



RESUMO: *O problema central da pesquisa gira em torno da seguinte questão: quais os impactos sociojurídicos da rápida disseminação das deepfakes e as formas de enfrentamento por múltiplos atores no Brasil? Foi realizada uma pesquisa bibliográfica, a fim de compreender o surgimento e as aplicações dessa inteligência artificial (IA), bem como analisar o grau da discussão da temática na literatura e no âmbito do legislativo. Além disso, foram analisados os potenciais danos ocasionados à jurisdição.*

PALAVRAS-CHAVES: *Deepfakes; Inteligência artificial; Jurisdição constitucional*



2

CAPÍTULO 2

DEEPPFAKES E SEUS IMPACTOS SOCIOJURÍDICOS NO BRASIL:

*UMA ANÁLISE HOLÍSTICA DOS EFEITOS MALÉFICOS E AS
FORMAS DE ENFRENTAMENTO POR MÚLTIPLOS ATORES.*

**GUSTAVO DE ASSIS SOUZA,
WILLIAM TIMÓTEO,
ARTHUR SOUSA BORGES DE MENEZES**

<https://doi.org/10.47658/20220102>

INTRODUÇÃO

Sabe-se que, na contemporaneidade, a evolução e o surgimento de novas tecnologias proporcionaram inúmeros benefícios à sociedade, em diferentes aspectos. Nesse sentido, pesquisas apontam para um decréscimo nos índices de pobreza extrema e um crescimento na expectativa de vida humana, além de indicar uma elevação no acesso à eletricidade e, como consequência, à tecnologia (ROSLING *et al*, 2019), contribuindo significativamente para a vida das pessoas.

Esse acesso, entretanto, também veio acompanhado por nefastas consequências, ocasionadas pela deturpação das imensuráveis benesses proporcionadas pelas inovações tecnológicas, as quais já foram sentidas, inclusive, em diversos contextos da sociedade e, em que pese a popularidade ganhada pelas *fakenews*, especialmente depois das eleições brasileiras de 2018, outro fenômeno tem ganhado espaço, as *deepfakes*.

Embora o termo ainda não esteja tão disseminado entre as pessoas, são temidos os seus potenciais efeitos maléficos, que podem adquirir dimensões ainda difíceis de mensurar, mas que, certamente, podem se constituir em ameaça ao Estado Democrático de Direito.

Desse modo, o surgimento e a rápida disseminação das *deepfakes* já é uma preocupação de diversos países. Muitos deles não conseguem lidar satisfatoriamente nem mesmo com as *fakenews*, e agora estão sendo obrigados a enfrentar um inimigo mais poderoso e dissimulado. Por isso, a presente pesquisa se justifica devido à escassez de estudos, sobretudo na literatura nacional, acerca do potencial danoso que essa tecnologia pode causar na vida de milhões de pessoas e até mesmo na própria Democracia.

Pretende-se com a investigação delinear caminhos para o enfrentamento da problemática e, quiçá, embasar futuras proposições legislativas em âmbito federal sobre o tema. No primeiro capítulo, foi realizada uma pesquisa bibliográfica qualitativa de artigos científicos internacionais, para uma melhor compreensão do tema, no tocante ao surgimento e às aplicações tanto benéficas quanto maléficas das *deepfakes*, bem como os seus reflexos e o possível comprometimento do Estado Democrático de Direito.

Em seguida, foi verificado o nível da discussão da referida temática na literatura científica nacional, por meio de buscas em bases de dados como o Google Acadêmico e Biblioteca Digital Brasileira de Teses e Dissertações. Também foi realizada pesquisa legislativa no site do Congresso Nacional, para identificar

possíveis proposituras legislativas em âmbito federal que versem sobre *deepfakes* e o nível da discussão da temática também na seara política.

Por fim, foram investigados os potenciais danos que poderão ser ocasionados à Jurisdição Constitucional brasileira por esse tipo de tecnologia e as formas de enfrentamento mais adequadas.

1. O SURGIMENTO DAS DEEPFAKES E SUAS IMPLICAÇÕES NO ESTADO DEMOCRÁTICO DE DIREITO

Não é novidade a crescente disseminação dos celulares *smartphones* e o conseqüente aumento na utilização de redes sociais pela população, inclusive como ferramenta de trabalho, o que tem facilitado a divulgação e compartilhamento de informações em tempo *recorde*, contribuindo assim com o alastramento de *fakenews* e *deepfakes*, por exemplo; situação agravada no Brasil, por estar classificado em quinto lugar, no *ranking* mundial dos países que mais utilizam esses aparelhos¹.

