



**UNIVERSIDADE ESTADUAL DA PARAÍBA  
CAMPUS I  
CENTRO DE CIÊNCIAS JURÍDICAS  
CURSO DE BACHARELADO EM DIREITO**

**JOSÉ JEFFERSON CUNHA SOARES**

**IMPLICAÇÕES JURÍDICAS NO USO DE IMAGENS MANIPULADAS NA  
PRODUÇÃO DE *FAKE NEWS* E DESINFORMAÇÃO**

**CAMPINA GRANDE  
2025**

JOSÉ JEFFERSON CUNHA SOARES

**IMPLICAÇÕES JURÍDICAS NO USO DE IMAGENS MANIPULADAS NA  
PRODUÇÃO DE *FAKE NEWS* E DESINFORMAÇÃO**

Trabalho de Conclusão de Curso (Artigo) apresentado à Coordenação do Curso de Direito da Universidade Estadual da Paraíba, como requisito parcial à obtenção do título de Bacharel em Direito.

**Área de Concentração:** Transformação das relações privadas em meios eletrônicos.

**Orientador:** Prof. Me. Caio José Arruda Amarante de Oliveira.

**CAMPINA GRANDE  
2025**

É expressamente proibida a comercialização deste documento, tanto em versão impressa como eletrônica. Sua reprodução total ou parcial é permitida exclusivamente para fins acadêmicos e científicos, desde que, na reprodução, figure a identificação do autor, título, instituição e ano do trabalho.

S676i Soares, Jose Jefferson Cunha.  
Implicações jurídicas no uso de imagens manipuladas na produção de fake news e desinformação [manuscrito] / Jose Jefferson Cunha Soares. - 2025.  
41 f. : il. color.

Digitado.

Trabalho de Conclusão de Curso (Graduação em Direito) - Universidade Estadual da Paraíba, Centro de Ciências Jurídicas, 2025.

"Orientação : Prof. Me. Caio José Arruda Amarante de Oliveira, Centro de Ciências Jurídicas".

1. Manipulação de imagens. 2. Fake news. 3. Deepfake. 4. Inteligência artificial. 5. Desinformação. I. Título

21. ed. CDD 070.4

JOSE JEFFERSON CUNHA SOARES

IMPLICAÇÕES JURÍDICAS NO USO DE IMAGENS MANIPULADAS NA  
PRODUÇÃO DE FAKE NEWS E DESINFORMAÇÃO

Trabalho de Conclusão de Curso  
apresentado à Coordenação do Curso  
de Direito da Universidade Estadual da  
Paraíba, como requisito parcial à  
obtenção do título de Bacharel em  
Direito

Aprovada em: 02/06/2025.

BANCA EXAMINADORA

Documento assinado eletronicamente por:

- **Joana Darc Araujo Ferreira** (\*\*\*.370.504-\*\*), em **10/06/2025 19:27:42** com chave **20054536464a11f0a4ba06adb0a3afce**.
- **Caio José Arruda Amarante de Oliveira** (\*\*\*.200.734-\*\*), em **10/06/2025 15:31:04** com chave **112da272462911f098b806adb0a3afce**.
- **Claudio Simao de Lucena Neto** (\*\*\*.185.374-\*\*), em **11/06/2025 17:47:13** com chave **40cc6fe0470511f086e306adb0a3afce**.

Documento emitido pelo SUAP. Para comprovar sua autenticidade, faça a leitura do QrCode ao lado ou acesse [https://suap.uepb.edu.br/comum/autenticar\\_documento/](https://suap.uepb.edu.br/comum/autenticar_documento/) e informe os dados a seguir.

**Tipo de Documento:** Folha de Aprovação do Projeto Final

**Data da Emissão:** 16/06/2025

**Código de Autenticação:** 3679b8



“O que fazemos na vida ecoa na eternidade”  
Máximus Décimus Meridius - Gladiador

## LISTA DE FIGURAS

<b>Figura 1 - Retrato Abraham Lincoln.....</b>	<b>9</b>
<b>Figura 2 - Imagem de Stalin e Kirov, 1925.....</b>	<b>10</b>
<b>Figura 3 - Mergulhador cercado por peixes coloridos em águas claras....</b>	<b>15</b>
<b>Figura 4 - Processo de criação dos deepfakes.....</b>	<b>19</b>
<b>Figura 5 - Processo de criação com os dados disponíveis.....</b>	<b>20</b>
<b>Figura 6 - Mulher com longos cabelos prateados e olhos azuis.....</b>	<b>23</b>

## SUMÁRIO

<b>1 INTRODUÇÃO.....</b>	<b>7</b>
<b>2 A EVOLUÇÃO DA MANIPULAÇÃO DE IMAGENS.....</b>	<b>8</b>
<b>3 O PAPEL DAS IMAGENS MANIPULADAS NA DESINFORMAÇÃO E FAKE NEWS.....</b>	<b>16</b>
<b>4 LEGISLAÇÃO APLICÁVEL E SEUS DESAFIOS PERANTE O AVANÇO TECNOLÓGICO.....</b>	<b>23</b>
<b>5 METODOLOGIA.....</b>	<b>30</b>
<b>6 CONCLUSÃO.....</b>	<b>31</b>
<b>REFERÊNCIAS.....</b>	<b>32</b>

# IMPLICAÇÕES JURÍDICAS NO USO DE IMAGENS MANIPULADAS NA PRODUÇÃO DE *FAKE NEWS* E DESINFORMAÇÃO

## LEGAL IMPLICATIONS OF THE USE OF MANIPULATED IMAGES IN THE PRODUCTION OF *FAKE NEWS* AND DISINFORMATION

SOARES, José Jefferson Cunha<sup>1</sup>

### RESUMO

O artigo trata das implicações jurídicas no uso de imagens manipuladas, com o fito de produzir *fake news* e desinformação, usando inteligência artificial. Abordou-se a evolução tecnológica que trouxe os *deepfakes* e seu papel na propagação de conteúdos falsos. Além disso, analisou-se os riscos à honra e privacidade, questionando a eficácia do ordenamento jurídico brasileiro. A metodologia adota o tipo de abordagem qualitativa, com pesquisa bibliográfica e documental, focando em legislações e jurisprudências, além de utilizar o método hipotético-dedutivo. Mostrou-se os tipos de uso dos *deepfakes* e como as normas são aplicadas frente às limitações legais atuais. Logo, normas como o Código Penal aplicam-se por analogia, porém não têm especificidade. Conclui-se que o ordenamento jurídico brasileiro não está preparado para o fenômeno, precisando se adaptar e formular leis específicas, além de tecnologias de detecção para proteger direitos fundamentais frente aos desafios tecnológicos.

**Palavras-chave:** Manipulação de imagens, *Fake news* e desinformação, *deepfakes*, inteligência artificial.

### ABSTRACT

This article discusses the legal implications of using manipulated images to produce fake news and misinformation using artificial intelligence. It addresses the technological evolution that led to deepfakes and their role in spreading false content. In addition, it analyzes the risks to honor and privacy, questioning the effectiveness of the Brazilian legal system. The methodology adopts a qualitative approach, with bibliographic and documentary research, focusing on legislation and case law, in addition to using the hypothetical-deductive method. It shows the types of use of deepfakes and how the rules are applied in light of current legal limitations. Therefore, rules such as the Penal Code apply by analogy, but are not specific. It is concluded that the Brazilian legal system is not prepared for the phenomenon, needing to adapt and formulate specific laws, in addition to detection technologies to protect fundamental rights in the face of technological challenges.

**Keywords:** Image manipulation, Fake News; Disinformation, deepfakes, Artificial intelligence.

---

<sup>1</sup> Graduando do curso de Direito do Centro de Ciências Jurídicas da Universidade Estadual da Paraíba. E-mail: josejeffersoncs3@gmail.com

## 1 INTRODUÇÃO

A presente pesquisa aborda a manipulação de imagens, sendo um fenômeno cada vez mais presente no cotidiano, impulsionado pelos avanços tecnológicos que tornaram esses recursos mais acessíveis e sofisticados. Inicialmente, essa manipulação se restringia a processos manuais e técnicas demoradas e caras, porém, houve uma evolução com o surgimento da fotografia e da computação gráfica. Assim, passou a ser algo muito utilizado em diversas situações, como na arte, na publicidade e na política. Porém, se por um lado essa evolução possibilitou novas formas de expressão e inovação, por outro, trouxe implicações preocupantes, principalmente diante do uso mal intencionado dessas imagens para prejudicar terceiros.

Nesse contexto, a presente pesquisa busca analisar as implicações jurídicas da manipulação de imagens, com destaque nos desafios decorrentes do uso dessas tecnologias na divulgação de desinformação, especialmente nas redes sociais, dado o seu grande alcance, atingindo rapidamente uma grande quantidade de pessoas. A facilidade com que se pode criar conteúdos visuais falsificados atualmente, inclusive por meio de inteligência artificial generativa, torna necessário a discussão sobre a eficácia da legislação brasileira para lidar com esse fenômeno. Portanto, frente à dualidade entre o avanço tecnológico e a evolução normativa, faz-se necessário uma análise crítica da legislação vigente, bem como a reflexão sobre a necessidade de novos instrumentos jurídicos que assegurem a proteção de direitos fundamentais para esses riscos emergentes.

Dessa forma, o trabalho parte das dificuldades jurídicas enfrentadas diante do uso cada vez mais sofisticado de imagens manipuladas, focando nas imagens criadas com uso de inteligência artificial, funcionando como instrumentos de disseminação de *fake news* e desinformação, assim, questiona-se em que medida o ordenamento jurídico brasileiro está preparado para lidar com essas condutas maliciosas e quais mecanismos legais são eficazes ou quais devem ser levados em pauta para criação, para proporcionar a proteção dos direitos fundamentais frente ao avanço tecnológico acelerado.

Com base nesse aspecto, os objetivos da pesquisa são compreender os mecanismos técnicos e sociais que envolvem a manipulação de imagens por meio da inteligência artificial e sua evolução, tornando-se atualmente algo extremamente realista, bem como identificar o papel das imagens manipuladas no processo de desinformação e *fake news*, principalmente nas redes sociais, além de analisar os principais impactos jurídicos decorrentes do uso indevido dessas imagens e os dispositivos legais atualmente aplicáveis e suas limitações.

Sendo assim, o trabalho adota o método hipotético-dedutivo, com uma abordagem qualitativa, com foco em pesquisa bibliográfica e documental, tendo em vista a natureza recente e tecnológica do tema. A análise se baseia em doutrinas, legislações, jurisprudências e estudos interdisciplinares que envolvem Direito, Computação e Arte. A partir dessa metodologia, busca-se compreender os impactos jurídicos e sociais da manipulação de imagens com inteligência artificial, sobretudo no contexto da desinformação e das *fake news*, avaliando também as limitações da legislação brasileira frente a esses desafios contemporâneos.

A escolha do tema surgiu após o autor ser surpreendido por um conteúdo falso em uma publicação nas redes sociais, só percebendo que tratava-se de um *deepfake* após ler os comentários. Com isso, a situação evidenciou a sofisticação dessas manipulações e os riscos que representam à honra e à privacidade. Diante disso, a pesquisa se justifica pela crescente relevância do tema devido à rapidez com que essa tecnologia passou a ser usada pela população em geral e pela insuficiência da legislação brasileira frente aos avanços tecnológicos, especialmente no uso da inteligência artificial para criar imagens falsas com potencial de causar graves danos sociais e jurídicos. Logo, o estudo busca contribuir para o debate sobre a proteção dos direitos fundamentais e a necessidade de atualização normativa diante dessa nova realidade.

Portanto, espera-se que o estudo contribua para a compreensão dos impactos jurídicos do uso de imagens manipuladas na disseminação de *fake news* e desinformação, estimulando o desenvolvimento legal, tendo em vista o destaque às lacunas da legislação brasileira, assim, sugerindo caminhos para uma regulamentação mais eficaz. Além disso, objetiva-se promover a reflexão sobre o equilíbrio entre liberdade de expressão e proteção de direitos fundamentais em um cenário social amplamente preenchido por redes sociais que são o caminho perfeito para que *deepfakes* se espalhem, assim, profissionais do Direito podem desenvolver soluções mais adequadas ao contexto digital atual.

## 2 A EVOLUÇÃO DA MANIPULAÇÃO DE IMAGENS

Inicialmente, nota-se que a manipulação de imagens é uma prática que se iniciou séculos antes da invenção da fotografia. Desde os primórdios da arte, artistas buscaram alterar ou idealizar representações visuais para atender a propósitos políticos, religiosos ou culturais. No entanto, foi com o advento da fotografia, no século XIX, que a manipulação de imagens começou a se tornar uma ferramenta amplamente utilizada para influenciar a percepção pública, seja para fins artísticos, propagandísticos ou até fraudulentos. Nesse sentido, a história da fotografia mostra que a "manipulação da imagem" - *lato sensu* - é inerente ao processo de produção de imagens fotográficas (Sousa, 1998; Mitchell, 1992; Wheeler, 2002; Lister, 1997; Fontcuberta, 1998).

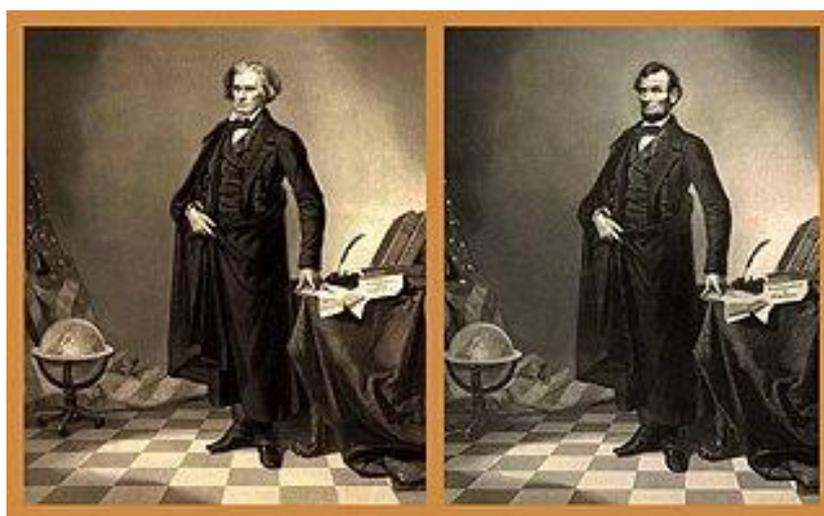
Antes da fotografia, a pintura e outras formas de arte visual já eram utilizadas para alterar ou distorcer a realidade. No campo das artes plásticas, as esculturas da Grécia Antiga, apresentavam linhas e contornos idealizados, sendo assim, antes mesmo da popularização de academias modernas e do estilo de vida saudável, as esculturas expunham corpos musculosos e bem definidos. Nesse sentido, representações como a cabeça do cavalo, no lado leste do Parthenon, são consideradas a expressão do próprio ser, nesse caso, chegou a ser vista por Goethe, como a ideia platônica de cavalo (Spivey, 1996). Portanto, nesse período já havia um certo grau de distorção do real, todavia, essas alterações não causavam prejuízos à terceiros.

