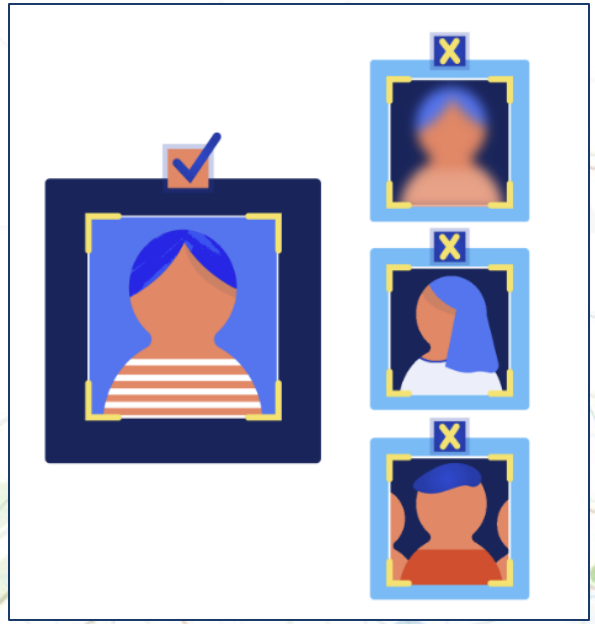
NIST FATE Quality Specific Image Defect Detection (SIDD)



Joyce Yang-NIST Image Group
IFPC 2025: Virtual Presentation, April 1 2025

FRTE/FATE Team: Patrick Grother, Mei Ngan, Kayee Hanaoka, Austin Hom, Joyce Yang, Jim Matey

NIST Face Benchmarks

FRTE
FACE **RECOGNITION**
TECHNOLOGY EVALUATION

RECOGNITION: **WHO** IS IN AN IMAGE

1:1 VERIFICATION	SAME PERSON OR NOT?
1:N SEARCH	WHO? WHERE? WHEN?
FACE IN VIDEO 2024	1:N ON NON-COOP PEOPLE
TWINS DISAMBIGUATION	SAME PERSON, OR TWIN?

FATE
FACE **ANALYSIS**
TECHNOLOGY EVALUATION

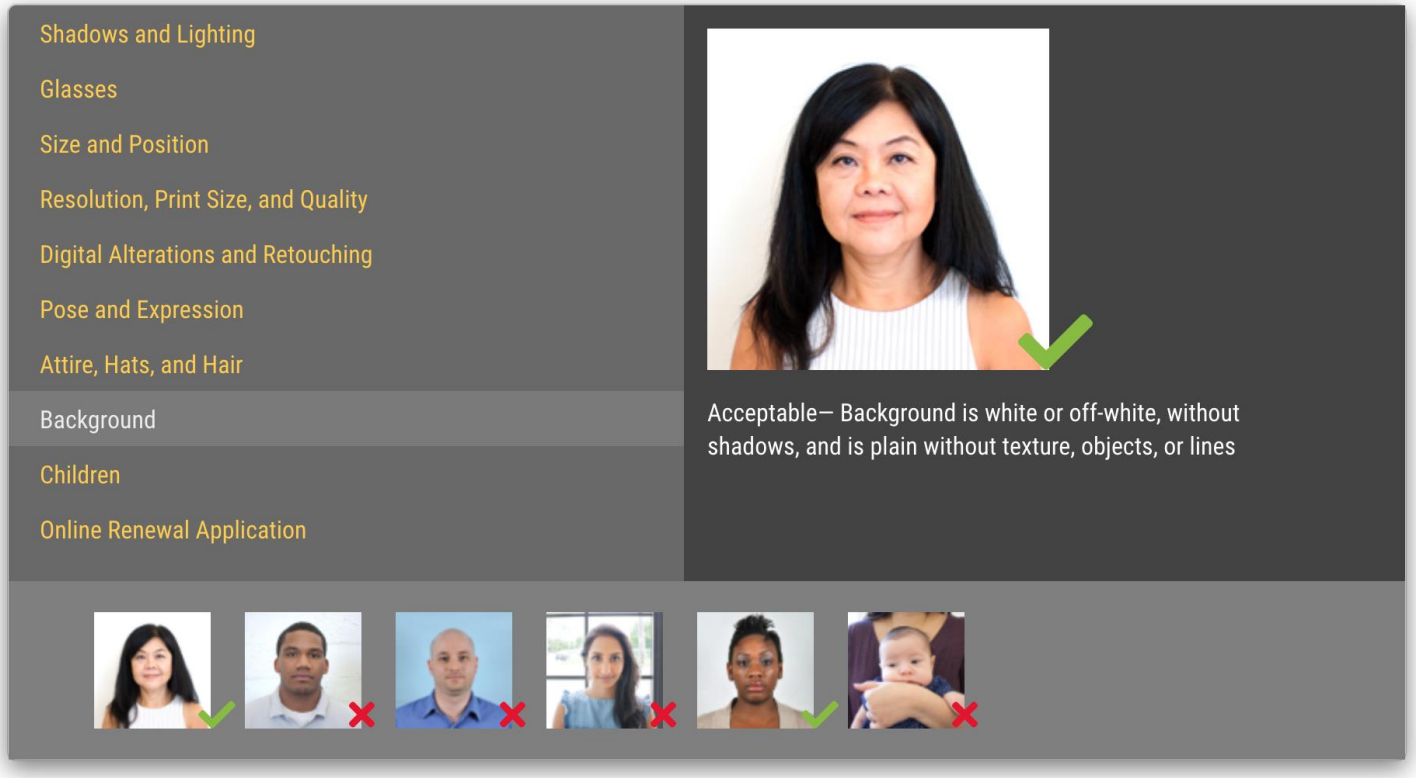
ANALYSIS: **WHAT** ABOUT AN IMAGE

MORPH DETECTION	TWO PEOPLE IN ONE FACE?
QUALITY + DIAGNOSTICS	**HOW IS THIS PHOTO BAD?**
PAD	SUBVERSIVE PHOTO?
AGE ESTIMATION	HOW OLD? OLD ENOUGH?

