

Evaluation of Face Recognition Accuracy for Subjects Potentially Wearing Face Masks

**Mei Ngan
Patrick Grother
Kayee Hanaoka**

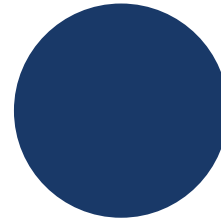
National Institute of Standards & Technology

International Face Performance Conference (IFPC) 2020
October 27, 2020

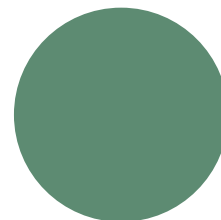
Motivation



Widespread requirement/practice of people wearing protective face masks in public places



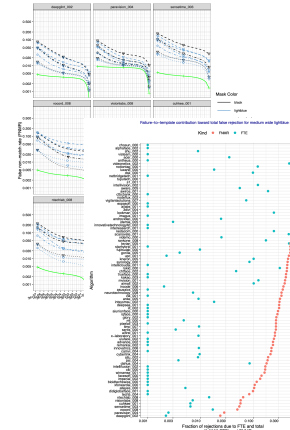
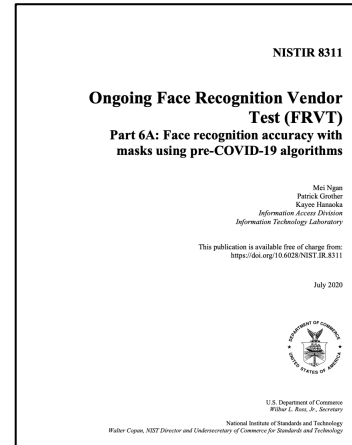
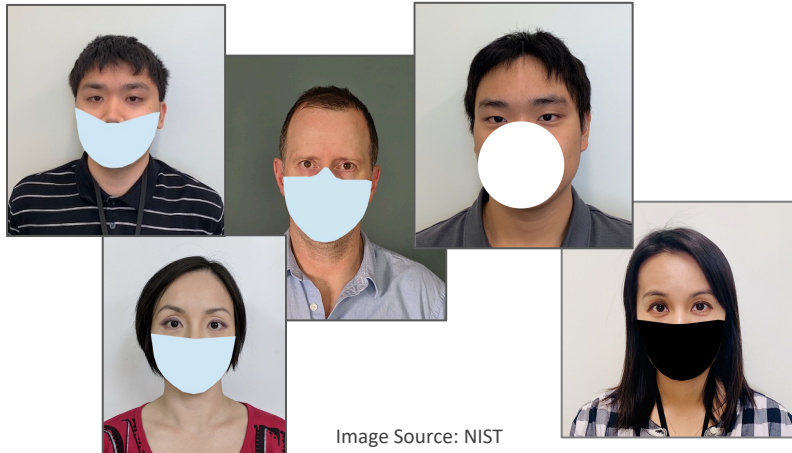
Operational need for face recognition to be applied to persons wearing masks



Masks present a problem for face recognition, because large regions of the face are occluded

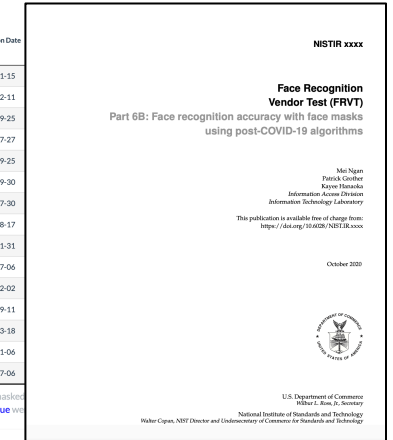
Image Source:
"Coronavirus (COVID-19) Sheffield, UK" by Tim Dennell is licensed under CC BY-NC 2.0.
"Street in Hong Kong during the COVID-19 pandemic.jpg" by Andrii Makukha is licensed under CC BY-SA 4.0.
"N95 mask while visiting a TB case.jpg" by Ashashyou is licensed under CC0 1.0.
"Passengers waiting for a flight to Shanghai in March 2020.jpg" by Mx. Granger is licensed under CC0 1.0.

What we did



Algorithms	VISABORDER Photos FMR @ FMR < 0.0001 (NOT MASKED)	VISABORDER Photos FMR @ FMR < 0.0001 (MASKED: NOSE) lightblue, white, medium coverage	Submission Date
deepfint-002	0.0039 ^[11]	0.0237 ^[1]	2019-11-15
paravision-004	0.0088 ^[27]	0.0281 ^[2]	2019-12-11
xforwardai-001	0.0041 ^[12]	0.0289 ^[3]	2020-09-25
visionlabs-009	0.0028 ^[1]	0.0355 ^[6]	2020-07-27
hr-003	0.0049 ^[21]	0.0419 ^[3]	2020-09-25
fujitsu-lab-001	0.0078 ^[25]	0.0427 ^[6]	2020-09-30
icface-002	0.0086 ^[25]	0.0445 ^[7]	2020-07-30
pensees-001	0.0106 ^[21]	0.0461 ^[3]	2020-08-17
vecord-008	0.0038 ^[3]	0.0500 ^[9]	2020-01-31
idemia-006	0.0048 ^[25]	0.0539 ^[25]	2020-07-06
sensetime-003	0.0045 ^[18]	0.0544 ^[18]	2019-12-02
teclian-006	0.0038 ^[25]	0.0547 ^[12]	2020-09-11
culkee-001	0.0041 ^[12]	0.0572 ^[12]	2020-03-18
visionlabs-008	0.0034 ^[3]	0.0579 ^[14]	2020-01-06
ercacat-001	0.0096 ^[25]	0.0616 ^[12]	2020-07-06

FMR values are reported at a fixed threshold calibrated to give FMR < 0.0001 on unmasked images. Algorithms in **black** were submitted prior to mid-March 2020, and algorithms in **blue** were submitted thereafter.



