

ANÁLISE DE IMAGENS GERADAS POR INTELIGÊNCIA ARTIFICIAL
PARA O ROMANCE “UM CÃO LATE NA NOITE”

ANALYSIS OF IMAGES GENERATED BY ARTIFICIAL INTELLIGENCE
FOR THE NOVEL “ONE DOG BARKS IN THE NIGHT”

Ives da Silva Duque-Pereira¹
Sérgio Arruda de Moura²

Resumo: O presente artigo explorou a capacidade da Inteligência Artificial Generativa (IAG) de interpretar e visualizar um trecho literário da obra *Um Cão late na noite* como *prompt* (comando), utilizando a ferramenta Midjourney. O artigo investigou como a IAG texto-para-imagem ressignifica o texto literário ao criar imagens, por meio da análise sógnica e o processo inverso de imagem-para-texto, percebendo a “leitura” que o Midjourney faz da imagem gerada por IA. Observou-se que, embora o trecho literário possa não seguir as diretrizes convencionais para *prompts* ideias, os resultados apresentados oferecem *insights* valiosos sobre a interseção entre tecnologia e literatura.

Palavras-chave: Inteligência Artificial Generativa (IAG), Midjourney, texto-para-imagem, imagem, comunicação.

Abstract: The present article explored the ability of Generative Artificial Intelligence (IAG) to interpret and visualize a literary excerpt from the work “Um Cão barks na noite” as a *prompt* (command), using the Midjourney tool. The article investigated how IAG text-to-image reframes the literary text by creating images, through sign analysis and the inverse process of image-to-text, realizing the “reading” that Midjourney does of the image generated by AI. It was noted that although the literary excerpt may not follow conventional guidelines for idea prompts, the results presented offer valuable insights into the intersection between technology and literature.

Keywords: Generative Artificial Intelligence (IAG), Midjourney, text-to-image, image, communication.

¹ Doutorando no Programa de Pós-Graduação em Cognição e Linguagem na Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes-RJ. ivesduque@gmail.com

² Doutor em Letras (UFRJ), professor no Programa de Pós-Graduação em Cognição e Linguagem na Universidade Estadual do Norte Fluminense Darcy Ribeiro (UENF), Campos dos Goytacazes-RJ. arruda@uenf.br

1 Introdução

Este estudo se situa na interseção da semiótica e da comunicação, explorando esta relação no contexto da Inteligência Artificial Generativa (IAG). A massificação do uso da IAG, a partir de 2022, com a disponibilização de ferramentas para uso comum na geração de textos e imagens, tem despertado preocupações e entusiasmo. Santaella (2023) afirma que se trata de uma quarta ferida narcísica na humanidade que tem provocado grande desconforto.

Uma parte das ferramentas de IAG de texto-para-imagem funciona a partir da evolução de uma arquitetura de dados chamada *diffusion*. Não é propósito do presente trabalho a compreensão do complexo funcionamento interno das ferramentas de IAG. Porém, consideramos importante o entendimento da existência de uma rede neural artificial formada pela conexão entre representações visuais e textuais de objetos, a partir de um treinamento por aprendizagem de máquina, tendo como base um vasto banco de dados.

De maneira geral, imagens que fazem parte deste treinamento são decodificadas visualmente em seus pixels e descritas textualmente, criando uma relação que possibilita a combinação e produção de novas imagens a partir de comandos textuais humanos (texto-para-imagem). O foco está na percepção e análise das imagens resultantes (*outputs*) e sua relação com o texto gerador (*inputs*). Outro ponto de entendimento necessário é o processo descritivo de geração imagem-para-texto.

Oppenlaender (2022), ao pesquisar a criatividade da geração de texto-para-imagem, concluiu que uma visão centrada no produto é insuficiente e falha, pois a criatividade está muito relacionada ao *prompt* (comando textual humano). Assim, deve-se prestar atenção nas já existentes comunidades online para o ecossistema criativo da produção de imagens geradas a partir de textos e o surgimento de um novo campo de prática textual chamado engenharia – ou *design* - de *prompt*.

O problema deste estudo é investigar como a Inteligência Artificial Generativa (IAG) compreende e interpreta um texto literário, a partir de fragmentos, quando utilizado como *prompt*. Além disso, também indaga de que maneira a IAG pode expandir ou ressignificar o texto literário e quais são as consequências desse processo. Nosso principal objetivo é explorar a geração de imagens por IAG a partir de trechos do romance *Um cão late na noite*, utilizando a ferramenta Midjourney como instrumento. Para alcançar esse propósito, selecionamos um trecho do romance mencionado e a análise dos resultados será conduzida sob a ótica da semiótica, fundamentado nesse campo teórico.

A presente investigação foi motivada pela identificação de uma lacuna significativa na literatura científica referente à análise da interação entre a Inteligência Artificial Generativa

(IAG) e a produção de imagens baseadas em textos literários. Uma revisão realizada por meio do Google Acadêmico, empregando os termos "Inteligência Artificial Generativa", "IA Generativa", "Inteligência Artificial", "Imagem", "Literatura", "literária" e "literário" como filtros de busca nos títulos dos trabalhos, revelou escassez de estudos pertinentes a esta temática. Curiosamente, a única incidência encontrada estava relacionada à análise de imagens odontológicas por meio de inteligência artificial, destacando-se, portanto, a ausência de investigações focadas na aplicação da IAG na interpretação e geração de imagens a partir de narrativas literárias.

