

**Eixo Temático: GT5 – Propostas de instrumentos para diagnóstico,
metodologias e ferramentas de gestão pública**

***Inpainting* no desenho urbano em Campos dos Goytacazes/RJ:
Inteligência Artificial como ferramenta para pensar cidades e
comunidades sustentáveis**

*Inpainting in Urban Design in Campos dos Goytacazes: Artificial Intelligence as
a tool for thinking about sustainable cities and communities*

Gabriel Barcelos e Silva¹
Sergio Rafael Cortes de Oliveira²

RESUMO

Os avanços recentes das Inteligências Artificiais Generativas de Imagens ampliam as possibilidades para pensar o planejamento urbano, em especial o *inpainting*, como potencial para a melhoria na comunicação no planejamento urbano. Portanto, o objetivo geral desta pesquisa é analisar as capacidades dessa ferramenta a fim de melhorar a comunicação do planejamento urbano na cidade de Campos dos Goytacazes, no Rio de Janeiro. A metodologia combina uma análise bibliográfica e pesquisa experimental. Pesquisas sobre o uso de Redes Neurais Generativas de imagens (RNGI) ainda são principiantes e escassas em escala nacional e internacional. Foram conduzidos sete experimentos empregando a técnica de *inpainting* para propor soluções visuais com o propósito de contribuir para um planejamento urbano mais sustentável. Cada experimento foi estruturado para responder a metas específicas do Objetivo de Desenvolvimento Sustentável 11, que visa promover cidades mais sustentáveis e inclusivas, aplicando tal tecnologia em diferentes cenários em Campos dos Goytacazes/RJ. Embora a democratização do uso das RNGI ainda esteja em fase inicial, o que acaba limitando sua aplicação no setor profissional, o estudo indica que essas ferramentas podem aprimorar a comunicação entre urbanistas e a população, alcançando um planejamento mais inclusivo e

¹ Mestrando em Arquitetura, Urbanismo e Tecnologias, IFFluminense. E-mail: gbarcelosesilva587@gmail.com

² Doutor em Engenharia Civil, UENF. Professor no mestrado em Arquitetura, Urbanismo e Tecnologias, IFFluminense. E-mail: s.rafaelcortes@yahoo.com.br

claro. Contudo, ainda ressalta a importância de uma interpretação especializada no desenho urbano técnico. O percurso metodológico busca alcançar contribuições na comunicação entre urbanistas e a população, promovendo um planejamento mais inclusivo e facilitando a participação cidadã, além de expandir a investigação para os diversos desafios urbanos, em busca de uma vivência mais sustentável nas cidades.

PALAVRAS-CHAVE: Inteligência Artificial; *inpainting*; planejamento urbano.

ABSTRACT

The recent advances in Generative Image Artificial Intelligences expand the possibilities for thinking about urban planning, particularly inpainting, as a potential tool to improve communication in urban planning. Therefore, the general objective of this research is to analyze the capabilities of this tool to enhance communication in urban planning in the city of Campos dos Goytacazes, Rio de Janeiro. The methodology combines a bibliographic analysis and experimental research. Studies on the use of Generative Neural Networks for images are still in their early stages and remain scarce both nationally and internationally. Seven experiments were conducted using the inpainting technique to propose visual solutions aimed at contributing to more sustainable urban planning. Each experiment was structured to address specific targets of Sustainable Development Goal 11, which seeks to promote more sustainable and inclusive cities, applying this technology in various scenarios in Campos dos Goytacazes/RJ. Although the democratization of this technology usage is still in its initial phase, limiting its application in the professional sector, the study indicates that these tools can enhance communication between urban planners and the population, leading to more inclusive and clear planning. However, it also emphasizes the importance of specialized interpretation in technical urban design. The methodological approach aims to contribute to improving communication between urban planners and the population, fostering more inclusive planning and facilitating citizen participation, while expanding the investigation to address various urban challenges in pursuit of more sustainable urban living.

KEYWORDS: Artificial Intelligence; *inpainting*; urban planning.

1 INTRODUÇÃO

No campo do planejamento urbano, a comunicação efetiva é fundamental para o sucesso de projetos e iniciativas. A utilização de imagens não apenas enriquece a apresentação das ideias, mas também desempenha um papel vital na criação de uma conexão mais profunda entre os diversos agentes da sociedade, como planejadores, gestores e membros da comunidade. Nesse contexto, a transmissão de informações de forma eficiente e clara potencializa a participação ativa de todos os envolvidos no processo.

As imagens desempenham um papel fundamental na simplificação de ideias complexas, facilitando o diálogo entre planejadores, gestores e a comunidade. Elas tornam os projetos mais claros e acessíveis, eliminando ambiguidades e permitindo uma interpretação intuitiva das propostas. O impacto

das imagens está intrinsecamente ligado ao funcionamento do cérebro humano na absorção de informações. Segundo Lynell Burmark (2002), em seu livro *Visual Literacy: Learn to See, See to Learn*, a cognição humana é profundamente influenciada por estímulos visuais. Enquanto palavras são processadas na memória de curto prazo, as imagens são armazenadas na memória de longo prazo, tornando-se mais duradouras.

