

Nikoleta Kerinska

O destino da imagem técnica: uma proposição de classificação

A aparente objetividade das imagens técnicas é ilusória, pois na realidade são tão simbólicas quanto são todas as imagens. Devem ser decifradas por quem deseja captar-lhes o significado. Com efeito, são elas símbolos extremamente abstratos: codificam textos em imagens, são metacódigos de textos. A imaginação, à qual devem sua origem, é capacidade de codificar textos em imagens. Decifrá-las é reconstituiu os textos que tais imagens significam. Quando as imagens técnicas são corretamente decifradas, surge um mundo conceitual como sendo o seu universo de significado. O que vemos ao contemplar as imagens técnicas não é “o mundo”, mas determinados conceitos relativos ao mundo, a despeito da automaticidade da impressão do mundo sobre a superfície da imagem.¹

Vilém Flusser

1. Introdução

Os anos setenta e oitenta do século passado foram marcados por uma expansão sem precedentes das possibilidades audiovisuais nas sociedades ocidentais, impulsionada sobretudo pela televisão a cabo, que permitiu uma segmentação de públicos muito mais refinada, e abriu espaço para canais especializados. Um dos casos mais emblemáticos foi o surgimento da MTV em 1981, cuja programação contínua de videoclipes transformou a relação entre música, imagem e consumo cultural, constituindo um vetor central de publicidade e de direcionamento de estilos de vida para públicos jovens. Paralelamente, tecnologias como *teletext*, *walkman*, *fax*, *telemática*, secretárias eletrônicas, e os primeiros computadores pessoais, tornaram-se parte do cotidiano, ampliando a circulação de conteúdos e a portabilidade da experiência audiovisual, num contexto marcado por transformações sociais e econômicas profundas, como o avanço do neoliberalismo.

A circulação das imagens eletrônicas acelerou-se graças à consolidação de infraestruturas televisivas, acompanhada pelo aumento da capacidade de arquivamento e distribuição de conteúdos. A televisão consolidou-se como o principal meio de difusão audiovisual, enquanto o crescimento de catálogos e de bases de dados começou a transformar a forma como conteúdos audiovisuais eram organizados e acessados, antecipando práticas de seleção algorítmica que só se tornariam dominantes décadas depois. A década de 80 também testemunhou a intensificação do consumo de imagens como forma de identidade cultural, como videoclipes, publicidade segmentada, e produtos midiáticos circulando em escala internacional. Surgiu assim um novo regime de comunicação visual, no qual a imagem técnica passou a estruturar comportamentos,

¹ Flusser, V. (2002). *A filosofia da caixa preta: ensaios para uma futura filosofia da fotografia*. Rio de Janeiro: Ed. Relume Dumará, Coleção Conexões, p. 14-15.

desejos e modos de participação cultural – um prenúncio direto da cultura audiovisual hiperconectada, que viria a seguir.

Neste período, Vilém Flusser questionava já o funcionamento e a circulação das imagens na nossa sociedade. Definir e entender a imagem técnica, e a emergência de uma “tecno-imaginação” era para este filósofo uma etapa essencial do entendimento da nossa própria civilização: “Se levarmos em conta a incrível variedade de superfícies multicoloridas que nos cercam desde a Segunda Guerra Mundial, qualquer tentativa de encontrar uma definição capaz de abranger tanta diversidade parece fadada ao fracasso. À primeira vista, parece não haver nada em comum entre os inúmeros símbolos coloridos que chamam nossa atenção nas ruas, vitrines, paredes de casas e praças públicas, livros e televisão, camisetas e rótulos de garrafas, ferramentas e gadgets.”²

Para Flusser, essa ‘selva visual’ que caracteriza o presente não é um fenômeno acidental, mas uma expressão profunda do “espírito do tempo”³, isto é, da forma como a cultura ocidental passou a organizar simbolicamente o mundo por meio de superfícies codificadas. O filósofo identificou nessas manifestações um padrão epistemológico que, apesar de sua aparente heterogeneidade visual, constitui aquilo que ele descreve como “o primeiro verdadeiro estilo do Ocidente, depois do período gótico”⁴. Trata-se de um ‘estilo’ que não se define por coerência formal, mas pela lógica comum determinada pelos dispositivos, que geram este mundo codificado. A força analítica dessa observação reside justamente na percepção de que a unidade do nosso tempo não está na forma, mas no modo de codificação, na estrutura programática que subjaz às imagens técnicas, e que as torna, simultaneamente, múltiplas na aparência e homogêneas na operação.

Contudo, a razão fundamental da importância dessas imagens, para Flusser, era seu poder de comunicabilidade: “(...) descobrimos, conhecemos e avaliamos o mundo em que vivemos através de mensagens recebidas das superfícies, e é por isso que agora estamos submersos num mar de cores e gadgets. Porque já não somos programados prioritariamente por textos impressos em preto e branco, mas por imagens técnicas em technicolor.”⁵

² Flusser Vilém, (2022) *Mutations dans les Relations Humaines : De la Communicologie*, Ed. HD/Le Bon Voisin / AKA, p. 253. Em original: « Si nous tenons compte, toutefois, de l’incroyable variété de surfaces multicolores qui n’ont cessées de nous environner depuis la seconde guerre mondiale, toute tentative de trouver une définition susceptible d’englober tant de diversité paraît vouée à l’échec. A première vue, il semble qu’il n’y ait rien de commun entre les innombrables symboles colorés qui s’imposent à notre attention dans les rues, les vitrines, sur les murs des maisons et des places publiques, dans les livres et sur le petit écran, sur les tee-shirts et les étiquettes des bouteilles, sur les outils et les gadgets. »

³ *Ibid.*, p. 253. Em original : « l’esprit du temps ».

⁴ *ibid.*, p. 254. Em original : « On peut dire que ce premier véritable style occidental depuis le gothique est ce qu’il y a de commun à tous les éléments hétérogènes dans notre monde codé ».

⁵ *Ibid.*, p. 256. Em original : « (...) nous découvrons, connaissons et évaluons le monde dans lequel nous vivons à travers des messages reçus des surfaces, et que c’est là la raison pour laquelle nous sommes à présent submergés par une mer de couleurs et de gadgets. Parce que nous ne sommes plus prioritairement programmés par des textes imprimés en noir et blanc, mais par des techno-images en technicolor. »

Tem-se a impressão de que hoje a voz do filósofo ressoa com uma nitidez ainda mais intensa, pois estamos mais do que nunca cercados por ‘superfícies multicoloridas’, que emergem de todo tipo de telas e de gadgets. Consequentemente, somos programados por elas – por essas imagens técnicas que moldam percepções, afetos e modos de estar no mundo, e que desde os meados do século passado não pararam de evoluir do ponto de vista de seus dispositivos de geração. Transpondo as palavras de Flusser para o contexto contemporâneo, elas se revelam de uma pertinência inquietante, e esta constatação é o ponto de partida da minha reflexão, que retoma o conceito de imagem técnica para examinar suas manifestações mais recentes.

