

## Vigilância e Viés sobre a Imagem: Visão Computacional sob Perspectivas Críticas

*Surveillance and Bias on Images: Computer Vision  
under Critical Perspectives*

Sergio José Venancio Júnior<sup>1</sup>

### Resumo

Este trabalho propõe uma análise sobre a imagem digital nos contextos ubíquos da Inteligência Artificial, Aprendizado de Máquina e Visão Computacional. Tal situação submete a imagem a redes e regimes de vigilância automatizada, que julgam a informação visual conforme critérios e interesses ocultos. São analisados percursos e processamentos de imagens na nuvem com base em estudos técnicos e críticos, que revelam um acúmulo de vieses estruturais e precarização de trabalho humano que reduzem a imagem a rótulos restritivos e parciais. Uma vez que sistemas de automação de julgamentos sobre imagens se difundem amplamente, tal situação acaba por amplificar preconceitos, desigualdades e injustiças sociais.

**Palavras-chave:** inteligência artificial, aprendizado de máquina, visão computacional, vigilância, datasets de imagem.

### Abstract

*The present study proposes an analysis of the digital image in the ubiquitous contexts of Artificial Intelligence, Machine Learning and Computer Vision. Such situation submits the image to networks of automated surveillance, which judge visual information according to hidden criteria and interests. Image processing and distribution are analysed based on technical and critical studies, revealing an accumulation of structural biases and precarisation of human labour that reduce images to restrictive and non neutral labels. Given that systems of automated judgments over images become widespread, such a situation amplifies prejudices, inequalities and social injustices.*

**Keywords:** artificial intelligence, machine learning, computer vision, surveillance, image datasets.

### Introdução

Em seus “ensaios para uma futura filosofia da fotografia” em 1983, Vilém Flusser (2009) reforça a ideia de aparelho fotográfico, cujo programa condiciona a maneira que o fotógrafo age para produzir, manipular e manifestar as ditas imagens técnicas. Já ciente dos amplos

---

<sup>1</sup> Sergio José Venancio Júnior é doutorando pelo Programa de Pós-Graduação em Artes Visuais (PPGAV) ECA USP, bolsista CAPES Proex nº 88887.488639/2020-00. Mestre em Artes Visuais pelo mesmo programa. Graduado em Artes Visuais e em Ciência da Computação pela UNICAMP. Professor nos cursos de Especialização em Design Gráfico UNICAMP, e Pós-Graduação em Arquitetura Digital e Projetos Paramétricos na Belas Artes SP. Participa do grupo de pesquisa Realidades (ECA USP), coordenado pela prof<sup>a</sup> dr<sup>a</sup> Sílvia Laurentiz.

avanços das tecnologias digitais, o autor indicava um futuro desta imagem submetido aos programas determinados não apenas por hardware e software, mas também pelos demais aparelhos que definem tais programas e seus metaprogramas: uma estrutura hierárquica e dispersa de aparelhos políticos, econômicos, culturais – uma estrutura de “caixas pretas”. Quase 40 anos depois, tais ideias seguem cabíveis e a imagem técnica tem seu estatuto condicionado ao programa ubíquo chamado Inteligência Artificial (IA).

Hoje é plausível dizer que toda imagem digital foi ou será submetida a processos de Visão Computacional e Aprendizado de Máquina (*Machine Learning*). Uma simples demonstração desse efeito se dá no fato que a grande maioria das câmeras fotográficas comercializadas atualmente integram modelos de detecção facial em seus sistemas de fotometria que, por exemplo, automatizam o ajuste de foco com base no rosto de uma pessoa. O movimento supostamente cognitivo de uma câmera digital ao identificar um rosto humano é, na verdade, uma operação de detecção de padrões estatísticos sobre as leituras numéricas de seu sensor fotoelétrico. Tais padrões foram previamente estabelecidos a partir do Aprendizado de Máquina sobre diversos exemplos de imagens com rostos, e o modelo de inferência resultante é integrado ao software da câmera, este que impõe ao fotógrafo uma compreensão da imagem carregada de ajustes prévios. Se o aparelho fotográfico de Flusser é um jogo e o fotógrafo é *homo ludens* (Flusser, 2009, p. 24), automações da visão e das regulagens da câmera removem algumas possibilidades de jogadas contra seu programa, enquanto o fotógrafo é condicionado a se distanciar do objeto fotografado e a confiar mais em um dispositivo que toma decisões por ele – um jogo mais complexo e desafiador. A quem interessaria este jogo? E quem define como serão estas decisões prévias sobre o funcionamento do aparelho?