Applied digital masks to existing test data

Leveraged open-source Dlib face landmark detector + supplemental eye coordinates to draw solid masks on 5.2M photos

Evaluated “pre-COVID-19” algorithms

Published results for 89 algorithms submitted to FRVT 1:1 between April 2019 to mid-March 2020

- Unmasked vs. masked probe
- Impact of mask nose coverage, shape, color

Evaluated “post-COVID-19” algorithms

Results for algorithms submitted to FRVT 1:1 after mid-March 2020

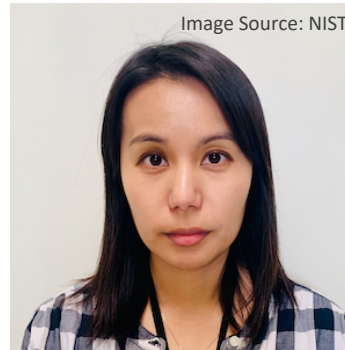
- Single vs. both images masked
- Cumulative results for all algorithms evaluated to date

Public report TO BE PUBLISHED

Datasets

A. Enrollment images:

- 1 019 232 DHS/CIS Application images
- Approximately ISO compliant



B. Verification images:

- 5 225 633 DHS/CBP Entry images



- 3 225 633 genuine comparisons
- 200 000 000 impostor comparisons

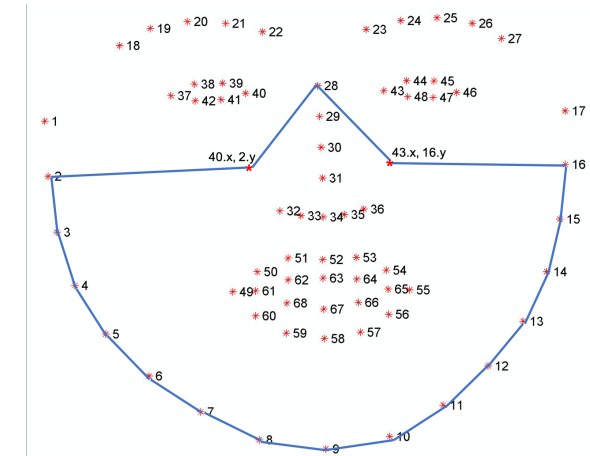
Image Source: NIST MEDS-II Dataset



DHS Logitech webcam images collected in operational settings using the same camera and procedure as the verification photos (examples are very similar, not actual)

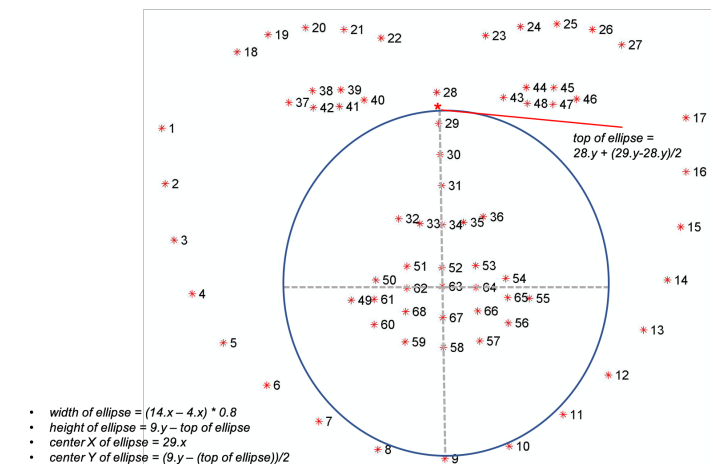
The masking process

1. Open-source Dlib toolkit used to detect face + generate 68 facial landmarks
2. Facial landmarks were used to generate closed solid polygon drawn over the face



Limitations:

- Not real masks
- No texture or patterns
- Digital masks are tailored to faces (real ones tend to be one size fit all)
- No camera-mask interactions (exposure)
- No eye occlusion



