INTELIGÊNCIAS ARTIFICIAIS GERADORAS DE IMAGENS: estudo com o sistema DALL-E 2 para a produção de conteúdos com base em descrições de texto

IMAGE GENERATING ARTIFICIAL INTELLIGENCES: study with the DALL-E 2 system for producing content based on text descriptions

Vitor de Lima Amaral – vitordelimaamarall2@gamil.com
Faculdade de Tecnologia de Taquaritinga – Taquaritinga – SP – Brasil

Gustavo Henrique Del Vechio – gustavo.vechio@fatec.sp.gov.br
Faculdade de Tecnologia de Taquaritinga (Fatec) – Taquaritinga – SP – Brasil

DOI: 10.31510/infa.v20i2.1805
Data de submissão: 06/09/2023
Data do aceite: 16/11/2023
Data da publicação: 20/12/2023

RESUMO

Inteligências Artificiais são cada vez mais utilizadas entre indivíduos e profissionais para a realização de diversas tarefas cotidianas, incluindo a capacidade de geração, edição ou transformação de imagens com base em descrições detalhadas de texto. Considerando-se tal cenário, este estudo tem o objetivo de explicar os fundamentos das inteligências artificiais, incluindo os conceitos de neurônios artificiais e redes neurais, aprofundando-se nas tecnologias capazes de trabalhar com a manipulação de imagens, quais sejam fotografias ou ilustrações, para avaliar os benefícios e possíveis implicações no uso destas tecnologias. Para isso, como metodologia, propõe a revisão de livros e materiais científicos sobre o assunto, em seguida a realização de um estudo de caso com um sistema em específico de criação de imagens por IA, chamado DALL-E 2. A escolha deste sistema em particular deve-se ao fato de ele ser gratuito, de fácil acesso, bastante popular, além de já ser aproveitado por empresas de grande porte, tais como a Microsoft. O que se percebe, a partir deste estudo, é que estes recursos geradores de imagens podem contribuir muito no processo criativo, sobretudo para profissionais de áreas como design, ilustração, fotografia e vídeo; no entanto, seu uso também pode desviar para outros propósitos, a partir da geração de conteúdos prejudiciais ou que contribuam para a propagação de notícias faltas, cada vez mais difíceis de lidar em uma sociedade com tantas tecnologias à disposição.

Palavras-chave: Inteligência Artificial. Geração de imagens. DALL-E 2. Evolução tecnológica.

ABSTRACT

Artificial Intelligence is increasingly used among individuals and professionals to perform various everyday tasks, including the ability to generate, edit or transform images based on detailed text descriptions. Considering this scenario, this study aims to explain the foundations of artificial intelligence, including the concepts of artificial neurons and neural networks, delving deeper into

technologies capable of working with the manipulation of images, whether photographs or illustrations, to evaluate the benefits and possible implications of using these technologies. To achieve this, as a methodology, it proposes the review of books and scientific materials on the subject, then carrying out a case study with a specific AI image creation system, called DALL-E 2. The choice of this system in This particular feature is due to the fact that it is free, easy to access, quite popular, and is already being used by large companies, such as Microsoft. What can be seen from this study is that these image-generating resources can contribute a lot to the creative process, especially for professionals in areas such as design, illustration, photography and video; however, its use can also be diverted to other purposes, such as generating harmful content or contributing to the spread of bad news, which is increasingly difficult to deal with in a society with so many technologies at our disposal.

Keywords: Artificial Intelligence. Image generation. DALL-E 2. Technological evolution.

1 INTRODUÇÃO

Ao longo da história, novas tecnologias permitiram ao homem aumentar sua eficiência, tanto física quanto cognitiva. Na década atual, a Inteligência Artificial (IA) é uma tecnologia que apresenta crescimento exponencial, sobretudo por envolver o processamento de informações a fim de gerar resultados lógicos e racionais, contribuindo em muitas tarefas que antes só poderiam ser desenvolvidas por indivíduos especializados. Atualmente, é possível identificar vários tipos de inteligências com propósitos específicos, que vão desde a geração de documentos e cálculos complexos, até a criação de músicas e imagens com base em descrições detalhadas de textos.

