



CENTRO UNIVERSITÁRIO DE GOIÁS – UNIGOIÁS
COMUNICAÇÃO SOCIAL – PUBLICIDADE E PROPAGANDA

ISABEL CECÍLIA GOMES CAETANO
JÚLIA CIRINO CAMARGO

**MÁSCARAS DIGITAIS: A REVOLUÇÃO DAS DEEPPFAKES NA
COMUNICAÇÃO MODERNA**

GOIÂNIA
2024

ISABEL CECÍLIA GOMES CAETANO
JÚLIA CIRINO CAMARGO

**MÁSCARAS DIGITAIS: A REVOLUÇÃO DAS DEEPFAKES NA
COMUNICAÇÃO MODERNA**

Estudo de caso apresentado ao Curso de Publicidade e Propaganda do Centro Universitário de Goiás – UniGoiás, como requisito para conclusão do Curso de Bacharelado em Publicidade e Propaganda.

Orientadora: Giovanna Adriana Tavares Gomes

GOIÂNIA
2024

RESUMO

Este estudo de caso examina os impactos dos deepfakes, uma tecnologia baseada em inteligência artificial e aprendizado de máquina que permite criar imagens, vídeos, áudios e textos manipulados de forma realista para simular eventos e discursos fictícios. Embora inicialmente utilizados como uma ferramenta de entretenimento, os deepfakes rapidamente se tornaram uma ameaça séria em contextos políticos, econômicos e sociais, comprometendo a confiança nas informações veiculadas. Neste trabalho, investigamos como essa tecnologia tem sido utilizada para enganar e manipular, expondo riscos significativos à sociedade. Por meio da análise de incidentes recentes nos contextos religioso, artístico e empresarial, e do aumento de campanhas de desinformação, o estudo demonstra a vulnerabilidade crescente das instituições e do público diante das novas técnicas de manipulação. Também são explorados os esforços contemporâneos na detecção de deepfakes e a evolução dos métodos para combater essas manipulações.

Palavras-chave: Inteligência Artificial; Desinformação; Manipulação; Comunicação Contemporânea; Informações Falsas.

1 INTRODUÇÃO

O estudo de caso é uma metodologia que permite ao pesquisador investigar um fenômeno em profundidade, considerando sua conjuntura. Este método é amplamente utilizado em diversas áreas do conhecimento e pode ser descritivo ou exploratório. Na modalidade descritiva, o objetivo é detalhar um fenômeno, enquanto na aplicação exploratória, o intuito é investigar um objeto de estudo para gerar hipóteses e teorias.

O estudo em questão se enquadra na abordagem exploratória e examina os efeitos provocados pelas tecnologias de deepfakes na comunicação. Apesar do avanço significativo nessa tecnologia, ainda é pouco compreendido como esses conteúdos manipulados impactam a percepção e a confiança nas informações consumidas, especialmente em contextos religiosos (objeto 1), artísticos (objeto 2) e empresariais (objeto 3). Considerando que os deepfakes utilizam algoritmos avançados para criar imagens, vídeos e áudios que se confundem com a realidade, é necessário investigar suas repercussões nos diferentes segmentos abordados. Diante disso, a questão central que norteia o estudo é: "Como as deepfakes afetam a comunicação e de que forma essas manipulações influenciam a confiança das pessoas nas informações que consomem?"

Sobre o objeto 1, imagens do Papa Francisco em 2023, vestindo uma jaqueta estilosa e contemporânea, viralizaram e foram criadas com a ferramenta de inteligência artificial Midjourney, levando muitos a acreditarem que se tratava de uma nova moda do pontífice. A rapidez com que essas imagens se espalharam nas redes sociais levantou questões sobre a responsabilidade dos meios de comunicação na verificação de informações, especialmente quando até mesmo a Vogue, uma revista mundialmente renomada, caiu na armadilha da desinformação.

Em relação ao objeto 2, a pesquisa analisa a situação da atriz Mel Maia, que recentemente enfrentou a disseminação de um vídeo íntimo falso atribuído a ela e associado ao traficante William Sousa Guedes, conhecido como Corolla. Mel esclareceu aos seguidores que o conteúdo era, na verdade, um deepfake, uma manipulação gerada por inteligência artificial que cria a ilusão de realismo. No objeto 3, será abordado um incidente alarmante envolvendo um funcionário de uma multinacional que, acreditando estar em uma videoconferência com seu diretor financeiro, foi vítima de um esquema de fraude utilizando deepfake. Os golpistas se

fizeram passar por membros da equipe, induzindo o empregado a transferir US\$ 25 milhões.

Dentro do assunto apresentado, faz-se uma delimitação do tema: esses exemplos ilustram como a tecnologia de deepfake não apenas engana, mas também instiga debates sobre a veracidade das informações e o impacto na confiança pública. Este estudo se propõe a examinar as diferentes maneiras pelas quais as deepfakes moldam o entendimento e a disseminação de informações.