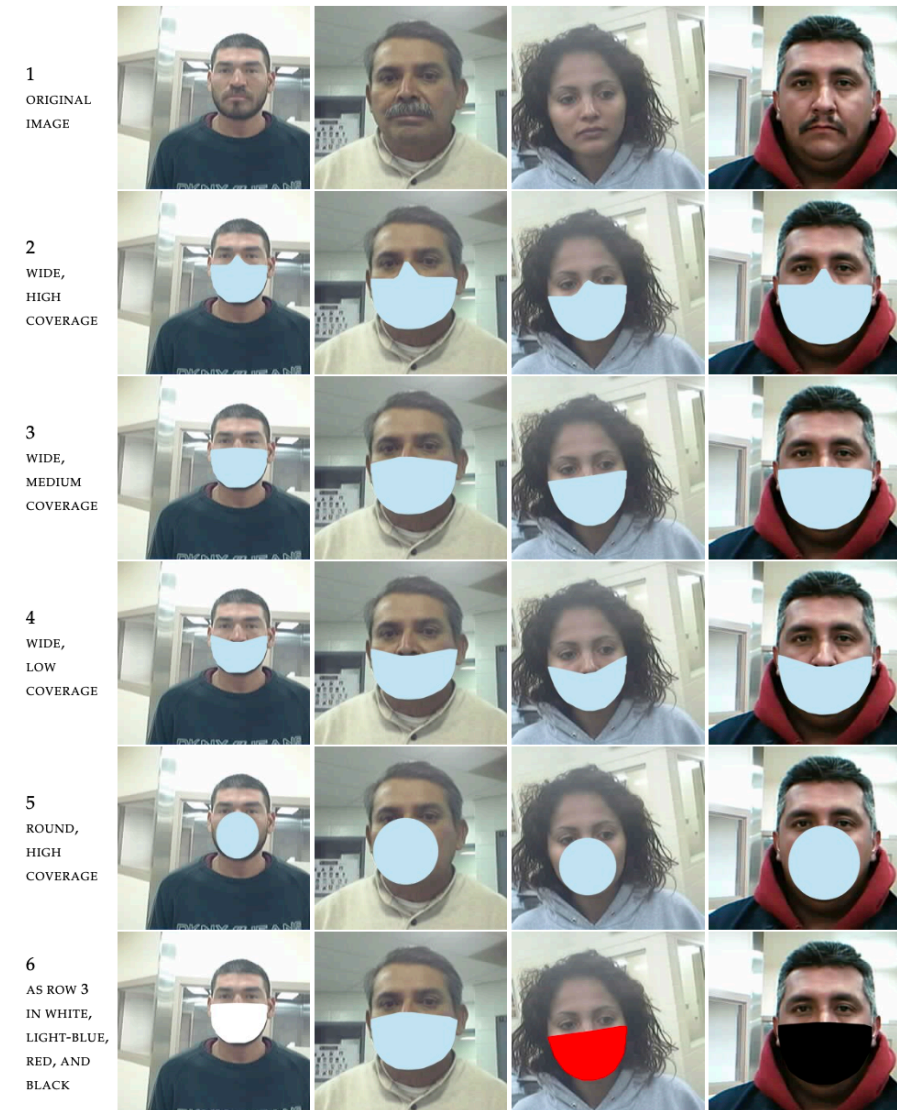
Synthetic masks

NIST varied

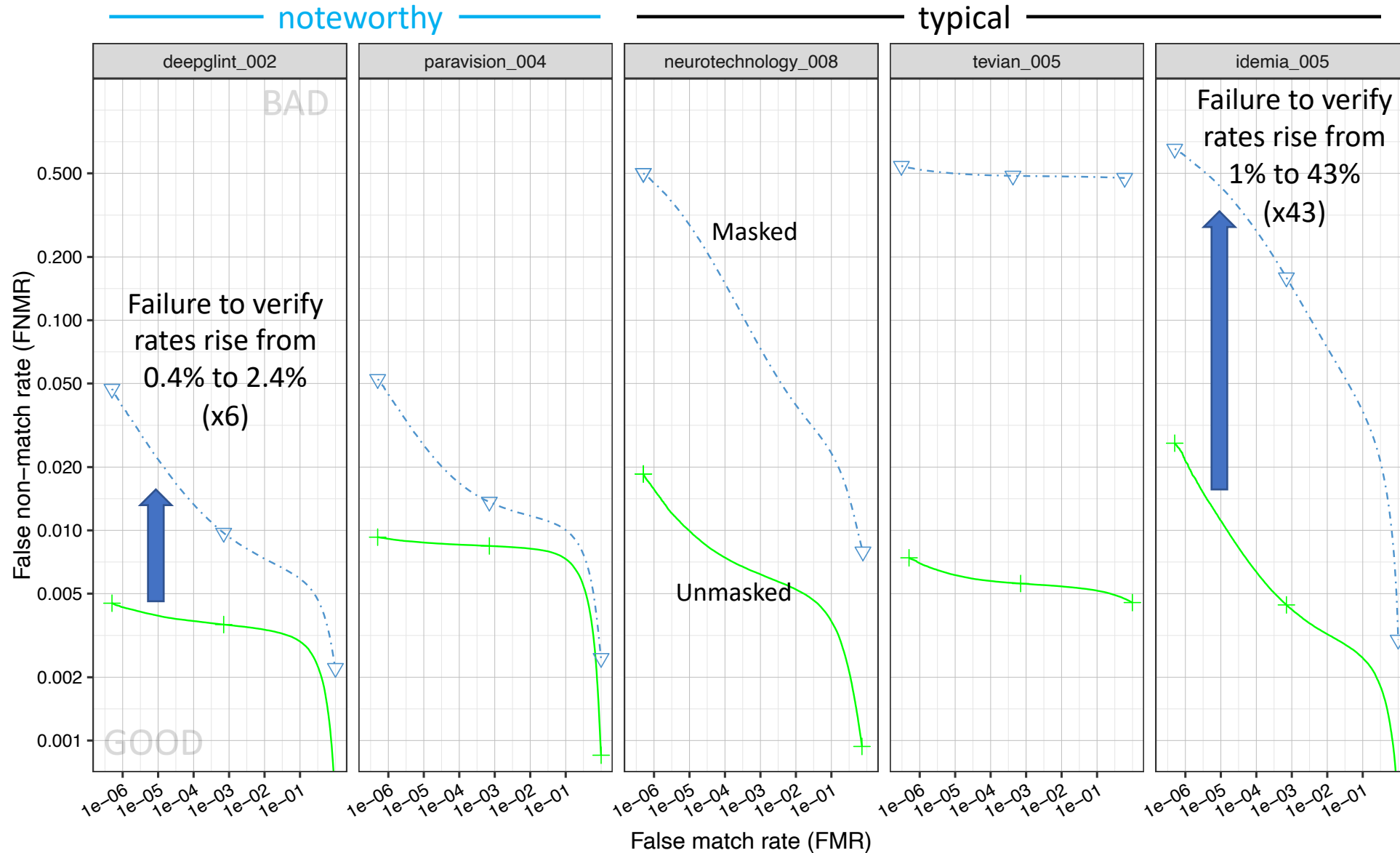
- Nose coverage, shape, color

Positioning

- Relative to landmarks reported by open-source Dlib toolkit
- If Dlib failed, then relative to detected eyes from good FRVT FR algorithms



Accuracy with and without masks: PRIOR to mid-March 2020



Unmasked probe



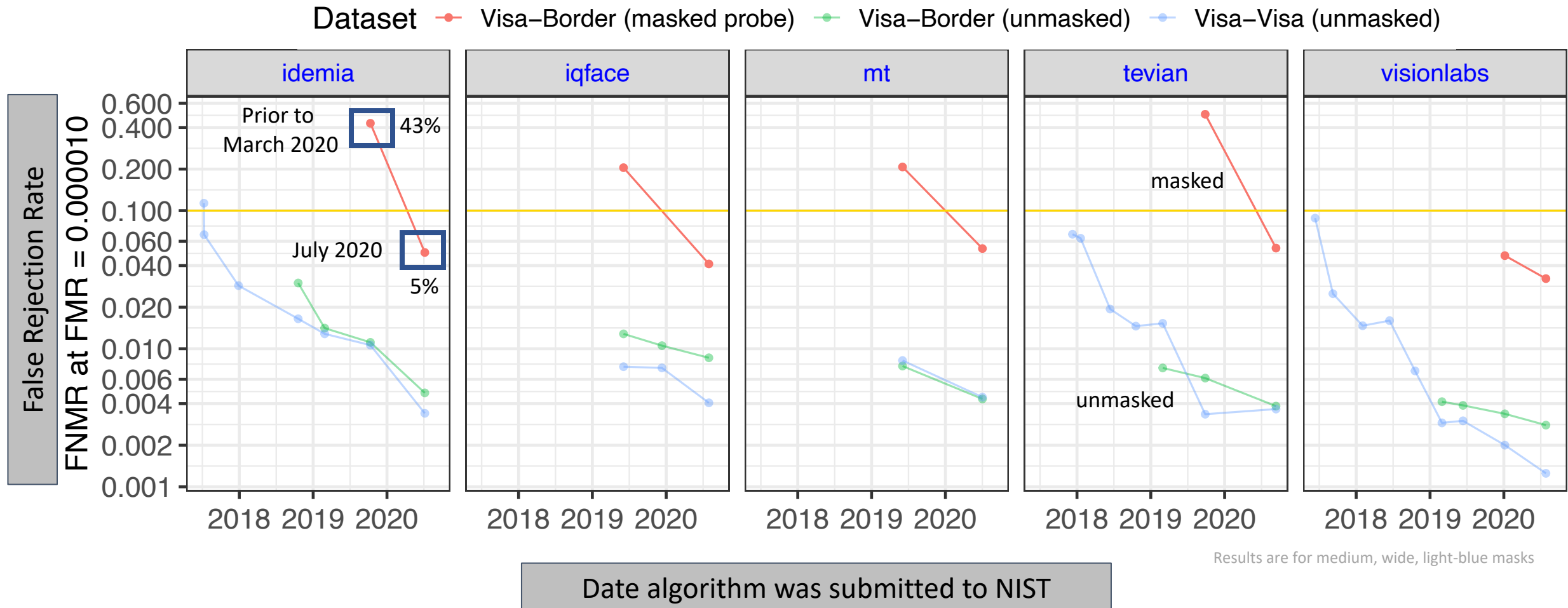
Masked probe
wide, lightblue, medium coverage

Accuracy with and without masks: (AFTER mid-March 2020)