No planejamento urbano, o uso de imagens atrai a atenção e gera respostas emocionais, estabelecendo uma conexão mais profunda e transparente entre as propostas urbanísticas e o público. Quando as propostas são apresentadas de forma visual e clara, os envolvidos conseguem analisar suas consequências e benefícios com maior facilidade, promovendo um ambiente de colaboração e diálogo.

Nesse sentido, manifesta-se a possibilidade de aplicar a Inteligência Artificial (IA) generativa, especialmente com os avanços recentes em sistemas baseados em aprendizado profundo (*deep learning*), como as Redes Neurais Generativas de Imagens (RNGI). Para compreender o potencial dessa tecnologia, é fundamental investigar o funcionamento das RNGI e como elas podem ser integradas ao planejamento urbano.

O atual cenário de transformação no planejamento urbano ressalta a necessidade de aprimorar a comunicação entre os diferentes atores envolvidos no processo. Nas últimas décadas, a necessidade de criar ambientes sociais mais inclusivos, resilientes, seguros e ecologicamente corretos revelou a importância de repensar criticamente como comunicamos nossas estratégias e intenções. Nesse contexto, os Objetivos de Desenvolvimento Sustentável (ODS) – um conjunto de 17 metas globais estabelecidas pela ONU em 2016, com o intuito de enfrentar os principais desafios globais – oferecem uma estrutura para um planejamento mais sustentável, considerando aspectos ambientais, econômicos, sociais e culturais. Dentre esses, o ODS 11, que trata de Cidades e Comunidades Sustentáveis, ganha destaque ao propor diretrizes que visam transformar os espaços urbanos e assentamentos humanos em ambientes mais inclusivos.

Esse debate emergente destaca a importância de um planejamento urbano que vá além das discussões técnicas e acadêmicas. Apesar de debates entre especialistas serem fundamentais para o desenvolvimento urbanístico, há a

crecente necessidade de uma abordagem mais acessível e atrativa para a sociedade, com a finalidade de influenciar a participação pública ampla e informada, facilitada pelo uso de representações visuais.

2 OBJETIVO

O objetivo geral desta pesquisa é analisar as capacidades da ferramenta de *inpainting*, resultante dos modelos de Redes Neurais Generativas de Imagens a fim de melhorar a comunicação do planejamento urbano na cidade de Campos dos Goytacazes/RJ.

3 METODOLOGIA / MÉTODO DE ANÁLISE

O percurso metodológico desta pesquisa baseia-se na obra de Gil (2019), que oferece um conjunto metodológico para a construção de projetos de pesquisa. Seguindo este autor, tem-se uma pesquisa aplicada, de caráter exploratório e descritivo e de abordagem qualitativa. Entre os métodos abordados pelo autor, destacam-se a pesquisa bibliográfica, que possibilita uma revisão crítica e detalhada da literatura existente, e a pesquisa experimental, que permite testar as hipóteses por meio de experimentos controlados. A combinação desses métodos forma a estrutura do desenvolvimento do estudo em questão.

Conforme enfatizado por Gil (2019), a pesquisa bibliográfica é uma etapa primordial, indispensável em quase todos os trabalhos científicos, pois dá suporte ao caráter exploratório do estudo, oferecendo uma visão abrangente sobre o conhecimento atual em determinada área. Dada a contemporaneidade do tema abordado no artigo, se fez crucial que a revisão fosse realizada a partir de fontes recentes, com a finalidade de assegurar que os dados e as teorias reflitam os avanços atuais. Nesse sentido, as publicações periódicas são especialmente importantes, principalmente revistas científicas e repositórios de artigos que facilitam o acesso eficiente e rápido a informações, em vez de publicações tradicionais, como livros. Isso torna a pesquisa bibliográfica uma ferramenta primordial para compreender e contextualizar os tópicos discutidos, especialmente em campos que evoluem rapidamente. Artigos científicos oferecem discussões importantes sobre ferramentas como o *inpainting* e

fornecem estratégias para a incorporação de tecnologias de IA no dia a dia dos urbanistas.

Com o arcabouço estruturado pela pesquisa bibliográfica, foi possível delimitar e estabelecer a pesquisa experimental. Para Gil (2019), a produção de um experimento passa pela identificação do objeto de estudo, das variáveis capazes de influenciar esse objeto e do modo de controle e a forma de observação. Ainda segundo o autor, essa delimitação é essencial para garantir a reprodutibilidade do experimento por outros pesquisadores, um passo fundamental na produção do conhecimento científico.

