NIST FRVT



Craig Watson
Image Group Manager
craig.watson@nist.gov



Face Recognition Vendor Test (FRVT)

1. FRVT 1:1
Core Biometric Operation

2. FRVT 1:N
Search Performance

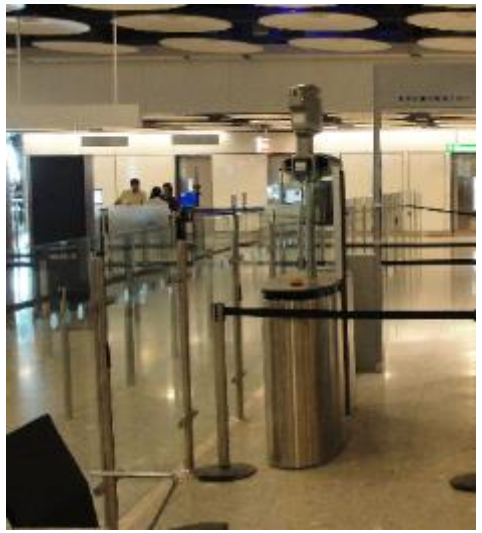
3. FRVT Q
Automated Quality Assessment

4. FRVT Morph
Detection

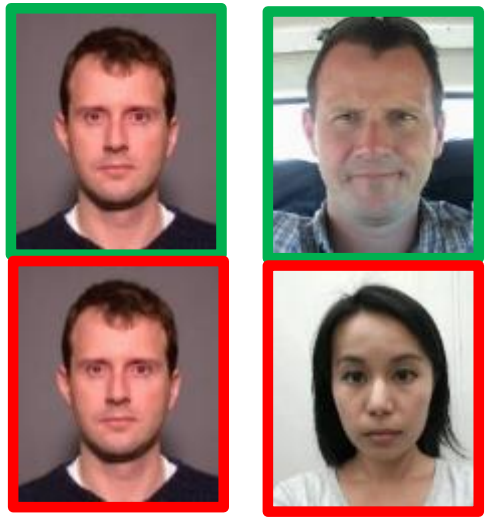
FRVT Demographic Effects

Verification of ID (1-to-1, Threshold)





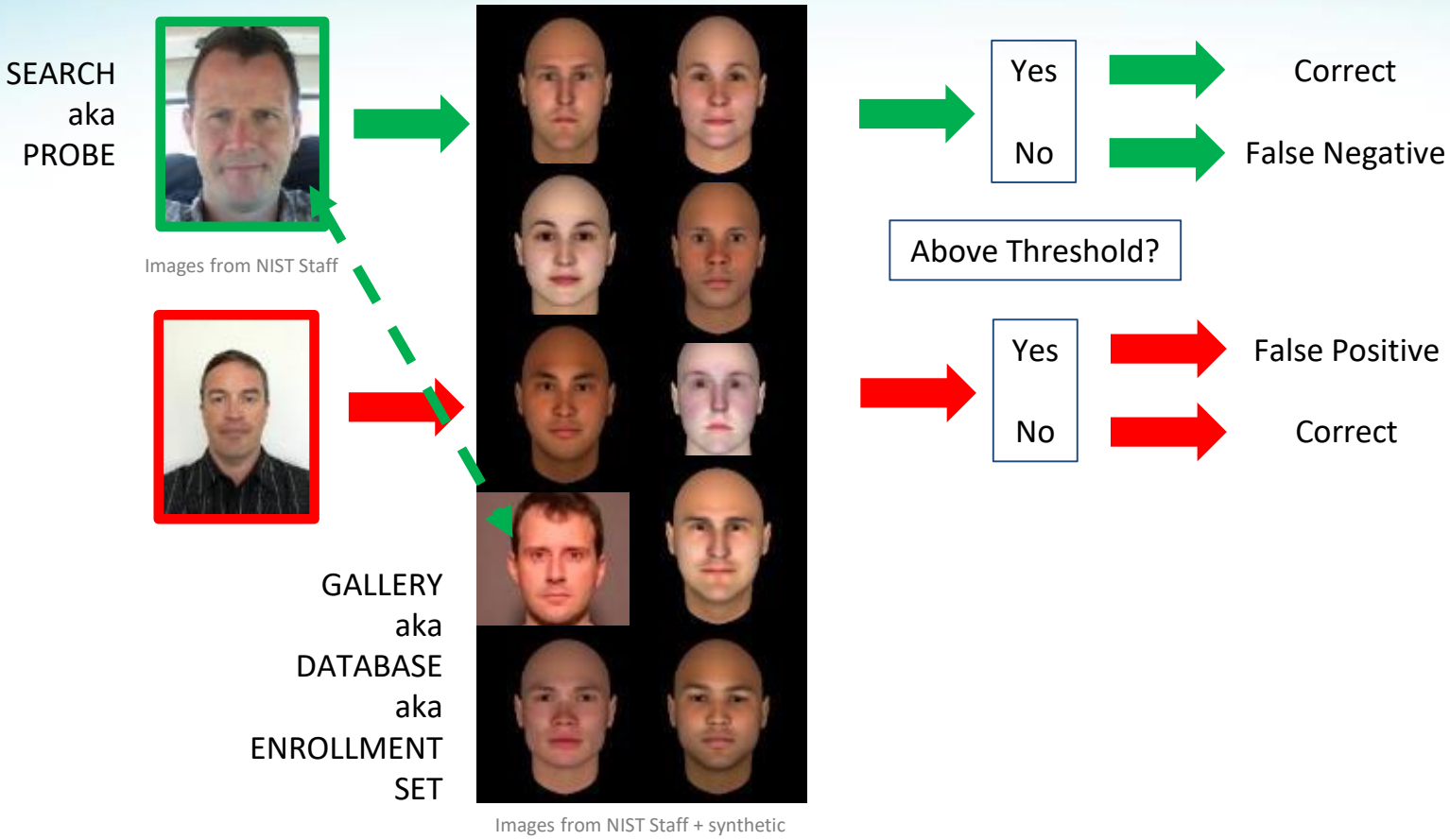
https://en.wikipedia.org/wiki/EPassport_gates
CC BY 2.0. File:Heathrow Terminal 5 ePassport gates.jpg
Created: 16 July 2010