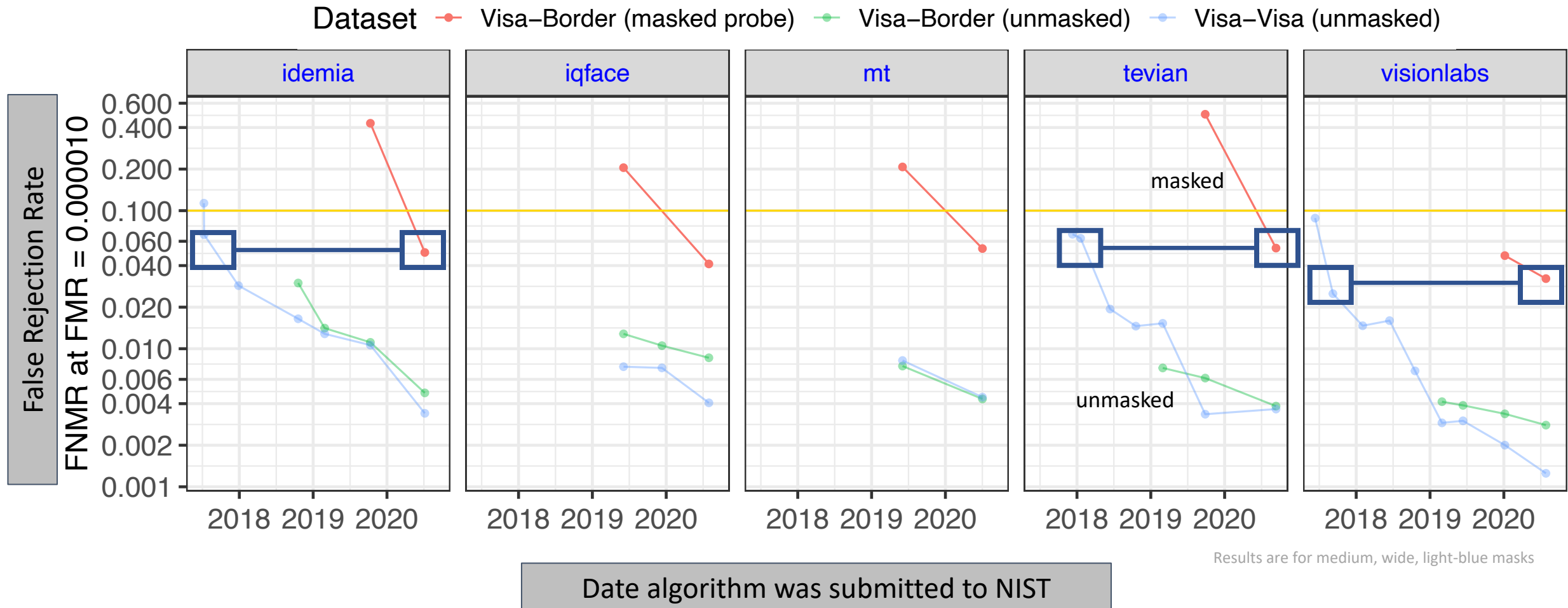
Algorithm	VISABORDER Photos FNMR @ FMR ≤ 0.00001 (NOT MASKED) UNMASKED	VISABORDER Photos FNMR @ FMR ≤ 0.00001 (MASKED PROBE) lightb MASKED average	Submission Date
deepglint-002	0.0039 ⁽¹¹⁾	0.0237 ⁽¹⁾	2019-11-15
paravision-004	0.0088 ⁽⁵⁷⁾	0.0281 ⁽²⁾	2019-12-11
xforwardai-001	0.0041 ⁽¹²⁾	0.0289 ⁽³⁾	2020-09-25
visionlabs-009	0.0028 ⁽¹⁾	0.0355 ⁽⁴⁾	2020-07-27
hr-003	0.0049 ⁽²¹⁾	0.0419 ⁽⁵⁾	2020-09-25
fujitsulab-001	0.0078 ⁽⁵⁰⁾	0.0427 ⁽⁶⁾	2020-09-30
iqface-002	0.0086 ⁽⁵⁵⁾	0.0445 ⁽⁷⁾	2020-07-30
pensees-001	0.0106 ⁽⁷¹⁾	0.0461 ⁽⁸⁾	2020-08-17
vocord-008	0.0038 ⁽⁸⁾	0.0500 ⁽⁹⁾	2020-01-31
idemia-006	0.0048 ⁽²⁰⁾	0.0539 ⁽¹⁰⁾	2020-07-06
sensetime-003	0.0045 ⁽¹⁹⁾	0.0544 ⁽¹¹⁾	2019-12-02
tevia-006	0.0038 ⁽¹⁰⁾	0.0547 ⁽¹²⁾	2020-09-11
cuhkee-001	0.0041 ⁽¹³⁾	0.0572 ⁽¹³⁾	2020-03-18
visionlabs-008	0.0034 ⁽³⁾	0.0579 ⁽¹⁴⁾	2020-01-06
ercacat-001	0.0096 ⁽⁶⁵⁾	0.0616 ⁽¹⁵⁾	2020-07-06

FNMR values are reported at a fixed threshold calibrated to give FMR = 0.00001 on unmasked images. Algorithms in black were submitted prior to mid-March 2020, and algorithms in blue were submitted thereafter.

