UNIVERISTY OF PENNSYLVANIA

SCHOOL OF SOCIAL POLICY AND PRACTICE

MASTER OF SCIENCE IN SOCIAL POLICY AND DATA ANALYTICS

SEMESTER LONG PROJECT - PART #2

The ethical conundrum of Facial Recognition Technologies powered by Artificial Intelligence

Autor: Barragan Barranco, Rose

Professor: Amoroso Latta, Gina

Course: MSSP 6280 Policy Analysis
of Issues, Strategy and Process

October 11, 2023

INDEX

1. PROBLEM DEFINITION 1
2. CONSTRUCT POLICY ALTERNATIVES 1
2.1. BASELINE CONDITION 1
2.2. NATIONAL AND INTERNATIONAL BIOMETRIC DATA PROTECTION TREATIES AND LAWS 3
2.3. HUMAN-CENTRIC MODELS THAT PROMOTE PRIVACY BY DESIGN 4
2.4. INDIVIDUAL CONSENT FORMS 4
3. DEFINE CRITERIA 5
3.1. EFFECTIVENESS 5
3.2. EFFICIENCY 6
3.3. EQUITY 6
3.4. SUSTAINABILITY 7
4. EXHIBITS 8
5. REFERENCES 11

1. PROBLEM DEFINITION

Facial Recognition Technology (FRT) presents an unprecedented threat to the individual's privacy and fundamental freedoms.

The global widespread surveillance of FRT - capable of identifying and tracking individuals' facial or biometric information - without adequate safeguards in place, the lack of awareness [1] and consent or knowledge [2], indiscriminate use case application [3], lack of democratic oversight, fairness, and reliability (significant racial bias) [4], lack of accountability and transparency measures, erosion of anonymity and privacy rights (which are crucial for safety and security), unwarranted intrusion into individuals' lives, and legal loopholes within the regulations are leading to scenarios in which human beings' civil rights and liberties are severely undermined, and even violated.

Individuals' facial or biometric information is sensitive, private, and highly valuable data as it is directly linked to personal identity, and its misuse can result in identity theft, fraud, and law enforcement abuse. Therefore, it is crucial to oversee FRT to safeguard individuals' biometric data and privacy, prevent security risks, and maintain the balance between national security and individual freedoms.

For those reasons, State and Federal Governments need to establish regulations and safety measures to ensure that Facial Recognition Technology - capable of tracking individuals' biometric data - is handled ethically, with informed consent and legal compliance, in order to maintain public trust in their institutions and safeguard citizens' fundamental rights.

2. CONSTRUCT POLICY ALTERNATIVES

2.1. BASELINE CONDITION

To address this alternative, we have to assume a scenario where we are in the year 2030, and facial recognition technologies are widespread in all nations, as government and military investments have resulted in the widespread adoption of FRT technologies and are being deployed leveraging homeland security, federal law enforcement departments and other security agencies for criminal investigation, identification of suspects, tracking persons of interest, finding missing persons, monitoring the movement of people at border crossings, banking, and retail for access control, patient identification, etc. [Exhibit #1] [5].

Moreover, technological advancements in artificial intelligence, deep learning (DL) algorithms, and computer vision drive the facial recognition market. Furthermore, the market is facing an expansion fostered by the increasing availability of high-resolution cameras, the increasing use of smartphones with facial recognition features, and the demand for contactless authentication techniques [6] .

According to the latest report published by Allied Market Research [6] , due to the widespread use of these technologies, the FRT market-valued at $5.5 billion in 2022 and $8.5 billion in 2025, now has a projected $24.3 billion by 2032, growing at a compound annual growth rate (CAGR) of 16.4% from 2023 to 2032 [Exhibit #2] [7].

However, this exponential growth of technologies has the opposite effect in addressing the ethical issues that this technology entails, reaching a critical point of concern in society, especially in reference to privacy, data security, the biases of the algorithms used for facial recognition, and the accuracy of the results, especially when groups of color and vulnerable communities come into play.

