
MAPEAMENTO DA DETECÇÃO DE *DEEPPAKES*: um trabalho terminológico

Mapping detection of Deepfakes: a terminological work

Gabriel Meneguelli Soella (1), Giovana Deliberali Maimone (2)

(1) Universidade de São Paulo (USP), Brasil, soellagm@usp.br
(2) Universidade de São Paulo (USP), Brasil, gdmaimone@usp.br



Resumo

Os documentos audiovisuais gerados/modificados artificialmente com *deepfakes* podem ser usados para propagar desinformação, e existem pesquisas voltadas para sua detecção. A fim de promover um mapeamento do domínio para detecção de *deepfakes*, coloca-se como uma possibilidade a análise bibliográfica (seleção de termos/conceitos) de base teórica referente à Organização do Conhecimento, Terminologia e Arquitetura da Informação. Os objetivos são: realizar mapeamento terminológico do domínio da detecção de *deepfakes* e verificar possibilidades de relação teórica dos materiais selecionados. O método indutivo foi selecionado para levantamento da bibliografia e mapeamento do domínio a partir de uma pesquisa exploratório-descritiva de abordagem qualitativa, sendo que o aporte teórico foi analisado sob uma perspectiva reflexiva e crítica. O levantamento bibliográfico se deu nas bases Web of Science, Scopus (Elsevier) e Association for Computing Machinery. O recorte teórico-epistemológico demonstrou possibilidades de aplicação em domínios-web. Foram identificadas 22 publicações em artigos de periódicos, expondo contexto e propostas de detecção de *deepfakes* e outras técnicas de detecção de manipulação/geração artificial de documentos audiovisuais resultando em lista com 78 termos candidatos. Conclui-se que há necessidade de novos estudos na área pois os resultados apontaram para um domínio mais amplo do que apenas uso de técnicas *deepfake* com uso de GAN (*Generative Adversarial Networks*).

Palavras-chave: Desinformação; *Deepfakes*; Redes Adversariais Generativas (GAN); Organização do Conhecimento; Terminologia.

Abstract

The audiovisual documents artificially generated/modifiers with *deepfakes* can be used by disinformation propagation, and have researches for your detection. In order to propose a *deepfake* detection domain mapping, this study propose the bibliographic analysis (terms/concepts selection) in the Knowledge Organization, Terminology, and Information Architecture theoretical basis. The objectives are: realize terminological mapping on *deepfake* detection domain and verifier possible theoretical relationship between interdisciplinary basis selected. The inductive method was selected by bibliographic survey and

domain mapping from exploratory-descriptive research of applied qualitative, and the theoretical basis was analyzed in reflexive and critical perspective. The bibliographic survey occurred in Web of Science, Scopus (Elsevier), and Association for Computing Machinery repositories. The theoretical-epistemological cut demonstrated application possibilities in web-domains. It was identified 22 papers with context and propositions of deepfake detection and others detection techniques of artificial manipulation/generation of audiovisual documents resulting in list with 78 candidate's terms. In conclusion, believe that are necessity of new studies in the domain because the results indicate to a largest domain than exclusively deepfake techniques with Generative Adversarial Networks use.

Keywords: Disinformation; Deepfakes; Generative Adversarial Networks (GAN); Knowledge Organization; Terminology.

1 Introdução

A *World Wide Web*, fundada em 1989 por Tim Berners-Lee com o objetivo de “[...] criar um meio de troca de dados utilizando a infraestrutura da internet, tendo como base o conceito do hipertexto [...]” (VIDOTTI et al., 2019, p. 197), trouxe consigo muitos desafios para a Organização do Conhecimento e a Representação/Recuperação da Informação. Entre eles, o surgimento de tecnologias de informação e comunicação (TIC) e a disseminação de documentos digitais na *web*. Após o desenvolvimento de plataformas digitais, com conteúdos informacionais circulando entre seus usuários, a disseminação digital de documentos e o desenvolvimento de novas TICs se acentuou, com destaque às plataformas de mídias sociais.

O número de perfis de usuários ativos em plataformas de mídias sociais – como *Facebook*, *Instagram*, *TikTok*, *WhatsApp*, *YouTube* etc. – referentes ao ano de 2020 foi próximo de 4.2 bilhões. Esse fenômeno representou um aumento de 13,2% em relação ao ano anterior e ascendente desde 2016. Entre perfis verdadeiros e falsos de indivíduos, grupos, instituições etc., as relações estabelecidas nas plataformas facilitaram e intensificaram os processos de interação e compartilhamento de informação e documentos entre usuários mediante postagens públicas ou privadas, mensagens enviadas por *chats*, *e-mail*, comentários em publicações, marcação de conteúdos com *hashtags* etc. (KEMP, 2021).

Na última década, as postagens em mídias sociais revelaram grande quantidade de afirmações falsas, discursos de ódio, “*fake news*”, testemunhal falso, negacionismos científico e histórico (ARAÚJO, 2021), conspirações e outras práticas de manipulação e modulação da informação na *web*. Interpretações racionais da realidade foram preteridas diante de afirmações

em favor do controle da percepção da realidade via apelos emocionais, no que foi chamado de “pós-verdade” (*post-truth*) (OXFORD LANGUAGES, 2016).

A disseminação de informação manipulada/falseada por vieses políticos e ideológicos em plataformas digitais se acentuou a partir de 2016, afetando a percepção da realidade com consequências potencialmente prejudiciais à sociedade. Bradshaw e Howard (2019) identificaram que a manipulação política da opinião pública em processos eleitorais por plataformas digitais de redes sociais foi progressiva em 2017 (presente em 28 países), em 2018 (em 48 países) e em 2019 (em 70 países). Inclusive no Brasil em 2018 onde, de acordo com os autores, se destacaram as plataformas *WhatsApp*, *YouTube* e *Facebook*. Desinformação e negacionismos que, como exemplificado por Araújo (2021), também foram massivamente empregados para negar a existência do aquecimento global, promover mentiras sobre vacinas e/ou recusar quaisquer informações e conhecimentos antagônicos aos projetos despóticos dos detentores do poder.

No contexto de práticas de desinformação e negacionismos, Froehlich (2020) apontou o uso de técnicas automatizadas, entre as quais destaca-se a técnica de *deepfake* em vídeos (embora esta também seja aplicável a imagens estáticas e áudios). Com as potencialidades de usos de inteligência artificial e o intenso fluxo de desenvolvimento de TICs na contemporaneidade, o mapeamento de técnicas voltadas para combater essas falsificações/adulterações se evidencia como um desafio para a comunidade científica e sociedade civil como um todo.

Em meio a uma intensa dinâmica digital, facilitar processos de recuperação da informação do domínio de detecção de *deepfakes*, por meio dos termos que compreendem o domínio, pode auxiliar especialistas do domínio e proporcionar celeridade e dinamicidade na tentativa de mapear e organizar conhecimentos relacionados. A identificação, a organização e o mapeamento terminológico do domínio, voltado à ação de detecção de *deepfakes* (que são geradas com *Generative Adversarial Networks*), pode se tornar um trajeto de investigação do qual emerge o seguinte problema: é possível mapear os termos do domínio da detecção de *deepfakes* mediante estado-da-arte recuperado e sob um aporte teórico de Organização do Conhecimento, Terminologia e Arquitetura da Informação?

Com a finalidade de investigar o problema proposto, a pesquisa transcorreu sob dois objetivos. O primeiro objetivo (geral) consiste em realizar um mapeamento terminológico do

domínio da detecção de *deepfakes*, a partir do estado-da-arte do tema em artigos de periódicos recuperados em bases de dados. A observação prévia do fenômeno (uso da técnica *deepfake* para desinformar) expôs indícios de que existem técnicas voltadas para sua detecção, ou seja, existe um domínio sobre o qual se objetiva proceder mapeamento terminológico sob base teórica de Organização do Conhecimento, Terminologia e Arquitetura da Informação. O segundo objetivo (específico) está atrelado ao primeiro, visto que intenta identificar as áreas de conhecimento em que os documentos recuperados se inserem e as relações que estabelecem entre si no processo de detecção de *deepfakes*.

Para a investigação utilizou-se o método indutivo (GIL, 2008), mediante pesquisa exploratório-descritiva de abordagem qualitativa aplicada. Um processo de pesquisa que ocorreu através de base teórico-metodológica da Organização do Conhecimento na Ciência da Informação em diálogo interdisciplinar com a Terminologia e a Arquitetura da Informação. Nesse sentido, o artigo foi organizado nas seguintes seções. A seção 2 Desenvolvimento, apresenta o aporte teórico-epistemológico interdisciplinar da Organização do Conhecimento, da Terminologia e da Arquitetura da Informação que legitima a pesquisa realizada. Os procedimentos metodológicos supracitados foram detalhados na seção 3; e na seção 4, os resultados do mapeamento terminológico foram explicitados com a ressalva de que a lista de termos criada a partir da pesquisa, por seu tamanho, foi inserida como Apêndice. Por fim, a seção 5 expõe considerações finais sobre o estudo, suas limitações e caminhos possíveis identificados para prosseguimentos do estudo.

2 Revisão de Literatura

O termo *deepfake* consiste numa combinação entre “*Deep Learning*” (termo em inglês para “Aprendizado Profundo” que nomeia um subdomínio do “Aprendizado de Máquina”) e “*fake*” (que designa a fabricação artificial dos dados). Seu uso designa uma técnica de Inteligência Artificial empregada para gerar/modificar documentos de imagem, vídeo e voz (GONG *et al.*, 2020). As *deepfakes* tem potencial de moldar a percepção da realidade, quando usada para disseminar documentos audiovisuais gerados/modificados artificialmente como captações reais de imagem/voz de pessoas. De acordo com Greengard (2020, p. 18, tradução nossa), a propagação da

manipulação dos elementos multimídia através desta técnica é desafiadora e tem potencial de ser “[...] a próxima fronteira nas guerras de propaganda.”.

O uso da técnica por usuários de plataformas digitais tem crescido devido a disponibilidade para *download* de programas que podem produzir *deepfakes*. O que, para Greengard (2020), desperta preocupação da sociedade civil pelo uso da técnica para manipular processos políticos e eleitorais, falsificar vídeos pornográficos, afetar relações comerciais, dificultar a identificação de crimes etc. Aliada às práticas de desinformação e à velocidade de circulação de informação em plataformas digitais, a expansão de *deepfakes* tem potencial de reduzir a credibilidade de documentos audiovisuais autênticos com prejuízo à percepção da realidade (GREENGARD, 2020).