Masks: Developer evolution



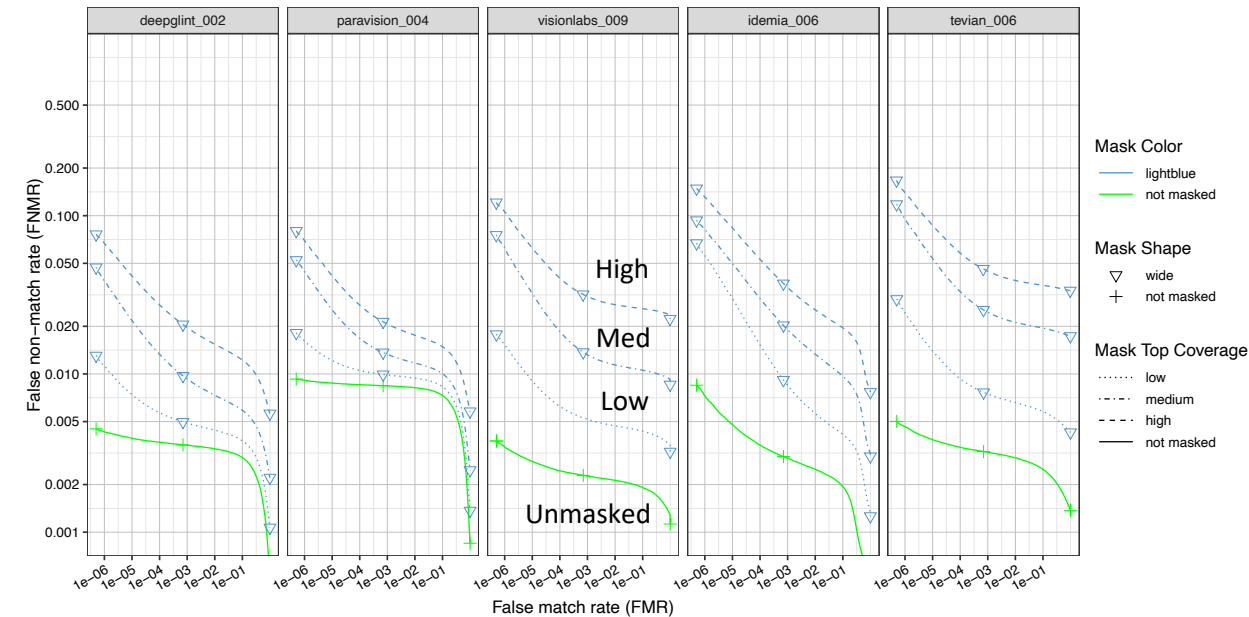
Masks: How much of a setback?



