



IV Encontro Virtual da ABCiber

Associação Brasileira de Pesquisadores em Cibercultura

Perspectivas Interdisciplinares e Reconfigurações na Cibercultura:

Dados, Algoritmos e Inteligência Artificial

Online — 20 e 21/06/2024

Desafios e riscos dos deepfakes: uma perspectiva crítica sobre as vítimas mulheres¹

Denise Frigo²

Resumo

Os *deepfakes* referem-se a técnicas de manipulação de vídeos e imagens para criar conteúdo falso e realista, geralmente envolvendo a substituição de rostos e vozes em vídeos existentes. Com o avanço da inteligência artificial (IA), essas técnicas tornaram-se mais acessíveis e sofisticadas, levantando preocupações significativas sobre seu potencial de uso malicioso. Embora os *deepfakes* possam ter uma variedade de aplicações legítimas, como no entretenimento e na dublagem de filmes, eles também podem ser usados para difamar, assediar ou manipular pessoas, especialmente mulheres. Como, o relato feito por uma vítima, Kate Isaac à BBC, em 2022 "Alguém pegou meu rosto, colocou em um vídeo pornográfico e fez parecer que era eu". Recentemente, a cantora Taylor Swift também foi vítima de *deepfakes*. Assim, este estudo visa explorar os desafios e riscos associados aos *deepfakes* e à multimídia generativa da inteligência artificial, com foco nas vítimas mulheres. O objetivo principal é analisar criticamente o impacto dessas tecnologias emergentes na segurança, privacidade e dignidade das mulheres, identificando os mecanismos subjacentes de criação, disseminação e mitigação dos *deepfakes*, e propor estratégias para proteger e empoderar as vítimas. Para embasar essa análise, são exploradas teorias de ética da IA, privacidade, justiça algorítmica e responsabilidade algorítmica, fornecendo um arcabouço para avaliar os impactos dos *deepfakes* na sociedade. Aliando-se a estudos de mídia e comunicação, a pesquisa oferece perspectivas críticas sobre a influência dos meios de comunicação na

¹ Trabalho apresentado no GT1 Regimes de verificação em tempos de IA do IV Encontro Virtual da ABCiber – Associação Brasileira de Pesquisadores em Cibercultura. Perspectivas Interdisciplinares e Reconfigurações na Cibercultura: Dados, Algoritmos e Inteligência Artificial. Realização da UNIFAE, nos dias 20 e 21 de junho de 2024.

² Doutora em História, Universidade Federal de Santa Maria e denise.frigo@ufsm.br.



IV Encontro Virtual da ABCiber

Associação Brasileira de Pesquisadores em Ciberultura

Perspectivas Interdisciplinares e Reconfigurações na Ciberultura:

Dados, Algoritmos e Inteligência Artificial

Online — 20 e 21/06/2024

construção da realidade e na disseminação de imagens e narrativas. Essas teorias ajudam a entender como os *deepfakes* podem distorcer a percepção pública e reforçar narrativas prejudiciais sobre as mulheres. O surgimento dos *deepfakes* e da multimídia generativa da IA representa uma ameaça significativa à integridade e dignidade das mulheres. Essas tecnologias permitem a manipulação digital convincente de imagens e vídeos, possibilitando a criação de conteúdo pornográfico não consensual, difamação, assédio e outras formas de violência de gênero. Além disso, os *deepfakes* têm o potencial de minar a confiança na veracidade das evidências visuais, comprometendo a capacidade das mulheres de buscar justiça em casos de abuso ou discriminação. Uma análise crítica desses desafios é crucial para desenvolver políticas, regulamentações e práticas que protejam os direitos e a segurança das mulheres em um mundo digital cada vez mais permeado pela IA. Em conclusão, os *deepfakes* representam uma tecnologia poderosa que, embora ofereça oportunidades legítimas em áreas como o entretenimento, carrega consigo um potencial destrutivo, especialmente para as mulheres. A manipulação realista de imagens e vídeos sem consentimento pode levar a graves violações de privacidade, difamação e violência de gênero. Este estudo destaca a necessidade urgente de abordar esses desafios por meio de uma análise crítica e multidisciplinar, que inclui teorias de ética da IA, justiça algorítmica e estudos de mídia. A compreensão das implicações sociais e éticas dos *deepfakes* é crucial para a formulação de políticas e regulamentações que protejam a dignidade e a segurança das mulheres. Além disso, o desenvolvimento de tecnologias eficazes de detecção e suporte contínuo às vítimas é imperativo para mitigar os danos causados por estas manipulações digitais. Promover a conscientização social e a implementação de medidas de proteção são passos essenciais para garantir a justiça digital e a igualdade de gênero em uma era cada vez mais dominada pela inteligência artificial. Assim, esta pesquisa busca não apenas diagnosticar os problemas, mas também propor soluções práticas e teóricas que contribuam para um ambiente digital mais seguro e equitativo para todas as mulheres.