FRTE/FATE evaluations are:

Why Specific Image Defect Detection?

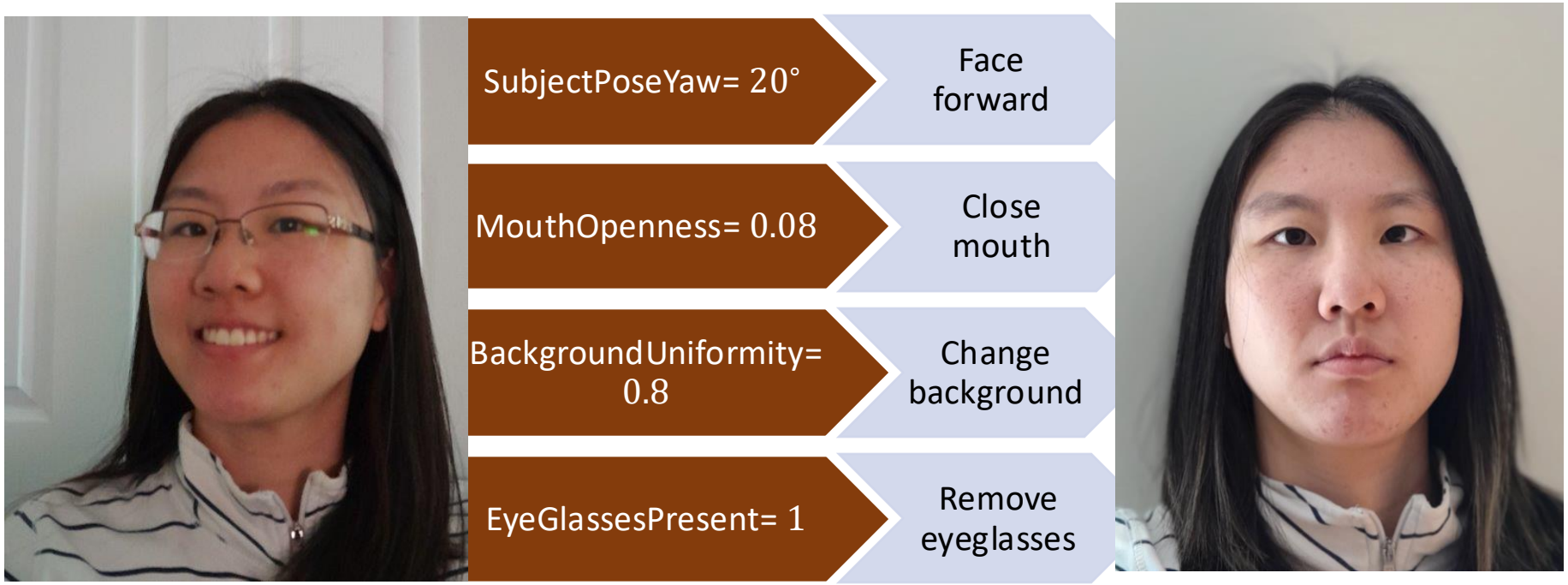
- Automated face recognition is sensitive to quality problems in images
- Standards set requirements on images
  - For MRTD, ISO/IEC:39794-5
  - For law enforcement, ANSI/NIST standard
- SIDD gives users a set of tools to check those requirements