Na Idade Média e no Renascimento, por exemplo, era comum que artistas retratassem monarcas e líderes religiosos de maneira idealizada, removendo imperfeições físicas ou inserindo símbolos de poder e divindade. Retratos encomendados por reis e nobres muitas vezes mostravam os retratados mais jovens, esbeltos e imponentes do que eram na realidade,

evidenciando o uso da manipulação visual como ferramenta de controle da imagem pública. Além disso, alguns artistas eram extremamente habilidosos desenvolvendo técnicas que permitiam passar a impressão de profundidade e realismo, nesse sentido, o pintor Vasari afirma que o quadro “Virgem”, de Leonardo da Vinci, parecia ser mais real do que a própria realidade (Lichtenstein, 2004), assim como a manipulação contemporânea busca transcender a realidade visual.

Com a invenção da fotografia no início do século XIX, a manipulação de imagens tornou-se mais sofisticada. A fotografia, inicialmente vista como um reflexo exato da realidade, logo foi utilizada para criar ilusões. Algumas das primeiras técnicas de manipulação fotográfica incluíam dupla exposição, montagem de negativos e retoques manuais nas cópias impressas. Um dos primeiros registros históricos de manipulação fotográfica remonta a meados do século XIX. Em 1860, um famoso retrato do presidente dos Estados Unidos, Abraham Lincoln, foi amplamente divulgado, no entanto, a imagem era, na verdade, uma montagem. O corpo pertencia ao político John C. Calhoun, e apenas a cabeça de Lincoln foi sobreposta, criando uma imagem mais imponente e adequada para fins propagandísticos, conforme ilustrado na Figura 1.

**Figura 1 - Retrato Abraham Lincoln.**



Fonte: Pate (1865); Ritchie (1852).

Além disso, outro exemplo clássico da manipulação de imagens é a remoção de figuras de fotografias, utilizada para fins políticos. Joseph Stalin, líder da União Soviética, foi um dos pioneiros a usar esse artifício, ele era conhecido por ordenar a exclusão de opositores de registros fotográficos oficiais. Imagens de eventos políticos e reuniões de Estado eram frequentemente alteradas para apagar figuras que haviam caído em desgraça, tornando-as “inexistentes” na memória coletiva, sendo possível encontrar diversas fotografias nas quais a imagem de Leon Trotsky, que era seu rival, foi removida. Além disso, esse objetivo de manipulação na era Stalin não se limitava apenas a imagens, se estendendo a documentos oficiais e secretos, discursos e registros históricos, com a premissa de criar uma versão distorcida da realidade. Na grande parte das vezes, a figura removida era a de Trotsky,

tendo em vista que ele era uma das principais figuras da Revolução Russa, porém, como virou rival de Stalin, foi-se necessário fazer com que ele “nunca tivesse existido”.

Sendo assim, para manipular a opinião pública, Stalin buscava, forjar proximidade com o antigo líder Lenin, ao mandar produzir estátuas, pinturas e fotos falsas, aparecendo ao seu lado, além de apagar as memórias de seus inimigos, nesse sentido, ao lado de Stalin estão Nikolai Antipov, Sergei Kirov e Nikolai Shvernik. Com os sucessivos expurgos, eles foram sumindo um a um, até que sobrasse apenas uma pintura a óleo inspirada nas fotografias, conforme retratado na Figura 2. Esse processo de reescrever a história visualmente, muitas vezes em nome da propaganda política, gerou uma verdadeira "realidade paralela", na qual os eventos eram distorcidos para reforçar a imagem do líder como infalível. O uso de tal manipulação não era apenas uma questão estética, mas parte de uma estratégia maior para moldar a percepção pública e consolidar o poder de Stalin. Ao fazer desaparecer fisicamente os opositores das imagens, ele também os apagava simbolicamente da memória do povo soviético. Esse controle sobre a imagem pública não se limitava à fotografia, mas cruzava toda a produção cultural, social e política do regime, consolidando uma versão oficial da história.

**Figura 2 - Imagem de Stalin e Kirov, 1925.**



Fonte: The David King Collection at Tate (1926).

Com o avanço das técnicas de fotografia e impressão, a manipulação de imagens tornou-se ainda mais refinada. Durante a Primeira e a Segunda Guerra Mundial, os governos utilizavam a edição de fotografias para propaganda, removendo elementos indesejados ou adicionando informações

que favorecessem sua narrativa. E, a partir desse momento, inúmeros governantes começaram a utilizar dessa manipulação para se autopromover, usando os diversos artifícios que foram surgindo ao longo dos anos, seja para criar narrativas ou multiplicar números, afinal, conteúdos que se propagam em larga escala tendem a alcançar certa legitimação social (Recuero, 2020).

Entretanto, essa manipulação não é utilizada apenas para fins políticos, posto que na era do cinema e da publicidade, período compreendido nas primeiras décadas do século XX, a manipulação de imagens também começou a ser explorada para fins comerciais. Empresas alteravam fotografias para tornar produtos mais atrativos, ajustando cores, proporções e detalhes para convencer consumidores, que só tomavam conhecimento da verdadeira forma dos produtos quando os recebiam. Essa prática passou a ter um certo controle com a edição do Código de Defesa do Consumidor em 1990, que, dentre outras normas, proibiu a propaganda enganosa, definindo-a como “qualquer modalidade de informação ou comunicação de caráter publicitário, inteira ou parcialmente falsa, ou, por qualquer outro modo, mesmo por omissão, capaz de induzir em erro o consumidor a respeito da natureza, características, qualidade, quantidade, propriedades, origem, preço e quaisquer outros dados sobre produtos e serviços” (Brasil, 1990). Todavia, essa prática não tem eficácia total, tendo em vista que gigantes empresas, de *fast food* por exemplo, continuam vendendo e divulgando seus lanches com imagens que não condizem com a realidade.

Ademais, verifica-se que ao longo do século XX, foi possível desenvolver a tecnologia de animação de faces que evoluiu de simples experimentos gráficos manuais para complexos sistemas digitais, sendo um dos marcos mais relevantes na história da arte animada. Em 1906, James Stuart Blacktin inaugurou essa tecnologia com o curta metragem *Humorous phases of funny faces*, no qual desenhos quadro a quadro simulavam sorrisos e caretas, porém ainda era uma animação bastante simples. Alguns anos depois, Winsor McCay aperfeiçoou essas expressões faciais em *Gertie the Dinosaur* (1914), sendo considerado um dos primeiros personagens a possuir personalidade. Todavia, foi na década de 1930 que Walt Disney revolucionou as animações trazendo a articulação facial em longas-metragens como *Branca de Neve e os Sete Anões* (1937), utilizando até mesmo de estudos de anatomia e fisiologia para garantir maior naturalismo às emoções. Com o advento da computação gráfica, foi possível desenvolver a captura de movimento facial (*facial motion capture*) sendo uma grande revolução ao registrar microexpressões reais, utilizadas em cenas marcantes de produções como *O Segredo do Abismo* (1989) e, posteriormente, em animações e videogames.

Com o surgimento da computação gráfica na segunda metade do século XX, a edição de imagens deixou de depender exclusivamente de métodos analógicos e passou a ser feita digitalmente. Esse avanço marcou o início de uma nova era para a manipulação de imagens, que se tornaria ainda mais sofisticada e acessível com a chegada dos softwares de edição na década de 1990, como o *Photoshop*. Nesse sentido, os conteúdos midiáticos na ecologia digital são criados, editados, organizados, remixados, experimentados por meio de *softwares*. A dimensão desses *softwares* inclui aplicativos de *design*, gerenciamento e edição, como *Photoshop*, *Illustrator*, *After Effects* e variados aplicativos específicos dos sistemas operacionais de *smartphones*, que

associam ferramentas para editar e compartilhar em plataformas de redes sociais, se tornando cada vez mais fáceis de se operar (Manovich, 2011). No entanto, novos desafios surgiram, principalmente no tocante à manipulação das imagens criadas digitalmente, ou edição das fotografias tiradas, permitindo a propagação de imagens falsas para fins fraudulentos.

Dessa forma, o aprimoramento dessas técnicas ao longo da história demonstra como a imagem nunca foi um reflexo absoluto da realidade, mas sim um meio passível de alteração e interpretação. Sendo assim, entre a fotografia e a flor (objeto) há um aparelho, logo, há algo de subjetivo na imagem, portanto uma fotografia é tão subjetiva como qualquer outra imagem, pois entre a coisa fotografada e a imagem intromete-se um aparelho (Flusser, 2014). Com o desenvolvimento das tecnologias digitais e, mais recentemente, da inteligência artificial generativa, a manipulação de imagens atingiu um nível sem precedentes de sofisticação e disseminação. Esse fenômeno levanta questões éticas e jurídicas complexas, especialmente no contexto da desinformação e das *fake news*, tema que será aprofundado ao longo deste trabalho.

Nesse aspecto, é importante destacar como surgiram os *softwares* de edição e como se deu seu avanço, posto que surgiram a partir da computação gráfica, que podemos conceituar como uma área da Ciência da Computação que se dedica ao estudo e desenvolvimento de técnicas e algoritmos para a geração (síntese) de imagens através do computador, estando presente em quase todas as áreas do conhecimento humano (Manssour; Cohen, 2006). Assim, o início da computação gráfica foi na década de 1960, quando as grandes universidades e centros de pesquisa, principalmente nos Estados Unidos, conduziram experimentos ao realizar simulações que utilizavam imagens digitais. Todavia o uso para fins comerciais ou criativos era bastante limitado, pois envolvia altos custos e havia uma certa complexidade. Porém, foi por volta de 1980 que a computação gráfica avançou de maneira significativa, já que os processadores e computadores avançaram e permitiram o desenvolvimento dos primeiros *softwares* específicos para essa função de edição. Pode-se considerar o pioneiro o aplicativo *Quantel Paintbox*, lançado em 1981, permitindo a edição digital de imagens e sua manipulação. Todavia, era mais utilizado no cenário de produções televisivas e na indústria da publicidade, permitindo a criação de efeitos visuais que não eram possíveis de serem feitos de forma natural.

Porém, foi em 1987 que aconteceu talvez o marco mais importante, com o lançamento do *Adobe Photoshop*, criando uma revolução na área, ao apresentar recursos como camadas, filtros e outras ferramentas, tornando-se rapidamente o aplicativo mais utilizado por fotógrafos e *designers*. Portanto, o *Photoshop* possibilitou que a edição de imagens se tornasse mais acessível ao público. Após esse período, começaram a surgir diversas ferramentas de edição como o *CorelDRAW*, *GIMP*, *Photoscape*, permitindo que qualquer pessoa pudesse fazer ajustes nas suas imagens com facilidade, atendendo desde o público com computadores domésticos até o público que produzia imagens de alta qualidade e mais refinadas. Dessa forma, conforme exposto, diversas áreas do cotidiano foram diretamente impactadas pela edição e manipulação de imagens, na publicidade e entretenimento passou-se a utilizar esses artifícios para criar campanhas impactantes utilizando do ajuste de cores e filtros, bem como modificando produtos e criando cenários; na indústria da

moda, contribuiu para a criação de um ideal de beleza, alterando a imagem de modelos que em várias situações se torna algo exacerbado, criando uma “falsa imagem” (Alves, 2008); no cinema foi possível criar universos visuais completamente fantasiosos, no jornalismo, muitas vezes acaba distorcendo a verdade sobre os fatos narrados. Logo, essa rápida difusão dos *softwares* de edição de imagem trouxe a manipulação de imagens uma prática comum na vida das pessoas, porém trouxe preocupações no que tange ao seu uso indevido. A capacidade de alterar imagens com perfeição e precisão, de forma cada vez mais rápida, trouxe questionamentos sobre a veracidade de uma fotografia, que no passado era uma forma autêntica da realidade.

Com isso, percebe-se que essa manipulação de imagens intensificou o problema da desinformação, posto que, no passado, as notícias se difundiam de forma lenta, principalmente por meio de jornais, que, costumeiramente, verificavam a fonte das suas informações. Entretanto, devido ao costume de geralmente receber informações verídicas das fontes de notícias, as pessoas tendem a considerar, nos dias atuais, as informações que encontram na internet como verdadeiras, principalmente quando essa notícia nos agrada, nesse sentido, assim que uma crença que pensamos ser verdadeira é alcançada, ficamos inteiramente satisfeitos, seja verdadeira ou seja falsa (Peirce, 2008). Dessa forma percebe-se que a opinião é moldada por crenças, mesmo que não seja necessariamente verdade, há um impacto na opinião. Nesse viés, casos mundialmente conhecidos demonstraram como a influência da manipulação de imagens pode afetar de diversas maneiras o cenário social e político. Como já demonstrado, em diversas campanhas eleitorais ao redor do mundo, no transcorrer dos anos, foram utilizadas montagens fotográficas para ligar candidatos a situações que não existiram, alimentando a desinformação.

