figure 8 shows for all algorithms the increase in false negative results for wide masks vs. narrower round masks, broken out by pre and post-COVID algorithms.
- ▷ **Mask nose coverage:** The scatter plot in Figure 9 shows for all algorithms the increase in false negative rates for masks that substantially cover the nose and those pulled beneath the nose, broken out by pre and post-COVID algorithms.
- ▷ **FTE:** Table 9 gives empirical failure-to-template results by color, shape, and nose coverage. The table was produced using 10 000 images of each kind of mask.
- ▷ **FTE as contributor to FNMR:** The FNMR results include failure-to-template rates (FTE). Figure 10 shows the proportion of template generation failures, broken out by pre and post-COVID algorithms.
- ▷ **DET - impact of mask nose coverage and shape:** This [section](#) of figures shows detection error tradeoff characteristics for each algorithm, across different mask nose coverages and shapes.
- ▷ **DET - impact of mask color:** This [section](#) of figures shows detection error tradeoff characteristics for each algorithm, across mask colors.
- ▷ **FNMR and FMR vs. threshold:** This [section](#) of figures shows the explicit dependence of false non-match rate and false match rate on threshold.

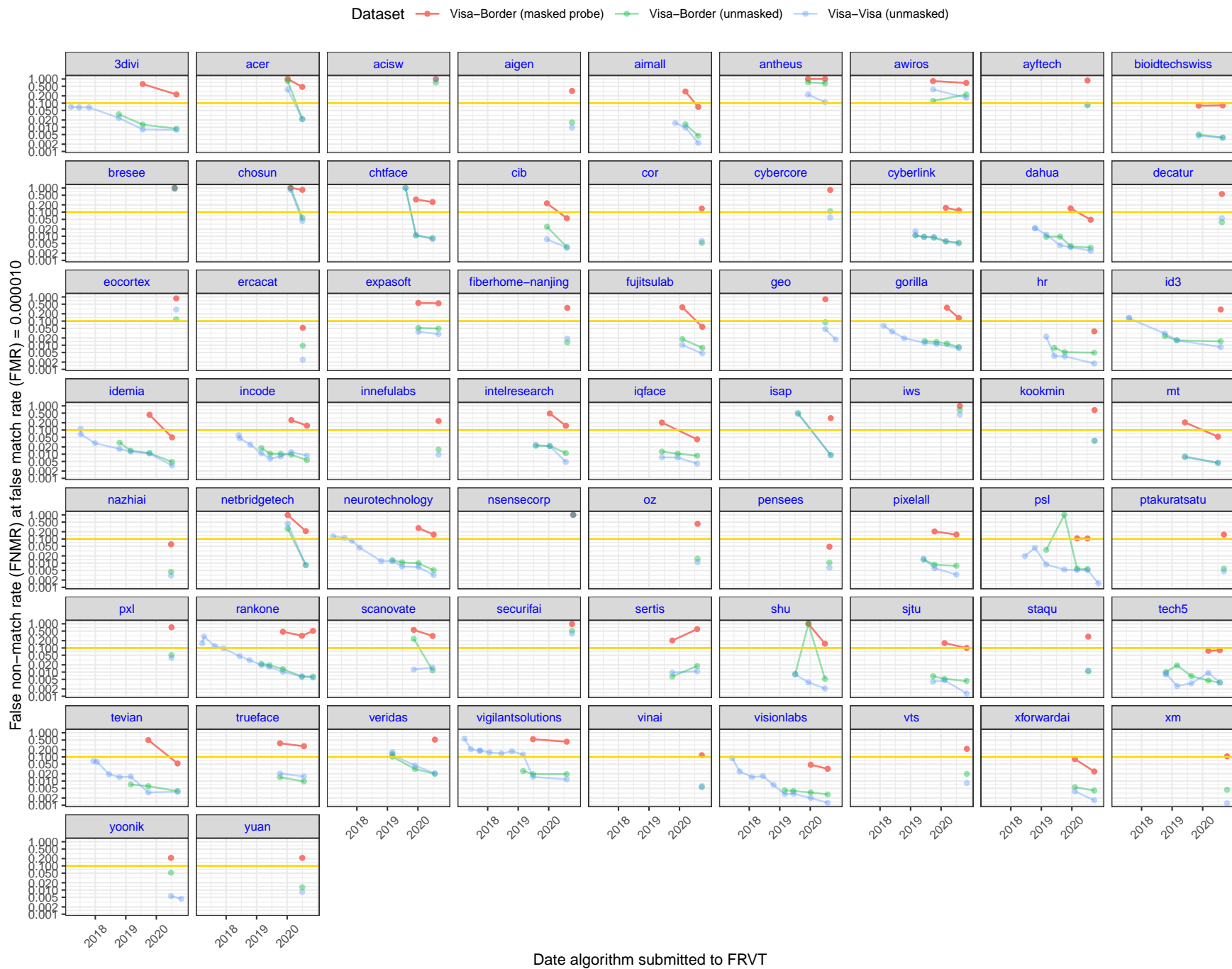
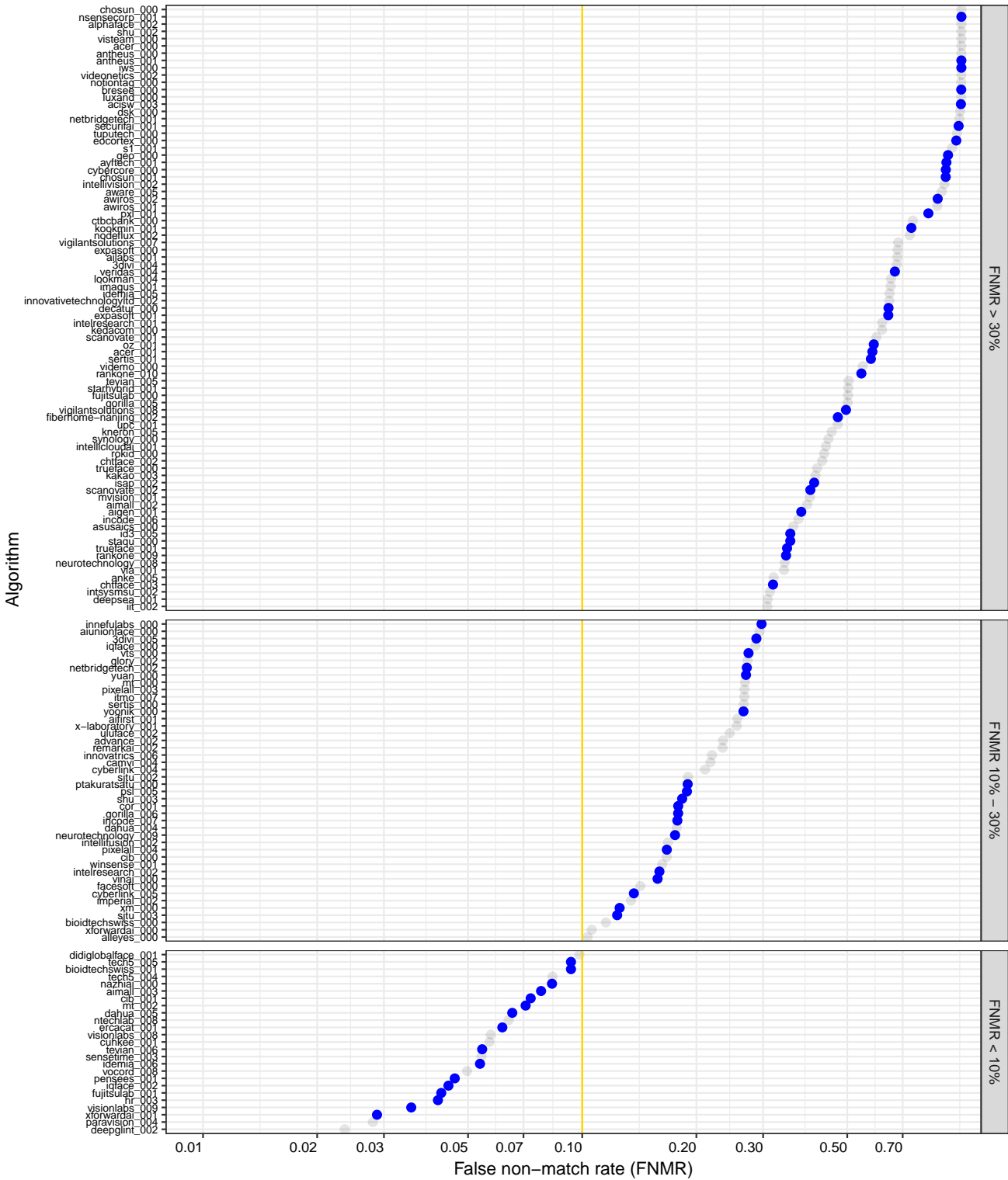


Figure 5: Evolution of accuracy with face masks for developers that have submitted algorithms since mid-March 2020. Red line represents results on a masked dataset (masked probe), and for reference, the blue and green lines are results for unmasked datasets. FNMR on masked dataset is for medium, wide, lightblue masks. The horizontal gold line shows where FNMR=10%.

FNMR for medium wide lightblue masks

Algorithm Submission ● Post-COVID ● Pre-COVID



This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

Figure 6: For each algorithm, each dot shows FNMR @ FMR=0.00001, where the threshold is set for FMR on unmasked probe images. The results are for when the probe is masked, and the enrollment image is unmasked. Gray dots represent results for algorithms submitted prior to mid-March 2020 (pre-COVID), and blue dots represent algorithms submitted thereafter (post-COVID).

	Algorithm Name	NOT MASKED	MASK COLOR = LIGHTBLUE						COLOR = BLACK			COLOR = RED	COLOR = WHITE
			SHAPE = WIDE			SHAPE = ROUND			SHAPE = WIDE			SHAPE = WIDE	SHAPE = WIDE
			COVERAGE	LO	MED	HI	LO	MED	HI	LO	MED	HI	MED
1	3divi-004	0.0130 ⁸⁷	0.4123 ⁹⁹	0.6760 ¹¹⁷	-	-	-	-	-	-	-	-	-
2	3divi-005	0.0088 ⁶²	0.1006 ⁷⁰	0.2895 ⁶⁷	0.3670 ⁴⁸	0.0783 ⁵⁰	0.1918 ⁴⁹	-	0.1185 ⁵⁰	0.3057 ⁵³	0.4272 ⁴⁹	-	-
3	acer-000	0.8432 ¹⁴⁷	0.9995 ¹²³	0.9999 ¹⁴⁷	-	-	-	-	-	-	-	-	-
4	acer-001	0.0219 ¹⁰⁸	0.2587 ⁹²	0.5835 ¹⁰⁵	0.6719 ⁶²	0.1536 ⁵⁷	0.4325 ⁵⁷	0.6831 ⁴⁸	0.3132 ⁵⁶	0.6304 ⁶⁴	0.7143 ⁵⁵	-	-
5	acisw-003	0.7160 ¹⁴⁵	0.9810 ¹¹⁶	0.9970 ¹³⁹	0.9970 ⁷²	-	-	-	0.9883 ⁶⁴	0.9997 ⁷⁶	-	-	-
6	advance-002	0.0328 ¹¹⁷	-	0.2351 ⁵²	-	-	-	-	-	-	-	0.2333 ⁴⁵	-
7	aifirst-001	0.0079 ⁵⁴	0.0778 ⁵⁶	0.2567 ⁵⁵	-	-	-	-	-	-	-	0.2624 ⁴⁹	-
8	aigen-001	0.0159 ⁹⁹	0.1268 ⁷⁸	0.3790 ⁸³	0.4432 ⁵⁶	-	0.2880 ⁵⁶	-	0.1956 ⁵⁴	0.4761 ⁵⁹	0.6329 ⁵⁴	0.3189 ⁵¹	0.3754 ³⁹
9	ailabs-001	0.0243 ¹¹⁰	-	0.6792 ¹¹⁸	-	-	-	-	-	-	-	-	-
10	aimall-002	0.0133 ⁹¹	-	0.3919 ⁸⁴	-	-	-	-	-	-	-	-	-
11	aimall-003	0.0045 ¹⁹	0.0188 ¹⁶	0.0781 ²⁰	0.1325 ²¹	0.0175 ¹⁵	0.0524 ²⁴	0.1021 ²³	0.0223 ¹⁶	0.0913 ²⁰	0.1577 ²⁵	0.0738 ¹³	0.0800 ¹⁷
12	aiunionface-000	0.0094 ⁶⁶	0.0917 ⁶⁶	0.2935 ⁶⁸	-	-	-	-	-	-	-	-	-
13	alleges-000	0.0044 ¹⁸	-	0.1038 ²⁶	-	0.0181 ¹⁹	0.0542 ²⁶	0.1050 ²⁴	0.0262 ²⁴	0.1287 ³¹	0.1991 ³¹	0.1066 ²¹	0.1098 ²³
14	alphaface-002	1.0000 ¹⁵²	1.0000 ¹²⁵	1.0000 ¹⁵⁰	-	-	-	-	-	-	-	-	-
15	anke-005	0.0062 ⁴⁴	0.0671 ⁴⁹	0.3207 ⁷⁴	-	-	-	-	-	-	-	-	-
16	antheus-000	0.7319 ¹⁴⁶	0.9994 ¹²²	0.9999 ¹⁴⁶	-	-	-	-	-	-	-	-	-
17	antheus-001	0.6608 ¹⁴²	0.9988 ¹²¹	0.9998 ¹⁴⁵	0.9998 ⁷³	0.9993 ⁶⁵	0.9998 ⁶⁷	0.9998 ⁵⁴	0.9993 ⁶⁶	0.9998 ⁷⁷	0.9998 ⁶⁴	-	-
18	asusaics-000	0.0090 ⁶⁴	-	0.3616 ⁸¹	-	-	-	-	-	-	-	-	-
19	aware-005	0.0308 ¹¹⁶	0.4962 ¹⁰³	0.8876 ¹²⁷	-	-	-	-	-	-	-	-	-
20	awiros-001	0.1233 ¹³³	0.6823 ¹¹⁰	0.8635 ¹²⁵	-	-	-	-	-	-	-	-	-
21	awiros-002	0.2283 ¹³⁷	0.6356 ¹⁰⁷	0.8671 ¹²⁶	0.9221 ⁶⁸	-	-	-	0.7703 ⁶²	0.9068 ⁷³	-	0.8628 ⁵⁷	0.8508 ⁴²
22	ayftech-001	0.0828 ¹²⁹	0.6740 ¹⁰⁸	0.9132 ¹³¹	0.9333 ⁷⁰	0.5865 ⁶³	0.8519 ⁶⁴	0.9538 ⁵³	0.7540 ⁶¹	0.9048 ⁷²	0.9705 ⁶²	-	-
23	bioidtechswiss-000	0.0050 ²⁴	0.0308 ²³	0.1155 ²⁸	0.1840 ³⁰	0.0223 ²⁴	0.0632 ²⁸	0.1207 ²⁸	0.0331 ²⁷	0.1163 ²⁷	0.1786 ²⁷	0.1167 ²³	0.1167 ²⁵
24	bioidtechswiss-001	0.0037 ⁶	0.0225 ²¹	0.0945 ²⁴	0.1519 ²⁵	0.0180 ¹⁸	0.0524 ²⁵	0.1070 ²⁶	0.0254 ²¹	0.0986 ²³	0.1571 ²³	0.0979 ¹⁹	0.0945 ²²
25	bresee-000	0.9802 ¹⁴⁸	0.9968 ¹²⁰	0.9989 ¹⁴¹	-	-	-	-	0.9975 ⁶⁵	-	-	-	-
26	camvi-004	0.0063 ⁴⁵	0.0697 ⁵¹	0.2179 ⁵⁰	-	-	-	-	-	-	-	0.3337 ⁵²	-
27	chosun-000	1.0000 ¹⁵³	1.0000 ¹³⁵	1.0000 ¹⁵²	-	-	-	-	-	-	-	-	-
28	chosun-001	0.0582 ¹²⁸	0.5759 ¹⁰⁴	0.9091 ¹²⁹	0.9263 ⁶⁹	0.5021 ⁶²	0.8531 ⁶⁵	0.9278 ⁵¹	0.5158 ⁵⁸	0.9031 ⁷¹	0.9409 ⁶⁰	-	-
29	chtface-002	0.0108 ⁷⁶	0.1423 ⁸⁰	0.4303 ⁹⁰	-	-	-	-	-	-	-	-	-
30	chtface-003	0.0084 ⁵⁶	0.1008 ⁷¹	0.3201 ⁷³	-	0.0774 ⁴⁹	0.2393 ⁵⁴	0.4387 ⁴⁴	0.1232 ⁵²	0.3309 ⁵⁶	0.5044 ⁵¹	-	-
31	cib-000	0.0249 ¹¹¹	0.0757 ⁵⁴	0.1670 ³⁷	-	-	-	-	-	-	-	0.3825 ⁵³	0.1859 ³⁶
32	cib-001	0.0036 ⁵	0.0149 ¹⁰	0.0731 ¹⁹	0.1095 ¹⁶	0.0125 ⁵	0.0336 ⁹	0.0690 ⁷	0.0175 ⁹	0.0731 ¹⁵	0.1137 ¹⁰	0.0863 ¹⁷	0.0779 ¹⁶
33	clova-000	0.0087 ⁶⁰	0.0881 ⁶⁴	-	0.5017 ⁵⁸	-	-	-	0.0950 ⁴⁷	0.3051 ⁵²	-	-	-
34	cor-001	0.0053 ²⁹	0.0504 ⁴²	0.1802 ⁴³	0.2470 ³⁹	0.0300 ³¹	0.0864 ³¹	-	0.0364 ³¹	0.1328 ³²	0.1828 ²⁸	0.1527 ²⁹	-
35	ctcbcbank-000	0.0133 ⁸⁹	0.1594 ⁸⁷	0.7448 ¹²³	-	-	-	-	-	-	-	-	-
36	cuhkee-001	0.0041 ¹³	0.0143 ⁹	0.0572 ¹³	0.0963 ¹³	0.0143 ⁹	0.0333 ⁸	0.0715 ⁸	0.0164 ⁸	0.0652 ¹⁰	0.1193 ¹¹	-	-
37	cybercore-000	0.1096 ¹³¹	-	0.9097 ¹³⁰	-	-	-	-	-	-	-	-	-
38	cyberlink-004	0.0061 ⁴²	0.0538 ⁴⁵	0.2115 ⁴⁹	-	-	-	-	-	-	-	0.2591 ⁴⁸	-
39	cyberlink-005	0.0055 ³⁰	0.0376 ³²	0.1374 ³²	0.1808 ²⁹	0.0388 ⁴²	0.1018 ³⁵	0.2695 ³⁹	0.0451 ³⁸	0.1610 ³⁷	0.3991 ⁴⁶	0.4285 ⁵⁵	0.1316 ²⁶
40	dahua-004	0.0038 ⁹	0.0328 ²⁷	0.1784 ⁴¹	0.2026 ³²	-	-	-	0.0226 ¹⁸	0.1186 ²⁸	0.1983 ³⁰	-	-
41	dahua-005	0.0034 ⁴	0.0155 ¹¹	0.0660 ¹⁷	0.1135 ¹⁹	0.0127 ⁷	0.0340 ¹⁰	0.0829 ¹³	0.0137 ⁶	0.0601 ⁸	0.1019 ⁸	0.0579 ⁶	0.0613 ¹⁰
42	decat-000	0.0384 ¹¹⁹	-	0.6440 ¹¹¹	-	-	-	-	-	-	-	-	-
43	deepglint-002	0.0039 ¹¹	0.0077 ¹	0.0237 ¹	0.0455 ¹	0.0078 ¹	0.0141 ¹	0.0292 ¹	0.0083 ¹	0.0254 ¹	0.0513 ¹	0.0234 ¹	0.0248 ¹
44	deepsea-001	0.0110 ⁷⁹	0.1218 ⁷⁵	0.3094 ⁷¹	0.3778 ⁵²	0.0922 ⁵⁴	0.2217 ⁵²	0.4469 ⁴⁷	-	-	-	-	-
45	didiglobalface-001	0.0050 ²⁵	-	0.0986 ²⁵	0.1517 ²³	0.0255 ²⁵	0.0515 ²²	0.0979 ²⁰	0.0291 ²⁵	0.1033 ²⁵	0.1558 ²²	0.1241 ²⁴	0.1655 ³²

Table 5: This table summarizes False Non-Match Rate (FNMR) on unmasked and masked probe images. FNMR is the proportion of mated comparisons below a threshold set to achieve FMR=1e-05 on unmasked probe images. False Match Rate (FMR) is the proportion of impostor comparisons at or above that threshold. The red superscripts give rank over all algorithms in that column. Missing entries generally mean the algorithm was not run on that particular mask variation due to time and resource constraints. Algorithms with FTE=1.00 were not run at all. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

	Algorithm Name	NOT MASKED	MASK COLOR = LIGHTBLUE						COLOR = BLACK			COLOR = RED	COLOR = WHITE
			SHAPE = WIDE			SHAPE = ROUND			SHAPE = WIDE			SHAPE = WIDE	SHAPE = WIDE
			COVERAGE	LO	MED	HI	LO	MED	HI	LO	MED	HI	MED
46	dsk-000	0.1961 ¹³⁴	0.9108 ¹¹³	0.9929 ¹³⁸	-	-	-	-	-	-	-	-	-
47	eocortex-000	0.1187 ¹³²	-	0.9694 ¹³⁴	-	-	-	-	-	-	-	-	-
48	ercacat-001	0.0096 ⁶⁹	0.0187 ¹⁵	0.0616 ¹⁵	0.0994 ¹⁴	0.0173 ¹⁴	0.0357 ¹¹	0.0728 ¹⁰	0.0200 ¹³	0.0663 ¹²	0.1114 ⁹	0.0679 ¹²	0.0648 ¹²
49	expasoft-000	0.0519 ¹²⁶	0.3186 ⁹⁶	0.6796 ¹¹⁹	-	-	-	-	-	-	-	-	-
50	expasoft-001	0.0492 ¹²⁴	-	0.6414 ¹¹⁰	-	-	-	-	-	-	-	-	-
51	facesoft-000	0.0057 ³⁶	0.0397 ³⁴	0.1428 ³³	-	-	-	-	-	0.1573 ³⁶	-	0.1446 ²⁸	0.1428 ²⁹
52	fiberhome-nanjing-002	0.0133 ⁹⁰	0.1554 ⁸⁶	0.4734 ⁹⁶	-	-	-	-	0.1641 ⁵³	-	-	-	-
53	fujitsulab-000	0.0180 ¹⁰³	-	0.5052 ¹⁰¹	-	-	-	-	-	-	-	-	-
54	fujitsulab-001	0.0078 ⁵³	0.0330 ²⁸	0.0427 ⁶	0.0608 ⁴	0.0311 ³²	0.0363 ¹²	0.0666 ⁵	0.0407 ³⁴	0.0985 ²¹	0.1525 ²¹	0.0898 ¹⁸	0.0776 ¹⁵
55	geo-000	0.0902 ¹³⁰	0.5765 ¹⁰⁵	0.9236 ¹³²	-	-	0.8082 ⁶³	0.9311 ⁵²	0.6214 ⁶⁰	0.9428 ⁷⁴	0.9541 ⁶¹	-	-
56	glory-002	0.0109 ⁷⁸	-	0.2729 ⁶⁴	-	-	-	-	-	-	-	-	-
57	gorilla-005	0.0117 ⁸³	0.1463 ⁸²	0.5037 ⁹⁹	-	-	-	-	-	-	-	-	-
58	gorilla-006	0.0085 ⁵⁷	0.0495 ⁴¹	0.1805 ⁴⁴	0.2883 ⁴³	0.0385 ⁴¹	0.1134 ⁴²	-	0.2260 ⁴⁶	0.3419 ⁴³	-	0.1980 ⁴⁰	-
59	hr-003	0.0049 ²³	0.0133 ⁵	0.0419 ⁵	0.0613 ⁵	0.0122 ³	0.0247 ⁴	0.0475 ⁴	0.0142 ⁷	0.0527 ⁶	0.0914 ⁵	0.0542 ⁵	0.0476 ⁷
60	id3-005	0.0145 ⁹⁶	0.1181 ⁷⁴	0.3540 ⁸⁰	-	-	-	-	-	-	-	-	-
61	idemia-005	0.0111 ⁸¹	0.2051 ⁹⁰	0.6469 ¹¹³	0.6968 ⁶⁴	0.1349 ⁵⁶	0.4387 ⁵⁸	-	0.2786 ⁵⁵	0.7402 ⁶⁹	0.8119 ⁵⁶	-	-
62	idemia-006	0.0048 ²²	0.0327 ²⁶	0.0539 ¹⁰	0.0911 ¹⁰	0.0322 ³³	0.0440 ¹⁹	0.0946 ¹⁷	0.0359 ²⁹	0.0714 ¹⁴	0.1489 ¹⁹	0.0607 ⁸	0.0564 ⁹
63	iit-002	0.0141 ⁹⁴	-	0.3078 ⁷⁰	-	-	-	-	-	-	-	-	-
64	imagus-001	0.0276 ¹¹³	0.3488 ⁹⁸	0.6510 ¹¹⁴	-	-	-	-	-	-	-	-	-
65	imperial-002	0.0055 ³¹	0.0320 ²⁵	0.1350 ³¹	0.1972 ³¹	0.0258 ²⁶	0.0775 ³⁰	0.1556 ²⁹	0.0359 ³⁰	0.1510 ³⁵	0.2302 ³⁶	0.1533 ³⁰	0.1432 ³⁰
66	incode-006	0.0095 ⁶⁸	-	0.3725 ⁸²	-	-	-	-	-	-	-	-	-
67	incode-007	0.0058 ³⁷	0.0470 ³⁸	0.1790 ⁴²	0.2393 ³⁸	0.0323 ³⁴	0.1040 ³⁷	0.1779 ³¹	0.0561 ⁴¹	0.1891 ⁴³	0.2754 ³⁹	0.1978 ³⁹	-
68	innefulabs-000	0.0155 ⁹⁸	-	0.2971 ⁶⁹	-	-	-	-	-	-	-	-	-
69	innovativetechnologyltd-002	0.0251 ¹¹²	0.2701 ⁹³	0.6454 ¹¹²	-	-	-	-	-	-	-	-	-
70	innovatrics-006	0.0059 ⁴⁰	0.0543 ⁴⁸	0.2210 ⁵¹	0.3118 ⁴⁴	0.0369 ³⁷	0.1109 ⁴⁰	0.1984 ³³	0.0557 ⁴⁰	0.1909 ⁴⁴	0.2764 ⁴⁰	0.1936 ³⁸	-
71	intellicloudai-001	0.0095 ⁶⁷	0.1044 ⁷²	0.4394 ⁹²	-	-	-	-	-	-	-	-	-
72	intellifusion-002	0.0056 ³⁵	0.0539 ⁴⁶	0.1690 ³⁹	-	-	-	-	0.1822 ⁴²	-	-	0.2556 ⁴⁷	0.2119 ³⁷
73	intellivision-002	0.0463 ¹²³	0.5999 ¹⁰⁶	0.9028 ¹²⁸	-	-	-	-	-	-	-	-	-
74	intelresearch-001	0.0220 ¹⁰⁹	0.2254 ⁹¹	0.6184 ¹⁰⁸	-	-	-	-	-	-	-	-	-
75	intelresearch-002	0.0111 ⁸⁰	0.0458 ³⁶	0.1598 ³⁵	0.2085 ³³	0.0384 ⁴⁰	0.1025 ³⁶	0.2568 ³⁸	0.0598 ⁴²	0.1764 ⁴¹	0.2583 ³⁸	0.1857 ³⁶	0.1595 ³¹
76	intsysmsu-002	0.0089 ⁶³	0.0827 ⁶⁰	0.3138 ⁷²	-	-	-	-	-	-	-	-	-
77	iqface-000	0.0128 ⁸⁶	0.0885 ⁶⁵	0.2867 ⁶⁶	-	-	-	-	-	-	-	-	-
78	iqface-002	0.0086 ⁵⁸	0.0193 ¹⁷	0.0445 ⁷	0.0725 ⁷	0.0184 ²¹	0.0317 ⁶	0.0671 ⁶	0.0215 ¹⁵	0.0643 ⁹	0.0997 ⁷	0.0633 ¹⁰	0.0448 ⁵
79	isap-002	0.0094 ⁶⁵	-	0.4090 ⁸⁷	-	-	-	-	-	-	-	-	-
80	itmo-007	0.0098 ⁷²	0.0840 ⁶¹	0.2685 ⁶⁰	-	-	-	-	-	-	-	-	-
81	iws-000	0.6797 ¹⁴³	0.9960 ¹¹⁸	0.9997 ¹⁴⁴	-	-	-	-	-	-	-	-	-
82	kakao-003	0.0170 ¹⁰¹	0.1541 ⁸⁵	0.4123 ⁸⁸	-	-	-	-	-	-	-	-	-
83	kedacom-000	0.0391 ¹²⁰	0.3444 ⁹⁷	0.6188 ¹⁰⁹	0.6848 ⁶³	0.2663 ⁵⁹	0.5975 ⁶⁰	-	-	-	-	-	-
84	kneron-005	0.0296 ¹¹⁵	-	0.4567 ⁹⁴	-	-	-	-	-	-	-	-	-
85	kookmin-001	0.0357 ¹¹⁸	0.4479 ¹⁰¹	0.7378 ¹²²	0.8497 ⁶⁶	0.2933 ⁶⁰	0.6006 ⁶¹	-	0.4160 ⁵⁷	0.7365 ⁶⁸	0.8917 ⁵⁸	0.8642 ⁵⁸	0.7432 ⁴¹
86	lookman-004	0.0398 ¹²¹	-	0.6520 ¹¹⁵	-	-	-	-	-	-	-	-	-
87	luxand-000	0.2167 ¹³⁶	0.9732 ¹¹⁵	0.9988 ¹⁴⁰	-	-	-	-	-	-	-	-	-
88	mt-000	0.0075 ⁵¹	0.0768 ⁵⁵	0.2700 ⁶¹	0.3736 ⁵⁰	0.0482 ⁴³	0.1746 ⁴⁶	-	0.0749 ⁴³	0.3084 ⁵⁴	0.4239 ⁴⁸	-	-
89	mt-002	0.0043 ¹⁵	0.0215 ¹⁸	0.0713 ¹⁸	0.1114 ¹⁷	0.0162 ¹²	0.0404 ¹⁵	0.0777 ¹²	0.0225 ¹⁷	0.0877 ¹⁹	0.1399 ¹⁸	0.0804 ¹⁵	0.0757 ¹⁴
90	mvision-001	0.0137 ⁹²	-	0.3987 ⁸⁵	-	-	-	-	-	-	-	-	-