A justificativa para a escolha do tema na área da comunicação se baseia no fato de que a deepfake é um assunto extremamente atual, utilizando técnicas avançadas de inteligência artificial, especialmente redes neurais, para criar conteúdos audiovisuais manipulados que parecem autênticos. Esses elementos tornam a deepfake uma ferramenta poderosa, mas também potencialmente prejudicial, necessitando de uma reflexão crítica sobre seu uso e suas consequências.

Para estudar melhor o tema, foram selecionadas as seguintes obras e autores: “Estudo, desenvolvimento e comparação de técnicas de detecção de Deepfake” por Felipe Manfio Barbosa (2020), “Deepfake: explorando técnicas de detecção de manipulação digital de imagens de faces” por Jorge de Jesus Gomes Leandro (2022), e o relatório “Increasing Threat of DeepFake Identities” do Homeland Security, que discutem a evolução das técnicas de deepfake e os desafios na sua detecção.

O objetivo geral desta pesquisa é analisar os efeitos provocados pelas deepfakes no mundo da comunicação, tendo como referências os casos mencionados. Os objetivos específicos são: 1) conceituar deepfakes, sua origem e o processo de criação dessas manipulações audiovisuais, destacando as ferramentas, aplicativos e plataformas utilizadas; 2) analisar as repercussões das deepfakes nos meios artístico, religioso e empresarial, focando nos casos da atriz Mel Maia, do “Papa Fashion” e da fraude na multinacional em Hong Kong; 3) investigar os impactos dessa tecnologia na confiança das pessoas e no modo como consumimos informações online, além de discutir a existência de recursos para combater essa desinformação.

A metodologia utilizada na pesquisa será por meio de revisão bibliográfica e incluirá análise documental, que se baseia em informações extraídas de redes sociais, sites de notícias e matérias jornalísticas. Isso permitirá um embasamento técnico-científico para o desenvolvimento deste estudo de caso, tornando seus resultados ainda mais confiáveis, ao empregar técnicas já conceituadas e bem aceitas

pela comunidade acadêmica. No contexto final, apresentar-se-á o campo de pesquisa e serão demonstrados os resultados, os principais problemas encontrados e a percepção crítica do pesquisador. Por fim, nas considerações finais, serão tecidas as análises.

2 DESENVOLVIMENTO TEÓRICO

2.1 Avanços em Inteligência Artificial e Aprendizado de Máquina

Nas últimas décadas, a Inteligência Artificial (IA) e o Aprendizado de Máquina têm revolucionado inúmeros setores, como a comunicação, a segurança cibernética e o entretenimento. Essas tecnologias evoluíram ao ponto de permitir a criação de sistemas capazes de aprender com grandes volumes de dados, criando representações e previsões que simulam comportamentos humanos com precisão. Como destaca Barbosa (2020), o uso de IA para gerar soluções cotidianas tem se tornado cada vez mais comum, facilitando o desenvolvimento de ferramentas que automatizam processos complexos.

A IA, especificamente as Redes Neurais Artificiais (RNA), permitiu que vídeos, áudios e imagens fossem manipulados com uma sofisticação nunca antes vista. Esse avanço, no entanto, não veio sem desafios. O potencial para o uso malicioso dessas tecnologias logo se tornou evidente, com deepfakes surgindo como um dos exemplos mais preocupantes. Robles-Lessa et al. (2020) ressaltam que a manipulação de conteúdo digital, embora inicialmente inofensiva, pode ser utilizada por grupos ou indivíduos com a intenção de enganar, difamar ou fraudar.

2.2 Deepfakes e Suas Implicações na Comunicação

O termo deepfake surgiu em 2017, quando um usuário com o nome “deepfake” compartilhou vários vídeos falsos e eróticos de celebridades na rede social Reddit. Muitas personalidades famosas da mídia começaram a ver sua imagem ligada a vídeos de sexo, mas obviamente sabiam que as imagens eram falsas.

A palavra “deepfake” deriva da junção de “deep learning” com “fake”, mesmo sendo inicialmente utilizada em contextos de entretenimento, como na criação de memes ou vídeos satíricos, a capacidade de criar vídeos falsos com pessoas

proferindo palavras ou realizando ações que nunca ocorreram introduz uma crise de confiança na era digital.

As deepfakes destacam-se como um dos produtos mais complexos e perigosos do avanço da Inteligência Artificial. Utilizando algoritmos de aprendizado profundo, essas manipulações estudam as expressões faciais dos indivíduos, sendo capazes de replicá-las em vídeos e imagens falsas de maneira quase imperceptível. Segundo Wang *et al.* (2020), essas tecnologias têm o potencial de enganar até mesmo os observadores mais atentos, já que replicam movimentos, expressões faciais e vozes com impressionante realismo.

O impacto das deepfakes não se restringe apenas ao entretenimento. Campanhas políticas e eleições têm sido alvo de vídeos falsos, deslegitimando candidatos ou disseminando informações enganosas, muitas vezes com consequências irreversíveis para o debate público. Wang *et al.* (2020) apontam que, com a crescente sofisticação dessas tecnologias, a dificuldade em detectar vídeos manipulados compromete a credibilidade de figuras públicas, jornalistas e até mesmo governos.