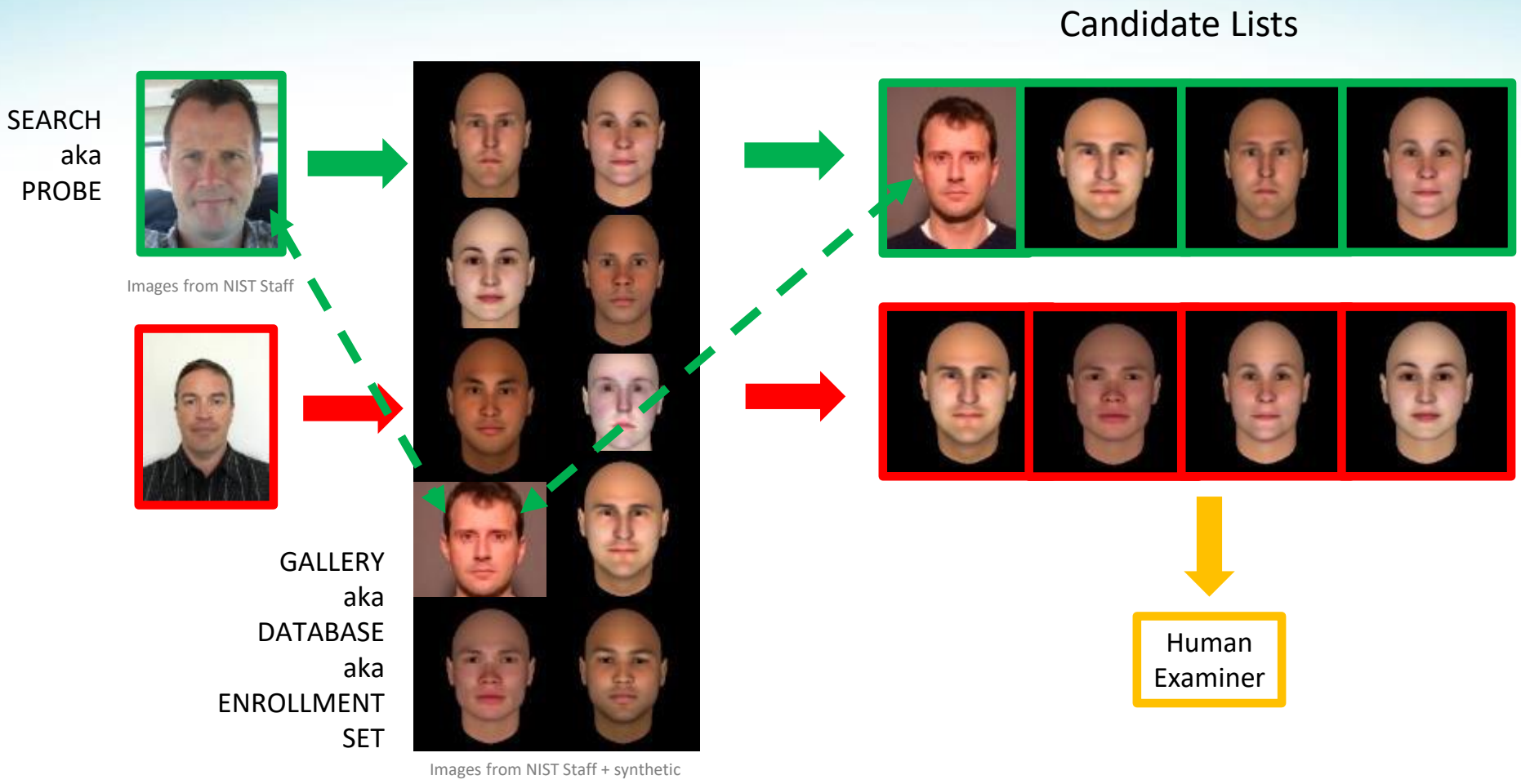
Above Threshold?