Entre os usos de Inteligência Artificial, ou mais especificamente de “Aprendizado de Máquina” (*Machine Learning*), para manipulação audiovisual destaca-se o subdomínio de “Aprendizado Profundo” (*Deep Learning*). Nadeem e outros (2019, p. 84003, tradução nossa) expuseram que os métodos de *Deep Learning* demonstraram eficácia em processar e analisar mídias visuais por terem “[...] a capacidade de extrair e aprender informação abstrata [...]” e, assim, “[...] eliminam a necessidade de recursos manuais para extração e representação [...]” de dados. Os autores (NADEEM et. al., 2019) pontuaram que o treinamento de uma rede neural profunda demanda tempo e especialistas do domínio para realizar esse treinamento. Assim como explicitaram que, quando treinados, os métodos *Deep Learning* têm poder de processamento de um número elevado de dados em segundos. (NADEEM et. al., 2019)

No *Deep Learning*, as “Redes Adversariais Generativas” (*Generative Adversarial Networks*, GAN) se destacam, de acordo com Greengard (2020), devido ao realismo da mídia gerada artificialmente. As GAN consistem em uma classe de modelos *Deep Learning*, e consequentemente *Machine Learning*, que possuem duas redes neurais: uma discriminadora, responsável por diferenciar dados reais de dados falsos; e uma geradora, voltada para gerar dados falsos que imitam os reais.

De acordo com Taheri e outros (2020), durante o treinamento de uma GAN um “vetor de ruído”, que parte de uma distribuição de dados reais, é recebido pela rede geradora responsável por gerar uma amostra com as mesmas dimensões dos dados reais, porém, sintetizada. A rede

discriminadora, por sua vez, consiste em “[...] um classificador binário que leva tanto amostras reais quanto sintetizadas como entrada e calcula a probabilidade de que seja real e não falsa [...]”. (TAHERI *et al.*, 2020, p. 4, tradução nossa).

Essas redes neurais são treinadas com uso de grandes conjuntos de imagens/vídeos reais e artificiais (formadas por *pixels* representados por valores numéricos específicos) oriundos de conjuntos de dados selecionadas. Nesse processo, um problema matemático de otimização característico é resolvido para que a rede neural voltada para geração de dados falsos consiga criar dados o mais próximo possível de uma aparência realista. Os vídeos gerados com GAN tem cada vez mais aparência de realidade, mas são simulacros que necessitam de captação de imagem e de som provenientes de documentos reais. Desse modo, Lyu (apud GREENGARD, 2020, p. 18-19, tradução nossa) afirmou que “[...] existem pistas e artefatos que geralmente podem ser detectados. Em muitos casos, há distorções de imagem, inconsistência em iluminação, suavidade em áreas e formação incomum de *pixels*.”.

Os documentos audiovisuais gerados com GAN podem ainda ser combinados com outros mecanismos para manipular atributos específicos e amplificar seu realismo, com potencial para desatualizar técnicas de sua detecção (LI, D. *et al.*, 2021). Assim, disputas entre algoritmos de pesquisas para geração de *deepfakes* cada vez mais realistas e de pesquisas para sua detecção tendem a se perpetuar (GREENGARD, 2020). O contexto evidencia disputas onde organizar as produções técnico-científicas voltadas à detecção de *deepfakes* se torna elemento estratégico e desafiador, devido ao aporte de informação disseminada na *web* e a velocidade com que as técnicas são modificadas e novos dispositivos tecnológicos são gerados/aperfeiçoados. Para este fim, o presente artigo aciona a Organização do Conhecimento, a Terminologia e a Arquitetura da Informação enquanto base teórica para proceder o mapeamento terminológico.

A Organização do Conhecimento, como frisou Maimone (2020), possui intensa relação com a Linguística e a Terminologia mediante os Sistemas de Organização do Conhecimento (SOC). Nesse contexto, o conhecimento “[...] é construído individualmente e advém da elaboração mental subjetiva do indivíduo [...]”, já a informação “[...] é de natureza social, circula entre pessoas e grupos, funciona como fluxo e permite a comunicação”. Assim, a Organização do Conhecimento tem por objeto de estudo os “[...] processos que possibilitem o acesso aos registros do

conhecimento (documentos), a fim de que se tenha uma recuperação informacional adequada dos itens buscados” (MAIMONE, 2020, p. 424).

O aporte teórico e prático da Organização do Conhecimento remete às civilizações antigas na invenção da escrita, em processos de catalogação, registro de livros, classificação de acervos de bibliotecas no decorrer das Eras etc. Com o acentuado desenvolvimento de tecnologias do século XIX a informação e o conhecimento registrado pela escrita dos livros encontraram outros meios de circulação e registro. Nesse contexto, também se destacaram a Classificação Decimal de Melvil Dewey (1876) e a Classificação Decimal Universal de Paul Otlet e Henri La Fontaine (1905). (GOMES, 2017).

O século XIX foi marcado pelo acentuado desenvolvimento de tecnologias como “[...] fotografia (1826), telefone (1876), filme fotográfico (1879), telégrafo sem fio (1890), rádio (1890), cinematógrafo (1895) [...]” (GOMES, 2017, p. 46). Nesse contexto, a autora destacou que emergiram também a Classificação Decimal de Melvil Dewey (1876) e a Classificação Decimal Universal de Paul Otlet e Henri La Fontaine (1905) já adentrando ao século XX.

A Classificação Facetada foi idealizada por Shiyali Ramamrita Ranganathan na *Colon Classification* (1933) reunindo as Categorias Fundamentais (personalidade, matéria, energia, espaço e tempo), que foram posteriormente detalhadas por Brian Vickery evidenciando a importância da obra de Ranganathan para construção de taxonomias e de interfaces de navegação na *web*. A Teoria do Conceito de Ingetraut Dahlberg tem grande importância para a Ciência da Informação e se mostrou compatível ao método de Facetas, contendo “[...] elementos para identificar um conceito como resultado da análise do referente, o que nos leva à definição. Esta, segundo Dahlberg, tem natureza estruturante, produzindo hierarquias (taxonomias) e contribuindo igualmente para as ontologias”. Dahlberg adotou uma abordagem onomasiológica para identificar e descrever elementos “verdadeiros” a fim de estruturar a “[...] representação de um domínio, de uma técnica, ou de qualquer atividade” (GOMES, 2017, p. 59).

Por abordagem onomasiológica entende-se o que, em Baldinger (2001, p. 29), é também nomeado de “campo das designações”. Nesse campo, o linguista e filólogo afirmou que é “[...] a designação que parte do conceito para atingir a forma (nome)”. A onomasiologia tem grande importância para estudos terminológicos, pois consiste em uma perspectiva que é interdependente

à abordagem semasiológica, “campo das significações”, em que se “[...] parte da forma (nome) para atingir o conceito [...]” (BALDINGER, 2001, p. 29). De acordo com o autor:

A estrutura **onomasiológica** é baseada na sinonímica, a estrutura **semasiológica** é baseada na polissemia. A onomasiologia visualiza os problemas sob o ângulo do que **fala**, daquele que deve escolher entre diferentes meios de expressão. A semasiologia focaliza os problemas sob o ângulo do que **ouve**, do interlocutor que deve determinar a significação da palavra que ele entende dentre todas as significações possíveis (BALDINGER, 2001, p. 30, grifo do autor).

Ambas as abordagens têm sua importância nas discussões teórico-metodológicas. O presente estudo não pretende se aprofundar nessa dimensão, mas apenas sublinhar a importância de ambas as abordagens, enquanto pontos de vista distintos e relevantes para uma leitura mais aproximada do domínio em análise. No entanto, sem adentrar em uma perspectiva universalista do que é considerado elemento “verdadeiro”.

A perspectiva de discurso científico aqui adotada se aproxima de Capurro e Hjørland (2007, p. 149) que expuseram que “[...] conceitos teóricos não são elementos verdadeiros ou falsos ou reflexos de algum outro elemento da realidade; em vez disso, são construções planejadas para desempenhar um papel, da melhor maneira possível”. Em outro trabalho, Hjørland (2007) frisou que um conceito é definido para expressar um significado para evitar potenciais ambiguidades das palavras e é representado por um descritor. Os descritores são “organizados” com base em terminologias padronizadas a fim de que se tenha um controle vocabular. Nesse prisma, um tesouro por exemplo consiste em “[...] uma seleção de conceitos complementados com informações sobre suas relações semânticas (como relações genéricas ou ‘relações associativas’) [...]”, que consiste em um Sistema de Organização do Conhecimento (SOC) (HJØRLAND, 2007, p. 367, tradução nossa).

Os SOC puderam ser dispostos como: listas de termos (arquivos de autoridade, glossários, dicionários e os específicos *gazetteers* que são dicionários geoespaciais), classificações e categorias (títulos de assunto, esquemas de classificação, taxonomias e esquemas de categorização) e listas de relacionamento (tesauros, redes semânticas e ontologias). (HODGE, 2000 apud HJØRLAND, 2007; SOUZA, et al., 2005). Para Hjørland (2007, p. 369, tradução nossa) existem outros tipos de SOC não considerados na taxonomia de Hodge, como enciclopédias, bases

de dados bibliográficos etc. ou ainda “[...] em um sentido amplo e geral estão relacionados à organização de literaturas, tradições, disciplinas e pessoas em diferentes culturas”.

Mazzochi (2018) concordou com a crítica de Hjørland acerca da perspectiva simplista que abordagens objetivistas da ciência adotam ao inferir neutralidade no processo de construção do conhecimento científico. A crítica endossa que o processo de construção de um SOC exige que em algum momento decisões sejam tomadas e ocorram ações como hipóteses de organização, escolha de fontes, periódicos, bases teóricas etc. Nesse sentido, o autor discorreu que os SOC são “entidades classificatórias” constituídas por termos, conceitos e relações inseridas em contextos teóricos, históricos e culturais específicos.

Porém, a existência de estruturas semânticas oriundas de diferentes bases teóricas, áreas temáticas e interesses pragmáticos não se configura em uma equiparação sem considerar seus contextos e suas potencialidades. Nesse sentido, a lógica tem sua importância reconhecida ao lado do “perspectivismo” para a Organização do Conhecimento. Enquanto a lógica permite a identificação de linguagens e estruturas conceituais comuns/estáveis, o perspectivismo possibilita identificar informações contextuais (MAZZOCHI, 2018).

