Rank One
COMPUTING

How a Global Pandemic Altered the Trajectory of Automated Face Recognition Technologies

2022 IAPR / IEEE
Winter School on Biometrics

Presented by: Brendan Klare
Jan. 13th, 2022

DISCLAIMER: *This presentation is for academic purposes only, and may include images without copyrights or attribution. Such use is based on "fair-use" and strictly due to the academic nature of this presentation.*

March 2020 – The World Came to a Standstill:

The Woodlands causeway between Singapore and Malaysia.

Global lockdowns and quarantines

How did they change the course of FR technology?

Presentation Overview

- Company background
- Face recognition in the pre-COVID era
- What changed due to COVID?
- FR reqts more in demand due to COVID
- FR reqts less in demand due to COVID
- Edge cases automated face recognition
- Summary

Get to Know Rank One Computing

INDUSTRY LEADING INNOVATORS IN BIOMETRICS AND MACHINE LEARNING

- Team of AI/ML Algorithm Developers from across the industry
  - Michigan State University
  - IARPA
  - DOJ
  - DOD
  - Noblis
- Designers and Engineers working in harmony
  - Put the human customer first
  - Embrace engineering industry best practices
  - Produce intuitive, powerful GUIs and APIs designed for end users

Headquartered in Denver, CO USA

Denver, Colorado: the Mile High City

The Rank One Difference

Industry-leading algorithms

Engineering first, design forward & nimble

Video / Real-Time
2-5x less CPU hardware

ID Applications
10-20x less RAM

Customer-friendly, trusted & proven

Bootstrapped & made in the U.S.

Enterprise
Significantly lower hardware needs

Mobile
193 milliseconds vs. 1 sec+

Flagship Product

ROC SDK

Cross-Platform Code Library

Multi Programming Language

Easy to Integrate

Integrates with UAS, HUD, LPR, and other 3rd party Platforms

25M+
Facial Verifications Per Year

40+
Integrator Customers

6
US Dept. of Defense Agencies

20+
Law Enforcement Agencies

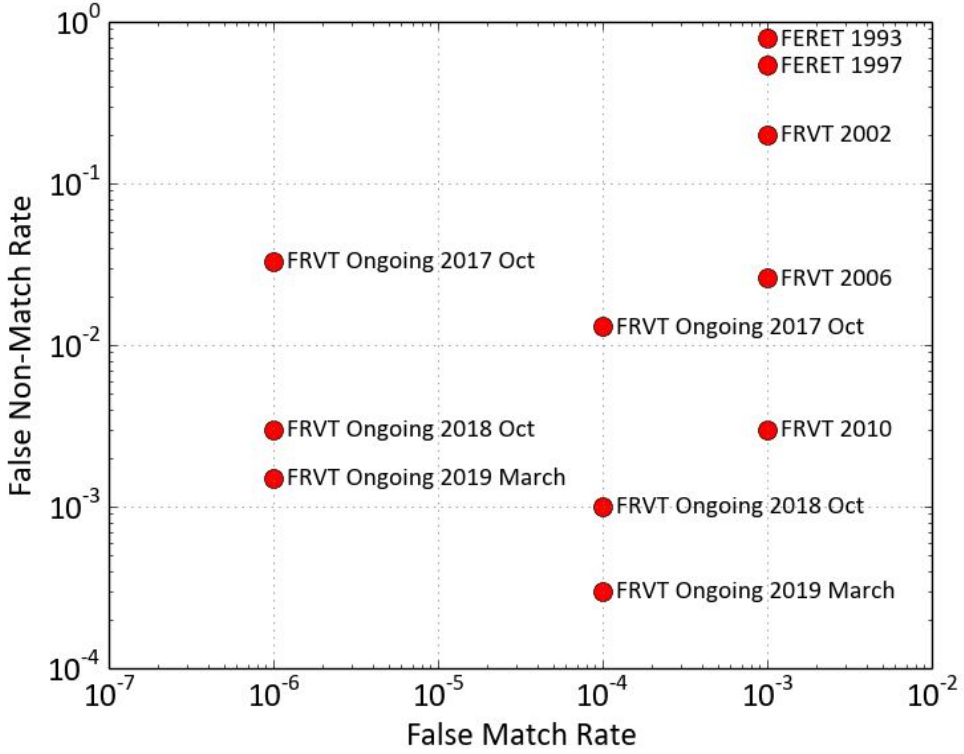
5+
Fortune 500 Financial Institutions

1 of 2
Major Global Credit Card Companies

Government | **Law Enforcement** | **Identity Services** | **Other**

Face Recognition in the pre-COVID era

Exponential Accuracy Improvements



FR Industries and Applications

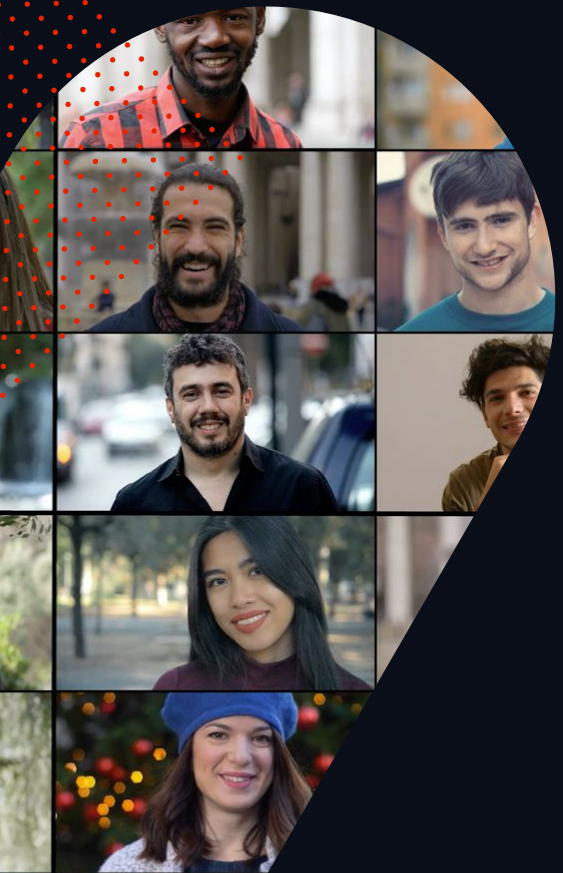
Legacy applications further strengthen:

- Identity deduplication (DMV's, Dept. of State)
- Forensic identification (FBI, state/local)
- Border screening (DHS)

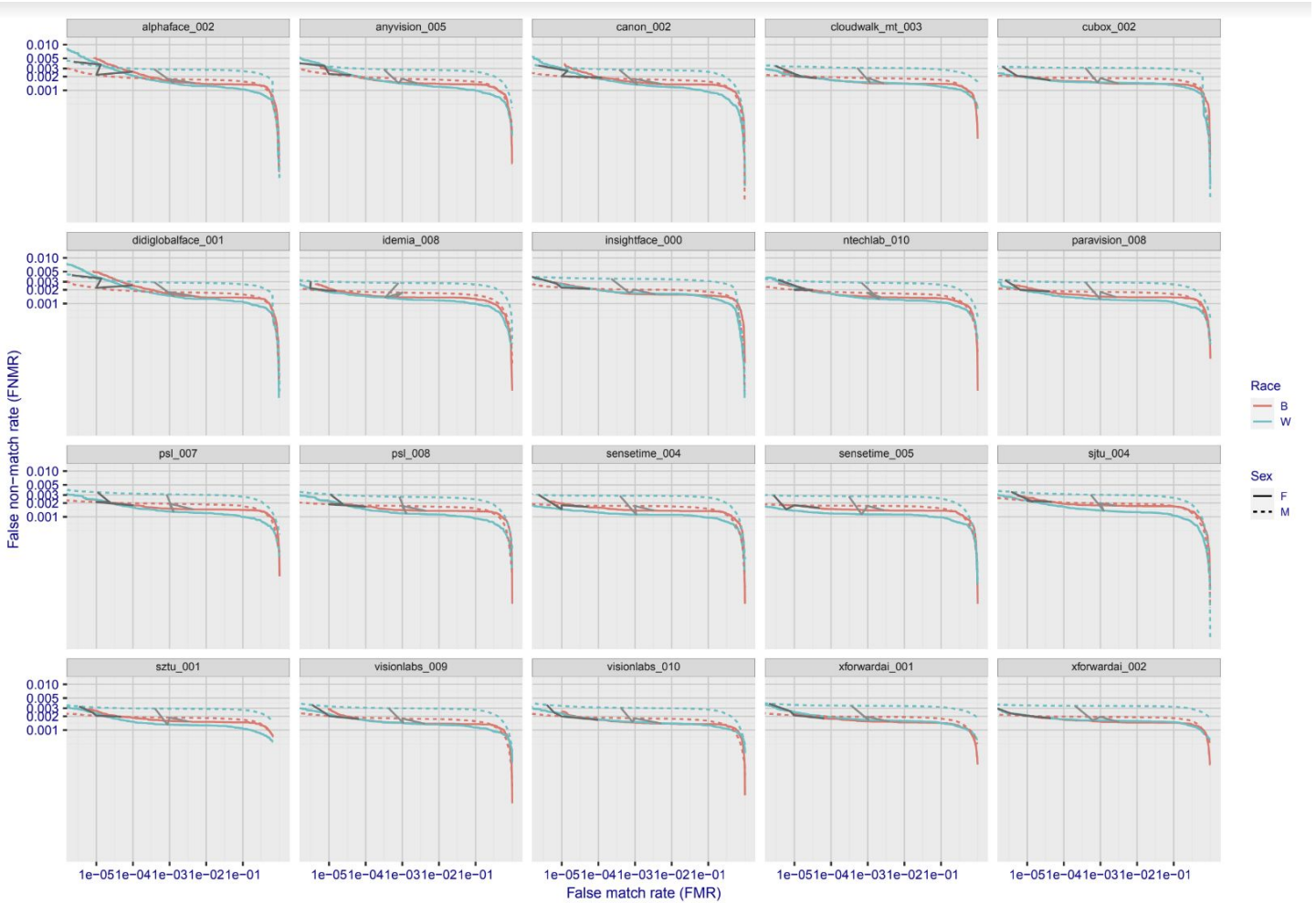
Nascent applications plant roots:

- Identity proofing
- Access control
- Real-time screening
- Customer experiences
- Security

