

UNIVERSIDADE FEDERAL DE MINAS GERAIS
Faculdade de Filosofia e Ciências Humanas
Programa de Pós-Graduação em Comunicação Social

André Goes Mintz

VISUALIDADES COMPUTACIONAIS E A IMAGEM-REDE:

reapropriações do aprendizado de máquina para
o estudo de imagens em plataformas *online*

Belo Horizonte
2019

André Goes Mintz

VISUALIDADES COMPUTACIONAIS E A IMAGEM-REDE:

reapropriações do aprendizado de máquina para
o estudo de imagens em plataformas *online*

Versão final

Tese apresentada ao Programa de Pós-Graduação em Comunicação Social da Universidade Federal de Minas Gerais, como requisito parcial para obtenção do título de Doutor em Comunicação Social.

Área de concentração: Comunicação e Sociabilidade Contemporânea.

Linha de pesquisa: Textualidades midiáticas.

Orientador: Prof. Dr. Carlos Frederico de Brito d'Andréa

Belo Horizonte
2019

301.16
M667v
2019

Mintz, André Goes.

Visualidades computacionais e a imagem-rede [manuscrito]:
reapropriações do aprendizado de máquina para o estudo de
imagens em plataformas online / André Goes Mintz. - 2019.
316 f.

Orientador: Carlos Frederico de Brito d Andréa.

Tese (doutorado) – Universidade Federal de Minas Gerais.
Faculdade de Filosofia e Ciências Humanas.

Inclui bibliografia.

1. Comunicação – Teses. 2. Tecnologia – Teses. 3. Máquinas –
Teses. I D'Andréa, Carlos Frederico de Brito. II. Universidade
Federal de Minas Gerais. Faculdade de Filosofia e Ciências
Humanas. III. Título.

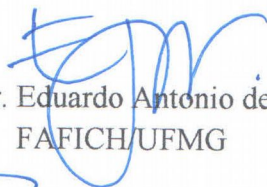
VISUALIDADES COMPUTACIONAIS E A IMAGEM-REDE:
reapropriações do aprendizado de máquina para o estudo de imagens em
plataformas *online*

André Goes Mintz

Tese aprovada pela banca examinadora constituída pelos Professores:



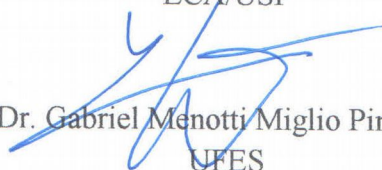
Prof. Dr. Carlos Frederico de Brito D Andrea
Orientador - FAFICH/UFMG



Prof. Dr. Eduardo Antonio de Jesus
FAFICH/UFMG



Profa. Dra. Patrícia Moran Fernandes
ECA/USP



Prof. Dr. Gabriel Menotti Miglio Pinto Gonring
UFES



Prof. Dr. Carlos Henrique Rezende Falci
EBA/UFMG

Programa de Pós-graduação em Comunicação Social
Faculdade de Filosofia e Ciências Humanas
Universidade Federal de Minas Gerais
Belo Horizonte, 22 de novembro de 2019

para naina

Agradecimentos

Tive muita sorte de contar com o privilégio da parceria e da orientação do Carlos d'Andréa. Seu acompanhamento atento, tranquilidade e generosidade foram marcas importantes da experiência do doutorado e suas contribuições à travessia desse difícil percurso foram maiores do que posso agradecer aqui. Aos professores Carlos Falci e Eduardo de Jesus agradeço, especialmente, pela leitura na fase de qualificação e pelas indicações que me ajudaram a concluir o trabalho. À professora Maria Aparecida Moura, agradeço pelo parecer elaborado ao projeto, em um momento ainda de muitas incertezas, diante do qual me ajudou a encontrar o eixo experimental de minhas indagações.

Aos demais professores do PPGCOM e do DCS, agradeço pelo diálogo, pelo acolhimento da pesquisa e por mais esse período de convivência, em especial a Fernanda Duarte, Joana Ziller, Geane Alzamora, Roberta Veiga, Ana Carolina Vimieiro, Carlos Alberto Carvalho, Ângela Marques e Elton Antunes. Às secretárias Elaine Martins e Tatiane Oliveira sou muito grato pelo apoio e orientação nos trâmites burocráticos. Como sempre, trago ainda a gratidão a outros professores e orientadores tão importantes em minha trajetória até aqui, em especial: André Brasil, Morten Søndergaard, Bernardo Jefferson Oliveira, Laura Guimarães Corrêa, Patrícia Moran e Paulo B.

Aos colegas do PPGCOM, do NucCon e do R-EST, sou muito grato pela partilha desse período. Pelas trocas, palavras de apoio e preciosa interlocução, agradeço especialmente a Leo Melgaço, Amanda Jurno, Ana Guerra, Thereza Nardelli, Maiana Abi, Sílvia Dalben, Tiago Salgado, Luciana Andrade, Daniel Loiola e Igor Lage. Esta pesquisa também não poderia ter acontecido sem as muitas oportunidades de colaboração e interlocução externas à UFMG. Agradeço em especial a Janna Joceli Omena, Elaine Rabello, Tarcízio Silva, Gabriel Menotti, Beatrice Gobbo e Bernhard Rieder. Por conversas mais pontuais, mas generosas, e que muito contribuíram a este percurso, sou também grato a Richard Rogers, Donato Ricci, Gabriele Colombo, Daniel Irrgang, Martin Guinard-Terrin, Markus Stauff, Judith Keilbach, Marc Tutters, Fabio Goveia e Fernanda Bruno.

Pelos convites e apoios, parciais ou integrais, concedidos à minha participação em eventos e outras atividades derivadas deste projeto, e que a ele muito contribuíram, agradeço: ao iNOVA Media Lab da Universidade Nova de Lisboa, no nome de Janna Joceli

Omena; à Association of Internet Researchers, em especial à comissão de bolsas de viagem; ao CODA.BR, nos nomes de Tarcízio Silva e Natália Mazzote; ao Instituto de Medicina Social da UERJ, no nome de Elaine Rabello; à organização da 7ª conferência MediaArtsHistories, no nome de Wendy Coones; à organização da Summer School 2017 da Digital Methods Initiative, no nome de Sabine Niederer; e ao ConJor, da UFOP, no nome de Débora Lopez.

Esta pesquisa foi desenvolvida com bolsa concedida pela CAPES, instituição cuja ameaçada existência foi fundamental para a realização deste doutorado. Agradeço tanto pela bolsa quanto pelos apoios concedidos ao PPGCOM/UFMG. Da mesma forma, sou imensamente grato à UFMG, instituição em que completo agora um terceiro ciclo de formação. Também gostaria de reconhecer aqui a importância do CNPq e da Fapemig – bem como de outras fundações de apoio do país – que, embora não tenham concedido apoio direto, foram fundamentais para a realização desta pesquisa. Mesmo em estado já fragilizado, elas são responsáveis pela constituição de um sistema de pesquisa e ambiente de trocas de valor inestimável a este doutorado.

Aos amigos e à minha família, agradeço pelo apoio e pelo cuidado durante este difícil período. Às amigas Hortência Abreu e Clarice Lacerda, especialmente, agradeço pela escuta, pelas conversas e pela amizade ao longo destes anos. À minha mãe, meu pai e minhas irmãs, à Tau e aos meus sogros, agradeço pela compreensão afetuosa de meu distanciamento e por me apoiar, de diferentes formas, neste desafio. À Janaina dedico a tese, mas não é suficiente. Quem quer este calhamaço em agradecimento? Espero, ao menos, ser capaz de retribuir à altura. Obrigado por me ajudar a segurar as pontas do mundo enquanto eu só podia me voltar a isto.

*A questão é sempre saber a que ponto
essa metamorfose ainda é consciente.
Não vivemos outra coisa senão
a metamorfose.*

Aby Warburg

Resumo

Esta pesquisa visa contribuir ao campo dos Métodos Digitais para estudos em Comunicação e Mídia, enfocando especificamente o problema da análise computacional de imagens por técnicas de aprendizado de máquina. Visando superar perspectivas metodológicas operacionais que se voltam a abordagens quantitativas, este estudo propõe refletir sobre o tratamento teórico concedido às imagens digitais nessas investigações e sobre como métodos computacionais conformam modos de ver particulares. O referencial teórico ampara-se principalmente nos Estudos de Ciência e Tecnologia (STS) e, em especial, na Teoria Ator-Rede (TAR). Por meio de conceitos desses domínios, as imagens são compreendidas como inscrições sociotécnicas, em uma condição de multiplicidade ontológica. Assume-se, portanto, a incerteza quanto à possibilidade de tratamento individual da imagem, considerando-a, de outro modo, como efeito de materialidades relacionais distribuídas. Estes são aspectos centrais da hipótese conceitual da *imagem-rede*, proposta por esta tese. Esta formulação é articulada a um esforço de descrição da operação de técnicas de reconhecimento de imagem por aprendizado de máquina baseadas em redes neurais artificiais. Em vista da vinculação desses modelos a amplas bases de treinamento colhidas da internet, bem como de sua tendência de infraestruturalização, considera-se que eles seriam parte importante do campo visual contemporâneo, engendrando visualidades computacionais com forte participação em processos de datificação e mediação algorítmica das imagens em plataformas *online*. A aplicação dessas técnicas como recursos metodológicos é proposta, portanto, como uma reapropriação crítica que considera os métodos como partes integrantes do objeto investigado. Esse gesto é exercitado em um estudo de caso voltado a imagens publicadas no Twitter durante um evento midiático. Em uma confluência das discussões teórico-metodológicas elaboradas ao longo do trabalho, propõe-se, ao final, um dispositivo metodológico denominado *Atlas para Imagens-Redes*. Ele visa proporcionar condições para práticas de navegação heurística pelas imagens, enquanto também busca resguardar a multiplicidade ontológica de suas instâncias.

Palavras-chave: imagem; métodos digitais; STS; aprendizado de máquina; visão computacional.

Abstract

Computational visualities and the image-network: repurposing machine learning for studying images on online platforms

This research seeks to contribute to the field of Digital Methods for studies in Communication and Media by focusing, specifically, the issue of computational analysis of images through machine learning techniques. Aiming to overcome operational methodological perspectives which turn to quantitative approaches, this study proposes to reflect upon how digital images are theoretically considered in those efforts and how computational methods conform particular modes of seeing. The theoretical framework is mainly derived from Science and Technology Studies (STS) and, especially, Actor-Network Theory (ANT). Through concepts from these domains, images are understood as sociotechnical inscriptions, in a condition of ontological multiplicity. An uncertainty is therefore assumed regarding the possibility of treating images individually and, in contrast, it is proposed to approach them as effects of distributed relational materialities. These are central aspects of the conceptual hypothesis of the *image-network*, which is proposed by this thesis. This formulation is articulated to an effort of describing the operation of machine learning image recognition techniques based on artificial neural networks. Considering the relations among these models and large training data sets harvested from the internet, as well as their infrastructuralization tendencies, it is considered that they are important components of the contemporary visual field, generating computational visualities with strong participation in datafication and algorithmic mediation processes to which images are subjected in online platforms. The application of these techniques as methodological resources is proposed, therefore, as a critical repurposing which considers methods as integral parts of the objects under scrutiny. This approach is exercised in a case study focusing on images published on Twitter during a media event. In a confluence of the theoretical and methodological discussions, a methodological device named *Atlas for Image-Networks* is finally proposed. It seeks to afford conditions for heuristic navigational practices through the images, while also attempting to preserve the ontological multiplicity of their instantiations.

Keywords: image; digital methods; STS; machine learning; computer vision.

Lista de figuras

Figura 1: Interface do Instagram sem imagens.....	16
Figura 2: Exemplo de 'cartão Shirley' de 1978.....	68
Figura 3: Reconstrução de pranchas do Atlas Mnemosyne de Aby Warburg.....	93
Figura 4: Prancha 46 do Atlas Mnemosyne de Aby Warburg.....	98
Figura 5: Esquema comparativo entre os paradigmas dedutivos e indutivos.....	130
Figura 6: Ilustração do programa desenvolvido por Lawrence G. Roberts.....	141
Figura 7: Ilustração do 'estímulo ótimo' para classes de um modelo de rede neural.....	142
Figura 8: Arquitetura da rede neural do modelo VGG19.....	145
Figura 9: Estímulos ótimos para 12 filtros da camada block1_conv2, modelo VGG19.....	147
Figura 10: Estímulos ótimos para 12 filtros da camada block5_conv1, modelo VGG19.....	148
Figura 11: Exemplo de imagem gerada pelo programa DeepDream.....	149
Figura 12: Impressão de tela do projeto Curator Table.....	162
Figura 13: Impressão de tela do projeto Neural Neighbors.....	163
Figura 14: Montagem justapondo fotografias do tecido estendido no SMABC.....	185
Figura 15: Protocolo de coleta e primeira fase de processamento.....	188
Figura 16: Impressão de tela do Twitter com imagem carregada na plataforma.....	203
Figura 17: Impressão de tela do Twitter com imagem de pré-visualização de link.....	204
Figura 18: Protocolo de agrupamento das instâncias de imagens 'únicas'.....	222
Figura 19: Protocolo de processamento das imagens pelo modelo VGG19.....	232
Figura 20: Esquema de processamento de uma imagem pelo VGG19.....	234
Figura 21: Detalhe da visualização em grade com layout pelo modelo VGG19.....	241
Figura 22: Protocolo de processamento das imagens pela GVAPI.....	244
Figura 23: Detalhe da visualização em grade com layout pelo modelo GVAPI.....	256
Figura 24: Impressão de tela do estado inicial do Atlas no layout VGG19.....	264
Figura 25: Impressão de tela do Atlas em layout VGG19 com mapa de calor.....	265
Figura 26: Impressão de tela do Atlas em layout VGG19 com linhas de grade.....	266
Figura 27: Impressão de tela do Atlas em visualização de detalhe.....	267
Figura 28: Impressão de tela com exemplo de agrupamento inconsistente.....	269
Figura 29: Montagem de matriz e instâncias de um agrupamento mal formado.....	271
Figura 30: Impressão de tela do Atlas com imagens similares consideradas distintas.....	271
Figura 31: Impressão de tela com imagem do personagem Lula Molusco.....	272

Figura 32: Impressão de tela do Atlas com fotografia de Proner Ramos.....	274
Figura 33: Memes derivados da fotografia de Proner Ramos.....	275
Figura 34: Montagens de Lula atrás das grades.....	277
Figura 35: Variações sobre o retrato de identificação de 1980.....	277
Figura 36: Depoimentos de testemunhas e réus da Lava Jato.....	277
Figura 37: Faixas da campanha pela liberdade de Lula.....	278
Figura 38: Chegada de Lula na sede da PF em Curitiba.....	278
Figura 39: Figurações de Lula na multidão.....	278

Lista de gráficos

Gráfico 1: Visualização de espaços semânticos de três APIs de Visão Computacional.....	171
Gráfico 2: Diagrama de Gantt de períodos de coleta para termos de busca.....	194
Gráfico 3: Visualização da rede de coocorrência de termos de coleta.....	197
Gráfico 4: Série temporal do número de tuítes coletados, agregados por hora.....	202
Gráfico 5: Visualização de grade das imagens com similaridade estimada pelo VGG19.	239
Gráfico 6: Visualização de rede bimodal imagem-etiqueta com dados da GVAPI.....	246
Gráfico 7: Visualização de rede imagem-etiqueta com imagens plotadas.....	251
Gráfico 8: Visualização de grade das imagens com similaridade estimada pela GVAPI..	254

Lista de tabelas

Tabela 1: Descrição geral dos tuítes do caso Lula.....	201
Tabela 2: Dados dos tuítes com media e URLs de media.....	206
Tabela 3: Dados das URLs de links.....	212
Tabela 4: Estatísticas descritivas das publicações de URLs de links.....	214
Tabela 5: Estatísticas descritivas das publicações com URLs de link selecionadas.....	216
Tabela 6: Estatísticas descritivas de publicação das 'imagens únicas'	229
Tabela 7: Estatísticas descritivas de publicação das 'imagens únicas' selecionadas.....	230
Tabela 8: Termos de busca e métricas de desempenho de coleta no caso Lula.....	310
Tabela 9: Sumário de camadas da rede neural do modelo VGG19.....	315

Lista de abreviações e siglas

API	<i>Application Programming Interface – Interface de Programação de Aplicações</i>
CBIR	<i>Content-based image retrieval – Recuperação de imagem baseada em conteúdo</i>
CSS	<i>Cascading Style Sheets – Folhas de Estilo em Cascata</i>
CSV	<i>Comma-separated values – Valores separados por vírgula</i>
DMI	<i>Digital Methods Initiative – Iniciativa de Métodos Digitais da Universidade de Amsterdã</i>
DMI-TCAT	<i>Twitter Capture and Analysis Toolset – Kit de Ferramentas para Captura e Análise do Twitter</i>
GEXF	<i>Graph Exchange XML Format – Formato XML de Intercâmbio de Grafo</i>
GIF	<i>Graphics Interchange Format – Formato de Intercâmbio de Gráficos</i>
GVAPI	<i>Google Cloud Vision API – API Cloud Vision da Google</i>
GPU	<i>Graphical Processing Unit – Unidade de Processamento Gráfico</i>
HTML	<i>Hypertext Markup Language – Linguagem de Marcação de Hipertexto</i>
HTTP	<i>Hypertext Transfer Protocol – Protocolo de Transferência de Hipertexto</i>
IA	<i>Inteligência Artificial</i>
ILSVRC	<i>Imagenet Large Scale Visual Recognition Challenge – Desafio de Reconhecimento Visual de Grande Escala Imagenet</i>
JPEG / JPG	<i>Joint Photographic Experts Group – Grupo Conjunto de Especialistas em Fotografia</i>
JSON	<i>JavaScript Object Notation – Notação de Objeto JavaScript</i>
ML4A	<i>Machine Learning for Artists – Aprendizado de Máquina para Artista</i>
NASA	<i>North-American Space Agency – Agência Espacial Norte-Americana</i>
PCA	<i>Principal Component Analysis – Análise de Componente Principal</i>
PF	<i>Polícia Federal</i>
PNG	<i>Portable Network Graphics – Gráfico de Rede Portátil</i>
SMABC	<i>Sindicato dos Metalúrgicos do ABC</i>
SSK	<i>Sociology of Scientific Knowledge – Sociologia do Conhecimento Científico</i>
STS	<i>Science and Technology Studies – Estudos de Ciência e Tecnologia</i>
SVG	<i>Scalable Vector Graphics – Gráfico de Vetor Escalável</i>
<i>t</i> -SNE	<i>t-Distributed Stochastic Neighbor Embedding</i>

TAR	<i>Teoria Ator-Rede</i>
TRF-4	<i>Tribunal Regional Federal da 4ª Região</i>
URL	<i>Uniform Resource Locator – Localizador Uniforme de Recursos</i>
UvA	<i>Universiteit van Amsterdam – Universidade de Amsterdã</i>
VSML	<i>Visual Social Media Lab – Laboratório de Mídia Social Visual das seguintes instituições britânicas: Universidade de Sheffield, Escola de Arte de Manchester, Universidade de Wolverhampton e Universidade de Londres</i>
URL	<i>Uniform Resource Locator – Localizador Uniforme de Recursos</i>
VGG	<i>Visual Geometry Group – Grupo de Geometria Visual da Universidade de Oxford</i>
ZKM	<i>Zentrum für Kunst und Medien Karlsruhe – Centro de Arte e Mídia de Karlsruhe</i>

Lista de notações e símbolos

fonte monoespçada

Quando utilizada no corpo do texto, a fonte tipográfica monoespçada identifica termos com sentido específico a uma linguagem de programação ou ao protocolo de uma API.

[]

Colchetes indicam a delimitação de um termo utilizado na busca (*query*) em uma coleta de dados, exceto quando utilizados para indicar supressões ou adições a citações diretas. Seguindo recomendação proposta por Richard Rogers (2017), o uso dos colchetes tem por objetivo evitar a ambiguidade que poderia decorrer do uso de aspas (“”), já que as aspas são geralmente utilizadas como parte dos termos de busca para limitar os resultados a ocorrências literais.

□ VGG19 C:9

□ GVAPI C:9

Coordenadas alfanuméricas precedidas por um quadrado (□) indicam o quadrante em que se localiza uma imagem ou grupo de imagens no *Atlas para Imagens-Redes* apresentado no capítulo 4. Instruções de acesso e navegação são apresentadas no Apêndice B.

A sigla VGG19 indica se tratar do mapa organizado pelo modelo de reconhecimento de imagens homônimo. A sigla GVAPI indica se tratar do mapa organizado segundo o modelo de reconhecimento de imagens da API Cloud Vision da Google.

A letra indica a coluna do quadrante referenciado (A-Q) e o numeral indica a linha (1-17).

Sumário

1	Introdução.....	17
	1.1 Definição do problema.....	25
	1.2 A questão do método.....	32
	1.3 Estrutura da tese.....	37
2	Imagem-rede.....	40
	2.1 A imagem e o <i>digital</i>	45
	2.2 A imagem múltipla.....	57
	2.3 A imagem e a rede.....	92
3	Visualidades computacionais.....	111
	3.1 O visual e o computacional.....	116
	3.2 Aprendizado de máquina.....	127
	3.3 Visualidade e infraestrutura.....	165
4	Estudo de caso: imagens de uma prisão.....	178
	4.1 Coleta e extração de dados.....	186
	4.2 Processamento das imagens.....	217
	4.3 Compondo imagens-redes.....	257
5	Considerações finais.....	281
	Referências.....	288
	Apêndices.....	309
	Anexo.....	314

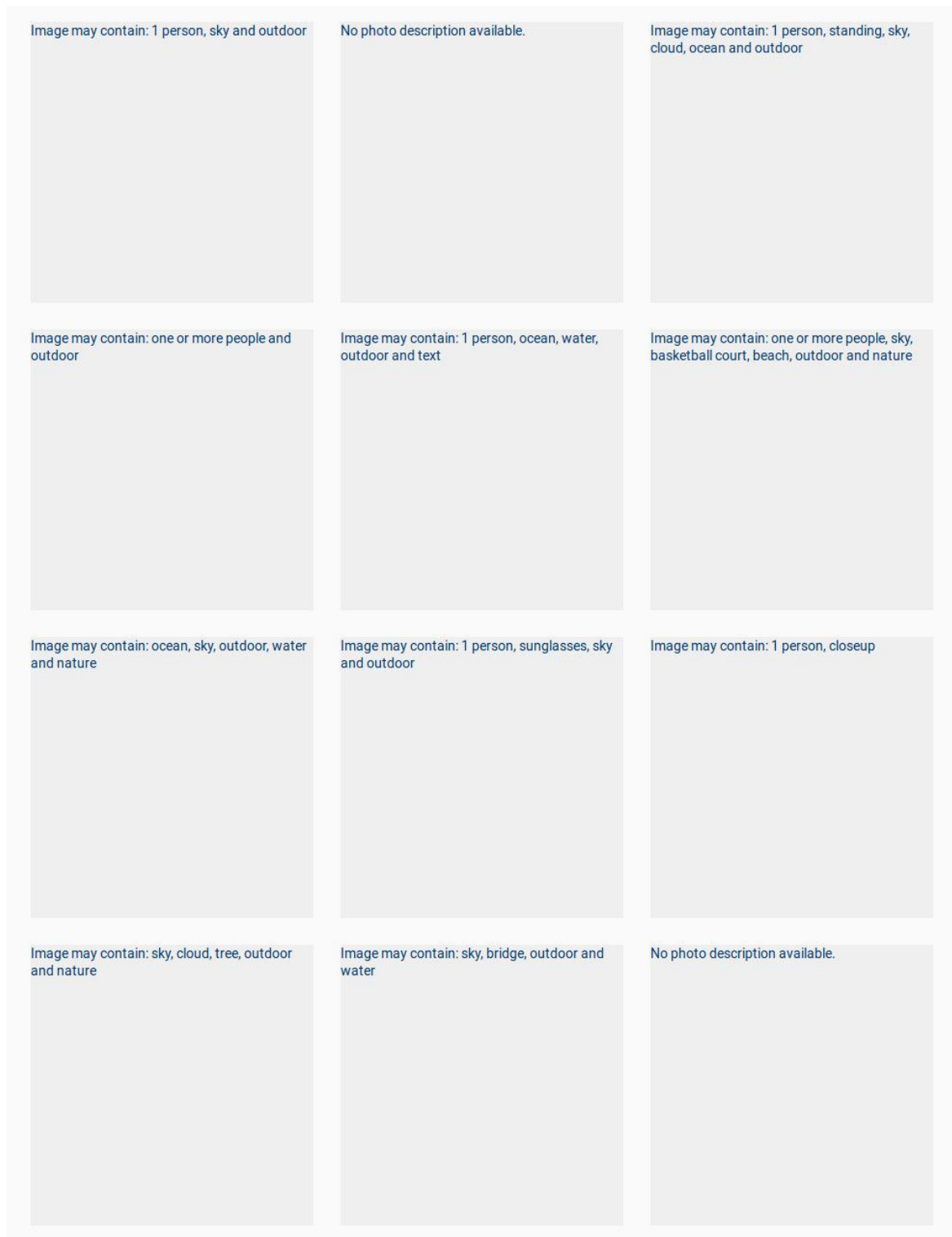


Figura 1: Interface do Instagram sem imagens.
Simulação da falha operacional ocorrida no dia 3 de julho de 2019.

Fonte: Elaboração própria por meio de intervenção no carregamento da interface da plataforma.

1 Introdução

Em 3 de julho de 2019, uma quarta-feira, *Facebook*, *Instagram* e *Whatsapp* tiveram seu funcionamento comprometido com a impossibilidade de acessar imagens, vídeos e outros arquivos, em todo o mundo (THE GUARDIAN, 2019). Embora compromettesse apenas parcialmente o uso das plataformas, o defeito não passou despercebido pelos usuários, inclusive com a repercussão, no *Twitter*, de *hashtags* sobre o estado inoperante das plataformas “vizinhas” – *#facebookdown*, *#instagramdown* e *#whatsappdown*. A ausência das imagens, cada vez mais centrais às práticas de uso das plataformas, tornava estes ambientes inutilizáveis. As interfaces, normalmente povoadas de fotografias, *memes* visuais e ilustrações, eram agora permeadas de quadrados cinzas. Ocupando as lacunas deixadas pelas imagens, emergiram descrições verbais do conteúdo ausente. Indicando os aspectos presumidos das imagens segundo o *software* de reconhecimento que integra as plataformas, apresentavam-se verbalmente formulações como: “A imagem pode conter: uma pessoa em *closeup*” (Figura.1).

Os componentes defeituosos que levaram à falha técnica não foram divulgados. O *Facebook* – empresa proprietária das três plataformas – apenas se desculpou aos usuários quando os problemas haviam sido resolvidos¹. Finda a instabilidade, reestabelecia-se a regularidade funcional das plataformas. No entanto, as descrições tornadas visíveis pela falha temporária tornaram-se objeto de discussão, ainda que esta também fosse transitória. Com reações que ora ressaltavam o caráter simplório e falho das descrições, ora especulavam sobre os seus usos potenciais para o perfilamento algorítmico e microdirecionamento publicitário, o caso chamou a atenção de usuários, pesquisadores² e imprensa especializada (VINCENT, 2019). Portadores de deficiências visuais que fazem uso de tecnologias assistivas em sua interação com a *Web* possivelmente já teriam tomado contato com tais descrições, já que uma de suas finalidades é a acessibilidade do

1 No *Twitter*, o *Facebook* se manifestou na noite daquele dia: “Mais cedo hoje, algumas pessoas e empresas experienciaram problemas no carregamento e envio de imagens, vídeos e outros arquivos em nossos aplicativos e plataformas. O problema foi resolvido desde então e deve estar 100% operante para todos. Nos desculpamos por qualquer inconveniência”. No original: “Earlier today, some people and businesses experienced trouble uploading or sending images, videos and other files on our apps and platforms. The issue has since been resolved and we should be back at 100% for everyone. We're sorry for any inconvenience”. Tradução minha. (FACEBOOK, 2019).

2 Agradeço ao Tarcízio Silva, que chamou minha atenção a esse acontecimento.

conteúdo visual a esse segmento do público (WU, 2016)³. No âmbito da pesquisa em mídias sociais, é também possível que as descrições fossem conhecidas como um dos metadados disponíveis para o estudo de publicações nas plataformas⁴. Para a maior parte dos usuários videntes, contudo, tratava-se de uma camada oculta que precisou da falha técnica do dia 3 de julho para vir à luz.

Situações de falha como esta constituem uma das oportunidades descritas por Bruno Latour para observar a agência de objetos técnicos: “subitamente, intermediários completamente silenciosos tornam-se mediadores plenos”⁵ (LATOURE, 2005, p. 81)⁶. Segundo o vocabulário do autor (LATOURE, 2001), diz-se que os objetos técnicos passam por um processo de obscurecimento ou “encaixapretamento” (*blackboxing*), em referência à figura cibernética da caixa-preta. Este processo consistiria uma transição dos objetos técnicos em que suas partes internas e dependências externas seriam invisibilizadas, passando a ser tomados como entidades integrais e estabilizadas. Remetendo a Gilbert Simondon (2007), poderíamos relacionar tal processo de “encaixapretamento” com o vetor de *concretização* que, segundo descreve, leva à integração eficiente dos componentes de um objeto técnico em seu processo de gênese. Tal processo, segundo o autor: “tende à coerência interna, ao encerramento do sistema de causas e efeitos que se exercem circularmente no interior de seu recinto” (SIMONDON, 2007, p. 67)⁷. A falha técnica, porém, tem o poder de reverter o processo de *concretização* e de revelar o interior, até então ignorado, da caixa-preta. Um ou mais componentes são então trazidos à superfície e, com ele, incertezas quanto a seu papel e consequências naquele arranjo sociotécnico.

Com as imagens fora de cena, tornaram-se visíveis indícios do modo pelo qual as plataformas as “veem”. Embora largamente ocultas pelo progressivo alisamento das

3 As descrições são inseridas no código fonte das páginas sob o atributo `alt` das imagens. Este atributo, segundo recomendações do *World Wide Web Consortium (W3C)* para a HTML (*Hypertext Markup Language*), têm como finalidade de descrever o conteúdo da imagem para um usuário que, por qualquer motivo não consiga vê-la ou compreendê-la, tais como usuários com conexões lentas ou que fazem uso de tecnologias assistivas (CALDWELL et al., 2008). Geralmente oculto, portanto, constitui um metadado utilizado por tais tecnologias.

4 Eu havia tomado conhecimento de sua existência apenas poucas semanas antes, durante uso da ferramenta *Instaloader* (GRAF et al., 2019), para coleta de conteúdo do *Instagram*. A ferramenta funciona por meio de raspagem, realizada por meio da extração de dados da interface gráfica da plataforma. O conteúdo do atributo `alt` é integrado aos resultados da ferramenta sob o identificador `accessibility-caption` (legenda para acessibilidade).

5 No original: “[...] *all of a sudden, completely silent intermediaries become full-blown mediators*”. Tradução minha.

6 A listagem das situações em que a agência dos objetos técnicos se torna perceptível é apresentada por Latour em 2005 é uma reedição expandida de listagem apresentada por ele e Madeleine Akrich em 1992 (cf. AKRICH; LATOURE, 1992, p. 260), como situações que tornam possível a *de-scrição* dos arranjos sociotécnicos.

7 No original: “[...] *tiende a la coherencia interna, a la cerrazón del sistema de causas y efectos que se ejercen circularmente em el interior de su recinto [...]*”. Tradução minha.

interfaces, na constituição de interações fluidas e intuitivas aos usuários, as tecnologias de reconhecimento de imagens encontram-se cada vez mais integradas às plataformas digitais. Elas vêm realizar, para as imagens, a demanda de *datificação*, central ao modelo de negócios e à própria operação destas formações midiáticas contemporâneas (VAN DIJCK, 2014, 2017). Compreendida como o processo de pretensa tradução de toda atividade social à forma dos dados computacionais, a datificação é o que permite às plataformas *online* tratar computacionalmente esta atividade e submetê-la a seus procedimentos *algorítmicos* de mediação.

José Van Dijck (2013, p. 13–14) reflete que, contrariamente aos *slogans* de plataformas como o *Facebook*, o que acompanhamos pelo menos desde o advento da chamada *Web 2.0* (cf. O'REILLY, 2005) não foi tanto a Web se tornar mais social quanto, inversamente, a sociabilidade tornar-se mais técnica. Segundo sugere a autora (VAN DIJCK, 2013, p. 51–52), a amizade no âmbito das plataformas não pode ser compreendida apenas segundo relações interpessoais espontâneas. Deve-se, de outro modo, compreendê-la também como uma sociabilidade programada (cf. BUCHER, 2013b), informada pelo modo como a plataforma codifica a amizade, inclusive por algoritmos de recomendação de novos amigos. A *mediação algorítmica do visível* evidenciada pelo caso em questão, parece nos autorizar a dizer que também as imagens são inseridas em outros arranjos técnicos no âmbito das plataformas, em algo como uma *visualidade programada*. Esta tecnicização das imagens relaciona-se, portanto, ao modo pelo qual as plataformas conseguem lidar computacionalmente com as imagens para identificar pessoas, locais e objetos; desenvolver perfis datificados de seus usuários; microdirecionar anúncios e publicações; e realizar buscas, entre outras operações – explícitas ou não.

A demanda por estes processos automatizados de tratamento das imagens evidencia-se quando se considera a crescente centralidade que os conteúdos visuais adquirem nos usos contemporâneos das mídias sociais. Nos últimos anos, diz-se que as plataformas passam por uma “virada visual” ou “pictórica” (FAULKNER; VIS; D'ORAZIO, 2018; GIBBS et al., 2015; NIEDERER, 2018), ou, ainda, que manifestam uma “ubiquidade do visual” (HIGHFIELD; LEAVER, 2016). Nota-se não apenas o crescimento de plataformas com foco específico em imagens (como *Instagram*, *YouTube*, *Pinterest* e *TikTok*) como, também, a priorização das imagens em outras plataformas (como *Facebook*, *Twitter* e *Whatsapp*). Diferentes estimativas referenciadas pelos autores que indicam esta “virada” sugerem números astronômicos, na ordem de *bilhões* de imagens por dia, em uma única

plataforma, como o *Instagram*. A repercussão da falha do dia 3 de julho de 2019 constitui um indicador ainda mais palpável da centralidade do visual: a indisponibilidade das imagens foi sentida como total inoperância das plataformas, para além de uma falha parcial e isolada.

A tendência de vinculação das imagens a processos de tratamento computacional não se manifesta apenas nas plataformas mas, também, nas pesquisas que se voltam à comunicação em mídias digitais. A “virada visual” coloca desafios metodológicos a um campo de pesquisa que, como destacado por Ana Carolina Vimieiro e Janine Bargas (2018), em diálogo com David Berry (2012), experiencia há alguns anos uma “virada computacional” em seus métodos de pesquisa. Segundo define Berry (2012, p. 4), esta “virada” seria uma terceira onda das chamadas humanidades digitais, interessada no tema da *computacionalidade* das formas digitais, nos seus aspectos em alguma medida específicos, e nos impactos epistêmicos das mudanças midiáticas. Os desafios do visual, como discutirei adiante, decorrem de aspectos relativos às características informacionais das imagens digitais, que elevam a complexidade técnica e, de modo mais importante, às incertezas metodológicas em seu tratamento computacional. Recursos como as tecnologias de reconhecimento de imagens mobilizadas pelas plataformas são, nesse contexto, instrumentos sedutores para a pesquisa já informada por métodos computacionais. Em particular, dada sua maior acessibilidade, nos últimos anos, em soluções comerciais ou de código aberto. Mas, embora cada vez mais mobilizadas como ferramentas metodológicas⁸, poucos estudos refletem acerca de suas implicações epistemológicas ou de seu lastreamento teórico nos estudos das imagens. Este é um dos aspectos que busco trabalhar nesta tese.

Em linhas gerais, esta pesquisa se volta a um problema marcadamente metodológico que, em seu aspecto mais operacional, se manifesta no desafio colocado pelas imagens às pesquisas contemporâneas da comunicação em plataformas *online*. Em particular, àquelas que se baseiam em métodos computacionais de coleta e análise. Tal aspecto operacional, contudo, desdobra-se em dois fenômenos que considero de forma entrelaçada: a já referida “virada visual” das plataformas de mídia social; e o desenvolvimento de *mediações algorítmicas do visível*, em especial por programas de reconhecimento de

8 Cf. HUSSAIN et al., 2017; JOO et al., 2014; NIEDERER; COLOMBO, 2019; RICCI et al., 2017; RYKOV et al., 2016; SILVA; BARCIELA; MEIRELLES, 2018; SILVA; ZANOTTI, 2018. Indico, também, alguns estudos em que participei diretamente: CICALI et al., 2018; D’ANDREA; MINTZ, 2019; MINTZ; SILVA et al., 2019; OMENA; RABELLO; MINTZ, 2017, no prelo.

imagens. Estes se manifestam, hoje, de muitas formas. Uma das mais controversas, em vista de suas graves premissas e implicações, é o reconhecimento de rostos. Outro caso é o reconhecimento e detecção de objetos, como no exemplo apresentado na [Figura 1](#), utilizado em programas de recuperação da informação, tecnologias assistivas e na moderação e seleção de conteúdos em plataformas *online*. Casos em que a vinculação é menos óbvia são, por exemplo, filtros do *Instagram* e *Snapchat* (muitos dos quais são sensíveis ao conteúdo das imagens) ou na tecnologia recente que possibilita os *deep fakes*.

O entrelaçamento entre estes fenômenos pode ser considerado em ao menos três níveis. Primeiro, tendo em vista que a profusão de imagens – em larga medida estimulada pelos arranjos tecnológicos contemporâneos – gera a demanda por técnicas para integrá-las à constituição sociotécnica das plataformas. Isto é: demanda processos de *datificação do visível*, realizados por programas de reconhecimento de imagens.

Em um segundo nível, o desenvolvimento recente das tecnologias de reconhecimento de imagens, pela via do *aprendizado de máquina*, encontra-se diretamente vinculado à profusão de imagens em plataformas *online*. Isto porque este paradigma de desenvolvimento, baseado em redes neurais artificiais, realiza-se por processos indutivos de “treinamento” com base em exemplos. Desse modo, são tecnologias dependentes da disponibilidade de grandes volumes de dados. Como ressaltado por autores da área da computação (ALPAYDIN, 2016), a *Web* participativa – outrora chamada *Web 2.0* e que hoje culmina nas plataformas de mídia social – teve papel determinante na construção destas bases de treinamento. Adrian Mackenzie (2017), em sua discussão do aprendizado de máquina, situa-o como tecnologia emergente que tem o acúmulo de dados de treinamento na *Web* uma de suas condições de existência. Não por acaso, o projeto *Imagenet*, que constituiu um dos principais bancos de imagens para treinamento destes programas, baseia-se inteiramente em imagens extraídas da *Web* (DENG et al., 2009). Nesse sentido, mais do que uma tecnologia de análise exógena, trata-se de tecnologia cuja constituição encontra-se já profundamente integrada ao contexto das plataformas *online* e sua virada visual recente.

Em um terceiro nível, o tratamento articulado entre a profusão das imagens e seu tratamento computacional se justifica pela consideração de como estes fenômenos constituiriam aspectos de uma *visualidade* contemporânea. Pouco acima me referi à ideia

de uma *visualidade programada*, em diálogo com a ideia de *sociabilidade programada* de Van Dijck e Taina Bucher. Mas elaborarei nesta tese a ideia de uma *visualidade computacional*, retomando noção que esbocei em trabalhos anteriores (MINTZ, 2015, 2018d). Não quero estabelecer, aqui, uma distinção muito precisa entre os adjetivos (*programada* e *computacional*), embora certamente haveria matizes por salientar. Um aspecto por ressaltar, em todo caso, é mais conotativo: *computacional* sugere uma discussão mais voltada ao aspecto salientado por Berry, das características de um meio computacional e suas implicações epistemológicas, enquanto *programada* sugere a execução de uma rotina predeterminada. Voltarei a algumas destas questões adiante (veja [3. Visualidades computacionais](#)), mas já queria indicar a que me refiro antes de avançar.

Segundo formulou Hal Foster (1999, p. ix), *visualidade* designa um dos elementos de uma diferença interna ao *visual*. Didaticamente, o termo sugere a constituição social e histórica da *visão*. Esta, por sua vez, é tipicamente considerada em seu aspecto físico ou fisiológico, com suposição de uma capacidade universal e a-histórica. Contudo, propõe o autor, isto não deveria levar à contraposição entre estes termos nem sua correspondência à oposição entre natureza e cultura, mas sim à sua articulação enquanto duas dimensões do tratamento do *visual*⁹. De modo importante, *visualidade* designa o caráter necessariamente histórico do campo visual. Escreve:

[a] visão também é social e histórica e [a] visualidade envolve o corpo e a psique. No entanto, elas tampouco são idênticas: aqui, a diferença entre os termos marca uma diferença interna ao visual – entre o mecanismo da visão e suas técnicas históricas, entre o dado da visão e suas determinações discursivas – uma diferença, muitas diferenças, entre como nós vemos, como nós somos capazes, permitidos, ou levados a ver, e como nós vemos esse ato de ver ou o que nele permanece não visto¹⁰ (FOSTER, 1999, p. ix).

9 John Walker e Sarah Chaplin (2002, citados por ABRIL, 2012, p. 17) sintetizam a noção de *visualidade* como “*visão socializada*”. Esta noção, retomada por Gonzalo Abril é, por sua vez, mobilizada por Carlos Mendonça e Bruno Leal (2018, p. 110) em sua conceituação da *visualidade*. WALKER, John. CHAPLIN, Sarah. *Una introducción a la cultura visual*. Barcelona: Octaedro-EUB, 2002. Hal Foster (1999, p. ix), contudo, sugere uma abordagem que parcialmente diverge dessa ao evitar a correlação imediata entre *visão* e *visualidade* ao par *natureza* e *cultura*. Foster não recusa a oposição entre natureza e cultura, mas apenas indica que ela não é adequada para este caso, como desenvolve na citação trazida ao texto. Neste trabalho, contudo, busco aprofundar esta perspectiva ao adotar a perspectiva simétrica sugerida por Bruno Latour (1994), que visa justamente desfazer esta distinção essencialista. Embora não se referenciem, considero que, no âmbito de seu tratamento do visual, a proposição de Foster parece ir ao encontro de Latour. A discussão sobre a visão em articulação às suas mediações técnicas, realizada por Jonathan Crary (1992), também ressoa nesta proposição, sem distinguir de antemão fisiologia e história. Por esse caminho, a oposição entre visão e visualidade, que supõe a definição de ‘visão socializada’ funciona apenas de forma didática, mas não conceitualmente. Pois segundo esta perspectiva não haveria algo como um conceito de ‘visão pré-social’ que pudesse ser então ‘socializada’.

10 No original: “*vision is social and historical too, and visibility involves the body and the psyche. Yet, neither are they identical: here, the difference between the terms signals a difference within the visual – between the mechanism of sight and its historical techniques, between the datum of vision and its discursive determinations – a difference, many differences, among how we see, how we are able, allowed or made to see, and how we see this seeing or the unseen therein*”. Tradução minha.

Visualidade sinaliza portanto uma demarcação contingente e situada do visual em sua constituição situada, enquanto *visão* aponta para sua compreensão essencialista, de pretensão universalizante. A *visão*, neste sentido, não existe senão como efeito retórico de um encerramento desta fratura constituinte do visual, ignorando as práticas, materialidades e formações discursivas que rompem com a possibilidade de uma *visão* universal (FOSTER, 1999). Por *visualidade computacional*, portanto, visio denominar um fator contemporâneo desta fratura, um *modo de ver* articulado pelas tecnologias de reconhecimento de imagens da chamada Visão Computacional e que, a seu modo, hoje participam da modulação do visual. Evidentemente, mesmo nesse caso seria forçoso dizer de *uma* *visualidade* computacional, como indicarei adiante. Um fator de complexificação é o fato de que estas tecnologias sejam desenvolvidas, como indiquei, a partir das próprias imagens que elas visam descrever. De certo modo, as tecnologias de reconhecimento de imagens são tanto *efeitos* da cultura visual contemporânea quanto parte de suas muitas *causas*. Isto porque a Visão Computacional é hoje produzida a partir destas imagens e é, também, constituinte dos modos pelos quais as imagens são integradas a plataformas digitais e a procedimentos metodológicos de pesquisa.

Recapitulando, sugiro ao menos três níveis de articulação entre a profusão de imagens *online* e o desenvolvimento de técnicas para o seu tratamento computacional. Primeiro, considerando que a explosão quantitativa das imagens demanda técnicas para seu tratamento computacional. Segundo, considerando que as próprias técnicas demandam um afluxo de muitas imagens para seu desenvolvimento. Terceiro, que a conjunção entre as imagens e as técnicas de seu tratamento computacional sugeriria modos particulares de *ver*. Este complexo entrelaçamento, quero argumentar, traz implicações teóricas e metodológicas à pesquisa. Sustentarei, nas páginas a seguir, que em linha com a tradição recente dos Estudos de Ciência e Tecnologia (*Science and Technology Studies – STS*), ambos fenômenos precisam ser tratados em uma perspectiva *simétrica*. O princípio da simetria, proposto inicialmente por David Bloor, passou por diferentes releituras ao longo das décadas¹¹. Noortje Marres e David Moats (2015) propõem uma inflexão particular ao princípio com o objetivo de superar tentativas de purificação das dinâmicas sociais observadas em mídias digitais, destilando-as dos arranjos técnicos em que elas se desenvolvem. Os autores propõem que as dinâmicas tecnológico-midiáticas sempre serão

11 Marres e Moats (2015) apontam que em sua proposição inicial por David Bloor, visava-se o tratamento simétrico entre explicações consideradas verdadeiras e falsas em determinado campo científico. Posteriormente, Callon e Latour propuseram o tratamento simétrico entre actantes humanos e não-humanos na descrição de arranjos sociotécnicos. A proposição de Marres e Moats configura, assim, uma terceira aceção.

parte das dinâmicas sociais observadas e, portanto, que ambas devem ser consideradas *simetricamente*. Para esta pesquisa, tal formulação sugere um caminho para compreender as imagens digitais em plataformas de mídia social de maneira articulada com as formas de sua mediação algorítmica.

Como ilustrado pela anedota trazida na abertura desta introdução, os programas de reconhecimento de imagens constituem uma instância importante da integração das imagens às plataformas de mídia social. Do ponto de vista dos pesquisadores de comunicação e mídia, tais programas também se apresentam como recursos para a análise das imagens no contexto de sua profusão e de sua deriva digital. Contudo, mais do que *instrumentos*, tais técnicas conformam um modo particular de olhar para as imagens. Um *modo de ver*. Baseando-se apenas nas descrições obtidas por tais programas, uma pesquisa sobre imagens no Instagram acabaria por se debruçar sobre uma realidade que a própria pesquisa produziu: um Instagram “sem imagens”, como ilustrado pela [Figura 1](#). Evidentemente, levando esta consideração às últimas consequências, cada método, a seu modo, com sua epistemologia, constrói sua própria realidade. Este não é, em si, o problema. Na linha do que discute Bruno Latour (2005), a questão é se esta construção é consistente ou não, algo que apenas se pode refletir se conhecemos as limitações de nosso método. Compreendendo que estas técnicas conformam uma *visualidade* particular, que constitui o campo visual, mas que nele também introduz diferença, seria preciso abordá-las dentro de seus próprios limites. Da mesma forma que esta *visualidade* não poderia ser tomada como uma *visão* essencial, tampouco pode seu modo particular de considerar as imagens sobrepor-se a outros modos possíveis. Conforme elaborarei adiante, isto demanda que compreendamos as imagens em suas múltiplas *inscrições*, que fazem delas o que Latour (LATOURE, 2001) se refere por uma *referência circulante*: não objetos precisos e de contornos definidos, mas como efeitos de uma multiplicidade ontológica que se diversifica por meio de diferentes manifestações sociotécnicas.

A pergunta “como comparar um milhão de imagens?” já foi feita por pesquisadores do campo (MANOVICH, 2012). Porém, quero argumentar que o problema não pode ser resumido a um *como* mas deve, de outro modo, mobilizar questões quanto ao *quê*. O *que* é comparar um milhão de imagens? O que significa este gesto de conhecimento e suas descobertas? Quais suas implicações? Pois a eventual solução do problema de ordem operacional apenas pode ser atestada se soubermos dizer o que é mesmo que estamos fazendo ao resolvê-lo. Os procedimentos computacionais a que as imagens são submeti-

das inevitavelmente transformam sua ontologia e o próprio ato de ver. Como venho sugerindo, esta transformação já ocorre em práticas contemporâneas de produção e consumo de imagens, em que a *quantidade* não é apenas um atributo dos *corpora* mobilizados nas análises, mas também de nossa experiência ao haver-mos com as imagens. Da mesma forma, para além de seu uso em pesquisas acadêmicas, mediações algorítmicas do visível já são incorporadas aos modos de ver contemporâneos, como partes integrantes das plataformas de mídia social. Contudo, esta pervasividade não significa uma completa sobredeterminação das imagens e do visual. O *como* vincula-se desse modo a um *quê*. O que são, então, as imagens que se compara?

1.1 Definição do problema

Esta tese visa contribuir ao desafio metodológico enfrentado por pesquisadores das imagens e da comunicação contemporânea quando, por exemplo, em um único dia de observações, se deparam com nada amigáveis 10 mil imagens como parte de seus *corpora*. Não falo aqui de um número hipotético. Uma coleta por tuítes mencionando os termos [*climate change*]¹² no dia 4 de julho de 2019 encontrou em torno de 10 mil imagens¹³. Diante deste conjunto, o que fazer? Em certos casos, seria suficiente voltar-se às publicações mais *curtidas* ou *retuitadas*. Isto é, pelo vocabulário das próprias plataformas, às publicações que alcançaram maior “engajamento” e que, portanto, poderiam ser tomadas, supostamente, como mais relevantes. Com base nestas métricas, seria possível obter um conjunto menor que poderia ser analisado de forma detida. Contudo, para além das várias críticas tecidas a tal abordagem¹⁴, é importante considerar que a *publicação* mais *curtida* ou *retuitada* não necessariamente corresponde à *imagem* mais relevante para o caso em questão, mesmo que o critério de relevância sejam tais métricas, pois uma ‘mesma’ imagem não é uma unidade de análise prontamente proporcionada pelos dados coletados. Geralmente, as imagens são identificadas por um arquivo de imagem que, por sua vez, é indicado por uma URL (Localizador Uniforme de Recursos, na sigla

12 Utilizo aqui a notação de termos de busca segundo proposta feita por Richard Rogers (2017). Cf. [Lista de notações e símbolos](#).

13 A coleta foi realizada por meio da API Streaming do *Twitter*, com o Twitter Capture and Analysis Toolset (RIE- DER; BORRA, 2014). A base coletada foi composta de 36.558 tuítes no período e incluiu 2.349 tuítes contendo imagem e/ou vídeo além de ao menos 7.848 tuítes com *links* externos, a maioria dos quais com imagens que são exibidas pelo próprio *Twitter* como pré-visualização das páginas referenciadas.

14 Junto com Janna Joceli Omena e Elaine Rabello, apresento uma alternativa crítica a esta perspectiva (OMENA; RABELLO; MINTZ, no prelo). Rafael Grohmann (2018) realiza uma importante revisão conceitual de engajamento. Richard Rogers (2018b), por sua vez, sugere uma abordagem crítica das métricas de engajamento, em contraposição ao que denomina *vanity metrics*.

em inglês). Mas uma imagem pode aparecer em muitas URLs diferentes se for carregada por diferentes atores ou em diferentes plataformas. Uma situação em que tal desacordo se manifesta será abordada no estudo de caso (veja [4 Estudo de caso: imagens de uma prisão](#)).

Uma ‘mesma’ imagem pode ser publicada e republicada por diferentes atores, inclusive com pequenas variações, em uma atividade difusa cujo rastreamento é dificultado pelo modo de representação informacional da imagem digital. O caso dos *memes* visuais e dos *memes* fotográficos (SHIFMAN, 2014), em particular, ilustra bem a questão. Uma ‘mesma’ fotografia serve, nestes casos, a múltiplas variações sobre o mesmo tema. Esta situação levanta uma questão de ordem metodológica: estas variações devem ser tratadas como uma ‘mesma’ imagem? Também, levanta uma questão de ordem operacional: como traçar relações de identidade ou de diferença entre estas múltiplas instâncias? Até uma ‘mesma’ fotografia, ou uma ‘mesma’ variação de um *meme* fotográfico, republicada por diferentes fontes, não será absolutamente idêntica em cada publicação. Processos de compressão da imagem aplicados pela plataforma em cada iteração (ou nos percursos da imagem para além das plataformas) gerarão resultados não idênticos a cada execução. Contrariamente ao senso comum sobre as mídias digitais, embora seja possível gerar cópias idênticas da informação, este raramente é o caso em situações do “mundo real”, pois algoritmos de compressão aplicados pelas plataformas, por exemplo, geram instâncias similares mas não idênticas dos arquivos originais. Não é possível, portanto, aferir uma vinculação absoluta entre estas diferentes instâncias, mas apenas relações probabilísticas – uma chance maior ou menor de se tratar de uma ‘mesma’ imagem. Mesmo a questão da *identidade* de uma imagem, portanto não é facilmente resolvida.

Outro desafio comumente enfrentado, diante destas imagens, é o de elaborar descrições gerais do conjunto. No caso de conteúdos verbais, técnicas como a contagem da ocorrência de palavras ou de pares de palavras (*bigramas*) oferecem recursos simples, embora limitados, evidentemente, para a sumarização. Para as imagens, este tipo de recurso não é tão simples, como discutirei, tanto em um nível técnico quanto teórico e conceitual. A identificação do conteúdo semântico de uma imagem, por exemplo, com atribuição de uma palavra representativa de seu ‘conteúdo’, não é tarefa simples. Em todo caso, mesmo quando conseguimos fazê-lo de forma minimamente satisfatória, com aplicações de reconhecimento de imagens, há muitas questões quanto a pertinência deste tipo de análise quando estamos tratando de imagens, em especial as fotográficas. A teoria da

imagem fotográfica consolidou sua compreensão como o registro *individual* e *indicial* de um acontecimento (BARTHES, 1984; DUBOIS, 2012; SONTAG, 2004), profundamente distinta de uma palavra, cujo significado tem caráter mais geral e convencional. Adiante retomarei criticamente esta vertente teórica, mas este aspecto seguirá relevante: uma imagem não é equivalente a uma ou mais de mil palavras e, por mais que possa ser descrita dessa forma, não pode ser a ela reduzida¹⁵. Práticas contemporâneas de uso das imagens tensionam esta consideração, tais como nos bancos de imagem e sua remissão alegórica a tipicidades conceituais (FROSH, 2001; MINTZ; SILVA et al., 2019). Mas tal tensionamento não autoriza que desconsideremos esta diferença.

Estudos em comunicação e mídia têm se debatido com alguns destes desafios em proposições metodológicas diversas. Uma das mais conhecidas é a chamada Analítica Cultural (*Cultural Analytics*), proposta inicialmente por Lev Manovich, que hoje atua na City University of New York (CUNY). Junto a colaboradores, Manovich desenvolveu abordagens e ferramentas para realizar leituras distantes de grandes conjuntos de imagens de tipos diversos – de páginas de mangás (MANOVICH, 2012) a *selfies* no Instagram (TIFENTALE; MANOVICH, 2015). Discutirei sua abordagem em maiores detalhes adiante (veja 3 Visualidades computacionais), mas seu aspecto característico é o tratamento das imagens como *dados culturais*, em agregados de valores cromáticos ou de brilho. Desse modo, as análises têm por objetivo descrever tendências estatísticas gerais a partir de tais dados e, tipicamente, os estudos se distanciam da imagem em sua individualidade para tratar de métricas descritivas de grandes conjuntos – reduzindo a complexidade dos *corpora* sob uma visualidade voltada aos dados descritivos das imagens¹⁶. A Analítica Cultural teve importante incidência no contexto brasileiro em pesquisas do Laboratório de Estudos sobre Imagem e Cibercultura da Universidade Federal do Espírito Santo (LABIC). Seus estudos se dedicaram especialmente ao entorno das manifestações políticas e grandes eventos esportivos de 2013 e 2014 no Brasil (HONORATO et al., 2014; HONORATO; CARREIRA; GOVEIA, 2014; MALINI et al., 2016).

Em outra vertente, o grupo interinstitucional britânico *Visual Social Media Lab* (VSML), liderado por Farida Vis, da Universidade de Manchester, pauta um olhar multiperspectivado do problema. Em 2015 o grupo realizou amplo estudo coletivo, derivando em múltiplas produções (DRAINVILLE, 2018; FAULKNER; VIS; D’ORAZIO, 2018; VIS; GO-

15 Gillian Rose (2016) elabora sobre este fator de diferenciação como a *agência* das imagens.

16 Embora esta seja uma tendência geral da analítica cultural, Alise Tifentale (2015), colaboradora de Manovich, sugere a necessidade de retorno à imagem, como também busco sustentar aqui.

RIUNOVA, 2015), em que se debruçam sobre *uma* imagem: a fotografia do corpo do garoto sírio Alan Kurdi, encontrado em uma praia na Turquia, em 2015, após tentativa de sua família de buscar refúgio na Europa. As questões lançadas pelo grupo tomam esta imagem como centro da investigação interessada em suas dinâmicas de apropriação por diferentes públicos e, também, seu papel como articuladora dos debates acerca da crise política gerada pelo afluxo de refugiados ao continente europeu. Há nesta abordagem, portanto, um interesse justamente pela *particularidade e individualidade* daquela imagem, buscando recuperar seu potencial e o caráter icônico que assume para o acontecimento, ao longo do percurso investigativo.

Outra proposição contemporânea é a do grupo *Visual Methodologies*, da Universidade de Amsterdam (UvA), liderado por Sabine Niederer. Articulando-se em torno da noção de “imagem em rede” (*networked images*)¹⁷, Niederer e seus colaboradores propõem uma compreensão da imagem articulada à gramática das plataformas de mídia social. As imagens seriam *enredadas*, neste sentido, à medida que fossem vinculadas à atividade de diferentes atores, em *curtidas*, comentários e *republicações* (NIEDERER, 2018). O grupo se interessa, deste modo, pelo desenvolvimento e sistematização de métodos analíticos e de visualização que permitam o estudo das imagens em grupo ao mesmo passo em que coloquem em relevo os aspectos visuais dos objetos estudados (NIEDERER; COLOMBO, 2019; PEARCE et al., 2018). De modo importante, suas abordagens buscam recuperar as imagens em sua materialidade visual – contrastando, assim, da indistinção dos *dados culturais* trabalhados por Manovich. A estratégia das *imagens compostas*, proposta no âmbito do grupo de pesquisa (COLOMBO, 2019), sugere estratégias de visualização de dados que proporcionem uma passagem entre os grupos de imagens e suas instâncias individuais.

Esta pesquisa busca somar esforços a estas iniciativas com um enfoque específico sobre as implicações do emprego de tecnologias de reconhecimento de imagens, baseadas em aprendizado de máquina. Técnicas deste tipo são mobilizadas em alguns estudos vinculados à perspectiva da Analítica Cultural (RYKOV et al., 2016; TIFENTALE; MANOVICH, 2015), embora boa parte das análises nesta perspectiva tendam a focar descrições estatísticas das imagens com base nos valores de cor dos *pixels*. O grupo *Visual*

17 Embora não seja mencionado por Niederer, esta formulação foi também proposta por Daniel Rubinstein e Katrina Sluis (2008), embora sem o mesmo peso para a definição de sua abordagem. A noção também nomeia centro de estudos da London South Bank University, o Centre for the Study of the Networked Image (<http://www.centreforthestudyof.net/>).

Methodologies também faz uso destas tecnologias, embora não desenvolvam uma discussão quanto a suas implicações (NIEDERER; COLOMBO, 2019). Um ponto que quero destacar, portanto, é que o uso destes instrumentos de conhecimento e descrição não é problematizado por estas pesquisas. Os estudos não discutem com maior foco ou densidade os efeitos destas técnicas na pesquisa ou aspectos de sua constituição. De certo modo, portanto, o problema permanece no plano operacional. Outro ponto importante diz respeito à articulação que os estudos realizam entre a imagem individual e os grandes conjuntos. A Analítica Cultural, como indiquei, tende a trabalhar o conjunto em detrimento das singularidades, dissolvendo as imagens em métricas estatísticas ou mesmo em massas de *pixels*. O VSML e o *Visual Methodologies* complexificam esta relação ao se voltar às duas escalas, mas até então se abstêm de tratar as escalas de forma articulada, isto é, oscilando entre o micro e o macro – ou, de outro modo, entre o *quali* e o *quanti* – que é outra ambição desta pesquisa.

Por mais que o tema da *quantidade* seja geralmente salientado como um dos desafios centrais à análise de imagens em mídias sociais, em perspectivas como a Analítica Cultural, sugiro que este problema precisaria ser também abordado *qualitativamente*. Não basta dizer que temos hoje imagens na ordem dos *bilhões* e que isto cria dificuldades operacionais aos estudos. É preciso considerar como esta transformação *quantitativa* é acompanhada de transformações de ordem *qualitativa*. Pois, como sugeri, este volume “sem precedentes” não constitui um desafio apenas à pesquisa mas também à nossa experiência destas imagens e ao papel que elas desempenham na vida social. Trata-se de questão que atravessa todo o circuito contemporâneo do visual, da produção à circulação e ao consumo das imagens, oferecendo múltiplos pontos de entrada à investigação. Como já antecipado, esta pesquisa se volta a um recorte particular deste contexto, definido pelas mediações algorítmicas que possibilitam a integração das imagens às plataformas de mídia social.

Este enfoque demanda que consideremos a tecnicidade¹⁸ (RIEDER et al., 2015; SIMONDON, 2007) tanto das imagens digitais quanto dos métodos computacionais que nos permitem estudá-las em grandes conjuntos. As técnicas investigadas nesta pesquisa são respostas à compreensão da imagem digital enquanto *inscrição* sociotécnica (AKRICH, 1992; LATOUR, 2001), o que coloca desafios computacionais específicos. Uma imagem digital é tipicamente constituída por uma sequência linear de valores numéricos que,

18 A tecnicidade das mídias sociais é tema de pesquisa desenvolvido atualmente por Janna Joceli Omena.

para a exibição, é organizada na forma de uma matriz de *pixels*. Em meio a essa massa de valores, mesmo a simples identificação de um objeto representado na imagem (um abacaxi, um rosto, uma árvore) demanda um processamento complexo. Não há nenhum segmento preestabelecido nos dados que seja pertencente a este ou aquele objeto representado. Devido à não correspondência entre a estrutura de sua representação computacional e a forma de seu “conteúdo” representacional, podemos compreender, como formula Golan Levin (2006), que as imagens são “computacionalmente opacas”. Tal *opacidade* é o que mobiliza a subdisciplina das Ciências da Computação denominada Visão Computacional, que visa desenvolver algoritmos e heurísticas para a interpretação computacional das imagens.

Pelo paradigma hoje dominante do aprendizado de máquina por redes neurais (CARDON; COINTET; MAZIÈRES, 2018), interpretações computacionais das imagens sempre produzirão representações probabilísticas, as quais elevam o grau de incerteza e o problema ontológico e epistemológico das *imagens*. Em qualquer paradigma de desenvolvimento, segundo o jargão das Ciências da Computação, diz-se que a descrição computacional de imagens é um problema “malposto” (*ill-posed*) (SMEULDERS et al., 2000). Isto significa que, independente da abordagem adotada, não haverá resposta única ao problema, mas apenas aproximações, que são validadas pragmaticamente segundo a finalidade das aplicações. Nenhum programa de computador poderia oferecer uma interpretação *definitiva* de uma imagem, não apenas porque sua representação do mundo, na forma visível, será sempre incompleta e ambígua mas, também, porque múltiplas interpretações sempre serão possíveis¹⁹. O paradigma objetivista ou positivista em que tipicamente se baseia a computação, portanto, não consegue dar conta do problema da imagem – algo que também pode ser afirmado, mas com outros matizes, à linguagem.

Portanto, superar a *opacidade computacional* não poderia levar a imagem a tornar-se, de todo, *transparente*. Seria como colapsar a imagem e subtender, de forma positivista, sua *omnitraduzibilidade*. Em outro contexto, o historiador da arte Georges Didi-Huberman (2013a, p. 11) questiona, nestes termos, a herança da iconologia de Erwin Panofsky a seu campo. A imagem ser a tudo tradutível seria, nesse caso, a premissa de que a imagem poderia ser reduzida ao conceito. Didi-Huberman (2013a, p. 163) diz de uma “tirania do conceito, da definição e, no fundo, do nomeável e do *legível*”, sugerindo uma operação

19 Smeulders et al. (2000) referem-se ao primeiro problema como “hiato sensorial” (*sensory gap*). Ao segundo, como “hiato semântico” (*semantic gap*).

reduzora da iconologia na subsunção da imagem e sua complexidade. Tomar a imagem como *omnitradutível* implica, então, apartá-la de sua própria condição de existência, como se fosse simples *intermediária*²⁰ entre intelecto e mundo. Seria, assim, uma inscrição *visível* que pudesse ser reduzida a uma condição *legível* sem que, em si mesma, produza qualquer diferença. Em última medida, elabora o autor, tal perspectiva levaria a “matar a imagem”, subjugando-a a uma condição de *certeza* (DIDI-HUBERMAN, 2013a, p. 283), quando, efetivamente, estar *diante da imagem* é estar diante de uma fonte de incerteza. Considerando a questão metodológica desta pesquisa, poderíamos compreender esta morte figurada da imagem como um efeito, também, de sua subsunção às descrições geradas por programas de reconhecimento ou, ainda, por sua dissolução em mensurações estatísticas que as tratam, irrefletidamente, como *dados visuais*. Teríamos por este caminho, novamente, um Instagram sem imagens (Figura 1).

O desafio enfrentado por esta tese, portanto, é o de evitar que o caminho seja trilhado sem a possibilidade de retorno, garantindo a possibilidade de recomposição do trajeto que leva as imagens, de uma disponibilidade ao *olhar* às inscrições geradas por seu tratamento computacional. Uma das táticas possíveis envolve redobrar a atenção a cada passo. Assim formulado, o problema da pesquisa diz respeito às complexas mediações sociotécnicas que precisam ser mobilizadas para lidar com essas imagens, sob a compreensão de que estas mediações não apenas *operacionalizam* o tratamento das imagens analisada mas, efetivamente, as transformam. Da imagem visível, oferecida ao nosso olhar, à matriz de valores de cor e, enfim, à *extração de características* ou à *classificação* por programas de aprendizado de máquina, a imagem passa por uma série de transformações. A cada passo, aquilo que é uma imagem se traduz a, progressivamente, outras formas. Indo do fenômeno visual à inscrição computacional, percebe-se que não há *como* que não transforme o *quê* da investigação. O problema enfrentado não pode ser, portanto, apenas de método. Deve ser um problema *metodológico*, que nos leve a reconsiderar a teorização sobre as imagens e a elaborar uma *crítica das ferramentas* (VAN ES; WIERINGA; SCHÄFER, 2018), em uma consideração de como condicionam os modos pelos quais podemos conhecer as imagens.

Diante deste desafio, a presente tese tem como objetivo específico investigar a aplicação de técnicas de aprendizado de máquina como recursos metodológicos para o estudo de

20 Didi-Huberman não utiliza esta palavra. Eu a introduzo aqui em referência ao sentido que lhe atribui Bruno Latour (2005) como contraponto à noção de *mediador*. Um *intermediário* seria, nesse sentido, um elemento que apenas transmite uma agência sem produzir nenhuma interferência – sem possuir uma agência própria.

imagens em plataformas *online*, voltando-se, em sua parte empírica, a um estudo de caso em que se propõe tensionar este referencial metodológico com a proposição de dispositivos de análise específicos. Em outras palavras, o método é parte substancial do objeto. Uma das motivações é a indagação acerca das potenciais transformações das imagens na sequência de translações a que são submetidas, em cada etapa do processo. Sem recusar as oportunidades analíticas introduzidas pela análise computacional, procuro considerá-las criticamente em um “trabalho de campo técnico” (RIEDER et al., 2015) atento tanto aos potenciais quanto às limitações dos procedimentos adotados. Para além do *como*, as perguntas geradoras da pesquisa assumem, principalmente, o pronome interrogativo *quê*. *O que fazemos ao estudar imagens por aprendizado de máquina? O que se tornam as imagens quando interpeladas por tais métodos? O que eles nos dão a conhecer?* Situadas no horizonte da investigação, estas perguntas orientam o estudo de caso, sem a pretensão de serem plenamente respondidas. Assumindo o estudo como uma articulação situada das técnicas e métodos de análise, objetiva-se um exercício reflexivo em que as circunstâncias específicas possam elucidar aspectos gerais do problema.

1.2 A questão do método

A elaboração de tais questões inspira-se fortemente na proposição dos *Métodos Digitais*, inicialmente formulada por Richard Rogers (2013). Trata-se de um programa investigativo que se volta ao estudo de objetos *nativos do digital* por meio de métodos também *nativos do digital*. Sua especificidade reside, portanto, em uma premissa de acoplamento entre as características tecnológico-midiáticas dos objetos estudados e também das técnicas e procedimentos empregados no estudo. Em comparação, o campo mais abrangente das chamadas Humanidades Digitais (BERRY, 2012) frequentemente se vale de técnicas analíticas computacionais para estudar objetos que não são nativos do digital – como obras literárias, obras de artes visuais e documentos históricos. Em outra comparação, estudos da Internet frequentemente se valem de metodologias não nativas do digital, como a etnografia ou a análise do discurso, que são então transpostas a ambientes digitais, geralmente em formulações híbridas como “etnografia virtual”. Os *Métodos Digitais*, de outro modo, se caracterizam pela busca de se *estudar o digital digitalmente*, por assim dizer, o que implica um outro modo de concepção de seus objetos, considerada a materialidade de sua constituição sociotécnica.

A respeito da noção de “objeto nativo do digital”, ela não deve ser compreendida na forma de uma essencialização do meio – questão que será discutida adiante (veja 2.1 [A imagem e o digital](#)). O próprio autor relativiza esta questão apontando que a “especificidade do meio” diz respeito menos à circunscrição de uma definição essencialista do digital e mais às implicações epistemológicas que os meios colocam ao estudo (ROGERS, 2013). Os objetos nativos do digital, nesse sentido, podem ser melhor compreendidos pela noção de *inscrição*, segundo elaborado por Madeleine Akrich (1992). Na formulação da autora, a pessoa que projeta um arranjo tecnológico “não apenas fixa a distribuição dos atores, mas ele ou ela também provê uma ‘chave’ que pode ser usada para interpretar todos eventos subsequentes”²¹ (AKRICH, 1992, p. 216). As inscrições, nesse sentido, oferecem algo como uma articulação empírica, situada, do princípio de relação entre formas de saber e poder analisadas por Foucault em boa parte de sua obra (cf. BRUNO, 2008, 2013; FOUCAULT, 1997). A noção também se relaciona com o que Venturini e Latour (2010) indicam por “traços digitais”, indicando como as mídias digitais produzem múltiplos registros que podem ser reaproveitados pela pesquisa para interpelar a realidade social (BRUNO, 2012). A noção de *nativo do digital*, portanto, deve ser considerada em perspectiva ampliada, observando, principalmente, as complexas mediações que conformam o objeto a cada instanciamento, mais do que sua ontologia essencial. Noutra ocasião, Rogers (2018a) enuncia este aspecto como uma *sensibilidade ao meio*, salientando a necessidade de não se ignorar a relação entre o método e o modo de constituição dos meios e objetos estudados.

Em vista destes princípios, um dos procedimentos chave dos Métodos Digitais é a “reapropriação” (no inglês, *repurposing*) de ferramentas computacionais e dos objetos digitais estudados em sua materialidade específica. Estes elementos são tomados criticamente como instâncias metodológicas. Por exemplo, o mecanismo de busca da *Google* é recorrentemente mobilizado por diferentes estudos como instância de mapeamento de determinado tema. Porém, longe de tomá-lo ingenuamente como uma *janela* neutra para a realidade estudada, a perspectiva dos Métodos Digitais busca realizar uma investigação que se volta simultaneamente aos *objetos digitais* e aos próprios *métodos* empregados para descrevê-los – movimento relacionado, como indiquei acima, ao princípio da *simetria*, em sua releitura por Marres e Moats (2015). Ferramentas como a busca da *Google*, portanto, são tomadas como mais do que um método para inquirir a realida-

21 No original: “... not only fixes the distribution of actors, he or she also provides a ‘key’ that can be used to interpret all subsequent events”. Tradução minha.

de. Eles são efetivamente inseridos como objetos das investigações, enquanto mediadores ativos da investigação e, também, da experiência de navegação da Internet (cf. RIEDER, 2012; RIEDER; SIRE, 2013).

Os Métodos Digitais são devedores da perspectiva teórico-metodológica dos STS (cf. VENTURINI; MUNK; JACOMY, 2018). Diante da crescente complexidade das mediações tecnológicas envolvendo os processos comunicacionais contemporâneos, os STS oferecem importantes subsídios conceituais e metodológicos para os estudos voltados a plataformas de mídia social, na medida em que investigam a constituição social da ciência e da tecnologia juntamente à constituição científica e tecnológica da sociedade. Os Métodos Digitais e os STS permitem a adoção da perspectiva dupla, como afirmei, volta-tanto às imagens em plataformas *online* quanto às mediações sociotécnicas que as conformam. Esta compreensão do problema se fundamenta, portanto, na não distinção, de antemão, entre práticas sociais e as materialidades técnicas que participam, fundamentalmente, de sua constituição. No âmbito desta pesquisa, além da sua incidência sobre os Métodos Digitais, os STS serão também mobilizados a partir da Teoria Ator-Rede (TAR), uma de suas vertentes mais conhecidas, formulada inicialmente por Michel Callon, Bruno Latour e John Law, nos anos 1980, com derivações e desdobramentos nas décadas subsequentes.

John Law (2017, p. 47) elabora que, pela perspectiva dos STS, os métodos não devem ser compreendidos simplesmente como técnicas, mas como um arranjo materialmente complexo de práticas que articulam entidades heterogêneas, como sujeitos, objetos, imaginários e instituições. Longe de dar acesso a uma realidade estável, portanto, esse arranjo descreve “espaços de conhecimento” (“*knowing spaces*”, no original em inglês), que definem fronteiras mais ou menos permeáveis entre o que é possível e o que é impossível de se conhecer. Há, desse modo, uma performatividade do método, que produz diferença no interior de seu próprio objeto (LAW, 2017, p. 45). Tal consideração, evidentemente, levanta questões de ordem ontológica, já que se o método produz diferença, a realidade é lançada em uma condição de instabilidade. Esta é, contudo, a circunstância que descrevo acerca do problema lançado aos estudos das imagens. Situação esta que, segundo defende John Law (2004), em perspectiva almejada por esta pesquisa, não deve ser simplificada mas, sim, abraçada pela pesquisa.

Em resposta a esta instabilidade ontológica, o procedimento que adoto nesta investigação pode ser compreendido pela ideia de *ontografia*, segundo propõe Michael Lynch (2013), como forma de *desinflar* a ontologia. Isto porque uma premissa dos STS envolve justamente o ceticismo quanto a afirmações de verdades sobre o mundo sem que estas sejam vinculadas às práticas que permitem revelá-las e que, em última medida, participam de sua constituição e manutenção. Segundo elabora Law (2017, p. 43), “ontologias são efeitos relacionais que emergem em práticas [...] [e] como práticas variam, também variam os objetos”²². Uma reflexão metodológica, como a que proponho fazer, talvez precise se haver com este ponto ao se indagar o que é, mesmo, que estamos fazendo ao observar as imagens de um ou de outro modo. Que imagem estamos, afinal, *produzindo* com nossa prática de saber? Importante ressaltar, especialmente em momentos de ataques contra o saber científico, que isto não implica um relativismo absoluto. Trata-se, de outro modo, do exercício de uma compreensão complexa das realidades produzidas pelos métodos, cuja validade depende justamente da força e consistência das práticas que as constroem²³. A perspectiva da *ontografia* visa, portanto, dar maior peso às formas de conhecimento empregadas, justamente ao reconhecer a necessidade de evitar distinções duras entre *ontologia* e *epistemologia* (LYNCH, 2013). Segundo a formulação de Lynch, a *ontografia* compreende uma descrição situada e não essencialista dos objetos, assumindo novamente uma perspectiva *simétrica* mas, desta vez, entre identidade e diferença.

Gabriel Menotti (2019), embora sem se referir a Lynch, mobiliza a noção de ontografia para fraturar a noção unitária do *dispositivo* cinematográfico. Tomo sua discussão como um ponto de apoio para também compreender a necessidade de revisar compreensões essencialistas da dimensão técnica das imagens. Segundo sugere Menotti (2019, p. 15), o dispositivo deve ser compreendido sempre no *plural* mesmo quando tratado no singular. Embora o dispositivo seja frequentemente subsumido por instâncias específicas da

22 No original: “*ontologies are relational effects that arise in practices [...] [and] since practices vary, so too do objects*”. Tradução minha.