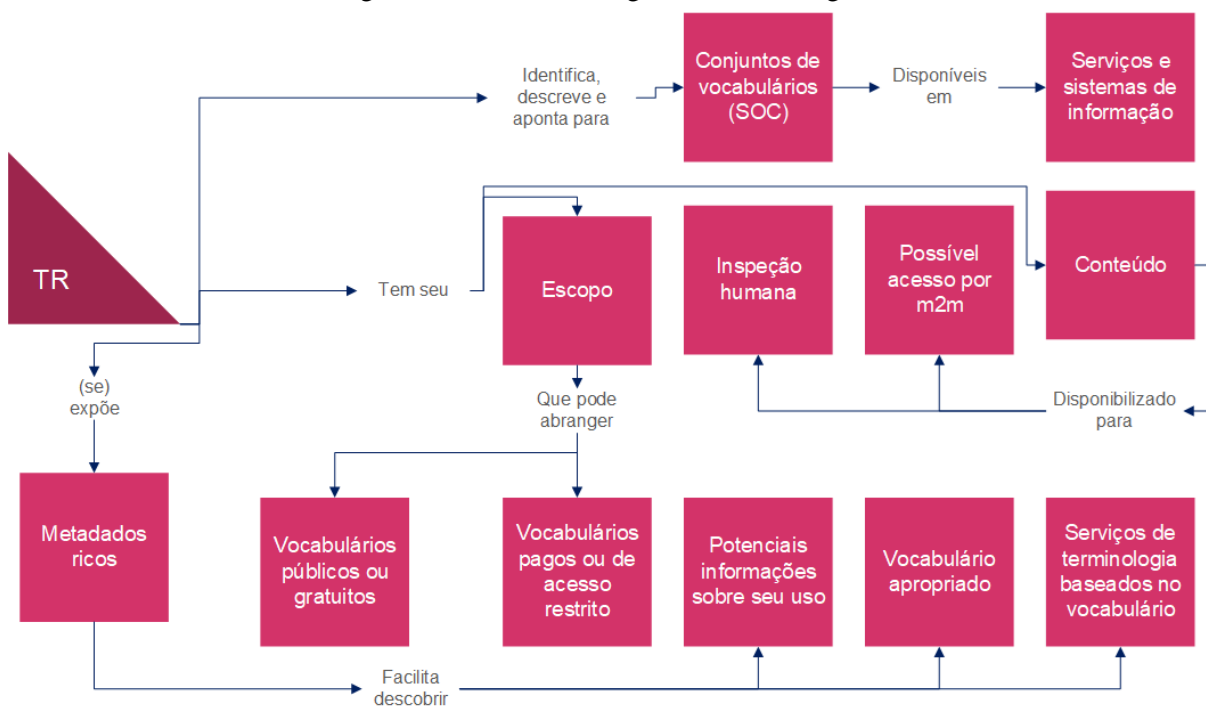
Em um domínio específico do conhecimento:

O trabalho com vocabulários controlados, [...] requer diálogo constante com a Terminologia, pois esta, além de auxiliar a comunicação especializada (técnica), [...] proporciona o trabalho com termos e conceitos unívocos em uma língua de especialidade, permitindo o direcionamento interpretativo da estrutura conceitual, fato que admite melhor representação e conseqüente recuperação dos conteúdos dos documentos [...] (MAIMONE, 2013, p. 85).

Desse modo, a proposição de conceituar *deepfakes* necessita da Terminologia com coleta e estudo dos termos do domínio. Na Terminologia voltada para a *web* com possibilidades de muitos SOC (com vasta potencialidade de aplicação), Golub *et al.* (2014) expuseram que os registros de terminologia (*Terminology Registries*, TR) foram empregados com padrões de *web* semântica. De acordo com as autoras, isso ocorreu para “[...] promover uma adoção mais ampla, padronização e interoperabilidade geral de metadados, facilitando reuso, harmonização e sinergia entre diversas disciplinas e comunidades de práticas [...]” (GOLUB *et al.*, 2014, p. 1901, tradução nossa). O fluxograma da figura 1 sintetiza o contexto dos registros terminológicos.

Os registros terminológicos (TR) na figura 1 foram organizados a partir de seu escopo, conteúdo e ação/potência: a abrangência do escopo envolve vocabulários públicos/gratuitos ou vocabulários pagos/restritos; seu conteúdo é disponibilizado tanto para inspeção humana quanto potencial acesso entre máquinas (m2m); sua ação em serviços e sistemas de informação consiste na disponibilização de conjuntos de vocabulários (SOC), para os quais os registros identificam, descrevem e apontam termos; e podem facilitar a descoberta de vocabulários apropriados, seu uso e serviços nele baseados caso forneçam metadados ricos (GOLUB *et al.*, 2014).

Figura 1 – Síntese dos registros terminológicos



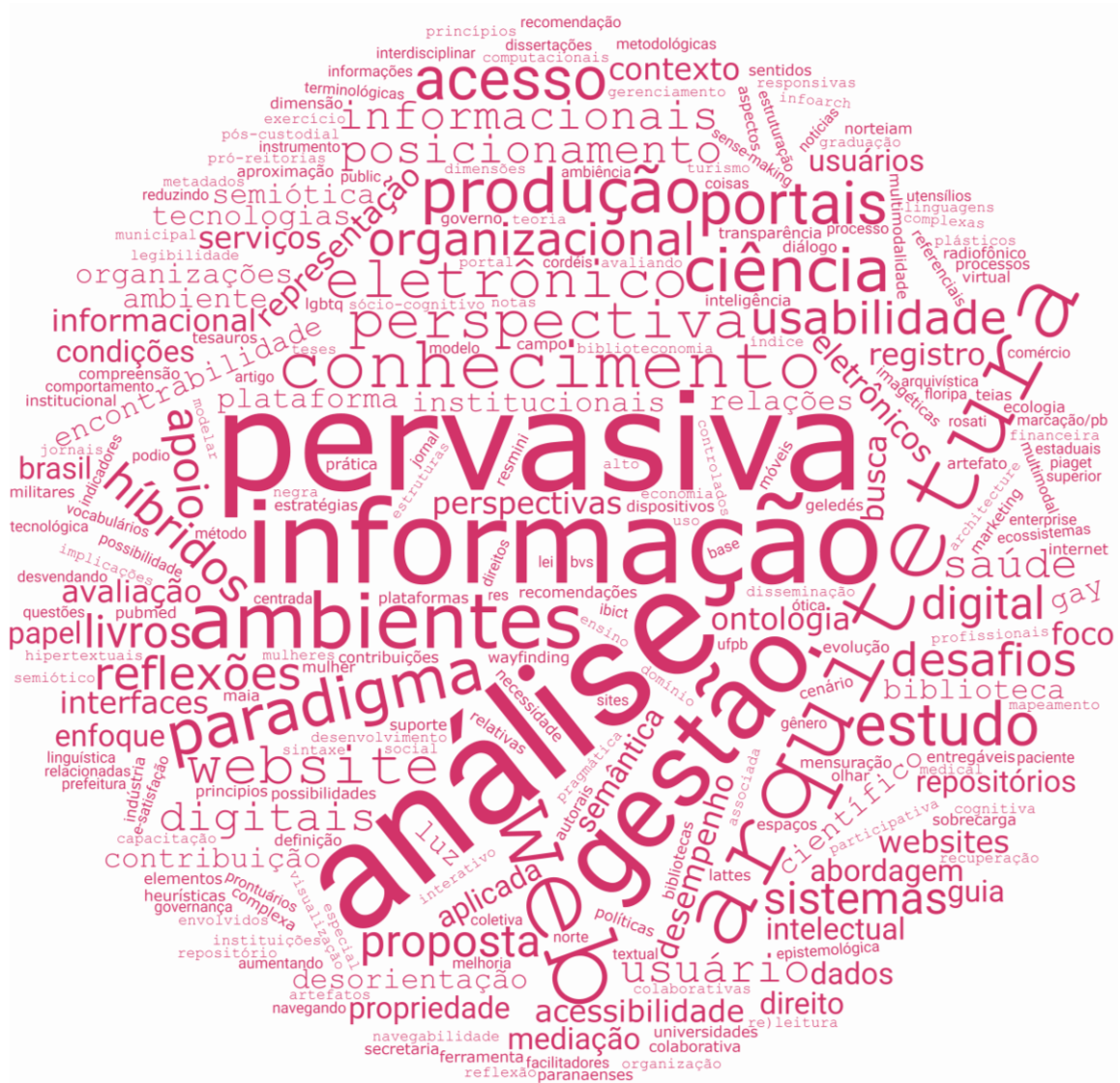
Fonte: Elaboração autoral com base em Golub *et. al.* (2014, p. 1903)

Quando os TR são aplicados a documentos multimídia verifica-se uma complexificação no processo de identificação e descrição diante da intensa produção de tecnologias para a *web*, conteúdo digital e sua veiculação em plataformas digitais. O gerenciamento desse conteúdo multimídia, conforme exposto por Lemos e Souza (2020, p. 301, tradução nossa), “[...] tornou-se caótico sem o suporte adequado de tecnologias para armazenamento, organização e recuperação de informação”.

A natureza multimídia – com “[...] vários tipos de objetos como vídeos, textos, sons, imagens, modelos 3D etc. [...]” – é complexa e de ampla potencialidade de segmentação em outras mídias como de vídeo para imagem, texto para som, imagem e som para vídeo etc. (LEMOS; SOUZA, 2020, p. 301, tradução nossa). A partir desse grandioso volume de documentos multimídia produzidos e veiculados na *web* as *deepfakes* são geradas e disseminadas. Esse fenômeno torna necessária a investigação de estratégias capazes de atender aos objetivos de um domínio *web* em constante modificação, como é o caso da detecção de *deepfakes*.

Outra disciplina que juntamente com a Organização do Conhecimento e a Terminologia aplicada no presente estudo é a “Arquitetura da informação” (AI). De acordo com Rocha, Pinto e David (2020), esse termo foi cunhado pelo arquiteto Richard Saul Wurman em 1975 em reação ao grande volume de informação que o contexto tecnológico já gerava na época. Porém, foi a partir da produção *Information Architecture for the World Wide Web* em 1998, dos bibliotecários Louis Rosenfeld e Peter Morville que a disciplina se consolidou. O estado-da-arte da Arquitetura da Informação em bases de dados da Ciência da Informação foi identificado por Rocha, Pinto e David (2020), a partir da qual foi elaborada a figura 2.

Figura 2 – Nuvem sobre Arquitetura da Informação em bases de dados da Ciência da Informação



Fonte: Elaborado com *wordclouds.com* referenciado em Rocha, Pinto e David (2020)

Os autores demonstraram através de uma revisão de literatura integrativa que o contexto de estudos da AI ultrapassou os domínios *web*. Contendo 73 artigos, a pesquisa indicou o caráter multidisciplinar/interdisciplinar da área e sua relação com a *web* (sejam ambientes digitais, estudos de uso/usuários, ontologias, metadados etc.). Porém, não se restringem ao meio digital visto que se relacionam à “ecologia da informação”, um conceito que em si evidencia uma relação entre “usuários”, “conteúdo” e “contexto” (ROSENFELD; MORVILLE; ARANGO, 2015 apud

ROCHA; PINTO; DAVID, 2020). Como exemplo dessa ampla relação entre usuários, conteúdo e contextos, alguns indícios da “ecologia da informação” podem ser visualizados na figura 2.

