Research, Society and Development, v. 14, n. 2, e3014248186, 2025
(CC BY 4.0) | ISSN 2525-3409 | DOI: http://dx.doi.org/10.33448/rsd-v14i2.48186

Influence of racial bias in the use of facial recognition applied to access control: A critical analysis

Influência do viés racial no uso do reconhecimento facial aplicado ao controle de acesso: Uma análise crítica

Influencia del sesgo racial en el uso del reconocimiento facial aplicado al control de acceso: Un análisis crítico

Received: 01/29/2025 | Revised: 02/06/2025 | Accepted: 02/07/2025 | Published: 02/10/2025

Alexandre Mundim de Oliveira
ORCID: https://orcid.org/0000-0002-7991-2883
University of Brasília, Brazil
E-mail: alexandremundimdeoliveira@gmail.com
Hugo Xavier Rodrigues
ORCID: https://orcid.org/0009-0006-9565-1720
University of Brasília, Brazil
E-mail: rx.oguh@gmail.com
Alexandre Solon Nery
ORCID: https://orcid.org/0000-0002-3199-4322
University of Brasília, Brazil
E-mail: anery@unb.br
Fábio Lúcio Lopes de Mendonça
ORCID: https://orcid.org/0000-0001-7100-7304
University of Brasília, Brazil
E-mail: fabio.mendonca@redes.unb.br
Luiz Antonio Ribeiro Junior
ORCID: https://orcid.org/0000-0001-7468-2946
University of Brasília, Brazil
E-mail: ribeirojr@unb.br

Abstract
Racial bias has been a persistent issue in facial recognition technologies, particularly within access control applications. This study aims to examine the widespread adoption of these technologies in the machine learning era, highlighting their integration into information security, cybersecurity, and data privacy frameworks. Despite their growing prevalence, the underlying datasets and algorithms frequently exhibit significant biases, disproportionately impacting individuals from marginalized racial groups. Through an extensive literature review, this research identifies critical gaps and proposes 14 targeted recommendations aimed at mitigating racial bias in facial recognition systems. These recommendations encompass diversifying training datasets, enhancing algorithmic transparency, and incorporating multidisciplinary teams to ensure ethical decision-making. The findings underscore the potential to improve both the equity and accuracy of these technologies, paving the way for more reliable and inclusive applications. By implementing the proposed measures, stakeholders can address ethical concerns, reduce discriminatory outcomes, and enhance public trust in the adoption of facial recognition for sensitive access control contexts. This critical analysis provides a roadmap for advancing fairness and accountability in artificial intelligence, fostering transformative impacts in the field.
Keywords: Face recognition; Machine learning; Artificial intelligence; Access control; Cybersecurity; Racial bias; Algorithmic racism.

Resumo
O viés racial tem sido uma questão persistente nas tecnologias de reconhecimento facial, particularmente em aplicações de controle de acesso. Este estudo tem como objetivo examinar a adoção generalizada dessas tecnologias na era do aprendizado de máquina, destacando sua integração em frameworks de segurança da informação, cibersegurança e privacidade de dados. Apesar de sua crescente prevalência, os conjuntos de dados e algoritmos subjacentes frequentemente apresentam vieses significativos, impactando desproporcionalmente indivíduos de grupos raciais marginalizados. Por meio de uma revisão extensa da literatura, esta pesquisa identifica lacunas críticas e propõe 14 recomendações direcionadas com o objetivo de mitigar o viés racial nos sistemas de reconhecimento facial. Essas recomendações abrangem a diversificação dos conjuntos de dados de treinamento, o aprimoramento da transparência algorítmica e a incorporação de equipes multidisciplinares para garantir a tomada de decisões éticas. Os

resultados destacam o potencial para melhorar tanto a equidade quanto a precisão dessas tecnologias, abrindo caminho para aplicações mais confiáveis e inclusivas. Ao implementar as medidas propostas, as partes interessadas podem abordar questões éticas, reduzir os resultados discriminatórios e aumentar a confiança pública na adoção do reconhecimento facial para contextos sensíveis de controle de acesso. Esta análise crítica fornece um roteiro para promover a justiça e a responsabilidade na inteligência artificial, fomentando impactos transformadores na área.
Palavras-chave: Reconhecimento facial; Aprendizado de máquina; Inteligência artificial; Controle de acesso; Cibersegurança; Viés racial; Racismo algorítmico.

Resumen
El sesgo racial ha sido un problema persistente en las tecnologías de reconocimiento facial, particularmente en las aplicaciones de control de acceso. Este estudio tiene como objetivo examinar la adopción generalizada de estas tecnologías en la era del aprendizaje automático, destacando su integración en los marcos de seguridad de la información, ciberseguridad y privacidad de datos. A pesar de su creciente prevalencia, los conjuntos de datos y algoritmos subyacentes frecuentemente exhiben sesgos significativos, impactando desproporcionadamente a individuos de grupos raciales marginados. A través de una extensa revisión de la literatura, esta investigación identifica brechas críticas y propone 14 recomendaciones específicas dirigidas a mitigar el sesgo racial en los sistemas de reconocimiento facial. Estas recomendaciones abarcan la diversificación de los conjuntos de datos de entrenamiento, el fortalecimiento de la transparencia algorítmica y la incorporación de equipos multidisciplinarios para garantizar la toma de decisiones éticas. Los hallazgos subrayan el potencial para mejorar tanto la equidad como la precisión de estas tecnologías, allanando el camino para aplicaciones más confiables e inclusivas. Al implementar las medidas propuestas, las partes interesadas pueden abordar las preocupaciones éticas, reducir los resultados discriminatorios y aumentar la confianza pública en la adopción del reconocimiento facial para contextos sensibles de control de acceso. Este análisis crítico proporciona una hoja de ruta para avanzar en la equidad y la responsabilidad en la inteligencia artificial, fomentando impactos transformadores en el campo.
Palabras clave: Reconocimiento facial; Aprendizaje automático; Inteligencia artificial; Control de acceso; Ciberseguridad; Sesgo racial; Racismo algorítmico.

