Impacto e Estratégias de Mitigação dos Riscos na Criação de Imagens e Vídeos por Inteligência Artificial

Lucas Eduardo Sasse[1], **Otavio Matheus Neves**[1], **MSc Tathiana Amarante Duarte**[2], **MSc Vitor Hugo Furtado**[2]

Centro Universitário UniSENAI/SC - R. Isidoro Pedri, 263 - Rio Molha, Jaraguá do Sul

[1]{lucas_sasse, otavio_neves}@estudante.sc.senai.br,
2{tathiana.amarante, vitor.furtado}@edu.sc.senai.br

1. Resumo

Este artigo investiga os potenciais riscos associados à criação de imagens e vídeos gerados por inteligência artificial e propõe estratégias de mitigação para minimizar seus impactos negativos. Através de uma fundamentação bibliográfica, o estudo examina as principais vulnerabilidades dessa tecnologia, destacando a falta de controle sobre o conteúdo gerado e sua disseminação potencialmente prejudicial. Ao abordar questões que têm gerado preocupação na sociedade, o artigo sugere medidas resolutivas para mitigar os danos, com o objetivo de reduzir o impacto sobre as vítimas e limitar a propagação de conteúdos problemáticos.

2. Introdução

Ao longo da história, o ser humano desenvolveu diversas ferramentas e tecnologias para aprimorar e facilitar seu modo de vida. Atualmente, a *internet*, redes sociais, filmes e séries desempenham um papel central na sociedade, tornando a criação de conteúdo uma profissão cada vez mais comum e desejada (LEMOS, 2023). Nesse contexto, a inteligência artificial (IA) emergiu como uma tecnologia revolucionária, capaz de criar, melhorar e editar imagens e vídeos com um nível de realismo sem precedentes.

Com o uso de descrições detalhadas, ou "scripts", a IA pode gerar imagens e vídeos que se assemelham à realidade de maneira impressionante. Além disso, a elaboração desses *scripts* pode ser automatizada por outras IA's, aumentando ainda mais a eficiência e a qualidade do conteúdo produzido. No entanto, essa facilidade e potencial de criação trazem à tona preocupações sobre o controle e o uso responsável dessas ferramentas, destacando a necessidade de uma discussão sobre seus impactos e estratégias para mitigar possíveis danos.

3. Problemática e Justificativa

É importante destacar que, apesar das vantagens proporcionadas pelas ferramentas generativas de IA, como a possibilidade de criação de conteúdo inovador de forma mais prática e econômica, elas também apresentam riscos significativos. Uma das principais preocupações é a produção em massa de *fake news*, que alimentam a desinformação e podem manipular a opinião pública. Além disso, tecnologias como *deep fakes* têm sido usadas para criar imagens e vídeos falsos extremamente realistas,

comprometendo a segurança e a reputação das pessoas. Essas falsificações podem levar a situações em que vítimas são retratadas como culpadas em eventos que nunca ocorreram, levantando sérias questões éticas e legais.

Apesar de a inteligência artificial oferecer benefícios significativos, facilitando a vida dos usuários e possibilitando a produção de conteúdo de maneira inovadora, seu uso na prática tem gerado receios em várias áreas. Uma das principais preocupações está relacionada ao risco de substituição de profissionais, especialmente em setores criativos. Artistas, *designers*, atores e dubladores têm manifestado preocupações quanto à possibilidade de serem substituídos por figuras ou vozes geradas por computador, levantando questões sobre o futuro dessas profissões e o valor da criatividade humana.

Sob essa perspectiva, há um crescente temor na sociedade quanto ao uso inadequado dessas tecnologias, especialmente na criação de conteúdo não verídico e na manipulação de vídeos. Essas práticas podem ser usadas para gerar sensacionalismo ou influenciar o pensamento e posicionamento político das pessoas. Campanhas políticas, por exemplo, podem se valer de conteúdo desinformativo para prejudicar a imagem de um candidato oponente, aproveitando-se da dificuldade que o público em geral tem para identificar falsificações sofisticadas. A incapacidade de distinguir conteúdo falso, principalmente por olhos não treinados ou desatentos, agrava a disseminação de informações incorretas e torna a correção dessas informações um desafio, perpetuando a desinformação e impactando negativamente processos democráticos.