Table 6: This table summarizes False Non-Match Rate (FNMR) on unmasked and masked probe images. FNMR is the proportion of mated comparisons below a threshold set to achieve FMR=1e-05 on unmasked probe images. False Match Rate (FMR) is the proportion of impostor comparisons at or above that threshold. The red superscripts give rank over all algorithms in that column. Missing entries generally mean the algorithm was not run on that particular mask variation due to time and resource constraints. Algorithms with FTE=1.00 were not run at all. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

	Algorithm Name	NOT MASKED	MASK COLOR = LIGHTBLUE						COLOR = BLACK			COLOR = RED	COLOR = WHITE
			SHAPE = WIDE			SHAPE = ROUND			SHAPE = WIDE			SHAPE = WIDE	SHAPE = WIDE
			COVERAGE	LO	MED	HI	LO	MED	HI	LO	MED	HI	MED
136	via-001	0.0097 ⁷⁰	0.1234 ⁷⁶	0.3406 ⁷⁵	-	-	-	-	-	-	-	-	-
137	videmo-000	0.0140 ⁹³	-	0.5509 ¹⁰³	-	-	-	-	-	-	-	-	-
138	videonetics-002	0.6032 ¹⁴¹	0.9941 ¹¹⁷	0.9996 ¹⁴³	-	-	-	-	-	-	-	-	-
139	vigilantsolutions-007	0.0194 ¹⁰⁴	0.2849 ⁹⁵	0.6839 ¹²⁰	-	-	-	-	-	-	-	-	-
140	vigilantsolutions-008	0.0195 ¹⁰⁵	-	0.4970 ⁹⁷	0.5521 ⁶⁰	-	-	-	-	0.6365 ⁶⁵	-	-	-
141	vinai-000	0.0056 ³³	0.0388 ³³	0.1587 ³⁴	0.2130 ³⁴	0.0286 ²⁹	0.0965 ³⁴	-	0.0364 ³²	0.1487 ³⁴	0.2024 ³²	0.1691 ³³	0.1710 ³³
142	visionlabs-008	0.0034 ³	0.0139 ⁶	0.0579 ¹⁴	0.1014 ¹⁵	0.0154 ¹⁰	0.0412 ¹⁶	0.1004 ²¹	0.0187 ¹²	0.0664 ¹³	0.1284 ¹⁵	0.0614 ⁹	0.0618 ¹¹
143	visionlabs-009	0.0028 ¹	0.0095 ³	0.0355 ⁴	0.0667 ⁶	0.0123 ⁴	0.0307 ⁵	0.0719 ⁹	0.0123 ³	0.0423 ⁴	0.0943 ⁶	0.0399 ⁴	0.0395 ⁴
144	visteam-000	0.9960 ¹⁴⁹	1.0000 ¹²⁴	1.0000 ¹⁴⁸	-	-	-	-	-	-	-	-	-
145	vocord-008	0.0038 ⁸	0.0140 ⁸	0.0500 ⁹	0.0762 ⁸	0.0176 ¹⁶	0.0393 ¹⁴	0.0892 ¹⁴	0.0135 ⁵	0.0459 ⁵	0.0771 ⁴	0.0607 ⁷	0.0482 ⁸
146	vts-000	0.0199 ¹⁰⁶	0.0870 ⁶²	0.2755 ⁶⁵	0.3566 ¹⁷	-	-	-	-	0.2584 ⁵⁰	-	0.2249 ⁴³	0.2976 ³⁸
147	winsense-001	0.0058 ³⁸	0.0473 ³⁹	0.1626 ³⁶	0.2244 ³⁷	0.0325 ³⁵	0.0946 ³²	0.1853 ³²	0.0406 ³³	0.1471 ³³	0.2231 ³⁵	0.1622 ³¹	0.1843 ³⁵
148	x-laboratory-001	0.0058 ³⁹	0.0517 ⁴³	0.2569 ⁵⁶	-	-	-	-	-	-	-	0.2333 ⁴⁴	-
149	xforwardai-000	0.0056 ³²	0.0235 ²²	0.1064 ²⁷	0.1615 ²⁷	0.0197 ²²	0.0606 ²⁷	0.1156 ²⁷	0.0255 ²³	0.1091 ²⁶	0.1608 ²⁶	0.1108 ²²	0.1101 ²⁴
150	xforwardai-001	0.0041 ¹²	0.0087 ²	0.0289 ³	0.0536 ³	0.0087 ²	0.0180 ²	0.0377 ³	0.0090 ²	0.0303 ²	0.0544 ²	0.0313 ²	0.0294 ²
151	xm-000	0.0044 ¹⁷	0.0334 ²⁹	0.1255 ³⁰	0.1682 ²⁸	0.0275 ²⁸	0.0774 ²⁹	0.1648 ³⁰	0.0324 ²⁶	0.1274 ²⁹	0.1839 ²⁹	0.1706 ³⁴	0.1381 ²⁸
152	yoonek-000	0.0520 ¹²⁷	0.1056 ⁷³	0.2675 ⁵⁷	0.3773 ⁵¹	0.0895 ⁵²	0.1823 ⁴⁷	0.3214 ⁴¹	0.1027 ⁴⁹	0.2582 ⁴⁹	0.3628 ⁴⁴	-	-
153	yuan-000	0.0130 ⁸⁸	0.0871 ⁶³	0.2704 ⁶²	0.3516 ⁴⁵	0.0654 ⁴⁶	0.1661 ⁴⁵	-	0.0805 ⁴⁴	0.2492 ⁴⁸	0.3645 ⁴⁵	-	-

Table 8: This table summarizes False Non-Match Rate (FNMR) on unmasked and masked probe images. FNMR is the proportion of mated comparisons below a threshold set to achieve FMR=1e-05 on unmasked probe images. False Match Rate (FMR) is the proportion of impostor comparisons at or above that threshold. The red superscripts give rank over all algorithms in that column. Missing entries generally mean the algorithm was not run on that particular mask variation due to time and resource constraints. Algorithms with FTE=1.00 were not run at all. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

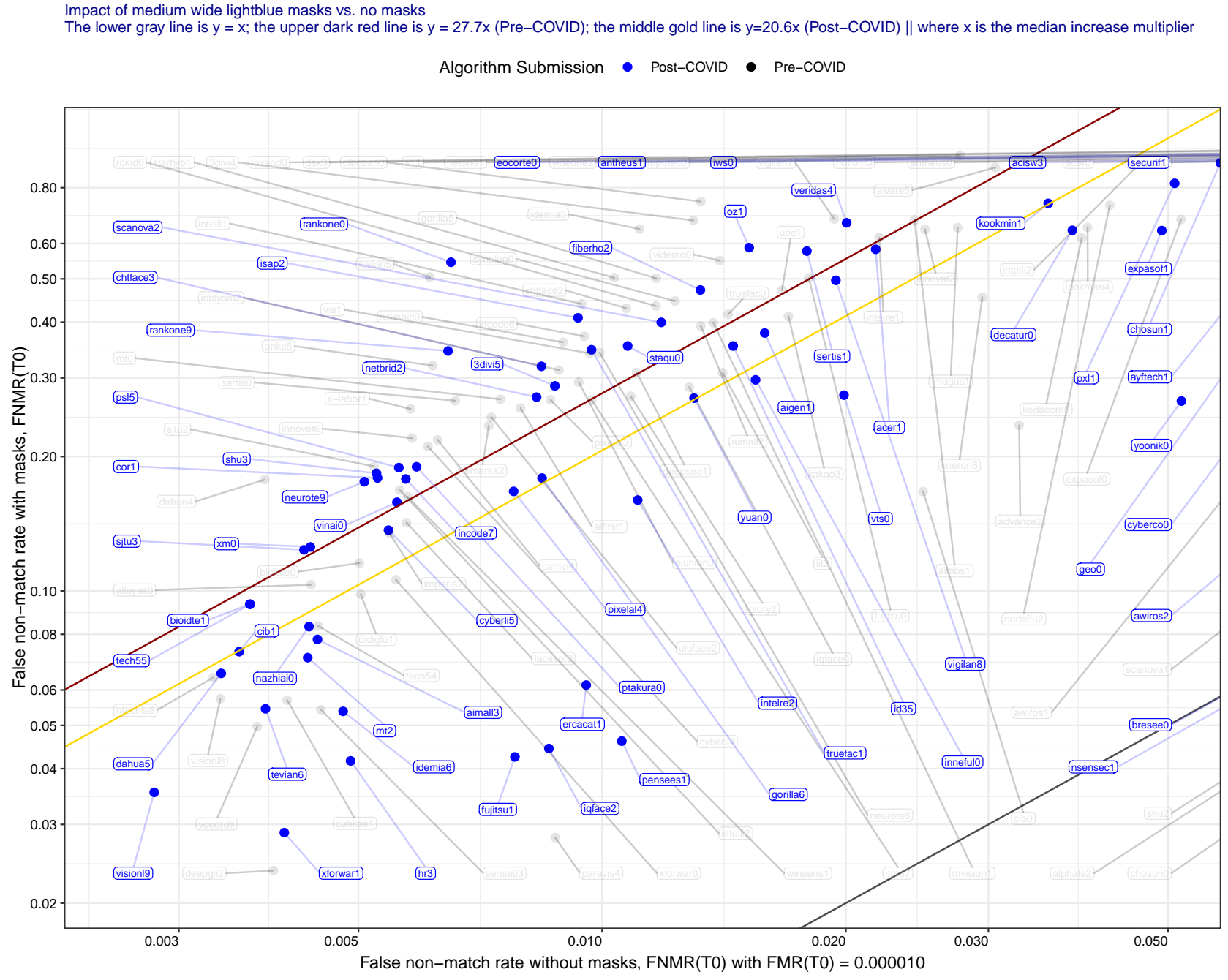


Figure 7: At a fixed threshold, a plot of FNMR with and without masks. The displacement of the dark red line relative to the black "parity" line shows a large increase in FNMR with masks for pre-COVID algorithms. The reduction in distance (relative to the black line) observed in the gold line indicates a reduction in median FNMR with masks for post-COVID algorithms. The value in the title is the median increase multiplier.

Impact of wide vs. round shape for medium lightblue masks
 The lower gray line is $y = x$; the upper dark red line is $y = 1.7x$ (Pre-COVID); the middle gold line is $y = 1.6x$ (Post-COVID) || where x is the median increase multiplier

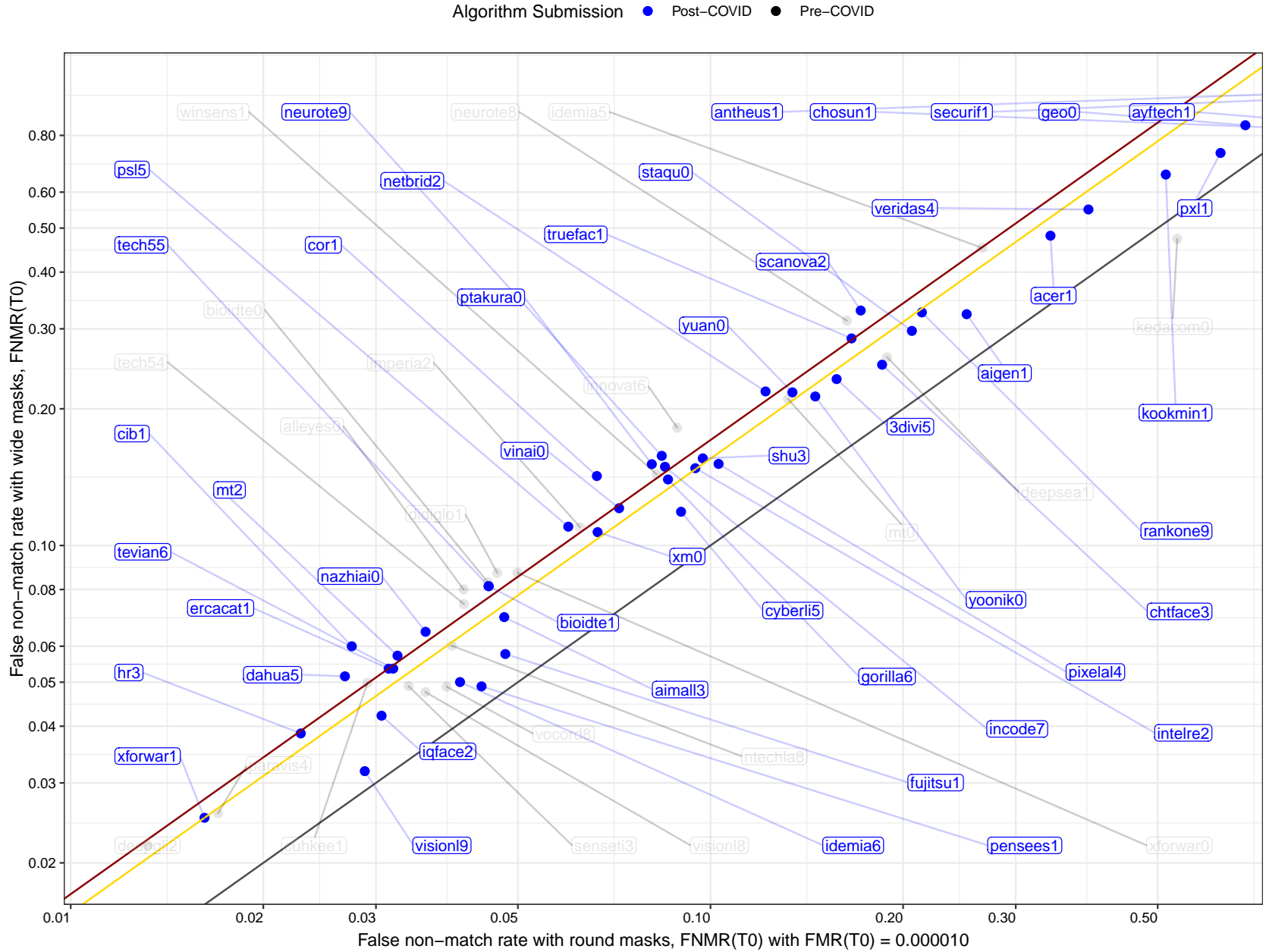


Figure 8: At a fixed threshold, a plot of FNMR with round versus wide masks. The displacement of the dark red (pre-COVID algorithms) and gold (post-COVID algorithms) lines relative to the black "parity" line shows a modest increase in FNMR with wide masks, with median post-COVID results showing nominal FNMR reductions. The value in the title is the median increase multiplier.

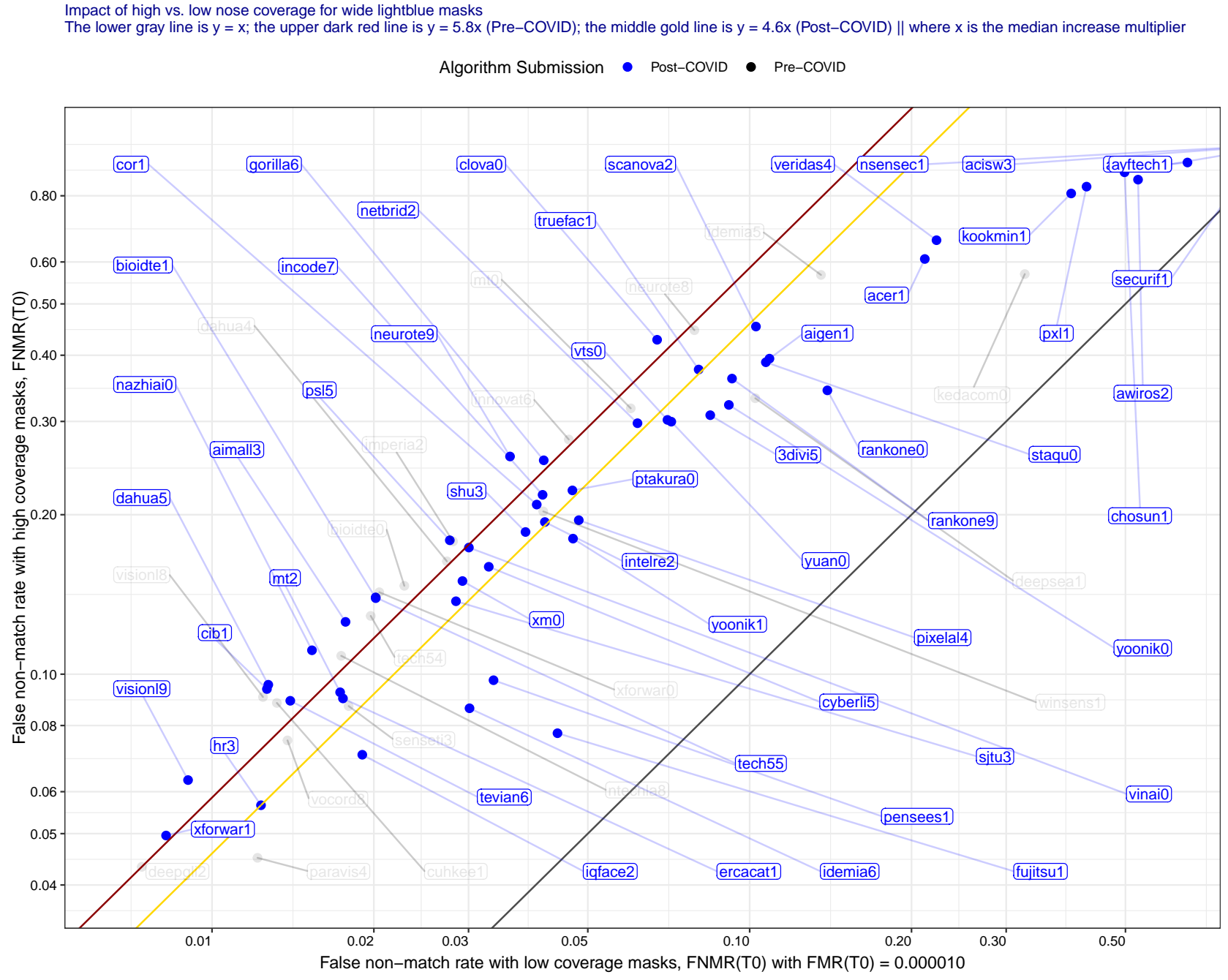


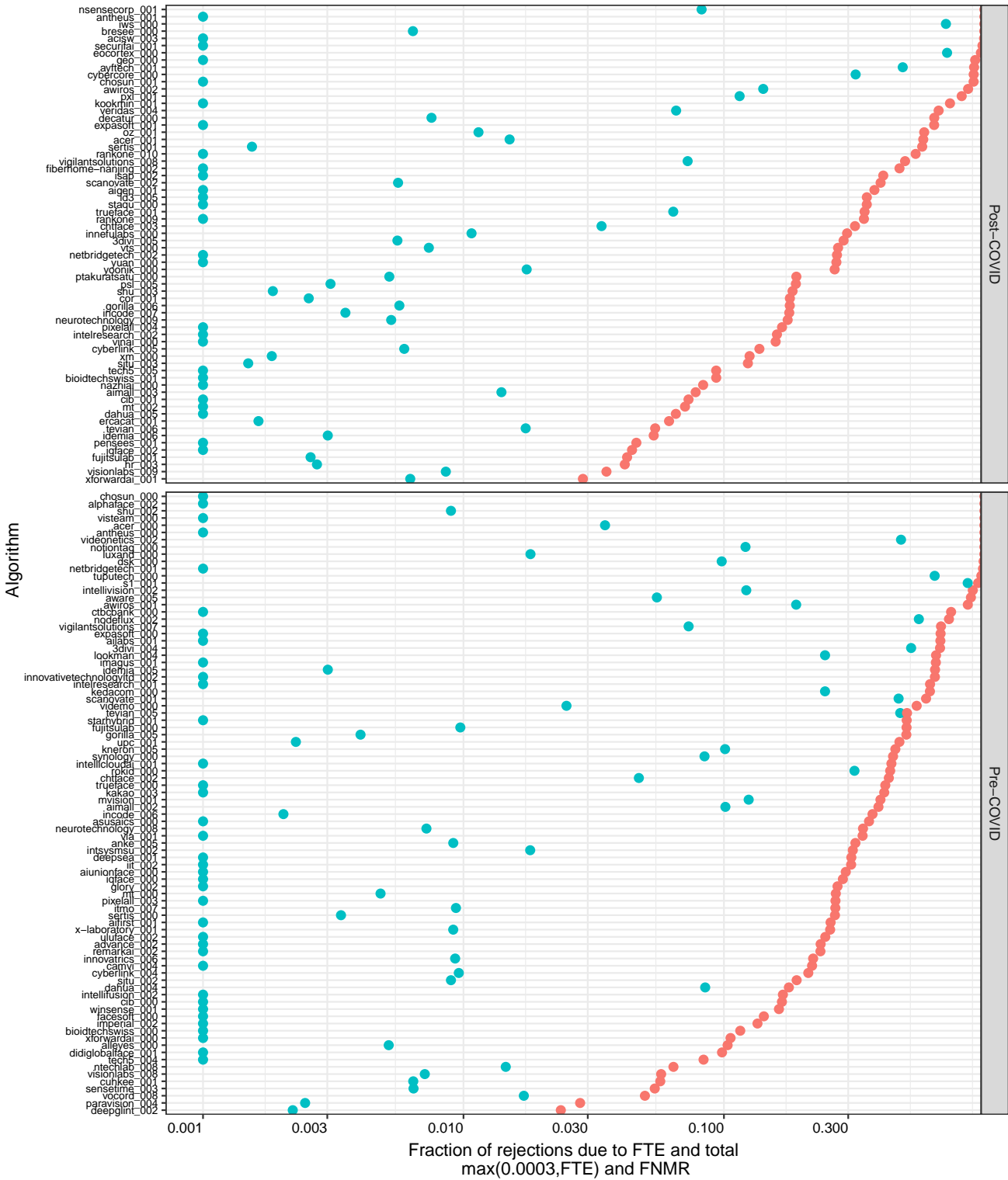
Figure 9: At a fixed threshold, a plot of FNMR with low versus high masks. The displacement of the dark red (pre-COVID algorithms) and gold (post-COVID algorithms) lines relative to the black "parity" lines shows a considerable increase in FNMR with high vs. low nose coverage masks, with median post-COVID results showing modest FNMR reductions. The value in the title is the median increase multiplier.

	Algorithm Name	COLOR = WHITE			COLOR = LIGHTBLUE						COLOR = RED			COLOR = BLACK		
		SHAPE = WIDE			SHAPE = WIDE			SHAPE = ROUND			SHAPE = WIDE			SHAPE = WIDE		
	COVERAGE	LO	MED	HI	LO	MED	HI	LO	MED	HI	LO	MED	HI	LO	MED	HI
136	tevian-006	0.012	0.089	0.160	0.010	0.066	0.124	0.028	0.041	0.146	0.034	0.266	0.408	0.016	0.256	0.307
137	trueface-000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
138	trueface-001	0.078	0.163	0.217	0.071	0.153	0.217	0.084	0.119	0.268	-	-	-	0.147	0.317	0.474
139	tuputech-000	0.517	0.679	0.684	0.626	0.758	0.765	0.502	0.619	0.714	0.652	0.893	0.926	0.661	0.904	0.933
140	uluface-002	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
141	upc-001	0.002	0.005	0.012	0.002	0.005	0.012	0.002	0.002	0.005	0.002	0.010	0.027	0.003	0.007	0.018
142	veridas-003	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000	1.000
143	veridas-004	0.047	0.139	0.202	0.041	0.112	0.164	0.048	0.085	0.154	0.106	0.401	0.568	0.054	0.213	0.383
144	via-001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
145	videmo-000	0.019	0.067	0.125	0.018	0.051	0.106	0.023	0.040	0.089	0.122	0.352	0.368	0.027	0.100	0.296
146	videonetics-002	0.338	0.581	0.557	0.330	0.569	0.542	0.378	0.559	0.785	0.344	0.693	0.735	0.396	0.702	0.848
147	vigilantsolutions-007	0.062	0.168	0.220	0.052	0.137	0.193	0.069	0.126	0.206	0.098	0.408	0.634	0.072	0.273	0.493
148	vigilantsolutions-008	0.062	0.168	0.218	0.051	0.136	0.189	0.068	0.124	0.202	0.096	0.407	0.633	0.072	0.268	0.492
149	vinai-000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
150	visionlabs-008	0.013	0.035	0.083	0.012	0.031	0.072	0.019	0.038	0.097	0.016	0.040	0.096	0.024	0.061	0.124
151	visionlabs-009	0.011	0.037	0.082	0.010	0.035	0.076	0.018	0.042	0.097	0.016	0.052	0.120	0.024	0.065	0.150
152	visteam-000	0.058	0.150	0.210	0.048	0.118	0.176	0.052	0.092	0.156	0.160	0.372	0.506	0.074	0.202	0.369
153	vocord-008	0.013	0.046	0.087	0.011	0.052	0.089	0.031	0.059	0.111	0.013	0.079	0.127	0.009	0.050	0.093
154	vtS-000	0.011	0.029	0.072	0.012	0.028	0.074	0.015	0.020	0.048	0.014	0.043	0.129	0.016	0.051	0.149
155	winsense-001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
156	xforwardai-000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
157	xforwardai-001	0.006	0.025	0.069	0.007	0.031	0.079	0.011	0.021	0.052	0.007	0.030	0.079	0.007	0.022	0.060
158	xm-000	0.003	0.006	0.013	0.002	0.004	0.008	0.003	0.004	0.009	0.008	0.031	0.070	0.005	0.013	0.038
159	yoonik-000	0.029	0.061	0.138	0.027	0.060	0.142	0.037	0.049	0.124	0.040	0.094	0.219	0.039	0.099	0.257
160	yoonik-001	0.015	0.052	0.110	0.017	0.055	0.113	0.024	0.035	0.072	0.031	0.093	0.170	0.026	0.086	0.175
161	yuan-000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Table 12: This table summarizes Failure to Enroll (FTE) rates surveyed over 10 000 images of each mask variant. FTE is the proportion of failed template generation attempts. Failures can occur because the software throws an exception, or because the software electively refuses to process the input image as would occur if the algorithms does not detect a face or determines that the face has insufficient information. FTE is measured as the number of function calls that give EITHER a non-zero error code OR that give a “small” template containing fewer than 60 bytes. This second rule is needed because some algorithms incorrectly fail to return a non-zero error code when template generation fails but do produce a skeletal template. The effects of FTE are included in the accuracy results of this report by regarding any template comparison involving a failed template to produce a low similarity score. Thus higher FTE results in higher FNMR and lower FMR. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

Failure-to-template contribution toward total false rejection for medium wide lightblue masks

Kind ● FNMR ● FTE



This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

Figure 10: For each algorithm the rightmost dot shows FNMR @ FMR=0.0001 (as reported throughout this report). The left most dot shows the failure-to-template (FTE) rate over the masked verification set of 5.2M images. The gap between the two dots is attributable to low similarity score. Some FTE rates are zero - rates below 0.001 are shown as 0.001.