No entanto, apesar de ser cada vez mais utilizada em vários aspectos da vida humana, a aplicação de inteligências artificiais para a criação de imagens tem gerado opiniões divergentes quanto à sua usabilidade e benefícios, pois embora possa ser usada para criar ou editar imagens muito mais facilmente, também acaba impactando consideravelmente em como os profissionais operam neste novo cenário, além de como a sociedade recebe e aproveita esta tecnologia.

Pensando-se neste contexto, este estudo procura apresentar, de maneira mais clara, como uma inteligência artificial geradora de imagens opera em suas criações, acarretando novas formas de geração de conteúdos e de disseminação de imagens. Quer-se, neste sentido, avaliar não apenas a premissa básica de inteligências artificiais, neurônios e rede neural artificial, mas também os benefícios e possíveis implicações com o uso desta tecnologia.

Para alcançar este objetivo, propõe-se, como metodologia, inicialmente a construção de uma fundamentação teórica que explore cada um destes conceitos, tendo como base a revisão

bibliográfica de livros, artigos e outros materiais de pesquisa científica. Em seguida, com o intuito de demonstrar na prática como estes conceitos se aplicam, propõe-se um estudo de caso com um famoso sistema de geração de imagens por inteligência artificial, chamado DALL-E 2, demonstrando como ele opera e quais as suas contribuições para os indivíduos e profissionais.

A justificativa para a elaboração deste estudo incide sobre o fato de que é cada vez mais recorrente o surgimento de novas inteligências artificiais que manipulam imagens e outros conteúdos multimídia; logo, compreender sua utilização e saber como aproveitar de seu potencial pode ser de enorme valia a indivíduos e sociedade como um todo.

2 FUNDAMENTAÇÃO TEÓRICA

2.1 Conceitos básicos de Inteligência Artificial

A Inteligência Artificial (popularmente conhecida pela sigla IA) não é uma invenção tão recente quanto se imagina. Na verdade, Gomes (2010) explica que uma busca atenta pela história permite demonstrar que este conceito remonta há milênios, tendo-se em vista a contínua busca por métodos que possam emular o raciocínio humano. No entanto, foi com o passar do tempo que este conceito evoluiu consideravelmente, dando origem a diversas abordagens conceituais, com destaque para as lógicas e pensamentos fundamentados em sistemas que emulam não apenas o pensamento humano, mas que desempenham funções comparáveis às ações humanas ou ainda sistemas dotados de capacidade de raciocínio autônomo. Para Gomes (2010, p. 239), aliás, a "IA também está relacionada com psicologia, biologia, lógica matemática, linguística, engenharia, filosofia, entre outras áreas científicas."

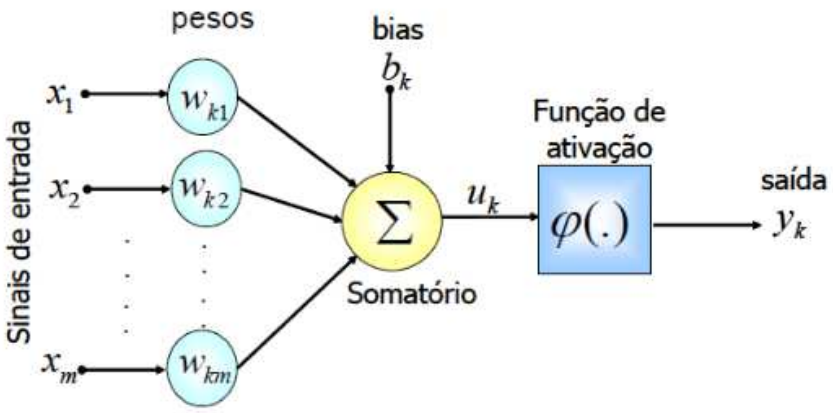
Uma matéria no site da Oracle (O QUE É IA..., s.d.) complementa que a Inteligência Artificial se fundamenta sobretudo na análise de extensos conjuntos de dados, nos quais algoritmos inteligentes operam em conformidade com a validação e interpretação destes respectivos dados. Isso resulta na geração de informações que permitem ao sistema identificar e interpretar padrões. Nesse contexto, a IA adquire a capacidade de aprendizado automático (*machine learning*), o que inclui a habilidade de ler e compreender informações de maneira mais automática.