Adicionalmente, ao explorar a presença acadêmica da ferramenta Midjourney nos títulos de trabalhos científicos, identificou-se um total de 153 publicações. Dentre estas, apenas sete eram redigidas em língua portuguesa. Este achado sublinha a emergente popularidade da IAG, especialmente a partir de 2023, e reforça a necessidade de um engajamento mais robusto por parte da comunidade acadêmica em pesquisas que abordem as potencialidades e desafios da integração entre inteligência artificial generativa e a criação de imagens inspiradas em textos literários. Desta forma, entendemos que há um vasto campo de investigação a ser explorado, que pode contribuir significativamente para o avanço do conhecimento nas áreas de inteligência artificial, literatura e estudos interdisciplinares.

Como será discutido adiante, o texto literário não se alinha perfeitamente com as diretrizes tradicionalmente associadas a um *prompt* ideal. No entanto, ao adotar essa abordagem, este estudo se aprofunda na maneira como a IAG interpreta textos literários. Esse processo estabelece conexões entre tecnologia e literatura, avaliando a adaptabilidade da IAG e buscando identificar suas possíveis limitações e potencialidades ao lidar com entradas "não convencionais". Com isso, almejamos enriquecer os campos da comunicação e da semiótica ao explorar o que pode ser considerada uma nova forma de semiose, refletindo sobre a formação, interpretação e transformação dos signos em diferentes contextos.

2 Midjourney e a Engenharia – Ou *Design* – de *Prompt*

Passado um ano da disponibilização e popularização dessas ferramentas, a tecnologia de IAG está sendo empregada em diversos campos, incluindo arte e *design*, tendo a capacidade de criar ilustrações de todos os tipos, inclusive para livros. Grandes empresas – e agências publicitárias – já estão explorando a capacidade das IA generativas para criar imagens e frases em suas campanhas. Mesmo aquelas empresas cuja boa parte do faturamento está ligado a distribuição de publicidade como a Bigtech Google/Alphabet, já

anunciaram funcionalidades para montagem de anúncios – um trabalho antes manual – com uso de IAG na geração de textos e imagens (DISCHLER, 2023).

Conforme Li *et al.* (2023), a IAG tem a capacidade de gerar imagens de qualquer objeto, em qualquer contexto, com alta fidelidade e qualidade técnica, reforçando que essa capacidade tem revolucionado a criatividade e causado um impacto significativo na produção de imagens para fins comerciais e artísticos.

No entanto, é importante destacar que ainda existem questões controversas não resolvidas relacionadas à produção de imagens por IAG. As mais relevantes são o direito de uso e autoria das imagens geradas e os direitos autorais das imagens do banco de dados que alimentam essas ferramentas.

Devido à complexidade dessas questões e à necessidade de uma discussão mais aprofundada, elas não serão o foco deste trabalho. Adotaremos a premissa de que as ferramentas de IAG para imagens, incluindo o Midjourney, não reivindicam direitos autorais sobre as imagens produzidas. Atualmente, há uma lacuna na legislação sobre o uso dessas imagens. No entanto, o US Copyright Office, responsável pelos direitos autorais nos EUA, já reconhece que os humanos que orientam a produção de imagens via IAG podem ter direitos autorais em certas circunstâncias (BRITAIN, 2023).

O Midjourney é uma dessas ferramentas de IAG que produz imagens a partir de *prompts*. Criado por um laboratório de pesquisa independente de São Francisco, EUA, o Midjourney tem se destacado como uma das ferramentas mais potentes no que se pretende. Frequentemente referenciado pela sua alta qualidade de imagens, chegando a ganhar prêmio em concursos de arte concorrendo com humanos (ROOSE, 2022), o Midjourney tem uma tendência a imagens esteticamente marcantes e é muito popular entre os artistas (BORJI, 2022).

O acesso à ferramenta paga *Midjourney* é realizado através da plataforma *Discord*, uma ferramenta de comunicação e rede social amplamente reconhecida no universo dos *gamers*. No *Discord*, é possível criar comunidades denominadas "servidores". Estes servidores, por sua vez, permitem a formação de *chats* segmentados, conhecidos como "canais", que funcionam como subgrupos especializados para conversas por texto, voz ou vídeo.

O Midjourney mantém um servidor próprio no *Discord* que abriga diversos canais. Estes canais facilitam a interação entre os usuários da ferramenta, consolidando uma comunidade *online* voltada para a criação, discussão e aprendizado sobre imagens geradas por IA. Dentre os canais disponíveis, há espaços para boas-vindas, suporte relacionado a pagamentos, esclarecimento de dúvidas, criação coletiva de imagens, compartilhamento de

dicas sobre *prompts*, *brainstorming* de novas ideias, estudos de caso, debates éticos e reflexões filosóficas sobre o uso e a natureza das imagens.

Além disso, a plataforma oferece a opção de utilizar um robô (*Bot*) de criação em conversas privadas em outros servidores, garantindo que os membros da comunidade *Midjourney* não tenham acesso a essas interações particulares.