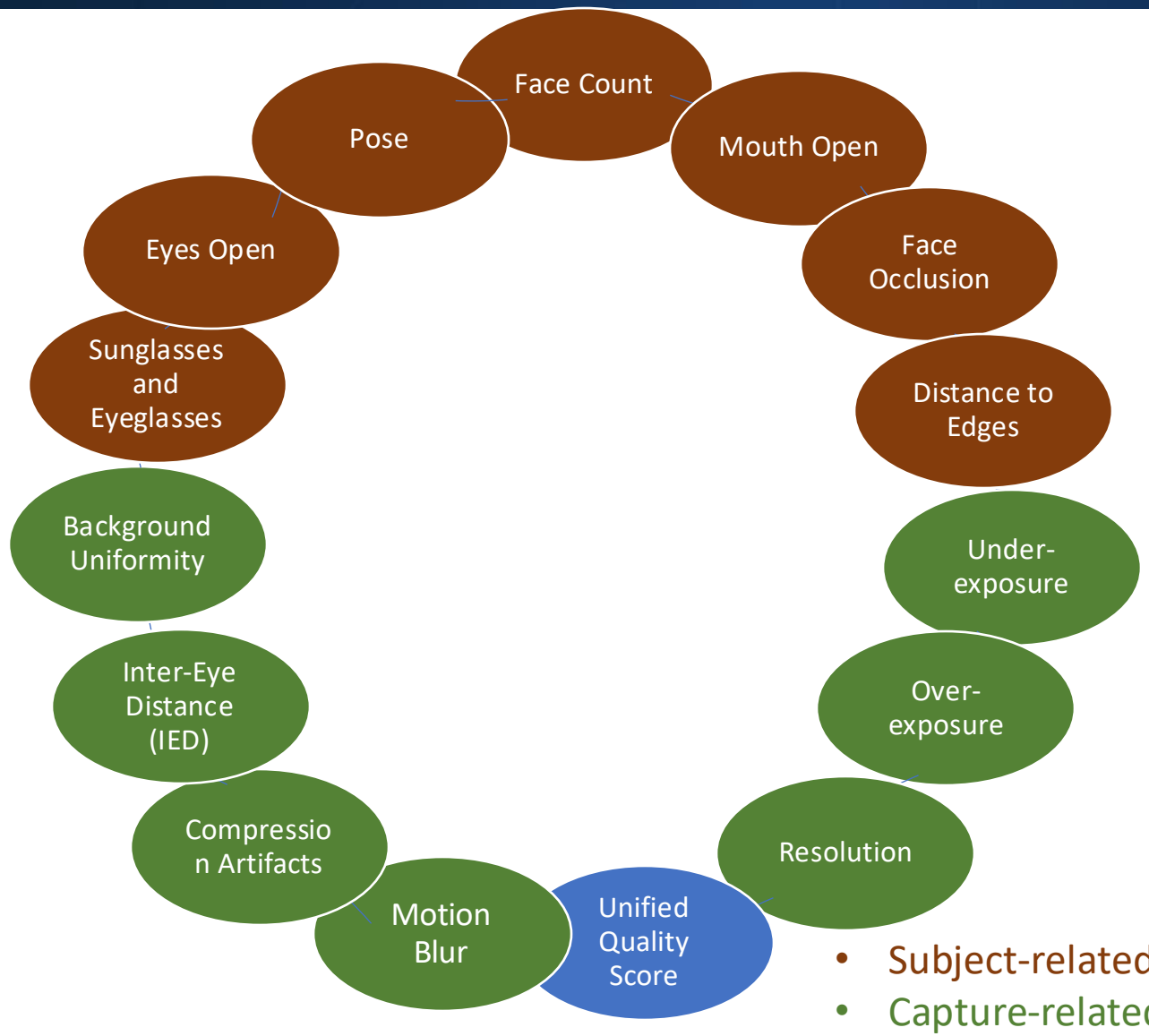
US Passport Photo Guidelines from
https://travel.state.gov/content/travel/en/passports/how-apply/photos.html

Why SIDD? Example: Passport Photo

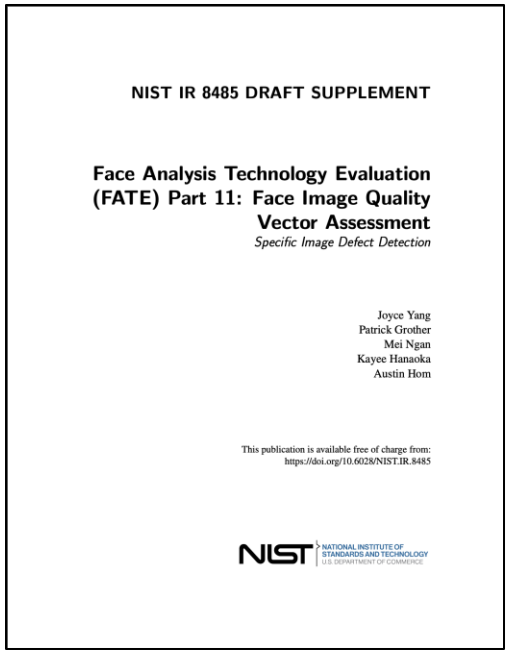
Face capture software should be able to point out an issue such as the mouth being open and give feedback in real time. A SIDD algorithm might provide the output below.



FATE SIDD At A Glance



- Subject-related
- Capture-related
- Other



NIST IR 8485 DRAFT SUPPLEMENT

Face Analysis Technology Evaluation (FATE) Part 11: Face Image Quality Vector Assessment

Specific Image Defect Detection

Joyce Yang
Patrick Grother
Mei Ngan
Kayee Hanaoka
Austin Hom

This publication is available free of charge from:
https://doi.org/10.6028/NIST.IR.8485

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

The two main goals of SIDD are to

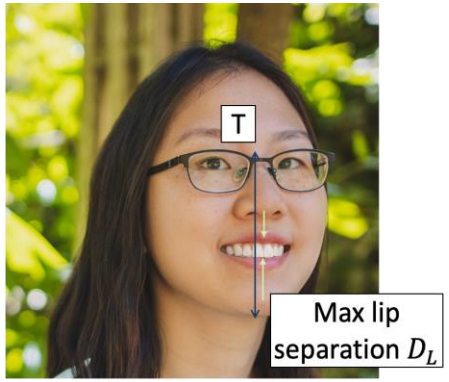
1) Promote quality defect detection algorithm development and
2) Support ISO/IEC 29794-5 on face image quality.