23 Esta discussão ultrapassa o escopo desta pesquisa, mas como se trata de tema hoje *quente*, gostaria de fazer breves indicações. Em tom de revisão, Latour (2013, p. 6), identifica uma virada recente no discurso científico que não se basearia mais nos fatos (*matters of fact*) ou na racionalidade para se defender da negação da ciência, como no caso das mudanças climáticas. Em contraste, a defesa passaria hoje a se basear em um discurso de *confiança* na *instituição* da ciência. Latour reflete que cientistas não apenas obscureceram as condições que garantiriam esta possibilidade de confiança quanto combateram os estudos do STS que se propuseram a investigar as práticas que garantiriam solidez à instituição. Entretanto, reconhece que o tempo deste debate já teria passado e estaríamos todos diante do desafio comum de defender os diagnósticos da ciência sobre o clima. Mais recentemente, Latour (2018, p. 23) elabora: “Nenhum conhecimento comprovado pode se sustentar sozinho, como sabemos bem. Fatos permanecem robustos apenas quando eles são suportados por uma cultura comum, por instituições que podem ser confiadas, por uma vida pública mais ou menos decente, por uma mídia mais ou menos confiável”. No original: “*No attested knowledge can stand on its own, as we know very well. Facts remain robust only when they are supported by a common culture, by institutions that can be trusted, by a more or less decent public life, by more or less reliable media*”. Tradução minha.

constelação que efetivamente constituem nossa compreensão do *cinema* (tais como sala de projeção, câmera etc.) (cf. BAUDRY, 1983), Menotti argumenta que este tipo de compreensão tende a priorizar determinadas instâncias em detrimento de outras e, assim, configurar uma visão parcial do meio. Por exemplo, tal perspectiva unitária tende a desconsiderar a importância da pirataria como parte da dinâmica de circulação contemporânea dos filmes e, portanto, da própria constituição do *cinema*. No caso desta tese, como sugeri, cada etapa do processamento computacional de uma imagem, para os fins da análise, implica uma transformação do seu estatuto ontológico. Nesse sentido, não haveria como reivindicar uma ontologia estável atravessando todo o processo – a imagem que se analisa ao final não pode ser tomada como equivalente àquela observada ao começo. Neste sentido, o que proponho é uma *ontografia* da imagem digital nos processos de translação a que é submetida ao longo da investigação. Espero que esta abordagem contribua para adensar aplicações de aprendizado de máquina aos estudos das imagens.

Dado o caráter necessariamente situacional deste tipo de abordagem, baseio-me em um estudo de caso como instância a partir da qual aspectos gerais poderiam ser elaborados. Nesse sentido, esta tese renova as apostas feitas no período de formação dos STS (cf. BIJKER; LAW, 1992), quando se buscava um caminho intermédio entre a contingência das narrativas singulares, no âmbito dos estudos históricos, e a busca por padrões e grandes sistematizações da perspectiva sociológica. Gostaria, nesse sentido, de também assumir o compromisso de me “debater com o desconcerto entre a exploração de estudos de caso em seu desalinhamento e a tentativa de construir modelos em alguma medida mais gerais ou modos de pensamento sobre a formação social da tecnologia”²⁴ (BIJKER; LAW, 1992, p. 7). Algo que, sugerem os autores, depende de modelos descritivos “empiricamente sensíveis”. Por esse motivo, para além da reflexão teórica e metodológica geral, esta tese tem, como ponto de chegada, reflexões desenvolvidas em um “corpo a corpo” com os objetos analisados e os métodos empregados. Especificamente, como ponto de confluência das discussões empreendidas neste trabalho, apresenta-se um estudo de caso voltado aos desafios metodológicos colocados para uma investigação interessada nas imagens que circularam no Twitter durante o acontecimento da prisão do ex-presidente Luiz Inácio Lula da Silva, em abril de 2018.

24 No original: “wrestle with the trade-off between the exploration of messy case studies and the attempt to built [sic] somewhat more general models or ways of thiking about the social shaping of technology”. Tradução minha.

1.3 Estrutura da tese

Após esta introdução, o texto se divide em três capítulos de desenvolvimento, seguidos das considerações finais. O capítulo 2 (*Imagem-rede*) busca estabelecer uma base teórica e conceitual para a compreensão das imagens segundo um princípio de “materialidade relacional” (LAW, 1999). Parte-se de uma revisão da compreensão das imagens digitais pelas teorias da imagem, buscando desenvolver uma crítica da hipótese de imaterialidade da imagem digital e, em especial, da desconsideração da constituição técnica como elemento pertinente ao seu tratamento teórico. Em seguida, uma via alternativa é sugerida a partir de elementos teóricos e conceituais dos STS e da Teoria Ator-Rede (TAR), em particular. Três eixos estruturam esta elaboração: o tema da *mediação técnica* (LATOUR, 2001, 2005), como forma de compreender o caráter distribuído das agências que participam na formação da imagem; o conceito de *inscrição* (AKRICH, 1992; AKRICH; LATOUR, 1992; LATOUR, 1986, 2001), em sua dimensão sociotécnica, a fim de situar as imagens em relação aos agenciamentos que as produzem mas, também, para compreender as implicações de deslocamentos a outros arranjos sociotécnicos; e a noção de *multiplicidade ontológica* (LAW, 2017; MOL, 1999), como forma de compreender as diferentes instanciações materiais de uma imagem. Em uma terceira seção, busco elaborar a hipótese conceitual desta tese, denominada *imagem-rede*. Primeiro, a abordagem da imagem pela via dos STS e da TAR é aproximada a uma leitura interessada do trabalho do historiador da arte Aby Warburg, do início do século XX. Aspectos de uma materialidade relacional das imagens são postos em relação com a noção de “veículos de imagem”, de Warburg (2015), e com seu projeto inacabado do *Atlas Mnemosyne*. A noção de *imagem-rede* é proposta, então, como forma de descrever as imagens como efeitos emergentes de formações distribuídas – noção que atravessará as discussões elaboradas na tese.

O capítulo 3 (*Visualidades computacionais*) volta-se mais especificamente ao tratamento computacional da imagem como um fator de sua multiplicidade ontológica e, especialmente, como elemento constituinte de “modos de ver” contemporâneos. A discussão enfoca, primeiro, a relação fricativa entre o *visual* e o *computacional*. A questão é articulada, em particular, em uma revisão crítica da Analítica Cultural (MANOVICH, 2009), vertente de estudos que propõem uma abordagem quantitativa das imagens como “dados culturais”. O principal argumento elaborado em contraste com a proposição de Manovich identifica em sua redução da imagem a um dado computacional aspectos do que Van Dijck (2014) elabora, criticamente como dataísmo: a suposição de

que toda atividade social pudesse ser mensurável e tradutível a *dados*. Em um segundo movimento, a discussão se volta ao aprendizado de máquina baseado em redes neurais como uma forma contemporânea de transladar o visual ao computacional. As máquinas indutivas (CARDON; COINTET; MAZIÈRES, 2018) constituídas por tal tecnologia são consideradas pelo desafio que colocam à inteligibilidade de seus procedimentos e segundo o modo com que “aprendem a ver”, com a elaboração de modelos preditivos a partir de grandes bases de dados. O aprendizado de máquina também é abordado pela natureza dos *vectores*, as inscrições que esta técnica produz a partir dos dados processados e que manipula em uma espacialidade abstrata, multidimensional. Ainda outro aspecto discutido a seu respeito são as reconfigurações humano-máquina (SUCHMAN, 2007) decorrentes da redistribuição agencial provocada no entorno de tais tecnologias. Na linha da proposição de Adrian Mackenzie (2017), sugiro a compreensão de *aprendizes maquínicos*, observando práticas de conhecimento distribuídas e compartilhadas entre humanos e máquinas. Por fim, a terceira seção do capítulo se volta a aspectos das dependências infraestruturais do aprendizado de máquina e como isto, para o caso específico das aplicações de Visão Computacional, descreveria uma tendência de *infraestruturalização* (PLANTIN et al., 2016) de determinados modelos e bases de treinamento. Esta situação levaria à prevalência de *visualidades computacionais* específicas que, inclusive reproduziriam assimetrias políticas e sociais em seus modos de ver (BUOLAMWINI; GEBRU, 2018; MINTZ; SILVA et al., 2019; SILVA, 2019). Estes aspectos são tensionados, então, em uma discussão do gesto de *reapropriação* destes modelos como recursos metodológicos para o estudo de imagens, em diálogo com o campo dos Métodos Digitais (RIEDER; RÖHLE, 2017; ROGERS, 2013).

O capítulo 4 (*Estudo de caso: imagens de uma prisão*) desenvolve uma discussão metodológica situada a partir das imagens que circularam no Twitter na época da prisão do ex-presidente Luiz Inácio Lula da Silva, em abril de 2018. São processados 7,3 milhões de tuítes publicados entre os dias 4 e 16 abril, desde a véspera da expedição do mandado de prisão até o décimo dia de sua efetivação, ocorrida no dia 7 de abril. Entre os motivos da escolha do caso, será ressaltada a relevância histórica do acontecimento e sua repercussão midiática em uma disputa pela produção de imagens representativas da prisão. Estes aspectos salientam tanto a relevância de uma abordagem centrada nas imagens quanto a dinâmica relacional que se estabelece entre as diferentes figurações produzidas. Centralmente, porém, busca-se discutir os procedimentos metodológicos

demandados para o desenvolvimento de um estudo voltado a tais aspectos. Em remissão aos debates teóricos dos capítulos precedentes, o objetivo é verificar o rendimento das noções de *imagem-rede* e *visualidade computacional*, tomadas como eixos conceituais desta tese. Como um ponto de confluência destas discussões, o capítulo apresenta um dispositivo metodológico desenvolvido a fim de proporcionar uma navegação heurística entre as imagens do acontecimento, denominado *Atlas para imagens-redes* (veja [4.3 Compondo imagens-redes](#)). Instruções de acesso e utilização são indicadas no [Apêndice B](#). O *Atlas* apresenta cerca de 18,4 mil imagens extraídas dos tuítes, dispostas a partir de dados extraídos por meio de dois modelos de reconhecimento de imagens baseados em aprendizado de máquina: VGG19 (SIMONYAN; ZISSERMAN, 2014) e API Cloud Vision da Google (GOOGLE, 2017). Ele também reúne métricas da circulação das imagens e alguns dos tuítes que as compartilharam. Conforme a discussão realizada adiante, o *Atlas* busca exercitar, em uma situação concreta, os desafios teóricos e metodológicos elaborados nesta tese, em um esforço de composição (LATOUR, 2010) não reducionista da multiplicidade ontológica dos objetos estudados.

As [Considerações finais](#) sintetizam, por fim, os principais argumentos desenvolvidos ao longo da tese e propõem respostas possíveis às perguntas geradoras da investigação.

2 Imagem-rede

Um primeiro desafio posto a esta pesquisa diz respeito ao tratamento teórico que será concedido à imagem e, de modo mais específico, à sua materialidade técnica. Este enfoque, por si só, demanda um esforço de definição conceitual, pois a imagem possui um problema ontológico de base que tende a se complexificar no contexto contemporâneo. Como sugere Emmanuel Alloa (2015, p. 7), é como se a crescente exposição às imagens a que somos submetidos fosse inversamente proporcional à nossa capacidade de descrever o que elas efetivamente são. Um complicador fundamental, que independe das inflexões contemporâneas, é a dificuldade em se estabelecer o lugar de sua ontologia. Sua condição de *coisa*, como descreve didaticamente Lavaud (1999, p. 13), tende a ser colocada entre parênteses para dar lugar ao fenômeno de sua percepção ou à coisa representada. Nesse sentido, as imagens parecem se apresentar como intermediárias fugazes que apenas teriam relevância segundo fatores que, efetivamente, as transcendem²⁵. Entre seus múltiplos estatutos, como *coisa*, *referente*, *percepção* e, ainda, como *desejo*, a questão do que é *uma imagem* é, por si só, desafiadora. No contexto de sua digitalização, o problema é amplificado, já que nem mesmo os contornos materiais que a definiriam como *coisa* não parecem, à primeira vista, ser assim tão claros.

Alloa elabora que a questão poderia estar mal colocada pois, longe de ser *uma*, “a imagem tende a se disseminar, declinar-se dela mesma em formas plurais e desmultiplicar em um devir-fluxo” (ALLOA, 2015, p. 7). Tudo pareceria indicar, nesse caminho, que a questão da *materialidade* da imagem, de seu ser enquanto *coisa*, estaria no cerne da mal colocação da questão. Como se fosse sua aparente *estabilidade* enquanto objeto que permitisse esta interrogação, contrastando com a experiência e o modo de significação das imagens, que sugerem, de outro modo, que a *imagem-coisa* seria apenas a manifestação tangível de um fenômeno múltiplo. No entanto, sem recusar a complexidade de tais movimentos, quero argumentar que mesmo a materialidade da imagem não deve ser tomada como estável ou determinada. Por um lado, é próprio à imagem constituir-se de forma relacional, seja entre si e um olhar externo, seja na remissão ao acontecimento ou

25 Descreveu, em dado momento, Jacques Aumont: “Com todo rigor, a *parte da imagem* pode ser completamente atribuída a um ou outro dos agentes da história social das imagens. Se a isolamos aqui, de modo um pouco artificial, é por pura comodidade, para apresentar um conjunto de pesquisas sobre a representação que, de certa forma, consideraram a imagem como dotada de valores imanentes” (AUMONT, 2002, p. 197).

objeto ausentes. Por outro, a questão que levanto é que mesmo em sua constituição material a imagem não se reduz a uma condição de imanência. Não é por a tratarmos, transitoriamente, como coisa ou, no caso da imagem digital, como dado informacional, que seria possível circunscrever sua ontologia. De outro modo, o caráter relacional que fundamenta a compreensão fenomenológica, semiótica ou psicanalítica da imagem deve se realizar, também, em um movimento que conecta a imagem-coisa às suas dependências de ordem material, técnica e infraestrutural. Neste sentido, meu argumento é de que não há um artefato que possa ser considerado individualmente, senão pelo gesto de um corte. Isto não só se aplica como se complexifica com as imagens digitais, colocando-se, inclusive, como um dos aspectos centrais do problema metodológico a que se volta esta pesquisa.

Uma primeira explicação para esta compreensão da materialidade dos meios poderia ser apresentada na linha do que vêm discutindo autores como Jussi Parikka (2015) e Sean Cubitt (2014, 2017). Segundo elabora Marcio Telles (2016) acerca de Parikka, teríamos na linha destes autores algo como uma passagem da *materialidade* às *matérias-primas* da comunicação. São proposições que se voltam concretamente às matérias que constituem as tecnologias midiáticas. Estas, salientam, não surgem do nada e têm uma temporalidade que deve ser remetida a uma escala geológica. Dos pigmentos da pintura aos minerais da indústria tecnológica contemporânea, as imagens seriam produtos que dependem de amplas cadeias de extração, transformação e descarte, as quais se inserem fortemente em disputas geopolíticas, inclusive como parte de sua valoração estética. Cubitt (2014, p. 118–119) destaca como na Renascença, por exemplo, o alto custo político e econômico do pigmento para a cor azul ultramar, proveniente do Afeganistão, fazia com que ele fosse reservado para a representação de figuras de alto valor simbólico no período, como a Virgem Maria. Sobre o contexto contemporâneo, Parikka (2015) e Cubitt (2017) destacam as dependências da cadeia produtiva das mídias digitais, com dinâmicas neocoloniais de exploração do trabalho e dependências de minerais e de produção de lixo ambientalmente desastrosos. Mesmo para o caso das imagens digitais, portanto, sua imaterialidade é um *mito*²⁶. Longe de supor uma ontologia estável, a dimensão da materialidade parece se situar justamente no campo do impensado da imagem²⁷, como uma

26 Como Christianne Paul (2007) já havia alertado, do ponto de vista da preservação de obras de arte digital.

No enfoque específico que assumirei aqui, o tratamento da imagem enquanto *coisa* não se voltará propriamente à matéria-prima das imagens, mas sim a como elas se integram a arranjos sociotécnicos complexos, como condicionantes de seu modo de *inscrição*. A relacionalidade da constituição material das imagens será considerada, então, de modo próximo ao que John Law sugeriu, em uma revisão da Teoria Ator-Rede (TAR), como uma “materialidade relacional” (LAW, 1999) ou uma “semiótica material” (LAW, 2009). A TAR, escreve Law: “toma a compreensão semiótica da relacionalidade das entidades, a noção de que elas são produzidas em relações, e a aplica impiedosamente a todos os materiais – e não apenas àqueles que são linguísticos”²⁸ (LAW, 1999, p. 4). Para o caso da imagem digital no contexto das plataformas *online*, esta relacionalidade deverá incluir aspectos dos modos de inscrição dessa imagem e os processos computacionais que participam de sua formação, transformação e circulação. Evidentemente, as imagens possuem um apelo semiótico no seu sentido mais tradicional. No entanto, o olhar informado por uma perspectiva mais claramente material implica que consideremos outros modos de sua relacionalidade que ultrapassam a questão do significado ou da semiose, em seu sentido mais estrito, para lidar com sua *ontologia*.

Nas teorias da imagem derivadas do cinema e da fotografia, este aspecto não é ignorado, sendo tipicamente articulado na noção de *dispositivo*. Geane Alzamora, Joana Ziller e Carlos d’Andréa (2018), voltando-se a um contexto mais amplo de aplicação do conceito, salientam a complexidade de se circunscrever uma definição precisa. A proposição mais proeminente seria aquela de Foucault, que mobiliza o *dispositivo* de forma difusa em sua obra, em especial em sua fase genealógica. Em seu pensamento, a noção visa nomear um conjunto ou uma rede de práticas que envolve entidades heterogêneas, materiais e discursivas (FOUCAULT, 1979, p. 244). Esta rede de relações configura, de modo articulado, formas de saber e de poder, socialmente disseminadas, em operações prescritivas e posicionais que não se caracterizam tanto pela censura mas, principalmente, pela positividade (ALZAMORA; ZILLER; D’ANDRÉA, 2018). Isto é, seria mais pelo que permitem do que pelo que proíbem que poderíamos compreender o modo de operação dos

27 Michel Callon, em texto que antecede, em muitos aspectos, sua proposição da teoria ator-rede, sugere que a dinâmica da constituição de problemas tecnocientíficos depende da distinção entre um campo de incertezas, submetido à análise, e um campo de certezas, que não será analisado. A respeito deste último, afirma: “sua estrutura se assemelha àquela do inconsciente. Ela representa o que é silenciado para que o resto possa ser afirmado” (CALLON, 1980, p. 213). No original: “its structure resembles that of the unconscious. It represents what is kept silent so that the rest may be stated”. Tradução minha.

28 No original: “takes the semiotic insight, that of relationality of entities, the notion that they are produced in relations, and applies this ruthlessly to all materials – and not simply to those that are linguistic”. Tradução minha.

dispositivos. Entretanto, muito embora tenha Foucault ao centro, a noção de dispositivo lhe antecede em outras concepções²⁹ e, também, lhe sucede no trabalho de comentadores (cf. AGAMBEN, 2009a; DELEUZE, 1999, 2006).

Especificamente no campo das imagens, a noção de *dispositivo* é anterior à consolidação da acepção foucaultiana, inclusive em denominações distintas como *aparelho* ou *aparato*. Um de seus principais antecedentes encontra-se na crítica de base marxista do *dispositivo cinematográfico*, também denominado *aparato cinematográfico*, metonimicamente referenciado pela sala de projeção (BAUDRY, 1983). Nos anos 1980, já sob influência de Foucault, mas sem dispensar as demais correntes de formulação, a noção adquire importância na reorientação pós-estruturalista das teorias da imagem, em especial no entorno da fotografia. Em artigo de revisão, Philippe Dubois (2017) salienta como a compreensão da *especificidade* da fotografia desenvolvida à época, em relação a outros estatutos da imagem, se baseava fundamentalmente na essencialização de seu *dispositivo*, o qual vinculava-se fortemente à constituição técnica da câmera e seu condicionamento da gênese das imagens. Com base nesta compreensão, autores como o próprio Dubois articularam uma teoria fundada na *indicialidade* e na *referencialidade* como aspectos distintivos da categoria do *fotográfico*, descrevendo imagens com vinculação ontológica com o real.

No âmbito destas teorias, a noção de dispositivo – que combina, em alguma medida, as duas heranças – demonstra-se produtiva para a articulação conceitual dos efeitos coletivos produzidos por elementos dispersos. Para o caso da fotografia, o efeito de indicialidade é compreendido como derivado da atuação conjunta de processos físico-químicos da película e do processo de revelação; da constituição óptico-mecânica da câmera; e, também, de uma elaboração discursiva que reconhece nesse registro a condição de inscrição do mundo ou de “emanação do real”, com efeitos ideológicos e psíquicos (AUMONT, 2002; DUBOIS, 2012). No entanto, como Dubois (2017) reconhece em retrospectiva, o tratamento conferido ao dispositivo na teoria da fotografia tendeu à sua *essencialização*, tomando-o, por vezes, como uma unidade.

29 Alzamora, Ziller e d’Andréa indicam, a partir de Raffnsøe, Gudmand-Høyer e Thaning (2014, citado por ALZAMORA; ZILLER; D’ANDRÉA, 2018), a existência de formulações anteriores de Louis Althusser, Jean-François Lyotard e Louis Baudry. RAFFNSØE, Sverre; GUDMAND-HØYER, Marius & THANING, Morten S. What is a dispositive? Foucault’s historical mappings of the networks of social reality (2014). Disponível em: <http://foucaultnews.com/2015/01/10/foucaults-dispositive/>. Acesso em fev. 2016. Outras formulações precursoras podem ser observadas em Dubois (2012, p. 317–318), que se baseia na noção de *aparelho* em Sigmund Freud, a qual toma como equivalente a dispositivo, para tratar dos efeitos psíquicos da fotografia. Ou em Aumont (2002), que se refere principalmente a Christian Metz, para os efeitos psíquicos do dispositivo cinematográfico e a Jean-Louis Baudry e Jean-Louis Comolli, para os efeitos ideológicos.

Em consequência, conforme argumentarei na primeira seção deste capítulo, a vertente de teorização da imagem fotográfica desenvolvida naquele momento enfrentou dificuldades no tratamento da imagem digital. A suposição de uma ontologia essencialista condensada em um dispositivo da imagem digital, a partir de uma transposição do modelo teórico elaborado para a fotografia, revelou-se míope diante das implicações trazidas pela digitalidade às imagens.

Na segunda seção do capítulo, busco delinear um caminho alternativo. Inspirado em aspectos da TAR, reviso a concepção teórica da mediação técnica e compreendo as imagens como *inscrições* sociotécnicas (AKRICH, 1992) em atendimento às prescrições dos sistemas técnicos em que se encontram integradas e, também, a demandas institucionais e políticas. Esta compreensão aponta, em particular, para uma abordagem teórica das imagens como uma ontologia instável e *múltipla* (MOL, 1999). Contrariamente à essencialização do dispositivo como condição determinante do modo de existência das imagens, o vocabulário teórico e a sensibilidade empírica que são constituídos pela TAR mostram-se capazes de descrever as múltiplas instanciações da imagem, segundo a rede de mediações técnicas e as diferentes *translações* (CALLON, 1980, 1984) a que é submetida. Desse modo, sugere-se uma teorização da imagem pelas suas dependências, no sentido da *materialidade relacional* indicada por John Law (1999).

Em um terceiro movimento, apresento a proposição da *imagem-rede*, tomada como hipótese conceitual desta tese. Como elaborarei, trata-se de uma formulação que visa reconhecer a tensão colocada pela multiplicidade ontológica das imagens. Busco dar um passo além de tentativas de resolver tal tensão e argumento que uma tal elaboração teórica é demandada pela condição digital contemporânea, ao colocar em crise a estabilidade ontológica outrora conferida à condição da imagem como *coisa*. Porém, também argumento que esta consideração não se limita ao presente. Como forma de desenvolvimento desta proposição, realizo uma aproximação livre entre a noção de *imagem-rede* e alguns aspectos da obra do historiador da arte Aby Warburg, cuja renovada relevância em períodos recentes parece se dever, como sugere Maurício Lissovsky (2014), à pertinência de suas observações para o contexto contemporâneo de deriva das imagens. Em particular, retomarei sua noção de “veículos de imagem” (*Bilderfahrzeuge*) e o seu projeto inacabado do *Atlas Mnemosyne*. Estes aspectos de sua obra serão tomados como indicações importantes acerca da materialidade da circulação das imagens e também das práticas desenvolvidas para conhecê-los. Busco, assim, verificar o rendimento

da proposição teórica da imagem-rede em contextos anteriores ao digital e mesmo ao fotográfico. Esse olhar para o passado também visa compreender possíveis aspectos transversais que o conectariam ao presente.

2.1 A imagem e o digital

Uma vertente dominante das teorizações produzidas a respeito da imagem digital nos anos 1990 e 2000, como discuto a seguir, salientou seus aspectos de simulação, virtualidade e perda do referente. Estas foram, de fato, algumas das principais “novidades” inauguradas pela tecnologia digital – “informática” – em um primeiro momento, quando câmeras fotográficas digitais, celulares com câmeras e a internet ainda não eram muito difundidas e a imagem digital a que se referia era, majoritariamente, a da computação gráfica. As imagens de síntese, como foram nomeadas, eram produzidas não mais a partir do gesto criativo manual ou por uma inscrição fotoquímica do mundo mas, sim, por meio do cálculo computacional. Em retrospecto, contudo, percebe-se um sobredimensionamento destes aspectos, em especial por sua elevação como traços definidores da tecnologia digital. Ao focar a ruptura, as abordagens teóricas desenvolvidas no período acabaram por perder de vista o que se oferecia como continuidade e, de modo mais importante, as diferenças que eram ocultadas pelo tratamento monolítico do *digital*. Sob esta perspectiva, um dos principais fatores que contribuíram a esse modo de compreender a imagem digital parece se dever não apenas às possibilidades tecnológicas da época mas, de modo mais importante, à tradição teórica que se desenvolveu na década anterior, nos anos 1980, no entorno da imagem fotográfica. As teorizações sobre a imagem digital refletem, em larga medida, as premissas teóricas desenvolvidas naquele período.

Em artigo de revisão, Philippe Dubois (2017) descreve os anos de 1980 como um período de efervescência em que a noção de *imagem* assumiu seu sentido mais amplo de “regime de visualidade” (DUBOIS, 2017, p. 34). Esta formulação ganhou corpo no entorno da constituição teórica da fotografia e da categoria derivada do *fotográfico*. O autor delimita a década da teoria francesa com dois marcos simbólicos, iniciando pela publicação, em 1980, de *A Câmara Clara* de Roland Barthes (1984) e concluindo com a publicação, em 1990, de *O fotográfico*, de Rosalind Krauss (2002). Outras obras-chave da periodização incluem as de Susan Sontag (2004), publicada na década anterior, em 1977, e de Jean-Marie Schaeffer (1996), publicada em 1987. Naquele mesmo período, em 1983, o próprio

Dubois publica seu influente *O ato fotográfico* (2012). Segundo elabora o autor, os teóricos daquela geração tinham em comum a busca por uma superação da abordagem semiológica estruturalista e, para isso, propunham uma visada em torno da *especificidade* da fotografia e de seu *dispositivo* técnico. Desenvolveram, assim, um pensamento *ontologizante* que buscava um “pensar próprio às imagens” (DUBOIS, 2017, p. 37). Este se realizava em uma vinculação profunda entre a imagem e o real e em uma perspectiva de irredutibilidade do visual à língua e sua racionalidade.

Entre os textos precursores retomados por aquela geração, há o conhecido “Ontologia da imagem fotográfica” publicado em 1945 pelo crítico de cinema André Bazin (2014). Notoriamente, Bazin destacara diferenças fundamentais entre a imagem fotográfica e a pintura, sob a chave da *objetividade essencial* que lhe conferia, justamente, a máquina: “Pela primeira vez, entre o objeto inicial e sua representação nada se interpõe, a não ser outro objeto. Pela primeira vez, uma imagem do mundo exterior se forma automaticamente, sem a intervenção criadora do homem, segundo rigoroso determinismo” (BAZIN, 2014, p. 31). A questão ontológica sobre a imagem era colocada, portanto, pela singularidade de uma imagem que era produzida pelo próprio mundo *não humano*, por assim dizer, em um distanciamento da subjetividade. Por um lado, este distanciamento entre a imagem e a “intervenção criadora do homem” ressoa a apropriação da gênese da imagem que no cristianismo, como descreve Marie-José Mondzain (2015), se deslocou da mão humana à mão de Deus, como no sudário de Turim. Pela constituição técnica da fotografia, contudo, a substituição não supunha uma divindade mas, de outro modo, a noção moderna de *objetividade*³⁰. Tratava-se, nesse sentido, de uma inscrição que a própria realidade objetiva do mundo realizava-se em si mesma.

Outra referência central à teoria fotográfica do período foi a semiótica de Charles S. Peirce. Em particular, a categoria do *índice*, segunda instância da tríade pela qual Peirce descreveu as relações entre os signos e seus objetos. O *índice* define-se por uma “correspondência de fato ou relação existencial” (SANTAELLA, 2001, p. 21) entre signo e objeto. Sua aplicabilidade à fotografia deve-se a que, diferentemente das imagens provenientes do gesto criador da mão humana, a fotografia pode ser compreendida como efeito de um processo químico, óptico e mecânico. Este aspecto, segundo desenvolveu Dubois (2012)

30 Lorraine Daston e Peter Galison (2010) desenvolvem como aspectos desta produção das imagens foram importantes na constituição do ideal de objetividade científica no século XIX.

em sua influente formulação, sobressaía até mesmo à relação de semelhança entre a imagem e o mundo. Mais do que um espelho do real, a imagem fotográfica seria, assim, um *traço* do real:

O ponto de partida é portanto a natureza técnica do processo fotográfico, o princípio elementar da *impressão luminosa* regida pelas leis da física e da química. Em primeiro lugar, o traço, a marca, o depósito [...]. Em termos tipológicos, isso significa que a fotografia aparenta-se com a categoria de “signos” em que encontramos igualmente a fumaça (indício de fogo), a sombra (indício de uma presença), a cicatriz (marca de um ferimento), a ruína (traço do que havia ali), o sintoma (de uma doença), a marca de passos etc. (DUBOIS, 2012, p. 50).

A imagem sob o *fotográfico* (enquanto um regime de visualidade) definia-se, assim, pela *referencialidade*, elevada à condição de essência pela conjunção entre o dispositivo técnico e a chave de leitura que ele proveu ao modo de significação da imagem.

Segundo elabora Dubois, em revisão, teria sido por causa deste enfoque ontológico que o surgimento da imagem digital se fez sentir como um corte “dramático” ou “trágico” entre a imagem e o mundo (DUBOIS, 2017, p. 42). Entre vertentes celebratórias (cf. COUCHOT, 2003; QUÉAU, 1993) e críticas (cf. VIRILIO, 1993, 1994) indicadas pelo autor, encontramos acepções em que a digitalidade da imagem se resume a aspectos da virtualidade e da simulação, sempre em vista do debate sobre a referencialidade, tomada como aspecto *essencial e ontológico* da fotografia e, por conseguinte, da *imagem* enquanto objeto teórico. Desse modo, a visada essencialista, focada na especificidade do *fotográfico*, acabou por enviesar a abordagem teórica da imagem digital. Esta limitação se fez sentir, em especial, em um tratamento monolítico do digital, que não apenas ignorou modos de sua diferenciação interna quanto supuseram um movimento de profunda ruptura, perdendo de vista as transformações graduais provenientes da digitalização da imagem. O foco voltava-se ao que se introduzia de novidade *específica*, como o caráter sintético de certas imagens de base digital, buscando elevá-lo à condição de essência de um novo meio.

Em texto anterior dedicado às “máquinas de imagens”, Dubois (2004) expressa alguns dos argumentos que ele viria revisar. Eles são ilustrativos da compreensão enviesada do digital pela perda de referencialidade. As imagens digitais, segundo descreve, seriam advindas de uma ‘maquinaria extrema’ que constituiria, ela própria, a *causa* da imagem. Nesse registro, a imagem seria imaterial, sem representação nem referente. Escreve ele:

De fato, com a imagem informática, pode-se dizer que é o próprio “Real” (o referente originário) que se torna maquínico, pois é gerado por computador. Isto produz uma transformação fundamental no estatuto desta “realidade”, entidade intrínseca que a câmara escura do pintor captava, que a química fotográfica inscrevia e que o cinema e a televisão podiam, em seguida, projetar ou transmitir. Não há mais necessidade destes instrumentos de captação e reprodução, pois de agora em diante o próprio objeto a se “representar” pertence à ordem das máquinas. Ele é gerado pelo programa de computador, e não existe fora dele. É o programa que o cria, forma e modela a seu gosto (DUBOIS, 2004, p. 47).

Desse modo, o autor sugere que a imagem digital causaria profundas inflexões nas categorias que até então fundamentaram a compreensão das imagens. Seria um caso extremo de maquinização, que torna obsoleta a discussão da semelhança. Um caso situado, também, no extremo da imaterialidade: “A imagem informática é menos uma imagem que uma abstração. Nem mesmo uma visão do espírito, mas o produto de um cálculo” (DUBOIS, 2004, p. 65). Assim como se afasta do mundo, a imagem digital se afastaria também da humanidade, pois o extremo da *maquinaria* produtora de imagens seria também seu antípoda no eixo *maquinismo-humanismo* que estrutura o percurso descrito por Dubois.

Em um tom celebratório que contrasta com o discurso da perda, Edmond Couchot (2003) reforça, a seu modo, compreensão similar. O digital, segundo sugere, engendraria um regime visual possuidor de características “totalmente novas” em que a imagem “se libera” (COUCHOT, 2003, p. 160–164). Embora reconheça, para além da imagem de síntese, a possibilidade de *digitalização* de uma imagem produzida sob outro regime (como o fotográfico), Couchot iguala a ambas quanto à perda de referencialidade:

Quer o computador tenha procedido a partir de objetos reais numerizados [digitalizados] ou de objetos descritos matematicamente, a imagem que aparece sobre a tela não possui mais, *tecnicamente*, nenhuma relação direta com qualquer realidade preexistente. Mesmo quando se trata de uma imagem ou objeto numerizado, pois a numerização rompe esta ligação – esta espécie de cordão umbilical – entre a imagem e o real. São números e somente números expressos sob a forma binária na memória e nos circuitos do computador que preexistem a esta imagem e a engendram (COUCHOT, 2003, p. 163).

Mais uma vez, a mudança no substrato tecnológico, em especial na forma da inscrição – do químico ao informacional – era tido como ponto de clivagem fundamental pelo qual se demandaria uma completa revisão ontológica da imagem e de seu modo de representação. Couchot aponta, assim, para uma espacialidade *utópica* e uma temporalidade *ucrônica* que a imagem digital conseguiria estabelecer por este corte fundamental que a separa do real e, por conseguinte, da história das imagens. O digital iria, assim, reencon-

trar-se com a “lógica da escrita alfabética que liberava o pensamento da materialidade sonora da língua” (COUCHOT, 2003, p. 164)³¹. É interessante notar na descrição de Couchot como a ideia de uma ruptura provocada pelo digital no estatuto da imagem a levaria de volta ao paradigma linguístico, do qual, como relata Dubois (2017), a geração dos anos 1980 buscava se distanciar.

Sob marcada influência da teoria francesa, Lucia Santaella e Winfried Nöth (2001) chegaram a formulação similar em sua proposição analítica que distribuiu as tecnologias de imagem segundo três paradigmas fundamentais, em analogia com a tríade semiótica de Peirce. Assumindo o fotográfico como ponto de referência fundamental, Santaella e Nöth sugerem como *pré-fotográficas* as imagens elaboradas manualmente; como *fotográficas* aquelas geradas pela projeção óptica da imagem sobre um suporte químico ou eletrônico (vídeo); e como *pós-fotográficas* aquelas de base informacional, referindo-se especificamente às imagens de síntese. Os autores descrevem cada paradigma segundo aspectos derivados da semiótica peirceana: o fotográfico seria caracterizado pela dominância diádica, ou indicial, como já sugeria Dubois (2012); o pré-fotográfico, pela dominância monádica, ou icônica; e o pós-fotográfico, a imagem digital, pela dominância triádica, ou simbólica, dada a codificação binária da imagem na linguagem da máquina. Com a dominância simbólica, os autores sugerem um caráter imaterial e informacional ao modo de representação, já distante de qualquer materialidade ou empiria: “O que preexiste um pixel? Um programa, linguagem e números. O que está implícito no programa? Um modelo. O ponto de partida da imagem sintética já é uma abstração, não existindo a presença do real empírico em nenhum momento do processo” (SANTAELLA; NÖTH, 2001, p. 167).

Em discussão realizada por César Guimarães (2002), em dado momento, a imagem digital foi também caracterizada, criticamente, no contexto de um novo regime do visível marcado pela perda de sua relação com a experiência. Retomando expressão de Alain Renaud-Alain (1994), Guimarães refere-se às imagens digitais (às de síntese, em especial, isto é, produzidas “no interior” das máquinas, sem o intermédio da câmera) como “imagens sem gravidade”. Porém, diferentemente dos autores supracitados, argumenta não se tratar de uma especificidade tecnológica mas, sim, de uma instância da experiência de sobrecarga informacional da pós-modernidade. Segundo sugere, a partir de

31 Abordando a interface sensorial da artemídia, Grazielle Lautenschlaeger (2016) elabora, pela via das teorias germânicas da mídia, uma boa crítica da sugestão de imaterialidade das mídias digitais por Edmond Couchot.

Deleuze, o estatuto da imagem nesse contexto de sobrecarga teria sido antecipado por certas obras do cinema moderno – em uma substituição do par Olho-Natureza pelo par Cérebro-Informação (GUIMARÃES, 2002, p. 153). Entretanto, ainda que sem o mesmo papel determinante, a natureza informacional da imagem digital ainda orienta a indagação do autor acerca da disponibilidade das imagens digitais à experiência sensível:

A questão que nos interessa aqui é: em que medida a ambiguidade da forma-imagem digital ainda solicita a experiência sensível? Lembremos que se as obras digitais são – em maior ou menor parte – orientadas pelo conceito (pois a imagem é, inicialmente, a atualização dos dados do programa informático) e como tal – aos olhos de Lyotard – ameaçadas pelas situações controladas e calculadas, a experiência estética define-se justamente por uma duração que não pode ser medida ou calculada (GUIMARÃES, 2002, p. 158).

Voltando-se a aspectos da ontologia da imagem digital enquanto código e computação, o autor salienta, a seu modo, a digitalidade no âmbito de um estatuto marcado pelo acúmulo de três perdas: da transcendência, do poder de fabulação, e do vínculo sensível com o mundo (GUIMARÃES, 2002, p. 147).

Esse conjunto heterogêneo de elaborações é exemplar de vertente significativa de um corpo teórico que, como venho discutindo, se desenvolveu acerca das imagens digitais em seu contexto de emergência, nos anos 1990 e 2000. O cotejamento entre estas diferentes formulações torna evidente como, em comum, expõem uma dramatização (seja pela lamentação, seja pela efusividade) das transformações da passagem do regime fotográfico ao que se identificava, então, como um regime digital. Estas proposições, embora hoje pareçam datadas, são compreensíveis diante da instabilidade criada pela tecnologia digital nos anos 1990, com especulações as mais diversas sobre suas implicações, entre promessas e as ameaças³². Para além do tom, contudo – que poderíamos até renovar em meio a aflições contemporâneas – o período também produziu elaborações conceituais acerca da imagem e do digital que incidem ainda hoje em um direcionamento epistemológico dos estudos do campo. Quero argumentar que a ênfase essencialista sobre a natureza do registro, herdada das teorias sobre a imagem fotográfica, achatou a compreensão das imagens digitais. Isto se torna ainda mais sensível se consideramos a digitalidade para além das possibilidades de síntese, que eram tipicamente colocadas em relevo nestas abordagens. Quero dizer com isto que boa parte das teorias sobre a imagem digital enfocaram o tema de sua relação com o real e como esta seria transformada (ou mesmo rompida) no âmbito das imagens de síntese, mas não consideraram outros as-

32 Wendy Chun (2006) e Geert Lovink (2009) elaboram boas revisões críticas dos discursos do período.

pectos desta transformação tecnológica – como as possibilidades de digitalização de imagens originalmente não digitais, ou as possibilidades de circulação e as práticas culturais delas decorrentes.

Deve-se reconhecer que as possibilidades de manipulação da imagem pelas tecnologias digitais apontam, efetivamente, para outros tipos de regime de veracidade e atestação. Hoje, o principal exemplo parecem ser os chamados *deep fakes*³³, que colocam em crise qualquer vestígio de credibilidade imanente da imagem de aparência fotográfica (estática ou em movimento) como registro do real. Entretanto, por mais que se amplifiquem as possibilidades de manipulação da fotografia e, com isso, sejam lançadas novas questões acerca da realidade e de seu referente, parece ser significativo que a imagem considerada realista ainda se manifeste em uma continuidade da estética fotográfica. Afinal, o desafio colocado hoje pelos *deep fakes* se deve a uma maior verossimilhança da manipulação que é obtida justamente em uma aproximação da imagem de base fotográfica. Como tratarei brevemente no capítulo seguinte (3. Visualidades computacionais), é igualmente relevante que esse efeito não seja obtido por uma síntese *pura* ou *absoluta* – sem “a presença do real empírico em nenhum momento do processo” (SANTAELLA; NÖTH, 2001, p. 197). Pelo contrário, mesmo a manipulação verossimilhante, hoje, baseia-se no que poderíamos caracterizar como um processo de *condensação* e *destilação* do real empírico, obtido *por meio* do registro fotográfico, mais do que por um cálculo matemático “puro”. Os *deep fakes* são exemplares de um aprendizado de máquina generativo, ponto a que retornarei adiante. O tensionamento da referencialidade no campo visual contemporâneo parece se dar, nesse sentido, muito mais por uma incorporação da visualidade da fotografia pelo computacional do que por sua ruptura ou substituição. Não cabem, portanto, oposições binárias simples, a questão é mais complexa e poderá ser melhor compreendida adiante, segundo o caráter indutivo do aprendizado de máquina. O *virtual* hoje opera, com frequência, menos no sentido do cálculo abstrato do que por meio de condensações de individualidades concretas.

Contudo, mesmo quando não se voltam à síntese como categoria definidora do regime digital das imagens, teorizações sobre a imagem digital buscam evidenciar efeitos de ruptura. São comuns formulações como a de Hoelzl e Marie (2015, p. 63) que em certa

33 Esta técnica de falseamento deriva de aplicações do aprendizado de máquina por redes neurais, na modalidade chamada de *deep learning* (aprendizado profundo). Por isso a denominação *deep fake*. Uma de suas primeiras aplicações conhecidas foi demonstrada em vídeo do ex-presidente estadunidense Barack Obama, produzido por pesquisadores da Universidade de Washington, dos EUA (SUWAJANAKORN; SEITZ; KEMELMACHER-SHLIZERMAN, 2017). O vídeo ganhou atenção *online* a partir de julho de 2017.

passagem postulam uma superação da projeção geométrica como aspecto constituinte da fotografia digital, em favor de sua constituição algorítmica. Um exagero, evidentemente, já que não se verifica uma transformação substancial na composição óptica das câmeras digitais. Também William Uricchio (2011) ao se voltar para as possibilidades de representação do espaço e de navegação pelas imagens de base fotográfica, aponta para uma “virada algorítmica” das imagens. Ele não chega a afirmar que esta virada tenha sido concluída, indicando de outro modo que as transformações que observa em tecnologias de realidade aumentada ou de fotografia 360° seriam “fissuras” em um regime ainda bastante resiliente da modernidade. Contudo, ele sugere o horizonte de uma transformação profunda pela via das imagens digitais.

Sem entrar no mérito do argumento teleológico, diria que estas formulações talvez exagerem as questões em jogo. Parece-me mais produtivo o argumento de Daniel Rubinstein e Katrina Sluis (2008, p. 11) que indicam como uma das principais consequências da digitalização a “amadorização em massa” da fotografia, em uma “visibilidade *online* renovada”. Com os celulares com câmera, encontramos mais imagens feitas para circular do que para permanecer³⁴. Seria também nesse sentido que Dubois (2017, p. 39) indica a emergência de estudos voltados não tanto a uma discussão ontológica, mas aos usos das imagens e suas manifestações vernaculares: “A questão ‘o que é a fotografia?’ é assim sucedida por uma outra questão de fundo: ‘o que pode a fotografia?’ (a que ela serve? Quais são os valores que ela veicula e que atribuímos a ela?)”.

Gostaria, em todo caso, de insistir na questão ontológica, porém em outro enquadramento. Como venho argumentando, trata-se de uma questão fundamental para uma consideração crítica do tratamento computacional conferido às imagens, tanto no âmbito dos processos de *datificação* das plataformas, quanto no âmbito das metodologias de pesquisa que se voltam à cultura visual contemporânea. Isto se torna mais evidente à medida que este modo de tratamento das imagens, com sua visualidade particular, passa a integrar ativamente o campo visual contemporâneo. Nesse sentido, trata-se de uma transformação pertinente à questão do *que é* a fotografia, que também coloca questões importantes quanto ao *que pode* a fotografia e *a que* ela serve. Parece-me especialmente importante essa insistência porque o distanciamento que se observa da questão ontoló-

34 Indicações destes outros aspectos podem ser encontradas na ideia de um *circulacionismo* das imagens contemporâneas, descrito pela artista e pesquisadora Hito Steyerl (2009, 2013)

gica, longe de significar sua resolução, sugere um abandono. As razões indicadas para este caminho, por sua vez, me parecem se dever justamente a uma manutenção do equívoco teórico que venho descrevendo, até aqui, acerca da caracterização do *digital*.

Voltando ao percurso do artigo de revisão de Dubois, o argumento central sustentado por ele é o de uma reconfiguração teórica, que passa da concepção de uma *imagem-traço* (conforme a teorização da fotografia dos anos 1980) a uma *imagem-ficção*. Esta transição implica, segundo discute, um abandono da ideia de um “universo de referência” a que a imagem se vincularia, por um “universo de ficção” (DUBOIS, 2017, p. 45). Ou seja, abandona-se a ideia da imagem como emanção do real para sua reformulação como inscrição fabuladora de um mundo possível. O que quero contestar não é a vertente de investigação assumida pelo autor, mas sim como, em aspectos secundários de sua argumentação, permanecem heranças do momento teórico precedente que talvez merecessem uma revisão mais profunda. Um primeiro ponto é que, embora se observe a redução de uma pretensão ontologizante, isto não se refletiu em um deslocamento da *referencialidade* como eixo central de sua compreensão das imagens ou em uma reconsideração do modo de constituição desta *referencialidade*. Ao contrapor o “universo de ficção” a um “universo de referência”, a questão permanece central, embora em negativo. Ou seja, sem a mesma ênfase ontologizante, a ideia de uma síntese ou de distanciamento referencial ainda organiza sua elaboração.

Ao sustentar essa questão como eixo de sua análise, contudo, Dubois não o articula a uma reconsideração do lugar da técnica na constituição da imagem e, com isto, chego ao segundo ponto, central à minha discussão. Pois, justamente com a questão ontológica, Dubois dispensa a técnica como categoria teórica. Este ponto, em particular, parece-me decorrer de um equívoco já presente na concepção assumida nos anos 1980 e que perdura na reconfiguração agora proposta. Isto porque o instrumental teórico que fundamentou a concepção da *imagem-traço* a partir do dispositivo da fotografia acabou por elevar a uma condição determinante o momento da inscrição fotográfica. Este gesto redutor teria sido suficiente, naquele momento, mas as tentativas de sua transposição às imagens digitais parecem revelar sua inadequação. A esta inadequação, Dubois responde por um abandono da questão, em vez de refinar seu tratamento teórico.

Essa dispensa seria justificada, para ele, porque a tecnologia digital tornaria a discussão sobre o *dispositivo* produtor das imagens, em certa medida, obsoleta. De fato, como ar-

gumentarei, a tecnologia digital demanda que revisemos certa visada essencialista sobre a técnica, não só acerca de suas configurações contemporâneas mas, também, como uma falha que deve ser considerada retrospectivamente. Contudo, não é esse o sentido da revisão de Dubois, que propõe um esvaziamento da questão. Segundo elabora, o digital viria aplainar a discussão do dispositivo na medida em que o código informacional único, da máquina, trataria imagens de diferentes naturezas, ou, mesmo, imagens e textos, de igual modo:

todos alojados sob a mesma insígnia digital indiferenciada da reprodução e da transmissão dos “sinais” da informação [...]. Do ponto de vista do digital, não há *diferença* entre um texto, uma imagem e sons; tudo é reduzido à base “informacional” dos *data*, ao mesmo substrato de sinais codificados digitalmente. [...] Essa mudança é fundamental, tanto para o pensamento da ontologia da imagem e de seus dispositivos quanto para o pensamento sobre os usos e as práticas da imagem. O campo teórico sob esse prisma se torna mais intenso, mais denso, mais complexo; mais vasto e diversificado, mas também menos claro, menos definido, menos estruturado (uma vez que tudo é, agora, “digital”) (DUBOIS, 2017, p. 41).

Percebe-se, então, que seu argumento se baseia na redução da tecnologia digital a seu aspecto mais elementar, o código numérico discreto. Pela universalidade presumida do registro – reduzindo qualquer tipo de informação a uma mesma codificação – Dubois argumenta que o código levaria a uma configuração menos definida e estruturada. Por isso, sugere, não haveria muito o que se dizer sobre a ontologia das imagens e seus dispositivos. Sem estrutura, não haveria o que se descrever.

Há pertinência na consideração que faz ao final da citação ao indicar que “tudo é, agora, ‘digital’”. Com a pervasividade das tecnologias digitais, perde hoje sentido a postulação do digital como indicação de uma especificidade essencial. Este é, inclusive, o mote de parte das reflexões contemporâneas no entorno de uma condição *pós-digital* (BISHOP; GANSING; PARIKKA, 2016; CRAMER, 2015; PAUL, 2015). No entanto, esta condição não deveria encaminhar para um abandono da discussão sobre a *digitalidade*. Pelo contrário, ela deveria motivar a que se adensassem as discussões acerca das diferenciações internas às tecnologias digitais, abandonando o digital enquanto categoria monolítica. Salta aos olhos, nesse sentido, que embora Dubois reconheça um aumento de intensidade e complexidade da dimensão técnica, esta constatação não seja acompanhada por um equivalente adensamento das questões acerca da implicação ontológica da técnica, da imagem e de seus dispositivos. No outro extremo, Dubois vê nesta situação um pretexto para dispensar a questão:

No fundo, eu diria que a chegada do digital permitiu justamente relativizar, recolocar no lugar essa teoria dos anos 1980 ao limitá-la à sua dimensão “genética”, a esse simples momento do processo de fabricação da imagem, e ao mostrar que sua “ontologização” foi uma extensão para o menos discutível, um tipo de cegueira epistemológica, uma tentativa de epifania teórica pela absolutização, pela glorificação, do que não é, em suma, nada mais do que um procedimento técnico (DUBOIS, 2017, p. 43–44).

Descrevo, deste modo, o que me parece o equívoco central da consideração da tecnologia digital nos estudos da imagem. Se, por um lado, o afã ontologizante que levava a uma visada *essencialista do fotográfico*, de fato deva ser revisado, parece-me equivocado que a digitalidade seja motivo para o abandono de uma reflexão acerca da materialidade técnica das imagens contemporâneas. Haveria, me parece, outros modos de conceber a tecnicidade das imagens digitais sem, com isto, derivar em uma compreensão essencialista, seja do digital, seja de suas imagens.

Possivelmente, um dos problemas colocados pela tecnologia digital seria, justamente, sua maior resistência a uma visada reducionista ou essencializante. Pois, se por um lado o registro da imagem se materializa em uma codificação binária, por outro essa codificação é apenas uma de suas camadas constituintes, que é suplantada por outras, em combinações com ampla variabilidade e que são irredutíveis ao tratamento essencialista por uma categoria ampla como *o digital*. De fato, a codificação dos sinais pela matemática discreta é um dos traços distintivos (se não o traço distintivo) das tecnologias digitais. São eles que permitem a manipulação simbólica da informação por procedimentos algorítmicos bem como a tradução de diferentes tipos de informação (imagem, som, números, textos verbais) a registros de uma mesma natureza. O problema não está nesta constatação mas, sim, na sua elevação a uma condição determinante das mídias digitais sem levar em conta as múltiplas mediações em jogo.

A argumentação de Dubois é abordada aqui pela sua relevância e influência no campo dos estudos das imagens, embora se trate, evidentemente, de um recorte limitado. Notavelmente, este enfoque tende a privilegiar autores franceses. Contudo, trata-se de vertente teórica com forte influência também no contexto brasileiro, como minha revisão indicou, e que ainda hoje reverbera nos estudos da imagem no país. Apesar desta limitação, é importante destacar que este ponto de divergência que discuto agora não é específico de sua abordagem. Trata-se de uma compreensão recorrente das tecnologias digitais que acaba por desconsiderar as particularidades técnicas ou ‘mitificá-las’ diante do desafio colocado à sua descrição. De certo modo, parece ser um sintoma da carência

de instrumentos teóricos e metodológicos para descrever estas tecnologias, no campo de estudo das imagens, na arte e na comunicação. Talvez indique, também, certa permeabilidade do campo a formulações pouco rigorosas propagadas pela indústria tecnológica ou no senso comum. O caso da codificação binária, ressaltada por Dubois, talvez seja hoje um aspecto menos saliente, mais claramente vinculado à emergência das tecnologias digitais. Porém, algo similar pode ser observado na hipervaloração do *algoritmo* como categoria conceitual sem considerar as diferenciações subsumidas por esta denominação. A esse respeito, Adrian Mackenzie (2017, p. 9), por exemplo, reivindica a necessidade de uma compreensão mais específica destes objetos: “Que algoritmo, que tipo de abstração, e que ‘processo matemático’ devemos focar? Como a automação e o cálculo, a abstração e a matemática são historicamente mutáveis”³⁵. Na mesma linha, Bernhard Rieder (2018) sugere a categoria de “técnica algorítmica” como uma unidade de estudo viável para considerar modos constituídos e replicáveis de elaboração de algoritmos que permitam certo grau de generalização sem perder de vista as especificidades. Observa-se, pois, a demanda que encontremos modos de tratar *dados* ou *algoritmos* evitando formas vagas que tendem à essencialização.

Com este percurso, busquei sustentar ao menos três pontos. Primeiro, seguindo a revisão de Dubois (2017), que as teorias da imagem desenvolvidas na década de 1980 acerca da imagem fotográfica demonstraram-se insuficientes ou mal direcionadas para lidar com o caso da imagem digital. Pois a transposição do modelo teórico entre estes dois objetos foi realizada a partir da premissa de uma distinção ontológica de base entre seus processos de gênese. Segundo, que a tecnologia digital colocou um problema epistemológico a esta tentativa de abordagem uma vez que não havia ainda um instrumental teórico ou metodológico para descrever a constituição da imagem digital. Como resultado, as teorias tenderam a essencializar o caráter simbólico e calculável do registro digital, sem observar as mediações que o estruturam para além de sua instância mais concreta de representação numérica. Avançando sobre estes desafios, na seção a seguir busco esboçar modos de elaborar estas distinções internas ao digital, sem derivar em seu tratamento monolítico.

35 No original: “Which algorithm, what kind of abstraction, and which ‘mathematical way’ should we focus on? Like automation and calculation, abstraction and mathematics are historically mutable”. Tradução minha.

2.2 A imagem múltipla

Contrariamente à concepção da técnica como determinante ou como sobredeterminada, busco adotar um caminho intermediário em que a constituição digital da imagem seria *desessencializada*. A justificativa para este movimento ficará mais compreensível ao longo desta seção, mas antecipo que ele se deve a uma compreensão mais complexa da mediação técnica, que busca evitar tanto sua anulação quanto sua acepção determinista. Para isso é preciso desenvolver um olhar mais atento às diferentes mediações que tomam parte dos dispositivos de imagem, tomando-os em seus vetores de diferenciação e não tanto no sentido de sua homogenização. Este modo de tratamento, conforme argui na seção anterior, levou a uma incompreensão acerca das implicações da tecnologia digital ao campo das imagens. *O digital* não pode ser reduzido à natureza do registro pois ‘zeros e uns’ podem ser arrançados de diferentes formas. Por sua vez, os modos de estruturação dos dados tampouco podem ser tomados como traços essenciais, sem observar os algoritmos que os processam. E mesmo estes não podem ser tratados por uma forma genérica – como *o algoritmo* – sem que neles reconheçamos seus diferentes pressupostos e modos de operação, ou suas diferentes vinculações institucionais e epistêmicas. Nesta seção, busco um vocabulário teórico para melhor descrever estas formas de mediação evitando elaborações essencialistas como as que permearam o *fotográfico* e, depois, *o digital*.

Um primeiro reenquadramento demandado pela tarefa diz respeito à compreensão unitária do dispositivo. Não se trata de questão exclusiva da imagem digital. Voltado especificamente ao cinema, Gabriel Menotti (2019) reivindica a necessidade de revisar a aplicação singular do *dispositivo cinematográfico* para dar lugar à sua multiplicidade. Sua proposta busca articular a perspectiva simondiana da individuação dos objetos técnicos (SIMONDON, 2007) como um ponto de tensionamento com a categoria do *dispositivo*. Ele propõe, por esta articulação teórica, a ideia de que uma *identidade* do cinema seria constituída performativamente³⁶. Projecionista, projetor, curadores, público, filme, arquivo digital, algoritmos de compressão/descompressão, tela, luzes, *torrent*, plataformas de *streaming*. Todos esses elementos estão em jogo na constituição contemporânea do objeto *filme*³⁷. De modo similar ao que observei na seção anterior acerca da categoria do

36 Não enfrentarei aqui as possíveis aproximações entre Simondon e a TAR, que mobilizo como referencial teórico principal. Há contudo, importantes aproximações, embora Latour em certo momento tente traçar diferenças fundamentais, conforme elabora Pedro Ferreira (2017), que contesta a posição de Latour.

37 No original em inglês, Menotti prefere se referir a estes objetos como *movies*. Diante da impossibilidade de tradução, utilizo aqui *filme*.

fotográfico, Menotti argumenta que haveria um reducionismo nas acepções clássicas que concebem o dispositivo cinematográfico por elementos proeminentes como a câmera, a sala de projeção, a tela ou a película. Concepções como estas, argumenta, se evidenciam inclusive nos editais de produção e premiação, os quais expressam diretrizes normativas quanto ao que é considerado um *filme*, com demandas relacionadas ao suporte, ao circuito de exibição e à duração, entre outros aspectos.

Em contraposição às abordagens reducionistas, Menotti (2019, p. 24) propõe um gesto analítico de *desindividuação* do cinema, perseguindo desde as dinâmicas heterodoxas ou clandestinas de distribuição possibilitadas pelo *videotape* até as dinâmicas de conformação do *filme* pelos formulários de submissão de obras cinematográficas a festivais. Contrariamente à ideia do *cinema* e do *filme* como possuidores de identidades estáveis e bem definidas, ele argumenta que ambos levam uma existência precária que se encontra sempre em processo de manutenção (MENOTTI, 2019, p. 61). A abordagem proposta por Menotti abre algumas possibilidades interessantes para esta pesquisa. Uma *desindividuação* da imagem digital demanda que consideremos, assim como no caso do *cinema*, que a categoria do *digital* seria um *efeito* performativo de práticas e materialidades dispersas e não uma categoria pré-definida. Neste sentido, se convocamos *o digital* por um pressuposto de *irreducibilidade*, como uma complexidade que não pode ser subsumida por descrições gerais, trata-se de um fator de complexificação do problema, não de sua solução.

Em linhas gerais, meu argumento é o de que a crise colocada pelo digital às teorias que propunham uma ontologia das imagens impõe a necessidade de revisarmos a compreensão da materialidade técnica e suas implicações. Não porque a tecnologia digital implique uma ruptura definitiva com os outros modos de existência da imagem mas, sim, porque sua introdução promove uma multiplicação dos seus modos de existência. Esta condição se coloca como oportunidade analítica para reconsiderarmos a suposição de uma identidade que outrora conferíamos aos demais 'regimes'. Como elaborei à Introdução, é justamente o trânsito entre esta *multiplicidade* que coloco como problema à pesquisa, compreendendo que as práticas metodológicas *performam* materialidades distintas para uma 'mesma' imagem. No cerne desta compreensão encontram-se concepções da realidade e das práticas de produção do conhecimento elaboradas no contexto da TAR.

Logo de início, é importante destacar que a TAR não constitui um corpo teórico coeso. Pelo contrário, seu estatuto enquanto *teoria* é contestado até mesmo por John Law, um dos membros da espécie de ‘triumvirato’ que é reconhecido pela formulação inicial dessa abordagem, formado também por Michel Callon e Bruno Latour. Law (2009, p. 141) afirma categoricamente que a TAR não é uma teoria. Entre outros motivos, porque sua força não reside propriamente em seu poder *explicativo* mas, sim, em seu poder *descritivo*. Mais do que isso, como também discute o autor, trata-se de um conjunto de métodos e sensibilidades analíticas sem uma identidade clara – “uma diáspora que se sobrepõe a outras tradições intelectuais”³⁸ (LAW, 2009, p. 142).

Seus aspectos teóricos e conceituais são melhor considerados de forma articulada, sem tomar a TAR como proposição artificialmente isolada mas, sim, como uma entre outras abordagens que lhe avizinham. Em uma das narrativas genealógicas mais consolidadas, trata-se de uma vertente dos Estudos de Ciência e Tecnologia (STS), precedida pela Sociologia do Conhecimento Científico (SSK, na sigla em inglês). Deve-se destacar, também nesse contexto, seus débitos à historiografia da ciência, da qual herda a preferência por estudos de caso e não a grandes sistematizações (cf. BIJKER; LAW, 1992). Como enfatiza Tiago Salgado (2019), trata-se também de uma de três vertentes das sociologias pragmáticas francesas, gestadas no mesmo período e compartilhando alguns princípios teóricos e premissas analíticas. Destaca-se, por exemplo, o compartilhamento de uma compreensão do social como efeito das ações e associações entre os atores, e não como entidade preestabelecida. Em ainda outro caminho, sem relação de exclusão com os demais, há as dívidas da TAR à filosofia pós-estruturalista, em especial de Foucault e Deleuze, pelas quais John Law (2009, p. 145) sugere que a TAR “pode ser compreendida como uma versão empírica do pós-estruturalismo”³⁹. Exatamente por esta *sensibilidade empírica*, a TAR funda-se menos em modelos e paradigmas teóricos abstratos do que em teorizações a partir de casos empíricos.

Arriscando uma síntese interessada pela abordagem que visou elaborar, diria que um dos principais traços da TAR é a sua compreensão da realidade como um efeito de práticas situadas. Segundo esta compreensão, a TAR parte da premissa de que não haveria nenhuma realidade social preexistente que se oferecesse como *explicação* para os fenômenos abordados. De outro modo, como elabora Latour (2005), é a própria consti-

38 No original: “a diaspora that overlaps with other intellectual traditions”. Tradução minha.

39 No original: “can be understood as an empirical version of post-structuralism”. Tradução minha.

tuição desta realidade que deve ser investigada. Ao se voltar às práticas como conformadoras da realidade, a TAR compreende uma ontologia performativa, já que, como formula Law (2009, p. 141), “nada tem realidade ou forma fora da performance destas relações”. A noção *ator-rede* vem nomear, nesse contexto, a constituição relacional e performativa dos próprios atores sociais, que não são compreendidos em isolamento. Pois a TAR assume um alto grau de incerteza quanto à fonte ou à origem da ação, concebendo o próprio *ator* como efeito de uma rede de práticas, e a *agência* como uma capacidade distribuída.

A proposição do termo *ator-rede*, nesse sentido, busca dar conta de uma ambiguidade ontológica, como indica seu verbete em volume fundador desta proposta (CALLON; LAW; RIP, 1986a, p. xvi): “o ator é tanto a rede quanto um ponto nela”⁴⁰. Ou, como posteriormente formulado por Latour (2011, p. 800): “É nessa completa *reversibilidade* – um ator não é mais do que uma rede, exceto que uma rede não é mais do que atores – que reside a principal originalidade desta teoria. Aqui, novamente, a rede é o conceito que lhe ajuda a redistribuir e realocar a ação”⁴¹. Por isto, o hífen em *ator-rede* pode ser compreendido como sinal de igualdade: *ator = rede* (VENTURINI; MUNK; JACOMY, 2018).

A mobilização da TAR nesta investigação deve-se principalmente à sua capacidade descritiva diante de dinâmicas relacionais complexas. Especialmente, diante casos em que as mediações técnicas são instâncias constitutivas fundamentais, demandando que sejam consideradas em sua capacidade de agência. Meus objetivos ao me apropriar da TAR são, portanto, bastante específicos. Primeiro, me volto ao modo com que a TAR concebe a mediação técnica, com o objetivo de *desessencializar* ou *desindividuar* a categoria do *digital* e dos dispositivos de imagem. Em um segundo momento, volto-me à noção de *inscrição*, segundo proposição de Madeleine Akrich (1992), para propor uma compreensão da imagem como inscrição sociotécnica, segundo os agenciamentos a que é submetida em cada momento. Em um terceiro aspecto, a partir de Annemarie Mol (1999), volto-me às implicações ontológicas da TAR, em especial sua acepção da realidade como *múltipla* e *performativa*. Atentando-me à *sensibilidade empírica* constituinte da TAR, busco desenvolver as discussões teóricas em proximidade com aspectos pertinentes à constituição da imagem digital em uma antecipação, em parte, da discussão do

40 No original: “*the actor is both the network and a point therein*”. Tradução minha.

41 No original: “*It is in this complete reversibility—an actor is nothing but a network, except that a network is nothing but actors—that resides the main originality of this theory. Here again, network is the concept that helps you redistribute and reallocate action*”. Tradução minha.

próximo capítulo. O instrumental teórico que busco elaborar aqui será importante para, adiante, dar conta das múltiplas instanciações das imagens em protocolos de análise por métodos computacionais.

2.2.1 Mediação técnica: da luz à imagem

Uma das proposições mais conhecidas da TAR é a de um princípio de *simetria generalizada*, desenvolvida inicialmente por Michel Callon (1984). Trata-se de uma extensão do princípio formulado por David Bloor, no contexto da SSK, que se referia à consideração de perspectivas divergentes em uma controvérsia científica sem distinções *a priori* acerca de sua validade ou falsidade. Segundo a proposição de Callon, levada adiante no contexto da TAR, “dado o princípio de simetria generalizada, a regra que devemos respeitar é a de não mudar de registro quando nos movermos dos aspectos técnicos aos aspectos sociais do problema estudado”⁴² (CALLON, 1984, p. 200). Em outras passagens, a questão é posta por Callon também no sentido de uma simetria entre o “mundo natural” e o “mundo social”, em aspecto posteriormente enfatizado por Bruno Latour (1994). Fundamentalmente, uma consequência do princípio de simetria generalizada é a sua permissividade ontológica quanto aos tipos de entidades consideradas em uma análise social. Denunciando um antropocentrismo excessivo da sociologia clássica, autores da TAR irão propor que, enquanto capacidade distribuída, a *agência* deve ser considerada tanto em entidades *humanas* e quanto em entidades *não-humanas*.

Essa consideração da *agência não-humana* viria ser enfatizada em perspectivas notoriamente influenciadas pela TAR, como a chamada *Ontologia Orientada aos Objetos* (OOO), em exercícios especulativos visando o descentramento do humano não apenas como agente mas, também como senciente do mundo (BRYANT; SRNICEK; HARMAN, 2011; GRUSIN, 2015; SALGADO, 2018). No âmbito do pensamento de Bruno Latour, contudo, algumas das principais consequências deste princípio vêm hoje orientar um esforço de revisão de aspectos da visão de mundo moderna em vista, especialmente, de seu fracasso diante do novo regime climático e ambiental⁴³ (LATOURE, 2017, 2018; LATOURE; LENTON, 2019). Para esta investigação, contudo, a questão da *agência não-humana* tem uma relevância mais restrita, relacionada às contribuições mais consolidadas da TAR

42 No original: “given the principle of generalized symmetry, the rule which we must respect is not to change registers when we move from the technical to the social aspects of the problem studied”. Tradução minha.

43 Nesse contexto, a consideração das entidades não-humanas é direcionada no sentido de sua descrição como participantes *ativas* da construção do mundo como realidade habitável, como no equilíbrio das distribuições de carbono nas diferentes camadas da chamada *zona crítica* (ARÈNES; LATOURE; GAILLARDET, 2018).

para os STS e as ciências sociais de modo geral. Ela importa, especificamente, como compreensão teórica do lugar dos objetos técnicos na descrição da imagem, como fator de multiplicidade de sua ontologia. Dois pontos, a esse respeito, serão fundamentais.

Primeiro, os objetos técnicos não podem ser tomados como simples ferramentas sobre-determinadas que apenas transmitiriam, sem transformação, a agência de outras entidades ontologicamente superiores. É ilustrativa dessa formulação a distinção conceitual elaborada por Latour entre *intermediários* e *mediadores* (LATOURE, 2001). Intermediários seriam entidades que poderiam ser plenamente compreendidas apenas segundo aquilo que as antecede ou as sucede na cadeia de relações, ou seja, que não produziriam diferença por si próprias. Mediadores, ao contrário, não podem ser definidos por tais fatores e, efetivamente, os ultrapassam. Como explica Tiago Salgado (2019), trata-se de distinção didática, já que a mediação enquanto fator de transformação constitui o cerne da formulação teórica do ator-rede. Isto porque o mediador modifica não apenas o curso da ação mas, também, a si mesmo: “O mediador [...] é aquele ator que não apenas transporta a ação de um lugar a outro, mas a partilha, a distribui junto a outros mediadores e, ao fazer isso, transforma a si mesmo, o que transporta e aquilo que possibilita o transporte” (SALGADO, 2019, p. 108–109). Portanto, uma primeira consequência da agência não-humana, segundo o princípio de simetria generalizada, é a consideração de que os objetos técnicos não podem ser reduzidos a intermediários neutros a serviço de um ator humano ou quaisquer outros. Como *mediadores*, os objetos técnicos transformam o curso da ação de modos que não podem ser resumidos por uma situação preexistente ou pelos demais atores a que se associam.

De modo importante, esta compreensão não significa que os objetos *determinam* o curso da ação. Como discute Latour (2005, p. 72):

podem existir muitas gradações metafísicas entre causalidade total e simples inexistência. Além de ‘determinar’ e servir como ‘pano de fundo para a ação humana’, coisas podem autorizar, conceder, proporcionar, encorajar, permitir, sugerir, influenciar, bloquear, tornar possível, proibir e assim por diante⁴⁴.

Posto de outro modo, deve-se compreender que a agência de um objeto não é imanente. Ele não pode determinar um programa de ação porque não age sozinho. Nos termos da TAR, sua força deve ser compreendida de forma situacional, segundo sua capacidade de

44 No original: “there might exist many metaphysical shades between full causality and sheer inexistence. In addition to ‘determining’ and serving as a ‘backdrop for human action’, things might authorize, allow, afford, encourage, permit, suggest, influence, block, render possible, forbid, and so on”. Tradução minha.

transladar (CALLON, 1984) os programas dos demais atores a os seus próprios termos. Trata-se de processo que pode ser melhor ou pior sucedido, mas que nunca ocorrerá no sentido de uma translação *completa*.

Um segundo aspecto que quero enfatizar, diz respeito à irredutibilidade do objeto a uma unidade estável. Em referência à noção cibernética da caixa-preta, a TAR sugere uma inflexão ao termo, compreendendo-o menos como coisa do que como processo: obscurecimento ou *blackboxing* (“encaixapretamento”). Assim, visa-se descrever a dinâmica que torna o trabalho técnico e científico invisível diante do bom funcionamento de seus produtos. Escreve Latour: “quando um fato é estabelecido, basta-nos enfatizar sua alimentação [(input)] e produção [(output)], deixando de lado sua complexidade interna. Assim, paradoxalmente, quanto mais a ciência e a tecnologia obtêm sucesso, mas opacas e obscuras se tornam” (LATOURE, 2001, p. 353). Como me referi à introdução (cf. p. 18), há similaridades entre este processo e aquilo que Simondon (2007) descreve como um vetor de *concretização* dos objetos técnicos. No âmbito da TAR, contudo, não se trata de um processo apenas relativo aos *objetos*, mas também a outras instâncias do trabalho tecnocientífico: fatos, conceitos, experimentos etc. O obscurecimento leva à invisibilização daquilo que sustenta, precariamente, a constituição destas instâncias. Ou, parafraseando John Law (2009), à ignorância do caráter relacional da materialidade tecnológica e científica.

Ou seja, este segundo ponto, articulado ao primeiro, diz respeito a uma multiplicidade constituinte de um objeto técnico, cujo tratamento como uma *unidade* seria um efeito de seu bom funcionamento e estabilidade. Longe de um *dado*, a unidade seria um estado transitório, que poderíamos relacionar à capacidade de translação das múltiplas instâncias que o compõem em um programa de ação único. Esta situação tende a invisibilizar componentes, premissas teóricas e outras dependências, embora elas precisem continuar atuando de forma coordenada para que a entidade *obscurecida* siga mantendo este estado. Há um trabalho de manutenção do objeto, portanto, para que ele siga igual a si mesmo⁴⁵.

Seria possível avançar em outros aspectos da teorização que a TAR elabora sobre a mediação técnica, mas estes dois pontos serão suficientes para os objetivos desta

45 Em formulação posterior, já se distanciando da TAR, Latour indicará a reprodução [REP] como um dos *modos de existência* segundo o qual as entidades precisariam estar constantemente se re-produzindo de modo a permanecer as mesmas (LATOURE, 2013, “A perilous change of correspondence”)

investigação. O sentido de sua mobilização, como indiquei, é a revisão do lugar dedicado à materialidade e às técnicas nas teorias da imagem, especialmente para as imagens digitais. Este trabalho se desdobra nas diferentes seções do capítulo, mas nesta subseção volto-me especificamente ao entendimento de que *o digital* implicaria um distanciamento ou, mais ainda, um *corte* da relação entre a realidade empírica e a imagem.

Um primeiro aspecto, como já indiquei, parece ter relação com o tipo de imagem digital a que se referia. Com efeito, muitos dos autores enfatizavam as imagens de síntese, isto é, as geradas “internamente” ao computador e não o registro fotográfico digital. Contudo, esta consideração raramente é feita pelos autores ao se referir ao digital ou ao regime de visualidade que ele instituiria. Em alguns casos, como em Couchot (2003), a diferença entre “imagens digitais” e “imagens digitalizadas” é até explicitamente tomada como irrelevante. De modo mais importante, contudo, o contraste entre *o fotográfico* e *o digital* se realiza em uma essencialização não apenas de um mas de ambos os termos, desfazendo diferenciações internas, também, ao próprio regime fotográfico. O aspecto essencializado é aquele da gênese da imagem, que no caso da fotografia é compreendida pelo modo de constituição da técnica, como uma “emanação do real”, ou como uma imagem formada “automaticamente” por uma impressão luminosa do mundo. A imagem digital, nesta dicotomia, é compreendida essencialmente por sua codificação digital e virtualidade, tomando forma visível apenas por meio de uma atualização algorítmica.

A questão da essencialização torna-se mais inteligível diante da compreensão conceitual da mediação técnica da TAR. A convocação das categorias *fotográfico* e *digital* opera, afinal, pelo obscurecimento de seus modos de operação específicos, sugerindo uma homogeneidade de cada tipo de imagem e uma diferença essencial que as separa. Uma segunda questão, articulada a esta, relaciona-se à fundamentação do fotográfico em uma proximidade ontológica com o real sem, contudo, pormenorizar o caráter transformador das mediações técnicas mobilizadas no registro da imagem. Esta compreensão subtende a mediação técnica como *intermediária*, algo que aparece, por exemplo, nas indicações de uma linguagem da imagem fotográfica ou cinematográfica, que se funda nas decisões assumidas pela pessoa que opera as máquinas, compreendidas como ferramentas à realização criativa dos sujeitos sem necessariamente implicar um fator fundamental de diferenciação. A inversão dos polos da relação sugerida por Flusser (2002) ao cristalizar o programa da máquina e designar os operadores como pertencentes à classe de *funcionários* das máquinas ou de seus *programadores*, nesse caso, não

resolve a questão. O espectro dicotômico do *maquinismo–humanismo* descrito por Dubois (2004) tampouco ajuda. Seria importante compreender como o ato do registro é distribuído entre os diferentes actantes, tornando a imagem uma inscrição híbrida, nem humana nem não-humana, que resulta destas múltiplas mediações.

Volto-me, então, especificamente, à questão do anteparo fotossensível que permite o registro das imagens. De certo modo, esta é a primeira diferença substancial entre cada tipo de aparato técnico, já que o principal componente logicamente anterior, a *objetiva*, é substancialmente idêntico seja na fotografia analógica, seja na digital⁴⁶. A mesma projeção geométrica, em princípio, é o ponto de partida para o registro da imagem. Tipicamente, contudo, a fotografia analógica utiliza como suporte uma película fotossensível, o *filme*, enquanto a fotografia digital utiliza um sensor eletrônico (CCD ou CMOS). Entretanto, como sustenta Sean Cubitt (2014) em sua genealogia das tecnologias de imagem, mesmo a diferença entre estes suportes não implica uma ruptura ontológica se o aspecto que nos interessa é, especificamente, o vínculo da imagem com o real. Isto porque tanto a fotografia analógica quanto a fotografia digital baseiam-se em princípios físico-químicos ontologicamente similares.

Como descreve Cubitt, na fotografia analógica, a incidência da luz sobre o filme fotossensível causa a oxidação do material (os grãos de prata), com a liberação de elétrons. O grau de oxidação irá, no processo de revelação, converter-se em tonalidades visíveis. Neste sentido, poderíamos dizer que as intensidades luminosas são convertidas a informação química. Na fotografia digital, a incidência da luz sobre o sensor também faz com que o material libere elétrons, mas em um processo fotolítico. A principal diferença é que enquanto na fotografia analógica os elétrons são dispensados, na fotografia digital eles são coletados e processados como sinais elétricos. A intensidade destes sinais é convertida à forma digital (discreta) e codificada digitalmente como significante de uma tonalidade. Esta informação digital, que pode ser armazenada em diferentes suportes (magnético, eletrônico etc.) será novamente convertida em intensidades luminosas no momento de exibição da imagem. Escreve o autor:

Analogico e digital dependem de reações químicas. Reveladores analógicos reduzem os haletos de prata, mudando o estado de sua oxidação por um processo que envolve a troca de íons entre o filme e o reagente: uma

⁴⁶ Existem hoje diferenças em câmeras que operam em um registro distinto, como as tecnologias de registro do “campo de luz”, também chamada de câmeras plenópticas, em que a objetiva já não tem a mesma constituição. Trata-se, contudo, de uma diferenciação minoritária que não tem interesse específico para esta argumentação.

transferência de elétrons que é ontologicamente pouco diferente dos elétrons liberados pelos processos fotolíticos diretos em chips digitais (CUBITT, 2014, p. 244)⁴⁷.