Entretanto, foi com o advento da inteligência artificial generativa, por volta de 2010, que a desinformação atingiu proporções que antes eram inimagináveis, sendo assim, o que antes exigia um alto nível de habilidade e conhecimento prévio, nos dias atuais, com apenas um computador ou celular, qualquer pessoa é capaz de criar imagens falsas de alta complexidade. Percebe-se que essa é uma ferramenta que se popularizou mais recentemente, com a criação do *ChatGPT* em 2022, que funciona como uma plataforma de interação, que rapidamente gera as respostas necessárias para as questões do usuário, dessa forma, ficou mundialmente conhecido, sendo usado para diversos fins, criando códigos de programação, tabelas, analisando dados, traduzindo idiomas ou até mesmo simples recomendações do dia a dia, por sua versatilidade, ficando no topo das listas de *downloads* das lojas dos *smartphones*. Assim, considera-se que a Inteligência Artificial, devido a seus grandes avanços e possibilidades, deixa de ser apenas um campo de pesquisa e passa a ser um determinante da nossa capacidade de categorizar o conhecimento, ao dar-lhe um significado e potencializando nossa capacidade de tomar decisões (Villani, 2018).

Dessa maneira, percebe-se que a Inteligência Artificial é um artifício que já vem sendo aplicado a bastante tempo, sendo um campo de estudo que surgiu após a Segunda Guerra Mundial, por volta de 1950 (Russel, 2013). O conceito de inteligência artificial veio a ser inicialmente conhecido por estar presente em jogos eletrônicos, tendo em vista que os jogos sempre lideraram os avanços tecnológicos da indústria criativa, assim, se fazia presente quando

o computador controlava os oponentes que o jogador viria a enfrentar, utilizando de movimentos que simulavam comportamento inteligente, buscando criar a ideia de jogar contra uma pessoa real, porém, ainda era bastante primitivo. Basicamente, permite que personagens, sejam inimigos ou não jogáveis (NPCs), se comportem como se tivesse mente própria, trazendo maior imersão ao jogo. Dessa forma, pode-se dizer que desde o princípio da Inteligência Artificial a mesma já foi sendo aplicada nos jogos, a exemplo do jogo *Nim*, um antigo jogo chinês que consiste em cada jogador retirar algumas peças desde que estejam na mesma coluna, ganhando quem retirar a última peça. Nesse contexto, é considerado como a primeira aparição dessa tecnologia, a máquina *Nimatron* (1939) já demonstrava o poder que esse tipo de ferramenta poderia proporcionar, já que no período de seis meses apenas algumas pessoas conseguiram vencer a máquina, os poucos vencedores eram presenteados com uma moeda que dizia "*Nim Champ*". A seguir, em 1951, Ferranti construiu um computador que jogava o jogo, apresentando-o no Festival da Grã-Bretanha, inspirando outras máquinas que surgiram posteriormente.

No entanto, foi no período conhecido como a era de ouro dos fliperamas, no final dos anos 70, que a ideia de oponentes controlados por Inteligência Artificial se popularizou, principalmente devido ao sucesso do jogo *Space Invaders* (1978) que apresentava um nível de dificuldade progressivamente crescente, utilizando de padrões de movimento distintos e eventos no jogo que eram determinados por funções específicas, baseadas nas entradas do jogador. Posteriormente, já no ano de 1979, o jogo *Galaxian* introduziu movimentações inimigas mais complexas e diversificadas, simulando, de forma primitiva, comportamento, ao incluir manobras individuais de inimigos que saíam de suas formações. Já *Pac-Man* (1980) inovou ao incorporar padrões de Inteligência Artificial em jogos de labirinto, caracterizando-se pela atribuição de personalidades distintas a cada um dos fantasmas inimigos, trazendo uma dinâmica única à jogatina. Nesse contexto, cada fantasma tem uma programação diferente para responder à localização do boneco "come-come", ao detalhar a dinâmica, pode-se notar que um dos fantasmas é projetado para vagar aleatoriamente pelo labirinto, enquanto outro adota uma abordagem mais agressiva, perseguindo *Pac-Man* de forma direta, o terceiro fantasma segue *Pac-Man* na mesma direção, a menos que esteja sendo perseguido pelo próprio jogador, alterando sua estratégia, por fim, o último inimigo seguirá *Pac-Man* quando ele estiver longe, mas se moverá para um local específico no mapa quando *Pac-Man* estiver perto.

Com isso, pode-se dividir os principais aspectos da Inteligência Artificial Generativa em jogos na sua aplicação em: comportamento de personagens não jogáveis, ao permitir interações mais realistas e adaptáveis a esses personagens propiciando uma imersão quando eles agem naturalmente, reagindo de forma dinâmica às decisões do jogador; ambientação dinâmica ao criar mundos vivos que evoluem com o decorrer do jogo e reagem ao que o jogador faz, como no jogo *Minecraft* (2009), no qual o mundo pode ser destruído e alterado de diversas formas; geração procedidas de conteúdos ao permitir a criação automática de cenários de jogo, fornecendo diversidade e saindo do método "engessado" utilizado em outras épocas, isso permite que todo o universo de um jogo seja gerado proceduralmente oferecendo possibilidades infinitas.

No contexto diário da vida humana, muitas vezes sem saber, a inteligência artificial está presente no cotidiano das pessoas, pois são utilizadas tecnologias baseadas nessa ferramenta, desde as simples palavras sugeridas pelo teclado nos aplicativos de mensagens, como acessar conteúdos que aparecem de forma personalizada nas redes sociais, adequando o algoritmo ao usuário (Kaufman; Santaella, 2020; Fernández Marcial; Gomes, 2022). Assim, todas as ferramentas mencionadas são potencializadas com o uso do *Machine Learning* (aprendizado de máquina), sendo uma área da inteligência artificial capaz de criar códigos que conseguem observar e aprender padrões dentro dos dados fornecidos, dessa forma, o *machine learning* dá à máquina a capacidade de aprender uma determinada tarefa sem ser explicitamente programada, assim como uma criança aprende a reconhecer um brinquedo, por exemplo, ao ser exposta ao objeto várias vezes, identificando, portanto, padrões que caracterizam o conceito daquele objeto (Homem, 2020), logo, é uma tecnologia que constrói a si mesma. Portanto, para se introduzir nessa evolução, as gigantes empresas de tecnologia desenvolveram suas próprias ferramentas categorizadas como *chatbots*, logo, o *Google* criou a *Gemini*, a *Meta* criou sua homônima que pode ser utilizada dentro do próprio *WhatsApp*, o *X (Twitter)* criou o *Grok*, a *Microsoft* criou a *Cortana*, para que dessa forma pudessem cativar ainda mais seus usuários.

Todavia, a Inteligência Artificial utilizada em *videogames* difere da utilizada para a manipulação de imagens, posto que para um jogo busca-se facilitar à programação ou melhorar a forma como um conteúdo é criado, todavia essas ferramentas de imagens permitem substituir rostos e vozes em vídeos com uma precisão gigantesca, tornando ainda mais difícil distinguir o que é real do que é manipulado. Com isso, surgiram diversas ferramentas de inteligência artificial utilizando do modelo conhecido como “*Stable Diffusion*”, uma ferramenta *open-source* que funciona por meio de *prompts* fornecidos pelo usuário e outras variáveis para obter uma imagem conforme suas especificações (Neto, 2022), basicamente você pode digitar “uma pessoa nadando no sol” e sairá uma imagem de alta qualidade, o grande benefício é que esse modelo possui licenciamento permissivo, ou seja, transfere todos os direitos da imagem gerada para o usuário que a criou, desde que não sejam ilegais ou possam prejudicar outros, o que, na prática, não ocorre e muitas das imagens acabam tendo esse uso. Assim, surgiram muitas outras ferramentas baseadas nesse modelo, sendo os mais famosos o *Midjourney*, e *Dall-e*, logo, as imagens são tão perfeitas que na Figura 3 não é possível saber se é uma foto real ou gerada por Inteligência Artificial, havendo uma alternativa ao analisar os metadados da imagem, porém uma pessoa comum não detém esse conhecimento.

**Figura 3 - Mergulhador cercado por peixes coloridos em águas claras.**



Fonte: Midjourney (2025).

Além disso, essa tecnologia surpreendente também é capaz de produzir vídeos, porém, ainda não está no seu ápice, devido a dificuldade de criar alguns recursos como expressões faciais, movimentos do cabelo e do corpo, bem como a iluminação, por isso, muitas vezes as imagens acabam saindo deformadas, criando figuras que inicialmente até parecem normais, mas abruptamente perdem a forma. Assim, o problema da desinformação visual se torna ainda mais complexo, exigindo um olhar atento do Direito para garantir a proteção da veracidade da informação e dos direitos individuais.

### **3 O PAPEL DAS IMAGENS MANIPULADAS NA DESINFORMAÇÃO E FAKE NEWS**

A princípio percebe-se que o fenômeno das *fake news* e desinformação não é algo recente, de fato o termo se popularizou nos últimos anos, podendo ser considerado como uma releitura do antigo fenômeno social da “mentira” (Guimarães; Silva, 2019), entretanto, desde a antiguidade podemos encontrar registros de informações falsas disseminadas em larga escala, como a caça às bruxas que ocasionou mais de 50 mil condenações à morte, sem nenhuma prova concreta ou quando o Rei Felipe II, da Espanha, soube da disseminação de um boato de que ele mesmo tivesse sido morto a tiros, tendo que mover todo o sistema de transmissão de mensagens para informar que se tratava de uma mentira. Dessa forma, observa-se que esse termo passou a ser utilizado por volta de 2016, durante as eleições estadunidenses, atacando e desqualificando os adversários, com isso, outros países também foram afetados por essa desinformação e, considerando seu potencial danoso, tomaram atitudes para minimizar os efeitos causados. Nesse sentido, segundo o dicionário Cambridge, o termo *fake news* indica histórias falsas que parecem ser notícias, espalhadas na internet ou usando outros meios de comunicação, geralmente criadas para influenciar as pessoas. Desse modo, partimos do conceito de desinformação, o qual não é o contrário de informação, mas uma espécie de distorção ética da informação, o que levaria à ampliação das incertezas, em vez de sua redução (Salgado; Mattos, 2022).

Antes de serem reveladas pela mídia, as *fake news* já causavam danos sociais. Em 2014, uma notícia falsa sobre uma mulher que sequestrava crianças para rituais de magia negra, espalhada por uma página no *Facebook*, levou ao linchamento e morte de uma mulher de 33 anos, confundida com a suposta sequestradora. Não havia registros de sequestros na cidade, segundo a polícia. Devido a isso, cinco pessoas foram condenadas, mas o criador da *fake news* ficou impune, pois a lei da época não punia incitação à violência *online*, esse caso serviu de inspiração para a novela “Travessia” (2022). Assim, verifica-se que mesmo com apenas um boato sendo espalhado sem nenhuma prova, já criou-se uma situação grave, com consequências severas, sendo letal para uma pessoa sem qualquer relação com o caso. No âmbito das redes sociais, verifica-se um fenômeno inexistente no mundo real, o uso de robôs com o objetivo de difundir informações em massa, se analisarmos um contexto excluindo as redes sociais, as informações se espalham no “boca a boca”, onde cada indivíduo passa para o outro a notícia, já no ambiente digital, não é

possível saber a identidade da pessoa que está do outro lado, nem há como saber nem se a pessoa de fato existe ou se é um robô ou perfil falso. Assim, é possível tomar como exemplo o experimento realizado pelos pesquisadores do Laboratório de Aplicação de Vídeo Digital da Universidade Federal da Paraíba, no qual uma *fake news* foi criada e propagada por robôs, alcançando 200 mil interações apenas no *Facebook*. Esses números são ainda mais interessantes, tendo em vista que se tratava de um boato na área do esporte, porém, quando o assunto é política, a proporção é muito maior, envolvendo massas, grandes investimentos e proporcionando brigas que envolvem ainda mais as pessoas.

Portanto, o uso de robôs mostra-se como mais uma das ferramentas que surgiram nos últimos anos que permitem a manipulação da opinião pública, posto que o universo semântico da desinformação envolve uma multiplicidade de perspectivas, como informação imprecisa, distorcida, descontextualizada, ressignificada, mal-intencionada, equivocada, boato, ruído e mentira (Alzamora; Mendes; Ribeiro, 2021; Rogers, 2023). Assim, esses robôs têm um papel significativo na produção desse tipo de conteúdo, pois conseguem compartilhar conteúdos falsos em grande escala e de forma automatizada, portanto, se um robô consegue fazer publicações a cada dois segundos, uma pessoa comum leva bem mais tempo para fazer apenas uma publicação, isso gera uma falsa sensação de que a opinião da maioria é a que está sendo veiculado pelos robôs, logo, manipula a percepção pública. Desse modo, em virtude da sua rápida ação, os robôs podem compartilhar, curtir e comentar notícias falsas em massa, tornando-as mais relevantes e confiáveis, validando-as, bem como criando tendências, ao tornar viral o tema em questão, já que por ser uma propagação em massa, enganam os algoritmos ao se passarem por pessoas normais, criando “*trending topics*”, além disso, nota-se que a maior parte das pessoas compartilham *fake news* sem verificar a fonte dos fatos, principalmente se perceberam que muitas pessoas estão compartilhando a mesma notícia ou se a informação agrada sua ideologia. Dessa forma, surgiu o termo “*junk news*” (notícias-lixo), abrangendo um universo de práticas variadas envolvendo conteúdos enganosos, sendo definido como informação ideologicamente extrema, enganosa e incorreta (Marchal *et al*, 2019).

No mais, existindo a possibilidade de anexar imagens a cada publicação, ao usar as imagens manipuladas, torna-se mais fácil criar uma narrativa plausível porém inexistente. Afinal, é mais fácil acreditar que uma pessoa fez alguma coisa vendo uma foto na qual supostamente a pessoa está fazendo o que está sendo alegado. Nesse sentido, nota-se também que as *fake news* e a desinformação se espalham de forma mais rápida do que nas décadas anteriores, tendo em vista que a inovação tecnológica e a globalização alteram a relação entre o espaço e o tempo, dessa forma o mundo parece “encolher” (Harvey, 1992), facilitando a difusão de informações. Atrelado a isso, essa problemática torna ainda mais necessária de se analisar quando se diz respeito a produção de provas em processos judiciais bem como sua repercussão penal nesse contexto, posto que a habilidade de criar evidências fictícias abre a possibilidade de contestar a autenticidade das provas apresentadas em um processo jurídico.