If there is one thing about technology, it is that it is not only able to mimic human behavior and bias, but even to exaggerate and magnify it, making the impact far worse than expected.

This is precisely what Pew Research Center experts Lee Rainie, Cary Funk, Monica Anderson and Alec Tyson indicate. Their field research [5] focuses on bringing Americans' perspectives on the widespread use of FRT by law enforcement, focusing in particular on the views of the most vulnerable groups in society.

Among the relevant research data, it is worth noting the potential risks of FRTs, where 69% of research participants strongly believe that law enforcement will definitely track everyone's location at all times, and 66% claim that police would use this technology to surveil black and Hispanic neighborhoods much more frequently than other neighborhoods [Exhibit #3] [5].

On the other hand, mention should be made of facial recognition's effects on false arrests. 53% of Americans say that police would definitely make more false arrests if the use of facial recognition technology were widespread among police. Of these, 28% are of African-American origin, followed by 19% of Hispanics [Exhibit #4] [5].

So, technology should be conceived to the extent that it creates security, equality, and protection of people. However, the widespread use of FRT without control and oversight will only generate more social inequality, insecurity, uncertainty, and injustice by the authorities and organizations that deploy and use such technologies.

2.2. NATIONAL AND INTERNATIONAL BIOMETRIC DATA PROTECTION TREATIES AND LAWS

This alternative seeks to promote international agreements and standards with transposition in each nation, specifying the sensitive nature of facial recognition data collected through FRT systems.

These compelling standards seek to regulate the development, deployment, and use of FRT as well as the collection, storage, processing, and use of biometric data. The resulting robust regulatory framework of protection and practical enforceability should address aspects such as:

- Scope of its use, limiting the use of FRT to restricted, specific, and well-defined use cases, and thus, establishing the prohibited use cases that must be banned from using FRT.
- Define harmonized standards and ethical guidelines reflecting specific essential requirements for the responsible use of these technologies, emphasizing aspects such as transparency, fairness, and accountability.
- Design training and evaluation programs for the technical teams that develop the artificial intelligence models that embed FRTs to prevent programmer biases from being transferred to the models, creating ethical and fair solutions.
- Implement an oversight system focused on monitoring and supervision through a network of Notifying Authorities and Notified Bodies, which allow for auditing (internally and externally) the models, technologies, and results.
- Place a whistleblower protection system where public and private organizations and individuals can lift the veil on FRTs and demand transparency and accountability.
- Support innovation through regulatory sandboxes where stakeholders can test technologies in safe environments without transferring risks and biases to society.

2.3. HUMAN-CENTRIC MODELS THAT PROMOTE PRIVACY BY DESIGN

On numerous occasions, technology is first developed in response to a problem that needs to be addressed, altogether leaving aside two fundamental considerations of ethical technologies: 1) the importance of people interacting (directly or indirectly) with these technologies and 2) ensuring privacy in the design and development of FRT systems from their very beginning.

Although at first glance they may seem to be two separate considerations, they appeal to the same paradigm since the privacy of the models places people and their data at the center of the FRT design, seeking to understand who will be the end user for whom the FRT is to be intended, and what privacy measures will be applied to the data, such as anonymization measures, de-identification, and biometric encryption.

Anonymization and de-identification of facial data involve the application of robust techniques to process and store information in a way that makes it extremely difficult to trace it back to specific individuals. In addition, biometric encryption plays a crucial role in exploring technologies that allow biometric data to be encrypted, ensuring that it can only be decrypted by authorized parties and designated bodies for specific purposes.

Typical methods[8] used for facial image anonymization include image processing with face blurring and noise adding [Exhibit #5][9], data masking techniques [Exhibit #6][10], K-Same anonymization algorithms[11], and adversarial generative networks [Exhibit #7] [12] [13].

These measures become cornerstones for the ethical development of facial recognition technologies, where each facial image identified by the FR model is anonymized while ensuring high action detection performance.

In other words, facial image anonymization methods will ensure both privacy by design and the protection of human rights in an ever-evolving technological environment.