O foco central dos experimentos, os objetos da pesquisa, foram imagens coletadas do programa Google *Earth*, na ferramenta Google *Street View*, com o objetivo de demonstrar o potencial de serem utilizadas como representações visuais para comunicar estratégias de planejamento urbano. A escolha por essa ferramenta se deu pela sua acessibilidade e abrangência no território de recorte da pesquisa, o município de Campos dos Goytacazes, no norte do Rio de Janeiro. Desse modo, é possível capturar uma variedade de contextos urbanos, sem necessitar visitas presenciais. Além disso, o Google *Street View* oferece uma perspectiva imersiva e cotidiana, possibilitando uma análise aprofundada de um dia comum de cada região estudada, sem manipulação de imagem.

As imagens destacam problemas urbanos em diferentes regiões de Campos dos Goytacazes/RJ, com situações que demandam intervenção do poder público. Para orientar o experimento de forma mais precisa, as metas do décimo primeiro Objetivo de Desenvolvimento Sustentável (ODS) – Cidades e Comunidades Sustentáveis (ONU, 2016) – serviram como referência na seleção de imagens. Foram escolhidas aquelas que ilustram de forma clara as fragilidades que cada meta busca solucionar. As metas, que vão de 11.1 a 11.7, estão listadas a seguir:

11.1 Até 2030, garantir o acesso de todos à habitação segura, adequada e a preço acessível, e aos serviços básicos e urbanizar as favelas

11.2 Até 2030, proporcionar o acesso a sistemas de transporte seguros, acessíveis, sustentáveis e a preço acessível para todos, melhorando a segurança rodoviária por meio da expansão dos transportes públicos, com especial atenção para as necessidades das pessoas em situação de vulnerabilidade, mulheres, crianças, pessoas com deficiência e idosos

11.3 Até 2030, aumentar a urbanização inclusiva e sustentável, e as capacidades para o planejamento e gestão de assentamentos humanos participativos, integrados e sustentáveis, em todos os países

11.4 Fortalecer esforços para proteger e salvaguardar o patrimônio cultural e natural do mundo

11.5 Até 2030, reduzir significativamente o número de mortes e o número de pessoas afetadas por catástrofes e substancialmente diminuir as perdas econômicas diretas causadas por elas em relação ao produto interno bruto global, incluindo os desastres relacionados à água, com o foco em proteger os pobres e as pessoas em situação de vulnerabilidade

11.6 Até 2030, reduzir o impacto ambiental negativo per capita das cidades, inclusive prestando especial atenção à qualidade do ar, gestão de resíduos municipais e outros

11.7 Até 2030, proporcionar o acesso universal a espaços públicos seguros, inclusivos, acessíveis e verdes, particularmente para as mulheres e crianças, pessoas idosas e pessoas com deficiência (ONU, 2016, s.p.).

Foram selecionadas imagens de diversas regiões do município, além do distrito-sede, a fim de registrar a diversidade geográfica e urbana de Campos dos Goytacazes/RJ, com o objetivo de ilustrar tanto os desafios enfrentados em termos de infraestrutura e sustentabilidade, quanto os cenários com potenciais de melhorias. A Figura 1 apresenta as localidades selecionadas, demarcando-se numericamente as metas do Objetivo 11 a qual estão relacionadas.

Figura 1 – Regiões selecionadas para o estudo no município de Campos dos Goytacazes/RJ



Fonte: Google Earth, modificada pelos autores (2024).

Na seção de Resultados, são apresentadas as imagens, as áreas mascaradas e os resultados obtidos com a técnica de *inpainting*, além dos *prompts* utilizados para gerar as novas imagens, ou seja, o texto de entrada para a geração. Em algumas simulações, foi necessário aplicar as máscaras em várias etapas para alcançar um resultado polido e melhor contextualizado. Nessas situações, as imagens com as áreas mascaradas possuem preenchimentos em cores distintas, numeradas para indicar a ordem das modificações feitas na foto original. Os *prompts* utilizados para cada geração de imagem estão listados na parte textual, próximos às figuras, tanto em inglês, idioma no qual foram feitas as requisições no algoritmo, associadas às suas respectivas traduções.

Para garantir a reprodutibilidade do experimento, foi utilizada a plataforma *online* Clipdrop (2024), da *Stability.AI*, por meio da ferramenta chamada *Generative Fill*, que permite a edição de imagens utilizando o método de *inpainting*. Nessa aba, basta carregar a imagem desejada, selecionar a área a ser mascarada com o cursor, inserir o *prompt* para a geração da nova imagem e iniciar o processo, que dura apenas alguns segundos. Após o processamento, um novo preenchimento é criado para a área selecionada, que pode ser aceito pelo usuário ou gerado novamente até que se atinja um resultado satisfatório. Devido à rapidez do processamento, é possível criar várias versões até obter um

resultado que se adeque à expectativa inicial. Outra abordagem é modificar a imagem em diferentes etapas, alterando partes individuais para facilitar a compreensão contextual do algoritmo, como realizado em alguns dos experimentos. Mesmo ao optar por esse processo em etapas, o tempo total de edição é significativamente mais rápido do que em outros métodos de manipulação de imagens, levando apenas alguns minutos de interação do usuário com a ferramenta.