Yes	Correct	
No	False Negative	
Yes	False Positive	
No	Correct	

Identification (1-to-N, Threshold)



Investigation (Ranked List, Human Examiner)



Face Recognition Vendor Test

Current Capabilities

Improvement since 2010

NEC
2010 FRVT
Rank 1 hit rate
~86.0%

NEC
2013 FRVT
Rank 1 hit rate
95.9%

NEC
2018 FRVT
Rank 1 hit rate
99.7%



2010 2013 2015 2017 2018

Face Recognition Vendor Test

Current Capabilities - Twins

Identical




Fraternal




Images from U. Notre Dame Twins Day Database

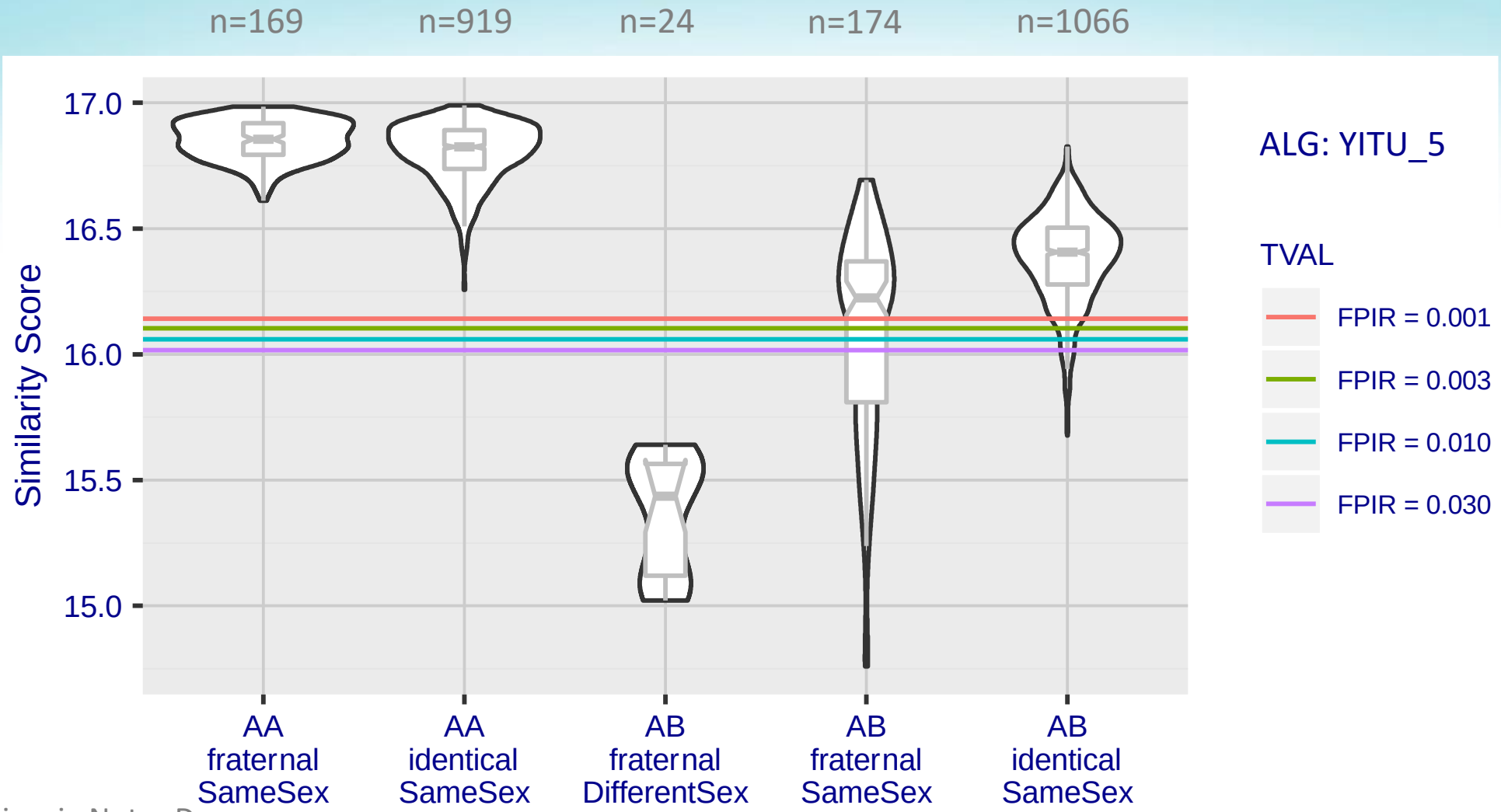
Twins: A demographic with very high FPIR

