

## Estratégias de combate ao *deepfake* nas eleições municipais de 2024<sup>1</sup>

Ana Carolina Thomé Jacome<sup>2</sup>

André Fabrício da Cunha Holanda<sup>3</sup>

Universidade Federal Rural do Rio de Janeiro - UFRRJ

### RESUMO

O artigo aborda o impacto dos *deepfakes* na desinformação política, a partir de revisão e estudo de casos durante as eleições municipais de 2024. Analisamos como essa tecnologia amplifica a disseminação de notícias falsas, afetando a opinião pública e a estabilidade democrática. Exemplos de manipulação nas eleições municipais brasileiras são investigados, evidenciando o papel da Inteligência Artificial na produção e viralização desses conteúdos. A pesquisa conclui que a regulamentação e o trabalho de fact-checking são essenciais para mitigar os efeitos dos *deepfakes*, mas enfatiza a necessidade de educação digital e aprimoramento das técnicas de detecção.

**PALAVRAS-CHAVE:** desinformação; deepfake; algoritmos; inteligência artificial; eleições.

### INTRODUÇÃO

Em fevereiro de 2025, o presidente dos Estados Unidos, Donald Trump, compartilhou um vídeo gerado por inteligência artificial que retrata a Faixa de Gaza transformada em um luxuoso destino turístico chamado “Trump Gaza”. O vídeo, criado pelos cineastas Solo Avital e Ariel Vromen, segundo os mesmos, como uma sátira política à megalomania do republicano, apresenta imagens surrealistas de Gaza como um resort cheio de arranha-céus, praias paradisíacas, estátuas douradas de Trump e personagens como Elon Musk esbanjando dinheiro e Benjamin Netanyahu, na beira da piscina junto a Trump. A mídia internacional reagiu com críticas à publicação, considerando-a de mau gosto e insensível ao contexto de conflito prolongado na região.

A naturalização, sem qualquer pudor, da disseminação deste tipo de conteúdo fictício, gerado por IA, levanta preocupações sobre a integridade da comunicação

---

<sup>1</sup> Trabalho apresentado no Grupo de Trabalho Inteligência Artificial: usos e perspectivas críticas evento integrante da programação do 28º Congresso de Ciências da Comunicação na Região Sudeste, realizado de 15 a 17 de maio de 2025.

<sup>2</sup> Bacharela em Jornalismo pela UFRRJ, e-mail: thomejacome@gmail.com

<sup>3</sup> Professor Adjunto da UFRRJ, Pesquisador do NECOM-UFRRJ; GJOL-UFBA; email: andreholanda73@gmail.com

política nas democracias contemporâneas, como sabemos, ameaçadas em várias frentes pela emergência de novas formas de desinformação associadas a fenômenos recentes como as redes sociais, a mediação algorítmica e a Inteligência Artificial.

Com a presença de um usuário-produtor-ativo ao invés de usuário-consumidor -passivo nas mídias digitais, o público não apenas consome a informação, mas expõe suas opiniões, compartilha vídeos e fotos, comenta nas publicações, produz e distribui conteúdo. Essa dinâmica propicia o cenário de hiperinformação nas redes, marcado pela proliferação de vozes e discursos sem uma hierarquia definida entre os emissores.

Em paralelo ao caráter potencialmente positivo desta transformação do ecossistema midiático, surgem resultados negativos. Principalmente, o usuário apresenta dificuldade em distinguir a fronteira entre a realidade e a ficção. Isso pode ser explicado porque, de acordo com Brisola e Bezerra (2018), os usuários, ao serem atingidos pela quantidade de informações, não têm tempo de checar as origens, credibilidade e veracidade dos conteúdos que recebem.