Logo, se no passado uma prova documental era vista como incontestável, nos dias atuais não é essa a realidade. No Direito, a inteligência artificial já vem sendo empregada para diferentes funções, seja para analisar documentos extensos, prever decisões judiciais ou realizar o reconhecimento

facial em investigações criminais. Por um lado, projetos como o “*Smart Sampa*” buscam utilizar dessa função para realizar a captura de foragidos da justiça por meio do monitoramento das 25 mil câmeras presentes no município, que são equipadas com a tecnologia, sendo realizadas mais de 900 prisões, desde sua implementação em julho de 2024, nota-se, portanto, que sem o auxílio da inteligência artificial para realizar o reconhecimento facial, essas prisões não seriam realizadas e talvez essas pessoas nunca fossem encontradas pela justiça, como é o caso do homem que passou 30 anos foragido e foi capturado graças a esse trabalho. Porém, por outro lado, essa tecnologia não é 100% precisa, podendo cometer erros, como por exemplo o erro cometido com o idoso de 80 anos, Francisco Ferreira da Silva, que passou mais de 10 horas detido na delegacia após ser confundido com um estuproador pelo reconhecimento facial, evidenciando como essas tecnologias podem ser perigosas, posto que esses algoritmos programados por seres humanos podem refletir seus próprios preconceitos no uso direcionado às tomadas de decisões.

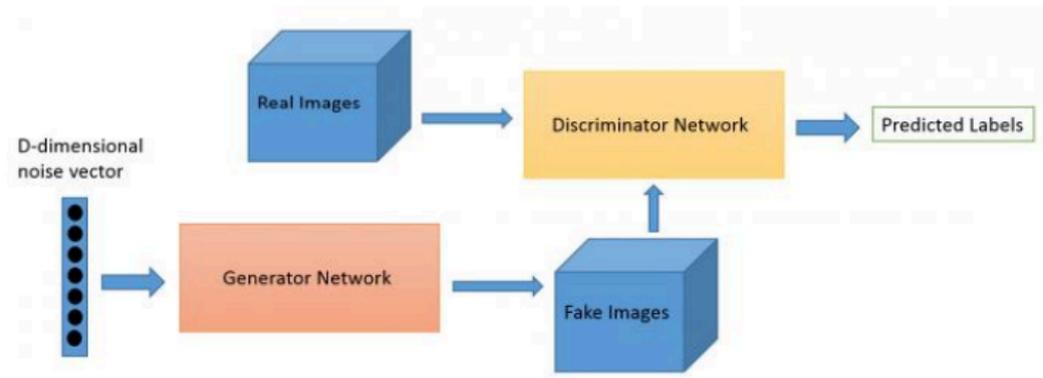
Logo, é mister ressaltar que o Direito busca alcançar a verdade, porém, mesmo sendo impossível alcançar a verdade absoluta, torna-se cada vez mais difícil criar uma verdade formal, dentro dos autos de um processo, posto que aquilo trazido pelas partes pode não refletir a realidade, pois é possível que tenha sido criado artificialmente, ou manipulada, tendo em vista que mesmo com apenas um *bit* tendo sido alterado, é como uma violação do DNA do arquivo, portanto, não devendo ser admitido e avaliado em juízo (Domingos, 2018). Nesse sentido, é essencial observar a cadeia de custódia da prova produzida, pois com o conjunto de procedimentos que devem ser observados, é possível garantir a integridade e confiabilidade da prova penal (Prado, 2019). Portanto, no caso de provas digitais, a cadeia de custódia é ainda mais importante, já que conforme exposto, nos dias atuais é fácil manipular uma imagem, principalmente se o manipulador detiver de um bom conhecimento técnico quanto as ferramentas disponíveis, tornando difícil categorizar a prova como falsa, sendo necessário a análise de especialistas. Dessa forma, para que a prova digital seja válida ela deve ser íntegra, portanto, deve ser completa e não ter sido alterada, sendo necessário sua verificação por meio de técnicas periciais como a identificação da assinatura digital do seu arquivo ou o código *hash* (Domingos, 2018).

Ainda, entre essas novas tecnologias, a que traz mais perigo são os *deepfakes*, posto que ao analisar e aprender padrões faciais, expressões e movimentos, apresentam uma grande dificuldade de distinguir o original do falso, ao criar vídeos e imagens, por meio de inteligência artificial, extremamente realistas, nos quais rostos são substituídos, roupas sejam removidas, ou gerando até mesmo faixas de áudio simulando a voz de uma pessoa, ou seja, criando uma situação onde a pessoa disse ou fez algo que na verdade nunca existiu. Assim, os *deepfakes* exercem um papel central na divulgação de desinformação e *fake news*, combinando a credibilidade que um vídeo propicia com a capacidade de gerar conteúdos completamente improcedentes, conseguindo, dessa maneira, enganar a população em geral de forma mais assertiva do que apenas um texto contando mentiras ou montagens fotográficas. Portanto, além do impacto na percepção da verdade, o uso de *deepfakes* representa uma grave ameaça aos direitos fundamentais, especialmente ao direito à honra, à imagem e à personalidade. Dessa maneira, pessoas públicas e anônimas podem ter suas identidades visual e vocal

indevidamente associadas a comportamentos, falas ou situações que jamais existiram, esses acontecimentos podem resultar em danos irreparáveis à sua reputação, dignidade e vida pessoal, se agravando diante da velocidade com que esse tipo de conteúdo é compartilhado nas redes sociais, sem qualquer verificação de autenticidade. Diante disso, os *deepfakes* não apenas ampliam o alcance da desinformação, mas também evidenciam a linha tênue entre os limites entre liberdade de expressão e responsabilidade jurídica, exigindo do Direito uma resposta mais célere e eficaz, tanto no âmbito legislativo quanto judicial, para proteger os direitos individuais e a integridade das informações compartilhadas no espaço digital.

Com isso, é importante conceituar o termo *deepfake*, que de acordo com o Guia Ilustrado Contra as *Deepfakes* (2024), do Supremo Tribunal Federal, é uma junção das palavras *deep learning* (aprendizado profundo) e *fake* (falso), baseando-se em redes neurais generativas, principalmente as GANs (Redes Generativas Adversariais), funcionando a partir da coleta de dados, ao analisar milhares de imagens ou vídeos da pessoa a ser falsificada, sendo diretamente proporcional a quantidade de material fornecido à precisão das imagens resultantes, além do treinamento ao aprender a identificar os traços individuais de cada alvo, recriando uma versão digital, assim, possibilitando a sobreposição de rosto ou áudio em um vídeo já existente, aumentando a credibilidade. Essas Redes Generativas Adversárias têm seu funcionamento baseadas nos modelos de *deep learning* (aprendizado profundo), sendo compostas por duas redes neurais, existindo um modelo gerador que gera novas instâncias de dados e um modelo discriminador que classifica se esses dados pertencem ou não à base de dados de treinamento real, conforme ilustrado pela Figura 4:

**Figura 4 - Processo de criação dos *deepfakes*.**



Fonte: Deep Learning Book (2022).

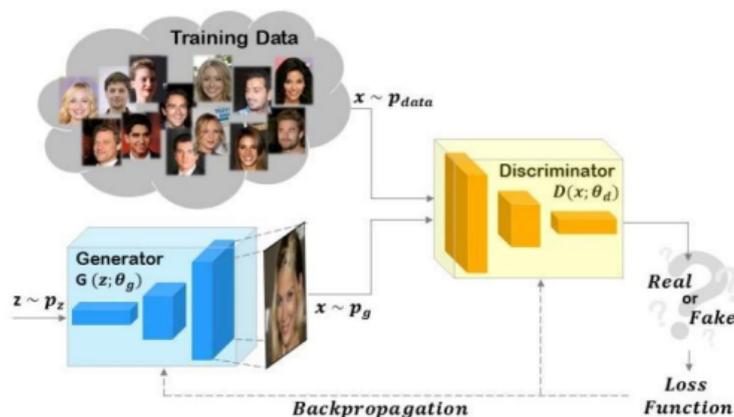
Sendo assim, o processo de criação dos *deepfakes* ganhou força com a evolução das Redes Generativas Adversárias, por meio dos conceitos de Arthur Samuel, criando um dos primeiros programas capazes de aprender com sucesso, para tanto, utiliza-se dos modelos de duas redes neurais que competem entre si, o gerador e o discriminador. O gerador tem como objetivo criar novas instâncias de dados, como imagens falsas, que sejam indistinguíveis dos dados reais. O discriminador, por sua vez, atua como um "avaliador", classificando se os dados que recebe são reais ou foram gerados

pelo modelo, dessa forma, os algoritmos buscam otimizar suas estratégias, produzindo conteúdos de alta qualidade, assim, a dualidade dos algoritmos se mostra extremamente eficiente para esse desenvolvimento, nesse sentido, conforme Wiggers:

A ideia de colocar dois algoritmos um contra o outro se originou com Arthur Samuel, um proeminente pesquisador no campo da ciência da computação que popularizou o termo "aprendizado de máquina". Enquanto estava na IBM, ele desenvolveu um jogo de damas - o Samuel *Checkers-playing Program* - que foi um dos primeiros a aprender com sucesso, em parte por estimar a chance de vitória de cada lado em uma determinada posição. Mas se Samuel é o avô dos GANs, Ian Goodfellow, ex-cientista pesquisador do *Google Brain* e diretor de aprendizado de máquina do Grupo de Projetos Especiais da Apple, pode ser o pai deles. Em um artigo de pesquisa intitulado "Redes adversárias generativas" de 2014, Goodfellow e seus colegas descrevem a primeira implementação prática de um modelo generativo baseado em redes adversárias (Wiggers, 2019).

Basicamente, a rede geradora pode ser assemelhada aos falsificadores que tentam produzir cédulas de dinheiro perfeitas e indetectáveis para seu uso, enquanto a rede discriminadora é análoga a polícia, buscando identificar se as notas são verdadeiras ou falsas, ambas buscam sempre melhorar seus métodos, até chegar ao ponto que as falsificações sejam indistinguíveis das notas reais (Goodfellow *et al*, 2020). Essa analogia ajuda a compreender como o funcionamento das GANs torna os *deepfakes* cada vez mais convincentes e difíceis de detectar. A interação contínua entre essas duas redes resulta em um aprimoramento progressivo da qualidade do conteúdo sintético, elevando significativamente o risco de manipulação maliciosa de vídeos e imagens. Esse processo coloca em evidência o enorme desafio enfrentado por autoridades, plataformas digitais e pelo sistema jurídico na identificação e contenção de conteúdos falsificados, sobretudo quando são utilizados para fins de desinformação, fraude ou violação de direitos fundamentais, dessa forma, esse processo ilustra-se pela Figura 5:

**Figura 5 - Processo de criação com os dados disponíveis.**



Fonte: Journal of Imaging (2022).

Embora a tecnologia permita criar diversos tipos de arquivos, seu uso mais comum é na produção de vídeos, especialmente *deepfakes* direcionados a celebridades, políticos e CEOs, devido à abundância de fotos e vídeos disponíveis na internet, que facilitam o treinamento das GANs (Westerlund, 2019). Contudo, isso não impede que ela seja usada para prejudicar qualquer pessoa, já que as redes sociais fornecem um vasto acervo de imagens e vídeos. Assim, não é difícil achar relatos de exemplos nos quais a tecnologia já foi usada para esses fins, mudando seu *modus operandi* para cada segmento fraudulento, como vingança, desinformação, crimes, fabricação de provas falsificadas, chantagens e *bullying*, por exemplo, sendo criados cada vez métodos mais criativos para conseguir enganar as vítimas, em um dos casos, a tecnologia foi usada para falsificar a voz de um dos CEOs de uma empresa britânica de energia para solicitar uma transferência bancária de €220.000, que foi prontamente realizada, pois a voz era tão convincente que enganou o outro CEO da empresa (Damiani, 2019). Entretanto, no tocante a golpes, um em cada cinco usa a imagens de famosos, além disso, das celebridades mais usadas, metade é jornalista (Tagiaroli, 2025), conforme já exposto, isso ocorre principalmente devido ao fato de que há um grande acervo de imagens dessas pessoas facilmente disponíveis para o treinamento da inteligência artificial, porém, isso também ocorre pois a imagem desses indivíduos passa credibilidade, possuindo um histórico de transmitir informações apuradas, logo, cria-se a ideia de que o que essa pessoa traz é a verdade, na maior parte das vezes são usados apenas trechos de participações em jornais ou podcasts tiradas de contexto, fazendo parecer que a pessoa está se referindo ao produto anunciado, quando na verdade era sobre outro assunto, esse golpe é mais fácil realizar pois utiliza apenas cortes de imagem, suficientes para enganar quem não verifica a informação, porém também são utilizados *deepfakes* para esse propósito, parecendo ainda mais verídico.

Além disso, é frequente o uso de inteligência artificial para burlar sistemas de reconhecimento facial em instituições financeiras, por meio da simulação do rosto da vítima, fraudando a prova de vida, realizando transações irregulares, contratando empréstimos que nunca serão pagos, acessando o cartão de crédito, ou seja, possuindo acesso livre às contas bancárias da vítima, esse golpe utiliza de técnicas que fazem o aplicativo entender que está diante de um ser humano de verdade, pois consegue fingir que está seguindo as orientações na tela do aplicativo, além de imitar o *smartphone* da vítima, incluindo seu sistema operacional, dados de rede e até mapa de satélites do GPS. Sendo assim, esse tipo de fraude demonstra o nível de sofisticação alcançado pelas tecnologias de manipulação, que, ao se combinarem com recursos avançados de engenharia social e falsificação digital, ultrapassam as barreiras tradicionais de segurança digital. A exploração de vulnerabilidades em sistemas de autenticação biométrica, que deveriam representar maior proteção para os usuários, revela o potencial destrutivo do uso indevido da inteligência artificial. Além dos prejuízos econômicos diretos, essas práticas ferem profundamente o direito à privacidade e à segurança dos dados pessoais, exigindo não apenas avanços na tecnologia de defesa, mas também uma resposta jurídica robusta, com normas específicas que regulem o uso dessas tecnologias em contextos sensíveis e responsabilizem os agentes por trás dessas condutas ilícitas.