Participation By Year

Year	Number of Algorithms	Unique Developers
2022	2	2
2023	15	9
2024	23	9
2025	1	1

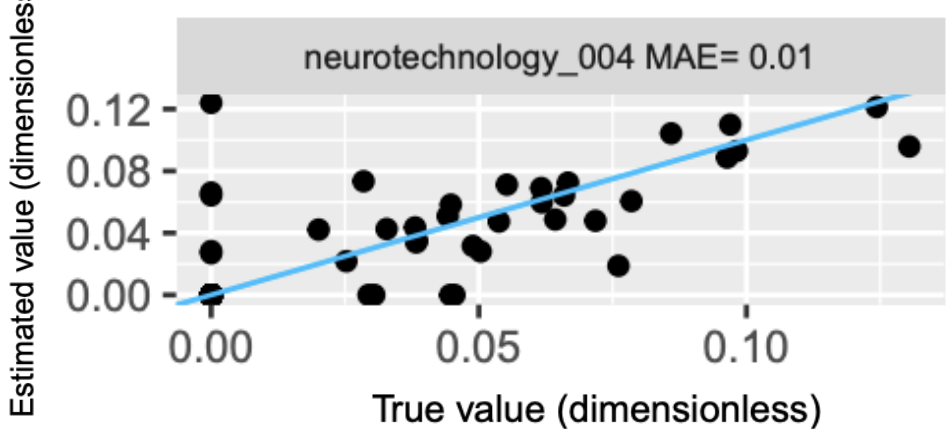
Subject-Related Quality Measure Example: Mouth Openness

$$D_{mouthOpen} = \frac{D_L}{T}$$



The Mouth Openness measure, MouthOpen2, is defined as the (vertical) maximum lip separation normalized by the chin-to-eyes distance.

Estimated vs. Known MouthOpenness



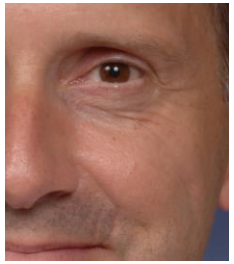
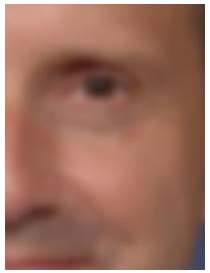
Mean Absolute Error (MAE) is calculated over all images.

The blue line y=x represents perfect performance.

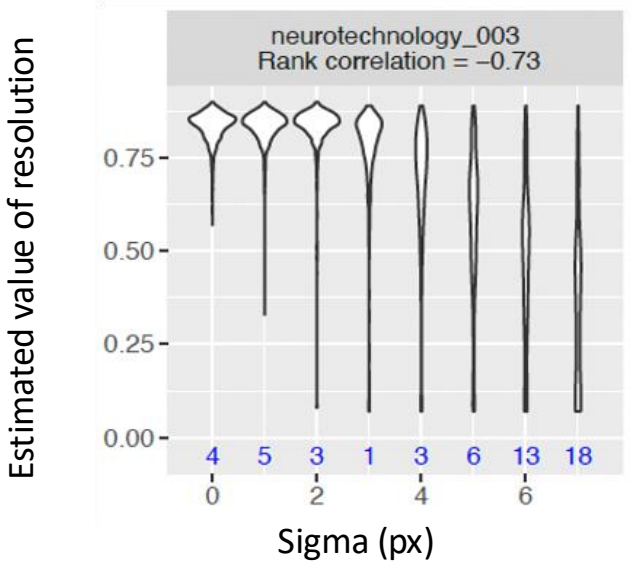
Algorithm	MAE (dimensionless)
neurotechnology_005	0.00
neurotechnology_004	0.01
mobbl_002	0.01
igd_003	0.01
secunet_003	0.01

Neurotechnology_005 performs best based on mean absolute error.

Capture-Related Quality Measure Example: Resolution

Standard deviation σ	0	4	7
Result of convert – gaussian-blur 0xσ			

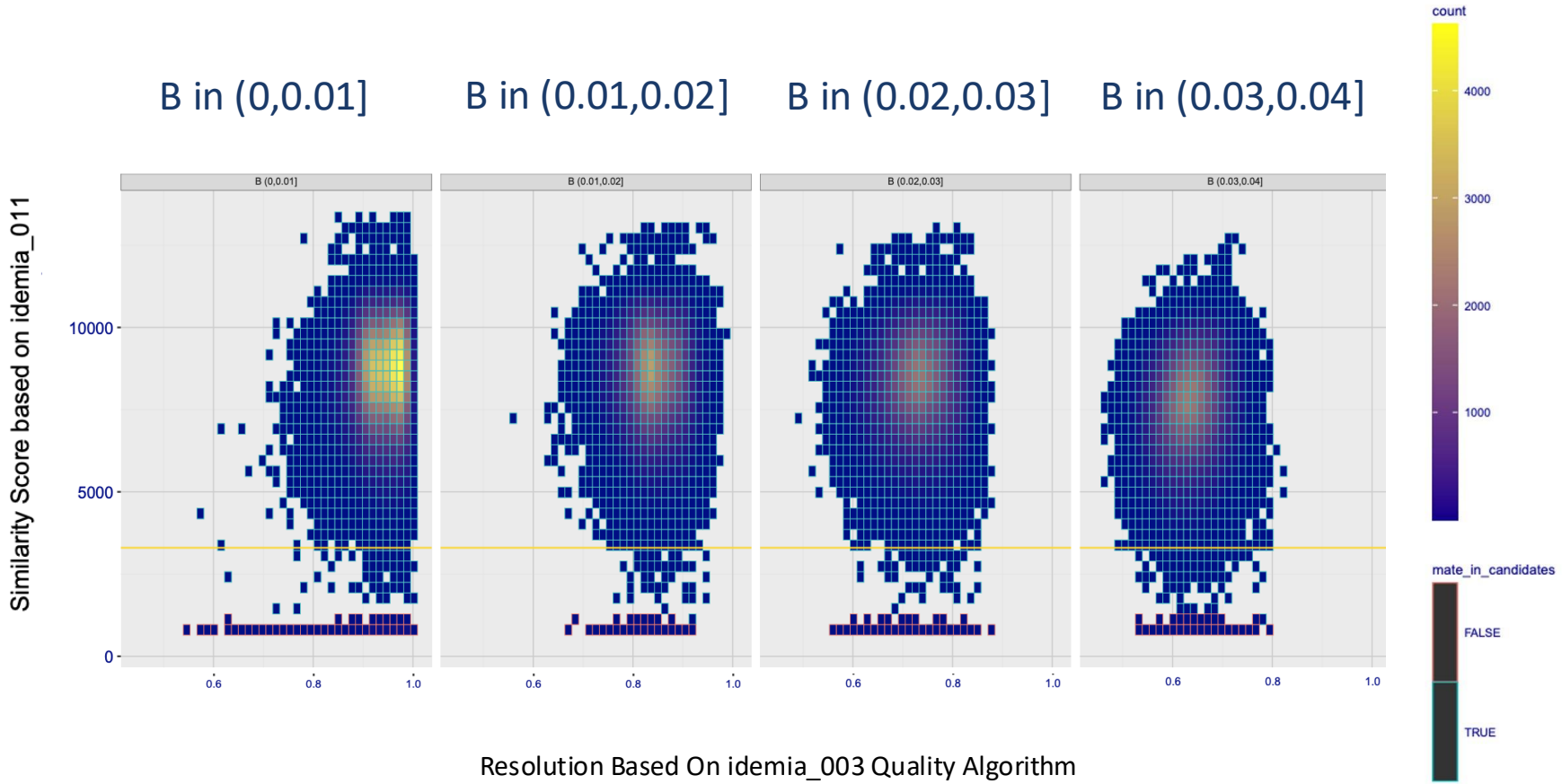
The resolution sets are made by applying synthetic blur. Higher values of sigma correspond to lower resolution.