Gallery:

- N = 640 000 portraits, and
- 104 images of Twin A

Probes:

- Second images of Twin A
- Images of Twin B



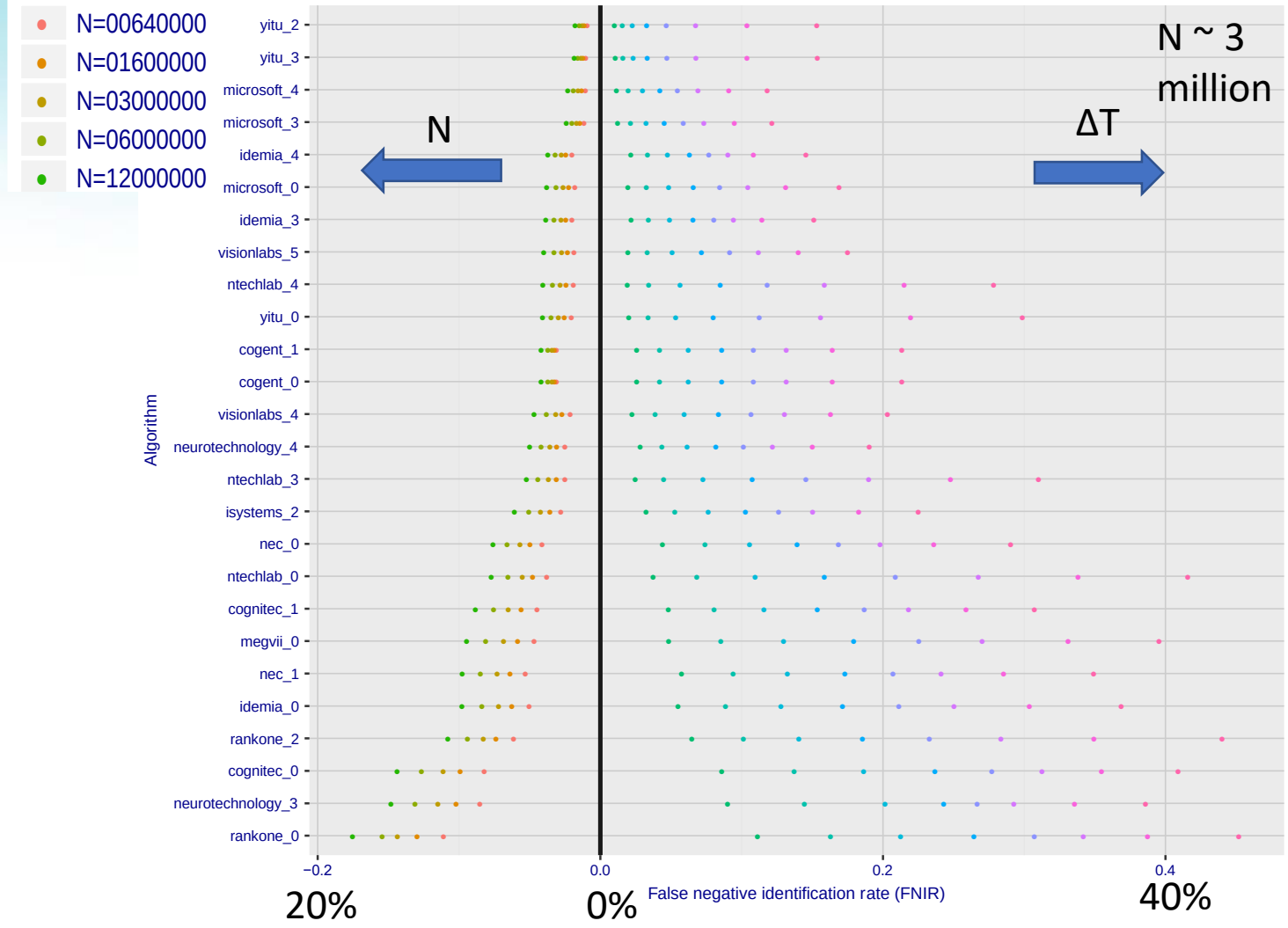
Images: Twins Day, Ohio, via Notre Dame
Algorithms: FRVT 1:N November 2018.