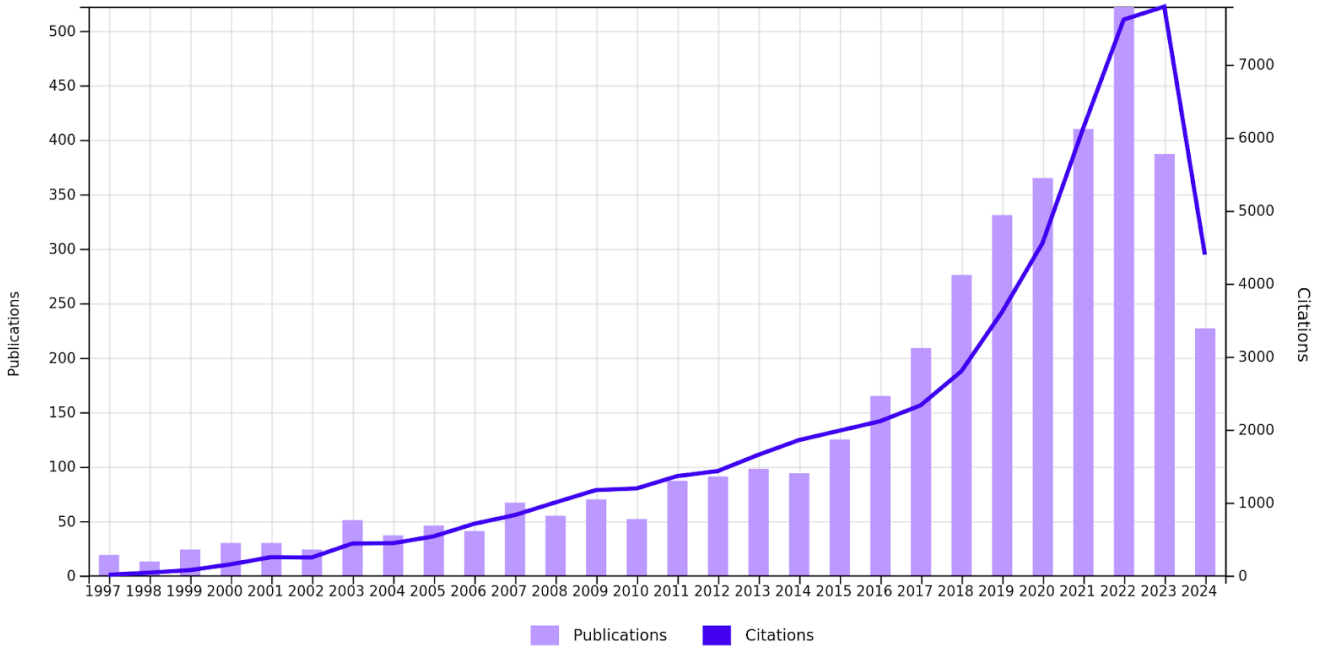
1. Introduction

Facial recognition, a technology that employs algorithms to analyze human facial features through cameras (Zhao et al., 2003), has found wide application in a variety of fields, such as security (Ratha et al., 2001), monitoring (Llauradó et al., 2023), and biometric authentication (Jain et al., 2004). Its use as a layer of access control plays a significant role in ensuring the three fundamental pillars of Information Security: integrity, availability, and authenticity (Caballero, 2017). These principles are essential for security in both physical and logical environments, spanning areas from border protection (Wayman, 2008) to cybersecurity (Cornish, 2021) and data privacy (Solove, 2006).

In the banking sector, for example, institutions such as Hong Kong and Shanghai Banking Corporation (HSBC) have adopted this technology in their mobile apps (HSBC, n.d.), providing customers with a safer and more convenient banking experience while actively combating fraud. Additionally, facial recognition is used in the cybersecurity sphere to secure access to devices and networks (Cantor, 2020). Microsoft, for example, launched Windows Hello (Microsoft, 2021), a facial recognition-based authentication system that allows users to unlock their devices with a simple glance.

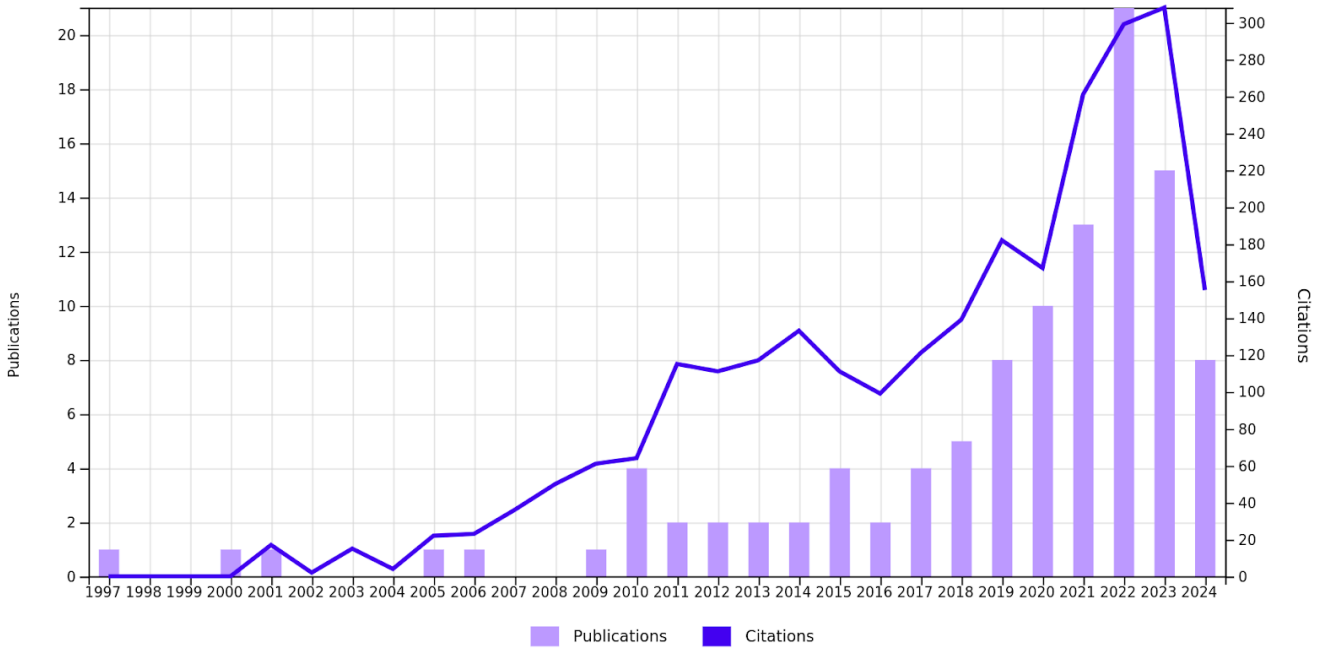
Facial recognition has emerged as a transformative technology in the contemporary landscape, driven by the rapid advancement of computer vision and machine learning, as highlighted by Jain et al. (2004). This approach stands out as one of the most actively studied areas in computer vision and pattern recognition, being applied in a variety of areas and sectors (Zhao et al., 2003). However, one of the recognized challenges in this field is the issue of racial bias, which has been the subject of in-depth studies and analysis (Buolamwini & Gebru, 2018; Whittaker, 2018; Kärkkäinen & Joo, 2021; Raji & Buolamwini, 2019; Najibi, 2020). This is a rapidly growing field of study, with the number of publications and citations growing each year, as shown in Figures 1 and 2.