4 RESULTADOS

Neste artigo, os resultados foram divididos em duas partes principais: revisão bibliográfica e pesquisa experimental. A primeira parte se reservou à reunião da base teórica essencial sobre o uso de Redes Neurais Generativas de Imagens (RNI), com destaque para os modelos de difusão, reconhecidos pela sua capacidade de modificar imagens utilizando a técnica de *inpainting*. A segunda parte se refere à pesquisa experimental, com a realização de sete ensaios práticos, nos quais as técnicas estudadas foram aplicadas a imagens de cenários urbanos reais na cidade de Campos dos Goytacazes/RJ. Esses experimentos foram organizados com o objetivo de testar e demonstrar o potencial e as limitações da ferramenta em situações reais do planejamento urbano.

4.1 Pesquisa bibliográfica

As Redes Neurais Artificiais frequentemente são comparadas ao modo de funcionamento do cérebro humano. A organização dessas redes, com suas camadas de processamento e nós de avaliação (chamados neurônios artificiais), é baseada no modelo de funcionamento cerebral. Quando se trata de tarefas mais complexas, que exigem um número maior de camadas, utilizam-se redes neurais profundas, sistemas mais complexos de nós, com a capacidade de gerar hierarquias de dados em processos como sintetização, processamento de informações e reconhecimento de padrões (Deng; Yu, 2014).

A chegada do *deep learning* à literatura sobre aprendizado de máquina, como discutido por Lecun, Bengio e Hinton (2015), marcou um avanço notável nas técnicas de IA, aprimorando a capacidade de gerar novos dados, seja em formato de texto, imagem, áudio ou qualquer modo de saída. Estruturas com tais

finalidades são chamadas de Redes Neurais Generativas (RNG). Entre os diversos dados sintéticos que podem ser criados por essas redes, as imagens se destacam.

Esses modelos são desenhados para aprender a distribuição probabilística dos dados com os quais foram treinados e, a partir dessa distribuição, criar novas amostras que se parecem com as características observadas nos dados originais. Os conjuntos de dados utilizados no treinamento costumam ser formados por pares de informações, criando uma relação entre imagens e textos. Em sistemas que utilizam entradas de informações textuais em linguagem natural é utilizado o termo *text-to-image*, ou “texto para imagem”. Żylińska (2020) explica que os sistemas *text-to-image* são hierarquias que transformam descrições em linguagem natural (não computacional), também chamadas de *prompts*, em representações visuais demarcadas por características como objetos destacados, cenários, composições e estilos. Já os sistemas *image-to-image* trabalham gerando uma nova imagem a partir de uma imagem inicial, capazes de alterar e estilizar a imagem original com base nas diretrizes previamente definidas durante o treinamento de dados ou na configuração do sistema. Grande parte das Redes Neurais Generativas de Imagens (RGNI) possuem uma estrutura que combina duas abordagens: modelos de difusão e redes adversárias.

Os modelos de difusão operam pelo processo de difusão probabilística, que ocorre em duas fases. Na primeira fase, que se chama difusão direta, adiciona-se gradualmente ruído gaussiano (um tipo de ruído estatístico que segue a distribuição normal em probabilidade) até que a imagem se degrade por completo, resultando em algo semelhante à estática de uma televisão sem sinal. Já na segunda fase, chamada de difusão reversa, o processo é invertido, removendo o ruído de maneira gradual para reconstruir a imagem original. A outra abordagem, as redes adversárias verificam a qualidade e o direcionamento do processo. Nesse sistema, duas redes opostas trabalham simultaneamente: uma é responsável por gerar as imagens e a outra por avaliá-las, garantindo assim a melhoria contínua e a qualidade do produto final (Ho; Jain; Abbeel, 2020).

Após a finalização do treinamento, o modelo de difusão desenvolve a habilidade de criar imagens diretamente, sem precisar passar pelo processo de

destruição dos dados. Começando a partir de um estado inicial de ruído, o modelo utiliza um número identificador, chamado de *seed*, que atua como uma semente para determinar a configuração exata de cada *pixel* no ruído inicial. Esse identificador guia o processo de refinamento gradual, em que o ruído é transformado, passo a passo, em uma imagem coerente.

Os modelos de difusão ganharam destaque tanto pela sua eficiência e alta qualidade nos resultados como pela disponibilidade em código aberto. O algoritmo pode ser acessado por meio da plataforma chamada *GitHub*, que permite o compartilhamento de código de forma gratuita. Um dos principais exemplos dessa tecnologia é o *Stable Diffusion* (SD), que está disponível como uma ferramenta de geração de imagens em várias plataformas digitais, como o *Clipdrop* (2024), desenvolvido pela *Stability.AI*, a empresa criadora do SD. Para utilização dessas plataformas, não se torna necessário instalar nenhum programa no computador, pois elas funcionam em servidores remotos. Além disso, elas oferecem sistemas de créditos gratuitos diários, permitindo o uso sem barreiras financeiras, mesmo em dispositivos móveis.