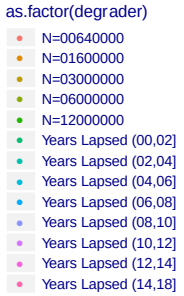
Face Recognition Vendor Test

Current Capabilities - Ageing



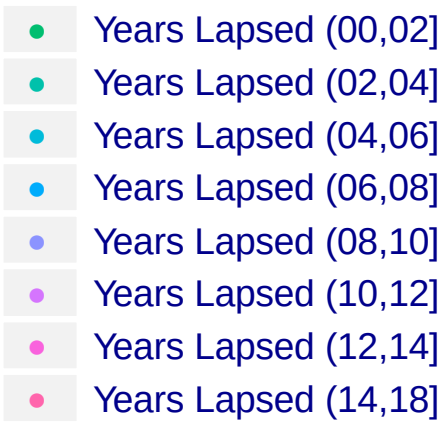
Images from NIST Staff





Performance in perspective: What matters more?

1. Algorithm
2. Population size
3. Ageing



Face Recognition Vendor Test

Quality



Identification (1-to-N, Threshold)

False Negative Rates

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

Table 3: Impact of image quality on accuracy.

Mugshot – Standard Quality Images, Constrained

Webcam – Non-Standard, Less-Constrained

Wild – Non-Standard, Unconstrained

Face Recognition Standardization: New Work

Quality standards

- Quality Scalar Standard — **ISO/IEC 24357 Approved 2019-08**
- Quality Vector Standard — **ISO/IEC 29794-5 REV Approved 2019-07**
- Next Generation Capture Subsystem Standard — **ISO/IEC 24358 Approved 2019-08**
- Quality Performance Testing Standard — Baseline metrics included in ISO/IEC 29794-1 Framework revision

ISO JTC 1 IEC
INFORMATION TECHNOLOGY STANDARDS

SC37 BIOMETRICS

Face Recognition Vendor Test

Morphing



Source: NIST Staff

Morphing: An attack on ID credentials





+







- Modern face recognition algorithms will authenticate both subjects against morph on ID credential
- Multiple people can now use the same fraudulent yet genuine document

Single-Image Morph Detection

Morph detection with single image in isolation (e.g. initial passport application)





Goal: Determine whether image being submitted is a morph or not.