2.4. INDIVIDUAL CONSENT FORMS

The fourth alternative focuses on the principle that individuals have the power and right to give explicit consent to disclose their biometric data before the FRTs can use it. Consent can be either paper or digital, but most importantly, it must address considerations such as:

- Indicate how the collection, storage, and specific purpose of the data will be carried out, as well as the identification of the agents who will have access to the data.
- Establish mechanisms to facilitate the user's access to his biometric data collected (even if he has given his consent to third parties). In the same way that access is granted, the individual should be able to limit access to third parties or even revoke access at any time and under any circumstances.
- Set fines for non-compliance and violation of the consent form when there is a breach or infringement of the agents in the non-consensual use of biometric data.
- Design a user-friendly consent form that is easy to access and simple to complete, allowing all stakeholders (regardless of their technological literacy and maturity) to understand what the consent form consists of and their role in the technological society.

As in the third alternative, this one also promotes a human-centric approach to technologies (from a different perspective), placing individuals at the center of the decision-making process regarding their facial information, which promotes the protection of privacy and individual autonomy.

3. DEFINE CRITERIA

3.1. EFFECTIVENESS

According to the paper *"A Framework to Model and Measure System Effectiveness"*[14], a baseline definition for effectiveness is:

"Measure of Effectiveness is a measure of the ability of a system to meet its specified needs from a particular viewpoint. This measure may be quantitative or qualitative, allowing comparable systems to be ranked". - Neill Smith and Thea Clark [14].

In other words, the criterion of effectiveness seeks to evaluate the extent to which an intervention has achieved the expected objectives and the undesired effects, i.e., it seeks to identify both positive and negative results. Moreover, the importance of achieving some (but not all) of the objectives should also be examined to draw conclusions about effectiveness.

Thus, effectiveness in the context of face recognition technology can be measured in several ways:

- Degree of success in the implementation of the identified measures through the percentage of regions that have successfully implemented the regulatory frameworks, number of legal and prohibited use cases collected in a registry, number of certifications of technical professionals trained for the ethical development, deployment, and use of FRT.
- Accuracy rate of FRT models for anonymizing biometric characteristics of individuals. The higher the accuracy rate of anonymization, the more effective the system is considered to be.

3.2. EFFICIENCY

The most prominent AI Research Laboratory in the United States, called OpenAI, defined efficiency as *"the capacity to reduce the compute needed to train a specific capability"* [15].

To this definition, I would like to add that efficiency must take into account as Key Performance Indicators (KPIs): 1) training efficiency improvement, 2) performance level, 3) runtime (processing time of the machines), 4) economic cost of development and implementation of the solution, and 5) the scalability of the solution.

For this purpose, a matrix will be created that collects each of the five KIPs to measure efficiency and will be visualized in the form of a spider graph, allowing the analysis and comparison of both the FRT models and the processes and choose those that reflect the highest total efficiency index expressed as a percentage (%).

Thus, efficiency looks to handle a high volume of processes without significant performance degradation or detriment to results.

3.3. EQUITY

The third criterion refers to the degree of equity and fairness of the solutions. It is essential that FRT is developed, deployed, and used in an impartial, fair, and non-discriminatory manner, as well as the bureaucratic and regulatory processes to implement and democratize it throughout society.

Misuse or bias can have negative impacts on marginalized or vulnerable groups. The evaluation of equity should consider both the technical and the ethical perspectives of the technological processes. To this end, the equity criterion should address issues such as:

- The inequality in the accuracy rate and percentage of bias, through quantifying false positives and false negatives, evaluating profiles of all ages, genders, and races.
- Percentage of transparency and explainability of models and processes, evaluating whether the FRT system is transparent and whether decision-making processes are explainable. An equitable technology should enable people to understand how decisions are made, how processes are implemented, and how their data are used.
- Societal awareness index: to properly assess equity, it is necessary to involve all societal stakeholders, but it is indispensable to involve the most affected communities, such as marginalized and vulnerable groups and people of color.