O objetivo geral da Inteligência Artificial será primeiramente definido como automação da cognição em larga escala, ou uma “supercognição” (Manovich, 2018) em rede. Paralela às narrativas ficcionais dos robôs sencientes que ainda dominam imaginários populares, a IA evoluiu a ponto de ser o mais poderoso meio para se extrair informação útil do Big Data, essa imensurável massa de dados digitais decorrente da popularização da internet e do uso cotidiano de computadores e *smartphones* para quantificar a vida. A escala é um aspecto chave: tornou-se humanamente impossível lidar com grandes quantidades de informação coletadas através destes dispositivos. Portanto qualquer movimento curatorial sobre dados,

desde seleções, combinações, classificações e filtragens, requer o uso de técnicas da IA em rede e computação distribuída. A capacidade de extração de informações úteis a partir dos numerosos dados digitais envolve poucas e grandes empresas multinacionais, detentoras de modelos de Aprendizado de Máquina de operação planetária, como Google, Facebook, Baidu e Tencent. Tais empresas oferecem serviços gratuitos como e-mail, redes sociais e aplicativos de conversação, exigindo em troca o direito de coletar informações de uso e do usuário, e oferecer a ele propagandas sob medida, de acordo com sua personalidade, gostos e hábitos, inferidos a partir de modelos de IA treinados sobre os dados coletados. Tal estratégia é pilar central do chamado Capitalismo de Vigilância, que “define unilateralmente que a experiência humana constitui material bruto livre para tradução em dados comportamentais” (Zuboff, 2019, n.p.). Também é possível falarmos em Estado de Vigilância (Docherty; Fanning, 2019), onde governos inferem informações sobre os cidadãos e os classificam de acordo com seus interesses de controle. Dados são a nova moeda de troca nas relações econômicas e geopolíticas, e as instituições que mais os acumularem terão vantagem, pois é da natureza do Aprendizado de Máquina se aperfeiçoar, quanto mais dados estiverem disponíveis e mais representativos eles forem. Pode-se dizer que o grande propósito que sustenta o desenvolvimento da IA é a vigilância, e seu melhor vigilante é o *smartphone*.

Estima-se que, desde 2016, mais de um trilhão de fotos digitais são feitas por ano (Cakebread, 2017). Trata-se de uma fonte inesgotável de dados, pois de uma imagem digital podem ser extraídas informações que identificam objetos, cenários, pessoas, locais. Qualquer imagem publicada em uma rede social passa por análises de Visão Computacional, que buscam identificar elementos concretos e abstratos que aperfeiçoem modelos comerciais de propaganda. Se alguém publica uma foto com amigos em um show de rock, é extremamente provável que esta pessoa e seus amigos sejam automaticamente identificados e passem a receber publicidade de outros shows nesta mesma rede social. Ou se esta mesma pessoa publicar outra foto participando de uma manifestação contrária a um governo, também é extremamente provável que esta pessoa seja automaticamente identificada e classificada como opositora por sistemas de vigilância de câmeras em vias públicas. Esta automação da inferência visual, capaz de processar informações do indivíduo e julgá-lo conforme critérios ocultos, adicionam uma camada importantíssima para a lógica da vigilância, mas extremamente preocupante para a sociedade. A automação do julgamento através de imagens, considerando-se as limitações da própria IA, configura ambiente propício a danos.

Há diversos aspectos invisíveis sobre as imagens digitais processadas pela IA, e este trabalho propõe um aprofundamento em algumas questões sobre este tipo de imagem e suas operações. Inicialmente serão analisados os percursos dos dados gerados a partir das interações humano-máquina em contextos de IA e Aprendizado de Máquina, com recorte sobre a imagem digital e o funcionamento da Visão Computacional. Serão também apresentadas diversas formas de viés incorporadas na automação dos julgamentos visuais, que em última instância condicionam comportamentos humanos diante de tais aparatos. A discussão busca um discurso questionador das forças políticas que disputam o lugar da imagem no contemporâneo.

