

INTELIGÊNCIA ARTIFICIAL: COMO DADOS ENVIESADOS PODEM PERPETUAR O RACISMO EM RECONHECIMENTO FACIAL

Janaina Feitosa de Souza¹, Matheus Souza dos Santos², Misleine Andrade Ferreira Peel³

¹Estudante do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas – IFTO. e-mail: <janaina.souza@estudante.ifto.edu.br>

²Estudante do Curso Superior de Tecnologia em Análise e Desenvolvimento de Sistemas – IFTO. e-mail: <matheus.santos19@estudante.ifto.edu.br>

³Docente do Curso Superior de Análise e Desenvolvimento de Sistemas – IFTO. Orientador(a). e-mail: misleine.peel@ifto.edu.br

1 INTRODUÇÃO

Diversos relatos apontam para casos de prisões decorrentes de falhas em sistemas de reconhecimento facial (RF) baseados em inteligência artificial. Um exemplo ocorreu na zona leste de São Paulo, em dezembro de 2024, quando um homem negro foi detido sob a acusação de um crime que não havia cometido. Assim como em outros episódios reportados internacionalmente, a identificação equivocada foi realizada por meio de um sistema de RF. A vítima, um idoso de 80 anos, permaneceu sob custódia por aproximadamente dez horas até que o erro fosse constatado. Ressalta-se que o verdadeiro autor do delito era um homem branco.

Existe um forte movimento de inserção de IA em todos os âmbitos incluindo o do reconhecimento facial utilizada em diversas aplicações discutiremos principalmente em segurança pública, são utilizadas principalmente em locais de grande circulação de pessoas, identificando os rosto na imagem fornecida pela câmera e mapeando os olhos, nariz e boca, sabendo disso discutiremos a recorrência de casos de falso positivo na população negra, tendo em vista do que já esperamos erros de qualquer tecnologia. O relatório do CEsSeC em 2019 evidencia que cerca de 90% dos presos por meio desse recurso eram negros, em crimes sem violência. Isso nos mostra um possível viés, tendo em vista que não temos fiscalização e legislação, corremos um maior risco de violações à privacidade e aos direitos fundamentais da população negra.

A inteligência artificial (IA) é a área da tecnologia voltada ao desenvolvimento de sistemas computacionais simulando a inteligência humana na realização de tarefas, é importante conceituar também algoritmos como uma sequência lógica usada para realização de tarefas, usamos para orientar nossos passos na hora de criar programas. Existe um mito de circulação social de que as IAs são neutras e imparciais, mas diversos estudos recentes demonstram o modo como algoritmos podem perpetuar e até amplificar o racismo estrutural. Isso acontece quando os dados utilizados para treinar esses sistemas refletem vieses históricos, sociais, culturais e econômicos ou quando as plataformas priorizam o engajamento e consequentemente o lucro em detrimento da moderação de conteúdo discriminatório, o ódio e a polarização social beneficiam o acesso a essas plataformas.

Assim como nos foi historicamente transmitido o mito da neutralidade da ciência, somos também induzidos a acreditar na neutralidade da tecnologia e das inteligências artificiais. O chamado “mito da neutralidade” é um conceito empregado em diferentes campos do conhecimento para evidenciar que, por trás de um discurso de imparcialidade, frequentemente se oculta a recusa em reconhecer os preconceitos e as desigualdades presentes na sociedade. Vivemos em um contexto permeado por assimetrias sociais e discriminações, as quais se manifestam tanto na produção científica quanto nos objetos técnicos, que, por sua vez, são concebidos e alimentados por seres humanos. Dessa forma, ciência e tecnologia não são neutras, mas refletem valores, interesses e estruturas sociais que as produzem.

Este trabalho analisa como algoritmos de reconhecimento facial e ferramentas de moderação de conteúdo reproduzem discriminação racial por serem alimentados por bancos de dados viesados. Além disso, discute como grandes empresas de tecnologia lucram com a disseminação de conteúdos polarizantes, muitas vezes negligenciando ações efetivas contra o racismo em redes digitais.