We calculate rank correlation over all images. Perfect performance corresponds to a value of –1.

Preliminary Study of Resolution Effect on Recognition

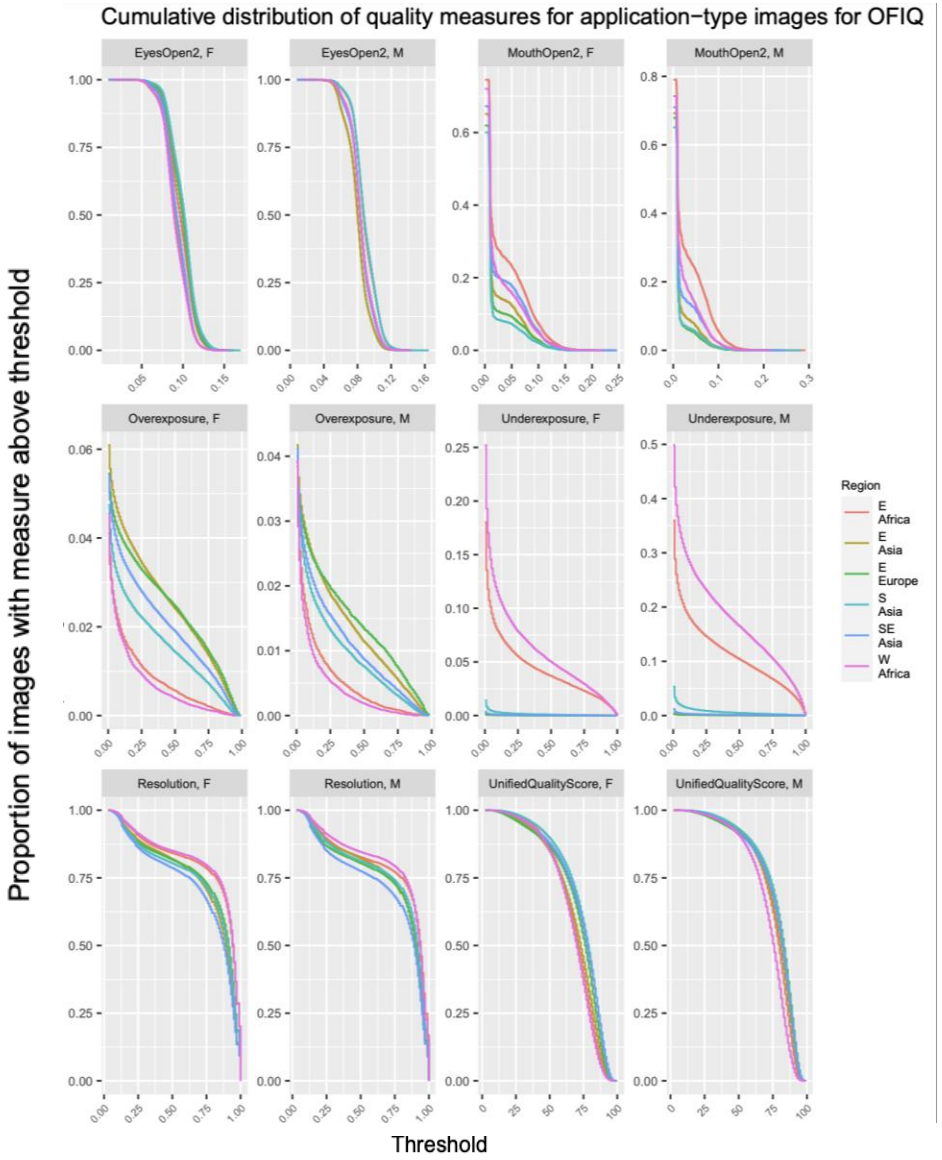
- Each panel corresponds to a range of B = sigma / IED i.e. the amount of blur “damage.”
- The mated similarity score cloud moves downward and leftward as B increases.