### **Bastidores da Inteligência Artificial**

A partir de uma simples interação com um dispositivo provido de sistema de Inteligência Artificial como um *smartphone*, toda uma cadeia planetária de outros dispositivos é acionada para que os dados envolvidos nesta interação trafeguem, até serem processados pelos modelos comerciais de *Deep Learning*. Tal técnica de processamento em rede se torna, na última década, uma das mais bem sucedidas em Aprendizado de Máquina, caracterizada pela grande escala do volume de dados que é capaz de processar e pela alta complexidade de suas operações. No entanto, a infraestrutura necessária para desenvolver e processar um modelo comercial de *Deep Learning* foge de qualquer escala doméstica. O infográfico *Anatomy of an AI System*, de Kate Crawford e Vladan Joler (2018), demonstra o caso das redes intercontinentais que viabilizam a chamada “nuvem”, com seus equipamentos, conexões materiais, teorias científicas e todo o trabalho humano necessário para viabilizar uma simples interação entre um consumidor e o *Amazon Echo*, um *smart speaker* doméstico que executa ações a partir de comandos de voz em tom imperativo como “ligue as luzes do quarto”, “telefone para fulano” ou “toque música clássica”. Tal dispositivo conta com o sistema assistente de voz *Alexa* que, assim como os sistemas *Siri*, *Cortana* e *Google Assistant*, captura continuamente os sons de um ambiente e reconhece a voz de seu usuário. A partir de uma captura de voz, o mapa evidencia toda a infraestrutura necessária para viabilizar seu percurso, bem como inúmeras relações de exploração ambiental e de trabalho humano envolvidas. São reveladas dimensões ocultas, típicas da lógica de terceirizações presentes nas cadeias de produção e descarte de dispositivos como *smartphones*, roteadores, repetidores de sinal, cabos, antenas, servidores, placas, circuitos, chips e o próprio *Echo*.

O mapa chama atenção para as relações exclusivas de trabalho humano inauguradas pelo uso comercial das tecnologias de Inteligência Artificial, especialmente de *Deep Learning*. Quando um trecho de voz do consumidor chega aos *datacenters* da Amazon, ele é avaliado por um modelo de *Deep Learning* treinado com milhões de exemplares de voz humana, classificados conforme o comando que expressam. Assim que o modelo consegue reconhecer o comando contido no trecho de voz, um código inequívoco é enviado de volta ao *Amazon Echo*, para que a tarefa solicitada seja executada. Mas quando não consegue reconhecer o comando, o aparelho entra em modo de treinamento e interage com o consumidor, até que descubra o que ele quis dizer. Neste momento todos os exemplares de voz trocados serão utilizados para aperfeiçoar o modelo de maneira generalizada. Este tipo de relação do consumidor com seu dispositivo de IA é considerada por Crawford e Joler como um “trabalho não reconhecido”, onde os dados de voz coletados são utilizados para aperfeiçoamento de um modelo privado. O usuário de um dispositivo de IA, além de pagar pelo dispositivo, se torna “simultaneamente um consumidor, um recurso, um trabalhador e um produto” (Crawford; Joler, 2018) do sistema, seu conhecimento e experiência se transformam em dados para o capitalismo de vigilância.