Todavia, embora seja frequente o uso dessa tecnologia para golpes, cerca de 96% dos *deepfakes* disponíveis na internet são de conteúdo pornográfico não consensual (Sensity, 2019), atingindo tanto atrizes famosas, como pessoas comuns alvos da pornografia de vingança, como por exemplo quando uma mãe utilizou dos *deepfakes* para favorecer sua filha que vinha tendo problemas com outras garotas da sua equipe, assim, foram divulgados conteúdos falsos onde as garotas menores de idade apareciam nuas, com bebidas alcoólicas e fumando, com o fito de constrangê-las e saírem da equipe (Gogoni, 2021). Essa prática é extremamente perigosa, pois, além de violar direitos fundamentais como a intimidade e a honra, pode desencadear graves consequências psicológicas, como depressão e ansiedade, ou até mesmo levar ao suicídio. O impacto vai além do financeiro, atingindo a dignidade humana de forma irreversível. Diante disso, fica evidente que os *deepfakes* não são apenas uma ferramenta de fraude, mas uma arma de violação em massa, amplificando a desinformação e a crueldade na era digital. Portanto, é urgente a implementação de leis mais rígidas e mecanismos de fiscalização eficazes para coibir esse tipo de abuso, protegendo as vítimas e preservando a ética no uso da inteligência artificial.

No mais, também são relevantes os golpes para fins de relacionamento, nos quais a vítima acredita estar conversando e se relacionando via *internet* com uma pessoa real, entretanto essa pessoa na verdade não existe, sendo criada por meio de inteligência artificial, esse golpe é conhecido por “*Romance Scam*”, que de acordo com o FBI (s.d.):

Em golpes de romance, um criminoso usa uma identidade online falsa para ganhar o afeto e a confiança de uma vítima. O golpista então usa a ilusão de um relacionamento romântico ou próximo para manipular e/ou roubar da vítima<sup>2</sup> (tradução nossa).

Esse golpe não surgiu com a inteligência artificial, tendo origem na facilidade de comunicação proporcionada pelas redes sociais, todavia, com o avanço dessas tecnologias, tornou-se possível criar uma pessoa fictícia que não existe, mas que aparenta ser real, o que facilita enormemente a manipulação das vítimas. Assim, o golpista trabalha para estabelecer um relacionamento rapidamente, utilizando promessas de encontros presenciais que conquistam a confiança cega da vítima, embora jamais se concretizem. A cada etapa, o criminoso inventa imprevistos e solicita dinheiro, podendo até mesmo produzir vídeos personalizados com mensagens dirigidas à vítima por meio de inteligência artificial, reforçando a ilusão de autenticidade, apesar de a pessoa retratada ser completamente fabricada. Esse cenário evidencia o poder alarmante da Inteligência Artificial em simular identidades convincentes a partir de um ponto zero, esse ponto inicial é apenas uma simples frase descrevendo as principais características da imagem que virá a ser criada, essas descrições incluem estilo da imagem (realismo, desenho), consistência de cores, filtros, qualidade da câmera, elementos da cena, basicamente, você deve ser o mais detalhado possível, porém de forma objetiva, para que o algoritmo entenda o

---

<sup>2</sup>No original: “In romance scams, a criminal uses a fake online identity to gain a victim's affection and trust. The scammer then uses the illusion of a romantic or close relationship to manipulate and/or steal from the victim”.

que você está tentando criar, assim é possível criar imagens capazes de enganar até os mais atentos, o que amplifica os danos emocionais e financeiros, exigindo maior vigilância digital e o desenvolvimento de ferramentas para identificar tais fraudes, a fim de proteger a sociedade dessa nova fronteira de crimes virtuais, dessa forma, é muito difícil distinguir uma pessoa real de uma falsa, conforme ilustrado pela Figura 6, que traz uma pessoa criada por inteligência artificial:

**Figura 6 - Mulher com longos cabelos prateados e olhos azuis.**



Fonte: Midjourney (2025).

Por fim, surge o dilema entre a liberdade de expressão e a violação de direitos decorrente do uso desenfreado de inteligência artificial na produção de *deepfakes*, refletindo uma tensão crescente na era digital, na qual tem-se a ideia de que tudo pode ser falado ou postado sem consequências. A liberdade de expressão, um pilar fundamental da democracia, garante o direito de compartilhar ideias e informações sem censura prévia. No entanto, o avanço de tecnologias como *deepfakes* e a disseminação de *fake news* têm desafiado esse princípio, ao permitir a criação e propagação de conteúdos falsos que, muitas vezes, ultrapassam os limites éticos e legais, violando direitos essenciais como a privacidade, a honra e a dignidade. Entretanto, destaca-se o caso em que o empresário Mark Zuckerberg foi alvo dos *deepfakes*, quando supostamente fez uma declaração polêmica acerca do roubo de dados, porém o *Facebook* surpreendeu ao informar que o vídeo manipulado não será removido do *Instagram*, com base na justificativa de que o material compartilhado não fere as diretrizes da companhia, gerando, assim, debates sobre os limites na moderação desses novos tipos de conteúdos. Assim cria-se um conflito sobre até que ponto a liberdade de expressão pode justificar a livre permanência desses conteúdos enganosos, circulando sem uma regulamentação que evite abusos mas mantenha o direito à liberdade de expressão.

#### **4 LEGISLAÇÃO APLICÁVEL E SEUS DESAFIOS PERANTE O AVANÇO TECNOLÓGICO**

Inicialmente, antes de adentrar no mérito da legislação aplicável é importante analisar as possibilidades de aplicação de métodos de combate às *deepfakes*, assim, com o uso da tecnologia é possível buscar impedir a disseminação de conteúdos criados por inteligência artificial, nesse sentido, em 2019 empresas como *Amazon*, *Facebook* e *Microsoft* se reuniram para estimular pesquisadores a criar tecnologias que ajudem a detectar esses *deepfakes* (Schorepfer, 2019). Também destaca-se a novidade criada pela rede social X (*Twitter*) com o recurso das notas da comunidade, onde as publicações são marcadas como potencialmente falsas, criando uma dúvida em quem lê. Além disso, outras empresas estão reconhecendo os perigos que surgiram com o desenvolvimento dos *deepfakes* e criando suas próprias tecnologias para identificá-los, como por exemplo a empresa *Sensity* que desenvolveu uma plataforma na qual diversos formatos de arquivo podem ser anexados possibilitando identificar vídeos e rostos gerados por GANs, ou seja, você coloca algo que acredita ser um *deepfake* e a plataforma analisa padrões e diz se é ou não um conteúdo desse tipo. Além disso, diversas outras possíveis soluções estão sendo testadas no combate ao uso desmedido de *deepfakes*, cientistas da Universidade de Buffalo, Estados Unidos, criaram um algoritmo capaz de analisar o reflexo presente nos olhos dos vídeos, pois, naturalmente, a luz emitida pelo cenário gera reflexos na córnea que não são perfeitamente copiados nos vídeos manipulados (Almenara, 2021). Essas investidas mostram que embora a tecnologia avance rapidamente, soluções sofisticadas surgem para frear esse avanço, basta que sejam devidamente aplicadas antes de atingir de fato algum indivíduo.

No mais, no que tange aos aspectos legais, inicialmente, é necessário destacar que há uma lacuna legal para reger especificamente esse tema, todavia, é importante destacar que na maioria das vezes essas situações ocorrem no ambiente digital, portanto, encaixa-se no conceito de cibercrime, sendo considerado como todo ato em que o computador serve de meio para atingir um objetivo criminoso ou em ato no qual o computador é objeto do crime (Marques; Martins, 2006). Em virtude disso, criou-se a Convenção de Budapeste em 23 de novembro de 2001 e a Decisão-Quadro 2005/222/JAI do Conselho 24 de fevereiro de 2005 são os dois principais instrumentos internacionais ao combate desse tipo de crimes, surgindo como resposta aos crimes cometidos no ambiente digital, servindo como base para o desenvolvimento de legislação para os países. Sendo assim, os crimes cibernéticos incluem todos os crimes praticados no ambiente da *internet*, como pornografia infantil, violação à segurança de redes, violações de direitos autorais e conexos, assim como fraudes (Ferrari; Senna, 2021).

Contudo, a velocidade com que as novas formas de manipulação de imagens e conteúdos surgem, como *deepfakes* e golpes, revela uma defasagem entre a evolução tecnológica e a resposta legal. No Brasil, a Lei Geral de Proteção de Dados (Lei nº 13.709/2018) foi criada para regulamentar o tratamento de dados pessoais, garantindo a proteção da privacidade, da intimidade, da honra e da imagem dos cidadãos. Além de estabelecer diretrizes para empresas e órgãos públicos, a legislação busca equilibrar o desenvolvimento tecnológico e econômico com a preservação dos direitos fundamentais dos titulares de dados. Dessa forma, a lei não apenas promove a segurança jurídica para organizações, mas também fortalece a autonomia individual frente ao uso de informações pessoais na era digital. Logo, embora

os marcos legais existentes ofereçam um certo embasamento, é necessário que os ordenamentos jurídicos se adaptem para tipificar condutas de forma mais específica, garantindo que as vítimas tenham proteção efetiva. Nesse viés, de acordo com Siqueira (2019):

A Legislação Brasileira não criminaliza especificamente o “*Deep Fake*”. Mas os intérpretes têm buscado amparo em tipos penais abertos descritos na Lei Federal n.º 12.735/2012 (Lei Azeredo), Lei Federal n.º 12.737/2012 popularmente conhecida como Lei Carolina Dieckmann, Lei Federal n.º 12.965/2014 (Marco Civil da Internet); Lei Federal n.º 13.718/2018 oriunda do Projeto de Lei n.º 5.555/2013, Lei Federal n.º 13.709/2018 - Lei Geral de Proteção de Dados Pessoais, Lei Federal n.º 13.853/2019. Além dos tipos penais descritos na Lei de Crimes Financeiro (Lei Federal n.º 7.492/86), Lei de Falências (Lei Federal n.º 11.101/2005), Código Eleitoral (Lei Federal n.º 4737/65) e principalmente nos crimes contra a honra (artigos 138/145 do Código Penal) e dignidade sexual (artigos 213/235 ‘c’ do Código Penal).

Ademais, nota-se que as imagens manipuladas, a depender de seu uso, podem violar diversos direitos, configurando crimes contra a honra, previstos no Código Penal brasileiro, posto que esses institutos visam proteger a dignidade subjetiva e objetiva da pessoa humana. Dessa forma, é possível analisar esses crimes, que estão elencados do artigo 138 ao 140 do Código Penal, ligando a manipulação de imagens às práticas desses delitos. Assim, a necessidade de resguardar o bem jurídico tutelado permanece irretocável, a proteção da honra não constitui interesse exclusivo do indivíduo, mas da própria coletividade, a fim de preservar a honra, a incolumidade moral e a intimidade, além de outros bens jurídicos indispensáveis para a harmonia social (Bitencourt, 2020). Nesse sentido, no que diz respeito à calúnia, nota-se que o crime se consuma ao imputar falsamente a outrem fato definido como crime, sendo assim, ao utilizar de imagens manipuladas para criar uma situação na qual comprova-se erroneamente que uma pessoa está cometendo um crime, induzindo e facilitando terceiros a acreditarem nisso, pode-se configurar o crime de calúnia, por exemplo, é possível facilmente criar um vídeo de uma celebridade cometendo um furto. Portanto, para exemplificar, no julgamento da Apelação nº 20130111822028 a 1ª Turma Criminal do Tribunal de Justiça do Distrito Federal reconheceu causa de aumento de pena em razão de o conteúdo calunioso ter sido disponibilizado em jornal de grande circulação diária, com site na internet, fato este que seguramente ocorre em casos imagens manipuladas.

Quanto ao crime de difamação, configura-se quando há a imputação de fato ofensivo à reputação de alguém, ainda quem não seja um fato criminoso, dessa forma, o uso de *deepfakes* que exponham a vítima a situações que desabonem sua conduta facilmente se enquadra nessa tipificação, assim, uma manipulação na qual a vítima está fazendo uso de drogas ilícitas por exemplo pode ser classificado como difamação. Nesse viés, o Superior Tribunal de Justiça julgou conflito de competência nº 173.458, caso de crime de difamação que envolvia uma empresa com sede na cidade de Barra Funda, em São Paulo, que imputou a determinado indivíduo, que reside em Blumenau, em Santa Catarina, a prática do crime de difamação por meio de publicação de

vídeos em rede social, evidenciando a aplicação legal no ambiente digital, que pode ter os danos ampliados por meio de *deepfakes*. Por fim, encerrando os crimes contra a honra tem-se o delito de injúria, no qual ofende-se a dignidade ou o decoro de alguém, assim, mesmo sem atribuir um fato específico publicamente à vítima, é possível que uma imagem manipulada seja usada para ridicularizar ou humilhar o indivíduo, podendo, assim, criar imagens que ofendem traços físicos da vítima. Aqui, pode-se observar o Conflito de Competência nº 184.269, julgado pelo Superior Tribunal de Justiça, caso no qual o crime de injúria foi praticado por meio das mensagens privadas do Instagram, ocasião em que o autor do delito enviou mensagem de áudio com conteúdo ofensivo à vítima, assim, o crime se consumou no momento e local onde a vítima tomou conhecimento do conteúdo ofensivo, podendo ocorrer o mesmo se no lugar de um áudio houvesse uma imagem manipulada.