Verifica-se, deste modo, que o processo de progressão tecnológica tem favorecido a manipulação de imagens e vídeos, dada especialmente a facilitação no acesso a esse tipo de ferramenta, inclusive por meio de aplicativos de celular, de fácil utilização e sem a exigência de qualquer conhecimento técnico ou profissionalizado.

Não diferente quanto às iniciativas de inteligência artificial (IA) aplicadas ao Direito, tem-se que em poucos anos estas cresceram exponencialmente² em vários países, incluindo o Brasil. Como consequência, muitos serviços mecânicos e repetitivos realizados pelos juristas foram substituídos pela alta performance das IA's. Atividades que demandavam inúmeras horas de trabalho intenso, hoje, são realizadas em segundos e com um nível de precisão maior, como o caso da Inteligência Victor³, implementada no Supremo Tribunal Federal (STF), em 2018.

1 *Compilados na pesquisa Global Mobile Market Report, os dados mostram que o Brasil só fica atrás da Indonésia, Estados Unidos, Índia e China.*

2 *Com a ampliação desse campo cria-se uma necessidade cada vez maior de discussões éticas (PEIXOTO; SILVA, 2019).*

3 *Segundo o STF seriam necessárias 22 mil horas de trabalho humano para analisar aproximadamente 42 mil processos que são recebidos por semestre pela corte. O objetivo do Victor é justamente acelerar esse tempo de análise por meio de ferramentas de processamento de linguagem natural (SILVA et. al, 2018).*

Os benefícios que foram propiciados à atividade jurídica, assim como em outros setores, são notórios. Entretanto, como se sabe, a utilização da IA de forma indevida também gera grande preocupação mundial na contemporaneidade. Tal fator suscita inúmeros debates na esfera jurídica entorno de questões éticas e normativas. Nesse contexto, destaca-se o surgimento das *deepfakes*, que consistem em redes neurais⁴ que examinam um vasto conjunto de dados, com o objetivo de aprender a imitar expressões faciais, vozes e jeitos dos seres humanos.

Esse tipo de IA consegue literalmente mudar a face de uma pessoa por outra em um vídeo ou imagem ou mesmo criar pessoas que não existem, de forma extremamente realista. Além disso, ela aprende a melhorar seus resultados continuamente. Tal fato resulta em *deepfakes* cada vez mais reais e de difícil detecção, até mesmo por pessoas ligadas à área de tecnologia.

Importante destacar que essa nomenclatura surgiu recentemente, no ano de 2017, por meio da plataforma Reddit. Além disso, ela não se confunde com as tão difundidas *fake news* e *shallowfakes*⁵. Os casos mais notórios até agora são aqueles relacionados à pornografia. De formas maliciosas, usuários desse tipo de tecnologia usam o rosto de celebridades famosas no corpo de atrizes pornô (WESTERLUND, 2019).

Segundo uma pesquisa recente, desenvolvida pela DeepTrace Labs, chegou-se ao resultado de que 96% das *deepfakes* são pornográficas ou envolvem imagens de sexo, sendo as mulheres o seu principal alvo (AJDER, 2019). Além das celebridades, os políticos também são alvo cada vez mais frequentes desse tipo de IA. Nesse último caso, o objetivo é afetar negativamente a imagem dessa pessoa pública perante a sociedade. Em casos mais graves, é possível até mesmo desencadear uma crise institucional.

Cumprе mencionar que se engana quem pensa que é necessário ser um profundo conhecedor de programação para a criação dessa tecnologia. Na realidade, com a difusão dos *softwares open source* e o surgimento de aplicativos como o FakeApp, qualquer pessoa mal-intencionada, sem

-
- 4 As *deepfakes* são fruto das Redes Adversariais Generativas (GAN's) que consistem em um sistema de aprendizado supervisionado. Elas se desenvolvem por meio de duas redes neurais artificiais que trabalham conjuntamente para gerar mídias com aparências hiper-realistas (WESTERLUND, 2019). Já as redes neurais são definidas por Peixoto e Silva (2019, p. 98) como "um tipo de sistema computacional inspirado pelas propriedades básicas de neurônios biológicos".
 - 5 Podem ser definidas como edições mais grosseiras e superficiais que não utilizam redes neurais. Em regra, os vídeos originais são descontextualizados por meio dessas edições, a fim de manipular o expectador. O termo foi cunhado pela primeira vez por Sam Gregory que é ativista de direitos humanos da organização Witness. (AJDER, 2019).

conhecimentos técnicos de programação, porém com um acervo de imagens de outra (vítima) consegue criar uma *deepfake* sem muitos esforços (WESTERLUND, 2019), mas, ainda assim, gerando grande grau de realismo e resultados exageradamente convincentes.