Nesse contexto, o crescimento das redes sociais não apenas acelerou a disseminação de notícias falsas, mas também originou novas formas de manipulação de informações, como o *deepfake*. A prática usa a inteligência artificial (IA) para alterar o rosto de qualquer personagem em vídeo, inclusive o movimento dos lábios, bem como transplantar um trecho de fala de um determinado vídeo para outro, mudando completamente o conteúdo de uma declaração, por exemplo. A tecnologia também possui capacidades como gerar discurso com voz humana e construir imagens originais a partir de modelos arquivados.

Mais do que uma ferramenta utilizada para entretenimento e humor, inclusive a sátira política, o *deepfake* se enquadra como uma técnica que acirra o quadro de desinformação. A situação piora muito quando agentes políticos lançam mão deste tipo de crime como recurso para interferir no processo eleitoral e na formação da opinião pública.

### **DEEPAKE: O NOVO FORMATO DAS FAKE NEWS**

De acordo com Westerlund (2019), a nova tecnologia afeta a sociedade como um todo, pois pressiona jornalistas a filtrarem notícias reais de falsas; ameaça à segurança nacional ao disseminar propaganda e interferir em eleições; prejudica a

confiança dos cidadãos em informações fornecidas pelas autoridades; e levanta questões de cibersegurança para pessoas e organizações.

No cenário político, vários exemplos ao redor do mundo demonstram o poder da disseminação de um *deepfake* durante e após o processo eleitoral. Um exemplo disso é o caso já citado do “Trump Gaza”. Lessa, Cabral e Silvestre (2020) ao abordarem as consequências do uso dessas ferramentas no campo político explicam que os conteúdos falsos podem ocasionar desestruturação institucional gerando, assim, um sentimento de instabilidade e ceticismo no público que se acostuma a consumir e compartilhar notícias manipuladas como se fossem reais.

A presença de algoritmos nas redes sociais amplifica esse fenômeno, pois a cada página visualizada, compartilhamento realizado ou postagem publicada, os usuários deixam rastros digitais. Esses dados são processados pelos algoritmos, que os utilizam para filtrar quais informações serão processadas e quais serão excluídas, prever passos e comportamentos dos usuários e determinar a relevância de conteúdo.

A mediação profissional exercida pelo jornalismo, antes central na sociedade, vem sendo deslocada por diversos fenômenos da contemporaneidade, a exemplo das redes sociais e dos algoritmos de recomendação, processo cada vez mais intensificado pela adoção da inteligência artificial generativa (Holanda, 2024). A tradicional curadoria editorial, baseada em critérios jornalísticos, cede espaço a sistemas automatizados que personalizam conteúdos conforme padrões de engajamento e preferências individuais que reforçam as bolhas de informação e priorizando materiais sensacionalistas ou polarizadores, maximizam a retenção da audiência em detrimento do interesse público.

Em um cenário de polarização entre dois grupos políticos, se determinado usuário demonstra interesse por informações relativas ao lado “A”, o processo de “filtro bolha” irá limitar o acesso a informações somente a respeito. Assim, eventuais “fake news” que explorem esse padrão para a sua disseminação terão maior êxito, já que ele não terá acesso a outras informações que poderiam contradizer ou até esclarecer os fatos. Conseqüentemente, o processo pode ganhar força ao ser compartilhado, já que a credibilidade deixa de ser da fonte original e passa a ser influenciada pelo responsável pelo compartilhamento da notícia (Sastre; Correio; Correio, 2018, p.8).

Ferramentas generativas em particular podem amplificar a desinformação criando e naturalizando o compartilhamento de conteúdo fictício ou simplesmente falso, ampliando os riscos da influência algorítmica sobre a esfera pública. Esta ameaça

ressalta a necessidade de políticas regulatórias, da mobilização da sociedade e de estratégias editoriais que preservem através da mediação humana os valores do jornalismo profissional, e a mediação responsável da agenda pública.

Com vistas a contribuir para com este debate, a pesquisa aqui apresentada aborda sobre a falsificação deliberada da imagem de concorrentes com a intenção de enganar o eleitor durante os processos eleitorais. A pesquisa (Jacome, 2024) identificou casos de manipulação através de *deepfake* e demonstra a eficácia de soluções técnicas, regulatórias e jurídicas no combate a esta prática durante pleito municipal de 2024.