Nesse contexto, os principais problemas resultantes giram em torno do potencial impacto do uso dessas ferramentas quando acessíveis ao público em geral. Assim, este estudo tem como objetivo não apenas avaliar os danos que podem ser causados pelo uso indiscriminado dessa tecnologia, mas também examinar como esses riscos são atualmente abordados e protegidos pela legislação.

4. Desenvolvimento

Em 2023, o *Sumsub Identity Fraud Report*, revelou que cerca de 224 países registraram mais de dois milhões de tentativas de fraude em diversos segmentos. No Brasil, as fraudes mais comuns são as sofisticadas, baseadas em Inteligência Artificial. Os *Deep Fakes* que envolvem a criação de conteúdos como notícias, documentos, áudios, vídeos e imagens por IA's estão no topo da lista de fraudes do país.

A Inteligência Artificial teve seu início por volta da Segunda Guerra Mundial e, nos dias atuais, abrange uma variedade de subcampos, como aprendizagem e percepção até tarefas específicas, como a prova de teoremas matemáticos e diagnósticos de doenças. A IA também é capaz de sistematizar e automatizar tarefas intelectuais. Ao simular a intelectualidade humana, a IA aplica dois princípios, o aprendizado de máquina (*Machine Learning*) e o aprendizado profundo (*Deep Learning*).

No processo de criação de imagens e vídeos, as ferramentas de IA podem gerar imagens realistas a partir de descrições de textos, além de editar fotos automaticamente, melhorar a qualidade de vídeos e até criar animações complexas baseadas em esboços simples. Embora esses avanços representem um progresso notável, eles também trazem consigo desafios significativos, sendo um dos principais o uso indevido dessas tecnologias, especialmente na criação de *deep fakes*. O surgimento dos *deep fakes* tem

se tornado uma ameaça à sociedade, aos sistemas políticos e às empresas (WESTERLUND, 2019). A exposição de um *Deep Fake* pode ameaçar a segurança nacional ao disseminar, por exemplo, uma propaganda que possa interferir numa eleição, comprometendo a integridade da informação.

A *Sumsub Identity Fraud Report*, também destacou o aumento global de 10 vezes o número de *Deep Fakes* detectados em todos os setores no período de um ano. Comparado com o aumento médio na América Latina (410% – menor região entre todas), o Brasil apresentou o maior crescimento de *Deep Fake* na região, com o aumento de 830% entre 2022 e 2023. A Espanha também se destacou, sendo o país mais atacado pelos *Deep Fakes* (a cada dez ataques que acontecem no mundo, um ocorre no país europeu). A mídia foi o segmento mais explorado pelos fraudadores, enquanto o passaporte dos Emirados Árabes Unidos foi o documento mais falsificado no mundo.

Em estudo divulgado no site do INDP (Instituto Nacional de Proteção de Dados), eventos atuais revelam o uso controverso da Inteligência Artificial na criação de conteúdos, com implicações significativas aos direitos individuais fundamentais destacados pela Constituição Federal (LEAL, 2024). No ano de 2023, um colégio particular do Rio de Janeiro foi cenário de um caso triste, onde um grupo de alunos utilizou um *software* de inteligência artificial para gerar imagens forjadas de colegas. Mais recentemente em Porto Alegre, alunos de uma escola particular criaram e divulgaram vídeos falsos, simulando nudez de dezesseis estudantes, utilizando IA.

No Brasil, o TSE (Tribunal Superior Eleitoral) tomou medidas para regular o uso dos *Deep Fakes*, proibindo seu uso e exigindo identificação de conteúdos multimídias sintéticas e responsabilizando as plataformas na identificação e remoção de temas que apresentam perigo. Apesar de ainda não haver uma regulamentação específica para o uso de IA's, o sistema jurídico oferece mecanismos para lidar com crimes cibernéticos que violem os direitos pessoais. O Código Civil protege os direitos de personalidade, como a imagem, assegurado pelo art. 20 o direito de exigir que cesse a divulgação de materiais que a ela se refiram, além da obrigação de reparação por parte de quem causou o dano, segundo art. 927 também do Código Civil.