Para jogar luz a essa realidade, um engenheiro chamado Phillip Wang criou um site chamado "*This Person Does Not Exist*", para gerar rostos artificiais, a partir da combinação de fotos reais. O *software* por detrás do site se chama StyleGAN e foi desenvolvido pelas Redes Adversariais Generativas (GAN's). Qualquer pessoa, ao entrar no site, gera automaticamente a imagem de uma pessoa não real, de forma aleatória. Observe-se o exemplo abaixo, que foi gerado ao acessar o site em 20/09/2021:



Fonte: Tero Karras, 2021.

Nesse sentido, questiona-se: como um ser humano poderá detectar que esse rosto acima, por exemplo, não é de uma pessoa real? Essa é uma reflexão que será aprofundada no último capítulo, quando a pesquisa adentrará nas tecnologias *antideepfakes*.

Destaca-se ainda que este artigo não desconsidera os benefícios que essa tecnologia trouxe para determinados setores da sociedade. Dentre os inúmeros exemplos de segmentos beneficiados, estão a indústria cinematográfica, de games, educacional e empresarial. Entretanto, o foco do

trabalho concerne em analisar as consequências danosas dessa tecnologia e os seus impactos sociojurídicos.

Como se viu, nas *deepfakes* inserem-se os rostos de pessoas em cenas nas quais elas nunca estiveram, com a possibilidade de alteração de falas etc., criando a falsa sensação de que são reais, o que, para Westerlund (2019), configura-se em temerário disfarce, no caminho da manipulação da sociedade e do Estado, em todas as suas esferas de poder.

A despeito disso, torna-se certo de que esse fenômeno constitui-se em arma potencial contra a ordem democrática, seja em um cenário individualizado, de ofensa a um direito humano; ou coletivo, no âmbito eleitoral (EMPOLI, 2020), especialmente porque, nesse caso, segundo pesquisa realizada pelo Instituto DataSenado, 79% dos brasileiros utilizam-se do WhatsApp como fonte de informação (BRASIL, 2019). Isso pode resultar, sem dúvidas, em uma deturpação da democracia representativa, por meio dessa nova tecnologia.

Desse modo, a criação de agências especializadas na verificação dessas informações torna-se relevante. Entretanto, a sua atuação é limitada, considerando a velocidade em que as notícias são transmitidas e circulam entre a população. Além disso, o perigo reside ainda em concentrar a chancela da validade da informação nas mãos do Poder Público, o que pode ensejar no comprometimento de princípios basilares da democracia liberal e da justiça social.

Nesse cenário, os jornalistas continuam sendo o principal mecanismo de combate à disseminação de informações inverídicas. Por isso, não há dúvidas de que a liberdade de imprensa representa uma das mais expressivas conquistas do sistema democrático brasileiro. É de se preocupar, no entanto, que a inovação trazida pelas *deepfakes* arruíne a solidificação dessa instituição, tão fundamental à sociedade e à manutenção da democracia.

Isso porque, os jornalistas estarão diante de um dilema que gira em torno de duas facetas: o compromisso com a verdade ou a manutenção da velocidade das informações. Neste último caso, o profissional correrá o risco de tornar-se ele próprio o disseminador de *fakenews*, *deepfakes* e discursos de ódio. Em qualquer dos casos, todavia, a confiabilidade mantida pela imprensa livre será estremecida.

Diante desse contexto, questiona-se se a completa proibição das *deepfakes*, através de medida legislativa, seria a alternativa mais adequada. Contudo, tal iniciativa esbarra na possibilidade de ofensa ao princípio constitucional da liberdade de expressão e, com isso, na restrição de um direito fundamental, descaracterizando a sua essência principal (BARROSO, 2018) e oferecendo uma brecha legal para o autoritarismo.

Por outro lado, o desempenho da atividade de controle pelo Estado pode resultar na incredulidade da população; e com razão, portanto, já que o monopólio sobre a veracidade das informações ficaria em suas mãos, ocasionando, em uma conjuntura mais extrema, mas não impensável, numa ditadura estatal velada.

Com isso, a partir de tudo aqui já delineado, torna-se perceptível que a vedação legislativa das *deepfakes* poderia resultar em adversidades ainda mais difíceis de ser enfrentadas. Ignorar os efeitos benéficos dessa tecnologia e o direito à utilização pela sociedade, pela indústria cinematográfica e na seara educacional, por sua vez, incorreria em manifesto atentado aos institutos da liberdade e da manutenção do progresso social.