No mais, é necessário comentar acerca das majorantes presentes nas disposições comuns dos crimes contra a honra, assim, o artigo 141 estabelece, entre outras, o aumento de pena quando o delito for cometido por meio que facilite a divulgação, destacando a preocupação do legislador, já em 1940, em trazer uma pena maior se o crime foi cometido perante um número considerável de pessoas. Nessa perspectiva, o bem jurídico protegido pelos dispositivos aqui em evidência, é conceituado como o prestígio ou a reputação que o sujeito goza perante a sociedade, faz todo sentido que a legislação estipule majorantes para situações em que forem constatadas maiores ameaças ao bem jurídico (Albuquerque, 2023). Nesse contexto, encaixa-se perfeitamente o ambiente virtual, tendo em vista que uma vez que um conteúdo é publicado na *internet*, o mesmo sempre estará presente na *internet*, posto que a publicação pode rapidamente se espalhar, alcançando pessoas que nunca seriam alcançadas nos meios tradicionais, gerando enorme prejuízos e, muito provavelmente, sempre existirá um cópia do conteúdo em algum dispositivo, podendo vir à tona novamente, mesmo após remoção. Ainda, seguindo esse aspecto, é relevante mencionar que entre os meios que facilitam a divulgação de conteúdos ofensivos, havia um respaldo legal, pela Lei nº 5.250, de 9 de fevereiro de 1967, conhecida como a Lei da Imprensa, todavia, isso foi alterado depois do julgamento da ADPF 130, conforme aduz Capez:

Observe-se que os meios de divulgação por excelência são o rádio, a televisão, os impressos em geral, porém os crimes contra a honra praticados através deles constituíam objeto da Lei de Imprensa (Lei n. 5.250/67). No entanto, a partir da ADPF 130, o aludido Diploma Legal deixou de integrar o ordenamento jurídico pátrio, por ser considerado incompatível com a Constituição Federal, de forma que o inciso III, a partir de agora, poderá também abarcar tais meios de divulgação (Capez, 2019).

Além disso, a depender da interpretação e do caso concreto, outros crimes podem ser imputados, como por exemplo o crime de estelionato, tipificado no artigo 171 do Código Penal, consistindo na obtenção de vantagem ilícita em prejuízo alheio, assim, trata-se de crime contra o patrimônio onde a legislação penal visa proteger a inviolabilidade patrimonial orientada pela prática de atos que visam enganar a vítima e beneficiar o agente (Cunha, 2019), logo, a utilização de *deepfakes* pode facilitar a simulação da

identidade de uma pessoa, com o objetivo de enganar os familiares a fim de obter valores em nome da vítima, dessa forma, induzindo a erro por meio ardiloso. Ainda, com relação aos casos mais recorrentes, como a pornografia não consensual, os *deepfakes* de nudez podem ser enquadrados no artigo 218-C do Código Penal, que tipifica condutas relacionadas à divulgação não consensual de conteúdo íntimo, incluindo *deepfakes* pornográficos e pornografia de vingança, sendo relevante devido ao fato de que imagens manipuladas podem causar graves danos à imagem, honra e privacidade da vítima, possuindo, inclusive, causa de aumento de pena se for realizado com o fim de vingança ou humilhação. Nesse sentido, relaciona-se com os *deepfakes* ao abranger vídeos ou imagens manipuladas que simulam pornografia sem autorização, enquadrando-se como "cena de sexo ou pornografia" não consensual, também se aplica a casos de vingança, por exemplo quando a divulgação ocorre por ex-companheiros. Além disso, em 2021 foi acrescentado ao Código Penal o artigo 147-A, prevendo o crime de perseguição, no qual a vítima é perseguida reiteradamente e por qualquer meio, tendo sua integridade física ou psicológica ameaçada, assim, *stalking* é a ação de quem molesta um sujeito (vítima) por meio de atos persecutórios e/ou intimidadores de forma obsessivamente repetitiva, deixando a vítima em estado de alerta e relevante preocupação, quando não em profunda angústia (Mazzola, 2008), então, pode ser considerado *stalking*, quando houver uso repetido de *deepfakes* para assediar ou ameaçar alguém. Portanto, embora a mera criação de um *deepfake*, por si só, não constitua crime, sua utilização indevida pode ser enquadrada em diversas tipificações penais, conforme o contexto e os danos causados.

Ademais, além das consequências na esfera penal, a utilização indevida de *deepfakes* pode acarretar responsabilidades na esfera civil, conforme disposto no artigo 927 do Código Civil Brasileiro. Este dispositivo estabelece que aquele que, por ato ilícito, causar dano a outrem, fica obrigado a repará-lo, abrangendo tanto os danos materiais quanto os morais. Então, o dano é elemento fundamental para o ato ilícito civil, trazendo como consequência a obrigação de indenizar, de reparar o dano (Tartuce, 2023). No contexto dos *deepfakes*, essa responsabilidade pode surgir quando a manipulação de imagens é utilizada para ofender a honra, privacidade, reputação ou até mesmo a vida profissional e pessoal. Já os danos materiais podem incluir perdas financeiras diretas, como a diminuição de renda decorrente de prejuízo à imagem, enquanto os danos morais se referem ao sofrimento psicológico ou ao constrangimento sofrido pela vítima. Assim, o autor do *deepfake* pode ser condenado judicialmente a retratação pública ou indenização à vítima, sendo o valor da reparação determinado com base na extensão do dano causado, na gravidade da conduta e nas circunstâncias do caso concreto, funcionando como um mecanismo de compensação em uma situação de vulnerabilidade ante essas práticas. Ocorre que o que é disseminado sem nenhum controle preventivo de conteúdo, ocasiona sérios problemas no que tange a responsabilidade civil (Braga *et al*, 2011).

Adicionalmente, o uso de imagens manipuladas com o objetivo de disseminar *fake news* e desinformação também representa uma grave violação aos direitos da personalidade, protegidos no ordenamento jurídico brasileiro, estando normatizado no Código Civil, em seus artigos 11 a 21. Esses direitos são inalienáveis, indisponíveis e irrenunciáveis, tendo como fundamento a

dignidade da pessoa humana, sendo assim, possuem um objetivo relacionado à vida, liberdade, honra, imagem, vida privada, intimidade, entre outros. Nesse aspecto, existem três importantes vertentes dentro dos direitos da personalidade, a saber: integridade física, integridade moral e a integridade intelectual (Domingues, 2016). Com o passar dos anos, a preocupação passou a ser a tutela dos direitos na esfera digital, tendo em vista que o próprio usuário utiliza sua voz e imagem para alcançar um grande número de pessoas, ficando vulnerável. No Brasil, a Lei Geral de Proteção de Dados (Lei. 13.709/18) foi a responsável por trazer o respeito à privacidade, à inviolabilidade da intimidade, da honra e da imagem, surgindo com o objetivo de promover o desenvolvimento econômico e tecnológico das empresas, ao passo que protege direitos e liberdades fundamentais dos titulares. Todavia, a problemática surge devido à extrema exposição presente nas redes sociais, onde as pessoas compartilhando tudo que fazem no dia, ou seja, fornecem muitas informações e imagens pessoais, facilitando para quem está com más intenções, nesse sentido, conforme Ávila:

A hiperexposição está ligada ao uso compulsivo das mídias sociais, expondo a vida cotidiana, sem levar em conta os riscos que podem causar, bem como a segurança da rede. Inicia-se o uso por um hábito de informar e acaba por um vício em se expor. Entre as tecnologias digitais que impulsionam os hábitos descritos anteriormente, o smartphone é a principal, vez que converge a informação em tempo real, a exposição, a conexão, o uso da tecnologia e o *multitasking* (Ávila, 2022).

Entre os principais direitos de personalidade relacionados ao tema em questão, destaca-se o direito à imagem, compreendendo não apenas o direito de controle sobre a reprodução da própria aparência física podendo dispor da mesma ou a reproduzir, seja de caráter comercial ou não, mas também a proteção contra usos indevidos, manipulações ou exposições que possam comprometer a honra, a moral ou a dignidade da pessoa, sendo um direito personalíssimo do indivíduo, logo, esse direito protege interesses existenciais do ser humano. Assim, quando uma imagem é manipulada para simular comportamentos ou contextos falsos, como em montagens pornográficas, vídeos forjados ou situações humilhantes, configura-se uma clara violação desse direito. A solução legal para resolver esse problema não se mostra eficaz, tendo em vista que não consegue suprir os danos frente à crescente violação dos direitos da personalidade no âmbito virtual. Desse modo, a reparação dada às vítimas não é capaz de amenizar ou superior o abuso cometido, independente de valor, pois não é possível desfazer a exposição sofrida (Oliveira; Murta, 2017). No aspecto da violação do direito de imagem, a Súmula 403 do Superior Tribunal de Justiça traz que é independente de prova ou prejuízo a indenização pela publicação não autorizada da imagem de pessoas com fins econômicos e comerciais. Além disso, o direito à honra e o direito à intimidade também são atingidos nessas situações. A manipulação de imagens pode criar um cenário falso que prejudica profundamente a reputação da vítima, expondo-a a julgamentos sociais, profissionais e até familiares baseados em conteúdos que não procedem. Ainda, nos casos de pornografia de vingança, por exemplo, ou de uso de *deepfakes* em contextos difamatórios, o abalo psicológico e emocional pode ser tão profundo quanto irreparável.

Dentre as soluções paliativas, o ordenamento jurídico prevê medidas de proteção e reparação para os danos causados nesses casos. No âmbito civil, é possível requerer a exclusão do conteúdo ofensivo e a indenização por danos morais. Ainda, existe a possibilidade na qual o juiz pode conceder medidas liminares para remoção imediata do conteúdo da *internet*, com base na urgência da proteção à dignidade da vítima ou até mesmo exclusão da conta violadora dos direitos. No mais, o Marco Civil da *Internet* (Lei nº 12.965/2014) também oferece instrumentos importantes para responsabilização dos provedores, como a responsabilidade civil de provedores caso não tomem as providências para, no âmbito do seu serviço e no prazo assinalado, não torne indisponível o conteúdo, mesmo sendo um dano advindo de um conteúdo criado por terceiros, estimulando uma ação positiva dos grandes provedores, haja vista que há um problema fica decorrente da dificuldade a respeito de quem punir em um caso desse tipo, afinal o conteúdo danoso foi criado pela plataforma, porém seu fim foi diversificando pelo sujeito que usa de má-fé, além de possibilitar aplicações e a rastreabilidade dos responsáveis, embora ainda haja lacunas práticas em sua aplicação, especialmente diante da velocidade de disseminação e do uso de perfis anônimos ou falsos.

Ademais, a inexistência de um marco legal específico sobre conteúdos gerados por inteligência artificial, como os *deepfakes*, torna ainda mais difícil a efetiva proteção dos direitos da personalidade. Entretanto, mesmo existindo a possibilidade de aplicar a legislação existente por analogia, muitas vezes os instrumentos atuais mostram-se insuficientes frente à complexidade tecnológica envolvida, não conseguindo se adequar de maneira satisfatória ao caso em questão. Dessa forma, a manipulação de imagens digitais não representa apenas um problema ético ou tecnológico, mas uma verdadeira guerra contra a dignidade humana, onde um lado avança rapidamente com o uso da tecnologia e o outro segue muitos passos atrás devido à morosidade legislativa. Nesse contexto, a proteção dos direitos da personalidade, exige não só a aplicação rigorosa das normas já existentes, bem como suas majorantes, como também o desenvolvimento de políticas públicas e legislações atualizadas, que contemplem os riscos e os danos provocados pelo uso abusivo da inteligência artificial, além do artifício que possivelmente trará mais resultados, que será uma ferramenta oposta que consiga neutralizar a validade dos *deepfakes*.

No mais, embora exista uma Lacuna legal no ordenamento jurídico brasileiro, tendo em vista que o Código Penal brasileiro é datado de 1940, mesmo que tenha sofrido alterações, não é rápido o suficiente para acompanhar os avanços da sociedade e da tecnologia, que permitiram que novas situações que antes não eram imaginadas começassem a ocorrer. Eventualmente, outros países criaram leis referentes ao uso de *deepfakes*, nos Estados Unidos, alguns estados, como Califórnia e Virgínia, atualizaram sua legislação contra a pornografia de vingança, incluindo a proibição da divulgação de materiais com nudez adulterada por meio de *deepfakes* (Parreira, 2021). No âmbito político, a Califórnia adotou medidas ainda mais rigorosas, vetando a veiculação de *deepfakes* difamatórios contra candidatos nos 60 dias que antecederam as eleições (Theobald, 2019). Por outro lado, a China, visando proteger a segurança nacional e evitar ameaças políticas, tipificou como crime a publicação e o compartilhamento de *deepfakes* sem a

devida identificação, sujeitando os infratores a penalidades legais (Teixeira, 2019).

Além disso, na Alemanha, foi criada a *Netzdurchsetzungsgesetz* (Lei de Aplicação de Redes) que provocou a denúncia de milhões de conteúdos com teor considerado criminoso nas redes sociais, posto que essa lei objetiva combater o discurso de ódio e o conteúdo ilegal nas redes sociais, obrigando as plataformas online a removerem conteúdo manifestamente ilegal em até 24 horas além de exigir que sejam reportados relatórios de transparência à Polícia Federal Alemã (BKA) sobre remoções de conteúdo, para investigar os itens excluídos, ainda, as empresas podem receber multas de até 50 milhões de euros caso desrespeitem a lei (BBC, 2018). Dessa forma, é considerada a primeira regulação mais dura sobre redes sociais em um país democrático, sendo referência mundial, todavia, sofreu críticas na sua apreciação posto que os opositores enxergavam graves ameaças à liberdade de expressão e à privacidade.

Por fim, no cenário político, destaca-se a Resolução nº 23.732 do Tribunal Superior Eleitoral que dispõe sobre a propaganda eleitoral, trazendo em seu art. 9º-C a proibição da utilização de conteúdo fabricado ou manipulado para difundir fatos notoriamente falsos ou descontextualizados, ademais, o referido dispositivo legal ganha mais destaque por ser ainda mais específico em seu primeiro parágrafo, pois dá ênfase aos conteúdos gerados por manipulação, incluindo a utilização de ferramentas como a inteligência artificial por meio de *deepfakes*:

§ 1º É proibido o uso, para prejudicar ou para favorecer candidatura, de conteúdo sintético em formato de áudio, vídeo ou combinação de ambos, que tenha sido gerado ou manipulado digitalmente, ainda que mediante autorização, para criar, substituir ou alterar imagem ou voz de pessoa viva, falecida ou fictícia (*deep fake*) (Brasil, 2024).

Entretanto, para evitar que as pessoas tenham sua percepção manipulada sem perceberem, nota-se que na referida Resolução, em seu art. 9º-B, traz-se a possibilidade de utilizar conteúdo sintético ou fabricado criado por inteligência artificial, dessa forma, é possível propagar conteúdos legítimos sem a imposição de multa ou a imediata remoção do conteúdo, desde que não cause prejuízos e seja informado explicitamente que se trata de um conteúdo manipulado, além de apontar qual foi a tecnologia utilizada, assim, trazendo as formas de apresentar essas informações, quais sejam:

I – no início das peças ou da comunicação feitas por áudio;  
II – por rótulo (marca d'água) e na audiodescrição, nas peças que consistam em imagens estáticas;  
III – na forma dos incisos I e II desse parágrafo, nas peças ou comunicações feitas por vídeo ou áudio e vídeo;  
IV – em cada página ou face de material impresso em que utilizado o conteúdo produzido por inteligência artificial. (Brasil, 2024).