O texto que segue adota o ponto de vista de uma artista, pesquisadora e professora de arte computacional, mais do que o de uma reflexão filosófica propriamente dita. Meu interesse se volta, em particular, para as imagens produzidas por inteligências artificiais generativas, cada vez mais presentes em plataformas digitais, redes sociais, motores de busca, ambientes de trabalho criativo, publicidade, interfaces de consumo cultural e, de modo crescente, em todos os espaços onde a visualidade se tornou um vetor dominante de comunicação. Esta é, portanto, uma tentativa de compreender as formas de vida dessas imagens, tentar penetrar sua natureza e sua origem, que permanecem opacas, quase mágicas, à olho nu.

Esta reflexão propõe examinar as imagens produzidas por inteligências artificiais a partir de seu método de geração específico, buscando no pensamento flusseriano antecipações e ecos proféticos que permitem elaborar um pensamento crítico sobre essas tecnologias emergentes. Trata-se de compreender como essas imagens constituem uma *nova geração de imagens técnicas*, cuja característica estrutural mais marcante é a *lógica de amostragem*, isto é, a *recombinação estatística de vastos conjuntos de dados visuais*, que substitui as referências visuais do mundo por um processo de síntese probabilística.

Tentarei por meio de uma análise comparativa cercar as principais características das imagens geradas por I.As., e, propondo um métodos de classificação do grande grupo das imagens técnicas, compreender como elas se situam neste universo. O termo ‘classificação’ me parece o mais apropriado para este estudo, pois ele se refere a um conjunto de imagens que compartilham características funcionais bem definidas, geralmente com base em critérios técnicos ou operacionais. Desde o início é importante demarcar uma escolha consciente, um posicionamento: as imagens fotográficas, as imagens de síntese, e as imagens generativas podem ser consideradas classes distintas da imagem técnica, pois cada um desses grupos corresponde a um modo específico de produção e a um regime técnico particular. Essas classes podem ser consideradas também como ‘gerações’ da imagem técnica de ponto de vista da evolução dos dispositivos. A noção de ‘geração’ remete à sucessão histórica dos dispositivos técnicos de produção de imagem, marcando rupturas nos modos de concepção e nos regimes de visualidade. Cada geração de dispositivos dá origem a

novas classes de imagens técnicas, definidas por suas condições materiais de produção, suas lógicas operatórias e seus modos de inscrição do visível.

Nesta linha de raciocínio, inicialmente mobilizarei a definição flusseriana de imagem técnica; depois abordarei a imagem de síntese⁶; em seguida, compararei suas especificidades às das imagens generativas, a fim de melhor elucidar as diferenças entre essas duas classes. Como conclusão, apresentarei algumas pistas que poderão orientar reflexões futuras.

2. O ponto de partida: a definição flusseriana de ‘imagem técnica’

Vilém Flusser define a imagem técnica com “uma imagem produzida por aparelhos”⁷. Tais aparelhos, por sua vez, operam como uma “caixa preta”⁸, isto é, um dispositivo que recebe um *input* para gerar a imagem, e cuja principal especificidade reside justamente no fato de que o processo de geração da imagem ocorre no interior dessa caixa preta, opaca ao usuário. A consequência deste fato é significativa: “Os aparelhos são produtos da técnica que, por sua vez, é texto científico aplicado. Imagens técnicas são, portanto, produtos indiretos de textos – o que lhe confere posição histórica e ontológica diferente das imagens tradicionais.”⁹ Neste sentido, as imagens técnicas não prolongam a mão humana, mas a própria racionalidade científica que as estrutura. Elas emergem de aparelhos programados por teorias, e é essa mediação que redefine o gesto de ver e de produzir visibilidade. A constatação do estatuto ontológico e genealógico radicalmente distinto das imagens técnicas em relação às imagens tradicionais ocupou consideravelmente intelectuais, artistas, pesquisadores e pensadores que questionaram seu poder expressivo, ideológico e artístico desde o seu surgimento com a fotografia.

A invenção da fotografia inaugura a era das imagens técnicas, que designa uma ruptura decisiva no universo da imagem: pela primeira vez, o real parece inscrever-se a si mesmo, dispensando a habilidade do artista. Esse realismo técnico é também um momento de liberação para as artes visuais, que deixam de carregar o peso da representação fiel e se abrem a outras formas de sensibilidade, mais subjetivas, simbólicas ou experimentais. Ao mesmo tempo, a imagem se democratiza, circula, e graças a sua reprodutibilidade técnica, infiltra-se no cotidiano e nas práticas científicas, jornalísticas e policiais, adquirindo o estatuto ambíguo de prova, testemunho e narrativa condensados num instante.

A fotografia também altera profundamente nossa relação com o tempo, com o olhar e com o próprio ato de conhecer. Ela captura o efêmero, fragmenta o contínuo, prepara o terreno para a

⁶ A imagem de síntese pode-se denominada também imagem computacional.

⁷ Flusser, Vilém (2002), *Op. Cit.* p. 13.

⁸ *Ibid.*, p. 15.

⁹ *Ibid.*, p. 13.

sensibilidade moderna e para a proliferação de grande variedade das imagens técnicas. Ela funda a genealogia da imagem automática, posteriormente cinematográfica, televisiva e videográfica.

A passagem do traço ao pixel e ao bit, tal como pode ser reconstruída a partir de *Filosofia da Caixa Preta*, descreve uma escalada de abstração na história das imagens técnicas. O traço representa o gesto humano originário, a inscrição manual que ancora a imagem na experiência corporal e na intenção subjetiva. Com o advento dos aparelhos, esse gesto é substituído pelo pixel, unidade mínima da imagem eletrônica, que já não deriva da mão, mas de cálculos que fragmentam o contínuo em pontos discretos. O bit, por sua vez, leva essa abstração ao limite: ele dissolve o pixel em pura informação, reduzindo a imagem a operações binárias que podem ser recombinadas indefinidamente por programas¹⁰. Assim, a imagem técnica deixa de ser uma superfície inscrita e passa a ser um evento informacional, cuja “materialidade” depende inteiramente de processos computacionais que permanecem invisíveis ao observador.

É justamente nesse ponto que a definição flusseriana de imagens técnicas ganha importância: ela permite compreender a fotografia não apenas como um novo meio, mas como o início de um regime imagético programado, no qual o aparelho passa a mediar, condicionar e orientar o gesto de criar. Flusser desloca a discussão da representação para a operacionalidade, mostrando que essas imagens não resultam de uma intenção expressiva direta, mas de um jogo entre o humano e o programa inscrito no aparelho.