2.3 Casos Reais de Uso das Deepfakes

Um dos exemplos mais alarmantes de uso malicioso da tecnologia deepfake ocorreu no setor empresarial, quando um funcionário de uma multinacional foi induzido a transferir US\$ 25 milhões a fraudadores. Usando um programa de inteligência artificial, os criminosos manipularam e se passaram por executivos da empresa durante uma videoconferência, criando uma ilusão quase perfeita (CNN Business, 2024).

Além disso, figuras públicas têm sido constantemente afetadas pelo uso inconsequente de deepfakes. Em 2023, uma série de imagens manipuladas do Papa Francisco vestindo uma jaqueta moderna e diferente do seu vestuário habitual, se espalhou rapidamente pelas redes sociais, gerando um debate global sobre a veracidade das informações disseminadas online. A gravidade da situação foi ampliada pelo fato de que até veículos de renome, como a revista Vogue, publicaram as imagens sem verificar sua autenticidade. Isso exemplifica como as deepfakes podem enganar até mesmo os meios de comunicação mais experientes, comprometendo a integridade do jornalismo (Lyu *et al.*, 2020).

Outro exemplo relevante é o caso da atriz Mel Maia, que enfrentou uma onda de desinformação quando um vídeo íntimo falsamente atribuído a ela foi amplamente compartilhado, tendo que vir a público esclarecer que não era ela nas cenas. Esse incidente ilustra os danos emocionais e reputacionais que podem resultar do uso de IA para criar conteúdos prejudiciais, mostrando como a tecnologia pode afetar diretamente a vida pessoal de indivíduos.

2.4 Métodos de Detecção e Iniciativas de Combate às Deepfakes

Dada a crescente ameaça representada pelas deepfakes, esforços substanciais têm sido feitos para desenvolver métodos de detecção eficazes. Inicialmente, a identificação de vídeos manipulados dependia de análises manuais, que se concentravam em irregularidades visíveis, como mudanças súbitas na iluminação ou imperfeições nos movimentos faciais. No entanto, com o aumento da complexidade das manipulações, novas abordagens, baseadas em redes neurais convolucionais, têm sido empregadas para detectar sinais sutis de manipulação (Barbosa, 2020).

Uma das iniciativas mais significativas nesse campo é o DeepFake Detection Challenge (DFDC), uma competição promovida por gigantes da tecnologia como Facebook, Microsoft e AWS. Essa colaboração entre o setor público e privado tem como objetivo estimular a criação de soluções inovadoras para identificar deepfakes, permitindo que tanto a academia quanto a indústria se unam no combate a essa ameaça. Segundo Robles-Lessa *et al.* (2020), a cooperação entre esses setores é fundamental para garantir que novas tecnologias de detecção sejam amplamente implementadas.

O uso de inteligência artificial para detectar deepfakes ainda está em desenvolvimento, e há muito espaço para melhorias. No entanto, essas iniciativas já representam um avanço significativo na luta contra a desinformação digital. Dolhansky *et al.* (2020) afirmam que, com o emprego de redes neurais e análises automatizadas, a detecção de manipulações tem se tornado mais precisa, oferecendo uma linha de defesa contra a proliferação dessas falsificações.

2.5 Desafios Éticos e o Futuro das Deepfakes na Comunicação

A ascensão das deepfakes também levanta questões éticas e sociais complexas. A facilidade com que esses vídeos podem ser criados e disseminados ameaça a privacidade e a reputação de indivíduos, enquanto, em uma escala maior, desafia a integridade das democracias. Robles-Lessa *et al.* (2020) destacam que, sem regulamentações claras, o uso irresponsável dessa tecnologia pode gerar um ambiente de desconfiança generalizada.

Além disso, há uma preocupação crescente com a forma como as deepfakes podem ser usadas para manipular eleições, desacreditar adversários políticos e influenciar a opinião pública. Gershgorn (2020) defende que é necessário um esforço conjunto entre governos, empresas de tecnologia e sociedade civil para estabelecer diretrizes que regulem o uso ético dessa tecnologia.

Para o futuro, é essencial que as pesquisas continuem avançando tanto na criação de soluções tecnológicas quanto na conscientização pública sobre os perigos das deepfakes. A educação midiática desempenhará um papel crucial, capacitando o público a questionar a veracidade dos conteúdos que consomem. Ao mesmo tempo, a regulamentação e o desenvolvimento de sistemas de detecção mais robustos continuarão sendo pilares na luta contra os malefícios dessa tecnologia.

3 METODO

A pesquisa qualitativa se distingue por seu foco no desenvolvimento de conceitos a partir de ideias, fatos e opiniões, bem como na compreensão indutiva e interpretativa dos dados relacionados ao problema de pesquisa. Ela busca interpretar características sociais por meio dos significados que as pessoas atribuem às suas experiências e ao mundo ao seu redor. Segundo Pope e Mays (2005), uma pesquisa qualitativa envolve uma tentativa de compreender essas interações e comportamentos em termos de como são percebidos pelos envolvidos:

"A pesquisa qualitativa (...) está relacionada aos significados que as pessoas atribuem às suas experiências do mundo social e como as pessoas compreendem esse mundo. Tenta, portanto, interpretar as características sociais (interações, comportamentos, etc.) em termos de sentidos que as pessoas lhes dão em função disso, é comumente referida como pesquisa interpretativa (POPE; MAYS, 2005, p.13).