A nuvem exibida permite identificar distintos enfoques voltados: à usuários, que foram identificados pelas palavras *gay*, *mulheres*, *militares*, *paranaenses*, *encontrabilidade*, *acessibilidade* etc.; ao conteúdo, com a identificação dos termos *dados*, *livros*, *websites*, *metadados*, *ontologias* etc.; e aos contextos, manifestos em termos como *web*, *universidades*, *repositórios*, *biblioteca*, *organizações*, *prefeitura* etc. Ressalta-se também que os temas dos estudos evidenciaram um acentuado caráter epistemológico, que não será aqui aprofundado.

Respeitadas as particularidades da Organização do Conhecimento, da Terminologia e da Arquitetura da Informação, alguns elementos comuns podem ser notados por auxiliarem na análise de domínios *web*. Eles indicam um caminho mediante o perspectivismo que caracteriza um domínio específico com suas particularidades, contextos e necessidades pragmáticas. Um processo no qual a univocidade de linguagem não possui um caráter universal, mas é possível apenas para representar pragmaticamente como uma comunidade de especialidade faz uso de termos/conceitos.

3 Procedimentos metodológicos

A base lógica da presente pesquisa, de acordo com as definições metodológicas em Gil (2008), consiste no método indutivo para descobrir as relações dos fenômenos investigados e identificar termos/conceitos do domínio de detecção de *deepfakes*. Um processo que se realiza mediante pesquisa exploratório-descritiva de abordagem qualitativa aplicada. Através de base teórica da Organização do Conhecimento na Ciência da Informação estabelece-se diálogo interdisciplinar com a Terminologia e a Arquitetura da Informação para propor caminhos de um domínio *web* e para a possibilidade de mapeamento de detecção de *deepfakes*.

Para o levantamento bibliográfico de publicações de periódicos revisados por pares foram selecionadas três bases de dados consolidadas: *Web of Science*, *Scopus (Elsevier)* e *Digital Library of Association for Computing Machinery*. Nesse processo buscou-se os termos “*generative adversarial network*” AND “*deep*fake*” AND “*detection*” em todo o conteúdo de artigos de periódicos em inglês, português ou espanhol de 2016 até 2021. A seleção dos termos para a busca

ocorreu após leitura preliminar sobre o uso de técnicas automatizadas em que as *deepfakes* (e seu principal método *Deep Learning* para geração, as *Generative Adversarial Networks*) se destacaram. O que foi complementado com o termo voltado para recuperar trabalhos que contemplassem a detecção da supracitada técnica. A fim de realizar o mapeamento terminológico do domínio, foi construído o Quadro 1:

Quadro 1 – Modelo de ficha para mapeamento terminológico

Observações	Referências			Definição síntese	Referências usadas
	Referência 1	...	Referência N		
	Áreas de conhecimento da referência				
	Áreas de conhecimento da Referência 1	...	Áreas de Conhecimento da Referência N		
Termo candidato	Contexto	Contexto	Contexto		
Termo candidato 1	Contexto do termo 1	...	Contexto termo 1	Explicação síntese do termo 1	AUTOR, data
...
Termo candidato N	0	Contexto do termo N	0	Explicação síntese do termo N	AUTOR, data

Fonte: Elaborado pelos autores

Nesse contexto, os termos candidatos consistem naqueles selecionados (ou preferidos) por serem considerados mais adequados ao trabalho que se pretende e, estão organizados em ordem alfabética. As áreas de conhecimento da referência indicam o campo de saber a que pertence o documento/referência identificadas a partir das biografias de autoras(es) presentes nos mesmos. Os contextos são citações diretas organizadas em colunas por referência consultada. A definição síntese é uma explicação sucinta dos termos/conceitos, e por fim, são expostas as referências usadas, visto que nem todos os termos candidatos estão presentes em todas as referências. No caso de ausência de contexto ao termo, à célula no quadro foi inserido o valor 0. A partir da informação coletada, um quadro simples contendo termo candidato, definição síntese, referência usada e área de conhecimento do documento recuperado forma a lista de termos no Apêndice.

A presente pesquisa fez uso dos seguintes materiais: um *Personal Computer* com sistema operacional Windows 11 e os softwares *Microsoft Excel 2019* e *Microsoft Word 2019*.

4 Resultados

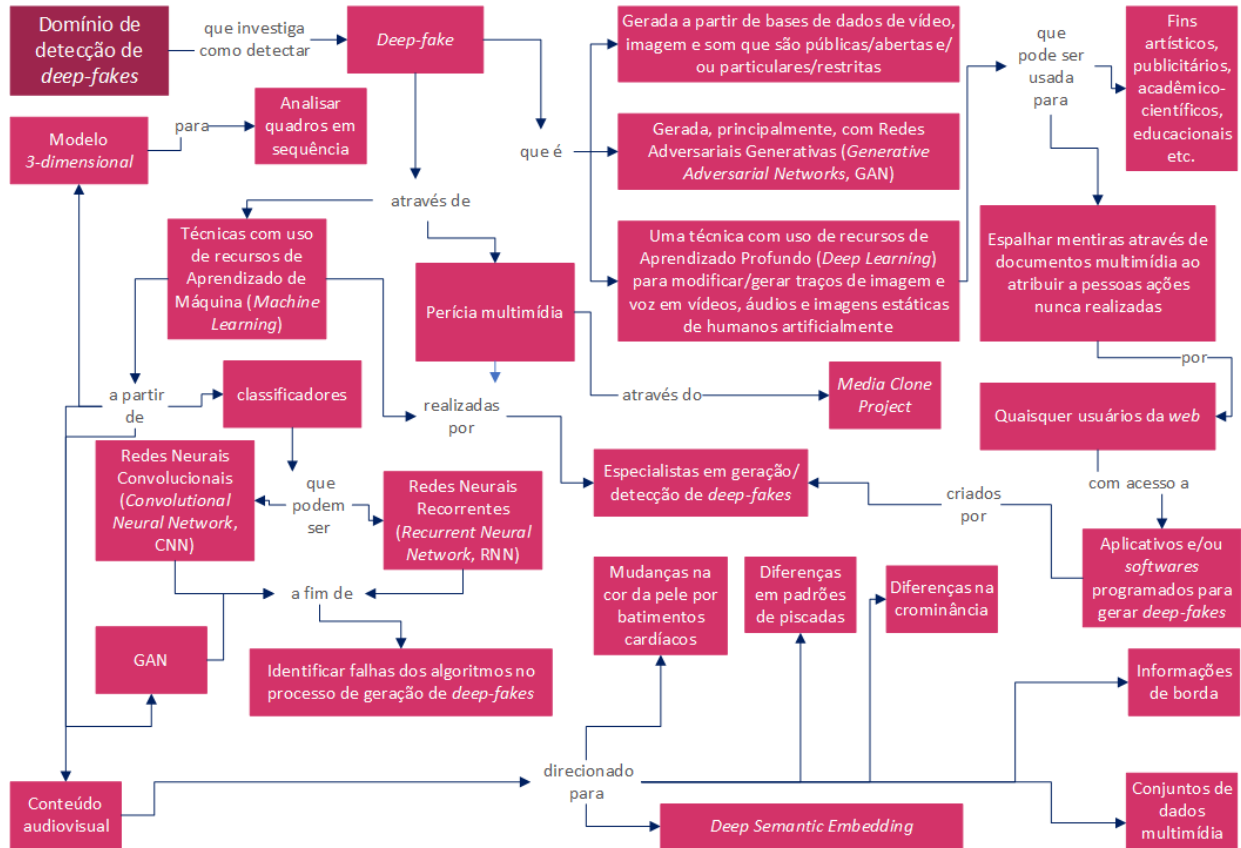
A partir do levantamento realizado nas bases de dados, foram obtidos 35 resultados (todos em inglês). Destes, 22 artigos pertinentes ao domínio foram recuperados, sendo 2 que expunham o contexto das *deepfakes* (GREENGARD, 2020; LI, D., et al., 2021). Os outros 20 continham estado-da-arte do domínio e/ou propostas de detecção automatizada ou pericial da técnica (de *deepfakes* geradas a partir de GAN) em documentos multimídia. A partir das referências recuperadas, o mapeamento do domínio foi organizado no fluxograma da figura 3.

O estado-da arte da detecção de *deepfakes* foi objeto de discussão em 3 artigos (GONG, et al., 2020; NEVES, et al., 2020; TOLOSANA, et al., 2020). Em comum, foram identificados métodos da CC baseados predominantemente: em recursos de *Deep Learning* e perícia multimídia; com classificadores *Convolutional Neural Network* (CNN) e *Recurrent Neural Network* (RNN); a partir de bases de dados próprias e/ou públicas. Entre as propostas de detecção automatizada recuperadas, 8 artigos (BAEK; YOO; BAE, 2020; CHINTHA, et al., 2020; GUARNERA; GIUDICE; BATTIATO, 2020; GUO, et al., 2021; HASHMI, et al., 2020; KOHLI; GUPTA, 2021; MI, et al., 2020; TAHERI, et al., 2020) indicavam uso de GAN, combinações com outros recursos de *Deep Learning* e classificadores CNN e RNN para identificar falhas dos algoritmos no processo de geração de *deepfakes*.

Os outros 9 artigos se relacionavam com recursos *Deep Learning*, mas focavam: em perícia multimídia com base no projeto de pesquisa *Media Clone* (ECHIZEN, et al., 2021); em análise automatizada de conteúdo multimídia (NADEEM, et al., 2019); em modelo *3-dimensional* para análise de quadros em sequência (NGUYEN, et al., 2021); ou em detecção a partir do conteúdo audiovisual. As pesquisas de detecção de *deepfakes* a partir do conteúdo audiovisual se direcionavam para: mudanças na cor da pele por batimentos cardíacos (QI et al. 2020); diferenças em padrões de piscadas (JUNG; KIM; KIM, 2020); diferenças na cromaticidade (LI, H., et al., 2020);

informações de borda (ZHANG, et al., 2019); conjuntos de dados multimídia (LI, X., et al., 2020); e semântica embedding (FERNANDO, et al., 2020).

