Rank One
COMPUTING

Edge Cases in Automated Face Recognition

Face Image Quality Workshop
Hosted by: European Association for Biometrics (EAB)

Presented by: Brendan Klare
Nov. 17th, 2021

Edge Cases in Automated Face Recognition

What are FR's breaking points, limiting factors, and unanswered questions?

Measurement of "quality" requires both known limits and known-unknowns

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.*

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

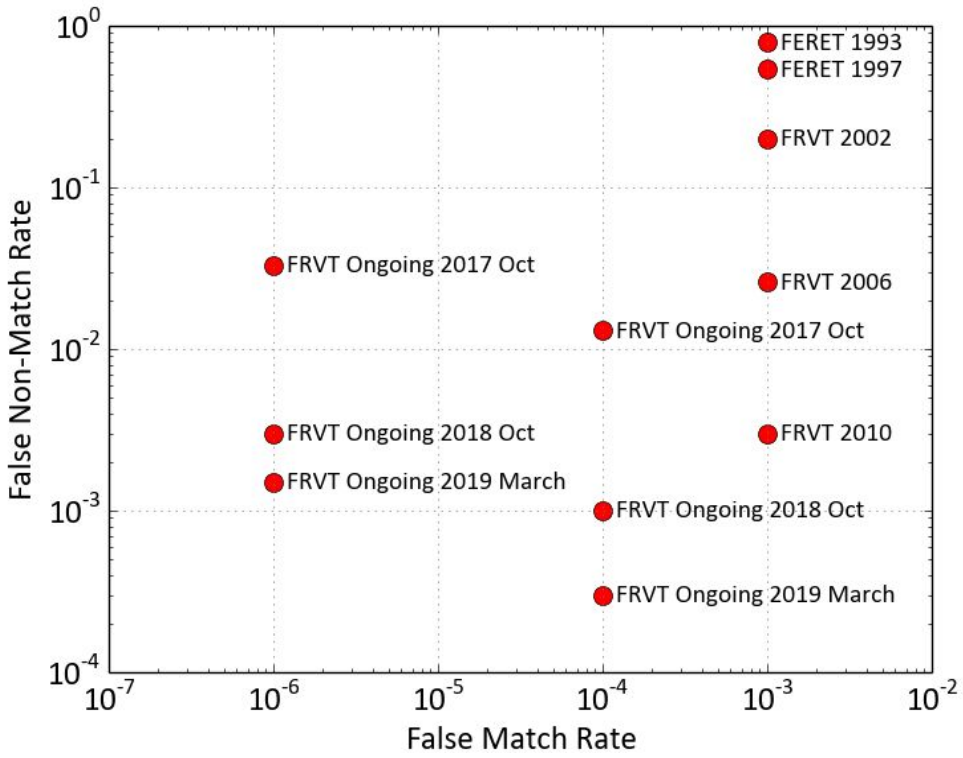
1 of 2
Major Global Credit Card Companies

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

Specific customer list is confidential. Contact bd@rankone.io for more information.

Where is the edge in automated face recognition?

Exponential Accuracy Improvements



Is face recognition becoming solved?

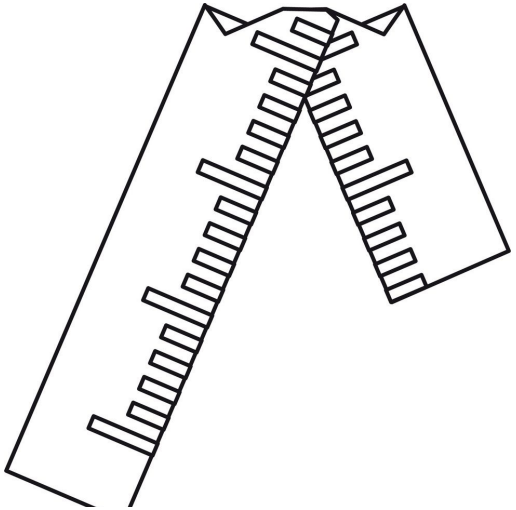
Where's the edge / when do we stop?

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

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.

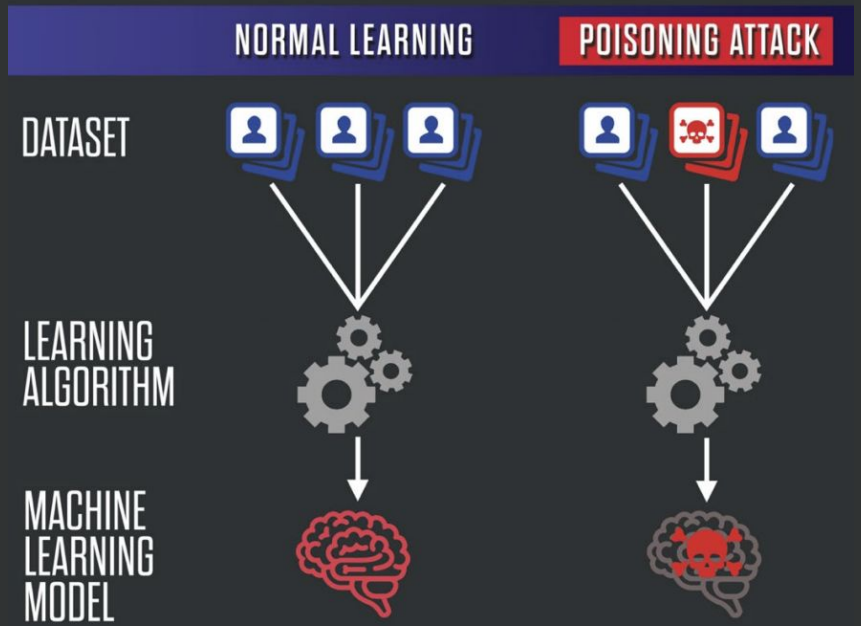
Among the ripest targets for the cybertheft have been jobless programs. The federal government cannot say for sure how much of the more than $900 billion in pandemic-related unemployment relief has been stolen, but credible estimates range from $87 billion to $400 billion – at least half of which went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