Ainda sobre o processo de treinamento de um modelo de *Deep Learning*, temos as redes neurais, que são estruturas de dados que mapeiam uma entrada, como uma imagem, som ou texto, a uma saída na forma de rótulo classificatório<sup>2</sup>. Um processo de treinamento deste tipo consiste em apresentar à esta rede diversas entradas e suas respectivas saídas, constituindo os chamados “*datasets* de treinamento”. Para se aperfeiçoar, uma rede neural típica de *Deep Learning* conta com muitas camadas de profundidade, e cada camada possui muitos nós. Um nó propaga ou não a informação recebida da combinação de todos os nós de uma camada anterior. Durante o treinamento, a rede consome o *dataset* de treinamento e ajusta progressivamente as permissividades de cada um de seus nós, para que as informações de cada entrada sejam corretamente propagadas por sua estrutura de nós e camadas até cada saída correspondente. Após o treinamento, a rede receberá novos estímulos e os mapeará para as saídas que mais se aproximarem de alguma correspondência específica do *dataset*

---

<sup>2</sup> A este tipo de treinamento que consiste em mapear entradas diversas a rótulos classificatórios damos o nome de “aprendizado supervisionado”. Há outros tipos de aprendizado na área de Aprendizado de Máquina (Goodfellow; Bengio, 2016).

de treinamento. Caso a resposta não apresente alta acurácia, a rede pode ser retreinada incluindo os novos estímulos, contanto que sejam rotulados corretamente.

Assim também funciona o modelo de *Deep Learning* da Amazon para reconhecer comandos de voz: cada exemplar digitalizado de voz humana fornecido é mapeado a um comando, formando e ampliando um *dataset* de treinamento imenso, dadas as inúmeras possibilidades de expressão vocal e a variedade de comandos. Tudo isso reforça a noção de grande escala que é natural deste tipo de objeto computacional, e justifica o emprego de uma infraestrutura gigantesca de datacenters e equipamentos de processamento paralelo, necessários para lidar com um modelo comercial como o da Amazon. É importante perceber que, antes de se tornar comercializável, o modelo de *Deep Learning* passa por intenso processo de construção de um *dataset* de treinamento em que, invariavelmente, se emprega trabalho cognitivo humano nas tarefas de rotulação dos dados de entrada. Eis um dos pontos mais controversos do *Deep Learning* – o labor dos chamados “*ghost workers*” (Gray; Suri, 2019) ou “*crowdworkers*”, realizados através de plataformas de terceirização de trabalho como a *Amazon Mechanical Turk*.

O trabalho de um *ghost worker* é repleto de tarefas simples como identificar objetos em fotos ou transcrever voz em texto, sendo extremamente repetitivas. Quando este tipo de trabalho é reconhecido, ele é mal remunerado (Crawford; Joler, 2018). Tais atividades são delegadas a milhares de trabalhadores informais que disputam tarefas na plataforma em troca de reputação e baixos pagamentos, quando não em troca de *vouchers* no próprio sistema de comércio eletrônico da Amazon. Alguns autores (Grohmann, 2020) (Moreschi; Pereira; Cozman, 2020) (Casilli, 2019) consideram essa nova forma de trabalho como um indício de que o trabalho humano, no futuro, não será substituído pelas máquinas, pois hoje ele é apenas deslocado para relações de maior exploração e precarização.

Quando a ação de classificar milhões de registros de dados com seus respectivos rótulos é delegada a milhares de trabalhadores informais, começamos a responder uma pergunta feita anteriormente: quem decide sobre as definições prévias de um modelo de IA que condiciona a operação de aparelhos como o fotográfico? A responsabilidade das grandes empresas sobre a forma de julgamento de seus modelos é diluída entre milhões de ações de rotulação executadas por trabalhadores invisíveis. Para além do problema da exploração humana,



como garantir que o julgamento de todos estes trabalhadores seja uniforme, justo e ético? Temos a partir deste ponto um imenso problema de viés (inclinações/predisposições) em IA.