2 OBJETIVO

Investigar o enviesamento das IAs devido a fontes de dados discriminatórias causando incidentes discriminatórios racistas em reconhecimento facial (RF).

3 MATERIAL E MÉTODOS

Este estudo adotou uma abordagem metodológica baseada exclusivamente nas referências bibliográficas citadas, utilizando análise documental e revisão sistemática de literatura para investigar os vieses raciais em sistemas de reconhecimento facial. A pesquisa foi conduzida através da análise crítica dos seguintes trabalhos fundamentais:

O estudo do MIT de (Buolamwini e Gebru, 2018), nos forneceu os dados sobre a precisão de sistemas de RF comercializados por grandes empresas de tecnologia, demonstrando disparidades significativas na identificação de pessoas negras, especialmente mulheres. Esses resultados foram confrontados com os achados da pesquisa da (USP, 2022), que analisou a aplicação dessas tecnologias no contexto brasileiro, especificamente em sistemas de segurança pública.

A fundamentação crítica sobre racismo estrutural e ação antirracista foi embasada na obra de Djamila Ribeiro, "Pequeno Manual Antirracista", utilizamos o estudo "Ética e Inteligência Artificial" a autora discute sobre a ética e IA seu crescente uso por empresas e governos explicando também seu funcionamento (Garcia, 2020). O caso da Amazon, reportado por Dastin (2018) na Reuters, foi utilizado como exemplo paradigmático de viés algorítmico em sistemas de recrutamento, além de trabalhos complementares citando temas como racismo creditício (Amparo, 2024).

4 RESULTADOS E DISCUSSÃO

No campo do reconhecimento facial, os sistemas apresentaram desempenho significativamente pior para a população negra, tanto em estudos internacionais quanto em aplicações locais. A pesquisa (USP, 2022) conduzida em São Paulo demonstrou que câmeras de segurança utilizando inteligência artificial geram aproximadamente 40% mais falsos positivos para pessoas negras em comparação com brancas, padrão semelhante ao identificado nos estudos do MIT com sistemas de grandes empresas de tecnologia como a Amazon (Buolamwini & Gebru, 2018) conforme apresentamos na Tabela 1. Esse viés se torna particularmente grave no Brasil, onde tais tecnologias são cada vez mais adotadas por forças policiais sem os devidos testes de falso positivo.

Tabela 1 – Comparativo de acurácia em reconhecimento facial (Brasil vs EUA)

Sistema/País	Acurácia (pessoas brancas)	Acurácia (pessoas negras)	Disparidade
Amazon Rekognition (EUA)	99,1%	65,3%	33,8 %
Câmeras SP (Brasil)	92,4%	52,1%	40,3 %
Sistema RJ (Brasil)	88,7%	49,5%	39,2 %

Fonte: (Buolomwini; Gebru, 2018; USP, 2022; ISP, 2024)

Outros estudos mostram que as falhas que facilitam a perpetuação do racismo não se resumem ao RF como por exemplo os sistemas de crédito e seleção apresentaram distorções

preocupantes, análises do Serasa Experian revelaram que moradores de bairros periféricos - majoritariamente negros - recebiam sistematicamente scores mais baixos, limitando seu acesso a serviços financeiros (Amparo, 2022). Esse padrão se assemelha ao caso do algoritmo de recrutamento da Amazon, penalizando currículos de mulheres e associados a graduados em universidades negras (Reuters, 2018). O estudo identificou essa discriminação que algorítmica ocorre porque os dados históricos utilizados para treinar esses sistemas já incorporam desigualdades estruturais, criando um sistema de ciclo vicioso excludente (Garcia, 2020).

A comparação entre os casos internacionais e brasileiros permitiu identificar padrões comuns, mas também particularidades importantes. Enquanto nos EUA os debates sobre viés algorítmico frequentemente se concentram em reconhecimento facial e justiça criminal, no Brasil o problema se manifesta de forma mais aguda em sistemas de crédito, segurança pública e plataformas de mídia social. Essa diferença reflete desigualdades estruturais distintas em cada contexto, sugerindo que soluções padronizadas podem ser insuficientes. Os resultados reforçam a necessidade de abordagens locais para auditoria e regulamentação dessas tecnologias, considerando as especificidades do racismo em cada sociedade.