Essa abordagem confere à pesquisa qualitativa uma flexibilidade que vai além de considerações previsíveis ou mensuráveis, permitindo uma análise mais profunda de características específicas. Em muitos casos, dados quantitativos também podem ser examinados por essa perspectiva qualitativa, unindo as abordagens para uma melhor compreensão do problema de pesquisa.

De acordo com Lüdke e André (2014), uma pesquisa qualitativa também se preocupa com o contexto em que os dados são coletados e com a experiência dos participantes. Para Bogdan e Biklen (1982), essa abordagem enfatiza a entrega de dados descritivos por meio de um contato direto com a situação científica, priorizando o processo ao invés do produto final e buscando retratar a perspectiva dos envolvidos:

A pesquisa qualitativa ou naturalística, segundo Bogdan e Biklen (1982), envolve a obtenção de dados descritivos, obtidos no contato direto do pesquisador com a situação científica, enfatiza mais o processo do que o produto e se preocupa em retratar a perspectiva dos participantes (LÜDKE; ANDRÉ, 2014, p.14).

Dentro desse contexto, a pesquisa bibliográfica é uma modalidade de pesquisa qualitativa, pois envolve a análise e interpretação de conceitos, teorias e ideias extraídas de materiais já publicados, como livros, artigos científicos e outras fontes documentais. Tal metodologia busca sistematizar o conhecimento já existente sobre o tema investigado, permitindo que o pesquisador construa novas interpretações e compreensões a partir de uma base de dados previamente acumulada.

Assim, uma pesquisa bibliográfica vai além de uma simples coleta de informações; ela é um processo interpretativo, no qual o pesquisador analisa criticamente o conteúdo existente para desenvolver insights e argumentações pertinentes ao seu objeto de estudo.

A presente pesquisa baseia-se no método de pesquisa bibliográfica, caracterizado por um conjunto planejado de procedimentos voltados à busca de soluções que se relacionam diretamente ao objeto de estudo. Conforme Lima e Mito (2007), uma revisão de literatura, muitas vezes confundida com uma pesquisa bibliográfica, representa apenas uma etapa preliminar de qualquer investigação. Embora a revisão de literatura seja uma etapa necessária, a pesquisa bibliográfica implica em uma sistematização organizada e direcionada, que visa responder às questões de pesquisa sem se perder em uma busca solicitada.

Minayo (1994, p. 23) reforça que a pesquisa é um processo contínuo e inacabado, no qual o pesquisador adota uma postura de constante busca teórica, aproximando-se da realidade de maneira crítica e histórica. Essa visão exige do pesquisador a conscientização de que a realidade social é composta por múltiplas camadas e que sua compreensão depende de sucessivas aproximações.

A elaboração do projeto de pesquisa iniciou-se com a escolha do tema, a formulação do problema de pesquisa e a criação de um plano detalhado para responder às questões formuladas. O problema definido buscou entender os impactos das deepfakes na percepção e confiança do público em relação às informações veiculadas. Com base nisso, foi elaborado um planejamento que incluiu a definição de objetivos gerais e específicos, orientando o estudo desde a compreensão do conceito e funcionamento das deepfakes até a análise de suas implicações em contextos específicos como os meios artístico, religioso e empresarial. Além disso, o plano de pesquisa abrangeu a investigação dos efeitos dessas manipulações na confiança dos indivíduos ao consumir conteúdos digitais e a discussão sobre possíveis ferramentas para enfrentar o desafio da desinformação.

Na fase de investigação das soluções, foi realizada uma coleta de dados baseada na análise de material bibliográfico relevante. Dois momentos distintos foram considerados: o levantamento de fontes e a análise do conteúdo dessas fontes. Conforme apontado por Lima e Miotto (2007), a qualidade dos resultados da pesquisa está diretamente associada à relevância e à profundidade dos dados encontrados. Foram consultadas obras como “Estudo, desenvolvimento e comparação de técnicas de detecção de Deepfake” de Felipe Manfio Barbosa (2020) e “Deepfake: explorando técnicas de detecção de manipulação digital de imagens de faces” de Jorge de Jesus Gomes Leandro (2022), além do relatório “Increasing Threat of DeepFake Identities” do Homeland Security, para melhor compreensão das deepfakes. Para a análise dos objetos, utilizaram-se redes sociais, sites de notícias e matérias jornalísticas de veículos como O GLOBO e CNN.