### Heranças e Amplificações de Viés

Um dos grandes desafios impostos aos detentores de grandes modelos de IA e de grandes *datasets* de treinamento diz respeito aos vieses que estes sistemas carregam e propagam. É bastante comum que um modelo de IA comercial cometa erros em suas inferências, detectando padrões sem sentido ou inexistentes. Isso ocorre muitas vezes porque o modelo pode não ter sido treinado suficientemente (*underfit*) ou por ter se especializado tanto a ponto de não ser mais capaz de generalizar ou relativizar (*overfit*). Mas há momentos em que os erros causam danos, como os inúmeros casos de discriminação racial ou de gênero em julgamentos por sistemas de Visão Computacional das grandes corporações, como demonstrado pelo estudo *Gender Shades* (Buolamwini; Gebru, 2018). Recentemente houveram casos como um falso positivo ter levado um cidadão negro norteamericano a ser preso injustamente (Hill, 2020), acusado de ter roubado uma loja de departamento após ter sido identificado erroneamente pelo sistema de reconhecimento facial da polícia de Detroit e suas câmeras de vigilância. O viés no julgamento de máquina é dos mais sérios problemas nestes contextos de segurança. Mas uma vez que entendemos como se dá a participação humana no processo de definição daquilo que um modelo de IA aprende, começamos a rastrear as origens destas inclinações e predisposições do julgamento de máquina.

Dois estudos propõem identificações da origem do viés em IA: *Excavating AI*, de Kate Crawford e Trevor Paglen (2019), e o *Nooscope*, de Vladan Joler e Matteo Pasquinelli (2020). *Nooscope* é um mapa que representa uma metáfora para definir o Aprendizado de Máquina como um instrumento ótico, utilizado para se enxergar e navegar pelo espaço do conhecimento. O estudo apresenta tal instrumento propondo que "entender o Aprendizado de Máquina e registrar seu impacto na sociedade é estudar o grau pelo qual dados sociais são difratados e distorcidos por estas lentes" (Pasquinelli; Joler, 2020, n.p.)

Segundo os autores, todo *dataset* de treinamento é uma "construção cultural", pois é construído a partir da quantificação da natureza – uma ideia que se alinha ao preceito do

capitalismo de vigilância, de que toda experiência humana pode ser traduzida a dados sociais. Quantifica-se o que um ser humano é, seu corpo, seu comportamento, e o que ele produz. As difrações e distorções que o *Nooscope* identifica são os vieses da IA, enquanto tal instrumento vem sendo considerado como um novo regime da verdade científica, de normatividade e racionalidade. Não se trata de uma nova metáfora: toda ciência carrega vieses, validados pela própria comunidade científica, que distorcem e difratam realidades. O diferencial fica novamente por conta da escala em que este instrumento opera, que já sabemos ser humanamente impossível de lidar. Trata-se de uma cadeia de filtragem de dados tão grande e complexa, que fica impossível explicar exatamente como um modelo de IA chega em determinado resultado, abrindo um precedente controverso para que empresas detentoras de modelos não se responsabilizem totalmente pelos julgamentos emitidos por suas tecnologias.

O *Nooscope* caracteriza o viés do Aprendizado de Máquina em três tipos:

- o viés histórico, que corresponde à vieses institucionalizados. Este tipo de viés determina muitas vezes que o material de aprendizado de uma IA inevitavelmente propagará as falhas daquilo que representa, como foi o caso do viés racial em um aplicativo que transformava *selfies* em pinturas clássicas da história da arte ocidental (Declercq, 2020). A falta de retratos de pessoas negras na própria história da arte, representada neste caso pelas coleções dos grandes museus, fez com que *selfies* de pessoas negras fossem convertidas em retratos de pessoas brancas;
- o viés de *dataset*, introduzido durante a rotulação dos dados pelos humanos, e aqui recordamos o problema do trabalho informal em plataformas como a *Amazon Mechanical Turk*, bem como a questão do controle sobre os julgamentos por parte dos trabalhadores terceirizados durante as tarefas de classificação. Nesta cadeia de precarização do trabalho, são estes trabalhadores que classificariam imagens com rótulos de raça e gênero;
- e o viés algorítmico, que neste caso corresponde à escolha e emprego de um paradigma estatístico na lógica dos modelos de inferência. O algoritmo de aprendizado por redes neurais se caracteriza por responder em termos de probabilidades. Por exemplo, ao lidar com uma foto de um rosto, um modelo de rede neural poderia classificar com 60% de chances desse rosto ser “branco”, 35% de ser “negro”, 5%



“amarelo”, uma forma improvável de caracterização racial. Mas o modelo sempre precisa oferecer uma decisão, aquela que tiver o mais alto nível de confiança. Por conta dessa natureza excludente entre categorias e seus níveis de confiança, o modelo pode errar ou exagerar na classificação de algo. Outra questão natural deste tipo de algoritmo é a compressão da informação, pois o processo de Aprendizado de Máquina lida melhor com dados livres de ruídos, limpos e padronizados. Toda compressão traz perdas em relação à coleta original de dados.