Figura 3 – Mapeamento do domínio de detecção de *deepfakes*



Fonte: elaboração autoral a partir do referencial recuperado

Em ambiente controlado, muitos dos métodos expostos alcançaram sucesso acima de 90% na detecção de *deepfakes* em documentos audiovisuais. Um resultado que ao ser confrontado com a célere evolução de técnicas *deepfake* exige constante inovação metodológica, introdução de novos métodos e/ou combinação de múltiplos métodos para detecção de *deepfakes*. O mapeamento do domínio a partir do referencial exposto resultaram em 79 termos com seus respectivos contextos e definições sínteses extraídas do referencial como mostra a lista de termos no Apêndice, da qual foi o extraído o Quadro 2.

Quadro 2 – Trecho de lista de termos do mapeamento do domínio

Termo candidato	Definição síntese	Referências	Áreas de Conhecimento
Deepfake Detection Challenge	Conjunto de dados de visualização lançado num desafio do <i>Facebook</i> com o mesmo nome contendo mais de 5000 vídeos, sendo mais de 1000 originais e mais de 4000 falsificados com diversificação etária, étnica e de gênero.	CHINTHA, et al., 2020; LI, X., et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
Deepfakes	Técnicas com uso de modelos de aprendizado de máquina capaz de gerar imagens, sons e vídeos artificialmente com efeitos realistas cada vez mais difíceis de serem detectados a olho nu e distinguidos de imagens naturais.	GONG, et al., 2020; GUARNERA; GIUDICE; BATTIATO, 2020; JUNG; KIM; KIM, 2020	Tecnologia da Informação e Comunicação; Ciência da Computação; Matemática
DeepForensics-1.0	Conjunto de dados em grande escala contendo diversificados vídeos falsos gerados com a técnica FaceSwap.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica
DeepNude	Aplicativo móvel capaz de gerar automaticamente imagens pornográficas a partir de inserção de uma fotografia feminina.	GONG, et al., 2020	Tecnologia da Informação e Comunicação

Fonte: Elaborado pelos autores

No processo de construção da lista de termos, muitos continham elementos explicativos e por isso não foi possível inseri-los. Outra questão identificada foram as explicações incompletas indicando elementos não explicitados nas produções, visto que muitos termos foram mencionados em pesquisa aplicada da Ciência/Engenharia da Computação e áreas afins, sem expor as

definições. Também foi verificado que o termo *Deepfake* não abarcava a totalidade de usos de *Generative Adversarial Networks* para manipulação/geração de documentos audiovisuais. Outrossim, ainda que seja predominante o uso de GAN para manipulação/geração de documentos com *Deepfakes*, outros termos como *Media Clone*, *Spoof Detection*, *Digital Media Forensics* etc. apontam para outros elementos passíveis de pesquisa bibliográfica para detecção de deepfakes a fim de compreender a abrangência do domínio.

As áreas de conhecimento mapeadas a partir das biografias referente à autoria informadas nos artigos indicaram uma predominante presença de profissionais de Ciência, Engenharia e Tecnologia da Computação. Nessa linha de observação, destaca-se a forte presença de outras Engenharias (Civil, Telecomunicações etc.) e Matemática na construção de proposições de detecção de fraudes ligadas ao uso de técnicas de Inteligência Artificial. Também foi identificada a participação de profissionais de áreas de Tecnologia, Segurança e Ciência da Informação, assim como de Ciências da Comunicação, Jornalismo e Relações Internacionais.

No uso dos documentos recuperados para a construção da lista de termos, as produções acadêmicas mais empregadas foram: Nadeem e outros (2019) para 11 termos candidatos; Guarnera, Giudice e Battiato (2020) para 10 termos; Chinthá e outros (2020) para 10 termos; Echizen e outros (2021) para 9 termos; e Tolosana e outros (2020) em 7 termos. Destes, foi observada a presença de biografias mais específicas (NADEEM, et al. 2019; GUARNERA; GIUDICE; BATTIATO, 2020; TOLOSANA, et al., 2020), delimitadas à Ciência da Computação e Matemática. Porém, outras manifestaram um recorte interdisciplinar que incluía profissionais de Tecnologias de Informação e Comunicação, Jornalismo e Relações Internacionais (CHINTHA, et al., 2020), Ciências de Segurança Social e de Dados (ECHIZEN, et al., 2021), por exemplo.

Em caráter preliminar foi possível mapear elementos relevantes para o domínio de detecção de manipulação/geração de documentos audiovisuais com técnicas automatizadas de *Machine Learning*. Em comum ao aporte bibliográfico analisado, processos de descrição e classificação a partir de conjunto de dados foram dispendidos para recuperar informações que possibilitassem a identificação dessas técnicas. Um processo em que o uso de algoritmos se verificou como relevante, mas insuficiente sem a ação humana para compreender os processos envolvidos, treinar

as redes neurais, gerar/indicar conjuntos de dados e possibilitar caminhos de resolução do problema.

5 Considerações finais

O aporte teórico-epistemológico interdisciplinar da Organização do Conhecimento, da Terminologia e da Arquitetura da Informação apresentado foi considerado suficiente para embasar o mapeamento terminológico realizado no presente estudo a partir da bibliografia levantada. Os resultados identificaram um domínio mais amplo de detecção de manipulação/geração automatizada de documentos textuais e audiovisuais digitais que necessita de aprofundamento para análise e conceituação; encontra-se circunscrito à Ciência da Computação e áreas afins.

O mapeamento terminológico-conceitual do domínio da detecção de *deepfakes* a partir dos documentos recuperados demonstrou-se parcialmente efetivo através da pesquisa bibliográfica, exploratório-descritiva e qualitativa realizada. De fato, foi possível identificar com certa amplitude elementos que constituem o domínio, expondo a partir da bibliografia levantada definições de termos candidatos e indícios de como se dá o processo de fazer científico da área explorada. Porém, os resultados apontaram para um domínio mais amplo do que apenas uso de técnicas *deepfake* com uso de GAN. Um conjunto de técnicas automatizadas e processos de detecção que envolvem outros saberes, e assim outras palavras-chave, que não foram contemplados no levantamento bibliográfico foram identificados e necessitam também serem aprofundados para um mapeamento terminológico mais completo e que possibilite conceituação e análise do domínio com mais aproximação.

Tendo em vista a dimensão do “perspectivismo” (MAZZOCHI, 2017), um caminho possível envolve agregar elementos contextuais na análise do domínio. Ou seja, compreender quais práticas, inclusive estruturais da realidade social em que o domínio se insere, estão atreladas aos processos de detecção de falsificações em documentos e, conseqüentemente, aos processos de manipulação/geração desses documentos. No entanto, sem abandonar a lógica do fazer científico que caracteriza o domínio estudado.

Acredita-se que os objetivos foram cumpridos, visto que a base teórico-metodológica empregada legitimou a pesquisa realizada, através da qual foi possível identificar lacunas e possibilidade de investigação de novas “fronteiras” do domínio, ou ainda, apontar a continuidade de levantamento de termos candidatos a uma arquitetura de informação diferente da observada a priori. Do mesmo modo, o processo de mapeamento identificou a área da Ciência da Computação como principal área do aporte do domínio em relação com outras áreas afins e/ou relacionadas nos recortes. Um processo que possibilitou considerar a Ciência da Informação, com suas relações interdisciplinares e transdisciplinares, um caminho para desbravar o domínio, seu aporte lógico-científico e seus contextos acadêmicos, tecnológicos e sociais.

Referências

- Araújo, Carlos Alberto Ávila. “Novos desafios epistemológicos para a Ciência da Informação.” *Palavra Clave (La Plata)*, vol. 10, no. 2, Apr./Sep. 2021, pp. 1-15.
<https://www.palabraclave.fahce.unlp.edu.ar/article/view/PCe116/13767>. Acessado 06 jul. 2022.
- Baek, Jae-Young. Yoo, Young-Sang. Bae, Seung-Hwan. “Generative adversarial ensemble learning for face forensics.” *IEEE Access*, vol. 8, Mar. 2020, pp. 45421-45431.
<https://doi.org/10.1109/ACCESS.2020.2968612>. Acessado 06 jul. 2022.
- Baldinger, Kurt. “Semasiologia e onomasiologia.” *ALFA: Revista de Linguística*, vol. 9, 2001, pp. 7-36.
<https://periodicos.fclar.unesp.br/alfa/article/view/3265>. Acessado 06 jul. 2022.
- Bradshaw, Samantha. Howard, Philip N. “The global disinformation order: 2019. Global inventory of organized social media manipulation.” *The Computational Propaganda Project at the Oxford Internet Institute*, 2019, pp. 1-23. <https://demtech.oii.ox.ac.uk/wp-content/uploads/sites/93/2019/09/CyberTroop-Report19.pdf>. Acessado 06 jul. 2022.
- Capurro, Rafael. Hjørland, Birger. “O conceito de informação.” *Perspectivas em Ciência da Informação*, vol. 12, no. 1, Jan./Apr. 2007, pp. 148-207.
<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54/47>. Acessado 06 jul. 2022.
- Chintha, Akash, et al. “Recurrent convolutional structures for audio spoof and video deepfake detection.” *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 5, Aug. 2020, pp. 1024-1037.
<https://doi.org/10.1109/JSTSP.2020.2999185>. Acessado 06 jul. 2022.

- Echizen, Isao, et al. "Generation and Detection of media clones." *IEICE Transactions on Information and Systems*, vol. 104, no. 1, Jan. 2021, pp. 12-23. <https://doi.org/10.1587/transinf.2020MUI0002> Acessado 06 jul. 2022.
- Fernando, Tharindu, et al. "Detection of fake and fraudulent faces via neural memory networks." *IEEE Transactions on Information Forensics and Security*, vol. 16, Dec. 2020, pp. 1973-1988. <https://doi.org/10.1109/TIFS.2020.3047768> Acessado 06 jul. 2022.
- Froehlich, Thomas J. "Ten lessons for the age of disinformation." *Navigating fake news, alternative facts, and misinformation in a post-truth world*, 2020, pp. 36-88. IGI Global. <https://www.igi-global.com/book/navigating-fake-news-alternative-facts/236998>. Acessado 06 jul. 2022.
- Gil, Antônio Carlos. *Métodos e técnicas de pesquisa social*. Atlas, 2008.
- Golub, Koraljka, et al. "Terminology registries for knowledge organization systems: Functionality, use, and attributes." *Journal of the Association for Information Science and Technology*, vol. 65, no. 9, 2014, pp. 1901-1916. Wiley, <https://doi.org/10.1002/asi.23090>. Acessado 06 jul. 2022.
- Gomes, Hagar Espanha. "Marcos históricos e teóricos da organização do conhecimento." *Informação & Informação*, vol. 22, no. 2, 29 Oct. 2017, pp. 33-66. <http://dx.doi.org/10.5433/1981-8920.2017v22n2p33>. Acessado 06 jul. 2022.
- Gong, Dafeng, et al. "Deepfake forensics, an ai-synthesized detection with deep convolutional generative adversarial networks." *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 3, 25 Jun. 2020, pp. 2861-2870. Warse, <https://doi.org/10.30534/ijatcse/2020/58932020>. Acessado 06 jul. 2022.
- Greengard, Samuel. "Will deepfakes do deep damage?" *Communications of the ACM*, vol. 63, no. 1, Jan. 2020, pp. 17-19. <https://dl.acm.org/doi/10.1145/3371409> Acessado 06 jul. 2022.
- Guarnera, Luca. Giudice, Oliver. Battiato, Sebastiano. "Fighting deepfake by exposing the convolutional traces on images." *IEEE Access*, vol. 8, Sep. 2020, pp. 165085-165098. <https://doi.org/10.1109/ACCESS.2020.3023037> Acessado 06 jul. 2022.
- Guo, Zhiqing, et al. "Fake face detection via adaptive manipulation traces extraction network." *Computer Vision and Image Understanding*, vol. 204, 21 Jan. 2021, pp. 1-12. *Scopus (Elsevier)*, <https://doi.org/10.1016/j.cviu.2021.103170> Acessado 06 jul. 2022.
- Hashmi, Mohammad Farukh, et al. "An exploratory analysis on visual counterfeits using Conv-LSTM hybrid architecture." *IEEE Access*, vol 8, May 2020, pp. 101293-101308. <https://doi.org/10.1109/ACCESS.2020.2998330>. Acessado 06 jul. 2022.