A análise explicativa das soluções coletadas envolveu a criação de um modelo explicativo onde a documentação selecionada foi examinada criticamente para obter instruções específicas. A fase analítica não se restringiu à exploração dos materiais, mas focou em uma reflexão aprofundada sobre as informações. Baseou-se na relevância dos estudos sobre os impactos das deepfakes na comunicação e na qualidade metodológica de cada fonte, priorizando obras que exploram técnicas de

detecção. Identificaram-se pontos de convergência e divergência entre os autores, considerando os impactos na confiança das pessoas e os efeitos da desinformação, além de incluir exemplos práticos como o caso do “Papa Fashion” e fraudes empresariais para contextualizar os efeitos das deepfakes.

A síntese integradora compreendeu a conclusão final dos achados, resultante da análise crítica e reflexiva dos documentos estudados. Essa síntese conectou o problema da pesquisa com as soluções encontradas durante a investigação, propondo respostas fundamentadas. A pesquisa esclareceu como as deepfakes afetam a confiança nas informações, especialmente nos contextos artístico, religioso e empresarial. O processo reflexivo destacou os desafios na identificação dessas manipulações e na conscientização pública sobre seus riscos, mesmo com as ferramentas de detecção existentes. As conclusões indicaram a necessidade de combinar avanços tecnológicos com educação digital para mitigar os efeitos negativos das deepfakes.

A coleta de dados iniciou-se com a definição de critérios que delimitam o universo de estudo, incluindo parâmetros temáticos, linguísticos e cronológicos. Foram selecionadas obras diretamente relacionadas ao objeto de estudo, abrangendo temas como deepfakes e seu impacto na comunicação, além de subtemas que exploram a influência dessa tecnologia em diferentes contextos e as estratégias para combater a desinformação gerada por essas manipulações. As fontes utilizadas incluíram livros, periódicos, teses, dissertações e materiais jornalísticos publicados entre 2020 e 2024, como as obras acadêmicas mencionadas e matérias do G1, CNN e O Globo que ilustram a aplicação prática das deepfakes.

Por fim, a técnica de leitura adotada foi realizada de forma sistemática, conforme sugerido por Salvador (1986), envolvendo etapas de leitura de reconhecimento, leitura seletiva e leitura crítica. Essa metodologia permitiu um aprofundamento teórico sobre o impacto das deepfakes na comunicação, a partir de uma revisão crítica e reflexiva das obras mais relevantes. O processo sistemático de coleta, análise e síntese das informações possibilitou investigar como essas manipulações influenciam a confiança das pessoas nas informações consumidas de maneira estruturada e embasada. A definição de critérios direcionou a seleção de fontes que refletissem a complexidade e relevância dos efeitos das deepfakes, proporcionando uma base teórica sólida para a análise e interpretação dos dados, e

integrando reflexões teóricas com soluções encontradas para compreender o impacto das deepfakes na credibilidade das informações e na confiança do público.

4 ANÁLISES DE DADOS

A análise dos dados está dividida em várias camadas que investigam os efeitos dessas tecnologias em diferentes áreas da sociedade, incluindo o meio religioso, artístico e empresarial. A pesquisa foca principalmente em como as deepfakes, que são vídeos, imagens ou áudios falsos manipulados por inteligência artificial (IA), impactam a confiança pública na veracidade das informações e, consequentemente, na comunicação.

4.1 Impacto das Deepfakes na Religião

O primeiro objeto aborda o caso das imagens falsas do Papa Francisco, que circularam nas redes sociais em 2023. Essas imagens mostravam o pontífice vestindo uma jaqueta estilosa e contemporânea, uma peça de vestuário que não fazia parte de sua tradicional indumentária religiosa. O estudo revela que essas imagens foram criadas utilizando a ferramenta de IA Midjourney, especializada na criação de imagens a partir de uma descrição textual.

O grande problema evidenciado aqui é a rápida disseminação dessas imagens, que enganaram milhões de pessoas e até mesmo veículos de comunicação estabelecidos, como a revista Vogue. A Vogue Brasil chegou a escrever na sua matéria o nome do estilista que teria vestido o Papa, mas as imagens nem eram reais então como poderia haver um estilista. Depois, a revista se retratou nas redes sociais, mas já era tarde demais, a falta de pesquisa da veracidade das imagens fez com que recebesse muitas críticas do público. Essa situação coloca em foco o papel da mídia e das plataformas de redes sociais na verificação de fatos, um veículo que deveria ser uma fonte confiável também foi enganado, questionando como essas manipulações podem afetar a percepção pública de o que é verdade e o que é mentira.

A análise do estudo demonstra que as deepfakes podem minar gravemente a confiança do público e a reputação de um líder religioso ao criar imagens que parecem realistas, mas que não têm base na realidade. Apesar do criador das imagens não ter a intenção de manchar a imagem do Papa, ele acabou gerando

muitas discussões pelo mundo, principalmente entre os que acreditaram, uns disseram que gostaram do novo estilo, enquanto outros criticaram a postura do líder religioso por pensarem que ele não deveria se vestir dessa forma, considerando um desrespeito a Igreja. O caso do “Papa Fashion” é um exemplo clássico de como essas tecnologias podem ser usadas para distorcer a imagem de líderes religiosos e enganar o público em larga escala.

Figura 1: Imagens do “Papa de Fashion” geradas pelo MidJourney



Fonte: Forbes (2024).

