



DO ÍNDICE AO MODELO: IA generativa, QR ControlNet e a produção de realidades sintéticas¹

FROM INDEX TO MODEL: generative AI, QR ControlNet and the production of synthetic realities

Marcio TELLES²

Resumo: Este artigo investiga como a inteligência artificial generativa transforma a relação entre imagem e objeto no contexto da história das tecnologias visuais, a partir da análise de uma série de imagens criadas com o protocolo QR ControlNet. Diferente da fotografia, que mantém um vínculo indexical com a realidade, as imagens geradas por IA são radicalmente antimiméticas, operando dentro de um circuito fechado de referências algorítmicas. Elas não representam um objeto externo, mas sintetizam padrões estatísticos de bancos de dados, criando uma visualidade autorreferencial. Argumentamos que essas imagens instauram uma ruptura na história das tecnologias visuais ao consolidar um regime de produção onde a plausibilidade não deriva do mundo material, mas da coerência interna do modelo algorítmico. Por meio de uma análise em três camadas – superficial, subficial e mídia-ecológica – demonstramos como essas imagens exemplificam um deslocamento epistemológico na cultura visual contemporânea.

Palavras-Chave: Inteligência Artificial. Visualidade Algorítmica. Fotografia.

Abstract: This paper investigates how generative artificial intelligence transforms the relationship between image and object in the context of the history of visual technologies, based on the analysis of a series of images created with the QR ControlNet protocol. Unlike photography, which retains an indexical link to reality, AI-generated images are radically antimimetic, operating within a closed circuit of algorithmic references. Rather than depicting external objects, they synthesize statistical patterns from databases, creating a self-referential visuality. We argue that these images introduce a rupture in the history of visual technologies by establishing a mode of production in which plausibility stems not from the material world, but from the internal coherence of the algorithmic model. Through a three-tiered analysis – surfaces, subsurfaces, and media ecology – we demonstrate how these images embody an epistemological shift in contemporary visual culture.

Keywords: Artificial Intelligence. Algorithmic Visualities. Photography.

¹ Trabalho apresentado ao Grupo de Trabalho Comunicação, Arte e Tecnologias da Imagem. 34º Encontro Anual da Compós, Universidade Federal do Paraná (UFPR). Curitiba – PR. 10 a 13 de junho de 2025.

² Professor Adjunto do Bacharelado em Cinema e Audiovisual e docente permanente do Programa de Pós-Graduação em Cinema e Artes do Vídeo, ambos da Unespar (Universidade Estadual do Paraná). Curitiba – PR. Doutor em Comunicação e Informação pela Universidade Federal do Rio Grande do Sul, com estágio doutoral na Winchester School of Art.

1. Introdução

Iniciemos com quatro imagens (fig. 1). À primeira vista, elas parecem fotografias em preto e branco de cenas cotidianas: paisagens urbanas, momentos carregados de emoção e situações de rua. No entanto, uma análise mais atenta revela distorções sutis – inconsistências surreais, erros anatômicos, textos sem sentido – que comprometem sua autenticidade. Estamos cada vez mais acostumados a reconhecer esses indícios característicos de imagens geradas por inteligência artificial (WASIELEWSKI, 2023).

Contudo, é ao nos afastarmos fisicamente dessas imagens que algo ainda mais estranho começa a emergir. Elementos como máscaras, multidões e arquitetura se combinam de modo a formar, ainda que de maneira pouco óbvia, a palavra “OBEY”, perceptível apenas à distância. O termo não aparece explicitamente; ele emerge da justaposição estratégica de elementos visuais. Embora tenham aparência de fotografias documentais plausíveis, as cenas retratadas são impossíveis de serem encenadas no mundo real.



FIGURA 1 – As quatro imagens que compõe a série OBEY, criada pelo *user* Dreaming Tulp com o protocolo QR ControlNet.

FONTE - <https://x.com/dreamingtulpa/status/1705123475738779904?t=sjt5ncyQm9aH8OJR3JYHEw&s=08>. Acesso 31 jan. 2025.

Essas imagens foram criadas utilizando o protocolo QR ControlNet, um sistema que permite manipular elementos para compor cenas altamente realistas a partir da incrustação de

uma imagem em outra. Embora inteiramente artificiais, elas ilustram como sistemas de IA generativa podem construir realidades visuais hiper-realistas e completamente sintéticas.

Historicamente, as imagens têm funcionado como mediadoras entre o material e o conceitual, moldando visões de mundo coletivas e influenciando a representação da verdade, da memória e da imaginação (BERGER, 1972; BELTING, 2011). Essa dinâmica fundamenta a maneira como os seres humanos documentam, interpretam e articulam as complexidades de seu ambiente (BERGER, 1972). Dada a profunda influência das imagens na construção da experiência do mundo concreto, compreender essa relação é essencial para analisar as transformações na cultura visual (MIRZOEFF, 1999; CRARY, 2012). A partir dessas observações, coloca-se a questão: de que maneira a IA generativa transforma os processos de produção imagética dentro da história das tecnologias visuais?

Nosso argumento é que a IA generativa instaura um novo regime de produção imagética baseado na modelagem estatística e na coerência algorítmica, reconfigurando a relação entre imagem, referente e significação na história das tecnologias visuais. Estas imagens oferecem um exemplo claro da tendência, então muito inicial, observada por Jonathan Crary na abertura de sua obra clássica, *Técnicas do Observador*:

A maioria das funções historicamente importantes do olho humano está sendo suplantada por práticas nas quais as imagens figurativas não mantêm mais uma relação predominante com a posição de um observado em um mundo ‘real’, opticamente percebido. Se é possível dizer que essas imagens se referem a algo, é, sobretudo, a milhões de bits em dados matemáticos eletrônicos. (CRARY, 2012, p.11-12, grifo nosso).

Como observam Montaña e Vagt (2017, p.5), “as imagens sintéticas não são manifestações físicas do mundo exterior”, mas sim produtos dos “bits matemáticos” que habitam o interior da máquina. Essa transformação na cultura visual demanda uma análise aprofundada de seus mecanismos e consequências. Diferentemente das imagens fotográficas, que – mesmo de forma sutil – mantêm um vínculo físico com o mundo material, as imagens geradas por IA se enquadram em uma categoria na qual a representação perde sua conexão com o “real”. Semelhantes às imagens operativas (FAROCKI, 2004; PAGLEN, 2014) usadas em sistemas de vigilância, mísseis teleguiados ou redes neurais de reconhecimento facial, elas funcionam em um sistema fechado de referências algorítmicas.

Sistemas ópticos, como a fotografia analógica e o cinema em seus primórdios, dependiam da interação direta entre a luz e os objetos materiais para criar imagens, vinculando-as ao mundo externo (BAZIN, 1991). Essas imagens mantinham, portanto, uma relação com a

realidade física, embora mediadas por perspectivas culturais e ideológicas (DUBOIS, 2012, p.23-56) e limitadas pelos processos tecnológicos da época (FLUSSER, 2011; KITTLER, 2016). Ao converter a luz em dados digitais e permitir manipulações extensivas em pós-produção, a fotografia digital já havia introduzido um nível de abstração que se distanciava do mundo material (MITCHELL, 1992; FLUSSER, 2011). No entanto, nessa modalidade ainda havia uma interação física inicial – a luz tocando um sensor eletrônico.

Em contraste, a IA generativa não documenta nem meramente distorce o mundo, mas constrói imagens integralmente dentro de um sistema algorítmico, baseando-se em padrões, probabilidades e estruturas de dados extraídas de outras imagens. Nessa lógica, conforme observam Montaña e Vagt (2017, p.5), a computação gráfica “não é sintoma de um mundo caótico que precise ser organizado,” pois se fundamenta em “cálculos puros” e independe de uma referência óptica inicial. Elas assinalam o “fim das mídias ópticas” (MONTAÑA; VAGT, 2017; KITTLER, 2016).