2.2 Neurônio artificial

Um neurônio artificial, como demonstra Haykin (2000), é uma unidade elementar de processamento, mais especificamente um subcampo da Inteligência Artificial. Ele recebe este nome devido à sua inspiração no funcionamento dos neurônios humanos, embora seja uma representação abstrata de um neurônio biológico.

Neurônios artificiais, ainda de acordo com Haykin (2000), podem aceitar múltiplas entradas ponderadas, que são tecnicamente somadas e submetidas a uma função de ativação, por sua vez responsável por determinar se o neurônio deve ou não gerar uma saída com base na soma das entradas ponderadas. O objetivo primordial de um neurônio artificial reside em sua capacidade de aprender e de reajustar os pesos associados às suas entradas, visando a execução de tarefas específicas. Esse processo de aprendizado é conduzido mediante a utilização de algoritmos de treinamento, a exemplo do algoritmo de retro propagação, que gradualmente otimiza os pesos das entradas com o intuito de minimizar a discrepância entre a saída desejada e a produzida pelo neurônio artificial:

Figura 1 - Neurônio artificial.



Fonte: retirado do estudo de Haykin (2000).

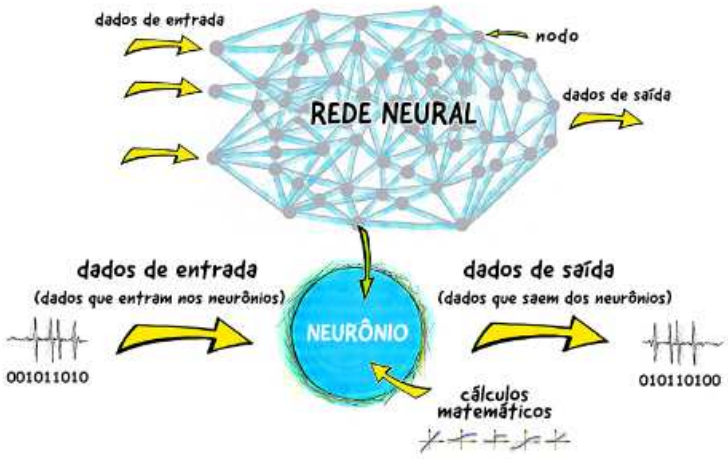
Em outras palavras, um neurônio artificial é uma unidade de processamento fundamental em redes neurais artificiais, desenvolvida para desempenhar tarefas de aprendizado de máquina e processamento de informações. Ele se baseia em conceitos simplificados dos neurônios biológicos

para operar de forma eficaz no contexto computacional. É, portanto, a base para o processamento da inteligência artificial, seja na geração de textos, imagens, cálculos e outras funções.

2.3 Rede neural artificial

De acordo com Pregowska e Osial (2021), uma rede neural artificial, assim como um neurônio artificial, corresponde a um modelo de processamento de informações inspirado no funcionamento do cérebro humano. A rede é de extrema importância para a aprendizagem da inteligência artificial, pois consiste em um conjunto interconectado de unidades de processamento (os já citados neurônios artificiais), que operam em conjunto a fim de realizar tarefas complexas, por meio do processamento de milhares de dados coletados ou gerados pela inteligência.

Figura 2 - Rede Neural Artificial.