Dentro da comunidade *Midjourney*, há um consenso de que o fornecimento de *prompts* genéricos e sem detalhes, que deixam a "imaginação" da IAG preencher as lacunas, não é a abordagem mais eficaz para utilizar a ferramenta. Conforme observado no estudo de Oppenlaender (2022), resultados oriundos de *prompts* mais genéricos e pouco detalhados não refletem a plena capacidade criativa e qualitativa da ferramenta de geração de texto-para-imagem.

De acordo com o *Midjourney Documentation* (2023) em seu Guia de Usuário, um *prompt* é definido como uma breve frase textual que o *Midjourney Bot*, ou apenas *Bot*, termo utilizado pelo Guia para se referir ao mecanismo responsável pelo processo, interpreta com o objetivo de gerar uma imagem. O *Midjourney Bot* fragmenta as palavras e frases do *prompt* em unidades menores, denominadas tokens. Esses tokens são então comparados com referências previamente treinadas em seu banco de dados. Assim, a elaboração cuidadosa de *prompts* pode influenciar diretamente a qualidade das imagens produzidas.

Como podemos perceber, esse processo não se resume em "procurar" imagens em um banco de dados a partir de palavras, pois não se apresenta como uma tradução direta de texto-para-imagem. Há uma série de operações matemáticas que interpretam textos como valores numéricos que representam características e padrões já previamente estabelecidas pelo treinamento anterior, chamado de aprendizagem de máquina, a partir do vasto banco de imagens. Outro ponto importante, já bem estabelecido entre a comunidade de usuários, é a ferramenta compreender outras línguas, embora com o seu treinamento tenha sido em inglês. Assim, os *prompts* mais assertivos são aqueles dados em língua inglesa.

O *Midjourney Bot* não processa palavras, gramática e estrutura de sentenças da mesma forma que humanos. Por essa razão, é aconselhável experimentar sinônimos em diferentes contextos para alcançar os melhores resultados. Palavras como "grande", "gigante", "enorme" e "imenso" podem ter efeitos distintos dependendo de como se relacionam com o contexto geral do *prompt*, resultando em imagens variadas. O principal foco ao elaborar um *prompt* deve ser detalhar "o que" e "onde". É essencial especificar o que se deseja, descrevendo aspectos como sujeitos, objetivos, iluminação, ambientes, cores, sentimentos e composição geral da cena (*MIDJOURNEY DOCUMENT*, 2023).

De acordo com o Guia de Usuário (2023), para um melhor funcionamento, recomenda-se o uso de sentenças descritivas, simples e curtas. No entanto, para *prompts* mais avançados, é possível incorporar elementos como imagens e emojis, que influenciarão o estilo da imagem resultante. Além do texto descritivo principal, pode-se adicionar parâmetros específicos, como estilos de arte, códigos que representam características de câmeras, incluindo modelo, abertura do obturador, velocidade do disparo, nomes de artistas de referência e configurações de diferentes versões do programa. Esses parâmetros têm o potencial de alterar significativamente o aspecto da imagem gerada.

É importante ressaltar que o processo de criação pelo *Midjourney* é guiado por diretrizes específicas, programadas para evitar a produção de imagens que possam ser consideradas desrespeitosas, promovam discurso de ódio, exibam violência explícita, nudez, sexualização de figuras públicas, conteúdo culturalmente ofensivo ou combinações de figuras públicas e eventos que possam gerar engano ou confusão.

Em abril de 2023, o *Midjourney* lançou um comando para descrever imagens em forma de *prompts*. Se utilizando da visão computacional, área da ciência da computação que extrai informações de imagens em uma espécie de leitura de imagens por computador, esse comando “permite que você faça o upload de uma imagem e gere quatro *prompts* possíveis baseados nesta imagem. Use o comando */describe* para explorar novos vocabulários e movimentos estéticos” (*MIDJOURNEY DOCUMENT*, 2023). O Manual de Usuário alerta que esses *prompts* gerados são inspiracionais e sugestivos, não devem ser entendidos como recriação exata da imagem descrita.

Neste estudo, ao empregar um trecho literário como *prompt*, desviando-se das diretrizes convencionais, buscamos explorar a capacidade interpretativa do *Midjourney Bot*. Nosso foco reside em compreender como esta ferramenta de Inteligência Artificial “lê” e “decodifica” nuances literárias, transformando-as em representações visuais. Através deste processo, pretendemos lançar luz sobre a interação entre literatura e tecnologia, bem como sobre a semiose emergente nesse cruzamento.

3 Método

Em um primeiro momento, houve a configuração da ferramenta, escolha do trecho literário, submissão do trecho como *prompt* e obtenção das imagens resultantes. Utilizamos a versão mais recente do *Midjourney*, o modelo 5.2, em um servidor privado. Optamos por configurações específicas: “Stylize very high”, “Public mode”, “High Variation Mode” e “Fast mode”. A intenção era fazer o mínimo de intervenções, pois nosso principal objetivo era

compreender a interpretação que o *Bot* faria do trecho literário. Por essa razão, não adicionamos imagens de referência ou outros parâmetros, confiando exclusivamente no excerto literário como instrução.