Dados são difratados e distorcidos cumulativamente por estas lentes de um instrumento ótico que acumula predisposições, preconceitos e reduções, que não acolhe toda a diversidade do mundo, da vida. Portanto é necessário continuar os questionamentos: quem define a escolha deste tipo de instrumento para julgar o indivíduo e a sociedade? Ou então, quem regula tal escolha, qual legislação existe? Estas perguntas de cunho político levam ao outro estudo, *Excavating AI*.

Olhando especificamente para *datasets* de imagem, o estudo de Crawford e Paglen foca nas políticas públicas e privadas que definem modelos de classificação de imagens de humanos através de Visão Computacional, e oferece exemplos de *datasets* famosos, como o *The Japanese Female Facial Expression (JAFFE) Database* (Lyons; Akamatsu; Kamachi; Gyoba, 1998). Tal conjunto de imagens pressupõe que emoções humanas são quantificáveis e classificáveis a partir de simples imagens digitais de rostos, correspondendo a expressões faciais femininas japonesas. As classes utilizadas eram apenas sete: felicidade, surpresa, tristeza, medo, nojo, raiva e um estado neutro. Qualquer uso deste *dataset* implicaria em assumir uma relação fixa entre expressão facial e emoção, assumir uma métrica a ser aplicada a qualquer mulher japonesa – tais considerações são bastante questionáveis do ponto de vista da própria experiência humana e sua diversidade.

Outro *dataset* famoso que *Excavating AI* discute é o *ImageNet* (Stanford Vision Lab, 2021), um projeto ambicioso concebido por pesquisadores das universidades norteamericanas de Stanford, Princeton e Carolina do Norte, cujo objetivo seria mapear visualmente tudo o que existe no mundo e ser o maior *dataset* público de imagens. Atualmente com mais de 14 milhões de imagens classificadas pelo trabalho de milhares de trabalhadores da *Amazon*

*Mechanical Turk*, o *ImageNet* toma como taxonomia o *WordNet* (Princeton University, 2021), uma grande base de termos e expressões para classificações das coisas do mundo, desenvolvida pela universidade de Princeton.

O problema do *ImageNet* reside no uso contraditório de uma lógica descritiva para classificar conceitos abstratos. O estudo de Crawford e Paglen cita uma ideia de gradientes de concretude-abstração de George Lakoff (2012), onde o conceito de "maçã" seria mais substantivo, mais concreto, do que o conceito de "luz", e "luz" por sua vez seria mais substantivo que "saúde", considerado mais abstrato. Representar uma maçã em termos de imagem implica relações visuais diretas – uma maçã possui forma, cor e textura características – mas o conceito abstrato de saúde dificilmente corresponderia a imagens específicas ou padrões visuais. O resultado do emprego desse esquema de classificação sobre imagens pode ser ilógico e cruel, especialmente em casos de rótulos críticos, como raça e gênero. Pode uma imagem por si só dizer sobre a raça de uma pessoa? Gênero? Emoções? Torna-se anticientífica a pretensão de reduzir identidades e experiências humanas a imagens ou termos inequívocos, portanto uma Inteligência Artificial não deveria ser capaz de generalizar conceitos apenas a partir do aprendizado por *datasets* de imagens, por maiores e mais diversos que sejam.

O *ImageNet* possuía uma grande categoria chamada "Pessoas", com associações extremamente controversas entre atributos pejorativos presentes no *WordNet* e imagens de humanos, fruto das rotulações incontroladas por parte dos trabalhadores invisíveis. Muitos modelos de *Deep Learning* foram treinados a partir do *ImageNet*, mas reproduziram seus preconceitos, e o *ImageNet* se viu obrigado a desativar esta categoria e suas respectivas imagens. Recentemente, Fei Fei Li, uma das idealizadoras do *ImageNet*, considerou recomeçar o projeto (Henpel, 2018).