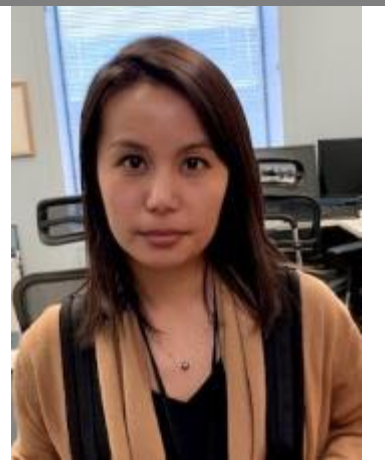
Two-Image Differential Morph Detection

Morph detection with additional live capture image (e.g. morph on passport + eGate webcam photo)

Morphed image is contained in a passport



This image represents a live capture during an eGate border crossing, say.



Goal: Determine whether image on passport is morphed by using the additional information available in the live capture image.

Face Recognition Vendor Test
Demographic Effects

- False Positives larger than False Negatives
- Demographic effects vary depending on application, data, and algorithm
- Impact of differentials vary application
- Most accurate algorithms have lower demographic differentials
- False Positives increase in oldest and youngest age groups
- False Positives in Females generally greater than Males
- False Positives greater in East/West African and East Asian than Eastern European
- Verification 1-to-1: False Positive variations of 10 to 100 times increase for Black/Asian/American Indian versus White
- Identification 1-to-Many: Some developers so False Positives are undetectable