Por padrão, a ferramenta gera uma única imagem formada por um conjunto de quatro cenas distintas. Optamos por tratar do resultado como uma única imagem e chamar as quatro representações – opções de imagem para o *prompt* dentro dessa única imagem – de cenas. Há *botões*, abaixo do resultado gerado, que permitem solicitar uma versão de qualidade superior (U1, U2, U3, U4) de cada uma das cenas individualmente ou optar por variações estilísticas específicas (V1, V2, V3, V4) também de forma individual para cada. Além disso, há uma opção para refazer a imagem, em uma nova interpretação do *prompt*, caso o resultado não tenha sido satisfatório. Essas funcionalidades não foram acionadas para o estudo.

O romance *Um Cão Late na Noite* narra a trajetória de dois jornalistas, um homem e uma mulher, que se unem na elaboração de uma reportagem que os afeta de maneiras distintas. A relação entre eles se fortalece gradualmente à medida que discutem a seleção de fontes, a abordagem sensível da notícia e os impactos que esperam gerar com sua publicação. A narrativa coloca em discussão os limites entre o jornalismo e a literatura, tudo isso ambientado no universo da imprensa e permeado por uma atmosfera erudita. O trecho selecionado como *prompts* é originário do primeiro capítulo, denominado "Tantas Histórias". Nesse capítulo, o leitor é apresentado ao personagem Joaquim de Assis, que, em meio a uma cena rotineira, é envolto em lembranças de um passado marcado por um trauma.

A escolha do romance *Um cão late na noite*, de autoria de Sérgio Arruda, para este trabalho fundamenta-se em critérios metodológicos e objetivos de pesquisa definidos. A inclusão desta obra literária, cujo autor é coautor do presente artigo, possibilita um acesso privilegiado ao entendimento profundo das nuances, intenções e contextos subjacentes ao texto, elementos esses cruciais para a adequada interpretação e posterior geração de imagens IAG. Tal proximidade com o material fonte é fundamental para garantir a fidelidade e a profundidade na análise das imagens geradas, permitindo uma avaliação mais precisa de como a ferramenta Midjourney interpreta e visualiza os elementos narrativos e simbólicos presentes no romance.

Na segunda etapa do estudo, conduzimos uma análise semiótica da imagem gerada, com ênfase na interpretação dos signos e na potencial expansão ou resignificação do texto original, tanto em português quanto em sua tradução para o inglês. Levando em conta que o Midjourney gera um único arquivo contendo quatro representações visuais distintas em

resposta ao *prompt*, decidimos avaliar essas imagens como um conjunto coeso dadas as suas semelhanças estilísticas e composicionais. No entanto, também nos dedicamos a explorar as especificidades de cada imagem, considerando suas inter-relações.

Para auxiliar na identificação e classificação dos signos presentes na imagem, recorreremos à tríade de Peirce (PEIRCE, 1974). Além disso, dedicamos especial atenção à análise das relações entre os signos, à maneira como eles se combinam para criar um significado global e à existência de significados culturais específicos (BARTHES, 1996, 2012; ECO, 1991; FONTANARI, 2016), sempre fundamentados no método de análise de Santaella (2015).

O método de análise de Santaella (2015) se desdobra em três dimensões principais. Primeiro, analisa-se o fundamento do signo, onde se observa a capacidade de contemplação, distinção e generalização do signo. Em seguida, a referencialidade do signo é explorada, focando nas qualidades concretas e abstratas, no contexto em que o signo opera e na sua coerência cultural e representatividade. Por fim, o processo interpretativo aborda o potencial interpretativo do signo, o efeito real que ele tem sobre um intérprete e uma concepção teórica do resultado interpretativo final. Esta abordagem permite uma compreensão profunda dos signos e suas inter-relações no contexto da semiótica.

4 Resultados

4.1 *Prompt* original em português

O trecho escolhido foi o primeiro parágrafo que abre o primeiro capítulo ao descrever um cenário e seu personagem. Portanto, o *prompt* dado ao *Midjourney Bot* foi:

Na praça circundada por prédios de apartamentos, certo homem acomodado em um banco, tão cedo da manhã, tenta ler um jornal. Seus esforços em dobrar as vastas páginas para melhor concentrar os olhos são arruinados pelo vento, o mesmo vento que faz balançar a folhagem das amendoeiras. Também soprava um vento de realismo mágico naquele dia fatal espalhando folhas em redemoinho. Joaquim de Assis afastou da cabeça a imagem do filho caído morto aos pés daquele banco e se concentrou na cena logo abaixo (MOURA, 2022, p. 7).

A seguir, apresentamos o resultado obtido.



Figura 1 - Resultado da geração de imagem por *Midjourney* (2023) e *prompt* de Arruda (2022). Numeração inserida pelos autores. Fonte: dos autores

As quatro cenas produzidas, presentes na Figura 1, trazem uma estética de realismo urbano, ao remeter a fotografia a um ambiente com calçada, carros, pessoas e paredes ao fundo, retratando indivíduos que, à primeira vista, parecem ser moradores em situação de rua em local de baixa temperatura evidenciado pelos casacos. É notável que três das cenas retratam homens que aparentam estar na meia-idade, enquanto um parece ser mais jovem. Há uma predominância de tons escuros, prevalecendo preto e cinza, nas imagens, tanto nas vestimentas dos retratados quanto nos objetos que compõem a cena.