A definição de Flusser permite, por um lado, considerar todas as imagens geradas por aparelhos ou máquinas como uma mesma família e, por outro, nos estimula, a cada novo dispositivo reativar a pergunta: “Que imagem é essa, e quais são suas relações com o dispositivo de produção?” Por exemplo, enquanto a fotografia analógica que registra a luz como uma inscrição contínua em suporte físico – um traço químico sensível ao mundo – a fotografia digital converte essa mesma luz em pixels e, por fim, em bits, transformando a imagem em informação calculável e programável. Seguindo as ideias e a intuição de Flusser, é possível responder a essa pergunta unicamente decifrando os modos de funcionamento dos dispositivos e seus métodos de automatização.

3. Da imagem de síntese a síntese estatística de informações

O termo “imagem de síntese” se refere às imagens produzidas por códigos e modelos matemáticos, consideradas neste texto como a segunda classe de imagens técnicas. As imagens de síntese foram amplamente teorizadas por autores como Edmond Couchot, Philippe Quéau e Paul Virilio entre

¹⁰ Flusser, Vilém (2002), *Op. Cit.* p. 63-65.

outros, que evidenciam a passagem do registro fotográfico para a simulação computacional. Uma breve análise da imagem de síntese é necessária para a compressão das imagens geradas por inteligências artificiais, assim como para situá-las como a terceira classe de imagens técnicas. As imagens de síntese precedem cronologicamente as imagens geradas por inteligências artificiais, cujo marco inicial pode ser situado nos primeiros experimentos de redes neurais generativas na década de 2010, e que constituem o foco central deste artigo.

No contexto da imagem de síntese, é importante reconhecer que o desenvolvimento das linguagens de programação e dos primeiros softwares de modelagem tridimensional instaurou esse *novum* na época regime de produção visual baseado na formalização matemática e na manipulação algorítmica da forma¹¹. Linguagens como C, Pascal ou FORTRAN, inicialmente concebidas para fins científicos e de engenharia, tornaram-se o suporte estrutural para a criação de imagens de síntese, enquanto os primeiros sistemas de modelagem 3D – inicialmente restritos a laboratórios, universidades e estúdios pioneiros – inauguraram a possibilidade de construir ambientes inteiramente virtuais. Esses dispositivos não apenas ampliam o campo da criação visual, mas consolidam a imagem técnica como entidade calculada, navegável e transformável.

O universo das imagens de síntese viveu uma expansão exponencial na década de 90, que foi acompanhada por uma euforia perceptível entre artistas, pensadores e pesquisadores que viam nesse novo regime visual a promessa de uma transformação profunda no campo da imagem, assim como no campo da arte. As imagens de síntese e sua manifestação mais complexa – a realidade virtual – foram proclamadas como a arte do século XXI¹².

Pierre Lévy imaginava os mundos virtuais como a expressão concreta da cibercultura: “O gênero canônico da cibercultura é o mundo virtual. Não devemos entender esse termo no sentido estrito da simulação computacional de um universo tridimensional explorado com um capacete estereoscópico e datagloves. Vamos antes apreender o conceito mais geral de uma reserva digital de virtualidades sensoriais e informacionais que só se atualizam na interação com os seres humanos. De acordo com os dispositivos, essa atualização é mais ou menos inventiva, imprevisível, e deixa uma parte variável para as iniciativas daqueles que nela mergulham. Os mundos virtuais podem eventualmente ser enriquecidos e percorridos coletivamente. Tornam-se, nesse caso, um lugar de encontro e um meio de comunicação entre seus participantes. O engenheiro de mundos surge, então, como o grande artista do século XXI. Ele provê as virtualidades, arquiteta os espaços de

¹¹ A imagem de síntese, chada também de ‘computação gráfica’ surgiu pela primeira vez no início da década de 1950, com imagens eletrônicas experimentais como os “Oscillons” de Laposky, e tornou-se um campo tecnológico formal na década de 1960 com o “Sketchpad” de Sutherland.

¹² Alguns exemplos de tais publicações: *Cyberarts: Exploring Art & Technology* de Linda Jacobson (1992); *Esthétique des arts médiatiques* de Louise Poissant (1995); *The Art of Human-Computer Interface Design* de Brenda Laurel e S. Joy Mountford (1990); *Arte no Século XXI: A Humanização das Tecnologias* de Diana Domingues (1997).

comunicação, organiza os equipamentos coletivos da cognição e da memória, estrutura a interação sensório-motora com o universo dos dados.”¹³

Para este filósofo, a arte do próximo milênio é concebida coletivamente por múltiplos criadores em ambientes computacionais compartilhados, onde a obra deixa de ser objeto para tornar-se processo, espaço e experiência¹⁴ - em outras palavras uma imagem espaço – temporal imersiva, conectada, compartilhada.

Nesse horizonte, a obra de arte enquanto processo se caracteriza por sua natureza evolutiva e aberta, continuamente atualizada pelas interações dos participantes e pelas transformações internas do sistema, de modo que a criação se realiza em fluxo permanente, sem ponto de conclusão. Enquanto espaço, os mundos virtuais configuram ambientes navegáveis e habitáveis, territórios simbólicos que organizam a percepção e a ação, instaurando formas de presença que, embora não materiais, estruturam modos de circulação, encontro e construção coletiva.

A experiência dos mundos virtuais se efetiva na imersão sensorial e cognitiva do usuário, que não apenas observa, mas participa, intervém e coproduz o sentido do ambiente, tornando cada vivência singular e irrepitível. Assim, para Lévy, os mundos virtuais constituem propostas artísticas conectadas e compartilhadas, nas quais criação, interação e vivência se entrelaçam, redefinindo profundamente o estatuto da obra, do autor e do público na cultura digital.

Com as práticas artísticas da década de 90, ressentiu-se a necessidade de um vocabulário específico para nomear as imagens de síntese. Expressões como *imagens 3D*, *imagens computacionais*, *infografias*, *gráficos computacionais*, *imagens virtuais* e, finalmente, *imagens de síntese* foram surgindo em busca de uma terminologia apropriada, que designava melhor sua natureza. A expressão ‘imagem de síntese’ é particularmente reveladora: ela designa imagens produzidas pela síntese de informações abstratas, pela construção de modelos matemáticos e pela manipulação algorítmica da forma. Afirma-se deste modo a identidade de uma imagem concebida computacionalmente, que deriva de procedimento de cálculo e de operações matemáticas.

Segundo Edmond Couchot, a imagem de síntese constitui uma espécie singular no universo das imagens técnicas, justamente porque se desconecta de modo inédito da realidade visível, e das operações de figuração pré-existentes: “de maneira geral, todas as operações de figuração fundadas na ótica geram imagens que ‘colam’ ao real, imagens das quais cada ponto está ligado ao real, pela lógica projetiva da representação.”¹⁵ Na imagem fotográfica analógica, cada ponto corresponde a um ponto do objeto real, que define uma continuidade física, uma filiação direta entre o mundo e sua inscrição visual. Já no domínio computacional, essa correspondência desaparece: “nenhum

¹³ Lévy P. (1999). *Cibercultura*, (Trad. de Carlos Irineu da Costa) São Paulo: Ed. 34, 1999, p. 144.