A imagem resultante já não remete a “algo necessariamente real”, conforme Barthes (1984) caracterizou no *noema* da fotografia, encapsulado na expressão “isso foi”. Não há mais conexão entre a imagem e um momento específico no tempo e espaço que indubitavelmente existiu. Em vez de se configurar como um índice de algo externo, a imagem gerada por IA opera de maneira puramente simbólica (WILF, 2013), baseando-se em estatísticas e padrões probabilísticos, sem estabelecer qualquer vínculo indexical com o mundo material. Ademais, ela exibe uma “dupla existência: uma matriz sensorial de formas e cores em uma tela e um código de computador em memória e processamento” (MONTAÑA; VAGT, 2017, p.2).

Entre as várias maneiras de gerar imagens por IA, este artigo concentra-se no protocolo QR ControlNet³. Essas imagens exemplificam uma mudança fundamental nos modos de produção imagética, na qual a aparência de familiaridade visual não decorre de um vínculo direto com processos materiais ou referenciais externos, mas de um modelo estatístico que organiza padrões dentro de um sistema algorítmico fechado. Apesar de sua verossimilhança, elas não são derivadas do mundo físico nem seguem as tradições indexicais das imagens técnicas anteriores. Ao examinar esse caso, este estudo busca situar as imagens geradas por IA dentro da evolução das tecnologias visuais, analisando como elas reconfiguram os princípios da representação e introduzem uma nova lógica na produção da visualidade.

³ Explicaremos o funcionamento do sistema abaixo.

Neste artigo, analisaremos uma série de quatro imagens criadas utilizando este protocolo que “representam” cenas de protestos e incorporam a palavra de ordem “OBEY” dentro de sua composição visual. Essas imagens serão examinadas por meio de três camadas analíticas. Primeiramente, a camada superficial, que se concentra no que é visível na superfície da imagem – os elementos, a composição e os sinais visuais que constroem um senso de plausibilidade e realismo. Em seguida, a camada *subficial*⁴, que explora os protocolos de IA e os sistemas computacionais que possibilitaram a criação dessas imagens, examinando como estruturas de dados e algoritmos internos orquestram essas realidades sintéticas. Por fim, analisaremos o sistema que dá origem a estas imagens dentro de uma ecologia midiática, situando essas imagens dentro de uma história mais ampla das mídias imagéticas.

2. Superfícies

Analisemos uma das imagens do conjunto *OBEY*. Esta (fig. 2) retrata uma cena urbana em preto e branco, ambientada em um grande centro urbano, com *outdoors* luminosos e edifícios altos ao fundo. Em primeiro plano, uma criança utilizando máscara facial, popularizada durante a pandemia de Covid-19, ocupa a maior parte do enquadramento. Seu rosto, parcialmente coberto pela máscara, exibe olhos grandes e cabelos negros ondulados. No segundo plano, outras duas crianças também usam máscaras e parecem complementar a composição com a criança no primeiro plano. A criança à esquerda olha diretamente para a câmera, e sua camiseta aparenta estar manchada de sangue, embora possa ser apenas a estampa da roupa. A menina à direita está em uma postura que remete a uma posição militar: seu corpo está ereto, os braços rigidamente posicionados ao lado do corpo, e sua expressão é séria, olhando para algo fora do enquadramento que somente ela observa.

A profundidade da cena indica uma grande cidade cosmopolita, com letreiros luminosos adornando as laterais dos edifícios, elementos característicos de uma paisagem urbana globalizada que pode sugerir Nova York, Tóquio, Pequim ou Londres. Os letreiros estão imersos em uma iluminação cintilante, produzindo uma oscilação visual que dificulta a leitura clara das palavras. Iluminados pela luz dos *outdoors*, pedestres transitam em diversas direções e ao menos um veículo circula pela larga avenida. A imagem transmite uma sensação de

⁴ “A superfície [surface] da imagem é visível, enquanto a subface [subsurface] é manipulável” (Frieder Nake, 2005, citado em MONTAÑA; VAGT, 2017, p.4, grifos no original).

urgência, similar à inquietação associada a distúrbios civis, e remete a fotos documentais de crises sanitárias ou emergências globais.



FIGURA 2 – As meninas de OBEY, pelo user Dreaming Tulpa.

FONTE - <https://x.com/dreamingtulpa/status/1705123475738779904?t=sjt5ncyQm9aH8OJR3JYHEw&s=08>.

Acesso 31 jan. 2025.

Vamos agora nos concentrar nos elementos que revelam a artificialidade das imagens geradas por IA, o *uncanny valley* característico da tecnologia. Este movimento analítico provém de uma tradição humanística consolidada de *close reading*: então devemos aproximar a imagem sobre os nossos olhos e, em vez de um olhar superficial, devemos analisar atentamente a superfície. As evidências são claras: a pele da menina em primeiro plano é excessivamente plastificada; os rostos das duas meninas no segundo plano são simétricos (são a mesma?) e os detalhes na camiseta da menina à esquerda não possuem coerência textual. Além disso, falta coerência também para os letreiros luminosos: as inscrições nos *outdoors* são irreconhecíveis, compostas por caracteres que imitam a tipografia comercial sem formarem palavras comprehensíveis. Mesmo que a profundidade de foco e a sobre-exposição da fotografia

dificultem a identificação desses elementos, algumas palavras deveriam ser reconhecíveis. Somados, esses elementos indicam a artificialidade da imagem: não é uma *fotografia*, mas uma *imagem* (hiper-)realista gerada com IA.

Façamos, então, um segundo movimento. Diferentemente da leitura atenta (*close reading*), posicionemo-nos à distância. Na figura 3, temos a mesma imagem, porém com tamanho reduzido em 70% comparado à figura 2. Agora percebemos detalhes que antes não eram visíveis: o tronco da menina à esquerda forma a letra *O*, com sua estampa ensanguentada centralizando essa letra. As mechas negras da menina em primeiro plano formam as letras *B* e *E*, enquanto seu rabo-de-cavalo se une à pose rígida da menina à direita para criar a letra *Y*. Assim, formam-se as letras *O-B-E-Y*, “Obedeça”⁵, uma palavra de ordem oculta dentro da imagem, inscrita apenas por elementos aparentemente naturais. Esse tipo de inscrição na *superfície* da imagem é característico do protocolo QR Control.Net, indicando a necessidade de analisarmos também a *subfície* da imagem – a camada manipulável de dados que afeta o que é superficialmente aparente. Antes, porém, investiguemos não apenas *o que* aparece na superfície, mas *como* isso se manifesta.



FIGURA 3 – Redução a 30% do tamanho anterior deixa mais clara a mensagem “secreta” criada pelo protocolo.
FONTE - <https://x.com/dreamingtulpa/status/1705123475738779904?t=sjt5ncyQm9aH8OJR3JYHEw&s=08>.
Acesso 31 jan. 2025.

Uma visão comum sobre modelos generativos de imagem é a de que a menina retratada teria realmente existido em alguma fotografia utilizada no treinamento da IA. Essa suposição se relaciona ao conceito ontológico do “isto foi” na fotografia, mantendo a ideia cultural de que “a fotografia significa que algo real (não importa o quê, seja ou não de natureza criminosa) deixa vestígios em uma mídia de arquivamento” (KITTLER, 2016, p.197). Nessa perspectiva, supõe-se que a garota posou para um fotógrafo em algum momento e lugar – talvez em algum parque em Chicago ou Hanói. Da mesma forma, as outras meninas teriam sido capturadas em

⁵ A justaposição de protestos e controle nas imagens evoca *Eles Vivem* (1987) de John Carpenter.

situações reais, como a da direita, que poderia estar participando de um evento pseudomilitar, cantando o hino nacional no pátio de uma escola. Assim, os elementos visíveis na imagem gerada seriam extraídos de diferentes fotografias, reorganizados em uma nova composição.