The following plots are detection error tradeoff (DET) characteristics for each algorithm, across different mask nose coverages and shapes.

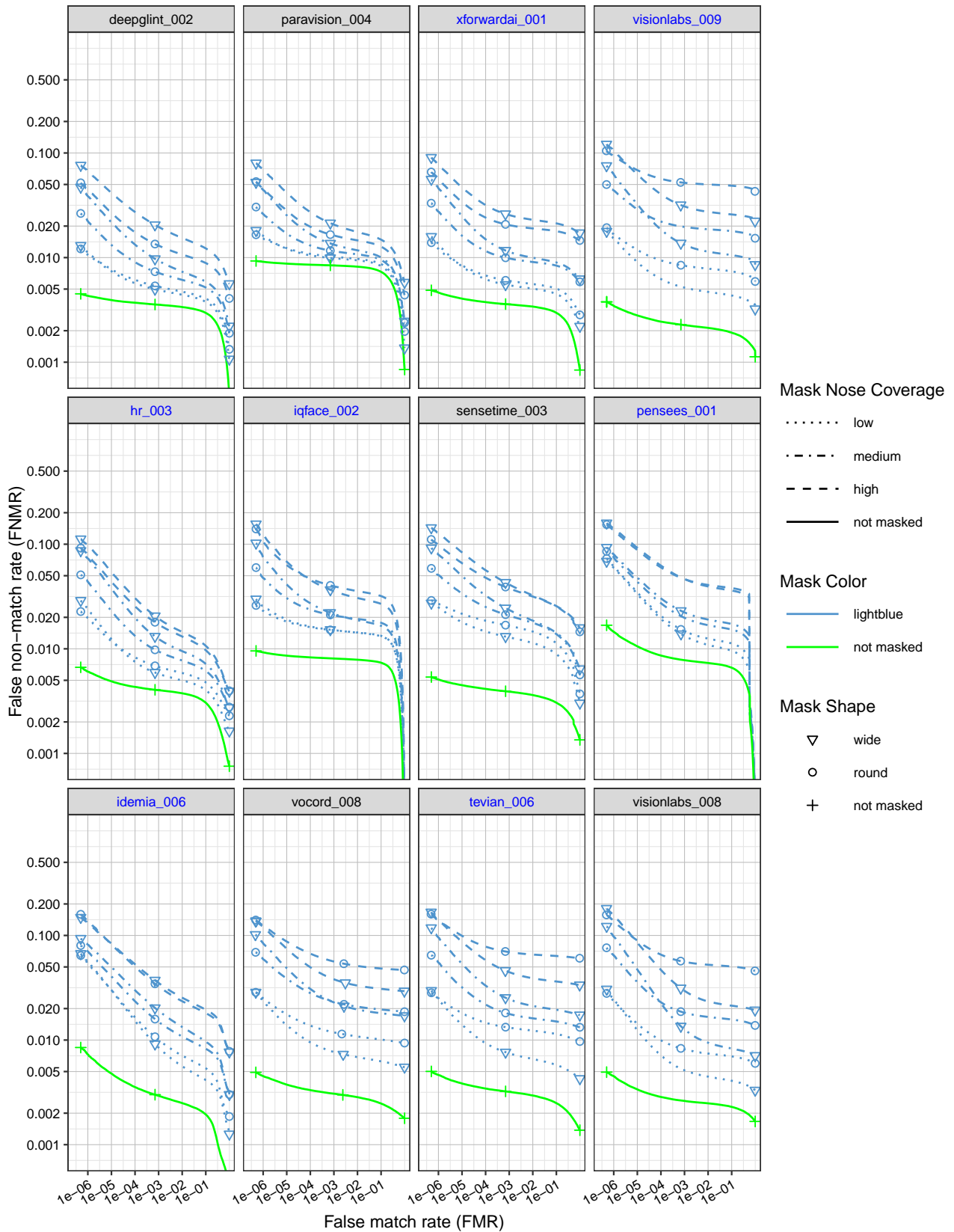


Figure 11: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

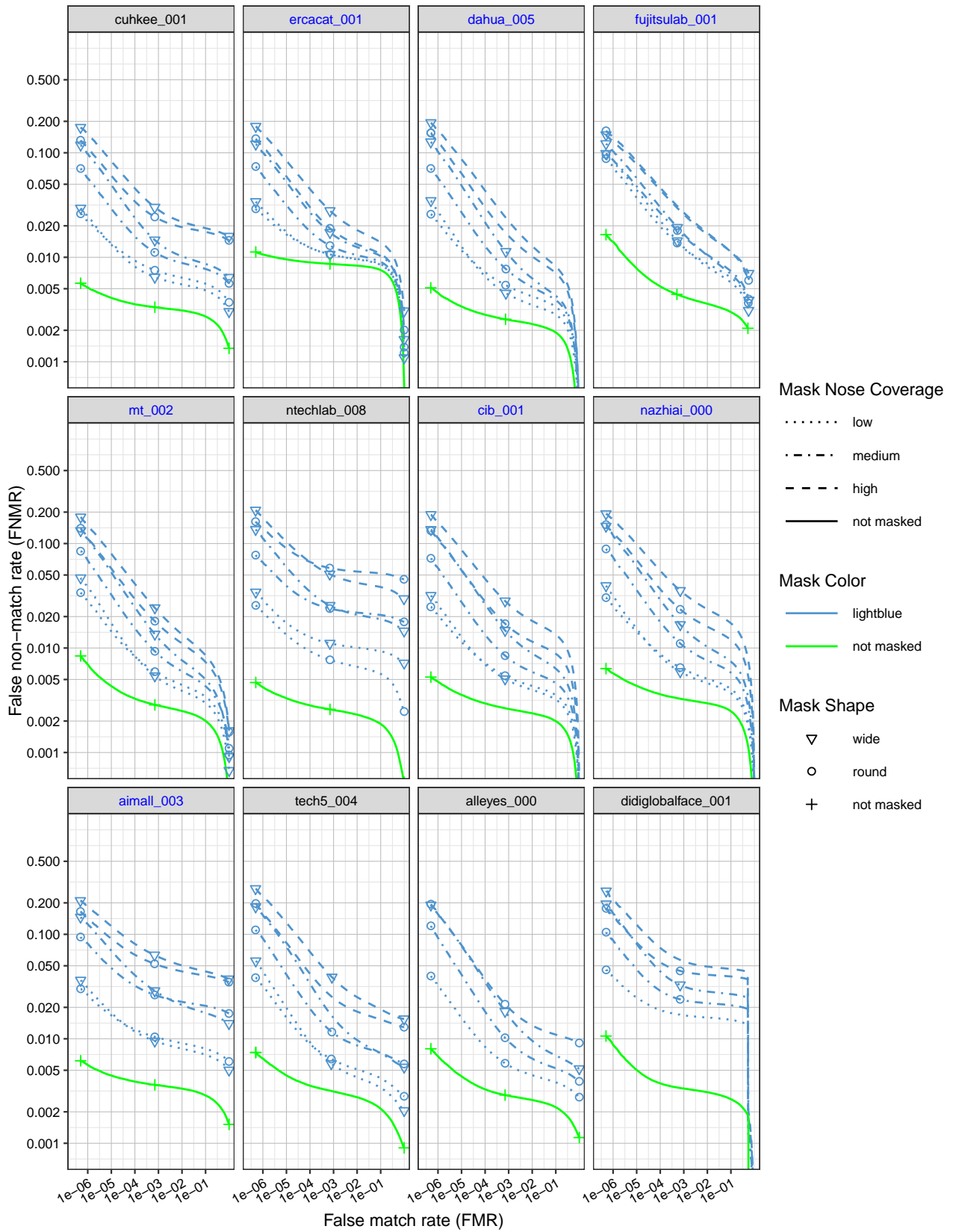


Figure 12: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

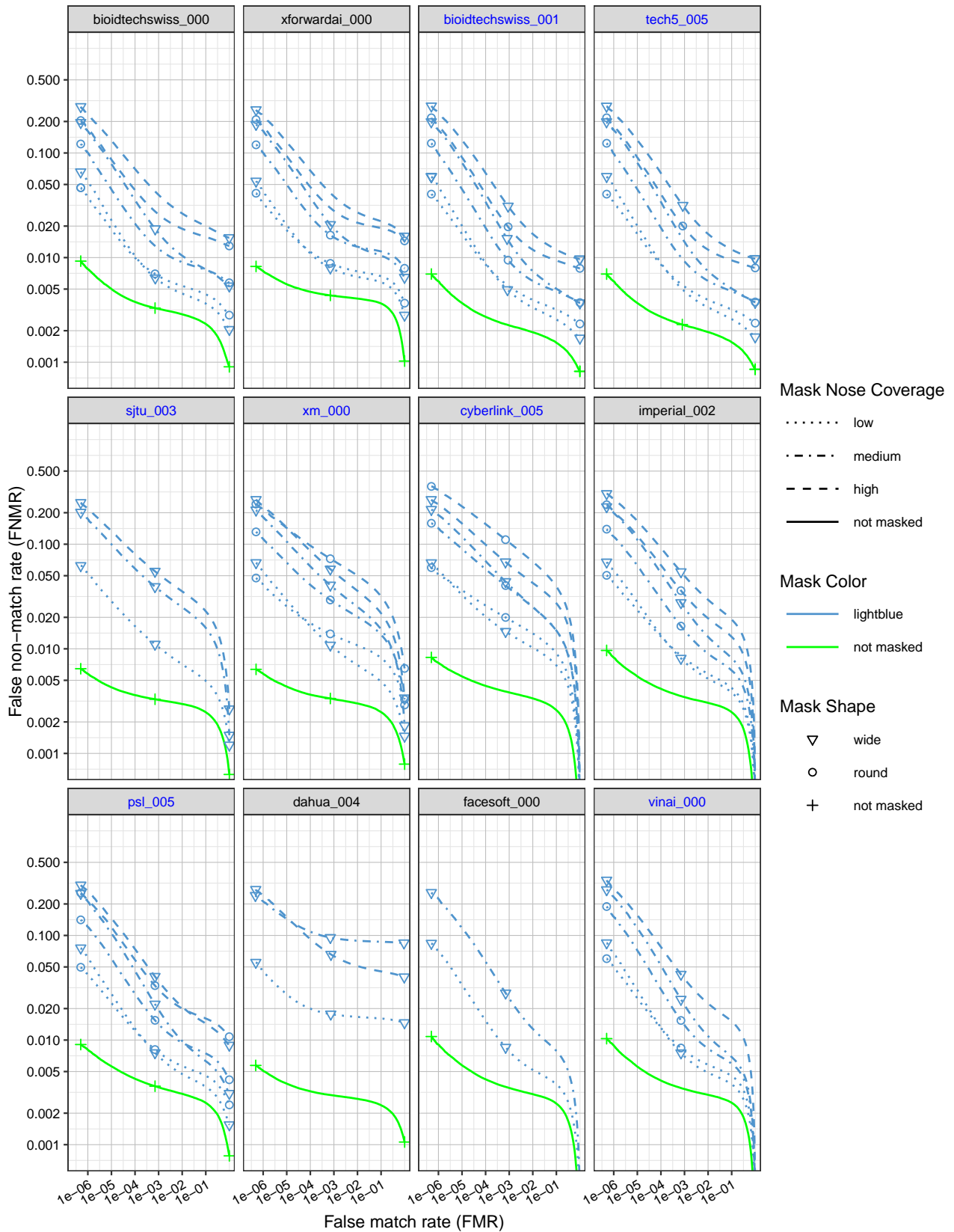


Figure 13: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

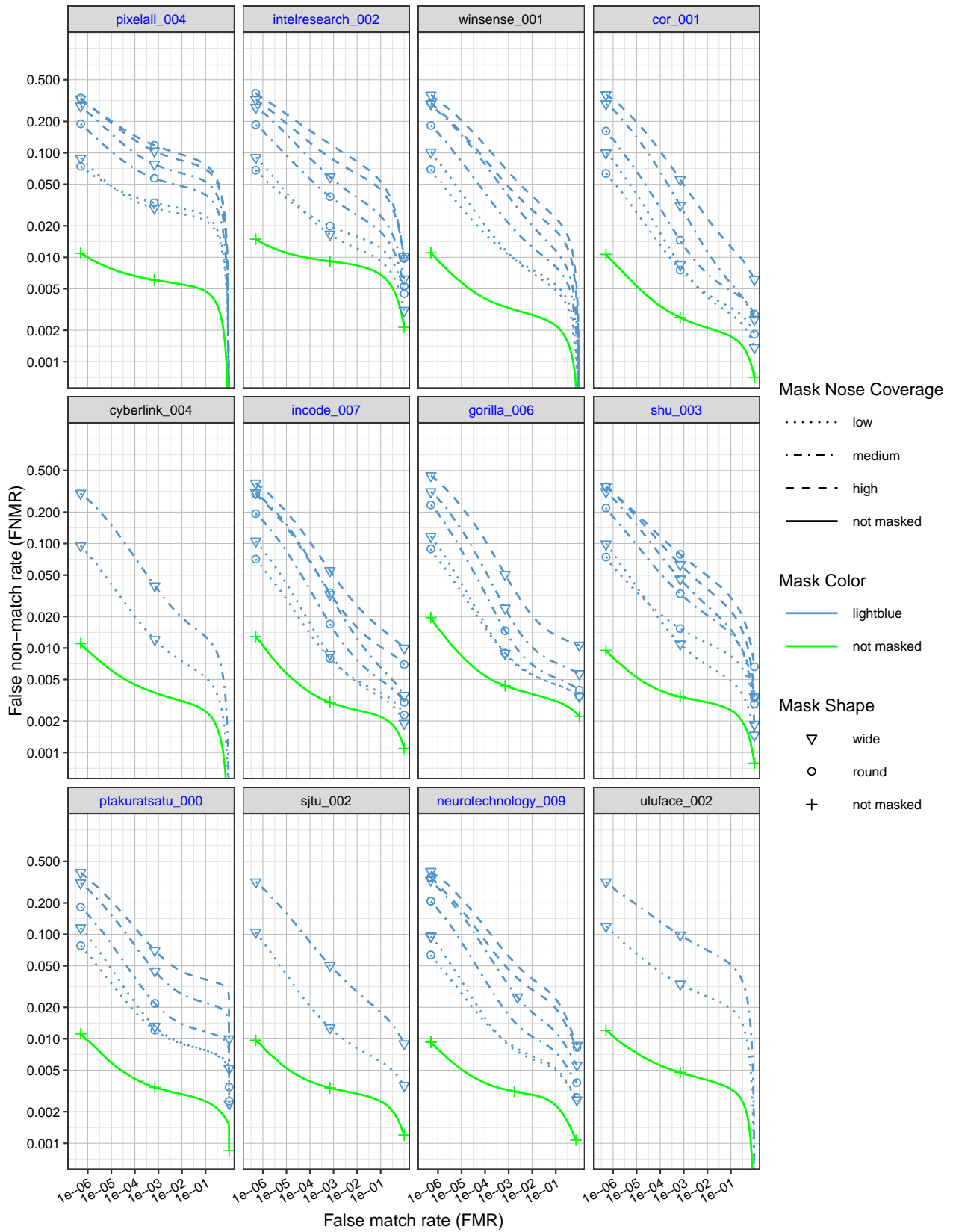


Figure 14: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

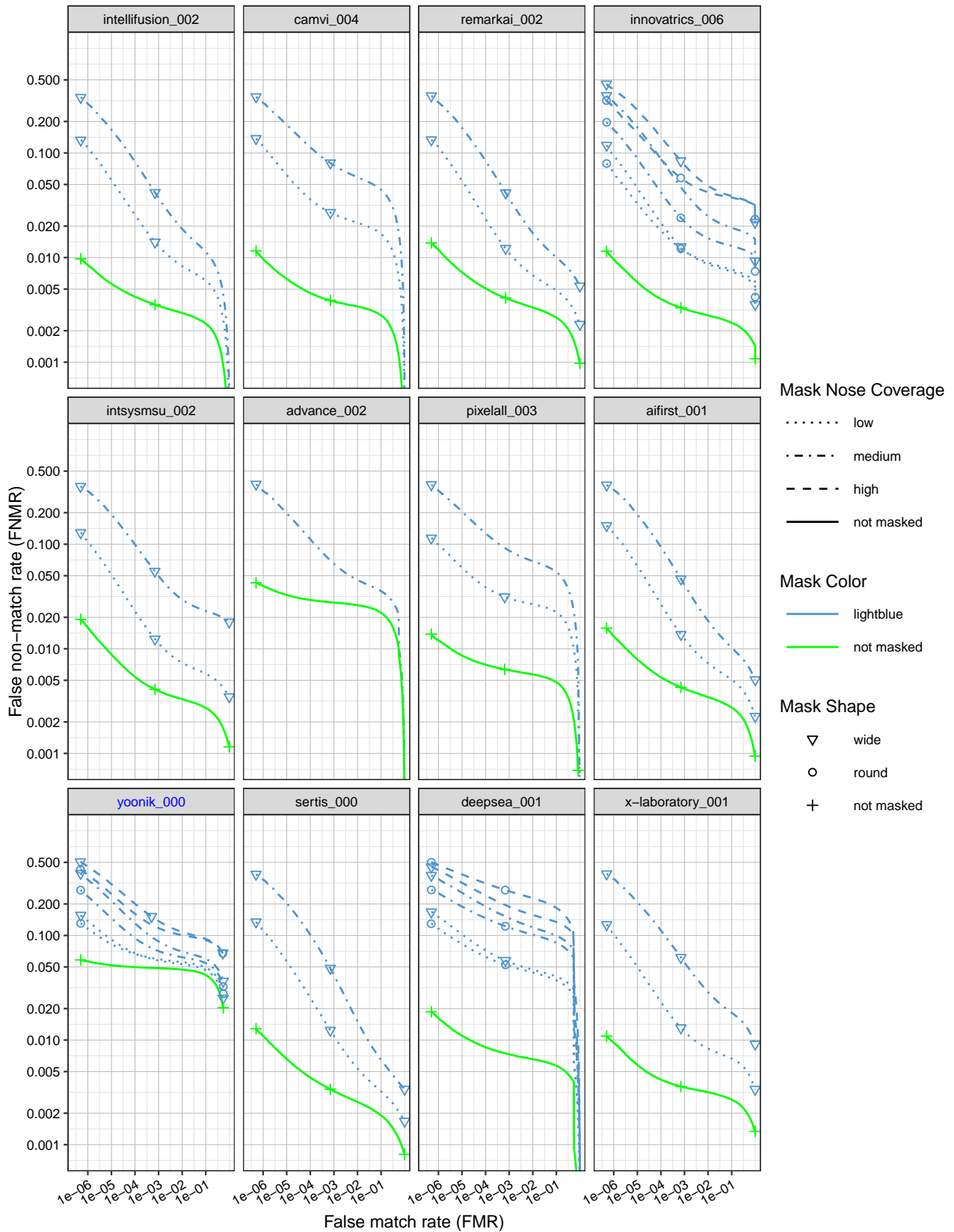


Figure 15: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

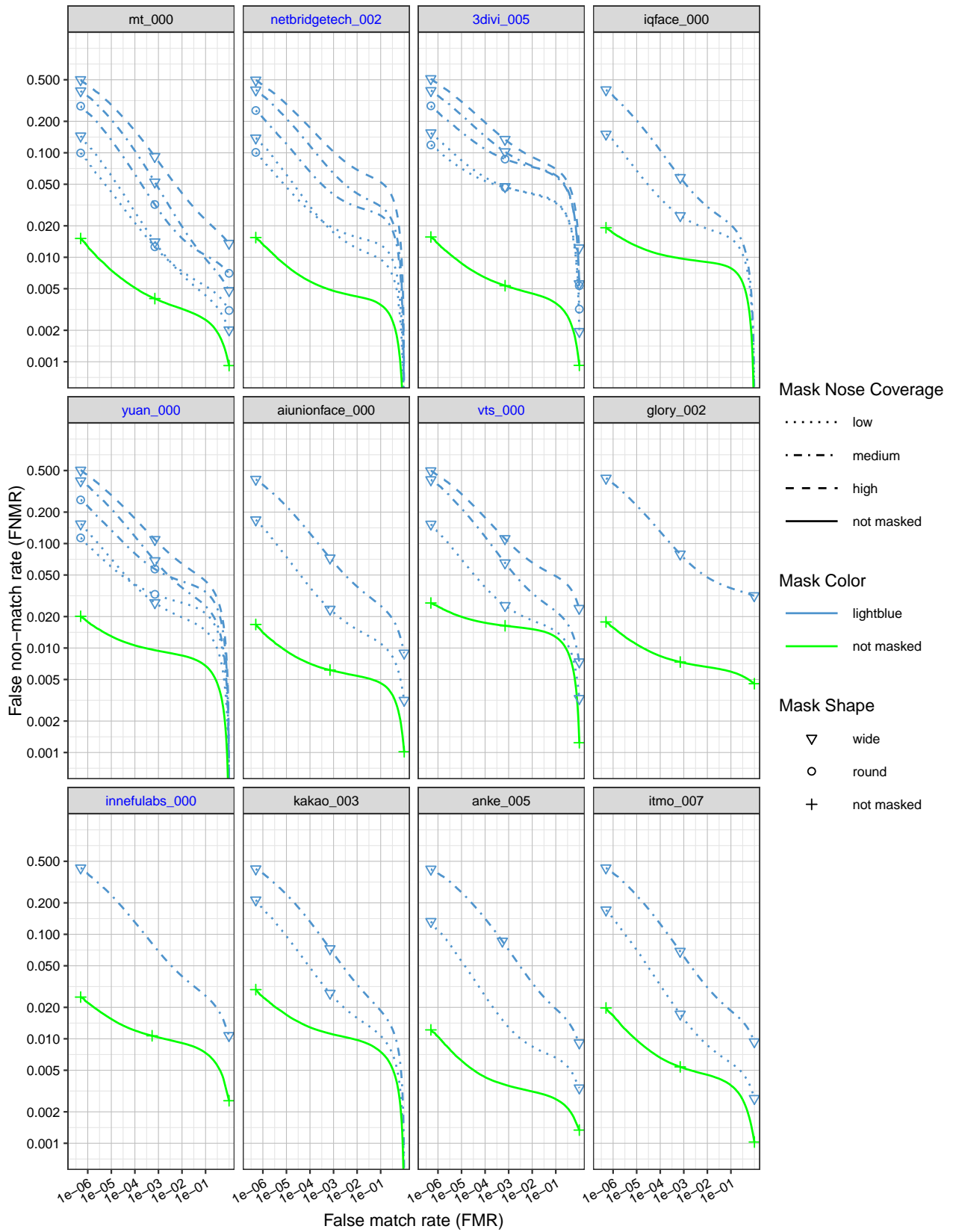


Figure 16: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

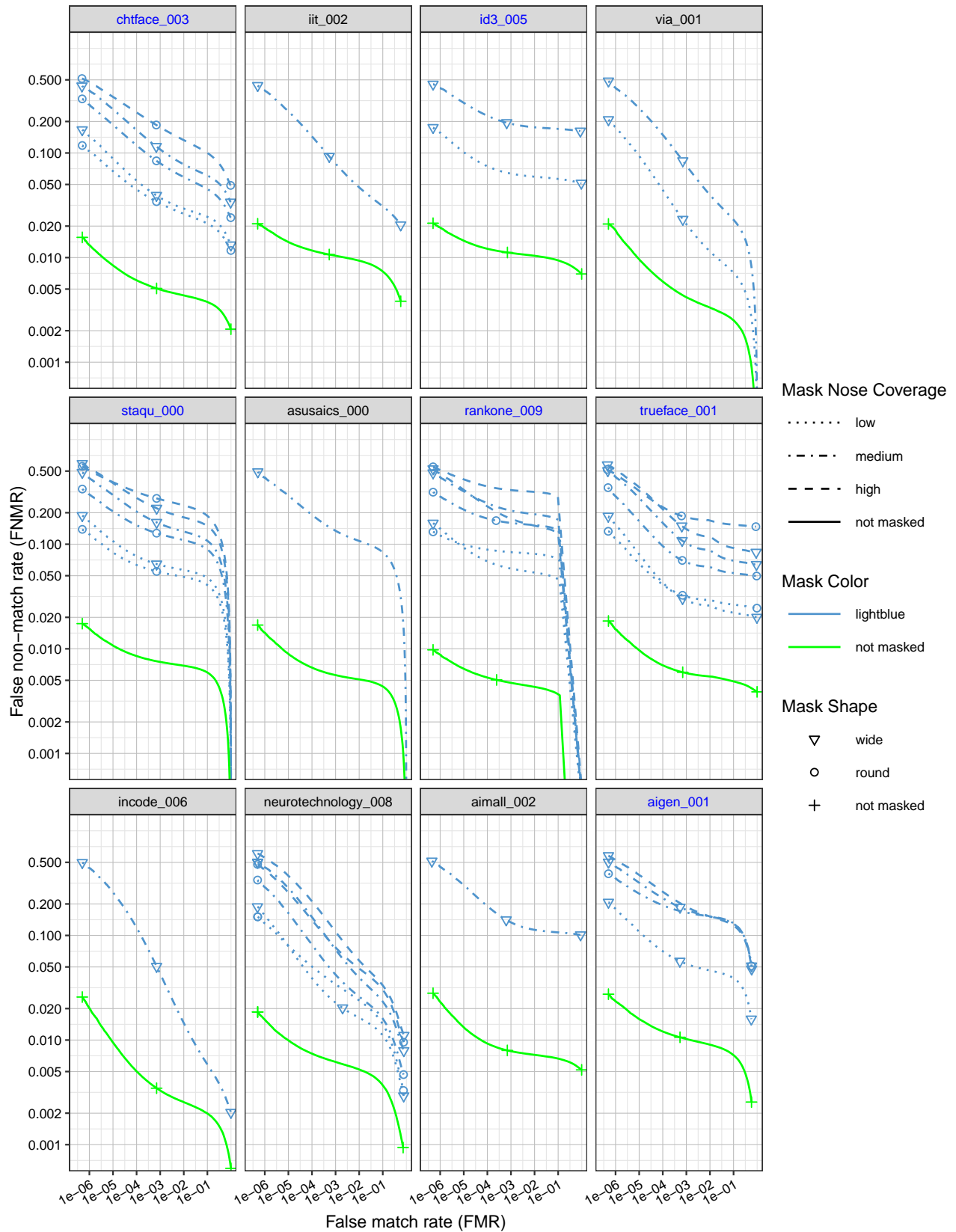


Figure 17: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

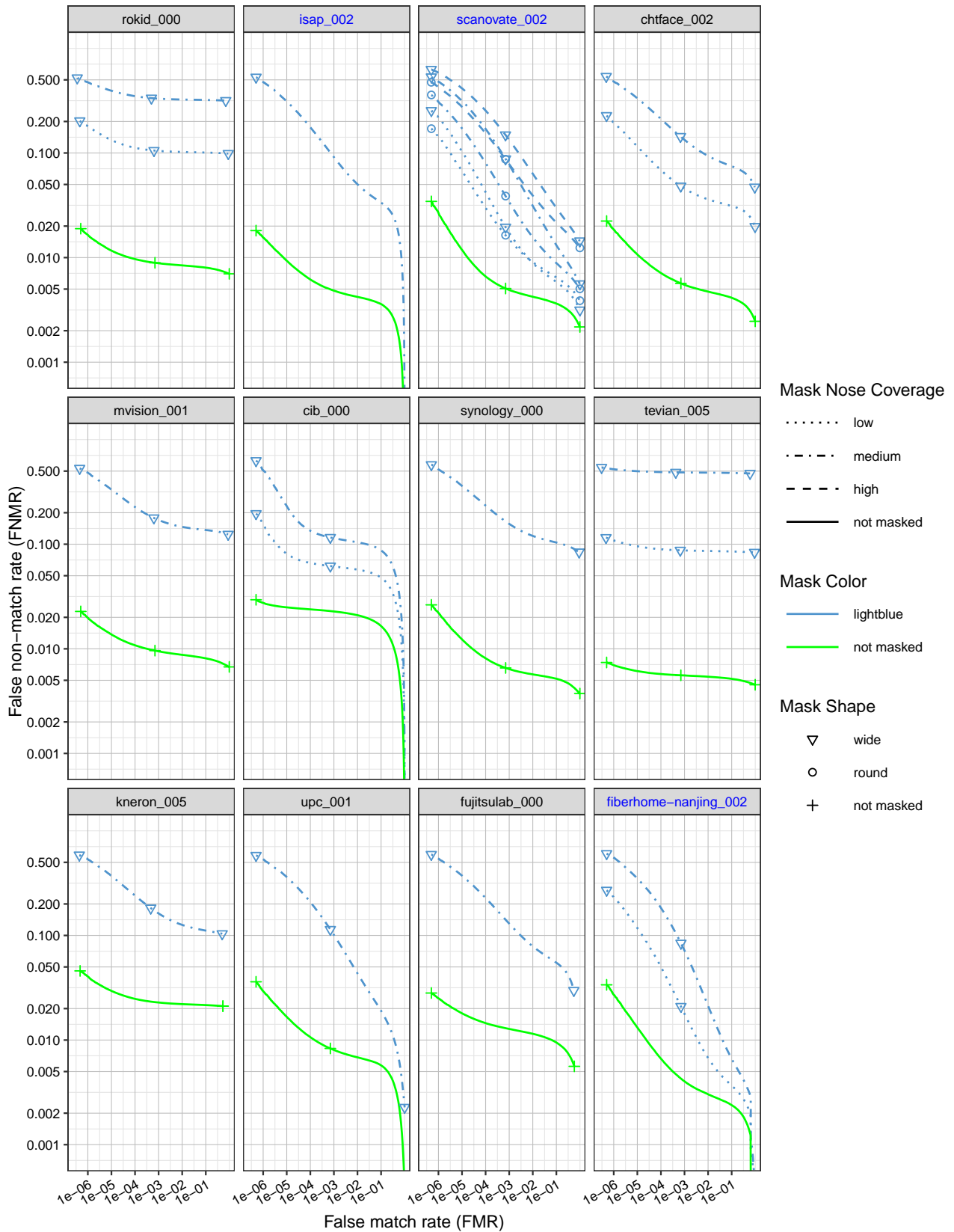