Figura 2: Retratação da Vogue Brasil pela plataforma X (antigo Twitter).



Fonte: Poder 360 (2024).

4.2 Impacto das Deepfakes na Arte e Cultura

O segundo objeto examina o caso da atriz Mel Maia, que foi vítima de um vídeo falso gerado por IA que contém cenas de sexo da atriz com um traficante. Este vídeo foi amplamente compartilhado nas redes sociais, gerando um impacto devastador na vida pessoal da atriz, além de trazer à tona questões de privacidade, ética e reputação. Mel Maia chegou a se pronunciar nas suas redes sociais afirmando que o vídeo é uma deepfake: “Todo mundo insiste em dizer que é real, mas todo mundo sabe que é gerado por Inteligência Artificial”, no vídeo ela se encontra revoltada com a situação.

Figura 3: Foto da Atriz Mel Maia.



Fonte: O Globo (2024).

O estudo destaca como as deepfakes têm sido utilizadas para manipular a imagem de figuras públicas, expondo-as a situações vexatórias e prejudiciais. O incidente revela não só a capacidade dessas tecnologias de enganar o público, mas também o potencial de causar danos emocionais e psicológicos.

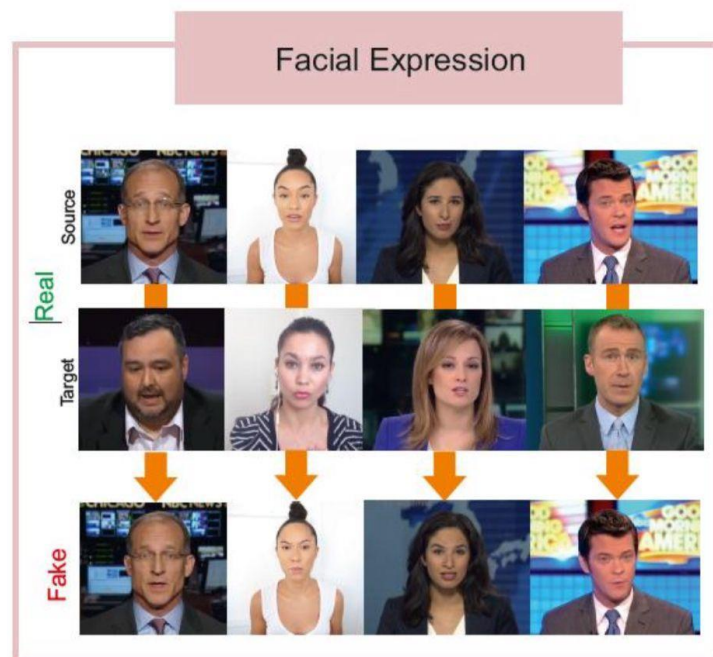
Essa análise vai além da exploração técnica das deepfakes, trazendo uma reflexão sobre as consequências emocionais e sociais que essas manipulações geram em indivíduos específicos, principalmente em uma sociedade em que as mulheres são sempre “objetificadas” e julgadas quando se trata de sensualidade. Esse caso chama atenção para a facilidade com que as deepfakes podem ser criadas e disseminadas, tornando difícil a distinção entre o real e o falso, e prejudicando a reputação de pessoas inocentes.

4.3 Impacto das Deepfakes no Setor Empresarial

O terceiro objeto analisa um caso empresarial alarmante, onde um funcionário de uma multinacional em Hong Kong foi vítima de um golpe utilizando deepfake. Durante uma videoconferência, os golpistas usaram deepfakes para se passar por diretores financeiros da empresa, convencendo o funcionário a transferir US\$ 25 milhões para uma conta fraudulenta. A polícia não revelou nomes e nem detalhes da empresa e do trabalhador para manter a reputação e privacidade das vítimas.

O caso demonstra como as deepfakes, além de causar danos reputacionais, podem ser usadas para esquemas de fraude altamente sofisticados. O estudo aponta que a sofisticação dessa tecnologia permite que golpes como esse sejam extremamente difíceis de detectar, colocando empresas e indivíduos em risco financeiro. Isso deixa um alerta, não só para as pessoas envolvidas no caso, como para outras empresas e trabalhadores, deixando claro a necessidade de medidas de segurança mais rigorosas e de novas tecnologias capazes de identificar e neutralizar esses golpes antes que causem danos irreparáveis. A fraude mostra que as deepfakes têm implicações que vão além da manipulação de imagens públicas e podem ter um impacto significativo nas operações empresariais e financeiras.

Figura 4: Manipulação de expressões faciais.



Fonte: Tolossana et al. (2016).

4.4 Ferramentas de Detecção e Soluções

O estudo também aborda as iniciativas em desenvolvimento para combater a ameaça das deepfakes. Uma das principais ferramentas discutidas é o uso de redes neurais convolucionais para detectar manipulações sutis em vídeos e imagens. Organizações como Facebook, Microsoft e AWS se uniram no “DeepFake Detection Challenge” (DFDC), uma iniciativa que busca desenvolver métodos mais eficazes para identificar deepfakes.