Durante o período eleitoral, não houve material suficiente para análise no Rio de Janeiro, devido à eficácia das regulamentações e ao combate à impunidade. Com isso, o estudo teve como foco a disputa pela prefeitura de São Paulo. A coleta de dados ocorreu entre 16 de agosto e 6 de outubro, fase em que as campanhas estavam em intensa mobilização nas redes sociais proporcionando, assim, condições ideais para a análise dos fenômenos de desinformação.

Os casos escolhidos são “Vídeo usa ‘*deepfake*’ de Bolsonaro para dar a entender apoio a Marçal” (Mangabeira, 2024); “É #FAKE foto de Tabata Amaral em pose sensual; trata-se de *deepfake*” (Dauer, 2024); e “Vídeos usam IA para criar ‘*deepfakes*’ em que Bonner pede voto em candidatos a vereador” (Mangabeira, 2024).

A análise de imagens e vídeos com sinais de manipulação, como ressalta Verdoliva (2022), pode ser realizada tanto por meio de métodos baseados em características manuais (“*methods based on handcrafted features*”) quanto por métodos baseados em deep learning - aprendizado profundo (“*methods based on deep learning*”), que permitem detectar padrões sutis e frequentemente imperceptíveis ao olho humano.

A primeira abordagem se concentra em identificar falhas específicas no processo de geração, que surgem devido à incapacidade das ferramentas de inteligência artificial em reproduzir todos os detalhes de um rosto humano, por exemplo. Exemplos dessas falhas incluem, mas não se limitam a incompatibilidades na cor dos olhos (como diferenças entre o olho esquerdo e o direito), assimetrias faciais (como um brinco visível de um único lado ou orelhas com características marcadamente diferentes), ou distorções na textura da pele e nos movimentos faciais (Verdoliva, 2022).

Já os métodos baseados em aprendizado profundo oferecem uma abordagem mais detalhada para a detecção de um *deepfake*. O *deep learning* permite que os modelos sejam treinados em grandes volumes de dados, o que favorece identificar padrões complexos e sutis que, muitas vezes, escapam à percepção humana ou aos métodos tradicionais de análise. Esse método tem sido utilizado nos programas especializados em detectar potenciais usos de inteligência artificial em vídeos e áudios. A integração dessas abordagens mencionadas forneceu a base metodológica que utilizamos para compreender o uso de *deepfakes* em ambientes digitais contemporâneos e chegar aos resultados obtidos.

## CONSIDERAÇÕES FINAIS

Uma das principais características da sociedade moderna é a utilização da internet e das redes sociais como ferramentas para a interação social, a disseminação de informações e a produção de conteúdo. Dentro deste espaço, usuários mal-intencionados encontram um ambiente perfeito para disseminar a desinformação ao seu favor, ou de quem o pague. Desse modo, afetando a percepção pública, a reputação das instituições democráticas e, muitas vezes, promovendo discursos de ódio ou polarização social.

Os avanços da inteligência artificial e o surgimento de novas técnicas ajudaram a fomentar o cenário da desinformação nas redes sociais. Entre as diversas formas de conteúdo falso que circulam, os *deepfakes* chamam atenção pela grande dificuldade de distinguir o que é falso e o que é verdadeiro no produto final. Aliada a essa técnica, a manipulação dos algoritmos também proporciona uma maior viralização dos conteúdos falsos. Como ressaltava Joseph Goebbels, ministro da propaganda nazista, “Uma mentira repetida mil vezes torna-se verdade”.