Por meio do SD, incluindo a plataforma *Clipdrop*, é possível utilizar uma variação do modelo de difusão voltada para a reconstrução das áreas selecionadas pelo usuário, chamada de *inpainting*. O processo funciona por meio da aplicação de máscaras sobre uma imagem original, marcando as regiões que devem ser modificadas. O modelo então é ajustado para interpretar o que está faltando e gerar conteúdo visual que se integre de forma harmoniosa ao restante da imagem, orientado pelo *prompt*. A rede neural analisa as características e os padrões ao redor da área mascarada e gera *pixels* que se encaixam esteticamente, preservando a coerência de texturas, cores e formas.

Essa ferramenta abre um amplo espaço de possibilidades, permitindo que a população interaja diretamente com necessidades emergentes e soluções propostas, por meio de reconstruções visuais que refletem suas condições reais e contextos cotidianos. Assim, o *inpainting* pode se tornar uma ferramenta estratégica no planejamento urbano, facilitando uma comunicação visual mais aberta que conecta as expectativas dos cidadãos com soluções projetadas, promovendo a criação de cidades mais inclusivas e sustentáveis.

4.2 Pesquisa experimental

A ONU definiu como primeira meta para alcançar cidades e comunidades mais sustentáveis a garantia de moradias seguras para todos, com acesso a serviços essenciais e à urbanização de áreas marginalizadas. A Figura 2, subdividida em três partes, mostra na Figura 2-a a Rua Arnaldo da Silva Vizela, localizada no bairro Parque Prazeres, em Guarus, na margem esquerda do Rio Paraíba do Sul. A imagem retrata um conjunto de casas periféricas situadas ao redor de um canal poluído, com acúmulo de lixo em sua margem. O canal deságua na Lagoa do Parque Prazeres, também em Guarus.

Figura 2 – *Inpainting* a respeito da meta 11.1



Fonte: Google Earth (2024), modificada pelos autores por meio do *Clipdrop*.

Para essa imagem, o *inpainting* foi feito em quatro etapas (Figura 2-b). A primeira e a segunda máscaras, representadas por regiões 1 e 2, foram aplicadas nas margens do canal, com o texto de entrada (*prompt*): “*park, vegetation, green grass, sustainable, concrete paths*” (em tradução para o português significa “parque, vegetação, grama verde, sustentável, caminhos de concreto”), com o intuito de recuperar ambientalmente a área ao redor do valão. As terceira e quarta máscaras, associadas às regiões 3 e 4, foram aplicadas na área das habitações, buscando alternativas mais sustentáveis e integradas, tirando proveito do canal após sua despoluição, utilizando o *prompt*: “*sustainable*

houses” (“casas sustentáveis”). O resultado final (Figura 2-c) mostra uma urbanização sustentável, com habitações que interagem de forma positiva e tiram proveito do ambiente natural recuperado.

A segunda meta está ligada às melhorias essenciais no transporte público, buscando tornar os meios de transporte mais seguros, sustentáveis e acessíveis. A Figura 3, subdividida em três partes, mostra a Rua Tenente Coronel Cardoso, próxima ao cruzamento com a Avenida José Alves de Azevedo, na área central de Campos dos Goytacazes/RJ. A Figura 3-a evidencia um engarrafamento no trecho, marcado pela ausência de transporte público e pela predominância de veículos particulares.

Figura 3 – *Inpainting* a respeito da meta 11.2



Fonte: Google *Earth* (2024), modificada pelos autores por meio do *Clipdrop*.

No processo de *inpainting*, duas etapas foram realizadas (Figura 3-b). A primeira máscara, representada pela região 1, foi aplicada com o *prompt*: “*asphalt, avenue, street*” (“asfalto, avenida, rua”), com o objetivo de retirar os carros e substituir o trecho por uma via livre de trânsito. A segunda máscara, associada à região 2, utilizou o *prompt*: “*bus*” (“ônibus”), adicionando um meio de transporte alternativo, capaz de transportar um número de pessoas equivalente ao dos veículos particulares removidos. O resultado final apresenta

uma transformação no fluxo viário, priorizando o transporte coletivo e reduzindo o congestionamento na área (Figura 3-c).

A terceira meta foca em promover uma urbanização mais inclusiva e sustentável. A Figura 4 mostra o Residencial Veneza, um conjunto habitacional localizado próximo à estrada do carvão, na margem esquerda do canal Campos-Macaé. A Figura 4-a retrata uma paisagem de casas padronizadas, sem arborização predominante, criando um ambiente urbanisticamente homogêneo.

Figura 4 – *Inpainting* a respeito da meta 11.3



Fonte: Google *Earth* (2024), modificada pelos autores por meio do *Clipdrop*.