Por isso, qualquer mecanismo desenvolvido com o objetivo de controlar a sua utilização indevida será profundamente complexo, do mesmo modo que o problema realmente o é.

2. PANORAMA DAS DEEPFAKES NA LITERATURA NACIONAL E LEGISLATIVO FEDERAL

Compreender o nível das discussões sobre *deepfakes* no âmbito da literatura nacional, bem como do legislativo federal, é de fundamental importância para delinear possíveis caminhos para lidar com essa problemática no Brasil.

Para tanto, no tocante às produções científicas, a presente pesquisa conduziu uma investigação, por meio do mecanismo de pesquisa avançada do Google Acadêmico, a fim de identificar quantos trabalhos científicos se encontravam disponíveis na referida plataforma e que versavam sobre a temática, em português. O período específico da busca foi de 2017⁶ a 2021⁷.

As palavras-chave utilizadas foram "*deepfake*" (9 resultados encontrados) "*deep fake*" (0 resultados encontrados) e suas respectivas formas no plural "*deepfakes*" (8 resultados encontrados) e "*deep fakes*" (1 resultado encontrado).

Foi excluída a palavra-chave *fake news*, pois em que pese a proximidade de nomenclatura com *deepfakes*, ambas exigem análises e propostas de intervenções distintas, uma vez que são geradas e disseminadas de formas diferentes. Portanto, não foram considerados os resultados nos quais *deepfakes* e *fakenews* apareciam expressamente no título e indicavam, deste modo, uma análise conjunta.

6 Ano do surgimento das *deepfakes*.

7 Os dados se encontram atualizados até 02 nov. 2021.

De todos os resultados encontrados, pôde-se observar que a grande maioria dos trabalhos são muito recentes e foram publicados entre 2020 e 2021, com exceção de um que foi publicado em 2018, sendo, portanto, a primeira produção brasileira sobre o assunto a constar no Google Acadêmico, a partir dos critérios de busca retromencionados.

Esse primeiro artigo possui como título "Implicações políticas, sociais e éticas da popularização de *deepfakes* por meio de *softwares* livres". Os autores defendem, em síntese, como proposta de enfrentamento às *deepfakes*, a necessidade de atualização legislativa criminal e a atuação conjunta entre sociedade e provedores de aplicação, para resguardar os direitos dos usuários, inclusive das mulheres, que são as mais afetadas. Além disso, no campo político, sustentam a necessidade de constante verificação e questionamento das fontes (KEHDY, 2018).

Tais propostas mostram-se adequadas, com exceção da atualização criminal, tendo em vista que leis penais mais severas ou a criação de uma nova tipificação penal para determinada conduta, considerada delituosa, não tem o condão de, por si só, diminuir a incidência dos crimes (GOMES, 2013). Na realidade, o discurso punitivo é utilizado com mera retórica eleitoral e não deve ser endossado nas pesquisas científicas, tendo em vista a já comprovada ineficácia.

Além do Google Acadêmico, foram realizadas buscas na Biblioteca Brasileira de Teses e Dissertações (BDTD), utilizando-se das mesmas palavras-chave, e foi encontrado apenas 1 resultado, a dissertação sob o título "detecção de *deepfakes* a partir de técnicas de visão computacional e aprendizado de máquina", na área da Ciência da Computação. O objetivo da pesquisa consistiu em analisar como as técnicas de aprendizado de máquina e de aprendizado profundo podem auxiliar na detecção das *deepfakes*. A pesquisadora analisou as principais bases de dados disponíveis e desenvolveu uma própria, composta de vídeos falsos envolvendo mulheres⁸.

Deste modo, ao olhar para a literatura acadêmica nacional, pode-se concluir que são necessários mais estudos críticos que envolvam as *deepfakes* nas suas várias nuances, não só no âmbito da Computação, mas também no das Ciências Humanas.

A baixa produção acadêmica sobre a temática deve servir de alerta para a comunidade científica. Além disso, mostra-se de suma importância que esse conhecimento não se restrinja ao âmbito acadêmico, pois é de interesse de

8 Principal público-alvo das *deepfakes*.

toda sociedade. Portanto, mostra-se adequado a educação e treinamento de toda a população, desde o período escolar⁹.

Já no tocante ao âmbito legislativo federal, a presente pesquisa também procedeu a uma investigação no Portal da Câmara dos Deputados¹⁰, a fim de identificar possíveis proposições legislativas que abordassem tal tecnologia. Foram utilizadas as mesmas palavras-chave da pesquisa realizada no Google Acadêmico.