Nossa pesquisa traz resultados encorajadores ao constatar que a regulamentação e medidas jurídicas prometidas e previamente adotadas funcionaram no caso da campanha municipal do Rio de Janeiro. Nos casos analisados, todos da campanha paulista, percebe-se que a mobilização de jornalistas e ativistas no fact-checking igualmente funciona na identificação das manipulações, ainda que para que tal ocorra, o conhecimento técnico tenha se mostrado essencial.

Considerando o avanço da tecnologia, os deepfakes da próxima geração serão capazes de superar as imperfeições atuais e criar falsificações visuais praticamente perfeitas. Como resposta, cabe às autoridades judiciais desenvolver outras medidas regulatórias, além de aprimorar as técnicas de identificação de conteúdos manipulados.

Ao mesmo tempo, a sociedade deve estabelecer estruturas educacionais que incentivem tanto o uso responsável e ético das inovações, quanto o desenvolvimento das competências críticas exigidas para lidar com responsabilidade em vez de esperar pela erradicação dessas tecnologias. Assim como no caso das redes sociais, o verdadeiro desafio não reside no uso da IA, mas na forma como decidimos utilizá-la.

## REFERÊNCIAS

BRISOLA, Anna; BEZERRA, Arthur Coelho. **Desinformação e circulação de “fake news”: distinções, diagnóstico e reação**. In: XIX Encontro Nacional de Pesquisa em Ciência da Informação (XIX ENANCIB). 2018.

DAUER, Letícia. **É #FAKE foto de Tabata Amaral em pose sensual; trata-se de deepfake**. Fato ou Fake (G1). 15 de setembro de 2024. Disponível em: <https://g1.globo.com/fato-ou-fake/sao-paulo/noticia/2024/09/15/e-fake-foto-de-tabata-amaral-e-m-pose-sensual-trata-se-de-deepfake.ghtml>

HOLANDA, André.F. C. **Avaliando efeitos dos agentes de busca de inteligência artificial generativa sobre a polarização política**. In: ANAIS DO 22º ENCONTRO NACIONAL DE PESQUISADORES EM JORNALISMO - VOL. 22, 2024, 2024, Belém. Anais eletrônicos, Galoá, 2024. Disponível em: <https://proceedings.science/encontros-sbpjor/sbpjor-2024/trabalhos/avaliando-efeitos-dos-agentes-de-busca-de-inteligencia-artificial-generativa-sob?lang=pt-br>. Acesso em: 29 Mar. 2025.

JACOME, Ana Carolina Thomé. **Realidade ou ficção? O uso de deepfake durante as eleições municipais da cidade de São Paulo**. 2024. Monografia (Graduação) – Universidade Federal Rural do Rio de Janeiro, Rio de Janeiro, 2024.

LESSA, Moyana Mariano Robles; CABRAL, Hideliza Boechat; SILVESTRE, Gilberto Fachetti. **Deepfake**. Derecho y Cambio social, n. 61, p. 475-487, 2020.

SASTRE, Angelo; CORREIO, CSP de O.; CORREIO, Francisco Rolfsen Belda. **A influência do “filtro bolha” na difusão de Fake News nas mídias sociais: reflexões sobre as mudanças nos algoritmos do Facebook**. Revista GEMINIS, v. 9, n. 1, p. 4-17, 2018.

MANGABEIRA, Milena. **Vídeo usa ‘deepfake’ de Bolsonaro para dar a entender apoio a Marçal**. Aos Fatos. 28 de agosto de 2024. Disponível em: <https://www.aosfatos.org/noticias/deepfake-bolsonaro-apoio-pablo-marcal/>

\_\_\_\_\_. **Vídeos usam IA para criar ‘deepfakes’ em que Bonner pede voto em candidatos a vereador**. Aos Fatos. 4 de outubro de 2024. Disponível em: <https://www.aosfatos.org/noticias/deepfakes-william-bonner-voto-candidatos-vereador/>

VERDOLIVA, Luisa. **Media forensics and deepfakes: an overview**. IEEE journal of selected topics in signal processing, v. 14, n. 5, p. 910-932, 2022.

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