Figure 1 - Evolution of publications and citations, by year, from 1997 to 2024, with titles containing the term “Facial Recognition”.



Source: Web of Science (2024, August 2).

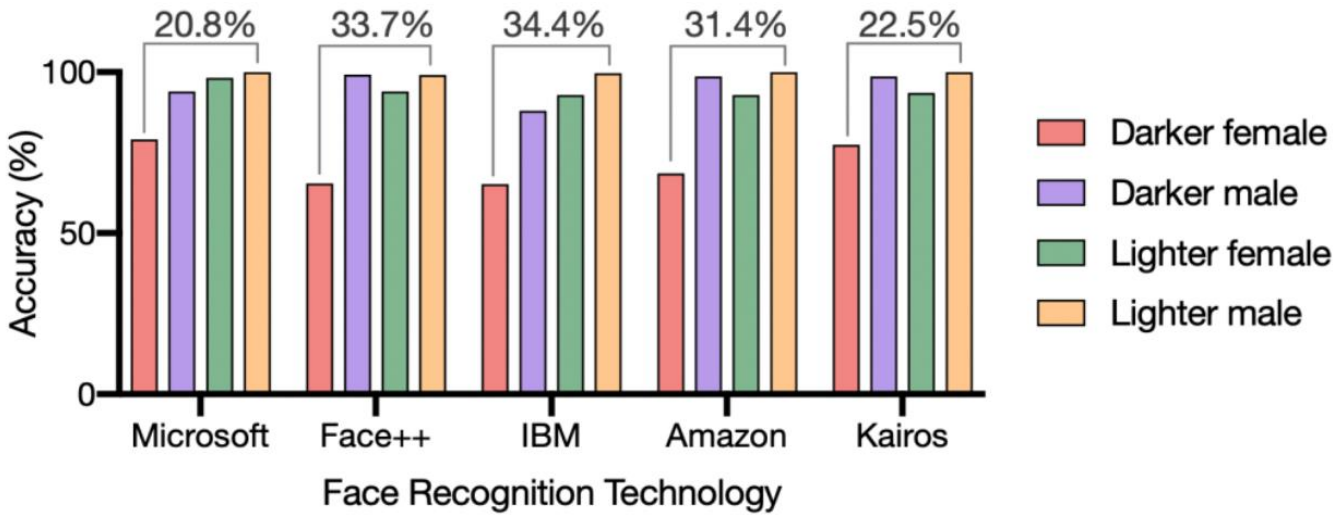
Figure 2 - Evolution of publications and citations, by year, from 1997 to 2024, with titles containing the terms "Facial Recognition" and "Racial Bias".



Source: Web of Science (2024, August 2).

Recently, Buolamwini and Gebru (2018) conducted a comprehensive analysis of accuracy disparities in gender classification systems that rely on facial recognition technology. Their findings highlighted how minority ethnic groups, especially women with darker skin tones, were frequently subjected to inaccurate classifications, as shown in Figure 3.

Figure 3 - Accuracy of face recognition technologies. An audit of five face recognition technologies revealed discrepancies in the classification accuracy of these technologies for different skin tones and sexes. The algorithms consistently demonstrated the poorest accuracy for darker-skinned females and better accuracy for lighter-skinned males. This groundbreaking study illuminated the inherent dangers of algorithmic biases and prejudices, highlighting the crucial importance of technological fairness. Additionally, it underscored the urgent need to adopt more equitable and inclusive approaches in the research and deployment of artificial intelligence systems.



Source: Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification.

As described by Buolamwini and Gebru (2018), racial bias refers to unconscious tendencies or prejudices that influence treatment and decisions regarding different racial groups. These biases can manifest themselves in a variety of ways, including negative stereotypes, discrimination, and unequal treatment based on race. In the context of facial recognition, these biases can be incorporated into systems due to several factors, such as imbalances in training datasets, biased algorithms, and a lack of diversity among developers of these technologies (Hua et al., 2011; Buolamwini & Gebru, 2018; Raji & Buolamwini, 2019; Najibi, 2020; Klare et al., 2012; Grother et al., 2019; Benjamin, 2019; Larson et al., 2016; Gordon, 2019; Stanley, 2019; Lynch, 2020). This scenario raises increasing concerns about the accuracy and potential discriminatory impact of these systems, as well as ethical and moral issues that must be considered in the development of Artificial Intelligence, as advocated by the AI Now Institute at New York University (Whittaker, 2018), to mitigate “algorithmic racism”.

Several additional studies have investigated the influence of racial bias in facial recognition. For example, a study led by Klare et al. (2012) analyzed the accuracy of facial recognition systems in different racial groups, concluding that there are significant disparities in the performance of these technologies based on race. Another study by Grother et al. (2019) examined the representation of race in the data sets used to train facial recognition algorithms and found that lack of diversity in these data sets can contribute to bias. According to Kärkkäinen and Joo (2021), existing public facial image datasets have a solid tendency to predominantly feature Caucasian faces, while other races are significantly underrepresented. This scenario leads to models trained with inconsistent classification, limiting the applicability of facial recognition analytics systems to non-white racial groups. Furthermore, a study by the MIT Media Lab (Raji & Buolamwini, 2019) identified that commercial facial recognition systems exhibited higher error rates when identifying individuals with darker skin and women than individuals with lighter skin and men. According to Eubanks (Gordon, 2019), the lack of diversity and representation in the development teams of these technologies contributes to the perpetuation of injustices.