The presence of masks has set face recognition technology back 2-3 years

Mask coverage + shape matters

“High” masks generate more errors than “low” masks

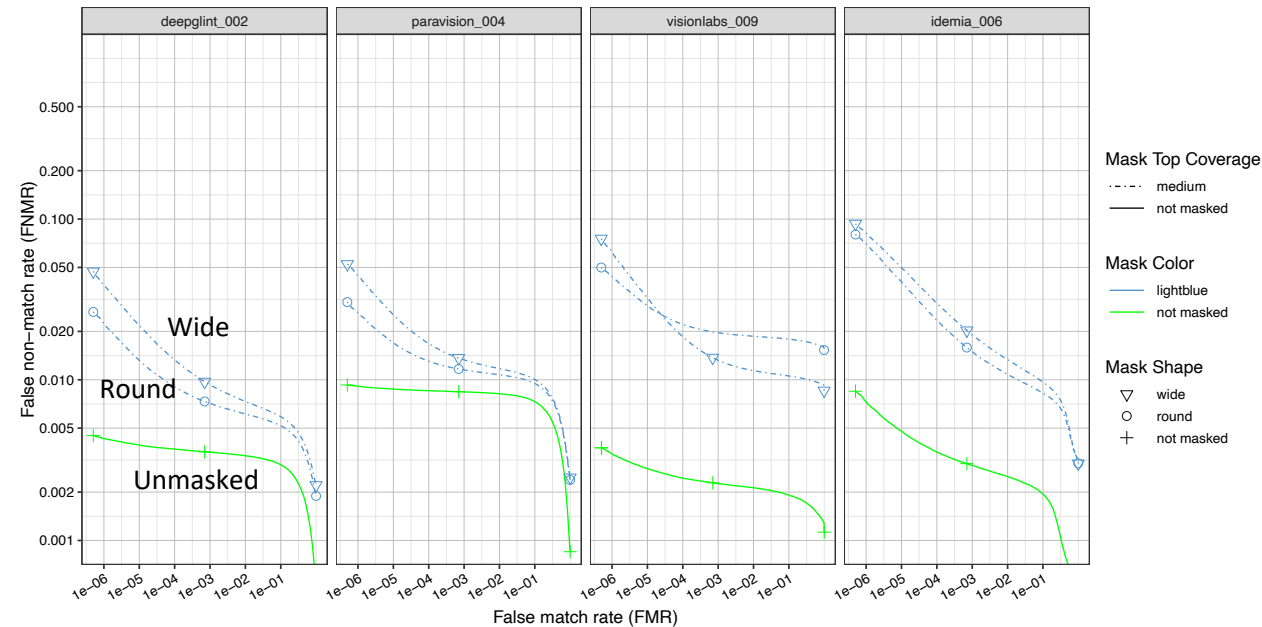


High

Medium

Low

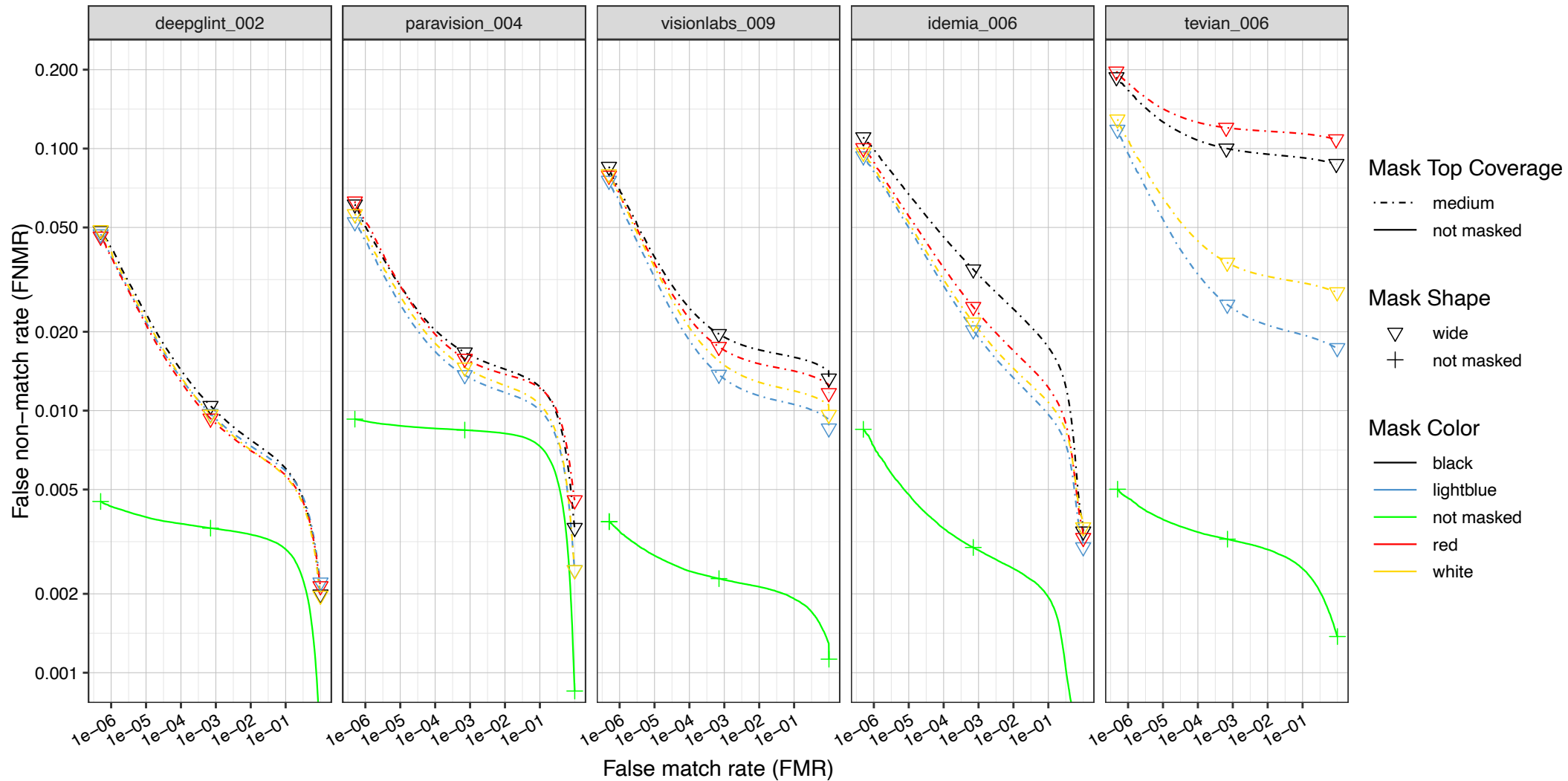