O ponto principal para o argumento é que embora se altere a natureza do registro, ambos os tipos de imagem são traduções da informação luminosa em outro tipo de informação. Também, ambos processos de tradução são reações químicas que diferem apenas no modo do aproveitamento dos elétrons liberados para a produção do registro.

O argumento de Cubitt se articula ainda com outro aspecto. Ele indica como a oposição de uma *atualidade* da imagem analógica, contraposta a uma *virtualidade* da imagem digital tampouco se sustenta. Segundo o autor, ambas formas possuiriam um estado de latência que supõe a virtualidade, embora este estado se constitua de modos distintos, em cada caso. Na fotografia analógica, a imagem permanece latente entre a exposição do filme e a sua revelação e fixação química. No caso da fotografia digital, a imagem permanece latente entre a ativação do sensor, seu processamento e armazenamento, até que ela seja atualizada em uma tela, projetada ou impressa⁴⁸. Não há nenhuma realidade essencial que atravesse intocada a estes processos, os quais são igualmente construídos e manipuláveis. Em ambos os casos, a luz é codificada em um suporte material (como informação química ou eletrônica) e permanece latente até que seja atualizada em um processo de revelação (químico ou computacional).

A principal distinção entre os dois registros, segundo este caminho, refere-se ao modo de codificação das imagens: espacial e contínuo, no caso do filme; temporal e discreto, no caso do digital. O aspecto temporal deve-se ao modo de tradução da superfície da imagem em código digital, o qual é, em última medida, linear. A codificação e a decodificação digital, por isso, não admite simultaneidade como no caso da reação da película ou do papel fotossensível. O sensor tipicamente registra a imagem por uma varredura ao longo do tempo e não na forma de um *instantâneo*⁴⁹.

O caráter contínuo ou discreto, por sua vez, diz respeito a pelo menos dois aspectos. Um é o modo de composição da superfície sensível que, no caso da imagem digital, estru-

47 No original: “Analog and digital rely on chemical reactions. Analog developers reduce the silve halides, changing the state of their oxidation through a process involving the exchange of ions between film stock and developer: an electron transfer, ontologically scarcely different from the electrons freed by direct photolytic processes in digital chips”. Tradução minha.

48 Haveria um sentido ainda mais específico de latência que se refere à persistência da carga no sensor fotossensível após a sua ativação. Isto é, relativo ao período em que o sensor permanece inutilizável, até que seja descarregado e preparado para uma nova exposição (CUBITT, 2014).

49 Este aspecto é compartilhado pelos registros digital e eletrônico (da televisão e do vídeo analógico) e tem implicações estéticas discutidas, em outro momento, por Arlindo Machado (1993).

tura-se na forma de uma matriz de *pixels* (*picture elements*), que são a menor unidade sensível. Trata-se de uma superfície cartesiana, com números de linhas e colunas predefinidos. No caso da fotografia analógica, a distribuição dos grãos fotossensíveis não é estruturada e tende à distribuição aleatória sugerindo, nesse sentido, o aspecto de uma superfície contínua.

Outro aspecto pertinente ao caráter contínuo ou discreto é o modo de tradução das intensidades luminosas ao registro químico ou informacional. Na fotografia analógica, haveria um registro mais próximo do contínuo na medida em que não há um limite predefinido às subdivisões entre tons, as quais são sujeitas às características da matéria. Embora a *latitude de exposição* do filme limite o espectro de gradações possíveis, não há limites predefinidos entre os tons possíveis, formando uma variação que não apresenta saltos quantitativos discretos. Na fotografia digital, por sua vez, os tons são discretos e com um limite de gradações definido segundo a *profundidade de cor utilizada* no registro. Há um limite explícito para o número de cores e tonalidades segundo o limite informacional estabelecido para cada *pixel* – a profundidade de 24 bits, utilizada em boa parte dos casos contemporâneos, permite 16.777.216 cores (calculado por 2 elevado à 24ª potência).

Esta descrição pormenorizada, contudo, não implica diretamente o sentido de *mediação* conforme sugerido a partir da TAR, pois há como compreender estes processos de ativação do material fotossensível e de registro técnico da imagem como processos lisos, decorrentes da relação entre componentes relativamente *inertes*. Efetivamente, este é o lugar ao qual tais processos são relegados na literatura técnica. A agência estaria inteiramente na parte humana da relação, que manipularia estes componentes segundo sua vontade. Entretanto, tais descrições são problematizadas e a agência dos materiais emerge quando observamos situações de crise ou de instabilidade que seriam capazes de levar estes componentes a ‘falar’. Madeleine Akrich e Bruno Latour (1992, p. 260) indicam que este seria o momento de uma ‘de-scrição’ (*de-description*), quando “um evento extraordinário – uma crise – modifica a direção da translação das coisas de volta às palavras e permite ao analista rastrear o movimento das palavras às coisas”⁵⁰. Entre os eventos que eles indicam, há situações de falha, de revisão histórica ou de uma brecha experimental explícita. Esta capacidade de *desobscurecer* as mediações e apreendê-las

50 No original: “if some extraordinary event – a crisis – modifies the direction of the translation from things back to words and allows the analyst to trace the movement from words to things”. Tradução minha.

em ação é um dos valores epistemológicos que os STS e a história da ciência atribuem às controvérsias (D'ANDRÉA, 2018; LATOUR, 2005; SHAPIN; SCHAFFER, 2011; VENTURINI, 2010). Elas permitem apreender as fricções e contradições inerentes a entidades que, de outro modo têm seu modo de operação invisibilizado.



Figura 2: Exemplo de 'cartão Shirley' de 1978

Fonte: Imagem da coleção de Hermann Zschiegner retirado de matéria da versão online do New York Times (LEWIS, 2019).

Um caso emblemático que se relaciona diretamente à instância de registro das imagens tornou-se conhecido pelos “cartões Shirley” (Figura 5). Este foi o apelido dado aos materiais impressos usados ao longo do século XX como referência para a calibragem da representação visual cromática – de tons de pele, em especial – em impressões fotográficas analógicas e, posteriormente, também na produção televisiva. Lorna Roth (2009) desenvolveu um amplo estudo destas representações, inclusive com entrevistas a químicos e gerentes de produtos de fabricantes de filmes e relata como os “cartões Shirley” são exemplares de um viés racial das tecnologias de imagem. Sistemáticamente, o parâmetro de calibragem hegemônico, guiando não apenas as práticas de estúdios de la-

boratórios fotográficos, mas da própria indústria de câmeras e emulsões, favorecia o registro da pele branca. Esta situação, segundo relata a autora, levaria décadas sem que fosse efetivamente problematizada no âmbito das empresas fabricantes, embora mobilizassem esforços de fotógrafos e produtores audiovisuais em táticas de compensação. Segundo relata, este atraso se deve não apenas à dimensão socialmente estruturante do racismo mas, também, a uma suposição geralmente compartilhada de que se tratava de um problema estritamente técnico ou científico e não de uma *escolha* por parte dos responsáveis pelo desenvolvimento tecnológico.

De modo crucial, contudo, Roth (2009) elabora em sua análise como este viés é resultado da constituição sociotécnica do desenvolvimento das emulsões e das demais tecnologias de registro⁵¹. Por um lado, características químicas das emulsões colocam desafios a demandas de ampla latitude cromática, isto é, que figuras de tonalidades e brilho distintas sejam representadas visualmente com igual nível de detalhe. Por outro, os modos como esta característica do material é trabalhada durante o desenvolvimento do produto vai muito além das propriedades físicas da matéria. Escreve a autora:

Acreditava-se à época que física era física, química era química e a ciência era baseada em decisões ponderadas sem considerações de sutilezas culturais ou raciais. Agora tem se tornado mais amplamente reconhecido na indústria que refinamentos à química das emulsões dos filmes nunca foram problemas de física ou química exclusivamente, mas sim resultado de escolhas culturais também⁵² (ROTH, 2009, p. 118).

Como resultado, as imagens produzidas, ao longo de décadas, tendiam à má qualidade da representação de tons de pele além do branco, entre pessoas negras, orientais e indígenas. Esta questão, relata a autora, apenas viria a ser diretamente problematizada no âmbito da indústria no final do século XX, com cartões de calibragem multirraciais tornando-se mais comuns, inclusive pela descentralização de sua produção.

Roth (2009) não mobiliza a TAR em seu relato, mas parece-me possível uma aproximação, na linha do que venho discutindo. O caso estudado por ela é elucidativo da condição *mediadora* das instâncias de registro da imagem que descrevi, pois, por meio de uma revisão histórica, alcança uma situação de de-scrição das mediações técnicas envolvidas.

51 No próximo capítulo desenvolverei a partir de trabalhos de Tarcízio Silva (2019) e Joy Buolamwini (BUOLAMWINI; GEBRU, 2018; RAJI; BUOLAMWINI, 2019) como se observa hoje uma situação similar no âmbito das tecnologias computacionais de aprendizado de máquina, com consequências ainda mais urgentes.

52 No original: “It was [...] believed at the time that physics was physics, chemistry was chemistry, and science was based on reasoned decisions without consideration of cultural or racial subtleties. It is now becoming acknowledged more widely within the industry that refinements to the chemistry of film emulsions have never been issues of physics or chemistry exclusively, but have been the result of cultural choices as well”. Tradução minha.

Sua análise, por meio da mobilização dos “cartões Shirley” e, também, por meio de entrevistas, oferece sucessivos enquadramentos pelos quais torna-se explícito o modo como a técnica institui determinada *prescrição* em seu modo de operação. Isto é: ela não apenas transporta de forma neutra, neste caso, um modo de apresentação da realidade para seu registro em imagem.

A necessidade de algum parâmetro de referência é já significativa, nesse sentido, do caráter *mediador* das tecnologias de representação visual, que não apenas *transportam* as características da luz a registros químicos, eletrônicos ou digitais mas, efetivamente, *criam e transformam* estes sinais no processo. Diante destas transformações, a representação não pode ser tomada como “natural”, “automática” ou “objetiva”, pois uma série de decisões precisam ser tomadas de modo a *transladar* a forma de mediação da câmera aos interesses de seus operadores e, nesta articulação, o problema não pode ser meramente técnico e torna-se, efetivamente, *sociotécnico*. Devido a propriedades físicas da matéria, *transladadas* por químicos e engenheiros a um determinado modo de funcionamento almejado, as emulsões fotográficas efetivamente *prescrevem* um modo como esta realidade se apresenta em imagem. Assim, a técnica não opera de forma neutra e sim, como bem apresenta a análise de Roth (2009), reforça dinâmicas culturais e sociais difusas, atuando como uma prática constituinte do racismo.

Outros casos poderiam ser indicativos de programas de ação distintos em que as tecnologias de registro de imagens também seriam problematizadas. Em um outro exemplo com implicações bem distintas do do anterior, mas que também contribui ao argumento geral, imagens produzidas da superfície de Marte pela NASA (Agência Espacial Norte-Americana, na sigla em inglês) constituem um caso interessante em que o modo de registro das imagens por meio de câmeras digitais não pode ser tomado como totalmente “objetivo” mas, sim, como dependente de uma série de *escolhas*. Se, por um lado, o modo de interpretação dos valores cromáticos registrados pelas câmeras depende sempre de um padrão de referência (como os “cartões Shirley”), não há como estabelecer esse padrão para uma condição de registro à qual o olhar humano não tem acesso direto, já que jamais alguém esteve em Marte. Postagem em blog da NASA acerca da sonda Spirit, enviada ao planeta em 2004, relata o esforço demandado para gerar uma imagem com cores que se aproximem daquelas que seriam percebidas por um olhar humano (NASA, 2004). Os cientistas envolvidos relatam que eles conseguem chegar a uma boa *aproximação* e apenas porque realizam muitos experimentos em condições simuladas na Terra e

porque as sondas enviadas possuem, acoplados, cartões de referência para operações de calibragem. Nestes relatos, também é indicado que alcançar este registro cromático ‘natural’ não constitui uma demanda propriamente científica, já que as análises se voltam frequentemente à luz não-visível. Ou seja, trata-se de uma demanda para esforços de divulgação, cruciais para a justificação dos investimentos nestes caros empreendimentos de pesquisa, por exemplo. Mais uma vez, portanto, longe de *automáticos*, os possíveis vínculos entre imagem e realidade empírica são produto de múltiplas mediações compreendidas sempre como entrelaçamento entre fatores ontologicamente heterogêneos (“sociais” e “técnicos”, “humanos” e “não-humanos”).

O caso dos “cartões Shirley” e o caso das fotografias de Marte ilustram, portanto, cada um a seu modo, que as tecnologias de registro da imagem – sejam analógicas (no sentido típico do *fotográfico*), sejam digitais – jamais operam como uma “emanação do real”, em um sentido estrito⁵³. Evidenciam, também, que os processos técnicos que permitem o registro da imagem tampouco são meros instrumentos ou *intermediários*. Dizer que eles são mediadores não significa que eles possuiriam “vontade própria” mas apenas que eles possuem modos de operação específicos que são irreduzíveis às vontades das pessoas que os operam. Eles oferecem resistências mas também ‘proporcionam’ (*afford*) seus próprios programas de ação. O desenvolvimento destes artefatos, por sua vez, não é feito no vácuo nem apenas com base em uma racionalidade etérea ou em um critério de eficiência universal. Eles respondem a, fundamentalmente, demandas políticas e institucionais de seu tempo e, neste sentido, possuem uma política⁵⁴. Esta, me parece, é a força principal do conceito de mediação técnica pela TAR, que implica uma compreensão não apenas técnica mas *sociotécnica* dos artefatos. Só assim seria possível articular modos de agir distintos, irreduzíveis a categorias generalistas ou a equivalências simplistas.

Sob a perspectiva da TAR torna-se mais compreensível, também, a sugestão de uma dispensa da categoria técnica diante do desafio colocado pelo digital. A elaboração em certa medida paradoxal que emerge no texto de Dubois (2017) é sintomática do processo de obscurecimento (*blackboxing*). Este, como apresentei, levaria à invisibilização do traba-

53 Parece-me importante destacar que esta observação não equivale a uma objeção à proposição de estéticas realistas, seja na fotografia ou no cinema, já que muitas delas não se baseiam em premissas acerca do aparato técnico, simplesmente, mas sim de efeitos estilísticos e de linguagem. Jean-Louis Commoli (2008), por exemplo, possui uma proposição influente em que o realismo se caracteriza por uma *abertura* às incertezas do momento da filmagem, mais do que um princípio essencialista do registro *em si*.

54 Uma boa referência a esse respeito, que não chegarei a desenvolver aqui, é o artigo *Artefatos têm política?*, de Langdon Winner ([1980], 2017). Sua aproximação à TAR demandaria alguns cuidados dado a discordâncias manifestas por Winner em relação a autores desta perspectiva. Mas trata-se, ainda assim, de texto fundamental para a discussão da política tecnológica.

lho técnico e científico na medida proporcional do seu sucesso. Dubois sugere, afinal, como indiquei mais acima (veja citação na p.54), que a tecnologia de imagem, no contexto digital, se complexificaria e se adensaria mas, simultaneamente, se desestruturaria e indiferenciaria. Ora, este aparente paradoxo parece ser justamente o obscurecimento em ação. Não é bem que a tecnologia se tornaria menos estruturada, mas apenas que esta estruturação pareceria menos relevante sob o ponto de vista de usuários finais, justamente pelo sucesso da operação. Além disso, a tecnologia digital se invisibiliza de uma forma bastante literal com a microeletrônica e com as camadas de processamento lógico que são inacessíveis ao usuário comum. Isto não significa que haja menos estrutura, nem que ela seja menos relevante, mas o desafio se amplia.

Antes de passar ao ponto seguinte, gostaria de mencionar brevemente um último aspecto pertinente à discussão das mediações técnicas. Trata-se das possíveis relações entre estas formulações que elaborei aqui e a noção, mencionada anteriormente, do *dispositivo*. Evidentemente, essa *breve menção* não tem como dar conta da complexidade desta tarefa de articulação, que segue como um desafio. Apenas parece ser relevante ensaiar algumas indicações dada a importância do *dispositivo* para os esforços teóricos precedentes. Centro-me, para tanto na menção ao *dispositivo* que Latour (2001) faz de passagem (mais breve do que eu). Em sua explanação acerca da mediação técnica, em certo momento ele responde a um questionamento hipotético que lhe indagaria acerca da *intencionalidade* dos objetos técnicos. Ele responde:

A ação intencional e a intencionalidade talvez não sejam propriedades de objetos; contudo, também não são propriedades de humanos. São propriedades de instituições, de aparatos, daquilo que Foucault chama de *dispositifs*. Somente pessoas jurídicas estão aptas a absorver a proliferação de mediadores, a regular sua expressão, a redistribuir habilidades, a forçar caixas a obscurecer-se e fechar-se. [...] Os artefatos reais são sempre partes de instituições, hesitantes em sua condição mista de mediadores, a mobilizar terras e povos remotos, prontos a transformar-se em pessoas ou coisas, sem saber se são compostos de um ou de muitos [...] (LATOURE, 2001, p. 221).

Neste sentido, os dispositivos podem ser compreendidos como uma espécie de individualização de coletivos de atores que realiza uma orquestração das diferentes agências, transladando seus programas de ação em um direcionamento comum. Seria igualmente sugestivo desta compreensão a indicação feita por John Law (2009) de uma analogia entre o conceito de *episteme* e a noção ator-rede. Segundo elaboram Alzamora, Ziller e d'Andréa (2018) o conceito foucaultiano de *episteme* seria uma formulação predecessora da noção de *dispositivo*, porém com uma conformação mais estritamente linguístico-

discursiva. O ator-rede compreende justamente o agregado que pode ser, às vezes, visto como unidade, em uma oscilação irresoluta. Latour sugere que alguma forma de ‘regulação’, ‘absorção’ ou ‘redistribuição’ das agências seria necessária para a constituição de um dispositivo. A leitura que Agamben (2009a) faz do conceito – tomando liberdade, explicitamente, para além da aceitação de Foucault – esboça linha de raciocínio similar sob a ideia de uma *economia* ou de *governo* do ser, como traço característico do dispositivo – “operação por meio da qual se realiza uma pura atividade de governo sem nenhum fundamento no ser” (AGAMBEN, 2009a, p. 38). Talvez então, nesse sentido, um dispositivo fotográfico poderia ser concebido segundo um programa comum resultante do agenciamento coletivo das múltiplas instâncias que o compõem. Algo que não se resume à emulsão fotográfica ou à câmera, mas que pressupõe uma articulação mais ampla: *instituições*. Considere-se, quanto a este ponto, o papel dos fabricantes de emulsões fotográficas na constituição de padrões como os materializados nos cartões Shirley, por exemplo. Seria uma instanciação condicionada, portanto, pela força de translação destas múltiplas agências, ocasionando a composição do que os proponentes da TAR outrora se referiram (em denominação logo abandonada) como um *ator-mundo* (CALLON; LAW; RIP, 1986a).

2.2.2 Inscrição: perspectiva e consistência óptica

Uma segunda noção formulada no âmbito da TAR, que gostaria de desenvolver, é a de *inscrição*. Haveria ao menos duas perspectivas articuladas pelas quais compreendê-la, ambas pertinentes a esta discussão. Primeiro, em um sentido aparentemente mais trivial (mas logo complexificado), uma *inscrição* é um modo pelo qual uma entidade se traduz, *materialmente*, em um signo, tipicamente no âmbito das práticas científicas (LATOURE, 2001). Ou seja, como um território se torna um mapa; como um objeto se torna um desenho; ou como as qualidades de uma pessoa tornam-se um conjunto de dados. Apenas aparentemente trivial, esta questão se complexifica quando consideramos os tipos de operações proporcionadas pelas inscrições enquanto o que Latour denomina como “móveis imutáveis” (LATOURE, 1986, 2001; LATOURE; HERMANDT, 2004). Segundo sugere, não seria apenas uma questão de significação, de signo a signo, mas fundamentalmente, de transformação da realidade, de mundo a inscrição (LATOURE; HERMANDT, 2004).

Um outro sentido, bastante articulado a este, relaciona-se ao que Madeleine Akrich (1992) identifica como um *script* (algo como um ‘roteiro’ ou ‘programa’) dos objetos téc-

nicos. Diferentes das inscrições, os *scripts* ofereceriam, contudo, chaves para sua leitura. Os objetos técnicos, elabora a autora, prescrevem papéis e ações para seus usuários, os quais são sempre sujeitos a dinâmicas de reapropriação. Contudo, como um de seus produtos, os objetos técnicos elaboram inscrições da realidade que são tanto parte de seu modo de funcionamento quanto formas de conhecimento sobre as realidades decorrentes de sua operação, reapropriados por analistas e por formas de exercício do poder. Esta compreensão será importante para considerarmos algumas implicações da mediação técnica da fotografia como ponto de partida para análises de base computacional, como as que realizo adiante nesta tese.

Acerca da primeira acepção, como comecei a indicar, o conceito de *inscrição* indica um modo peculiar de tratamento do que a semiótica denomina como signos, porém com um interesse particular na configuração da realidade pelas práticas científicas. A diferença é tênue. Em uma de suas discussões do termo, Latour (2001) contextualiza sua pertinência em um esforço para superar o que descreve como um “antigo acordo” modernista, que supôs uma cisão fundamental entre realidade (“fora”) e a sua percepção e representações humanas (“dentro”). A solução desta cisão pela modernidade teria se dado, precariamente, pelo que Latour (2001, p. 39) descreve como uma “estreita pinguela”: um vínculo *epistemológico* que forçaria a correspondência entre *linguagem* e *natureza*. Ele sugere, alternativamente, uma atenção à materialidade das *inscrições* como outro modo de conceber esta relação, compreendendo o modo como articulam “cadeias de translações” entre a realidade e suas representações. Seriam estas cadeias, localizadas em práticas, instrumentos e materiais – mais do que em uma epistemologia etérea – que garantiriam a manutenção de certa continuidade entre mundo e suas representações, como condição dos esforços de conhecimento e atuação sobre a realidade.

Uma das primeiras elaborações adensadas sobre esta questão aparece em texto que Latour (1986) dedica a um conjunto de desenvolvimentos da modernidade que ele considera como fatores fundamentais à constituição da ciência moderna, entre os quais o desenho em perspectiva, a escrita tipográfica, a moeda e o desenho de mapas (cartografia). Estes desenvolvimentos, sugere ele, oferecem elementos descritivos mais consistentes para compreendermos a emergência da ciência moderna do que a suposição de uma racionalidade etérea – que Latour vincula a explicações ‘mentalistas’. Práticas representacionais que emergem da modernidade, materializadas em textos e imagens, sugere Latour, seriam distintivas na medida em que proporcionariam às inscrições a

condição de “móveis imutáveis”. Isto é: ao mesmo tempo em que desenvolvem modos de garantir consistência às representações, transladando de forma sistemática aspectos do mundo às representações e entre suas múltiplas cópias e instanciações, estas inscrições são também recombinaíveis e mutuamente legíveis (comparáveis), mesmo em contextos distintos daqueles em que foram elaboradas e para os quais se dirigem (LATOUR, 1986).

Estas qualidades não são triviais e, ao final, são fatores fundamentais para se compreender o poder da ciência e também de muitos outros “centros de cálculo” (LATOUR; HERMANDT, 2004) da modernidade e da contemporaneidade, como laboratórios, bibliotecas e coleções e também mercados, escritórios e órgãos do Estado, poderíamos acrescentar. No caso específico desta tese, considerarei também sob esta chave os modelos de rede neural para o reconhecimento de imagens, em especial por sua tendência contemporânea de *infraestruturalização* (veja 3.3. *Visualidade e infraestrutura*). Estes são todos lugares cuja qualidade fundamental é a reunião de múltiplas inscrições que os conectam a realidades distantes, as quais, são, naquele local, analisadas, permutadas e combinadas. Neste sentido, as descrições dos modos de conhecimento e ação sobre o mundo na modernidade deveriam se atentar aos vários aspectos no entorno destes objetos peculiares: o modo de elaboração destas imagens e textos, em uma translação do mundo às inscrições; o modo de sua circulação, de sua mobilidade; e as práticas de sua coleção, análise e replicação.

Estas considerações ganham corpo em uma situação anedótica narrada por Latour (2001) durante estudo em que acompanhou o trabalho de cientistas que investigavam a floresta amazônica. O estudo que desenvolviam buscava compreender se evidências do solo e da vegetação indicavam tendências de avanço ou de retraimento da floresta sobre a savana. Em meio aos trabalhos, Latour descreve uma situação curiosa em que quatro cientistas, dois franceses e duas brasileiras, reuniam-se em torno de um mapa disposto improvisadamente sobre a mesa de um restaurante em Boa Vista, Rondônia. Aquela inscrição, descreve, era o que permitia que observassem juntos as evidências e que mobilizassem, naquela situação precária, uma ampla rede de atores que incluíam satélites e instituições científicas, além da própria floresta. Escreve:

Removam-se os mapas, confundam-se as convenções cartográficas, eliminem-se as dezenas de milhares de horas investidas no atlas Radambrasil, interfira-se com o radar dos aeroplanos e nossos quatro cientistas ficarão perdidos na paisagem, obrigados a reiniciar todo o trabalho de exploração,

referenciação, triangulação e quadriculação feito por centenas de predecessores. Sim, os cientistas dominam o mundo – mas desde que o mundo venha até eles sob a forma de inscrições bidimensionais, superpostas e combinadas. É sempre a mesma história, desde que Tales se postou ao pé das Pirâmides (LATOUR, 2001, p. 44).

O estudo que Latour abre com esta anedota se desenvolve, então, com um amplo conjunto de observações situadas das práticas científicas a qual é repleta de situações similares a esta. Cientistas vão a campo, colhem materiais, mensuram, analisam, observam e, em algum momento, reúnem-se em uma sala em que as múltiplas inscrições decorrentes destas atividades são dispostas em uma tela, um quadro ou uma mesa. Tabuladas e processadas, podem ser agregadas em um gráfico. Sistematizadas e simplificadas, apresentadas em um diagrama. Em todo caso, a reunião dos cientistas em torno de um mapa é representativa do adensamento das práticas na medida em que a inscrição assume, ali, o lugar da realidade observada⁵⁵. As noções de significação ou de representação são indicativas do processo, mas insuficientes e, por isso, Latour busca na noção de “móveis imutáveis” um modo de distinguir o sentido que atribui às inscrições de seu tratamento semiótico clássico. Ele enfatiza sobretudo a condição material das inscrições e as práticas que permitem seu acúmulo e alinhamento, ampliando seu poder de convencimento e de translação dos atores.

Segundo esta conceituação, compreender as imagens como inscrições sociotécnicas demanda que voltemos nossa atenção ao modo com que elas seriam articuladas em uma cadeia de translações, segundo demandas específicas de produção de conhecimento, mas também como elas poderiam alcançar um plano comum, permitindo que transitassem entre circuitos distintos. Em todos estes movimentos, a condição de translação do mundo à imagem seria fundamental, assim como os aspectos que as constituiriam como *móveis imutáveis* – isto é, como manteriam a consistência de sua representação mesmo em trânsito e permitindo práticas de recombinação, comparação e observação sinóptica.

Um dos princípios que permitem às imagens serem trabalhadas desta forma já se encontra na proposição conceitual de Latour (1986), em sua consideração do desenho em perspectiva em um contexto de emergência da ciência moderna. Em um sentido geral, a

55 Tecendo considerações acerca da fotografia da mão de um cientista apontando, com o dedo indicador, um aspecto de um diagrama, Latour (2001, p. 82) escreve: “A menos que seja o prelúdio rancoroso de um soco, a extensão do indicador revela sempre um acesso à realidade, até quando tem por alvo um simples pedaço de papel – acesso que, neste caso, engloba a totalidade do sítio, o qual paradoxalmente desapareceu por completo, embora estejamos suando no meio dele. Temos aí a inversão de espaço e tempo a que já assistimos inúmeras vezes: graças às inscrições, podemos superintender e controlar uma situação na qual estamos mergulhados, tornamo-nos superiores àquilo que é maior que nós e conseguimos reunir sinoticamente todas as ações empreendidas no curso de vários dias, desde então esquecidas”.

perspectiva linear, desenvolvida no Renascimento, integra um percurso histórico mais extenso de *racionalização da visão*. Conforme desenvolvido por William Ivins (1975) e depois retomado por Latour (1986) a formalização do desenho em perspectiva é uma construção com implicações profundas para as práticas de produção do conhecimento. Trata-se de desenvolvimento vinculado principalmente a demandas da arquitetura, cuja principal qualidade era a redução de dimensionalidade da representação espacial. A formalização da perspectiva em princípios geométricos permitia a transposição de um espaço tridimensional a uma figuração plana, na tela ou no papel. De modo crucial, contudo, o desenho em perspectiva abria a possibilidade de recomposição do espaço representado, pois o caráter formal do procedimento adotado permitia extrapolar o espaço tridimensional a partir do plano, em uma correspondência ponto a ponto. Desse modo, a imagem consistia um móvel imutável, considerando que a imutabilidade diria respeito não apenas a uma estabilidade da própria inscrição em circulação mas, principalmente, a uma consistência de seus vínculos com a realidade inscrita. Por meio do desenho em perspectiva o espaço poderia, ele próprio, circular.

Com um desenvolvimento concomitante a dispositivos ópticos como a câmara escura, a formalização da perspectiva também viria a ser transladada à operação óptica da câmera fotográfica. As dívidas da câmera fotográfica à câmera escura são bastante conhecidas pois, efetivamente, ambas se aproximam em seu princípio de projeção geométrica do mundo sobre um anteparo⁵⁶. A câmera escura, grosso modo, é uma técnica em que a projeção geométrica fundante do desenho em perspectiva é realizada por uma máquina óptica. Com a câmera fotográfica, o registro desta projeção é, por sua vez, também automatizado, como vimos, por meio da emulsão fotossensível do filme ou pelo sensor eletrônico – além, claro, da operação de outros componentes como obturadores e diafragmas. Nesse sentido, a mediação da projeção óptico-geométrica da câmera seria um plano de continuidade a conectar desde as imagens em perspectiva da Renascença até as imagens fotográficas analógicas e, então, as digitais. O modo de translação do espaço à representação plana é, afinal, um outro eixo de continuidade a conectar *o fotográfico ao digital*, já que não configura uma diferença ontológica de princípio entre cada ‘regime’⁵⁷.

56 Jonathan Crary (1992) contesta, em certa medida, a relação direta entre a câmara escura e as imagens fotográficas, não por recusar por completo esta relação, mas por reivindicar um entrelaçamento mais complexo de linhas temporais e desenvolvimentos concomitantes. Esta questão não tem, contudo, implicações diretas para a relação traçada aqui, cujas motivações são mais restritas.

57 Este ponto contraria o que Ingrid Hoelzl e Rémi Marie (2015, p. 63) sugerem, em certa passagem, ao sobrevalorar a dimensão algorítmica da imagem digital. Voltarei a sua reivindicação da natureza algorítmica das imagens adiante, em uma consideração ponderada de seus argumentos (cf. [2.2.3 Multiplicidade ontológica: declinações materiais da imagem](#)).

Esta constituição sociotécnica das imagens de base fotográfica (analógicas ou digitais) lhes confere, portanto, um modo de existência bastante peculiar, em que elas não se resumem a práticas culturais e afetivas. A câmera fotográfica configura, afinal, uma máquina de *inscrições*. As imagens que produz se encontram na interseção de múltiplas cadeias de translação, informada por séculos de desenvolvimentos de óptica, mecânica, química e, então, eletrônica e computação. Enquanto condensação destas redes, as câmeras produzem, desse modo, “automaticamente”, móveis imutáveis. As imagens não seriam, portanto, emanações do real, mas materializações de uma complexa translação da realidade a inscrições planas. Os próprios princípios que regem o funcionamento das câmeras, por sua vez, oferecem chaves interpretativas pelas quais a realidade poderia ser recomposta, ainda que parcialmente, a partir da imagem. Diante da estabilidade da cadeia de translações que constituem esse registro, teríamos na imagem fotográfica uma inscrição cujo contexto de aplicação não se resume apenas a este. Na sequência, busco em Madeleine Akrich (1992) dois outros aspectos das *inscrições*: seu potencial de “exportação” para além de uma cadeia de translações específica; e seu papel em articulações da produção do conhecimento com o exercício do poder.

No tratamento que Akrich (1992) confere ao tema das inscrições, elas não constituem somente uma forma de mediação entre linguagem e natureza, ou entre mundo e cognição, mas, também, entre formas de conhecimento e formas de poder. Seu interesse específico é pelas dinâmicas de negociação entre os usos presumidos de objetos técnicos em seu processo de desenvolvimento e os casos reais de sua apropriação pelos usuários. Nesse sentido, para Akrich (1992), o trabalho da análise dos objetos técnicos deve buscar interpretar os *scripts* dos objetos técnicos (em alusão ao roteiro de um filme ou peça teatral) e, ao mesmo passo, observar situacionalmente como este roteiro ou programa é deslocado e reapropriado pelos usuários finais e outros atores implicados. O vocabulário mobilizado pela autora desdobra-se, assim, em um amplo leque de termos que visam descrever as dinâmicas de conformação que operam entre: o objeto e o programa assumido pelos engenheiros; a realidade e o programa de ação assumido pelo objeto; e entre o objeto e o programa de ação assumido pelos usuários.

Efetivamente, em Akrich, a noção de *inscrição* é mais abrangente do que aquela mobilizada por Latour. Por vezes ela equivale à noção de *script*, como ‘conteúdo’ de um objeto técnico. Em outros momentos, ela sugere diferentes formas de instanciação de um mediador. Em um dos estudos de caso que aborda em seu texto, ela analisa como esta

dinâmica opera na implantação de redes de distribuição elétrica na Costa do Marfim. Sobre este caso, por exemplo, Akrich (1992) se refere aos medidores individuais de consumo de energia como inscrição material do contrato de serviços. Permanece nesta acepção de inscrição a ideia de uma translação consistente entre instâncias ontologicamente distintas (do contrato ao medidor), mas esta translação não visa necessariamente a uma mobilidade, como sugere Latour, nem tem em vista a produção de conhecimento. O medidor é, antes de tudo, um mediador político. Porém, Akrich também articula esta mediação a uma forma de produção de conhecimento que dela seria indissociável. Pois, escreve a autora, a pessoa que projeta um arranjo tecnológico “não apenas fixa a distribuição dos atores, mas ele ou ela também provê uma ‘chave’ que pode ser usada para interpretar todos eventos subsequentes”⁵⁸ (AKRICH, 1992, p. 216). Os mediadores individuais de consumo, nesse sentido, não apenas permitem a cobrança pelo serviço como também fornecem subsídios a outras inferências por parte da companhia elétrica, tais como o cálculo de padrões coletivos de consumo e a identificação de possíveis ligações clandestinas, denunciadas por padrões individuais anormais. Desse modo, Akrich descreve uma articulação, com os medidores ao centro, entre formas de conhecimento e formas de poder.

Esta articulação se aprofunda se consideramos que, além da própria empresa, que toma as inscrições como instrumento regulador de seus processos técnicos, haveria também possibilidades de “exportação” das inscrições, como observa Akrich (1992). Pois as inscrições também podem ser reapropriadas por agentes que são, em princípio, externos à distribuição de energia (para seguir no exemplo da autora). Os *dados* de consumo de energia interessam, assim, a sociólogos e economistas que podem reapropriar estas inscrições como forma de conhecer aquela realidade (AKRICH, 1992, p. 221). Inscrições sociotécnicas seriam pontos de articulação não apenas entre mundo e cognição mas, também, entre a técnica, a política e o saber. Como se sabe, e como Akrich brevemente menciona, estes termos são notoriamente articulados por Michel Foucault. Em sua análise das sociedades disciplinares, Foucault (1997) demonstra esta relação em operações distribuídas observadas em hospitais, prisões e escolas que vão deste a construção arquitetônica às práticas de monitoramento e mensuração comportamental dos indivíduos. Na análise de Akrich, tais dinâmicas são descritas em uma observação empírica situada.

58 No original: “... not only fixes the distribution of actors, he or she also provides a ‘key’ that can be used to interpret all subsequent events”. Tradução minha.

Indo além do que compreendemos logo acima com Latour, portanto, Akrich indica uma transversalidade das inscrições para além de contextos específicos da produção do conhecimento em que, de certo modo, a “cadeia de translações” é regida por uma motivação científica comum. Desse modo, as inscrições possuem um sentido específico no arranjo técnico em que se inserem, mas também encontram sentido em um plano geral, no qual circulam amplamente, em atendimento a demandas imprevistas pelo sistema de inscrição inicial. Uma condição para esta circulação ampla, contudo, como salienta Akrich, é que sejam mobilizados termos em comum. Isto é, que alguma chave de transposição seja mobilizada em cada contexto, permitindo o compartilhamento de inscrições. Para o caso abordado pela autora, estas poderiam ser classes ou categorias econômicas ou sociológicas que permitiriam que os dados gerados pela companhia elétrica fossem apropriados por cientistas sociais.

A dinâmica de datificação, que conforma o contexto de comunicação plataformizada abordado por esta tese, tende a multiplicar estas possibilidades. Com efeito, este é o sentido principal da perspectiva de reapropriação (*repurposing*) sugerida pelos Métodos Digitais (ROGERS, 2013). Também é este o fator que atrai as ciências sociais aos “rastros digitais” do ambiente midiático contemporâneo (BRUNO, 2012; VENTURINI; LATOUR, 2010). Para o caso específico das imagens, uma primeira aproximação possível deriva do que indiquei anteriormente acerca da projeção geométrica, ou seja, da representação visual perspectivada, como fator de consistência óptica da imagem como inscrição. Este aspecto não garante apenas a “imutabilidade” das inscrições, no sentido latouriano, mas também parece constituir um fator da sua “exportabilidade” para além do contexto específico do fotográfico. Isto é, permite que o registro fotográfico seja integrado em outras cadeias de translação.

Parece-me ilustrativo desta “exportabilidade” do registro fotográfico o estudo realizado pelo filósofo e fotógrafo Alan Sekula (1986) acerca dos usos repressivos da fotografia. Percorrendo aplicações do registro fotográfico na criminalística e na criminologia do século XIX, Sekula descreve sistemas de identificação policial e de investigações eugenistas sobre *tipos* criminosos em que o registro fotográfico desempenha um tal papel mediador. Os sistemas em questão se formalizam em operações que aprofundam o fator de consistência dos registros fotográficos, com padronizações dos parâmetros de tomada das imagens dos retratos policiais. Muito além do registro fotográfico, contudo, eles também se materializam em fichas arquivísticas, técnicas de sobreimpressão foto-

gráfica, medidas biométricas e estatística descritiva⁵⁹. O registro fotográfico, neste sentido, seria exportado da cadeia de translações específica da câmera. Isto apenas seria possível, contudo, porque a câmera se tornou um objeto estável – um requisito, segundo elabora Akrich (1992, p. 221), para que o objeto técnico se obscureça e, assim, que “fatos sociotécnicos” sejam tomados como “fatos”, simplesmente. Acerca deste arranjo de objetos e técnicas, escreve Sekula (1986, p. 16), em certa passagem: “A câmera é integrada em um arranjo maior: um sistema burocrático-administrativo-estatístico de ‘inteligência’. Este sistema pode ser descrito como uma forma sofisticada do arquivo. O artefato central deste sistema não é a câmera, mas o armário de arquivo”⁶⁰. Segundo os termos que venho discutindo acerca da TAR, teríamos o arquivo como “centro de cálculo” ao qual convergiriam as cadeias de translação integradas por múltiplas mediações, entre as quais a câmera fotográfica. Teríamos, as imagens como *inscrições*, como *móveis imutáveis* rearranjados e combinados nestes arquivos.

Os casos investigados por Sekula são importantes antecedentes para os contextos contemporâneos de vigilância e monitoramento. Eles são casos de aplicação em que se observa uma mobilização instrumental da imagem de base fotográfica que viria a se ampliar sobremaneira com o barateamento e consequente multiplicação dos dispositivos produtores de imagens. O cineasta Harun Farocki (2004) refere-se por “imagens operativas”, ou “operacionais”, às imagens produzidas com esta finalidade, distantes daquelas tipicamente encontradas nas teorias humanísticas sobre as imagens⁶¹. O que definiria, para Farocki, o caráter operativo ou não de uma imagem seria sua destinação, ou sua finalidade. Imagens operativas não são feitas “nem para entreter nem para informar”, “não representam um objeto mas, sim, compõem parte de uma operação”⁶² (FAROCKI, 2004, p. 17).

Levando adiante a proposição de Farocki, o artista estadunidense Trevor Paglen diz, hoje, de “imagens invisíveis” que, já além de seu objetivo estritamente operacional, se

59 Segundo desenvolve o autor, a criminalística tem um objetivo mais diretamente operacional, como sistematização do trabalho desempenhado por forças policiais. Sekula analisa, a esse respeito, o sistema de identificação desenvolvido pelo chefe de polícia de Paris Alphonse Bertillon. A criminologia, em contraste, visa produzir conhecimento científico sobre a prática criminosa, abordagem examinada por Sekula no abjeto trabalho do estatístico e eugenista britânico Francis Galton.

60 No original: “*The camera is integrated into a larger ensemble: a bureaucratic-clerical-statistical system of ‘intelligence’.* This system can be described as a sophisticated form of the archive. The central artifact of this system is not the camera but the filing cabinet”. Tradução minha.

61 Para artigos que elaboram leituras desta noção, cf. FLORES, 2016; PAGLEN, 2014; VIDAL JUNIOR, 2016.

62 No original: “*neither to entertain nor to inform*”; “*do not represent an object, but rather are part of an operation*”. Tradução minha.

encontrariam inclusive restritas a circuitos computacionais de uma visão não-humana (PAGLEN, 2016)⁶³. Gostaria de indicar, porém, um aspecto complementar ao sugerido por Paglen. Pois se, por um lado, o caráter operativo das imagens possa ser definido segundo os arranjos sociotécnicos em que se encontram integradas, estes limites nem sempre são tão claros e esta situação se complexifica em um contexto de digitalização. De modo mais crucial, considerando que a constituição sociotécnica da câmera leva a fotografia a inevitavelmente se constituir como uma inscrição sociotécnica, em maior ou menor grau, gostaria de sugerir que elas poderiam ser consideradas como *imagens operacionais flutuantes*.

Meu argumento central, a esse respeito, é de que justamente pela “exportabilidade” das imagens enquanto inscrições sociotécnicas, não caberia considerar as *imagens operativas* segundo sua destinação mas, sim, segundo seus usos, decorrentes das dinâmicas de circulação a que são submetidas. A representação perspectivada constitui, nesse sentido, apenas um dos fatores que articulam a imagem a esse plano de transversalidade que permite o reaproveitamento das inscrições. Outro fator que também contribui contemporaneamente a este processo é a digitalização, ao potencializar a circulação e a reapropriação das imagens mas, principalmente, ao tornar a imagem disponível à interpretabilidade computacional. Nesse contexto, teríamos, portanto, que mesmo imagens de cunho afetivo ou feitas para “entreter e informar” poderiam ser tomadas como *operativas* no sentido sugerido por Farocki. Isto é, mesmo elas podem compor, hoje, “parte de uma operação”. Um exemplo são as imagens utilizadas em bases de treinamento de redes neurais de reconhecimento de imagens e de rostos que, embora não produzidas para este fim, são coletadas em sua circulação *online* e apropriadas para este treinamento (cf. CRAWFORD; PAGLEN, 2019; HARVEY, 2019). Diria serem imagens operacionais *flutuantes*, portanto, porque a finalidade operacional das imagens não precisa estar definida de antemão. Parafraseando Agamben em sua famosa asserção sobre o Estado securitário contemporâneo⁶⁴: nada se assemelha melhor a uma imagem operacional do que uma fotografia de família. Voltarei a este ponto no capítulo seguinte (3 Visualidades computacionais).

63 Em texto escrito durante o doutorado elaborei algumas discussões acerca do trabalho recente de Paglen (MINTZ, 2018d).

64 “Aos olhos da autoridade – e, talvez, esta tenha razão – nada se assemelha melhor ao terrorista do que o homem comum” (AGAMBEN, 2009a, p. 50).

Quero destacar, contudo, que mesmo antes das imagens digitais, justamente pela perspectiva e pela câmera fotográfica como fatores de consistência da inscrição, as fotografias já se disponibilizavam a uma visada operativa. Um caso certamente curioso que ilustra esta possibilidade é o exercício de uma espécie de “historiografia forense” pelo engenheiro e artista Billy Kluver (2003) diante de um conjunto de fotografias tomadas por Jean Cocteau em um passeio por Paris com Pablo Picasso. As imagens são registros afetivos de um grupo de amigos boêmios caminhando pela cidade e seus cafés. Porém, interessado no estudo da história daquele momento cultural de Paris, Kluver coleciona aquelas imagens e as aborda de um modo peculiar, em um trabalho de investigação que visa não apenas identificar os personagens e recompor a sequência original das fotografias mas, também, situá-las precisamente no tempo e no espaço. Para isto, ele faz uma análise detida de cada imagem, atenta-se aos marcadores geográficos e até mesmo à projeção das sombras dos prédios. Este olhar minucioso, amparado pela consistência óptica do registro, permite a Kluver recompor detalhes de cada imagem e, desse modo, adensar a narrativa que reconstrói daquele dia.

Trago o exemplo de Kluver para ilustrar minha sugestão de uma *imagem operativa flutuante* por um possível antecedente. Esta sugestão irá se complexificar no capítulo seguinte. Mas parece-me relevante destacar outro aspecto deste exemplo que é como as fotografias, neste caso, prestam-se a múltiplos olhares e desse modo, parecem se declinar em múltiplas manifestações. Com Akrich (1992), vimos como uma inscrição pode prestar-se a múltiplas apropriações, por diferentes atores. Porém, não se indicava, ali, uma multiplicidade própria às inscrições. O dado de um medidor de consumo de energia segue sendo apenas ele próprio. As fotografias de Jean Cocteau, contudo, não são apenas inscrições que se prestam a múltiplas aplicações, mas objetos múltiplos, que podem se declinar em inscrições para um olhar forense, mas que têm, ao que parece, outra destinação. Na próxima subseção, desdobro esta questão.

2.2.3 Multiplicidade ontológica: declinações materiais da imagem

Um terceiro e último aspecto que gostaria de desenvolver em uma aproximação à TAR diz respeito mais a uma consequência de suas operações conceituais do que uma formulação intrínseca a seu delineamento “clássico”, propriamente. O conceito de mediação, como vimos, envolve uma sensibilidade ao caráter distribuído da agência que lança a uma condição de incerteza os contornos precisos dos atores. O conceito de inscrição, por

sua vez, nomeia instâncias materiais desta distribuição. Porém, mais do que uma distribuição da ação, as inscrições apontam para uma distribuição dos próprios objetos. Afinal, uma inscrição não é apenas um signo que se vincula a um referente previamente conhecido mas, de outro modo, ele constitui um vetor de expansão material daquele referente ou, de modo ainda mais transformador, pode se tratar da única manifestação cognoscível daquele referente – como no caso do mapa de um território, por exemplo. Neste sentido, Latour (2001) sugere que não seria propriamente um referente “externo”, mas um “referente circulante” que seria conhecido por meio das inscrições. As inscrições servem como um modo de conhecer a realidade mas também, de modo importante, como um modo de agir sobre esta realidade.

Esta condição levaria, como observa Annemarie Mol (1999), a uma *multiplicidade ontológica* que abre possibilidades de ação política sobre a realidade. Não se trata de uma proposição metafórica ou especulativa. Tampouco se trata de elaboração relacionada ao tema do perspectivismo ou do construtivismo social, como reivindica a autora. Ao dizer de uma ontologia múltipla, Mol sustenta que as inscrições seriam efetivamente (pois pragmaticamente) um fator de variabilidade dos entes. Esta variabilidade não levaria a uma situação de *pluralidade ontológica*, como sugerem, por diferentes caminhos, o perspectivismo e o construtivismo, pois não se trata de realidades distintas, incomunicáveis ou em disputa. De outro modo, sob a chave da *multiplicidade*, teríamos linhas de coexistência ou mesmo de inter-relação e dependência entre diferentes constituições dos entes.

Um exemplo desenvolvido por Mol (1999), cujas investigações voltam-se às ciências da saúde, é o da anemia. A autora apresenta que haveria (àquele momento) ao menos três diferentes definições para a anemia, a depender do tipo de exame utilizado no diagnóstico e das chaves interpretativas às inscrições. Uma seria a clínica, em que anemia seria constituída segundo um conjunto de sintomas e queixas dos pacientes. Outra seria a laboratorial, em que a anemia seria constituída segundo a medida do nível de hemoglobina no sangue do paciente em comparação a uma medida estatística do nível considerado normal para uma população. A terceira seria a patofisiológica, em que a anemia seria constituída segundo o nível de hemoglobina de um paciente comparado ao seu próprio nível quando saudável. Todas estas definições configurariam, conforme elabora, a multiplicidade ontológica da anemia e reforça: “Estas não são perspectivas vistas por

peças diferentes [...]. Nem são construções antigas, alternativas, das quais apenas uma emergiu do passado [...]. Então elas são diferentes versões, diferentes performances, diferentes realidades que coexistem no presente”⁶⁵ (MOL, 1999, p. 79).

A autora sugere, portanto, uma compreensão *performativa* da realidade. Este é um aspecto subtendido da TAR, ressaltado principalmente em exercícios de sistematização posterior, como os de John Law (1999, 2009, 2017). Em certo sentido, esta proposição se encontra já contida na primazia das *práticas* e das observações situadas como foco privilegiado da TAR. Formulada conceitualmente, contudo, a ideia de uma performatividade sugere uma incerteza fundamental às descrições e às análises: “entidades alcançam suas formas como consequência das relações em que elas estão localizadas [...]. Uma consequência é que tudo é incerto e reversível, ao menos em princípio”⁶⁶ (LAW, 1999). Em um aprofundamento da ideia de que verdades ou afirmações seriam dependentes de um contexto, sugere-se assim que mesmo a realidade alcançaria este status por meio de práticas situadas.

A anemia, no exemplo de Mol (1999), é muitas e nenhuma delas, em isolamento. Inclusive porque, como a autora observa, elas dependem umas das outras. Por exemplo, a medida de hemoglobina estatisticamente normal, presumida pela anemia laboratorial, depende de um exame clínico para distinguir pacientes saudáveis dos anêmicos antes da tomada das medidas. Além disso, ela elabora, a construção da norma laboratorial também necessita que sejam consideradas diferenças nos valores entre populações de homens e de mulheres, uma distinção que não é relevante para as anemias clínica e patofisiológica. Por relações de inclusão e interferência, portanto, a autora descreve uma compreensão da realidade em que diferentes constituições ontológicas da anemia articulam-se umas às outras e mobilizam, ainda, outras categorizações não necessariamente vinculadas ao objeto em questão. Esta complexidade de relações explica, em larga medida, porque, embora seja *múltipla*, a realidade não é *flexível*. Há escolhas que podemos tomar na configuração desta realidade e, desse modo, podemos falar, como sugere Mol (1999) de uma política ontológica. Para o exemplo dela, haveria uma escolha acerca de qual ontologia da anemia seria a mais adequada e a mais viável por se adotar em um programa de saúde pública, por exemplo. Contudo, estas escolhas não se

65 No original: “These are not perspectives seen by different people [...]. Neither are they alternative, bygone constructions of which only one has emerged from the past [...]. So they are different versions, different performances, different realities, that co-exist in the present”. Tradução minha.

66 No original: “entities achieve their form as a consequence of the relations in which they are located. [...] A consequence is that everything is uncertain and reversible, at least in principle”. Tradução minha.

fazem no vácuo e relações já fortemente estabelecidas em práticas reiteradas não são fáceis de se desfazer. Como elabora Law (2017, p. 44): “Performar objetos é difícil, mesmo nesse mundo relacional. É algo difícil e custoso (pense nos departamentos hospitalares d[*o trabalho de*] Mol). Nós não podemos simplesmente sonhar novas realidades”⁶⁷.

Podemos compreender esta elaboração teórica de maneira articulada com os aspectos priorizados nas subseções anteriores. A multiplicidade ontológica da anemia, no exemplo de Mol, deriva das performances que elaboram suas diferentes inscrições, cada qual com sua chave interpretativa particular. Estas, por sua vez, não podem ser compreendidas fora das mediações que as conformam, inclusive as técnicas: os instrumentos laboratoriais, as ferramentas de cálculo estatístico, os instrumentos da anamnese e dos exames clínicos. Imaginar “outra” anemia demandaria, portanto, reconfigurar boa parte destas relações, em uma extensa redistribuição dos papéis atoriais e translação de seus programas de ação. De certo modo, grandes transições tecnológicas, como a experienciada a partir da introdução das tecnologias digitais, podem ser fatores de alguma transformação desse tipo.

Ingrid Hoelzl e René Marie (2015), em sua proposição teórica acerca da imagem digital, sugerem uma reorganização profunda deste tipo. Provocativamente, eles sugerem que a categoria do *fotográfico*, outrora vinculada a um dispositivo sintetizado, metonimicamente, pela câmera fotográfica, deveria ser redefinido no contexto digital pela categoria presumida pelo algoritmo de compressão e descompressão (*codec*) conhecido pela sigla JPEG (*Joint Photographic Experts Group*, “Grupo Conjunto de Especialistas em Fotografia”). Considerando a instituição normativa deste algoritmo como formato padrão assumido pelas imagens digitais, eles sugerem que o traço definidor da categoria do fotográfico não seria, hoje, vinculada a um aspecto de sua gênese, como pretendiam os teóricos dos anos 1980. Em contrapartida, sugerem que a imagem fotográfica seria definida pelos atributos visuais do tipo de imagens para as quais o *codec* JPEG foi elaborado. Gostaria de retomar a proposição destes autores em uma aproximação livre com a ideia de multiplicidade ontológica em um último gesto de revisão do posicionamento do digital no domínio das teorias da imagem.

Deve-se compreender, a esse respeito, que uma estratégia algorítmica de compressão da informação não serve igualmente a todos e quaisquer tipos de dados. Determinadas ca-

67 No original: “Performing objects is tough, even in this relational world. It is difficult and costly (think of Mol’s hospital departments). We can’t just dream up new realities”. Tradução minha.

racterísticas informacionais precisam ser priorizadas para que se possa obter uma melhor razão entre a eficiência da compressão e as perdas de qualidade. Isto também se aplica para o caso das imagens, em que diferentes distribuições cromáticas dos *pixels* direcionam estratégias de compressão distintas. O JPEG não é a melhor opção, por exemplo, para uma imagem contendo caracteres tipográficos ou um logotipo, especialmente quando a imagem apresenta limites com transições tonais abruptas. Para estas imagens, esse *codec* costuma gerar imagens com aspecto ruidoso próximo às bordas das letras ou das formas gráficas. Isto porque o JPEG funciona melhor quando as variações de cor são graduais ao longo da superfície da imagem, como tipicamente ocorre em uma fotografia. Seu algoritmo trabalha com blocos de 8 por 8 *pixels* de forma conjunta, partindo do pressuposto de que a variação cromática entre eles não deve ser muito aguda. Trata-se, nesse caso, de uma imagem de ‘tom contínuo’, em contraposição a imagens de ‘tom discreto’. Para estas últimas, o padrão PNG (Gráfico de Rede Portátil, na sigla em inglês) seria mais adequado, pois este se baseia no tratamento agregado de áreas que possuem o mesmo valor de cor ‘chapada’, sem perdas. Seu uso em fotografias, contudo, tende a compressões menos eficientes, gerando arquivos maiores.

Com base nesta especificidade e considerando a forte padronização no entorno do *codec* JPEG, Hoelzl e Marie sugerem que haveria uma redefinição da categoria de *fotográfico*:

Para o grupo JPEG, o ‘fotográfico’ não é mais vinculado a uma tecnologia específica de registro e impressão; de outro modo, o termo designa um conjunto de imagens digitais que podem ser comprimidas da mesma maneira. Posto de outra forma, ‘fotográfico’ designa uma distribuição estética particular (tom contínuo) de *pixels* que podem ser correlacionados uns aos outros durante o processo de compressão (HOELZL; MARIE, 2015, p. 69)⁶⁸.

Importante notar que se tal proposição fosse levada ao pé da letra, estaríamos supondo a substituição de uma definição ontológica essencialista por outra. Do *fotográfico* como inicialidade do registro fotoquímico ao *fotográfico* como distribuição de *pixels* em tom contínuo, teríamos duas formulações radicalmente distintas, porém igualmente redutoras. Não é este meu argumento – nem me parece ser o sentido pretendido por Hoelzl e Marie, embora eles sugiram outras distinções fortes como essa (cf. nota 57). Compreen-

⁶⁸ No original: “For the JPEG group, the ‘photographic’ is no longer tied to a specific recording and printing technology; instead, the term designates an array of digital images that can be compressed in the same manner. Put differently, ‘photographic’ designates a particular aesthetic distribution (continuous-tone) of *pixels* that can be correlated with each other during the compression process”. Tradução minha.

do, de outro modo, que a provocação dos autores pode ser assumida transitoriamente como um passo intermédio para reorientar o olhar teórico que, ao final, pretendo dedicar às imagens.

Minha intenção principal, com o argumento do fotográfico como JPEG, é contestar a compreensão que supõe ao digital a absoluta maleabilidade de um registro “menos claro, menos definido, menos estruturado” (DUBOIS, 2017, p. 41). Neste sentido, a proeminência desse *codec* não é sem razão. Trata-se do padrão que hoje modela a constituição informacional de uma parcela majoritária das imagens digitais. Por sua vez, essa força parece se dever, em grande medida, à formação institucional que lhe deu origem: um comitê formado por membros da Organização Internacional para Padronização (ISO, na sigla em inglês), da Comissão Eletrotécnica Internacional (IEC, na sigla em inglês) e da União Internacional de Telecomunicações (ITU, na sigla em inglês) – este último, uma agência da Organização das Nações Unidas (ONU) (Cf. JPEG, [s.d.]). Como apontam Hoelzl e Marie (2015, p. 69), apenas o fato de que este grupo tenha sido formado já é uma evidência da importância da imagem fotográfica no âmbito das tecnologias de informação e comunicação (TICs). A necessidade de padronização da forma de sua representação informacional faz contraponto, também, à ideia de que o digital anularia as distinções entre tipos de mídia e tipos de informação. Por sua vez, a mobilização de tamanho esforço institucional para fazê-lo é indicativo do equívoco que fundamenta a suposição de que com o *digital* se tornaria obsoleta a questão da materialidade tecnológica das imagens. O *codec* JPEG e o arranjo institucional que permite sua elaboração e manutenção devem ser reconhecidos como componentes fundamentais da imagem fotográfica digital, embora não signifique que sejam traços essenciais de sua ontologia.

Na linha do que discuti a partir de Mol (1999), o *codec* JPEG constituiria uma das instâncias performativas contemporâneas das imagens fotográficas. Contudo, seria apenas uma dentre muitas outras ontologias possíveis. Há pouco, mencionei a compressão PNG, mas teríamos também, para ficar apenas em alguns: o GIF (Formato de Intercâmbio de Gráficos, na sigla em inglês) ou os diferentes formatos do tipo *Raw*, sem compressão e sem predefinição de um espaço de cor, por exemplo. Cada uma destas versões deste objeto incerto que é a imagem digital implica *prescrições* específicas segundo a mediação dos *codecs* e protocolos envolvidos, e seus programas de ação subtendidos. A força do JPEG, nesse sentido, está na eficiência da compressão, com um nível de perdas conside-

rado aceitável, mas também está na amplitude da articulação institucional que lhe conformou. Mas uma imagem digital poderia (como muitas vezes acontece) apresentar-se em muitas outras versões, sem se reduzir a nenhuma delas, individualmente.

Não se trata de uma filigrana técnica. Essas diferentes ontologias são condicionantes de um modo de existência contemporâneo das imagens. Cada uma revela uma concepção particular de como tratar o visual: quais aspectos priorizar, quais podem ter mais perdas, quão circulável deve ser a imagem etc. Não são por acaso as siglas PNG ou GIF, que indicam formatos portáteis ou intercambiáveis de imagem. São compressões que priorizam a circulação das imagens e que efetivamente proporcionam esta possibilidade para imagens gráficas. O JPEG, com sua prescrição peculiar, busca atender a uma compreensão específica da imagem fotográfica. Ao favorecerem certos usos e registros, estes formatos excluem outros e são, em última medida, condicionantes do modo de existência contemporâneo das imagens. Arlindo Machado (2007) ressalta, neste sentido, como a produção artística contemporânea frequentemente se debate com estas limitações, como condicionantes de determinadas estéticas visuais. Menotti (2019) descreve, também, um filme de sua autoria que se debate com as características dos *codecs*. No contexto desta investigação, é importante compreender como estes modos de estruturação da existência informacional das imagens respondem a demandas pragmáticas e institucionais diversas e não são, de modo algum, flexíveis ou desestruturadas como a categoria essencializada do *digital* em algum momento supôs.

Em ainda outra articulação, quero indicar que seriam estes modos de estruturação que permitem às imagens digitais operar como *móveis imutáveis* do campo visual contemporâneo. Isto é, como inscrições que permitem a circulação e o agenciamento de múltiplas realidades sobre um plano sinótico (LATOUR, 1986, 2001). Atendendo não só ao roteiro ou aos programas de ação “originais” pelos quais foram concebidas mas, também, podendo ser “exportadas” (AKRICH, 1992), justamente por sua relativa estabilidade, que leva a obscurecer sua estruturação interna. Exportadas, desse modo, as inscrições tornam-se mais amplamente articuláveis em grandes “centros de cálculo” como laboratórios e coleções (LATOUR; HERMANDT, 2004) nos quais a realidade é performada com um alcance ainda maior. Para o caso das imagens digitais, bancos de dados, bases de treinamento de sistemas de aprendizado de máquina e plataformas *online* seriam alguns hoje destes grandes centros de cálculo do visível.

Como discuti a partir de Akrich, esta exportabilidade das inscrições deve-se também a seu alcance de um plano de sentido geral. Isto é, quando sua chave de interpretação não seria mais específica a este ou aquele sistema de inscrições e seria, de outro modo, transversal a múltiplos domínios. Conforme indiquei na subseção anterior (veja [2.2.2 Inscrição: perspectiva e consistência óptica](#)), a automatização da inscrição em perspectiva da câmera fotográfica faria este trabalho para as imagens feitas sob este registro. Toda fotografia seria, nesse sentido, uma inscrição sociotécnica disponível a olhares forenses que nela buscassem reconstituir o espaço retratado. Por esse motivo sugeri que toda imagem fotográfica, mesmo as feitas “para entreter e informar” poderiam ser consideradas, em uma releitura de Farocki (2004), como *imagens operativas flutuantes*. Esta condição que já era própria às imagens fotográficas se aprofundaria em uma perspectiva de digitalização, na medida em que não apenas sua mobilidade se dinamiza e se amplia, como também sua acessibilidade a outros olhares e possibilidades interpretativas, proporcionadas por sua *computacionalidade*. Como indiquei à [Introdução](#), estas são as condições fundamentais que permitem a integração das imagens a dinâmicas de *datificação* do visível. São também estas as condições que possibilitam o tratamento metodológico que dedicarei às imagens, por meio das técnicas de aprendizado de máquina. Voltarei a me aprofundar nesta questão no capítulo seguinte (veja [3 Visualidades computacionais](#)).

Em vista da discussão elaborada por Annemarie Mol (1999), é possível, então, sugerir uma revisão mais profunda da teorização das imagens digitais anteriormente discutida. De modo fundamental, a noção de *multiplicidade* desfaz tanto a suposição de uma estabilidade ontológica que seria conferida às imagens por um dispositivo técnico individualizado, quanto a ideia de que a mediação técnica seria fator desimportante em sua consideração. A tecnologia digital é, como comecei a elaborar, um fator de variabilidade que contribui à multiplicidade ontológica das manifestações das imagens. A essencialização do código informacional como traço definidor de uma imagem digital puramente de síntese ou algorítmica não seria, por isso, adequada à sua descrição. Em um contato mais próximo com a experiência contemporânea das imagens, é necessário considerar que a imagem digital é código *e também* é visível. É fotográfica *e também* é digital. Isto sem entrar no detalhamento das múltiplas versões das imagens subsumidas pelas categorias do código, da informação ou do algoritmo. Esta multiplicidade, longe de dispensar a questão ontológica dirigida às imagens, a complexifica e, como sugere Mol

(1999), também a politiza pois admite alguma margem de escolha por priorizar alguma versão específica dentre as muitas em que as imagens se materializam – aspecto que será tensionado adiante.

Embora esta condição sugira uma revisão do vínculo ontológico ou essencial da imagem com uma realidade empírica, este vínculo é apenas reconfigurado mas não totalmente rompido. Como discuti a partir de Sean Cubitt (2014), do filme *ao sensor*, temos apenas dois modos distintos de translação em que a mesma projeção luminosa é convertida ora em informação química, ora em informação digital. Em ambos os casos temos inscrições derivadas de uma realidade empírica mas, nem por isso, idênticas a ela. Com a disponibilidade das inscrições digitais a procedimentos computacionais, teríamos, de outro modo, a potencialização do estatuto da imagem como *referência circulante* e, como sugeri, *imagem operativa flutuante*. Assim, teríamos uma ampliação das consequências pragmáticas deste vínculo e, portanto, das imagens.

Evidentemente, esta elaboração teórica não “resolve”, por assim dizer, a questão acerca do que seriam as imagens digitais mas, de outro modo, a reenquadra. Pois se o que se observa é uma maior variabilidade das inscrições, levando a um quadro de *multiplicidade ontológica*, parece que não haveria uma resposta única possível. De outro modo, trata-se de questão que apenas pode ser respondida de maneira circunstancial e transitória em um momento específico das cadeias de translação. Entretanto, este parece ser justamente o desafio colocado para esta pesquisa em seu problema metodológico. Proponho que os procedimentos de análise a que me volto adiante sejam compreendidos como estas cadeias de translação em que o visível se declina em múltiplas inscrições, as quais são então justapostas, combinadas e analisadas. Porém, buscando um caminho de retorno ou uma reversibilidade dos processos, como sugeri à introdução, o desafio é constituir um protocolo de pesquisa em que esta multiplicidade ontológica seja abraçada pela investigação, em vez de tomar descrições computacionais como o ponto de chegada da investigação.

Como sugeri algumas vezes neste capítulo, quero argumentar que embora a tecnologia digital venha colocar em crise as elaborações essencialistas que visaram em algum momento *fixar* o estatuto ontológico da imagem segundo seu dispositivo, ao demonstrar a inadequação desse enquadramento teórico a um novo paradigma tecnológico, trata-se de crise que não afetaria apenas o presente das imagens mas que, de outro modo, deveria

incidir também, retrospectivamente, em imagens anteriores ao digital. Na subseção anterior, ao retomar o ensaio de Alan Sekula (1986) acerca dos usos repressivos da fotografia, indiquei como sua discussão já apontava para uma reconfiguração da noção de dispositivo fotográfico ao colocar no centro não tanto o aparato da câmera e sim o armário de arquivo. Na seção seguinte, junto à formulação da hipótese conceitual da *imagem-rede*, volto-me a um outro percurso analítico, anterior ao digital, desenvolvido pelo historiador da arte alemão Aby Warburg, no início do século XX, para sugerir ainda outras articulações para a abordagem teórica proposta.

2.3 A imagem e a rede

Visando constituir um modo de lidar com a instabilidade ontológica que descrevi para a imagem, no nível de sua materialidade, quero nesta seção sugerir uma hipótese conceitual que, acredito, conseguiria nomear de forma conjunta alguns dos aspectos descritos. Proporei, portanto, a noção de *imagem-rede* para me referir a uma acepção que levaria em conta as relações de dependência material das imagens, considerando não apenas seu modo concreto de inscrição como, também, suas condições de circulação e, por conseguinte, de formação. Conceber a imagem como *rede* permitiria, conforme argumentarei, compreender sua multiplicidade ontológica. Permitiria também compreendê-la como inscrição que ganha corpo e sentido no âmbito de agenciamentos sociotécnicos mais amplos. Vários destes aspectos já estariam sugeridos pela evidente citação que faço à TAR da proposição da noção de *imagem-rede*. No entanto, gostaria de acrescentar uma outra perspectiva à discussão, que pode ajudar a melhor articular esta vertente teórica com uma consideração específica à dimensão cultural das imagens, em um outro registro. Volto-me, para isso, a aspectos da obra de Aby Warburg.

2.3.1 O *Atlas Mnemosyne* e os veículos de imagem de Aby Warburg

A leitura que proponho possui um recorte decididamente interessado. Não busco realizar uma recuperação detalhada seja da história de sua obra, seja de seus conceitos mais reconhecidos, como o *Pathosformel* (“fórmula de páthos”) ou o *Nachleben* (“vida póstuma”). Esforços neste sentido têm sido empreendidos por outros autores que, inclusive, me amparam nesta reapropriação (cf. ABREU, 2015; AGAMBEN, 2009b; DIDI-HUBERMAN, 2013b; LISSOVSKY, 2014; MACIEL, 2018; MICHAUD, 2013). Warburg incide

neste trabalho como um motor de invenção e inspiração metodológica, sem um compromisso específico de sustentar uma articulação metodológica ou teórica mais profunda. Apoio-me, em todo caso, em autores recentes que argumentam que haveria certo grau de pertinência na mobilização do “método warburguiano” no âmbito das imagens digitais e em rede (LISSOVSKY, 2014; MACIEL, 2018). Minha motivação para esta retomada considera reverberações entre a elaboração teórica que venho apresentando, que orienta minha investigação, e as práticas em que se basearam a investigação warburguiana sobre as imagens. Nesse sentido, interessa-me especificamente o modo com que Warburg observou as condições materiais de circulação das imagens como um componente de sua investigação e, também, como esta materialidade incide sobre sua concepção de “espaços de pensamento” no âmbito de sua biblioteca e, em especial, do seu projeto inacabado do *Atlas Mnemosyne* (Figura 3).



Figura 3: Reconstrução de pranchas do *Atlas Mnemosyne* de Aby Warburg. Registro de exposição realizada em 2016 no Zentrum für Kunst und Medien (ZKM) em Karlsruhe, Alemanha.

Fonte: ZKM, 2016. Fotografia de Tobias Wootton.

Embora tenha sido personagem central no desenvolvimento de linhagens canônicas da história da arte do século XX, tendo como seus “seguidores” dois dos principais nomes da vertente germânica da disciplina – como Ernst Gombrich e Erwin Panofsky – a obra de Warburg assumiu um lugar secundário na literatura, por boa parte do século. Este relativo “esquecimento” motivou, contudo, diferentes esforços de recuperação. Ainda na década de 1960, o historiador italiano Carlo Ginzburg (1999) revisou aspectos de sua obra em um ensaio. Nos anos 1970, foi a vez de Giorgio Agamben (2009b), em ensaio que revisaria posteriormente nos anos 1980 (cf. LISSOVSKY, 2014). Nos anos 2000, teríamos ainda outro momento desse ciclo, com sua retomada no centro do cânone francófono da

história da arte, com obras dedicadas a ele por Georges Didi-Huberman (2013b, 2015) e Philippe-Alain Michaud (2013), entre outros autores. No contexto brasileiro, uma exposição realizada no Museu de Arte do Rio em 2013, com curadoria de Didi-Huberman, foi importante na divulgação local da obra de Warburg, inclusive com a tradução e publicação no país de alguns de seus principais textos (WARBURG, 2013, 2015).

Escrevendo nesse contexto, Maurício Lissovsky (2014), sugere compreender Warburg nos próprios termos de sua teoria das imagens: como uma “imagem sobrevivente” que encontraria hoje uma “vida póstuma”. Em particular, porque a retomada contemporânea atém-se justamente a aspectos de uma sugerida *atualidade* de seu pensamento. Esta atualidade contrasta, contudo, com o lugar hoje reservado a seus principais herdeiros, Gombrich e Panofsky que, segundo elabora Lissovsky (2014, p. 311): “parecem estar do lado ‘conservador’, ‘erudito’ da história da arte, imune às necessidades e urgências da atualidade”. Em sentido similar, Hortência Abreu (2015, p. 14–15) indica como a retomada contemporânea de Warburg tem algum nível de ambiguidade, vinculando-se ora a uma atenção à tradição, ora a um interesse pela sensação de frescor de alguns de seus gestos inaugurais que, ainda hoje, “chega[m] até nós com um sabor de coisa nova”. Efetivamente, há em sua perspectiva sobre as imagens um alto grau de liberdade no tratamento histórico e crítico. Tanto que, em geral, seus leitores são francamente relutantes a sugerir algo como uma “teoria” ou um “método” warburguiano ou, quando o fazem, se abstêm de defini-lo com precisão – até porque o próprio Warburg deixou poucos subsídios para que o fizéssemos.

A respeito de sua atualidade, Lissovsky (2014, p. 306) identifica no primeiro ciclo de retomada, de Carlo Ginzburg, três aspectos que seriam pertinentes à recepção contemporânea de Warburg: a preocupação com “os limites epistemológicos das disciplinas historiográficas, o problema da semelhança na história e o da ‘atemporalidade’ ou ‘anacronicidade’ das imagens”. Com efeito, estes são alguns dos traços da leitura que dele faz Didi-Huberman, no entorno do problema do anacronismo (DIDI-HUBERMAN, 2015) e dos limites da história da arte em sua pretensão quase positivista de certeza (DIDI-HUBERMAN, 2013a). Lissovsky descreve em Warburg uma “concepção impura da imagem”, como um dos desafios que teria colocado ao cânone da história da arte, mas que hoje ganham relevância no contexto de uma “nova era global de migração das imagens”. As tecnologias digitais, após o ciclo da fotografia, sugere Lissovsky (2014, p. 321) “propiciaram uma expansão exponencial dos recursos de manipulação, processamento e

distribuição”, bem como de “apropriação, hibridação e transformação das imagens produzidas hoje e, junto com elas, de todas aquelas produzidas outrora”. Percebe-se, assim, a estranha circularidade pela qual a compreensão warburgiana da imagem, fundada há um século, com um olhar voltado para outros séculos antes dela, retorna ao presente com uma assustadora pertinência. Também hoje – ou sobretudo hoje – nos parece ser demandado olhar para a imagem como entidade impura e em constante migração. É justamente este um dos aspectos inspiradores que busco na remissão a Warburg.

Entre os desafios que levaram a uma retomada algo tardia da produção warburgiana, Lissovsky (2014) ressalta o caráter marcadamente enigmático do autor, que é exacerbado por episódios de sua biografia, como o período de sua internação psiquiátrica, ou pela apenas aparente desarticulação de seu percurso intelectual, que vai de estudos do Renascimento europeu a um interesse pontual e aparentemente circunstancial pelos rituais de povos indígenas do Novo México⁶⁹. Em larga medida, seguindo com Lissovsky (2014), alguns dos aspectos mais ousados do trabalho de Warburg teriam sido domesticados ou esvaziados na disciplina da história da arte, caminho contra o qual Georges Didi-Huberman visa fazer frente. Mais além dos conceitos, o próprio modo de fazer do “método warburgiano”, fundado em operações de *montagem*, teria se perdido pelos riscos metodológicos que assume:

Para que este conhecimento-montagem fosse possível, era preciso rejeitar as matrizes da inteligibilidade causal e criar a possibilidade de uma vertigem. Aceitar que a imagem não é um campo de conhecimento fechado, mas é centrífuga, vertiginosa. Vertigem a que Warburg nos convida, não apenas por meio de seus escritos, mas, sobretudo, por meio de seus silêncios (LISSOVSKY, 2014, p. 315).

Em especial no derradeiro projeto do *Atlas Mnemosyne*, Warburg propõe realizar a história da arte em uma “forma não discursiva”, como uma história sem texto (MICHAUD, 2013, p. 237). Seria justamente esse projeto que mais insuflaria ânimo ao olhar contemporâneo às imagens, concedendo uma maior liberdade à história e à crítica.

O *Atlas* é uma espécie de culminância da trajetória intelectual de Warburg, devendo ser lida, como sugere Michaud (2013), à luz de sua obra precedente. Trata-se de um arquivo de reproduções imagéticas pelo qual Warburg propunha um estudo *visual* da história das imagens. Em seu *Atlas*, ele dispunha, em uma sequência de pranchas, diversas reprodu-

69 Philippe-Alain Michaud (2013) indica, a esse respeito, como leituras da trajetória intelectual de Warburg tenderam a situar suas expedições ao oeste norte-americano como uma espécie de interrupção momentânea, um parêntese. Ele busca contestar, em seu livro, esta compreensão.

ções em papel (fotografias, desenhos e gravuras) de um amplo conjunto heterogêneo de imagens, originalmente disperso, espacial e temporalmente. Por meio desse dispositivo, articulava relações de semelhança e reverberações de formas visuais que atravessavam estes múltiplos objetos. Sobre a maior parte destas articulações, Warburg não deixou registros escritos, fiando-se no potencial das próprias imagens e das associações manifestas na superfície das pranchas de falarem por si próprias. Especificamente, as montagens seriam modos de dar visibilidade à emergência daquilo que o autor anteriormente se referira por *Pathosformeln* (“fórmulas de páthos”). Segundo elabora Hortência Abreu (2015, p. 35), o *Pathosformel* seria algo como a “cristalização de um movimento afetado por uma emoção, um corpo afetado”. Na leitura de Agamben (2009b, p. 132), teríamos uma espécie de unidade analítica que “designa o indissolúvel entrelaçamento de uma carga emotiva e de uma fórmula iconográfica”. Por isso, Agamben sugere que os *Pathosformeln* cristalizariam, de forma indissociável, forma e conteúdo. Embora trate-se de conceito que Warburg elabora ao longo de seus escritos, no *Atlas* o *Pathosformel* se materializaria na forma de percepções tornadas possíveis a partir dos gestos de justaposição, comparação e tensionamento realizados sobre a superfície sinóptica das pranchas.