Algorithm bias



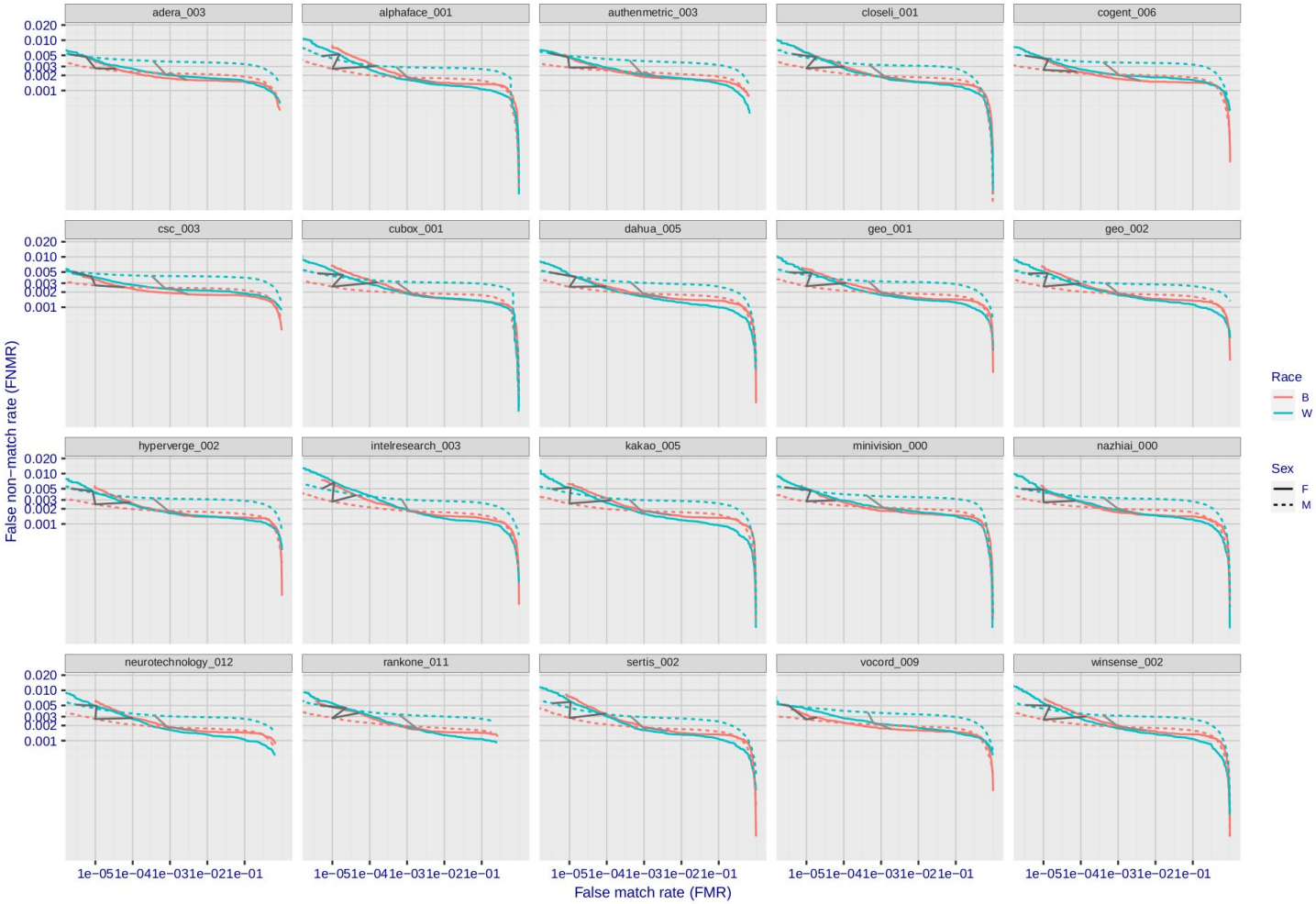
- A furor of media misinformation skewed even scientists perceptions about FR bias
  - Deep misunderstanding about FR capabilities
  - Laws being codified based on misinformation
- Do FR algorithms exhibit different degrees of bias?
  - Yes
- Are top-tier FR algorithms deeply biased against certain demographic groups?
  - No
- Are we measuring bias with broken rulers?
  - Probably



DET curves for top performing FRVT Vendors on "Black" vs. "White" and "Male" vs. "Female"

"Black" subjects are frequently recognized with higher accuracy

Source: Figure 111, NIST FRVT Ongoing, Sep. 8, 2021

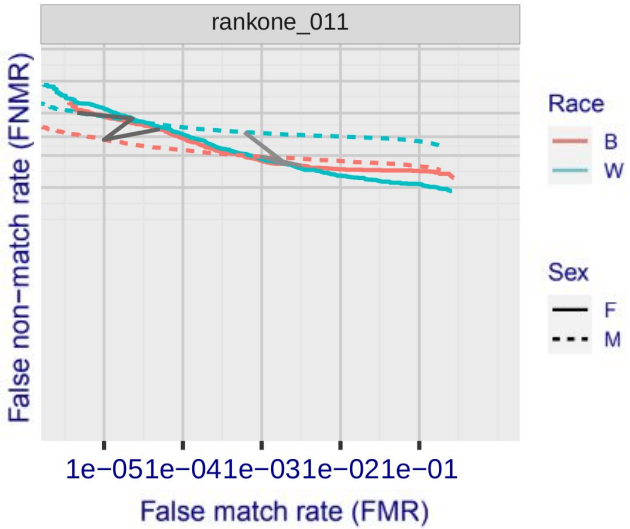


DET curves for top performing FRVT Vendors on "Black" vs. "White" and "Male" vs. "Female"