¹⁴ *Ibid.* p.135 - 145.

¹⁵ Couchot, Ed. (1999). « Da representação à simulação » in *Imagem Máquina*, André Parente (org.), Rio de Janeiro: Éd. 34, p. 41.

ponto de qualquer objeto real pré-existente/corresponde ao pixel. O pixel é a expressão visual, materializada na tela, de um cálculo efetuado pelo computador, conforme as instruções do programa.”¹⁶ Se algo preexiste à imagem de síntese, não é o objeto, mas a linguagem de programação e o cálculo que a engendram. Trata-se, portanto, de uma lógica figurativa distinta, na qual a imagem não representa o real, mas o *simula*.

A noção de simulação é igualmente central nos trabalhos de Philippe Quéau¹⁷, que sublinha o papel determinante dos modelos matemáticos na geração das imagens de síntese: “A imagem de síntese remete necessariamente ao seu modelo. Ela não pode compreender-se senão através da sua relação com ele”¹⁸. Aqui, o modelo não é um referente externo, mas uma estrutura formal que antecede e organiza a imagem. A imagem de síntese é a visualização de seu próprio sistema de regras.

Evidencia-se assim uma ruptura epistemológica profunda com as primeiras imagens técnicas: a imagem deixa de ser uma captura do real para tornar-se um produto de operações abstratas. Ela não testemunha, mas projeta; não registra, mas reconstitui. A consequência é uma distinção no modo de funcionamento entre a primeira geração de imagens técnicas e a segunda, na qual o campo da representação é substituído pela modelização, inaugurando um regime de visualidade em que a imagem é sempre o resultado de um cálculo.

A imagem de síntese distingue-se também pela sua capacidade de incorporar qualquer tipo de imagem pré-existente, e transcrita em código binário. Isto determina uma de suas características mais importantes: a flexibilidade estrutural e a consequente abertura ao hibridismo. Para a imagem transcrita em dados numéricos manipuláveis, a montagem torna-se um procedimento quase natural, pois é justamente por meio dela que diferentes materiais visuais podem ser combinados, sobrepostos e reconfigurados. Embora a montagem de imagens já fosse possível na fotografia analógica, ela se torna incomparavelmente mais simples, precisa e difundida com a fotografia digital e, sobretudo, com as imagens de síntese. Fotografias, vídeos, desenhos, modelos tridimensionais ou sinais captados por sensores podem ser convertidos em estruturas numéricas, permitindo que elementos antes separados por suportes, técnicas ou lógicas de produção sejam integrados numa mesma proposta visual. Essa convergência amplia radicalmente o campo da imagem, abrindo-o a transformações, recombinações e multiplicações praticamente ilimitadas, nas quais o hibridismo não é apenas uma possibilidade, mas uma condição constitutiva do próprio meio digital.

¹⁶ *Ibid.*, p. 42. Obs.: Neste texto Edmond Couchot se refere às imagens geradas diretamente no computador, por meio de programas e linguagens de programação específicos. Aparentemente o autor não leva em consideração a fotografia digital.

¹⁷ Ver Quéau, Ph. (1989) *Metaxu : théorie de l'art intermédiaire*, Ed. Camp Vallon et *Éloge de la simulation : de la vie des langages à la synthèse des image* (1986), Ed. Camp Vallon.

¹⁸ Quéau, Ph. (1999) «O tempo do virtual» in *Imagem Máquina* André Parente (ord.), Éd. 34, Rio de Janeiro, p.92.

Nesse contexto, a imagem deixa de ser um objeto estável para tornar-se um conjunto de operações possíveis: ela pode ser decomposta, recombinada, simulada, animada, deformada ou expandida segundo lógicas que não pertencem unicamente ao domínio óptico, mas ao domínio algorítmico. A imagem gerada por síntese de dados não apenas produz novas imagens; ela absorve e transforma todas as outras, instaurando um regime de *virtualidade*¹⁹ total.

É justamente nesse terreno da simulação computacional que se inscreve a passagem para as imagens generativas produzidas por inteligências artificiais. Se, como mostram Couchot e Quéau, a imagem de síntese já não representa o real, mas o simula a partir de modelos matemáticos, as imagens geradas por I.As. deslocam ainda mais esse paradigma: elas não dependem de um modelo matemático ou geométrico explícito, mas de redes neurais treinadas e modeladas sobre corpora de dados em larga escala, capazes de inferir padrões e produzir novas variações. A simulação deixa de ser a execução de um modelo formal para tornar-se uma *síntese estatística de informações*, cujos resultados são objetos visuais denominados *imagens generativas*. A expressão *síntese estatística de informações* descreve bem o funcionamento dos modelos generativos: eles não criam imagens a partir do nada, nem recuperam imagens pré-existentes, mas sintetizam combinações inéditas a partir de padrões estatísticos treinados em grandes conjuntos de dados.

4. As imagens geradas por inteligências artificiais

As imagens geradas por inteligências artificiais configuram um campo de produção recente²⁰, e os estudos sobre elas e seu impacto na indústrias criativas e na sociedade ainda está por vir. Para Lev Manovich, essas imagens iniciam um momento decisivo na nossa história: “A nova revolução da “mídia generativa” das I.As. parece ser tão significativa quanto qualquer uma das revoluções precedentes. Na verdade, é possível que seja tão significativa quanto a invenção da fotografia no século XIX, ou o uso da perspectiva linear na arte ocidental no século XVI.”²¹ Antes de discutir as recorrências das imagens geradas por I.As., sugiro uma breve análise dos princípios de funcionamento dos programas responsáveis pela sua existência.

É importante notar que as modalidades de geração de imagens por I.As. configuram um campo heterogêneo, no qual diferentes formas de entrada – texto, esboço, imagem, layout ou estruturas semânticas – orientam modos distintos de produção. Do *text-to-image*, que traduz

¹⁹ O termo ‘virtualidade’ é usado para designar possibilidades futuras, apontando para tudo aquilo que ainda não se atualizou, mas que existe como potência pronta para emergir em diferentes direções.

²⁰ Destacamos lançamentos significativos de software e os algoritmos subjacentes que moldaram o panorama das I.As. generativas de 2021 a 2024 por empresas como Google, Meta, Microsoft e OpenAI.