Fonte: adaptado do estudo de Pregowska e Osial (2021).

Bittencourt e Osório (2000) complementam que os conhecimentos das redes neurais artificiais se definem por um conjunto de valores numéricos que descreve as conexões estabelecidas; por consequência, descreve também o comportamento da respectiva rede. Logo, a estrutura de uma rede neural artificial pode ser organizada em algumas etapas principais, como:

- **Camada de dados de entrada (*input layer*):** esta é a etapa responsável por receber os dados brutos que servirão de base para a rede neural funcionar e processar informações.

- **Camada intermediaria (*hidden layer*):** esta segunda etapa, mais conhecida como camada intermediaria, é composta por neurônios artificiais que processam as informações de entrada de maneira sequencial e de acordo com a hierarquia; cada camada pode fazer transformações complexas nos dados de saída, à medida que os dados passam pela rede.
- **Camada de dados de saída (*output layer*):** esta é a terceira e última etapa, responsável por produzir a saída final da rede neural artificial, que pode ser uma previsão, classificação ou resposta a um problema específico.

De acordo com Rauber (2005, p. 28), "a vantagem das redes é a aprendizagem por um conjunto de treino, junto com um caráter universal da arquitetura da rede. Isso significa que o projetista de uma rede não tem que se preocupar muito com parâmetros internos da rede."

Ora, este aspecto se deve porque a conexão de uma etapa para outra é realizada a partir de pesos sendo que, durante a fase de aprendizagem da rede neural artificial, esses pesos são ajustados a fim de reduzir as chances de erros no processo de saída dos dados.

Segundo Bender e Osório (2003), as redes neurais artificiais podem ser muito úteis e utilizadas em diversas tarefas, incluindo o reconhecimento de imagem, processamento de linguagem natural, previsão de séries temporais, classificação de dados e muito mais. Seu potencial de aprendizagem é baseado na quantidade de dados que recebe, tornando-se uma ferramenta poderosa em visão computacional, processamento de linguagem natural e análise de dados.

2.4 Inteligência Artificial geradora de imagens

Uma IA geradora de imagens opera a partir de uma descrição detalhada de texto, ou seja, é uma ferramenta alimentada por inteligência artificial que recebe um texto descritivo, processa-o e produz uma imagem que melhor corresponde à descrição fornecida.

No entanto, não se deve confundir esta tecnologia com certas ferramentas de busca por imagens: a geração de imagens por IA é um processo distinto, pois enquanto as ferramentas de busca são utilizadas para pesquisar e recuperar imagens existentes disponíveis em diversas plataformas, uma geradora de imagens cria imagens totalmente novas, com base nas solicitações dos usuários ou em certas imagens de exemplo. Por isso, é importante ao usuário ter a clara ideia

do tipo de imagem que deseja obter, ou seja, decidir o assunto da imagem, o sentimento que deseja transmitir e como deve ser composta a cena, considerando os objetos envolvidos.

Alguns princípios essenciais para a obtenção de imagens de alta qualidade geradas por IA incluem a elaboração de descrições bem detalhadas e precisas. A qualidade das imagens geradas depende muito da habilidade em formular um *prompt* de descrição claro. A chave para melhorar as imagens de IA encontra-se na escolha de palavras adequadas, sendo que, muitas vezes, os termos utilizados por profissionais de design, fotografia ou arte podem ser importantíssimos para se obter resultados satisfatórios e que descrevem de forma eficaz as imagens a serem geradas.

Para Rola (2022), o "processamento de imagens digitais consiste em utilizar a capacidade dos computadores para identificar, comparar e projetar informações baseadas no espólio de dados disponíveis internamente nas organizações, assim como os publicados na Internet."

Vale ressaltar, ainda, que a qualidade de uma imagem gerada artificialmente depende não apenas da qualidade e quantidade dos dados de treinamento da máquina, mas também da complexidade do modelo requisitado pelo usuário na etapa de descrição da imagem. Quanto mais uma inteligência artificial é treinada, mais suas capacidades de gerar imagens de alta qualidade e fidelidade tendem a melhorar.