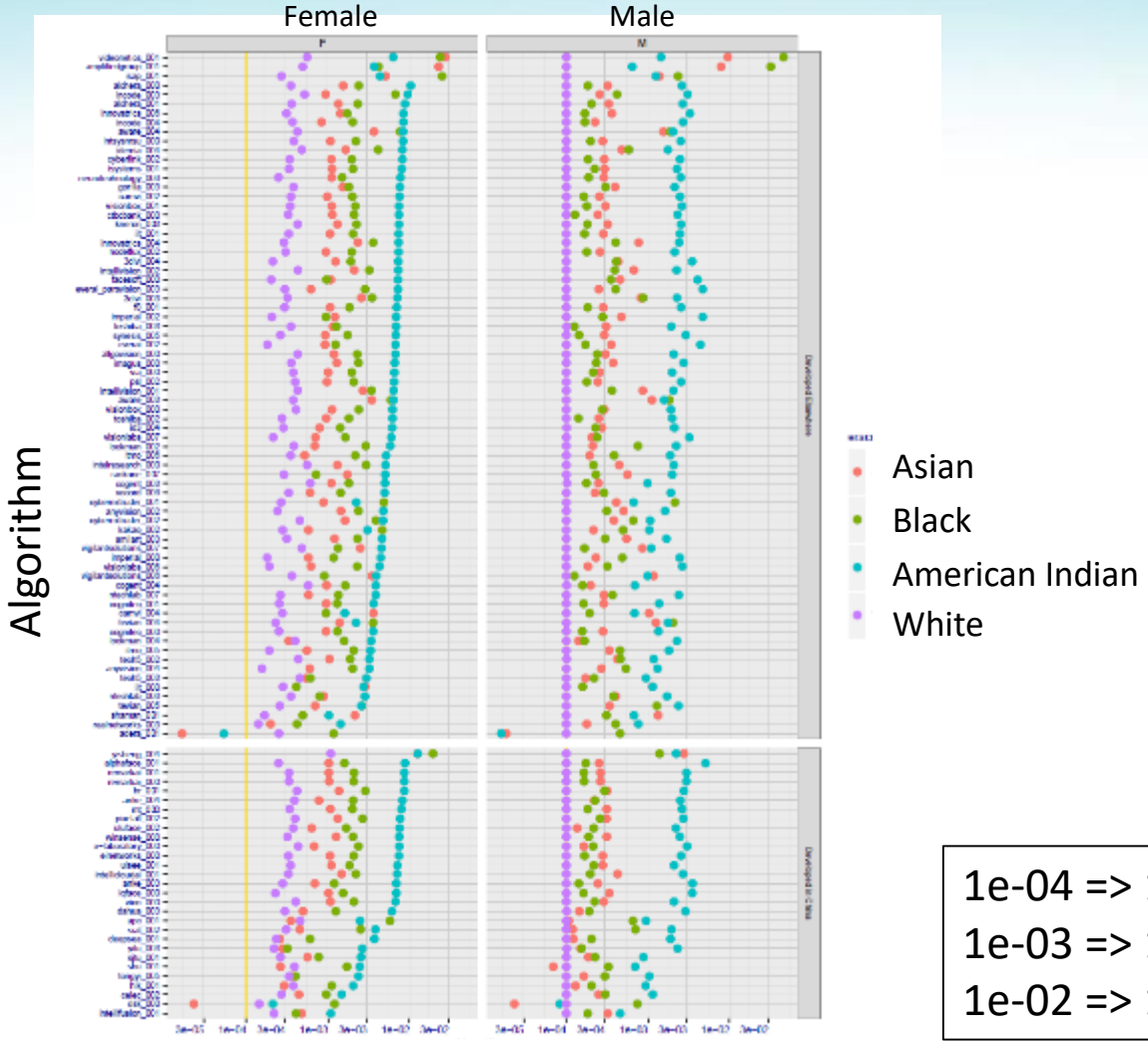
NIST FRVT



Craig Watson
Image Group Manager
craig.watson@nist.gov

Verification (1-to-1)

Mugshot Data



1e-04 => 1 in 10,000
1e-03 => 1 in 1,000
1e-02 => 1in 100

Verification (1-to-1)

Applicant Data

7 – East Asia

6 – South Asia

5 – East Africa

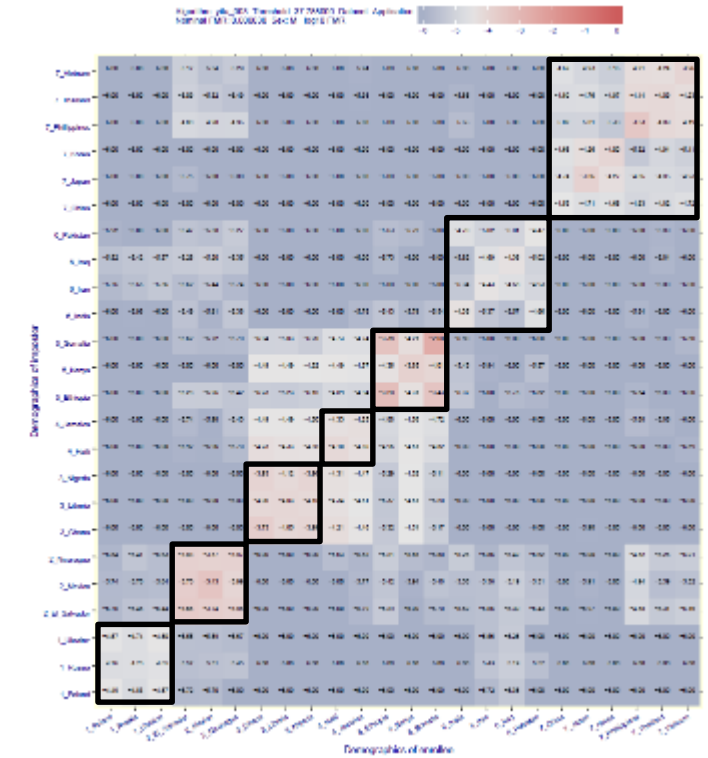
4 – The Caribbean

3 – West Africa

2 – Central America

1 – Eastern Europe

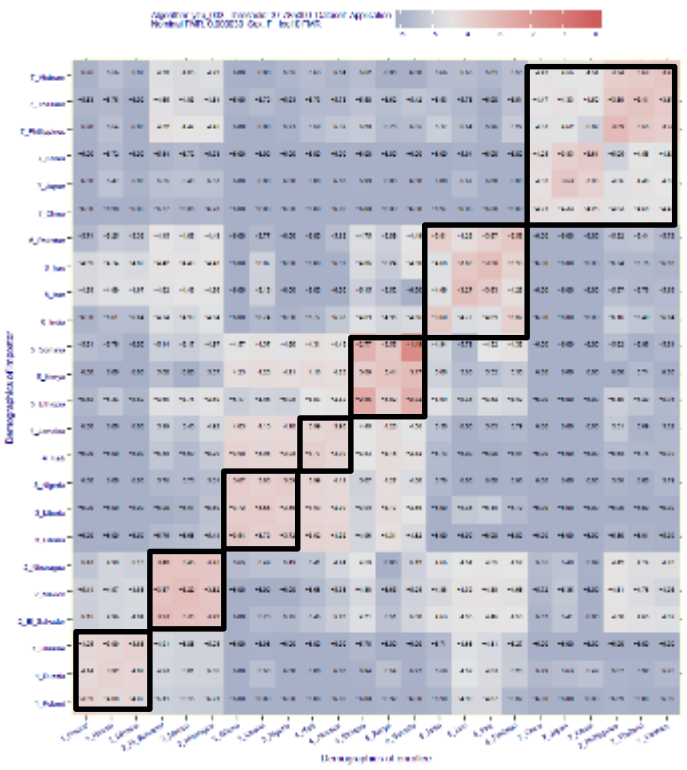
Male



7 6 5 4 3 2 1

1 2 3 4 5 6 7

Female



1 2 3 4 5 6 7

Verification (1-to-1)