- Hjørland, Birger. "Semantics and knowledge organization." *Annual Review of Information Science and Technology*, vol. 41, 2007, p. 367-405. WILEY, <https://doi.org/10.1002/aris.2007.1440410115> Acessado 06 jul. 2022.
- Jung, Tachkyun. Kim, Sangwon. Kim, Keecheon. "DeepVision: deepfakes detection using human eye blinking pattern." *IEEE Access*, vol. 8, Apr. 2020, pp. 83144-83154. <https://doi.org/10.1109/ACCESS.2020.2988660> Acessado 06 jul. 2022.
- Kemp, Simon. "Digital 2021: Global Overview Report." *Data Reportal*, We Are Social, Hootsuite, 27 Jan. 2021. KEPIOS, <https://datareportal.com/reports/digital-2021-global-overview-report> Acessado 06 jul. 2022.
- Kohli, Aditi. Gupta, Abhinav. "Detecting deepfake, FaceSwap and Face2Face facial forgeries using frequency CNN." *Multimedia Tools and Applications*, 18 Feb. 2021, pp. 1-18. Springer, <https://doi.org/10.1007/s11042-020-10420-8> Acessado 06 jul. 2022.
- Lemos, Daniela Lucas da Silva. Souza, Renato Rocha. "Knowledge organization systems for the representation of multimedia resources on the web: a comparative analysis." *Knowledge Organization*, vol. 47, no. 4, 2020, pp. 300-319. Nomos e-Library, <https://doi.org/10.5771/0943-7444-2020-4-300>. Acessado 06 jul. 2022.
- Li, Defang, et al. "A novel attribute-based generation architecture for facial image editing." *Multimedia Tools and Applications*, vol. 80, no. 4, 1 Feb. 2021, pp. 4881-4902. <https://doi.org/10.1007/s11042-020-09858-7>. Acessado 06 jul. 2022.
- Li, Haodong, et al. "Identification of deep network generated images using disparities in color components." *Signal Processing*, vol. 174, Sep. 2020, pp. 1-12. Scopus (Elsevier), <https://doi.org/10.1016/j.sigpro.2020.107616>. Acessado 06 jul. 2022.
- Li, Xiaodan, et al. "Sharp multiple instance learning for deepfake video detection." *Proceedings of the 28th ACM International Conference on Multimedia*, 12 Oct. 2020. Association for Computing Machinery, <http://doi.org/10.1145/3394171.3414034>. Acessado 06 jul. 2022.
- Maimone, Giovana Deliberati. "Informação e conhecimento: insumos para o trabalho organizativo em ciência da informação." *Informação & Informação*, vol. 25, no. 3, 31 Oct. 2020, pp. 423-436. <https://doi.org/10.5433/1981-8920.2020v25n3p423>. Acessado 06 jul. 2022.
- Maimone, Giovana Deliberati. *Organização da informação e do conhecimento de documentos artísticos à luz da terminologia*, 2013, University of São Paulo, DSc dissertation. <https://doi.org/10.11606/T.27.2013.tde-23082013-112839>. Acessado 06 jul. 2022.

- Mazzocchi, Fulvio. “Knowledge organization system (KOS)”. *Knowledge Organization*, vol. 45, no. 1, 2018, pp. 54-78. Also available in ISKO Encyclopedia of Knowledge Organization, eds. Birger Hjørland and Claudio Gnoli, <http://www.isko.org/cyclo/kos>. Acessado 06 jul. 2022.
- Mi, Zhongjie, et al. “GAN-Generated image detection with self-attention mechanism against GAN generator defect.” *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 5, 14 May 2020, pp. 969-981. <https://doi.org/10.1109/JSTSP.2020.2994523>. Acessado 06 jul. 2022.
- Nadeem, Muhammad Shahroz, et al. “A Survey of deep learning solutions for multimedia visual content analysis.” *IEEE Access*, vol. 7, 24 Jun. 2019, pp. 84003-84019. <https://doi.org/10.1109/ACCESS.2019.2924733>. Acessado 06 jul. 2022.
- Neves, João C., et al. “GANprintR: improved fakes and evaluation of the state of the art in face manipulation detection.” *IEEE Journal of Selected Topics in Signal Processing*, vol. 14, no. 5, 6 Jul. 2020, pp. 1038-1048. <https://doi.org/10.1109/JSTSP.2020.3007250>. Acessado 06 jul. 2022.
- Nguyen, Xuan Hau, et al. “Learning spatio-temporal features to detect manipulated facial videos created by the deepfake techniques.” *Forensic Science International: Digital Investigation*, vol. 36, Mar. 2021, pp. 1-8. *Scopus (Elsevier)*, <https://doi.org/10.1016/j.fsidi.2021.301108>. Acessado 06 jul. 2022.
- Oxford Languages. *Oxford Word of the Year 2016*. Oxford University Press, 2016. <https://languages.oup.com/word-of-the-year/2016/>. Acessado 06 jul. 2022.
- Qi, Hua, et al. “DeepRhythm: exposing deepfakes with attentional visual heartbeat rhythms. Proceedings of the 28th ACM International Conference on Multimedia.” *Proceedings of the 28th ACM International Conference on Multimedia*, 12 Oct. 2020. *Association for Computing Machinery*, <https://doi.org/10.48550/arXiv.2006.07634>. Acessado 06 jul. 2022.
- Rocha, Carin Cunha. Pinto, Virgínia Bentes. David, Priscila Barros. “Arquitetura da informação: revisão integrativa em bases de dados de ciência da informação.” *Informação & Informação*, vol. 25, no. 2, Jul. 2020, pp. 49-73. <http://dx.doi.org/10.5433/1981-8920.2020v25n2p49>. Acessado 06 jul. 2022.
- Souza, Ligiane A., et al. “The Role of Gazetteers in Geographic Knowledge Discovery on the Web.” *Proceedings of the Third Latin American Web Congress*, 2005, pp. 157-165.
- Taheri, Shayan, et al. “Developing a robust defensive system against adversarial examples using generative adversarial networks.” *Big Data and Cognitive Computing*, vol. 4, no. 2, 22 May 2020, pp. 1-15. *MDPI*, <https://doi.org/10.3390/bdcc4020011>. Acessado 06 jul. 2022.
- Tolosana, Ruben, et al. “Deepfakes and beyond: a survey of face manipulation and fake detection.” *Information Fusion*, vol. 64, Dec. 2020, pp. 131-148. *Scopus (Elsevier)*, <https://doi.org/10.1016/j.inffus.2020.06.014>. Acessado 06 jul. 2022.
-
- Soella, Gabriel Meneguelli, and Maimone, Giovana Deliberali. Mapeamento da Detecção de *Deepfakes*: um trabalho terminológico. *Brazilian Journal of Information Studies: Research trends*, vol.16, publicação contínua 2022, e02153. DOI: 10.36311/1981-1640.2022.v16.e02153

- Vidotti, Silvana Aparecida Borsetti Gregorio, et al. “Web, Web Semântica e Web Pragmática: um posicionamento da Arquitetura da Informação.” *Informação & Sociedade: Estudos*, vol 29, no. 1, Jan./Mar. 2019, pp. 195-214. <https://periodicos.ufpb.br/ojs/index.php/ies/article/view/44358>. Acessado 06 jul. 2022.
- Zhang, Kejun, et al. “No one can escape: a general approach to detect tampered and generated image.” *IEEE Access*, vol. 7, 23 Sep. 2019, pp. 129494-129503. <https://doi.org/10.1109/ACCESS.2019.2939812>. Acessado 06 jul. 2022.

Apêndice

Quadro 3 – Lista completa de termos do domínio para elaboração do mapeamento realizado

Termo candidato	Definição síntese	Referências	Áreas de Conhecimento
100K-Faces	Conjunto de dados com 100.000 imagens sintéticas geradas a partir de StyleGAN.	NEVES, et al., 2020	Ciência e Engenharia da Computação; Engenharia de Telecomunicações
10-layer CNN network	Rede Neural Convolutacional (<i>Convolutional Neural Network, CNN</i>) que auxiliam na detecção de manipulação de cópia de movimento e emenda de imagem.	NADEEM, et al. 2019	Ciência da Computação
AlexNet	Rede Neural que foi treinada em dois dias, venceu a competição "2012 <i>ImageNet Competition</i> " e contém 650.000 neurônios e 60 milhões de parâmetros.	NADEEM, et al. 2019	Ciência da Computação
ATTGAN	Método de <i>Generative Adversarial Network</i> que atua na representação do espaço latente para estabelecer restrições aos atributos de rosto objeto de modificação.	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
Attribute manipulation	A manipulação de atributo ou edição de face modifica atributos como cor de pele/cabelo, sexo, idade, adicionar/remover óculos etc. que ocorre por meio de GAN e tem como exemplo o aplicativo FaceApp onde o público em geral podia fazer essas modificações de atributos.	TOLOSANA, et al., 2020	Ciência da Computação
Audio spoofing attacks Logical Access (LA)	Os ataques de falsificação de áudio por Acesso Lógico (<i>Logical Access, LA</i>) ocorre mediante uso de software de conversão/síntese de uma entrada de texto para uma saída de voz semelhante ao alvo da falsificação.	CHINTHA, et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
Audio spoofing attacks Physical Access (PA)	Os ataques de falsificação de áudio por Acesso Físico (<i>Physical Access, PA</i>) volta-se para sistemas de segurança detectores de voz em que uma fala gravada do alvo é reorganizada/reproduzida para enganar o sistema com dados reais do alvo.	CHINTHA, et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
Audio-to-Video and text-to-Video	Também conhecidos como <i>DeepFakes de Lip-sinc</i> , consiste na troca da expressão facial em vídeo a partir de áudio ou texto para que os movimentos da face pareçam falar o conteúdo desejado.	TOLOSANA, et al., 2020	Ciência da Computação