3 PROCEDIMENTOS METODOLÓGICOS

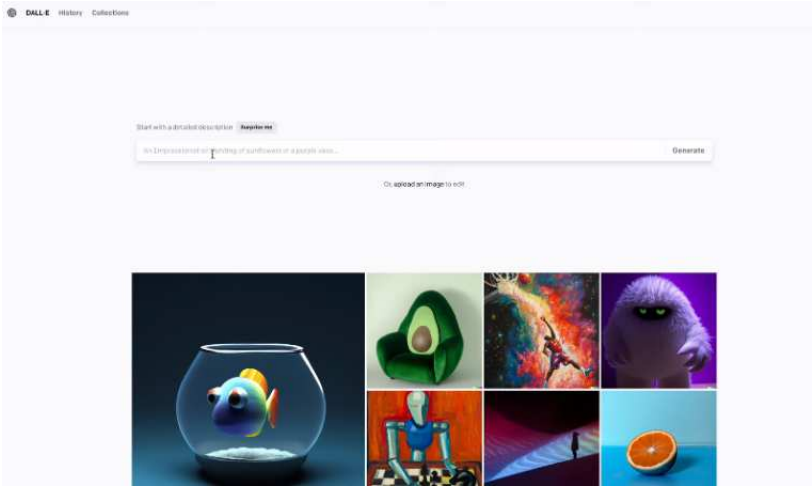
Compreendido o conceito de Inteligência Artificial, bem como a estrutura de neurônios e redes neurais artificiais, apresenta-se, a seguir, um estudo de caso com o intuito de demonstrar, na prática, as etapas de geração artificial de imagens por meio de descrições detalhadas de texto. Neste estudo, optou-se pela escolha do sistema chamado DALL-E 2 (https://openai.com/dall-e-2), em função de este ser um dos sistemas mais utilizados por usuários no mundo, bem como ele oferece a criação de imagens livres de direitos autorais (JOHNSON, 2023). Outro fato interessante sobre este sistema é que ele é originalmente criado pelo OPEN AI, mesmo desenvolvedor do famoso Chat GPT. Aliás, a própria Microsoft já incorporou recursos do DALL-E 2 em algumas das suas plataformas, incluindo seu mecanismo oficial de busca na Internet, o Bing.

O funcionamento do DALL-E 2 começa a partir do momento em que o usuário entra no site e aceita os termos e condições de uso do serviço, estando apto a iniciar a geração de imagens.

A figura 3 demonstra a tela inicial da ferramenta, que contém, além do campo de inserção de *prompts* de texto, uma belíssima galeria de imagens já geradas pela tecnologia.

Ao usuário, cabe inicialmente inserir, no campo superior (*prompt*), a descrição detalhada de texto, tendo em vista a imagem que se imagina ou pretende obter. É importante destacar que, não apenas o DALL-E 2, mas a maioria de outras ferramentas conhecidas e geradoras de imagens por IA, tais como (apenas para citar algumas) Midjourney (https://www.midjourney.com/), DreamStudio (https://dreamstudio.ai/), Adobe Firefly (https://firefly.adobe.com/) ou Jasper Art (https://www.jasper.ai/), operam desta mesma forma. A criação por meio de detalhamento de texto é, portanto, um comportamento comum a todas estas ferramentas.