Foram encontrados quatro resultados¹¹, mas apenas um deles realmente versava sobre *deepfakes*, que foi a proposição legislativa nº 1429. Ela foi apresentada em 01/04/2020 e possuía como autores os deputados federais Felipe Rigoni e Tabata Amaral, mas cerca de dois meses depois, foi arquivada, a pedido dos próprios deputados¹².

Em síntese, a referida propositura pretendia instituir a Lei Brasileira de Liberdade, Responsabilidade e Transparência na Internet¹³. A mesma estabelece diretrizes de atuação obrigatória para os provedores de aplicação em relação a contas inautênticas, disseminação de desinformação, conteúdos, anúncios online e propagandas políticas patrocinadas.

Em que pese ser uma iniciativa legislativa com um cunho muito mais direcionada ao combate às *fakenews*, o art. 5º, §3º, também privilegiava as *deepfakes*. No dispositivo, foi estabelecida a necessidade de os provedores de aplicação desenvolverem procedimentos de acompanhamento, para melhorar a proteção dos usuários contra comportamentos ilícitos, incluindo a proteção contra o uso de imagens manipuladas para imitar a realidade (*deepfakes*), entre outros.

Tal previsão é tímida, devido à magnitude e às nuances de tal problemática. Entretanto, o caráter defensivo, ainda assim, mostra-se mais adequado do que medidas populistas e ineficazes, como as atualizações legislativas criminais.

9 WESTERLUND (2019) recomenda a inclusão de pensamento crítico e alfabetização digital nas escolas.

10 A página específica consultada foi: <https://www.camara.leg.br/buscaProposicoesWeb/pesquisaSimplificada>. Acesso em: 20 set. 2020.

11 A palavra-chave utilizada que teve êxito na descoberta de quatro resultados foi "deep fake".

12 Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1898026&filename=Tramitacao-PL+1429/2020. Acesso em: 26 nov. 2020.

13 Disponível em: https://www.camara.leg.br/proposicoesWeb/prop_mostrarintegra?codteor=1872575&filename=PL+1429/2020. Acesso em: 25 nov. 2020.

3. OS IMPACTOS NA JURISDIÇÃO CONSTITUCIONAL E POSSÍVEIS CAMINHOS DE ENFRETEAMENTO

A Jurisdição Constitucional também é uma das potenciais áreas afetadas pela disseminação das *deepfakes*, especialmente no que diz respeito à prova. Sabe-se que o sistema judicial brasileiro adota a possibilidade de os magistrados serem assistidos por peritos, quando a prova do fato depender de conhecimento técnico ou científico, conforme artigo 156 do Código de Processo Civil.

Mas, as provas que são juntadas pelas partes em ações judiciais em diversos ramos do direito não contam com uma validação pericial e, portanto, estão sujeitas a fraude, especialmente com a rápida disseminação das *deepfakes*. Assim, como um magistrado poderá verificar se determinada prova apresenta pelas partes é uma *deepfake* ou não?

Destaca-se que não se tem notícias ainda da utilização de *deepfakes* no judiciário brasileiro. Para essa constatação, foram empreendidos esforços, por meio da pesquisa jurisprudencial nos websites dos órgãos que compõem o Poder Judiciário Brasileiro¹⁴, com as palavras-chave "*deepfake*", "*deep fake*" e suas respectivas formas no plural "*deepfakes*" e "*deep fakes*", na data de 14/09/2021.

Após a pesquisa, constatou-se resultado em 7 deles. Esses tribunais foram: o Supremo Tribunal Federal (2 resultados), o Tribunal Regional Federal da 3ª Região (2 resultados), o Tribunal Regional do Trabalho da 6ª Região (1 resultado), o Tribunal Regional do Trabalho da 22ª Região (3 resultados), o Tribunal Superior Eleitoral (2 resultados), o Tribunal de Justiça do Estado de Minas Gerais (1 resultado) e o Tribunal de Justiça do Estado do Pará (2 resultados).

Para a comprovação de que realmente se tratava de caso de *deepfake*, foi realizada a abertura do inteiro teor da jurisprudência e, a partir do comando "Ctrl+f", foi escrito cada um dos modos da palavra *deepfake* descritos

14 esses órgãos estão especificados no art. 92 da Constituição da República Federativa do Brasil de 1988:

"Art. 92. São órgãos do Poder Judiciário:

I - o Supremo Tribunal Federal;

I-A o Conselho Nacional de Justiça;

II - o Superior Tribunal de Justiça;

II-A - o Tribunal Superior do Trabalho;

III - os Tribunais Regionais Federais e Juízes Federais;

IV - os Tribunais e Juízes do Trabalho;

V - os Tribunais e Juízes Eleitorais;

VI - os Tribunais e Juízes Militares;

VII - os Tribunais e Juízes dos Estados e do Distrito Federal e Territórios."

anteriormente. Constatou-se assim que não se trata de *deepfake*, mas de outras palavras parecidas que tinham o termo "deep" ou "fake".