De maneira geral, os personagens retratados nas cenas transmitem uma aparência de descaso com a higiene, especialmente devido ao cenário circundante, que apresenta elementos que lembram lixo se acumulando. Nas duas cenas superiores, os personagens seguram objetos associados à leitura: na primeira, um jornal e, na segunda, um celular. As mãos retratadas nas cenas inferiores apresentam deformidades, com dedos em excesso ou em falta. Em todas as representações, há um contraste marcante entre a predominância de cores escuras e detalhes em vermelho, seja nas roupas, objetos ou no cenário ao fundo. O semblante dos personagens é uniformemente triste ou desanimado.

Apesar de seu hiper-realismo fotográfico, a deformidade nas mãos é um problema já identificado em IAG de imagens. Uma dica comum, enquanto essa questão não é completamente resolvida, é observar as mãos para identificar imagens geradas por IA. Esse

desafio surge devido ao baixo número de cenas de mãos nos bancos de dados de treinamento das IAG. Contudo, acreditamos que avanços técnicos futuros resolverão essa questão.

O resultado gerado divergiu das expectativas iniciais, uma vez que as cenas são significativamente diferentes das concepções mentais formadas pelos autores do presente estudo, ao ler o trecho do livro e apresentarem cenas, numa primeira vista, nada relacionadas com o que foi descrito. Após superar essa surpresa inicial e ao detalhar a imagem e seus signos, identificamos elementos do texto que podem ter sido reinterpretados e reconstruídos no conjunto de imagens apresentado.

De alguma forma, a descrição "praça circundada por prédios de apartamentos" foi traduzida como um cenário urbano. Embora o "jornal" apareça em apenas uma das cenas, o ato de "ler" é representado na segunda cena, onde o personagem observa um dispositivo eletrônico. É interessante observar que o jornal está associado ao personagem de meia-idade, enquanto o dispositivo eletrônico está com o homem mais jovem.

Podemos traçar uma analogia entre o trecho "concentrar os olhos" e os olhares "distantes" e focados, quase que perdidos, em duas cenas (1 e 3) com personagens de meia-idade que olham para fora da cena, enquanto na cena 4 o personagem quebra a "quarta parede", encarando diretamente o espectador com semblante de tristeza.

Outra relação possível é entre o trecho "filho caído morto" e as cores vermelho e preto presentes em todas as cenas. O preto, na cultura ocidental, é associado ao luto, e o vermelho evoca o sangue, o que pode estar reforçando a memória da morte e a imagem mental do filho falecido como derramamento de sangue. Como as diretrizes internas do Midjourney impedem a representação de violência explícita, esse caminho representativo pode ter sido adotado pelo *Bot*.

Fazendo o caminho inverso, ao dar o comando para descrever a imagem (/describe), fizemos o *upload* desta imagem gerada e recebemos a seguinte descrição (em inglês):

1. Quatro fotos de um morador de rua em uma calçada sentado em um banco, no estilo cenas indianas, nikon d850. 2. Seis fotos de um sem-teto sentado, no estilo cenas de rua sobrepostas, cultura pop indiana, fotos espontâneas de figuras famosas, cinza e vermelho, telas múltiplas. 3. Três pessoas sentadas em frente a dois vendedores ambulantes com sacolas na cabeça, no estilo de composições de vários painéis, fotografia espontânea, imagem uhd, Pierce Brosnan, gestos emocionais, saurabh jethni, metrópoles sombrias. 4. Homens na estrada ouvindo seus celulares, no estilo de dípticos narrativos, nikon d850. **[Tradução dos autores]** (MIDJOURNEY BOT, 2023)

Ao analisar a descrição da imagem feita pelo *Midjourney bot*, a primeira impressão é de objetividade, alinhada ao modelo descrito no manual do usuário sobre *prompts*. A descrição apresenta sentenças curtas, especificando "o que" e "onde", complementadas por parâmetros estilísticos, com características estéticas como "cultura pop indiana" e "estilo dísticos", referências humanas como o ator Pierce Brosnan e o artista gráfico Saurabh Jethni e modelo de câmera nikon d85.

Embora o *bot* demonstre dificuldade em identificar elementos factuais como o número exato de pessoas na imagem ou inventando elementos como "dois vendedores ambulantes com sacolas na cabeça", ele surpreende ao interpretar aspectos mais subjetivos, como identificar "morador de rua" – uma leitura que coincidiu com a nossa análise.

Conforme o *Midjourney Document (2023)*, o objetivo desse comando é estimular "possibilidades de *prompts*" e auxiliar na exploração de "novo vocabulário e estilos". Assim, fica evidente que a intenção não é fornecer uma descrição exata da imagem apresentada, mas sim sugestões provocativas que inspirem a criação de novas imagens.

Embora seja esta nossa conclusão, certos elementos da descrição se destacam por se alinharem de alguma maneira com nossa análise feita anteriormente. A ênfase nas cores "cinza e vermelho" e a unidade perceptiva sugerida pelos "gestos emocionais" coincidem com nossa interpretação de olhares perdidos e tristes, barbas por fazer, cabelo despenteado e uma certa poluição visual talvez exteriorizando uma confusão interna dos personagens. A menção a "metrópoles sombrias" também se alinha à nossa leitura de um cenário escuro e abandonado, evidenciado pela presença de suposto lixo acumulado em uma calçada.