Essa linha de raciocínio assemelha-se à concepção de que modelos de linguagem natural, como o ChatGPT, seriam meros reorganizadores de fragmentos textuais preexistentes – ou seja, plagiadores robóticos. Em outras palavras, se os modelos de IA se limitassem a “recortar e colar” partes de fotografias já existentes, cada elemento visual manteria seu vínculo indicial com o objeto original, de maneira análoga às colagens ou às fotografias de “fantasmas” do início do século XX, produzidas por dupla exposição. Nessa hipótese, a relação indexical entre o que foi capturado pela câmera e o que é visto na imagem permaneceria intacta em cada fragmento, embora rearranjada em uma nova composição.

Convém lembrar que, mesmo na fotografia tradicional, o vínculo com o “real” jamais foi absoluto. Segundo Lissovsky (2021), ao analisar a obra de Barthes, a indicialidade fotográfica se liga ao corpo e ao vestígio de seu desaparecimento. Inspirando-se em Derrida, Lissovsky (2021, p.21-22) mostra que a fotografia não captura a coisa em si, mas apenas o seu espectro, sua ausência. É por isso que Derrida propõe o termo “referencial”, que não é nem “referente” nem “referência”, mas um rastro do “ter estado aí” – sempre diferido, sempre espectral. Barthes (1984) já assinalava essa dimensão fantasmagórica da fotografia ao sustentar que o fotografado se transforma em *spectrum*, um simulacro daquilo que foi. Derrida aprofunda a ideia ao mostrar como o referente fotográfico se dissipa no acontecimento passado, deixando apenas a referência: uma presença que nunca se torna plenamente apreensível.

A noção de uma relação espectral entre a fotografia e o fotografado parte da ideia de que a fotografia não constitui evidência do referente, mas integra uma estrutura de referência que “assinala a absoluta *singularidade* do outro” (SAGHAFI, 2000, p.101 apud LISSOVSKY, 2021, p.22, grifo nosso). No entanto, o que se vê na superfície desta imagem difere completamente de qualquer singularidade; há apenas *generalidade* e *uniformidade*. Nenhuma das garotas ou qualquer outro elemento existiu previamente em outro contexto. Sem um instante físico de interação com a luz, desaparece o último resquício do “ter estado aí”, rompendo-se, assim, o elo com a realidade material. Isso ocorre porque cada pixel, cada forma e cada detalhe foram gerados estatisticamente no momento da criação da imagem.

Assim como um modelo de linguagem não reproduz, palavra por palavra, as frases presentes em seu conjunto de treinamento, a imagem gerada por IA não constitui uma mera

recombinação de fragmentos fotográficos. Trata-se, antes, de uma *síntese probabilística* de traços comuns de imagens previamente analisadas, o que justifica o caráter “estocástico” (BENDER ET AL., 2021) do modelo. A aparência da menina exibida não reproduz o rosto de qualquer indivíduo singular (não há um “referencial”), mas resulta de uma *convergência estatística* de traços faciais aprendidos a partir de um grande volume de dados. Desse modo, ela representa uma *média*, aproximando-se dos padrões majoritários presentes em seu banco de treinamento (*generalidade*). Essa aparência não apenas é gerada, mas também “normalizada” (STAUFF, 2024), pois tende a refletir padrões dominantes de representação visual, reforçando ideais estéticos e estereótipos implícitos no conjunto de dados de treinamento (*uniformidade*). Assim, desfaz-se qualquer noção, ainda que mínima, de relação entre a imagem e um referente, inclusive a uma “estrutura de referência” que apontaria uma singularidade.

É nesse sentido que Salvaggio (2023, p.84, grifo no original) considera as imagens geradas por IA “um *infográfico* sobre o conjunto de dados subjacente”. Para o autor, elas refletem a percepção coletiva do consenso da internet em relação ao comando requisitado (*prompt*). As imagens tendem a se alinhar a tendências centralizadoras, produzindo representações “fortes” (mais realistas) e bem definidas quando se aproximam desse ideal consensual. Todavia, ao se desviarem desse consenso, tornam-se mais difíceis de categorizar e podem parecer “fracas” (mais surreais), surgindo com menor frequência ou apresentando qualidade e clareza reduzidas.

Curiosamente, outro livro de Roland Barthes se mostra mais pertinente para a análise de imagens geradas por IA. Em *Mitologias*, Barthes (2001) define mito como uma forma de significação secundária que converte um significado histórico e contingente em algo aparentemente universal. Trata-se de um sistema de comunicação que não apenas representa o mundo, mas o reconstrói dentro de uma lógica que esvazia a historicidade dos signos, reapresentando-os como evidências naturais.

Sob essa perspectiva, as imagens geradas por IA podem ser compreendidas como mitos, não apenas pela ausência de um referente específico, mas também por funcionarem em uma estrutura de significação que normatiza e universaliza padrões visuais. Ao expressarem um consenso estatístico sobre a aparência de certo tema (“menina em ambiente urbano” ou “inquietação civil”), essas imagens reforçam códigos culturais e estéticos.

Assim, a IA opera como um “sistema mitológico” ao transformar um processo essencialmente técnico – modelagem estatística baseada em dados – em uma representação que se apresenta como natural, neutra e inevitável. Com isso, o procedimento estatístico subjacente converte-se em uma visualidade normalizada, exibida como uma verdade consensual sobre o mundo, quando na realidade reflete apenas padrões hegemônicos codificados nos dados de treinamento. É, portanto, justamente para além da superfície que precisamos direcionar nosso olhar.

3. Subfícies

Novamente, começemos de outra imagem da série *OBEY*. Nesta (fig. 4), três pedestres caminham na calçada com roupas de inverno, ao fundo há uma vista de uma cidade cosmopolita com arranha-céus e letreiros luminosos, semelhante a Times Square em Nova Iorque. A composição da imagem é organizada por linhas geométricas fortes, com a sobreposição do viaduto e das colunas metálicas no topo, criando um efeito de moldura para a mensagem principal da imagem: *OBEY*. À esquerda, um painel publicitário exibe um grande “O”, dentro do qual há uma imagem em miniatura de um arranha-céu (?). O prédio com placas publicitárias à esquerda forma os contornos internos da letra “B”, enquanto a perspectiva dos edifícios ao fundo e à direita delimitam as letras “E” e “Y”. Nesta imagem, os indícios de artificialidade são mais evidentes, como letreiros com mensagens confusas e prédios impossíveis. No entanto, visualmente, é pertinente considerarmos como essa imagem foi gerada.

As imagens geradas com o protocolo ControlNet são, em termos técnicos, QR (*Quick Response Codes* em inglês) Codes, códigos bidimensionais que armazenam informações de forma compacta e são rapidamente lidos por dispositivos eletrônicos. Diferente dos códigos de barras unidimensionais, os QR Codes utilizam um padrão de módulos preto e branco em um quadrado, permitindo incorporar dados complexos como URLs, textos, números de telefone, informações de pagamento e coordenações geográficas. Criados pela empresa japonesa Denso Wave em 1994 para melhorar a rastreabilidade na indústria automotiva, sua capacidade de armazenar informações de forma redundante e ser lido mesmo parcialmente danificado os tornou populares em outros setores.



FIGURA 4 – Pedestres caminhando na “Times Square”.

FONTE - <https://x.com/dreamingtulpa/status/1705123475738779904?t=sjt5ncyQm9aH8OJR3JYHEw&s=08>.