²¹ Manovich, L. (2023). “AI Image Media through the Lens of Art and Media History” In: IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft. Generative Imagery: Towards a ‘New Paradigm’ of Machine Learning-Based Image Production, Jg. 19, Nr. 1, S. 34–41. DOI: <https://doi.org/10.25969/mediarep/22323>

descrições verbais em superfícies imagéticas, ao amplo conjunto de operações *image-to-image* (como *sketch-to-image*, *style transfer*, *inpainting* ou *super-resolution*), cada procedimento revela uma articulação específica entre código, modelo e dados de *input*. A essas modalidades se somam abordagens estruturais, como *layout-to-image*²² e *scene-graph-to-image*,²³ que organizam a síntese visual a partir de mapas espaciais ou relações semânticas, além de modalidades híbridas que combinam múltiplos vetores de orientação. Contudo, embora os métodos mencionados operem por vias distintas, todos se alinham à concepção de Flusser segundo a qual “funcionar é permutar símbolos programados”²⁴. Essa convergência evidencia uma genealogia técnica contínua, na qual a lógica das imagens programadas descritas por Flusser se prolonga e se transforma, agora potencializada pela plasticidade algorítmica dos modelos generativos, que expandem e reconfiguram o próprio regime de programação.

Todas as modalidades mencionadas merecem um exame minucioso, mas neste texto a minha atenção dirige-se ao método *text-to-image*. Analisar o método *text-to-image* não é uma escolha arbitrária, mas estratégica: ele concentra, de forma exemplar, as tensões que atravessam a geração de imagens por I.As. É nesse procedimento que a relação entre linguagem e visualidade se torna mais explícita, revelando como os modelos traduzem descrições verbais em superfícies imagéticas e, portanto, como operam suas gramáticas internas, seus vieses e suas potências semânticas. Outras modalidades – *sketch-to-image*, *layout-to-image*, *style transfer* – são igualmente relevantes, mas distribuem as funções operacionais entre humano e sistema computacional de forma segmentada e contextualmente delimitada. Por sua vez, o *text-to-image* condensa o gesto generativo em sua forma mais radical, pois transforma diretamente um enunciado em imagem, simulando uma espécie de *imaginação algorítmica*.

Além disso, o *text-to-image* tornou-se o paradigma dominante na pesquisa recente, tanto pela sua difusão cultural quanto pelo impacto que produz nas discussões sobre autoria, criatividade, veracidade e estética. Ele funciona como um laboratório privilegiado para observar como os modelos articulam referências, sintetizam estilos e produzem coerência visual a partir de instruções abstratas. Em outras palavras, ao estudar o *text-to-image*, analisamos o ponto em que a imagem generativa se mostra mais explícita em seus mecanismos e mais decisiva em suas implicações.

Para entrar no mérito da questão das imagens geradas por I.As., proponho examinar o funcionamento do Midjourney. Esta plataforma é considerada entre as melhores na geração de

²² Transforma esboços estruturais (caixas, posições, tamanhos e rótulos de objetos) em uma imagem completa e realista. O modelo recebe um *layout* - por exemplo, “uma pessoa aqui, uma árvore ali, um carro ao fundo” - e gera uma imagem respeitando essa organização espacial.

²³ Converte um grafo semântico de cena (objetos + relações entre eles, como “gato em cima da mesa”, “janela atrás da cadeira”) em uma imagem coerente. O foco está nas relações e na estrutura semântica, não apenas na posição. Se quiser, preparo versões mais formais, mais didáticas para estudantes, ou mais concisas para slides.

²⁴ Flusser, Vilém (2002), *Op. Cit.* p. 25.

imagens e conteúdos visuais. Seus mecanismos generativos são baseados em modelos de difusão: uma tecnologia que transforma gradualmente ruído aleatório em uma imagem coerente guiada por um texto descritivo. Uma vez que o *prompt* é escrito e fornecido, o Midjourney começa convertendo cada palavra em uma representação matemática que captura seu significado, estilo e possíveis associações visuais. A partir de uma ‘nuvem’ de pixels totalmente caótica, a I.A. aplica uma série de etapas sucessivas para remover o ruído e fazer surgir formas, texturas e cores correspondentes à sua descrição. Esse processo é iterativo: a imagem se torna mais nítida aos poucos, como se estivesse se materializando ponto por ponto feito um quebra cabeça no qual cada pixel encontra seu lugar, depois de tentativas repetidas.

O outro aspecto fundamental do funcionamento do Midjourney reside no seu treinamento em imensos conjuntos de imagens e textos. Graças a esse *aprendizado*, o modelo desenvolve uma compreensão estatística dos estilos artísticos, objetos, ambientes luminosos, composições visuais e até mesmo tendências estéticas. Ele não copia imagens existentes, mas combina os padrões que aprendeu para produzir *algo novo*. É essa capacidade de fundir conceitos, interpretar estilos e extrapolar detalhes plausíveis que dá ao Midjourney seu *caráter criativo* e muitas vezes surpreendente.

No caso de Midjourney, como das outras plataformas que usam modos de criação semelhantes, os algoritmos de geração de imagens são especialmente aqueles baseados em redes neurais profundas. Eles funcionam segundo um princípio de seleção estatística a partir de vastos conjuntos de dados visuais. Em vez de representar um objeto pré-existente ou simular um modelo geométrico explícito, esses sistemas aprendem regularidades estatísticas de elementos de imagens e, em seguida, geram novas configurações visuais recombinaando as distribuições iniciais. Cada imagem produzida é, portanto, o resultado de um cálculo probabilístico, em que cada pixel ou fragmento formal é resultado de uma estimativa, média ou variação dentro de um espaço latente aprendido. Nessa perspectiva, torna-se pertinente qualificar essas produções como ‘imagens de amostragem estatística’: elas não derivam de um referente real, nem de um modelo formal estável, mas de um processo de geração baseado na densidade das ocorrências passadas. Essa definição destaca sua natureza probabilística, seu enraizamento em dinâmicas de dados e seu status ontológico singular – ao mesmo tempo sintéticas e sem origem concreta atribuível (a não ser uma imensa quantidade de dados).

No início do seu livro *O universo das imagens técnicas: elogio da superficialidade*, Flusser expõe uma tese: “as novas imagens não ocupam o mesmo nível ontológico das imagens tradicionais, porque são fenômenos sem paralelos no passado. As imagens tradicionais são superfícies abstraídas de volumes, enquanto as imagens técnicas são superfícies construídas com pontos.”²⁵ Ao distinguir as imagens técnicas das imagens tradicionais em termos ontológicos, Flusser fornece uma chave

²⁵ Flusser, V. (2012) *O universo das imagens técnicas : elogio da superficialidade*, Coimbra : Ed. AnnaBlume, p.16.

decisiva para compreender as imagens generativas. Se as imagens tradicionais são ‘superfícies abstraídas de volumes’, elas mantêm ainda uma relação de continuidade com o mundo sensível, baseada na projeção de um gesto, de um ponto de vista ou de uma experiência perceptiva. As imagens técnicas, por sua vez, sendo ‘superfícies construídas com pontos’, inauguram um regime inteiramente distinto: não resultam da redução de um real pré-existente, mas da composição de unidades discretas calculáveis, cuja organização decorre de programas e operações abstratas.