Independente da natureza do *dataset*, a escala do *Deep Learning* força o uso massivo de trabalho remoto humano, o que facilita situações em que não se garante consistência de julgamentos e rotulações, especialmente nos casos que lidam com classificações abstratas e simbólicas, menos concretas. Validar rotulações feitas por tantas pessoas seria outro trabalho hercúleo, o que torna a auditoria de modelos de *Deep Learning* praticamente impossível.

## Considerações Finais

Os grandes sistemas de Inteligência Artificial das empresas de tecnologias e dos governos estão operando com nossos dados de formas enviesadas, com normatizações que não condizem com a realidade da diversidade do mundo. A representação para os grandes modelos do Aprendizado de Máquina é uma questão estatística, reducionista, quantitativa, ao invés de qualitativa. E ainda, tais modelos não podem ser validados ou auditados, pois seus métodos de treinamento em escala diluem responsabilidades entre milhares, milhões de trabalhadores precarizados. Estes modelos estão sendo usados para promover ou discriminar, aprovar ou desaprovar, tornar visível ou invisível. Os regimes de vigilância, através do julgamento de máquina, transpõem limites éticos e de regulamentação política.

As imagens que passam por processos de IA estão subordinadas à ideia de que existem correspondências fixas entre imagens e conceitos, de que existem vínculos quantitativamente mensuráveis entre objeto, signo e interpretante. Embora seja plausível este tipo de relação, o problema se dá quando ele se torna norma e se institucionaliza, se transforma em política pública ou privada que vigia e julga. Dados os problemas de viés, impor o uso de tais sistemas estabelece uma espécie de eugenia visual, onde o mundo passa a ser condicionado por aproximação estatística a um conjunto limitado, ainda que numeroso, de imagens e seus rótulos. Vemos isso hoje claramente em aplicativos de redes sociais com seus populares filtros de *selfies*, normatizando faces, apagando identidades.

Um objetivo comum entre *Anatomy of AI*, *Nooscope* e *Excavating AI* é desmistificar a ideia de que a tecnologia da IA seria neutra. Toda tecnologia é definida por ideologias e projetos de poder que afastam e subjagam o utilizador comum, pois a ele cabe cada vez mais o papel de funcionário do aparelho e menos de jogador, aquele que talvez seria capaz de esgotar seu programa. No que se compreende pela Inteligência Artificial hoje, a escala impede o lidar humano, este é tarefa acessível apenas aos financeiramente poderosos e suas máquinas distribuídas por uma nuvem que nada tem de leveza – ela pesa na forma de crises humanitárias e ambientais, como nos mostra o infográfico de Crawford e Joler.

Hoje o propósito da imagem digital é ser pretexto para a vigilância. É preciso mudar esse estatuto e reivindicar o controle daquilo que se infere sobre a imagem, pois é privado e carrega vieses que não correspondem às dimensões da verdade. É preciso literacia digital, é preciso

entender a lógica dos dispositivos eletrônicos e inteligências artificiais que fazem parte do cotidiano. Abrir “caixas pretas” é sempre um movimento importante na conscientização sobre o complexo progresso da humanidade e suas tecnologias.

### Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001. Agradecemos a esta instituição e ao Programa de Pós-Graduação em Artes Visuais da Escola de Comunicações e Artes (ECA) da Universidade de São Paulo (USP).

### Referências

Buolamwini, J; Gebru, T. (2018) Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification. **Proceedings of Machine Learning Research**, v.81, p.77-91. Recuperado 29 mai. 2021 em <http://proceedings.mlr.press/v81/buolamwini18a/buolamwini18a.pdf>

Cakebread, C. (2017) People will take 1.2 trillion digital photos this year — thanks to smartphones. **Business Insider**. 31 ago 2017. Recuperado 29 mai. 2021 em <https://www.businessinsider.com/12-trillion-photos-to-be-taken-in-2017-thanks-to-smartphones-chart-2017-8>

Casilli, A. A. (2019) **En Attendant les Robots**: enquête sur le travail du clic. Paris: Seuil.