“Wide” masks generate more errors than “round” masks



Wide

Round

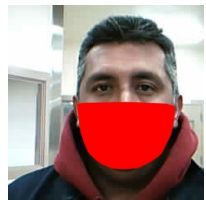
Mask color matters



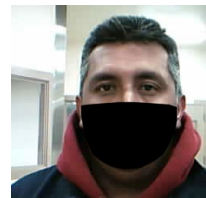
White



Lightblue

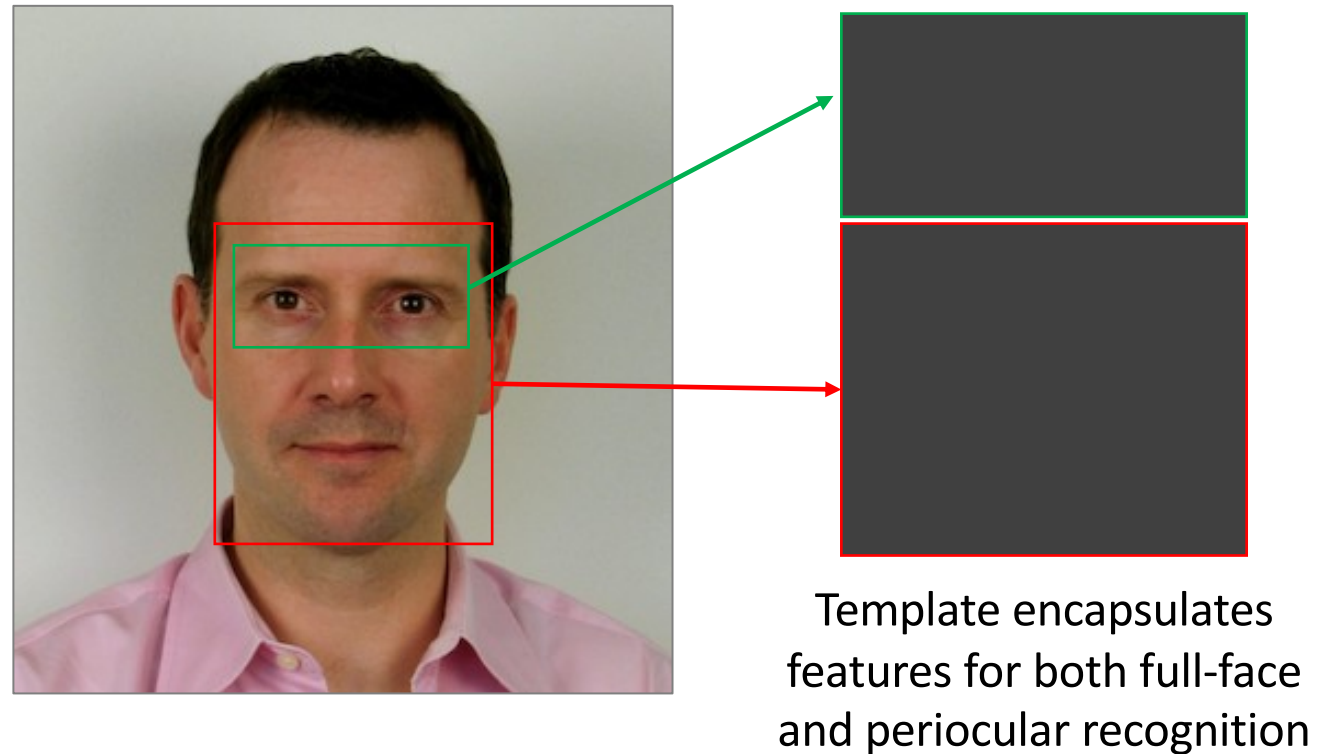
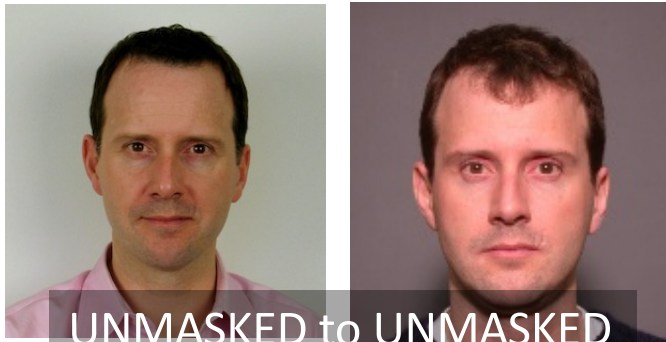