Mesmo assim, não se pode descartar a possibilidade da utilização de *deepfakes* no judiciário brasileiro, uma vez que a análise de jurisprudências é limitada. Não é possível, nessa análise, abarcar, por exemplo, todos os processos que tramitam em primeiro grau. Soma-se a isso o fato de que poucas pessoas conhecem a nomenclatura e as diferenciações em relação a *fakenews* e *shallowfakes*.

Em alguns países como os Estados Unidos, já foram estabelecidas restrições bem rígidas para que os magistrados aceitem uma determinada prova em juízo. A *Rule 901(b) (9) of the US Federal Rules of Evidence* estabelece, por exemplo, que haja o testemunho de alguma pessoa (geralmente um perito) com conhecimento técnico ou científico que demonstre evidências da validade e confiabilidade da prova, antes de ela ser admitida. Até mesmo as qualificações e competências do perito são levadas em consideração para o convencimento judicial.

No âmbito criminal, as *deepfakes* podem resultar em condenações injustas, o que leva a conclusão da necessidade das testemunhas oculares, bem como de peritos especializados para corroborarem a validade da prova (imagem ou vídeo) (MARAS; ALEXANDROU, 2019).

Mas, mesmo com essas precauções, essa tecnologia aprende constantemente. Logo, os vídeos e imagens falsas, com o decorrer do tempo, tornam-se cada vez mais realistas e até mesmo os especialistas poderão não conseguir identificar a manipulação (MARAS; ALEXANDROU, 2019).

Nesse caso, apenas as tecnologias *antideepfakes*, que incluem detecção, autenticação de conteúdo e prevenção, poderão sanar quaisquer vieses interpretativos.

No Brasil, a Justiça Eleitoral vem alertando a população, por meio de seus portais na rede mundial de computadores¹⁵, sobre o que são as *deepfakes*, o que fazem e algumas maneiras de identificá-las. Mas, mesmo com esses alertas e com campanhas para a sensibilização da população, é preciso ter muita cautela, pois as *deepfakes* vêm a cada dia se desenvolvendo e assim dificultando mais a sua identificação. Além disso, deve-se destacar que milhões de brasileiros não tem acesso à internet para acessar esse tipo de informação.

15 Tribunal Superior Eleitoral (TSE). Programa Minuto da Checagem explica o que é "deepfake". Disponível em: <<https://www.tse.jus.br/videos/programa-minuto-da-checagem-explica-o-que-e-deepfake>>. Acesso em: 13 set. 2021.

Até a data da realização desse artigo, ainda não existe nenhum *software* que consiga diferenciar uma *deepfake* de um vídeo ou imagem real com exatidão. Mas, já existe no mercado alguns *softwares*, como por exemplo o que foi desenvolvido pela *Microsoft Research*, que se chama *Microsoft Video Authenticator*. Dentre as suas funções, estão:

“[...] analisar uma foto ou vídeo para fornecer um percentual, ou pontuação de confiança, de indica se a mídia foi – ou não – manipulada artificialmente. No caso de um vídeo, ele pode fornecer essa porcentagem em tempo real em cada quadro à medida que o conteúdo é reproduzido. Ele funciona detectando o limite de mesclagem dos elementos do *deepfake* e o sutil desvanecimento ou tons de cinza que podem não ser detectados pelo olho humano.” (MICROSOFT NEWS CENTER BRASIL, 2020).

Convém destacar que esse *software* ainda não está liberado para a utilização pela população em geral, podendo ser usado somente por alguns grupos específicos, como organizações envolvidas no processo democrático, incluindo veículos de notícias e campanhas políticas.

No ano de 2019, o Facebook reuniu-se com outras empresas, como a *Partnership on AI*, Microsoft e acadêmicos da *Cornell Tech*, MIT, *University of Oxford*, UC Berkeley, *University of Maryland*, *College Park* e *University at Albany-SUNY*, para construir o *Deepfake Detection Challenge (DFDC)*.

O objetivo desse encontro foi a produção de *softwares* que pudessem detectar melhor a utilização de *deepfakes*. Esse projeto conteve a participação de especialistas do mundo inteiro e houve compartilhamento dos projetos entre os participantes, para que pudessem comparar com os trabalhos que estavam realizando.