3.4. SUSTAINABILITY

In terms of sustainability, Deep Learning (DL) technology, and more specifically, the Convolutional Neural Network (CNN) that drives Facial Recognition Technologies, consume a high volume of energy, which makes their development and deployment limited to specific platforms that embed Graphics Processing Units (GPUs) (used to accelerate computational processes) capable of managing a high consumption and a long runtime of the computing machines.

Therefore, runtime and power consumption tests are critical for both users and cloud service providers. To identify energy-efficient CNNs, it is crucial to use accurate runtime, power, and energy models.

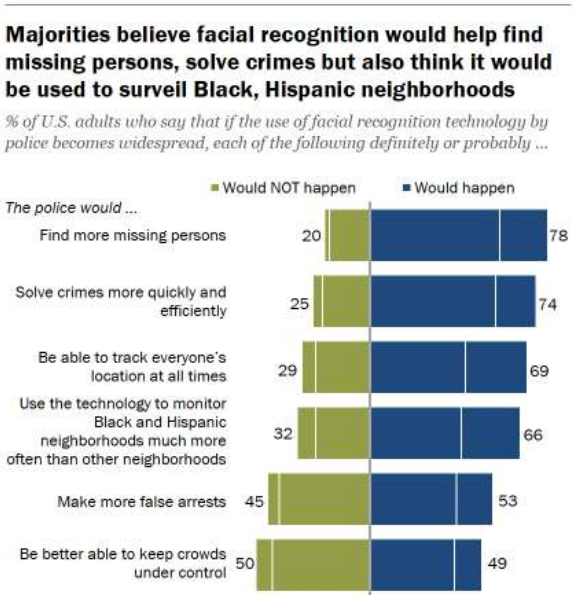
Therefore, as a reference framework to predict the energy consumption of CNNs running on GPUs, we take Ermao Cai's project *"NeuralPower - Learning-based Power and Runtime Modeling for Convolutional Neural Networks"* [16] where he has been able to predict runtime, power and energy of the most advanced CNN architectures, achieving an average accuracy of 88.24% in runtime, 88.34% in power and 97.21% in energy.

Therefore, to measure the sustainability criterion, we start from the reference framework presented in "NeuralPower" [analyzing the processing time of the machines, the power, and the energy resources used], where the threshold for each KPI are the percentages

presented in the study. Thus, above the threshold, the FRT model is considered a sustainable solution, and below it, an non-sustainable solution.

4. EXHIBITS

EXHIBIT #1 – PEW RESEARCH CENTER I



Source: Nadeem (2023)