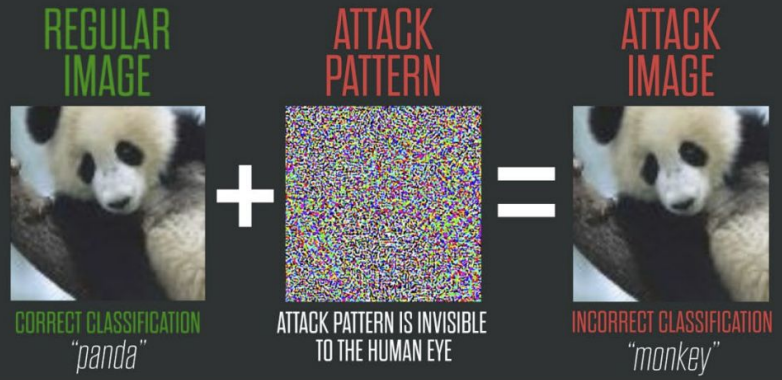
https://www.nbcnews.com/news/us-news/easy-money-how-international-scam-artists-pulled-epic-theft-covid-n1276789

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

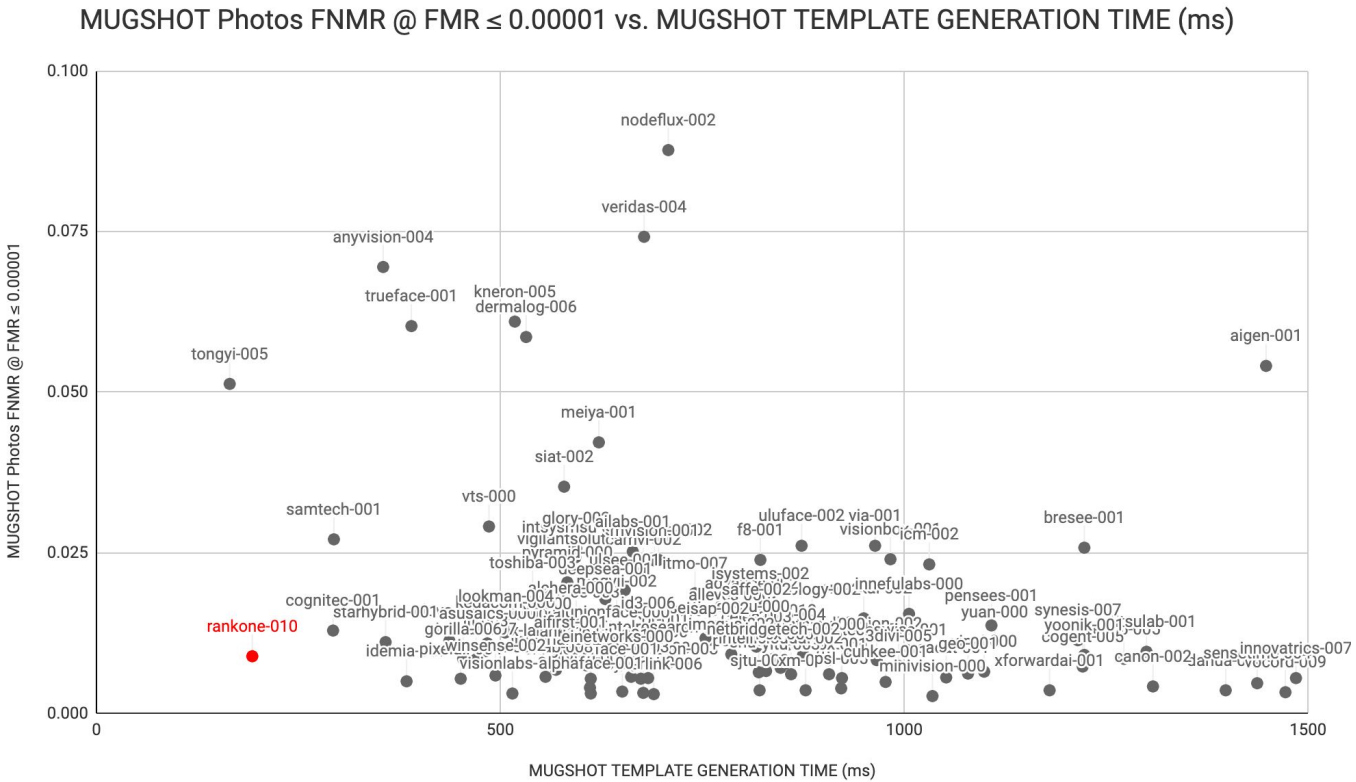
Template size – the number of bytes required to represent a face

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)

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**

Rank One
COMPUTING

Thank you!

Questions?

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