In another study, Benjamin (2019) highlights that modern technologies, including algorithms and artificial intelligence, can perpetuate racial biases and social inequalities. It is argued that technology is not neutral and often reflects

and amplifies biases that already exist in society. The study advocates for a "technological abolitionism" movement as new technologies emerge to dismantle discriminatory structures embedded in this technology. Additionally, the study emphasizes the importance of inclusion and diversity in technology development, highlighting that the lack of diversity between developers and decision-makers contributes to racial bias and discrimination in algorithms and automated systems.

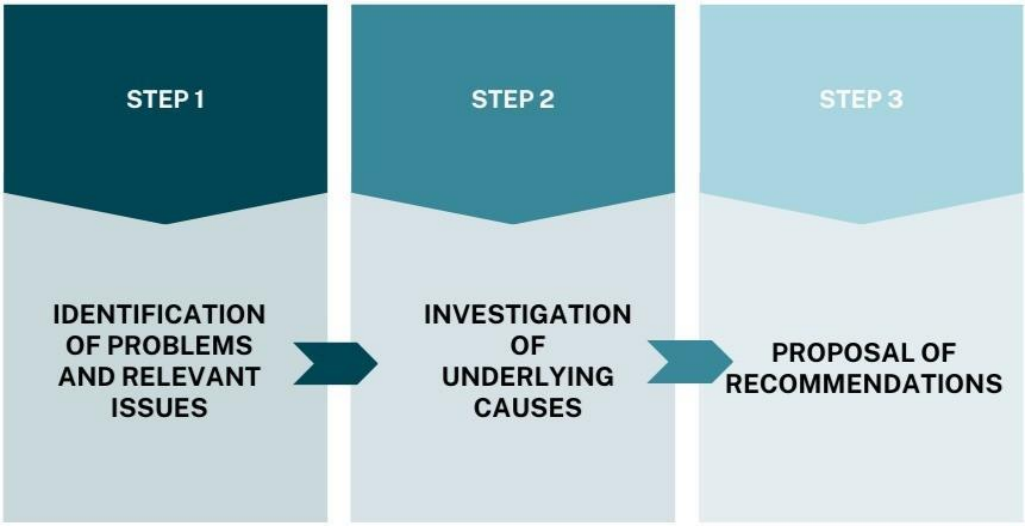
The work of Larson et al. (2016) suggests that the COMPAS algorithm, used to make judicial decisions in the United States, exhibited substantial racial bias. The algorithm disproportionately predicted that black criminals were more likely to re-offend than white criminals, even when the characteristics of the cases were similar. This trend raises concerns about how these algorithms can influence sentencing and prison terms, potentially exacerbating racial disparities in the criminal justice system. The analysis also highlights the lack of transparency regarding the functioning of COMPAS. The authors found it difficult to obtain detailed information about how the algorithm was developed and how it made its predictions, which raises questions about transparency, accountability, and responsibility in the use of such systems in society. In this context, it is essential to study the impact of racial bias on facial recognition systems to improve the efficiency of these applications and eliminate this harmful bias in using this technology.

This study aims to examine the widespread adoption of these technologies in the machine learning era, highlighting their integration into information security, cybersecurity, and data privacy frameworks. Through a critical analysis of artificial intelligence approaches focused on data analysis and engineering, the study led to the development of 14 recommendations aimed at mitigating racial bias in facial recognition systems.

2. Methodology

Our work involved a careful and in-depth analysis through a review of recent scientific studies that explore the problem of racial bias in the use of facial recognition (Hua et al., 2011; Buolamwini & Gebru, 2018; Klare et al., 2012; Grother et al., 2019; Benjamin, 2019; Larson et al., 2016; Raji & Buolamwini, 2019; Najibi, 2020; Gordon, 2019; Stanley, 2019; Lynch, 2020). Initially, the search for publications was conducted in the Web of Science database using the following filtering criteria: the publication period from January 1, 1997, to August 2, 2024, and the expressions of "Facial Recognition" and "Racial Bias" in all fields. This search yielded 108 publications (Figure 2). Refining the search of the titles and abstracts of these 108 publications allowed the selection to be narrowed down to 28 articles. The selection considered publications with the highest number of citations as an indicator of scientific relevance, and the full-text reading of these articles resulted in the final selection of 12 works; the final selection of articles was based on adherence to the research objective: investigating racial bias in the use of facial recognition in access control systems. Furthermore, we included reports from prominent organizations and institutions, such as ACLU (Stanley, 2019), Electronic Frontier Foundation (Lynch, 2020), and NIST (Grother et al., 2019), which conducted investigations on the impacts of racial bias in facial recognition systems. Figure 4 shows the flow chart carried out in this investigation, followed by the details of the respective steps:

Figure 4 - Flowchart with the steps process conducted in this work.



Source: Authors.

Step 1: Identification of problems and relevant issues: Initially, the issues related to facial recognition and racial bias were identified through a review of scientific literature, reports from relevant organizations, and previous research;

Step 2: Investigation of underlying causes: In addition to identifying the bias, it was crucial to investigate the underlying causes of this bias. This included analysis of how facial recognition algorithms are developed, the training datasets used, and the data engineering techniques applied; and,