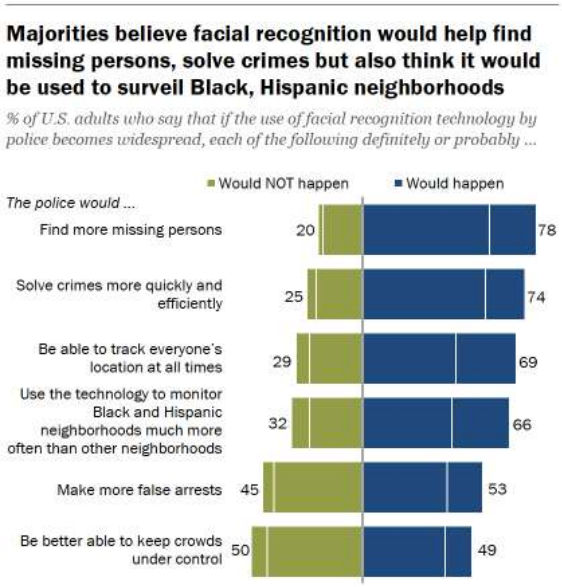
EXHIBIT #2 – FRT TRENDS AND HIGHLIGHTS

Facial Recognition Market Report Highlights

Aspects	Details
Market Size By 2032	USD 24.3 billion
Growth Rate	CAGR of 16.4%
Forecast period	2022 - 2032
Report Pages	278
By Application	• Access Control • Attendance Tracking and Monitoring • Emotion Recognition • Security and Surveillance • Others
By Technology	• 3D • 2D • Facial Analytics
By End User	• Retail and E-commerce • Media and Entertainment • BFSI • Automobile and Transportation • IT and Telecom • Government • Healthcare • Others
By Region	• **North America** (U.S., Canada) • **Europe** (UK, Germany, France, Italy, Spain, Rest of Europe) • **Asia-Pacific** (China, Japan, India, Australia, South Korea, Rest of Asia-Pacific) • **LAMEA** (Latin America, Middle East, Africa)
Key Market Players	NVISO, Cognitec Systems GmbH., FacePhi, Fujitsu, NEC Corporation , Aware, Inc., Thales, Daon, Inc., FaceFirst, Onfido

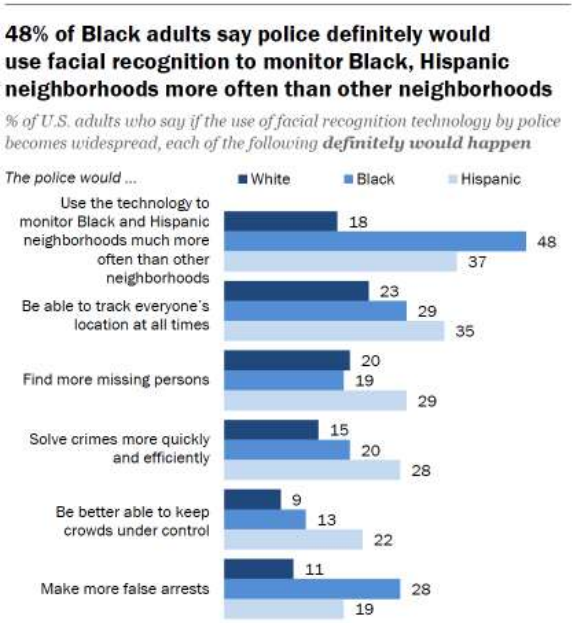
Source: Facial Recognition Market size, share, statistics, global trends, revenue forecast & opportunities | MarketsandMarketsTM. (s. f.).

EXHIBIT #3 - PEW RESEARCH CENTER II



Source: Nadeem (2023)

EXHIBIT #4 - PEW RESEARCH CENTER III



Source: Nadeem (2023)

EXHIBIT 5 - BLUR AND ANONYMIZE FACES

Source: Rosebrock (2021)

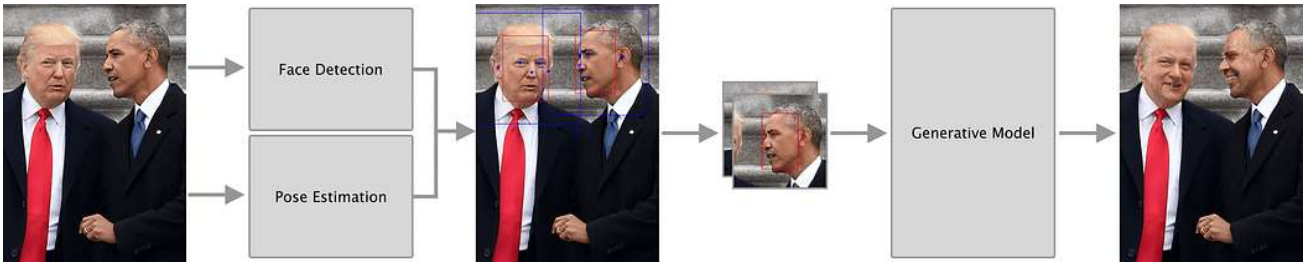
EXHIBIT 6 - DATA MASKING TECHNIQUES

Original Data		
Name	**SSN**	**Age**
Brown	180-80-0724	54
Duncan	252-38-1786	24

Masked Data		
Name	**SSN**	**Age**
Michaels	111-22-3333	52
Lee	111-22-3334	22

Source: Bales, A., & Fritsch, J. (2023)

EXHIBIT 7 - ADVERSARIAL GENERATIVE NETWORKS (GANS)