IV Encontro Virtual da ABCiber

Associação Brasileira de Pesquisadores em Cibercultura

Perspectivas Interdisciplinares e Reconfigurações na Cibercultura:

Dados, Algoritmos e Inteligência Artificial

Online — 20 e 21/06/2024

Palavras-chave

Deepfakes; inteligência artificial; ética; vulnerabilidade feminina; impactos sociais.

Referências

AWADALLAH, H. A. et al. (2020) **Leveraging multi-source weak social supervision for early detection of fake news**. ArXiv preprint, 3 abr. 2020. Disponível em: <https://arxiv.org/abs/2004.01732>. Acesso em: 14 set. 2023.

BRASIL. Projeto de Lei nº 2639, de 2020. **Institui a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet**. Disponível em: <https://www25.senado.leg.br/web/atividade/materias/-/materia/141944>. Acesso em: 14 set. 2023.

BONTRIDDER, N. and Pouillet, Y. (2021) **'The role of artificial intelligence in disinformation'**, Data & Policy, 3, p. e32. doi:10.1017/dap.2021.20.

CHADHA, Anupama et al. Deepfake: an overview. In: **Proceedings of Second International Conference on Computing, Communications, and Cyber-Security: IC4S 2020**. Springer Singapore, 2021. p. 557-566.

DA SILVEIRA, Paulo Antônio Caliendo Velloso. **Ética e Inteligência Artificial: da possibilidade filosófica de Agentes Morais Artificiais**. 2021.

GU, R. et al. (2023). **Unsupervised fake news detection on social media: a generative approach**. In: Proceedings of the AAAI Conference on Artificial Intelligence, v. 33, julho de 2019. Disponível em: <https://www.aaai.org/ojs/index.php/AAAI/article/view/4508>. Acesso em: 14 set. 2023.

FILHO, R. N. I.; MARRAFON, M. A.; MEDÓN, F. (2022) **A inteligência artificial a serviço da desinformação: como as deepfakes e as redes automatizadas abalam a liberdade de ideias no debate público e a democracia constitucional e deliberativa**. In: Economic Analysis of Law Review, v. 13, n. 3, p-32-47, out-dez.



IV Encontro Virtual da ABCiber

Associação Brasileira de Pesquisadores em Ciberultura

Perspectivas Interdisciplinares e Reconfigurações na Ciberultura:
Dados, Algoritmos e Inteligência Artificial
Online — 20 e 21/06/2024

HERN, A; SOLON, O (2018). **Facebook closed 583m fake accounts in first three months of 2018.** The Guardian, 15 mai. 2018. Disponível em: <https://www.theguardian.com/technology/2018/may/15/facebook-closed-583m-fake-accounts-in-first-three-months-of-2018>. Acesso: 14 set. 2023.

KERTYSOVA, K. (2018). **Artificial intelligence and disinformation.** In: **Security and Human Rights**, v. 29, pp. 55-81.

KLEIN, David. **Report: Minorities and Women are More Likely Victims of Cyber Crime.** Disponível em: <https://www.occrp.org/en/daily/15343-report-minorities-and-women-are-more-likely-victims-of-cyber-crime>. Acesso em: 29 maio 2024.

KWOK, Andrei OJ; KOH, Sharon GM. Deepfake: a social construction of technology perspective. **Current Issues in Tourism**, v. 24, n. 13, p. 1798-1802, 2021.

MCDERMOTT, Sarah; DAVIES, Jess. **'Deepfake': 'Colocaram meu rosto em um vídeo pornô'.** BBC News, 21 out. 2022. Disponível em: <https://www.bbc.com/portuguese/geral-63212234>. Acesso em: 23 maio 2024.

RANA, Md Shohel et al. Deepfake detection: A systematic literature review. **IEEE access**, v. 10, p. 25494-25513, 2022.

VILLASENOR, J., 2020. **How to deal with AI-enabled disinformation**, Brookings Institution. United States of America. Disponível em: <https://policycommons.net/artifacts/4144071/how-to-deal-with-ai-enabled-disinformation/4952692/>. Acesso em: 25 maio 2024.

WALLACH, Wendell; ALLEN, Colin. **Moral machines: Teaching robots right from wrong.** Oxford University Press, 2008.

WESTERLUND, Mika. The emergence of deepfake technology: A review. **Technology innovation management review**, v. 9, n. 11, 2019.