Apesar desses avanços, o estudo destaca que as tecnologias de detecção ainda estão em fase inicial e precisam de mais refinamento. Algumas redes sociais, como Instagram e TikTok, já possuem selos para adicionar a vídeos e imagens que foram criados por Inteligência Artificial. No entanto, só isso não basta, a dificuldade em identificar deepfakes mais sofisticados demonstra a urgência de desenvolver ferramentas mais robustas e acessíveis que possam ser utilizadas não apenas por grandes corporações, mas também por jornalistas e usuários comuns de redes sociais.

4.5 Desafios Éticos e o Futuro das Deepfakes

Sem regulamentações claras, o uso irresponsável dessa tecnologia pode se espalhar rapidamente, minando a confiança pública nas informações disseminadas. Além de avanços tecnológicos, é necessário um esforço conjunto entre governos, empresas de tecnologia e a sociedade para criar diretrizes que regulem o uso ético das deepfakes. Há propostas de regulamentação do uso de IA para alguns países, como as que a União Europeia desenvolveu para garantir que os sistemas dessa nova tecnologia sejam seguros. Campanhas educativas também são essenciais para alertar o público sobre os perigos dessas manipulações e para aumentar a conscientização sobre a necessidade de verificação de informações online.

Além disso, o estudo prevê que as deepfakes continuarão a evoluir, tornando-se mais difíceis de detectar, o que exigirá avanços contínuos em tecnologia de segurança e uma maior alfabetização midiática entre os consumidores de conteúdo digital. Por enquanto, pode-se aplicar algumas técnicas humanas para reconhecer

uma deepfakes, como: qualidade do áudio, movimento da boca e dos olhos e a velocidade do vídeo, algumas tecnologias ainda não conseguem reproduzir esses aspectos de forma realista.

A análise dos dados destaca que os deepfakes representam uma ameaça significativa à comunicação, minando a confiança pública e afetando diferentes setores da sociedade. Embora essa tecnologia traga inovações, ela também expõe vulnerabilidades éticas e institucionais, exigindo uma resposta urgente. Casos como o do Papa Francisco e da atriz Mel Maia ilustram como os deepfakes podem manipular a imagem de líderes e figuras públicas, gerando impactos na percepção pública e na privacidade das pessoas, especialmente no contexto.

No setor empresarial, os deepfakes têm sido usados em fraudes financeiras, evidenciando a necessidade de segurança digital avançada. A tecnologia pode simular reuniões e enganar profissionais, o que exige tanto o desenvolvimento de tecnologias de detecção mais eficientes quanto a implementação de protocolos e treinamentos para reduzir riscos. O estudo enfatiza que, embora existam ferramentas como o “DeepFake Detection Challenge”, uma tecnologia de combate ainda é insuficiente para acompanhar a sofisticação dos deepfakes mais avançados, exigindo um esforço colaborativo de instituições, empresas e do público.

Por fim, o estudo alerta para a urgência de regulamentações e de uma alfabetização midiática mais robusta para combater o uso irresponsável dos deepfakes. O fortalecimento da verificação de informações e a criação de políticas éticas são essenciais para preservar a confiança pública. A conclusão aponta que uma abordagem integrada, que combina tecnologia, legislação e educação, é crucial para enfrentar os desafios dos deepfakes e proteger a comunicação transparente e a veracidade das informações.

5 CONCLUSÃO

A presente pesquisa abordou o impacto das deepfakes na comunicação contemporânea, destacando os desafios que essa tecnologia apresenta para a confiança pública e para a credibilidade da informação em contextos variados, incluindo o religioso, o artístico e o empresarial. Ao analisar casos reais como as imagens falsas do Papa Francisco, o vídeo deepfake de Mel Maia e a fraude empresarial, fica evidente que as deepfakes não apenas desafiam a percepção do

que é real, mas também acarretam consequências graves para indivíduos e organizações, expondo vulnerabilidades.

Os resultados deste estudo reforçam que a disseminação de deepfakes coloca em risco a integridade das informações e potencializa a desinformação em larga escala. O fenômeno observado no caso do “Papa Fashion”, onde uma imagem manipulada do pontífice viralizou nas redes sociais, evidenciou a rapidez com que conteúdos falsos podem enganar o público, inclusive veículos de imprensa renomados. Esse exemplo ilustra a fragilidade das barreiras de verificação e a facilidade com que a manipulação digital pode iludir até mesmo aqueles mais atentos. A situação se repete no âmbito artístico, como observado no caso de Mel Maia, onde um vídeo falso comprometeu a imagem da atriz e provocou um impacto emocional, levantando questões éticas sobre a proteção da privacidade e a responsabilidade dos criadores de conteúdo digital.

No setor empresarial, a pesquisa revelou que as deepfakes possuem o potencial de causar prejuízos financeiros substanciais, como no caso da multinacional que sofreu um golpe milionária com deepfakes. Esse exemplo demonstra como a tecnologia, quando utilizada de forma maliciosa, pode comprometer a segurança corporativa e afetar negativamente a confiança nos meios de comunicação internos e externos das organizações.