No Código Penal, o art. 171 define o estelionato como crime e prevê que os *Deep Fakes* possam ser utilizados para obter vantagem ilícita ou causar prejuízo a terceiros. A Lei Geral de Proteção de Dados (LGPD), traz diretrizes sobre a manipulação de dados pessoais. Embora não mencione diretamente, a LGPD é aplicável quando tais técnicas envolvem a utilização de modo malicioso das informações.

5. Resultado e Conclusão

O cenário digital atual é caracterizado por um turbilhão de inovações tecnológicas, especialmente no campo da Inteligência Artificial. Embora essa rápida evolução abra inúmeras oportunidades para o desenvolvimento de ferramentas, ela também levanta preocupações sobre os impactos sociais e éticos que tais tecnologias podem gerar.

Este estudo mapeou as principais ferramentas de IA que estão transformando a

criação e manipulação de imagens e vídeos, com um enfoque especial nos *deep fakes*. Confirmou-se que, apesar do grande potencial da IA, seu uso indevido pode levar à geração de conteúdo falso, manipulação da opinião pública, violação de direitos individuais e até ameaças à segurança nacional.

Diante desses riscos, é crucial que sejam estabelecidas medidas regulatórias e educativas para promover um uso ético da IA. Investir em educação digital é fundamental para capacitar os cidadãos a identificar e resistir à manipulação e à desinformação. Além disso, fomentar uma cultura de consciência ética sobre a aplicação dessas tecnologias ajudará a fortalecer a integridade e a confiança social.

Embora este estudo não tenha a pretensão de oferecer soluções definitivas, ele busca contribuir para um debate mais aprofundado e consciente sobre os desafios e implicações da IA. Há, certamente, a necessidade de pesquisas futuras para acompanhar a evolução dessa área, buscando formas de equilibrar inovação tecnológica e responsabilidade social.

Referências

BRASIL. **Lei nº 10.406, de 10 de janeiro de 2002**. Institui o Código Civil. Diário Oficial da União: seção 1, Brasília, DF, ano 139, n. 8, p. 1-74, 11 jan. 2002.

BRASIL. **Lei nº 13.709, de 14 de agosto de 2018**. Lei Geral de Proteção de Dados. Diário Oficial da União: seção 1, Brasília, DF, p. 59, col. 2, 15 ago. 2018.

BRASIL. **Decreto-Lei nº 2848, de 7 de dezembro de 1940**. Código Penal. Diário Oficial da União, Rio de Janeiro, 31 dez. 1940.

IBGE. **O impacto transformador da inteligência artificial na geração de conteúdo e imagem: uma jornada evolutiva**. 09 jan. 2024. Disponível em: <https://www.ibge.gov.br/ibge-digital/38980-o-impacto-transformador-da-inteligencia-artificial-na-geracao-de-conteudo-e-imagem-uma-jornada-evolutiva.html>. Acesso em: 12 jun. 2024.

LEAL, Marta. **O uso controverso da IA na criação de conteúdos e as implicações para os direitos individuais**. INDP. Disponível em: <https://www.inpd.com.br/post/o-uso-controverso-da-ia-na-cria%C3%A7%C3%A3o-de-conte%C3%BAdos-e-as-implica%C3%A7%C3%B5es-para-os-direitos-individuais>. Acesso em: 26 jun. 2024.

LEMOS, Ronaldo. **'Influenciador' virou profissão das mais desejadas**. Disponível em: <https://www1.folha.uol.com.br/colunas/ronaldolemos/2023/05/influenciador-virou-profissao-das-mais-desejadas.shtml>. Acesso em: 20 jun. 2024.

REPORT, Sumsub Identity Fraud. **A comprehensive, data-driven report on identity fraud dynamics and innovative prevention methods**. Disponível em: <https://www.cybersource.com/en-us/solutions/fraud-and-risk-management/fraud-report.html>. Acesso em: 19 jun. 2024.

WESTERLUND, Mika. **The emergence of deepfake technology: a review**. Technology Innovation Management Review, v. 9, n. 11, 40-53 2019. Disponível em: <timreview.ca/article/1282>. Acesso em: 27 jun. 2024.