"Black" subjects are frequently recognized with higher accuracy

Source: Figure 113, NIST FRVT Ongoing, Sep. 8, 2021

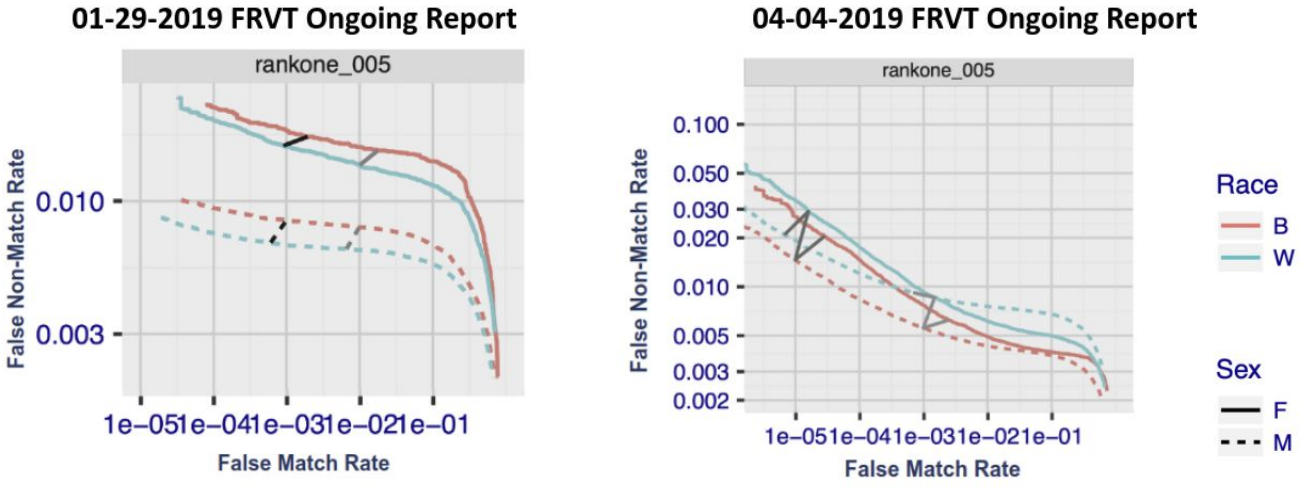
Why the misperceptions about bias?



- Seed:
  - First large-scale study to measure error rates across demographics groups:

    Face recognition performance: Role of demographic information
    BF Klare, MJ Burge, JC Klontz, RWV Bruegge, AK Jain
    IEEE Transactions on Information Forensics and Security 7 (6), 1789-1801
  - Relied on ground truth labels from law enforcement agencies
- Amplification:
  - Georgetown Privacy Center cites lone source above to amplify the findings
- Deception:
  - MIT Gendershades study (did not study FR)
- Validation?
  - NIST FRVT?

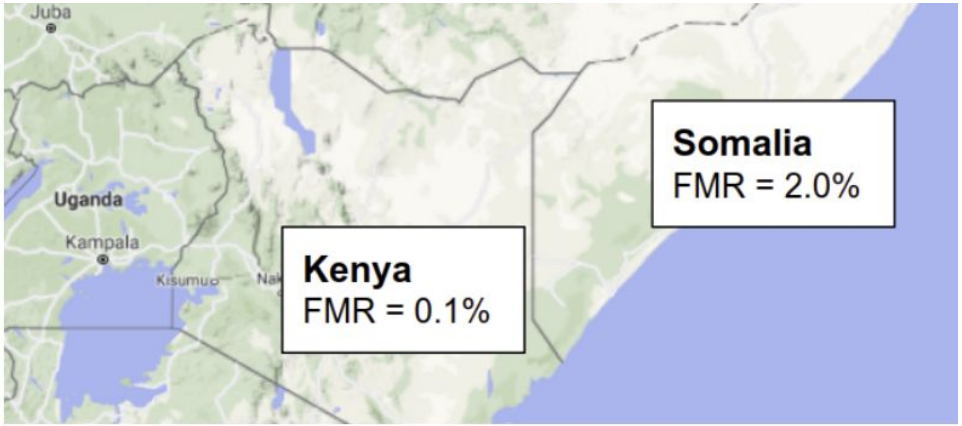
Non-uniform ground truth errors?



Shown are results from the exact same algorithm submission, rankone_005, in two different versions of the NIST FRVT Ongoing report. The results in (a) were published using a dataset with ground truth errors. The results in (b) were published on a dataset with ground truth errors corrected. Without correcting ground truth errors in datasets, algorithm performance cannot be properly measured. In this case, the ground errors falsely demonstrated an algorithm was less algorithm on "Black" persons than "White" persons.

Same algorithm, relative accuracy on race flipped after ground truth label corrections

Non-uniform ground truth errors?



False Match Rates measured in the NISTIR 8280 "Demographic Effects" report for the countries Somalia and Kenya are listed. The report largely attributed the high error rates measured in Somalia to be due to racial effects, despite being geographically neighboring and homogeneous in race to countries with much smaller error rates. The high error rate in Somalia has since been attributed to ground truth errors in the Somalia dataset, though the report was never revised.

- 2019 NIST Demographic Effects paper resulted in **media headlines stating 100x higher error rates on African persons**
- Lost in coverage a ***20x difference in error between African nations***
- Ground truth errors, often due to fraud, have skewed our ability to measure FR error
- Key benchmark datasets need to be inspected for accuracy
- Ironically, FR algorithms are the solution to finding GT errors

Summary of FR pre-COVID:

- Deep learning fuels exponential reduction in error rates
- Legacy applications further strengthen; novel applications plant roots
- Misinformation on FR algorithm bias proliferates
- Major investments in core FR technology:
    - 16 organizations submitted to NIST FRVT in 2014
    - Roughly 100 organizations algorithms submitted Jan 2020 (pre-COVID) FRVT Ongoing
    - Roughly 200 organizations algorithms submitted to most recent FRVT Ongoing

What changed with COVID?