Figure 18: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

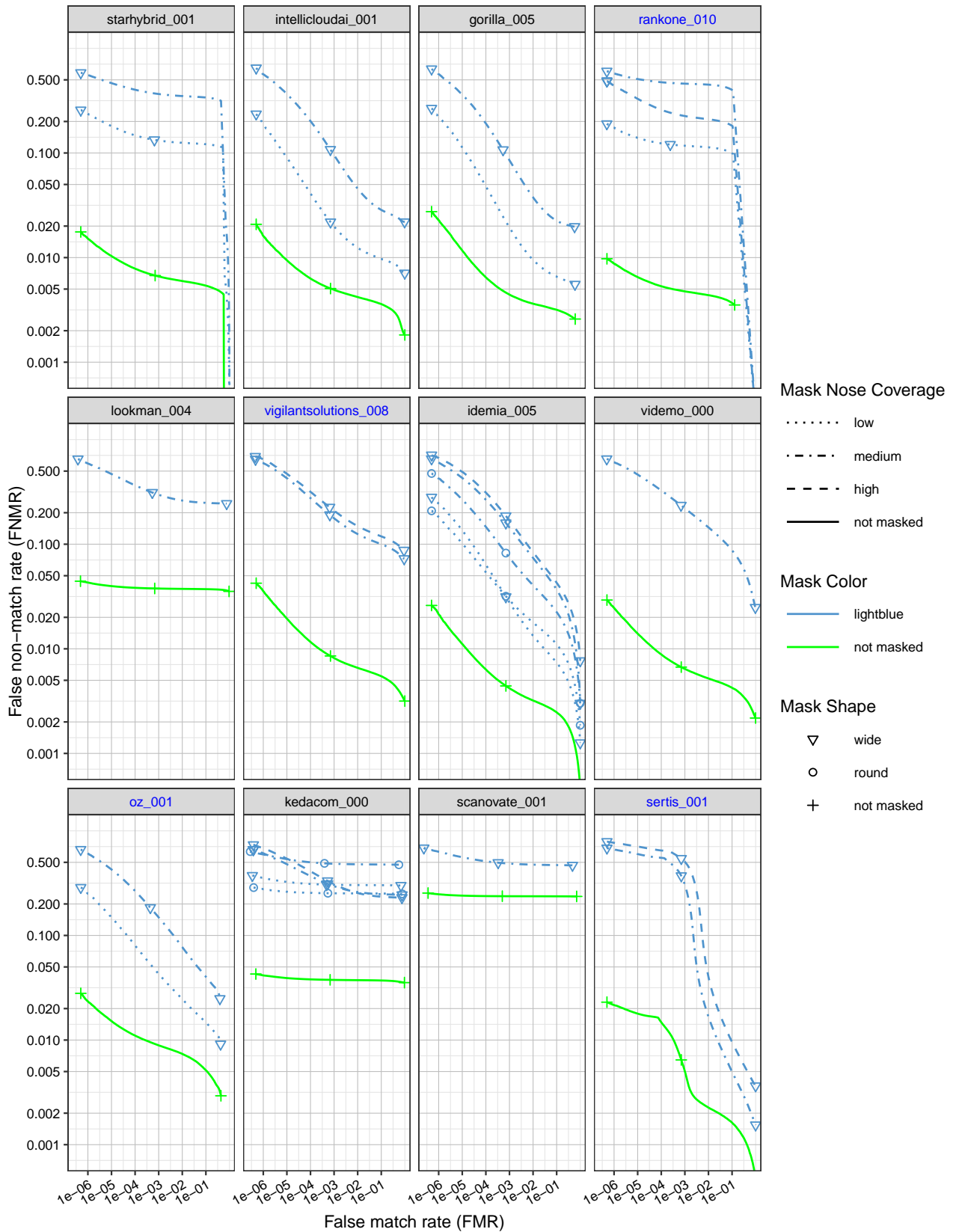


Figure 19: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

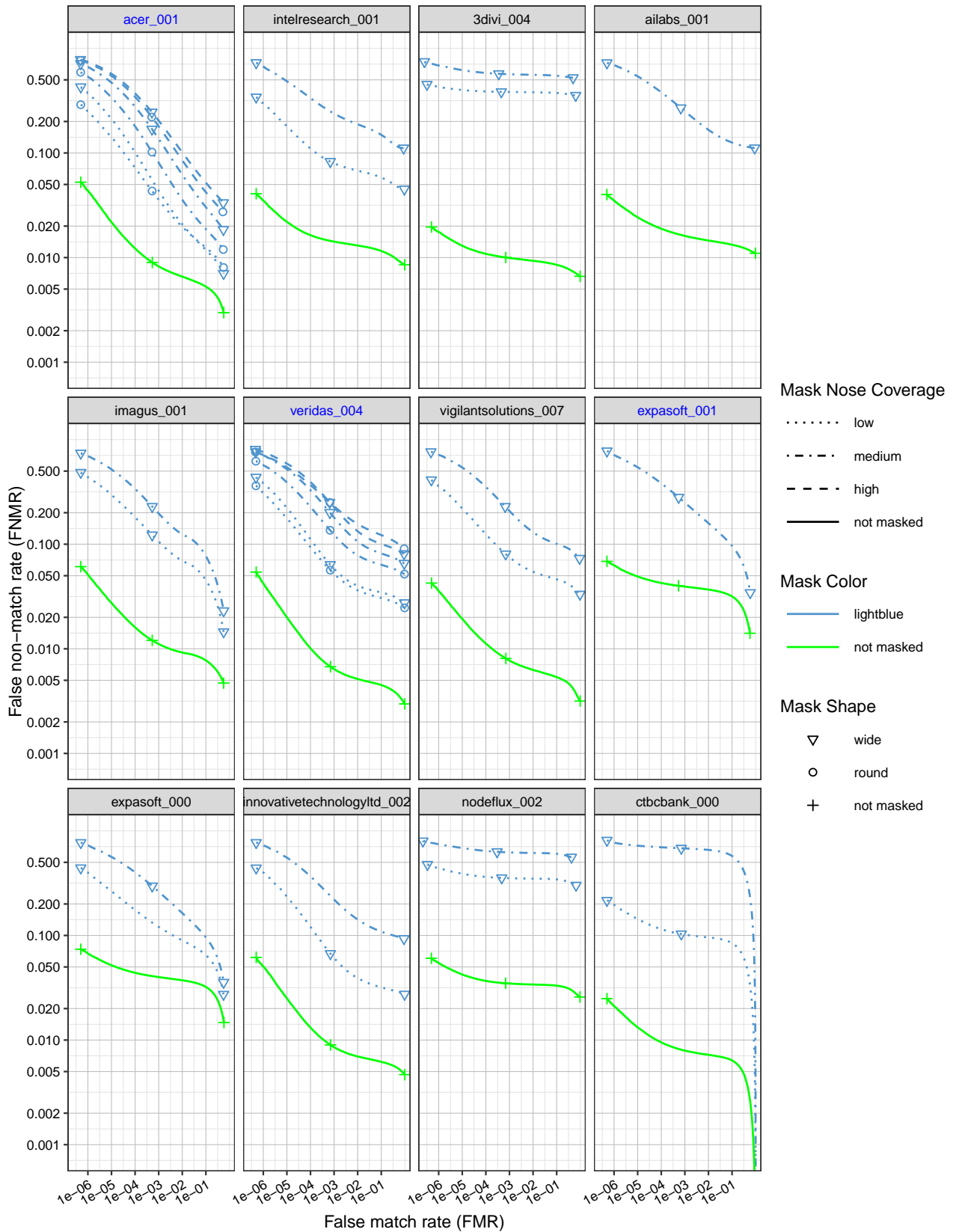


Figure 20: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

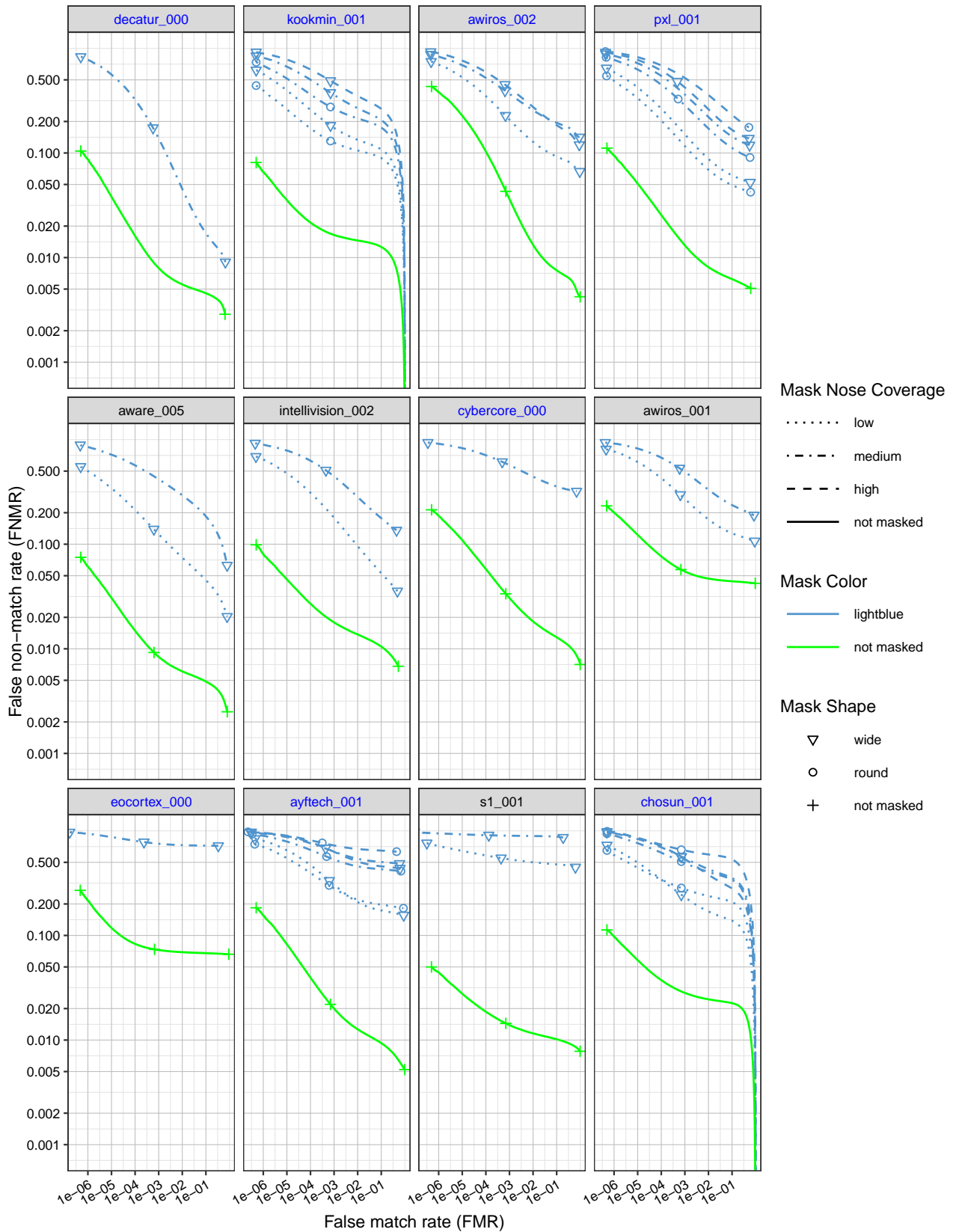


Figure 21: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

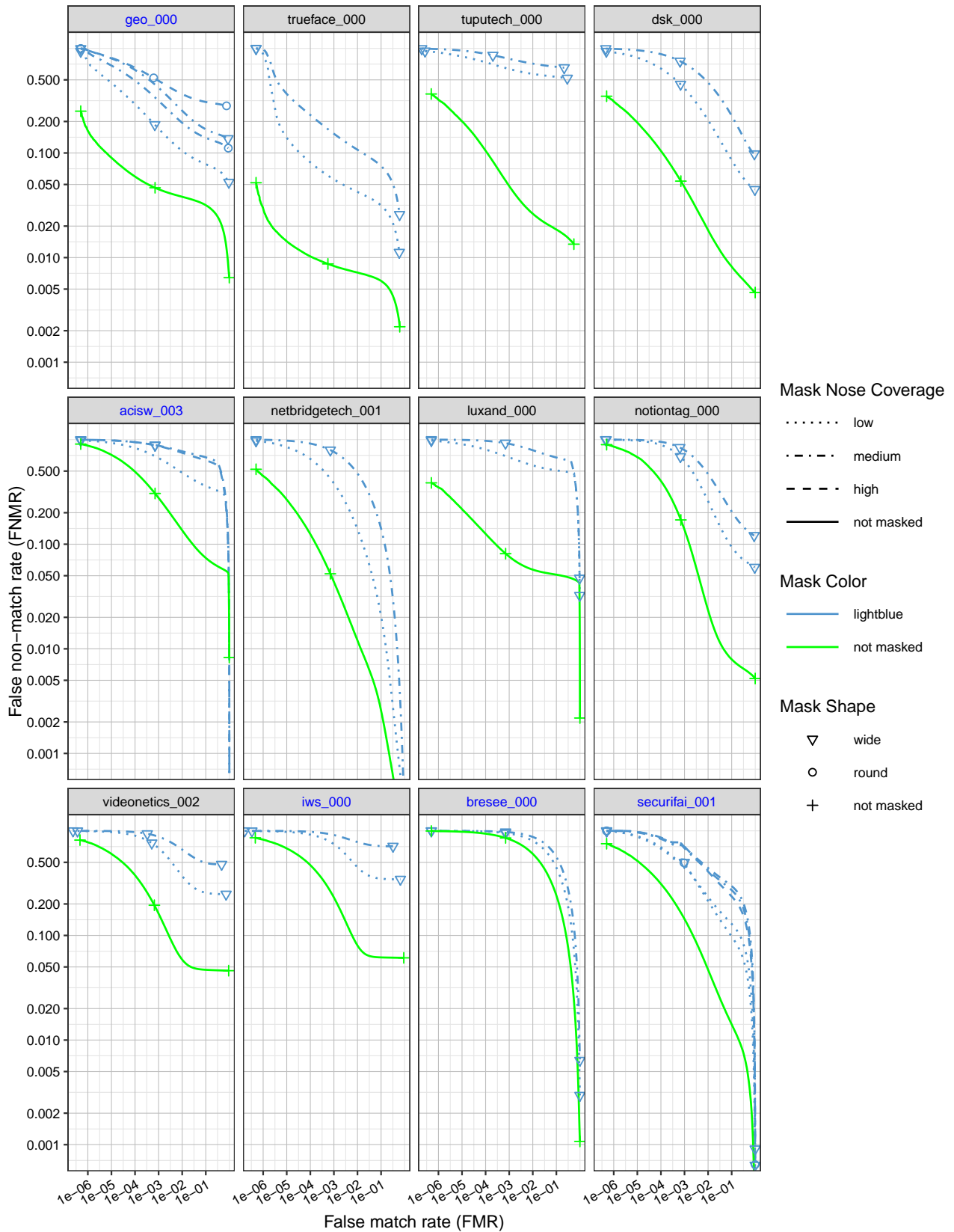
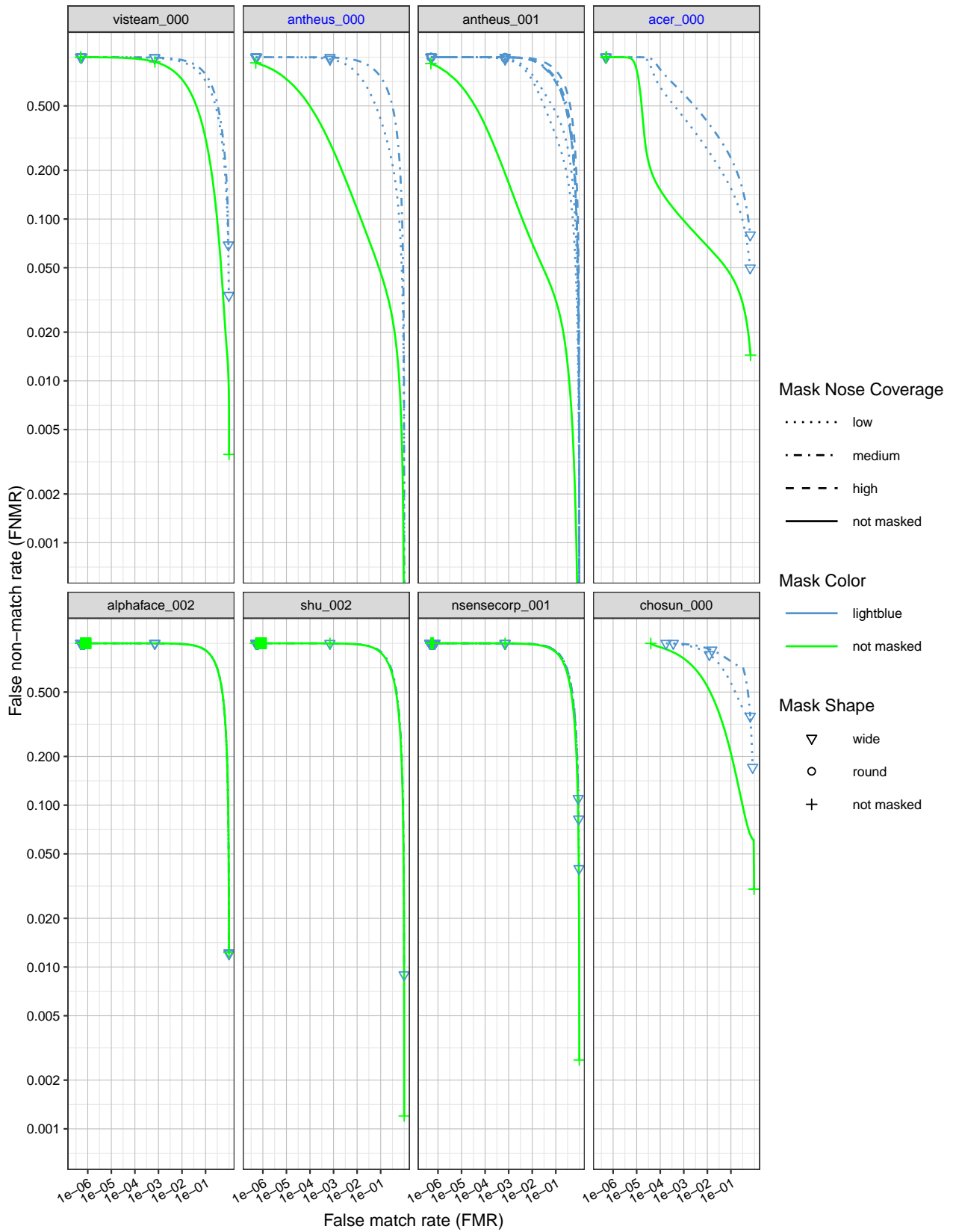


Figure 22: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>



This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

Figure 23: DET curves showing error rates on unmasked and masked probe images, broken out by mask shape and nose coverage. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

The following plots are detection error tradeoff (DET) characteristics for each algorithm, across different mask colors.