Nas imagens generativas, essa condição é levada ao limite, pois os pontos já não correspondem sequer a traços físicos ou ópticos, mas a valores probabilísticos compilados em um espaço de virtualidades. A imagem não é extraída do mundo, mas sintetizada a partir de distribuições estatísticas aprendidas, emergindo como atualização provisória de um campo latente multidimensional. Assim, sua ontologia deixa de se fundar na representação ou na semelhança, e passa a operar segundo uma lógica de amostragem, variação e otimização, em que cada imagem é ao mesmo tempo singular e estatisticamente provável. Nesse sentido, as imagens generativas confirmam e radicalizam a tese flusseriana: não apenas pertencem a um outro nível ontológico, mas instauram uma visualidade na qual a imagem não é vestígio do visível, e sim manifestação momentânea de uma lógica abstrata – uma estética programada da probabilidade.

Um outro termo que permite qualificar essas imagens é sugerido por Lev Manovich. Segundo o pesquisador, ‘previsão’ é de fato o termo técnico mais recorrente nas pesquisas que descrevem os métodos dos programas de geração de imagens. Assim, mesmo que a palavra possa soar metafórica ou evocativa, ela nomeia precisamente o processo científico que ocorre quando utilizamos essas ferramentas: “Ao trabalhar com um modelo *text-to-image*, a rede neural tenta prever as imagens que melhor correspondem à entrada de texto. Certamente não estou sugerindo que o uso de todos os outros termos já aceitos, como “mídia generativa”, seja inadequado. Mas se quisermos compreender melhor a diferença entre os métodos de síntese de mídia visual e outros métodos representacionais desenvolvidos na história da humanidade, empregar o conceito de ‘previsão’ e, assim, referir-se a esses sistemas como ‘mídia preditiva’ captura bem essa diferença.”²⁶

Nessa lógica, a imagem não é um objeto singular no sentido clássico, mas a realização provisória de uma possibilidade estatística. Cada imagem é simultaneamente inédita e provável, singular e redundante, configurando uma estética da variação e da proliferação, na qual – segundo a definição flusseriana – a informação não é o que se repete, mas aquilo que surge como situação

²⁶ (Manovich, 2023, *Op.Cit.*) Em original: “‘Prediction’ is the actual term often used by researchers in their publications describing visual generative media methods. So, while this term can be used figuratively and evocatively, this is also what actually happens scientifically when you use image generative tools. When working with a text-to-image model, the neural network attempts to predict the images that correspond best to your text input. I am certainly not suggesting that using all other already accepted terms such as ‘generative media’ is inappropriate. But if we want to better understand the difference between visual media synthesis methods and other representational methods developed in human history, employing the concept of ‘prediction’ and thus referring to these systems as ‘predictive media’ captures this difference well.”

provável no interior de um programa. Longe de uma estética da representação ou da simulação, tais imagens pertencem a uma estética generativa, cuja natureza paradoxal reside precisamente na tensão entre novidade e repetição.

Seguindo a lógica flusseriana, se a fotografia traduz o mundo em imagem via máquina, o método *text-to-image* traduz um texto em imagem também via máquina, produzindo uma imagem técnica radicalizada, pois não deriva do mundo concreto, mas de modelos estatísticos treinados sobre outras imagens técnicas. Trata-se de uma visualidade que emerge de distribuições probabilísticas aprendidas, um nível de abstração ainda mais distante do gesto humano. Nesse sentido, a operação confirma o diagnóstico de Flusser de que “o propósito de toda abstração é tomar distância do concreto para poder agarrá-lo melhor”²⁷, mas também de que tal movimento é regressivo – “um *reculer pour mieux sauter*”²⁸. A cultura, diz ele, dança em torno do concreto, mas cada novo estágio torna paradoxalmente mais difícil o retorno a ele. O *text-to-image* encarna esse ‘último estágio’ de que Flusser fala: uma concretização do absurdo da abstração, em que a imagem já não representa o mundo, mas sim o cálculo que o substitui²⁹, e, onde o aparelho produz superfícies visuais cada vez mais autônomas, enquanto o agente humano se limita a acionar o programa.

Em outras palavras, as imagens resultantes de síntese estatística de informações, que são constituídas como uma emergência probabilística, sem referencial estável ou formalizado, podem ser definidas como *imagens de amostragem estatística*. Elas não transpõem um olhar ou um gesto, mas se configuram como ‘projeções de conceitos’, dentro da análise flusseriana. Elas emergem de um processo de aprendizagem distribuído, em que a forma é obtida por operações de interpolação ou otimização num espaço latente multidimensional³⁰. Pode-se considerar, este espaço latente como a forma extrema da caixa preta, sendo assim a versão contemporânea e expandida do ‘programa’ flusseriano: não mais um conjunto explícito de regras, mas um campo estatístico opaco, inferido por treinamento acessível apenas por entradas e saídas. O resultado desta lógica operacional é uma estética da variação, da proliferação e da ambiguidade: cada imagem é ao mesmo tempo singular e estatisticamente provável, inédita e saturada de reminiscências. Longe de uma estética da representação ou da simulação, essas imagens pertencem a uma estética generativa, cuja natureza é paradoxal.

²⁷ Flusser (2012), *Op. Cit.* p. 20

²⁸ *Ibid.* p. 20.

²⁹ *Ibid.* p. 20-21.

³⁰ Isto significa que a imagem, ou o resultado gerado, surge porque o modelo percorre e combina coordenadas dentro de um espaço matemático muito amplo - o chamado espaço latente - onde cada dimensão representa alguma variação possível dos dados.

O paradoxo vem do fato de que as imagens geradas por I.As. são constituídas por elementos que predominam estatisticamente, e, portanto, são resultados de inferências matemáticas, deduzidas de conclusões lógicas a partir de premissas, dados e modelos formais. Sendo assim, para o sistema que as gera, elas são totalmente previsíveis, e qualquer originalidade é inadmissível. Por uma razão que nos escapa, para nós elas podem ser surpreendentes. Este paradoxo é definido também pelo regime de figuração inédito que elas instauram, e, que é talvez o motivo principal da curiosidade que elas despertam.

Nesse sentido, as I.As. generativas radicalizam a lógica inaugurada pela computação gráfica: elas não simulam um modelo, elas recompõem estatisticamente possíveis modelos visuais. O resultado é este regime de figuração inédito e recente (neste momento) em que o hibridismo se intensifica, a relação com o real é mais abstrata e indeterminada, talvez mesmo inexistente, – um regime no qual a imagem deriva de um processo generativo, que reconfigura continuamente o seu próprio campo visual.