Multitude of changed societal factors due to COVID

- Social isolation for some lead to depression and vulnerability
- Some industries shuttered, which lead to bankruptcy, unemployment, etc.
- Other industries boomed, which lead to massive increases digitally connected industries (retail, food, payments, telecommunications, gaming, etc)
- Economies stuttered, but still performed
- Limited in person experiences
- Steep rise in digital media experiences

Rise of masks

- Many societies rapidly pivoted from anti-mask laws to laws mandating facial masks
- Significant challenges introduced for both human facial perception and automated systems
- Several adjustments required:
  - Biometric services needed to be contactless and support masks
  - Anonymity in public places
  - Societal adjustment for human cognitive facial analysis

Danger of hidden facial appearance:

Civil unrest: the double-edged sword

Peaceful demonstrations:

- Global concerns about potential use of FR to violate human right for peaceful protest
- FR being used to violate civil liberties is less of in a viable concern in Western Europe and North America

Riots and Looting:

- Physical and financial harm caused by mob violence
- Some jurisdictions prevented from using FR due to local bans

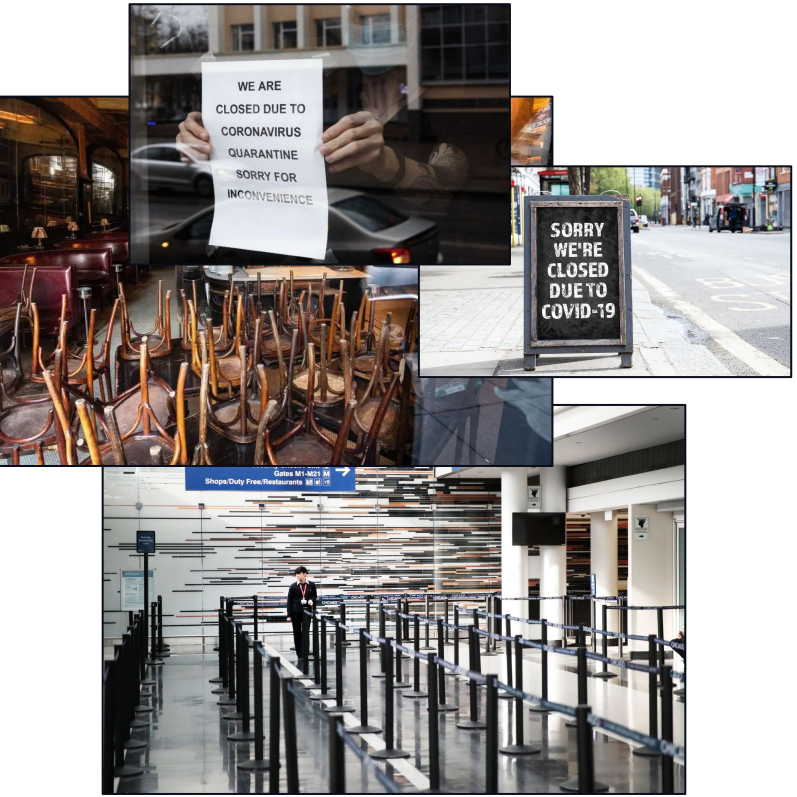
Is our facial appearance private?

Historically: **No**

Our facial appearance is the single least private piece of information about ourselves

Privacy-by-design must recognize this fact and focus on **what information is linked to facial appearance**

COVID: Industry Recessions vs. Booms



VS.

shopify
Trulioo
DocuSign
zoom
Square

U.S. NEWS

NEWS

'Easy money': How international scam artists pulled off an epic theft of Covid benefits

Russian mobsters, Chinese hackers and Nigerian scammers have used stolen identities to plunder tens of billions of dollars in pandemic aid, officials say.

Among the ripest targets for the cybertheft have been jobless programs. The federal government cannot say for sure how much of the more than $900 billion in pandemic-related unemployment relief has been stolen, but credible estimates range from $87 billion to $400 billion – at least half of which went to foreign criminals, law enforcement officials say.

https://www.nbcnews.com/news/us-news/easy-money-how-international-scam-artists-pulled-epic-theft-covid-n1276789

Telehealth Fraud

- **Number of telehealth visits increased from about 10,000 per week to 300,000 per week** in late March of 2020, according to CMS" [1]
- Significant challenges validating doctors rendered services billed to gov't, or patients match identities on insurance plan

[1] https://www.natlawreview.com/article/qui-tam-quarterly-hunting-telehealth-fraud-under-covid-19-waivers-and-expansion

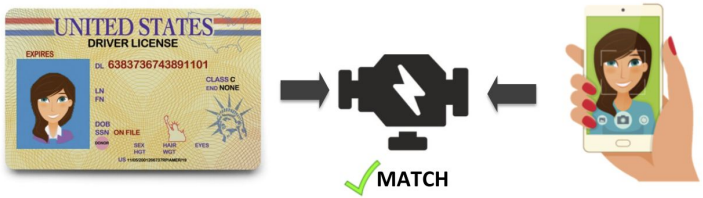
Proof of vaccination

Emerging FR Requirements

How have all the societal changes altered FR requirements?