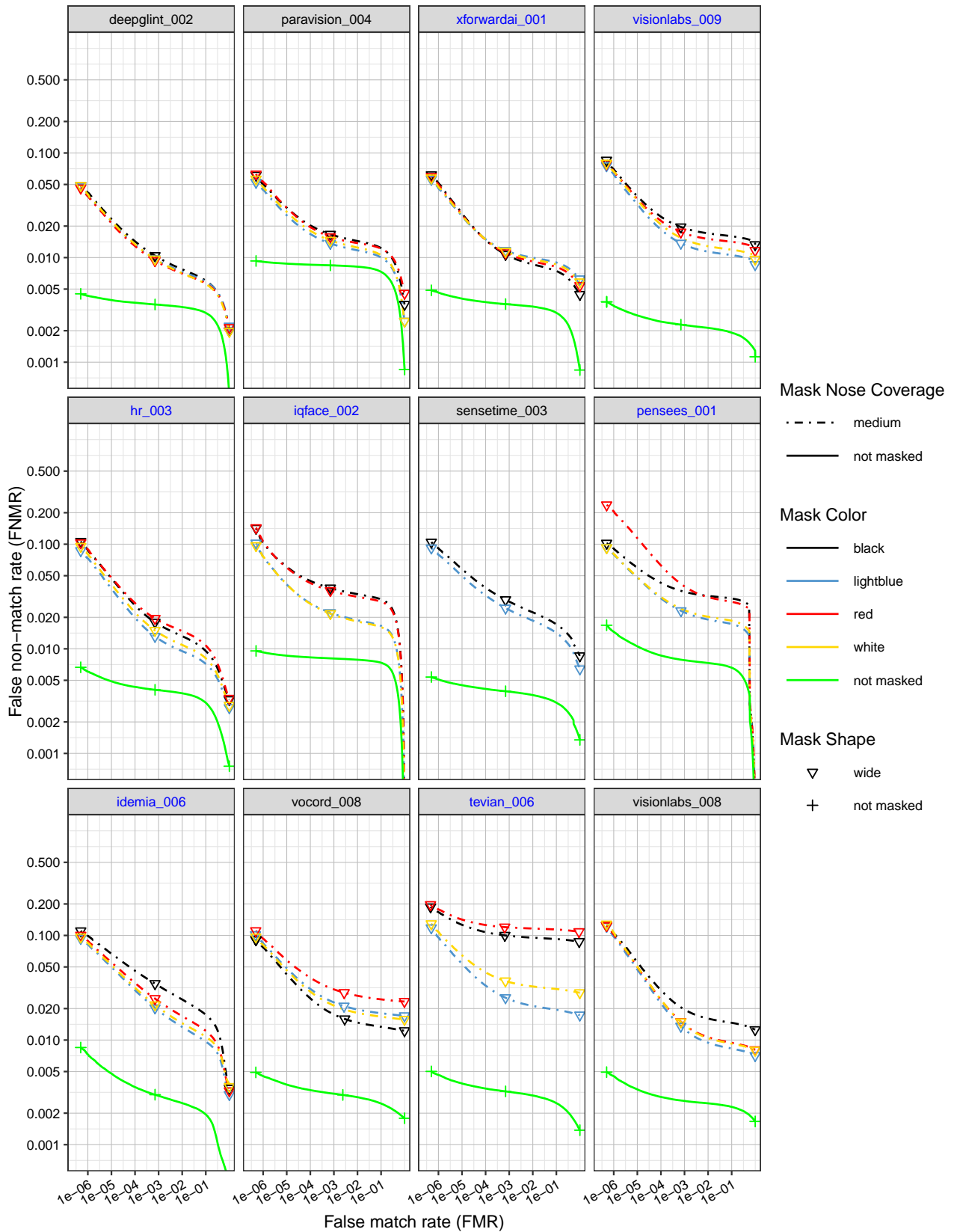


Figure 24: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

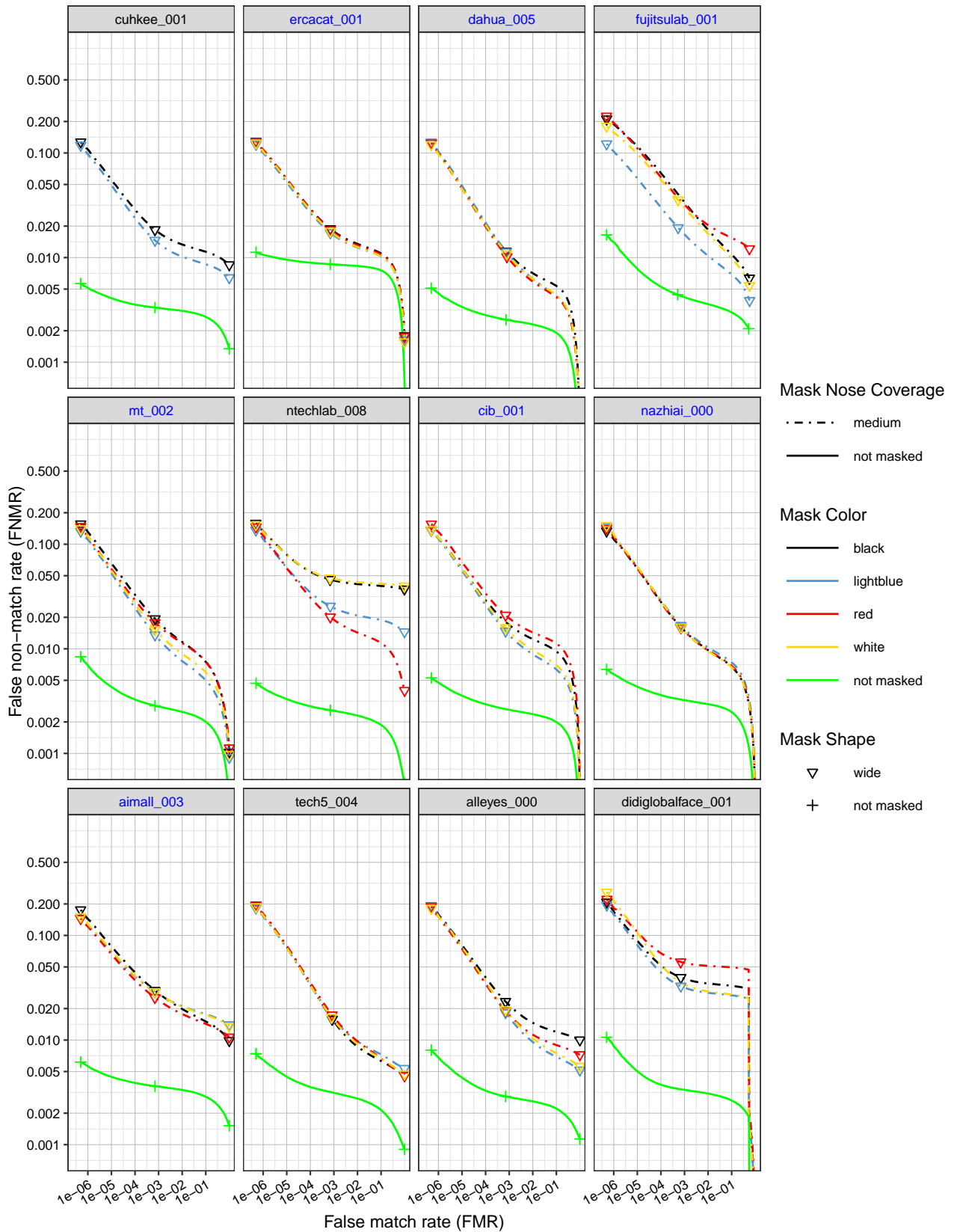


Figure 25: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

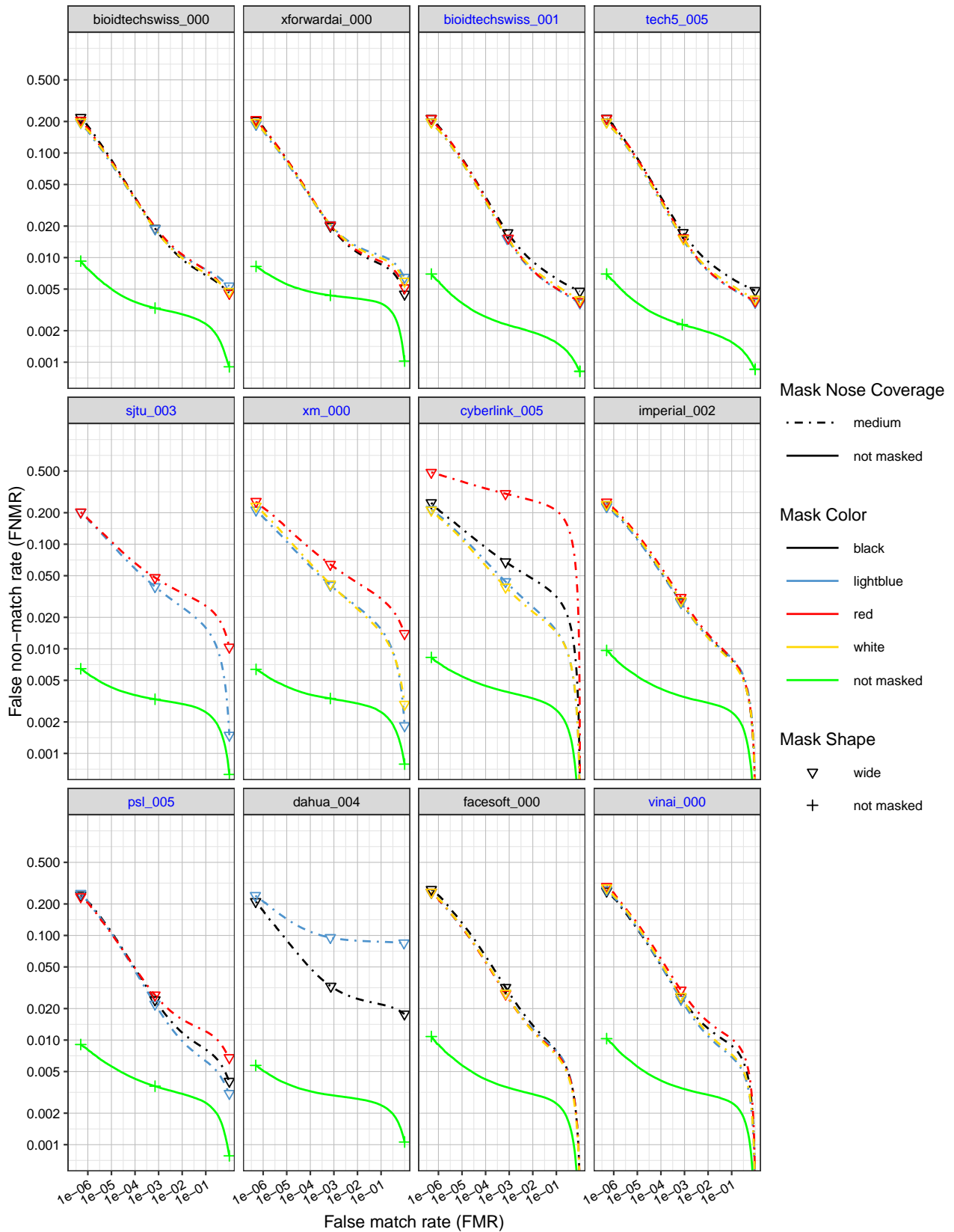


Figure 26: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

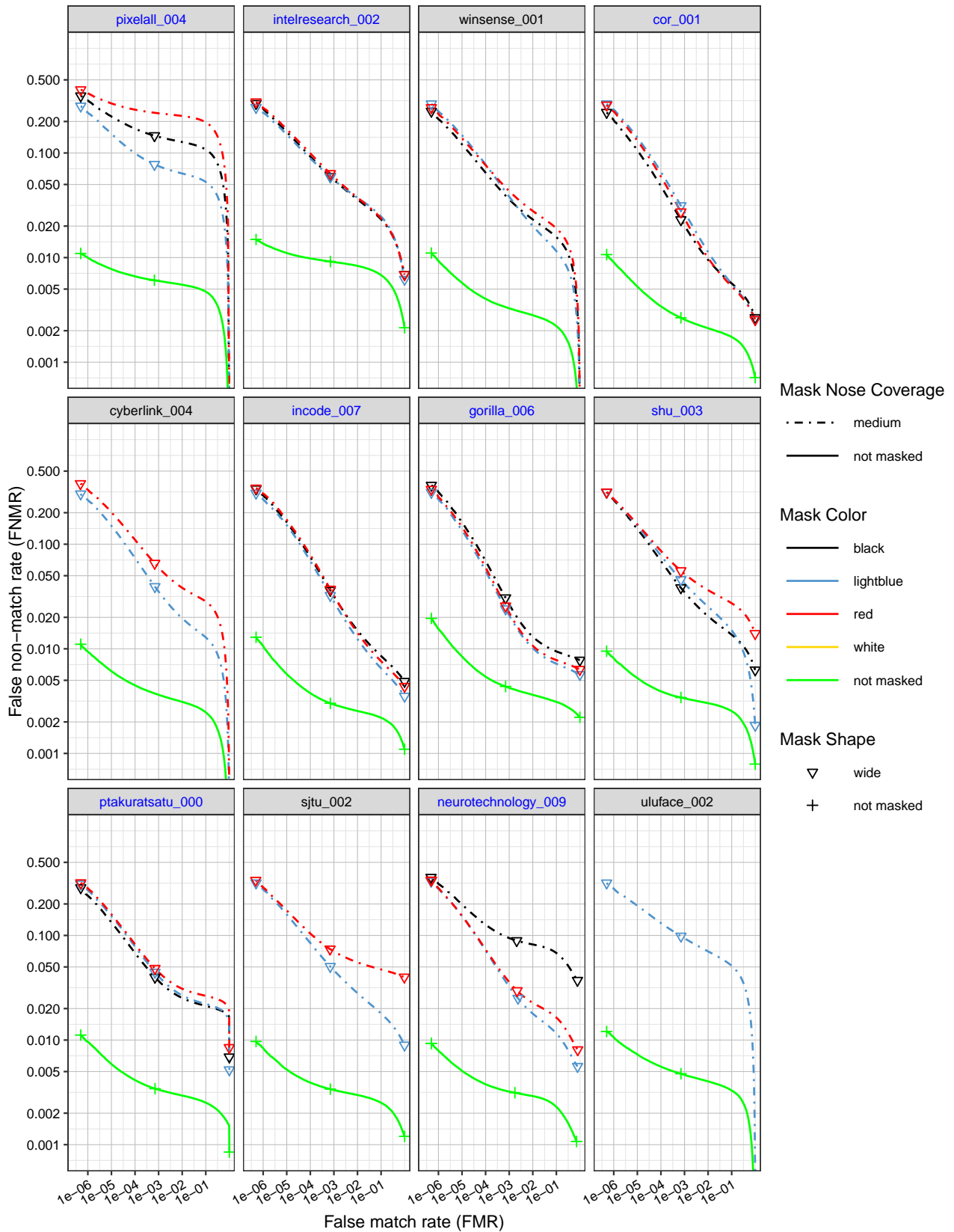


Figure 27: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

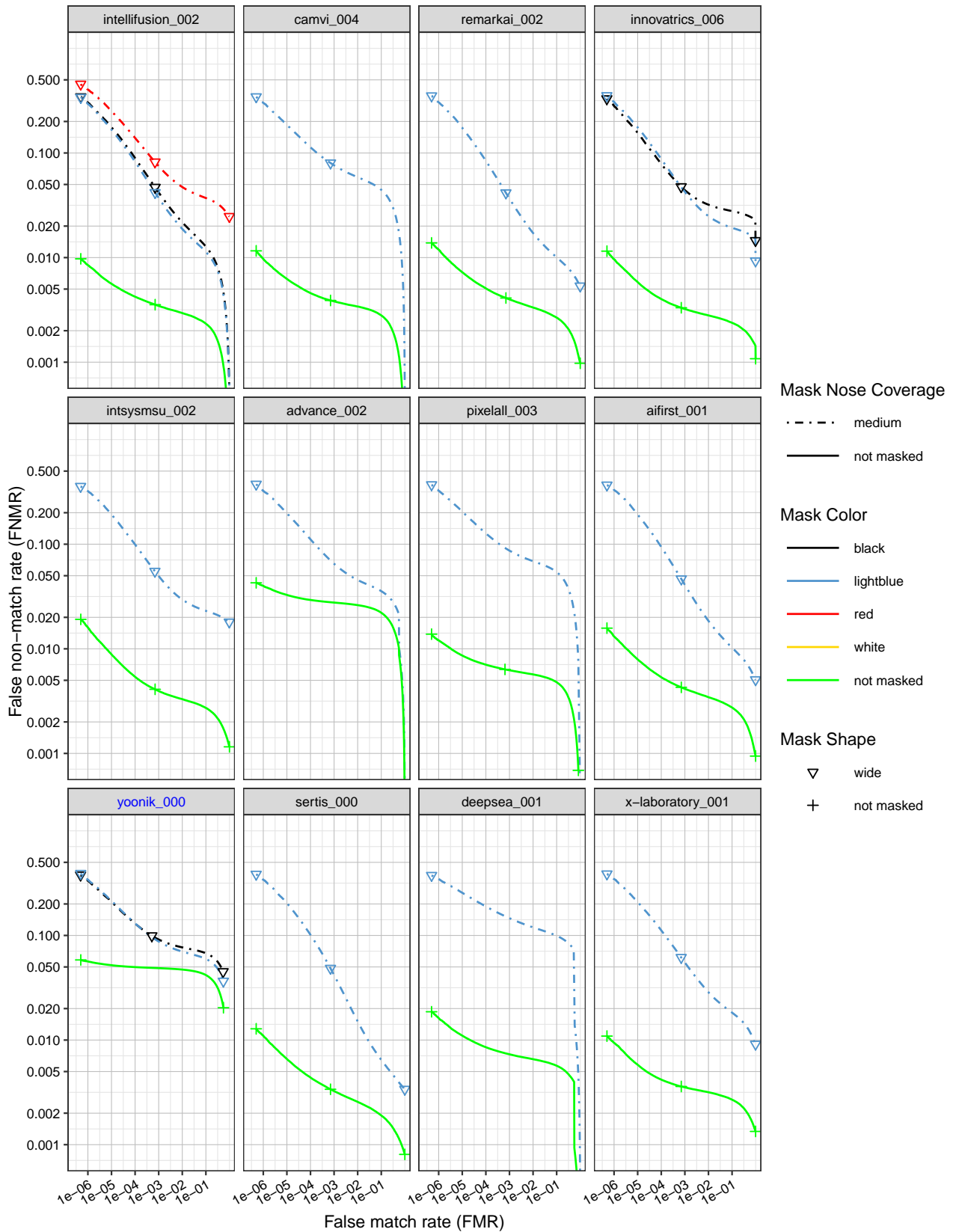


Figure 28: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

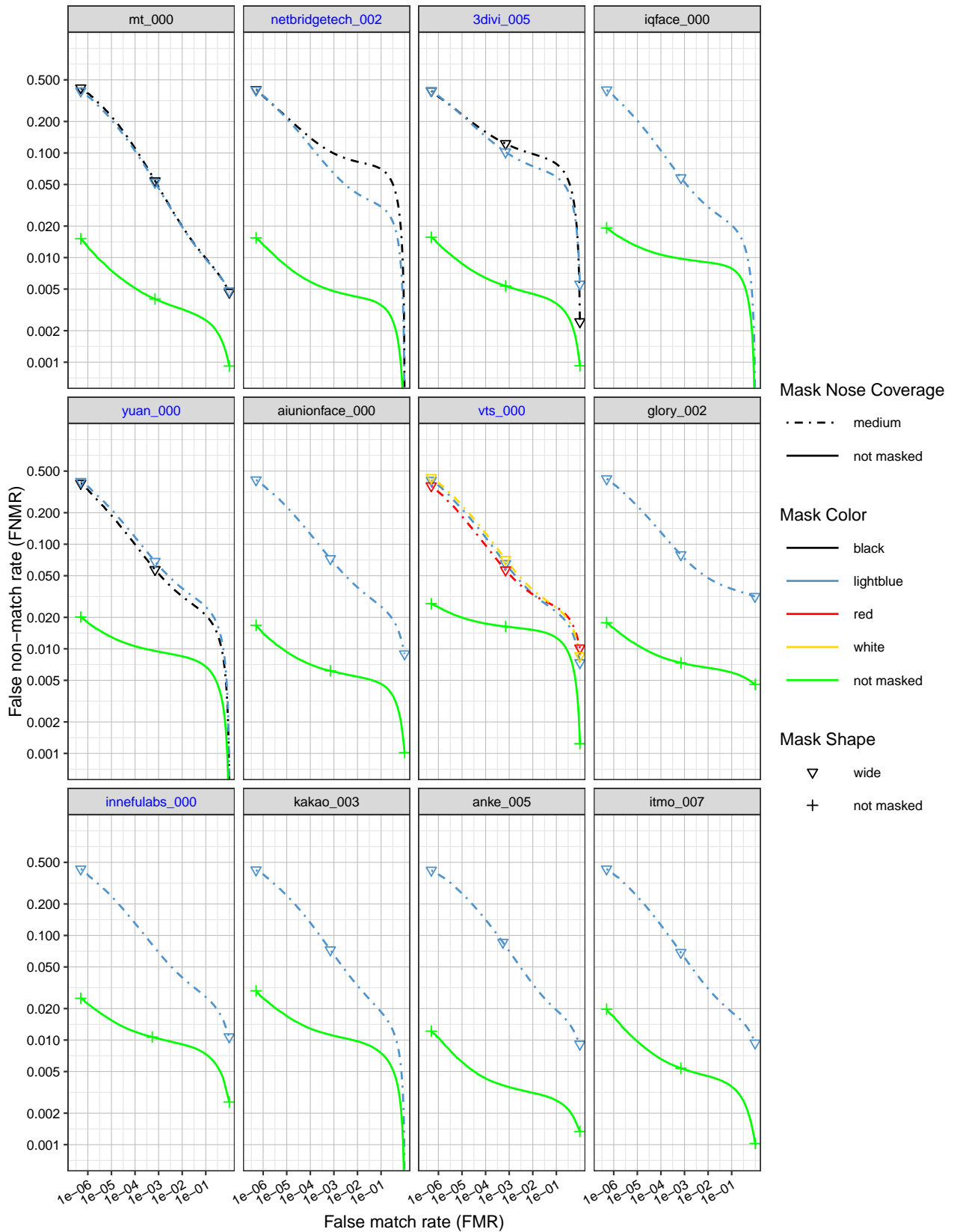


Figure 29: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

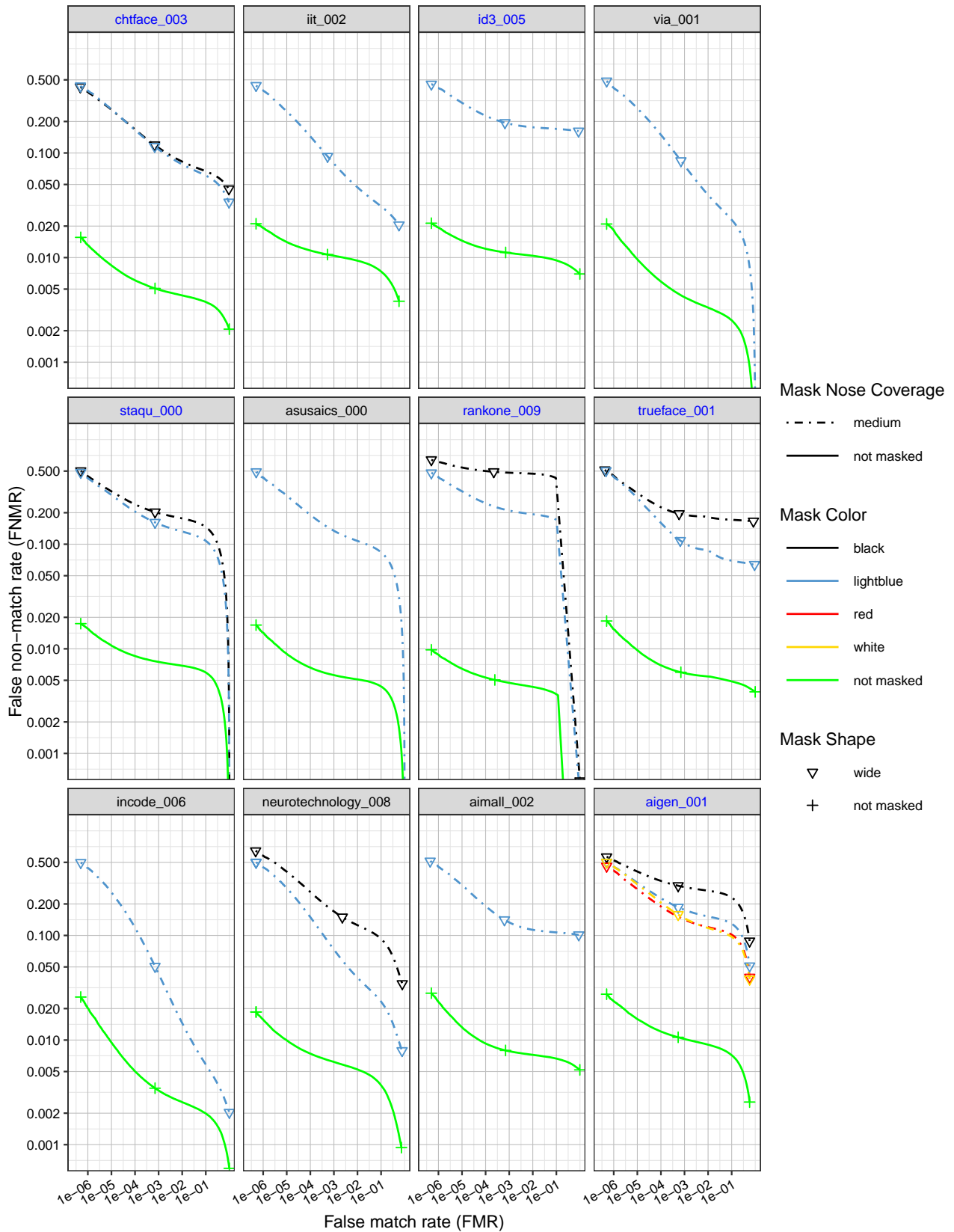


Figure 30: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

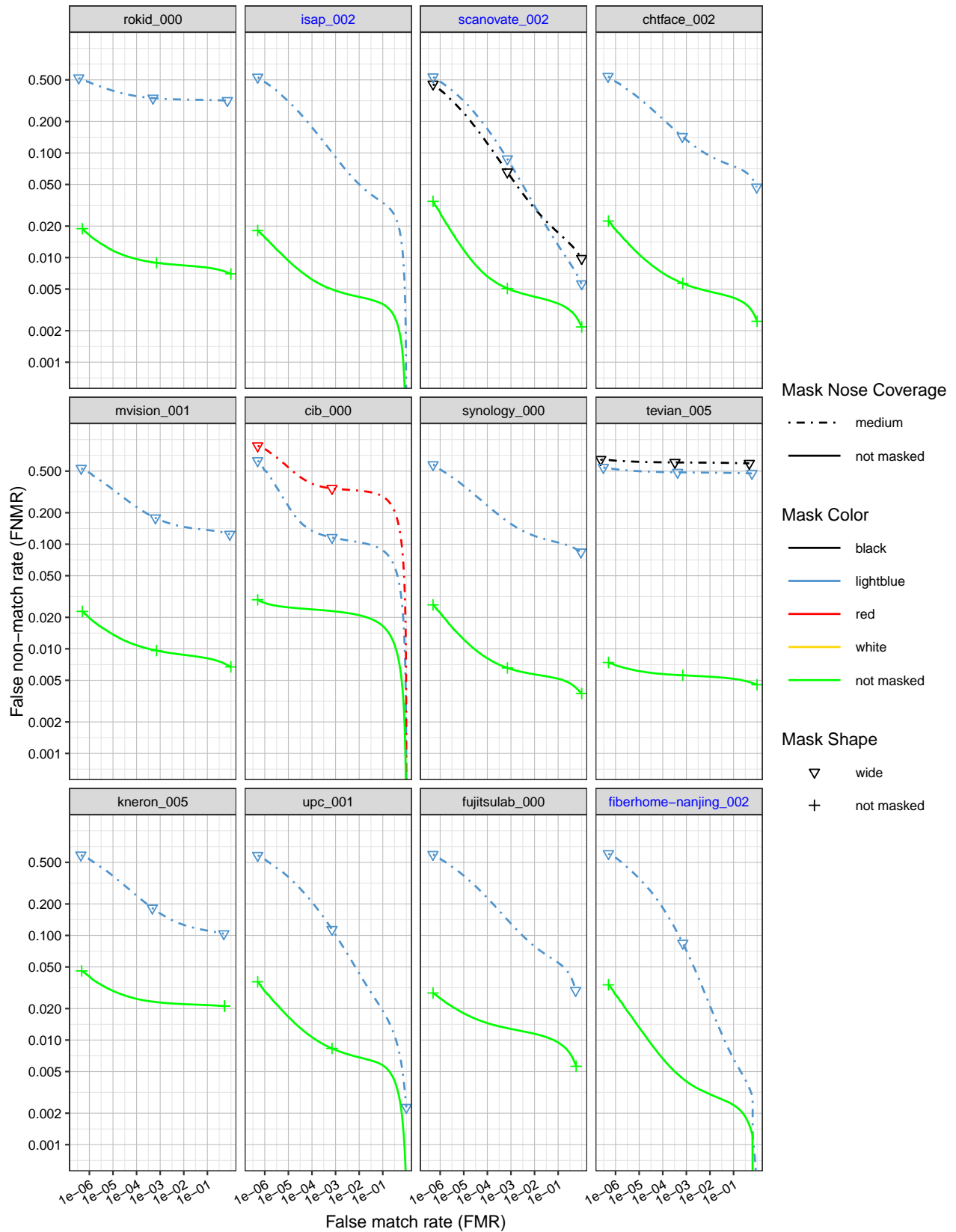


Figure 31: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