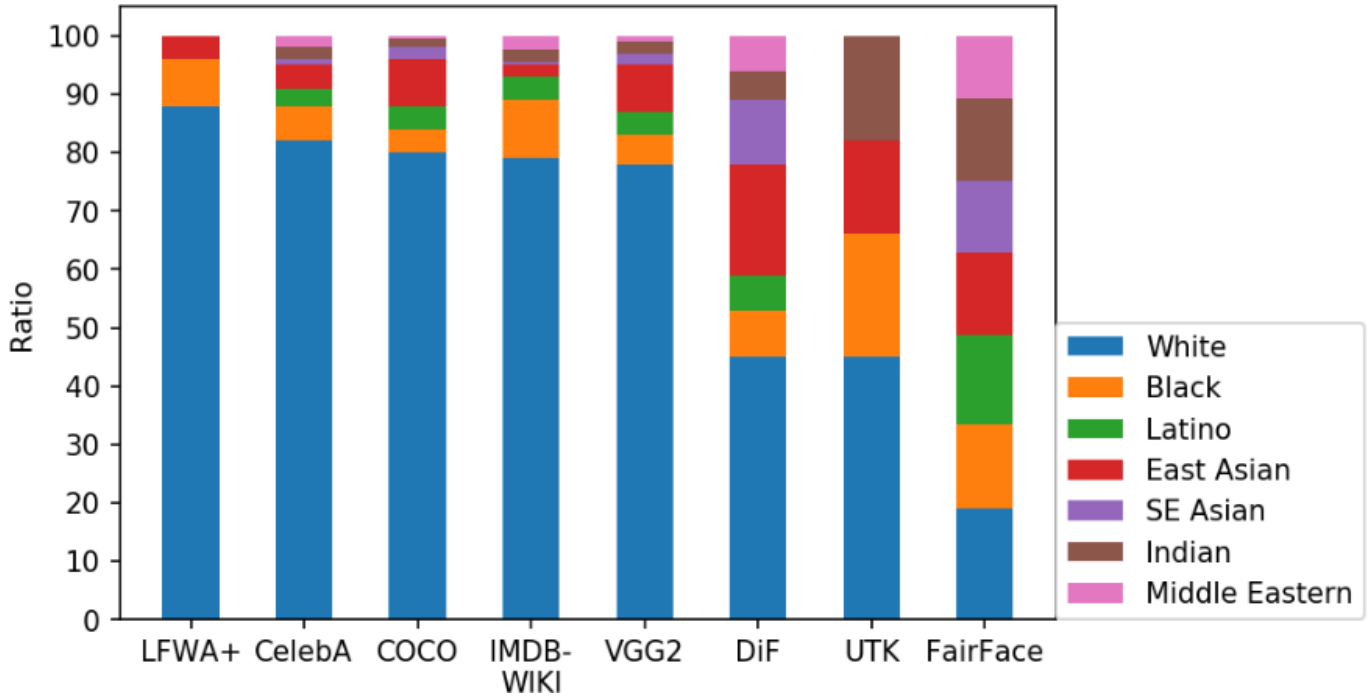
Step 3: Proposition of recommendations: After identifying the underlying causes, specific recommendations were proposed to mitigate racial bias in facial recognition systems.

3. Results and Discussion

Numerous studies have found clear evidence of racial bias in facial recognition systems (Hua et al., 2011; Buolamwini & Gebru, 2018; Klare et al., 2012; Grother et al., 2019; Benjamin, 2019; Raji & Buolamwini, 2019; Najibi, 2020; Gordon, 2019; Stanley, 2019; Lynch, 2020; Larson et al., 2016). For example, individuals with darker skin are more likely to be falsely identified, which can lead to unfair treatment, such as being denied access to services (Kärkkäinen & Joo, 2021). The lack of diversity in the training data used to train facial recognition algorithms is a significant cause of racial bias in this technology. When training data does not adequately represent the diversity of the population, systems can misidentify the faces of underrepresented groups (Grother et al., 2019).

In recent research (Kärkkäinen & Joo, 2021), the goal was to create a facial attribute dataset balanced in terms of race, gender, and age. The researchers argue that existing facial attribute datasets are often biased, which can lead to biases in machine learning systems that are trained with these data. The researchers evaluated the FairFace dataset using various methods and found that it is significantly more balanced than existing datasets. Additionally, they found that machine learning systems trained with the FairFace dataset are less likely to be biased. Imbalances in commonly used facial attribute datasets (LFWA+, CelebA, COCO, IMDB-WIKI, VGG2, DiF, and UTK) and the balanced nature of the FairFace dataset are demonstrated in Figure 5.

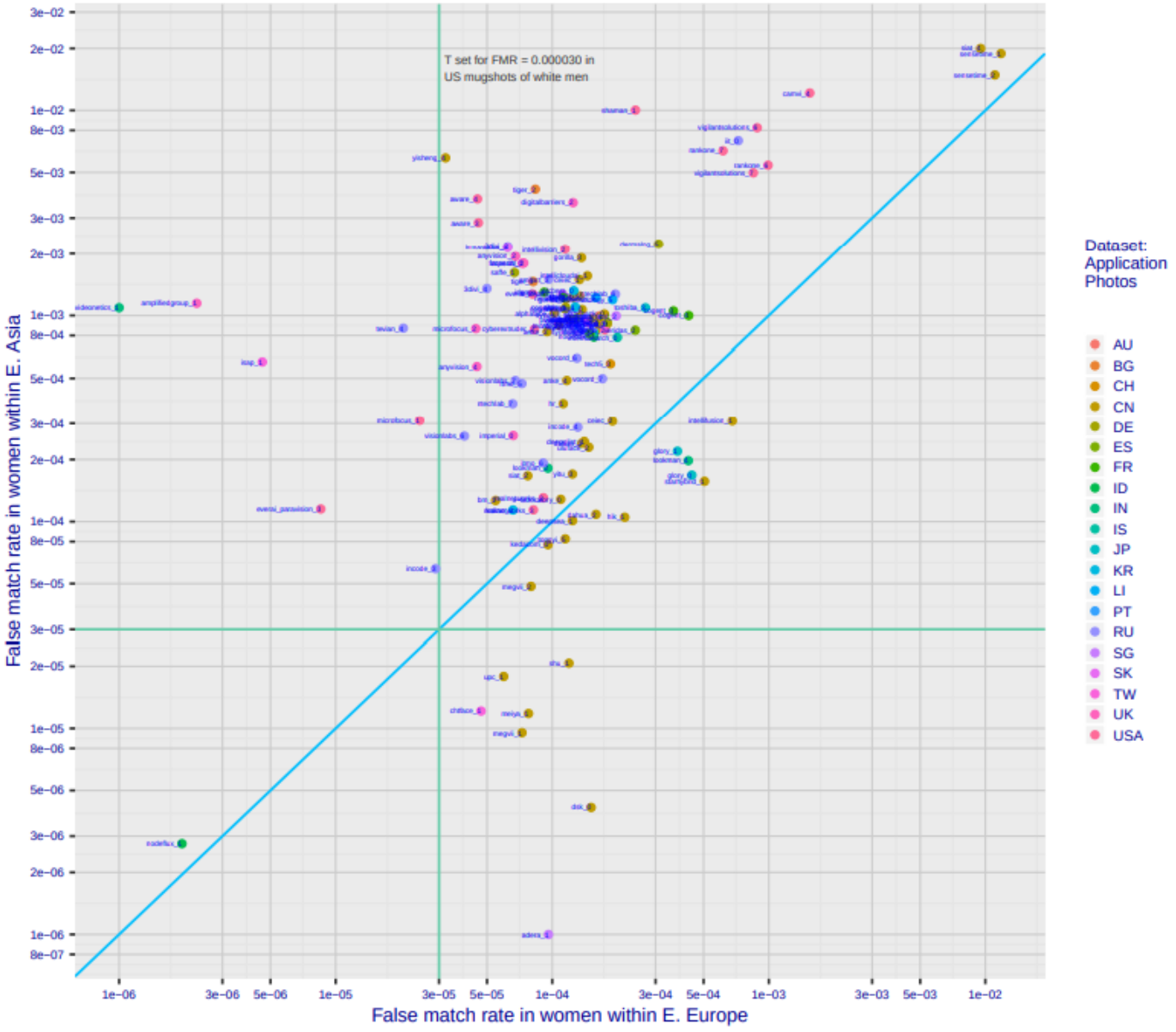
Figure 5 - Racial compositions in face datasets.