OFIQ Evaluation

OFIQ is the reference implementation of ISO/IEC 29794-5.

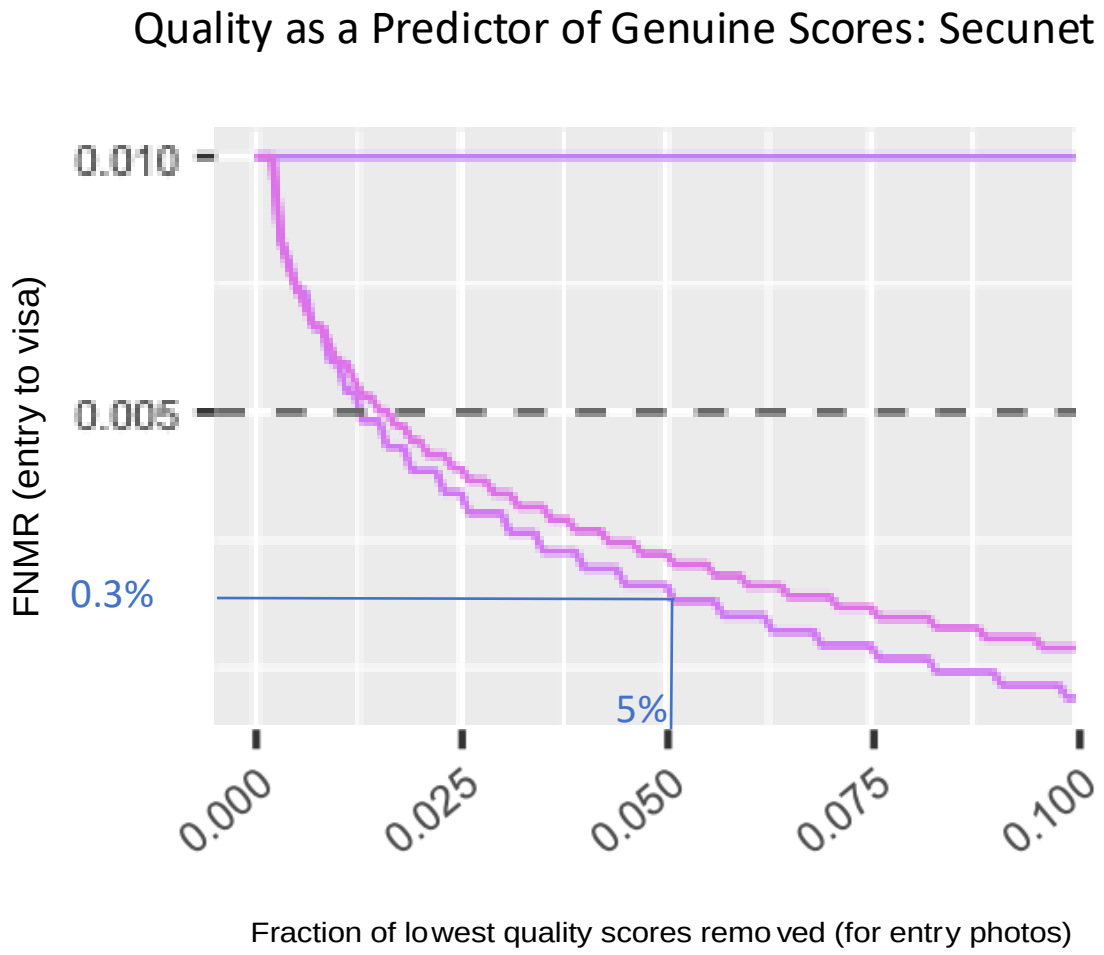
We contributed analysis of OFIQ to the WG3 Ad Hoc Group Report.



Measure	Component Algorithm
EyesOpen2	Secunet_003
MouthOpen2	Secunet_003
Overexposure	Secunet_005
Underexposure	Secunet_005
Resolution	Secunet_003
Unified Quality Score	Secunet_003

Evaluation of Unified Quality Score

- For Unified Quality Score (UQS), quality prediction of false non-match rates is shown by the Error Discard Characteristic (EDC).
- This has been used for face quality in our 2019 quality evaluation and for fingerprint prior to that.
- False non-match rate (FNMR) is calculated based on consensus of 15 recognition algorithms. Lower FNMR is better.
- EDC curves are shown for secunet_002, secunet_004, secunet_003 (from top to bottom)
- Secunet_003 UQS is used in OFIQ.