BigGANs dataset	Conjunto de dados proposto em 2018 considerado um dos modelos mais diversos e fidedignos de imagens integradas.	ZHANG, et al., 2019	Ciência e Tecnologia Eletrônica; Segurança da Informação; Segurança da Computação; Ciência e Tecnologia da Computação
Birajdar and Mankar taxonomy	A Taxonomia de Birajdar e Mankar consiste num conjunto de técnicas para falsificar imagens digitais, como clonagem, retoque etc.	NADEEM, et al., 2019	Ciência da Computação
CACIA 2.0 dataset	Conjunto de dados que contém 7491 imagens coloridas reais e 5123 modificadas em tamanhos e formatos diferentes, de diversas cenas e diferentes áreas adulteradas.	ZHANG, et al., 2019	Ciência e Tecnologia Eletrônica; Segurança da Informação; Segurança da Computação; Ciência e Tecnologia da Computação
Capsule-Forensics method	Método que usa redes de capsulas para detecção de ataques de repetição e imagens/vídeos gerados por computador.	CHINTHA, et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
CASIA-WebFace	Base de dados contendo 494.414 imagens de rostos de 10.575 de artistas do IMDb, com variações aleatórias de pose, iluminação, expressão facial e resolução.	NEVES, et al., 2020	Ciência e Engenharia da Computação; Engenharia de Telecomunicações;
DF	Conjunto de dados construído inicialmente de 400 vídeos reais do YouTube e 800 vídeos sintetizados com refinamentos para sanar inconsistências de cores, suavização de baixa frequência, oscilação temporal e síntese de falsificação de alta resolução. Conta com mais de 5600 vídeos do YouTube baseado em 59 celebridades e contém mais de 2 milhões de quadros de vídeo adulterados.	CHINTHA, et al., 2020; LI, X., et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
CGFace	Método baseado em <i>Deep Learning</i> direcionado para detecção de imagem em formato DNG onde uma rede neural convolucional (CNN) é personalizada para detecção de faces geradas por computador.	LI, H., et al., 2020	Processamento Inteligente da Informação; Segurança de Mídia; Inteligência Artificial e Robótica para a Sociedade

CNN	As Redes Neurais Convolucionais (<i>Convolutional Neural Networks</i>) são ferramentas de <i>Machine Learning</i> eficazes para análise de visão computacional e reconhecimento de padrões que aprende hierarquias espaciais de recursos de padrões de baixo e de alto nível.	KOHLI; GUPTA, 2021; MI, et. al., 2020; NADEEM, et al., 2019	Tecnologia da Informação; Ciência e Engenharia da Computação; Segurança da Informação; Segurança no Ciberespaço; Engenharia Elétrica, Eletrônica e de Informação; Sistemas de Informação e Comunicação;
CoALBP+LPQ	Método desenvolvido para detectar falsificação de faces cujo conjunto de recursos é composto por <i>Co-Ocurrence of Adjacent Local Binary Patterns</i> e <i>Local Phase Quantization</i> .	LI, H., et al., 2020	Processamento Inteligente da Informação; Segurança de Mídia; Inteligência Artificial e Robótica para a Sociedade
CoGAN	As Redes Adversariais Generativas Acopladas (<i>Coupled Generative Adversarial Networks</i>) faz uso de um par de GAN, de dois domínios diferentes, responsáveis por sintetizar imagens em seus respectivos domínios, com aplicação de restrição de compartilhamento de peso às camadas.	LI, D., et al., 2021	Ciência e Tecnologia da Computação; Matemática
Color filter array (CFA) analysis method	Método de análise de matriz de filtro de cores (<i>Color Filter Array</i>) que usa diferença entre padrões de matriz de filtro de cores para identificar imagens reais e falsas.	BAEK; YOO; BAE, 2020	Ciência da Computação; Engenharia da Computação
Convolutional Trace (CT)	O Traço Convolutivo é um vetor de características de Camadas de Convolução de Transposição utilizadas para gerar uma imagem a fim de identificar se a imagem consiste em uma DeepFake ou não.	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática;
Convolutional-LSTM	Rede Neural projetada para detecção de vídeos <i>DeepFake</i> que considera piscadas de olhos no processo.	NADEEM, et al., 2019	Ciência da Computação
COVERAGE dataset	Conjunto de dados com 100 pares de imagens original/sintetizada com seis tipos de falsificação e vários objetos semelhantes nas imagens originais a fim de aumentar a dificuldade de identificação pelo discriminador.	ZHANG, et al., 2019	Ciência e Tecnologia Eletrônica; Segurança da Informação; Segurança da Computação; Ciência e Tecnologia da Computação
CycleGAN	Duas <i>Generative Adversarial Networks</i> básicas conectadas por duas perdas de consistência de ciclo que mantém a informação linguística inalterada.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica

Deep Learning (DL)	Subdomínio de <i>Machine Learning</i> inspirado no sistema nervoso humano, que é eficaz para análise de conteúdo visual capaz de extrair recursos de alto nível para modelagem de conceitos abstratos. Predominantemente de métodos supervisionados, mas com abordagens semi-supervisionadas e sem supervisão. Seu treinamento é complexo e requer volume massivo de dados e alto poder de computação, mas ao serem treinados eliminam a necessidade de extração e representação manual de recursos e são capazes de processamento elevado de dados em poucos segundos.	NADEEM, et al., 2019	Ciência da Computação
Deep Stacked Auto-Encoder (SAE)	Rede Neural criada para detecção de manipulação de imagem em diferentes formatos de imagem em duas etapas, onde a primeira aprende recursos complexos no modelo SAE e a segunda identifica regiões adulteradas pelo aprendizado de contexto.	NADEEM, et al., 2019	Ciência da Computação
Deepfake Detection Challenge	Conjunto de dados de visualização lançado num desafio do <i>Facebook</i> com o mesmo nome contendo mais de 5000 vídeos, sendo mais de 1000 originais e mais de 4000 falsificados com diversificação etária, étnica e de gênero.	CHINTHA, et al., 2020; LI, X., et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
Deepfakes	Técnicas com uso de modelos de aprendizado de máquina capaz de gerar imagens, sons e vídeos artificialmente com efeitos realistas cada vez mais difíceis de serem detectados a olho nu e distinguidos de imagens naturais.	GONG, et al., 2020; GUARNERA; GIUDICE; BATTIATO, 2020; JUNG; KIM; KIM, 2020	Tecnologia da Informação e Comunicação; Ciência da Computação; Matemática;
DeepForensics-1.0	Conjunto de dados em grande escala contendo diversificados vídeos falsos gerados com a técnica FaceSwap.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica;
DeepNude	Aplicativo móvel capaz de gerar automaticamente imagens pornográficas a partir de inserção de uma fotografia feminina.	GONG, et al., 2020	Tecnologia da Informação e Comunicação

DeepVision	Método abrangente elaborado a partir de pesquisas de medicina, biologia, engenharia cerebral, aprendizado de máquina e algoritmos baseados em engenharia e estatística, que verifica integridade de vídeos a partir do rastreamento de mudanças nas piscadas de olhos em <i>DeepFakes</i> .	JUNG; KIM; KIM, 2020	Tecnologia da Informação; Ciência e Engenharia da Computação; Segurança da Informação
DENSENET discriminator	Rede Neural em que os mapas de recursos extraídos em camadas anteriores são conectados a camadas posteriores.	BAEK; YOO; BAE, 2020	Ciência da Computação; Engenharia da Computação
Dictionary learning based method	Método baseado em aprendizado de dicionário que treina um dicionário linear para captura de padrões de imagens de faces reais, seguido de tentativa de representação de imagens falsas com o dicionário aprendido.	FERNANDO, et al., 2020	Engenharia Elétrica; Ciência da Computação; Engenharia Aeroespacial e Aeronáutica; Engenharia da Comunicação
Digital Media Forensics	Processo que abarca etapas de investigação, preservação, análise e apresentação de evidências digitais.	HASHMI, et al., 2020	Engenharia de Eletrônica e Comunicação; Tecnologia da Informação; Matemática
Double JPEG localization method	Método de localização JPEG que detecta imagens manipuladas usando compactação JPEG.	BAEK; YOO; BAE, 2020	Ciência da Computação; Engenharia da Computação
Entire Face Synthesis manipulation	Manipulação com <i>Generative Adversarial Networks</i> que cria imagens de rosto inteiros inexistentes com alta qualidade e realismo.	TOLOSANA, et al., 2020	Ciência da Computação
Expectation-Maximization (EM) algorithm	Algoritmo empregado para definir modelo matemático-conceitual capaz de identificar a correlação de pixels diferenciando imagens reais de imagens possivelmente manipuladas.	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
Expression Swap	Manipulação que troca expressões faciais por meio de <i>Generative Adversarial Networks</i> .	TOLOSANA, et al., 2020	Ciência da Computação
Face de-identification	Processo de remoção de informações de identidade facial presentes em imagem e vídeo.	TOLOSANA, et al., 2020	Ciência da Computação
Face morphing	Manipulação facial que pode criar amostras biométricas artificiais semelhantes a faces de alvos possibilitando acessar sistemas que usem reconhecimento facial dos alvos.	TOLOSANA, et al., 2020	Ciência da Computação
Face Spoof detection	Semelhante à detecção de DeepFake, a detecção de falsidade facial objetiva identificar rostos vivos em vídeo diferenciando-os de imagens geradas artificialmente.	QI et al. 2020	Tecnologia