Acesso 31 jan. 2025.

Em 5 de junho de 2023, na comunidade Stable Diffusion⁶, um usuário divulgou uma série de QR Codes singulares (fig. 5), criados a partir do Stable Diffusion em conjunto com um protocolo que chamou de ControlNet. Segundo Maameri (2023), o ControlNet é “uma rede neural que controla um modelo de difusão⁷ de imagem pré-treinado”, permitindo inserir uma imagem de condicionamento que manipula o processo de geração.

Nas imagens da figura 5, os QR Codes funcionam como as “imagens de controle”: a IA gera uma imagem que se integra visualmente ao próprio QR, mas mantém a capacidade de ser escaneada. O ControlNet garante que, apesar de a IA seguir o processo generativo de síntese de texto em imagem, ela obedeça às restrições estruturais do padrão QR, que fica “embutido”,

⁶ <https://www.reddit.com/r/StableDiffusion/comments/141hg9x/controlnet_for_qr_code/> Acesso: 1 fev. 2025.

⁷ Não retomarei aqui, por questão de espaço, explanações sobre redes neurais e modelos de difusão na geração de imagens. Para uma revisão destes processos, cf. Telles, 2023.

de maneira implícita, na imagem. Essa técnica viabiliza a criação de QR Codes que se distanciam da estética convencional de blocos em preto e branco, incorporando-os a composições mais complexas. Como resultado, o que inicialmente parece uma imagem comum mantém sua funcionalidade como QR Code digitalizável.

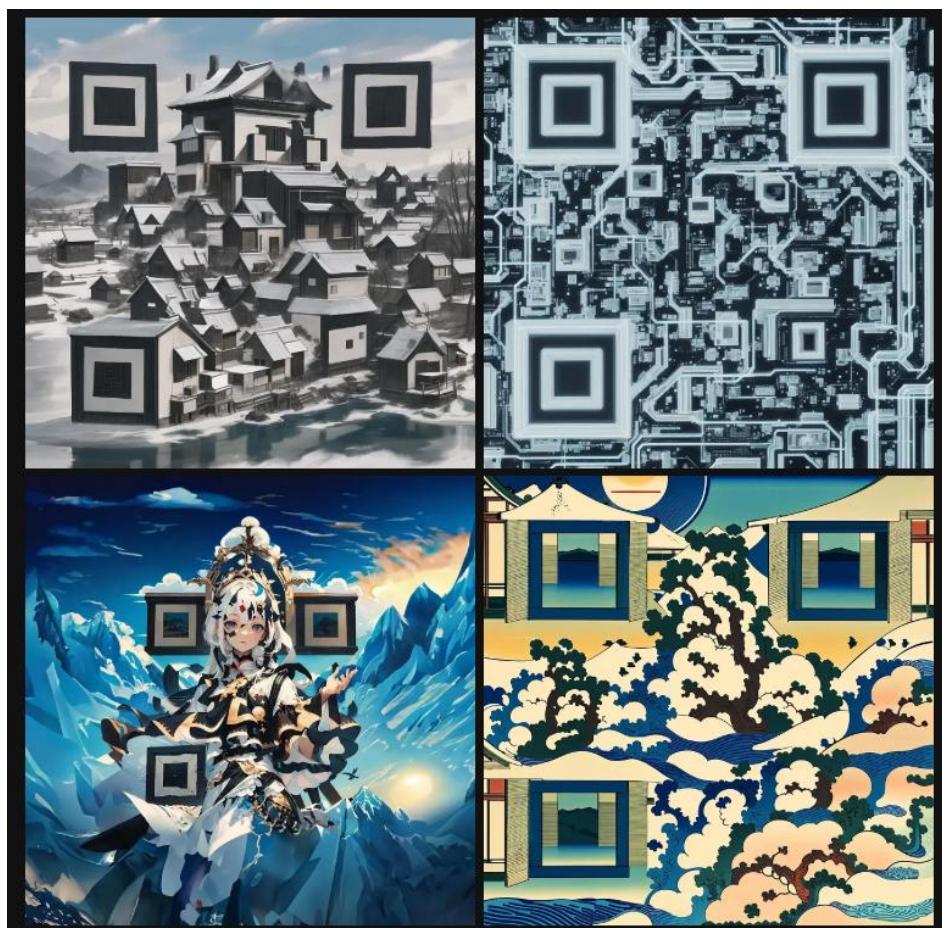


FIGURA 5 – Os QRCode gerados pelo usuário nhciao e postados no Reddit. Alguns destes *links* funcionam.
FONTE - https://www.reddit.com/r/StableDiffusion/comments/141hg9x/controlnet_for_qr_code/ Acesso 2 fev. 2025.

Os criadores das imagens da figura 5 desenvolveram posteriormente um site chamado QRBTF⁸, que permite aos usuários criarem seus próprios QR Codes imagéticos (MAAMERI, 2023). O aspecto mais relevante para nossa discussão surge a seguir: o usuário Dion Timmer desenvolveu um modelo derivado do original e o disponibilizou no repositório HuggingFace⁹. Na versão de Timmer, outra imagem pode ser utilizada para gerar o QR Code. Embora a ideia original fosse permitir que empresas e usuários incorporassem seu logotipo ou marca em

⁸ <<https://qrbtf.com/en>> Acesso 1 fev. 2025.

⁹ <https://huggingface.co/DionTimmer/controlnet_qrcode-control_v11p_sd21> Acesso 1 fev. 2025.

códigos QR, outros usuários perceberam a possibilidade de mesclar duas imagens e inserir, em vez de códigos QR, palavras¹⁰ (como no caso da série “OBEY”), ou criar imagens surrealistas (fig. 6).

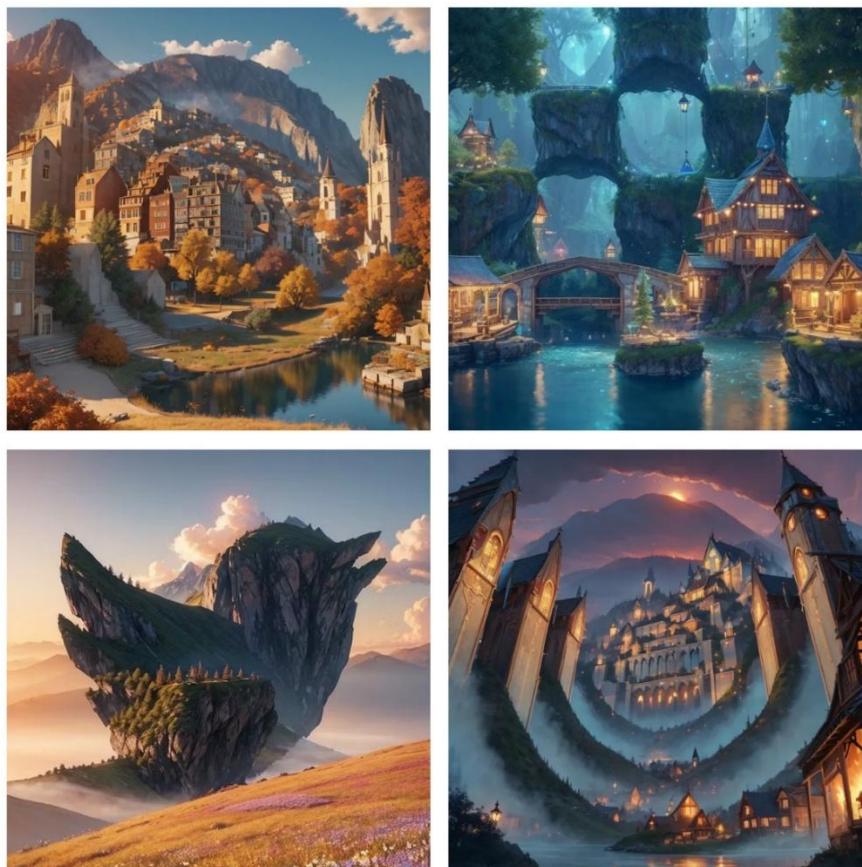


FIGURA 6 – QR ControlNet utilizado para criar imagens surrealistas ao utilizar, em vez de QR Code ou texto, outra imagem como controle pré-geração.