Red

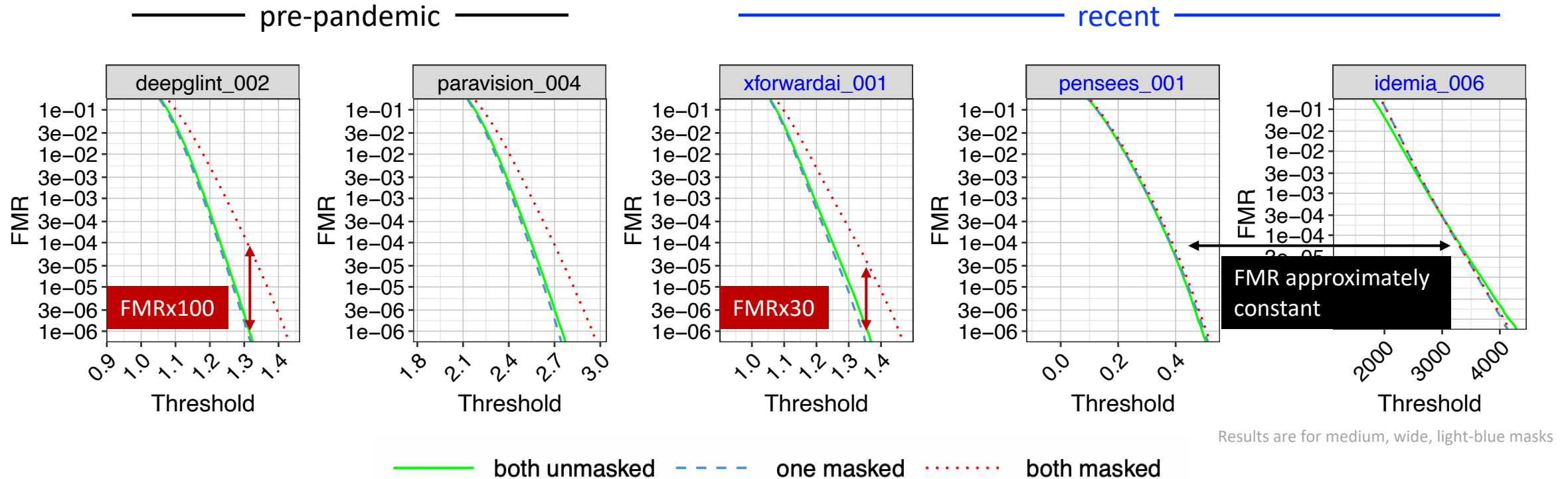


Black

Mask-agnostic face recognition: Some transactions involve a mask (and some do not):



False Match Rates



RESULTS

- Face masks have set face recognition accuracy ~2-3 years back
 - But some algorithms may still be usable, depending on the application
- Better accuracy with
 - Low masks (vs. high)
 - Round masks (vs. wide)
 - Light colors (vs. black)
- Elevated false match rates when both enrollment and verification images are masked
 - With the exception of idemia_006 and pensees_001

LIMITATIONS/CONTEXT

- No real masks, no textured masks
- No eye-occlusion
- No camera-mask interactions (exposure)

WHAT'S NEXT

- Report #2: Face recognition accuracy with face masks using “post-pandemic” algorithms (Nov. 2020 est.)
- Report #3: Evaluation of 1:N algorithms with masks (1Q 2021 est.)

Thank you!

Mei Ngan

National Institute of Standards and Technology (NIST)

mei@nist.gov | frvt@nist.gov