Source: PrivacyNet: Semi-Adversarial Networks for Multi-Attribute Face Privacy. (2020)

5. REFERENCES

[1] Harwell, D. (2019, December 19). Federal study confirms racial bias of many facial-recognition systems, casts doubt on their expanding use. *The Washington Post.* . https://www.washingtonpost.com/technology/2019/12/19/federal-study-confirms-racial-bias-many-facial-recognition-systems-casts-doubt-their-expanding-use/

[2] Wessler, N. F. (2020, February 5). *We're Taking Clearview A.I. to Court to End its Privacy-Destroying Face Surveillance Activities*. https://www.aclu.org/news/privacy-technology/were-taking-clearview-ai-to-court-to-end-its-privacy-destroying-face-surveillance-activities

[3] New Jersey v. Francisco Arteaga (June 7, 2023). A-3078-21. Superior Court Of New Jersey Appellate Division. https://www.njcourts.gov/system/files/court-opinions/2023/a3078-21.pdf

[4] Cagle, M., & Ozer, N. (2018, May 22). *Amazon Teams Up With Government to Deploy Dangerous New Facial Recognition Technology*. https://www.aclu.org/news/privacy-technology/amazon-teams-government-deploy-dangerous-new

[5] Nadeem, R. (2023, 1 marzo). *Public more likely to see facial recognition technology use by police as good for society | Pew Research Center*. Pew Research Center: Internet, Science & Tech. https://www.pewresearch.org/internet/2022/03/17/public-more-likely-to-see-facial-recognition-use-by-police-as-good-rather-than-bad-for-society/

[6] *Facial Recognition Market Size, Share | Forecast - 2032*. (s. f.). Allied Market Research. https://www.alliedmarketresearch.com/facial-recognition-market

[7] *Facial Recognition Market size, share, statistics, global trends, revenue forecast & opportunities | MarketsandMarketsTM*. (s. f.). MarketsandMarkets. https://www.marketsandmarkets.com/Market-Reports/facial-recognition-market-995.html

[8] *T*argeted anonymization: a face image anonymization method for unauthorized models. (2022, 18 julio). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/9859898

[9] Rosebrock, A. (2021, 17 abril). *Blur and anonymize faces with OpenCV and Python - PyImageSearch*. PyImageSearch. https://pyimagesearch.com/2020/04/06/blur-and-anonymize-faces-with-opencv-and-python/

[10] Bales, A., & Fritsch, J. (2023). *Market Guide for Data Masking*. https://www.gartner.com/doc/reprints?id=1-2CKW1WNW&ct=230215&st=sb&utm_campaign=TY%20Mailers&utm_medium=email&

_hsmi=231288347&_hsenc=p2ANqtz--7g0FRvcp1is5IIAGoStFUZdcQfqEOfp9ikAh_i1QYaCCgsc5GI8-vGFjENiz7T8GTxvhH2d561avKt43821tXvjIpug&utm_content=231288347&utm_source=hs_automation

[11] *Model-Based Face De-Identification*. (2006). IEEE Conference Publication | IEEE Xplore. https://ieeexplore.ieee.org/document/1640608

[12] *PrivacyNet: Semi-Adversarial Networks for Multi-Attribute Face Privacy*. (2020). IEEE Journals & Magazine | IEEE Xplore. https://ieeexplore.ieee.org/document/9201364

[13] Ren, Z., Lee, Y. J., & Ryoo, M. S. (2018). *Learning to Anonymize Faces for Privacy Preserving Action Detection*. https://par.nsf.gov/servlets/purl/10100521

[14] Smith, N., & Clark, T. (s. f.). *A Framework to Model and Measure System Effectiveness*. http://www.dodccrp.org/events/11th_ICCRTS/html/papers/054.pdf

[15] *AI and efficiency*. (s. f.). https://openai.com/research/ai-and-efficiency

[16] Cai, E. (2017). *Learning-based Power and Runtime Modeling for Convolutional Neural Networks*. https://www.ml.cmu.edu/research/dap-papers/S18/dap-cai-ermao.pdf