Figura 3 – Página inicial do sistema DALL-E 2



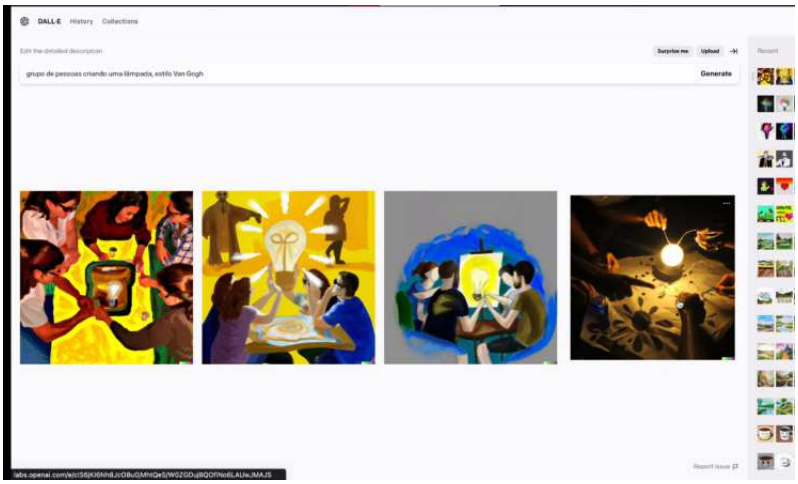
Fonte: tela do site DALL-E 2, capturada pelos autores (2023).

Como exemplo, a seguinte descrição de texto foi submetida ao DALL E-2: “grupo de pessoas criando uma lâmpada, estilo Van Gogh”. A partir deste *prompt*, a tecnologia gerou não apenas 1, mas 4variações de imagens, de acordo com as características requisitadas (figura 4).

Do lado direito do site do DALL-E 2, é possível encontrar diversos filtros que permitem refinar ou adequar, por exemplo, as imagens obtidas a um estilo artístico ou fotográfico em específico, além de ajustar métricas gráficas, tais como o controle de cores, dimensões e proporção original da imagem gerada. Aliás, a qualquer momento pode-se pedir novas variações, forçando a IA a gerar ainda mais imagens com base nas descrições iniciais de texto (este é um comportamento não apenas do DALL-E 2, mas de todas as IAs geradoras de imagens). Aliás, corroborando a

fundamentação teórica previamente apresentada, o fato de solicitar mais variações ao sistema não é prejudicial; pelo contrário, permite ao algoritmo de inteligência artificial aprender cada vez mais, para, no futuro, ser ainda mais assertivo na geração de novas imagens.

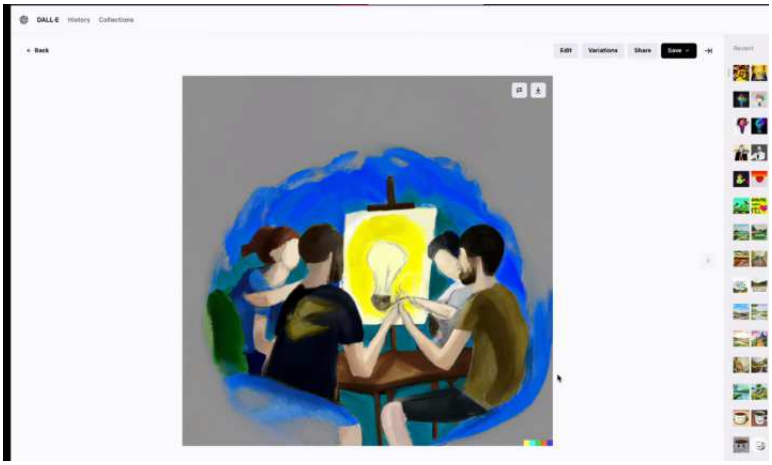
Figura 4 – Dall-E 2 criando imagem a partir de um *prompt* de texto



Fonte: tela do site DALL-E 2, capturada pelos autores (2023).

No momento em que o usuário estiver satisfeito com o resultado, isto é, com a imagem resultante, pode baixá-la para seu uso próprio. Um detalhe a ser notado é que, assim como os outros sistemas, no rodapé da imagem é adicionada uma pequena assinatura ou marca d'água, a fim de registrar que o respectivo conteúdo foi criado utilizando a inteligência artificial DALL-E 2.