Para que houvesse incentivo das pessoas e empresas que participaram do DFDC, foram incluídos concessões e prêmios, de acordo com a classificação alcançada pelos participantes. Foram utilizados mais de 100.000 vídeos, dentre *deepfakes* e não *deepfakes*, para que os participantes do demonstrassem a eficácia de seus *softwares*, na identificação de vídeos manipulados por IA.

Mesmo com a participação de várias pessoas e o desenvolvimento de vários modelos de *software* e compartilhamento de seus trabalhos entre os participantes, não houve nenhum que conseguiu diferenciar uma *deepfake* de um vídeo real com 100% de precisão.

Assim, é de grande importância mundial que os governos, pesquisadores, jornalistas, organizações e as empresas trabalhem em conjunto para

combater as *deepfakes* e desenvolver *softwares* que consigam detectá-las com extrema exatidão.

Assim como as *fakenews*, as *deepfakes* têm um grande poder de trazer desinformação e é um grande perigo, tanto para a iniciativa privada quanto para a iniciativa pública, para a imagem das pessoas perante a sociedade e, em casos mais extremos, até mesmo aos regimes políticos democráticos de um país.

CONSIDERAÇÕES FINAIS

A presente pesquisa analisou os impactos danosos das *deepfakes* na sociedade contemporânea. Identificou-se que as mulheres são as mais afetadas com a disseminação das *deepfakes* de cunho sexual (96% dos resultados) e que os políticos também são alvo cada vez mais frequentes.

Buscou-se responder, no decorrer da investigação, a seguinte problemática: quais os impactos sociojurídicos da rápida disseminação das *deepfakes* e as formas de enfrentamento por múltiplos atores no Brasil? Em síntese, chegou-se ao resultado de que os principais atores (pesquisadores, jornalistas, legisladores e o judiciário brasileiro) ainda não estão preparados para lidar e combater de forma efetiva as *deepfakes* e resguardar a sociedade.

Foram apontados os seguintes caminhos para a resolução dessa problemática: 1- a necessidade do desenvolvimento de mais estudos e pesquisas críticas acerca da temática e que debatam suas várias nuances; 2- a disseminação desse conhecimento científico e técnico para lidar com as *deepfakes*, por meio da educação e treinamento da sociedade; 3- a participação da mídia e o reforço da sua independência; 4- a necessidade de peritos especializados e tecnologias anti *deepfakes* (o mais recomendado) como forma de assegurar a fidelidade de determinada prova apresentada em juízo.

Com a adoção de tais propostas, será possível combater, ou ao menos minimizar, essa problemática. Esse rol de soluções não é taxativo e deverá ser incrementado, de acordo com a evolução dos debates e pesquisas.

REFERÊNCIAS

AJDER, Henry. *et al.* **State of Deepfakes: Landscape, Threats, and Impact.** Amsterdã: Deeptrace, 2019.

BARROSO, Luís Roberto. **Curso de direito constitucional contemporâneo: os conceitos fundamentais e a construção do novo modelo.** 7.ed. São Paulo: Saraiva, 2018.

BRASIL. Instituto de Pesquisa DataSenado. **Redes sociais, notícias falsas e privacidade de dados na internet.** Novembro/2019. Disponível em: <https://www12.senado.leg.br/institucional/datasenado/arquivos/mais-de-80-dos-brasileiros-acreditam-que-redes-sociais-influenciam-muito-a-opinio-das-pessoas>. Acesso em: 18 set. 2021.

EMPOLI, Giuliano da. **Os engenheiros do caos: Como as fake news, as teorias da conspiração e os algoritmos estão sendo utilizados para disseminar ódio, medo e influenciar eleições.** São Paulo: Vestígio, 2020.

Facebook AI. **Creating a dataset and a challenge for deepfakes.** Publicado em: 05 de setembro de 2019. Disponível em: <https://ai.facebook.com/blog/deepfake-detection-challenge/>. Acesso em: 19 set. 2021.

Facebook AI. **Deepfake Detection Challenge Dataset.** Publicado em: 25 de junho de 2020. Disponível em: <https://ai.facebook.com/datasets/dfdc/>. Acesso em: 19 set. 2021.

Facebook AI. **Deepfake Detection Challenge launches with new dataset and Kaggle site.** Publicado em: 11 de dezembro de 2019. Disponível em: <https://ai.facebook.com/blog/deepfake-detection-challenge-launches-with-new-data-set-and-kaggle-site/>. Acesso em: 19 set. 2021.

GOMES, Luiz Flávio. ALMEIDA, Débora de Souza. **Populismo penal midiático: caso mensalão, mídia disruptiva e direito penal crítico.** São Paulo: Saraiva, 2013.