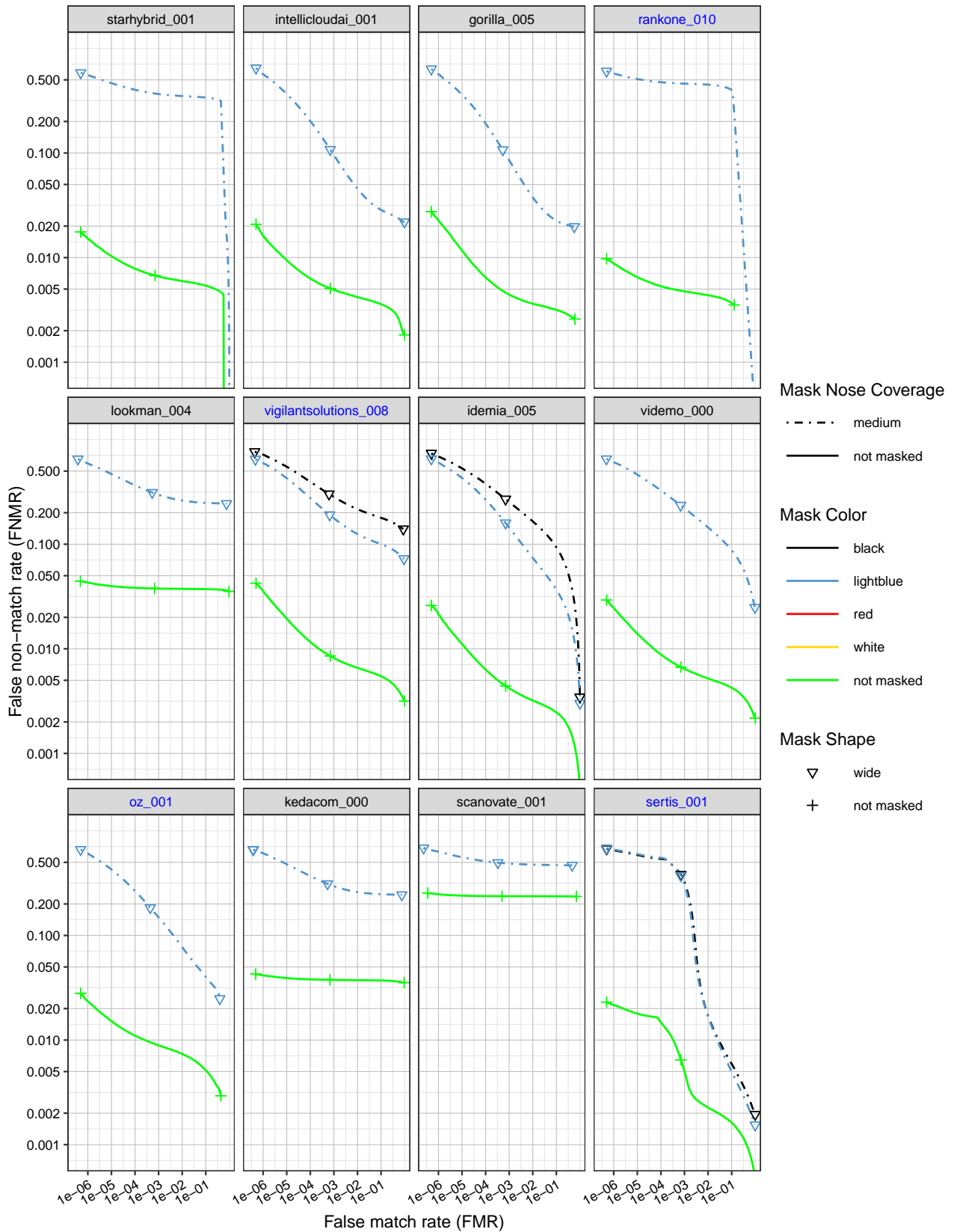


Figure 32: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

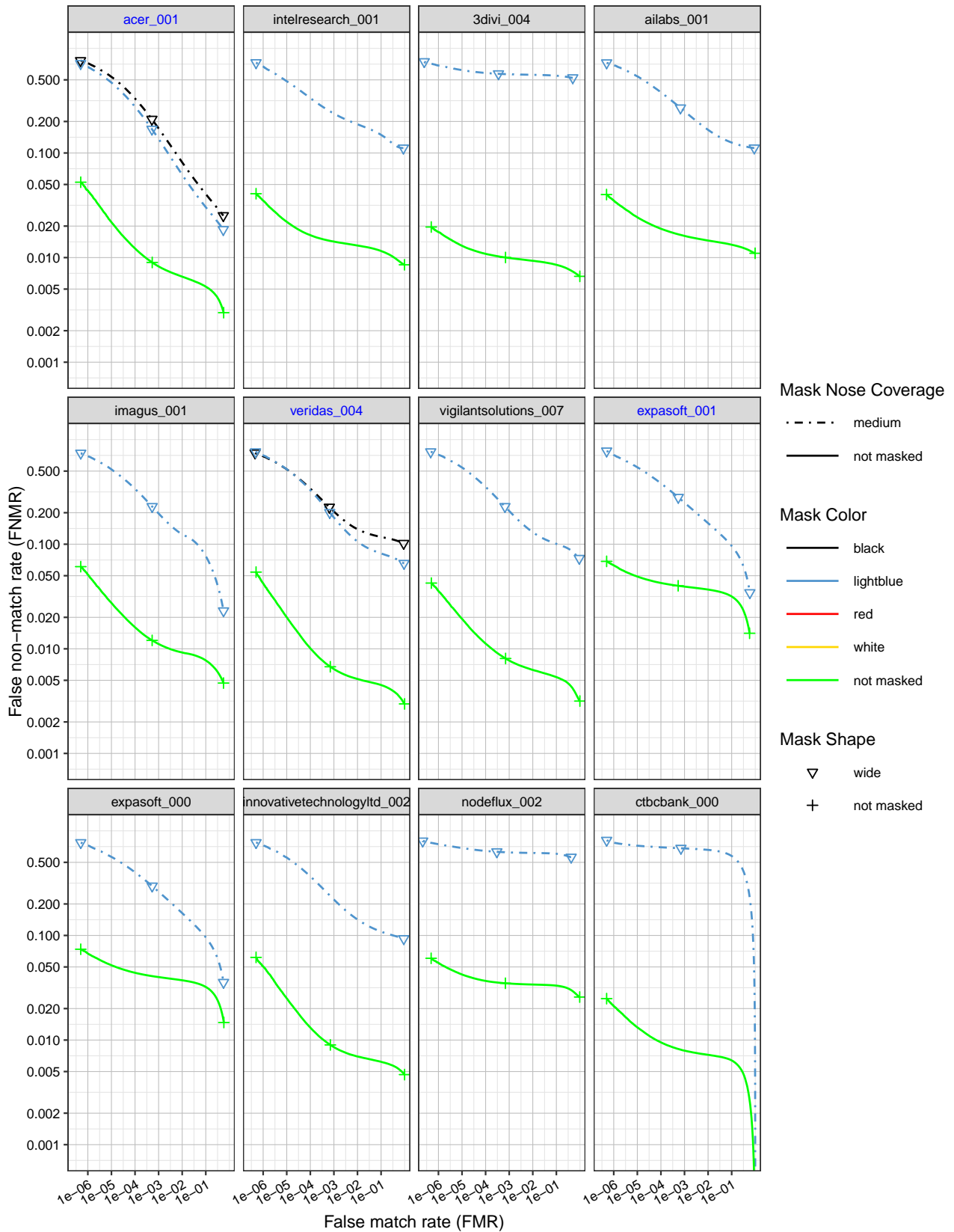


Figure 33: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

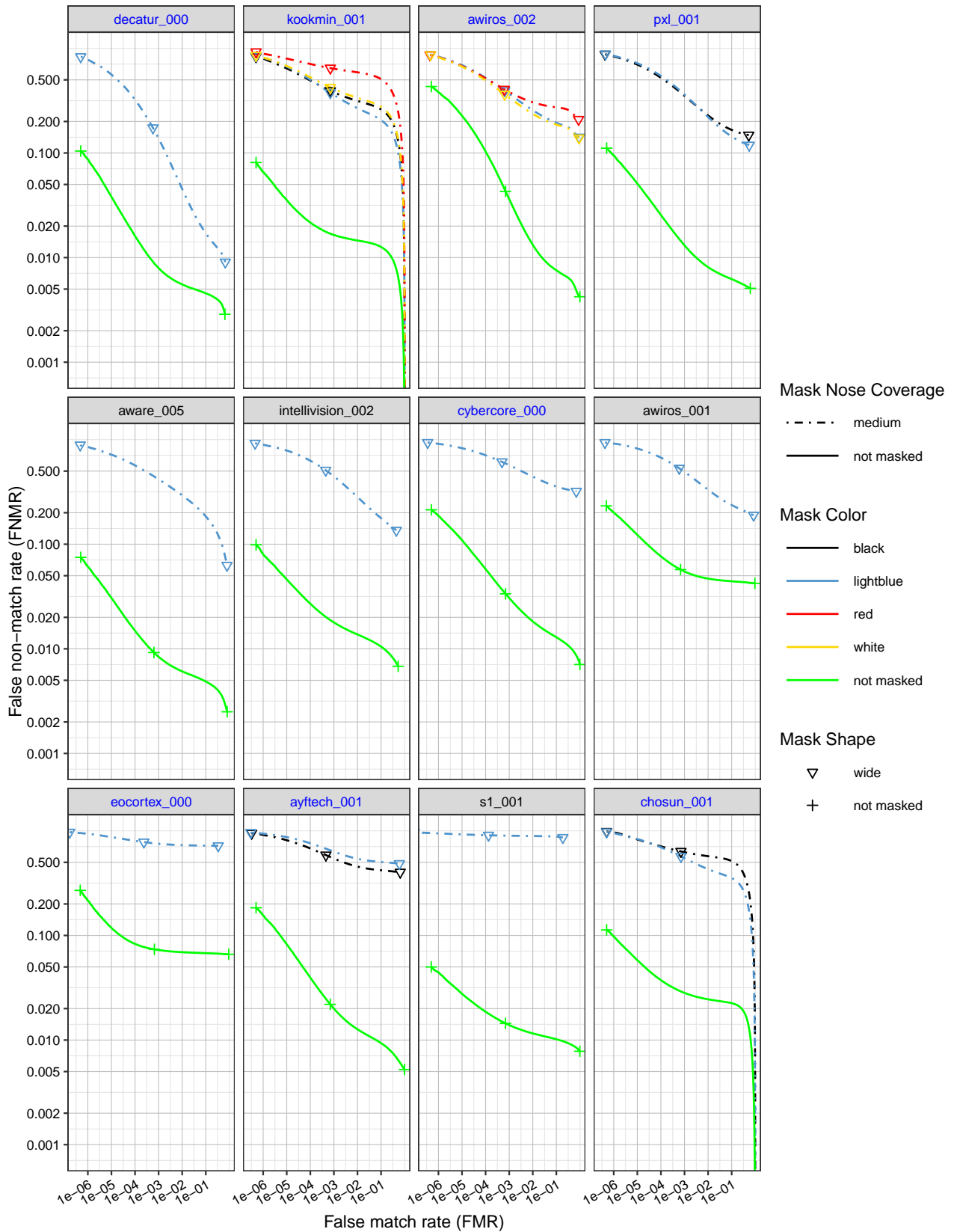


Figure 34: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

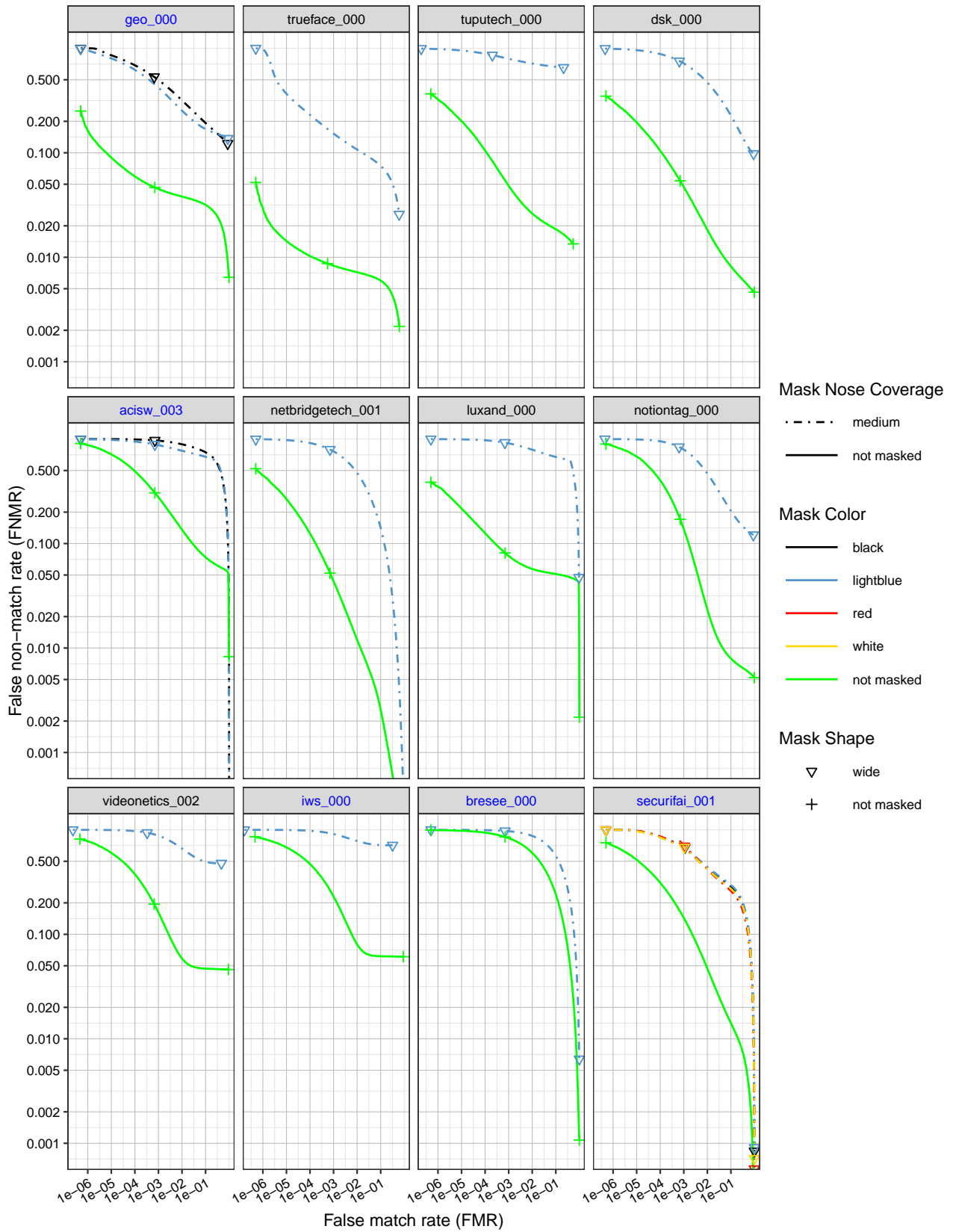


Figure 35: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

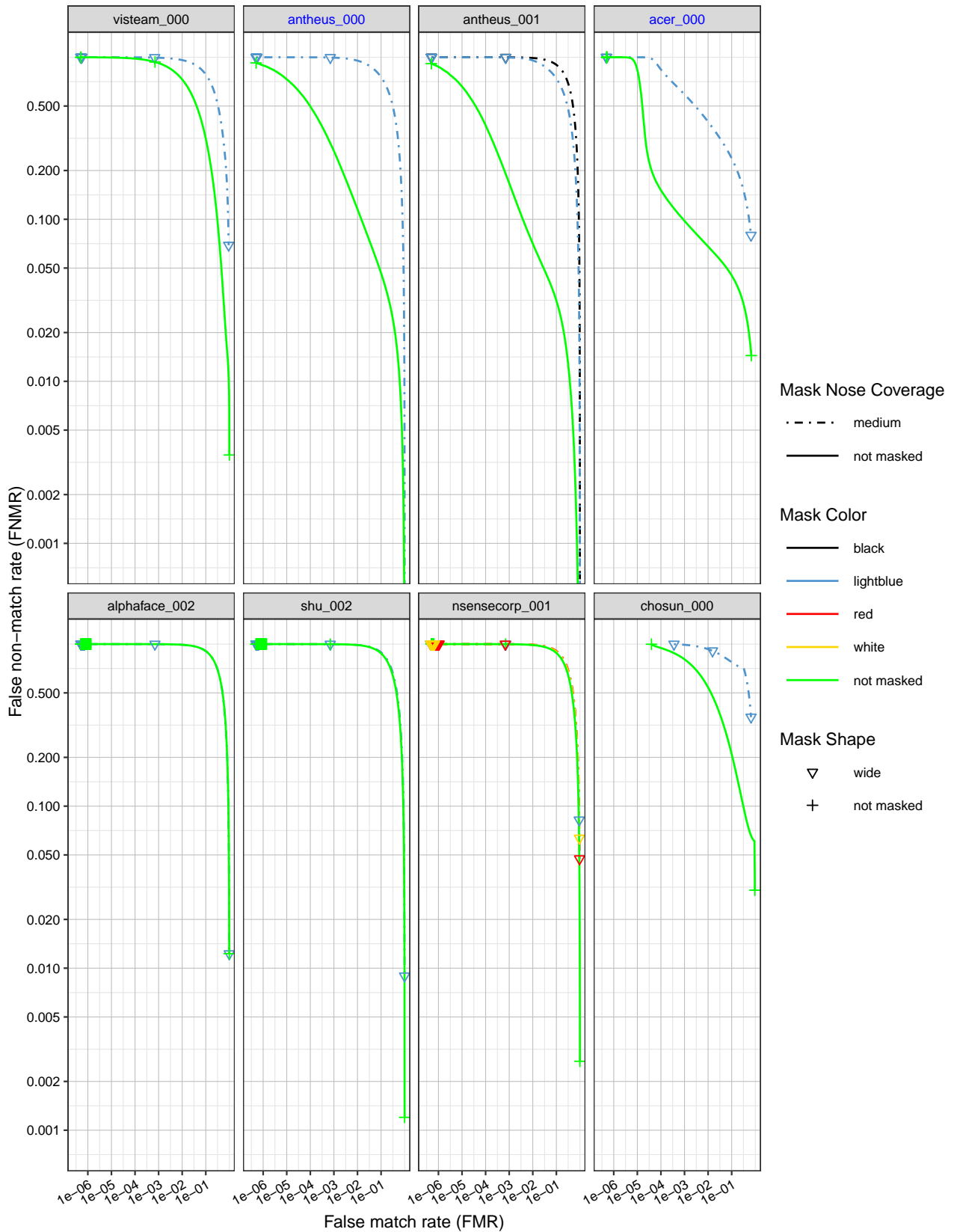


Figure 36: DET curves showing error rates on unmasked and masked probe images, broken out by mask color. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

The following plots show the explicit dependence of false non-match rate (FNMR) and false match rate (FMR) on score threshold for each algorithm, across different masked/unmasked combinations.