FONTE - https://twitter.com/dr_cintas/status/1704121911263854814 Acesso 19 set. 2023.

Este desdobramento evidencia uma característica fundamental das imagens geradas pelo protocolo QR ControlNet: em vez de se basearem em um vínculo externo, elas operam a partir de um modelo interno que organiza e determina sua estrutura. Diferentemente de outras imagens de IA, que simulam a realidade a partir de padrões estatísticos extraídos de bancos de dados visuais, aqui um segundo elemento – outra imagem – não atua simplesmente como componente gráfico, mas como operador que define tanto a forma quanto a função. Nesse processo, a referência externa perde relevância, pois o que define a imagem não é a correspondência com o real, mas a coerência interna de seu modelo algorítmico.

¹⁰ É bom notar que, tecnicamente, “OBEY” não é um texto, mas uma *imagem* (ou *word art*) inserida no protocolo, manipulada em conjunto com o *output* desejado no *prompt*.

Em consequência, a imagem torna-se um sistema autorreferencial de significação, em que significante e significado se fundem em um circuito regido pela lógica computacional. Dessa forma, a IA generativa – especialmente com o uso do ControlNet – suprime a distinção entre representação e objeto. O “significado” passa a existir apenas no interior de sua estrutura algorítmica, transformando a imagem em uma entidade autônoma produzida em um universo fechado de criação sintética. Esse aspecto ressalta a “natureza autorreflexiva” das imagens técnicas, descrita por Flusser (2011).

Enquanto Dondero et al. (2024) propõem uma “semiótica da inteligência artificial” centrada sobretudo na dimensão simbólica das imagens – como gestos, posturas corporais ou estilos de artistas (indexicais) –, sua *nulodimensionalidade* (FLUSSER, 2008) sugere que “nada é significado, nada é representado, não mais. Em vez disso, a *superfície está completamente livre para a apresentação*” (MONTAÑA; VAGT, 2017, p.19, ênfase nossa).

Livres para apresentarem, e não representarem, elas se tornam simulacros. Para Baudrillard (1991), o conceito de simulacro refere-se a representações que não fazem referência a qualquer realidade externa. O interesse de Baudrillard são as consequências políticas e ontológicas do que denomina, então, *hiper-realidade*. Nesse regime, os signos perdem sua ancoragem no mundo material e passam a se referir apenas a outros signos, formando um circuito fechado de significação que impede a distinção entre representação e representado.

Embora Baudrillard não trate especificamente de técnicas ou materialidades de produção imagética, seu argumento ecoa, de modo sub-reptício, em Mitchell (1992, p.17), ao afirmar que a manipulação fotográfica digital “borra a fronteira entre as representações e seus objetos.” Entretanto, mesmo quando alterada, a fotografia digital ainda mantém vestígios de seu referente, ao passo que as imagens geradas por IA rompem por completo esse vínculo, produzindo-se inteiramente por processos algorítmicos, sem “re-interpretação” de algo pré-existente. Elas não são meros simulacros, mas sim *simulações*, pois, conforme ressalta Krapp (2024, p. 19), “a simulação é uma ciência do artificial,” em que fatos se confundem com seus modelos (BAUDRILLARD, 1991).

Edwards (1996) vê a simulação como um avanço técnico que viabilizou a modelagem de fenômenos complexos, alterando a racionalidade científica por meio da quantificação e aplicando-se a áreas como meteorologia e estratégia militar. Já Krapp (2024) define a simulação como uma técnica cultural presente em diversos contextos – de modelos

aeronáuticos a videogames. Nessa perspectiva, a simulação não apenas “modela o real” (em consonância com Baudrillard), mas também transforma nossa compreensão e interação com ele por processos algorítmicos e estatísticos (segundo Edwards e Krapp). Para estes últimos, a simulação não é apenas um signo que remete a outro signo, mas sim um modelo capaz de criar representações funcionais e abstratas voltadas para análise, previsão e visualização de fenômenos.

Para nossa discussão, é fundamental ressaltar que a simulação é *antimimética* – “mina qualquer distinção entre original e cópia” (Krapp, 2024, p. 29) –, mas não se torna *antirreferencial*. Em *Simulacros e Simulação*, Baudrillard (1981) defende que o real é substituído por signos que remetem apenas a si mesmos, rompendo qualquer ligação com um referente externo. Em contraste, autores como Edwards (1996) e Krapp (2024) consideram a simulação uma técnica que ainda mantém um vínculo com o real, mesmo que de maneira abstrata ou parcial, voltada à análise ou previsão. Assim, se para Baudrillard a simulação elimina a conexão com o real, do ponto de vista de uma técnica cultural ela não perde o referente por completo: preserva um “ato” fundamental, mesmo sem copiar cada detalhe daquilo que representa.

Se a IA não mimetiza o real, mas constrói plausibilidade interna, podemos dizer que ela, também, simula sem representar. Compreender a IA generativa como uma simulação a situa em uma posição híbrida e singular na história das tecnologias de imagem. Historicamente, a comunicação visual foi moldada por um modelo representacional enraizado na tradição ocidental da mimese (AUERBACH, 1971). Esse princípio tem sido fundamental tanto na representação artística quanto literária, consolidando o realismo como valor estético e cultural central na cultura visual ocidental. Durante o Renascimento italiano, os artistas aperfeiçoaram técnicas miméticas, como a perspectiva geométrica, para simular a tridimensionalidade, valendo-se de dispositivos como a *camera obscura* a fim de intensificar a ilusão de realidade (PANOFSKY, 1997; CRARY, 2012; KITTLER, 2016).

O advento da fotografia e do cinema – primeiras formas de imagem técnica – não só expandiu as técnicas renascentistas, mas também introduziu uma nova dimensão à comunicação visual ao integrar convenções miméticas a um aparato mecanizado (FLUSSER, 2011; MIRZOEFF, 1999). Para Dubois (2012), a fotografia foi recebida pelo olhar ingênuo do século XIX como uma reprodução objetiva do real, sendo considerada mimética em sua essência. Até meados do século XX, a fotografia foi amplamente considerada a forma mais

perfeita de imitação da realidade¹¹, graças à sua natureza técnica e ao procedimento mecânico, que possibilitavam a criação de imagens de maneira “automática” e “objetiva”, sem a intervenção direta do artista (BAZIN, 1991). Essa característica a distinguiu das obras de arte tradicionais, vistas como fruto do trabalho manual e do talento criativo (KITTNER, 2016). A dicotomia entre fotografia e arte reforçava a separação entre indústria e pintura, influenciando os discursos da época, que ora exaltavam sua precisão mimética, ora a criticavam por sua suposta falta de valor artístico (DUBOIS, 2012).