4.2 *Prompt* traduzido para o inglês

Como já mencionado, há um consenso que os *prompts* são mais assertivos quando utilizados na língua inglesa. Assim, traduzimos o trecho escolhido para o inglês para utilizá-lo como comando.

In the square surrounded by apartment buildings, a certain man seated on a bench, so early in the morning, tries to read a newspaper. His efforts to fold the vast pages to better focus his eyes are ruined by the wind, the same wind that makes the foliage of almond trees sway. A wind of magical realism also blew on that fateful day, scattering swirling leaves. Joaquim de Assis took the image of his son who had fallen dead at the foot of that bench out of his head and concentrated on the scene below (MOURA, 2022, p. 7).

A Figura 2 apresenta o resultado obtido.



Figura 2 - Resultado da geração de imagem por *Midjourney* (2023) e *prompt* de Arruda (2022). Numeração inserida pelos autores. Fonte: dos autores

Assim como no primeiro conjunto de cenas, percebe-se uma coesão visual nas cenas apresentadas na Figura 2, principalmente pela presença constante de árvores com folhas ou flores destacadas. Estas cores, que variam entre amarelo, laranja, rosa, marrom e ocre, contrastam harmoniosamente com a iluminação solar que incide diretamente sobre os personagens nas três primeiras cenas. Na quarta cena, o personagem é iluminado de forma contrária, com a luz vindo do fundo, gerando uma sombra que confere um tom dramático à composição. Todas as imagens remetem a uma praça-jardim, e podemos notar que, nas três primeiras, os personagens seguram objetos que parecem ser de leitura: um jornal na primeira, folhas soltas na segunda e um livro na terceira. Na última cena, a personagem segura algo que parece ser um copo.

Há também um equilíbrio na aparência dos personagens. Enquanto duas cenas (1 e 4) retratam homens de meia-idade, as outras duas (2 e 3) apresentam jovens. As representações dos homens mais velhos tendem ao realismo fotográfico, enquanto as dos jovens se assemelham mais a ilustrações. Esta diferença é reforçada pelas vestimentas: os homens mais velhos usam ternos ajustados, óculos e têm cabelos brancos, enquanto os jovens vestem roupas mais largas e casuais, com cabelos ondulados que parecem ser movidos pelo vento.

Nas cenas 2, 3 e 4, observamos construções residenciais com mais de dois andares, possivelmente aludindo ao trecho “uma praça cercada de prédios”. Notamos que os homens mais velhos estão acomodados em bancos de praça facilmente identificáveis, enquanto os

locais onde os jovens estão sentados geram dúvidas – poderiam ser escadas, beiradas de alguma construção ou outros espaços. A menção no *prompt* de que a cena ocorre “tão cedo de manhã” é evidenciada pela iluminação solar, que cobre a cena com tons amarelados. Este conjunto de imagens captura com mais precisão o conceito de “leitura”, presente em três das cenas (1, 2 e 3), e as flores de amendoeira nas cenas 1 e 4 fazem uma clara conexão com a referência a essa árvore no *prompt*.

Na cena 2, o personagem parece estar tentando organizar páginas soltas, o que pode ser uma alusão ao trecho “esforço para dobrar as páginas para melhor concentrar os olhos”. Além disso, o cabelo deste personagem e do personagem da cena 3 parecem ser movidos pelo vento, possivelmente remetendo ao trecho “arruinados pelo vento, o mesmo vento que fez balançar a folhagem das amendoeiras”. É interessante observar que as cenas 2 e 3, que possuem uma estética mais ilustrativa, podem estar relacionadas à menção do “vento de realismo mágico” no *prompt*. Em todas as cenas, os personagens parecem focados em um ponto específico, refletindo o trecho “concentrar os olhos” do *prompt*.

Embora as cenas sejam banhadas pela luz matinal, há um tom melancólico, evidenciado pelas sombras nas cenas com os personagens mais velhos (1 e 4) e pela postura corporal mais retraída e vestimentas despojadas dos personagens mais jovens (2 e 3). A sensação de solidão é uma constante em todas as cenas, reforçada pela composição, expressões faciais e pela ausência de outras pessoas, exceto na cena 1, onde algumas pessoas desfocadas aparecem ao fundo. Em um esforço interpretativo, podemos relacionar essa melancolia com a imagem do filho morto que invade a mente do personagem descrito no trecho.

Fazendo o caminho inverso, ao dar o comando para descrever imagens (/describa), fizemos o *upload* da imagem gerada e recebemos a seguinte descrição (em inglês):

1. Quatro fotos com um homem lendo em um banco do lado de fora, no estilo de arte conceitual, ultrarrealista, tons quentes, não acredito como isso é lindo, ilustrações detalhadas de personagens, ferramentas de pintura e escrita, cenários cinematográficos. 2. Dois homens de óculos sentados nas folhas da rua, no estilo renderizações realistas e hiper detalhadas, e ventrue, composições de vários painéis, shilin Huang, cenas cinematográficas, laranja e marrom, foto de alta qualidade. 3. Quatro fotos de homens lendo jornais, no estilo cenário realista, kawacy, motor irreal 5, cenas naturais detalhadas, laranja, geração de imagens, yeong-hao han. 4. Um dos muitos homens sentados a ler livros e jornal, ao estilo paisagens foto realistas, personagens de anime, castanho e âmbar, flores de cerejeira, cenas urbanas foto realistas, laranja e bronze, cenários cinematográficos. **[Tradução dos autores]** (*Midjourney Bot*, 2023)

Os *prompts* gerados após o *upload* da imagem mantêm a objetividade e padronização observadas na análise da primeira imagem. Eles incorporam referências a movimentos artísticos, como "arte conceitual", e destacam parâmetros estilísticos específicos, como qualidade desejada em "cenas cinematográficas" e "foto realistas" e referências a artistas digitais, como "Eve Ventrue", "Shilin Huang" e "Yeong-hao Han".