FaceApp	Aplicativo móvel com capacidade de gerar mudanças etárias, de gênero etc. realistas em fotografias de rostos humanas.	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
FaceForensics dataset	Conjunto de dados com vídeos coletados do <i>YouTube</i> contendo mais de 1000 vídeos de treinamento, de validação e de teste.	FERNANDO, et al., 2020	Engenharia Elétrica; Ciência da Computação; Engenharia Aeroespacial e Aeronáutica; Engenharia da Comunicação
FaceForensics++	Conjunto de dados contendo mais de 5000 vídeos originais e manipulados com diferentes métodos e contém quatro subconjuntos de vídeos manipulados: <i>DeepFake Detection</i> , <i>DeepFake</i> , <i>Face2Face</i> e <i>FaceSwap</i> .	KOHLI; GUPTA, 2021; QI et al. 2020; LI, X., et al., 2020	Engenharia Eletrônica e da Comunicação; Tecnologia da Computação; Tecnologia
Facial image analysis	Análise focada no reconhecimento de padrões e visão computacional para identificação forense de pessoas, análise de imagens faciais em 3D etc.	LI, D., et al., 2021	Ciência e Tecnologia da Computação; Matemática
Facial reenactment	A reconstituição facial consiste na mudança de expressões faciais em um vídeo a fim de gerar expressões falsas dessa pessoa, assim como transferir a expressão facial para outras pessoas.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica
FakeApp	Programa para Windows que permite mudança de rostos em vídeo por usuários não-especialistas.	GONG, et al., 2020	Tecnologia da Informação e Comunicação
FakeFace in the Wild (FFW) dataset	Conjunto de dados construído com vídeos do <i>YouTube</i> com grande variedade de vídeos falsos gerados por várias técnicas de computação gráfica, de <i>Generative Adversarial Networks</i> , sejam elas manuais, automáticas ou combinações.	FERNANDO, et al., 2020	Engenharia Elétrica; Ciência da Computação; Engenharia Aeroespacial e Aeronáutica; Engenharia da Comunicação
FFPMS	Conjunto de dados FaceForensics Plus com amostras de mixagem para detectar vídeos com <i>DeepFake</i> parcialmente manipulados.	LI, X., et al., 2020	Tecnologia da Computação
GAN	As <i>Generative Adversarial Networks</i> consistem em uma classe de modelos de <i>Deep Learning</i> , e consequentemente de <i>Machine Learning</i> , que possuem duas redes neurais: uma discriminadora, responsável por diferenciar dados reais de dados falsos; e uma geradora, voltada para gerar dados falsos que imitam os reais.	GREENGARD, 2020; TAHERI, et al., 2020	Jornalismo; Engenharia Elétrica e da Computação;

GDWCT	Método capaz de melhorar a qualidade de estilo de imagens geradas com <i>Generative Adversarial Networks</i> .	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
Generative model	O modelo generativo pode ser treinado com dados fornecidos e produzem amostras que seguem a mesma distribuição de dados do treinamento. Os mais conhecidos são baseados em redes neurais profundas como as <i>Generative Adversarial Networks</i> e as <i>Variational Autoencoders</i> .	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
GPIR dataset	Conjunto de dados com 80 imagens com cópia-movimento realista e em alta definição (768x1024 <i>pixels</i>).	ZHANG, et al., 2019	Ciência e Tecnologia Eletrônica; Segurança da Informação; Segurança da Computação; Ciência e Tecnologia da Computação;
Identity Swap manipulation	Manipulação que substitui o rosto de uma pessoa em vídeo pelo rosto de outra por meio de técnica de computação gráfica ou técnica de <i>Deep Learning</i> .	TOLOSANA, et al., 2020	Ciência da Computação
Illumination inconsistency method	Método que investiga inconsistências de iluminação em imagens e vídeos manipulados.	BAEK; YOO; BAE, 2020	Ciência da Computação; Engenharia da Computação
LSUN Bedroom dataset	Conjunto de dados contendo 10 categorias de cena e 20 categorias de objeto com aproximadamente um milhão de imagens rotuladas para cada categoria.	ZHANG, et al., 2019	Ciência e Tecnologia Eletrônica; Segurança da Informação; Segurança da Computação; Ciência e Tecnologia da Computação
Media clone	Outro termo que designa o uso de técnicas de <i>Machine Learning</i> para manipulação e geração de vozes, imagens faciais, vídeos e textos em linguagem natural.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica;
NeuralTextures	Falsificação de reencenação facial que pode alterar uma expressão facial de uma imagem de origem com expressão de uma imagem de destino.	BAEK; YOO; BAE, 2020; CHINTHA, et al., 2020	Ciência da Computação; Engenharia da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais

PGGAN dataset	Conjunto de dados composto por 80.000 imagens sintéticas de faces geradas através da Rede PGGAN.	NEVES, et al., 2020	Ciência e Engenharia da Computação; Engenharia de Telecomunicações
Presentation attack	Ataque que contorna sistemas de autenticação com informações biométricas.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica
RESNET discriminator	Rede Neural baseada em aprendizado residual profundo de alta precisão.	BAEK; YOO; BAE, 2020	Ciência da Computação; Engenharia da Computação
Sat-Cues	Recurso para detecção de imagens DNG que usa dicas de saturação para classificar imagens reais e imagens geradas com <i>Generative Adversarial Networks</i> .	LI, H. et al., 2020	Processamento Inteligente da Informação; Segurança de Mídia; Inteligência Artificial e Robótica para a Sociedade
STARGAN	Método que realiza traduções de imagem-para-imagem em vários domínios como mudança de cor de expressão facial, mudança de cor de cabelo etc. usando um único modelo.	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
StarGAN-based model	Modelo baseado no método STARGAN para conversão de muitas vezes sem treinar paralelamente os dados, onde as imagens foram substituídas por espectrogramas e as identidades por vetores de domínio, com uso do algoritmo Griffin-Lim.	CHINTHA, et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
StyleGAN	Conjunto de dados proposto em 2019 para gerar imagens com <i>Style Generative Adversarial Networks</i> , que transfere conteúdo semântico entre domínios a partir de um estilo e geram imagens falsas de rostos em alta qualidade como as imagens geradas no site " <i>this person does not exist</i> ".	ECHIZEN, et al., 2021; GUARNERA; GIUDICE; BATTIATO, 2020; ZHANG, et al., 2019	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica; Matemática; Ciência e Tecnologia Eletrônica; Segurança da Informação; Segurança da Computação; Ciência e Tecnologia da Computação;

StyleGAN2	Melhoria na qualidade da STYLEGAN através do gerador que cria imagens ainda mais realistas.	GUARNERA; GIUDICE; BATTIATO, 2020	Ciência da Computação; Matemática
Sub-SRM	Subconjunto de imagens de SRM com bom desempenho forense por seu refinamento.	LI, H. et al., 2020	Processamento Inteligente da Informação; Segurança de Mídia; Inteligência Artificial e Robótica para a Sociedade
Tampering Detection	Designa os estudos voltados para detectar adulterações em mídias e geração de mídia falsa por meio de <i>DeepFakes</i> , cujo conteúdo é empregado para desinformação, propaganda e influência em mídias sociais.	NADEEM, et al., 2019	Ciência da Computação
Text-to-speech synthesis	Método que gera falas falsas a partir de texto com uso de Redes Neurais profundas.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica
TPDNE	Conjunto de dados de imagens contendo 150.000 rostos únicos sintetizados a partir de <i>StyleGAN</i> e treinados com FFHQ.	NEVES, et al., 2020	Ciência e Engenharia da Computação; Engenharia de Telecomunicações
TP-GAN	Método em duas vias, nomeado <i>Two-Pathway Generative Adversarial Network</i> , que pode gerar uma imagem frontal a partir de outra imagem da visão lateral de um rosto.	NADEEM, et al., 2019	Ciência da Computação
TS-CDNN	Método baseado em <i>Deep Learning</i> que detecta imagem de rosto gerada por <i>Generative Adversarial Networks</i> por meio de uma Rede Neural Profunda acoplada (<i>Coupled Deep Neural Network</i> , CDNN) com duas etapas (<i>Two-Steps</i> , TS) de aprendizado.	LI, H. et al., 2020	Processamento Inteligente da Informação; Segurança de Mídia; Inteligência Artificial e Robótica para a Sociedade
Two-Stream faster R-CNN	Modelo com uma Rede Neural inspirada em R-CNN de fluxo RGB (que analisa as pistas visuais da imagem) e fluxo de ruído (que localiza áreas de possíveis adulterações).	NADEEM, et al., 2019	Ciência da Computação
VAE	No <i>Deep Learning</i> , os <i>Variational Autoencoders</i> (VAE) são usados para construção de modelos generativos por métodos de inferência variacional que produzem imagens de baixa qualidade limitando suas aplicações. O que abriu espaço para os modelos generativos construídos com <i>Generative Adversarial Networks</i> .	LI, H. et al., 2020	Processamento Inteligente da Informação; Segurança de Mídia; Inteligência Artificial e Robótica para a Sociedade

VGG-19 + capsules	Método <i>Deep Learning</i> com uso de cápsulas combinadas com VGG-19 para detecção de falsificação em imagens e vídeos.	NADEEM, et al., 2019	Ciência da Computação
VGGFace2	Conjunto de dados com 3.31 milhões de imagens muito variadas em pose, idade, iluminação, etnia etc. organizadas em 9.131 assuntos diferentes e uma média de 363 imagens por assunto.	NEVES, et al., 2020	Ciência e Engenharia da Computação; Engenharia de Telecomunicações
Voice cloning model	Modelo de clonagem de voz baseado em adaptação e codificação de fala de um alvo que pode gerar outra fala com voz parecida a partir de amostras de treinamento.	CHINTHA, et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais
Voice conversion (VC)	Técnica de processamento de voz que converte uma fala para que a voz se pareça com a voz de uma alvo.	ECHIZEN, et al., 2021	Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciência da Computação; Ciências da Segurança Social; Engenharia da Computação e Eletrônica;
WaveNet	Modelo de Rede Neural Profunda capaz de gerar formas de onda de fala a partir de texto e condicioná-las a identidade de voz de um falante específico.	CHINTHA, et al., 2020; ECHIZEN, et al., 2021	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais; Informática; Engenharia de Informação e Comunicação; Ciência de Dados; Ciências da Segurança Social

XceptionNet	Modelo de Rede Neural Profunda que apresenta altíssimo resultado de detecção apontado em publicações.	CHINTHA, et al., 2020; GUO, et al., 2021; NEVES, et al., 2020	Ciência da Computação; Engenharia Elétrica/Eletrônica; Tecnologia da Informação e Comunicação; Jornalismo e Relações Internacionais; Ciência e Engenharia da Informação; Ciência e Tecnologia da Informação; Engenharia da Computação; Engenharia de Telecomunicações
-------------	---	---	--

Fonte: Elaborado pelos autores

Copyright: © 2022 Soella, Gabriel Meneguelli, and Maimone, Giovana Deliberali. This is an open-access article distributed under the terms of the Creative Commons CC Attribution-ShareAlike (CC BY-SA), which permits use, distribution, and reproduction in any medium, under the identical terms, and provided the original author and source are credited.

Received: 22/04/2022

Accepted: 16/06/2022