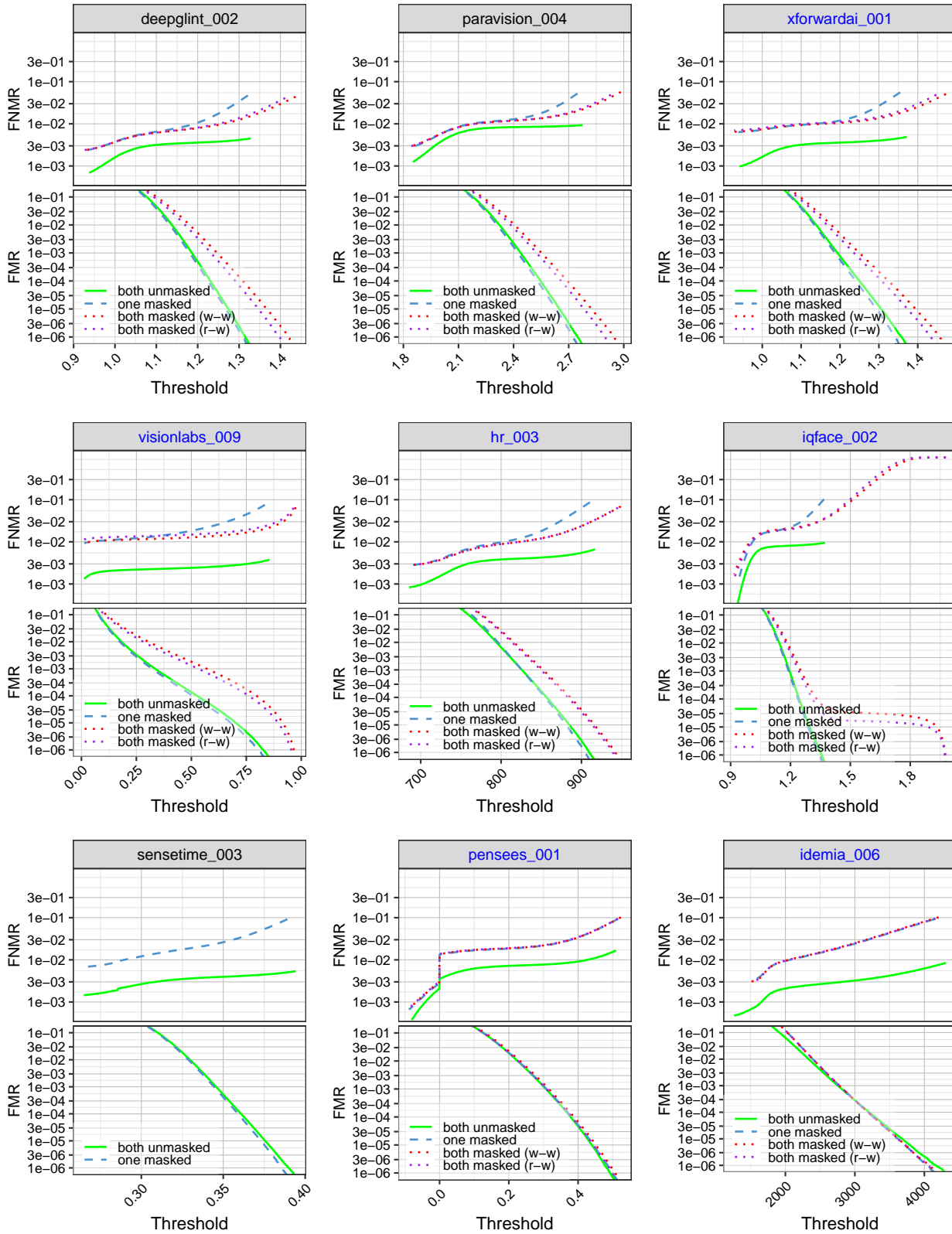


Figure 37: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

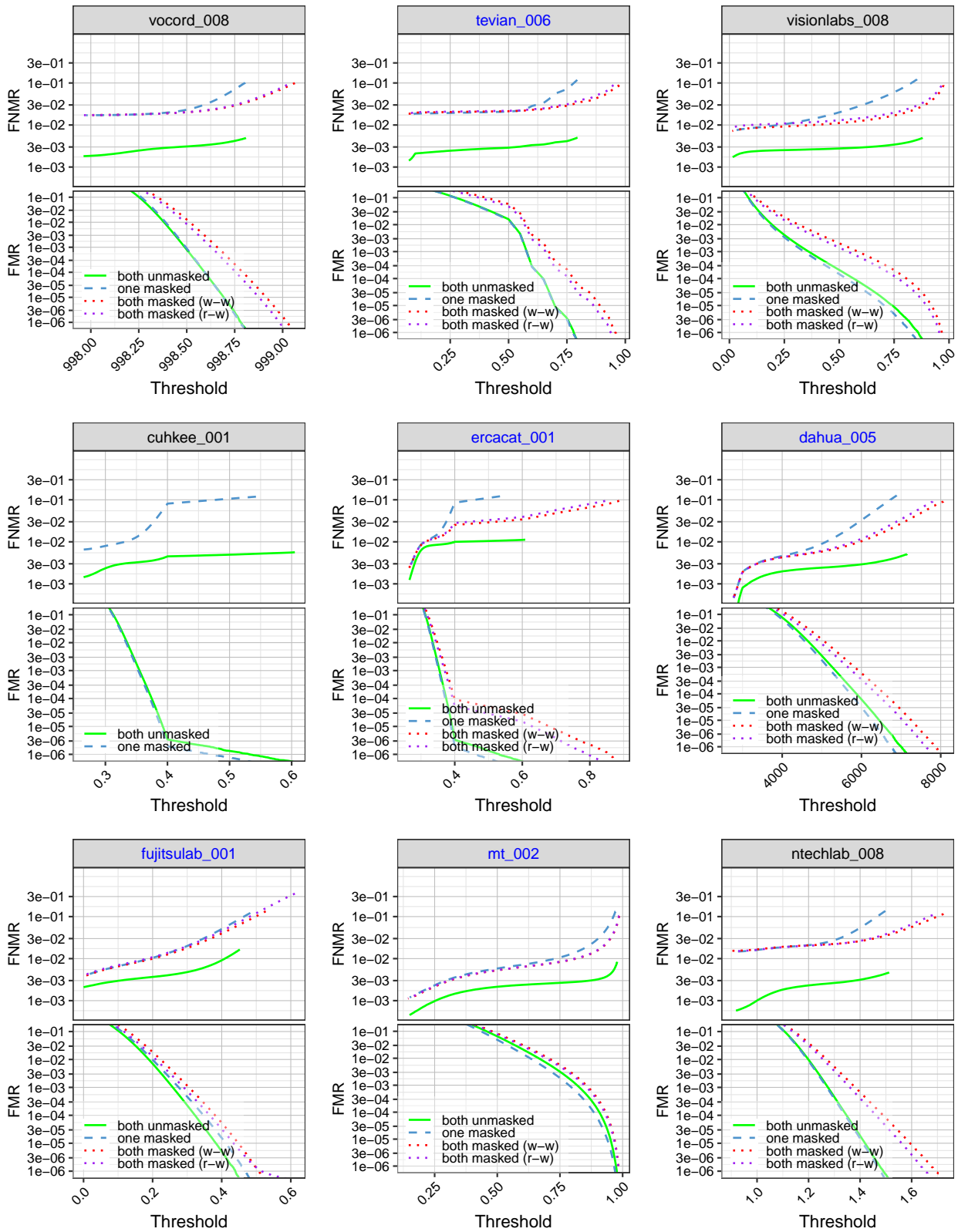


Figure 38: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

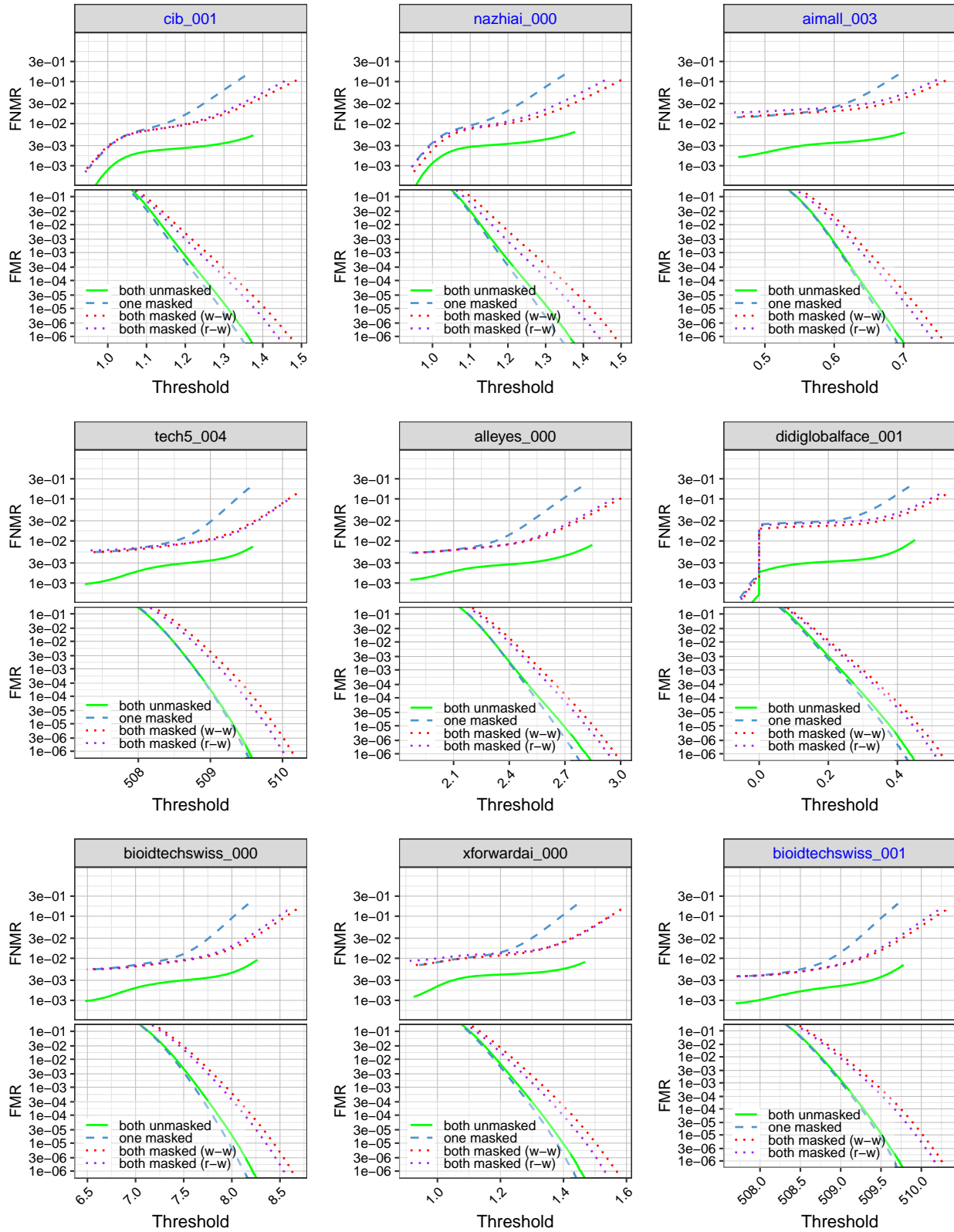


Figure 39: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

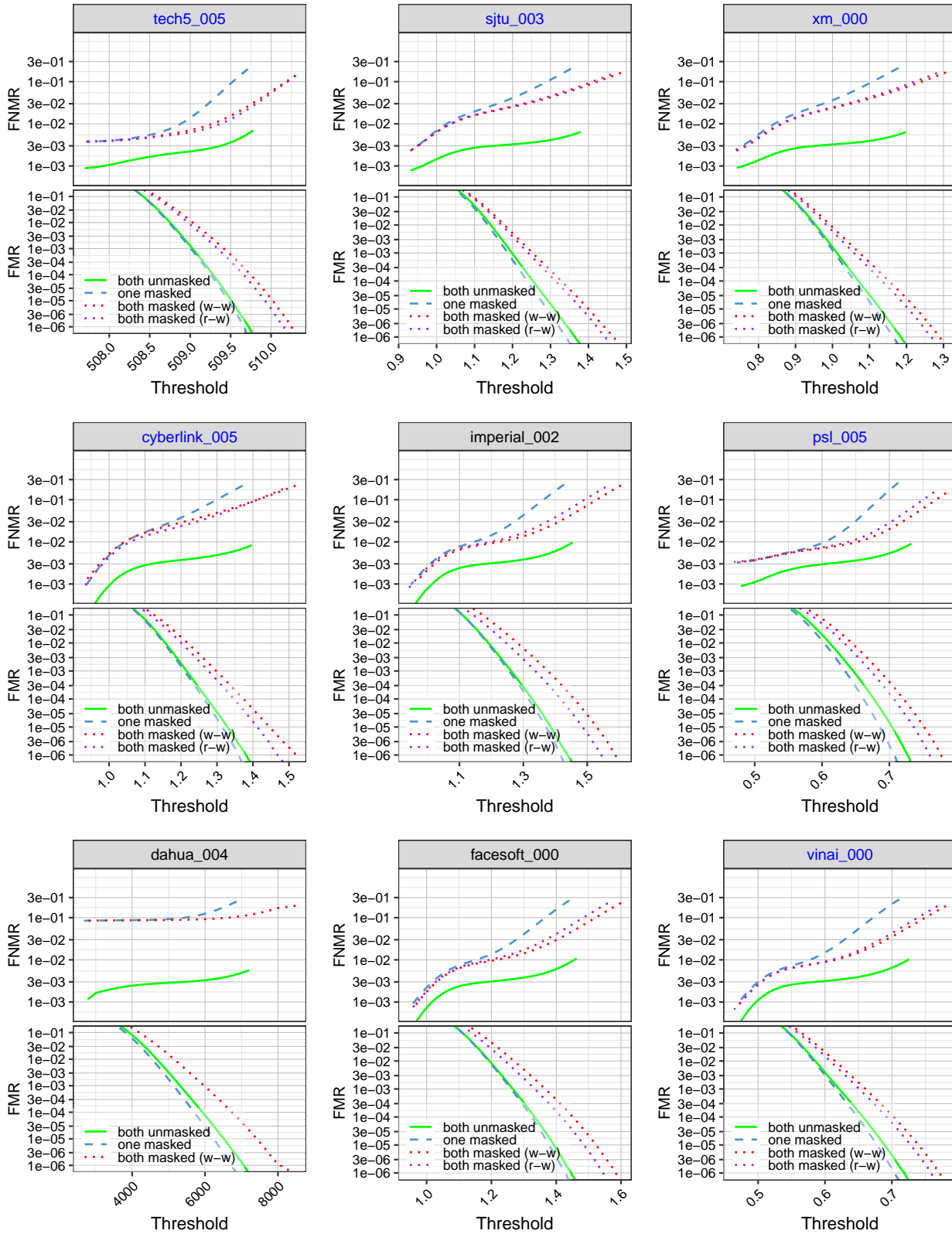


Figure 40: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

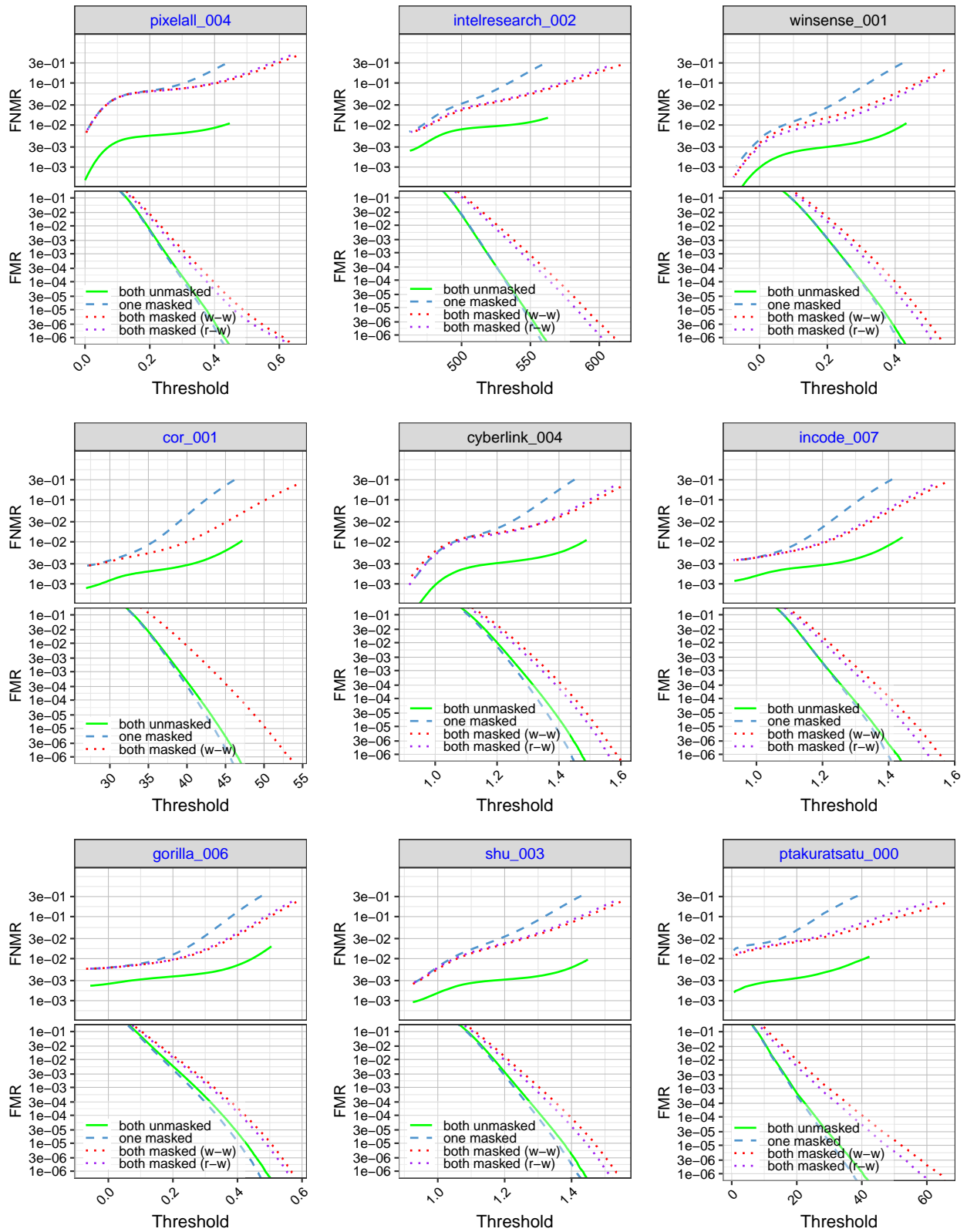


Figure 41: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

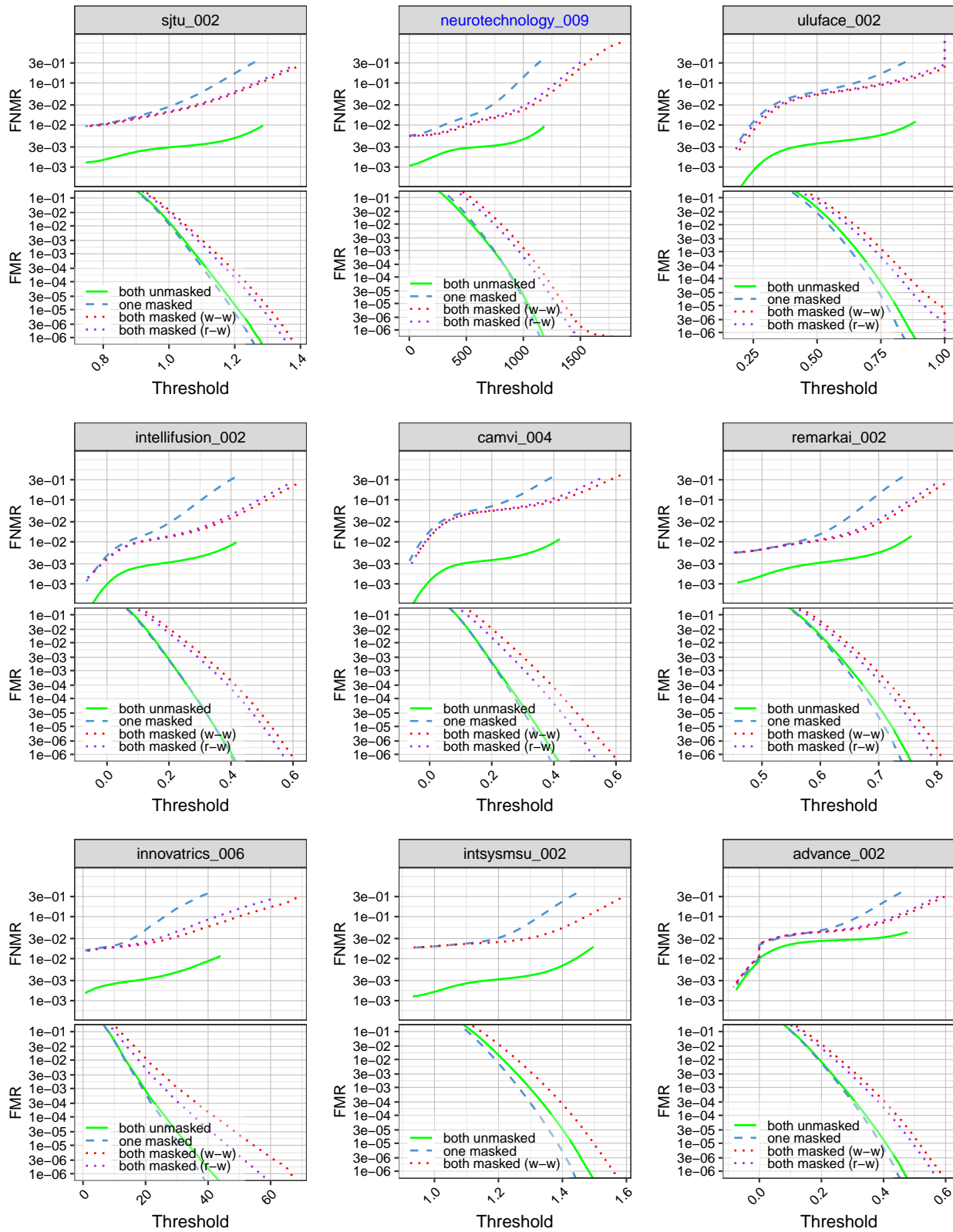


Figure 42: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

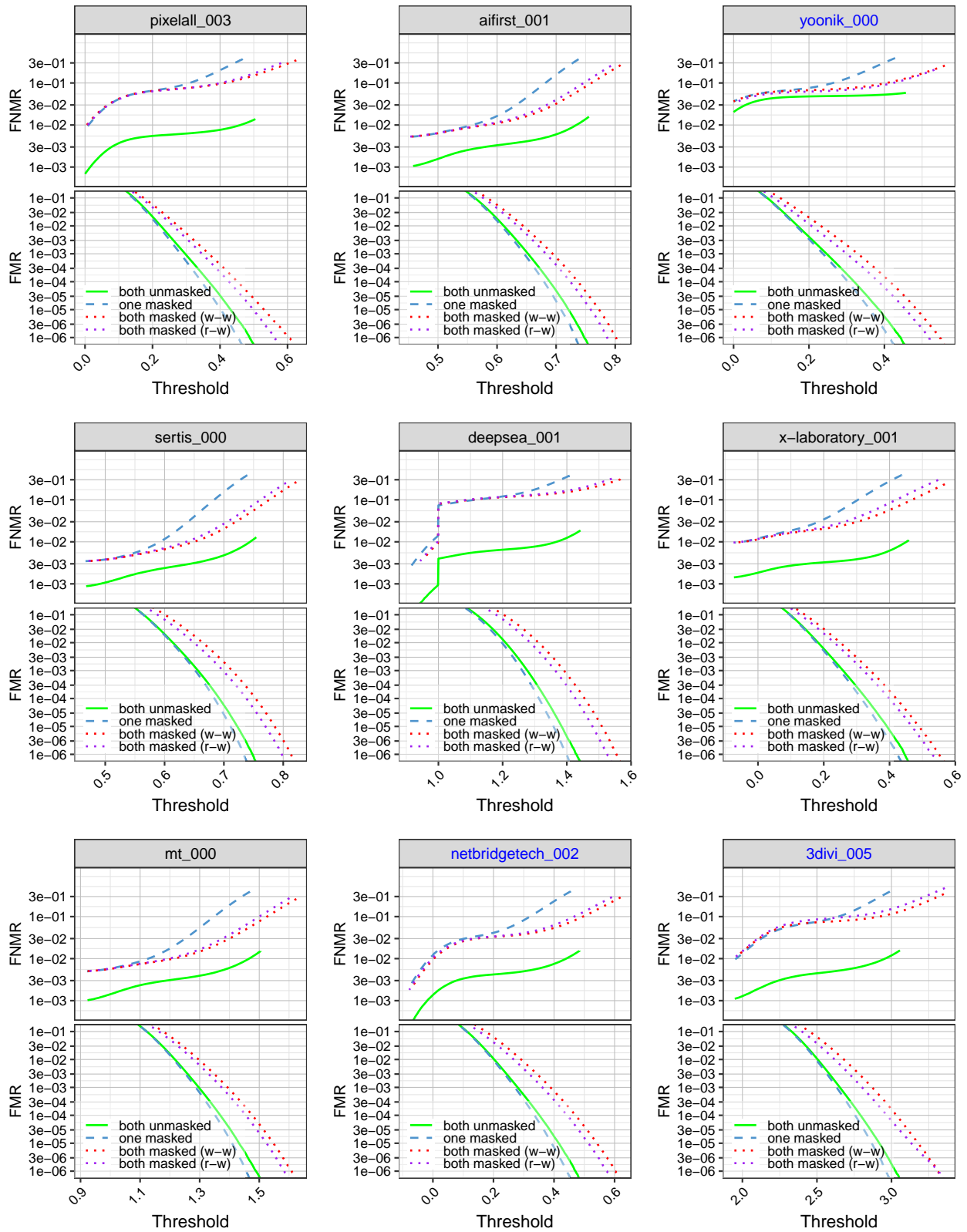


Figure 43: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

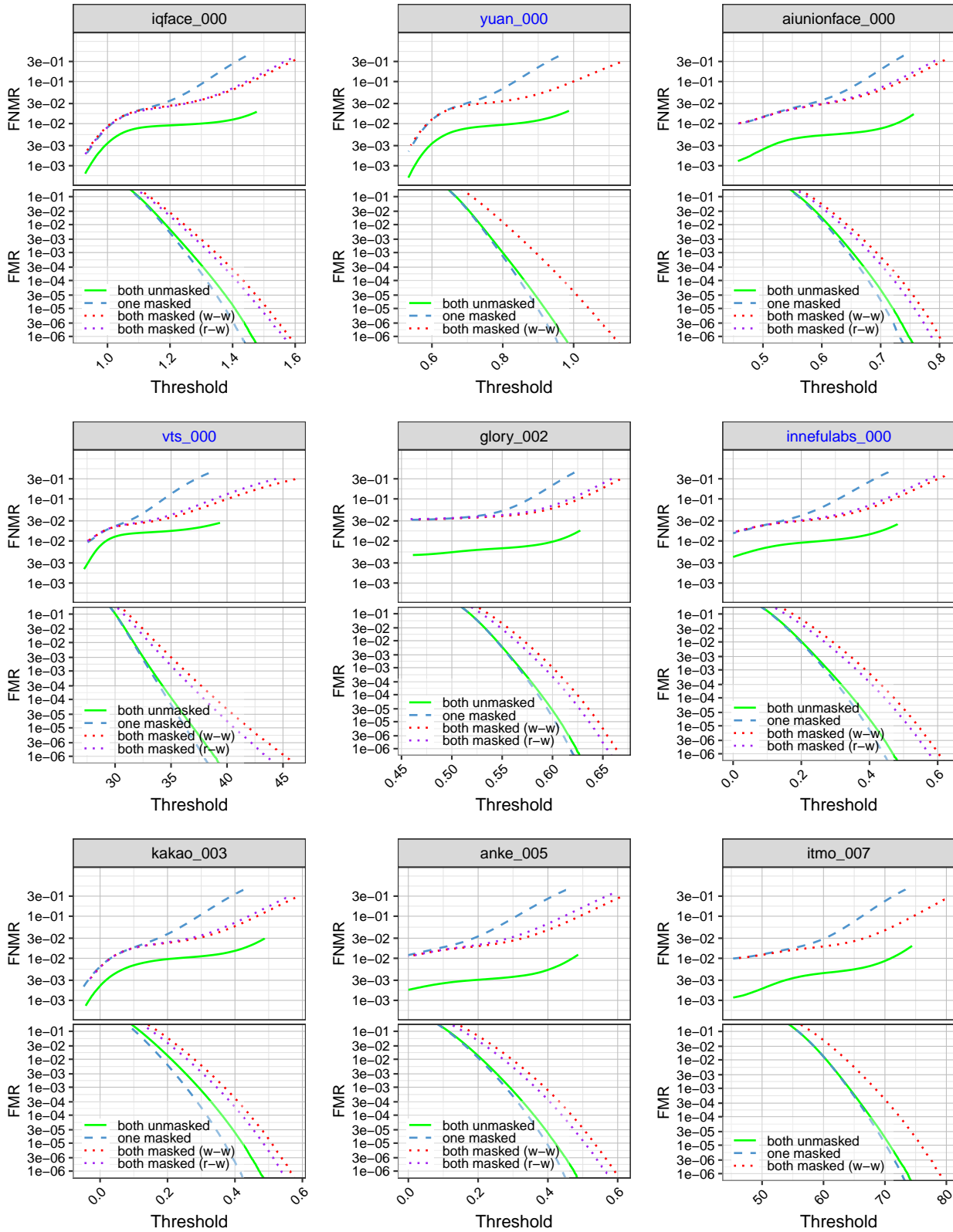


Figure 44: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

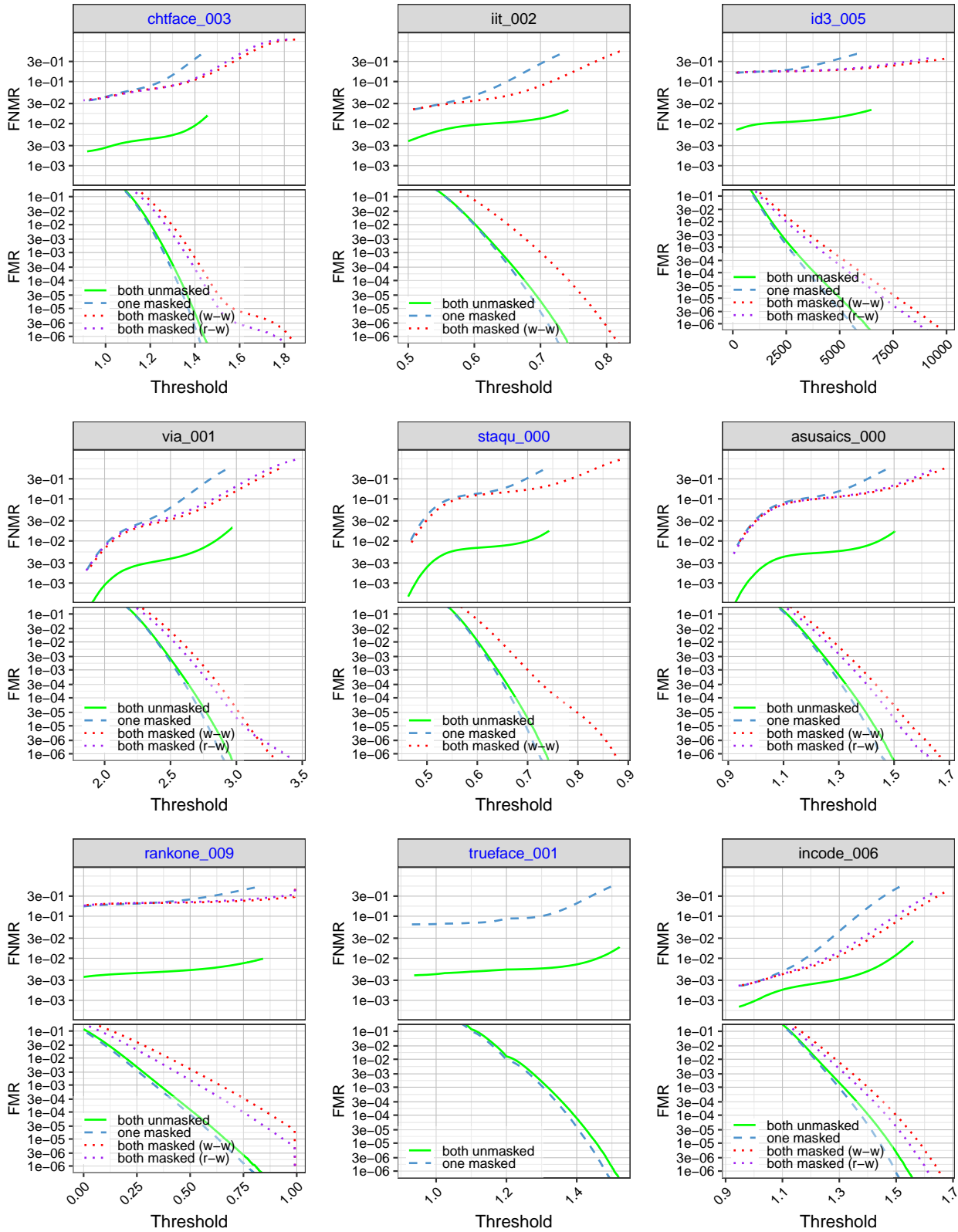


Figure 45: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

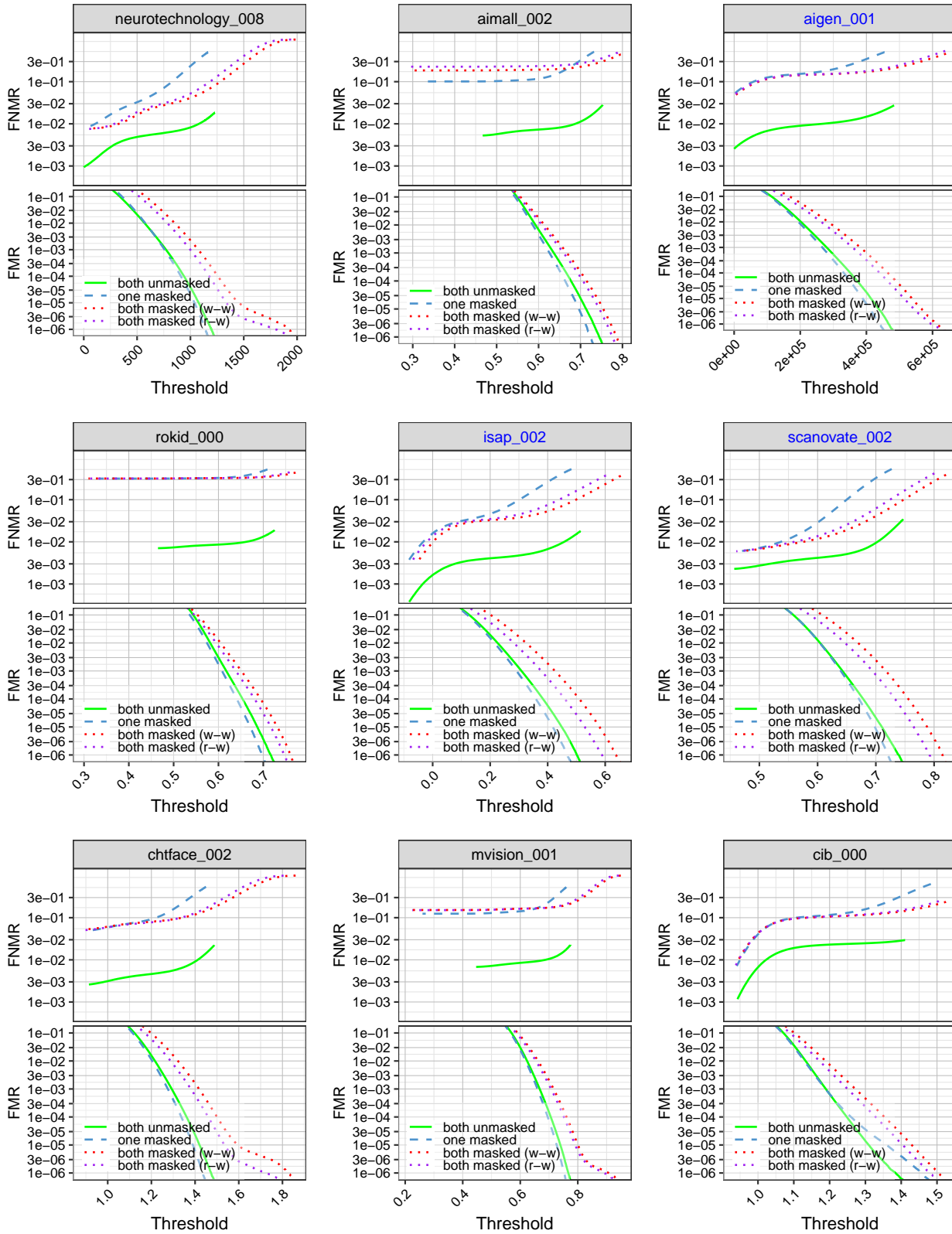


Figure 46: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

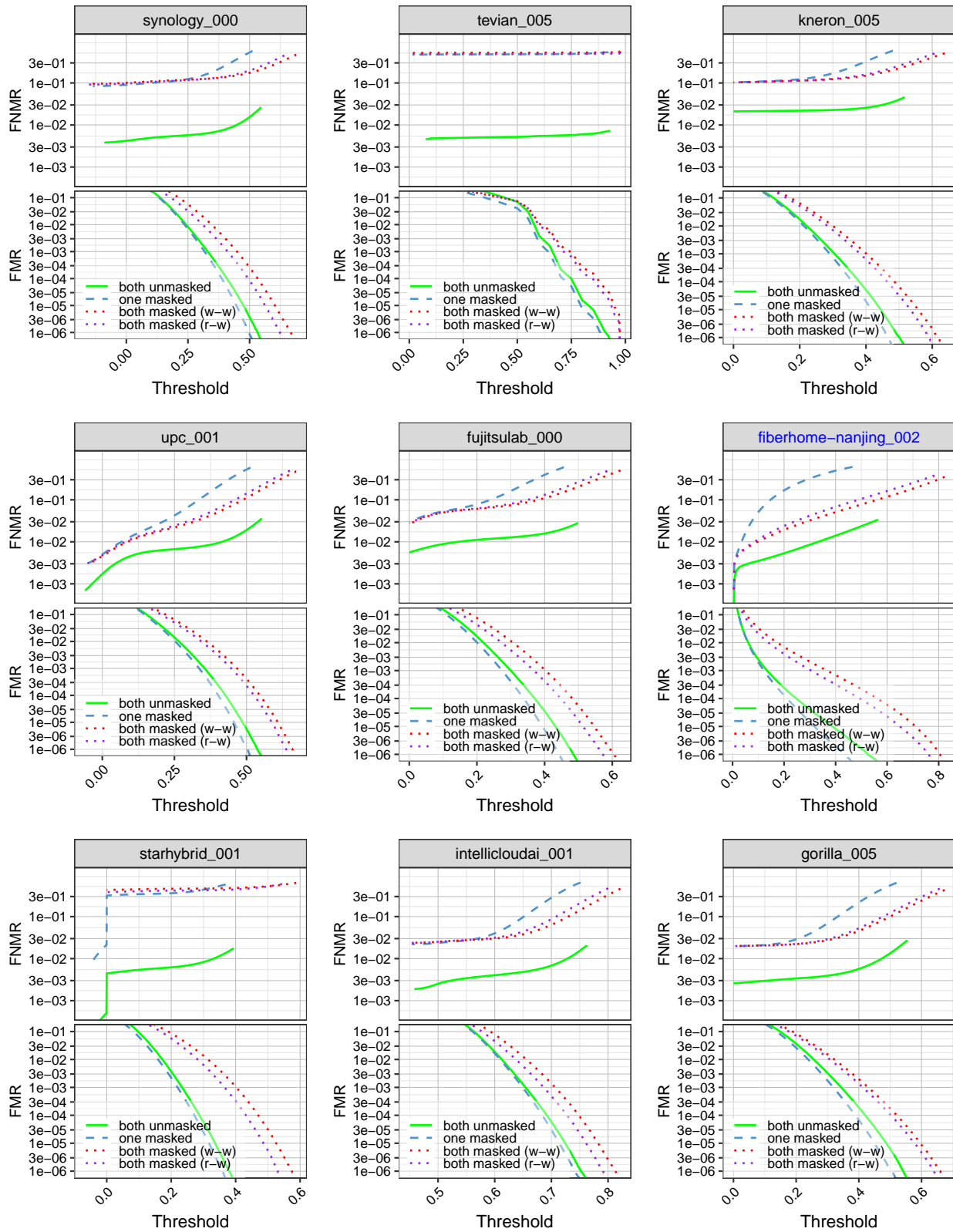


Figure 47: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

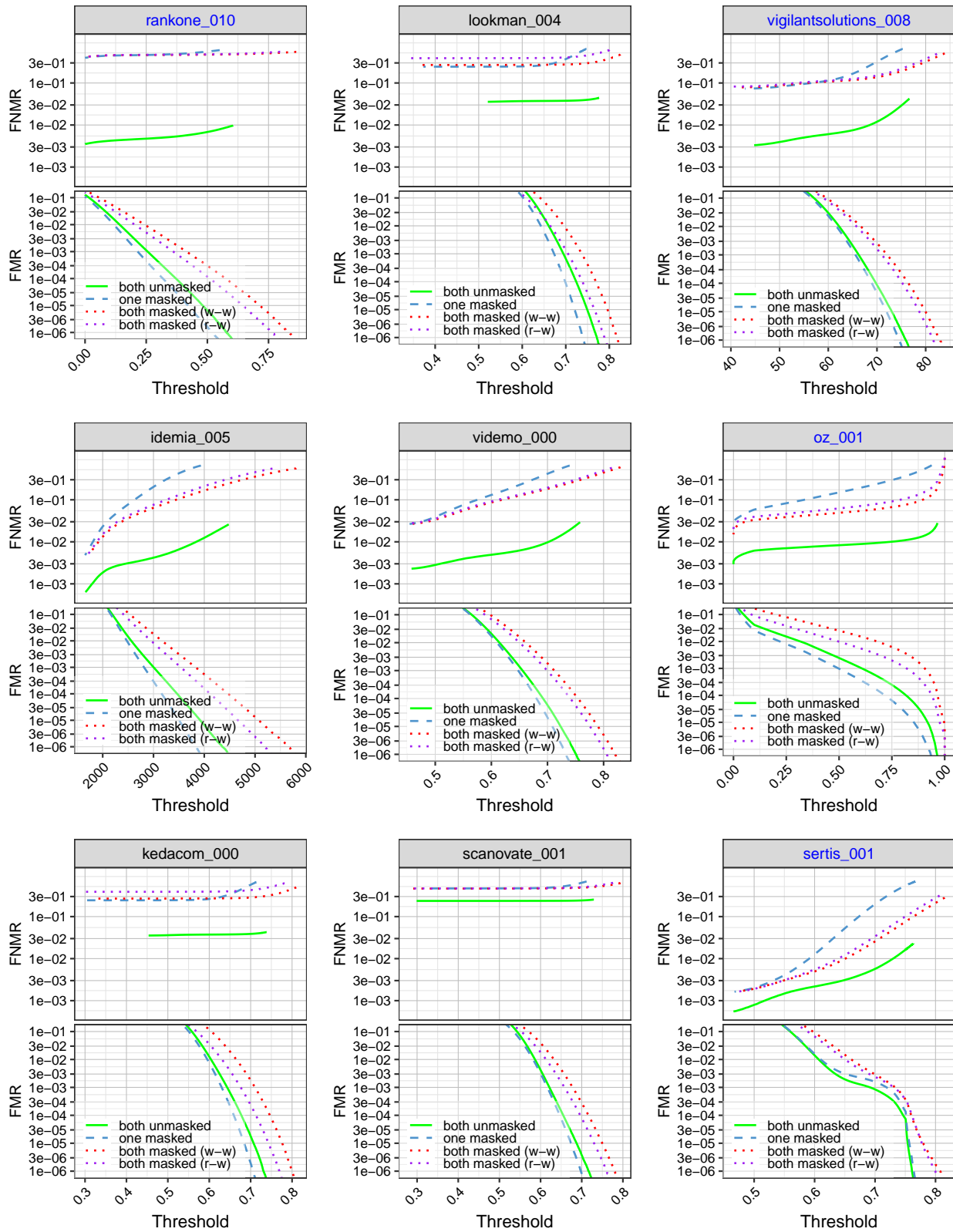


Figure 48: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

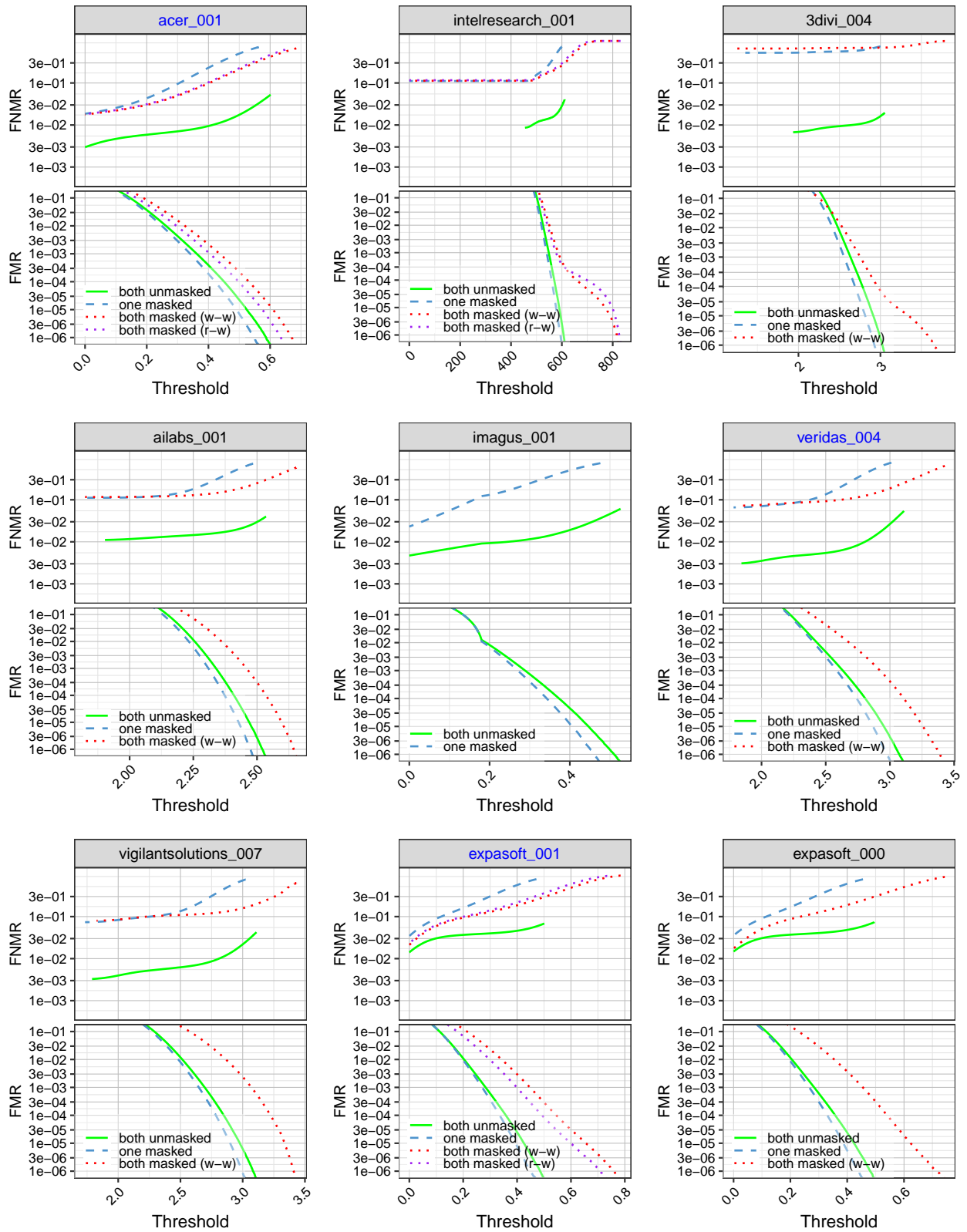


Figure 49: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

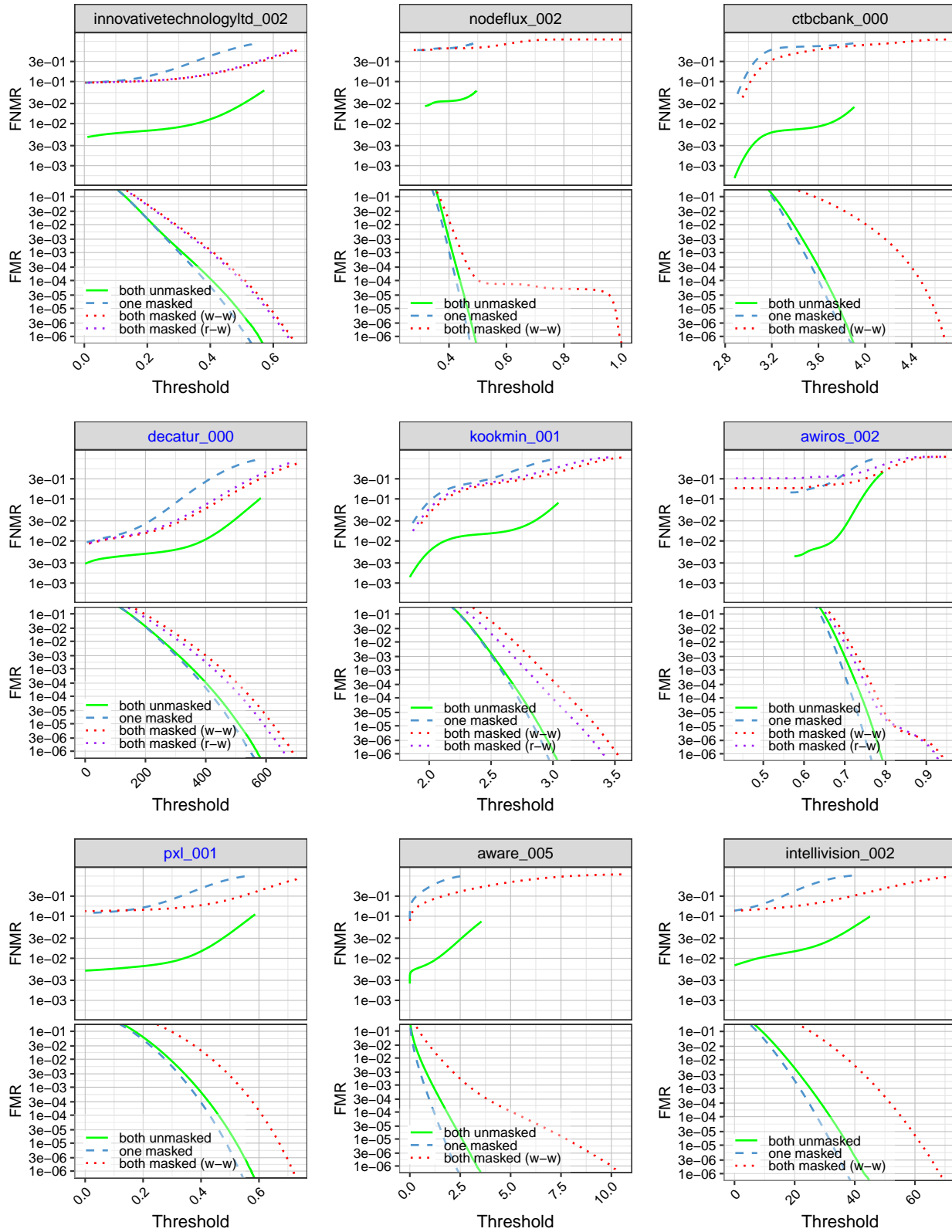


Figure 50: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

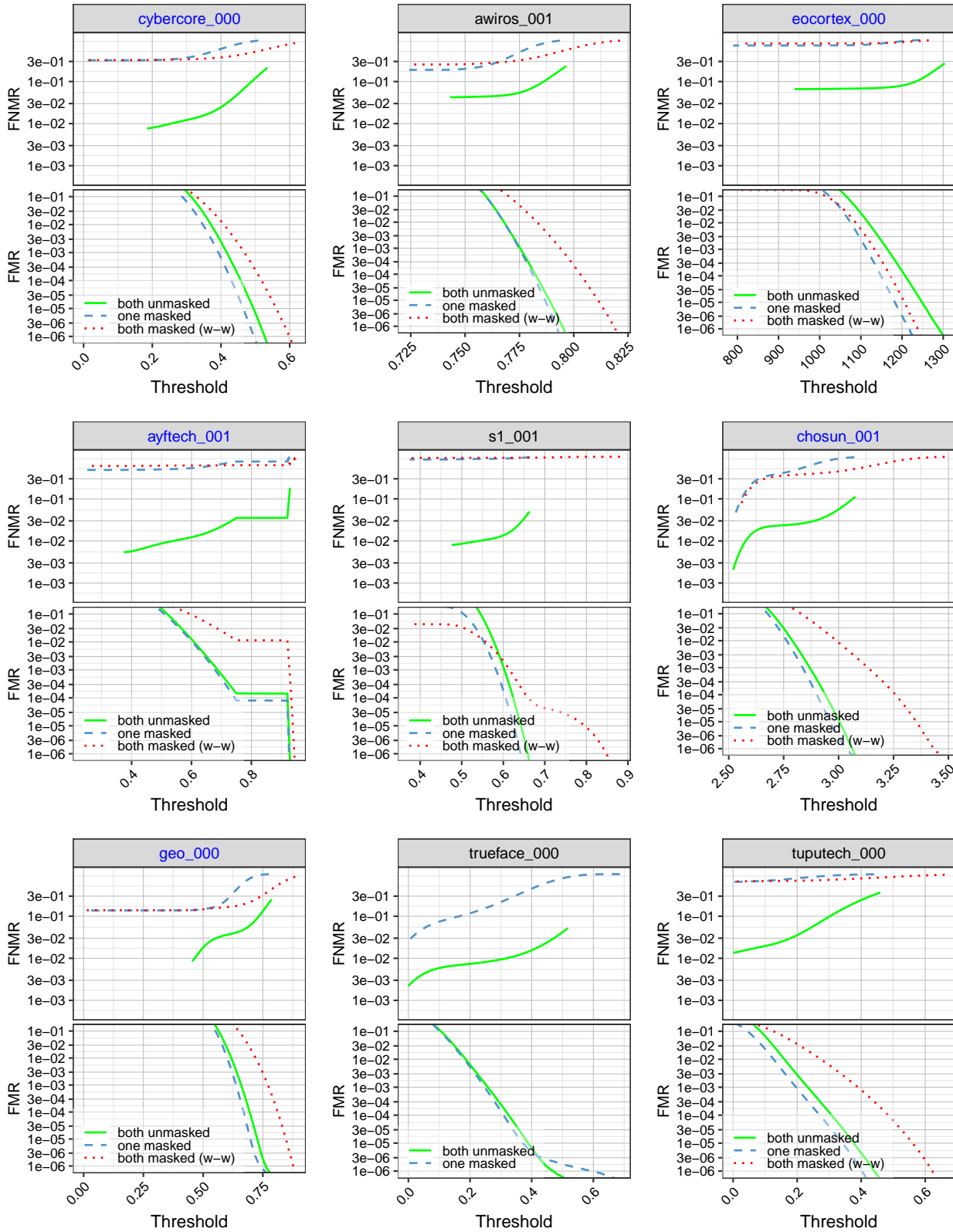


Figure 51: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

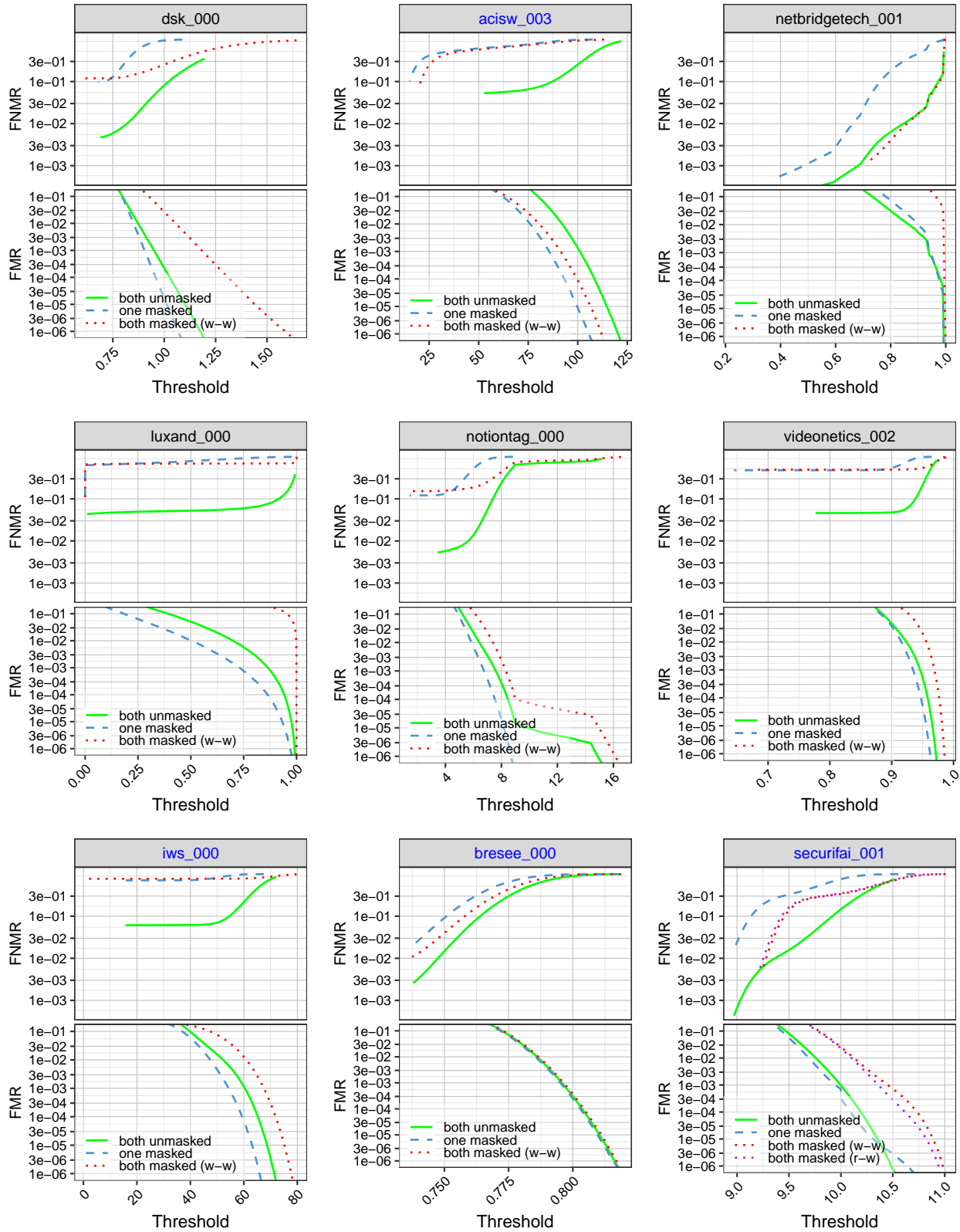


Figure 52: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

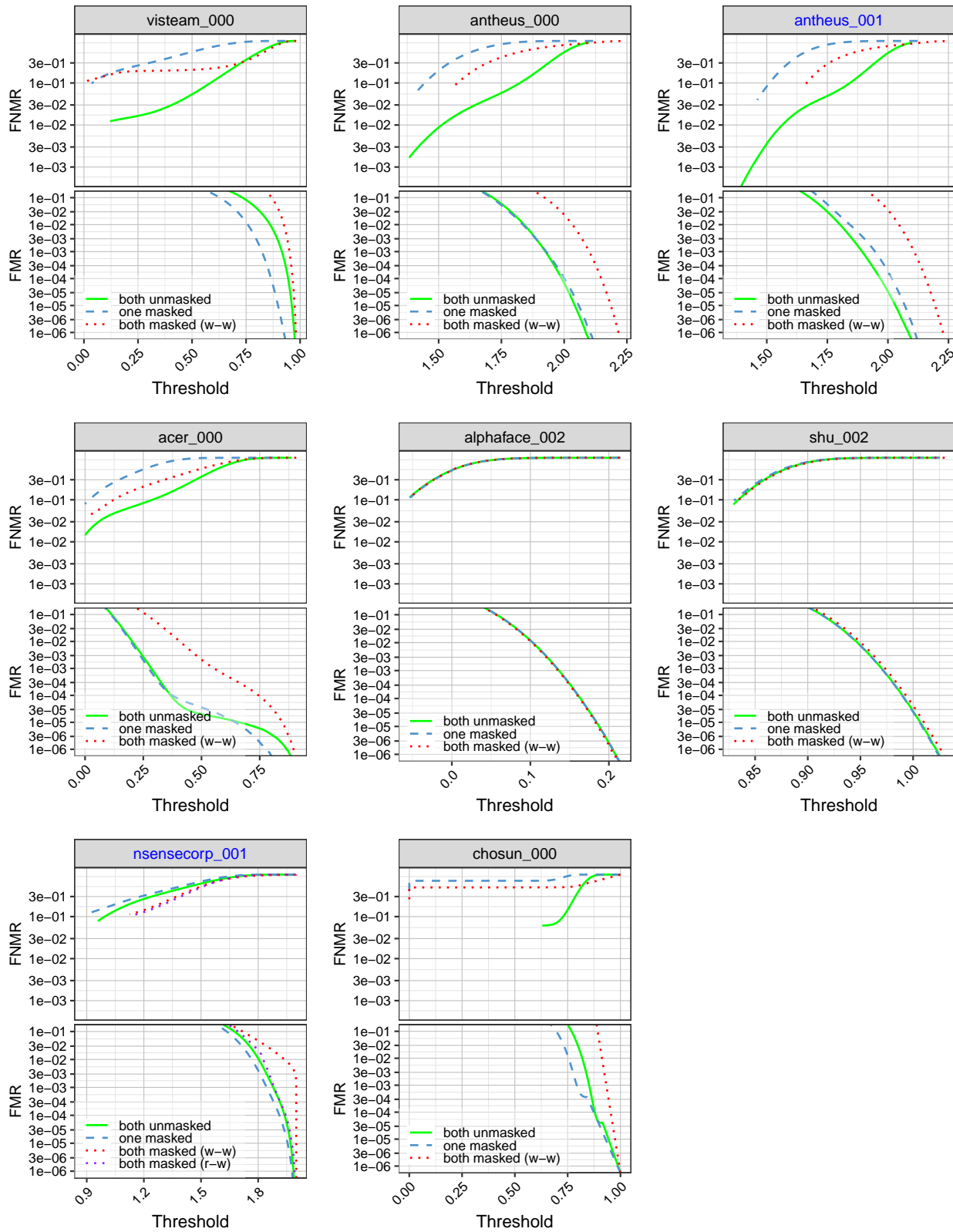


Figure 53: FNMR and FMR calibration curves on unmasked-to-unmasked (both unmasked), masked-to-unmasked with medium, wide, lightblue masks (one masked), masked-to-masked with medium, wide, lightblue masks (both masked (w-w)), and masked-to-masked with medium, round, white masks (enrollment image) and medium, wide, lightblue masks (verification image) (both masked (r-w)). Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

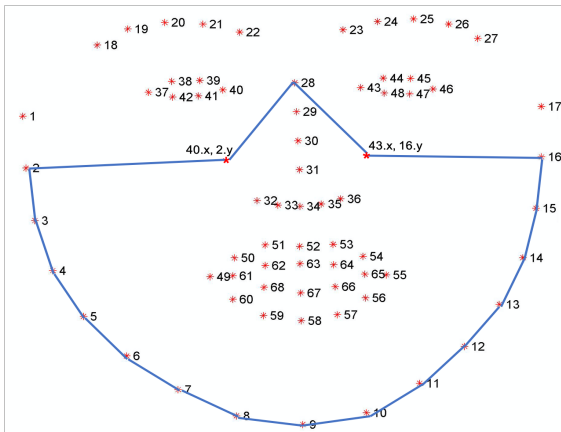
This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>

References

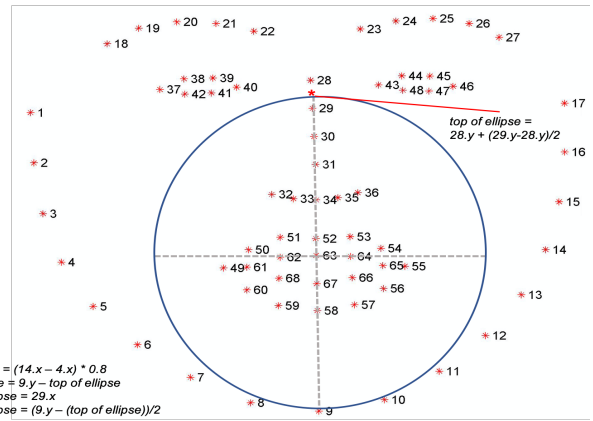
- [1] ISO/IEC 19794-5:2011 - Information technology - Biometric data interchange formats - Part 5: Face image data.
- [2] N95 Respirators, Surgical Masks, and Face Masks. <https://www.fda.gov/medical-devices/personal-protective-equipment-infection-control/n95-respirators-surgical-masks-and-face-masks>.