Diante desses exemplos, torna-se claro que a tecnologia de deepfake exige atenção e ações proativas para mitigar seus impactos. Iniciativas voltadas para a criação de ferramentas de detecção de deepfakes, como o DeepFake Detection Challenge, liderado por empresas como Facebook e Microsoft, representam avanços importantes, mas ainda insuficientes para lidar com a crescente sofisticação dessas manipulações. As redes neurais e outras técnicas de inteligência artificial desenvolvidas para identificar manipulações digitais são essenciais, mas a complexidade da detecção aumenta à medida que as deepfakes se tornam mais avançadas e indistinguíveis da realidade.

Além do desenvolvimento tecnológico, há uma demanda urgente por regulamentação e ética digital. A ausência de diretrizes claras e a falta de regulamentação específica para o uso de deepfakes criam um ambiente de incerteza e risco. Governos, empresas de tecnologia e a sociedade civil precisam se unir para estabelecer normas que garantam o uso ético da tecnologia, protegendo tanto a privacidade quanto a integridade das informações. A conscientização do público sobre

os perigos das deepfakes e a promoção da educação digital são igualmente cruciais para capacitar as pessoas a identificarem conteúdos falsos e adotarem uma postura crítica frente à informação consumida.

Por fim, este estudo sugere que a combinação de tecnologia, regulamentação e educação pode ser um caminho promissor para enfrentar os desafios impostos pelas deepfakes na comunicação. A criação de uma cultura de verificação de informações e o fortalecimento da educação midiática são essenciais para que os usuários desenvolvam uma postura de desconfiança saudável, capaz de distinguir o conteúdo real do manipulado. Somente com a união desses esforços será possível mitigar os efeitos negativos das deepfakes, promovendo um ambiente digital mais seguro e confiável para a comunicação.

Concluimos que as deepfakes são uma tecnologia poderosa e de grande impacto, e embora possam ter aplicações positivas, o seu potencial para uso mal-intencionado exigem atenção contínua e adaptativa. Esse estudo lança as bases para futuras investigações, que poderão focar na evolução das técnicas de deepfake, nos novos métodos de detecção e nas estratégias educacionais mais eficazes para lidar com as ameaças que essa tecnologia representa.

REFERÊNCIAS BIBLIOGRÁFICAS

BARBOSA, Felipe. **Estudo, desenvolvimento e comparação de técnicas de detecção de Deepfake**. 2020. Projeto de Formatura I do Departamento de Sistemas de Computação do Instituto de Ciências Matemáticas e de Computação – ICMC-USP. Acesso em: 10 nov. 2024.

CAMPOS, Claudia Abreu. FELIPE, Pedro. Usina da Comunicação. **Deepfake**. [sd]. Disponível em: <https://usinadacomunicacao.com.br/deepfake>. Acesso em: 10 nov. 2024.

FORBES. **Imagens do "Papa de Fashion" geradas pelo MidJourney**. [Imagem]. 2024. Disponível em: <https://forbes.com.br/forbes-tech/2024/01/papa-francisco-alerta-sobre-perigos-da-ia-a>. Acesso em: 10 nov. 2024.

G1, Globo. **Imagens falsas mostram Papa Francisco de casaco branco 'estiloso' e viralizam**. 2023. Disponível em: <https://g1.globo.com/tecnologia/noticia/2023/03/26/imagens-falsas-mostram-papa-francisco-de-casaco-branco-estiloso-e-viralizam.ghtml>. Acesso em: 10 nov. 2024.

LEANDRO, Jorge de Jesus Gomes. **Deepfake: Explorando técnicas de detecção de manipulação digital de imagem de faces**. 2022. Departamento de Ciências de Computação do Instituto de Ciências Matemáticas e de Computação, Universidade de São Paulo - ICMC/USP. Acesso em: 10 nov. 2024.

O GLOBO, Globo. **Deepfake: entenda técnica de manipulação que expôs Mel Maia na web**. 2024. Disponível em: <https://oglobo.globo.com/ela/noticia/2024/07/05/deepfake-entenda-tecnica-de-manipulacao-de-imagem-que-expos-mel-maia-na-web.ghtml>. Acesso em: 10 nov. 2024.

PACETE, Luiz Gustavo. Forbes Tech. **O que a imagem do Papa diz sobre o futuro da IA?** 2023. Disponível em: <https://forbes.com.br/forbes-tech/2023/03/o-que-a-imagem-fake-do-papa-diz-sobre-o-futuro-da-ia/>. Acesso em: 10 nov. 2024.

PODER360. **Retratação da Vogue Brasil pelo X, antigo Twitter**. [Imagem]. 2024. Disponível em: <https://www.poder360.com.br/midia/imagem-falsa-de-papa-francisco-com-casaco-volumoso-viraliza/>. Acesso em: 10 nov. 2024.

SECURITY, Homeland. **Increasing Threats of Deepfake Identities.** [sd]. United States Government or the Public-Private Analytic Exchange Program. Acesso em: 10 nov. 2024.