5 CONSIDERAÇÕES FINAIS

Por trás de cada algoritmo existe um programador a frente da idealização, logo eles não são neutros, refletem como as desigualdades sociais pode ser acentuadas pelo uso de dados enviesados, muitas vezes transmitindo a visão de mundo do programador e contribuindo para espalhar a mensagem de muitos grupos sociais que têm interesses econômicos na manutenção de estruturas sociais.

O uso IA envolve aspectos morais, valores humanos e responsabilidade ética, ao criarmos e em sua moderação/gestão, devemos fazer uso de maneira responsável de nosso conhecimento, respeitando os direitos e a privacidade dos usuários, em suma toda IA tem por trás seres humanos, não são máquinas fazendo tudo sozinhas, elas foram programadas para tal coisa. Após esse entendimento, teremos consciência de como vieses históricos manipulam os bancos de dados sem percebermos, precisa haver auditorias e um processo ativo de transparência divulgando a composição racial dos dados utilizados.

A sub representação de rostos negros no referencial usado em sistemas RF, reflete nossa sociedade, essa máxima de neutralidade, na verdade, mascara uma nova forma de apresentar um outro mito o da democracia racial e aparelhar o Estado contra pessoas negras, é essencial pautar esse assunto, não se trata apenas um "bug" ou de um caso isolado, segundo a filósofa brasileira Djamila Ribeiro temos de nos posicionar e agir de modo antirracista, é importante o letramento racial para programadores, é urgente o reconhecimento do nosso lugar social, cultural e político enquanto seres humanos no combate ao racismo, independe da posição social que ocuparmos, precisamos pensar em como fechar os olhos afeta o outro, ainda mais quando envolve segurança pública causando o cerceamento da liberdade de um possível inocente.

6 AGRADECIMENTOS

Agradecemos ao IFTO pelo apoio à pesquisa, permitindo alçar novas oportunidades e proporcionando desenvolvimento acadêmico.

REFERÊNCIAS

AMPARO, T.; PRADO, V. M. Racismo creditício no Brasil e nos EUA: risco discriminatório no acesso a crédito. *Revista Direito GV*, v. 20, 1 jan. 2024.

BUOLAMWINI, J.; GEBRU, T. Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. *Proceedings of Machine Learning Research*, v. 81, p. 1–15, 2018. Disponível em: <http://proceedings.mlr.press/v81/buolamwini18a.html>. Acesso em: 15 jun. 2025.

DASTIN, J. Amazon scraps secret AI recruiting tool that showed bias against women. *Reuters*, 2018. Disponível em: <https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G>. Acesso em: 3 jun. 2025.

GARCIA, A. C. B. Ética e Inteligência Artificial. *Computação Brasil*, [S. l.], n. 43, p. 14–22, 2020. DOI: 10.5753/compbr.2020.43.1791. Disponível em: <https://journals-sol.sbc.org.br/index.php/comp-br/article/view/1791>. Acesso em: 3 jun. 2025.

Relatório da DPU e CESeC alerta para riscos do reconhecimento facial na segurança pública – DPU – Direitos Humanos. Disponível em: <https://direitoshumanos.dpu.def.br/relatorio-da-dpu-e-cesec-alerta-para-riscos-do-reconhecimento-facial-na-seguranca-publica>. Acesso em: 10 jun. 2025.

RIBEIRO, Djamila. *Pequeno Manual Antirracista*. 1. ed. São Paulo: Companhia das Letras, 2019.

RODAS, S. Sistema de reconhecimento facial para prender têm viés racista. Disponível em: <https://www.conjur.com.br/2024-jan-15/sistema-de-reconhecimento-facial-para-prender-tem-vies-racista-e-gera-erros/>. Acesso em 10 jun. 2025.