Para Bazin (1991), a fotografia abriu caminho para a abstração e as formas não representacionais nos movimentos de vanguarda. Já na década de 1960, a videoarte e a arte computacional inauguraram uma mudança ainda mais profunda (MACHADO, 1996), rejeitando a mimese do mundo visível e subvertendo os objetivos tradicionais das imagens técnicas (MACHADO, 1993). Em vez de reproduzirem realidades externas, essas obras tratam a imagem como matéria-prima a ser distorcida, sintetizada e reconfigurada, enfatizando os processos internos da máquina. Em outras palavras, o vínculo com o “real” deixa de ser estabelecido por uma relação direta com o mundo externo, passando a ser determinado pelo próprio aparelho e pelo ato de produção. Mesmo as obras que capturam apenas a interferência eletrônica carregam o indício da concretude de seu aparelho. Dessa forma, a tela transcende sua função tradicional de “janela” para o mundo e se transforma em uma superfície dinâmica e interativa, onde as formas se dissolvem, fragmentam e se fundem, evidenciando a natureza construída e mediada das imagens (MACHADO, 1996). Ainda assim,

Enquanto na TV, no cinema e na fotografia o *hardware* óptico simplesmente faz o que tem de fazer, a óptica no computador é virtual e, portanto, limitada apenas pelos recursos de armazenamento e processamento. Antes da camada sensual da percepção, os gráficos do computador simulam uma óptica (*hardware*), onde apenas os algoritmos (*software*) reinam (MONTAÑA; VAGT, 2017, p.2, tradução nossa).

Logo, as imagens de IA generativa não se enquadram nem na tradição mimética da fotografia e do cinema, nem na abordagem material da videoarte, pois emergem de um sistema estatístico fechado, baseado em padrões extraídos de grandes bases de dados (DONDERO ET AL., 2024). Dessa forma, elas não referenciam diretamente o mundo externo a elas, mas constituem sínteses algorítmicas que apenas parecem plausíveis dentro de um modelo matemático – isto é, dentro de uma simulação.

¹¹ Em *Non-Human Photography*, Joana Zylinska (2017) desenvolve uma contra-história da fotografia a partir de seu viés não mimético. Como exemplo, a autora argumenta que “Vista do Boulevard do Templo” (Daguerre, 1839) pode ser considerada não mimética, pois o longo tempo de exposição converteu uma cena urbana em um deserto irreal.

Embora não imitem o real – como evidenciado pela pele plástica da menina na Figura 2 ou pelos letreiros impossíveis na Figura 4 – as imagens geradas por IA mantêm algum tipo de relação com referentes externos, ainda que de forma estocástica. São, portanto, antimiméticas, mas não deixam de ser referenciais. Em outras palavras, elas se fundamentam em modelos, conceitos ou estruturas preexistentes que as ancoram em um sistema de referência, o qual descrevemos como mítico.

Assim como simuladores de voo não se preocupam em reproduzir exatamente a aparência da pilotagem, mas sim sua lógica interna, as imagens criadas por IA também operam em um nível em que a verdadeira representação ocorre no modelo subjacente, não na superfície. Em um simulador, o que importa é a precisão do modelo matemático e físico que define aerodinâmica, resposta dos controles e condições de voo – não a fidelidade visual. Por exemplo, as diferenças visuais entre o *Microsoft Flight Simulator* de 1982 e o de 2023 (fig. 7) são apenas interfaces estéticas que não alteram a qualidade do modelo interno. Em suma, a representação está na modelagem subjacente, e não na aparência superficial.



FIGURA 7 – Diferenças na interface do *MS Flight Simulator* 1982 (esquerda) e 2023 (direita).
FONTE - Pesquisa no *Google Images*.

O mesmo princípio se aplica às imagens geradas por IA: o que vemos na superfície não é uma representação em seu sentido tradicional, mas o efeito de um modelo. Enquanto o simulador de voo se ancora nas leis da física para criar uma experiência funcional, a IA generativa baseia-se em um “referente estatístico” – ou seja, um padrão derivado da distribuição probabilística dos dados de seu banco de treinamento. Esse regime reflete tanto as tendências normativas do conjunto de dados quanto as inferências do modelo, desafiando as noções clássicas de representação e mimese. Logo, a IA não mimetiza a realidade, mas produz uma realidade plausível.

Esse fenômeno pode ser ainda mais complexo do que a confusão entre cópia e referente proposta por Baudrillard, pois não apenas dissolve a distinção entre representação e uma cena

pré-existente, mas também desloca a própria relação entre significante e significado. Aqui, o significado não é fixo nem referencial, mas sim uma convergência estatística que busca coerência visual dentro de um modelo fechado de dados. Embora essas imagens aparentem coerência e familiaridade, sua plausibilidade visual não decorre de um vínculo óptico com o mundo externo, mas de padrões probabilísticos inferidos de grandes bases de dados. Assim, a IA generativa não apenas rompe com a tradição mimética, mas inaugura um regime onde a referência cede lugar a uma lógica algorítmica interna (CRARY, 2012).

4. Ecologias

Segundo Dobson (2023), as tecnologias de visão computacional transformam a percepção visual ao tornar o invisível visível, detectar objetos ocultos, rastrear movimentos e, no nosso caso, estruturar dados visuais a partir do passado. Essa capacidade tem borrado a distinção entre visão humana e computacional, visto que esses sistemas foram originalmente desenvolvidos para modelar a percepção humana (KITTLER, 2016). Assim, mais do que meras ferramentas, a visão computacional instituiu novos regimes de visualidade e governança, redefinindo os conceitos de percepção e observação (DOBSON, 2023).

Em um contexto em que as imagens geradas por IA não mais se vinculam diretamente ao mundo material, é preciso questionar qual *modelo de simulação* fundamenta sua lógica de produção. Como argumentaremos a seguir, essas imagens não representam um mundo concreto, mas emergem de um processo de modelagem estatística que se insere em uma tradição computacional consolidada ao longo do século XX. Para investigar essa dinâmica, é necessário situá-las dentro da ecologia midiática computacional, que estruturou a visualidade tecnocultural.

Em *The Closed World: Computers and the Politics of Discourse in Cold War America* (1996), um dos principais estudos sobre as origens do pensamento computacional contemporâneo, Paul N. Edwards demonstra como, durante a Guerra Fria, prioridades como contenção e vigilância incentivaram a criação de sistemas centralizados de comando e controle militar, a exemplo do SAGE¹². Tais projetos avançaram a tecnologia computacional e

¹² O Semi-Automatic Ground Environment (SAGE) foi um sistema computacional de grande porte projetado para integrar e processar dados de múltiplos radares, criando uma visão unificada do espaço aéreo em larga escala. Desenvolvido para o Comando de Defesa Aeroespacial da América do Norte (NORAD), o SAGE coordenava a resposta a possíveis ataques aéreos soviéticos, desempenhando esse papel estratégico entre o final dos anos 1950 e a década de 1980.

consolidaram uma visão em que os conflitos globais passariam a ser gerenciados pela supremacia tecnológica e pelo controle centralizado da informação.

Nesse cenário, os computadores tiveram um papel duplo: ao mesmo tempo em que gerenciavam dinâmicas de poder global, tornaram-se metáforas em teorias sociológicas e narrativas culturais. Segundo Edwards (1996), esse discurso extrapolou o uso militar, incorporando a ideologia do controle tecnológico em práticas culturais e sociais mais amplas da Guerra Fria. Com isso, governos passaram a confiar cada vez mais em dados computacionais para decisões estratégicas, mesmo quando os números não refletiam fielmente as realidades que deveriam representar. Especialmente na estratégia nuclear, sistemas simplificavam questões geopolíticas complexas em conjuntos de dados, de modo que as simulações se tornavam a base para a ação, ainda que não capturassem todo o contexto concreto.

Esse fenômeno consolidou a supremacia ideológica dos sistemas tecnológicos, alicerçada na ideia de que informações geradas por computador seriam inherentemente objetivas. No entanto, Edwards (1996) mostra como essa confiança exacerbava os vieses, limitações e pressupostos embutidos nesses sistemas. Autoridades e líderes militares acabavam operando em um “mundo fechado”, em que sua compreensão dos eventos globais era filtrada pelas abstrações dos dados, afastando-os ainda mais das realidades materiais que pretendiam influenciar.