As imagens geradas por I.As. são perfeitamente contempladas no pensamento de Flusser. Quando ele afirma que somos programados por imagens técnicas, Flusser antecipa precisamente o deslocamento que hoje observamos: a imagem deixa de ser uma construção explícita e passa a ser o resultado de processos automáticos que escapam à intencionalidade humana direta. Nesse sentido, as imagens realizadas por modelos generativos atualizam de modo definitivo a tese flusseriana de que vivemos submersos em superfícies codificadas: se a imagem de síntese ainda permitia rastrear o modelo, o algoritmo, a geometria, a imagem gerada por I.As. opera num nível de opacidade muito maior: ela emerge de correlações aprendidas, de amostragens estatísticas, de padrões invisíveis e evolutivos.

Para Flusser as imagens técnicas são superfícies que escondem seus processos internos, e essa formulação é extremamente atual. A imagem generativa é inacessível para a pessoa que a cria, como também para pessoa que a contempla – sua mensagem (ou verdade) se dissolve entre a amostragem, o cálculo, e a inferência. Ela também não se refere à um tempo-espço e nem dispõe de um tempo-espço próprio. Ela opera um deslocamento no campo das lógicas figurativas que foram inventadas até hoje. Esse deslocamento convoca diretamente a noção flusseriana de tecno-imaginação, que ele descreve como uma ruptura entre nós e o mundo, mediada por aparelhos que passam a imaginar por nós. A imagem realizada por I.As. intensifica essa ruptura: não se trata apenas de recompilar imagens, mas de calcular e combinar uma quantidade de dados inimagináveis, pretendendo tornar visível os mecanismos de ‘imaginação’ da caixa preta.

O argumento da caixa preta, tal como formulado por Vilém Flusser, ganha um novo vigor no caso das imagens geradas por I.As., porque a opacidade do aparelho não é meramente técnica, mas sim epistémica e ontológica. No início da sua existência, a fotografia era opaca para a maioria

dos usuários, pois as reações químicas na câmera obscura eram invisíveis, e, o aparelho codificava processos físicos e óticos que poucos conseguiam explicar na íntegra, mas a sua opacidade era limitada: embora oculta, a cadeia causal entre o mundo, a luz, a lente e a imagem era materialmente rastreável. Da mesma forma, as imagens de síntese produzidas por softwares dependem de modelos matemáticos que podem ser difíceis de interpretar para não especialistas, mas os parâmetros, algoritmos e fluxos de renderização permanecem, em princípio, inspecionáveis e reversíveis.

Em contrapartida, os sistemas de I.A. generativa introduzem uma forma qualitativamente diferente de opacidade, porque as suas operações internas não estão simplesmente ocultas, mas são fundamentalmente ininterpretáveis. Os valores numéricos de uma rede neural profunda não correspondem a categorias legíveis para o ser humano, e os dados usados para a aprendizagem – frequentemente proprietários, mas as vezes desconhecidos e em forma bruta – não podem ser reconstituídos a partir do resultado. Isto significa que a cadeia causal entre o prompt de entrada e a imagem gerada não é apenas oculta, mas *irrecuperável*, produzindo uma imagem cujas condições de possibilidade não podem ser rastreadas até um referente estável, um processo físico ou um modelo abstrato. A correspondência entre a caixa preta da fotografia e esta das imagens geradas por I.As. é rompida no nível de acesso epistémico: embora ambas envolvam processos invisíveis, a invisibilidade da química ou da ótica assentava em leis físicas determinísticas, enquanto a invisibilidade da I.A. decorre da abstração estatística, da escala e da não linearidade, às quais podemos somar sequências de cálculos evolutivos de complexidade extrema e de resultados imprevisíveis.

O que não conseguimos compreender nas imagens geradas por I.A. não é apenas ‘como a máquina funciona’, mas *o que a máquina viu, como o codificou e por que razão uma determinada configuração visual emerge* de milhares de parâmetros. Isto cria um novo tipo de caixa preta, na qual a opacidade não é uma limitação temporária do conhecimento do usuário, mas uma característica estrutural da própria tecnologia. Consequentemente, a opacidade da I.As. generativas não é apenas mais profunda, mas de natureza diferente: ela desloca a autoria, a referencialidade e a responsabilização de maneira que nem a fotografia, nem a imagem de síntese fizeram, porque o aparato generativo não pode ser totalmente inspecionado, reconstruído ou mesmo mapeado conceptualmente em termos de categorias humanas de intenção, representação ou causalidade.

Neste sentido, a questão deixa de ser apenas estética ou técnica e torna-se política, como Flusser já sugeria ao afirmar que a luta do futuro seria entre programadores e programados. Esta dimensão política adquire uma nova urgência no contexto dos sistemas generativos. Para Flusser, os aparelhos não são ferramentas neutras, mas máquinas tecno-culturais que moldam o campo de gestos possíveis, e as plataformas de I.As. radicalizam esta assimetria ao concentrar o poder nas mãos daqueles que concebem, treinam e controlam os modelos. Quando os usuários geram

imagens sem acesso aos dados de treino, à lógica interna ou aos processos de tomada de decisão do sistema, não são coautores, mas operadores de um aparelho pré-estruturado, cujas regras permanecem inacessíveis.

Hoje, perguntar “O que significa gerar imagens via I.As.?” é perguntar quem controla os programas, quem define os padrões e quem orienta a imaginação coletiva num mundo em que a visibilidade é cada vez mais automatizada. Isto resulta numa economia política das visibilidades, cuja autoria é detida pelas empresas, engenheiros e infraestruturas opacas – em vez de se dirigir para fora, em direção aos usuários. O comando do usuário torna-se uma interface superficial, uma pseudoliberalidade que mascara o facto de que o espaço generativo já está delimitado por decisões tomadas noutro lugar: quais os dados que foram extraídos, quais os preconceitos que foram codificados, quais as normas estéticas que foram estatisticamente reforçadas, quais as exclusões que foram silenciosamente reproduzidas.

Neste sentido, o ‘mestre’ da criação não é o usuário, mas o próprio aparelho, apoiado pelas instituições que o configuram. Os desafios políticos não se prendem, portanto, apenas com a representação, mas também com a governação: quem define os parâmetros das possibilidades visuais, quem controla as condições epistémicas sob as quais as imagens surgem, e quem se beneficia da opacidade que protege esses processos de amostragem estatística. Ao contrário das tecnologias fotográficas ou de imagem de síntese anteriores, em que os usuários podiam, em princípio, compreender ou manipular os mecanismos subjacentes, as I.As. generativas deslocam a agência para um domínio estruturalmente quase inacessível³¹, transformando o usuário naquilo que Flusser chamaria um ‘funcionário’ do aparelho. O resultado é um novo regime de produção de imagens em que o poder é exercido não através de coerção visível, mas através de arquiteturas computacionais invisíveis, tornando a questão política da opacidade das I.As. central nos debates contemporâneos sobre autoria, responsabilização e redistribuição dos agentes criativos.