Identity Verification Services

- When physical store fronts close, customer onboarding and issuance of services are needed
- Accurate face recognition in unconstrained, cooperative, frontal capture environments:
  - Selfie to ID document verification
  - Minimal demographic bias
- Anti-fraud
  - Liveness / Anti-Spoof
    - Sensor agnostic
    - Network bandwidth vs. edge processing
  - 1:N search
  - Deep fake detection?
    - Easier to solve issue via liveness validation



Biometrics: Contactless and mask-compatible

- Masks will be around for a while
- Germs and sanitization will be important for a while
- These both lead to the need for biometrics systems that can perform **contactless** identification **while masks are worn**

Privacy-by-design

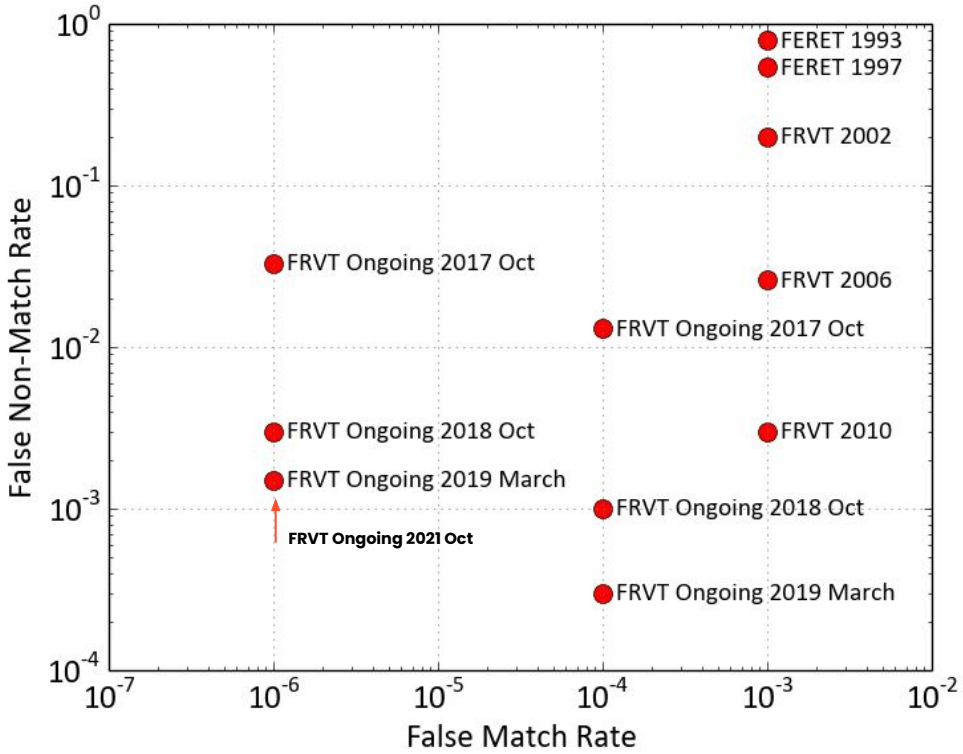
- **Data security is constantly being compromised**
- Privacy-by-design is increasingly a requirement as we switch to a digital world
- GDPR, CCPR and other privacy laws are in effect
- FR systems need to provide ability to easily remove templates & metadata, enforce data retention policies, allow for easy response to data subject requests
- Obligations reside with the data controller not the product vendor
- Much of the burden is at the design, implementation, and installation stage

Applications less in-demand due to COVID

- Growth of the roots planted for these applications temporarily stalled:
  - Security
  - Real-time screening
  - Retail analytics
- Demand will resume, but for now applications supporting the digital world are rapidly outpacing those supporting the physical world

Where is the edge in automated face recognition?

Exponential Accuracy Improvements



Is face recognition becoming solved?

Only a 1.5x error rate reduction in the last 2.5 years

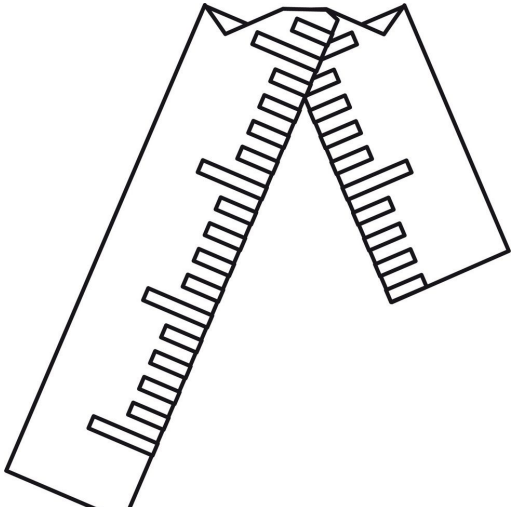
Where's the edge / when do we stop?

Major nations range in population from ~100M (10e8) to 1B (10e9)

Earth population in 2050 estimated to be ~ 10B (1e10)

Goal needs to measurement of False Match rates at population sized samples (10e-8 to 10e-10)

How do we measure accuracy at the edge?



- How do we test algorithms on population size data?
- How do we measure accuracy when the databases have some degree of error in them?
    - Errors could have originated from fraud or human or machine mistakes
- We are measuring accuracy with broken rulers unless we have means to accurately cleanse testing sets

How accurate are operational systems?

- Accuracy of face recognition algorithms is typically measured in isolation of the operational system
- Most sensitive of use-cases involve human decision making (e.g., forensic face recognition)
- How accurate is the combined system?
    - We have started to answer this question [1] but a long ways to go
- E.g.: *Does the "other race effect" exist when using the morphological facial comparison process?*
- Is the algorithm deployed what is benchmarked in NIST FRVT?

[1] Phillips, P. Jonathon, et al. "Face recognition accuracy of forensic examiners, superrecognizers, and face recognition algorithms." Proceedings of the National Academy of Sciences 115.24 (2018): 6171-6176.