Michaud relaciona esse procedimento com a *montagem* cinematográfica, remetendo a teoria do cineasta soviético Sergei Eisenstein. Por este caminho, ele elabora que a operação não visaria a articulação de significações, mas sim a produção de efeitos (MICHAUD, 2013, p. 322). Ele então sugere que, como na montagem de Eisenstein, seria possível compreender a montagem warburgiana como uma sintaxe ideogramática que repousaria sobre o *intervalo* como espaço de potência – isto é, como um afastamento do *conteúdo* da imagem para aquilo que se manifesta *entre* as imagens (MICHAUD, 2013, p. 326). Nesse sentido, o autor sugere uma circularidade pela qual a temática do movimento que teria inspirado os primeiros trabalhos de Warburg, voltados à representação do movimento das vestes nas pinturas renascentistas, retornaria, enfim, ao projeto do *Atlas*: “O movimento não mais se manifesta [...] pelo tratamento dos atributos externos, dobras da roupa, cabelos, mas pela aproximação entre dois elementos visuais heterogêneos e pela utilização da montagem no plano” (MICHAUD, 2013, p. 240). Lissovsky (2014, p. 319), a partir de Agamben, indica, porém, que não se trata apenas de algo como a recomposição de uma trajetória histórica linear. De modo fundamental, sua leitura aproxima Warburg de Foucault em uma compreensão não-linear da história, trasladada a uma dimensão espacial. Referindo-se ao painel 46 do *Atlas* (Figura 4), dedicado às nin-

fas, Lissovsky (2014, p. 319) retoma a leitura que lhe dedica Agamben e destaca que, na superfície da prancha, “nenhuma imagem precede às demais”, sendo “impossível distinguir ali entre criação e performance, entre original e cópia”. O princípio seria o do *paradigma*, segundo aquilo que Agamben (2002) elabora, a partir de Foucault: um gesto metodológico que opera pela singularidade como forma de oferecer inteligibilidade ao conjunto. As ninfas seriam expressas, nesse sentido, pelo conjunto de imagens e por cada uma delas, de modo indissociável, “em um misto de primariedade e repetição e não simplesmente cópias umas das outras”, como elabora Abreu (2015, p. 53).



Figura 4: Prancha 46 do Atlas Mnemosyne de Aby Warburg.
 Registro da montagem da prancha original em versão de outubro de 1929, na Biblioteca Warburg.

Fonte: Instituto Aby Warburg.

No projeto warburgiano de uma *história das imagens*, seu interesse não recai sobre a imagem como objeto único de elevação estética mas sim, como sugere Lissovsky (2014, p. 317) – a partir de Agamben, Didi-Huberman e Ginzburg – como “marca” ou “vestígio”, isto é, como processo que incorpora, inclusive, uma temporalidade complexa no interior de si própria. O potencial de vertigem que Lissovsky aponta para o método warburgiano residiria, então, na operação da montagem e nesta acepção complexificada da ontologia das imagens. Por sua lente, a imagem não se fixa nem se localiza. O método, pelo contrário, tende a desestabilizá-la e a colocá-la em movimento. Em um plano teórico, isto se manifesta no distanciamento que Warburg assume em relação à tradição estetizante da história da arte que concebia a imagem em seu aspecto plástico-formal ou segundo sua vinculação à psicologia ou ao gênio criativo do artista. De outro modo, como salienta Agamben (2009b), Warburg localiza a imagem como localizada em algum ponto entre a arte e a religião, em relação estreita com a memória em sua dimensão coletiva.

Evidentemente, a aproximação entre esta vertente do estudo das imagens e a investigação proposta por esta tese não é automática. Meu gesto de aproximação, aqui, talvez deva ser compreendido de modo similar ao próprio método warburgiano: “não é apenas o deslocamento de um ponto a outro, mas salto, montagem, repetição e diferença” (LISSOVSKY, 2014, p. 315). Como alertei ao início desta seção, não visio sugerir aqui uma herança warburgiana à Teoria Ator-Rede, nem postular uma leitura de Warburg à luz da TAR⁷⁰. A montagem teórica que realizo visa encontrar um caminho entre estas perspectivas que as complexifique mutuamente. Isto é, enfatizando aspectos em Warburg que em alguma medida remeteriam (anacronicamente) à TAR, e inflexionando os aspectos da TAR, que desenvolvi, com uma compreensão adensada da imagem.

Um aspecto seria a possível aproximação da compreensão warburgiana da imagem de uma *multiplicidade ontológica*, embora de forma distinta da elaborada por Mol (1999). A imagem seria em Warburg um “processo” e não “resultado de processos” (LISSOVSKY, 2014, p. 317). Seria um efeito decorrente da montagem e não uma simples associação de sentidos (MICHAUD, 2013, p. 322). A imagem da ninfa (para além da imagem-coisa) emerge no *Atlas* como o efeito da justaposição e do tensionamento entre suas múltiplas e

70 Um aspecto curioso, em todo caso, é o recurso a montagens de estilo warburgiano como dispositivo expográfico da exposição *Reset Modernity!*, com curadoria liderada por Bruno Latour, realizada no ZKM em 2016 (LATOURE, 2016). A pertinência desta aproximação não seria, ao que parece, tão improvável.

singulares manifestações, o que não significa que sejam manifestações em competição nem perspectivas plurais sobre uma imagem (MOL, 1999). A ninfa seria múltipla e a montagem da prancha teria a potência de colocá-la em evidência (Figura 4).

Um segundo aspecto poderia ser sugerido no tratamento da imagem (imagem-coisa) como *inscrição*. A montagem, em si mesma, é possibilitada pela mobilização de reproduções em papel (fotografias, desenhos, gravuras) de pinturas, esculturas, selos, moedas e outros objetos. Estas reproduções têm sentido bem próximo daquilo que Latour (1986, 2001) descreveu como “móveis imutáveis”. Guardando uma relação material com as coisas do mundo que representam, as inscrições em papel possibilitam o acúmulo, a justaposição, a sobreposição e a comparação entre si, reunidas e dispostas sobre um mesmo plano comum. A projeção plana das reproduções – seja pelo registro em perspectiva do objeto tridimensional, seja pela natureza bidimensional do ‘original’ – as colocam em um local que permite a comparação (AKRICH, 1992), embora tenham sido produzidas em contextos e técnicas diversas. Auxilia-nos nesta aproximação a descrição de elaborada Michaud (2013, p. 321):

Em *Mnemosyne*, a reprodução fotográfica não é mais um suplemento, e sim um equivalente plástico geral a que são remetidas todas as figuras, antes de serem dispostas no espaço da prancha. Assistimos a três operações sucessivas de transformação do material de origem: os objetos de diversas naturezas (pinturas, relevos, desenhos, impressos etc.) são unificados pela fotografia, antes de serem reunidos nas pranchas, as quais, por sua vez, são refotografadas para criar uma imagem única.

Temos, então, a fotografia como chave que traduz as diferentes imagens a um “equivalente plástico geral”. Nos termos de Madeleine Akrich (1992), diríamos da fotografia como chave comum que condiciona a exportabilidade destas inscrições e sua comparação em um mesmo plano. Reunidas na Biblioteca Warburg e dispostas sobre a superfície das pranchas, as imagens encontram ali um “centro de cálculo” (LATOUR; HERMANDT, 2004), capaz de produzir, a partir delas, outras ontologias.

Seria importante observar, a esse respeito, como as dependências materiais da investigação foram conscientemente articuladas por Warburg em seu pensamento acerca da biblioteca como um espaço de pensamento constituinte de seu método. Como sugere Philippe-Alain Michaud, a Biblioteca Warburg – ele lhe deu seu próprio nome – era como uma objetivação de seu pensamento. A disposição e a disponibilidade dos livros, documentos e imagens eram consideradas condições fundamentais do trabalho. Em ci-

tação de Warburg por Michaud (2013, p. 233), a partir de Stockhausen (1992, p. 51, citado em MICHAUD, 2013), temos uma descrição da articulação entre o método e a disposição espacial da biblioteca:

A novidade do meu método prende-se a que, para explicar a psicologia da criação artística, reúno documentos provenientes do campo da linguagem, bem como das artes plásticas ou do mundo do drama religioso. Para consegui-lo, eu e meus companheiros de pesquisa precisamos ter diante de nós os documentos, isto é, livros e imagens, dispostos em grandes mesas, a fim de podermos compará-los, e esses livros e imagens devem estar ao alcance da mão, sem dificuldade e instantaneamente. Por isso necessito de uma verdadeira *arena* com mesas, para ter à mão os livros comuns e o material iconográfico.

O pensamento warburguiano era, assim, debitário de sua biblioteca. À configuração desta, Warburg creditava a inovação de seu método. Latour (1986) elabora argumento similar em sua abordagem das inovações técnicas que teriam possibilitado a conformação da ciência moderna. Diante da descrição feita por Warburg das grandes mesas nas quais dispor dos livros e imagens, nos retorna a situação, narrada por Latour (2001), dos cientistas que analisavam um mapa da região amazônica sobre a mesa de um restaurante em Boa Vista. Embora os objetos estudados e as condições espaciais e institucionais sejam marcadamente distintas entre um caso e outro, há uma analogia possível entre a “arena” de mesas de Hamburgo e a mesa de Boa Vista. A operação é bastante similar: o momento da produção de conhecimento e do diálogo entre pesquisadores é constituído em uma relação direta com inscrições coletivamente manipuladas. Ainda a respeito da Biblioteca Warburg e de suas inquietas sessões de trabalho, escreve Michaud (2013, p. 233–234):

A coleção inteira de livros era, ao mesmo tempo, a objetivação de seu pensamento e uma alegoria do mundo e dos corpos que nele se movem. A classificação das fichas e o deslocamento ao longo das estantes eram um rito de orientação, cujo modelo Warburg tinha concebido no correr de sua viagem e que reproduzia de maneira incansável na cosmologia confinada da biblioteca.

Michaud sugere que a constante reorganização dos livros, documentos e imagens seriam como um “rito de orientação”, o que sugere que a própria biblioteca poderia ser compreendida como dispositivo de navegação. Michaud (2013, p. 237–238) articula, em outra passagem, uma continuidade entre o projeto da biblioteca e o *Atlas*⁷¹, que, assim, poderia cumprir função similar. Jane Maciel (2018, p. 198), interessada na força da “for-

71 “Tal como fazia com os textos no espaço de sua biblioteca, ele atualizou, dessa vez com imagens, os conflitos sem resolução dos quais extraía o sentido da história das formas, bem como o sentido de seu próprio pensamento” (MICHAUD, 2013, p. 237–238).

ma Atlas” para estudar as imagens no contexto das redes digitais, enfatiza nas pranchas do *Atlas* a configuração de constelações de imagens, paralisadas no transcurso de seu movimento e servindo, elas próprias, para operações de navegação e orientação “entre as imagens do mundo e o mundo das imagens”. O saber das imagens, em Warburg, realiza-se, assim, por meio de inscrições materiais catalisadas com sua disposição sobre o plano.

A partir de Didi-Huberman (2013c, p. 21) e Philippe-Alain Michaud (2013, p. 9), parece-me também relevante considerar como a materialidade dos processos fotográficos de reprodução e circulação das imagens impactou não apenas o projeto warburgiano, mas também o ensino acadêmico de arte na virada do século XIX ao XX. A respeito de Warburg, Lissovsky (2014) salienta como a fotografia teria sido central ao projeto do *Atlas*, como vimos também, acima, na citação a Michaud. A descrição de Trevor Fawcett (1983, p. 450) permite, contudo, que ampliemos a consideração sobre o impacto da fotografia para os estudos da imagem em uma perspectiva mais ampla:

As obras de arte discutidas em uma palestra normalmente não estavam disponíveis em toda sua imediatez; os originais eram inacessíveis, espalhados por muitos países, frequentemente conhecidos apenas de segunda ou terceira mão. Asserções verbais, se sequer fossem ser checadas, tinham que ser verificadas a partir de transcrições não confiáveis, provavelmente em escala reduzida e realizadas em um outro meio, produzidas por intermediários humanos que inevitavelmente reinterpretavam os originais. A ilustração das palestras era menos um fac-símile que um diagrama, que era o que alguns palestrantes do século XIX o chamavam; útil ou suficiente para explicar pontos gerais do estilo, iconografia e composição, mas não para análises mais sutis e comparações detalhadas de um trabalho com o outro⁷².

As reproduções fotográficas, tornadas progressivamente disponíveis a partir de meados daquele século, apareceram enquanto alternativas promissoras a tais inscrições que eram, ao mesmo tempo, dispendiosas e insatisfatórias. Elas substituíam gravuras realizadas como versões de circulação de grandes pinturas, por exemplo, que eram a única *inscrição* colecionável por grandes bibliotecas⁷³. Contudo, enquanto na forma de diaposi-

72 No original: “*The works of art discussed in a lecture were usually not available in all their immediacy; the originals were inaccessible, scattered through many countries, often known only at second or third remove. Verbal assertions, if they were to be checked at all, had to be measured against unreliable transcripts, probably on reduced scale and executed in an alien medium, produced by human intermediaries who inevitably reinterpreted the originals. The lecture illustration was less a facsimile than a diagram, which is what some nineteenth-century lecturers called it; useful enough for explaining general points of style, iconography and composition but not for subtler analysis and detailed comparison of one work with another*”. Tradução minha.

73 Um exemplo desta prática é a coleção de gravuras da abadia de Göttweig, na Áustria, que foi, em seu tempo, uma das maiores coleções particulares de gravuras da Europa com a finalidade de subsidiar a formação artística na instituição. Hoje, o monastério abriga o departamento de *Image Science* da Universidade do Danúbio em Krems, inclusive com projetos de digitalização desse acervo. Cf. <http://www.stiftgoettweig.at/>.

tivos ou grandes reproduções permitia-se reconstituir aspectos da uma experiência única da obra analisada, Warburg levou a um extremo as possibilidades de recombinação destes registros:

Por intermédio dela, Warburg criava à sua volta um universo cósmico onde todas as imagens-astros se equivaliam, independente de seu tamanho, distância e natureza, agrupando-se contra o fundo escuro do céu zodiacal como constelações cintilantes em torno destes estranhos atratores, que denominava fórmulas do patético (LISSOVSKY, 2014, p. 320–321).

Avançando em outro aspecto, em uma dobra da materialidade do método sobre seu próprio objeto, encontramos mais uma circularidade do percurso warburguiano. Pois a importância das dependências materiais não são observadas apenas como condicionantes dos dispositivos epistêmicos da biblioteca ou do *Atlas*, mas também são parte importante da própria concepção teórica que Warburg dedica às imagens. Notadamente, Warburg observara a relevância da tapeçaria, com sua mobilidade e reproduzibilidade características, para a constituição do estilo do período renascentista ao longo do continente europeu. Ou seja, não só no âmbito da investigação, mas da própria realidade investigada, Warburg dedicava atenção às condições materiais de sua efetuação. Nos fragmentos que compõem sua introdução ao projeto do *Atlas*, Warburg (2015, p. 372, ênfase minha) escreveu:

O tempo entre Piero della Francesca e a escola de Rafael é uma época na qual começa a *perambulação internacional intensiva das imagens* entre norte e sul, cuja veemência elementar envolve tanto o ímpeto do impacto como a abrangência do domínio por onde perambula — algo que se furtou ao historiador europeu dos estilos sob a “vitória” oficial do alto Renascimento romano. A *tapeçaria de Flandres é o primeiro tipo, ainda colossal, de veículo automotivo para o transporte de imagens, que, desprendido da parede — e não só pela mobilidade, mas também pela técnica, voltada à reprodução multiplicadora do conteúdo da imagem —, foi um precursor da folha de papel impressa com imagens, isto é, das gravuras em cobre e xilogravuras, que mormente tornariam o intercâmbio de valores expressivos entre norte e sul uma ocorrência vital no processo de circulação da formação do estilo na Europa.*

Nesta passagem, o estudioso faz menção ao que Heil e Ohrt (2016) indicam como um dos conceitos-chave do pensamento warburguiano: os veículos de imagem (*Bilderfahrzeuge*). A tapeçaria, ele observa, permitiu que as imagens então realizadas sobre mídias fixas, pintadas diretamente sobre a superfície das paredes, destas se desprendessem e circulassem. O estilo artístico do período, como formação transversal à Europa renascentista, demandaria que considerássemos este substrato como um dos condicionantes de sua emergência. Isto é, indo além das questões tipicamente colocadas à história dos estilos, Warburg estabelece uma relação fundamental entre uma economia material da circula-

ção das imagens e o seu processo de constituição. Seria justamente esta economia material, que permitiria, séculos depois, a deriva das imagens em que se fundamenta seu método.

Lissofsky (2014) articula sua discussão da obra de Warburg com uma indagação quanto aos motivos pelos quais o seu pensamento seduz os pensadores contemporâneos da imagem. Posso responder por mim. A aproximação que faço tem por objetivo inspirar o gesto metodológico e uma concepção das imagens contemporâneas que, como já indiquei à introdução, compreendo em um processo articulado entre: o volume e a ampla circulação das imagens, no âmbito da chamada “virada visual” ou “virada pictórica” das plataformas de mídia social; as mediações técnicas que modulam as dinâmicas de circulação das imagens nesse contexto; e as mediações técnicas que possibilitam seu estudo. A respeito destas últimas, volto-me em particular às técnicas de aprendizado de máquina por redes neurais, aplicadas às imagens, como forma de lidar com a proliferação contemporânea do visível. À luz de Warburg e do modo como ele mobilizou os “veículos de imagem” de seu tempo e do passado em sua investigação, diria, então, que um dos objetos de minha pesquisa seria, também, a operação dos veículos de imagem contemporâneos e os desafios que colocam a um tal estudo.

Hortência Abreu (2015, p. 15) indica como a aproximação do trabalho de Warburg e a noção mais recente do *hipertexto* não são raras nos esforços de atualização de sua obra. Martin Warnke, um dos principais especialistas na obra de Warburg, professor da Universidade de Leuphana, nos arredores de Hamburgo, desenvolveu um projeto denominado *HyperImage*⁷⁴, em que busca transpor a um *software* computacional as operações de montagem do método warburguiano, potencializado pelas possibilidades de *linkar* conteúdos a partir da superfície da tela. Contudo, minha apropriação de Warburg não se resume à sua operação de montagem cristalizada no *Atlas*. Interesse-me, sobretudo, na mobilização consciente de *inscrições*, de *veículos de imagem* contemporâneos tempo, para investigar as formações imagéticas produzidas, elas próprias, em sua circulação. Isto é, volto-me a Warburg como inspiração não apenas metodológica, mas também teórica, em seu modo de compreensão da imagem como efeito dessa migração ou perambulação das imagens, e como unidade paradigmática, formada em um processo que talvez possamos descrever como uma *emergência* a partir de singularidades irreduzíveis.

74 Cf. <http://www.uni-lueneburg.de/hyperimage/hyperimage/ebsKart.htm>.

veis. Junto ao que vim elaborando acerca de inspirações da TAR a este trabalho, a obra de Warburg sugere, assim, ainda outros aspectos para o que proponho denominar como *imagem-rede*.

2.3.2 Imagem-rede: uma hipótese conceitual

Pelo modo como enuncio o conceito – *imagem-rede* – trata-se, obviamente, de uma apropriação direta da formulação proposta inicialmente por Michel Callon (1986) e que veio a nomear a vertente teórica que ele ajudou a fundar: a Teoria Ator-Rede (TAR)⁷⁵. Ao assumir basicamente a mesma forma em minha proposição, busco incorporar uma reversibilidade similar à subsumida pela conjunção entre ator e rede, sugerindo algo como uma dupla concepção para a imagem ou, ainda, uma tensão irresoluta. Por um lado, a *imagem é rede*, pois constitui-se de forma relacional, como vim salientando. Trata-se de um objeto ontologicamente fraturado, múltiplo e amplamente dependente, não apenas de agenciamentos sociotécnicos como, também, de modos de *inscrição*, *veículos de imagens* e de outras imagens junto às quais produz efeitos duradouros na memória coletiva. Por outro lado, a *rede é também imagem*, pois é apenas enquanto imagem que ela pode ser apreendida, adquirindo uma forma minimamente reconhecível. Isto não equivale, como alertam Tommaso Venturini, Anders Munk e Mathieu Jacomy (2018), a dizer que a rede seria sua *visualização* – estas são instâncias distintas de sua manifestação. A imagem da rede, nesse sentido, tampouco deveria ser compreendida em sua forma gráfica. A rede é imagem porque ela não é um dado empírico imediato. Como ressalta Latour (2005, p. 131): “[A] rede é um conceito, não algo no mundo. Ela é uma ferramenta para descrever algo, não o que é descrito”⁷⁶. Nesse sentido que afirmo, portanto, que a *rede é apenas concebível quando imaginada, ou projetada*.

75 Para além desta inspiração, a conjunção vocabular hifenizada faz também ressoar outras elaborações canônicas dos estudos da imagem. Provavelmente o primeiro a fazê-lo, Deleuze sugeriu que a passagem do cinema clássico ao cinema moderno poderia ser compreendida pela passagem da *imagem-movimento* à *imagem-tempo* (DELEUZE, 1985, 1990). A primeira noção advém, em particular, da leitura que ele faz de Henri Bergson, a partir do qual concebe uma relação de identidade entre imagem e movimento. Como no caso do *ator-rede*, portanto, também para a *imagem-movimento* o hífen poderia ser compreendido como sinal de igualdade, como Deleuze efetivamente faz em uma passagem: “IMAGEM = MOVIMENTO” (DELEUZE, 1985, p. 78). Sugestivamente, o filósofo (DELEUZE, 1985, p. 81) também descreve uma analogia pela qual a identidade entre imagem e movimento poderia ser compreendida como a identidade entre matéria e luz, logo antes de indicar algumas das implicações da teoria da relatividade geral de Einstein para a filosofia de Bergson. Em uma coincidência que reforça a topologia enredada da própria formulação da *imagem-rede*, esta é justamente a analogia mobilizada por Latour em ao menos dois momentos (1999, p. 18–19, 2011, p. 800) ao discutir a proposição *ator-rede*: “Tentar seguir um ator-rede é um pouco como definir uma onda-corpúsculo nos anos 1930: qualquer entidade pode ser tomada ou como um ator (um corpúsculo) ou como uma rede (uma onda)” (LATOUR, 2011, p. 800). Temos em ambos os casos, nesse sentido, proposições conceituais fundadas não na certeza, mas na oscilação.

76 No original: “Network is a concept, not a thing out there. It is a tool to help describe something, not what is being described”. Tradução minha.

Efetivamente, bem antes da TAR, a rede é uma imagem filosófica recorrente, que nos permite conceber relações e sistemas complexos. Como sugere o filósofo Pierre Musso (2004, p. 17), a rede é antes de tudo um “receptor epistêmico”. Uma de suas primeiras formulações, elabora Musso, seria, em uma aproximação da *renda* para compreensão do corpo – para dizer da pele e, já mesmo em Descartes, para descrever a estrutura do cérebro. Do século XVIII ao XIX, segundo Musso (2004, p. 20), haveria uma passagem fundamental, que chega ao nosso presente, em que a rede deixa de dizer de um elemento interno ao corpo para se exteriorizar como forma construída – “como artefato técnico sobre o território”. Porém, elabora: “Para sair de sua relação com o corpo físico, a rede devia, primeiramente, ser pensada como conceito para se tornar operacional como artefato” (MUSSO, 2004, p. 20). Enquanto conceito, falaríamos menos da rede como coisa, mas como um “efeito de rede”, que emerge na interseção entre domínios díspares como a medicina, a economia política da circulação e as engenharias. Segundo desdobra Musso, haveria nesse processo uma dinâmica que vai da imaginação conceitual da rede ao artefato-rede em um processo de contaminação mútua. Enquanto figuração imaginada, a rede sugere utopias de associação universal, que projetos como a internet almejavam em alguma medida concretizar. Enquanto formação técnica e tecnológica, contudo, temos materializações singulares que não se igualam à proposição conceitual mas que incorporam, imaginariamente, alguns de seus aspectos.

Em outra fase de sua obra, posterior a seu engajamento com a TAR, Bruno Latour (2013) retoma a figura da rede sob outro registro que me parece interessante para adensar esta consideração da relação entre *imagem* e *rede*. Latour identifica a rede, nesse contexto, como o primeiro de quinze modos de existência que compõem seu ambicioso ensaio acerca da antropologia dos modernos. Nesse trabalho, Latour propõe voltar-se ao que denomina “erros de categoria” que teriam fundado a visão de mundo moderna. Sua abordagem busca a multiplicidade dos modos de existência que compõem o mundo moderno, ou seja, a diversidade de categorias de seres que o habitam – tais como seres de religião [REL], lei [LAW], metamorfose [MET], tecnologia [TEC] e ficção [FIC]⁷⁷. Nessa proposição, os seres da rede [NET] compreendem um primeiro modo de existência, que institui uma maneira para a pessoa investigadora percorrer os demais modos de existência em seus processos de translação. A rede seria pois, também nesse sentido, uma ferramenta ou um ente operador da análise (LATOUR, 2013, p. 33). Escreve:

77 A fim de traçar uma distinção entre o uso comum das palavras e a denominação conceitual dos modos de existência, Latour vale-se destas abreviações de três letras entre colchetes.

A ESSÊNCIA de uma situação, por assim dizer, será, para um [NET], a lista de outros seres pelos quais faz-se necessário passar para que tal situação possa perdurar, ser prolongada, mantida, ou estendida. *Traçar uma rede* é, então, sempre reconstituir por uma PROVAÇÃO (uma investigação é uma provação, mas também o é uma inovação, como também uma crise) os antecedentes e as consequências, os precursores e seus herdeiros, as entradas e as saídas, por assim dizer, de um ser. Ou, para colocá-lo em termos mais filosóficos, os outros pelos quais um ente deve passar de modo a tornar-se ou permanecer o mesmo – o que pressupõe, como logo mais veremos, que ninguém pode simplesmente “permanecer o mesmo”, por assim dizer, “sem fazer nada”. Para permanecer, um ente deve passar – ou, em todos os casos, “passar por” – algo que chamaremos TRANSLAÇÃO (LATOUR, 2013, p. 41, ênfases no original)⁷⁸.

Em um tom mais claramente filosófico, interessado na ontologia dos seres modernos, Latour retoma noções centrais à TAR – além da *rede*, há as noções de *translação* e de *provação*⁷⁹. Porém, indo além da rede como ferramenta conceitual para a descrição de um arranjo atorial específico, a *rede*, na forma [NET], aponta também para um modo de existência que funda a possibilidade de uma *identidade* estável (metaestável, talvez) por meio de um trabalho de constante *translação*. Isto é, sob a noção da [NET], o ser apenas teria algo a que pudéssemos chamar de uma essência através de instâncias efetivamente inessenciais, que o ultrapassam. Trata-se, me parece, de uma forma mais condensada e articulada de conceber aquilo que indiquei, ao longo do capítulo, acerca das categorias da *mediação*, *inscrição* e *multiplicidade ontológica*. O modo de existência [NET], ao perpassar os demais, seria uma qualidade dos seres e não uma categoria analítica como sugere-se, inicialmente, com a proposição *ator-rede*.

Conceber a imagem como [NET] ou como *imagem-rede*, portanto, implica compreender a instabilidade constitutiva de sua ontologia, como um ente inessencial. Ou, de outro modo, compreendendo sua essência a partir dos “outros seres pelos quais faz-se necessário passar para que tal situação possa perdurar”. A imagem enquanto rede, nesse sentido, poderia nos remeter à sua constituição como *inscrição*, como *móvel imutável*, que ganha sentido e agência por dinâmicas de circulação e associação. Poderia, também, assumir a figura do *Pathosformel* warburgiano, como “formação” – e não tanto como “forma” – que emerge como efeito das relações e tensões estabelecidas nas pranchas do

78 No original: “The essence of a situation, as it were, will be, for a [NET], the list of the other beings through which it is necessary to pass so that this situation can endure, can be prolonged, maintained, or extended. To trace a network is thus always to reconstitute by a TRIAL (an investigation is a trial, but so is an innovation, and so is a crisis) the antecedents and the consequences, the precursors and the heirs, the ins and outs, as it were, of a being. Or, to put it more philosophically, the others through which one has to pass in order to become or remain the same – which presupposes, as we shall see later on, that no one can simply ‘remain the same,’ as it were, ‘without doing anything.’ To remain, one needs to pass – or at all events to ‘pass through’ – something we shall call a TRANSLATION”. Tradução minha.

79 Tiago Salgado (2019) localiza a noção de *provação* (no francês, *épreuve*) como um dos eixos pertinentes às sociologias pragmáticas francesas, ramo integrado pela TAR.

Atlas. Nesse sentido, a rede ativa um movimento centrífugo em relação às imagens. Mas ela também incide sobre elas em sentido centrípeto, em um jogo de tensões que as imagens internalizam em seu processo de formação. A rede enquanto imagem, por sua vez, remete à sua condição de “receptor epistêmico”, nos termos de Musso (2004), organizando uma percepção das relações entre elementos heterogêneos. A rede seria uma instância imaginada, pela qual as relações entre os entes é traçada. Outra característica da rede como imagem pode ser sugerida por uma consideração de Philippe-Alain Michaud (2013, p. 240) acerca do *Atlas*:

Os painéis cobertos de tecido preto [...] não eram feitos para ser expostos como tais, e sim para serem fotografados, a fim de formarem uma nova entidade complexa. Portanto, não devem ser apenas apreendidos em seu conteúdo [...]. Também é preciso vê-los em sua configuração material [...].

Warburg constantemente alterava as pranchas, ao que autores como Michaud sugerem se tratar de uma metamorfose constante do arquivo, mas que Heil e Ohrt (2016, p. 26) contestam ao dizer que, embora o tenha deixado inacabado, Warburg buscava alcançar em seu *Atlas* configurações precisas. Em todo caso, independente da pretensão ou não de acabamento, as formações seriam decorrentes de processos transitórios. A rede como imagem, nesse sentido, seria como a rede que é mediada por registros fotográficos tomados como estâncias transitórias de um *devenir*, que não se encerra no momento da observação.

Esta concepção implica, então, um enquadramento específico para o problema ontológico que venho trazendo desde o início deste capítulo. Pela formulação *imagem-rede*, busco abraçar a instabilidade que, em outra perspectiva, poderíamos ser levados a *fixar*. Como argumentei ao início, esta compreensão não se volta apenas a um sentido fenomenológico, semiótico ou psicanalítico da imagem, pois incide sobre sua própria materialidade. Mesmo o ser da imagem enquanto coisa, portanto, não é facilmente resolvido. Dois aspectos fundamentais, a esse respeito, como vim desenvolvendo, são a sua condição como *inscrição sociotécnica*, pela qual a imagem deve ser compreendida em cadeias de translação mais extensas; e a *multiplicidade ontológica* pela qual o ser da imagem seria um efeito de suas múltiplas manifestações. A inflexão trazida a esta compreensão pela perspectiva warburguiana, para além desta cadeia de dependências materiais, sugere uma complexidade temporal que, aparentemente, escapa à TAR. Segundo desenvolve Agamben (2009b, p. 136), para Warburg a imagem seria como uma “herança transmitida pela memória social”, na qualidade de uma carga energética cristalizada.

Pela mediação da imagem, portanto, talvez possamos compreender como as cadeias de ação e translação podem se estender por temporalidades diversas que não se resumem ao presente.

Uma consequência deste enquadramento à questão ontológica é que a indagação quanto ao que são as imagens que vemos e que investigamos não pode ser respondida de forma unitária e, nem mesmo, de forma definitiva. Parece-me sugestiva, nesse sentido, a proposição feita por Gabriel Menotti (2019, p. 25), acerca do cinema, de compreender suas imagens não como “formas que circulam, mas [como] formas que resultam da circulação”⁸⁰. Circulação esta, salienta, que não se resume aos espaços canônicos do dispositivo cinematográfico, mas que inclui também seus “espaços negativos” – redes ilegais de cópia e distribuição, mídias de armazenamento, instâncias logísticas da exibição em cinemas e festivais. Ou seja, o ser da imagem não pode ser localizado. Ele precisa, de outro modo, ser mapeado, performativamente, na análise. Menotti (2019, p. 28), a esse respeito, nomeia seu esforço metodológico de investigação como uma *ontografia*, concebendo seu estudo do cinema como “uma entidade em contínua transformação”, em que a própria investigação encontra-se implicada. No seu estudo, as “ferramentas ontográficas” que mobiliza são oriundas da prática curatorial, concebendo a exposição e a exibição cinematográficas como laboratórios de investigação em que os objetos expostos performam um vir-a-ser (MENOTTI, 2019, p. 100). Há, nesta operação, algo próximo daquilo que descrevi em Warburg, em que a montagem produz efeitos de emergência dos objetos visuais. Indo de um enquadramento *ontológico* a um enquadramento *ontográfico*, a questão torna-se menos aquilo que a imagem é, portanto, do que aquilo que ela *se torna*. A investigação se voltaria, desse modo, ao mapeamento situado destes processos de translação, criando dispositivos capazes de ativá-los e registrá-los.

Michael Lynch (2013), no âmbito dos STS, sugere a ontografia como uma estratégia para desinflar a questão ontológica. Sua proposição se situa em um momento em que a que a ontologia se tornou um foco das atenções de estudiosos dos STS. Em síntese, Lynch propõe o termo *ontografia* para nomear esforços de investigação que se voltam a práticas pertinentes a uma compreensão ontológica do mundo – práticas de ‘produção do mundo’ (*world-making*) e ‘sustentação do mundo’ (*world-sustaining*) – que assumem uma perspectiva simétrica, sem definir de antemão, o que contaria como identidade e como diferença. Ele propõe, portanto, um outro sentido ao princípio da simetria, que in-

80 No original: “forms that circulate, but rather forms resulting from circulation”. Tradução minha.

diquei anteriormente. O que faz do ser igual a ele mesmo, desse modo, não poderia ser definido de antemão e, sim, investigado em um esforço situado. Em linha com o trabalho teórico de Mol (1999), Lynch sugere uma compreensão performativa da ontologia, que, inclusive, torna difusas as fronteiras entre a questão epistemológica e a questão ontológica, pois, em última medida, o processo de produção do saber seria também um processo de produção de seus objetos. A *ontografia* seria uma investigação que leva em conta a proeminência da questão ontológica para determinado estudo de caso e busca compreender as diferentes maneiras pelas quais a ciência e a tecnologia buscam estabilizar esta questão.

Como buscarei explicitar nas páginas a seguir, um dos grandes desafios ao estudo das imagens em plataformas *online* pode ser localizado justamente neste ponto: o que vale como uma imagem em diferentes abordagens. Não é uma questão facilmente respondida pois, mesmo em cada caso específico, a imagem, mesmo em seu nível material, declina-se em múltiplas instâncias e demanda que articulemos várias entidades para que possa ser recomposta. Em especial quando recorremos a técnicas computacionais para dar conta de grandes corpora visuais, o estudo multiplica as formas de inscrição das imagens, que não são as mesmas em cada etapa do processo. Esforços como os de Warburg, no entorno de uma montagem sinóptica destas variações, são inspiradores, me parece, da tarefa ontográfica demandada. O problema do método pode ser então caracterizado como o de constituir “centros de cálculo” em que estas múltiplas inscrições podem ser justapostas, comparadas e tensionadas a fim de compor *imagens-redes*, em suas múltiplas dependências e declinações.

3 Visualidades computacionais

No capítulo anterior, o argumento que persegui foi o de que não seria possível postular uma ontologia estável para as imagens, especialmente com base em seus processos de gênese. Trata-se de algo que tem particular pertinência para as imagens digitais, mas que também se aplica para imagens anteriores a esse registro. Distanciando-me de um olhar que visa a essência ou a unicidade das imagens, busquei, em uma aproximação a aspectos da TAR e, em menor medida, do pensamento warburgiano sobre as imagens, uma compreensão que as considere em sua multiplicidade. Isto é, como efeito de uma rede de mediações técnicas e de diferentes formas de inscrição que operam em conjunto. Nesse sentido, propus a noção de *imagem-rede*, que precisaria ser mapeada no exercício situado de uma *ontografia*. Neste capítulo, avanço sobre aspectos teóricos e metodológicos pertinentes às imagens digitais sob esta compreensão. Em particular, busco observar as mediações computacionais que pelo menos desde a última década vêm sendo mobilizadas em esforços de pesquisa dedicados a estes objetos. Estas mediações também vêm sendo implementadas em plataformas digitais como forma de integrar a dimensão visível das interações *online* aos processos de datificação e mediação algorítmica que lhe são constitutivos.

Situo esta etapa da investigação no entorno da noção de *visualidade*. Indiquei, à Introdução, uma definição para este termo a partir de Hal Foster (1999), para quem se trata da nomeação de uma diferença interna ao *visual*, que enfatizaria sua constituição social e histórica, contra uma tendência de essencialização presumida pelo termo *visão*. Este sugeriria uma ênfase mais física ou fisiológica que, por sua vez, tenderia a um ‘universal’. A reivindicação de Foster pelo termo *visualidade*, portanto, é também a demarcação de um programa de investigação que compreende *modos de ver* – talvez seja possível dizer, também, *visões de mundo* – social e historicamente situados. A esse respeito, Foster sugere que a tarefa dos estudos da imagem e da arte é buscar, a todo tempo, desfazer essa superposição que tenderia à naturalização dos processos históricos do campo visual. Escreve ele:

Com sua própria retórica e representações, cada regime escópico busca desfazer essas diferenças: fazer de suas muitas visualidades sociais uma visão essencial, ou ordená-las numa hierarquia natural da visão. É importante, en-

tão, deslocar estas superposições para fora de foco, perturbar o arranjo dado de fatos visuais (talvez seja, absolutamente, a única forma de vê-los)...⁸¹ (FOSTER, 1999, p. ix).

Nesse sentido, a dinâmica entre *visão* e *visualidade* não é bem uma de complementariedade, como se uma visão ‘natural’ fosse em dado momento ‘socializada’. Em sentido oposto, a *visão* seria decorrente de um movimento de cristalização e estabilização do campo visual que subsumiria os atritos entre diferentes modos de ver, a cada momento. O engajamento com as mediações técnicas que permitem a integração de imagens às plataformas digitais seria, portanto, uma questão pertinente aos muitos modos de ver dirigidos, hoje, às imagens. Ao sugerir a ideia de *visualidades computacionais*, quero observar a incidência de tecnologias computacionais de mediação do visível na constituição do campo visual, tomando estas como elementos que seriam pertinentes a uma discussão sobre as imagens e as visualidades contemporâneas.

Em minha dissertação de mestrado (MINTZ, 2015) desenvolvi aspectos desta hipótese de uma vinculação entre aparatos técnicos à constituição de visualidades. Voltei-me, em especial, ao trabalho de Jonathan Crary (1992), em sua revisão histórica da figura do *observador*. Sua abordagem é informada pela compreensão foucaultiana da técnica como instância de subjetivação. Por esse ponto de vista, Crary articula transformações teóricas e práticas estéticas, que emergiram a partir da primeira metade do século XIX, em relação com alguns brinquedos ópticos que eram populares naquele período, como o praxinoscópio, o zootrópio e o estereoscópio. Um aspecto característico destes objetos era sua vinculação a processos subjetivos e fisiológicos de formação da imagem. O praxinoscópio e o zootrópio são pequenas máquinas que, por seu movimento, levam a que percebamos curtas sequências de imagens estáticas como se estivessem se movendo. Seu funcionamento depende, portanto, de um processo fisiológico e cognitivo de síntese do movimento – o mesmo princípio em que se baseia o cinema. O estereoscópio, por sua vez, apresenta uma cena por dois pontos de vista distintos, cada qual oferecido a um olho do observador. Estas duas imagens, fundidas no ato da percepção, permitem a apreensão de um espaço tridimensional – o mesmo princípio em que se baseia o cinema 3D. São casos, portanto, que põem em relevo aspectos subjetivos da visão, em que ela não opera apenas como uma sensação de um mundo externo – objetivo – mas, sim, produz, ela própria, informações sobre o mundo. Crary indica como foram desenvolvidas

81 No original: “With its own rhetoric and representations, each scopic regime seeks to close out these differences: to make of its many social visualities one essential vision, or to order them in a natural hierarchy of sight. It is important, then, to slip these superpositions out of focus, to disturb the given array of visual facts (it may be the only way to see them at all)...”. Tradução minha.

no mesmo período, por Schopenhauer e Goethe, entre outros autores, elaborações acerca da percepção visual que sublinhavam estes aspectos. Indica, também, como estas qualidades da percepção seriam posteriormente elaboradas em práticas estéticas do impressionismo e como, antes dele, já incidiam nas pinturas de William Turner.

Esta compreensão da visão fundada no corpo contrasta, portanto, com a ideia de objetividade do processo de formação da imagem que era sugerida pela câmara escura e nas reflexões de Descartes sobre a visão (CRARY, 1992). A aposta metodológica de Crary é a de que seria possível inferir a constituição de sujeitos observadores no cruzamento entre estes elementos – técnicos, discursivos e estéticos – que constituiriam um campo de possibilidades para o que significaria, àquele momento, *ver*. Em um trecho bastante citado, Crary (1992, p. 6) propõe aquilo que entende por *observador* da seguinte forma:

Embora obviamente alguém que vê, um observador é, de modo mais importante, alguém que vê em um arranjo prescrito de possibilidades, alguém que está inserido em um sistema de convenções e limitações. E por ‘convenções’ eu quero sugerir bem mais do que práticas representacionais. Se pode ser dito que há um observador específico para o século XIX, ou para qualquer período, é apenas como um *efeito* de um sistema heterogêneo irreduzível de relações discursivas, sociais, tecnológicas e institucionais. Não há sujeito observador anterior a este campo continuamente em transformação⁸².

Ou seja, a partir de Crary, a visualidade deve ser compreendida como uma instância historicamente situada, na qual tomam parte elementos diversos como formações discursivas, práticas de produção e fruição das imagens, e também técnicas de produção e mediação do visível. Em certa medida – deve-se reconhecer – o percurso elaborado por Crary se aproxima de aspectos daquilo que critiquei no capítulo anterior, de uma tendência de essencialização da operação dos dispositivos, que o leva a traçar periodizações muito estritas e, de modo importante, distanciadas de descrições mais situadas de práticas observacionais empiricamente manifestas. Efetivamente, algumas das críticas dirigidas a ele ressaltam estes aspectos⁸³. Contudo, acredito que as considerações macro-históricas que propõe oferecem intuições produtivas para pensar como estas articulações podem operar em contextos específicos.

82 No original: “*Though obviously one who sees, an observer is, more importantly, one who sees within a prescribed set of possibilities, one who is embedded in a system of conventions and limitations. And by ‘conventions’ I mean to suggest far more than representational practices. If it can be said there is a specific observer for the nineteenth century it is only as an effect of an irreducibly heterogeneous system of discursive, social, technological and institutional relations. There is no observing subject prior to this continually shifting field*”. Tradução minha.

83 Cf. BATCHEN, 1993; HUHTAMO, 2013; MITCHELL, 1995.

Quando pensamos acerca das imagens digitais, é preciso levar em consideração como esta rede de mediações encontra-se configurada em cada caso. Meu argumento, neste capítulo, desdobra-se a partir da observação dos diferentes modos de inscrição da imagem em sistemas computacionais contemporâneos, em uma consideração acerca de como estas inscrições modulam os modos de ver encampados por tais sistemas. Isto tem relevância não apenas para compreender a constituição das visualidades no modo de sua expressão no mundo mas, também, para compreender os modos de ver adotados – conscientemente ou não – pelas investigações que mobilizam sistemas similares. Desse modo, a questão que lanço aqui vincula-se aos processos de translação a que as imagens são submetidas nesses contextos. Embora seja tentador sugerir que, em alguma medida, o modo de constituição das inscrições seria definidor de uma visão ‘essencial’ relacionada a tais sistemas, argumento, na sequência do que discuti no capítulo anterior, que devemos compreender tais inscrições no sentido de uma multiplicidade ontológica das imagens. Ou seja, que o objeto da investigação deve incluir, reflexivamente, as diferentes formas de ver mobilizadas em cada prática observacional, em sua multiplicidade irreduzível.

O foco principal de minha abordagem (veja [3.2 Aprendizado de máquina](#)) serão os programas de Visão Computacional desenvolvidos a partir de técnicas de aprendizado de máquina por redes neurais artificiais. Geralmente definido como um subdomínio ou como um domínio relacionado à inteligência artificial (IA), uma acepção corrente destes sistemas – que eu mesmo cheguei a sugerir em outros trabalhos – é de serem um domínio não-humano da visão, assumindo uma contraposição entre máquinas e humanos como esferas ontologicamente distintas. Contudo, também na minha dissertação de mestrado (MINTZ, 2015), discuti a partir de autoras como Lucy Suchman (2007) e Donna Haraway (2013) como esta oposição seria uma forma de essencializar os termos que, de outro modo, podem ser compreendidos como mutuamente constituídos. Na perspectiva de Suchman (2007), as definições de humano e máquina são tomadas como efeitos performativos instáveis e que são reiterados frequentemente – em particular, nos casos de interação humano-máquina. As visualidades computacionais, nesse sentido, não seriam pertinentes a um domínio exclusivo da máquina mas, sim, a modos de ver situados no entorno das máquinas computacionais mas que são afetados pelo olhar humano. Este também, por sua vez, seria afetado pelo modo de ver da máquina, em processos nos quais o que pertence a cada uma dessas categorias é constantemente *reconfigurado*.

Esta proposição contrasta, em alguma medida, com asserções elaboradas pelo artista estadunidense Trevor Paglen (2016). Em um texto dedicado ao que chamou de “imagens invisíveis”, Paglen articula ideias que perpassam alguns de seus trabalhos recentes, nos quais propõe um engajamento com tecnologias de visão computacional. Seu interesse recai, segundo elabora, sobre um domínio da cultura visual que teria se constituído à parte de um circuito humano-humano de produção e apreensão do visível. Ele sugere, neste sentido, que com a ação de máquinas capazes de ver e com a produção de imagens voltadas especificamente a seu ‘olhar’, hoje se constituiria uma visualidade cindida em duas. Efetivamente, muitos de seus trabalhos, inclusive o mais recente, realizado em parceria com Kate Crawford (CRAWFORD; PAGLEN, 2019), volta-se a imagens situadas fora de um circuito comum de apreensão do visível. No projeto recente, Paglen e Crawford voltam-se às bases de treinamento que são usadas para gerar os algoritmos de classificação de imagens. Uma massa visual que não encontra, nessa forma, o olhar humano. Em projeto anterior, Paglen (2016) volta-se às representações internas que estes programas constituem a partir do treinamento. O gesto de expor estas imagens como parte de projetos artísticos tem, para Paglen (2014), um sentido similar à operação realizada pelo cineasta alemão Harun Farocki, de trazer à luz imagens que integram aparatos de vigilância e operações securitárias. Nesse sentido, a ideia de uma cisão entre campos visuais distintos talvez se sustente, em alguma medida.

Porém, meu argumento será de que mesmo estes sistemas supostamente fechados, ao agirem sobre o mundo, não podem ser isolados de dinâmicas que lhes seriam “externas”. O ponto de vista do vigilante do panóptico de Bentham, analisado por Foucault (1997), encontra-se, evidentemente, inacessível para o prisioneiro. No entanto, esse olhar ainda assim é internalizado na subjetivação desse prisioneiro, que passa a agir sob a presunção desse olhar externo. De forma similar, quero argumentar que visualidades computacionais não têm relevância apenas a um domínio supostamente alheio ao olhar humano. Na medida em que os modos de ver, que abordo aqui, mediam aquilo que vemos, seja como usuários de plataformas *online*, seja como pesquisadores destas mesmas plataformas e suas imagens, a visualidade que eles elaboram é compartilhada conosco.

O presente capítulo desdobra-se, portanto, em três seções. Primeiro, volto-me às fricções entre as categorias do visual e do computacional. O eixo central da discussão será uma consideração crítica da Analítica Cultural e sua compreensão da imagem como *dado* cultural em uma abordagem marcadamente quantitativa. Na segunda seção, abordo as-

pectos do aprendizado de máquina por redes neurais, em especial em suas aplicações para as imagens. A seu respeito, indico como operam por múltiplos processos de translação das imagens, que tendem à condensação de visualidades difusas em imagens singulares. O problema da computacionalidade das imagens assume, então, a forma do *padrão*. Por fim, remeto a um processo contemporâneo de *infraestruturalização* dos sistemas de aprendizado de máquina, que tende a consolidar percursos relativamente estáveis de aplicação destas técnicas, formando espécies de “pontos de passagem obrigatórios” (CALLON; LAW; RIP, 1986a) para o processamento computacional das imagens em larga escala.

3.1 O visual e o computacional

Como indiquei à introdução, o tratamento computacional do visível enfrenta como desafio aquilo que Golan Levin (2006) descreveu, em dado momento, como a opacidade computacional da imagem. Este problema, que se manifesta de modo característico nas Ciências da Computação, voltadas aos computadores modernos, se aplica, também, em um sentido mais amplo, às resistências da imagem ao *cálculo* ou ao pensamento lógico-matemático. Isto porque as imagens são, mesmo antes do digital, tomadas por uma incerteza epistemológica. Afirmar, com precisão, o que uma imagem nos dá a conhecer, sob a demanda de um sistema de tomada de decisões, ou de cálculo, não é uma tarefa trivial. No capítulo anterior, a partir de Sekula (1986), discuti como, na criminalística, a integração da fotografia à tarefa de reconhecimento de pessoas demandou que o dispositivo fotográfico fosse suplementado, a fim de padronizar a tomada da imagem e, também, de permitir sua recuperação posterior. Estas são medidas que visam reduzir a margem de incerteza, mas que não são totalmente confiáveis – algo que se torna mais grave e urgente, como discutirei adiante (veja [3.3 Visualidade e infraestrutura](#)), no contexto do reconhecimento de rostos computacional.

A fotografia, embora automatize, em um processo relativamente padronizado, um modo de registro de imagens do mundo, não é uma máquina de certezas. Um primeiro desafio é, a partir da imagem, inferir características do espaço representado, reconstituir sua tridimensionalidade a partir da representação plana. Outro desafio diz respeito à interpretação, propriamente, da situação retratada – considerando, inclusive, as escolhas feitas no âmbito de seu registro. Técnicas de fotogrametria e de análise da representação

em perspectiva podem auxiliar no enfrentamento do primeiro desafio, embora não o resolvam, mas dificilmente conseguem ir muito longe no enfrentamento do segundo. Pois este demanda um esforço *interpretativo* que não se resume a manipulação de registros simbólicos, como a que realiza a computação. A interpretação será, afinal, sempre situada, em corpos e olhares específicos.

Evidentemente, muitas destas características não são exclusivas às imagens e também seriam extensíveis, por exemplo, aos sons ou à escrita. No entanto, comparada a estes (especialmente à escrita), a imagem pareceria ser muito mais resistente a esforços de estruturação e padronização. Ela frequentemente escapa a esquemas formais muito rígidos, apesar de muitas iniciativas que tentaram encontrar algo como uma “linguagem visual” em sentido estrito. No entanto, qualquer tentativa de integração da imagem a um processo computacional demandará enquadrar o visível em algum esquema, do qual a imagem tenderá a sempre escapar. No caso específico da computação moderna, o texto verbal é mais passível de tratamento computacional porque ele é formado por cadeias de caracteres permutáveis, organizados em unidades semânticas discretas e em uma estrutura sintática minimamente regular. Isto facilita que ele seja traduzido à linguagem estruturada da máquina e que sua estrutura seja analisada computacionalmente. Uma imagem, de outro modo, apresenta-se, em nível informacional, como um conjunto de dados de cor. O desafio de tratar esses dados, somado à dificuldade que descrevi acima – estendida às imagens não digitais e a um sentido ampliado de computacionalidade – descreve o tamanho do problema.

Esta incerteza como aspecto definidor das imagens seria o que, nas Ciências da Computação, qualifica a visão computacional como um “problema malposto”. Tratando de um subdomínio específico, da Recuperação de Imagens Baseada em Conteúdo (CBIR, na sigla em inglês), Smeulders *et al.* (2000) tratam a questão por um conjunto de “hiatos” (*gaps*), que separam a eficiência real e possível dos algoritmos de sua eficiência desejável ou ideal. Eles indicam, por exemplo, um “hiato sensorial” (*sensory gap*) que se refere à capacidade apenas parcial de reconstrução computacional do objeto e do espaço que a imagem representa. Um dos fatores que constituem esse hiato se relaciona à insuficiência de um ponto de vista único sobre o espaço para reconstruir o ambiente retratado. A perspectiva, como conhecidamente demonstram as gravuras de M. C. Escher, permite realizar truques: projetar espaços em si mesmos incoerentes, ou enganar nosso olhar quanto à sua disposição. A correspondência entre o espaço que é o referente da imagem e

aquele que pensamos apreender a partir delas não é garantida. Também, um objeto pode esconder-se atrás de outro e um único ponto de vista não teria como incluí-lo nas representações. Outro problema seria o “hiato semântico” (*semantic gap*), que se refere a um desacordo entre a interpretação algorítmicamente produzida, que tenderá a uma interpretação universalizante ou, em algum nível, genérica; e a interpretação elaborada por um sujeito particular. Este hiato, em especial, seria ainda mais profundo pois simplesmente não há como estabelecer um significado geral para qualquer imagem que seja. Embora seja em alguma medida possível indicar aspectos denotativos, como categorias dos objetos representados, como discutirei adiante (veja [3.3 Visualidade e infraestrutura](#)), tampouco se trata de algo trivial ou imune aos riscos de uma ambição universalizante.

Com esta breve revisão, busco situar melhor os desafios colocados ao tratamento computacional das imagens. Compreendo ser este um primeiro aspecto a introduzir nesta seção pois ele permite compreender as diferentes camadas implicadas quando digo da possibilidade de tratar a imagem como dado. Um enquadramento didático desta questão é a noção de níveis de abstração que é comumente mobilizada no âmbito da computação. Uma representação de *baixo nível* seria, nesse sentido, aquela que mais se aproxima da linguagem da máquina – estejamos falando de comandos ao processador ou da estrutura de armazenamento de dados na memória. Uma representação de *alto nível*, de outro modo, seria uma representação mais abstrata, decorrente do processamento ou tratamento das representações de *baixo nível*. Entre estes níveis, temos sempre camadas em que interpretadores computacionais traduzem as informações de um nível a outro. Quando dizemos da “opacidade computacional” das imagens, portanto, estamos tratando da dificuldade de se traduzir as representações de *baixo nível* das imagens digitais a outras representações de *alto nível*. Aserções semânticas sobre as imagens ou a reconstituição dos espaços retratados, embora façam parte do modo como nós tipicamente olhamos uma fotografia, por exemplo, estão muito distantes da representação de *baixo nível* desta mesma fotografia, em um computador. Esta distância, atravessada pelos hiatos sensorial e semântico, é o que torna a imagem, nesse sentido, opaca.

Um desafio suplementar a estes envolve o relativo alto custo computacional de algumas das técnicas que visam uma superação parcial desta distância. Como discutirei nas seções seguintes, foram desenvolvidas algumas alternativas, nos últimos anos – seja na forma de ferramentas comerciais, seja na forma de recursos de programação em código

aberto – que facilitam a aplicação destas técnicas. Importante dizer, contudo, que esta facilitação não vem sem custos e voltarei a alguns dos revezes desse processo adiante. De todo modo, falo aqui de desenvolvimentos bastante recentes, de três ou quatro anos atrás, cuja incidência nas pesquisas apenas agora começa a aparecer. Antes deles, pessoas interessadas em aplicar sistemas de análise de grandes quantidades de imagens precisariam ou desenvolver seus próprios sistemas de visão computacional, ou recorrer a abordagens mais diretamente acessíveis ao tratamento das imagens.

Os formatos de imagem tipicamente dividem-se em dois: imagens de tipo *vetor* e imagens de tipo *bitmap*. As primeiras são geralmente utilizadas para formas gráficas. As segundas são as mais comuns para fotografias e relacionam-se ao que quero me referir. Como o nome sugere, *bitmaps* funcionam como mapas de *bits*, ou mapas de informação digital. Para cada *pixel*, que seria o menor componente desta imagem, é atribuído um valor de cor. A variabilidade permitida aos valores de cor em um formato de imagem incidirá diretamente sobre a quantidade de informação necessária para representar aquela imagem. A isto se chama *profundidade de cor*. Uma profundidade de *24 bits*, por exemplo, indica a quantidade de informação que será utilizada para cada *pixel*. Para compreendermos quantas cores são possíveis, basta pensar de forma combinatória. Considerando que cada *bit* permite duas possibilidades (desligado ou ligado, 0 ou 1). *24 bits* permitem, portanto, 2^{24} possibilidades, o que significa que em uma imagem com esta profundidade de cor, a cada *pixel* pode-se atribuir um valor dentre 16,7 milhões de cores possíveis. Esse valor assume uma forma numérica, geralmente segmentado em canais de cores primárias, embora outras decomposições sejam possíveis. O ponto importante é que, para cada *pixel* temos um valor numérico que pode ser utilizado em cálculos aritméticos e, portanto, em medidas estatísticas e outros processos lógico-computacionais. Esta seria, grosso modo, a forma mais simples de tratamento computacional das imagens.

Um dos primeiros e mais conhecidos esforços de aplicação de métodos computacionais para o estudo de imagens no contexto das humanidades volta-se justamente a esta estratégia. A chamada Analítica Cultural (*Cultural Analytics*) baseia-se na disponibilidade destes valores na inscrição das imagens como *bitmaps*. Deve-se reconhecer que, em perspectiva ampliada, a Analítica Cultural não se limita ao baixo nível das representações das imagens. Algumas das pesquisas vinculadas a esta perspectiva mobilizam processos mistos que incluem categorias extraídas por programas de aprendizado de

máquina e, também, codificação manual das imagens por meio de serviços como o *Mechanical Turk*⁸⁴ (TIFENTALE; MANOVICH, 2015). Contudo, sua vertente de trabalho mais consolidada se fundamenta nos dados de baixo nível, que inclusive é o foco dos *software* de análise desenvolvidos pelo grupo, como o *ImagePlot* (MANOVICH et al., 2014). Em todo caso, tipicamente, a proposição da Analítica Cultural articula análises computacionais de larga escala com uma ênfase marcadamente *quantitativa* – mesmo quando mobilizam representações de nível mais alto. Em diversas ocasiões, Manovich sugere se tratar de uma “análise quantitativa de dados culturais” (MANOVICH, 2009, 2016). Com efeito, os esforços parecem se voltar sempre a quantidades expressivas de conteúdos e, principalmente, a uma análise enfocada em métricas extraídas dos atributos visuais destes conjuntos.

O método característico consiste em processar conjuntos de imagens com técnicas de estatística descritiva, por meio das quais definem-se parâmetros como cor dominante, variância cromática, saturação média e entropia. Estes valores são então utilizados para dispor as imagens que compõem o *corpus* analisado em um gráfico de dispersão. Neste tipo de visualização, cada item dos dados é representado por um ponto disposto em um plano cartesiano, com as coordenadas X e Y definidas segundo os valores de duas variáveis consideradas para a análise. Nas visualizações da Analítica Cultural, em procedimento proporcionado pelo *software* *ImagePlot*, (MANOVICH et al., 2014), os pontos são substituídos pelas próprias imagens, que são posicionadas segundo estas coordenadas.

Em um esforço despuoradamente ambicioso, Manovich (2012) sugere, em um estudo, que esse protocolo de pesquisa seria uma resposta à sua pergunta sobre “como comparar um milhão de imagens”. No caso, a análise volta-se a páginas de mangá, que são analisadas segundo duas variáveis: desvio padrão e entropia dos valores de brilho. Embora trate-se de um esforço valioso em seu potencial inovador, parece-me discutível em que medida sua abordagem alcançou um estágio verdadeiramente produtivo na investigação sobre o objeto abordado. Parece-me bastante problemático, em especial, a proposição feita por Manovich, nesse estudo, de definir uma categoria de *estilo* a partir da combinação destas variáveis. Trata-se, acredito, de manifestação sintomática de um dos

84 O *Mechanical Turk* é uma plataforma da empresa de tecnologia *Amazon*, por meio da qual pode-se terceirizar a realização de pequenas tarefas de forma difusa, por muitos trabalhadores que recebem uma pequena quantia por tarefa realizada. Trata-se de plataforma bastante utilizada para fins de codificação de bases de treinamento de sistemas de aprendizado de máquina – como discutirei adiante.

problemas que esta pesquisa visa enfrentar. Pois ao passar diretamente destas variáveis de baixo nível a uma categoria de alto nível, sem atentar-se às translações necessárias entre dois enquadramentos bastante distintos, Manovich realiza um salto mais amplo do que lhe permite seu referencial teórico. Seria como se os atributos analisados fossem forçados a se enquadrar em uma categoria culturalmente pertinente levando, desse modo, a uma redução, com perdas significativas, de um conceito complexo como *estilo*.

Em volume panorâmico dedicado a metodologias de pesquisa com materiais visuais, Gillian Rose (2016) situa a Analítica Cultural como uma vertente derivada de métodos de Análise de Conteúdo, cuja abordagem se caracteriza pela quantificação da ocorrência de determinadas categorias em meio ao *corpus* analisado. Para o caso da Analítica Cultural, esta análise geralmente é realizada de forma multivariada, considerando diversas métricas como características descritivas das imagens. Por exemplo, para o estudo *Selficity* (TIFENTALE; MANOVICH, 2015), que buscou analisar os *selfies* postados no *Instagram* por usuários de diferentes metrópoles globais, as métricas consideradas incluíam: idade presumida, gênero, inclinação lateral da cabeça, inclinação da câmera em relação ao rosto e expressão facial. Como na Análise de Conteúdo, efetivamente, estas métricas são abordadas de forma quantitativa, buscando interpretar as imagens – e, especialmente, seu contexto cultural – segundo medidas de estatística descritiva que foram calculadas a partir das métricas. De modo distinto da Análise de Conteúdo, contudo, Rose indica as visualizações produzidas pela análise. Estas se apresentam na forma de uma colagem das imagens em grade ou por meio de gráficos de dispersão nos quais as próprias imagens servem como marcadores dos dados – como indiquei, acima, sobre o *software* ImagePlot. Estas visualizações teriam a vantagem, sugere a autora (ROSE, 2016, p. 103), de evitar a “fragmentação” das imagens após sua análise por procedimentos típicos da Análise de Conteúdo, em que as imagens uma vez codificadas, seriam totalmente abandonadas, voltando-se apenas às métricas derivadas. As visualizações da Analítica Cultural, de modo distinto, retêm a imagem original que seguiria, portanto, disponível ao olhar do analista, para outros tipos de observações, sem se dispersar completamente em dados quantificáveis. Trata-se de um esforço na elaboração dos produtos da pesquisa que seria coerente com um aspecto rotineiramente salientado por Manovich e seu grupo como um dos aspectos distintos da Analítica Cultural: o centramento da análise em aspectos visuais, em oposição ao privilégio de aspectos verbais que seria típico nas humanidades digitais.

Entretanto, embora também considere acertada a manutenção da imagem como elemento observável nestas visualizações, eu diria que esta disponibilidade das imagens individuais não rende, nas análises realizadas, um olhar específico a elas, após a visada de conjunto. Ainda que sobrevivam às análises e à visualização quantitativa, o olhar dedicado às imagens não parece alternar, efetivamente, entre os diferentes registros incrustados nas visualizações – isto é, entre a estatística descritiva geral e as imagens individuais. Tipicamente, as conclusões das análises enfocam tendências gerais e se furtam a considerar aspectos da constituição das imagens que não são compreendidos por tais métricas. Nos termos do que venho discutindo até agora, na tese, trata-se de um gesto que parece buscar uma nova essencialização ao dar prioridade a uma manifestação específica da multiplicidade ontológica das imagens, em detrimento de outras. A inscrição digital das imagens enquanto valores de cor, ou a quantificação de atributos de maior nível de abstração, são compreendidos, desse modo, como os próprios objetos sob análise e não como apenas uma de suas manifestações.

Sean Cubitt (2013), embora não cite nominalmente Manovich ou a Analítica Cultural, parece endereçar a ele algumas de suas críticas a uma ênfase quantitativa e estatística nas humanidades digitais. Como contraponto a estas abordagens, Cubitt elabora o elogio a uma *metodologia anedotal*. Embora o uso de anedotas sejam objeto de forte crítica metodológica dada a possibilidade de que venha a distorcer possíveis conclusões por meio de escolhas localizadas de objetos que favoreçam determinada preconcepção, Cubitt retoma, em outro enquadramento, a importância da consideração crítica de objetos específicos nos estudos em artes e humanidades. A força da anedota, elabora o autor, reside no “princípio de que a instância única [*unique*] pode ensinar pesquisadores tanto quanto amostragens estatísticas ou aquelas abstrações que nos chegam seja como axiomas [...] seja como hipóteses, mapas e diagramas” (CUBITT, 2013, p. 5). Ele continua: “o núcleo da anedota não é sua tipicidade, mas sua especificidade”⁸⁵ (2013, p. 5). Trata-se de uma defesa que tem especial pertinência para o estudo da arte, que é um dos focos da argumentação do autor. Com frequência, nesse domínio, um objeto tem relevância não por ser típico, igual aos demais, mas por ser específico, único, “fora da curva”. Esta dis-

85 No original: “... principle is that the unique instance can teach researcher as much as statistical samples or those abstractions that arrive either as axioms [...] or hypotheses, maps or diagrams. [...] The core of the anecdote is not its typicality but its specificity”. Tradução minha.

tinção teria, por exemplo, um potencial de não apenas de expressar qualidades inexistentes no conjunto mas, também, de permitir, em negativo, compreender melhor os aspectos dos outros objetos, dos quais se distingue.

De modo ainda mais importante, a anedota, para Cubitt, não significa o abandono de outros métodos ou mesmo dos esforços de elaborar asserções gerais. Ela seria uma abordagem complementar por meio dos quais testar, no caso individual, hipóteses projetadas sobre ele por um olhar generalista. Escreve: “A altíssima resolução do método anedotal provê de profundidade e cor as descobertas generalistas de métodos que lidam com múltiplas instâncias e tendências de larga escala”⁸⁶ (CUBITT, 2013, p. 6). Seria, nesse sentido, um apelo para recuperar olhares mais detidos e próximos aos objetos como parte das análises, compreendendo a *experiência* como uma instância tão pertinente quanto as visadas generalistas oferecidas pela estatística descritiva e outras técnicas relacionadas. No encaminhamento de seu argumento, Cubitt retoma o problema da imagem como um caso central a esta discussão, em particular indicando como no livro *A câmara clara*, de Roland Barthes (1984), parte substancial do argumento sustenta-se em uma única foto e na relação afetiva que a ele lhe dedica o autor.

O percurso argumentativo de Cubitt vai mais longe, mas podemos nos bastar com isto: reconhecer o valor da anedota não equivale a desacreditar asserções gerais, mas, sim, a reconhecer a existência de singularidades que resistem a tais abstrações. O risco da evidência anedotal, em um plano metodológico estrito, remete a práticas de “*cherry picking*” – isto é, de seleção interessada dos dados para sustentar uma visão falaciosa da realidade. Não é, evidentemente, a estas práticas que se alude aqui. De outro modo, Cubitt (2013, p. 8) argumenta que o valor de uma anedota, para uma investigação, deve ser avaliado a partir de dois fatores principais: a profundidade daquilo que consegue revelar e a amplitude das ‘causas’ – Cubitt (2013, p. 8) diz de “*Goods*”, isto é, das causas moralmente boas – que ela coloca em jogo. “Furar as estatísticas” seria, em um sentido comum, uma das potências das anedotas. Um movimento que, evidentemente, não conseguiria reverter situações dadas mas que permitiria vislumbrar matizes em meio às certezas e horizontes de uma transformação possível.

Que a imagem surja como um caso privilegiado para a consideração das anedotas por Cubitt não é uma discussão inaudita. Historiadores da ciência como Peter Galison e Lor-

86 No original: “*The extremely high resolution of the anecdotal method provides depth and colour to the generalist findings of methods that deal with multiple instances and large-scale tendencies*”. Tradução minha.

raine Daston notoriamente enveredaram-se pela relação fricativa desenvolvida, ao longo dos anos, entre as imagens e as práticas científicas. Em obra conjunta (DASTON; GALISON, 2010), os autores percorrem atlas científicos buscando compreender como as imagens da ciência, em particular com o advento da fotografia e, depois, de visualizações computacionais, foram articuladas com a noção de *objetividade*. Galison voltou-se a, também, considerações sobre os atritos entre *imagem e lógica* no âmbito dos estudos da física. Em texto incluído no catálogo de exposição centrada sobre a noção de *iconoclash*, proposta por Latour, Galison (2002) sugere, na esteira da linha curatorial da exposição, que a relação das ciências com as imagens teria uma configuração similar. A noção de *iconoclash* é sugerida por Latour (2008; 2002) para lidar com conflitos centrados nas imagens e geralmente abordados de forma dicotômica, como entre uma *iconofilia* e uma *iconoclastia*. Um *iconoclash*, de outro modo, salientaria que mesmo o gesto iconoclasta pode ser um modo de produzir uma outra imagem ou que, tipicamente, seria sucedido pela geração de outros ícones em substituição aos destruídos. Nesta linha, Galison sugere uma relação similar, nas ciências, acerca do conflito entre dados e imagens.

Galison (2002) recupera desenvolvimentos em diferentes campos científicos – matemática, física, medicina, geologia – indicando como em cada um destes, o valor epistemológico da imagem é constantemente recolocado em debate. Por um lado, as imagens são vistas positivamente por seu potencial didático e intuitivo, oferecendo uma especificidade e uma materialidade sobre a qual repousar o pensamento. Por outro, as imagens são limitadas justamente por sua materialidade e especificidade, que impedem a abstração que, em certa acepção da ciência, configuraria a única “estrada nobre para o conhecimento”, pavimentada pela razão ‘pura’, lógica e não-intuitiva (GALISON, 2002, p. 300). No entanto, o autor aponta como, em diversos momentos, a prática científica vê-se às voltas com as imagens, seja como parte de um instrumento de medição e investigação dos fenômenos, seja como representação das conclusões elaboradas a partir da análise. Hoje, com a já bastante consolidada digitalização das práticas de análise computacional de dados, este debate pareceria ter, em princípio, esfriado. Por meio de técnicas de *visualização*, temos, afinal, representações visuais derivadas dos dados por processos lógicos definidos, de modo que podemos ter *dados*, *lógica* e, também, *imagens* em um

mesmo esforço investigativo e de modo, aparentemente, não contraditório. Galison (2002) sugere, neste sentido, um movimento pendular: “imagens dispersam-se em dados; dados agregam-se em imagens”.

O autor sugere, em conclusão, que o ponto principal para se aprender destes debates talvez seja questionar justamente a purificação destes dois domínios – da imagem, de um lado, dos dados e da lógica, de outro – como profundamente cindidos. Escreve:

Nós nunca podemos falar (ou pintar ou calcular) sem abstração metafórica. Ao mesmo tempo, o abstrato nunca o é, completamente; mesmo nos confins mais frios da física matemática sempre (tomando de Lutero) encontraremos a imagem de nosso rosto em águas calmas. Em vez de abstrato contra o concreto, realizações históricas mutáveis da abstração-concreta e concretude-abstrata⁸⁷ (GALISON, 2002, p. 323).

Este, me parece, é o cerne da questão. Não se trata de escolher, necessariamente, uma posição em meio a este debate, mas de questionar, justamente, a oposição. Ao menos, diria, de um ponto de vista epistemológico ou ontológico. Quero dizer que, na linha de Galison, seria fútil debatermo-nos se a realidade é melhor expressa em imagens (concretude) ou dados (abstração). A realidade conjuga ambos. Podemos, claro, discutir em casos específicos a adequação, ou não, de tender a um lado ou outro desse espectro, mas devemos sempre compreender que se trata de um espectro e que os termos não são, em si, contraditórios ou oponentes; eles podem ser conciliados. Nesse sentido, o problema do tratamento da imagem como dado, em particular pela Analítica Cultural, seria perder de vista a concretude da imagem e de sua experiência como um dos elementos pertinentes à análise. A ênfase em uma “análise *quantitativa* de *dados* culturais” postularia, a seu modo, uma proeminência da abstração dos dados em detrimento da concretude da experiência das imagens.

A promessa de conciliação entre visadas quantitativas e qualitativas, embora constitua parte do discurso de Manovich, parece obstruída por um deslumbramento com o *big data* que chega a parecer merecedor da denominação crítica do “dataísmo” proposta pela pesquisadora José Van Dijck (2014). A autora se vale desta denominação para o que identifica como uma “ideologia” que marca premissas ontológicas e epistemológicas problemáticas das aplicações contemporâneas do *big data*:

87 No original: “We cannot ever speak (or paint or calculate) without metaphoric abstraction. At the same time the abstract is never completely so; even in the coldest reaches of mathematical physics we will always (borrowing from Luther), find the image of our face in still water. Not abstract against the concrete, but rather shifting historical realizations of concrete-abstraction or abstract-concreteness”. Tradução minha.

a ideologia do *dataísmo* mostra características de uma *crença* amplamente difundida na quantificação objetiva e no potencial rastreamento de toda forma de comportamento humano e sociabilidade através de tecnologias midiáticas *online*. Além disso, o *dataísmo* implica *confiança* nos agentes (institucionais) que coletam, interpretam e compartilham (meta)dados colhidos das mídias sociais, plataformas de internet e outras tecnologias de comunicação⁸⁸ (VAN DIJCK, 2014, p. 198).

O discurso que acompanha as inovações técnicas e metodológicas propostas por Manovich se aproximam muito do alvo das críticas de Van Dijck uma vez que suas análises raramente vêm acompanhadas de um esforço reflexivo acerca dos vieses epistemológicos introduzidos pelas técnicas analíticas empregadas. Além disso, suas análises tendem a propor saltos demasiadamente largos entre os “dados culturais” e suas consequências teórico-conceituais.

Gillian Rose (2016, p. 104) critica a pretensão de objetividade presente na proposição da Analítica Cultural que, segundo a autora: “remove toda necessidade de uma postura reflexiva por parte dos pesquisadores, para além de reportar seu método em detalhe; qualquer outra forma de reflexividade não é parte destes métodos porque eles se supõem objetivos”⁸⁹. Outras críticas feitas por Rose direcionam-se: a seu marcado positivismo, dada sua insensibilidade ao que está ausente da imagem e às diferentes qualidades que podem compor uma ocorrência quantificável; e, também, à insensibilidade do método ao contexto e a outros fatores externos à superfície visível das imagens.

Podemos aproximar aspectos da crítica de Rose à ácida crítica elaborada por Alexander Galloway (2014) acerca do crescente uso de métodos quantitativos nas humanidades, o qual atribui ao que chama de um “ecumenismo metodológico” (GALLOWAY, 2014, p. 108) de cunho liberal, segundo o qual a escolha do método aplicado seria apenas uma questão de preferência do pesquisador, definida por um senso do que é ou não apropriado, apenas. Isto é, sem uma consideração reflexiva sobre a historicidade dos métodos empregados ou a que finalidades cumprem, agindo apenas segundo aquilo que parece funcionar para o propósito pretendido. De modo frequente, sugere, estes estudos realizam apenas uma reprodução daquilo que fazem, hoje, grandes corporações empresariais – “os titãs corporativos de hoje consistem pouco mais do que modos altamente desen-

88 No original: “the ideology of dataism shows characteristics of a widespread belief in the objective quantification and potential tracking of all kinds of human behavior and sociality through online media technologies. Besides, dataism also involves trust in the (institutional) agents that collect, interpret, and share (meta)data culled from social media, internet platforms, and other communication technologies”. Tradução minha.

89 No original: “...removes any need on the part of the researchers to be reflexive in any way other than by reporting their method in detail; any other form of reflexivity is not part of these methods because they assume they are objective”. Tradução minha.

volvidos de pesquisa quantitativa”⁹⁰ (GALLOWAY, 2014, p. 109). Voltarei adiante a aspectos de sua crítica, inclusive discutindo pontos divergentes. Em todo caso, o cerne da questão é que não basta a justificativa de aparente adequação dos métodos – a resposta ao *como* – sem que efetivamente compreendamos o *que* estamos fazendo.

Encontrar alguma possibilidade de equilíbrio entre o computacional e o visual no tratamento conferido às imagens constitui, efetivamente, um arco mais abrangente do desafio colocado para esta pesquisa, que não chegarei a enfrentar nesse momento. Espera-se que os esforços empreendidos ao final do percurso, a partir do estudo de caso, apontem para caminhos possíveis, nesse sentido. O aspecto importante por ressaltar na discussão desta seção e da Analítica Cultural, em particular, é a tensão inevitavelmente colocada quando as imagens são tratadas em grandes conjuntos e quando a experiência visual é transladada por sua datificação. Sugiro tomarmos estes casos como exemplares da complexidade da tarefa e de como a sedução de uma perspectiva *dataísta* pode levar a possíveis equívocos de categoria no tratamento analítico das imagens. Não é, portanto, que tratar as imagens como dados seria sempre, necessariamente, um problema. Porém, haveria que se considerar cuidadosamente o que esta translação implica para os objetos e para a questão de pesquisa. Ter clareza, enfim, quanto ao que se perde e o que se ganha em tal movimento e quanto ao que significa tomar as imagens analisadas, especificamente, sob tal visualidade.