## 5 METODOLOGIA

O trabalho, a fim de atender ao objetivo proposto, foi desenvolvido por meio de uma abordagem qualitativa, com foco na análise bibliográfica e

documental, que consiste na leitura, análise e interpretação de diferentes fontes disponíveis sobre o tema escolhido, qualquer contribuição científica, seja ela impressa ou eletrônica pode se tornar uma fonte de consulta (Martins *et al*, 2022). Ainda, foi adotado o método hipotético-dedutivo, partindo de um problema no conhecimento, formulando hipóteses para explicá-lo, dessa forma, parte-se da hipótese de que a manipulação de imagens por inteligência artificial contribui para a disseminação de *fake news* e viola direitos fundamentais, como a honra e a imagem. A partir disso, analisa-se o ordenamento jurídico, casos concretos e elementos doutrinários para verificar a validade dessa hipótese e apontar possíveis lacunas legais. Essas escolhas se justificam pela natureza do tema proposto, que envolve fenômenos tecnológicos recentes, buscando compreender os impactos jurídicos e sociais do uso de imagens manipuladas, especialmente com o advento da inteligência artificial generativa, no contexto da desinformação e das *fake news*. Assim, é possível interpretar fenômenos complexos, como a manipulação digital de imagens e suas implicações jurídicas, levando em consideração visões doutrinárias, legislações, jurisprudências e estudos interdisciplinares, que envolvem áreas como Ciências da Computação, Arte e Direito.

Como técnica principal, foi adotada a pesquisa bibliográfica, com a consulta de livros, artigos científicos, legislações nacionais e internacionais, doutrinas e publicações institucionais. Ainda, o trabalho analisa exemplos práticos e decisões judiciais, que ilustram o uso de redes sociais e imagens manipuladas como ferramenta de desinformação e violação da honra, possibilitando uma relação entre a atualização do ordenamento jurídico brasileiro e as decisões que estão sendo prolatadas com base nas normas disponíveis, possuindo caráter descritivo e interpretativo. Com base nesse aspecto, os objetivos da pesquisa são compreender os mecanismos técnicos e sociais que envolvem a manipulação de imagens por meio da inteligência artificial e sua evolução, bem como identificar o papel das imagens manipuladas no processo de desinformação e *fake news*, além de analisar os principais impactos jurídicos decorrentes do uso indevido dessas imagens e os dispositivos legais atualmente aplicáveis e suas limitações.

## 6 CONCLUSÃO

O trabalho abordou as implicações jurídicas do uso de imagens manipuladas, principalmente ao aplicar inteligência artificial na produção e disseminação de *fake news* e desinformação. Para tanto, foi explorada a evolução histórica da manipulação de imagens, desde técnicas manuais até o uso de ferramentas digitais mais modernas, destacando como essas tecnologias ampliaram as possibilidades de violação à honra, privacidade e dignidade. Assim, analisou-se o papel das imagens manipuladas na disseminação de desinformação, bem como os desafios enfrentados pelo ordenamento jurídico brasileiro e as limitações das normas atuais frente a esse fenômeno. A escolha do tema, motivada pela experiência pessoal do autor, se dá pela crescente influência das redes sociais e da inteligência artificial no cotidiano, em diversas imagens em que sequer desconfia-se da sua falsidade. Nesse aspecto nota-se que a legislação vigente não consegue enfrentar esses desafios de maneira satisfatória, gerando tanto danos individuais como

coletivos, atingindo desde à reputação até a interferência em processos eleitorais.

Dessa forma, ao questionar em que medida o sistema normativo brasileiro está preparado para lidar com condutas maliciosas envolvendo imagens manipuladas e quais mecanismos legais são eficazes notou-se que, embora existam instrumentos legais aplicáveis por analogia (como crimes contra a honra, estelionato e violação de direitos da personalidade), o sistema jurídico brasileiro apresenta limitações, já que as normas foram criadas a tempos, como o Código Penal de 1940, que mesmo com atualizações não consegue acompanhar os avanços da sociedade em alguns aspectos. Assim, a ausência de uma legislação específica para *deepfakes* e a morosidade legislativa dificultam a resposta ágil e efetiva aos danos causados, confirmando que o ordenamento não está plenamente preparado para esse desafio.

No mais, o trabalho detalhou a trajetória técnica da manipulação, desde a fotografia até os *deepfakes*, destacando o uso tecnologias modernas como as redes generativas adversariais (GANs) que estão disponíveis para a população em geral, nos aplicativos de criação de imagens. Também identificou como essas imagens amplificam a desinformação, criando narrativas falsas com alto potencial de dano. Por fim, examinou a aplicação de normas como o Código Penal, o Código Civil e a Lei Geral de Proteção de Dados, analisando sua aplicação e apontando suas lacunas frente à complexidade do tema.

Ainda, a pesquisa justifica-se pela urgência de discutir a adequação da legislação para acompanhar os avanços digitais além da necessidade de compreender esses impactos, posto que o uso indevido dessas ferramentas ameaça direitos fundamentais, comprometendo a veracidade das informações e desafiando a eficácia do sistema jurídico na proteção à sociedade. Assim, foi demonstrado que, embora a tecnologia avance em ritmo acelerado, a resposta jurídica e social precisa ser igualmente ágil. Este trabalho espera contribuir para o debate, destacando a urgência de medidas que equilibrem inovação tecnológica e garantias jurídicas, preservando a democracia, a liberdade de expressão e a dignidade humana, esperando contribuir para essa recente discussão, tendo em vista que, sem ações coordenadas, o abismo entre tecnologia e Direito aumentará cada vez mais rápido, continuando a apresentar riscos à democracia e aos direitos individuais.

## REFERÊNCIAS

ALBUQUERQUE, Beatriz Chaves Bittencourt de. **Crimes cibernéticos: uma análise sobre a tipificação dos crimes contra honra na internet**. 2023. 113f Fortaleza. Disponível em: <https://biblioteca.sophia.com.br/terminal/9575/acervo/detalhe/581973>. Acesso em: 13 abr. 2025.

ALMENARA, Igor. **Algoritmo é capaz de desmascarar deepfakes analisando o movimento dos olhos**. 2021. Disponível em: <https://canaltech.com.br/inteligencia-artificial/algoritmo-e-capaz-de-desmascara-r-deepfakes-analisando-o-movimento-dos-olhos-180574/>. Acesso em: 12 abr. 2025.

ALVES, Juliana Mendes. **Software de Edição de Imagem e O Conceito de Beleza na Sociedade Brasileira**. 2008. Disponível em:

<https://repositorio.uniceub.br/jspui/bitstream/123456789/2032/2/20475408.pdf>. Acesso em: 15 dez. 2024.

ALZAMORA, G.; MENDES, C.; RIBEIRO D. M. (Orgs.). **Sociedade da desinformação e infodemia**. 1ed. Belo Horizonte: Fafich/Selo PPGCOM/UFMG, 2022.

ÁVILA, G. N.; CORAZZA, T. A. M. **A Hiperexposição pessoal e seus reflexos nos direitos da personalidade: necessidade de uma tutela transversal do direito à privacidade, com enfoque no âmbito penal**. *JurisPoiesis*, v. 25, p. 144-177, 2022.

BBC. **Germany starts enforcing hate speech law**, 2018. Disponível em: <https://www.bbc.com/news/technology-42510868>. Acesso em: 12 abr. 2025.

BITENCOURT, Cezar Roberto. **Tratado de Direito penal – parte especial: dos crimes contra a pessoa**. 20. ed. rev. e ampl. São Paulo: Saraiva, 2020. v. 2.

BORGES, Luiz Antônio Dias. História da animação: uso da técnica e estética. **Revista Livre de Cinema**, v. 6, n.2, p. 63-82, mai-ago, 2019. Disponível em: <https://www.relici.org.br/index.php/relici/article/view/220/252>. Acesso em: 10 jun. 2025.

BRAGA, Diogo de Melo; BRAGA, Marcus de Melo; ROVER, Aires José. **Responsabilidade Civil das Redes Sociais no Direito Brasileiro**, 2011.

BRASIL. **Decreto-lei nº 2.848, de 7 de dezembro de 1940**. Código Penal. Rio de Janeiro: Presidência da República [1940]. Disponível em: <https://www2.camara.leg.br/legin/fed/declei/1940-1949/decreto-lei-2848-7-deze-mbro-1940-412868-publicacaooriginal-1-pe.html>. Acesso em: 22 mar. 2025.

BRASIL. **Lei nº 8.078, de 11 de setembro de 1990**. Dispõe sobre a proteção do consumidor e dá outras providências. Brasília: Presidência da República [1991]. Disponível em: [https://www.planalto.gov.br/ccivil\\_03/leis/l8078compilado.htm](https://www.planalto.gov.br/ccivil_03/leis/l8078compilado.htm). Acesso em: 22 mar. 2025.

BRASIL. Supremo Tribunal Federal. **Arguição de descumprimento de preceito fundamental 130 (ADPF)**. Lei de Imprensa. Disponível em: <https://redir.stf.jus.br/paginadorpub/paginador.jsp?docTP=AC&docID=605411>. Acesso: 10 abr. 2025.

BRASIL. Tribunal de Justiça do Distrito Federal. (1. Câmara Criminal). **Embargos de Declaração Criminal no(a) Apelação Criminal**. Relator: Desembargador Mário Machado, 09 de maio de 2019. Disponível em: <https://bitlybr.com/rsfzM>. Acesso em: 10 abr. 2025.

BRASIL. Superior Tribunal de Justiça. (3. Seção). **Conflito de competência nº 173458 SC 2020/0171971-7**. Crime contra a honra praticado pela internet. natureza formal. consumação no local da publicação do conteúdo ofensivo.

competência do juízo suscitante para o conhecimento e julgamento do feito. Relator: Ministro João Otávio de Noronha, 25 de novembro de 2020. Disponível em: <https://www.jusbrasil.com.br/jurisprudencia/stj/1206264534/inteiro-teor-1206264544>. Acesso em: 10 abr. 2025.

BRASIL. Superior Tribunal de Justiça. (3. Seção). **Conflito de competência nº 184269 PB 2021/0363685-3**. Relator: Min. Laurita Vaz, 09 de fevereiro de 2022. Disponível em: [https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ATC&sequencial=143940148&num\\_registro=202103636853&data=20220215&tipo=5&formato=PDF](https://processo.stj.jus.br/processo/revista/documento/mediado/?componente=ATC&sequencial=143940148&num_registro=202103636853&data=20220215&tipo=5&formato=PDF). Acesso em: 15 abr. 2025.

BRASIL. **Resolução nº 23.732, de 27 de fevereiro de 2024**. Altera a Res.-TSE nº 23.610, de 18 de dezembro de 2019, dispendo sobre a propaganda eleitoral. Brasília: Tribunal Superior Eleitoral [2024]. Disponível em: <https://www.tse.jus.br/legislacao/compilada/res/2024/resolucao-no-23-732-de-27-de-fevereiro-de-2024>. Acesso em: 24 abr. 2025.

BRASIL. **Guia Ilustrado Contra as Deepfakes**. Supremo Tribunal Federal; Data Privacy Brasil. Brasília: STF, Coordenadoria de Combate à Desinformação, 2024.

Cambridge Dictionary. (s.d.). **fake news**. In Cambridge Dictionary. Disponível em: <https://dictionary.cambridge.org/pt/dicionario/ingles/fake-news>. Acesso em: 20 fev. 2025.

CAPEZ, Fernando. **Curso de Direito penal: parte especial: arts. 121 a 212**. 19. ed. São Paulo: Saraiva Educação, 2019. v. 2

CAPÍTULO 54 - INTRODUÇÃO ÀS REDES ADVERSÁRIAS GENERATIVAS (GANS - Generative Adversarial Networks). **Deep Learning Book**. Disponível em: <https://www.deeplearningbook.com.br/introducao-as-redes-adversarias-generativas-gans-generative-adversarial-networks/>. Acesso em: 10 abr. 2025.

CUNHA, Rogério Sanches. **Manual de direito penal: parte especial (arts. 121 ao 361)**. 11.ed. rev., ampl. e atual. Salvador: JusPODIVM, 2019.

DAMIANI, Jesse. **A Voice Deepfake Was Used To Scam A CEO Out Of \$243,000**. Forbes, 2019. Disponível em: <https://www.forbes.com/sites/jessedamiani/2019/09/03/a-voice-deepfake-was-used-to-scam-a-ceo-out-of-243000/?sh=63e99eda2241>. Acesso em: 12 abr. 2025.

DOMINGOS, Fernanda Teixeira Souza. **As provas digitais nos delitos de pornografia infantil na internet**. In: SALGADO, Daniel de Resende; QUEIROZ, Ronaldo Pinheiro (orgs.). A prova no enfrentamento à macrocriminalidade. 3ª ed. rev., atual. e ampl. Salvador: Juspodvim, 2018.

DOMINGUES, Diego Sígoli. **Marco Civil da Internet e Consolidação dos Direitos de Personalidade**, 2016. Disponível em: <https://diegosigoli.jusbrasil.com.br/artigos/311886244/marco-civil-da-internet-e-consolidacao-dos-direitos-de-personalidade>. Acesso em: 12 abr. 2025.

**FANTÁSTICO**. Apresentado por Poliana Abritta e Maju Coutinho. Rede Globo de Televisão, 09 de outubro de 2022, 9 minutos. Noticiário Informativo. Disponível em: <https://g1.globo.com/fantastico/noticia/2022/10/09/familia-de-mulher-linchada-que-inspirou-a-novela-travessia-diz-como-e-seguir-a-vida-com-esse-trauma.ghtml>. Acesso em: 20 fev. 2025.

**FANTÁSTICO**. Apresentado por Felipe Santana. Nova Iorque: Rede Globo de Televisão, 25 de fevereiro de 2018, 20h, 17 minutos. Noticiário Informativo.