Attack the breach

- Opt-in Facial Verification Systems used to provide access to resources (e.g., banking, subsidies, etc.) are being exploited by fraudsters
- Loopholes and “hidden biases” - vulnerabilities in the algorithm that are inconsequential for normal use, but a vulnerability that can be exploited by fraudsters

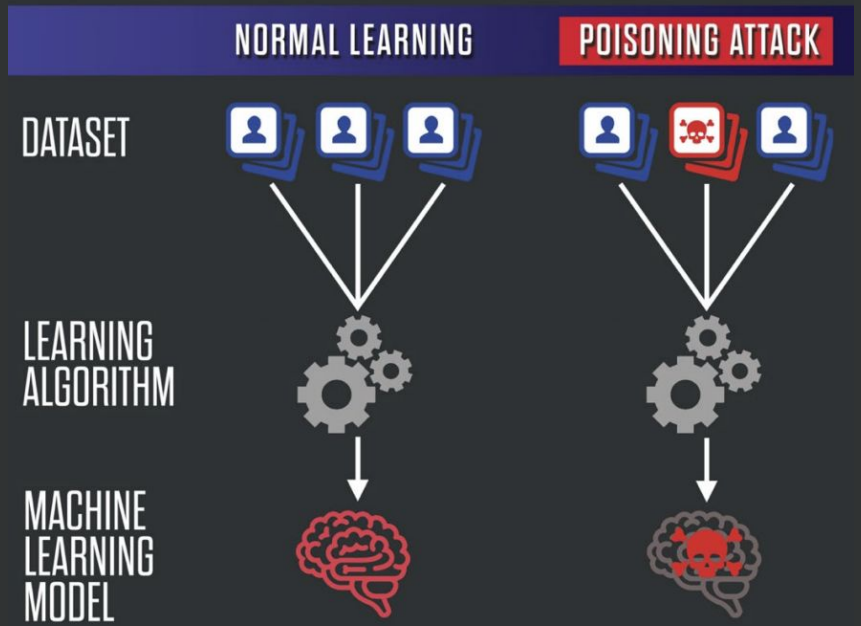
Credit: Klim Kireev/YouTube

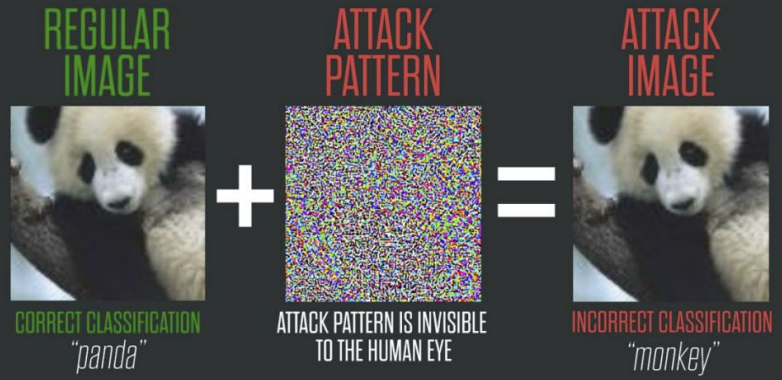
Credit: Amanda Dave of Dazzle Club. Photograph: Cocoa Laney/The Observer

Livescan and Liveness algorithms to the rescue

- Liveness / anti-spoof:
  - Algorithms to determine if a person is in front of camera or using a spoof image (e.g., a printed photo or phone screen)
  - Different procedures used to solve the problem:
    - **Passive vs. Active**
      - Passive methods: More convenient, Less secure
      - Active methods : Less convenient, More secure, but… “Deep” Fake is a particular threat for active liveness methods
    - **Sensors:** agnostic or custom hardware?
- Livescan:
  - Ensure adherence to ICAO or ISO standards - E.g., no facial ornamentation

Poison AI





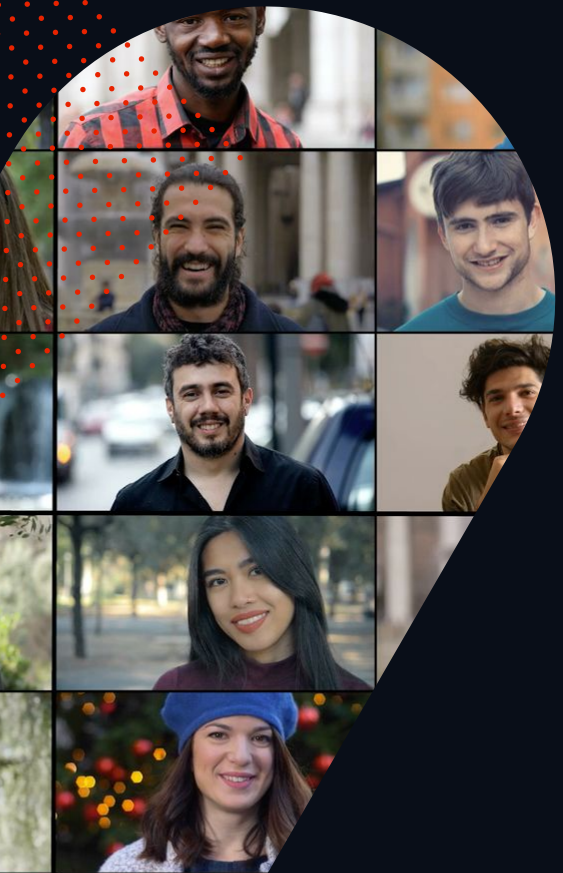
How to test for poisoned face recognition models?