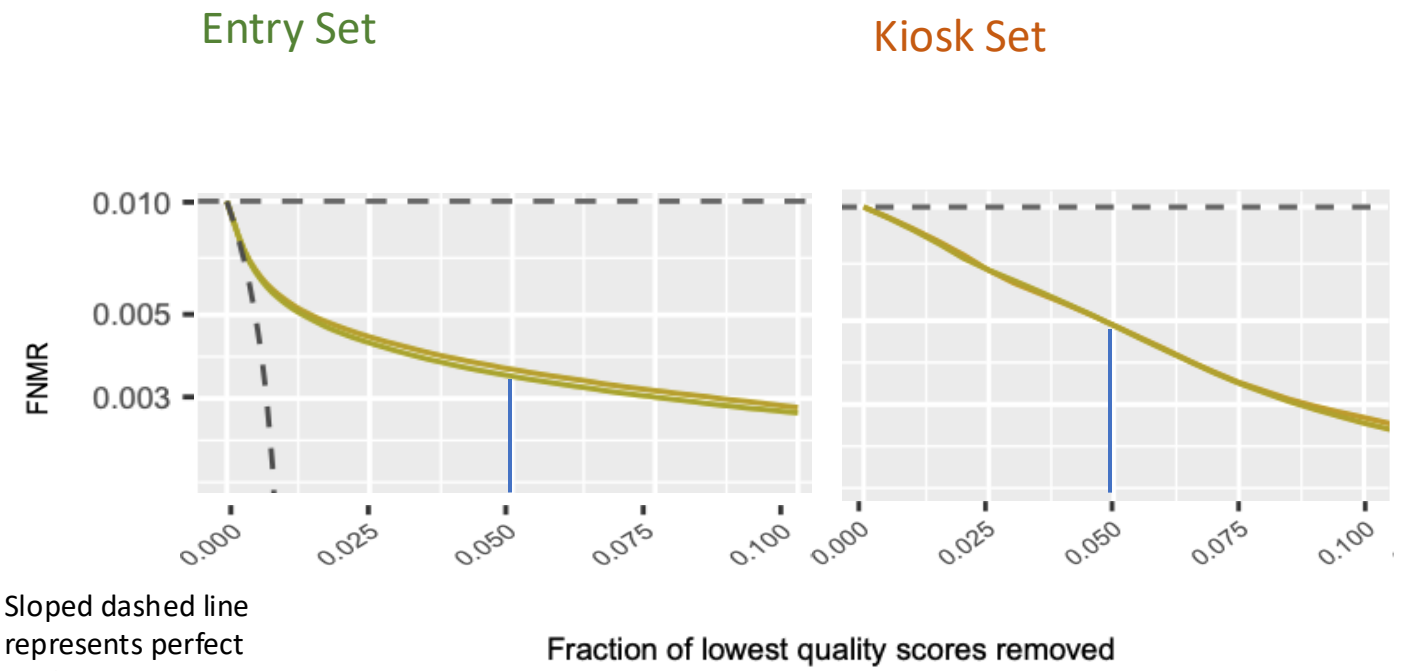


Kiosk Set For Unified Quality Score: Error Curves

Algorithms perform worse for the Kiosk set than for the Entry set.

Kiosk generally yields higher FNMR than Entry for low discard fractions.

Images have varying pose, illumination, and background clutter.



EDC curves for Entry (left; higher quality) and Kiosk (right; lower quality) for idemia_002,_003.

FNMR is calculated based on consensus of 15 recognition algorithms.

Kiosk Set for Unified Quality Score: Leaderboard

Kiosk set generally yields higher FNMR than the Entry set for discard fraction of 5%.

Entry Set

Quality Algorithm	FNMR after removal of 5% lowest quality Entry images. Initial FNMR = 0.01. Lower is better.
secunet_003	0.0031(1)
igd_005	0.0032(2)
qazsmartvisionai_000	0.0033(3)
secunet_004	0.0034(4)
intema_000	0.0034(5)
idemia_003	0.0034(6)
idemia_001	0.0034(7)
frpkauai_001	0.0034(8)
idemia_002	0.0036(9)
datech_000	0.0036(10)

Kiosk Set

Quality Algorithm	FNMR after removal of 5% lowest quality Kiosk images. Initial FNMR = 0.01. Lower is better.
idemia_003	0.0049(1)
idemia_002	0.0049(2)
datech_000	0.0051(3)
neurotechnology_003	0.0057(4)
roc_008	0.0062(5)
papil11_000	0.0062(6)
igd_005	0.0063(7)
neurotechnology_005	0.0064(8)
neurotechnology_004	0.0064(9)
roc_007	0.0065(10)

Recap

- We are in the process of studying effect of SIDD resolution on recognition
- We evaluated OFIQ, the reference implementation for ISO/IEC 29794-5.
- We are in the process of adding a challenging set for Unified Quality Score
- FATE SIDD remains open; 4-month frequency for participation

- For more information and updates, see https://pages.nist.gov/frvt/html/frvt_quality.html or scan the QR code.

Questions?

Contact frvt@nist.gov
Visit https://pages.nist.gov/frvt/html/frvt_quality.html
View SIDD report at
https://pages.nist.gov/frvt/reports/quality_sidd/frvt_quality_sidd_report.pdf

Additional Information

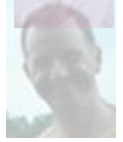
Complete Mapping of OFIQ to Secunet algorithms