FBI. **Romance Scams**. [s.d.] Disponível em: <https://www.fbi.gov/how-we-can-help-you/scams-and-safety/common-frauds-and-scams/romance-scams>. Acesso em: 12 abr. 2025.

FERRARI, D; SENNA, F. **Convenção de Budapeste e crimes cibernéticos no Brasil**. Migalhas, 2020. Disponível em: <https://www.migalhas.com.br/depeso/335230/convencao-de-budapeste-e-crimes-ciberneticos-no-brasil>. Acesso em: 12 abr. 2025.

FERNÁNDEZ MARCIAL, V. F.; GOMES, L. I. E. Impacto de la inteligencia artificial en el comportamiento informacional: elementos para el debate. **Revista Bibliotecas**. Anales de Investigación, Cuba, v. 18, n. 3, 2022.

FLUSSER, V. **Comunicologia: reflexões sobre o futuro**. São Paulo: Martins Fontes, 2014.

FONTCUBERTA, Joan. **El Beso de Judas Fotografia y verdad**. Barcelona, Editorial Gustavo Gili S.A, 1998.

GOGONI, R. **Era inevitável: deepfake usado como ferramenta de bullying**. Meio Bit, 2021. Disponível em: <https://tecnoblog.net/meiobit/434802/mae-usa-deepfake-prejudica-concorrentes-da-filha/>. Acesso em: 12 abr. 2025.

GOODFELLOW, I., POUGET-ABADIE, J., MIRZA, M., XU, B., WARDE-FARLEY, D., OZAIR, S., COURVILLE, A., BENGIO, Y. **Generative adversarial networks**. Communications of the ACM63, 2020, P. 139–144. Disponível em: <https://doi.org/10.1145/3422622>. Acesso em: 10 abr. 2025.

GUARNERA, L.; et al. **The Face Deepfake Detection Challenge**. Journal of Imaging, 2022. p. 8-10. Disponível em: <https://www.mdpi.com/2313-433X/8/10/263>. Acesso em: 10 abr. 2025.

GUIMARÃES, Glayder Daywerth Pereira, SILVA, Michael César. Fake news à luz da responsabilidade civil digital: O surgimento de um novo dano social.

**Revista Jurídica FA7**. 2019; 16(2): p. 99-114. Disponível em: [https://scholar.google.com/scholar?hl=pt-BR&as\\_sdt=0%2C5&q=fake+News+urgimento&btnG=&lr=lang\\_pt#d=gs\\_qabs&t=1744644959687&u=%23p%3Dig6YxTr7UDwJ](https://scholar.google.com/scholar?hl=pt-BR&as_sdt=0%2C5&q=fake+News+urgimento&btnG=&lr=lang_pt#d=gs_qabs&t=1744644959687&u=%23p%3Dig6YxTr7UDwJ). Acesso em: 04 jan. 2025.

HARVEY, David. **Condição pós-moderna: Uma pesquisa sobre as origens da mudança cultural**. 22<sup>a</sup> ed. São Paulo: Loyola, 2012.

HOMEM, William Ludovico. **Apostila de Machine Learning**. PET Engenharia Mecânica. Universidade Federal do Espírito Santo, 2020. Disponível em: [https://petengenhariamecanica.ufes.br/sites/petengenhariamecanica.ufes.br/files/field/anexo/apostila\\_do\\_minicurso\\_de\\_machine\\_learning.pdf](https://petengenhariamecanica.ufes.br/sites/petengenhariamecanica.ufes.br/files/field/anexo/apostila_do_minicurso_de_machine_learning.pdf). Acesso em: 04 jan. 2025.

JOTADETC. **Midjourney**. [S.l.: s.n.], 2025. Disponível em: <https://www.midjourney.com/jobs/ea5c19b4-cf01-454d-a886-0344809af20f?index=0>. Acesso em: 04 abr. 2025.

KAUFMAN, D.: SANTAELLA, L. O papel dos algoritmos de Inteligência Artificial nas redes sociais. **Revista Famecos**, Porto Alegre, v. 27, [s.n.], p. 1-10, mai. 2020.

KBUCZERZA1. **Midjourney**. [S.l.: s.n.], 2025. Disponível em: <https://www.midjourney.com/jobs/4924d4a6-7793-49b3-9239-900b1d3f6c0d>. Acesso em: 04 abr. 2025.

LICHTENSTEIN, Jaqueline. **A pintura: textos essenciais**. São Paulo, Ed. 34, 2004.

MAGRANI, E. **Entre dados e robôs: ética e privacidade na era da hiperconectividade**. 2 ed. Porto Alegre: Arquipélago Editorial, 2019.

MANOVICH, L. **Inside Photoshop. Computational Culture**. v. 1, Novembro, 2011. Disponível em: <http://computationalculture.net/inside-photoshop/>. Acesso em: 25 nov. 2024.

MANSSOUR, Isabel Harb; COHEN, Marcelo. **Introdução à Computação Gráfica**. RITA, Volume XIII, Número 2, 2006. p. 01. Disponível em: <http://www.inf.pucrs.br/manssour/Publicacoes/TutorialSib2006.pdf>. Acesso em: 25 nov. 2024.

MARCHAL, N., KOLLANYI, B., NEUDERT, L. M., HOWARD, P. N. **Junk news during the EU parliamentary elections: Lessons from a seven-language study of Twitter and Facebook**. Oxford Internet Institute, 2019.

MARQUES, G., MARTINS, L. **Direito da Informática**. Coimbra, Almedina, 2006.

MARTINS, J. P., SILVA, M. P. d., SOUZA, M. Q. P. d. **O ensino híbrido e os desafios da sua implementação em escolas públicas**, 2022. Disponível em:

[https://revistas.icesp.br/index.php/FINOM\\_Humanidade\\_Tecnologia/article/view/2755/1643](https://revistas.icesp.br/index.php/FINOM_Humanidade_Tecnologia/article/view/2755/1643). Acesso em: 17 abr. 2025.

MAZZOLA, Marcello Adriano. **IL Novo Danni**. Padova. Dott: Antonio Milani. 2008.

MITCHELL, William J. **The Reconfigured Eye: Visual Truth in the Post-Photographic Era**. Londres, The MIT Press, 1992.

NETO, Geraldo de Medeiros Galvão. **Ferramenta para criação de pixel art através de modelo generativo**. Disponível em: [https://www.cin.ufpe.br/~tg/2023-1/propostas\\_CC/prop\\_gmgn.pdf](https://www.cin.ufpe.br/~tg/2023-1/propostas_CC/prop_gmgn.pdf) Acesso em: 04 jan. 2025.

OLIVEIRA, Bruna Thacianne de Araújo; MURTA, Diego Nobre. **O direito de imagem nas redes sociais**. Portal Jurídico Investidura, Florianópolis/SC, 2017. Disponível em: <http://investidura.com.br/biblioteca-juridica/artigos/direito-civil/335465-o-direito-d-a-imagem-nas-redes-sociais>. Acesso em: 22 abr. 2025.

PARREIRA, Rui. **80% dos legisladores não sabem o que são deepfakes e as vítimas acumulam-se. Sobretudo mulheres**. Sapo, 2021. Disponível em: <https://tek.sapo.pt/noticias/internet/artigos/80-dos-legisladores-nao-sabe-o-que-sao-deepfakes-e-as-vitimas-acumulam-se-sobretudo-mulheres>. Acesso em: 12 abr. 2025.

PATE, William. **Abraham Lincoln**. [S.l.: s.n.], 1865. 1 fotografia. Disponível em: <https://www.loc.gov/item/2003654314/>. Acesso em: 22 mar. 2025.

PEIRCE, C S. A fixação da crença. In: PEIRCE, C. **Ilustrações da lógica da ciência**. Tradução: Renato Rodrigues Kinouhci. Aparecida, Ideias & Letras, 2008. p. 35-58.

PRADO, Geraldo. **A cadeia de custódia da prova no processo penal**. Marcial Pons, 2019.

RECUERO, R. **Fraude nas urnas: estratégias discursivas de desinformação no Twitter nas eleições 2018**. Revista Brasileira de Linguística Aplicada, v. 20, p. 1, 2020. Disponível em: <https://www.scielo.br/j/rbla/a/vKngghPRMJxbypBVRLYN3YTB>. Acesso em: 25 nov. 2024.

RITCHIE, Alexander Hay; HICKS, Thomas; BRADY, Mathew B. **John C. Calhoun**. Nova York: A.H. Ritchie & Co. [ca. 1852]. 1 fotografia. Disponível em: <https://www.loc.gov/item/2003679757/>. Acesso em: 20 mar. 2025.

RUSSELL, S. J. **Inteligência artificial**. Rio de Janeiro: Elsevier, 2013.

SALGADO, T. B. P.; MATTOS, M. Â. Da informação à desinformação: conceitos e abordagens das teorias acerca da comunicação. In: ALZAMORA, G.;

MENDES, C.; RIBEIRO D. M. (Orgs.). **Sociedade da desinformação e infodemia** 1º ed. Belo Horizonte: Fafich/Selo PPGCOM/UFMG, 2022, v. 1, p. 39-62.

SCHROEPFER, Mike. **Creating a dataset and a challenge for deepfakes**. Facebook A.I., 2019. Disponível em: <https://ai.facebook.com/blog/deepfake-detection-challenge/>. Acesso em: 12 abr. 2025.

SENSITY. **Mapping the Deepfake Landscape**. Sensity, 2019. Disponível em: <https://sensity.ai/blog/deepfake-detection/mapping-the-deepfake-landscape/>. Acesso em: 12 abr. 2025.

SIQUEIRA, P. A. R. d. **O 'Deep Fake' e a Legislação Brasileira -utilização de instrumentos legais para a proteção à imagem**. Conteúdo Jurídico, 2019. Disponível em: <https://www.conteudojuridico.com.br/consulta/artigo/53256/o-deep-fake-e-a-legislao-brasileira-utilizacao-de-instrumentos-legais-para-a-proteo-imagem>. Acesso em: 12 abr. 2025.

Prefeitura de São Paulo. **SMART SAMPA**, 2025. Disponível em: <https://smartsampa.prefeitura.sp.gov.br/>. Acesso em: 12 abr. 2025.

SOUSA, Jorge Pedro. **A tolerância dos fotojornalistas portugueses à alteração digital de fotografias jornalísticas**. Covilhã - BOCC (Biblioteca On-line de Ciências da Comunicação) - Universidade da Beira Interior, 1998. Disponível em: [http://www.bocc.ubi.pt/pag/\\_texto.php?html2=sousa-pedro-jorge-Alteracao-Fotografias.html](http://www.bocc.ubi.pt/pag/_texto.php?html2=sousa-pedro-jorge-Alteracao-Fotografias.html). Acesso em: 25 nov. 2024.

SPIVEY, Nigel. **Understantig greek sculpture: ancient meanings, modern readings**. Londres, Thames and Hudson, 1996.

TAGIAROLI, Guilherme. **Das celebridades mais usadas em golpes, metade é jornalista**. Disponível em: <https://www.uol.com.br/tilt/noticias/redacao/2025/04/13/golpes-deepfake-jornalista-e-apresentador-de-tv.htm>. Acesso em: 14 abr. 2025.

TARTUCE, Flávio. **Manual de Direito Civil: Volume Único**. São Paulo: Método, 2023.

TATE. **The David King Collection**. [S.l.: s.n.], 1926. Disponível em: <https://www.tate.org.uk/search?q=dauid-king-collection>. Acesso em: 22 mar. 2025.

TEIXEIRA, Leonardo Ávila. **Nova lei na China criminaliza deepfakes**. GQ, 2019. Disponível em: <https://gq.globo.com/Prazeres/Poder/noticia/2019/12/nova-lei-na-china-criminaliza-deepfakes.html>. Acesso em: 12 abr. 2025.

THEOBALD, Bill. **Deepfakers beware: Do it in California or Texas and you'll be in deep trouble.** Fulcrum, 2019. Disponível em: <https://thefulcrum.us/deepfake-political-video>. Acesso em: 12 abr. 2025.

VILLANI, Cédric. **For a meaningful artificial intelligence towards a French and European strategy.** 2018. Disponível em: [https://www.aiforhumanity.fr/pdfs/Mission\\_Villani\\_Report\\_ENG-VF.pdf](https://www.aiforhumanity.fr/pdfs/Mission_Villani_Report_ENG-VF.pdf). Acesso em: 15 dez. 2024.

WESTERLUND, Mika. **The Emergence of Deepfake Technology: A Review.** *Technology Innovation Management Review*, 2019. 9(11): p.40-53. Disponível em: <http://doi.org/10.22215/timreview/1282>. Acesso em: 12 abr. 2025.

WHEELER, Thomas H. **Phototruth or photofiction? Ethics and Media Imagery in the Digital Age.** Mahwah, Lawrence Erlbaum Associates, Inc. 2002.

WIGGERS, Kyle. **Generative adversarial networks: What GANs are and how they've evolved.** VentureBeat, 2019. Disponível em: <https://venturebeat.com/2019/12/26/gan-generative-adversarial-network-explain-er-ai-machine-learning/>. Acesso em: 12 abr. 2025.

### AGRADECIMENTOS

Agradeço, primeiramente, a Deus, por nunca me abandonar e por me conceder força, sabedoria e discernimento ao longo desta trajetória, guiando-me sempre pelo caminho do bem.

Aos meus pais, Jaqueline e Júnior, pelo apoio constante e por todo o empenho em me proporcionar uma educação de qualidade e o suporte necessário em cada etapa da minha vida.

Ao meu orientador, Me. Caio, amigo de longa data, cuja orientação atenciosa, disponibilidade e compreensão frente a minha rotina foram fundamentais para a realização deste trabalho.

À banca examinadora, composta pelos professores Drs. Joana e Cláudio, pela dedicação, disponibilidade e pelas valiosas contribuições que enriqueceram significativamente esta pesquisa.

À minha namorada Maria Paula e a todos aqueles que, de alguma forma, contribuíram com minha caminhada acadêmica, sejam em grandes gestos ou em simples atitudes, como emprestar uma caneta, expresso minha sincera gratidão. Cada gesto teve seu valor e colaborou para que eu pudesse chegar até aqui.