Source: FairFace dataset (Kärkkäinen & Joo, 2019).

The LFWA+ dataset, with its labeled facial attributes, provides a foundation for evaluating face recognition systems in real-world scenarios. CelebA, with its extensive collection of over 200,000 celebrity images and 40 attribute labels, offers a rich dataset for facial attribute analysis, contributing to the model's generalization capabilities. COCO, known for its large-scale object detection and segmentation tasks, enhances the robustness of the model by introducing a diverse range of objects and contexts, which are crucial for non-facial recognition tasks. The IMDB-WIKI dataset, focusing on age and gender labels across a broad range of celebrity faces, aids in developing comprehensive facial recognition models. VGG2, with its high variability in pose, age, illumination, and ethnicity, is pivotal in improving face recognition accuracy under diverse conditions. The DiF dataset, emphasizing diversity, is critical in reducing biases and enhancing fairness in facial recognition across different demographic groups. Lastly, UTKFace, with its balanced representation of age, gender, and ethnicity, complements the other datasets by providing a robust foundation for demographic analysis. In Figure 6, false match rate compared between women in countries of Eastern Europe and East Asia.

Figure 6 - Schematic representation for false match rates (FMR) comparison of women of the same age. The countries considered are Poland, Russia, Ukraine, China, Japan, Korea, Philippines, Thailand, and Vietnam. The green lines are the threshold for each algorithm regarding the FMR data collected for white men in the U.S. mugshot database. The blue line (y = x) indicates parity. The color code denotes the developer's domicile, as research and training data may stem from different locations.



Source: NIST (2019), Face Recognition Vendor Test Part 3: Demographic Effects.

In information and cybersecurity, inaccurate facial recognition can be a gateway to unauthorized access to sensitive data and systems (Wayman, 2008). Cybercriminals could exploit biased algorithms in security systems, potentially bypassing checks or impersonating authorized users through deepfakes. It is crucial to note that deepfake generation systems may also inherit biases if trained on imbalanced datasets like those discussed in this work. By prioritizing accuracy, we can build robust defenses against such attacks, safeguarding confidential information and critical infrastructure. Privacy concerns also come into sharp focus. Misidentification in biased systems can lead to unwarranted surveillance, data breaches, and even discriminatory treatment. Ensuring accurate identification protects individuals from infringements and upholds crucial data

privacy regulations. Accurate facial recognition is equally essential for effective access control. Imprecise systems risk granting access to unauthorized individuals and jeopardizing physical assets, intellectual property, and confidential information. In contrast, accurate identification strengthens security measures, protects restricted areas, and streamlines access processes without compromising safety.

In this job, we focused on understanding the root causes of the identified racial bias in each reference, seeking to understand the factors contributing to bias in facial recognition systems. Table 1 presents the results of Steps 1 and 2 described in the methodology section. Step 1 involved identifying relevant problems and issues through a literature review, while Step 2 involved investigating the underlying causes of these problems and issues.

Table 1 - Underlying causes listed in a literature review.

Underlying causes	Reference	Problems and relevant issues
1. Imbalance in training datasets	(Buolamwini & Gebru, 2018; Grother et al., 2019; Kärkkäinen & Joo, 2021; Raji & Buolamwini, 2019; Najibi, 2020)	Highlighted the importance of diverse training datasets. The lack of representation of various racial groups in the training data contributes to biased algorithms.
2. Development of algorithms without systematic evaluations over time	(Benjamin, 2019; Klare et al., 2012; Hua et al., 2011)	Pointed out issues with algorithms lacking continuous evaluations. Algorithms developed without ongoing assessments can perpetuate biases over time.
3. Lack of continuous monitoring systems	(Klare et al., 2012; Hua et al., 2011)	Emphasizing the absence of continuous monitoring systems. This gap allows biases to persist without timely intervention.
4. Inadequate configuration of system parameters	(Benjamin, 2019; Buolamwini & Gebru, 2018)	Discussed the impact of poorly configured system parameters on biased performance among ethnic groups.
5. Absence of a multidisciplinary team	(Raji & Buolamwini, 2019; Gordon, 2019)	Highlighted the importance of a diverse team with expertise in ethics, diversity, machine learning, and civil rights for comprehensive decision making.
6. Lack of transparency in system decisions and the absence of activity logs	(Grother et al., 2019; Raji & Buolamwini, 2019)	Emphasized the importance of transparent systems with activity logs for effective bias detection.
7. Absence of external testing and audits by independent organizations	(Grother et al., 2019; Raji & Buolamwini, 2019)	Highlighted the need for external testing and audits for impartial assessment of racial bias.
8. Lack of constant user feedback	(Raji & Buolamwini, 2019)	Emphasized the importance of continuous feedback, especially from users susceptible to racial bias.
9. Exclusive reliance on facial recognition in critical decisions	(Larson et al., 2016; Klare et al., 2012)	Discussed the risks associated with exclusive reliance on facial recognition in critical decisions, contributing to discriminatory outcomes.
10. Lack of regular updates	(Benjamin, 2019; Hua et al., 2011)	Highlighted the importance of staying updated to align facial recognition systems with the best bias mitigation approaches.
11. Absence of clear policies and guidelines	(Wayman, 2008; Raji & Buolamwini, 2019)	Pointed out the lack of clear policies for the ethical use of facial recognition, considering bias.