Figura 5 – Etapa final de download da imagem no Dall-E 2.



Fonte: tela do site DALL-E 2, capturada pelos autores (2023).

Como é possível observar, embora a inteligência artificial seja uma tecnologia complexa, dotada de uma rede neural capaz de processar muitos dados de maneira lógica, seu uso é extremamente amigável para permitir a qualquer usuário, seja profissional ou não, gerar imagens antes nunca vistas de maneira simples e rápida. Não se trata, neste estudo, de comparar diferentes sistemas a fim de esclarecer seus prós e contras; o que se pretende mostrar, com base em um deles, o Dall-E 2, é que a inteligência artificial está disponível e ao alcance de qualquer indivíduo que queira utilizá-la, por meio de diversos dispositivos, sejam computadores, celulares ou tablets.

4 RESULTADOS E DISCUSSÃO

O funcionamento de uma IA geradora de imagens, tal como o DALL-E 2, deixa muito claro o impressionante potencial desta tecnologia e como ela pode ser capaz de gerar imagens completamente inéditas a partir de uma descrição detalhada de texto.

Analisando-se com atenção o funcionamento desta tecnologia, pode-se elencar diversos benefícios ou pontos positivos, a começar pela eficiência, já que a IA tem o potencial de acelerar o processo de criação visual, economizando a indivíduos e profissionais tempo e recursos. Outro fato marcante diz respeito à personalização e persistência, ou seja, a habilidade de a inteligência gerar diversas variações do que é solicitado, sendo que o usuário pode sempre pedir mais e mais variações até obter um resultado que faz jus ao que se espera. É importante considerar, ainda, a redução de custos em alguns casos, pois dependendo daquilo que se queira, as empresas podem optar por não terceirizar certos serviços, tais como a elaboração de uma ilustração, ou uma edição de imagens que envolva alterar objetos ou criar um cenário em específico.

Seja como for, é de se notar o caráter vantajoso de este tipo de inteligência artificial poder ser usado cada vez mais em uma enorme variedade de segmentos de mercado, agregando mais recursos sobretudo aos profissionais de design, artes, ilustrações e fotografia e vídeos.

Contudo, embora estes sejam benefícios interessantes, nunca se pode esquecer que toda tecnologia também carrega consigo certos pontos negativos. No caso desta tecnologia geradora de imagens por IA, vale destacar a qualidade ou detalhamento de imagens geradas, que ainda pode apresentar certas inconsistências e apresentar resultados diferentes do esperado, sobretudo tratando-se de novidades. Em outras palavras, é claro que o detalhamento está muito ligado ao

nível de descrição textual fornecido pelo usuário, no entanto, tudo o que esta tecnologia é capaz de criar está ligado ao que ela já aprendeu e analisou de outras imagens já existentes; tentar pedir a ela para desenvolver, por exemplo, um novo conceito de produto, pode ainda não ser algo tão satisfatório se comparado à capacidade humana de criar, inovar e renovar conceitos.

Por fim, também é preciso considerar questões relacionadas à segurança e uso social de imagens, já que tal inteligência pode ser usada para a geração de conteúdos perigosos e enganosos, contribuindo para a disseminação de *fake news*, cada vez mais difíceis de identificar. É, portanto, uma questão legal a se considerar, que pode envolver mais esclarecimentos em relação às políticas de autoria, plágio e de direitos autorais dos conteúdos gerados pelas inteligências artificiais.

5 CONSIDERAÇÕES FINAIS

Ao longo da história, a tecnologia tem caminhado lado a lado com a humanidade, desempenhando um papel crucial no processo evolutivo não apenas de indivíduos em particular, mas das sociedades como um todo. Desde os tempos mais longínquos, as pessoas usam ferramentas e recursos tecnológicos para facilitar suas tarefas diárias e estender o potencial físico e cognitivo de seus corpos. No século XXI, é possível observar diversas tecnologias de alto impacto e que são parte integrante da vida em sociedade, quais sejam os computadores e dispositivos eletrônicos, sem os quais grande parte dos humanos sequer imaginam viver sem.