3.2 Aprendizado de máquina

Dando sequência à discussão, volto-me agora às técnicas de aprendizado de máquina, que compreendem processos mais complexos de tratamento computacional das imagens. Estes buscam aceder a representações com nível de abstração mais alto, que vão além dos dados cromáticos que compõem o nível mais bruto da representação computacional para encontrar categorias ou padrões nas imagens. Estas técnicas possuem especial proeminência no modo como se busca, hoje, superar a “opacidade computacional” das imagens. O campo da visão computacional constitui uma das vitrines deste paradigma contemporâneo da computação, por ser um dos domínios em que teria conseguido aportar maiores avanços. Técnicas compreendidas pelo amplo espectro do aprendizado de máquina, em particular as chamadas Redes Neurais Profundas (DNN, na

90 No original: “today’s corporate titans consist of little more than highly evolved modes of quantitative research”. Tradução minha.

sigla em inglês), obtêm altos índices de eficácia, segundo parâmetros adotados na indústria e parte da academia, em tarefas como a classificação de imagens ou o reconhecimento óptico de caracteres, para ficar em dois exemplos. Trata-se, neste sentido, de uma perspectiva tecnológica sedutora para esforços metodológicos como o empreendido nesta tese.

Como discutirei ao longo desta seção e da seguinte, contudo, os resultados apresentados pelas técnicas de aprendizado de máquina devem ser observados criticamente. Um dos principais reveses já identificados diz respeito ao modo com que sistemas de aprendizado de máquina reificam discriminações por meio de seu tratamento estatístico do mundo. Especificamente, estudos contemporâneos têm identificado e discutido como estes sistemas tendem a reproduzir vieses de raça, gênero e cultura manifestos nos dados utilizados como base de treinamento e, também, no contexto social de seu desenvolvimento (cf. BROUSSARD, 2018; BUOLAMWINI; GEBRU, 2018; SILVA, 2019). Outro problema tipicamente apontado com relação ao aprendizado de máquina diz respeito ao modo de estruturação da arquitetura destes sistemas, que leva a um obscurecimento de suas representações e operações internas, dificultando esforços de auditoria e desafiando que resultados eventualmente acertados possam ser sustentados em um plano teórico. Apesar destas questões, estes sistemas têm assumido papéis importantes como mediadores algorítmicos difusos de múltiplas instâncias da vida contemporânea. Em particular, para o caso desta investigação, técnicas de aprendizado de máquina têm sido aplicadas ao tratamento computacional conferido às imagens nas plataformas *online* e, segundo quero argumentar, informam aspectos das visualidades contemporâneas. Em vista desta difusão argumentarei, portanto, acerca do valor de esforços de reapropriação destas tecnologias como metodologia de análise de imagens em aplicações críticas, conscientes do papel mediador destas técnicas e de seus riscos.

Em continuidade mais próxima da discussão que realizei na seção anterior, outro ponto importante por destacar é que o refinamento técnico do aprendizado de máquina, em comparação ao tratamento da imagem por dados de cor, não aporta nenhum avanço, por si só, sobre a questão do *dataísmo*, que discuti acerca de vertentes dos estudos computacionais da imagem. Efetivamente, não se trata de um problema para o qual a solução seria apenas técnica. Como indiquei, alguns dos estudos da Analítica Cultural (TIFENTALE; MANOVICH, 2015), inclusive, vão além dos dados cromáticos e aplicam sistemas baseados em aprendizado de máquina. Mas nem por isto se distanciam de uma visada

estritamente quantitativa de análise de grandes tendências. Os dados gerados por aprendizado de máquina podem ser, nesse sentido, facilmente integrados a proposições investigativas similarmente ‘dataístas’. Isto, sem entrar a fundo na própria fundamentação quantitativa e estatística inerente ao aprendizado de máquina. Logo, na linha do que discuti anteriormente, a consideração das inscrições geradas por técnicas de aprendizado de máquina no âmbito desta pesquisa deverá abraçar a multiplicidade ontológica que emerge neste processo, em vez de substituir uma manifestação do objeto investigado por outra.

Efetivamente, o caráter sedutor destas tecnologias e sua promessa de eficácia tendem a complexificar a reflexão metodológica proposta. Complementando as críticas de Van Dijck, Galloway e Cubitt, podemos ainda indicar o que Meredith Broussard (2018) chama de *tecnochauvinismo*. Segundo sugere a autora, trata-se de um regime de crença de que “mais tecnologia” seria sempre parte da solução. De modo similar ao *dataísmo* descrito por Van Dijck (2014), o *tecnochauvinismo* teria fortes contornos ideológicos, sugerindo uma postura socialmente irresponsável no desenvolvimento e na aplicação tecnológica fundados na inovação como valor absoluto. A aceitação de ‘danos colaterais’ como parte imprescindível do desenvolvimento tecnológico é, efetivamente, aspecto corrente de discursos da indústria tecnológica. Como elabora a autora, esta aceitação seria também sintomática da falta de diversidade desse setor, composto majoritariamente por homens brancos – um segmento, evidentemente, menos vulnerável aos “danos colaterais” desse desenvolvimento. Esta situação tende a se complexificar com técnicas de aprendizado de máquina que não oferecem um caminho fácil para explicação de seu modo de funcionamento e tendem a ser justificadas por resultados considerados aceitáveis, por critérios – eles próprios – enviesados. Com isto, sua aplicação em pesquisas e mesmo na indústria frequentemente se deve, como elabora Adrian Mackenzie, em referência a Matthew Jockers⁹¹ (2013, citado por MACKENZIE, 2017, p. 11–12), apenas à “beleza dos resultados”. O aprendizado de máquina demanda, portanto, uma atitude atenta não apenas a certo pendor positivista dos dados computacionais mas, também, à crescente demanda por métodos computacionais como se fossem sempre soluções para os problemas. Assumo nesta pesquisa que os métodos computacionais seriam parte do problema investigado.

91 JOCKERS, Matthew L. *Macroanalysis: digital methods and literary history*. Urbana: University of Illinois Press, 2013.

Em linhas gerais, o aprendizado de máquina pode ser definido como um paradigma de desenvolvimento computacional, composto por diversas técnicas (cf. ALPAYDIN, 2016; MACKENZIE, 2017) que, em comum, elaboram programas que se adaptam, dinamicamente, de modo a otimizar a realização de determinada tarefa. Diferentemente do desenvolvimento de *software* tradicional, a pessoa que desenvolve um sistema de aprendizado de máquina não se ocupa em projetar algoritmos em seu sentido típico, isto é, como uma sequência finita de procedimentos para realizar uma tarefa. De outro modo, seu papel é constituir uma arquitetura computacional adaptativa, coletar e tratar um conjunto de dados de treinamento, e ajustar o modelo gerado a partir do processamento destes dados. Esta diferenciação é elaborada por Cardon, Cointet e Mazières (2018) a partir das categorias de *máquinas dedutivas*, que correspondem ao desenvolvimento de *software* tradicional; e *máquinas indutivas*, que seriam características do aprendizado de máquina (Figura 5). No paradigma dedutivo, a máquina recebe um *programa*, na forma de um conjunto de regras finitas (um *algoritmo*); e recebe, também, um conjunto de *dados de entrada* (*input*). Sua operação produz, como resultado, um conjunto de *dados de saída* (*output*). Para que a máquina cumpra adequadamente determinada tarefa, portanto, é preciso que a pessoa desenvolvedora do *programa* o elabore segundo uma compreensão teórica prévia sobre como proceder. Esta compreensão teórica é que lhe permite *determinar* a sequência de procedimentos, visando tanto a eficácia quanto a eficiência da operação.

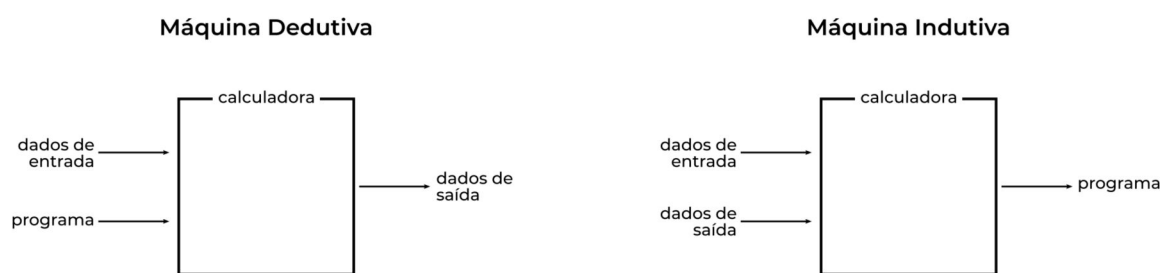


Figura 5: Esquema comparativo entre os paradigmas dedutivos e indutivos. O esquema da esquerda representa a abordagem tradicional da programação computacional. O esquema da esquerda representa o paradigma indutivo do aprendizado de máquina.

Fonte: Refação e tradução própria de diagrama de Cardon, Cointet e Mazières (2018, p. 3).

Nas máquinas indutivas, de outro modo, o *programa*, no sentido utilizado na descrição das máquinas dedutivas, é um *resultado* da operação da máquina. Dois conjuntos de dados são introduzidos no sistema. Estes seriam, para o paradigma dedutivo, os *dados de*

entrada e os *dados de saída*. Pareceria, em princípio, paradoxal que *dados de saída* sejam introduzidos no sistema, mas efetivamente, estes correspondem a exemplos de resultados esperados do *programa* que cabe à máquina elaborar. O conjunto destes “dados de entrada” e “dados de saída” alimentado às máquinas indutivas é o que se considera, para o aprendizado de máquina, *dados de treinamento*, com base nos quais a máquina deverá *aprender*⁹². Para um programa de classificação de imagens, por exemplo, estes dados poderiam ser um conjunto de imagens (*dados de entrada*) e classes que devem ser atribuídas a estas imagens (*dados de saída*). Isto é, um conjunto de imagens e indicações quanto a quais delas contêm cachorros e quais não contêm, por exemplo. A partir das possíveis correlações que poderiam ser identificadas nos dados de treinamento, a máquina elabora, então, um *programa* que permitiria realizar a mesma operação para outros conjuntos de dados. Fundamentalmente, portanto, para as máquinas dedutivas, não há a elaboração de regras explícitas a partir de um conhecimento prévio sobre a tarefa a se realizar. Estas regras são indutivamente derivadas a partir de exemplos, tomados como dados de *treinamento* para estas máquinas.

Em última medida, as máquinas indutivas são construídas para produzir *previsões*. Efetivamente, a última camada de uma rede neural empregada nesta pesquisa (veja [4.2.4 VGG19](#)) e que oferece classificações para uma imagem, é denominada **predictions** (veja [Anexo A](#)). Os resultados produzidos por um programa classificador seriam, desse modo, previsões das classes que seriam atribuídas a determinada imagem pelo agente que produziu os dados de treinamento. Adrian Mackenzie (2015), portanto, situa o aprendizado de máquinas no contexto de uma “generalização da previsão”, observando suas aplicações para prever a possibilidade de retorno de um investimento, prever o desenvolvimento de um tumor, prever os conteúdos que reterão a atenção de usuários etc. Mais do que a produção de interpretações “novas” sobre o mundo as máquinas indutivas visam prever como um dado seria enquadrado em um modelo implícito nos dados de treinamento.

Muito da literatura que se volta recentemente às técnicas de aprendizado de máquina tem ressaltado a importância de evitar perspectivas mitificadoras de sua operação (BROUSSARD, 2018; MACKENZIE, 2015, 2017; RIEDER, 2018). A ideia de um *aprendizado*

92 Há uma categoria específica de programas de aprendizado de máquina em que o modelo é elaborado apenas a partir dos ‘dados de entrada’, sem oferecer resultados específicos esperados. Chama-se a este tipo de desenvolvimento *aprendizado não-supervisionado*. Nesse caso, a máquina apenas identifica características recorrentes nos dados, sem partir de categorias predefinidas.

de máquina, bem como a vinculação do campo à *inteligência artificial* – ainda que conflituosa (cf. CARDON; COINTET; MAZIÈRES, 2018) – sugerem, afinal, um ponto de virada para narrativas futuristas ficcionais. Especialmente em vista de resultados impressionantes obtidos por tais sistemas, em diversas áreas de aplicação, não é raro encontrar no debate público discursos que descrevem – em posturas ora eufóricas, ora alarmistas – cenários em que as máquinas se tornariam sencientes e tão ou mais inteligentes que seres humanos. Broussard (2018) ressalta, a este respeito, uma diferença didática importante entre uma *inteligência artificial geral* (*general artificial intelligence*) e uma *inteligência artificial restrita* (*narrow artificial intelligence*). A primeira seria do tipo que nutre a imaginação da *singularidade* ou da máquina como ser em “equivalência” com o humano. Contudo, como elabora a autora, nunca se chegou nem perto de se realizar esta ideia e, além disso, trata-se de um paradigma de desenvolvimento abandonado, ao menos por ora, desde os anos 1990. A IA restrita, por outro lado, seria aquela a que se voltam as técnicas de aprendizado de máquina. São programas que visam atender a tarefas específicas. Embora as técnicas, em si, sejam aplicáveis em muitos campos e tarefas distintos, cada aplicação demanda um tipo de arquitetura e uma base de treinamento. Compreender esta diferença e a realidade dos desenvolvimentos em jogo é fundamental para que consigamos nos ater aos problemas e riscos reais destas tecnologias, bem como aos seus reais potenciais.

Outra definição importante, a fim de evitar linhas mitificadoras, é definir exatamente o que se compreende por *aprendizado* no âmbito do *aprendizado de máquina*. Não se trata de conceito equivalente ao que compreenderíamos como aprendizado em termos pedagógicos, por exemplo. Tecnicamente (cf. MACKENZIE, 2017), considera-se que um programa de computador *aprende* quando seu desempenho na realização de determinada tarefa torna-se mais eficiente à medida em que aumenta sua *experiência*. Porém, nem o conceito de *experiência* poderia ser tomado de forma ingênua. Trata-se, neste contexto específico, de uma medida da quantidade de dados de treinamento que teria sido processada por aquele programa. Posto de outra forma, portanto, *aprendizado*, para o *aprendizado de máquina* define-se por uma taxa de otimização algorítmica (MACKENZIE, 2017, p. 92). Um programa de aprendizado de máquina não realiza outra coisa que derivar uma *função* com base em sua adequação a um conjunto de dados. Não se trata do aprendizado como a produção de algo novo, mas apenas à reprodução de padrões implícitos na estrutura dos dados.

Nesta pesquisa, o aprendizado de máquina é abordado com um interesse específico em sua incidência no campo da Visão Computacional – isto é: por suas aplicações às imagens. Neste domínio, as técnicas de aprendizado de máquina têm tido, ao menos desde 2012⁹³, um impacto bastante significativo. Aplicações baseadas em aprendizado de máquina são hoje utilizadas para o reconhecimento e classificação de imagens (objetos, locais, temas etc.); reconhecimento óptico de caracteres; reconhecimento facial; reconhecimento de expressões faciais; recuperação da informação baseada em conteúdo; entre outras tarefas. De modo importante, como indiquei à introdução, o enfrentamento do problema da computacionalidade das imagens por meio do aprendizado de máquina implica em uma dupla articulação com a cultura visual contemporânea. Por um lado, a demanda por tais sistemas deve-se a um aumento significativo da quantidade de imagens produzidas e postas em circulação contemporaneamente – ao que Daniel Rubinstein e Katrina Sluis (2008) se referiram por uma “amadorização em massa” da fotografia digital. Uma produção que, por sua vez, cada vez mais toma forma em plataformas *online*, para as quais procedimentos de datificação e mediação algorítmica são fundamentais. Por outro, pela forte dependência de dados de treinamento por parte dos sistemas de aprendizado de máquina, é esta mesma produção massiva de imagens que alimenta a constituição destes programas. A base Imagenet (2009) é emblemática, entre outras bases abertas de dados de treinamento, em sua coleção de milhões de imagens colhidas da internet. Outras bases públicas poderiam ser igualmente consideradas (cf. HARVEY, 2019), mas talvez sejam ainda mais relevantes as bases privadas, mantidas por grandes corporações de tecnologia para o desenvolvimento de seus próprios sistemas de aprendizado de máquina (Google, Amazon, IBM, Microsoft, Apple etc.).

A esta dupla articulação, soma-se a disponibilidade recente destas técnicas em modelos abertos e serviços comerciais, impactando em uma maior pervasividade da interpretação de imagens por aprendizado de máquina, inclusive em pesquisas acadêmicas, como recurso analítico para o estudo de imagens em contextos *online*. Experimentos situados no abrangente domínio das humanidades digitais incluem, por exemplo, o uso de técnicas de aprendizado de máquina como recurso de navegação fortuita por arquivos fotográficos históricos (cf. YALE DIGITAL HUMANITIES LAB, 2017) ou em acervos museológicos

93 Neste ano, relatam Cardon, Cointet e Mazières (2018), equipe liderada por Geoffrey Hinton, pesquisador renomado na área de redes neurais ganhou o concurso anual *Imagenet Large Scale Visual Recognition Challenge* (ILSVRC) utilizando a técnica de DNN. O fator mais surpreendente era que nem Hinton nem o pesquisador mais diretamente envolvido no projeto, Alex Krizhevsky, possuíam um conhecimento sólido em Visão Computacional. Sua especialidade era redes neurais.

(cf. DIAGNE; BARRADEAU; DOURY, 2018). Em ainda outro eixo de aplicação, técnicas de aprendizado de máquina têm sido utilizadas por artistas como objeto de experimentação e investigação estética⁹⁴. Por este crescente entranhamento do aprendizado de máquina em diferentes instâncias de mediação e produção contemporânea de imagens, portanto, situo estas técnicas como objetos pertinentes ao tema das visualidades contemporâneas.

O objetivo desta seção é situar as técnicas de aprendizado de máquina considerando como elas poderiam compor abordagens reflexivas sobre a imagem. O desafio principal é articular estas técnicas a uma visada não-reducionista do visual que, como discuti na seção anterior, deveria evitar subsumi-lo ao seu tratamento computacional datificado. Distribuo esta linha argumentativa geral em três subseções. Primeiro, volto-me a uma discussão focada especificamente, a uma consideração de como o aprendizado de máquina por redes neurais aprendem a *ver*. Nessa primeira seção, interessa-me constituir uma descrição geral, *diagramática* (MACKENZIE, 2017), de seu modo de operação, em especial considerando suas implicações epistêmicas. Na segunda seção, volto-me a uma consideração sobre as inscrições sociotécnicas em que se baseiam estas técnicas em seu modo de operação. Tipicamente (cf. CARDON; COINTET; MAZIÈRES, 2018; MACKENZIE, 2017), são *vetores* de coordenadas, que descrevem os dados segundo sua disposição em um espaço vetorial multidimensional. Em linha com as discussões do capítulo precedente, interessa compreender o que este modo de inscrição proporciona à operação do aprendizado de máquina e, também, as dependências destas técnicas ao acúmulo e manipulação destes registros. Por fim, discuto como o aprendizado de máquina pode ser compreendido sob a dinâmica de reconfigurações humano-máquina (SUCHMAN, 2007), considerando a inflexão particular a que submete o domínio ampliado da IA. Na seção que conclui este capítulo (3.3 Visualidade e infraestrutura) desdubro ainda outros aspectos desta discussão, considerando o empacotamento de modelos de aprendizado de máquina em serviços comerciais e formatos de aplicação generalista.

3.2.1 Aprender a ver

Um desafio inerente à questão colocada para esta subseção diz respeito à ininteligibilidade das redes neurais diante de esforços de descrição pormenorizada de seu modo de funcionamento. As técnicas de aprendizado de máquina e as redes neurais, em particu-

94 Uma listagem não-exaustiva de artistas que lidam contemporaneamente com o aprendizado de máquina inclui: Trevor Paglen, Kate Crawford, Adam Harvey, Forensic Architecture, Caroline Sindors, Mario Klingemann, Refik Anadol, Memo Akten, Helena Sarin, Mimi Onuoha, Scott Eaton e Tom White.

lar, colocam-se como objetos críticos que escapam a vertentes de estudo como o *software studies* (FULLER, 2008). Como elabora Mackenzie (2017, p. 22) o modo de operação de programas de aprendizado de máquina não pode ser lido na forma de código. Isto não significa que o código, em si, seja opaco, mas apenas que seu modo de operação não assume a forma procedural típica de um programa de computador e articula, de outro modo, relações intrincadas com efeitos emergentes. Mackenzie (2017, p. 23) sugere, a esse respeito, que a transição poderia ser expressa pela passagem de *diagramas lógico-simbólicos* a *diagramas estatístico-algorítmicos*⁹⁵ – uma formulação análoga àquela entre máquinas dedutivas e máquinas indutivas de Cardon, Cointet e Mazières (2018). Enquanto um algoritmo tradicional pode ser descrito por sequências causais simples, programas de aprendizado de máquina articulam-se aos dados e assumem uma arquitetura distribuída. Nesta, a lógica *booleana* do tipo *se... então...* se dilui em *pesos* (*weights*) que definem os processos de transformação dos dados em uma rede distribuída de pequenas agências algorítmicas (“neurônios”) de operação probabilística. Por esse motivo, uma rede neural não pode ser compreendida ou descrita da mesma forma que um fluxograma descreve o modo de operação de um algoritmo tradicional. Seus efeitos são emergentes.

Segundo relatam Cardon, Cointet e Mazières (2018), esta característica dos sistemas de aprendizado de máquina, em especial do subdomínio das redes neurais, constitui um dos pontos centrais da controvérsia estabelecida entre dois paradigmas conflitantes do desenvolvimento da IA ao longo dos últimos 80 anos. O aprendizado de máquina constituiria, segundo descrevem, a corrente *conexionista*, cujo lastreamento teórico remonta a aspectos da cibernética de primeira ordem, de Norbert Wiener ([1948] 2007). Ainda com relação à corrente *conexionista*, outro antecedente fundamental seria a proposição do *Perceptron*, por Frank Rosenblatt (1958), uma formulação – ainda rudimentar, para os parâmetros contemporâneos – de uma rede neural artificial como modelo de desenvolvimento da IA. A outra vertente seria a corrente *simbólica*, que ganha tração a partir da década de 1960 colocando o paradigma *conexionista* em segundo plano até meados da década de 1990. Um caso exemplar de tecnologia sob o paradigma simbólico são os chamados *sistemas especialistas*, em que o conhecimento e o processo decisório de áreas específicas do conhecimento eram codificados em lógica *booleana*

95 Mackenzie adota em seu estudo o conceito de *diagrama* sugerido por Deleuze (2006) em sua leitura de Foucault.

(ALPAYDIN, 2016, p. 50) como forma de automação. Diferentemente do aprendizado de máquina e das redes neurais, portanto, são sistemas em que a lógica do programa encontra-se explicitamente descrita em código.

O principal ponto de divergência entre estes dois paradigmas diz respeito ao modo de compreensão da inteligência e de sua modelagem computacional. Enquanto o paradigma *simbólico* apoia-se em uma compreensão dedutiva, o paradigma *conexionista* apoia-se em uma compreensão indutiva. Pelo paradigma *simbólico*, desse modo, a inteligibilidade do sistema, ou seja, a possibilidade de descrever explicitamente o processo decisório, é uma qualidade *a priori*, já que é por meio desta elaboração que a máquina é programada. Além disso, a base dedutiva sugere que um horizonte de máxima eficácia e eficiência do sistema, que seria capaz de chegar sempre ao custo mínimo dos processos (ao que se chama, no jargão da área, de *convexidade*). Para o paradigma *conexionista*, de outro modo, não apenas a cadeia de decisões é definida *a posteriori* pela operação do sistema, como, também, não se apresenta nenhuma garantia de que este processo possa ser explicitamente explicado. Além disso, tampouco há garantia de *convexidade* a um ponto ótimo de eficiência ou eficácia, já que as máquinas indutivas operam por uma aproximação probabilística dos resultados. Seguindo o tortuoso debate recuperado por Cardon, Cointet e Mazières, o presente momento representa o auge do soerguimento do paradigma *conexionista*.

Como descrevem os autores (CARDON; COINTET; MAZIÈRES, 2018), a controvérsia entre os dois paradigmas envolve muitas camadas além de argumentos estritamente técnicos – incluindo políticas de financiamento de pesquisas e sucessos e fracassos circunstanciais de expoentes de uma ou outra vertente. Porém, os temas da *inteligibilidade* dos sistemas conexionistas e a não-garantia de uma *otimização convexa*, segundo descrevem, foram alguns dos principais focos do debate. A relativa vitória do paradigma *conexionista*, observada pelos autores desde pelo menos 15 anos não se deve, contudo, à definitiva superação destes problemas mas, de outro modo, a um argumento pragmático que considera que as limitações do paradigma *conexionista* seriam compensadas por outros benefícios, em particular devido a demandas circunstanciais do presente. Especificamente, as máquinas de base indutiva teriam prevalecido neste período recente por ao menos dois motivos: seu relativo sucesso em tarefas não superadas pelo paradigma *simbólico*; e a sua capacidade de operar em escala, atendendo às demandas emergentes das mídias digitais e da comunicação em rede.

Cardon, Cointet e Mazières (2018, p. 22) identificam um contraste discursivo operado pelos defensores do conexionismo entre a *largura* da arquitetura *rasa* de modelos precedentes, e a *profundidade* das arquiteturas baseadas em redes neurais – implicada, inclusive, na denominação das *redes neurais profundas* (DNN). Segundo o vocabulário da TAR (não adotado pelos autores) poderíamos dizer que os defensores do conexionismo conseguiram, por esta estratégia discursiva *transladar* os demais pesquisadores e fazer prevalecer seu enquadramento do problema. Em última medida, o debate passou a ser articulado nos termos de uma preferência por um ou outro atributo: *largura* ou *profundidade*. Escrevem:

Os cruzados do conexionismo conseguiram assim convencer as pessoas que era preferível sacrificar a inteligibilidade da calculadora e a otimização rigorosamente controlada por uma melhor percepção da complexidade de dimensões presentes nesta nova forma de dados. Quando o volume de dados de treinamento aumenta consideravelmente, muitos mínimos locais existem, mas há redundâncias e simetrias suficientes para que as representações aprendidas pela rede sejam robustas e tolerantes a erros nos dados de aprendizado. No centro do debate da comunidade de aprendizado de máquina, uma coisa era compreendida sem precisar ser dita: apenas laboratórios usavam modelos lineares; o mundo, o ‘mundo real’ onde os dados são produzidos pela digitalização de imagens, sons, fala, e texto, é não-linear. Ele é ruidoso; a informação nele contida é redundante; os fluxos de dados não são categorizados de acordo com atributos de variáveis homogêneas, claras e construídas de forma inteligível; exemplos às vezes são falsos⁹⁶ (CARDON; COINTET; MAZIÈRES, 2018, p. 22).

Esta citação contém muitos aspectos para desempacotar. Primeiro, sustentando o que eu indiquei logo acima, observa-se a construção do argumento pragmático de uma *compensação*: maior complexidade em detrimento da inteligibilidade e de uma eficiência e eficácia ótimas. Segundo, a indicação de uma ambivalência com respeito às consequências do treinamento extensivo de máquinas dedutivas em grandes quantidades de dados. A referência a “muitos mínimos locais” remete ao tema da otimização⁹⁷: mais dados tornam mais difícil aferir com exatidão o processo mais eficaz e eficiente, demandando o recurso a métodos probabilísticos que indicariam apenas aproximações do ponto ótimo. Porém, com mais dados, acredita-se que eventuais erros decorrentes desta limitação se-

96 No original: “*The crusaders of connectionism thus managed to convince people that it was preferable to sacrifice the intelligibility of the calculator and rigorously controlled optimization for better perception of the complexity of dimensions present in this new form of data. When the volume of training data increases considerably, many local minimums exist, but there are enough redundancies and symmetries for the representations learned by the network to be robust and tolerant to errors in learning data. At the heart of the debate with the machine learning community, one thing went without saying: only laboratories used linear models; the world, the ‘real world’ where data are produced by the digitization of images, sounds, speech, and text, is non-linear. It is noisy; the information contained in it is redundant; data flows are not categorized according to the attributes of homogeneous, clear, and intelligibly constructed variables; examples are sometimes false*”. Tradução minha.

97 Em uma topografia de dados ampla, um *mínimo local* seria um ponto que, por métodos probabilísticos de otimização, poderia ser equivocadamente identificado como o mais eficiente, ignorando o *mínimo global*.

riam minimizados por redundâncias e simetrias. Este ponto, em especial, abre uma ampla frente de discussão acerca da reprodução de vieses e discriminações sociais por tais sistemas: *mais dados* apenas tenderiam a nivelar assimetrias se estas não se manifestarem de forma sistêmica como, efetivamente, é o caso do racismo, do machismo e das assimetrias culturais. Longe de se anularem quando acumula-se mais dados, tais vieses tornam-se, nestes casos, ainda mais evidentes, pelo acúmulo reiterado de assimetrias difusas. Por fim, a citação discute o objetivo manifesto, por desenvolvedores de tais sistemas, de trabalhar dados *reais*, em um contraste entre a abstração lógico-matemática da computação aos aspectos concretos e ambivalentes do *mundo real*. Este ponto elabora, enfim, acerca da qualidade complexa dos dados que se busca abordar por meio das técnicas de aprendizado de máquina. Estes pontos sumarizam boa parte das questões que gostaria de lançar para as técnicas de aprendizado de máquina em seu modo de apreensão das imagens.

A relação entre o paradigma conexionista e o problema da computacionalidade das imagens é profunda. O *Perceptron* de Rosenblatt (1958) visava enfrentar justamente o problema da percepção visual, que se colocava como um dos desafios para a hipótese da IA, dado o grau de complexidade do tratamento da informação visual. Contudo, uma das narrativas acerca do desenvolvimento da Visão Computacional tende a situá-la como herdeira apenas de procedimentos lógico-dedutivos da *racionalização da visão*. Como indiquei anteriormente (veja 2.2.2 *Inscrição: perspectiva e consistência óptica*) este é o termo utilizado por Ivins (1975) para descrever o advento da perspectiva, no Renascimento, e seu modo de representação analítica do espaço. Depois dele, Latour (1986) mobiliza esta noção para dar conta de propriedades das imagens em perspectiva como inscrições científicas, conferindo às representações uma consistência óptica. Lev Manovich (1993) apoia-se, então, em ambos autores para descrever o advento da Visão Computacional como uma derivação da perspectiva tomada como instrumento analítico a ser codificado computacionalmente como forma de descrever corpos no espaço. Manovich recorre à noção de *nominalismo visual*, proposta por Alan Sekula (1986), para caracterizar este modo de ‘ver’ que teria sido precedido pela perspectiva geométrica e sistematizado pela visão computacional. O desenvolvimento fundador da versão algorítmica do nominalismo visual, segundo Manovich, teria sido a tese de doutorado de Lawrence G. Roberts (1963), desenvolvida na década de 1960 no MIT. Seu trabalho codificou, na máquina, um protocolo lógico-dedutivo de percepção de imagens

perspectivadas por um esforço de engenharia reversa, recompondo a tridimensionalidade do espaço e das formas a partir de aspectos de sua representação plana. Em minha dissertação de mestrado (MINTZ, 2015) também me apoiei nesta narrativa para caracterizar o desenvolvimento histórico da Visão Computacional. Mas, efetivamente, trata-se de apenas uma parte da história.

O percurso traçado por Manovich, situando a Visão Computacional em uma trajetória mais extensa do nominalismo visual, apoia-se, afinal, em uma compreensão dedutiva da percepção visual que pouco diz respeito às técnicas contemporâneas de aprendizado de máquina. Já no âmbito de minha dissertação, indiquei, em vista destas técnicas, que além de uma operação *nominalista*, haveria também operações que se aproximariam de um *idealismo* visual. Estas não seriam vinculadas a uma reconstrução espacial precisa, mas ao reconhecimento de padrões e à classificação de imagens – tarefas tipicamente realizadas por técnicas de aprendizado de máquina baseadas em redes neurais. Às tarefas nominalistas, indicava em meu texto um modo de operação descrito pelo par *localização-acionamento*. Às idealistas, descrevia por *reconhecimento-conexão*⁹⁸. Efetivamente, contudo, se consideramos a extensa controvérsia dos paradigmas *simbólico* e *conexionista* descrita por Cardon, Cointet e Mazières (2018), esta não seria uma bifurcação recente da Visão Computacional, mas pareceria ser uma ambivalência própria do campo, acompanhando o desenvolvimento destas correntes no campo ampliado da IA. Deste modo, um possível marco inicial da questão não seria Lawrence G. Roberts, mas a proposição do Perceptron por Rosenblatt.

Recompôr a narrativa histórica a partir destes desenvolvimentos vai além dos objetivos desta tese⁹⁹, mas gostaria de reter, acerca deste ponto, o vínculo profundo entre o paradigma *conexionista* e o problema da Visão Computacional, embora esta vertente de enfrentamento do problema tenha sido esquecida, em parte, pela prevalência do paradigma *simbólico* na segunda metade do século passado. Também, quero ressaltar a dupla tarefa empreendida no tratamento computacional das imagens e como cada paradigma proporcionou ênfases em operações diferentes. Sob o paradigma *simbólico*, a ‘interpretação’ computacional das imagens demanda compreensão sobre os modos de formação e constituição da imagem que devem ser, então, codificados na máquina. Demanda um modelo explícito da câmera, por exemplo, como fator de tradução do mundo

98 Agradeço a meu orientador de mestrado, André Brasil, por me ajudar a chegar nestas indicações.

99 Matteo Pasquinelli, professor da HfG Karlsruhe, tem desenvolvido um percurso histórico dos desenvolvimentos tecnológicos e teóricos da Visão Computacional por redes neurais (PASQUINELLI, 2017)

em imagem, para que se possa realizar o movimento inverso, da imagem ao mundo. Manovich (1993) destaca, inclusive, como Roberts (1963) teria voltado a formalizações da perspectiva geométrica elaboradas em tratados renascentistas, como o de Leon Battista Alberti, para desenvolver seu protótipo de uma percepção maquínica. Já para o paradigma conexionista, o modelo deriva de uma população de exemplos, a partir dos quais um modo de percepção é constituído por princípios estatísticos e probabilísticos. Trata-se de uma abordagem que tende a ser melhor sucedida para modelar categorias de entidades representadas nas imagens mas, talvez, não tanto a recompor, com precisão, aspectos geométricos do espaço retratado.

Portanto, a incidência de técnicas de aprendizado de máquina sobre a Visão Computacional impacta tanto os modos com que o problema foi enfrentado ao longo das décadas precedentes quanto aspectos da visualidade informada por tais sistemas. De uma visualidade analítica exemplificada pela identificação de objetos a partir de seus detalhados contornos (Figura 6), passa-se a uma visualidade sintética em que contornos dizem menos do que uma miríade de texturas superpostas, no entrecruzamento das quais os objetos emergem como *padrões* (Figura 7). Em um caso, busca-se deduzir, a partir da imagem, aspectos do espaço e dos objetos retratados, concretamente – *nominalismo visual*. Em outro, a imagem é vista como instanciação concreta de uma categoria abstrata, com a qual se relaciona por uma relação de *tipicidade* – *idealismo visual*¹⁰⁰. Embora problemas pertinentes ao *nominalismo* seguem centrais em diversas áreas de desenvolvimento contemporâneo da Visão Computacional – em especial na robótica, vigilância e ciências forenses; seriam operações vinculadas à vertente *idealista*, de base indutiva, que teriam maior proeminência em campos difusos como a mediação algorítmica do visível em plataformas digitais e mesmo em tecnologias de reconhecimento de rostos, por exemplo. Para estas, o *valor operativo* de uma imagem¹⁰¹ não é tanto como inscrição que permite atuar sobre um espaço concreto mas, sim, como inscrição que permite atuar sobre um campo semântico relacional, por seus vínculos a categorias e identidades. As categorias constituídas pelo aprendizado de máquina, contudo, não são formulações abstratas, em si mesmas, mas operam como condensações de instâncias concretas e individuais colhidas a partir de dados do ‘mundo real’. Este movimento

100 Em minha dissertação de mestrado e em outras oportunidades (MINTZ, 2015, 2016, 2018d) indiquei as relações destas operações com os casos exemplares, abordados por Alan Sekula (1986), no contexto da criminalística e da criminologia do século XIX. Não aprofundarei esta discussão aqui.

101 Veja que, como indiquei antes, vale qualquer imagem, não apenas aquelas destinadas, em sua origem, a uma finalidade operativa, segundo a formulação das imagens operativas de Harun Farocki (2004).

emergente de constituição das categorias, junto à ininteligibilidade deste processo, fazem do aprendizado de máquina por redes neurais um objeto crítico para investigações contemporâneas dos STS.

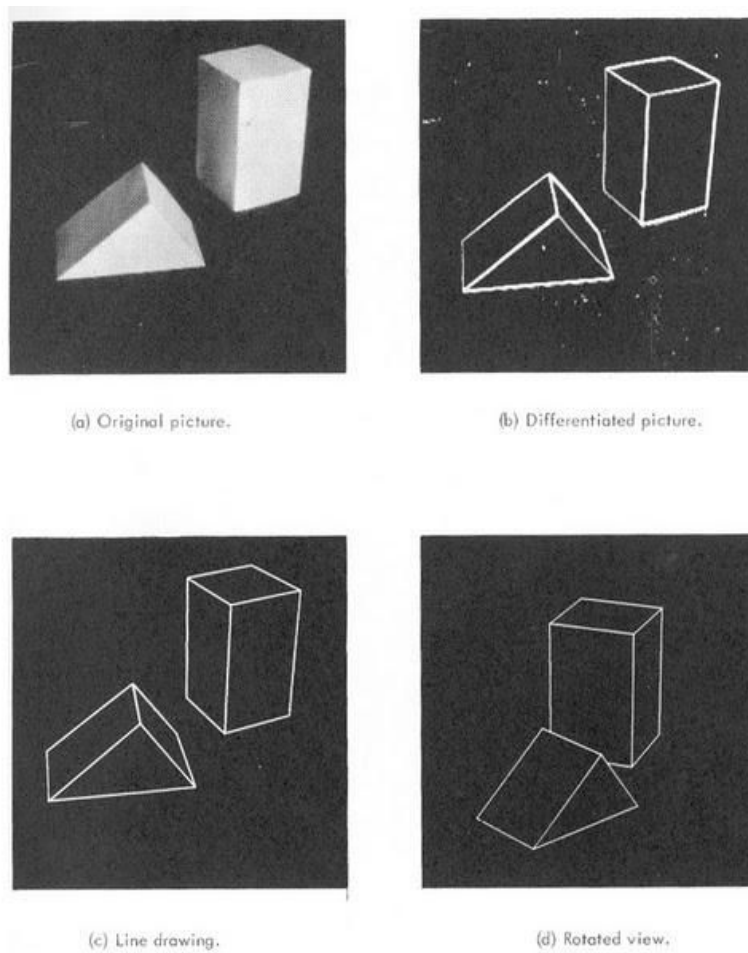


Figura 6: Ilustração do programa desenvolvido por Lawrence G. Roberts. Exemplo de implementação do modelo nominalista de base simbólica de Visão Computacional.

Fonte: ROBERTS, 1963.

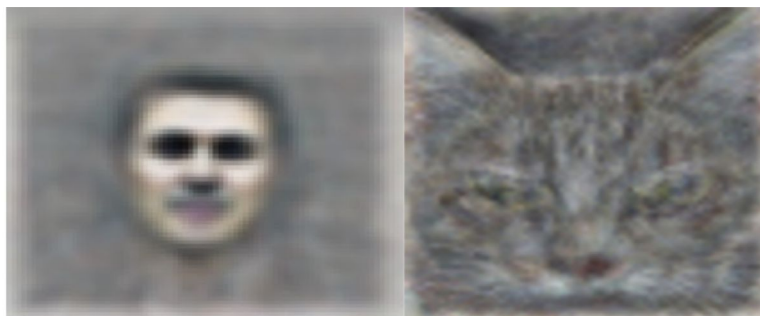


Figura 7: Ilustração do ‘estímulo ótimo’ para classes de um modelo de rede neural. Figuras incluídas como parte de um estudo realizado em uma parceria entre Google e a Universidade de Stanford, com técnicas de aprendizado não-supervisionado a partir de imagens colhidas do YouTube. À esquerda é apresentado o ‘estímulo ótimo’ para a categoria ‘humano’. À direita, para a categoria ‘gato’.

Fonte: LE et al., 2012.

Cardon, Cointet e Mazières (2018) sugerem um conjunto de categorias analíticas que podem auxiliar a compreender tal reconfiguração. Eles indicam que os diferentes estágios das máquinas preditivas compreendidas pelo amplo campo da IA poderiam ser comparados segundo as posições relativas que atribuem ao que chamam de *mundo*, *calculadora* e *alvo*. O *mundo* seria aquilo sobre o qual as máquinas operam – sua *exterioridade*, por assim dizer, que adentra a máquina na forma de *dados de entrada (inputs)* representando um ambiente, uma realidade específica ou de uma base de conhecimento, por exemplo, a depender do caso abordado. A *calculadora*, por sua vez, seria o programa, aquilo que irá operar com o *mundo* para chegar a determinado *alvo*. Didaticamente, os autores sugerem que o paradigma conexionista realizaria uma inversão:

enquanto os desenvolvedores de máquinas simbólicas buscaram inserir na calculadora tanto o mundo quanto o alvo, o sucesso atual das máquinas conexionistas relaciona-se ao fato de que, quase em contraste, seus criadores esvaziam a *calculadora* para que o *mundo* possa adotar o seu próprio *alvo*¹⁰² (CARDON; COINTET; MAZIÈRES, 2018, p. 6).

Mais adiante em seu texto (CARDON; COINTET; MAZIÈRES, 2018, p. 27), os autores detalham melhor a efetiva disposição destas categorias no aprendizado de máquina, indicando que as DNNs operariam como a calculadora e que o *mundo* seriam representações datificadas do mundo na forma de *vetores de big data*. Também, que o alvo seria a minimização dos erros segundo os objetivos traçados pela tarefa. Ou seja, não é que a *calculadora* seja de fato esvaziada ou que o *alvo* derive diretamente do mundo. Mas a for-

102 No original: “while the designers of symbolic machines sought to insert in the calculator both the world and the target, the current success of connectionist machines is related to the fact that, almost in contrast, their creators empty the calculator so that the world can adopt its own target”. Tradução minha.

mulação anterior dos autores tem um valor didático ao capturar um aspecto central à sustentação discursiva do aprendizado de máquina. A justificativa do paradigma conexionista, especialmente considerando seu valor epistêmico, seria fundado na ideia de que seus modelos de operação seriam derivados do próprio *mundo*.

Para o caso das técnicas de aprendizado de máquina aplicadas às imagens, seria como dizer, por tal asserção, que as máquinas aprendem a ver por meio das próprias imagens que compõem o seu *mundo*. Que estas imagens definiriam, portanto, a visualidade do aprendizado de máquina. Esta afirmação tem validade relativa, desde que se compreenda a abrangência apenas parcial deste conjunto de imagens utilizados para o treinamento, bem como sua pre-estruturação. Por mais amplo que seja, nenhuma base de dados pode se equivaler à totalidade dos dados existentes, evidentemente, e, mesmo que pudesse, nem mesmo a totalidade dos dados poderia ser tomada como uma representação *objetiva* ou *imparcial* do mundo. Nem todo o mundo se representa em dados – contra as premissas do *dataísmo* (cf. VAN DIJCK, 2014). Os dados tampouco seriam meras impressões do mundo em si mesmo, como discuti, anteriormente, na recuperação teórica do conceito de *mediação*. Neste sentido, portanto, modelos de reconhecimento de imagens baseados em aprendizado de máquina oferecem-se como instâncias de visualidades particulares, longe de qualquer suposição de uma visão universal. *Visualidades* computacionais, contra o subtendido universalismo de de uma *visão* computacional. Uma compreensão não positivista do aprendizado de máquina seria possível, portanto, compreendendo-o não como um instrumento neutro de conhecimento do mundo, mas como instância que também expressa aspectos do mundo sobre o qual opera. Este é, para esta pesquisa, o valor metodológico atribuído ao aprendizado de máquina.

Contra a ideia de que o *alvo* derivaria do próprio *mundo* por uma simples intermediação (em contraste com uma mediação plena) de uma *calculadora*, portanto, seria preciso observar criticamente como estas três instâncias estruturam a operação das máquinas indutivas. O caráter estruturante do *mundo* poderia ser descrito em uma investigação dos dados de treinamento, como nos esforços de auditoria realizados por Joy Buolamwini, Timnit Gebru (2018), que observaram vieses discriminatórios, de gênero e raça, em sistemas comerciais de análise facial em uma perspectiva interseccional. Também poderíamos citar os trabalhos de ativismo artístico de Adam Harvey (2019) e de Kate Crawford e Trevor Paglen (2019), que se voltam à proveniência e à constituição de bases

de treinamento massivas de ampla disponibilidade – como a Imagenet (DENG et al., 2009). O caráter estruturante do *alvo* pode ser observado em esforços de mapeamento dos espaços semânticos das APIs e dos vieses expressos em seus resultados, seja pelas categorias atribuídas às imagens, seja pelos índices de probabilidade que as acompanham. Um esforço deste tipo foi empreendido em estudo do qual participei, compartilhando a coordenação com Tarcízio Silva (MINTZ; SILVA et al., 2019), no qual realizamos uma investigação comparativa de três APIs comerciais de reconhecimento de imagens – Google, IBM e Microsoft.

Observar o caráter estruturante da *calculadora*, contudo, tende a ser mais desafiador, dada a ininteligibilidade que é característica das redes neurais. Um caminho possível é sugerido, quero argumentar, pela sugestão de Bernhard Rieder (2018) da noção de “técnica algorítmica”, em substituição tanto a remissões genéricas e, em alguma medida, essencialistas a um “*algoritmo*” indiferenciado; quanto a esforços de investigação de implementações concretas de um algoritmo específico. Escreve Rieder (2018, p. 126–127):

Cada técnica gira em torno de uma ideia central, um núcleo conceitual que normalmente é estabelecido através da combinação de linguagem natural e notação matemática. A técnica fornece uma lógica geral e especificações de cálculo formal, mas para que a implementação efetiva desses elementos em um sistema funcione efetivamente, requer que muitas decisões sejam tomadas [...]. Codificar [Programar], então, é uma forma de expressar essas técnicas em termos que um computador possa entender, e os algoritmos concretos são resultado de encontros situados entre ambientes de computação, técnicas algorítmicas e requisitos locais.

Sua elaboração não se endereça diretamente ao problema da inteligibilidade de redes neurais, mas parece-me que ela também pode auxiliar a compreender este contexto. Embora tenham implicações distintas, o tema da ininteligibilidade da operação das redes neurais e a generalização de implementações concretas de técnicas algorítmicas compartilham desafios metodológicos. Afinal, um dos sentidos da proposição da *técnica algorítmica* como objeto de análise por Rieder é justamente a inacessibilidade de implementações concretas em códigos proprietários. Os algoritmos “que importam” são tipicamente tratados como segredos industriais. Isto não impede, porém, que sejam estudadas as técnicas sobre as quais tais algoritmos, muito provavelmente, são baseados. Estas podem ser estudadas por sua fundamentação teórica e podem, também, render descrições *diagramáticas* que, como elabora Mackenzie (2017, p. 17) – a partir da leitura deleuziana de Foucault (DELEUZE, 2006) – seriam uma forma de de desenho que suavi-

za variações em favor da descrição de relações de força que constituem uma máquina, ou agenciamento. Ainda que não consigamos apreender detalhes da operação de uma rede neural, portanto, podemos ainda abordá-la de forma similar.

Para tratar de um exemplo, um dos modelos de rede neural que serão empregados no estudo de caso desta tese, denominado VGG19 (a sigla indica *Visual Geometry Group*, da Universidade de Oxford), é composto por 19 camadas com pesos¹⁰³ para transformação dos dados (SIMONYAN; ZISSERMAN, 2014) (veja Anexo A). Estas camadas distribuem-se em milhares de ‘neurônios’, cada qual com um peso específico – um fator numérico de transformação dos dados – definido a partir das rodadas de treinamento, com base em imagens e classes da base *Imagenet* (DENG et al., 2009). Considerando a atuação conjunta das camadas e neurônios, o modelo VGG19 tem 144 milhões de parâmetros para a realização da tarefa de classificação. A ininteligibilidade do modelo decorre, portanto, da impossibilidade de estabelecer vínculos precisos entre cada um destes parâmetros e as classes atribuídas ao final. Não há como descrever uma cadeia causal simples que conecta determinada imagem a sua classe e não há como intervir cirurgicamente na rede neural de modo a modificar um padrão de classificação observado (Figura 8).

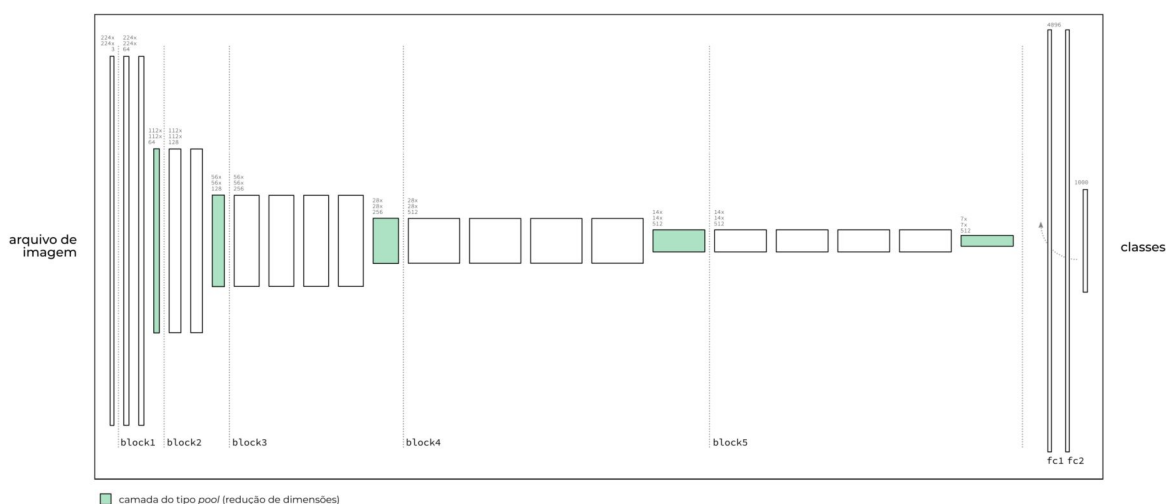


Figura 8: Arquitetura da rede neural do modelo VGG19.

Fonte: Elaboração própria com base na descrição do modelo (SIMONYAN; ZISSERMAN, 2014).

¹⁰³ Nem todas as camadas de uma rede neural realizam transformações efetivamente derivadas dos dados de treinamento. Algumas delas realizam transformações de redução de dimensionalidade, por exemplo, que independem do treinamento. As camadas “com pesos” (*weighted*) seriam as treinadas.

Uma abordagem possível do problema, contudo, é utilizar os modelos de rede neural para *gerar* imagens e não apenas para interpretá-las. No exemplo da [Figura 7](#), os pesquisadores (LE et al., 2012) o fizeram a fim de obter uma representação visual do estímulo ótimo vinculado a duas classes da rede neural que construíram em sua investigação. Similarmente, pode-se gerar estes estímulos ótimos para etapas intermédias da rede neural, que precedem à classificação. Apresento nas [Figuras 9](#) e [10](#) visualizações dos estímulos ótimos para doze dos ‘neurônios’ de duas camadas distintas do modelo VGG19. As visualizações foram geradas a partir de código e tutorial elaborado por François Chollet (2016). Especificamente, são apresentados estímulos ótimos para doze neurônios pertencentes a duas camadas do modelo VGG19¹⁰⁴. Estes dois extratos da rede neural permitem visualizar, portanto, o crescente nível de complexidade das formas que ativam neurônios específicos da rede e que compõem, ao final, os atributos que fundamentam tarefas de classificação. Os aspectos formais iniciais indicam apenas texturas, relativamente uniformes [Figura 9](#). Em um nível mais profundo de processamento, estas texturas se complexificam em formas mais elaboradas mas que, ainda assim, não são facilmente reconhecíveis como pertencentes a classes específicas [Figura 10](#). Uma *classe* ou *categoria* reconhecida pelo modelo deriva, efetivamente, de uma combinação específica de atributos identificados por filtros como estes. Elas são efeitos emergentes de atributos difusos, distribuídos entre os nós da rede neural. Embora a atribuição de classes aconteça apenas na última camada do modelo, *predictions* (veja [Figura 8](#) e [Anexo A](#)), as camadas anteriores incorporam aspectos das classes durante o processo de treinamento, cujo objetivo é otimizar o funcionamento da rede neural de modo obter uma melhor convergência entre as previsões e as classes conhecidas.

¹⁰⁴ `block1_conv2`, a segunda camada convoluta utilizada no processamento das imagens; e `block5_conv1`, a primeira do tipo do último bloco de processamento, que antecede a redução da dimensionalidade para a tarefa de classificação

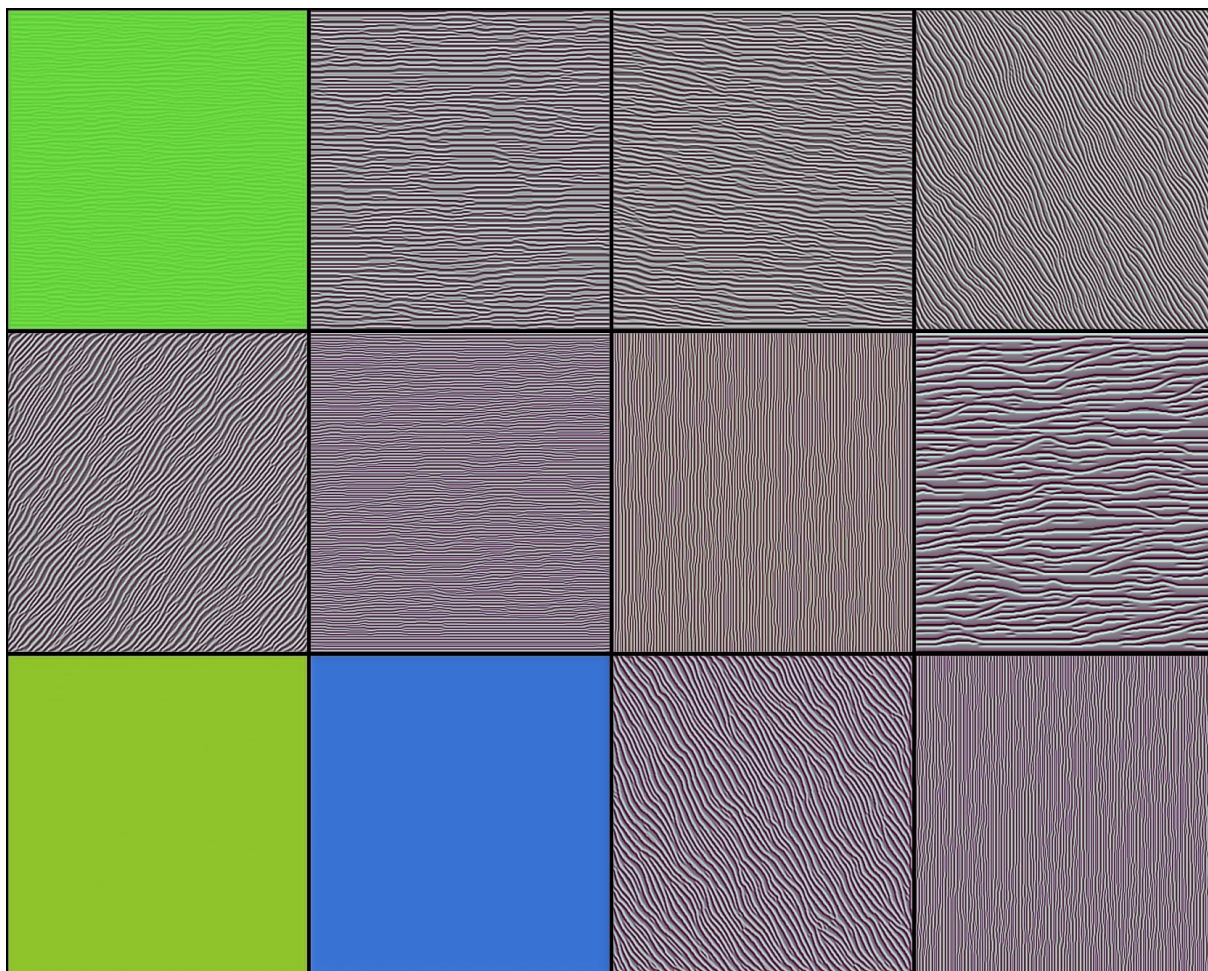


Figura 9: Estímulos ótimos para 12 filtros da camada `block1_conv2`, modelo VGG19. Seleção de 12 dos 64 neurônios da camada.

Fonte: Elaboração própria a partir do modelo VGG19 (SIMONYAN; ZISSERMAN, 2014) e tutorial de Chollet (2016).

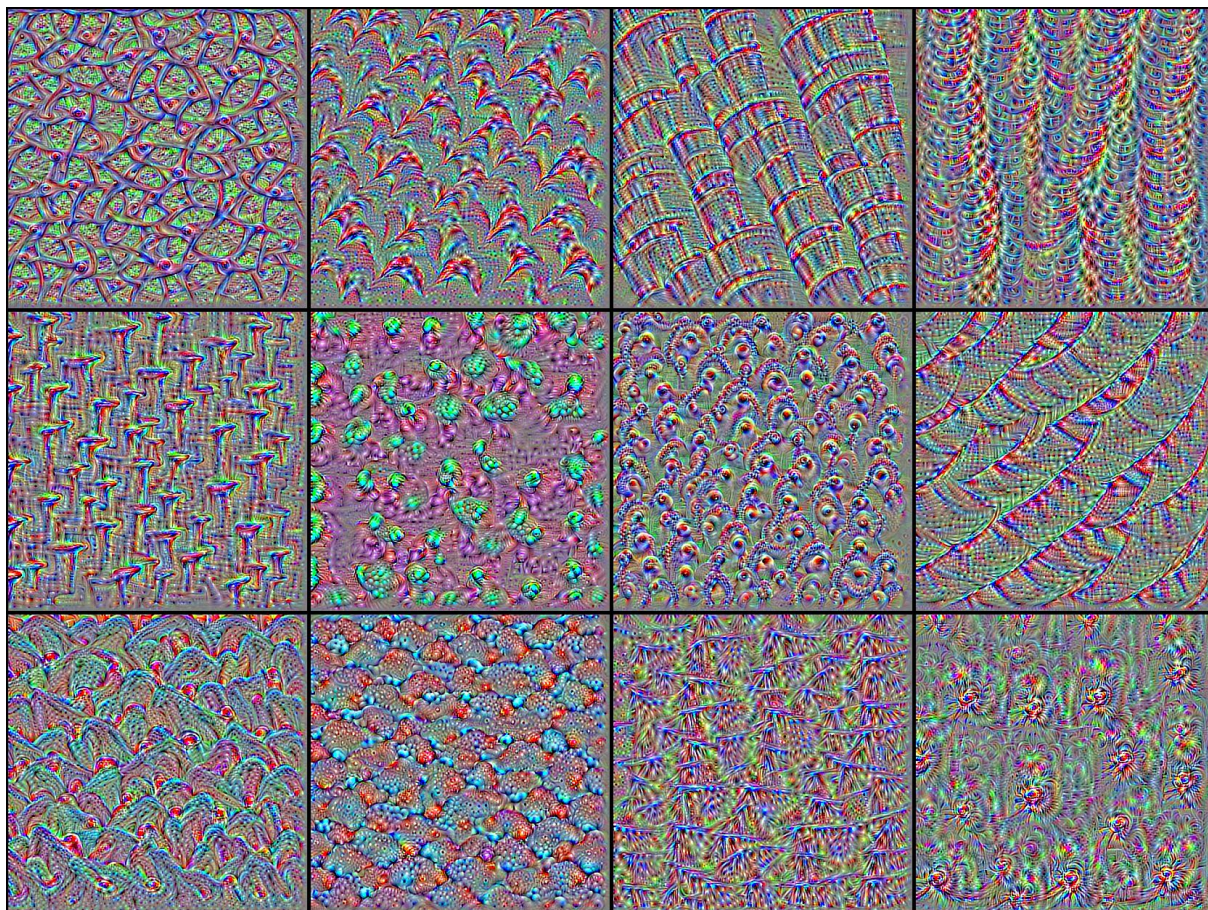


Figura 10: Estímulos ótimos para 12 filtros da camada `block5_conv1`, modelo VGG19. Seleção de 12 dos 512 neurônios da camada.

Fonte: Elaboração própria a partir do modelo VGG19 (SIMONYAN; ZISSERMAN, 2014) e tutorial de Chollet (2016).

Experimentos como este, de geração de imagens a partir de redes neurais, popularizaram-se em 2015 a partir de técnicas desenvolvidas por engenheiros da Google e divulgadas sob os nomes sugestivos de *DeepDream* (“sonho profundo”) e *Inceptionism*¹⁰⁵ (MORDVINTSEV; OLAH; TYKA, 2015a, 2015b). De modo similar às visualizações apresentadas, estes experimentos exploram o potencial generativo das redes neurais. Porém, ao invés de amplificar os estímulos ótimos de neurônios específicos a partir de padrões aleatórios, a técnica desenvolvida pelos autores parte de imagens específicas e amplifica os padrões que, nelas, ativam tais neurônios. Deste modo, os aspectos destas imagens que as relacionam com os padrões visados pela rede neural são explicitados. Por exemplo, características apenas sugeridas em uma imagem, que pudessem ativar a categoria de

¹⁰⁵ A tradução não seria simples neste caso, pois trata-se de uma remissão ao filme *Inception* (2010), dirigido por Christopher Nolan, cujo título adotado no Brasil foi *A Origem*. Um dos elementos principais da narrativa do filme é uma técnica utilizada por ladrões para invadir os sonhos das vítimas. *Inceptionism* para redes neurais, portanto, seria um modo de invadir os “sonhos” das redes neurais.

um templo de arquitetura asiática (*pagode*), são desse modo amplificados em uma fotografia de paisagem em que, efetivamente, tais construções não se encontram (Figura 11). Boa parte das apropriações artísticas de redes neurais se realizam, hoje, em versões sofisticadas destes mecanismos. Especificamente, pelo que Ian Goodfellow e colaboradores (GOODFELLOW et al., 2014) propuseram como *Redes Adversárias Generativas* (GAN, na sigla em inglês). Um exemplo perturbador deste tipo de aplicação é a página Web *This Person Does Not Exist*¹⁰⁶ (“Esta pessoa não existe”), que utiliza destas técnicas para gerar imagens de rostos humanos derivadas de modelos de reconhecimento de imagens¹⁰⁷.

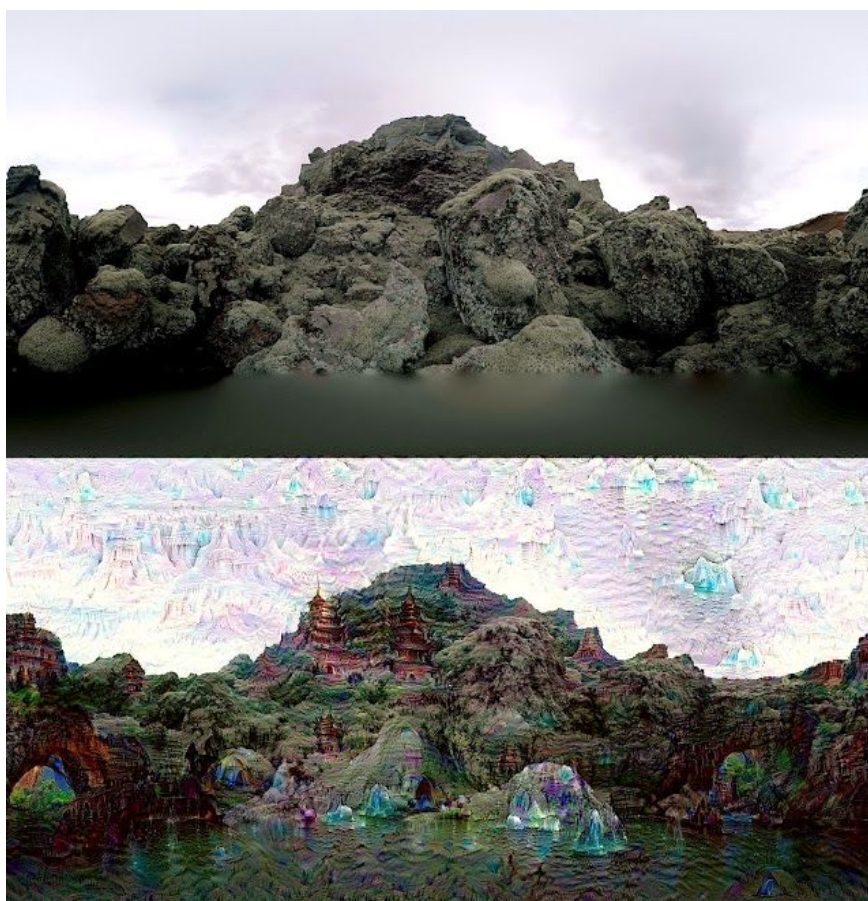


Figura 11: Exemplo de imagem gerada pelo programa DeepDream.

Fonte: MORDVINTSEV; OLAH; TYKA, 2015b.

A principal questão que gostaria de reter da discussão desta subseção refere-se ao modo de constituição dos modelos de redes neurais que, embora escapem descrições simplis-

106 Cf. <https://thispersondoesnotexist.com>

107 Embora seja um tema hoje fundamental, a discussão das implicações deste tipo de desenvolvimento vai além dos objetivos desta pesquisa.

tas, podem ser compreendidos como condensações de amplos conjuntos de imagens. *Aprender a ver*, para estas máquinas, significa encontrar um ponto de otimização de processos de classificação em que uma arquitetura difusa progressivamente se adéqua às relações entre imagens e categorias. Na linha do que discuti no capítulo anterior (veja [2. Imagem-rede](#)), gostaria de argumentar, portanto, que as redes neurais poderiam ser compreendidas como “centros de cálculo” (LATOURE; HERMANDT, 2004), constituídos a partir da coleção e recombinação de múltiplas inscrições colhidas do “mundo”. Neste caso, de milhares ou milhões de imagens capturadas em sua circulação *online*. Os modelos são, então, derivados de processos de condensação e destilação destas inscrições em descritores específicos. Não seria exagero dizer que cada operação de classificação realizada por uma máquina destas deriva, indiretamente, do extenso acervo de imagens utilizadas em seu treinamento. Deste modo, cada imagem individual que compõe estas bases de treinamento teria seu papel na constituição do modelo e, ao final, na interpretação que este modelo produz de uma imagem.

Figuras produzidas pela técnica do *DeepDream* ([Figura 11](#)), embora sejam decorrentes de uma reapropriação de modelos de reconhecimento, são ilustrativas de como estes múltiplos registros incidem sobre uma imagem submetida à interpretação da máquina. Sugerem algo como uma memória de padrões que se oferece como sintoma de uma visualidade computacional particular, inscrita naquele modelo de reconhecimento de imagens. Trata-se de um modo de tornar visível uma dimensão latente das relações traçadas entre as imagens por um “olhar” computacional específico. Uma brecha para compreendermos como aquele modelo *aprendeu a ver*. Esta seria, quero sugerir, uma das possíveis materializações contemporâneas da hipótese conceitual que sugeri, no capítulo anterior, por *imagem-rede*. Modelos de reconhecimento de imagens baseados em redes neurais são pontos de articulação entre imagens que, sob seu domínio, declinam-se em múltiplas inscrições que são progressivamente “comparadas”, “superpostas”, “recombinadas”. As imagens são decompostas e reconstituídas, neste processo, como efeitos emergentes das práticas de reconhecimento. Categorias são elaboradas como referências circulantes derivadas destas práticas.

Estes processos tornam-se tão mais relevantes quanto se difundem como principal modo de mediação algorítmica das imagens, em plataformas *online*, aplicativos de registro e organização de coleções fotográficas, em arquivos, em museus, em bibliotecas. Por meio destas aplicações, esta visualidade computacional particular incide diretamen-

te sobre a constituição de visualidades contemporâneas que extravasam a suposição de um domínio específico da máquina. Este enredamento, por sua vez, pode proporcionar uma operação metodológica para percorrer outros enredamentos. Os modelos de reconhecimento inscritos em redes neurais funcionam como expressões e como instrumentos de uma visualidade difusa das imagens em circulação. São objetos passíveis de reapropriação. É preciso uma rede para compreender uma rede: este é, em outros termos, a aposta metodológica desta tese.

3.2.2 Vetorização

Como discuti na seção anterior, uma das características mais proeminentes das técnicas de aprendizado de máquina é sua dependência de grandes bases de dados de treinamento. Simetricamente, por assim dizer, a demanda por tais sistemas é também decorrente da explosão exponencial dos volumes de dados por se analisar, nas mais diversas áreas do conhecimento e aplicações comerciais. Uma compreensão do aprendizado de máquina e das redes neurais deve também considerar, portanto, o modo como tais dados são incorporados a estas máquinas: as *inscrições* por meio das quais realizam suas operações. Tipicamente, os dados processados pelo aprendizado de máquinas assumem a forma de *vetores* e, em vista do percurso teórico elaborado a partir do capítulo anterior, trata-se de um tema fundamental para que possamos compreender as declinações ontológicas da imagem digital em seu processamento por redes neurais, em um dos caminhos por se perseguir em sua ontografia.

Adrian Mackenzie (2017, p. 51) indica que a *vetorização* seria uma operação fundamental que, inclusive, condicionaria a possibilidade de o aprendizado de máquina se situar em um “espaço epistêmico em expansão”. A centralidade da operação de *vetorização* seria tamanha, para o autor, que ele elabora que “aprendizes maquínicos” – como ele se refere às máquinas, junto a seus desenvolvedores e operadores – “habitam um espaço vetorizado”. Cardon, Cointet e Mazières (2018), embora expressem-se em outros termos, não divergem desta compreensão e descrevem o *vetor* como a forma assumida pelo *mundo* em seu contato com a *calculadora*. Citando indiretamente Yann LeCun, um cientista da computação de destaque no desenvolvimento contemporâneo do aprendizado de máquina, os autores dizem que “o objetivo dos desenvolvedores de máquinas conexionistas é colocar o mundo em um vetor”¹⁰⁸ (CARDON; COINTET; MAZIÈRES, 2018, p. 24).

¹⁰⁸ No original: “the goal of the designers of connectionist machines is to put the world in a vector”. Tradução minha.

Para recuperar termos utilizados no capítulo anterior, tomados de Madeleine Akrich (1992), o aprendizado de máquina teria como *prescrição* que um conjunto de dados, para que seja processado ou tomado como base de treinamento, precisaria antes ser transformado na forma *vetor*. Escrevem Cardon, Cointet e Mazières (2018, p. 23): “o mundo deve ser codificado de antemão na forma de uma representação digital vetorial pura”¹⁰⁹.

Compreender tais asserções demanda considerar as diferenças internas aos modos de representação da informação digital, aspecto que – como discuti no capítulo anterior – é frequentemente relevado em abordagens teóricas das mídias digitais que resumem a inscrição digital a uma simples “numerização”, em que *tudo* assumiria as formas de zeros e uns. Efetivamente, como já exemplifiquei com o caso do *codec* JPEG (veja 2.2.3 Multiplicidade ontológica: declinações materiais da imagem), há muitas formas pelas quais uma informação pode ser representada digitalmente. Embora todas elas, no nível mais baixo de abstração, sejam codificadas na forma de *bits* – os tais zeros e uns – em níveis mais elevados os dados podem assumir formas distintas que, por sua vez, favorecem determinadas operações em detrimento de outras. Um *vetor* é uma destas formas de nível intermediário, e sua estruturação dos dados é relativamente simples: trata-se de uma sequência linear de dados numéricos, de comprimento variável.

Para o aprendizado de máquina, os vetores funcionam como coordenadas em um espaço multidimensional. Efetivamente, cada número da sequência corresponde à coordenada de uma dimensão específica. Um vetor de 4.096 dimensões, por exemplo, descreveria a posição de um ponto em um espaço de 4.096 dimensões. Uma das consequências desta compreensão dos vetores no aprendizado de máquina é abordada por Mackenzie (2017) em contraste com outra estruturação de dados paradigmática: a *tabela*. O principal aspecto que gostaria de desenvolver, a respeito desta comparação, diz respeito ao que Mackenzie (2017, p. 56) sugere como um “colapso ou liquidação de bases de dados tabulares”¹¹⁰ pela vetorização.

A partir de Foucault¹¹¹, Mackenzie descreve como as tabelas foram formações de dados especialmente afeitas a certas operações da ciência e da política modernas, como a taxonomias. As tabelas favorecem uma *redução* de objetos complexos, realizando sua

109 No original: “*the world must be coded in advance in the form of a purely digital vectorial representation*”. Tradução minha.

110 No original: “*collapse or liquidation of tabular datasets*”. Tradução minha.

111 O autor se refere a diferentes obras do filósofo, mas especialmente à *Arqueologia do conhecimento* e *As palavras e as coisas*.

distribuição em planos regulares, divididos em linhas e colunas. Escreve Mackenzie (2017, p. 56–57): “De modo importante, a tabela como espaço de ordem era um espaço de imaginação, no qual uma pessoa poderia começar a ver continuidades e diferenças entre coisas (organismos, palavras, nações) por meio do ordenamento e do exame cuidadoso da tabela”¹¹². A estruturação do mundo na forma da tabela tende, afinal, a um ordenamento espacial cartesiano, em que identidades e diferenças operam de modo preciso. A cada ser, sua linha. A cada qualidade, sua coluna. As operações biopolíticas proporcionadas pela tabela também tendem a seguir estes eixos. Seguindo as linhas, adentramos no nível do indivíduo. Seguindo as colunas, acercamo-nos das populações e suas distribuições estatísticas. Mackenzie (2017) sugere que a vetorização, de outro modo, proporcionaria movimentos *diagonais*.

Um modo de compreender esta *diagonalização*, sugere Mackenzie, seriam estruturas de tabelas associadas, que compõem, por exemplo, mecanismos de bancos de dados. Como os vetores, também elas comportam uma complexidade maior do que as proporcionadas pelas tabelas – itens podem se distribuir em muitos espaços tabulares e, assim, ser submetidos a diferentes escalas e bases de comparação. Porém, seria importante observar como a expansão dimensional operada pelos vetores sugere um colapso estrutural mais profundo do tipo de ordem constituída pelas tabelas. À medida que todas as qualidades de um item individual são codificadas numericamente e condensadas na forma vetorial, todas elas contribuem ao posicionamento daquele item em um espaço multidimensional que, ao menos em princípio, aproxima-se de uma espacialidade lisa.

Um programa de aprendizado de máquina opera, então, pela navegação deste espaço. As *diferenças*, que proporcionam operações de *classificação* são identificadas pelos programas à medida que, por diversas operações, eles conseguem perseguir *estrias* no espaço vetorial – como Mackenzie formula a partir de Alfred North Whitehead¹¹³ (1960, citado por MACKENZIE, 2017). Elabora o autor:

Uma vez que esta distribuição de elementos no espaço existe – oculta, expansível e transformável (por rotação, deslocamento, ou mudanças de escala) – esforços vigorosos serão feitos para trazer *loci* [locais] à luz. Aprendizes maquínicos buscam estes *loci* ou tateiam *estrias* nos dados, para usar o termo de Whitehead, ao longo de diferentes linhas¹¹⁴ (MACKENZIE, 2017, p. 63).

112 No original: “Importantly, the table as a space of order was a space of imagination, in that one could begin to see continuities and differences between things (organisms, words, nations) by carefully ordering and scanning the table”. Tradução minha.

113 WHITEHEAD, Alfred North. *Process and reality, an essay in cosmology*. New York: Macmillan, 1960.

Tipicamente, as operações do aprendizado de máquina seriam descritas pela demarcação de porções do espaço vetorial que pudessem descrever *classes* ou *categorias* em meio aos dados. Isto porque a vetorização performa identidades e diferenças como proximidades relativas entre os entes no espaço vetorial. “No espaço vetorial, identidades e diferenças mudam de natureza. Similaridade e pertencimento não se fiam mais na aparência ou em uma gênese comum, mas em medidas de proximidade ou distância”¹¹⁵ (MACKENZIE, 2017, p. 73). Estas operações seriam proporcionadas pelos vetores em sua qualidade de expansão dimensional do espaço das representações, e também, pela *diagonalização* dos movimentos possíveis por este espaço, transpondo os limites estritos das linhas e colunas da tabela.

A vetorização também possui consequências pragmáticas para o processamento computacional e as práticas de programação sobre estes dados. Como elabora Mackenzie (2017, p. 67–68), a vetorização dos dados é acompanhada de transformações das linguagens de programação e da própria infraestrutura de processamento. Estruturas de dados lineares ou de poucas dimensões são tipicamente processadas por códigos recursivos, organizados em *loops*, que realizam operações individualmente sobre cada célula de uma tabela ou item de uma lista, por exemplo. Linguagens de programação que operam sobre vetores, de outro modo, recebem comandos que se endereçam simultaneamente a todos os elementos da estrutura de dados. Este é o caso da linguagem *R* e, também, do módulo *Pandas* da linguagem *Python*. Escreve Mackenzie (2017, p. 69): “O ponto verdadeiramente crucial na vetorização de dados não é a velocidade, mas a transformação da prática de dados. Ela torna o trabalho com dados menos parecida com a iteração por estruturas de dados [...] e mais parecida com a dobra de um material flexível”¹¹⁶. A indicação de que esta mudança não aumenta necessariamente a rapidez dos processos é importante. A execução dos comandos não é necessariamente mais veloz, mas a forma de pensar sobre os dados na prática de programação e processamento transforma-se substancialmente.