Verificamos que a descrição fornecida não é exata ao falhar em identificar o número correto de pessoas nas cenas. No entanto, algumas características que consideramos marcantes, como "homem lendo em um banco", "homens lendo jornais" e "homens sentados a ler livros e jornal", permaneceram consistentes. Essa repetição sugere que tal característica é central na interpretação do *bot*, destacando-se como o principal elemento narrativo neste contexto.

Temos elementos de composição comuns, como as cores predominantes descritas como "tons quentes", "laranja e marrom" e "castanho a âmbar". A descrição "ultrarrealista", "foto realista", "cenas realistas" e "cena cinematográfica" da imagem, indica que as cenas que simulam fotografias foram as mais evidenciadas. No entanto, o quarto *prompt* faz referência a cenas no estilo de ilustração, mencionando "personagens de anime".

O que se destaca nas descrições geradas é a forte conexão com a cultura digital, aludindo a artistas associados a videogames e terminologias típicas desse universo, como "renderizações". É evidente a influência de dois grandes pilares culturais da internet: *videogames* e *animes*. A menção a esses artistas e estilos de ilustração não parece ser coincidência. Vale ressaltar que, embora as flores de amendoeira e cerejeira sejam semelhantes e ambas originárias da Ásia, o *bot* associou a imagem das flores descrevendo-a como cerejeira após mencionar *animes*. A cerejeira, com sua flor sakura, é um ícone da cultura japonesa, chegando a ser considerada um símbolo nacional.

5 Considerações Finais

A interseção entre literatura e inteligência artificial, especificamente no contexto de imagens a partir de textos literários, é um campo fértil e intrigante de investigação. O estudo aqui apresentando, ao utilizar trechos do romance *Um Cão late na noite* com *prompt* para o *Midjourney Bot*, revelou nuances surpreendentes sobre como as IAG interpretam e visualizam textos literários.

Primeiramente, é evidente que a literatura, com as nuances de linguagem, não se enquadra perfeitamente nas diretrizes tradicionais para criação de *prompts* eficazes. No entanto, essa "inadequação" se mostrou uma força, não uma fraqueza. Ao desafiar a IAG

com um trecho de texto literário, pudemos explorar a flexibilidade e a adaptabilidade da tecnologia, revelando tanto suas limitações quanto seus potenciais inesperados.

A análise semiótica das imagens geradas, baseadas no método proposto por Santaella (2015), permitiu uma compreensão dos signos presentes e de como eles se relacionam com o texto original. Ficou claro que, enquanto a IAG pode não capturar todas as nuances e intenções do autor original, ela oferece uma interpretação única, filtrada através de sua vasta base de dados e algoritmos.

Além disso, a descrição objetiva gerada pelo *Midjourney Bot* sobre as imagens produzidas revelou uma tentativa da IAG de se alinhar às diretrizes de um “bom” *prompt*, mesmo quando confrontada com um texto literário. Isso sugere uma tentativa da máquina de encontrar um equilíbrio entre programação e o input fornecido.

A colaboração com o autor do romance proporcionou uma perspectiva valiosa, destacando a distância entre a imaginação humana e a interpretação da máquina. No entanto, essa distância não é necessariamente negativa. Em vez de ver a IAG como uma ferramenta que simplesmente reproduz a intenção humana, podemos começar a compreendê-la como uma entidade colaborativa que traz sua própria perspectiva e contribui com nossa criatividade.

Não podemos negligenciar o debate em curso sobre as implicações éticas e legais do uso de IAG no campo da criação artística. Há preocupações legítimas sobre a possibilidade de essas ferramentas reproduzirem estéticas de artistas sem autorização e remuneração, levantando questões sobre originalidade e plágio. Além disso, a utilização de bancos de dados privados para treinar essas máquinas intensifica o debate, pois muitas vezes esses dados contêm imagens e obras de terceiros, que são usadas sem permissão e sem compensação.

Essa apropriação da produção visual de outros para treinar máquinas levanta questões cruciais sobre direitos autorais, propriedade intelectual e a ética da criação assistida por IA. É essencial que, à medida que avançamos na exploração das capacidades das IAG, também nos aprofundemos nas discussões sobre como garantir práticas justas e éticas no domínio da criação digital.

Este estudo destaca o potencial da IAG não apenas como uma ferramenta de reprodução, mas como uma entidade criativa em si. Ao desafiar os limites tradicionais do que constitui um “bom” *prompt* e ao explorar a interação entre literatura e tecnologia, abrimos a porta para novas formas de expressão e compreensão. A jornada entre texto e imagem, entre autor e máquina, é repleta de descobertas, e este estudo serve como uma

demonstração do vasto território ainda a ser explorado nesse cruzamento entre humanidade e inteligência artificial.