Para esse experimento, foi aplicada uma única máscara, como mostrado na Figura 4-b, cobrindo a região onde as casas estão localizadas. O *prompt* utilizado foi: “*diverse sustainable houses*” (“casas sustentáveis diversas”), buscando introduzir maior diversidade arquitetônica e soluções mais sustentáveis. O resultado final (Figura 4-c) apresenta uma paisagem mais variada, com tipologias residenciais distintas, integração de áreas verdes e elementos que promovem uma urbanização mais ecológica e inclusiva. Provavelmente pelo método de treinamento centrado em regiões europeias e norte-americanas da rede neural utilizada no *Clipdrop*, o resultado retrata uma

realidade arquitetônica que foge da experienciada na maior parte do Brasil, inclusive no município de Campos dos Goytacazes/RJ.

A quarta meta está focada na importância da proteção do patrimônio cultural e natural. A Figura 5 apresenta o Solar dos Airizes, um patrimônio histórico localizado na Estrada Campos-Atafona, no segundo subdistrito de Campos dos Goytacazes/RJ. Na Figura 5-a, o patrimônio é mostrado nas condições atuais do imóvel, em estado precário de conservação, com partes de sua alvenaria colapsadas e esquadrias vedadas com paredes de tijolos, descaracterizando a estrutura original.

Figura 5 – *Inpainting* a respeito da meta 11.4



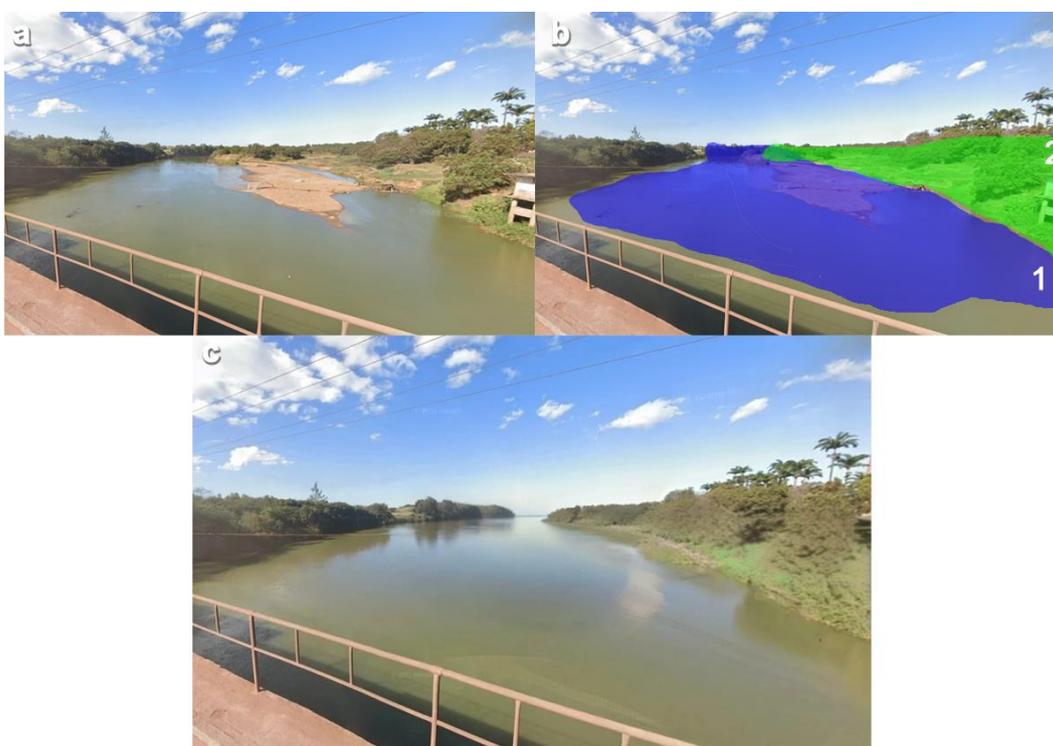
Fonte: Google *Earth* (2024), modificada pelos autores por meio do *Clipdrop*.

Para o processo de *inpainting*, foram aplicadas duas máscaras (Figura 5-b). A primeira, na região 1, tornou mais eficiente a utilização apenas do *prompt*: “*wall*” (“parede”), para a recuperação das alvenarias e do reboco, restaurando a integridade estrutural do casarão. A segunda máscara, na região 2, foi associada ao *prompt*: “*wooden frames*” (“esquadrias de madeira”), com o objetivo de restabelecer janelas e portas no estilo original, proporcionando uma requalificação fiel às características históricas do edifício. O resultado final

(Figura 5-c) revela uma restauração respeitosa e completa do Solar dos Airizes, devolvendo ao patrimônio sua dignidade arquitetônica e ressaltando seu valor histórico para a cidade.