O uso de inteligências artificiais cresce a cada dia, desempenhando um papel significativo neste cenário altamente tecnológico. Como foco deste estudo, as inteligências geradoras de imagens oferecem inúmeras vantagens a indivíduos e profissionais, como economia de tempo, estímulo à criatividade e suporte na elaboração de projetos, o que abre um leque de novas oportunidades ao trabalho criativo. No entanto, o potencial de resultado destas ferramentas é totalmente baseado no interesse de cada usuário, o que significa que uma tecnologia como essa, embora seja de fácil acesso e utilização, pode ser utilizada para a geração de imagens perigosas e contribuir para a geração de notícias faltas. É, portanto, uma tecnologia com crescimento exponencial, mas que se não usada de forma adequada, poderá, no futuro, causar certos problemas sociais por agravar a disseminação de notícias faltas e outros conteúdos maliciosos.

REFERÊNCIAS

BENDER, Túlio C.; OSÓRIO, Fernando S. Reconhecimento e recuperação de imagens: utilizando redes neurais artificiais do tipo MLP. **Unisinos**, 2003. Disponível em: < https://osorio.wait4.org/publications/Bender-Enia2003.pdf >. Acesso em: 16 set. 2023.

BITTENCOURT, João R. OSÓRIO, Fernando. **Processamento de imagens inteligente utilizando redes neurais artificiais**. XIII Brazilian Symposium on Computer Graphics and Image Processing, Gramado, out. 2000. Disponível em: < https://osorio.wait4.org/publications/Bittencourt-Osorio-SIC-Book2002.pdf >. Acesso em: 12 set. 2023.

DALL-E 2. Disponível em: < https://openai.com/dall-e-2 >. Acesso em: 20 set. 2023.

GOMES, Dennis S. Inteligência Artificial: conceitos e aplicações. **Revista Olhar Científico**, v. 1, n. 2, ago-dez 2010. Disponível em: < https://www.professores.uff.br/screspo/wp-content/uploads/sites/127/2017/09/ia_intro.pdf >. Acesso em: 21 set 2023.

HAYKIN, Simon S. **Redes Neurais**. 2 ed. Porto Alegre: Bookman, 2000.

JOHNSON, Ariana. Here are the best AI image generators. **Forbes**, 28 abr. 2023. Disponível em: < https://www.forbes.com/sites/ariannajohnson/2023/04/28/here-are-the-best-ai-image-generators/ >. Acesso em 20 set. 2023.

O QUE È IA? Saiba mais sobre inteligência artificial. **Oracle Brasil**, s.d. Disponível em: < https://www.oracle.com/br/artificial-intelligence/what-is-ai/ >. Acesso em: 18 set. 2023.

PREGOWSKA, Agnieszka; OSIAL, Magdalena. O que é uma rede neural e para que serve? **Frontiers**, 17 ago. 2021. Disponível em: < https://kids.frontiersin.org/articles/10.3389/frym.2021.560631 >. Acesso em: 14 set. 2023.

RAUBER, Thomas W. **Redes neurais artificiais**. Universidade Federal do Espírito Santo, 2005. Disponível em: < https://www.researchgate.net/profile/Thomas-Rauber-2/publication/228686464_Redes_neurais_artificiais/links/02e7e521381602f2bd000000/Redes-neurais-artificiais.pdf >. Acesso em: 14 set. 2023.

ROLA, Eulálio C. S. Os principais contributos da inteligência artificial para o processamento de imagens digitais a utilizar na segurança pública. Dissertação de mestrado em Segurança e Justiça. Faculdade de Direito da Universidade Lusíada, Lisboa, 2022. Disponível em: < https://repositorio.ulusiada.pt/handle/11067/6548 >. Acesso em: 13 set. 2023.