- [3] NIST Special Database 32 - Multiple Encounter Dataset (MEDS-II). <https://www.nist.gov/itl/iad/image-group/special-database-32-multiple-encounter-dataset-meds>.
- [4] Patrick Grother, Mei Ngan, and Kayee Hanaoka. NISTIR 8311 - Ongoing FRVT Part 6A: Face recognition accuracy with face masks using pre-COVID-19 algorithms. Interagency Report 8311, National Institute of Standards and Technology, July 2017. <https://pages.nist.gov/frvt/reports/11/frvt.11.report.2017.07.29.pdf>.
- [5] 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>.
- [6] J. Hanley and A. Lippman-Hand. If nothing goes wrong, is everything all right? interpreting zero numerators. *JAMA*, 249 13:1743–5, 1983.
- [7] Davis E. King. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, 10:1755–1758, 2009. <http://dlib.net>.
- [8] Mei Ngan, Patrick Grother, and Kayee Hanaoka. NISTIR 8311 - Ongoing FRVT Part 6A: Face recognition accuracy with face masks using pre-COVID-19 algorithms. Interagency Report 8311, National Institute of Standards and Technology, July 2020. <https://doi.org/10.6028/NIST.IR.8311>.
- [9] Patrick Grother and Mei Ngan and Kayee Hanaoka. NIST Ongoing Face Recognition Vendor Test (FRVT) 1:1 Verification Application Programming Interface (API), April 2019. https://pages.nist.gov/frvt/api/FRVT_ongoing_11_api_4.0.pdf.
- [10] Zhongyuan Wang, Guangcheng Wang, Baojin Huang, Zhangyang Xiong, Qi Hong, Hao Wu, Peng Yi, Kui Jiang, Nanxi Wang, Yingjiao Pei, Heling Chen, Yu Miao, Zhibing Huang, and Jinbi Liang. Masked face recognition dataset and application, 2020.

Appendix A Dlib Masking Methodology

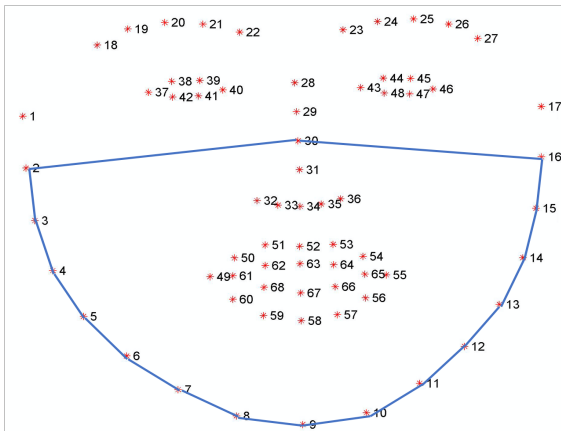
This publication is available free of charge from: <https://doi.org/10.6028/NIST.IR.8331>



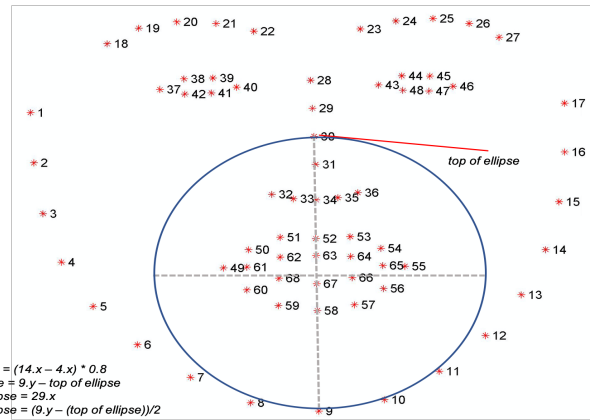
(a) wide, high coverage



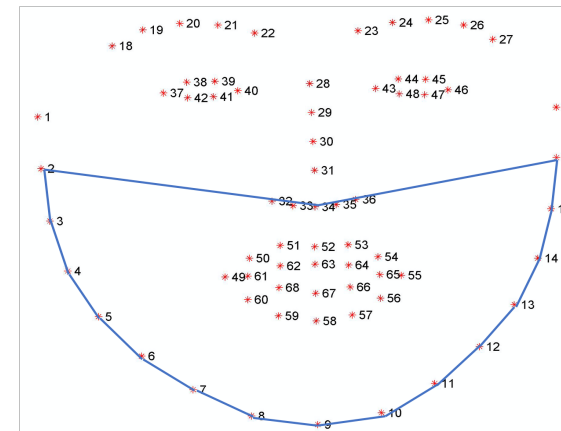
(b) round, high coverage



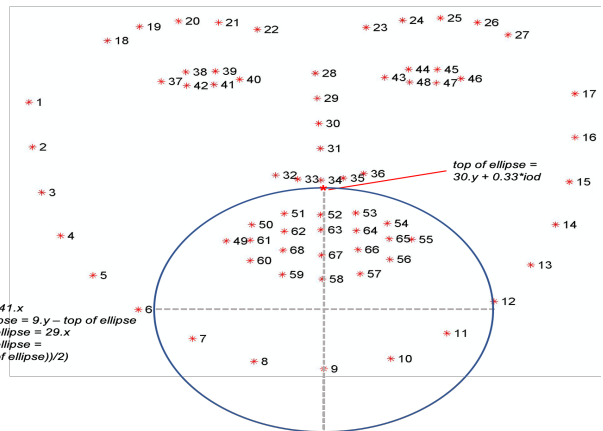
(c) wide, medium coverage



(d) round, medium coverage



(e) wide, low coverage



(f) round, low coverage

Figure 54: This figure shows the Dlib facial points used to create the various synthetic masks used in this report. For wide masks, the specified Dlib facial points were used to generate a closed polygon and two additional points were interpolated between each dlib facial point used for smoothing purposes. For round masks, the specified Dlib facial points were used to generate an ellipse. The Dlib C++ toolkit version 19.19, configured with the common histogram of gradients (HoG)-based face detector and 68 face landmark shape predictor was used to generate the 68 facial landmarks.