Crawford, K.; Joler, V. (2018) **Anatomy of an AI System**: The Amazon Echo as an Anatomical Map of Human Labor, Data and Planetary Resources. Recuperado 29 mai. 2021 em <https://anatomyof.ai>

Crawford, K.; Paglen, T. (2019) **Excavating AI**: The Politics of Training Sets for Machine Learning. Recuperado 29 mai. 2021 em <https://excavating.ai>

Declercq, M. (2020) Inteligência Artificial que 'pinta retratos' revela viés racial. **UOL Tab**. 15 abril 2020. Recuperado 29 mai. 2021 em <https://tab.uol.com.br/noticias/redacao/2020/04/15/inteligencia-artificial-que-pinta-retratos-releva-vies-racial.htm>

Docherty, N.; Fanning, D. (2019) **In the Age of AI**. Estados Unidos. Recuperado 29 mai. 2021 em [https://www.youtube.com/watch?v=5dZ\\_lvDgevk](https://www.youtube.com/watch?v=5dZ_lvDgevk)

Flusser, V. (2009) **Filosofia da Caixa Preta**: Ensaios para uma futura filosofia da fotografia. Rio De Janeiro: Sinergia Relume Dumará.

Goodfellow, I.; Bengio, Y. (2016) **Deep Learning**. Massachusetts: The MIT Press.

Gray, M.; Suri, S. (2019) **Ghost Work**: how to Stop Silicon Valley from Building a New Global Underclass. Boston: Houghton Mifflin Harcourt.

Grohmann, R. (2020) Plataformização do trabalho: entre a dataficação, a financeirização e a racionalidade neoliberal. **Revista Eptic**, v. 22, n. 1. Recuperado 29 mai. 2021 em <https://seer.ufs.br/index.php/epitic/article/view/12188>

Henpel, J. (2018) Fei Fei Li's Quest to Make AI Better for Humanity. **Wired**. 13 novembro 2018. Recuperado 29 mai. 2021 em <https://www.wired.com/story/fei-fei-li-artificial-intelligence-humanity/>

Hill, K. (2020) Wrongfully Accused by an Algorithm. **New York Times**. 24 junho 2020. Recuperado 29 mai. 2021 em <https://www.nytimes.com/2020/06/24/technology/facial-recognition-arrest.html>

Lakoff, G. (2012) **Women, Fire, and Dangerous Things**: What Categories Reveal about the Mind. Chicago: University of Chicago Press.

Lyons, M.; Akamatsu, S.; Kamachi, M.; Gyoba, J. (1998) Coding Facial Expressions with Gabor Wavelets. **Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition**, p. 200-205. Recuperado 29 mai. 2021 em <https://doi.org/10.1109/AFGR.1998.670949>>. Acesso em 09 ago 2020.

Manovich, L. (2018) **AI Aesthetics**. Moscou: Strelka Press.

Moreschi, B.; Pereira, G.; Cozman, F. G. (2020) The Brazilian Workers in Amazon Mechanical Turk: Dreams and realities of ghost workers. **Contracampo**, v. 39, nº 1. Recuperado 29 mai. 2021 em <https://periodicos.uff.br/contracampo/article/view/38252>

# PANORAMAS

## 2021

15 AL 18 JUNIO  
UNIVERSITAT POLITÈCNICA  
DE VALÈNCIA/ESPAÑA

VIII simposio internacional  
de innovación en medios  
interactivos

#20. ART

8vo. Balance-Unbalance  
arte+ciência x tecnologia =  
medioambiente/  
responsabilidad

Pasquinelli, M.; Joler, V. (2020) **The Nooscope Manifested**: Artificial Intelligence as Instrument of Knowledge Extractivism. Recuperado 29 mai. 2021 em <http://nooscope.ai>

Princeton University. (2021) **WordNet**. Recuperado 29 mai. 2021 em <https://wordnet.princeton.edu/>

Stanford Vision Lab. (2021) **ImageNet**. Recuperado 29 mai. 2021 em <http://image-net.org>

Zuboff, S. (2019) **The Age of Surveillance Capitalism**: The Fight for a Human Future at the New Frontier of Power. Nova Iorque: Public Affairs Books.