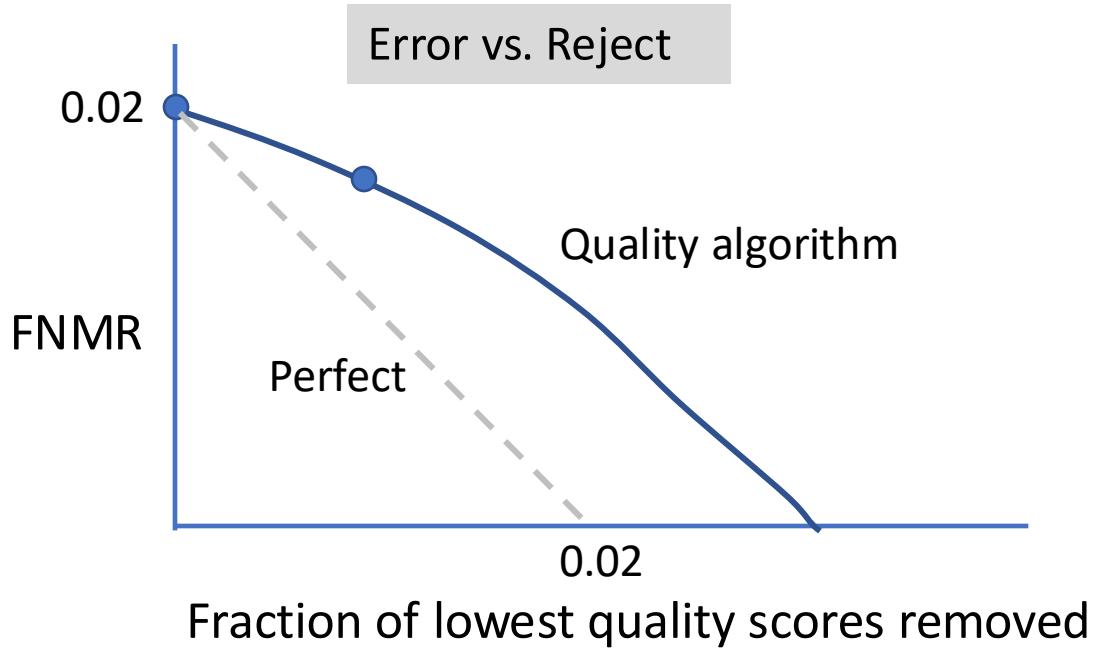
Measure	Component Algorithm
EyesOpen2	Secunet_003
MouthOpen2	Secunet_003
Overexposure	Secunet_005
Underexposure	Secunet_005
Resolution	Secunet_003
Unified Quality Score	Secunet_003
BackgroundUniformity	Secunet_003
Compression Artifacts	Secunet_004
Inter-Eye Distance	Secunet_004
Face occlusion	Secunet_003
Head Pose	Secunet _004

Quality Values as Predictors of Genuine Scores– Error vs. Reject

Quality algorithm

Recognition algorithm

Ref	Probe	Quality	MATCH SCORE	MATCH YES?
		97	0.97	TRUE
		94	0.91	TRUE
		74	0.60	FALSE
		68	0.85	TRUE
		57	0.81	TRUE
		32	0.72	FALSE
		29	0.90	TRUE
		27	0.57	FALSE

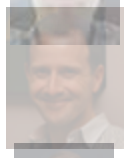
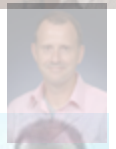
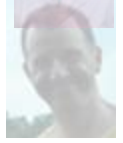


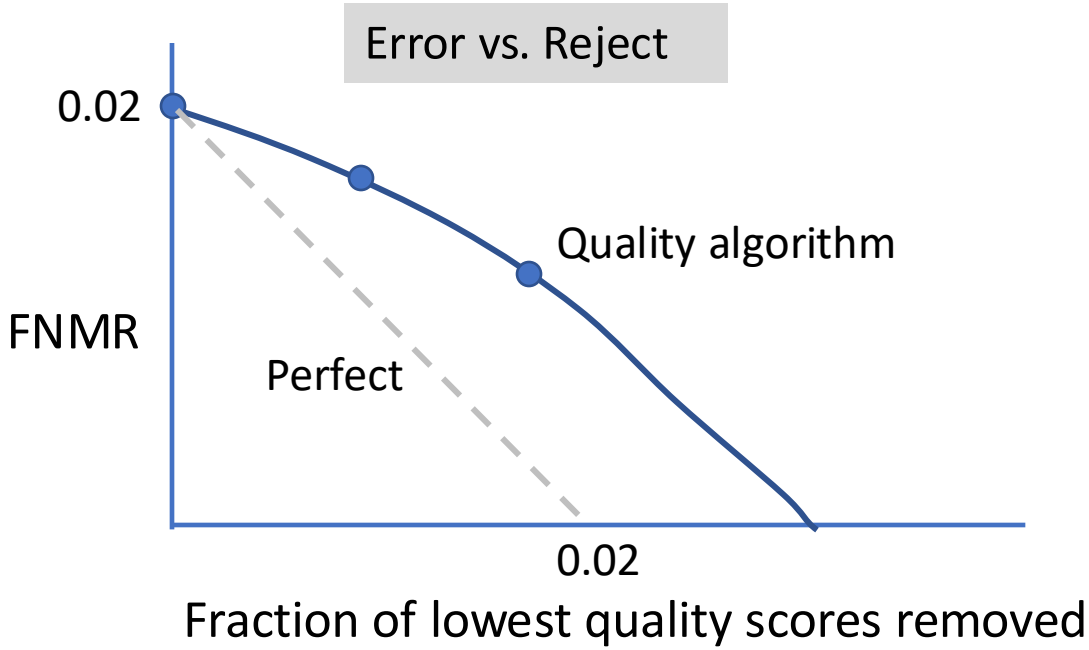
Discard n% of lowest quality probes

Quality Values as Predictors of Genuine Scores– Error vs. Reject

Quality algorithm

Recognition algorithm

Ref	Probe	Quality	MATCH SCORE	MATCH YES?
		97	0.97	TRUE
		94	0.91	TRUE
		74	0.60	FALSE
		68	0.85	TRUE
		57	0.81	TRUE
		32	0.72	FALSE
		29	0.90	TRUE
		27	0.57	FALSE



FNMR is ideally reduced as quality algorithm is used to discard low quality probes.