We generally can't, so vendor trustworthiness is vital

Image source: Goodfellow, Ian J., Jonathon Shlens, and Christian Szegedy. "Explaining and harnessing adversarial examples." arXiv preprint arXiv:1412.6572 (2014).

Image source: Comiter, Marcus. Attacking Artificial Intelligence: AI's Security Vulnerability and what Policymakers Can Do about it. Belfer Center for Science and International Affairs, 2019.

Demographic bias



- A furor of media misinformation has skewed even scientists perceptions about FR bias
  - Deep misunderstanding about FR capabilities
  - Laws being codified based on misinformation
- Do FR algorithms exhibit different degrees of bias?
  - Yes
- Are top-tier FR algorithms deeply biased against certain demographic groups?
  - No

The edge of this problem:

- **Are we measuring bias with broken rulers?**
- **How to even define racial and other cohorts?**

Cosmetics

- FR is incredibly accurate on women, though typically slightly less accurate than with men
- Lower accuracy with females likely a **latent effect** and not due to physiological differences between men and women
- Instead, ***likely due to cultural use of cosmetics***

Algorithm Efficiency

Enrollment speed – the amount of time it takes to detect and templatize all faces in an image

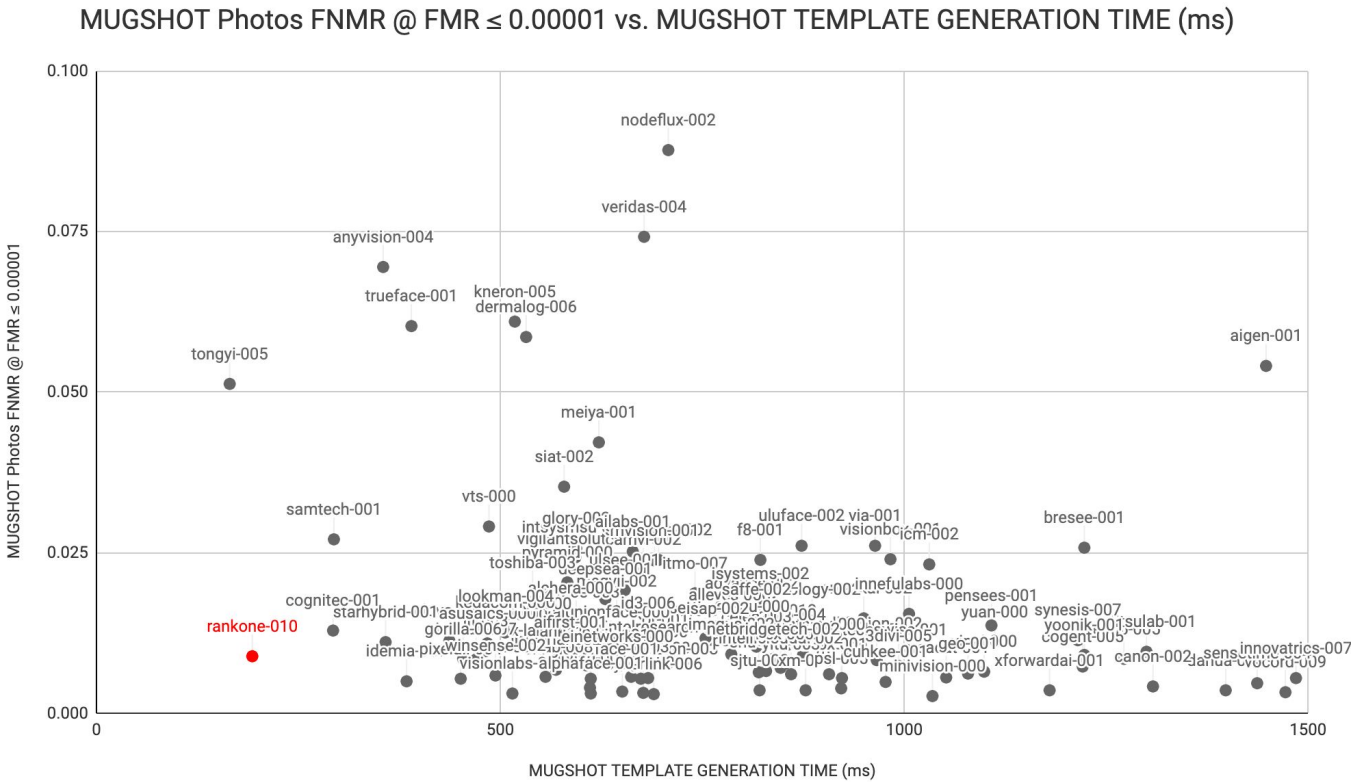
Template size – the number of bytes required to represent a face

Comparison speed – the amount of time it takes to compare two templates and generate a threshold

Binary size – the amount of computer memory required to run all FR models and libraries

Different applications have different efficiency requirements; most FRVT submitted algorithms cannot meet those requirements

Algorithm efficiency varies dramatically:



Source: NIST FRVT 1:1 Ongoing Report (July 27, 2020)

Summary

- Face recognition in the pre-COVID era
  - Major accuracy gains, legacy deployments grew, nascent applications incubated
- What changed due to COVID?
  - Online services, Masks, Subsidies, Vaccine checks, Civil unrest
- How COVID altered FR requirements
  - More in demand: ID proofing, contactless access control / identity verification, medical
  - Less in demand: real-time screening, security surveillance, retail / customer experience
- Many edge cases and challenging research problems remain

Rank One
COMPUTING

Thank you!

Questions?

Brendan F. Klare, Ph.D.
brendan@rankone.io