A quinta meta está centrada na redução de impactos causados por desastres relacionados à água, com foco na preservação e recuperação de recursos hídricos. A Figura 6 ilustra o Rio Muriaé, afluente do Rio Paraíba do Sul, no trecho próximo ao bairro Sapucaia. A Figura 6-a evidencia o severo assoreamento desse rio, causado pela ação antrópica, o que acarreta uma série de danos sociais e ambientais, como a escassez de água e o risco aumentado de enchentes.

Figura 6 – *Inpainting* a respeito da meta 11.5



Fonte: Google *Earth* (2024), modificada pelos autores por meio do *Clipdrop*.

Na Figura 6-b, duas máscaras foram aplicadas no processo de *inpainting*. A primeira, na região 1, utilizou o *prompt*: “*river*” (“rio”), para a recuperação das áreas assoreadas, simulando o fluxo adequado de água. A segunda máscara, na região 2, foi associada ao *prompt*: “*vegetated riverside*” (“margem vegetada”), com o objetivo de introduzir vegetação nativa ao longo das margens, uma estratégia eficaz para inibir o processo de assoreamento e garantir a

preservação sustentável do rio. O resultado final (Figura 6-c) propõe um ambiente restaurado, com o rio recuperado e protegido contra novas degradações.

No sexto experimento, a meta número seis trata de reduzir o impacto ambiental negativo das cidades, dando enfoque na melhoria da qualidade do ar e na gestão de resíduos. A Figura 7 retrata a Usina de Açúcar Sapucaia, localizada próximo ao local do ensaio anterior. Na Figura 7-a, a imagem mostra a usina ao fundo, enquanto no primeiro plano há uma pilha de troncos de madeira, aparentemente destinados à queima para a produção de calor, um processo responsável por emissões e o uso de recursos não renováveis.

Figura 7 – *Inpainting* a respeito da meta 11.6



Fonte: Google Earth (2024), modificada pelos autores por meio do *Clipdrop*.

No *inpainting* foram aplicadas duas máscaras (Figura 7-b). A primeira, na região 1, utilizou o *prompt*: “*sustainable industry*” (“usina sustentável”), com o objetivo de transformar a usina em uma instalação que adota práticas industriais mais limpas e sustentáveis. A segunda máscara, na região 2, foi associada ao

prompt: “park” (“parque”), visando a criação de um espaço verde que possa ser utilizado pela comunidade local, oferecendo lazer e convivência em uma área anteriormente ocupada por uma atividade industrial poluente. O resultado final (Figura 7-c) apresenta uma usina modernizada e integrada a uma área livre comunitária, equilibrando a produção industrial com o bem-estar ambiental e social. Um problema evidenciado nesse ensaio está nas características formais da usina apresentada no resultado final, que podem destoar da função original do empreendimento.

A sétima meta enfatiza a importância de garantir acesso universal a espaços públicos que sejam seguros, acessíveis e ambientalmente sustentáveis. A Figura 8 mostra um terreno sem função social, localizado na Rua Murilo Peixoto, no Parque São Silvestre.

Figura 8 – *Inpainting* a respeito da meta 11.7



Fonte: Google *Earth* (2024), modificada pelos autores por meio do *Clipdrop*.

A Figura 8-a retrata um local abandonado, onde lixo foi depositado, representando um risco e insegurança para a comunidade ao redor. A Figura 8-b apresenta a região de *inpainting*, aplicado em uma única máscara sobre a área do terreno, utilizando o *prompt*: “accessible square, sustainable and wooded

park” (“praça acessível, parque sustentável e arborizado”), com o objetivo de transformar o terreno em um espaço público requalificado. Apesar de uma visível deformação de escala, o resultado final (Figura 8-c) apresenta de forma positiva um parque acessível e verde, com árvores e áreas de lazer, oferecendo um ambiente seguro e agradável para a convivência da comunidade local, além de promover a sustentabilidade urbana e o uso eficiente dos espaços subutilizados.

5 CONCLUSÃO

O estudo em questão, além de apresentar uma breve revisão bibliográfica, realizou sete experimentos utilizando a técnica de *inpainting*, um método de geração de imagens por Inteligência Artificial, com o intuito de investigar soluções visuais que possam contribuir para um planejamento urbano mais sustentável e claro. Cada experimento foi projetado para abordar uma meta específica do Objetivo de Desenvolvimento Sustentável 11, que busca tornar as cidades e assentamentos humanos mais inclusivos, seguros, resilientes e sustentáveis.

No experimento, foram utilizadas diversas imagens de cenários urbanos da cidade de Campos dos Goytacazes, município do estado do Rio de Janeiro, que apresentavam desafios a serem superados. Os sete ensaios foram submetidos ao *inpainting*. Tal experimento foi feito através de diferentes *prompts* definidas pelos pesquisadores, visando a criação de uma reconstrução digital desses locais, sugerindo melhorias que poderiam ser base para propostas específicas no planejamento urbano. As sete metas do ODS 11 foram utilizadas como guias na escolha das imagens e para o desenvolvimento dos *prompts*.