Knorr-Cetina (2014) desenvolve o argumento de Edwards ao introduzir o conceito de “mídia escópica”. Um escópio é um dispositivo reflexivo de observação e projeção, no qual a audiência se orienta pela realidade projetada, reagindo ao próprio dispositivo como mediador e centro de atenção. Quando tal observador cria e veicula uma representação visual de um fenômeno, o público responde às características refletidas na superfície, em vez de reagir ao que aconteceu antes dessa reflexão (Knorr-Cetina, 2014, p. 40–42).

Quando empregadas sistematicamente, tais mídias causam efeitos de *world-making*, formando realidades paralelas onde os participantes se envolvem ou se absorvem completamente. Desse processo emerge uma “situação sintética” (KNORR-CETINA, 2014, p.48), fluida e em constante monitoramento, que supera barreiras de tempo e espaço, unificando todos diante das mesmas telas – como ocorre, por exemplo, entre *traders* em mercados financeiros fortemente computadorizados (não por acaso, parecidos com simuladores de guerra).

Para Edwards (1996), a metáfora do “mundo fechado”¹³ torna-se, então, espécie de uma profecia autorrealizável, influenciando não só o discurso, mas também o funcionamento das tecnologias computacionais. Na prática, representar a realidade por meio de sistemas computacionais acabou subvertendo o próprio conceito de representação: ao longo do tempo, as simulações passaram a definir e substituir essa realidade, eliminando a diferença entre representação e representado. Dessa forma, a lógica computacional, ao abstrair e internalizar o mundo, nega a ideia de algo externo a seus modelos.

Operando dentro dessa epistemologia computacional, os sistemas de IA generativa são autorreferenciais: funcionam exclusivamente a partir de bancos de dados preexistentes e algoritmos, sem interação direta com o mundo externo. A única entrada possível que vem do usuário é por meio de *prompts*, ou seja, solicitações ao sistema. Embora possa haver uma impressão de autonomia criativa, o usuário não orienta o modelo para *criar* algo específico. Ao inserir comandos textuais, ele solicita ao modelo que *busque* – dentro de sua base de dados – combinações no “espaço latente das imagens potenciais” (MEYER, 2023) que correspondam à descrição fornecida. Derivadas de estruturas de dados internas, essas imagens estabelecem um circuito fechado no qual as saídas algorítmicas são percebidas como realidades plausíveis.

Por isso, argumentamos que a IA generativa marca uma mudança radical nas tecnologias de criação de imagens ao expandir a lógica do mundo fechado para a esfera da representação visual. Em vez de manter uma conexão direta com a realidade material, esses sistemas constroem visualidades que reorganizam a relação entre contexto e significado. Aqui, a principal preocupação não é a correspondência com o exterior, mas a coerência interna dos modelos de dados e algoritmos que os estruturam. Com isso, a produção imagética se afasta tanto da fidelidade óptica quanto da indexicalidade fotográfica, substituindo esses critérios por um paradigma autorreferencial baseado em estatísticas e probabilidades (fig. 8), no qual a aparência de uma captação óptica é simulada por meio de cálculos algorítmicos (MONTAÑA; VAGT, 2017). Essa mudança reflete um movimento mais amplo na cultura contemporânea, no qual realidades sintéticas não apenas se tornam centrais na produção cultural, mas também estabelecem novos regimes normativos de verossimilhança, onde significado e referência são determinados pelos próprios processos computacionais e pela lógica estatística do modelo.

¹³ Para Edwards (1996), o oposto do mundo fechado não é um “mundo aberto”, mas o “mundo verde”, conceito inspirado em Northrop Frye. Esse espaço natural e ilimitado é regido por forças mágicas e naturais, contrastando com a racionalidade e o controle tecnológico do mundo fechado. Seu arquétipo central é a jornada, na qual os personagens buscam integrar-se à complexidade do mundo, e não dominá-lo.

Retomando Crary (2012, p. 21), “o problema da mimese [é] de poder social: um poder fundamentado na capacidade de produzir equivalências”. Se a mimese não se restringe à esfera artística, mas também reflete relações de poder, então um sistema visual antimimético não deixa de ter implicações políticas – possivelmente até mesmo anti-políticas. Afinal, na atual situação, estas imagens, “comandadas por códigos tecnológicos, delineiam e estruturam o código dominante com o qual imaginamos” (MONTAÑA; VAGT, 2017, p.1).



FIGURA 8 – Darth Vader vai à Copacabana nos anos 1970.

FONTE - Autor (2023), a partir do Midjourney v4.

5. Considerações finais

Neste artigo, adotamos e adaptamos a metodologia proposta por Munn et al. (2023) para investigar *outputs* midiáticos (texto, áudio e imagem) gerados por inteligências artificiais generativas (IAGs). Em um breve texto publicado ainda em 2023, Munn et al. sugeriram três eixos de análise: “Desfazer o ecossistema”, que aborda as infraestruturas técnicas, econômicas e corporativas; “Desfazer os dados”, dedicado aos conjuntos de dados de treinamento, com ênfase em suas características e vieses; e “Desfazer o output”, centrado no produto gerado pela IA e nas lógicas subjacentes que se revelam na interação com o sistema.

Todavia, invertemos a ordem desses eixos. Em vez de partir da estrutura macro para o produto final, começamos pela análise das superfícies visíveis das imagens criadas com o protocolo QR ControlNet, seguimos para o exame dos códigos e dados subjacentes e, então, inserimos essas produções em um ecossistema midiático mais amplo. Essa inversão se justifica porque buscamos realçar exemplos específicos capazes de demonstrar como tais imagens

funcionam como sintomas de uma estrutura técnico-cultural mais ampla, que impacta profundamente tanto os modos de produção imagética quanto nossas ideias sobre representação e significação.

Nos últimos dois anos, a produção de imagens por IA Generativa (IAG) tem recebido crescente atenção em diversos campos – desde a semiótica visual (DONDERO ET AL., 2024) até a história da arte (WASIELEWSKI, 2023), estudos de mídia (SALVAGGIO, 2023) e cultura visual (MEYER, 2023). Nossa ponto de vista, fundamentado nas “materialidades da comunicação” (FELINTO, 2001), argumenta que essas imagens estabelecem um regime de produção desvinculado do mundo material, operando em um sistema fechado de algoritmos e dados. Em vez de qualquer conexão indexical, por mais tênue e controversa que esta fosse, a IA generativa baseia-se em padrões estatísticos, tornando a verificação externa irrelevante e deslocando o eixo de autenticidade para a coerência interna dos próprios algoritmos. Sem um “original” a representar, surgem artefatos visuais plausíveis ao observador, mas sintetizados integralmente por processos computacionais.

Esse modelo não funciona como um índice (Peirce) nem como uma representação mimética (Platão, Aristóteles), mas sim como uma simulação operativa (Baudrillard, Edwards, Krapp). Flusser (2008) argumenta que as imagens técnicas não representam, mas operam dentro de uma lógica programática. Em sua discussão, ele descreve um deslocamento em que as imagens deixam de ser janelas para o mundo e se tornam programas de instrução visual – exatamente o que ocorre nas imagens de IA. Essa transformação reforça a autorreferencialidade das mídias digitais e desafia categorias epistemológicas estabelecidas, uma vez que as imagens passam a existir como mundos em si mesmas. Mais do que isso, esse processo conduz à naturalização do objeto imaginado e à reafirmação de valores dominantes na sociedade.