As consequências pragmáticas também incidem em um nível infraestrutural. Pois a obtenção de maior eficiência nestas operações vetoriais demanda arquiteturas de processamento específicas, que favoreçam o processamento paralelo e o trabalho simul-

114 No original: “Once this hidden, expandable, and transformable (by rotation, displacement, or scaling) distribution of elements in space exists, strenuous efforts will be made to bring loci to light. Machine learners search for these loci or feel for data strains, to use Whitehead’s term, along different lines”. Tradução minha.

115 No original: “In vector space, identities and differences change in nature. Similarity and belonging no longer rely on resemblance or a common genesis but on measures of proximity or distance”. Tradução minha.

116 No original: “The real stake in vectorizing data is not speed but a transformation in data practice. It makes working with data less like iteration through data structures [...] and more like folding a pliable material”. Tradução minha.

tâneo em múltiplas dimensões (MACKENZIE, 2017, p. 70). O processamento computacional demandado pelo aprendizado de máquina em redes neurais é, por isso, comumente realizado em *Unidades de Processamento Gráfico* (GPUs, na sigla em inglês), em uma espécie de reapropriação tecnológica no nível do *hardware*. Pois a finalidade que inicialmente motivou o desenvolvimento das GPUs foi o processamento de dados para computação gráfica, calculando a exibição de gráficos 3D de jogos digitais, por exemplo. Hoje, a demanda por tais unidades de processamento é também fortemente vinculada ao desenvolvimento de sistemas de aprendizado de máquina, em mais um exemplo das relações constitutivas entre este domínio e o problema da computacionalidade das imagens.

Na esteira da discussão do capítulo anterior (veja 2. Imagem-rede), estas considerações evidenciam a intrincada relação entre agenciamentos sociotécnicos e suas inscrições. A descrição da vetorização como processo de produção de inscrições e, de modo mais específico a esta discussão, de multiplicação ontológica das imagens traz, de arrasto, uma ampla rede de elementos que não podem ser considerados em isolamento. O vetor, enquanto forma de estruturação de dados *prescrita* pelo aprendizado de máquina indica como o processamento realizado por estas técnicas indutivas não ocorre sobre um domínio indiferenciado, como sugerem acepções essencialistas do *digital*. Os vetores proporcionam práticas de processamento e gestos de conhecimento específicos que, se não são exatamente impossíveis, tampouco são facilitadas por outras formas de estruturação. Como descreve Mackenzie (2017, p. 73), a vetorização “produz um espaço comum que justapõe e mistura realidades localizadas complexas”¹¹⁷. Vetores, ao definirem posições relativas em uma espacialidade multidimensional, favorecem um movimento diagonal pelos dados, em contraste com estruturas tabulares ou lineares típicas.

A compreensão destas possibilidades, porém, apenas ganha tração quando os vetores são considerados como parte do agenciamento constituído pelas técnicas de aprendizado de máquina e segundo as demandas que este agenciamento, por sua vez, visa atender. Segundo Mackenzie (2015), estas seriam demandas de *previsão* sobre realidades datificadas complexas. Para o caso específico das imagens, o aprendizado de máquina realiza tarefas de reconhecimento e classificação, ambas intimamente relacionadas, que atendem ao desafio de *datificação do visível*, tomando as imagens como instâncias operacionais de conhecimento e de ação sobre o mundo. As mediações implicadas por

117 No original: “*produces a common space that juxtaposes and mixes complex localized realities*”. Tradução minha.

tais operações distribuem-se entre a arquitetura destes sistemas, a estruturação de suas inscrições e os dados fornecidos como bases de treinamento. Por sua vez, as demandas geradas por tais mediações ampliam-se para além de qualquer possibilidade de circunscrição das máquinas indutivas, em direção às infraestruturas de processamento e práticas de programação e trabalho com dados.

Como consequência deste modo de operar sobre o mundo, haveria, no caso específico das imagens, um modo bastante peculiar de produção de inscrições sobre o visível. Uma rede neural como a esquematicamente descrita na [Figura 8](#) compreende uma máquina de transformação de vetores, que progressivamente desloca a imagem por um espaço vetorial até que, ao final, delimita *regiões* ou *locais* deste espaço que corresponderiam àquelas pertinentes às classes utilizadas no treinamento. Este deslocamento é orientado pelos filtros e pesos definidos para as camadas durante o período de “aprendizado”. Entre uma ponta e outra do processo, portanto, não se tem uma imagem *visível*. Estes vetores e os filtros que condicionam a ativação de determinados “neurônios” da rede podem ser *visualizados* de forma similar à produção de um gráfico mas eles não são, em si, visíveis. Talvez possamos compreendê-los a partir da noção que Paul Virilio (1993) propôs, certa vez, de “imagem virtual instrumental”, em uma analogia com as “imagens mentais”. Estaríamos no plano de representações internas ao processo de percepção maquínica do mundo que, para o caso do aprendizado de máquina por redes neurais, apresenta-se na forma destes vetores¹¹⁸.

Este complexo arranjo, por sua vez, produziria, como parte de sua operação, um “modo de ver” particular. Ao mesmo passo em que estas inscrições descrevem uma posição relativa para as imagens em um *espaço vetorial* multidimensional, a operação destas máquinas também constitui uma espacialidade específica segundo os dados de treinamento mobilizados. O escopo das imagens utilizadas para o treinamento configuram, afinal, um horizonte de possibilidades para o campo visual elaborado por estas máquinas. As estrias percorridas pelos sistemas de aprendizado em meio a este espaço podem ser descritas como variações de densidade. Cada base de treinamento e cada arquitetura de rede neural tenderia a produzir algumas regiões mais povoadas de imagens, em que características ou temas recorrentes formariam polos de atração para grandes agrupamentos; mas produziria também amplos desertos, em que características menos

¹¹⁸ Adiante, no estudo de caso apresentado no capítulo 4, voltarei a uma abordagem mais direta destas inscrições em uma operação de “extração de características”.

recorrentes se distribuem em uma espacialidade gasosa, esparsa. As características desta distribuição e, também, o perímetro descrito por seus limites seriam, quero sugerir, indicativos da *visualidade computacional* constituída por cada rede neural em sua operação. Cabe lembrar, aqui, o sentido específico do *aprendizado* das máquinas, que longe da apreensão do *novo*, tenderia sempre à reprodução de padrões implícitos na estrutura dos dados de treinamento. Justamente por esse motivo, o olhar que se volta às bases de treinamento como instâncias analíticas centrais dos modelos indutivos tem hoje imensa relevância (cf. BUOLAMWINI; GEBRU, 2018; CRAWFORD; PAGLEN, 2019; HARVEY, 2019).

3.2.3 Aprendizes maquínicos

Traduzo por “aprendizes maquínicos” a formulação adotada por Adrian Mackenzie (2017, p. 6) logo no título de seu volume dedicado a uma abordagem arqueológica do aprendizado de máquina – *Machine Learners*. Esta formulação visa se referir indistintamente a humanos e máquinas – ou a relações humano-máquina – como partes implicadas pelas técnicas de aprendizado de máquina. O sentido não é o de uma equivalência – como se a técnica produzisse máquinas plenamente inteligentes – mas sim de uma mútua afetação. Nesse sentido, a adjetivação da palavra máquina no português como “maquínic”, embora se afaste da formulação original de Mackenzie (que não fala, afinal de *machinic learners*), abre uma possibilidade de aproximação que, a meu ver, potencializa sua força conceitual se considerarmos o significado do termo *maquínic* na visada filosófica de Deleuze e Guattari. Pois, para estes autores, a noção de *máquina* não corresponde apenas às máquinas tecnológicas, mas a um problema que as antecede – algo que Guattari (2003) enuncia em um texto de revisão. De outro modo, para ele, a técnica seria um subdomínio da problemática maquínica que seria uma “categoria que engloba tudo o que se desenvolve como máquinas nos diferentes registros e suportes ontológicos” (GUATTARI, 2003, p. 42). Esta acepção possibilita, então, a descrição de diferentes fenômenos como *maquínicos*: desde a cidade até processos psíquicos, como o desejo.

Em comum, estes *agenciamentos maquínicos* articulariam movimentos complementares de *autopoiese* e *alopoiese*, isto é, um modo de funcionamento que produz e articula uma dimensão interior a uma dimensão exterior. De modo importante para o que gostaria de desenvolver aqui, as máquinas são compreendidas por Guattari (a partir de Leibniz)

como articuladas em um desenho fractal: “aquém e além da máquina, o ambiente da máquina faz parte de agenciamentos maquínicos” (GUATTARI, 2003, p. 41). Em suma, quero com esta aproximação sugerir que a noção de *aprendizes maquínicos* – esta corruptela que proponho, a pretexto de uma tradução – permite compreender o aprendizado de máquina como agenciamento maquínico que se articula, para além de todo isolamento, como um processo pertinente tanto às máquinas tecnológicas quanto aos humanos que desenvolvem tais máquinas e que elas, por sua vez, alcançam.

Um dos primeiros modos de compreender esta relação deriva do deslocamento que o paradigma do aprendizado de máquina provoca para o papel da pessoa desenvolvedora dos sistemas. Cardon, Cointet e Mazières (2018, p. 3) sintetizam: “O que era previamente concebido como o componente ‘humano’ na criação de calculadoras, o programa, as regras ou o modelo, não era mais um elemento inserido no sistema mas, sim, o seu resultado”¹¹⁹. Adrian Mackenzie (2017, p. 21) retoma de Pedro Domingos (2012) a metáfora que concebe o desenvolvimento de sistemas de aprendizado de máquina como “agricultura” (*farming*)¹²⁰. Isto é, enquanto a programação tradicional demanda a *construção* dos programas, o aprendizado de máquina permite que eles sejam “cultivados”. Como discuti, este deslocamento se encontra já presumido pela transição de um paradigma dedutivo, ou simbólico, a um paradigma indutivo, ou conexionista. Pois a função da programação, de uma abordagem à outra, deixa de assumir um papel tão claro quanto é a definição das operações que serão realizadas pela máquina em sua tarefa de produzir classificações e previsões. Ou seja, a pessoa programadora não sabe, de antemão, como o programa gerado por meio do aprendizado de máquina irá se comportar e nem precisa conhecer exatamente as características dos dados para que este processo gere algum resultado. Neste sentido, não é apenas a máquina que aprende, mas também a pessoa desenvolvedora. O modelo produzido pela máquina seria tão útil para produzir asserções sobre outros dados, além da base de treinamento, quanto para melhor compreender a constituição dos dados de treinamento.

Um dos efeitos desta mudança são as concepções equivocadas de que tais processos poderiam conferir uma *objetividade* aos programas, como se os dados “falassem por si

119 No original: “What was previously thought of as the ‘human’ component in the creation of calculators, program, the rules, or the model was no longer input into the system but rather its result”. Tradução minha.

120 Ironicamente, esta aceção reverbera com o fato de que boa parte do processamento realizado para o desenvolvimento destes programas ocorre em “fazendas de servidores” (*server farms*): grandes conjuntos de computadores, geralmente localizados em zonas remotas, que constituem um dos elementos fundamentais da infraestrutura da computação “em nuvem”.

próprios”. Efetivamente, como discuti, o paradigma do aprendizado de máquina opera por meio da condensação de enormes quantidades de exemplos em modelos progressivamente otimizados de modo a se adequar a estes dados. Mas, partir desta observação para supor tal *objetividade* é tomar um atalho argumentativo facilmente contestável. Pois conceber tais programas como *objetivos* ou *neutros* pressupõe que tais qualidades se estenderiam tanto para a constituição da máquina quanto aos dados e classificações prévias utilizadas para o treinamento. No entanto, deve-se reconhecer como o deslocamento da posição da pessoa desenvolvedora complexifica qualquer presunção de controle que essa pessoa poderia ter sobre o programa resultante. Contestar a objetividade ou neutralidade do programa não deve se equivaler a supor que o programa fosse um produto da *subjetividade* de quem o programou. Mesmo atribuir um vínculo muito estreito entre a operação do programa e uma intencionalidade institucional que o engendrou seria um movimento arriscado, dado o caráter emergente dos processos indutivos implicados. Há, evidentemente, uma margem de agência sobre o desenvolvimento dos programas, no desenho da arquitetura, na preparação dos dados de treinamento e em outras etapas do processo. Mas este controle não é completo e deve-se reconhecer o caráter *mediador* da técnica, que não se dobra, como discuti (veja [2.2.1 Mediação técnica: da luz à imagem](#)), a partir de Latour (2001), como uma simples ferramenta sob o comando de um super-ator. Contra a tentação de atribuir uma agência prioritária a um ou outro ator, humano ou máquina, é preciso enquadrar a questão sob a perspectiva de uma associação, em um construto sociotécnico. Este seria o primeiro sentido da ideia de *aprendizes maquínicos*.

Um segundo sentido poderia ser articulado considerando a partir do quê ou de quem a máquina aprenderia – se há máquinas que “aprendem”, haveria “educadores” de máquina? Se consideramos os dados de treinamento como uma construção realizada por (geralmente) muitos atores, sim, poderíamos encontrar múltiplas instanciações destes “educadores”. O “ensino”, tipicamente, ocorreria de forma distribuída e, muitas vezes, sem que os “educadores” tenham consciência de sua atividade. No contexto das plataformas *online*, trata-se de atividade tanto inconsciente quanto involuntária. Os limites da datificação nestes contextos nunca é claro mas, potencialmente, toda atividade realizada sob seu domínio seria passível de integração a um sistema de aprendizado de máquina – se já não como treinamento, certamente como um dado utilizado para gerar uma previsão a partir de um modelo pré-construído.

Em uma situação já mais consciente, mas nem por isso opcional, há o curioso caso do projeto *ReCAPTCHA*, hoje de propriedade da Google, que articula de maneira intrincada duas tarefas aparentemente paradoxais. Por um lado, identificar se a navegação em uma página da Web está sendo realizada por um humano ou por um “robô”, por meio de uma tarefa – tipicamente de Visão Computacional – que uma máquina geralmente teria dificuldade em realizar. Por outro, produzir dados de treinamento para aprendizado de máquina exatamente para que uma máquina seja capaz de realizar aquela tarefa. A sigla CAPTCHA significa *Completely Automated Public Turing Test to Tell Computers and Humans Apart* (algo como “Teste de Turing público completamente automatizado para diferenciar humanos de computadores”). O nome faz menção à famosa proposição do cientista da computação britânico Alan Turing ([1950] 2003), que sugeriu um teste como forma de declarar uma máquina como *inteligente*. Os detalhes de seu teste não vêm ao caso aqui, mas importa compreender que Turing o propôs como um parâmetro de referência para a inteligência, enquanto, de outro modo, os CAPTCHAs são testes de valor pragmático que visam *distinguir* e não *equivaler* humanos e máquinas. A forma prototípica do CAPTCHA envolvia a exibição de uma imagem ruidosa de letras e números para que o agente que buscava acessar uma página Web reconhecesse os caracteres. Os ReCAPTCHAs, por sua vez, utilizam a informação produzida nesse reconhecimento, pelos agentes humanos, para gerar bases de treinamento para o aprendizado de máquina. A Google inicialmente utilizou deste dispositivo para aperfeiçoar o reconhecimento de caracteres utilizado em sua base de livros. Hoje, o treinamento parece se dirigir especialmente ao reconhecimento de imagens do produto Google Street View, que registra imagens ao nível da rua de cidades de várias partes do mundo.

O caso do ReCAPTCHA é interessante por não se tratar apenas de um modo de produção de dados de treinamento mas também, simultaneamente, de um dispositivo que realiza uma distinção performativa entre as categorias de humano e de máquina. Por um lado, a dependência do aprendizado de máquina do olhar humano para produzir os dados utilizados em seu treinamento sugere um movimento com tendência de equiparação em que a visualidade constituída por meio da máquina tomaria o olhar humano como parâmetro. Por outro, ao também realizar uma distinção entre estas categorias, o ReCAPTCHA seguiria apostando na impossibilidade desta equivalência. Em todo caso, contudo, é preciso observar como o modo com que a distinção opera assume um parâmetro

performativo, em que as categorias de humano e máquina não são presumidas de antemão e sim constantemente *reconfiguradas* segundo aspectos dos comportamentos dos agentes, que são mutuamente constituídos.

Este argumento é elaborado pela antropóloga Lucy Suchman (2007) em sua discussão do desenvolvimento da IA e de suas figurações antropomórficas. Sua elaboração articula aspectos do ciberfeminismo de Donna Haraway e, também, a teoria performativa de gênero de Judith Butler, para compreender que humano e máquina não seriam categorias *essenciais* ou *naturais*, mas sim em constante mutação, sendo reiteradamente materializadas em práticas e comportamentos. De modo importante, enquanto categorias tipicamente consideradas em oposição, Suchman sugere a necessidade de considerá-las de maneira relacional, como mutuamente constituídas. Para esta investigação¹²¹, tal compreensão deve nos levar a reconhecer a impossibilidade da consideração de visualidades “humanas” ou “da máquina” em isolamento. Em larga medida, ao ensinarmos as máquinas a ver, estamos também nos ensinando a ver de determinada maneira. Isto não ocorre apenas quando estamos diante dos resultados da operação de um programa de aprendizado de máquina, como quando analisamos os dados categóricos produzidos por uma rede neural a partir de uma imagem. Isto também ocorre quando, diante do ReCAPTCHA, por exemplo, nos é solicitado olhar uma imagem como uma máquina o faria – demarcando a área ocupada por um hidrante ou um semáforo; identificando quais imagens contêm pontes ou quais contêm cachorros. Reconhecer e classificar é, enfim, a operação característica desta visualidade que nos interpela.

Aprendermos a ver com e como as máquinas não implica, evidentemente, que assumiríamos integralmente tal modo de ver. Mas, sim, que aprenderíamos a reconhecê-lo e, em alguma medida, a ver *com* ele. O reconhecimento se torna inclusive uma necessidade imperativa diante das imagens geradas por tais máquinas, pela técnica das GANs, e seu uso potencial em *fake news*, por exemplo: reconhecer na imagem aquilo que os modelos compreendem mal e que, por isso, também gerariam mal (cf. MCDONALD, 2018). Em outros momentos, este aprendizado maquínico se reflete em experiências estéticas realizadas diretamente com as máquinas, com o desenvolvimento de sistema de

121 Trata-se de argumento que também desenvolvi em minha dissertação de mestrado (MINTZ, 2015).

aprendizado de máquina para fins artísticos (veja nota 94); ou apenas pela sugestão do olhar da máquina, em uma reapropriação de seu modo de ver, como no trabalho *I'm Google* da artista estadunidense Dina Kelberman¹²².

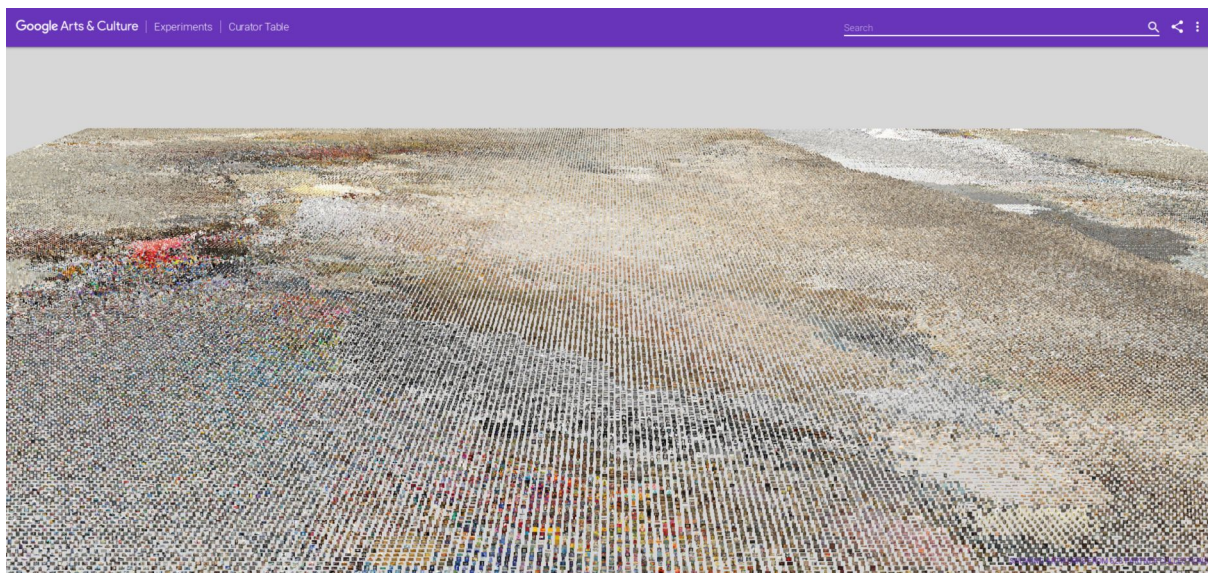


Figura 12: Impressão de tela do projeto Curator Table

Fonte: DIAGNE; BARRADEAU; DOURY, 2018.

Em ainda outra sugestão da noção de *aprendizes maquinaicos*, teríamos as implicações do uso destes sistemas em contextos de produção e organização do conhecimento. Indo além da constituição das redes neurais como “centros de cálculo”, como sugeri, teríamos nestes casos a integração das redes neurais a centros de cálculo mais abrangentes, como em laboratórios, museus, arquivos e bibliotecas. Neste sentido, aprenderíamos junto a máquinas modos de navegar por documentos, imagens, *inscrições*. Um exemplo deste tipo de operação pode ser encontrado no experimento *Curator Table* (DIAGNE; BARRADEAU; DOURY, 2018), desenvolvido por artistas residentes da Google a partir da base de dados de obras de arte que a empresa mantém em parceria com museus de várias partes do mundo (Figura 12). Trata-se de um dispositivo de navegação pelas milhares de obras que compõem esse acervo, que utiliza dos dados gerados por um modelo de reconhecimento de imagens para dispor as imagens em uma representação plana do espaço vetorial produzido pela rede neural. A planificação do espaço multidimensional utiliza o algoritmo t-SNE (MAATEN; HINTON, 2008), que realiza uma redução de dimensionalidade buscando preservar as relações de proximidades estabelecidas no espaço vetorial

122 Cf. <https://dinakelberman.tumblr.com/>

em sua projeção para espaços de duas ou três dimensões. Em um âmbito mais restrito, o projeto *Neural Neighbors*, do Digital Humanities Lab, da Universidade de Yale (Figura 13), utiliza metodologia similar para construir um dispositivo de navegação heurística pela coleção de fotografias Meserve-Kunhardt (YALE DIGITAL HUMANITIES LAB, 2017). A visualização de cada imagem vem acompanhada de seus ‘vizinhos neurais’ mais próximos, segundo 2.048 dimensões dos atributos visuais computados por meio de uma rede neural. Estes são exemplos, portanto, em que o aprendizado de máquina opera não como uma instância final de datificação das imagens que, a partir de então seriam tratadas exclusivamente como *dados*. De outro modo, as técnicas de aprendizado de máquina operam como uma interface ativa para orientar um observador em meio a milhares de imagens.

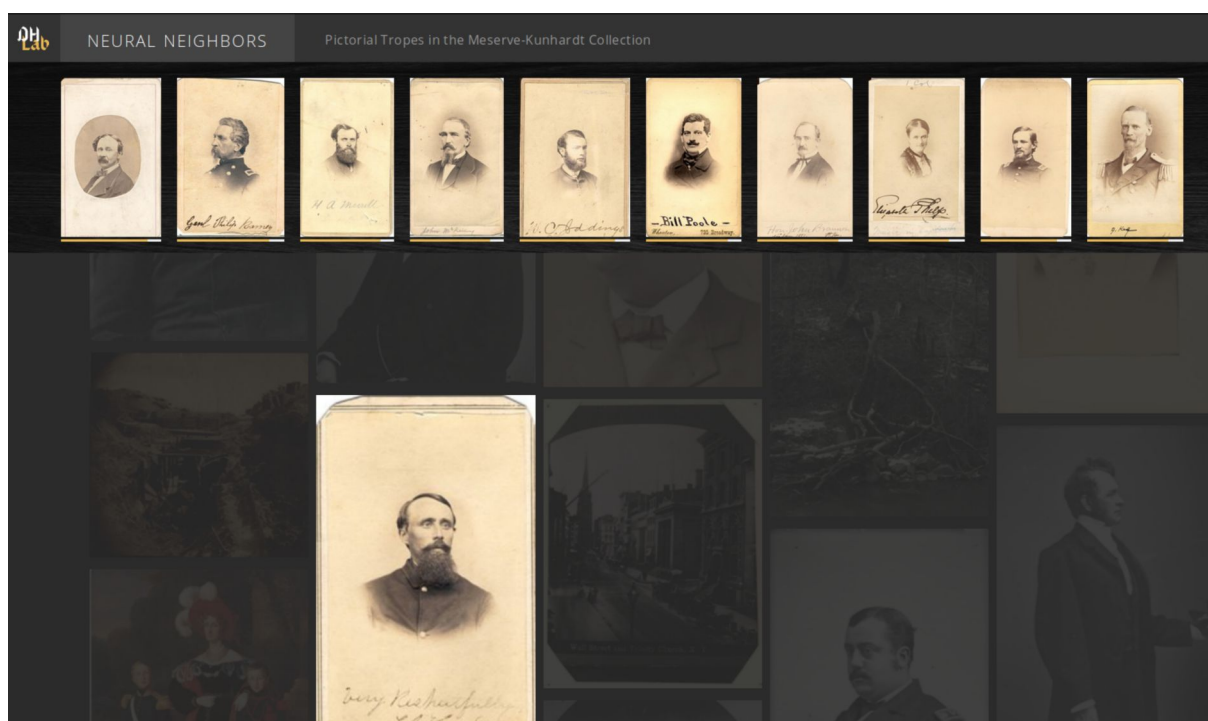


Figura 13: Impressão de tela do projeto *Neural Neighbors*.

Fonte: YALE DIGITAL HUMANITIES LAB, 2017.

A estratégia adotada no estudo de caso desta tese (veja [4 Estudo de caso: imagens de uma prisão](#)) baseia-se, em larga medida, nos procedimentos adotados nestes casos. Efetivamente, visualizações como essas têm se tornado parte de um modo cada vez mais consolidado de lidar com a operação de modelos de reconhecimento de imagens. Já na minha dissertação de mestrado (MINTZ, 2015), que propunha uma abordagem teórica

para a Visão Computacional pelo ponto de vista da imagem, expus na capa uma visualização bastante similar, produzida pelo então doutorando em Ciências da Computação, por Stanford, Andrej Karpathy (2012). Os procedimentos adotados por Karpathy são bastante similares aos que desenvolvo aqui e, também, aos adotados pelos casos citados. Os modelos de aprendizado de máquina são utilizados a fim de dispor as imagens em um espaço vetorial que, posteriormente, tem suas dimensões reduzidas por meio do algoritmo t-SNE. As imagens utilizadas por Karpathy eram, então, 50 mil figuras retiradas da base de dados Imagenet (DENG et al., 2009). Os procedimentos que adoto aqui, baseados no tutorial ML4A (REFSGAARD; TSENG; KOGAN, 2019), baseiam-se nas mesmas operações e permitem, como nestes outros casos, apreender um conjunto de imagens por relações de similaridade constituídas a partir da visualidade inscrita nos modelos de reconhecimento aplicados.

Como discuti na subseção anterior (veja 3.2.2 Vetorização), a distribuição dos dados analisados em um espaço vetorial constitui uma das operações fundamentais realizadas por sistemas de aprendizado de máquina. Porém, a derradeira operação, tendo em vista o uso típico destes sistemas para tarefas de classificação, seria a demarcação de regiões ou *locais* em meio a esse espaço, seguindo, como sugere Mackenzie (2017, p. 63), as *estrias* nos dados. Algumas visualizações realizadas por meio do algoritmo t-SNE, por outro lado, detêm-se em um estágio ainda anterior, apresentando uma representação plana do espaço vetorial antes de que se definam as regiões específicas às classes utilizadas no treinamento. Trata-se, portanto, de operação que interrompe, em certa medida, a operação típica realizada por esses modelos. Este gesto, quero sugerir, abriria uma margem para sua reapropriação crítica, uma brecha para lançar sobre a operação da máquina um olhar autorreflexivo.

Entre as técnicas pertencentes ao domínio do aprendizado de máquina, chama-se “aprendizado por transferência” (*transfer learning*) (REFSGAARD; TSENG; KOGAN, 2019) a apropriação de um modelo pré-treinado para a realização de uma tarefa específica não prevista por esse modelo. A apropriação é realizada por um novo processo de treinamento das camadas posteriores do modelo com base em uma nova base de treinamento. A vantagem desta técnica é que não seria preciso passar pelo complexo e custoso procedimento de construir um modelo “do zero”. Os recursos de processamento e a quantidade de imagens necessária para este novo treinamento seria bastante inferior. Por exemplo, seria possível tomar um modelo de classificação de imagens treinado a

partir dos dados do projeto *Imagenet* (DENG et al., 2009) para que ele seja capaz de identificar um objeto específico não previsto na taxonomia do modelo original. Digamos: para reconhecer fotografias do prédio da Faculdade de Filosofia e Ciências Humanas (Fafich) da UFMG. Pela técnica do aprendizado por transferência, utilizaríamos um segmento das camadas anteriores da rede neural, com seus pesos definidos pelo treinamento na base *Imagenet*, e o acoplaríamos a novas camadas, ainda sem pesos definidos, que seriam treinadas com imagens selecionadas especificamente para esta tarefa. Deste modo, são aproveitadas algumas das características “aprendidas” pelo modelo original na construção de um novo modelo, cujo objetivo será, grosso modo, desenhar outras linhas de demarcação no espaço vetorial anteriormente produzido, a fim de demarcar as porções pertencentes a essa classe específica – o prédio da Fafich.

Retomo essa descrição técnica para sugerir algo como uma metáfora, mas que talvez possa ser tomada de forma mais literal. As visualizações que indiquei acima, geradas a partir do algoritmo t-SNE, oferecem um ponto de vista para representações intermédias de um modelo de reconhecimento de imagens. Elas exibem, afinal, o espaço vetorial constituído pelo modelo, com as imagens distribuídas segundo suas posições relativas. Estas visualizações seriam, nesse sentido, provenientes de um segmento das camadas anteriores do modelo de reconhecimento, antes da definição de classes específicas para cada imagem. Trata-se, nesse sentido, de uma representação interna do modelo que se assemelha àquela que seria reapropriada na aplicação da técnica de aprendizado por transferência que descrevi no logo acima. A diferença deste caso, contudo, é que em vez de apresentar esta representação interna a um outro sistema de aprendizado de máquina, ela é apresentada a observadores humanos que podem então, eles mesmos, percorrer as *estrias dos dados*, representados no espaço vetorial. Quero sugerir este como mais um dos sentidos possíveis da noção de *aprendizes máqunicos*, compreendendo as múltiplas possibilidades de reconfiguração das posições e das definições de observadores humanos e não-humanos em um mesmo circuito. Voltarei a abordar este tipo de operação no estudo de caso apresentado adiante.

3.3 Visibilidade e infraestrutura

Mackenzie (2017, p. 3) propõe compreender o aprendizado de máquina na interseção de três grandes acúmulos. Primeiro, relativo a campos ou domínios de aplicação desta téc-

nica. Pois, efetivamente, não são apenas os dados que precisam ser constantemente deslocados e acumulados, mas também as técnicas algorítmicas do aprendizado de máquina constituem-se em um contexto de ampla mobilidade, sendo aplicada às mais diversas disciplinas científicas e, também, a inúmeros contextos cotidianos – do microdirecionamento de anúncios em plataformas digitais ao diagnóstico médico, passando pela vigilância distribuída e a elaboração de perfis de crédito, entre outros. Este intenso trânsito seria um fator importante de sua consolidação e desenvolvimento, conferindo-lhe uma relevância difusa e, com isto, um ecossistema de maturação e concretização técnica. Um segundo acúmulo seria relativo aos dados. Este aspecto seria mais evidente, diante da dependência destas técnicas de amplas bases de treinamento e, também, uma vez que seu desenvolvimento volta-se exatamente às demandas de processamento decorrentes do crescimento exponencial do chamado *big data*. Por fim, um terceiro acúmulo seria relativo às diferentes técnicas de aprendizado de máquina e aos aparatos – isto é, o *hardware* – demandados para seu funcionamento. O aprendizado de máquina cria uma alta demanda de capacidade de processamento e, como indiquei na seção anterior, estas demandas se voltam especialmente a GPUs. Boa parte do suprimento desta necessidade articula, hoje, amplas cadeias de “computação em nuvem”, com recursos de processamento vendidos como um serviço *online* por grandes corporações da área de tecnologia – notavelmente Amazon e Google.

Kate Crawford e Vladan Joler (2017) realizaram uma cartografia da ampla rede mobilizada por sistemas de IA baseados em aprendizado de máquina, percorrendo os meandros da ativação de um aparentemente banal assistente doméstico responsivo à voz – as “caixas de som inteligentes”, prototípicas de investimentos recentes na chamada Internet das Coisas. A partir do levantamento destes autores, poderíamos acrescentar ainda outros acúmulos que seriam conformadores do aprendizado de máquina, indo da infraestrutura de comunicação global em rede, às cadeias de extração mineral e de descarte que circundam a indústria tecnológica. Outro aspecto destacado por eles e que não chega a constar da descrição de Mackenzie é o trabalho humano de classificação prévia dos dados utilizados para o treinamento destes sistemas – aspecto que indiquei na seção anterior (veja 3.2.3 Aprendizes maquinaicos), embora sem uma ênfase na condição de *trabalho* destas relações. Estratégias utilizadas na codificação de dados de treinamento

para as máquinas indutivas envolvem desde plataformas de *crowdsourcing*, como a Mechanical Turk, de propriedade da Amazon; até o trabalho não remunerado de usuários, por meio de sua própria atividade em plataformas de mídia social, por exemplo.

Deve-se observar, portanto, que os sistemas de aprendizado de máquina são construídos sobre uma ampla camada infraestrutural, em um agenciamento que é tanto complexo quanto *caro*. Boa parte desta camada, inclusive, antecede ao aprendizado de máquina e empresas desenvolvedoras destes sistemas frequentemente contam com a vantagem competitiva de acesso a estas estruturas, ou ao capital para custeá-las. Pois o desenvolvimento contemporâneo desta tecnologia também constitui um motor de ampliação e complexificação infraestrutural, em especial pela integração destes sistemas à heterogênea ecologia das plataformas *online*.

Ao ressaltar a dinâmica de acúmulos que teriam conformado o aprendizado de máquina, Mackenzie (2017) – como também Cardon, Cointet e Mazières (2018) – visa, especificamente, ir contra a ideia de que estas técnicas representariam uma transformação radical em um plano epistemológico. Isto é, seu argumento arqueológico (no sentido foucaultiano do termo) reconstitui as relações desta técnica “nova” com práticas e tecnologias que lhe antecedem. Para o caso de Cardon, Cointet e Mazières (2018), o interesse são desenvolvimentos anteriores do paradigma conexcionista, que remontam à cibernética dos anos 1940 e 1950, bem como a controvérsia com o paradigma simbólico da IA. O ponto que quero salientar nesta seção, contudo, volta-se menos a estas raízes epistemológicas e pragmáticas do que aos acúmulos materiais e às assimetrias geradas pelo alto custo de arregimentação das amplas dependências dos sistemas de aprendizado de máquina. Esta questão, quero argumentar, tem fundamental importância para a discussão proposta para esta tese quanto às visualidades computacionais engendradas por estes sistemas.

O tema da infraestrutura adentra os estudos de mídia contemporâneos por diferentes caminhos. Estudos críticos da internet e certas perspectivas da arte mídia voltam-se ao tema da infraestrutura como forma de se contrapor às narrativas que postularam o caráter imaterial ou etéreo do *ciberespaço* ou da *nuvem* (cf. DRAGONA, 2016; VELDEN; KRUK; METAHAVEN, 2015). Nestes casos, a ideia de infraestrutura assume com frequência o sentido de *hardware*, em estudos interessados na geopolítica dos cabos submarinos ou dos satélites, por exemplo¹²³. Embora de forma não-exclusiva, outra vertente de consi-

¹²³ Estudo ainda não publicado de Judith Keilbach sobre a infraestrutura de transmissão “ao vivo” do julgamento de Adolf Eichmann descreveria outra vertente deste tipo de investigação.

deração sobre a infraestrutura relaciona-se a desenvolvimentos recentes dos chamados estudos de plataforma. Traçando uma perspectiva crítica à denominação autoindulgente de *plataforma*, sugerida pela própria indústria para o modelo de negócios assumido por *sites* da internet a partir da chamada Web 2.0 (O'REILLY, 2005), os estudos de plataforma visam descrever as mediações ativas das plataformas, contra a ideia de que seriam meros sustentáculos neutros para as ações de seus usuários (GILLESPIE, 2010). Fazendo um uso crítico da noção de plataforma, portanto, autores dedicados a esta perspectiva consideram dinâmicas de *plataformização* em que uma miríade de práticas sociais passariam a ser mediadas e, efetivamente, transformadas pelas plataformas *online* (cf. BUCHER; HELMOND, 2018; D'ANDRÉA, 2018; HELMOND, 2015; NIEBORG; POELL, 2018). O tema da infraestrutura adentra esta discussão em particular pelo que veio a ser descrito como uma *infraestruturalização* das plataformas (cf. DIJCK; POELL; WAAL, 2018; HELMOND; NIEBORG; VLIST, 2019; PLANTIN et al., 2016). Neste caso, a noção de infraestrutura não tem um sentido tão literalmente *material*, mas volta-se, em especial, à transição pela qual as plataformas *online* passariam prover serviços essenciais.

Segundo elaboram Plantin e colaboradores (2016), infraestruturas possuem uma série de características que as distinguiriam de plataformas. A partir do referencial teórico dos estudos de infraestrutura, um subdomínio dos STS, os autores (PLANTIN et al., 2016, p. 294) indicam entre as características centrais destas formações sociotécnicas: “ubiquidade, confiabilidade [e] invisibilidade”¹²⁴. A ubiquidade sugere o caráter efetivamente distribuído e fundante das infraestruturas enquanto condições de existência de um coletivo social. Também, a configuração das infraestruturas como agenciamentos sociotécnicos amplos que conjugam elementos heterogêneos e difusos. A confiabilidade seria, por sua vez, um atributo necessário para que infraestruturas sejam vistas como tais, uma vez que assumem papéis fundantes na organização coletiva em uma relação de dependência. Por fim, a invisibilidade sugere um processo próximo à noção de *obscurecimento*, decorrente do funcionamento eficaz das infraestruturas, que passam a ser largamente ignoradas – a não ser quando falham. Os autores sugerem um movimento de aproximação entre plataformas e infraestruturas – *plataformização das infraestruturas* e *infraestruturalização das plataformas* – à medida que as plataformas passam a ser provedoras de serviços essenciais. Isto, embora não estejam submetidas ao escrutínio público, como costuma ser o caso das infraestruturas tradicionais.

124 No original: “ubiquity, reliability, invisibility ...”. Tradução minha.

Exemplos voltados a domínios infraestruturais tradicionais são mais facilmente apreendidos, como a plataformização do transporte (Uber e afins); da habitação (AirBNB e afins); ou da educação (Coursera e afins). Haveria, contudo, outros movimentos de infraestruturalização menos evidentes, quando a plataforma constitui-se enquanto uma infraestrutura em um novo ramo de atividades. Helmond, Nieborg e van der Vlist (2019), por exemplo, sugerem que o desenvolvimento histórico do Facebook poderia ser compreendido como uma transição a uma “plataforma como infraestrutura”. Isto teria ocorrido à medida que, por diversas estratégias, ele teria conseguido se estabelecer como mediador fundamental de diferentes mercados, tais como o publicitário e o de vendas *online* por exemplo, além de se constituir como meio de identificação principal para milhões de usuários em diferentes serviços da Web que, a princípio, não seriam relacionados com o Facebook. Acerca deste último caso, trata-se do uso do Facebook como mediador do registro e do acesso (*log in*) dos usuários. Neste sentido, o Facebook teria assumido uma posição infraestrutural na internet, sobrepondo suas lógicas de governança interna, inclusive, às políticas públicas, em nível global.

A aproximação que quero propor entre o tema das infraestruturas e a discussão das visualidades computacionais e do aprendizado de máquina seria um destes casos a princípio menos evidentes. Quero sugerir que o reconhecimento de imagens operado por um pequeno conjunto de atores tenderia a assumir uma posição quase infraestrutural como forma de mediação algorítmica do visível. Esta tendência decorreria da relativa escassez e concentração dos modelos de reconhecimento de imagens disponíveis para pronta aplicação – um efeito das assimetrias infraestruturais sobre as quais o aprendizado de máquina é constituído. Para desenvolver um classificador de imagens equiparável em eficácia aos hoje disponíveis – porém “do zero” – a obtenção de dados de treinamento talvez constituiria um dos maiores custos iniciais. Seria preciso realizar a coleta, seleção e categorização (esta, manual) de centenas de milhares ou milhões de imagens. O processo de treinamento, por sua vez, pode demandar dias de processamento em um arranjo de GPUs com custo elevado, além de um alto consumo de eletricidade. A alternativa mais viável, portanto, seria apenas aplicar algum dos modelos disponíveis ou, em alguns casos, utilizar a técnica de “aprendizado por transferência” (*transfer learning*) para tornar o modelo mais especializado para um tipo de aplicação específica. Neste últi-

mo caso, contudo, as características inscritas nas camadas anteriores da rede neural permaneceriam atuando como parte de seu funcionamento inculindo ao modelo aspectos do treinamento geral.

Dentre as opções de modelos prontamente aplicáveis hoje disponíveis, as mais robustas talvez sejam as APIs comerciais, que oferecem o processamento sob demanda de imagens em diferentes modelos de aprendizado de máquina aplicados à Visão Computacional. Há APIs oferecidas por algumas das grandes corporações tecnológicas contemporâneas: Google, IBM, Amazon, Microsoft. Além de outras iniciativas relativamente menores, como a da empresa Clarifai. Os custos de utilização são relativamente baixos, em torno entre 1 e 2 dólares estadunidenses para cada mil imagens. Em todos os casos, tratam-se de modelos proprietários que não divulgam detalhes acerca da arquitetura das redes neurais; das bases de treinamento utilizadas; ou sobre as categorias atribuídas nos processos de classificação. No entanto, sua pronta disponibilidade e baixo custo tornam estas opções alternativas viáveis para o desenvolvedor de um aplicativo ou mantenedor de um serviço *online* para, por exemplo, moderar conteúdos postados pelos usuários ou organizar um arquivo fotográfico institucional, entre outras aplicações possíveis. Por tais caminhos, estes modelos tenderiam a difundir modos particulares de *ver* a diferentes contextos e a se tornarem “pontos de passagem obrigatórios”¹²⁵ para a mediação algorítmica do visível.

Em estudo de que participei, dividindo a coordenação com Tarcízio Silva (MINTZ; SILVA et al., 2019), comparamos os resultados obtidos por três destas APIs – da Google, Microsoft e IBM. A comparação foi realizada segundo o processamento de alguns conjuntos de imagens colhidos de dois bancos de imagens comerciais, Adobe Stock e Shutterstock. Em cada um destes *sites*, coletamos 2 mil imagens obtidas a partir da busca por adjetivos pátrios de quatro países: Brasil, Portugal, Áustria e Nigéria. Por meio de diferentes esforços metodológicos, investigamos algumas das diferenças entre os resultados obtidos e, também, o potencial de aplicação destas APIs como ferramentas analíticas para estudos voltados a representações nacionais nesses contextos. Um dos resultados encontrados, dos que mais importam a esta discussão, é a variação entre os

¹²⁵ Esta expressão (*obligatory point of passage*) remonta a textos fundadores da TAR e descreve um ator com tal força de translação que se se torna um mediador central da rede, um ponto de passagem obrigatório para se realizar determinada ação, por exemplo.

“espaços semânticos” das APIs, que indicam aspectos gerais da dinâmica de classificação das imagens. Observamos estes espaços por meio de visualizações de dados relacionais, em redes bimodais imagem-etiqueta¹²⁶ (Gráfico 1).

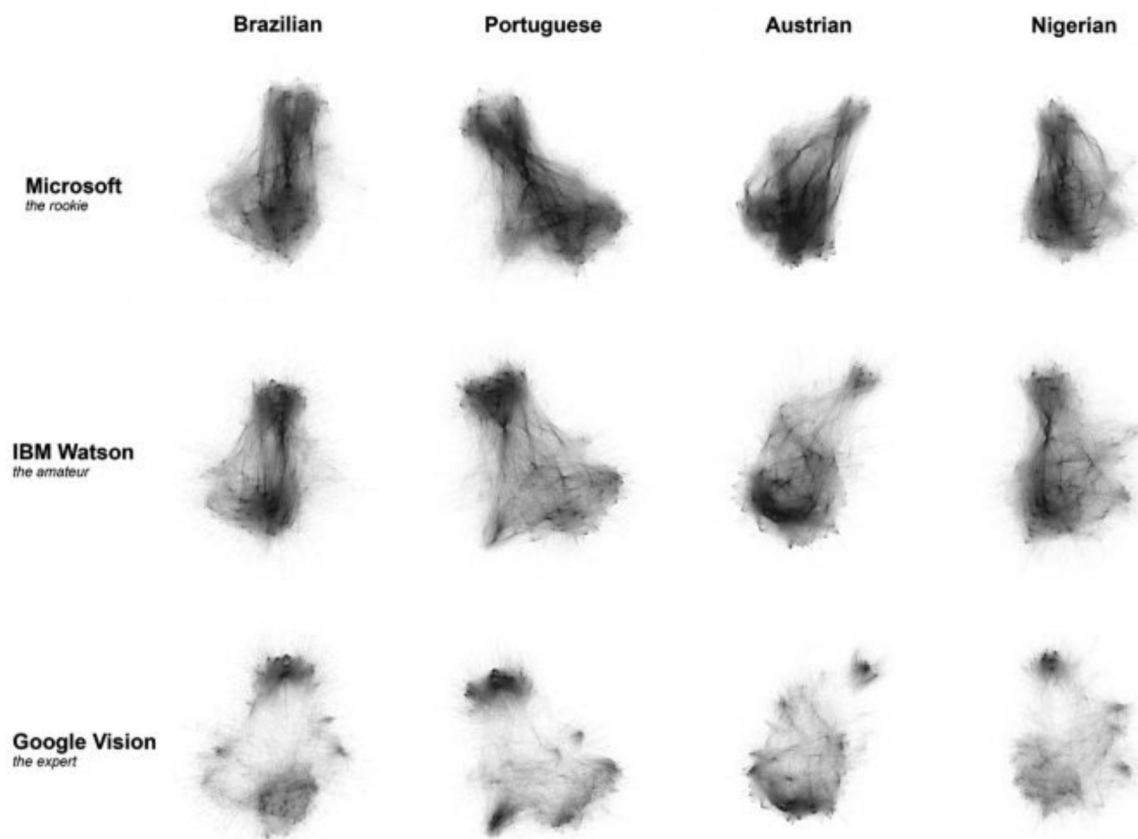


Gráfico 1: Visualização de espaços semânticos de três APIs de Visão Computacional. Cada visualização é uma rede imagem-etiqueta produzida a partir de imagens relativas à nacionalidade da coluna, com dados da API indicada na linha. Para cada nacionalidade, as redes das três APIs foram mescladas e espacializadas conjuntamente, a fim de comparar a distribuição das densidades das conexões pela rede.

Fonte: MINTZ; SILVA et al., 2019.

Busco simplificar as análises e conclusões desse estudo, apenas segundo o interesse específico de meu argumento, aqui. O aspecto que gostaria de salientar destas visualizações é como a distribuição de densidades das redes constituídas a partir de cada API varia consideravelmente quando comparadas em relação a um mesmo conjunto de imagens. Enquanto aquela elaborada a partir da API Cloud Vision da Google (GVAPI) apresenta densidades mais localizadas em agrupamentos específicos, as demais APIs são mais densamente conectadas, inclusive entre porções extremas da rede. Estas características sugerem que a GVAPI possuiria um maior grau de especialização, ou granularidade, das categorias. A baixa densidade da rede ao centro indica haver poucas

¹²⁶ Redes bimodais são aquelas em que os nós representam entidades de duas naturezas. Nesse caso, parte dos nós da rede representa imagens e parte representa as etiquetas categorias atribuídas às imagens.

categorias gerais o suficiente para conectar partes extremas – isto é, imagens consideradas muito díspares. Por outro lado, ela apresenta uma densidade maior no interior dos agrupamentos, sugerindo a existência de categorias mais específicas. As redes construídas a partir da API da Microsoft são bem contrastantes, apresentando uma alta densidade das conexões, mesmo entre porções extremas. As visualizações produzidas a partir de dados da API da IBM apresentam um aspecto intermediário, mas ainda densamente conectada ao centro.

Em uma outra perspectiva metodológica, também realizamos descrições densas a partir de casos específicos, em que foi possível observar como a maior granularidade da API da Google reflete-se em uma maior sensibilidade à representação de objetos (tipos de comida e instrumentos musicais, por exemplo) culturalmente específicos. Em contraste, a API da Microsoft apresentou categorias não apenas mais genéricas como também, em alguns casos, tendenciosos para representações da cultura estadunidense. Um caso específico em que isto foi observado foi nas categorias indicadas para o doce português pastel de nata, que a API da Microsoft classificou como *donut*. Porém, mesmo para a API da Google, as descrições densas indicaram situações em que as categorias reproduziam estereótipos racistas. Especificamente, imagens de mulheres negras com cabelos crespos, provenientes dos dados do Brasil e da Nigéria, foram lidas pela GVAPI, reiteradamente, com a categoria “peruca”. Este conjunto de observações indica portanto, como APIs comerciais constituem visualidades específicas em decorrência, inclusive, de assimetrias geopolíticas, raciais e de gênero das instâncias infraestruturais sobre as quais tais sistemas são construídos¹²⁷.

Em outra vertente, opções de modelos gratuitos e provenientes de um contexto de pesquisa acadêmica, como as disponíveis na biblioteca de programação em *Python Keras* (CHOLLET et al., 2018), embora mais numerosas, são todas treinadas a partir de uma mesma base de treinamento, a *Imagenet* (DENG et al., 2009). Trata-se de projeto mantido por grupo da Universidade de Stanford e composta por milhões de imagens de treinamento classificadas segundo mil categorias. Seu uso difundido entre modelos de reconhecimento gratuito deve-se, em larga medida, ao concurso ILSVRC, realizado pela comunidade científica de Visão Computacional e que toma a base *Imagenet* como parâmetro de comparação da eficácia de diferentes modelos de reconhecimento. Apesar de

¹²⁷ Agradeço aos coautores do estudo, em especial a Helen Takamitsu e Taís Oliveira, que realizaram as análises indicadas.

constituírem recursos valiosos, a base *Imagenet* e, conseqüentemente, os modelos nela treinados possuem diversos problemas tanto relativos à baixa diversidade racial e cultural – similarmente aos casos descritos acima – quanto à inclusão de subcategorias claramente ofensivas como parte de seus dados (cf. CRAWFORD; PAGLEN, 2019). O impacto destes problemas pode ser maior ou menor a depender do contexto de aplicação dos modelos. O problema das subcategorias ofensivas abordado por Crawford e Paglen, particularmente, não incide diretamente sobre os modelos indicados¹²⁸. Em todo caso, tais problemas são indicativos da já esperada parcialidade dos modelos.

Além da base *Imagenet*, cujo foco é o reconhecimento de imagens, uma abordagem compreensiva da questão também deve observar bases voltadas a outras tarefas de Visão Computacional por aprendizado de máquina, como a análise de rostos, que deriva em programas de reconhecimento facial. Neste domínio, uma referência fundamental é o estudo de Joy Buolamwini e Timnit Gebru (2018) realizou uma investigação dos vieses raciais e de gênero de duas bases de referência para a tarefa de análise facial – IJB-A e Adience. As pesquisadoras analisaram a composição fenotípica destas bases de imagens considerando um sistema de classificação dermatológica de tons de pele e observaram enormes disparidades na distribuição dessas características para as bases analisadas. Para ambos os casos, cerca de 80% das imagens eram de pessoas de pele clara. As autoras também comparam a eficácia de sistemas de classificação de gênero e indicam como a disparidade observada nas bases de referência repercute nas margens de erro encontradas em tais sistemas. Mulheres de pele mais escura obtêm índices de erro muito mais altos do que homens de pele mais clara – 34,7% contra 0,8%¹²⁹.

Segundo o que discuti acerca da noção de *visualidade*, a partir de Hal Foster (1999), todo “modo de ver” seria circunscrito a uma historicidade e à particularidade de uma “visão de mundo”. Seria bastante ingênuo esperar, portanto, que um programa de computador pudesse materializar a hipótese de uma *visão* universal, fora de uma realidade histórica e social. As limitações das APIs de Visão Computacional, da base *Imagenet* e de seus modelos derivados, ou de bases e programas de análise facial seriam, portanto, bastante previsíveis. Para além do caráter marcadamente injusto destas assimetrias, contudo,

128 Especificamente, um dos trabalhos realizados por Crawford e Paglen (2019) a partir da base *Imagenet* explora subcategorias da classe “pessoa” (*person*). Estas não fazem parte das mil categorias utilizadas por padrão para os principais modelos derivados, uma vez que elas não integram o concurso ILSVRC. Exemplos são os modelos disponíveis para pronta implementação na biblioteca Keras (CHOLLET et al., 2018). Tais subcategorias são, ainda assim, disponibilizadas pela base *Imagenet* como possíveis recursos de treinamento.

129 Estudo posterior de Buolamwini e Inioluwa Deborah Raji (2019) identificou que o estudo de Buolamwini e Gebru (2018) resultou em atualizações mudanças significativas de alguns dos casos analisados.

outro ponto que gostaria de indicar é, além da pouca diversidade interna aos modelos, a pouca diversidade *entre* os modelos, que é estrangida pelas demandas infraestruturais do desenvolvimento de modelos alternativos que constituam outras visadas sobre as imagens. Para além da classificação, do reconhecimento de caracteres, da análise facial, ou da detecção de conteúdo adulto – algumas das funções mais proeminentes de APIs comerciais – haveria ainda muitas outras perspectivas possíveis para uma análise computacional das imagens. Para estudos de comunicação e mídia, por exemplo, outras formas de codificação podem ser relevantes, como uma que se ocupe de certas estratégias formais persuasivas, ou de tipos de enquadramento das imagens¹³⁰. Esta limitação leva, evidentemente, a uma baixa diversidade das *visualidades computacionais* constituídas por tais sistemas. Estas tendem a ser, pelos motivos apresentados, bastante homogêneas. À medida que constituem as poucas opções disponíveis para o desenvolvimento de aplicações derivadas ou para esforços de pesquisa, esta visualidade tende, por sua vez, a se replicar e disseminar em contextos diversos.

À pervasividade de modelos como estes soma-se a tendência de sua essencialização sob uma suposta *objetividade* da máquina e, em especial, do aprendizado de máquinas – essa técnica “intermediária” por meio da qual o próprio mundo, supõe-se, poderia declinar-se a suas próprias categorias. Quero argumentar, portanto, que tal contexto coloca um problema crítico para o estudo das visualidades contemporâneas. Para Hal Foster (1999, p. ix), a tarefa dos estudos de visualidades seria fazer um movimento contrário a este que observo com relação ao aprendizado de máquina e que compreendo, nos termos do autor, como a constituição contemporânea de um regime escópico. Em citação que trago mais acima neste capítulo (veja p. 111), Foster (1999, p. ix) argumenta em favor da necessidade de “deslocar as superposições” construídas pelos regimes escópicos, a fim de desfazer seu movimento de naturalização de muitas visualidades em torno de uma “visão essencial”. A caracterização desta suposta “visão essencial” deve observar, como propõe Tarcízio Silva (2019), seu claro viés racial que se manifesta pela reprodução maquínica do ponto de vista da branquitude. Esta seria, conforme elabora Silva, uma instância contemporânea do racismo *online*, que vai além de “casos aberracionais de injúrias isoladas” (SILVA, 2019, p. 3) e deve ser observado em perspectiva sistêmica. De modo fundamental, seguindo Silva, faz-se importante considerar que: “Comumente aplicativos que buscam alcance global projetam consumidores apenas como brancos”

¹³⁰ Joo et al. (2014) elaboraram abordagens neste sentido. Algumas destas ideias vêm de uma conversa com Bernhard Rieder, a quem agradeço pela interlocução.

(SILVA, 2019, p. 10). Seria também necessário observar como, por um lado, limitações infraestruturais dificultam a diversificação deste olhar na forma da pluralização dos modelos de visão para diferentes públicos e culturas visuais. Por outro, uma abordagem criticamente informada destas tecnologias poderia assumir a tarefa de deslocar o campo da *visão computacional* ao domínio das *visualidades*. Por isto dizer de *visualidades computacionais*. Este é, me parece, um ponto crítico de uma investigação voltada às imagens contemporâneas.

Um modo de conduzir um tal esforço de aplicação crítica destas tecnologias é sugerido pela proposta dos Métodos Digitais, elaborada inicialmente por Richard Rogers (2013). Como indiquei em outro momento, ela se baseia em um gesto de *reapropriação* crítica de “métodos nativos do digital”, assumindo um alto grau de reflexividade acerca de suas implicações epistêmicas. De modo similar, o que proponho aqui é uma reapropriação crítica do aprendizado de máquina para o estudo de imagens. Trata-se de um modo de voltar nosso olhar simultaneamente às técnicas de análise e àquilo que elas nos permitem conhecer do mundo. A mediação algorítmica do visível imprime, segundo quero argumentar, uma dinâmica particular à circulação das imagens. Por isso, sua aplicação metodológica permite compreender aspectos emergentes de grandes conjuntos de imagens em modos que se aproximam do modo de operação destas mediações. Esta tentativa de investigar “objetos nativos do digital” *in situ*, por meio das mesmas mediações com que eles se constituem no âmbito da Web e das plataformas, é o principal aspecto que busco tomar dos Métodos Digitais para esta tese.

Em sua crítica à adoção de métodos quantitativos nas ciências humanas e sociais, Alexander Galloway (2014) indica que, para boa parte dos casos, trata-se apenas de tentativas fracassadas de pesquisadores de se equipararem às grandes corporações de tecnologia. Estas, ele sugere, possuiriam amplo domínio sobre os meios de produção de um conhecimento que se ancora em uma perspectiva quantitativa. Além disso, ele questiona as implicações políticas dos esforços que apenas buscassem se equiparar a tais iniciativas, indagando-se se as ferramentas produzidas por tais corporações seriam recursos epistemologicamente válidos ou se, de outro modo, apenas reproduziriam os interesses daquelas corporações. Uma das alternativas, sugere Galloway (2014, p. 128), seria desenhada pela proposta dos Métodos Digitais, entre outras que, segundo elabora: “apresentam o tipo de criatividade e cuidado necessário para compreender e responder à

crecente industrialização da mente e do corpo”¹³¹. Contra a suposição de que a tarefa das humanidades, pela via das humanidades digitais, seria “desafiar os mineradores de dados em seu próprio campo”, Galloway argumenta que as humanidades deveriam recorrer à sua “força fraca”, termo que toma do filósofo francês François Laruelle. Escreve: “O ponto é sair do jogo totalmente e continuar a perseguir aquelas mesmas questões com as quais a tecnociência sempre lidou mal, deslumbrada como é com mandamentos ideológicos e industriais específicos”¹³² (GALLOWAY, 2014, p. 128). Este “sair do jogo”, contudo, mesmo para Galloway, não significa abandonar a questão da técnica ou o fazer técnico como parte dessa reflexão. A questão é como articular a reflexão teórica, em nível estético, histórico, político e cultural, com as práticas e dispositivos tecnológicos que habitam nosso tempo.

Um par de textos de Bernhard Rieder e Theo Röhle (2012, 2017) apontam para alguns caminhos do enfrentamento deste desafio. Interessados em aprofundar a discussão das chamadas humanidades digitais, sob a perspectiva dos Métodos Digitais, os autores elencam um conjunto de desafios metodológicos colocados para as humanidades se elas querem se engajar criticamente com os métodos computacionais e suas implicações epistêmicas. Em linhas gerais, a discussão elaborada neste capítulo, embora não tomasse os autores como referência direta, acompanha as questões que eles colocam sobre: a suposição de objetividade das máquinas e dos dados; o poder retórico das visualizações de dados; a tendência de obscurecimento (ou “encaixapretamento”) das ferramentas e processos; os desafios transdisciplinares colocados ao se mobilizar conhecimento técnico da computação junto a questões provenientes das ciências humanas e sociais; e a busca problemática por um universalismo que estaria subtendida em muitos dos desenvolvimentos quantitativos e computacionais (RIEDER; RÖHLE, 2012).

Em um segundo momento, Rieder e Röhle (2017) buscam passar, como descrevem, dos desafios à “formação” ou à “construção”¹³³ de uma abordagem, isto é, constituir um campo de reflexão coletiva em constante revisão. Para isto, eles sugerem a necessidade de que os Métodos Digitais encarem as formas de conhecimento encapsuladas pelos métodos computacionais que empregam, tais como a estatística e a análise de redes. Posto

131 No original: “*exhibit the kind of creativity and care necessary for understanding and responding to the growing industrialization of mind and body*”. Tradução minha.

132 No original: “*The point is to withdraw from the game altogether and continue to pursue the very questions that technoscience has always bungled, beholden as it is to specific ideological and industrial mandates*”. Tradução minha.

133 Mesmo no texto em inglês, os autores utilizam uma palavra do alemão *Bildung*, traduzida, aqui, livremente. Eles tomam o termo de David Berry.

simplesmente, o argumento dos autores é quanto à necessidade de “encontrar a tecnologia como tecnologia, ao menos por um momento” (RIEDER; RÖHLE, 2017, p. 122). Segundo escrevem:

Embora nossa evidência seja apenas anedotal, nós notamos que em muito das humanidades há um desejo de explicar tecnologia o mais rápido possível através de outra coisa. Racionalidade instrumental, utopias cibernéticas, neoliberalismo, fantasias de controle perfeito, positivismo, maximização de receita, e assim por diante. Estas considerações podem ser iluminadoras ao final e significativas em um nível bem amplo de análise. Mas se quisermos estar à altura dos desafios dos métodos computacionais, temos que encontrar a tecnologia como tecnologia ao menos por um momento. [...] O problema do obscurecimento [blackboxing] não começa com a opacidade do código computacional, mas com o desejo de banir a tecnologia do ‘mundo da significação’¹³⁴ (RIEDER; RÖHLE, 2017, p. 122).

Os autores argumentam, portanto, no sentido de uma desessencialização da discussão teórica sobre a técnica por meio de um “corpo a corpo” com estas mediações e com as formações de conhecimento nelas inscritas. Portanto, as prováveis elucubrações sobre as técnicas em um nível *macro*, nas formas enumeradas pelos autores, apenas teriam consistência se articuladas por descrições que se atêm aos modos de operação da técnica e sua fundamentação.

Quero compreender sob esta chave a prática de *reapropriação* que é tão central à proposta dos Métodos Digitais. Reapropriar criticamente os modelos de reconhecimento de imagens seria um modo de investigá-los sob uma cláusula de não-competição, para encararmos a provocação de Galloway. Despossuídos que somos da infraestrutura necessária para desenvolver alternativas, nos restaria a possibilidade de aplicar tais modelos porém sem tomar seus resultados como válidos em si mesmos e, sim, como expressões políticas e culturais de nosso tempo. Contudo, para fazê-lo, não há atalhos simples nem suposição de uma sobredeterminação da técnica por fatores externos, nem pela suposição de uma sobredeterminação do mundo pela técnica. Seria preciso experimentar a técnica e as ontologias que produzem em abordagens situadas. “Encarar a tecnologia enquanto tecnologia”, como sugerem Rieder e Röhle (2017) significa ir ao nível de sua constituição e de suas formas de conhecimento. Este é o olhar que busco constituir, aqui, sobre o aprendizado de máquina enquanto um motor contemporâneo das articulações entre o visual e o computacional e da consolidação de visualidades computacionais particulares. Enquanto tais, estes mecanismos podem servir como elementos constituintes de um dispositivo de orientação ou de navegação pelas imagens contemporâneas. Este é o esforço que busco realizar na parte seguinte desta tese.

4 Estudo de caso: imagens de uma prisão

Neste capítulo, a discussão se concretiza em um estudo de caso, com o objetivo de tensionar as elaborações dos capítulos precedentes. O caso em questão é o evento da prisão do ex-presidente Luís Inácio Lula da Silva, ocorrido em abril de 2018, com forte repercussão na mídia e em plataformas *online*. Porém, o objetivo deste estudo não é, nesse momento, elaborar asserções analíticas sobre o caso em si. Principalmente, este estudo busca explorar as implicações metodológicas da discussão empreendida nesta tese, verificando o rendimento da reapropriação de técnicas de aprendizado de máquina para o estudo de imagens e, também, da hipótese conceitual da imagem-rede. Este esforço culmina, ao final, na proposição de um dispositivo navegação pelas imagens que busca proporcionar um percurso analítico que reconheça as diferentes visualidades mobilizadas e, também, a multiplicidade ontológica das imagens. Chamo a este dispositivo analítico de *Atlas para imagens-redes* (veja 4.3 Compondo imagens-redes) e ele é anexado a esta tese na forma de um documento de hipertexto navegável (veja instruções de acesso e navegação no Apêndice B.). As imagens do caso indicadas neste capítulo vêm acompanhadas de coordenadas para sua localização em duas visualizações (mapas) que constituem o *Atlas*.

Ao longo deste capítulo, descrevo as etapas adotadas no estudo de caso, em uma série de operações de coleta e processamento de dados que, na linha do que desenvolvi até agora, devem ser compreendidas como práticas por meio das quais as imagens do acontecimento em questão puderam *existir* e se fazer *visíveis*, efetivamente, no âmbito da análise. Busco evidenciar as *inscrições* mobilizadas em cada etapa bem como as *mediações técnicas* que as conformam e transformam ao longo do processo, elevando, na maior parte das vezes, as incertezas do método. Considerando-as parte integrante da discussão metodológica, volto-me frequentemente tanto aos potenciais analíticos das técnicas empregadas quanto às falhas e sombras que projetam sobre os objetos. Tanto àquilo que positivamente dão a conhecer do caso, quanto às negatividades que apenas podemos, quando muito, intuir. Gostaria, assim, de dar concretude às discussões empreendidas nos capítulos precedentes, considerando como elas podem sustentar uma perspectiva metodológica de pesquisa e análise das imagens em plataformas *online*.

O caso abordado neste capítulo tem como período mais intenso três dias de abril de 2018, logo após a expedição do mandado de prisão contra Lula, ex-presidente do Brasil por dois mandatos (2003-2008 e 2009-2012). A ordem foi assinada pelo então juiz federal Sérgio Moro, da 13ª Vara Criminal Federal de Curitiba, às 18 horas do dia 5 de abril, em cumprimento da pena de doze anos e um mês de prisão, decorrente da condenação do ex-presidente em processo por corrupção e lavagem de dinheiro, no que ficou conhecido como o caso do triplex do Guarujá, incluído como parte da força-tarefa da Lava Jato. Na noite do dia anterior, 4 de abril, o Supremo Tribunal Federal (STF) havia indeferido o pedido de *habeas corpus* preventivo impetrado pela defesa de Lula, abrindo caminho para sua prisão. A decisão de Moro oferecia pouco menos de 24 horas para o ex-presidente se entregar voluntariamente à Polícia Federal (PF). Imediatamente, Lula deslocou-se de São Paulo, capital, à sede do Sindicato dos Metalúrgicos do ABC (SMABC), em São Bernardo do Campo, onde, havia mais de 40 anos, iniciara sua carreira política como presidente sindical e liderança das grandes greves de 1980 e 1981. Por conta de seu papel nestas greves, Lula havia sido preso uma primeira vez, em 1980, pelo Departamento de Ordem Política e Social (DOPS), órgão de repressão da ditadura militar. Ele permaneceu confinado por 31 dias, enquadrado na Lei de Segurança Nacional, instituída pelo regime (SCHWARCZ; STARLING, 2015). Pelo prazo concedido no mandado de prisão, em 2018, Lula teria até as 17 horas de sexta-feira, dia 6 de abril, para se entregar. Contudo, ele apenas iria fazê-lo mais de 24 horas depois, no dia 7 de abril. Nas pouco mais de 48 horas que se estenderam entre a expedição do mandado e a realização da prisão, a vigília realizada por Lula e seus apoiadores no SMAB tornou-se um evento midiático, atraindo atenção constante da imprensa e do público – com forte repercussão nas plataformas *online*.

Antes mesmo da expedição do mandado de prisão, havia grande expectativa sobre o desenrolar destes acontecimentos que os tornavam históricos mesmo antes de acontecerem. Esta previsibilidade se devia a uma conjunção de fatores. Embora não fosse a primeira vez que se ordenava a prisão de um ex-presidente no Brasil¹³⁴, trata-se de caso sempre grave para a vida política do país. Em particular, como no caso, por se tratar de político que ainda gozava de enorme popularidade, apesar das denúncias e condenações que pesavam contra si e outros membros do Partido dos Trabalhadores (PT), do

134 Antes de Lula, outros quatro presidentes também foram presos sendo que um deles, Washington Luís (1926-1930), foi deposto e detido durante o exercício do cargo (Cf. <https://web.archive.org/web/20190215011200/https://acervo.oglobo.globo.com/em-destaque/ex-presidentes-presos-de-hermes-da-fonseca-washington-luis-jk-lula-21582019>).

qual foi um dos fundadores. O acontecimento também era a culminância de um período de instabilidade política que se estendia desde pelo menos 2013, com os grandes protestos que tomaram as ruas do país; que atravessou a conturbada eleição de 2014, que reconduziu ao cargo a então presidenta Dilma Rousseff, sucessora de Lula, com uma votação pouco superior ao segundo colocado; e também o controverso processo de *impeachment* de Rousseff em 2016, cuja legitimidade é fortemente contestada por críticos que qualificam o processo como golpe parlamentar¹³⁵. A esta complexa conjuntura, somava-se o fato de que a legitimidade da investigação e do julgamento eram fortemente contestadas por Lula e por observadores internacionais, que viam no processo um caso de perseguição política¹³⁶ que visava tornar o ex-presidente inelegível e, com isto, impedir que pudesse se candidatar à presidência nas eleições de 2018¹³⁷. Todos estes fatores tornavam o acontecimento, além de histórico, controverso.

Não é meu objetivo, aqui, explorar ou ponderar argumentos contrários ou favoráveis à condenação e à prisão, pois nem teria competência para tanto. Menciono alguns aspectos desta disputa por se tratarem de elementos contextuais que fundamentam a relevância do caso escolhido e, em particular, a relevância de dedicar a ele um esforço de análise focado nas suas imagens. Sua escolha deve-se a alguns fatores circunstanciais mas a, também, algumas de suas características específicas, que tornam o caso bastante adequado a uma investigação voltada às imagens.

Os fatores circunstanciais estão relacionados à disponibilidade dos dados a esta investigação. Como parte dos meus estudos no doutorado, desde meados de 2017 vinha realizando coletas experimentais de dados do Twitter como forma de desenvolver habilidades de trabalho com a ferramenta de coleta empregada, o DMI-TCAT (RIEDER; BORRA, 2014), e também de manutenção de um servidor de coleta. Tais exercícios volta-

135 Em entrevista próxima ao momento de escrita da tese, o processo de *impeachment* foi criticado até mesmo por Aloysio Nunes, ex-senador pelo Partido da Social Democracia Brasileira (PSDB). Durante o processo, opositor de Rousseff, Nunes foi um forte apoiador da remoção da presidenta (Cf. <https://web.archive.org/web/20190927215228/https://www1.folha.uol.com.br/poder/2019/09/lava-jato-manipulou-impeachment-de-dilma-diz-aloysio-nunes-do-psdb.shtml>)

136 No momento de escrita da tese, a questão da validade do julgamento ainda é alvo de intenso debate, em particular após junho de 2019, quando teve início de uma série de reportagens, denominada *Vaza Jato*, capitaneada pelo site jornalístico *The Intercept Brasil*, com o qual colaboram outros grandes veículos nacionais, como o jornal *Folha de São Paulo* e a revista *Veja*. As reportagens baseiam-se no vazamento de mensagens trocadas no aplicativo *Telegram* entre procuradores federais envolvidos no caso e também o então juiz Sérgio Moro. Entre outros pontos, as reportagens indicaram, nas mensagens, indícios de colaboração do juiz com o MPF e de quebra das normas do processo legal. Procuradores e juiz declaram não reconhecer a veracidade das mensagens.

137 De fato, mesmo após a prisão, a candidatura de Lula à presidência foi registrada pelo PT e sustentada até o momento em que o Tribunal Superior Eleitoral indeferiu o pedido de registro, em 31 de agosto de 2019 (Cf. <https://web.archive.org/web/20190826020711/http://www.tse.jus.br/imprensa/noticias-tse/2018/Setembro/tse-indeferiu-pedido-de-registro-de-candidatura-de-lula-a-presidencia-da-republica>).

ram-se a diferentes acontecimentos que propiciaram amplas discussões na plataforma no período, muitos dos quais vinculados ao momento político vivido pelo país – que constituíram alguns dos temas mais acalorados das discussões nas plataformas. Ao longo destes anos realizei coletas que combinavam a vontade de exercício técnico e metodológico e, também, um interesse de registro histórico e memorialístico – ou seja, nem sempre com um objetivo claro e predefinido de análise. A coleta de dados sobre este caso seguiu o mesmo caminho, começando como um experimento de coleta e culminando, já no âmbito desta tese, em um estudo de caso.

Quanto às características que tornam esse acontecimento particularmente adequado a uma investigação sobre as imagens e, também, realizada a partir de dados coletados em plataformas *online*, um primeiro aspecto diz respeito a sua previsibilidade. Embora a prisão viesse a ocorrer apenas no mês de abril de 2018, vinha coletando tuítes com termos relacionados ao caso desde 18 de janeiro, quando o recurso da defesa de Lula ao tribunal de segunda instância (TRF-4) fora julgado. Esta previsibilidade é importante para coletas realizadas em plataformas *online* porque, tipicamente, as modalidades de acesso aos dados limitam o acesso retrospectivo a postagens. No caso do Twitter, em particular, a melhor porta de entrada aos dados da plataforma, a API Streaming (de que depende a ferramenta DMI-TCAT) permite apenas a coleta em “tempo real”, o que demanda que os termos utilizados na busca do tuítes sejam, de algum modo, conhecidos de antemão. A previsibilidade é fundamental, portanto, para a consistência e amplitude da coleta.

Outro aspecto, decorrente da previsibilidade, diz respeito à performance dos atores na constituição do acontecimento, já que eles têm a condição de se antecipar e agir, em alguma medida, estrategicamente. Para o caso da prisão de Lula, em particular, o acontecimento de sua prisão já vinha sendo *premediado* (GRUSIN, 2010) há meses, pelo menos. Quando da expedição do mandado de prisão, em 5 de abril, embora os acontecimentos subsequentes ainda tenham sido marcados por um alto nível de improviso, já havia a clareza de uma disputa pela narrativa. Uma disputa que não visava apenas efeitos mais imediatos – como de fato visava, no entorno das eleições nacionais daquele ano – mas também efeitos históricos. Tratava-se desde o início de um acontecimento chave para uma temporalidade mais profunda da história nacional e, por esse motivo, para além do desenrolar do acontecimento em si, as imagens que dele se produziriam teriam um enorme valor, no longo prazo. Ou seja, como aspecto de adequação do caso a um es-

tudo voltado especificamente às imagens, como argumentarei mais adiante, trata-se de um acontecimento que poderíamos compreender de maneira próxima ao que Latour (2008; 2002) – em um raro texto dedicado às imagens, não apenas da ciência – descreveu como um *iconoclash*.

Indiquei brevemente esta noção na contextualização de algumas ideias de Peter Galison (veja 3.1 O visual e o computacional). Um *iconoclash*, segundo elabora Latour, seria uma situação de embate em que a dicotomia entre gestos de *iconoclastia* e *iconofilia*¹³⁸ seria diluída. Sua tese é a de que, frequentemente, gestos de destruição de imagens são acompanhados por uma substituição, isto é, pela constituição de outra imagem de “devoção”. Um olhar inicial às imagens deste acontecimento permite que observemos uma tal situação na medida em que boa parte do conflito manifesto no entorno da prisão do Lula pode ser descrito segundo as imagens estrategicamente produzidas ou interdidas, por diferentes partes envolvidas no caso. Embora seja demasiado simplista traçar uma dicotomia ou polarização simples, talvez seja possível arriscar um esquema com valor didático. De um lado ‘lulista’, teríamos esforços de recuperação ou reforço da imagem do Lula sindicalista e líder popular, inclusive com a justaposição entre o Lula daquele presente com imagens das greves que liderara em 1980. Haveria, desse modo, um culto à imagem do Lula popular e ‘revolucionário’. De um lado ‘lavajatista’, por sua vez, teríamos esforços de desconstrução desta imagem do ex-presidente, que era também acompanhada da produção de outros ícones: imagens premediadas da prisão, como o boneco do Lula presidiário (chamado de *pixuleco*); imagens de juízes e procuradores responsáveis por seu caso; etc.