Para cada imagem foram necessárias diferentes estratégias para atingir o resultado final. Foi observado que a criação das imagens e para adquirir um resultado final satisfatório, pode-se necessitar da divisão do experimento em uma ou mais etapas. O experimento também revelou que máscaras maiores, *prompts* mais detalhados resultam em produtos finais melhores, enquanto para máscaras mais específicas, entradas de texto mais abrangentes foram mais eficazes. Em todos os casos, a ferramenta demonstrou ser prática e ágil, evidenciando seu potencial no uso para a comunicação visual.

O processo de produção de imagens por meio de *inpainting* apresenta alguns desafios práticos e técnicos que podem limitar sua aplicação cotidiana,

especialmente quando se utiliza plataformas como o *Clipdrop* em sua versão gratuita. O número diário de gerações é limitado, o que pode representar um problema se associado à constante necessidade de repetir o processo de geração para conseguir resultados que atinjam o nível esperado de qualidade. Para uma prática profissional da ferramenta, possivelmente ocorrerá a necessidade da obtenção do plano pago da plataforma, nos valores de R\$79,00/mensal ou R\$745,00/anual (dados do mês de publicação do presente artigo).

Outro desafio é que, em alguns casos, os produtos finais apresentam problemas em relação à proporção e escala, dificultando a integração harmoniosa com o ambiente. As gerações também podem destoar do contexto natural e sociocultural original, provavelmente pela origem da base de treinamento da rede neural. Isso pode resultar na criação de estruturas que não se ajustam ao papel esperado ou ao contexto brasileiro, causando uma desconexão entre forma, função e localização.

Além disso, a aceitação da tecnologia por profissionais do setor ainda enfrenta resistência. Arquitetos e urbanistas podem hesitar em adotar novas ferramentas de IA por falta de familiaridade ou pelo receio de que essas tecnologias venham a substituir o trabalho humano. Isso é agravado pela falta de precisão entre os objetivos do projeto e os resultados gerados pela ferramenta, o que pode desmotivar seu uso em contextos profissionais mais exigentes.

O uso dessa ferramenta apresenta grande potencial para facilitar a comunicação entre técnicos e a comunidade, reduzindo ambiguidades e perdas de significados nas expressões e publicações de ideias. É válido destacar que, apesar do grande potencial da ferramenta, ela ainda possui limitações significativas e o papel de um urbanista especializado continua sendo insubstituível. A criação de espaços urbanos exige um olhar crítico e atento, além de uma abordagem multidisciplinar, que só profissionais qualificados podem proporcionar. Sendo assim, a ferramenta complementa o trabalho do profissional especializado, aprimorando a comunicação e facilitando a visualização de ideias, mas não substitui o conhecimento técnico e a experiência no campo do planejamento urbano.

6 REFERENCIAL BIBLIOGRÁFICO

BURMARK, L. **Visual Literacy**: Learn to see, see to learn. Alexandria, VA: Association for Supervision and Curriculum Development, 2002.

CLIPDROP. Reino Unido: Stability.ai, 2024. Disponível em: <https://clipdrop.co/>. Acesso em: 08 out. 2024.

DENG, L.; YU, D. Deep Learning: Methods and Applications. **Foundations and Trends in Signal Processing**, Boston, Delft, v. 7, n.3-4, p. 197-387, jun. 2014. Disponível em: <http://dx.doi.org/10.1561/20000000039>. Acesso em: 13 jun. 2023.

GIL, A. C. **Como elaborar projetos de pesquisa**. 6 ed. São Paulo: Atlas, 2019.

GOOGLE. **Google Earth**. [S.l.]: Google, 2024. Disponível em: <https://earth.google.com>. Acesso em: 10 out. 2024.

HO, J.; JAIN, A.; ABBEEL, P. **NeurIPS**, Vancouver, v. 34, p. 1-16, dez. 2020. Disponível em: <https://arxiv.org/pdf/2006.11239.pdf>. Acesso em: 02 ago. 2023.

LECUN, Y.; BENGIO, Y.; HILTON, G. Deep Learning. **Nature**, Reino Unido, v. 521, p. 436-444, maio 2015. Disponível em: <https://www.nature.com/articles/nature14539>. Acesso em: 03 jun. 2023.

ONU. Organização das Nações Unidas. Programa de Desenvolvimento das Nações Unidas. **Transformando nosso mundo**: a Agenda 2030 para o desenvolvimento sustentável. São Paulo: ONU, 2016. Tradução: Centro de Informação das Nações Unidas para o Brasil (UNIC Rio). 2. ed. Rio de Janeiro: UNIC Rio, 2016. 54 p. Disponível em: https://www.mds.gov.br/webarquivos/publicacao/Brasil_Amigo_Pesso_Idosa/Agenda2030.pdf. Acesso em: 14 jun. 2024.

ŻYLIŃSKA, J. **AI Art**: Machine Visions and Warped Dreams. 1. ed. Londres: Open Humanities Press, 2020. 178 p.