Dada a natureza exploratória e os resultados deste estudo preliminar, torna-se importante o desenvolvimento de trabalhos futuros que ampliem o escopo de investigação e testem a replicabilidade das descobertas em contextos variados e com diferentes obras literárias. A pesquisa atual, ao delinear um método de análise semiótica aplicado às imagens geradas por IAG a partir de textos literários, estabelece um marco inicial para futuras explorações. No entanto, é crucial que outros estudiosos da área possam aplicar, adaptar e expandir este método para verificar a consistência dos resultados e explorar novas dimensões da relação entre literatura e inteligência artificial.

A reprodutibilidade desta pesquisa por outros acadêmicos não apenas validaria as descobertas apresentadas, mas também enriqueceria o diálogo interdisciplinar, contribuindo com novas perspectivas e *insights* sobre as capacidades criativas das IAG e as complexas interações entre texto e imagem. Além disso, estudos futuros deveriam abordar as limitações observadas neste trabalho, explorando, por exemplo, técnicas avançadas de engenharia/*design* de *prompt* ou o uso de diferentes ferramentas de IAG para a geração de imagens, a fim de obter uma compreensão mais profunda das potencialidades e restrições dessa tecnologia.

Essencialmente, a continuidade da pesquisa neste campo deverá também considerar as implicações éticas e legais que emergem do uso de inteligência artificial na criação artística, propondo diretrizes para práticas responsáveis e éticas. A colaboração entre artistas, cientistas da computação, juristas e filósofos será fundamental para desenvolver um entendimento compartilhado e normas que respeitem os direitos de autor e promovam uma inovação consciente e ética.

O caminho adiante requer uma abordagem colaborativa e multidisciplinar, onde a tecnologia da inteligência artificial não seja vista meramente como uma ferramenta de auxílio à criação humana, mas como um campo de estudos em si, com potencial para transformar radicalmente nossa percepção de criatividade, autoria e arte. Este estudo contribui para essa jornada de descoberta, convidando a comunidade acadêmica a participar ativamente na moldagem do futuro da criação artística mediada por IA.

Referências

BARTHES, Roland. **Elementos de semiologia**. Trad. Izidoro Blikstein. São Paulo: Cultrix, 1996.

BRITAIN, Blake. **U.S. Copyright Office says some AI-assisted works may be copyrighted.** Disponível em: < <https://www.reuters.com/world/us/us-copyright-office-says-some-ai-assisted-works-may-be-copyrighted-2023-03-15/>>. Acesso em: 06 de jul de 2023.

BORJI, Ali. **Generated faces in the wild: Quantitative comparison of stable diffusion, Midjourney and dall-e 2.** Disponível em: <arXiv preprint arXiv:2210.00586>. Acesso em: 06 de jul de 2022.

DISCHLER, Jerry. Introducing a new era of AI-powered ads with Google. Disponível em: < https://blog.google/products/ads-commerce/ai-powered-ads-google-marketing-live/?utm_campaign=the_brief_newsletter_-_28_de_julho_de_2023&utm_medium=email&utm_source=RD+Station>. Acesso em 30 de jul de 2023.

ECO, Umberto. **Tratado geral de semiótica.** São Paulo: Perspectiva, 1991.

ÉPOCA NEGÓCIOS. Maior agência de publicidade do mundo adota IA para criar campanhas. Disponível em: < <https://epocanegocios.globo.com/empresas/marketing/noticia/2023/05/major-agencia-de-publicidade-do-mundo-adota-ia-para-criar-campanhas.ghtml>>. Acesso em 30 de jul de 2023.

FONTANARI, R. **Como ler imagens? A lição de Roland Barthes.** Galáxia (São Paulo), p. 144–155, 2016.

LI, Yuheng et al. **Generate Anything Anywhere in Any Scene.** Disponível em: <arXiv preprint arXiv:2306.17154>. Acesso em: 06 de jul de 2023.

MIDJOURNEY DOCUMENTATION. User Guide. Disponível em: < <https://docs.Midjourney.com/docs/quick-start>>. Acesso em 30 de jul de 2023.

MOURA, Sérgio Arruda de. **Um cão late na noite.** Guaratinguetá, SP: Editora Penalux, 2022.

OPPENLAENDER, Jonas. **The creativity of text-to-image generation.** In: Proceedings of the 25th International Academic Mindtrek Conference. 2022. p. 192-202.

PEIRCE, C. S. **Collected Papers of Charles Sanders Peirce.** [s.l.] Harvard University Press, 1974.

ROOSE, Kevin. **An AI-generated picture won an art prize. Artists aren't happy.** The New York Times, v. 2, n. September, 2022.

SANTAELLA, Lucia. **A IA e a quarta ferida da humanidade.** SBC Horizontes, 20 abr. 2023. ISSN 2175-9235. Disponível em: <<http://horizontes.sbc.org.br/index.php/2023/04/quarta-ferida>>. Acesso em: 06 de jul. de 2023.

SANTAELLA, L. **Semiótica aplicada.** São Paulo: Cengage Learning, 2015.