Compreender o acontecimento da prisão sob a chave do *iconoclash* proporciona, portanto, um movimento analítico acerca do papel das imagens como atores políticos do acontecimento que devem ser tomados, fundamentalmente, pelas tensões que introduzem e pelas relações que estabelecem entre si. Segundo elabora Latour (2008; 2002), a noção de *iconoclash* visa suspender o gesto *iconoclasta* como um recurso da crítica, compreendida como o esforço de expor aquilo que constitui as imagens contra as suposições de que seriam feitas pela ‘mão divina’. Contra os ícones religiosos, expor a fabricação humana de suas imagens. Escreve o autor (LATOURE, 2008, p. 116): “a mente crítica é a que mostra as *mãos* dos humanos agindo em todos os lugares, a fim de trucidar a santi-

¹³⁸ A *iconoclastia* seria um movimento de destruição de imagens e símbolos, tipicamente de cunho religioso, mas podem ser também de valor político ou estético. A *iconofilia*, em contraste, seria o culto destas imagens e símbolos.

dade da religião, a crença nos fetiches, o culto ao transcendente, os ícones mandados do céu, a força das ideologias”. Ele sugere que isto também valeria para a ciência e o discurso de transcendência das verdades objetivas: contra as inscrições científicas, expor as práticas de sua construção. Ao que o discurso defensor diria que: “A única maneira de defender a ciência das acusações de falsificação [...] é, aparentemente, insistir em que não alguma jamais tocou na imagem que produziu” (LATOUR, 2008, p. 117). Ao não insistir na iconoclastia como recurso, o autor propõe tomá-la como um tópico da investigação. Ou seja, se afastando do olhar que busca revelar uma verdade por trás da imagem, volta seu interesse ao trabalho que produz tanto a iconoclastia como a iconofilia, em seus movimentos no entorno das imagens:

Poderíamos dizer, contra o ímpeto crítico, que quanto mais humanos há, mais o trabalho humano se mostra, melhor a apreensão da realidade, da santidade, da devoção. Que quanto mais imagens, mediadores, intermediários e ícones se multiplicam e são abertamente fabricados, quanto mais eles são explícita e publicamente construídos, mais respeito temos por sua capacidade de acolher, reunir, recolher a verdade e a santidade (LATOUR, 2008, p. 117).

As imagens são, então, dispostas em um campo de tensões em que agem umas sobre e com as outras, modulando o acontecimento midiático e político, inclusive em sua constituição histórica. Latour (2008, p. 117) sugere que um *iconoclash* seria: “aquilo que ocorre quando há incerteza a respeito do papel exato da mão que trabalha na produção de um mediador”. Quando não se sabe, em suma, se a mão visa destruir a imagem e expor sua fabricação ou se, de outro modo, ela visa justamente produzir uma imagem, como forma de revelação da “verdade”.

Volto-me ao caso da prisão de Lula com um olhar informado por estas proposições, ao mesmo passo em que me interesso por explorar a força das questões que desenvolvi nos capítulos anteriores. Realizo, deste modo, um enquadramento específico às questões gerais da tese, de cunho metodológico, segundo aquilo que observo como uma demanda do próprio acontecimento. Um aspecto da pertinência da articulação entre as questões metodológicas gerais e as especificidades do caso emerge, contudo, na medida em que a noção de *iconoclash* situa a imagem como mediador ativo daquilo que representa e que precisa ser considerada no vínculo com outros mediadores e outras imagens. Os apontamentos teóricos e metodológicos realizados nos capítulos precedentes voltam a incidir aqui, portanto, em uma tentativa de mapear o *iconoclash* da prisão por uma descrição dos fluxos imagéticos que operam como constituintes do acontecimento. O desafio desta descrição, por sua vez, busco enfrentar considerando a multiplicidade ontológica das

imagens tomadas como referências circulantes do acontecimento. Esta descrição somente seria possível pela mobilização das cadeias de translação que compõem as visualidades computacionais contemporâneas, com os dispositivos analíticos disponíveis.

Um elemento que explicita bem a tensão colocada sobre as imagens do acontecimento apareceu figurado na cobertura jornalística e nos registros difusos produzidos durante o desenrolar da vigília em São Bernardo. Como forma de controlar parte dos registros imagéticos produzidos sobre o acontecimento, foi instalado no SMABC um dispositivo improvisado de interdição às imagens. Um tecido amarelo foi estendido cobrindo a entrada da garagem do sindicato, como forma de impedir que se avistasse, do exterior, as movimentações internas – como, supõe-se, do próprio ex-presidente. Isto apareceu na cobertura da imprensa, chamando minha atenção ainda durante o desenrolar da vigília. A [Figura 14](#) exibe uma montagem com algumas das fotografias que foi possível encontrar entre as imagens selecionadas para a análise – em processo que discutirei adiante. Este pano pode ser compreendido como uma materialização da disputa então em jogo, interditando a produção de certas imagens para que outras pudessem ser produzidas e prevalecessem como figurações do acontecimento. Não se trata de um gesto iconoclasta clássico, como o do martelo que destrói uma imagem pagã. Objetivou-se, de outro modo, frustrar a imagem em sua gênese – interditá-la – em estratégia que produziu, por si própria, uma imagem e que, além disso, compôs um esforço em larga medida bem-sucedido de controlar as imagens produzidas e postas em circulação a partir da vigília em São Bernardo. Na linha do que discuti acerca de uma abordagem relacional para as imagens e, em particular, da configuração deste caso como um *iconoclash*, esta circunstância exemplifica meus argumentos e também apresenta-se, introdutoriamente, como um ponto tensionador das relações entre as imagens produzidas no decurso da vigília e após a prisão.

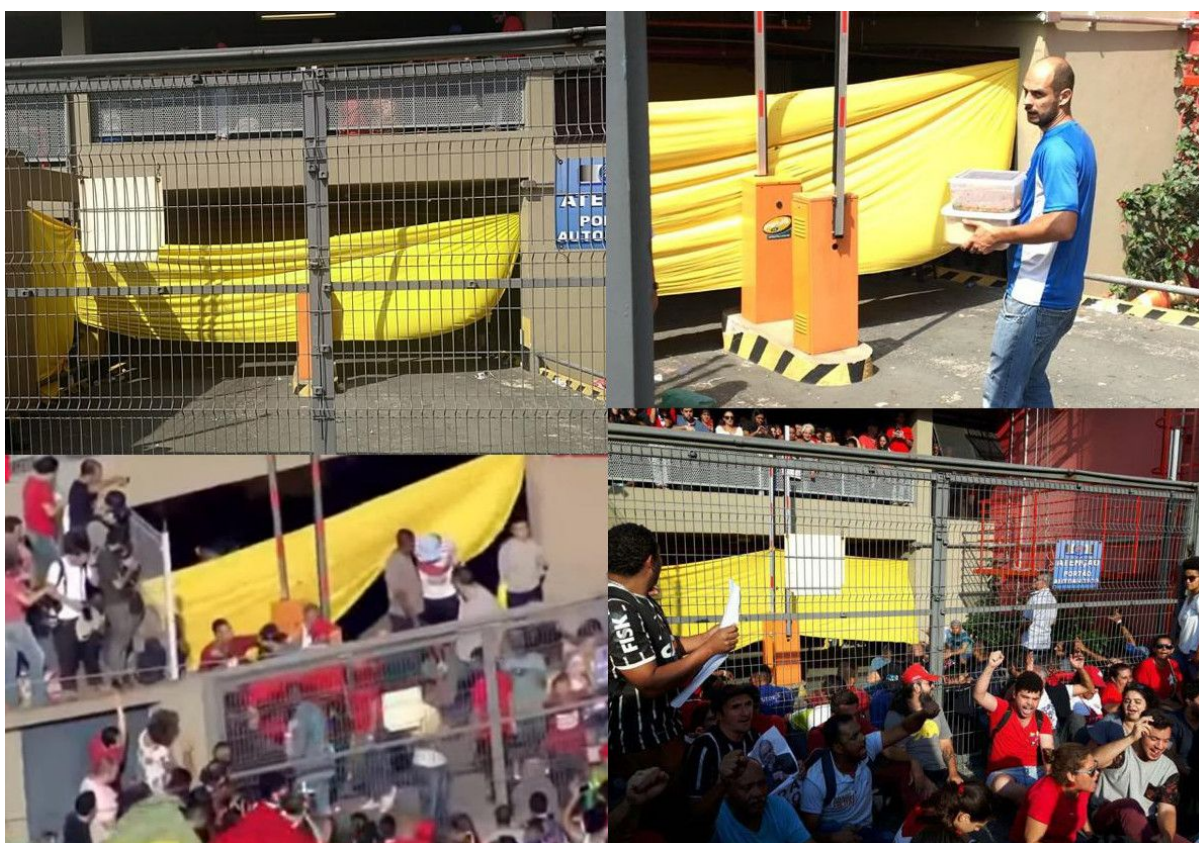


Figura 14: Montagem justapondo fotografias do tecido estendido no SMABC. As quatro fotografias foram publicadas no Twitter por contas vinculadas a veículos jornalísticos brasileiros e estrangeiros, entre o momento em que Lula discursou em frente ao SMABC e sua prisão, no dia 7 de abril.

Fonte: Montagem de elaboração própria a partir de fotografias obtidas por meio da API Streaming do Twitter. Em sentido horário, a partir do canto esquerdo superior, o quadrante das imagens na visualização de referência e as fontes encontradas para as imagens são: □ VGG19 0:7 | □ GVAPI M:6, conta da revista brasileira Carta Capital (@cartacapital) em tuíte publicado 7 de abril às 18h40 UTC; □ VGG19 0:10 | □ GVAPI M:6, conta do jornal brasileiro O Tempo (@otempo) em tuíte publicado 7 de abril às 16h26 UTC; □ VGG19 M:13 | □ GVAPI G:5, conta do jornal venezuelano A Todo Momento (@AtodoMomento) em tuíte publicado 7 de abril às 19h22 UTC, sendo que a mesma imagem foi publicada às 20h55 UTC por conta do jornal argentino El Tribuno (@eltribuno); □ VGG19 L:8 | □ GVAPI E:4, conta do canal de televisão uruguaio El Subrayado (@Subrayado) em tuíte publicado 7 de abril às 18h46 UTC. Nenhum dos tuítes indicavam fonte primária ou autoria para as fotografias publicadas.

Chegar até as imagens do acontecimento, contudo, constitui um dos primeiros desafios metodológicos desta investigação. Em mais uma consequência do desafio do tratamento computacional das imagens, as possibilidades proporcionadas pelas plataformas para acessar dados em uma pesquisa como esta são, tipicamente, centradas em elementos verbais de seus conteúdos. Deste modo, orientar a investigação por um viés imagético constitui uma primeira torção que precisa ser aplicada aos métodos de coleta e processamento. Os procedimentos adotados neste esforço serão, portanto, abordados na

primeira seção do capítulo. A segunda seção indicará, por sua vez, as técnicas de processamento empregadas uma vez que se conseguiu chegar a arquivos de imagens relativos ao acontecimento. Desafios específicos desta fase incluem o agrupamento de arquivos segundo relações construídas de *identidade e diferença* e, também, o esforço de traduzi-los em informações computáveis acerca de seu conteúdo, como auxílios ao desafio de navegar em meio ao conjunto. Por fim, a terceira seção apresenta o *Atlas para imagens-redes*, o dispositivo analítico constituído a partir destas etapas, indicando algumas possibilidades de ativação de seus potenciais.

4.1 Coleta e extração de dados

Por vezes relevada em descrições metodológicas, em particular quando subtende-se certo grau de objetividade para os dados coletados, considero, de outro modo, fundamental que a coleta seja descrita em detalhes. Mais do que isso, considero importante que seja abordada reflexivamente em termos daquilo que consegue revelar e daquilo que pode ter permanecido nas sombras ou nas brechas do dispositivo de captura constituído para a pesquisa. Nesse intuito, apresento nesta seção não apenas uma descrição do passo a passo que levou à constituição da base de dados em que se fundamenta a análise mas, também, uma análise em retrospectiva desse mesmo procedimento.

Assumo, nesse movimento, a proposição feita por Richard Rogers (2013, 2016) acerca das condições de “fundamentação” (*grounding*) da pesquisa em mídias digitais. Ele sustenta que uma pesquisa pode fundamentar-se em dados *online* à medida que o online se torna referência sobre a questão que se busca investigar. Se, no entanto, as mediações envolvidas no acesso a esses dados necessariamente transformam as observações possíveis, tais transformações não se aplicam somente à pesquisa, mas também ao modo com que a realidade observada é constituída no âmbito da Internet. Escreve o autor (ROGERS, 2017, p. 75–76):

Os métodos digitais desse modo reelaboram as condições de atestação, primeiro por considerar o *online* como base de fundamentação, mas também em um segundo sentido. Um pesquisador realiza descobertas *online* e, ao invés de deixar esse meio de lado para conferir maior consistência às descobertas, investiga, mais, em que medida o meio está afetando as descobertas. A pesquisa sobre o meio, então, serve a um propósito que é distinto do estudo apenas da cultura *online*¹³⁹.

¹³⁹ No original: “Digital methods thereby rething conditions of proof, first by considering the online as a site of grounding, but also in a second sense. One makes social research findings online, and, rather than leaving the medium to

Ou seja, é justamente pelo aprofundamento das questões acerca do método e suas mediações técnicas que se alcança, *não* uma positividade do conhecimento sobre o caso, mas, *sim*, uma compreensão mais ampla acerca das próprias condições de conhecimento. Aqui, proponho ir além das mediações técnicas, apenas, e considerar, reflexivamente, os limites inerentes aos métodos de coleta empregados e em que medida eles condicionam os dados disponíveis para análise.

Um primeiro aspecto por se considerar, portanto, é que mesmo o mais computacional dos métodos, ao se voltar para uma realidade social, demanda engajamento com o tema em questão, inclusive com certo caráter etnográfico. Afinal, antes de ter dados para pesquisar, uma pessoa interessada em determinado tema precisará constituir uma primeira compreensão a seu respeito, nem que seja apenas para selecionar um conjunto de termos de busca e critérios a partir dos quais constituir sua primeira base de dados. Isto apenas pode ser constituído em um contato direto com o tema, a partir da experiência do mundo. Para este caso específico, a elaboração do dispositivo de captura foi fundamentalmente informado pelo contato que tinha com o caso em meu envolvimento pessoal no acompanhamento do noticiário e das discussões em plataformas *online*, como Twitter e Facebook. Este primeiro contato, em larga medida assistemático, adensou-se com o ‘aquecimento’ da discussão, quando todas as possibilidades de recurso jurídico à prisão se esgotaram e, portanto, quando a prisão de Lula tornava-se iminente.

harden them, one subsequently inquires into the extent to which the medium is affecting the findings. Medium research thus serves a purpose that is distinct from the study of online culture alone”. Tradução minha.

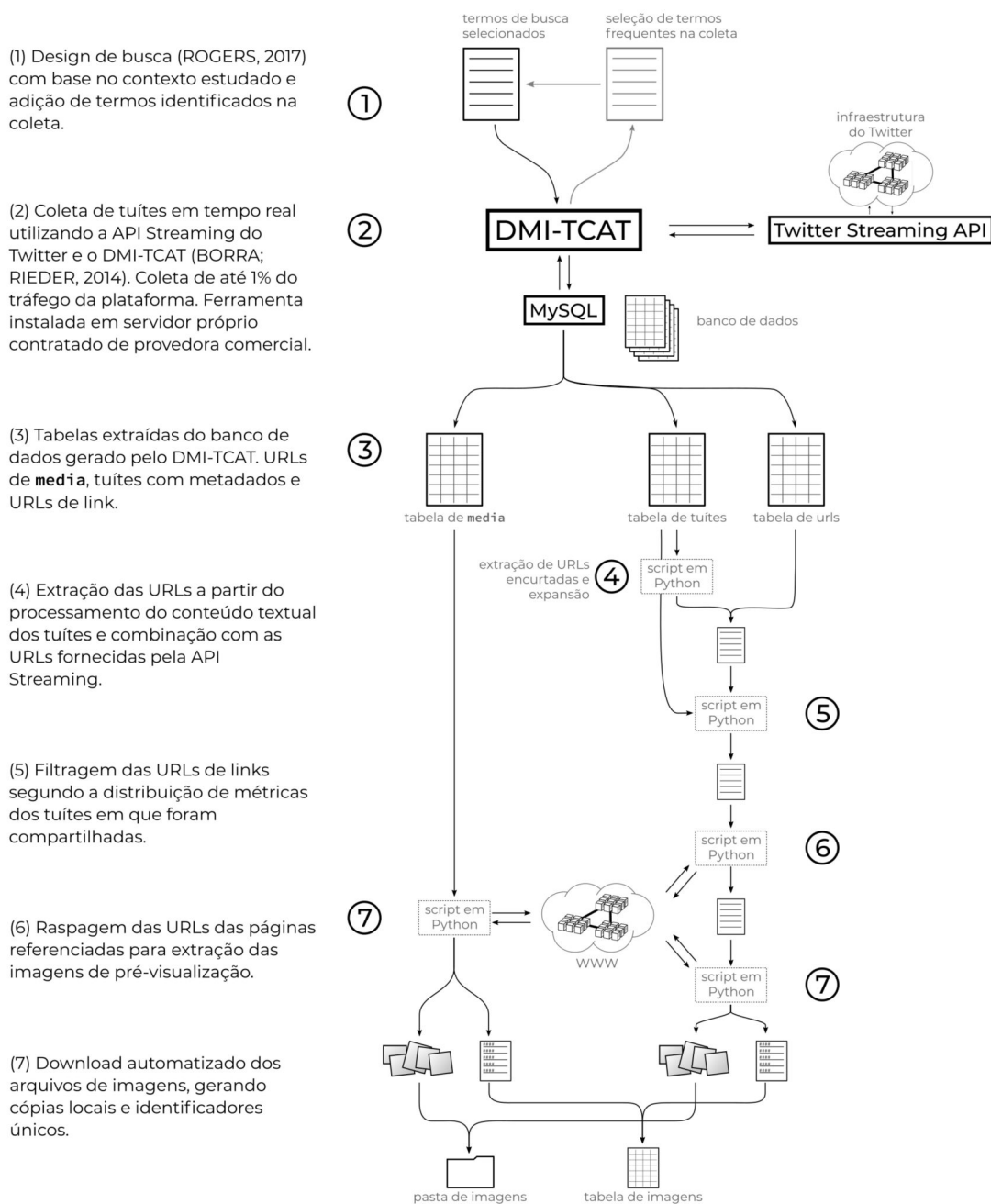


Figura 15: Protocolo de coleta e primeira fase de processamento.

Fonte: Elaboração própria.

Na [Figura 15](#) apresento um diagrama que representa esquematicamente os procedimentos de coleta e as tarefas de processamento que abordarei nesta seção do capítulo. Esta primeira fase compreende, portanto, desde o *design* de busca, com a escolha dos termos adotados para coletar dados do Twitter; até a extração e carregamento local das imagens a serem trabalhadas nas etapas subsequentes.

4.1.1 Construindo um dispositivo de captura

Como indiquei na introdução deste capítulo, os dados utilizados para analisar o caso foram inicialmente coletados sem que se visasse um objetivo de análise previamente estabelecido. Em todo caso, as coletas feitas desse modo voltavam-se ao Twitter como fonte prioritária dos dados, como acontece para muitas das pesquisas hoje realizadas em mídias sociais por métodos digitais. Este ponto merece consideração pois, como se sabe, o Twitter não está entre as plataformas mais utilizadas no Brasil¹⁴⁰. O critério de escolha desta mídia social não foi, efetivamente, a popularidade ou relevância específica para o caso em questão. Um fator crucial que justifica a coleta de dados do Twitter neste estudo, como também em outros casos, deve-se a se tratar de uma das plataformas mais permissivas quanto à possibilidade de acesso aos dados de publicações por meio de suas Interfaces de Programação de Aplicações (APIs), aspecto que elaboro a seguir. Um segundo ponto, talvez mais específico a este estudo, deve-se a que o objetivo perseguido não se volta tanto à plataforma de mídia social em si, às particularidades de sua constituição ou da conversação realizada por meio dela. De outro modo, interessam-me as imagens relacionadas ao acontecimento da prisão e o Twitter oferece-se como um meio privilegiado, dada a permissividade de sua API, para obter estas imagens e dados relacionados à sua circulação. Abordagem similar poderia ser realizada, por técnicas de raspagem, em outras plataformas como o Instagram ou mesmo – em outros protocolos de pesquisa e com outras implicações – em grupos de Whatsapp.

O Twitter, em particular, como elabora van Dijck (2014, p. 199–200), é frequentemente tomado por pesquisadores como um “termômetro” datificado das multidões. Como uma plataforma que reuniria as reações espontâneas de usuários em traços coletáveis e analisáveis em massa. Ou, ainda, como uma “ferramenta gigante de sondagem em tempo real”. Estas compreensões, sugere a autora, seriam sintomáticas do *dataísmo* – que dis-

¹⁴⁰ Embora tenha ganhado relevância recente, em especial pelo uso frequente desta plataforma por membros do atual governo federal – uma tendência também observada no contexto estadunidense.

cuti anteriormente – à medida em que não consideram a mediação da plataforma e os vieses que introduz nos dados e, portanto, nas análises. Em larga medida, a abordagem que faço do Twitter tampouco se aprofunda na investigação destes aspectos da plataforma, que é utilizada como uma porta de entrada para imagens em circulação. Contudo, gostaria de ressaltar a crítica de van Dijck e salientar que se problematizo este aspecto em profundidade não é por uma abordagem ingênua dos dados coletados mas, sim, pela escolha de um enfoque analítico. Justamente pelo não aprofundamento nestes aspectos, contudo, não chego nesta tese a elaborar asserções precisas sobre a realidade do caso abordado – não busco, por exemplo, definir qual instância do *iconoclash* teria saído “vitoriosa”. Estas questões fogem ao escopo deste estudo. Ao menos no âmbito dos procedimentos de coleta empregados, contudo, busco evidenciar as mediações em jogo, a começar pela porta de entrada oferecida pela plataforma, sua API.

Uma API, segundo o jargão da computação, é um tipo de estruturação de um programa de computador que permite algum grau de interoperabilidade com outros programas. Por meio de uma API, por exemplo, a pessoa que desenvolve um *software* pode habilitar que programas de terceiros façam uso de certas funcionalidades ou obtenham dados geridos pelo seu programa. No âmbito das pesquisas de internet e dos métodos digitais, as pesquisas frequentemente operam por meio das APIs constituídas pelas plataformas que permitem, entre outras funcionalidades, a extração de certos tipos de dados das publicações e atividades de usuários para fins diversos como pesquisas acadêmicas e comerciais. Esta ‘porta de entrada’ constitui, em larga medida, uma primeira forma de mediação que dispõe algumas das possibilidades analíticas de uma investigação. Taina Bucher (2013a), discutindo o caso específico das APIs do Twitter, salienta como estas não são objetos neutros, mas, sim, instâncias mediadoras que transformam os dados e cuja constituição não é nunca estável pois, embora participem das transações efetuadas por seus usuários, elas tanto constituem a relação quanto são por ela constituídas. Isto é: uma API define certas possibilidades, mas é, também, definida por seus usos. Transpondo esta indicação às APIs do Facebook, por exemplo, podemos compreender como elas permitiram uma realização como a da empresa Cambridge Analytica, que construiu perfis psicológicos de milhões de pessoas com base em dados obtidos da plataforma¹⁴¹. Compreendemos, também, como justamente esse caso pôde catalisar o progressivo fe-

¹⁴¹ Em artigo escrito com Olga Lukyanova em que apresentamos obra de arte que realizamos por meio da API do Facebook, desenvolvemos alguns aspectos relacionados (LUKYANOVA; MINTZ, 2018).

chamento das APIs como fonte de dados para a pesquisa em métodos digitais, criando uma crise atualmente vivida no campo, que diz-se entrar em um momento “pós-API” (cf. VENTURINI; ROGERS, 2019).

A Streaming API do Twitter, que utilizo para a coleta neste caso, permite a obter dados da plataforma “em tempo real” (item 2 da [Figura 15](#)). Isto é, os tuítes são coletados, minuto a minuto, assim que são publicados. Porém, ela não permite a coleta retroativa. Apesar desta limitação, os resultados oferecidos pela API são bastante robustos e permitem a coleta de volumes substanciais de dados. Até o momento da escrita, os critérios de requisição oferecidos pela Streaming API permanecem os mesmos (TWITTER, 2019a) daqueles disponíveis no período da coleta deste estudo e também conforme referenciado na bibliografia a seu respeito (PUSCHMANN; GAFFNEY, 2013). Os modos permitidos de coleta são:

- **track** (Rastrear): busca por tuítes contendo até 400 palavras-chave;
- **follow** (Seguir): tuítes postados por até 5.000 contas de usuário identificadas por seus números identificadores;
- **location** (Localização): busca por tuítes segundo sua geolocalização, com escopos definidos por até 25 quadrantes de coordenadas geográficas.

Historicamente, a API Streaming oferece a usuários comuns acesso limitado a até 1% do tráfego momentâneo da plataforma (PUSCHMANN; GAFFNEY, 2013; RIEDER; BORRA, 2014). Embora a documentação atual da API não apresente esta indicação específica, o desempenho da ferramenta não sugere grandes mudanças em relação a esta limitação. Evidentemente, contudo, a ausência de dados quanto ao volume total do tráfego na plataforma dificulta avaliar o limite com precisão.

Como indica a sigla API, trata-se de uma interface de *programação de aplicações*, o que significa que não é possível acessar diretamente os dados por meio de uma interface gráfica de usuário (GUI, na sigla em inglês). O acesso se dá por meio de comandos, tipicamente automatizados por meio de um programa. Para obter os dados da API Streaming, portanto, utilizei o *Kit de Ferramentas para Captura e Análise do Twitter (Twitter Capture and Analysis Toolset – DMI-TCAT)*. Este é um conjunto de ferramentas computacionais desenvolvido no âmbito da *Digital Methods Initiative (DMI)*, da Universi-

dade de Amsterdam (RIEDER; BORRA, 2014). A escolha desta ferramenta e API¹⁴², se deu principalmente em função da robustez da API Streaming e da fundamentação metodológica do DMI-TCAT, segundo elaboram seus desenvolvedores (RIEDER; BORRA, 2014).

Considerando a ferramenta um “método programado”, Rieder e Borra argumentam que a constituição do DMI-TCAT não pode ser resumida a um conjunto de proposições técnicas e que, de outro modo, é preciso considerá-la segundo suas orientações epistêmicas e as repercussões, destas, na pesquisa acadêmica derivada (RIEDER; BORRA, 2014, p. 263). Nesse sentido, os autores propõem uma série de diretrizes para a ferramenta que vão desde a gratuidade do programa e a abertura do código, até a opção por armazenar os dados da forma mais próxima possível da estruturação oferecida pela API do Twitter. Ao tornar o *software* gratuito e de código aberto, seu acesso é facilitado, além de se permitir sua expansão e escrutínio por parte de pessoas que venham a empregá-la em suas pesquisas. Por sua vez, a priorização do modo de estruturação de dados da própria API do Twitter torna a mediação analítica da ferramenta sempre reversível e aberta a múltiplas abordagens – salvaguardando, como afirmam os autores, a pluralidade epistêmica (RIEDER; BORRA, 2014, p. 266).

A coleta de termos relacionados ao Lula teve início, experimentalmente, no dia 18 de janeiro de 2018, às vésperas do julgamento de seu caso em segunda instância, no TRF-4. Contudo, o recorte escolhido para a análise compreende o período entre 4 e 16 de abril de 2018, quando se desenrolou a prisão do ex-presidente e em que a coleta foi acompanhada de modo mais próximo. Especificamente, foram extraídos, do conjunto mais amplo de dados, todos os tuítes obtidos desde as 3h00 do dia 4 até as 3h00 do dia 17 de abril de 2018, no horário UTC (Tempo Universal Coordenado, na sigla em inglês), segundo o fuso horário padrão utilizado pela API do Twitter¹⁴³. Deste modo, cobre-se desde a véspera da expedição do mandado de prisão até o 10º dia da prisão do Lula.

Nesse período específico, sem prejuízo para os termos de coleta cadastrados ainda em janeiro, os filtros da coleta foram revisados e atualizados diariamente, de modo a acompanhar os termos e *hashtags* cuja utilização era percebida em meio à conversação sobre o

142 A coleta de dados do Twitter pode ser realizada também via a API *Search* e ferramentas que a ela se vinculam, com acesso possível a tuítes postados no passado a partir de termos de busca. Contudo, trata-se de interface que tipicamente oferece bem menos resultados do que a API *Streaming*, além de não haver indicação tão clara acerca de quais os limites e a forma de mediação dos resultados.

143 À época, o UTC diferia em +3 horas do horário de Brasília. Portanto, o período estipulado equivale, para o fuso horário brasileiro, à faixa entre as 0h00 do dia 4 e 0h00 do dia 17 (ou 24h00 do dia 16), respectivamente. Entretanto, o escopo efetivo para o início da coleta foi prejudicado por uma falha temporária da coleta, de modo que o primeiro tuíte efetivamente coletado no dia 4 foi publicado às 12h20 UTC (09h20 em Brasília).

acontecimento. Chama-se a esta técnica de *snowballing* (bola de neve): a partir dos resultados obtidos pela coleta, observar possíveis termos que permitiriam sua ampliação para além dos limites inicialmente definidos (item 1 da Figura 15). Para além dos resultados da coleta, contudo, busquei também observar, nas notícias divulgadas por portais noticiosos e no próprio Twitter, os termos utilizados por diferentes atores. Este acompanhamento se deu por meio da interface *Tweetdeck*, do próprio Twitter, que permite várias visualizações alternativas ao *feed* tradicional, como o acompanhamento da menção de termos em tuítes em tempo real ou o acompanhamento simultâneo de múltiplas listas de contas de usuário de interesse, por exemplo. Para este fim, procurei estabelecer listas de atores ‘lulistas’, ‘lavajatistas’ e de veículos de notícias, além de acompanhar ocorrências do termo [lula], em especial de tuítes com imagens.

Richard Rogers (2017) denomina *design de busca* (*query design*) esta tarefa de escolha e refinamento das palavras-chave e filtros utilizados nas coletas de dados. O autor salienta, por exemplo, como a escolha deve buscar refletir o vocabulário utilizado por diferentes perspectivas sobre o tema estudado. Sugere, também, que a seleção dos termos deve se dar por um processo de revisão em que os resultados da busca são analisados a fim de aprimorar a qualidade da coleta, em um ciclo de retroalimentação. Assim se deu o processo de coleta do caso em questão, embora nem sempre da forma mais racional e dedutiva. Como já indiquei, devido à temporalidade peculiar das APIs do Twitter, dados de melhor qualidade – em maior quantidade, com mais metadados e com maior precisão sobre o escopo de coleta – são obtidos em “tempo real”. Por isso, é necessário que os termos de busca pertinentes sejam cadastrados na busca *antes* de os tuítes contendo estes termos sejam publicados. Procurei, nesse sentido, antecipar alguns tópicos de conversação como parte do esforço de coleta.

Em certo sentido, trata-se de processo análogo ao de se construir uma *armadilha* – tal como um dos sentidos primeiros da palavra *rede* (MUSSO, 2004, p. 18)¹⁴⁴. Deve-se antever aspectos da conversação, prever possíveis formulações verbais e acompanhar atentamente a evolução dos vernaculares. Por isso, a escolha e refinamento dos termos envolve planejamento e improvisação em igual medida, em tentativas de coletar vertentes imprevistas do acontecimento em curso, quando já se observa a repercussão de certos termos. Em especial ao longo dos dias entre 5 de abril, quando houve a expedição

¹⁴⁴ Segundo descreve o autor, a palavra francesa *réseau* (rede) apenas surge no século XI designando “redes de caça ou pesca e tecidos, uma malhagem têxtil que envolve o corpo” (MUSSO, 2004, p. 18).

do mandado de prisão, e 7 de abril, quando a prisão se concretizou, acompanhei constantemente a conversação no entorno de palavras-chave já antevistas para a coleta. Esta observação permitiu compreender aspectos da conversação em tempo real e intuir outros termos relevantes que talvez representassem lacunas da coleta, os quais foram acrescentados durante esse período (veja Gráfico 2).

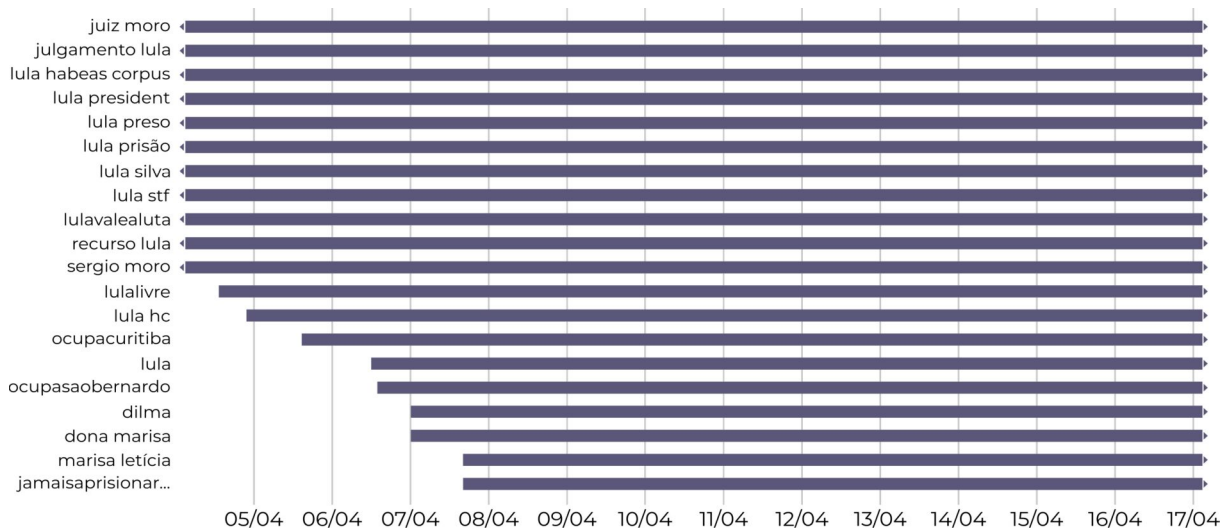


Gráfico 2: Diagrama de Gantt de períodos de coleta para termos de busca. Seleção dos 20 termos com mais ocorrências isoladas.

Fonte: Visualização de elaboração própria criada a partir dos dados da coleta por meio da ferramenta RAW Graphs (MAURI et al., 2017) com adaptações e ajustes no Inkscape (ALBERT et al., 2018).

Justamente pelo imprevisto, nem sempre as adições de termos consideraram a possível redundância com outros já cadastrados, além de, por vezes, visarem vertentes pouco expressivas em meio ao conjunto. Entre os termos acrescentados apenas durante o desenrolar dos acontecimentos está o mais óbvio: [lula]. Em experimentos de coleta anteriores, havia-se verificado que o termo era demasiado inespecífico e tendia a coletar tuítes não-relacionados ao tema visado – tanto pelo animal de mesmo nome quanto pelo uso do termo como apelido, inclusive em outros idiomas. Por isso, a coleta inicialmente incluiu o termo apenas em junto a outros, de modo a restringir o escopo de menções, como [lula prisão]. Nesses casos, apenas quando ambos os termos fossem mencionados o tuíte seria coletado. Contudo, à medida que o fluxo da coleta se intensificou, experimentos de busca pelo termo na interface *Tweetdeck*, do Twitter, indicaram que a inespecificidade já não constituía problema e o ‘ruído’ poderia ser desprezado. O termo [lula] passou a ser coletado às 14:42 do dia 5 de abril, sem que os termos anteriores, a ele redundantes, fossem excluídos. Um ruído gerado por esta inclusão, que foi notado apenas posteriormente à coleta, deve-se a que, além de outros contextos de uso

para o termo Lula, também foram coletados tuítes em que o termo aparecia como parte de uma palavra (como ‘celular’). Entretanto, como se verá, isto não chegou a comprometer a análise, já que o volume destas ocorrências foi muito inferior àquele efetivamente pertinentes ao caso¹⁴⁵.

No total, a coleta se baseou em 109 termos de busca cadastrados no DMI-TCAT. Contudo, considerando o desempenho real da coleta, é possível restringir este número a 68 termos relevantes, considerando apenas aqueles que, para ao menos um tuíte coletado, não foram redundantes com nenhum outro. Quero dizer de termos que tenham sido determinantes para coletar ao menos um tuíte em meio ao montante efetivamente coletado. Para isto, basta considerar se o termo foi encontrado em um tuíte no qual nenhum outro termo foi encontrado. Esta análise é possível tomando por base a tabela `tcat_captured_phrases`, que o DMI-TCAT cria no banco de dados da ferramenta. Esta tabela armazena, para cada tuíte coletado, qual ou quais dos termos cadastrados foram encontrados. Desconsidera-se, desse modo, termos que foram sempre encontrados em conjunto com outros termos e que, por isso, são apenas redundantes. Entretanto, como a lista de termos utilizados na coleta foi dinamicamente alterada ao longo do período considerado, os termos relevantes não excluem toda forma de redundância. As análises subsequentes desta seção considerarão apenas esse subconjunto de 68 termos relevantes. O objetivo desta filtragem é apenas reduzir o escopo considerado neste olhar retrospectivo ao esforço de coleta. Apresento no [Apêndice A](#) tabela contendo todos os termos utilizados na coleta, inclusive os depois considerados irrelevantes, junto à indicação do momento de início da coleta, bem como dos índices ocorrência total e em isolamento.

O [Gráfico 3](#) apresenta os termos de busca utilizados na coleta na forma de um grafo, construído segundo as coocorrências verificadas entre eles nos tuítes coletados. Ele serve como um instrumento para compreender o desempenho das estratégias de coleta diante do caso observado. Embora se trate de um grafo denso e muito conectado – a que se costuma referir por “bola de cabelo” – é possível descrever alguns aspectos da correspondência entre os termos segundo a espacialização obtida na visualização.

¹⁴⁵ Exemplo deste tipo de ruído pode ser encontrado em [VGG19 N:8](#) | [GVAPI P:1-2](#). Nesses quadrantes, encontram-se algumas imagens do carro em que a vereadora Marielle Franco foi assassinada em março de 2018. Os tuítes coletados para esta figura, como pode ser observado na página de detalhamento para o agrupamento, não têm relação com o caso Lula, nem contêm o termo como uma palavra. A coleta se deu, contudo, por causa da palavra ‘celular’. No dia 6 de abril, foi noticiado que a polícia havia descoberto celular usado pelo motorista do carro usado no assassinato.

Tipicamente, os algoritmos de espacialização dirigidos por força, como o ForceAtlas2 (JACOMY et al., 2014), utilizado nesta visualização, tendem a aproximar nós conectados e distanciar nós desconectados, gerando uma forma que pode ser analisada visualmente (VENTURINI; JACOMY; JENSEN, 2019). Trata-se de uma abordagem distinta, mas nem por isso incompatível, da análise computacional de grafos (cf. GRANDJEAN; JACOMY, 2019). Para este caso, o grafo oferece um modo de compreender como a lista de termos selecionados ao longo do processo de coleta relacionam-se entre si, de modo a compor a *teia* para a captura das conversações em curso na plataforma. Áreas mais densas da rede sugerem campos semânticos mais amplamente cobertos, inclusive com redundâncias, e que tendem a possuir um melhor desempenho de captura. Áreas menos densas, mais à margem da rede, sugerem campos semânticos menos cobertos que podem possuir “buracos” e que, por isso, podem ter deixado passar tuítes marginalmente relevantes pelo não cadastramento de termos específicos correspondentes.

na rede, em que podemos observar expressões distintas sobre o acontecimento. Em direção à parte inferior do grafo, encontramos termos tipicamente relacionados a perspectivas celebratórias da prisão, com referências pejorativas ao ex-presidente – como as variantes do termo ‘molusco’ – e com diferentes verbalizações da prisão, como [lulaprajaula], [lula cadeia] e [lulanapapuda]. Em direção à margem direita superior, encontramos termos relacionados à campanha pela liberdade do ex-presidente, como [lulavalealuta], [jamaisaprisionaráonossossonhos] e [lulainocente]; à sua candidatura às eleições presidenciais naquele ano, [lula2018]; e a chamadas pela ocupação de locais em que se desenrolavam o acontecimento, como [ocupasaobernardo], [ocupacuritiba] e [ocupaportoalegre]. Em direção à borda superior, encontramos termos mais relacionados ao julgamento e a condenação em si do ex-presidente, com menções ao apartamento triplex que foi objeto da denúncia, ao nome do advogado de defesa, Cristiano Zanin, e ao delator Léo Pinheiro, da construtora OAS. Também nesse campo encontra-se um termo específico tipicamente usado na campanha pela liberdade de Lula, mas que se direciona ao questionamento da legitimidade da condenação: [cadeaprova]. Em direção ao lado esquerdo, há termos relacionados ao juiz e desembargadores que julgaram o caso bem como a instâncias jurídicas que se envolveram no processo, como [trf 4], [lula stf] e [lula stj].

Podemos descrever esses diferentes campos semânticos segundo a conceituação de *programas* e *antiprogramas* proposta por Rogers (2017) na sistematização de metodologias de coleta. Rogers se baseia, nessa proposição, no trabalho de Madeleine Akrich e Bruno Latour (AKRICH, 1992; AKRICH; LATOUR, 1992; LATOUR, 1992) – outros aspectos destes textos foram já discutidos no capítulo 2, *Imagem-rede*. Os setores inferior e direito superior seriam, nesse sentido, correspondentes ao par *programa* e *antiprograma* do acontecimento. Se tomamos por referência a consumação da prisão durante aqueles dias, a parte inferior poderia ser considerada como termos relacionados ao *programa*, em perspectivas marcadamente concordantes e mesmo celebratórias da prisão. O setor direito superior já apontaria para um *antiprograma*, com a defesa da liberdade de Lula e o questionamento da legitimidade da condenação. Os setores esquerdo e superior, por sua vez, apontam, em parte, para o que Rogers descreve como “esforços de neutralidade”, pois são termos que, em si mesmos, não são necessariamente partidários e que, de outro modo, tendem a um esforço de descrição factual do acontecimento.

Um aspecto por se salientar sobre o processo de coleta, a partir destas análises, é que por mais que estejam contidas diferentes perspectivas do caso e por mais que o volume de tuítes pareça bastante grande, não se pode sugerir, de modo algum, que os dados dizem de uma *totalidade* da conversação. A única totalidade que se delimita, aqui, é a dos próprios dados coletados, os quais apenas podem representar o caso abordado por uma amostragem incerta. Por entre as teias do dispositivo de captura construído pelo *design* de busca certamente escoaram diversas formas de enunciação sobre o acontecimento que não foram antecipadas pelo dispositivo.

Esse escape pode se dever a diversos motivos como: a necessária insuficiência dos termos de busca; a impossibilidade de prever palavras-chave emergentes; os limites e mediações opacas da API do Twitter; ou, ainda, a contingências e falhas das múltiplas camadas técnicas mobilizadas para a coleta. Mais ainda, haveria um *hiato semântico*, pelo qual não há nem mesmo a possibilidade de um léxico finito por meio do qual detectar menções ao caso. Certamente, múltiplas postagens relacionadas se deram sem menção a qualquer dos termos mapeados ou mesmo a qualquer termo *mapeável*. Ainda mais tendo em vista a abrangência do caso em análise, há certamente tuítes cuja relação com a prisão do Lula apenas poderia ser compreendida contextualmente. Isto é, há toda uma dimensão insondável do acontecimento que deve ser reconhecida.

De modo importante para o tema desta tese, há que se considerar, ainda, como a coleta se limita a busca por *termos*, isto é, por chaves de busca verbais (cf. D'ANDREA; MINTZ, 2019). Não há a possibilidade de coletar tuítes segundo a imagem que contêm, aspecto que apenas pode ser considerado uma vez constituído o conjunto dos dados para análise. Esta seria uma questão para a qual da hipótese conceitual da *imagem-rede* viria ao auxílio da discussão. Para o estudo deste caso com um enfoque nas imagens, estas seriam profundamente dependentes de sua vinculação a palavras. Não há acesso direto à dimensão visual do acontecimento que não passe por sua dimensão verbal, portanto. A conexão entre as imagens e os termos de busca seria uma daquelas por perseguir na constituição dos objetos analíticos do estudo que, por isso, podem ser compreendidos como imagens-redes. A inclusão de *uma* imagem no escopo da pesquisa somente seria autorizada por meio da ocorrência de termos verbais e, com isso, certamente, muitas imagens *escapam*. Ainda que não tenhamos como precisar as dimensões e as caracte-

rísticas daquilo que se ausenta dos dados coletados, compreender a existência desta negatividade do método é fundamental para resistir à presunção positivista que metodologias baseadas em dados computacionais tendem a inspirar (cf. GALLOWAY, 2014).

Todos estes aspectos constituem, entretanto, ainda apenas uma dimensão da *negatividade* do método, que chamaria de uma *negatividade positivável*. Afinal, tratam-se de lacunas da coleta que podemos, ainda assim, identificar e reconhecer enquanto um limite conhecido do método adotado, muito embora sua extensão e impacto sobre a pesquisa não pode ser de todo precisado, mas apenas estimado. Uma negatividade não positivável, contudo, diria respeito ainda às sombras projetadas pelo método e que não podem ser identificadas com igual precisão, que derivam da *episteme* que ele constitui e que apenas poderíamos reconhecer por outra matriz metodológica. Ou, ainda, como sugere Sean Cubitt (2013), por evidências anedóticas que teriam precisamente a capacidade de furar o encapsulamento epistêmico dos métodos de observação e análise empregados computacionalmente e em larga escala. Quero, portanto, salientar que o cuidado com que busco evidenciar as lacunas e incompletudes do método não tem o objetivo de conter ou neutralizar a *negatividade* que viria a corroer os dados e números trazidos à análise. Gostaria, de outro modo, de salientar a circunstancialidade que ainda constitui esse processo, contra premissas de objetividade que ele tende a inspirar.

4.1.2 Tuítes

Os dados obtidos por meio dos procedimentos de coleta descritos acima têm como principal unidade de análise os *tuítes* – as publicações feitas na plataforma. Destes será possível extrair, na sequência, *metadados* e as imagens consideradas na análise. A base constituída pela coleta é composta por 7,3 milhões de tuítes (Tabela 1), dos quais uma ampla maioria (5,4 milhões, ou 75%) são retuítes. Este dado sugere inicialmente que a conversação teria uma dinâmica relativamente concentrada, com poucos atores postando conteúdos originais e um grupo mais extenso apenas replicando esses conteúdos. Entretanto, como também indica a Tabela 1, do total de 1,16 milhão de contas que publicaram conteúdos, cerca de 77% publicou ao menos um tuíte ‘original’, ou seja, que não era uma republicação direta do tuíte de outra conta. Portanto, embora a ampla maioria dos tuítes coletados sejam republicações, a também ampla maioria das contas de usuários postou ao menos um tuíte original no período observado.

Tabela 1: Descrição geral dos tuítes do caso Lula.

Fonte dos dados: API Streaming do Twitter e DMI-TCAT.

total de tuítes	7.265.147	100,00%
retuítes	5.401.941	74,35%
tuítes 'originais'	1.863.206	25,65%
contas de usuário	1.159.747	100,00%
contas com tuítes 'originais'	896.260	77,28%
contas sem tuítes 'originais'	263.487	22,72%

Outro aspecto destacado em uma visão geral dos dados é a já esperada concentração temporal da atividade em torno dos dias da vigília de Lula no SMABC. A maior parte dos 7,3 milhões de tuítes coletados foram publicados nos primeiros quatro dias do período, entre 4 e 8 de abril (Gráfico 4). Nos dias seguintes, a atividade foi progressivamente decaindo. Os períodos com mais publicações relacionam-se a momentos-chave para o acontecimento. Primeiro, quando da rejeição do habeas corpus preventivo solicitado pela defesa do ex-presidente (4 de abril, próximo da meia-noite no horário local). No dia seguinte, quando houve a expedição do mandado de prisão (5 de abril, próximo das 18h00 no horário local). No dia 6 de abril, quando expira o período concedido à entrega voluntária de Lula. E, enfim, no dia 7 de abril, observa-se o ponto mais alto da taxa de publicações, próximo às 19h00 no horário local, logo após a realização da prisão, com 128,7 mil tuítes ao longo daquela hora.

Reforçando observações feitas na seção anterior a respeito da *negatividade positivável* do processo de coleta, o Gráfico 4 exibe na parte inferior uma visualização dos períodos em que a coleta foi restringida pela API Streaming do Twitter por exceder o volume de captura permitido. Sem entrarmos no nível das estimativas numéricas absolutas¹⁴⁷, calculadas pelo DMI-TCAT, os períodos e a magnitude relativa dessas limitações indicam as janelas de fragilidade do processo de coleta. Estas coincidem justamente com os momentos em que os acontecimentos e a dinâmica de publicações são mais intensos.

¹⁴⁷ Evito entrar nesse ponto, especialmente, em função do dado não ser exclusivo a esta coleta específica realizada no DMI-TCAT, mas ao conjunto delas. O servidor que utilizei nesta coleta possuía cadastrados outros termos além dos relativos ao caso do Lula.

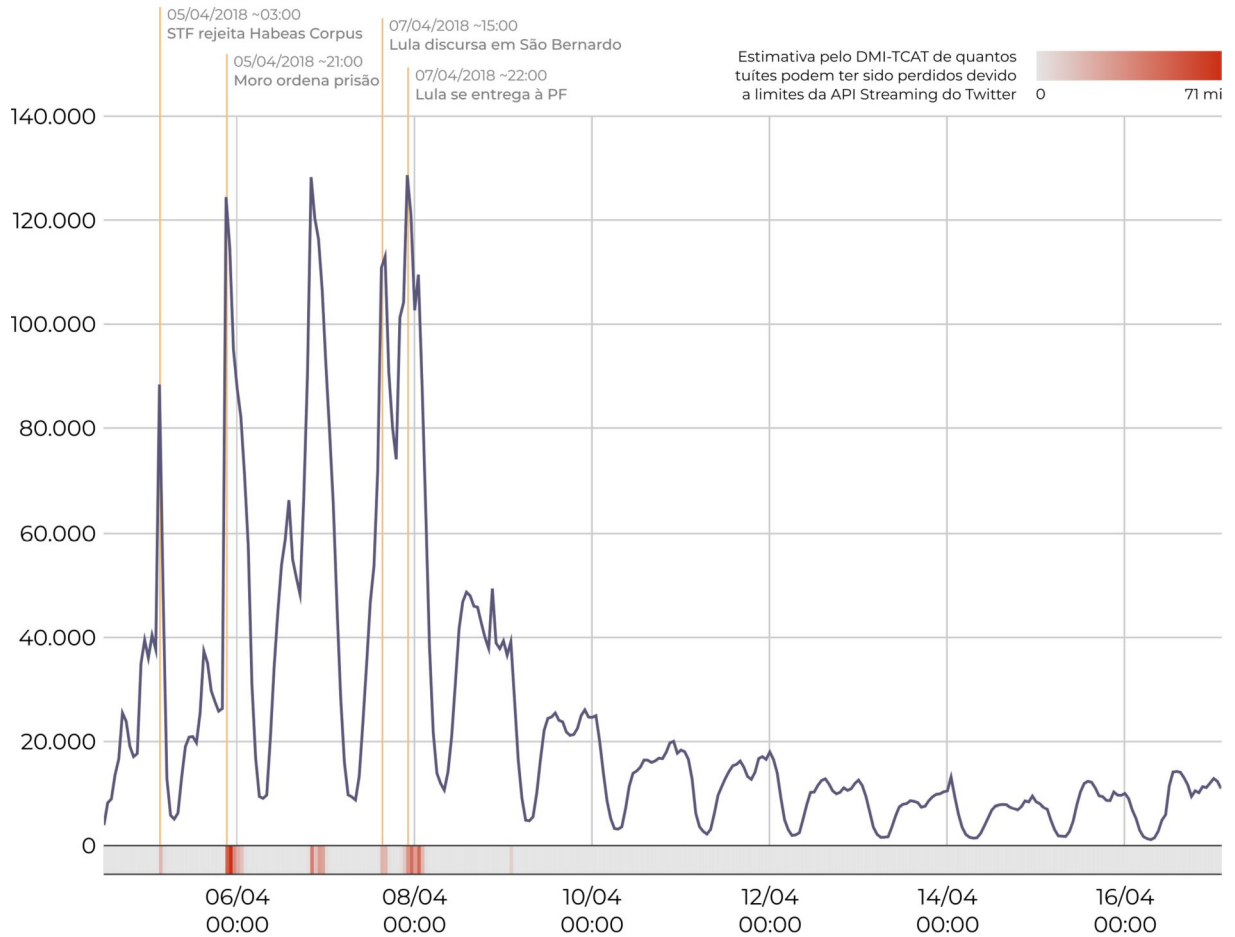


Gráfico 4: Série temporal do número de tuítes coletados, agregados por hora.
Período: de 04/04/2018 12:00 a 17/04/2018 03:00, UTC.

Fonte: Visualização de elaboração própria, com diferentes ferramentas para análise e representação gráfica dos dados (ALBERT et al., 2018; FRY et al., 2019; GOOGLE, 2019b; TIMM et al., 2018).

Fonte dos dados:
API Streaming do Twitter e DMI-TCAT.

Considerando o interesse específico desta investigação pelas imagens do acontecimento, um dos primeiros desafios colocados ao processamento dos dados diz respeito à extração das imagens por se analisar. Mesmo a definição de quantos ou quais destes tuítes foram apresentados com imagens não é algo que seja possível determinar com exatidão. Pelo modo com que o Twitter exibe as publicações, há ao menos dois modos pelos quais um tuíte é apresentado com imagem. Primeiro, quando a publicação incluiu uma imagem como seu conteúdo, isto é, quando quem publica faz *upload* de um arquivo de imagem como parte de sua publicação (Figura 16). Nesse caso, o Twitter armazena o arquivo e gera uma URL única para seu acesso.



Figura 16: Impressão de tela do Twitter com imagem carregada na plataforma.

Fonte: Twitter. Conta da UFMG.

Uma outra possibilidade é a exibição de imagens de pré-visualização de *links* contidos nos tuítes (Figura 17). Nesses casos, quem publica o tuíte apenas indica a URL. A plataforma, por sua vez, busca na página referenciada a indicação do *metadado* indicando a URL de uma imagem de pré-visualização, que é, então, exibida na interface do próprio Twitter. Questões sobre o uso desse metadado como recurso para a pesquisa serão discutidos adiante. O que é importante compreender, nesse momento, é a necessidade de desdobrar a busca por imagens para a investigação em duas frentes complementares, uma voltada aos ‘objetos de mídia’ (imagens estáticas ou em movimento) carregadas diretamente na plataforma. Chamarei a este caso de URLs de *media*, já que, efetivamente, esta é a materialidade da inscrição com a qual é preciso lidar durante o processamento. A outra frente se volta às imagens possivelmente exibidas como pré-visualização. Para este caso, estamos falando de URLs inseridas nas publicações na forma de *links* vinculando a publicação a outro endereço na web – chamo a estas de URLs de *links*.



Figura 17: Impressão de tela do Twitter com imagem de pré-visualização de link.

Fonte: Twitter. Conta da UFMG.

4.1.3 URLs de mídia

Quando a postagem vem acompanhada de uma imagem ou de um vídeo, o arquivo é carregado no Twitter e a plataforma gera uma URL para seu acesso, que é utilizada para exibição junto à postagem. Cada tuíte, por limitações colocadas pela própria plataforma, pode conter até quatro arquivos de **mídia**, categoria que abarca três tipos de arquivo, segundo distinções estabelecidas pelo Twitter:

- **photo**: qualquer imagem estática, não necessariamente fotográfica, apesar do nome da categoria. Seu armazenamento se dá pelo formato de uso difundido JPEG (veja discussão iniciada na p. 86).
- **animated_gif**: vídeo em *loop* de breve duração. Embora seja referido como um GIF, um formato de imagem que permite armazenar múltiplos quadros exibidos como animação, o Twitter na realidade converte este formato a um arquivo de vídeo no formato MP4.

- **video**: vídeos com duração de até 2 minutos e 20 segundos, segundo especificações da plataforma (TWITTER, 2019b). Esta limitação não se aplica a contas selecionadas pela plataforma, que podem postar vídeos mais longos.

Um aspecto por se observar acerca desta tipologia, reproduzida nos dados obtidos via API, são desacordos entre as categorias e os tipos de arquivo e de mídia a que se referem. A categoria `animated_gif` faz menção a um formato de arquivo (GIF) quando não se trata, efetivamente, do tipo de arquivo utilizado para o objeto de `media`, que é carregado como um arquivo de vídeo no formato MP4¹⁴⁸. Desse modo, a menção ao GIF tem menos um sentido estritamente técnico do que sugere algo como um ‘gênero’ de imagens em movimento, caracterizadas pela curta duração e exibição em *loop*. Contudo, em especial, interessa a esta pesquisa a peculiaridade daquilo que é considerado `photo`. A categoria indica, afinal, não uma foto, necessariamente, mas qualquer imagem estática, apresentada no formato JPEG. Esta subsunção de toda imagem estática a um tipo chamado de `photo` parece apontar, em um primeiro momento, pra uma expectativa da plataforma quanto ao tipo de imagem estática mais comum. Para além de uma questão de expectativa, contudo, esta indicação parece reforçar a compreensão sugerida por Ingrid Hoelzl e Rémi Marie (2015) de que no âmbito das imagens digitais a ontologia da imagem fotográfica seria reconfigurada: “não mais denotando imagens produzidas de uma determinada maneira, mas uma categoria de imagens que exibem uma distribuição estética específica (tom contínuo) com a qual o algoritmo JPEG funciona melhor” (HOELZL; MARIE, 2015, p. 3)¹⁴⁹. De certa maneira, portanto, não se trata necessariamente de uma questão de expectativa, apenas, mas de uma efetiva reconfiguração da categoria, por meio da qual tudo o que é JPEG seria, em certos contextos, uma foto.

Outro aspecto importante sobre os arquivos de `media` diz respeito ao modo de vinculação das URLs às imagens. De modo importante, não existe garantia de que para uma imagem haverá apenas uma URL. Pelo contrário, tipicamente, haverá várias. Isto porque as URLs são geradas a cada novo carregamento de um arquivo para publicação, mesmo que se trate exatamente do mesmo conteúdo de outra publicação já realizada. Há casos, entretanto, em que uma mesma URL de `media` é referida por diferentes tuítes. Para o caso de imagens estáticas, isto ocorre quando se trata de um retuíte ou, de modo mais

¹⁴⁸ Provavelmente, a conversão ao formato de vídeo visa melhor comprimir os arquivos, já que o algoritmo de compressão MP4 não salva os *frames* inteiros, diferentemente do formato GIF.

¹⁴⁹ No original: “no longer denotes images produced in a certain way, but a category of images that display a specific aesthetic distribution (continuous-tone) with which the JPEG algorithm works best”. Tradução minha.

improvável, quando um tuíte insere, de forma literal, uma URL da imagem que tivesse sido publicada por um outro tuíte. No caso de imagens em movimento, além do retuíte, há casos em que o arquivo é apenas publicado novamente sem vir acompanhado do tuíte original, possibilidade permitida para arquivos do tipo *video*, em particular. Uma consequência disto para a análise será abordada adiante (veja [4.2 Processamento das imagens](#)): para falarmos de *uma* imagem, será preciso encontrar um modo de agrupar, a partir de seu conteúdo, os diferentes arquivos que seriam pertinentes a uma mesma imagem. Não se trata de problema que possa ser resolvido de forma definitiva.

Dos 7,3 milhões de tuítes que compõem a base deste estudo, 1,2 milhão (16%) contém algum tipo de *media*. Esta indicação é oferecida pela própria API Streaming do Twitter, nos dados recuperados pelo DMI-TCAT (item 3 da [Figura 15](#)). Os tuítes são representados, nos dados obtidos da API, como um *objeto*, que nada mais é do que uma coleção estruturada de dados referentes ao tuíte. Estes dados são armazenados em um arquivo JSON (Notação de Objeto Javascript, na sigla em inglês), interpretados pelo DMI-TCAT e armazenados em seu banco de dados. Os itens de *media* são identificados como parte dos tuítes coletados (TWITTER, 2019c). Estes dados, armazenados pelo DMI-TCAT na tabela de terminação *_media*¹⁵⁰ embasam as análises realizadas nesta subseção.

Tabela 2: Dados dos tuítes com *media* e URLs de *media*.

Fonte dos dados: API Streaming do Twitter.

total de tuítes		7.265.147	100,00%	
tuítes com media		1.182.077	16,27%	100,00%
tuítes por tipo de media	photo	843.063	11,60%	71,32%
	animated_gif	105.395	1,45%	8,92%
	video	233.623	3,22%	19,76%
urls de media únicas		184.252	100,00%	
urls por tipo de media	photo	160.302	87,00%	
	animated_gif	13.721	7,45%	
	video	10.229	5,55%	

* A soma dos tuítes por tipo de *media* é superior ao número de tuítes com *media*, pois um tuíte pode conter mais de um tipo de *media*.

¹⁵⁰ Esta tabela é específica para cada coleta realizada no DMI-TCAT, de modo que a terminação é adicionada ao nome dado pelo usuário da ferramenta à coleta cadastrada. No caso, como dei à coleta o nome *juulgamento_lula*, a tabela recebeu o nome *juulgamento_lula_media*.

Estes 1,2 milhão de tuítes que contêm algum tipo de **media** compartilharam 184 mil URLs de **media** únicas. Ou seja, cada URL de **media** foi publicada, em média, 6,4 vezes por meio de retuítes ou compartilhamentos de um mesmo carregamento. Destas URLs de **media**, a ampla maioria, composta por cerca de 160 mil (87%), referencia arquivos do tipo **photo**. Estas URLs foram publicadas por cerca de 843 mil tuítes, com uma média de 5,3 tuítes por URL do tipo **photo**. Estes tuítes representam cerca de 12% de todos os tuítes coletados e cerca de 71% dos tuítes com **media** (Tabela 2).

Nesta pesquisa, o recorte da análise buscou excluir imagens em movimento (**animated_gif** e **video**) para focar apenas as imagens estáticas subsumidas pela categoria **photo**. O motivo principal para este recorte se devia a limitações do método empregado para lidar com imagens em movimento. Abarcá-las demandaria procedimentos mais complexos, além de muito mais recursos de armazenamento e processamento. Entretanto, devido a um aspecto que será discutido em maiores detalhes na subseção seguinte (4.1.4 URLs de links), uma imprecisão do protocolo de processamento fez com que ao menos uma parte dos conteúdos de imagem em movimento fosse inadvertidamente incluída no escopo da análise. Nesses casos, contudo, as imagens processadas não foram os vídeos em si e, sim, um quadro estático, extraído do vídeo, o qual é utilizado pela plataforma como forma de pré-visualização do conteúdo, antes do vídeo ser executado. Ou seja, parte dos objetos do tipo **video** e **animated_gif** foram incluídos na forma de imagens estáticas.

Considerando os dados obtidos pela tabela de terminação **_media** do DMI-TCAT, o último passo pertinente a esse tópico foi o carregamento local (*download*) das imagens indicadas pelas URLs de **media** (item 7 da Figura 15). Este carregamento foi realizado por meio de *script* na linguagem Python elaborado especificamente para este fim, que recursivamente percorreu a lista de URLs e armazenava as imagens em uma pasta local, em formato JPEG. Trata-se de processo relativamente simples, mas que tem um ponto importante para as tarefas subsequentes, relativo ao modo de nomeação dos arquivos. A relevância deste aspecto deve-se ao estabelecimento de um primeiro princípio de *identidade* para as imagens analisadas. Neste caso, *uma* imagem é compreendida como um *arquivo de imagem*. Como é a URL que identifica a individualidade destes arquivos até esta etapa, os arquivos foram nomeados de acordo com esta URL. Para este fim, o *script*

que utilizei para o carregamento aplicava o mesmo método utilizado por Bernhard Rieder no *script Memespector* (RIEDER; DEN TEX; MINTZ, 2018), baseado no algoritmo de criptografia SHA-1.

O algoritmo de criptografia SHA-1 é implementado como parte da biblioteca padrão do Python para funções de dispersão (*hash*). Usar as próprias URLs como nomes para os arquivos seria inviável, já que são tipicamente extensas, de comprimento variável e porque possuem caracteres especiais que não são permitidos em nomes de arquivos. Por meio deste algoritmo, porém, as URLs são codificadas como uma sequência finita de 40 caracteres alfanuméricos. Esta sequência comporta 2^{160} possibilidades e, portanto, tem baixíssima probabilidade de colisão (isto é, de que duas URLs distintas produzam a mesma sequência alfanumérica). Por exemplo:

`http://www.ppgcom.fafich.ufmg.br/ → 3f0faece295839635b0b7c1a110288f6f30223ce`

Deste modo, atribui-se um identificador único para o endereço e, portanto, para os arquivos de imagem. Isto permite garantir um bom nível de consistência durante o processamento dos dados, evitando a duplicação do carregamento da imagem e, ao mesmo tempo, distinguindo com alto grau de precisão as diferentes URLs de origem.

Diante do que foi apresentado e discutido nos capítulos precedentes, estas sequências de caracteres podem ser compreendidas como *inscrições* na medida em que, para diversas instâncias das tarefas de processamento, são elas que contam *como* as imagens. O identificador único opera como um elemento que garante que a vinculação da imagem à URL de origem seja sempre recuperável, uma vez que uma tabela vinculando cada *ID* às URLs é mantida como parte da base de dados. De fato, é por meio desta inscrição que os arquivos JPEG armazenados em um diretório único – isolados e autossuficientes enquanto tais – podem ser articulados aos dados armazenados nas tabelas de tuítes, por exemplo, com indicações quanto ao momento em que foram publicados e dados das contas que os compartilharam. Indicarei em seção adiante (veja [4.2 Processamento das imagens](#)) como esta inscrição será recuperada para identificar *matrizes* de conjuntos de imagens identificadas como referentes a um mesmo registro, embora sejam provenientes de diferentes URLs e, portanto, possuem diferentes sequências identificadoras.

4.1.4 URLs de *links*

Chamo URLs de *links* os endereços incluídos nos tuítes que não são pertinentes a arquivos de imagem estática ou em movimento (*media*) e que, de outro modo, remetem a outros conteúdos disponíveis na Web. A inclusão destas URLs na análise, como indiquei em subseção anterior, visa dar conta de um segundo modo pelo qual as imagens podem aparecer no Twitter (veja [Figura 17](#)). Quando um tuíte inclui uma URL, o Twitter busca no endereço a indicação de uma imagem de pré-visualização para incluir na exibição do tuíte. Para extrair estas imagens e considerá-las na análise, portanto, é preciso que o protocolo metodológico inclua uma etapa em que as URLs são acessadas e o dado indicativo da imagem de pré-visualização é extraído e a imagem é carregada localmente. Efetivamente, automatizei este processo por meio de um *script* em Python mas, ainda assim, trata-se de processo lento que não é viável realizar para a totalidade das URLs compartilhadas nos tuítes coletados. Nesta subseção, busco explicar, portanto, o processo de extração das URLs, a seleção daquelas que seriam efetivamente processadas e, enfim, o processo de carregamento das imagens. Indicarei como estes procedimentos são permeados de incertezas, muitas das quais fogem ao escopo desta pesquisa investigar a fundo, em especial porque não afetam substancialmente os objetivos perseguidos.

Um primeiro ponto por se abordar é a identificação das URLs nos tuítes. Esta tarefa foi realizada por dois procedimentos complementares. O primeiro procedimento é o método já integrado à própria API Streaming do Twitter. As URLs incluídas nos tuítes são discriminadas nos dados fornecidos, por meio da entidade `url`, compreendida como parte do objeto `tweet` (TWITTER, 2019c). O DMI-TCAT armazena estas URLs em tabela específica, identificada pela terminação `_urls`, que vincula cada tuíte à(s) URL(s) compartilhadas (item 3 da [Figura 15](#)). Como as URLs são retornadas em forma encurtada – isto é, indicando um endereço intermediário mais curto, para otimizar o uso de caracteres – a identificação do endereço final do *link* demanda que a URL encurtada seja *seguida* para obter o endereço completo. O DMI-TCAT também possui função para automaticamente expandir as URLs, armazenando já a versão completa delas. O segundo procedimento envolveu o processamento do conteúdo textual dos tuítes em busca de segmentos iniciados por ‘http’ e que, portanto poderiam indicar uma URL (item 4 da [Figura 15](#)). A adoção desse método se deveu a observações preliminares que sugeriam que nem todas as URLs incluídas nos tuítes eram efetivamente indicadas nos dados que o

DMI-TCAT obtém da API Streaming. Este processamento se deu por meio de script em Python, que também recuperava as URLs completas no caso de terem sido compartilhadas em forma encurtada.

A combinação dos métodos, embora tenha ampliado o escopo das URLs consideradas, também gerou alguns problemas para a pesquisa. Um primeiro problema foi a inclusão, como URL de *link*, de URLs internas ao Twitter que remetem, por exemplo, a publicações retuitadas com comentário. Estas URLs não são incluídas nos dados de URLs fornecidos pela API Streaming, mas o método de processamento do conteúdo textual, acaba por incluí-las, uma vez que não há como resolver, de antemão, o tipo de URL que foi encontrada. Este problema pode causar a duplicação da inclusão de uma imagem nos dados, embora, pelos procedimentos adotados em etapas posteriores, esta possível duplicação não gerar distorções na análise. A principal consequência deste problema foi, como antecipei na subseção anterior, a inclusão das imagens estáticas utilizadas como pré-visualização para imagens em movimento (`animated_gif` e `video`). Em alguns tipos de republicação permitidas pelo Twitter, as URLs destas imagens são inseridas como metadados das imagens de pré-visualização da página e, por isso, algumas URLs de `media` que inicialmente foram excluídas acabaram sendo incluídas nos dados.

Outro problema diz respeito aos momentos em que os processamentos foram realizados. O DMI-TCAT obtém as URLs encurtadas informadas pelo Twitter e logo as processa de modo a obter as URLs extensas. O processamento que realizei, contudo, dividiu-se em dois momentos. Um a poucas semanas do período de coleta e outro bem posterior (pouco mais de um ano após), de modo que houve uma maior incidência de erros causados por indisponibilidade dos endereços referenciados, já que algumas páginas poderiam já ter sido tiradas do ar. Há casos, também, em que a imagem de pré-visualização de uma página Web foi alterada em momento posterior a um tuíte que a compartilhou. Por esse motivo, por exemplo, o dispositivo analítico apresentado adiante (veja [4.3. Compondo imagens-redes](#)) indica ocorrências de uma imagem no Twitter que são, efetivamente, anteriores à tomada da fotografia. Isto ocorre para a fotografia de Francisco Proner Ramos, que tornou-se icônica do evento (□ VGG19 P-Q:15-16 | □ GVAPI A-B:1-2). Tuítes que foram publicados antes do acontecimento que propiciou o registro foram indicados como se contivessem a fotografia, mas efetivamente foi a página referenciada no tuíte

que, posteriormente, alterou a imagem de pré-visualização. As páginas da Web não são, efetivamente, documentos estáveis e podem mudar com o tempo, um aspecto que a pesquisa não tem como contornar.

Apesar destes problemas, a combinação de procedimentos permitiu uma substancial ampliação do número de URLs encontradas. A despeito da suspeita inicial de que o tratamento das URLs sem o “filtro” da API Streaming do Twitter poderia levar a uma supernotificação de URLs internas do Twitter que cumprem funções operacionais sem impacto no conteúdo, os dados obtidos pelo reprocessamento, quando comparados aos dados obtidos da API, sugerem cenário inverso. Em ambos os métodos empregados, o domínio do Twitter (`twitter.com`) é o que possui maior ocorrência entre as URLs. Nos dados obtidos pelo DMI-TCAT, via API Streaming, 73,7% das URLs estão nesse domínio. Nos obtidos por reprocessamento, são 63,6%. Quando consideramos apenas as URLs encontradas exclusivamente por um ou outro método, a diferença se explicita. Para o caso das URLs obtidas exclusivamente via API Streaming, 99,3% delas são do domínio do Twitter. Para aquelas obtidas via reprocessamento, são 85,9%. Por um lado, isto indica que as contribuições específicas do reprocessamento do conteúdo textual tinham mais URLs externas ao Twitter do que os dados obtidos via API. Isto, por si só, torna os resultados importantes no sentido de ampliação da abrangência da coleta. Por outro, ainda é possível que as URLs trazidas pelo método de reprocessamento tenham contribuído de forma substancial à supernotificação das URLs internas. Sem uma investigação mais aprofundada para verificar possíveis diferenças entre as URLs do Twitter notificadas por um e outro método, não é possível afirmar ao certo¹⁵¹. Trata-se, contudo, de questão que ultrapassa os objetivos desta pesquisa e que, portanto, optei por não perseguir. A combinação entre os métodos foi mantida, em todo caso, em função de um número substancial de URLs externas ao Twitter que foi possível obter apenas por meio do processamento do conteúdo textual dos tuítes.

Tanto a extração de URLs pela API, realizada pelo DMI-TCAT, quanto a realizada a partir do conteúdo verbal dos tuítes tiveram desempenho similar ao se considerar a comparação entre os números de URLs identificadas por cada um deles e o número total, obtido pela união entre os conjuntos. Cada um deles encontrou em torno de 70% do montante

151 Um indício de problema, em todo caso, é que a proporção de URLs relativas ao domínio do Twitter cresce quando vamos da interseção ou da união entre os métodos para a diferença entre eles. Isto sugere que cada método traz URLs internas à plataforma qualitativamente distintas, já que poucas destas são trazidas por ambos, em sua interseção.

total. Combinados, os métodos produziram um número total de cerca de 1,45 milhão de URLs encurtadas distintas. Destas, foi possível obter a forma completa de cerca de 93,14% (~1,35 milhão), obtendo, para as demais, falhas na requisição¹⁵². Considerando que mais de uma URL encurtada poderia se referir a uma mesma URL em formato completo, o número de URLs expandidas únicas é um pouco menor que o número de URLs encurtadas. Das cerca de 1,35 milhão que foi possível obter o formato expandido, chegou-se a cerca de 988 mil URLs expandidas únicas.

Tabela 3: Dados das URLs de links.

Fonte dos dados: API Streaming do Twitter, DMI-TCAT e processamento próprio.

total de urls encurtadas	1.452.021	
total de urls expandidas	988.448	100,00%
segundo API Streaming	726.112	73,46%
apenas segundo API Streaming	342.192	34,62%
segundo o reprocessamento	661.569	66,93%
apenas segundo o reprocessamento	262.336	26,54%
interseção	383.920	38,84%

Para os objetivos desta pesquisa, estas URLs servem como formas de acessar imagens que podem ter sido exibidas no Twitter, em tuítes que as compartilharam. Este acesso é possibilitado por protocolos estabelecidos pelo Twitter e pelo Facebook, adotados por desenvolvedores de páginas na Web. A imagem que é exibida depende da inserção de metadado pelo desenvolvedor da página referenciada. O Twitter reconhece padrões de marcação de metadados definidos por ele próprio (*Twitter Cards Markup*) e pelo Facebook (*Open Graph Markup*) (TWITTER, 2019d). Estas marcações definem o modo de inclusão de metadados por meio dos quais os conteúdos de páginas da Web tornam-se interpretáveis pelas plataformas de mídia social. O objetivo é melhorar o modo de exibição de *links*, quando compartilhados nas mídias sociais. Para tanto, desenvolvedores das páginas devem inserir no cabeçalho do documento *tags* HTML do tipo *meta* com informações como título da página, autor, data, brevíssimo resumo e URL da imagem que deve ser exibida como pré-visualização do conteúdo. Os padrões definidos pelo Twitter e pelo Fa-

¹⁵² A maior parte dos casos parecem se dever às páginas encontrarem-se indisponíveis. Mas uma parte também decorreu de má estruturação do endereço no conteúdo da mensagem. A expansão das URLs foi feita por meio de *script* em Python utilizando o módulo *urllib* para fazer requisições à Web pelas URLs encurtadas e então recuperando a URL final ao seguir o percurso até o ponto final. Outra parte majoritária das URLs foi expandida no processamento automático integrado ao DMI-TCAT.

cebook indicam como estas *tags meta* devem ser identificadas para que consigam ser interpretadas pelas plataformas. Que o Twitter também reconheça o padrão definido pelo Facebook sugere certa hierarquia, ainda que transitória, entre as plataformas. Efetivamente, observando os dados obtidos, é bem mais frequente que páginas insiram metadados direcionados ao Facebook do que ao Twitter, sugerindo uma maior capacidade daquela empresa, em comparação a essa última, em conseguir transladar os esforços de desenvolvimento de páginas a seus objetivos.

Nesta forma de padronização, encontramos um aspecto do que Anne Helmond (2015) descreve como o processo de “plataformização da *web*”: um processo de difusão das plataformas enquanto modo de organização do conteúdo, incidindo inclusive sobre o modo de organização dos dados em páginas da *web* externas às plataformas. Para tornar-se mais visível no Twitter ou no Facebook, produtores de conteúdo e desenvolvedores precisam estruturar os dados de forma que as plataformas consigam interpretá-los adequadamente e exibí-los de acordo com os requisitos de suas interfaces. Do mesmo modo, esta incidência externa que orienta a organização das páginas da *web* acaba por padronizar a massa difusa desses conteúdos, criando marcadores possíveis de acesso e mapeamento por parte de pesquisadores. As escolhas feitas quanto ao que exibir das páginas em plataformas *online* são relevantes para a pesquisa em métodos digitais compreender o modo de exibição e circulação desses conteúdos. Para os fins do presente estudo, esses metadados são utilizados como referência para se definir as imagens por analisar em relação à publicação de URLs externas.

Para extrair as imagens das páginas referenciadas pelas URLs, é necessário, portanto, acessar o conteúdo destas páginas e encontrar no seu código indicações feitas nas *tags meta* segundo os padrões estabelecidos pelo Twitter e pelo Facebook. Uma vez que isto demanda acessar e carregar o conteúdo de cada uma das páginas processadas, trata-se de processo muito dispendioso que, mesmo automatizado por um *script* em Python, como fiz, não poderia ser feito para a totalidade das URLs encontradas. Além do tempo que seria gasto neste acesso e carregamento de informações, o volume de dados seria excessivo. Por isso, métricas relativas ao número de referências a cada URL bem como ao número de seguidores das contas que compartilharam as URLs foram utilizadas para delimitar um escopo menor para a investigação.