As palavras de Flusser têm um impacto particularmente forte no contexto das imagens geradas por I.As. Originalmente concebidas como ferramentas de orientação, capazes de ajudar a explorar ideias, visualizar conceitos ou apoiar a criação, essas imagens tornam-se frequentemente telas que substituem a realidade. Em certos casos, sua aparente fidelidade visual pode dar a ilusão de um mundo coerente, quando na verdade se baseiam em estatísticas e não em experiências do mundo. Elas não representam a realidade: elas a recompõem, suavizam, normalizam, até torná-la irreconhecível. Assim, o que deveria ser um instrumento de compreensão se torna um dispositivo que molda nosso imaginário e nossa relação com o mundo, exatamente como Flusser havia escrito: “As imagens são mediações entre o homem e o mundo. (...) Imagens tem propósito de representar

³¹ Podemos admitir que um criador de imagens é capaz de desenvolver seu algoritmo e alimentar uma base de dados pessoal. Neste caso ele teria controle no processo de criação e se tornara programador de suas imagens.

o mundo. Mas ao fazê-lo, interpõem-se entre mundo e homem. Seu propósito é serem mapas do mundo, mas passam a ser biombos. O homem invés de se servir das imagens em função do mundo, passa a viver em função das imagens.”³²

Com as imagens generativas o deslizamento das imagens de *mapas* para *biombos* atinge um grau inédito de intensidade, elas se comportam mais do que nunca como filtros que reorganizam o próprio acesso ao real. O ‘biombo’ flusseriano deixa de ser apenas uma metáfora e torna-se um princípio estrutural: o usuário não vê o mundo através da imagem, mas vê o mundo através de um aparato que já filtrou, selecionou e sintetizou o visível antes mesmo que qualquer experiência sensível ocorra. As imagens geradas por I.As. não apenas desviam o olhar do mundo, elas reconstituem um mundo nunca visto ou habitado, instaurando assim um regime de visibilidade continuamente reconstruídas segundo padrões estatísticos, interesses econômicos e decisões algorítmicas invisíveis. Nesse sentido, a teoria de Flusser não apenas se aplica às I.As. generativas, mas encontra nelas sua realização mais extrema: o triunfo do biombo sobre o mapa, da programação sobre a experiência, da imagem sobre o mundo.

5. Conclusão provisória

Em termos de conclusão, retomo a ideia de Flusser segundo a qual ‘todas as imagens técnicas são metacódigos de textos’. Desse modo, cada geração de imagens técnicas redefine sua própria gramática de mediação: a fotografia codificava textos científicos sobre óptica e química; as imagens de síntese traduziam modelos matemáticos e descrições formais do espaço; contudo, as imagens geradas por I.As. introduzem uma ruptura decisiva, pois são as primeiras cujo texto gerador não é uma fórmula científica, mas a própria linguagem natural. Nesse sentido, elas representam não apenas uma continuidade, mas o ponto de saturação da tese flusseriana: o momento em que a imagem deixa de ser derivada de um saber especializado, e passa a ser produzida pela linguagem humana mediada por uma tecnologia. Provavelmente, esta relação entre o texto, que configura o prompt sem traduz o imaginário em palavras, e sua função para a geração de propostas visuais é o ponto mais importante. Nesse cenário, o mundo exterior, que outrora servia de referência para a imagem, é substituído por um regime textual que organiza e produz o visível. A imagem torna-se, assim, a realização plena do programa flusseriano: um metacódigo que já não remete ao mundo, mas ao próprio universo linguístico que o antecede.

Essa transformação abre um campo de questões que ultrapassa a técnica e alcança o domínio estético, ético e epistemológico. Em primeiro lugar, é necessário investigar como a reciprocidade e o intercâmbio entre linguagem e imaginário são operacionalizados pelos prompts;

³² Flusser (2002), *Op. Cit.*, p. 7.

como se reconfigura a relação entre texto e imagem no processo de criação, e como a disparidade entre os dois sistemas simbólicos interfere na criação visual. Em segundo lugar, torna-se crucial distinguir a interpretação humana do prompt – carregada de intenção, contexto e subjetividade – da leitura computacional, fundada em correlações estatísticas e regras formais, e compreender como essa assimetria afeta os processos de significação. Em terceiro lugar, é preciso investigar as consequências de um cenário em que o artista passa a orientar a máquina por meio de instruções, ao mesmo tempo em que o resultado tende a exceder e a deslocar sua própria imaginação, instaurando uma economia da autoria.

Em última instância, com as imagens generativas a caixa preta flusseriana torna-se uma verdadeira caixa de Pandora, que liberou forças que excedem o controle humano e propicia uma irrupção irreversível do inesperado. Estamos perante uma etapa tecnológica, cujos efeitos ainda devem ser decifrados. A sensação de que abrimos uma nova caixa de Pandora é reforçada também pela sensação de que enfrentamos tecnologias cujas possibilidades de criação de imagens são inesgotáveis, fato que provoca também ansiedades profundas sobre autoria, veracidade, poder e controle. Como no mito, não se trata apenas de forças inesperadas que escapam ao domínio humano, mas de reconhecer que cada inovação tecnológica traz consigo tanto expansões do visível, quanto novas zonas de opacidade. As imagens técnicas já haviam deslocado o humano para dentro de um jogo programado; as imagens generativas radicalizam esse deslocamento, multiplicando incertezas e potências que escapam à supervisão direta. E, no entanto, permanece – como no fundo da caixa da Pandora – um sopro de esperança: a possibilidade de reinventar nossa relação com os modelos generativos, de personalizá-los, de submetê-los a finalidades humanas, e de reinscrever a imaginação dentro de um horizonte em que cálculo e criação não sejam forças opostas, mas tensões produtivas continuamente negociadas e produtoras de experimentações críticas.

Referências

- Couchot, Ed. (1999) «Da representação à simulação» in *Imagem Máquina*, André Parente (org.), Rio de Janeiro : Éd. 34.
- Flusser, V. (2002) *A filosofia da caixa preta: ensaios para uma futura filosofia da fotografia*, Rio de Janeiro: Ed. Relume Dumará, Coleção Conexões.
- Flusser, V. (2012) *O Universo das imagens técnicas: elogio a superficialidade*, Coimbra: Anablume.
- Flusser, V. (2022). *Mutations dans les Relations Humaines : De la Communicologie*, Ed. HD/Le Bon Voisin/AKA.
- Lévy, P. (1999), *Cibercultura*, (Trad. de Carlos Irineu da Costa) São Paulo: Ed. 34, 1999.

- Manovich, L. (2023) : AI Image Media through the Lens of Art and Media History. In: IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft. Generative Imagery: Towards a 'New Paradigm' of Machine Learning-Based Image Production, Jg. 19 (2023), Nr. 1, S. 34–41. DOI: <https://doi.org/10.25969/mediarep/22323>.
- Quéau, Ph. (1999). « O tempo do virtual » in Imagem Maquina, André Parente (org.), Rio de Janeiro : Éd. 34.