Em resumo, ao substituir a experiência visual baseada na indexicalidade por uma visualidade modelada estatisticamente, as IAGs reconfiguram nossa percepção e questionam os limites entre simulação, simulacro e realidade. Seu propósito não é “capturar”, mas “gerar”, criando mundos visuais que se alinham à percepção humana, embora sejam fundamentalmente sintéticos.

Restam uma série de questões sem respostas. Se a autenticidade das imagens não está mais vinculada a um referente, como se reconfiguram as noções de verdade, memória coletiva e documentação visual? Além disso, é essencial compreender como os algoritmos e a modelagem visual das IAGs influenciam a normatização estética, reforçam vieses e

potencializam a reprodução de padrões hegemônicos. Outra vertente de pesquisa de grande relevância diz respeito à cultura visual e à sociedade, investigando, por exemplo, os impactos dessas imagens no fotojornalismo. Por fim, no campo da ecologia midiática, é necessário investigar as implicações políticas, econômicas e ambientais da dependência crescente de IA na produção imagética, considerando desde os impactos na indústria criativa até a concentração desse poder tecnológico em grandes corporações. Juntas, essas investigações podem oferecer uma compreensão mais ampla dos efeitos da IA generativa na cultura visual contemporânea.

Referências

AUERBACH, E. **Mimesis**: a representação da realidade na literatura ocidental. São Paulo: Perspectiva, 1971.

BARTHES, R. **A Câmara Clara**: notas sobre a fotografia. Rio de Janeiro: Nova Fronteira, 1984.

_____. **Mitologias**. Rio de Janeiro: Bertrand Brasil, 2001.

BAUDRILLARD, J. **Simulacros e Simulação**. Lisboa: Relógio d'Água, 1991.

BAZIN, A. Ontologia da imagem fotográfica. In: BAZIN, A. **O Cinema**: ensaios. São Paulo: Brasiliense, 1991, p.19-25.

BELTING, H. **An Anthropology of Images**: Picture, Medium, Body. Princeton: Princeton University Press, 2011.

BENDER, E.M.; GEBRU, T.; MCMILLAN-MAJOR, A.; SHMITCHELL, S. On the Dangers of Stochastic Parrots: Can Language Models Be Too Big? In: FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, 2021. p. 610–623.

BENJAMIN, W. Pequena história da fotografia. In: _____. **Magia e Técnica, Arte e Política**. São Paulo: Brasiliense, 1996, p. 91-107.

BERGER, J. **Ways of Seeing**. Londres: Penguin Books, 1972.

CRARY, J. **Técnicas do Observador**: visão e modernidade no século XIX. Rio de Janeiro: Contraponto, 2012.

DOBSON, J.E. **The Birth of Computer Vision**. Minneapolis: The University of Minnesota Press, 2023.



DONDERO, M.G.; DE CASTRO, G.H.R.; SCHWARTZMANN, M.N.; PEREIRA, D. Semiótica da inteligência artificial: análise computacional de grandes bases de dados e geração automática de imagens. **MATRIZes**, v.18, n.3, p.29-54, set./dez. 2024.

DUBOIS, P. **O Ato Fotográfico**. Campinas, SP: Papirus, 2012.

EDWARDS, P. N. **The Closed World**: Computers and the politics of discourse in Cold War America. Cambridge: The MIT Press, 1996.

FAROCKI, H. Phantom Images. **Public**, v. 29, 2004. Disponível em: <https://public.journals.yorku.ca/index.php/public/article/view/30354> Acesso 29 jan. 2020.

FELINTO, Erick. Materialidades da Comunicação: Por um Novo Lugar da Matéria na Teoria da Comunicação. **Ciberlegenda**, n. 5, Rio de Janeiro, RJ: 2001.

FLUSSER, V. **O universo das imagens técnicas**: ou elogio da superficialidade. São Paulo: Annablume, 2008.
_____. **Filosofia da Caixa Preta**. São Paulo: Annablume, 2011.

KITTLER, Friedrich. **Mídias ópticas**: Curso em Berlim, 1999. Rio de Janeiro: Contraponto, 2016.

KNORR CETINA, K. Scopic media and global coordination: the mediatization of face-to-face encounters. In: LUNDBY, K. (Org.). **Mediatization of Communication**. Berlim: De Gruyter, 2014, p. 39-62.

KRAPP, P. **Computer Legacies**: digital cultures of simulation. Cambridge: The MIT Press, 2024.

LISSOVSKY, M. **O Phi da Fotografia**. Rio de Janeiro: UFRJ, 2021.

MAAMERI, S. AI Generated QR codes with ControlNet, HuggingFace and Google Colab. **Medium**, 6 ago. 2023. Disponível em: <https://medium.com/@ssmaameri/ai-generated-qr-codes-with-controlnet-huggingface-and-google-colab-a99ffeee2210#>. Acesso 25 jan. 2025.

MACHADO, A. **A Arte do Vídeo**. 2.ed. São Paulo: Brasiliense, 1993.

_____. **Máquina e Imaginário**: O desafio das poéticas tecnológicas. São Paulo: Edusp, 1996.

MEYER, R. The New Value of the Archive: AI Image Generation and the Visual Economy of 'Style'. **Image**, v. 37, n. 1, p. 100-111, 2023.

MIRZOEFF, N. **An Introduction to Visual Culture**. Londres e Nova Iorque: Routledge, 1999.

MITCHELL, W.J. **The Reconfigured Eye**: visual truth in the post-photographic era. Cambridge: The MIT Press, 1992.

MONTAÑA, R.C.; VAGT, C. Constructing the invisible: computer graphics and the end of optical media. **communication+1**, v. 7, n. 1, p.1-23, 2017.

MUNN, L.; MAGEE, L.; ARORA, V. Unmaking AI Imagemaking: A Methodological Toolkit for Critical Investigation. Julho 2023. Disponível em: <[arXiv:2307.09753v1](https://arxiv.org/abs/2307.09753v1)>. Acesso em: 27 jul. 2023.

NAKE, F. Das doppelte Bild. In: PRATSCHKE, M. (org.). **Digitale Form**. (Bildwelten des Wissens, v. 3, n. 2). Berlin: Akademie-Verlag, 2005. p. 40–50.

PAGLEN, T. Operational Images. **e-flux journal**, n.59, nov. 2014.

PANOFSKY, E. **Perspective as Symbolic Form**. Nova Iorque: Zone Books, 1997.

SALVAGGIO, E. How to Read an AI Image: Toward a Media Studies Methodology for the Analysis of Synthetic Images. **Image**, v. 37, n. 1, p. 83-99, 2023.

STAUFF, M. Uma cultura da competição: a contribuição histórica do esporte para a dataficação. Recorde: Revista de História do Esporte, v.17, n.2, p.1-42, jul./dez. 2024.

TELLES, M. Estilo Artístico na Arte Gerada por Inteligência Artificial: um estudo de caso de Jim Lee. In: **ANAIS DO 32º ENCONTRO ANUAL DA COMPÓS**, 2023, São Paulo. Disponível em: <<https://proceedings.science/compos/compos-2023/trabalhos/estilo-artistico-na-arte-gerada-por-inteligencia-artificial-um-estudo-de-caso-de?lang=pt-br>>. Acesso em: 10 Fev. 2025.

WASIELEWSKI, A. “Midjourney Can’t Count”: Questions of Representation and Meaning for Text-to-Image Generators. **Image**, v. 37, n. 1, p. 71-82, 2023.

WILF, E. Y. From media technologies that reproduce seconds to media technologies that reproduce thirds: a Peircean perspective on stylistic fidelity and style-reproducing computerized algorithms. **Signs and Society**, v. 1, n. 2, p. 185-211, 2013.

ZYLINSKA, J. **Non-Human Photography**. Cambridge: The MIT Press, 2017.