Source: Research data.

This section has comprehensively analyzed the intricate factors underlying racial bias in facial recognition systems. Building on the critical issues identified in Step 2, it highlights key drivers, including unbalanced training data, inadequate algorithm evaluation over time, and lack of continuous monitoring. These revelations demand proactive measures to combat bias throughout the development and deployment of this technology. In the next section, "Recommendations," we focus on actionable solutions. We present a carefully curated set of 14 recommendations based on our findings. These recommendations guide developers, policymakers, and industry stakeholders, propelling the development of a more inclusive, transparent, and equitable future for facial recognition technology.

3.1 Recommendations

Based on the identification of 11 (eleven) underlying causes carried out through a literature review and in the light of the results and discussions, it was possible to suggest a comprehensive set of 14 (fourteen) recommendations to mitigate the causes as detailed in Table 2:

Table 2 - Correlation between recommendations and underlying causes.

Recommendation	Underlying causes
1. Diversity in training data	1. Imbalance in training datasets
2. Regular bias assessments	2. Development of algorithms without systematic evaluations over time
3. Continuous monitoring	2. Development of algorithms without systematic evaluations over time 3. Lack of continuous monitoring systems
4. Parameter adjustment	1. Imbalance in training datasets 3. Lack of continuous monitoring systems 4. Inadequate configuration of system parameters
5. Multidisciplinary team	5. Absence of a multidisciplinary team
6. Transparency	6. Lack of transparency in system decisions and a lack of activity logs
7. External testing and audits	3. Lack of continuous monitoring systems 6. Lack of transparency in system decisions and a lack of activity logs 7. Absence of external tests and audits by independent organizations
8. User feedback	6. Lack of transparency in system decisions and lack of activity logs 7. Absence of external tests and audits by independent organizations 8. Lack of constant feedback from system users
9. Limited use in critical decisions	4. Inadequate configuration of system parameters 9. Exclusive reliance on facial recognition in critical decisions
10. Regular update	5. Absence of a multidisciplinary team 10. Lack of regular updates
11. Policies and guidelines	8. Lack of constant feedback from system users 10. Lack of regular updates 11. Absence of clear policies and guidelines
12. Scope of Operation	7. Absence of external tests and audits by independent organizations 8. Lack of constant feedback from system users
13. User consent	11. Absence of clear policies and guidelines
14. Education and awareness	9. Exclusive reliance on facial recognition in critical decisions 10. Lack of regular updates

Source: Research data.

Table 3 summarizes the expected positive impacts of implementing the suggested recommendations. The table provides an overview of the benefits that can be achieved by following the proposed guidelines. By implementing these recommendations, improvements are expected in mitigating racial bias in systems that use facial recognition for physical or logical access control. Overall, the table is helpful for organizations looking to improve their operations and enhance their security posture. Importantly, the supplementary material addresses the impact of other recommendations in mitigating racial bias in facial recognition systems, such as increasing diversity in training data and fine-tuning parameters. This material includes an experiment with various machine learning models trained to classify face images, demonstrating that greater diversity in datasets, coupled with proper parameter adjustment, can effectively reduce bias. By incorporating more diverse datasets, models can more accurately categorize face images by race, leading to better identification of these images and their unique characteristics.

Table 3 - Recommendations and their positive impacts.

ID - Recommendation	Positive impact
1. Diversity in training data	Ensuring the training dataset is diverse in terms of race, ethnicity, gender, age, and other relevant characteristics. Enhance system inclusivity and accuracy, avoiding the dominance of a single demographic
2. Regular bias assessment	Conduct regular system assessments to identify any racial bias that may emerge over time Allow regular system evaluations to ensure timely identification and correction of emerging racial bias, improving equity over time
3. Continuous monitoring	Implement continuous monitoring systems to identify and address possible bias issues as they arise Help ensure the system remains fair over time and avoids the amplification of existing bias.
4. Parameter adjustment	Configure facial recognition system parameters appropriately for balanced performance across different ethnic groups, including adjusting decision thresholds Ensuring fair results to avoid discrimination
5. Multidisciplinary team	Form a multidisciplinary team, including experts in ethics, diversity, machine learning, and civil rights. Ensure a comprehensive approach to addressing racial bias issues and making assertive decisions.
6. Transparency	Ensuring the system is transparent, making its decisions understandable and logging activities Enable effective detection and correction of biases
7. External testing and audits	Subject the system to external testing and audits by independent organizations to impartially assess the presence of racial bias Ensure reliable and fair results
8. User feedback	Solicit constant feedback from system users, especially those susceptible to racial bias Help identify issues and continually improve the system
9. Limited use in critical decisions	Avoid relying solely on facial recognition in critical decisions, such as arrests or hiring. Use it as a supplementary tool combined with other information for more assertive decision-making. Reduce the risk of discrimination and serious errors.
10. Regular updates	Stay updated with the latest research and recommended practices to align the facial recognition system with the best bias mitigation approaches Continuous improvement to mitigate bias through error correction, application updates, algorithm enhancements, and dataset improvements
11. Policies and guidelines	Establish clear policies and guidelines for the ethical use of facial recognition, considering bias Promote nondiscriminatory use of facial recognition
12. Scope of Operation	Limit the use of facial recognition in sensitive contexts, such as law enforcement Avoid human rights violations
13. User consent	Ensuring informed consent from individuals regarding the use of facial recognition, whether in public or private settings Respect individuals' autonomy regarding the use of their facial images
14. Education and awareness	Promotion of education and awareness to address racial bias and ensure ethical use of facial recognition. Enhance understanding of the risks and benefits of these technologies.

Source: Research data.

4. Final Considerations

This study critically examined the impact of racial bias in applying facial recognition for access control. The findings underscored the existence of biases, which could lead to unequal and unjust treatment of specific racial groups. This research contributes significantly to our understanding of racial bias in facial recognition. It highlights the imperative of addressing this issue to ensure fairness and equity in implementing this technology.