Applicant Data

7 – East Asia

6 – South Asia

5 – East Africa

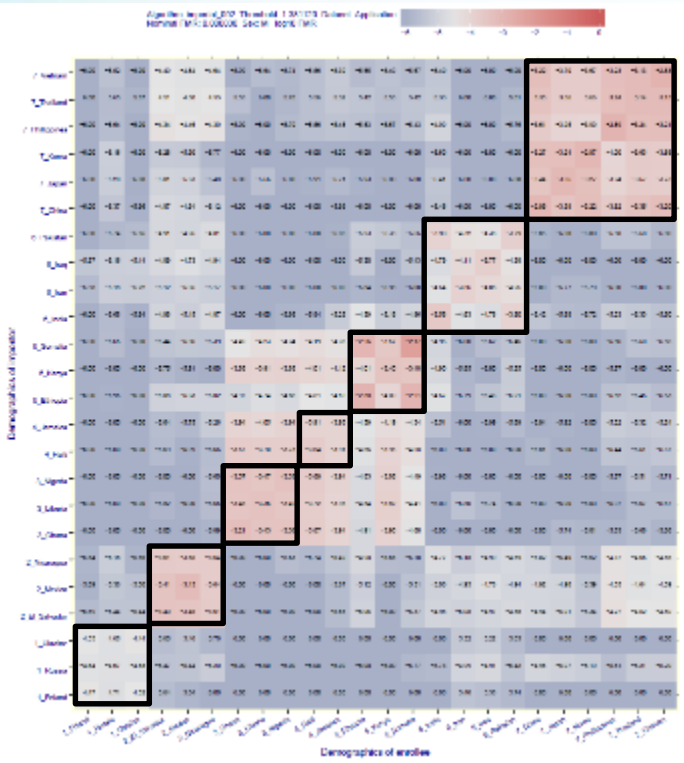
4 – The Caribbean

3 – West Africa

2 – Central America

1 – Eastern Europe

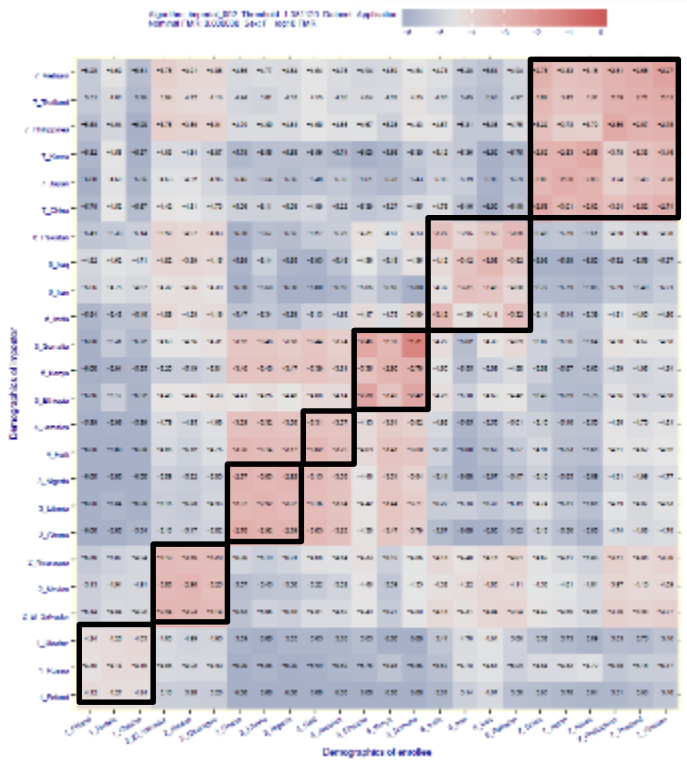
Male



7 6 5 4 3 2 1

1 2 3 4 5 6 7

Female



1 2 3 4 5 6 7