KARRAS, Tero. *et al.* **Imagem aleatória de uma pessoa que não existe gerada com redes neurais adversariais.** Disponível em: This Person Does Not Exist. Acesso em: 20 set. 2021.

KEHDY, Bernardo Okazaki. *et al.* Implicações políticas, sociais e éticas da popularização de deepfakes por meio de softwares livres. In: **Anais do Congresso Nacional Universidade, EAD e Software Livre.** v.2, n. 9, 2018. Disponível em: <http://www.periodicos.letras.ufmg.br/index.php/ueadsl/article/view/14431>. Acesso em: 25 nov. 2020.

MARAS, Marie-Helen; ALEXANDROU, Alex. **Determining authenticity of video evidence in the age of artificial intelligence and in the wake of Deepfake videos.** The International Journal of Evidence & Proof, v. 23, n. 3, p. 255-262, 2019.

Microsoft News Center Brasil. **Novos passos para combater a desinformação. Publicado em: 22 de setembro de 2020.** Disponível em: <https://news.microsoft.com/pt-br/novos-passos-para-combater-a-desinformacao/>. Acesso em: 13 set. 2021.

MOURA, Camila Stéffane Fernandes Teixeira de. **Detecção de Deepfakes a partir de técnicas de visão computacional e aprendizado de máquina.** 2021. 1 recurso online (105 p.) Dissertação (mestrado) - Universidade Estadual de Campinas, Instituto de Computação, Campinas, SP.

PEIXOTO, Fabiano Hartmann; SILVA, Roberta Zumblick Martins da. **Inteligência artificial e direito**. 1.ed. Curitiba: Alteridade Editora, 2019.

ROSLING, Hans; ROSLING, Ola; RÖNNLUND, Anna Rosling. **Factfulness**: o hábito libertador de só ter opiniões baseadas em fatos. Tradução de Vitor Paolozzi. 1ª ed., Rio de Janeiro: Record, 2019.

SILVA, Nilton Correia da. *et al.* **Document type classification for Brazil's Supreme Court using a Convolutional Neural Network**. The tenth international conference on forensic computer science and cyber law - ICoFCS 2018. p. 7-11, 2018.

Tribunal Superior Eleitoral. **Jurisprudência**. Disponível em: <<https://www.tse.jus.br/jurisprudencia/decisooes/jurisprudencia>>. Acesso em: 14 set. 2021.

Supremo Tribunal Federal. **Jurisprudência**. Disponível em: <<https://portal.stf.jus.br/>>. Acesso em: 14 set. 2021.

Tribunal de Justiça do Estado de Minas Gerais. **Jurisprudência**. Disponível em: <<https://www5.tjmg.jus.br/jurisprudencia/formEspelhoAcordao.do>>, <<https://www5.tjmg.jus.br/jurisprudencia/decisao.do>>, <<https://www5.tjmg.jus.br/jurisprudencia/decisaoVice.do>>, <<https://www5.tjmg.jus.br/jurisprudencia/decisaoTurmaRecursal.do>> e <<https://www5.tjmg.jus.br/jurisprudencia/sentenca.do>> Acesso em 11 set. 2021.

Tribunal de Justiça do Estado do Pará. **Jurisprudência**. Disponível em: <<http://gsa-index.tjpa.jus.br/consultas/search?q=&client=consultas&proxystylesheet=consultas&site=jurisprudencia&sort=date%3AD%3AS%3Ad1&aba=JP>>. Acesso em 11 set. 2021.

Tribunal Regional do Trabalho da 22ª Região. **Jurisprudência**. Disponível em: <<https://siscle.trt22.jus.br/jurisprudencia-beta/sentenca/pesquisa-sentenca.jsf>> e <<https://siscle.trt22.jus.br/jurisprudencia-beta/acordao/pesquisa-acordao.jsf>>. Acesso em: 12 set. 2021.

Tribunal Regional do Trabalho da 6ª Região. **Jurisprudência**. Disponível em: <<https://apps.trt6.jus.br/consultaAcordaos/>>. Acesso em: 12 set. 2021.

Tribunal Regional Federal da 3ª Região. **Jurisprudência**. Disponível em: <<http://web.trf3.jus.br/base-textual>>. Acesso em 14 set. 2021.

Tribunal Superior Eleitoral. **Programa Minuto da Checagem explica o que é "deepfake"**. Disponível em: <<https://www.tse.jus.br/videos/programa-minuto-da-checagem-explica-o-que-e-deepfake>>. Acesso em: 13 set. 2021.

WESTERLUND, Mika. **The Emergence of Deepfake Technology**: A Review. *Technology Innovation Management Review*, v. 9, n. 11, 2019.