It is crucial to acknowledge the limitations of this study regarding result interpretation, as subjectivity may influence the analysis of selected studies. However, these limitations were mitigated by using a rigorous selection approach. To further substantiate the potential of our recommendations, we suggest incorporating an additional research stage focused on their real-world impact (Stage 4 in Methodology Figure 4). This would involve evaluating the effectiveness of the proposed measures in mitigating bias after their implementation.

The results of this research have significant ethical and social implications. Racially biased facial recognition systems can lead to discriminatory treatment, privacy violations, and reinforcement of harmful stereotypes. These issues have negative impacts on marginalized and disadvantaged communities by perpetuating systems of oppression and exacerbating existing racial inequalities in society. Therefore, it is crucial to take measures to reduce bias and promote its ethical use.

Acknowledgments

The authors gratefully acknowledge financial support from the Brazilian Research Council (FAP-DF) under grants 00193.00001808/2022-71, 00193.00001857/2023-95, and 00193.00000857/2021-14, as well as the National Council for Scientific and Technological Development (CNPq) under grant 350176/2022-1, and the FAP-DF-PRONEM grant 00193.00001247/2021-20. Additional funding was provided by the PDPG-FAPDF-CAPES Centro-Oeste grant (00193-00000867/2024-94). The authors also acknowledge the technical and computational support provided by the LATITUDE Laboratory at the University of Brasília. This research was supported by grants from the Advocacia Geral da União (TED 01/2019, AGU Grant 697.935/2019), the Administrative Council for Economic Defense (CADE; TED 08700.000047/2019-14), and the National Social Assistance Secretariat (TED 01/2021, SNAS/DGSUAS/CGRS). Additional institutional support was provided by the Dean of Research and Innovation (DPI) at the University of Brasília (UnB), the SISTER City Project – Safe and Real-Time Effective Intelligent Systems for Smart Cities (Grant 625/2022), and the Foundation for Research Support of the Federal District (FAP-DF).

References

Benjamin, R. (2019). Review of "Race After Technology: Abolitionist Tools for the New Jim Code*." Social Forces*, 98(1), 1–3.

Buolamwini, J., & Gebru, T. (2018). Gender shades: Intersectional accuracy disparities in commercial gender classification. *Proceedings of Machine Learning Research*, 81, 1–15.

Caballero, A. (2017). *Computer and information security handbook*. Elsevier.

Cantor, J. R. (2020). Privacy impact assessment for the homeland advanced recognition technology system (HART) increment 1 PIA. Retrieved from https://www.dhs.gov/sites/default/files/publications/privacy-pia-obim004-hartincrement1-february2020_0.pdf [Accessed August 27, 2024].

Cornish, P. (2021). *The Oxford handbook of cyber security*. Oxford University Press

Gordon, F. (2019). Review of Virginia Eubanks (2018) Automating inequality: How high-tech tools profile, police, and punish the poor. New York: Picador/St. Martin's Press. *Law, Technology, and Humans*, 1(1), 162–164.

Grother, P., Ngan, M., & Hanaoka, K. (2019). Face recognition vendor test part 3: Demographic effects. NIST Interagency/Internal Report (NISTIR). Gaithersburg, MD: National Institute of Standards and Technology.

HSBC. (n.d.). Touch ID and Face ID: A simple and secure alternative to your digital security device passcode. Retrieved from https://www.us.hsbc.com/mobile-banking/biometrics/ [Accessed August 27, 2024].

Hua, G., Nasrabadi, N. M., Chellappa, R., & Aggarwal, G. (2011). Introduction to the special section on real-world face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(10), 1921–1924.

Jain, A., Ross, A., & Prabhakar, S. (2004). An introduction to biometric recognition. *IEEE Transactions on Circuits and Systems for Video Technology*, 14(1), 4–20

Kärkkäinen, K., & Joo, J. (2021). FairFace: Face attribute dataset for balanced race, gender, and age for bias measurement and mitigation. 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), 1547–1557.

Klare, B. F., Burge, M. J., Klontz, J. C., Vorder Bruegge, R. W., & Jain, A. K. (2012). Face recognition performance: Role of demographic information. *IEEE Transactions on Information Forensics and Security*, 7(6), 1789–1801.

Larson, J., Mattu, S., Kirchner, L., & Angwin, J. (2016). How we analyzed the COMPAS recidivism algorithm. ProPublica. Retrieved from https://www.propublica.org

Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278–2324.

Llauradó, J., Pujol, F., & Tomás, D. (2023). Study of image sensors for enhanced face recognition at a distance in the smart city context. Scientific Reports, 13, Article 14713.

Lynch, J. (2020). Face off: Law enforcement use of face recognition technology. Electronic Frontier Foundation (EFF), 10, 78–95.

Microsoft. (2021). Windows Hello face authentication. Retrieved from https://learn.microsoft.com/en-us/windows-hardware/design/device-experiences/windows-hello-face-authentication [Accessed August 27, 2024].

Najibi, A. (2020). Racial discrimination in face recognition technology. Retrieved from https://sitn.hms.harvard.edu/flash/2020/racial-discrimination-in-face-recognition-technology/ [Accessed August 27, 2024].

Oloyede, M., Hancke, G., & Myburgh, H. (2020). A review on face recognition systems: Recent approaches and challenges. *Multimedia Tools and Applications*, 279(22), 27891–27922.

Raji, I. D., & Buolamwini, J. (2019). Actionable auditing: Investigating the impact of publicly naming biased performance results of commercial AI products. Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society, 429–435.

Ratha, N. K., Connell, J. H., & Bolle, R. M. (2001). Enhancing security and privacy in biometrics-based authentication systems. IBM Systems Journal, 40(3), 614–634.

Solove, D. J. (2006). A taxonomy of privacy. *University of Pennsylvania Law Review*, 154, 477–564.

Stanley, J. (2019). The dawn of robot surveillance: AI, video analytics, and privacy. *American Civil Liberties Union*, 38–48.

Turk, M., & Pentland, A. (1991). Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 71–86.

Wayman, J. L. (2008). Biometrics in identity management systems. *IEEE Security & Privacy*, 6(3), 30–37.

Whittaker, M. (2018). AI Now Report 2018. AI Now Institute, New York University.

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *Association for Computing Machinery*, 35(1), 399–458.