A seleção das URLs observou, portanto, a distribuição de algumas das métricas relativas a cada endereço segundo os dados relativos aos tuítes que o compartilharam, fornecidos pela API Streaming. A Tabela 4 apresenta uma síntese da distribuição de cinco medidas realizadas a partir desses dados: o número de tuítes que compartilharam a URL; o número de contas de usuário que a compartilharam em seus tuítes; a soma dos seguidores destas contas; a média dos seguidores destas contas; e o número de “contas verificadas” pelo Twitter que compartilharam a URL. Contas verificadas são geralmente vinculadas a instituições ou pessoas públicas cuja autenticidade é verificada pela plataforma segundo sua identificação de que seria de “interesse público”¹⁵³ (TWITTER, 2019e). Uma parte significativa das URLs encontradas, segundo pode-se observar na tabela, possuem alcance extremamente reduzido e, por isso, poderiam ser desconsideradas sem prejuízo à abrangência dos dados em relação ao caso. Encontramos, por exemplo: que pelo menos 75% das URLs foram publicadas apenas uma vez e por apenas uma conta de usuário; para 50% delas o público potencial era de menos de mil seguidores; e para menos de 10% delas houve compartilhamento por “contas verificadas” pelo Twitter.

Tabela 4: Estatísticas descritivas das publicações de URLs de links.

Fonte dos dados: API Streaming do Twitter.

	núm tuítes	núm contas	soma seguidores contas	média seguidores contas	núm 'contas verif.'
média	5,32	5,22	68.830,32	25.577,65	0,06
desvio	70,80	70,50	539.857,13	274.890,57	0,29
mín	1	1	0	0,00	0
10%	1	1	30	30,00	0
25%	1	1	168	164,00	0
50%	1	1	924	778,00	0
75%	1	1	5.920	2.964,00	0
90%	4	4	38.499	12.917,00	0
máx	16.505	16.477	53.073.328	41.683.637,00	53

Com o objetivo principal de viabilizar a análise, portanto, as URLs foram selecionadas por um conjunto de critérios (item 5 da Figura 15). Especificamente, segundo o atendimento de *ao menos uma* das três primeiras condições abaixo e, *necessariamente*, a quarta condição:

1. que a URL tivesse sido publicada por ao menos 3 contas – 16,20% das URLs;

¹⁵³ O Twitter (2019e) explica-se assim: “Normalmente, verificamos contas de usuários nas áreas de música, teatro/cinema/TV, moda, governo, política, religião, jornalismo, mídia, esportes e negócios, entre outras.”

2. ou que a soma de seguidores das contas que a publicaram fosse igual ou superior a 20.000 – 14,94% das URLs;
3. ou que tivesse sido publicada por ao menos 1 ‘conta verificada’ – 3,31% das URLs;
4. e, satisfeita ao menos uma das anteriores, que, necessariamente, a soma de seguidores das contas que publicaram fosse igual ou superior a 100 – 79,81% das URLs e 99,65% do grupo das condições anteriores.

A seleção pela soma de seguidores das contas se justifica pela consideração do alcance potencial da publicação. Mesmo que apenas uma conta tivesse feito a publicação, se ela possuísse muitos seguidores, seria um fator que confere relevância à publicação. Similamente, uma URL publicada por muitas contas com poucos seguidores, cada uma, porém alcançando um grande público potencial, também teria relevância¹⁵⁴. A seleção com base na URL ter sido publicada por uma ‘conta verificada’ justifica-se pelo reconhecimento institucional por parte da plataforma de que aquela conta é, ao menos segundo esse critério, relevante, mesmo que não tenha muitos seguidores. A seleção pelo número de contas que publicaram a URL se justifica por ser um indicador da relevância geral da URL para o universo estudado, independentemente do nível de visibilidade da conta que a publicou. Por fim, a condição eliminatória, considerando a soma de seguidores mínima de 100, visa evitar considerar, na análise, URLs com visibilidade mínima ou mesmo nula que poderiam ter sido incluídas, excepcionalmente, apesar dos demais critérios.

A combinação destes fatores visou alcançar um recorte que atendesse aos diferentes critérios estipulados para a análise e que, ao mesmo tempo, tornasse a análise minimamente viável. Os limites para cada fator foram atribuídos experimentalmente a partir da observação da distribuição destes atributos ao longo dos dados, visando delimitar uma amostra de em torno de 20% da lista completa de URLs encontradas nos dados. Chegou-se, assim, ao número de 234.184 URLs, representando 23,69% da lista completa de URLs. A distribuição das mesmas métricas, agora para o subconjunto selecionado, é apresentada na [Tabela 5](#).

¹⁵⁴ Esta estratégia não chega a excluir URLs compartilhadas com contas com atividade automatizada (chamadas de ‘robôs’ ou *bots*), mas este tampouco é o objetivo. Em certo sentido, se as publicações impulsionadas por esse tipo de recurso alcançaram um grande público potencial, elas passam a ter relevância em uma compreensão da configuração midiática do evento. Não faria sentido purificar os dados por esse critério se não é meu objetivo elaborar asserções, por exemplo, quanto à opinião pública manifesta no Twitter.

Tabela 5: Estatísticas descritivas das publicações com URLs de link selecionadas.

Fonte dos dados: Twitter Streaming API.

	núm tuítes	núm contas	soma seguidores contas	média seguidores contas	núm 'contas verif.'
média	20,64	20,28	312.915,31	112.854,76	0,26
desvio	151,51	150,89	1.129.662,88	583.807,01	0,59
mín	1	1	101	17,50	0
10%	1	1	3.904	883,69	0
25%	2	1	11.357	1.862,00	0
50%	4	4	33.280	5.034,40	0
75%	10	9	129.255	38.416,00	0
90%	30	29	590.681	175.075,70	1
máx	16.505	16.477	53.073.328	41.683.637,00	53

O passo seguinte consistiu, então, em acessar cada uma destas URLs selecionadas a fim de extrair as imagens que elas indicavam, por meio da *tag meta* e dos padrões aceitos pelo Twitter, como pré-visualização de seu conteúdo (item 6 da Figura 15). Como indiquei anteriormente, o Twitter reconhece dois padrões para esse metadado, um que ele mesmo estabelece (*Twitter Card Markup*) e outro estabelecido pelo Facebook (*Open Graph Markup*). Para o primeiro caso, a *tag meta* deve ter um formato similar ao indicado abaixo, retirado do site da UFMG:

```
<meta name="twitter:image" content="http://ufmg.br/assets/img/ufmg-social.jpg" />
```

O atributo *name* indica como o metadado deve ser interpretado – como imagem de pré-visualização no Twitter. O atributo *content* indica a URL da imagem que deve ser exibida. Para o caso do *Open Graph Markup*, o atributo que indica a interpretação é o *property* e sua forma seria similar à seguinte:

```
<meta property="og:image" content="http://ufmg.br/assets/img/ufmg-social.jpg" />
```

Elaborei um *script* em Python para esta finalidade específica, portanto: recuperar o código HTML de cada uma das páginas referenciadas pelas URLs e nele buscar *tags* como estas. A esta técnica de extrair informações do documento de uma página Web, gerando dados estruturados (como na forma de tabela), dá-se o nome de ‘raspagem’ (*scraping*). As URLs indicadas como conteúdo nas *tags* identificadas foram, então, utilizadas para

carregar as imagens pertinentes a cada página. Evidentemente, contudo, nem todas as 234 mil URLs estavam acessíveis durante este processo e, das que estavam acessíveis, nem todas tinham imagens indicadas para pré-visualização. Especificamente, 3.662 URLs não puderam ser acessadas – cerca de 1,6%. Das que puderam ser acessadas, 28.016 (cerca de 12%) não possuíam imagem de pré-visualização configurada ou ela não pôde ser carregada. De 234 mil URLs, portanto, apenas cerca de 202,5 mil serviram como fontes para as imagens.

Em mais uma redução, contudo, as imagens usadas como pré-visualização destas 202,5 mil páginas da Web por vezes se repetiam em múltiplas páginas. Posto de outro modo, na configuração da *tag meta*, múltiplas páginas indicavam uma mesma URL para o arquivo de imagem de pré-visualização. Considerando apenas as URLs únicas encontradas para estas imagens, portanto, chega-se ao número exato de 129.278 URLs de arquivos de imagens extraídos das URLs referenciadas nos tuítes e selecionados para esse processamento. Assim como o que foi indicado na subseção anterior (veja descrição na p. 208), acerca das URLs de *media*, as URLs dos arquivos de imagem que foram obtidas por meio deste processo foram também a base para gerar identificadores das imagens obtidas, por meio do algoritmo de criptografia SHA1 (item 7 da [Figura 15](#)).

4.2 Processamento das imagens

A seção anterior voltou-se à descrição dos procedimentos que foram adotados para constituir a base de imagens a que se volta a análise. Como representado no procolo daquela primeira fase ([Figura 15](#)), o resultado obtido é um conjunto de arquivos de imagem, identificados por sequências alfanuméricas únicas segundo suas URLs de origem. Também foi gerada, no mesmo processo, uma tabela, isto é, um conjunto de dados estruturados em que estas imagens são relacionadas aos tuítes que as publicaram e seus respectivos metadados. Efetivamente, embora se trate de processo complexo e já permeado de incertezas, trata-se apenas de uma etapa preliminar ao estudo de caso pretendido. Chegamos, enfim, a um primeiro conjunto de arquivos de imagens – ou dados visuais, se assim se preferir – com os quais nos haver.

Na seção anterior, alguns esforços específicos de um estudo voltado às imagens já foram realizados quando, por exemplo, expandi o escopo da coleta para além dos dados diretamente fornecidos pelo Twitter para ir também às páginas referenciadas nos tuítes para

delas extrair as imagens de pré-visualização. Porém, é nesta etapa do processo que passo a me engajar com questões mais diretamente específicas às imagens e os desafios que colocam à investigação em métodos digitais. Gabriele Colombo (2019) situa na “pasta de imagens” a primeira instância desse desafio. Sua proposta, vinculada ao grupo *Visual Methodologies*, de Sabine Niederer, volta-se à demanda de estudar imagens em grupos e não apenas individualmente. Trata-se, claro, de um dos objetivos também compartilhado por esta investigação. Ao se referir à “pasta de imagens” Colombo parece aludir a uma entidade mais ou menos conhecida de pesquisadores de métodos digitais. Por um lado, APIs e técnicas de raspagem oferecem saídas estruturadas e “prontas para usar” de dados extraídos de plataformas *online*. Pelas tabelas, pode-se filtrar os dados, calcular estatísticas descritivas e observar aspectos pertinentes a cada publicação. A “pasta de imagens”, por outro lado, pode se apresentar como um estorvo.

Neste estudo de caso, tratou-se inicialmente de uma pasta com 279.490 imagens. Este é o conjunto de união das imagens obtidas por cada um dos dois procedimentos descritos na seção anterior – provenientes, portanto, das URLs de *media* do tipo *photo* (160.302) e das imagens de pré-visualização das URLs de *links* (129.278)¹⁵⁵. Alguma redução, em relação ao escopo completo dos dados coletados, já foi realizada quando, na fase anterior, as URLs de *link* foram selecionadas segundo as métricas relativas às contas que as publicaram. Nesta fase, contudo, novas reduções buscarão protocolos que, em alguma medida, não ignorem o “conteúdo” dos arquivos de imagem. Quero dizer que na seção anterior, as imagens foram tratadas, efetivamente, como URLs. Não falava de *imagens únicas* mas de *arquivos de imagens únicos*, pois provenientes de endereços específicos na Web. Nesta fase, um primeiro objetivo será encontrar, entre estes múltiplos arquivos, imagens “únicas” que neles se repetem. Esta tarefa demandará, então, um primeiro movimento de produzir *inscrições* computacionalmente comparáveis acerca destas imagens – um modo particular de realizar, para as imagens digitais, as operações de recombinação e sobreposição que Latour (1986, 2001) sugere para as inscrições científicas. Este é o primeiro desafio. A partir da identificação destas imagens “únicas” será possível, então, efetuar uma redução não mais do número de arquivos de imagens considerado mas, sim,

155 Descartaram-se alguns arquivos no processo devido à sua indisponibilidade *online* ou ao corrompimento dos arquivos que foi possível carregar. Sobre este último caso, cerca de 5 mil arquivos carregados estavam corrompidos.

do número de agrupamentos de imagens considerado – estes concernindo grandes conjuntos de arquivos que provavelmente se vinculam a um mesmo registro. Implicações e limitações deste processo serão elaborados adiante.

Em um segundo momento, volto-me à aplicação de dois modelos distintos de reconhecimento de imagens por redes neurais a fim de organizar estes agrupamentos de imagens. Um dos objetivos é conseguir traçar observações gerais acerca das imagens postas em circulação no contexto da prisão do Lula – isto é, elaborar descrições distantes das imagens, tomando-as como um “todo”. Outro objetivo é oferecer condições para nos orientarmos em meio a estas imagens em um esforço de observá-las em conjunto sem reduzi-las a métricas ou a “dados visuais”. Observá-los, portanto, como algo próximo de uma materialidade relacional destas figurações, considerando como os registros fazem reverberar uns aos outros e produzem, deste modo, uma visualidade para o acontecimento, que não pode se reduzir a nenhuma imagem individualmente, nem se diluir em métricas agregadas do “todo”. Avançarei nos modos de análise proporcionados por este esforço na seção posterior ([4.3.Compondo imagens-redes](#)).

Retomando questões discutidas nos capítulos anteriores, esta seção se volta, em outros termos, a um esforço de performar ontologias múltiplas para as imagens coletadas, ao produzir inscrições com propriedades distintas e que permitem, assim, operações de agrupamento, comparação e disposição em um “espaço de pensamento” como os que buscava constituir, em outro contexto, Warburg. As etapas descritas a seguir operam, portanto, como etapas de uma cadeia de translações heterogêneas a que as imagens são submetidas e que busco, ao final, situar em um plano comum, em um esforço ontográfico de *composição*.

4.2.1 Identificando imagens “únicas”

Como descrito anteriormente, a *identidade* dos arquivos de imagens coletados foi estabelecida, inicialmente, com base nas URLs utilizadas para realizar seu carregamento (*download*). Que dois arquivos sejam provenientes de URLs diferentes não implica, contudo, que as imagens sejam, necessariamente, diferentes. A única afirmativa possível é a de que se tratam apenas arquivos diferentes, mas que podem ser instanciações de uma mesma fotografia, por exemplo, seja em arquivos informacionalmente idênticos (embora carregados em endereços diferentes), seja em arquivos similares em que a imagem

pode ter sido submetida a pequenas alterações (corte, ajuste de cores, sobreposição de texto etc.). Contudo, como esta explanação já começa a sugerir, mesmo considerando o ‘conteúdo’ dos arquivos, não é simples estabelecer uma definição inequívoca do que contaria como *identidade* ou *diferença* entre dois arquivos.

Evidentemente, seria possível adotar um critério estrito e elevar o nível de exigência a uma identidade *absoluta* entre os arquivos – *bit a bit*. Este provavelmente seria o único modo de estabelecer esta identidade de modo definitivo, porém ele pouco contribui a um olhar que não se prenda ao arquivo de imagem apenas como *dado*. A identidade informacional absoluta, quero dizer, não possui lastro na nossa experiência das imagens – em como nós as *vemos* – e, se quero constituir uma abordagem do tema por uma perspectiva dos estudos de comunicação e mídia, o critério da identidade absoluta não apresenta grandes contribuições. Mesmo em uma perspectiva informacional ou computacional, a identidade absoluta seria algo raro já que na maior parte dos contextos reais de publicação e circulação das imagens, os arquivos de imagem raramente são apenas copiados, tal e qual. Qualquer forma de compressão informacional ou redução no tamanho da imagem (em *pixels*) levaria a uma transformação dos dados que impediria a correspondência absoluta entre as cópias.

Para além desta abordagem informacional, contudo, os limites daquilo que conta como *uma mesma* imagem não são facilmente definidos quando se considera, por exemplo, que versões de um mesmo registro fotográfico – grosso modo, de um mesmo *clique* – tendem a se multiplicar, contemporaneamente. Não apenas pela produção de *memes*, estas formações culturais contemporâneas que se definem, justamente, como variações em torno de um mesmo tema (SHIFMAN, 2013) ou, especificamente, de uma mesma fotografia (SHIFMAN, 2014). As variações também se multiplicam pela adição de filtros, por diferentes recortes efetuados sobre a mesma imagem, ou por compartilhamentos indiretos que, por exemplo, republicam uma impressão de tela da imagem e não o arquivo de imagem em si. Para um estudo interessado no impacto de uma fotografia específica, ainda que as peculiaridades das variações devam ser consideradas, a relação entre estas muitas variações também deve ser, de algum modo, mantida. Entre identidade e diferença, portanto, identifica-se mais um *contínuo* do que uma *fronteira*. Trata-se de caso em que o princípio de simetria entre *identidade* e *diferença*, sugerido por Michael Lynch (2013), torna-se fundamental. Indo além da proveniência dos arquivos de imagem, das URLs que foram codificadas em sequências de 40 caracteres alfanuméricos

pelo algoritmo SHA1, faz-se necessário buscar um outro referencial para agrupar os registros, em que a identidade seria estabelecida por uma margem de *similaridade* e que, portanto, não seria *absoluta* mas *probabilística*.

Para este fim, utilizei uma técnica de identificação de imagens baseada em técnica chamada de *Hash Perceptivo (Perceptual Hashing, ou pHash)*¹⁵⁶. Como no caso do algoritmo SHA1, utilizado para gerar os identificadores dos arquivos de imagem com base nas suas URLs, trata-se de uma função *hash*, que visa codificar dados a um formato de comprimento predefinido (a que se chama *hash*). Porém, a técnica de *pHash* (KLINGER; STARKWEATHER, 2010) realiza esta codificação de tal forma que dados visualmente similares produzam *hashes* também similares, de modo que as informações que deram origem aos *hashes* possam ser comparadas por meio destes. Obtém-se, assim, um método que é computacionalmente bem mais eficiente para tarefas de comparação de grandes conjuntos de dados. Funções *hash* típicas, devido a seu objetivo de produzir identificadores únicos e com baixa probabilidade de colisão, costumam produzir um efeito de cadeia para um único *bit* diferente, que faz com que *hashes* produzidos a partir dos dados de duas imagens similares acabem sendo muito distintos. O *Hash Perceptivo*, de outro modo, evita estes efeitos e busca produzir identificadores que guardem algum grau de correspondência com as qualidades *aparentes* dos dados.

¹⁵⁶ Seria possível utilizar os próprios modelos de reconhecimento de imagem por redes neurais para realizar esta identificação. Mas, tipicamente, trata-se de processo computacionalmente mais pesado. A técnica do *pHash*, via o módulo *Image-Match* é bastante eficiente na realização destas operações.

(1) Extração de identificadores dos conteúdos das imagens por meio da biblioteca de Python Image-Match (Edjo Labs et al., 2018), baseada na técnica pHash (Klinger e Starkweather, 2010).

(2) Armazenamento dos identificadores em banco de dados Elastic Search, segundo implementação nativa do Image-Match.

(3) Utilizando o Image-Match, busca reversa de cada identificador das imagens, por outros registros similares.

(4) A partir dos agrupamentos formados, identificar a imagem de maiores dimensões do grupo para usar como 'matriz'. Os demais arquivos são tomados como 'instâncias' dos agrupamentos.

(5) Registro dos dados gerados na tabela de imagens, identificando o grupo a que cada uma pertence.

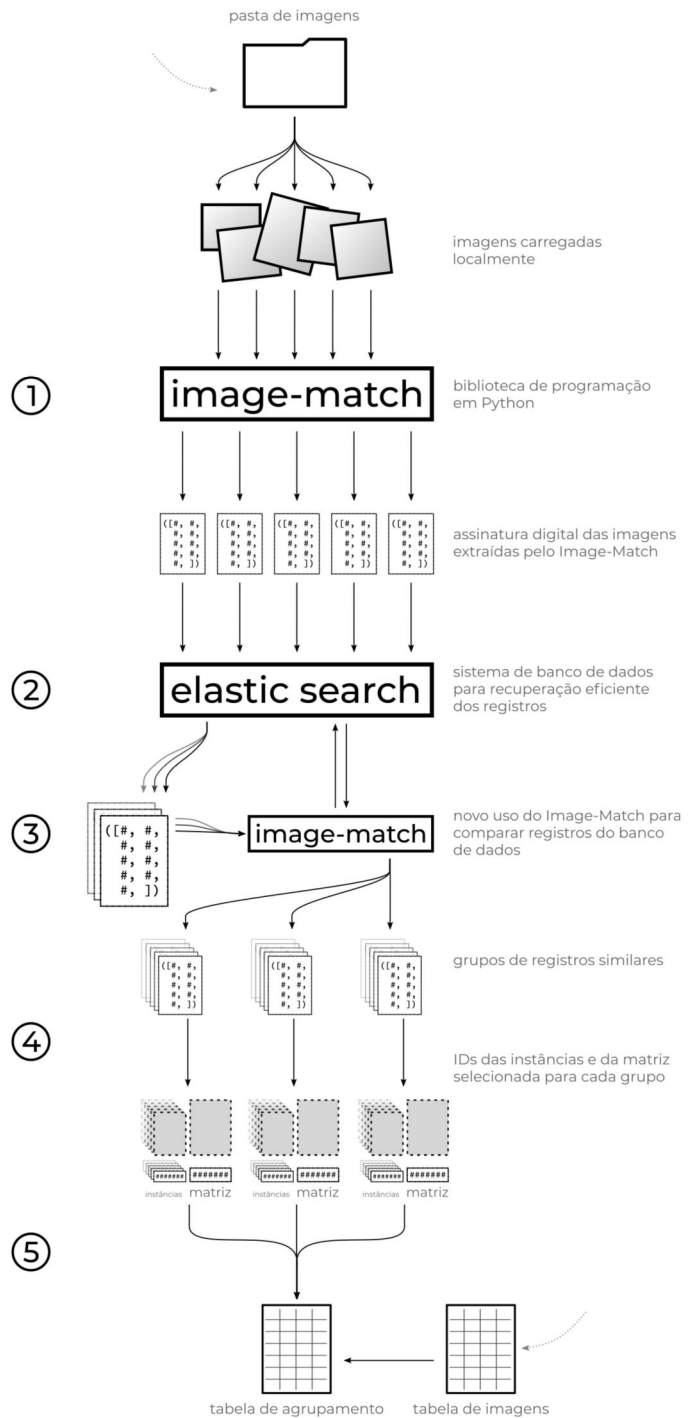


Figura 18: Protocolo de agrupamento das instâncias de imagens 'únicas'.

Fonte: Elaboração própria.

Especificamente, o procedimento adotado neste estudo utilizou a biblioteca de programação para a linguagem Python chamada Image-Match (EDJO LABS et al., 2018). Trata-se de uma implementação que se assemelha à técnica *pHash* e que, além de gerar estas “assinaturas” ou “impressões digitais” das imagens processadas, também possui funções de integração com sistemas de bancos de dados para permitir realizar CBIR com alta eficiência. Gerar o *hash* para as imagens seria, afinal, apenas a primeira operação à qual devem seguir esforços de comparação entre as imagens a fim de estabelecer agrupamentos de imagens *provavelmente* idênticas.

A [Figura 18](#) apresenta uma representação esquemática dos procedimentos adotados nesta etapa. Um *script* em Python elaborado especificamente para este protocolo automatizou o conjunto de processos descritos. Partindo da pasta de imagens obtida pelos procedimentos de coleta e processamento descritos na seção anterior, as etapas foram as seguintes. Primeiro, recursivamente, cada um dos 279,5 mil arquivos de imagem da pasta foi submetido à função de *hashing* perceptivo implementada pelo Image-Match (item 1 da [Figura 18](#)). Por este processamento, os dados que compõem cada imagem são codificados de modo a produzir uma inscrição resumida de seu conteúdo, que pode ser comparada a outras inscrições similares. Cada uma destas inscrições foi, então, armazenada em um banco de dados no mecanismo *ElasticSearch* (ELASTICSEARCH B.V., 2019) (item 2 da [Figura 18](#)), utilizando funções de integração do próprio Image-Match. Esse mecanismo de banco de dados permite maior eficiência na tarefa de comparação entre as inscrições.

O próximo passo envolveu a comparação de cada inscrição, isto é, de cada *hash* perceptivo das imagens coletadas, às de todas as demais (item 3 da [Figura 18](#)). Efetivamente, esta comparação se deu por um procedimento de busca em que, para cada inscrição, buscou-se outras inscrições similares. A similaridade é estabelecida, nesse caso, por uma margem de tolerância para diferenças. A diferença, indicada como *distance*, é computada pelo Image-Match como um índice normalizado entre 0 e 1. Após testar diferentes limites, optei por manter aquele indicado como referência pelos desenvolvedores: 0,4. Isto significa que quando a comparação entre os *hashes* de dois arquivos de imagem gera um valor de *distance* menor do que 0,4, estes dois arquivos de imagem são considerados, para todos os efeitos, instâncias da “mesma” imagem.

Gostaria de salientar este ponto pois, na sequência do que elaborei nos capítulos anteriores, ele estabelece um momento claro, no protocolo metodológico adotado neste estudo, em que a identidade de uma imagem é performada pelas práticas adotadas em seu processamento. Como abordei há pouco, a questão da identidade e diferença de uma imagem não tem como ser resolvida de forma definitiva. Isto significa que a definição do que conta como *uma mesma* imagem não tem como ser estabelecida *a priori*, sem atentarmos às práticas que produzem esta identidade como um *efeito*. Se esta discussão, elaborada no capítulo 2 (Imagem-rede) possa ter soado como um exercício teórico abstrato, aqui ela ganha concretude. Pois, como discutirei adiante, esta definição do que conta como *uma mesma* imagem não resolve completamente o problema quando passamos destas inscrições informacionais à aparência destas imagens. Se nos fosse solicitado encontrar tais correspondências, certamente chegaríamos a agrupamentos diferentes. Mas, diante desta escala de centenas de milhares de arquivos em uma pasta de imagens, esta não é uma alternativa viável.

É importante observar, a esse respeito, que o uso destinado a esta técnica se volta especialmente ao monitoramento de violação de direitos autorais (EDJO LABS et al., 2018)¹⁵⁷. Ao aplicar esta técnica para agrupar instanciações que podem ser tomadas como *uma mesma* imagem, realizo uma reapropriação que, efetivamente, desvia a finalidade da técnica com objetivos analíticos. Não é necessário que nos satisfaçamos plenamente com os agrupamentos feitos mas, diante do volume de imagens abarcado, alguma medida similar se faz necessária para ir dos arquivos de imagem a agrupamentos *possíveis* que conectem estas diferentes instâncias e reduza a redundância entre os arquivos. Entretanto, quero constituir uma abordagem que não toma os resultados desta operação como uma *realidade objetiva* dos objetos analisados mas, sim, de outro modo, como uma construção performativa instável, porém valiosa por nos permitir avançar na observação do caso sob estudo. Por mais esse motivo, portanto, os objetos (as imagens) não podem ter uma ontologia estável como poderíamos desejar. Ao definir sua identidade por um critério computacional e probabilístico, há agrupamentos que contêm arquivos de imagem que claramente, sob outro critério, não estariam ali. Também, os diferentes registros que talvez contássemos como *uma* imagem poderão ser encontrados de forma dispersa.

157 Técnicas similares são também utilizadas para rastrear a circulação de imagens específicas em demandas como o combate à pedofilia e outras práticas criminosas associadas a imagens. Pouco antes da escrita desta tese, o Facebook divulgou que colocaria sob código aberto a tecnologia utilizada pela empresa neste tipo de tarefa (DAVIS; ROSEN, 2019). Para estas aplicações, evidentemente, as correspondências calculadas computacionalmente não podem ser tomadas de forma definitiva, sendo passíveis de revisão caso a caso, embora provavelmente sejam fator determinante em medidas automáticas como a retirada de conteúdo do ar.

Em subseção adiante (veja [4.3. Compondo imagens-redes](#)) indicarei alguns exemplos e discutirei implicações deste modo de constituição da identidade das imagens para a análise.

Dos 279,5 mil arquivos de imagem coletados, chegou-se, inicialmente, a 99,7 mil agrupamentos. Para cada grupo, o arquivo de imagem com maior tamanho (em número de *pixels*) foi selecionado como *matriz* – como o arquivo de imagem que seria utilizado para representar as múltiplas *instâncias* que compõem o agrupamento (item 4 da [Figura 18](#)). A escolha da maior imagem visa, por um lado, um arquivo que potencialmente possui melhor qualidade e que, portanto, poderia servir melhor à exibição em detalhe. Por outro, embora o critério seja impreciso, supõe-se a possibilidade de chegar à versão mais próxima de um registro “original” – já que o tamanho da imagem pode ser sempre reduzido, mas uma vez que é isto é feito há perdas definitivas de informação que não podem ser recuperadas. Desse modo, quando disser da *matriz* de um agrupamento, refiro-me ao arquivo de maiores dimensões que o compõe. Todos os arquivos serão chamados de *instâncias*. O código de identificação da imagem *matriz* foi utilizado, por sua vez, como código de identificação do grupo que ela representa.

4.2.2 Medidas de privacidade

O grupo funciona como uma forma de tratamento agregado dos dados, já que diferentes arquivos de imagem passam a ser tratados em conjunto, representados por uma *matriz*. Neste processo, instâncias individuais de compartilhamento são subsumidas. Esta operação já contribui, portanto, a uma anonimização destas instâncias e, portanto, à proteção da privacidade dos usuários que compartilharam a imagem. Embora todos os tuítes coletados sejam provenientes de contas de usuário que mantêm suas publicações publicamente acessíveis, pois apenas estas são retornadas pela API Streaming do Twitter, é razoável compreender que nem todos usuários têm clareza do amplo acesso de suas publicações ou da possibilidade de que sejam posteriormente recuperadas em um esforço de pesquisa como este. Portanto, medidas que protejam a identidade destes usuários são consideradas como uma preocupação ética desta pesquisa.

Se, por um lado, o agrupamento já favoreceria esta anonimização, alguns cuidados específicos com o agrupamento foram também adotados, de modo a minimizar a possibilidade de exposição inadvertida de publicações que pudessem ter a intenção de

ser privadas. Este risco existe em, pelo menos, duas hipóteses, considerando os procedimentos descritos na subseção anterior. A primeira hipótese é a de que grupos inteiros, que tenham sido formados de forma eficaz, pudessem ser constituídos por imagens de cunho privado. Nesse caso, o grupo como um todo deveria ser excluído do processo de análise. A segunda hipótese é a de que, dado o caráter probabilístico com que a técnica de *hashing* perceptivo define relações de identidade e diferença, imagens bastante díspares fossem integradas em um mesmo agrupamento e, com isto, uma imagem de cunho privado fosse incluída como instância em um agrupamento de imagens de caráter público. Sobre esta hipótese, a inclusão dos dados desta imagem em representações agregadas do grupo introduz uma possível imprecisão aos dados, mas não fere, por si só, a privacidade de quem publicou. Pois a publicação em si ou a conta de usuário apenas incidiriam quantitativamente sobre os dados finais, sem serem necessariamente identificadas ou exibidas individualmente. Haveria problema, contudo, se esta imagem ou o tuíte que a compartilhou fossem exibidos em alguma etapa ou resultado da análise, seja como matriz do grupo, seja como uma de suas instâncias.

Como forma de evitar os casos supostos pela primeira hipótese, uma primeira medida adotada foi filtrar os dados de modo a excluir os arquivos de imagem que eram identificadas, por sua URL, como imagens de perfil de contas de usuário do Twitter. De fato, foi verificado que uma boa parte das URLs de imagens assumiam a estrutura típica destes arquivos, iniciadas com:

`https://pbs.twimg.com/profile_images/...`

Com base nesta estrutura padrão, foram excluídas todas as instâncias de grupos cuja URL de origem assumia este formato. A exclusão destes arquivos ocasionou, também, a exclusão de 24 mil grupos de ‘imagens únicas’, reduzindo o número final a 77,3 mil.

Uma segunda medida buscou estabelecer quais seriam os critérios razoáveis para que um tuíte e, portanto, a imagem que ele contém, fossem considerados passíveis de inclusão na análise. Ou seja, a questão era dos critérios para supor uma expectativa de publicidade de um tuíte por parte de seu autor. Dois critérios foram adotados, a esse respeito. Primeiro, se a conta de usuário responsável pela publicação fosse uma ‘conta verificada’ pelo Twitter (TWITTER, 2019e). Nestes casos, parece bastante claro que o autor tem plena consciência da publicidade de seus tuítes, até porque ter sua conta verificada depende

de uma demanda ativa do proprietário da conta. Segundo, se a conta de usuário possuía mais de 5 mil seguidores – 4,4% das contas que publicaram as imagens consideradas. Nestes casos seria também razoável supor que a pessoa responsável pela conta também possuiria consciência da publicidade de seus tuítes.

A aplicação desta medida se deu, então, pela filtragem dos tuítes e arquivos de imagem segundo estes critérios. Grupos de imagens únicas apenas poderiam ser considerados para a análise se ao menos uma de suas instâncias atendesse aos critérios estabelecidos. Também, instâncias que não atendessem aos critérios não poderiam servir como matriz do grupo, nem poderiam ser incluídas em formas individualizadas de exibição – seja a imagem, seja o texto do tuíte ou os dados sobre a publicação e sobre a conta que publicou. Tuítes de contas que não se enquadravam em nenhum dos dois critérios integram os dados apenas em forma agregada, isto é: na produção de métricas e representações dos dados em que nenhum elemento identificador da publicação ou da conta são exibidos – tais como nome de usuário, conteúdo verbal do tuíte e o arquivo de imagem (se este fosse publicado exclusivamente por contas que não atenderam aos critérios estabelecidos).

A aplicação destes critérios foi realizada após o processo de seleção de quais agrupamentos, dentre os 77,3 mil obtidos, seriam efetivamente considerados na análise. Esta inversão ocorreu, na prática da pesquisa, de forma contingente. Pois os riscos à privacidade de usuários apenas foi identificado após tal seleção. Contudo, esta ordem de operações foi mantida ao se perceber que foram pouquíssimos grupos de imagens únicas os que seriam inteiramente comprometidos pelos critérios adotados. Isto é: os critérios de seleção das imagens únicas – descrito na próxima subseção (4.2.3 Seleção das ‘imagens únicas’) – possuíam um alto grau de correlação com os critérios de privacidade estabelecidos. Dos 18,5 mil agrupamentos selecionados, segundo critérios que apresento adiante, apenas três eram compostos inteiramente por publicações que não atendiam a nenhum dos dois critérios elaborados acima. Diante do número reduzido de casos, foi possível considerar as imagens e publicações individualmente a fim de analisar, contextualmente, se sua inclusão das suas imagens poderia configurar uma exposição indevida. Para dois dos casos, as imagens exibiam políticos, portanto pessoas públicas para as quais as preocupações relatadas aqui não se aplicariam. O terceiro caso, por sua vez, trazia uma fotografia que não retratava nenhuma pessoa ou ambiente sensível, considerando implicações de privacidade, assemelhando-se a uma figura extraída de

banco de imagens. Portanto, optou-se por não excluir nenhuma destas três imagens. Contudo, os tuítes que as publicaram, bem como os dados passíveis de sua identificação, foram preservados de toda situação de exibição individualizada.

Além destes casos, houve outros em que as matrizes dos grupos de imagens únicas foram substituídas a fim de não exibir arquivos de imagem provenientes de tuítes que não atendem aos critérios estabelecidos. Dos 18,5 mil grupos de imagens selecionados, 2,1 mil tiveram suas matrizes substituídas por arquivos de imagem integrantes dos mesmos grupos, mas cujas publicações adequavam-se aos critérios estabelecidos. Segundo verificação manual em uma amostra de 10% destas substituições, observou-se que, quase sempre, a troca se deu entre arquivos aparentemente idênticos. Entretanto, considerando o caráter probabilístico da técnica de *hashing* perceptivo adotado, nem sempre foi este o caso. Nestes casos, o próprio grupo era heterogêneo.

4.2.3 Seleção das ‘imagens únicas’

Um passo posterior envolveu a seleção de um subconjunto destes agrupamentos para considerar na análise. De modo similar à seleção das URLs de *links* que seriam submetidas ao processo de raspagem para extração das imagens de pré-visualização, a seleção dos agrupamentos de imagem baseou-se na distribuição das métricas relativas aos tuítes que continham instâncias do agrupamento. Efetuar um novo recorte mostrava-se necessário para chegar a um número viável para proceder à análise e, também, evitar trabalhar com imagens de circulação minoritária. Diferentemente do recorte efetuado pelas URLs, contudo, nesta etapa a seleção tomava como unidade da análise os grupos de arquivos de imagens que foram considerados, nesta etapa, como *imagens únicas*. A [Tabela 6](#) apresenta a síntese destes dados.

Tabela 6: Estatísticas descritivas de publicação das ‘imagens únicas’.

Fonte: API Streaming do Twitter.

	núm tuítes	núm contas	soma seguidores contas	média seguidores contas	núm 'contas verif.'
média	41,75	38,26	468.713,72	25.716,26	0,42
desvio	405,46	309,21	2.607.918,55	125.621,24	2,25
mín	1	1	0	0,00	0
10%	1	1	118	104,00	0
25%	1	1	845	545,50	0
50%	3	3	10.929	2.192,48	0
75%	13	12	103.933	7.153,80	0
90%	56	53	664.379	44.754,54	1
máx	56.274	38.446	159.161.499	10.088.317,00	289

Com base nestas distribuições, os critérios do corte foram escolhidos tentativamente, visando chegar a um número total de agrupamentos entre 15 e 20 mil. Os critérios foram os seguintes:

- que o número de tuítes que continham uma instância do grupo fosse igual ou superior a 56 – 7,7 mil agrupamentos (10%);
- ou que a soma de seguidores das contas que a publicaram fosse igual ou superior a 664.379 – 7,7 mil agrupamentos (10%)
- ou que houvesse sido publicada por ao menos 1 conta verificada – 15,6 mil agrupamentos (20%)

A união dos grupos filtrados com base nestes critérios levou à seleção de 18.475 para conduzir as análises subsequentes. A [Tabela 7](#) apresenta a distribuição das métricas para os grupos selecionados.

Tabela 7: Estatísticas descritivas de publicação das ‘imagens únicas’ selecionadas.

Fonte: API Streaming do Twitter.

	núm tuítes	núm contas	soma seguidores contas	média seguidores contas	núm 'contas verif.'
média	156,44	143,13	1.876.768,30	85.040,07	1,74
desvio	819,10	621,06	5.085.639,07	243.111,25	4,35
mín	1	1	322	172,91	0
10%	2	2	60.061	2.151,79	0
25%	7	7	167.509	3.685,87	1
50%	34	32	481.605	12.705,50	1
75%	108	103	1.495.500	70.249,98	2
90%	292	275	4.918.656	210.304,51	3
máx	56.274	38.446	159.161.499	10.088.317,00	289

Neste ponto, o estudo bifurca-se em duas abordagens distintas, mas que quero considerar como complementares em um esforço reflexivo acerca dos problemas metodológicos para o estudo de imagens por métodos digitais. Em um primeiro momento, volto-me à análise das imagens por meio do modelo VGG19 (SIMONYAN; ZISSERMAN, 2014), uma rede neural para classificação de imagens disponibilizada como parte da biblioteca de programação em Python *Keras* (CHOLLET et al., 2018). Em um segundo movimento, abordo o estudo das imagens por meio de uma API comercial de reconhecimento de imagens: a API Cloud Vision, da Google (GVAPI) (GOOGLE, 2017). Abordarei os procedimentos adotados em cada caso, bem como suas oportunidades analíticas, antes de me voltar à constituição de um dispositivo de análise na forma de um *Atlas para imagens-redes*, discutido em seção posterior (4.3 Compondo imagens-redes).

4.2.4 VGG19

O modelo VGG19 (SIMONYAN; ZISSERMAN, 2014) é desenvolvido pelo *Visual Geometry Group* da Universidade de Oxford e disponibilizado gratuitamente. Trata-se de uma das gerações de um modelo de classificação de imagens desenvolvido pelo grupo, treinado a partir da base de treinamento do projeto *Imagenet* (DENG et al., 2009). Trata-se, também, de um entre diferentes modelos integrados à biblioteca de programação *Keras* (CHOLLET et al., 2018), um dos principais recursos para a aplicação de técnicas de aprendizado de máquina na linguagem Python hoje disponíveis. Diferentemente da GVAPI – um projeto comercial de código fechado, que abordo na seção seguinte – o

VGG19 é um modelo aberto, permitindo o acesso às camadas que compõem a rede neural e, com isso, não apenas um melhor conhecimento sobre sua arquitetura, mas, também, diferentes formas de aplicação em projetos específicos. Evidentemente, como discuti em capítulo anterior (veja 3.2 *Aprendizado de máquina*), uma das características das redes neurais é a sua ‘inescrutabilidade’, significando a impossibilidade de precisar, a partir do código, o percurso classificatório assumido pelo programa. No entanto, como discuti-rei nesta seção, a relativa abertura do modelo VGG19 permitiu um modo de aplicação particular que não é possível em modelos fechados como o GVAPI.

A aplicação do modelo VGG19 neste projeto baseou-se em tutorial que integra o projeto *Machine Learning for Artists (ML4A)* (REFSGAARD; TSENG; KOGAN, 2019). Composto por vários tutoriais e textos introdutórios, este projeto – ainda inconcluso – oferece uma entrada simplificada a aplicações de aprendizado de máquina que possam ter interesse em projetos estéticos experimentais. Alguns dos autores do projeto são também membros ativos de outros projetos de programação criativa como o *OpenFrameworks* (LIEBERMAN et al., 2019), que integra um conjunto já razoavelmente estabelecido de projetos do tipo na comunidade de arte e tecnologia. Entre os autores, há também artistas em residência na Google, aplicando algumas das técnicas apresentadas no ML4A, especialmente, ao programa *Google Arts and Culture*. Estas relações auxiliam a compreender as mediações em jogo na transposição destas técnicas ao estudo. Também ajuda a compreender que a relativa abertura dos procedimentos adotados nesse caso não se realiza em uma completa oposição a iniciativas comerciais como as encampadas pela Google – de modo mais frequente, estas iniciativas tendem a ser absorvidas por empresas como ela¹⁵⁸. Especificamente, os tutoriais em que se baseiam os procedimentos adotados nesta etapa são os “Feature extraction and reverse image search” (KOGAN; MATHEWSON, 2018); e “Image t-SNE” (KOGAN; OLDFIELD, 2018).

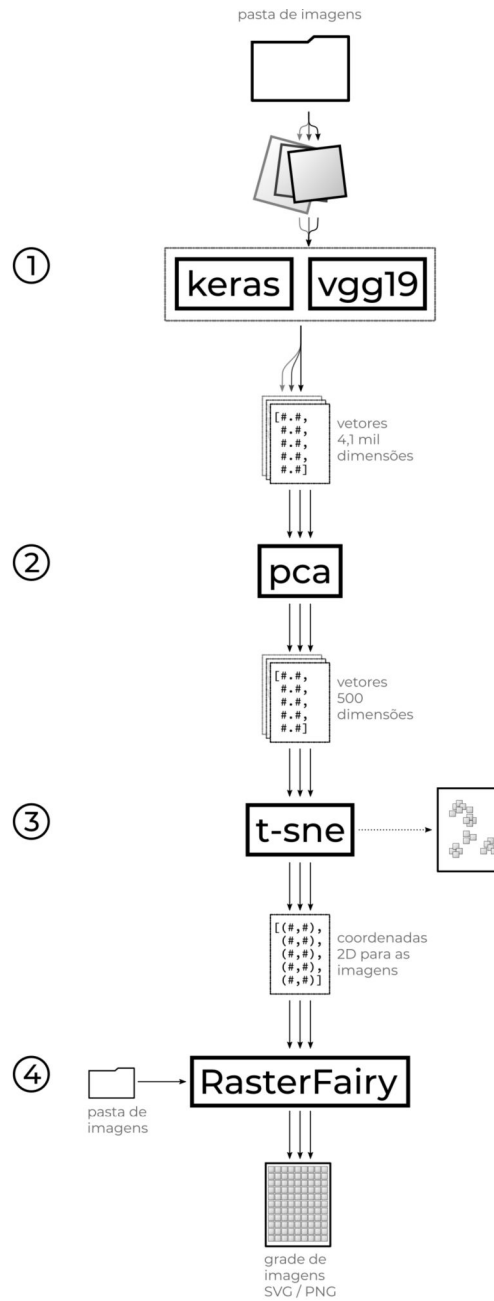
¹⁵⁸ Outro aspecto relevante, a esse respeito, é que François Chollet, proponente inicial da biblioteca *Keras*, também encontra-se, hoje, vinculado à Google.

(1) Submissão recursiva de cada imagem ao modelo VGG19 (SIMONYAN; ZISSERMAN, 2014) via biblioteca de programação Keras (CHOLLET et al., 2015). Modelo limitado até a penúltima camada da rede neural (fc2) operando com 4.096 dimensões e 16,8 milhões de parâmetros.

(2) Aplicação de algoritmo de Análise de Componente Principal para reduzir redundâncias e simplificar os vetores para processamento posterior. Limite de 500 componentes.

(3) Aplicação do algoritmo t-SNE (MAATEN; HINTON, 2008) para redução da dimensionalidade do espaço vetorial para coordenadas planas.

(4) Ajuste das coordenadas para disposição em uma grade regular, sem sobreposição, pelo algoritmo RasterFairy (KLINGEMANN, 2019).



Procedimentos baseados no tutorial Machine Learning for Artists (ML4A) (REFSGAARD; TSENG; KOGAN; 2019)

Figura 19: Protocolo de processamento das imagens pelo modelo VGG19.

Fonte: Elaboração própria. Baseado no tutorial ML4A (REFSGAARD; TSENG; KOGAN, 2019).

Os procedimentos adotados nesta etapa são esquematizados no diagrama da [Figura 19](#). Um aspecto importante para se compreender do processo como um todo é que os dados gerados pelo modelo VGG19 não são tomados por seu valor intrínseco mas, sim, como elementos de articulação relacional entre as imagens. Os dados gerados pelo modelo, como discutirei nesta subseção, operam como inscrições que permitem apreender relações de similaridade entre as imagens segundo suas características visuais. No entanto, embora o VGG19 seja um modelo de classificação de imagens, a sua aplicação neste estudo faz uso apenas de uma parte da rede neural que o constitui, algo que é proporcionado pela disponibilidade do modelo em código aberto. Em vez de se valer das classes atribuídas pelo VGG19, portanto, o protocolo adotado aqui utiliza as representações internas que o modelo produz para as imagens, extraíndo os dados do processamento de uma camada que antecede ao final da rede neural. A última camada (veja [Anexo A](#)) corresponde à camada de previsões (*predictions*), que mapeia as características extraídas da imagem, por meio das camadas antecedentes, às 1.000 classes utilizadas, por padrão, para a base *Imagenet*. Os dados considerados para a análise das imagens nesta etapa são extraídos da penúltima camada de processamento do modelo VGG19, denominada *fc2* ([Figura 20](#)).

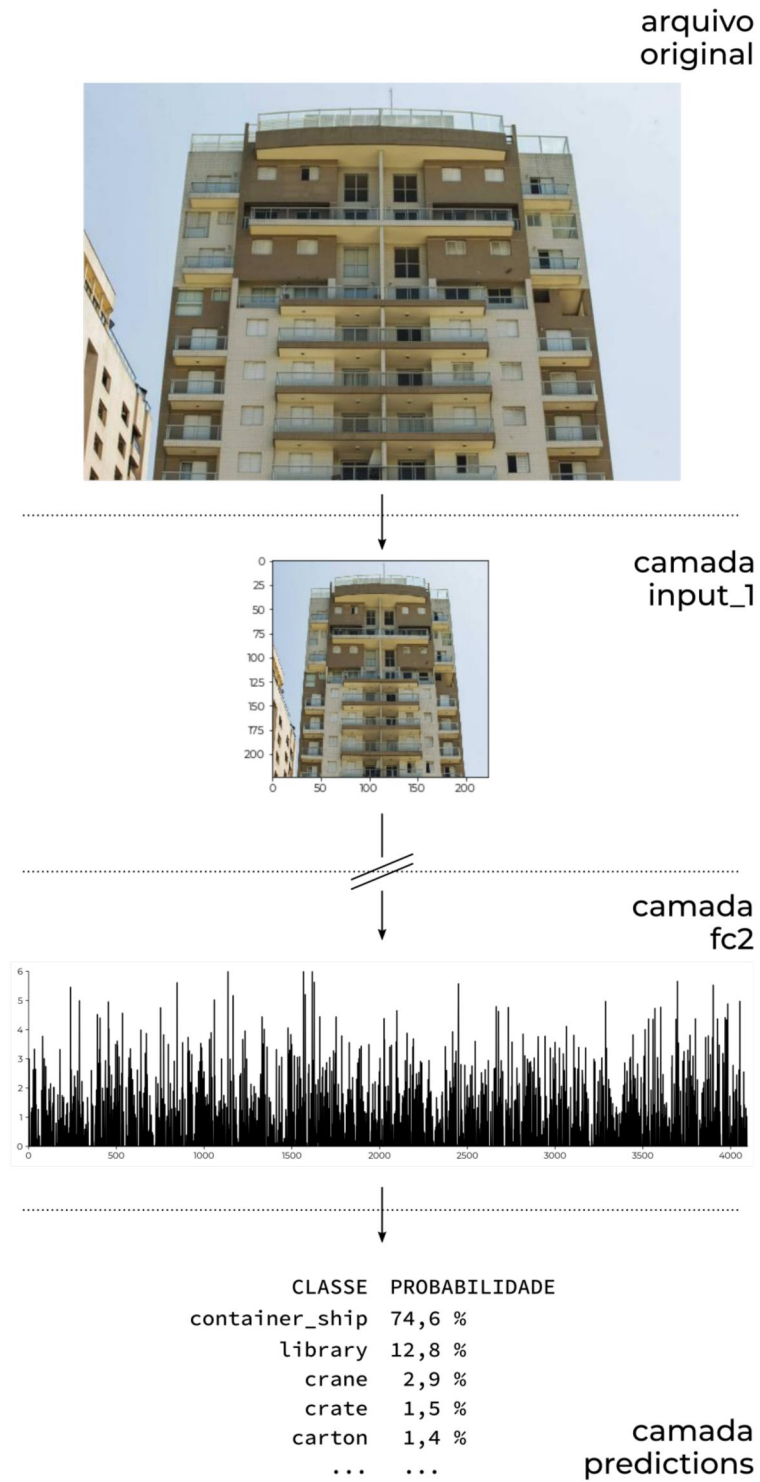


Figura 20: Esquema de processamento de uma imagem pelo VGG19.
 □ VGG19 Q:6 | □ GVAPI Q:3

Fonte: Elaboração própria.

Como descrito anteriormente (veja [3.2. Aprendizado de máquina](#)), as redes neurais operam por uma sequência de transformações que visam, em última instância, produzir uma convergência aos dados de saída estipulados para aquela tarefa. Para o caso do modelo VGG19, quaisquer imagens alimentadas à rede neural devem, ao final do processamento, expressar-se na forma de índices probabilísticos de seu pertencimento a determinadas classes. Embora se trate de tarefa distribuída ao longo das muitas camadas que compõem a rede VGG19, a última camada cumpre um papel decisivo ao reduzir as 4.096 dimensões da camada precedente para 1.000, correspondentes às classes. Esta redução constitui, portanto, uma perda informacional, já que as classes semânticas podem reunir imagens com características distintas, desde que sejam pertinentes a uma mesma classe. Uma vez o que interessa a esta análise é produzir, por meio desse modelo, inscrições alternativas, que proporcionem um modo de orientação em meio aos diferentes registros do acontecimento estudado, a última camada e as classes atribuídas tornam-se prescindíveis. Este procedimento corresponde ao que no jargão do aprendizado de máquina chama-se de ‘extração de características’ (*feature extraction*). Como escrevem Kyle Mathewson e Gene Kogan (2018), trata-se de uma reapropriação do modelo de classificação, aproveitando a relativa consistência da representação da imagem construída pelas anteriores para uma tarefa distinta da classificação.

Para uma comparação, o modelo comercial proprietário da GVAPI, abordado na subseção seguinte, não nos permite acessar esta instância do modelo e oferece apenas as classes – “etiquetas” (*labels*) – atribuídas ao final do processo. Por esse motivo, as relações de similaridade entre as imagens estabelecidas por meio desse outro modelo nem sempre correspondem a similaridades formais e, com frequência, remetem a relações semânticas – voltarei a esse ponto adiante. Embora o número de classes da GVAPI seja – dentro dos limites conhecidos (veja [3.3. Visualidade e infraestrutura](#)) – bastante superior às 1.000 do VGG19, elas são, ainda assim, fruto de uma convergência de características formais das imagens segundo classes semânticas predefinidas. A extração de características realizada a partir do modelo VGG19, por sua vez, permite observar não apenas uma similaridade semântica entre as imagens mas, sim, entre algumas características aparentes das imagens.

Considerando a arquitetura do modelo VGG19 (veja [Anexo A](#)), um primeiro aspecto por observar é *prescrição* da primeira camada da rede (*input_1*), que especifica o formato a que as imagens precisam ser adequadas para serem submetidas ao modelo. Akrich e La-

tour (AKRICH, 1992; AKRICH; LATOUR, 1992) indicam, como discuti, esta primeira instância de conformação da relação entre um objeto técnico e o mundo. Evidentemente, qualquer programa de Visão Computacional demandaria uma inscrição digital das imagens que constitui uma prescrição para a análise visual. O modelo VGG19, em particular, especifica um tamanho para as imagens, que precisam ser ajustadas segundo sua prescrição: 224 por 224 pixels, em três canais de cor. Frequentemente, portanto, será preciso reduzir a imagem, já que estas dimensões são menores do que as imagens que tipicamente encontramos em circulação. Esta primeira inscrição será, então, progressivamente transformada pelas camadas posteriores, segundo os pesos (*weights*) definidos a partir do treinamento com a base *Imagenet*. Ao chegar na camada *fc2*, da qual extraí as características das imagens, a inscrição inicial já assumiu a forma de uma sequência de 4.096 valores numéricos (item 1 da Figura 19).

Embora as classes que seriam obtidas pela camada final não estejam explicitamente expressas nesta representação, é importante lembrar que, pelo próprio caráter difuso das redes neurais, elas tampouco foram excluídas do processo. O treinamento da rede neural, como discutido anteriormente (veja 3.2 Aprendizado de máquina), ocorre por uma otimização de toda a rede de acordo com o alvo estabelecido para a tarefa. Embora o processo de convergência a esse alvo apenas se completaria com a camada final, a penúltima camada da rede já cumpriu uma boa parte deste objetivo e, desse modo, aspectos destas classes estariam implícitas nas relações que podemos estabelecer entre as imagens com base nestes 4.096 valores. Deste modo, tampouco são excluídos os vieses inscritos na base de treinamento ou nas classes estabelecidas como objetivo. Segundo o processo de vetorização empreendido pelo aprendizado de máquina (veja 3.2.2 Vetorização), estes parâmetros são compreendidos como coordenadas espaciais para a localização desta imagem em um *espaço vetorial* de 4.096 dimensões. Tarefas de classificação compreenderiam, grosso modo, a demarcação de limites em meio a este espaço, que seriam pertinentes a cada classe.

Seguindo os procedimentos propostos pelos tutoriais do ML4A (REFSGAARD; TSENG; KOGAN, 2019), os próximos passos adotados visam reduzir as dimensões da inscrição obtida por meio do modelo VGG19, a fim de tornar visualizáveis as relações entre as imagens nesse espaço multidimensional. Três algoritmos são aplicados para este fim. Primeiro, a Análise de Componente Principal (PCA) considera o conjunto das inscrições obtidas para o conjunto completo das 18,5 mil imagens para identificar características

redundantes dentre as 4.096 extraídas a partir do modelo. Este processamento, conforme elaboram Kogan e Mathewson (2018) visa não apenas diminuir as demandas de memória e processamento das etapas subsequentes como, também, evitar sobrerrepresentar certas características do grupo, em função de sua redundância nos dados. No caso desta pesquisa, as 4.096 dimensões originais foram reduzidas a 500 (item 2 da [Figura 19](#)).

Em seguida, o algoritmo t-SNE (MAATEN; HINTON, 2008) foi aplicado a fim de realizar mais uma redução da dimensionalidade dos dados (item 3 da [Figura 19](#)). Tipicamente o t-SNE é utilizado para traduzir coordenadas de espaços vetoriais de muitas dimensões a coordenadas em três ou duas dimensões que podem ser, assim, visualizadas graficamente. Pois um espaço de 4 mil ou 500 dimensões vai além de nossa capacidade de imaginação e, claro, dos dispositivos de representação de que dispomos. Uma analogia possível, me parece, são as projeções cartográficas, que visam transpor uma espacialidade esférica ao plano. Neste caso, a proximidade relativa entre dois pontos em uma visualização t-SNE sugere uma relação de similaridade entre estes pontos, simplificando o que se expressaria, na etapa anterior, em 500 dimensões. Como nas projeções cartográficas, contudo, esta transformação não se realiza sem distorções e a interpretação destas visualizações deve levar em conta esta complexidade. Martin Wattenberg, Fernanda Viegas e Ian Johnson (2016) destacam que visualizações t-SNE possuem várias ‘armadilhas’ que demandam uma análise cuidadosa antes de fazer asserções específicas acerca dos dados representados. Aspectos como *clusters*, tipicamente observados em análise visual de redes, por exemplo (GRANDJEAN; JACOMY, 2019; VENTURINI; JACOMY; JENSEN, 2019), não têm significado expressivo neste caso. Pequenas variações nos parâmetros adotados também podem gerar resultados bastante diversos. Para o caso desta investigação, contudo, estes riscos são menos graves já que não se busca fazer asserções sobre o conjunto de imagens mas, sim, encontrar um auxílio ao esforço de navegação pelos dados. Basta, para isto, a disposição espacial por similaridade relativa, que pôde ser experimentada tentativamente de acordo com os objetivos da pesquisa.

Por fim, um último algoritmo aplicado aos dados foi o *Rasterfairy*, criado pelo artista Mario Klingemann (2019), que transforma as coordenadas espaciais obtidas por meio do algoritmo t-SNE para ajustar as posições dos pontos em uma grade regular (item 4 da [Figura 19](#)). Embora também esta etapa esteja sujeita a distorções, este processo cumpre pelo menos três objetivos. Primeiro, evita a sobreposição entre as imagens, que dificul-

taria a visualização sinótica do conjunto. Segundo, otimiza a utilização do espaço para representar o conjunto das imagens, eliminando espaços vazios. Embora estes espaços tenham valor interpretativo nas análises que se voltam a identificação de agrupamentos (*clusters*), como na análise visual de redes – operando como “buracos estruturais” – eles não têm tanto valor quando o objetivo não é obter descrições gerais. Terceiro, esta disposição tende a uniformizar a representação das imagens, dando-lhes igual valor na espacialização e, também, favorecendo a localização de imagens específicas, uma vez dispostas em um plano cartesiano. A visualização obtida ao final deste processo é apresentada pelo Gráfico 5.



Gráfico 5: Visualização de grade das imagens com similaridade estimada pelo VGG19. A composição apresenta 18.475 matrizes das ‘imagens únicas’ identificadas em processamento anterior. A similaridade entre as matrizes foi computada segundo modelo VGG19 (SIMONYAN; ZISSERMAN, 2014), com pesos atribuídos com treinamento na base Imagenet (DENG et al., 2009). O VGG19 foi utilizado por meio de sua implementação na biblioteca de programação Keras (CHOLLET et al., 2018). O modelo foi ativado até a penúltima camada, com 4.096 dimensões posteriormente reduzidos por Análise de Componente Principal (PCA) e pelo algoritmo t-SNE (MAATEN; HINTON, 2008). Posicionamento em grade pelo algoritmo RasterFairy (KLINGEMANN, 2019). Método baseado no tutorial Machine Learning for Artists (ML4A) (REFSGAARD; TSENG; KOGAN, 2019).

Fonte: Visualização de elaboração própria.

Fonte dos dados:

Imagens – API Streaming do Twitter e Web.

Descrições – VGG19 (SIMONYAN; ZISSERMAN, 2014).

Por meio desta disposição das imagens a visualização permite observar formas visuais recorrentes. Esta seria a principal oportunidade analítica proporcionada por esse procedimento. A aproximação de imagens visualmente similares, como discuti no capítulo anterior (veja [3. Visualidades computacionais](#)), deriva do próprio modo de funcionamento do aprendizado de máquina, que opera pela atribuição de posições relativas para os dados analisados em um espaço vetorial multidimensional. Dois fatores contribuem a esta espacialização. Por um lado, padrões inscritos nos modelos por sua arquitetura computacional e pelos dados de treinamento. Por outro, aspectos dos próprios dados analisados. Pois apenas quando há um volume substancial de imagens relacionadas àquele padrão poderemos apreendê-lo na visualização. A incidência específica destes padrões no caso sob análise, por sua vez, abre possibilidades interpretativas acerca dos sentidos daquelas imagens e padrões naquele contexto.

Por exemplo, a [Figura 21](#) apresenta uma seção do [Gráfico 5](#) que reúne um conjunto heterogêneo de imagens que, em comum, enquadram rostos e mãos em primeiro plano. As características extraídas do modelo VGG19 reúnem nesta seção do gráfico, portanto, imagens formalmente similares, embora retratem situações diversas. A predominância de retratos do Lula é esperada tendo em vista o acontecimento investigado, mas além dele encontramos: ministros do STF; o ex-chanceler Celso Amorim (canto direito superior); o senador Cristovam Buarque; o então presidente Michel Temer (canto direito inferior); a ex-presidenta Dilma Rousseff (canto esquerdo superior); o ex-deputado Roberto Jefferson (próximo ao centro); e o hoje presidente Jair Bolsonaro (canto direito inferior), entre outras pessoas. Uma área mais ampla da visualização apresenta figuras similares, mas esse pequeno segmento já permite apreender a variedade dos sentidos produzidos por um padrão de construção formal das imagens. Em muitos casos, as expressões sugerem vergonha e tristeza, com a mão cobrindo os olhos, a boca, ou toda a face. Em outros, sugere apreensão, com as mãos apoiando o queixo em pose introspectiva. As imagens de ministros do STF, no canto direito inferior, em particular, sugere segredo e conspiração, com a mão cobrindo uma boca que fala.

Um caso interessante do exemplo da [Figura 21](#) é constituído por um par de imagens em preto e branco, posicionadas na margem direita, próximo ao centro vertical da figura. Uma é a capa de um álbum da banda britânica Artic Monkeys; outra é uma fotografia do Lula em pose similar ao do homem retratado na capa do álbum, segurando um cigarro na boca. O nome da banda foi sobreposto à fotografia do Lula, vinculando-a à foto do ál-

bum. Os dados relativos a cada imagem, exibidos na página de detalhamento, indicam que sua inclusão na coleta ocorreu por meio de tuítes publicados por atores distintos, inclusive separados no tempo. A imagem da capa original teria sido majoritariamente publicada a partir do dia 9 de abril, enquanto a montagem teria sido publicada principalmente no dia 7 de abril. É exclusivamente por meio de sua similaridade formal que estas duas figurações são justapostas na visualização que indica, assim, seu potencial para identificar *memes* visuais em meio aos dados. A mão sobre o rosto, nesse caso, sugere uma expressão contrastante com a maior parte do conjunto, uma vez que não expressa vergonha ou apreensão e, sim, altivez.

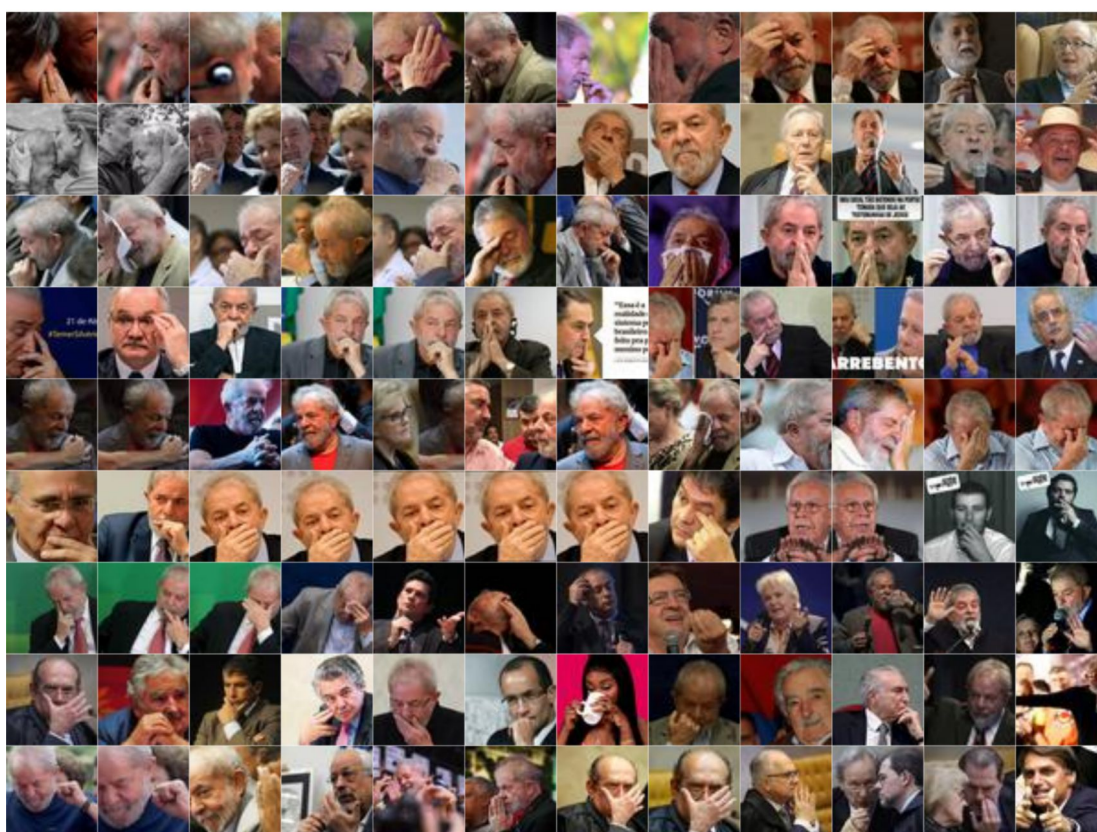


Figura 21: Detalhe da visualização em grade com layout pelo modelo VGG19.
□ VGG19 A:15-B:16

Fonte: Elaboração própria a partir do Gráfico 5.

Segundo o referencial teórico constituído para esta pesquisa, esta visualização pode ser considerada na mesma na mesma linhagem de outras que indiquei anteriormente (veja 3.2.3 Aprendizagem máquina) e que caracterizei como interfaces que nos dão a ver aspectos de uma visibilidade computacional elaborada a partir das redes neurais. Porém,

por meio de visualizações como esta, esta visualidade computacional extravasa qualquer possibilidade de autocontenção e passa a operar, também, sobre nosso modo de ver aquelas imagens. Efetivamente, como também discuti, esta mediação não é exclusiva de contextos de pesquisa, quando ela é deliberadamente reapropriada como forma de navegação ou orientação por grandes quantidades de imagens. Trata-se de instância integrada a plataformas digitais e, portanto, às dinâmicas de circulação e manifestação contemporânea das imagens, como dispositivos que informam nosso olhar. Sugeri em capítulo anterior (veja [3.2.3. Aprendizes maquínicos](#)) considerar este movimento em uma analogia com a técnica de “aprendizado por transferência” (*transfer learning*).

Seria este, quero sugerir, o gesto epistêmico realizado, que também pode ser compreendido como uma materialização específica do que discuti, a partir de Mackenzie (2017) e Suchman (2007) como *aprendizes maquínicos*. Pois visualizações como a exibida no [Gráfico 5](#) não são resultado da simples automatização de uma tarefa de classificação, por exemplo – embora fosse possível aplicar o modelo de classificação para tal fim. De outro modo, a rede neural de classificação é inserida como um mediador algorítmico para auxiliar a apreensão destas imagens, constituindo um dispositivo de orientação entre muitos outros possíveis. Trata-se, desse modo, de um exercício de cooperação, mais do que de automatização. Haveria um “aprendizado por transferência” na medida em que um certo modo de olhar para estas imagens nos é apresentado para que, então, possamos aprender tanto sobre este olhar quanto sobre as imagens às quais olhamos, conjuntamente. Trata-se, desse modo, de mais um sentido para a noção de *imagem-rede*, se consideramos a *imagem* em um sentido fenomenológico – não apenas a *imagem-coisa*, mas a *imagem* que formamos de tal coisa. Esta não se constitui por um olhar único, individual, mas sim pela conjugação de diferentes olhares, como a ideia “visão socializada” que é por vezes mobilizada para definir a visualidade (cf. ABRIL, 2012; MENDONÇA; LEAL, 2018). De modo mais importante, quero ressaltar a constituição do *olhar* como um ato *efetivamente* distribuído. Como discutirei adiante, potencializar este sentido da visualização demandará ainda outros esforços. Em particular, segundo a compreensão da multiplicidade ontológica da *imagem* que é constituída nestes processos. Antes de chegar a tal consolidação, volto a expandir o espectro das inscrições produzidas sobre estas imagens.

4.2.5 API Cloud Vision da Google

Nesta seção, volto-me à descrição dos procedimentos adotados para a análise das imagens por meio da API Cloud Vision da Google (GVAPI). Especificamente, a aplicação da GVAPI recorre ao seu módulo *Label Detection*, que extrai, das imagens, “etiquetas” (*labels*) correspondentes a categorias semânticas. Embora os dados relativos às etiquetas permitam um amplo leque de possibilidades de análise, inclusive por abordagens quantitativas acerca da incidência de determinados termos e categorias, elas são tomadas aqui, como para o caso da reapropriação do modelo VGG19, como mais um tipo de *inscrições* produzidas sobre as imagens como forma de possibilitar outras formas de comparação, justaposição e sobreposição.

O protocolo desta etapa do processamento das imagens encontra-se descrito na [Figura 22](#). Partindo da pasta de imagens e da lista das matrizes dos agrupamentos, foi utilizado um *script* em Python para submeter as imagens à API Cloud Vision (item 1 da [Figura 22](#)). Especificamente, o *script* foi o *Memespector Python*, de minha autoria (MINTZ, 2018b), mas que é uma transposição a essa linguagem de programação do *script Memespector*, desenvolvido inicialmente por Bernhard Rieder para a linguagem PHP (RIEDER; DENTEX; MINTZ, 2018). O *Memespector* realiza submissões sucessivas de uma lista ou tabela contendo nomes de arquivo ou URLs de imagens à GVAPI. Para cada submissão, o programa recebe a resposta da API no formato JSON, interpreta esta resposta e a organiza em diferentes arquivos de saída. Para o caso deste estudo, das 18.475 matrizes ‘imagens únicas’ identificadas, 149 não foram adequadamente processadas pela GVAPI, retornando mensagens de erro. Estas foram descartadas desta fase do processamento, que considera, portanto, 18.326 matrizes.

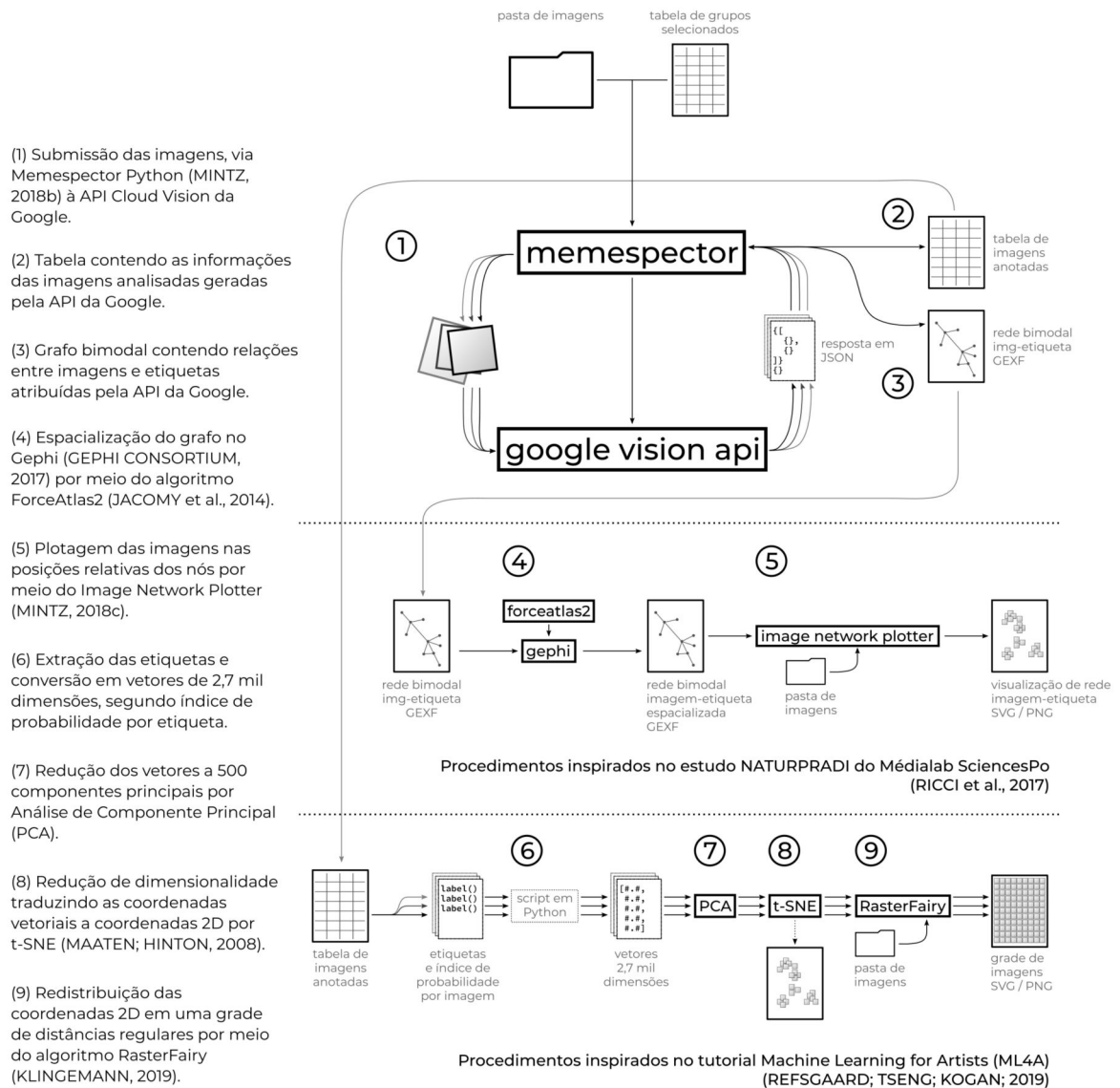


Figura 22: Protocolo de processamento das imagens pela GVAPI.

Fonte: Elaboração própria.

A resposta da GVAPI é transposta pelo script *Memespector* em dois formatos, um já incluído na versão originalmente elaborada por Bernhard Rieder e outra que incluí em minha versão expandida. O formato principal é uma tabela no formato CSV (Valores Separados por Vírgula, na sigla em inglês), que reúne boa parte dos dados do arquivo JSON dispostos como colunas adicionais à lista ou tabela de imagens utilizada como arquivo de entrada (item 2 da Figura 22). Esta forma tabular seria especialmente adequada a análises que visam, por exemplo, quantificar a ocorrência das etiquetas, ou localizar imagens segundo a atribuição de etiquetas específicas. O outro arquivo que adicionei ao *script* é

um grafo bimodal imagem-etiqueta no formato GEXF (Formato XML de Intercâmbio de Grafos, na sigla em inglês) (item 3 da [Figura 22](#)). Um grafo bimodal é uma estrutura de dados relacionais que possui dois tipos de *nós*. Neste caso, um tipo de nó representa as imagens e o outro representa as etiquetas que foram atribuídas às imagens. Quando uma etiqueta é atribuída a uma imagem, estabelece-se entre elas uma conexão, ou uma *aresta*. Deste modo, a atribuição de etiquetas, realizada individualmente a cada imagem, opera como um elemento de articulação entre as imagens. À medida que diferentes imagens são conectadas a um mesmo conjunto de etiquetas, pode-se depreender que elas também teriam algum grau de proximidade semântica segundo o modelo de reconhecimento de imagens que as analisou.

O grafo assim gerado foi, então, processado por meio do *Gephi* (GEPHI CONSORTIUM, 2017), um *software* livre de análise de redes (item 4 da [Figura 22](#)). No *Gephi*, aplicou-se à rede um algoritmo de *layout* dirigido por força, o *ForceAtlas2* (JACOMY et al., 2014), que faz com que as relações estabelecidas entre imagens e etiquetas sejam traduzidas a relações espaciais com propriedades particulares. Grosso modo, o *ForceAtlas2* usa as relações de conexão entre os nós para calcular forças, tratando o grafo como um sistema físico. Conexões operam como forças de atração, de modo que nós conectados entre si tendem a aproximar-se uns dos outros. A ausência de conexão, por sua vez, opera como força de repulsão, que leva a afastar os nós daqueles com os quais não compartilha conexões. Aplicando esse algoritmo de *layout* à rede bimodal construída a partir dos dados da API Cloud Vision, as imagens que compartilham etiquetas, e que, presumivelmente, seriam similares “aos olhos” da API, têm os nós que as representam posicionados próximos uns dos outros. Imagens que não compartilham etiquetas, por outro lado, que presumivelmente seriam dissimilares, têm seus nós posicionados afastados entre si. Por esta complexa dinâmica de forças, portanto, a disposição da rede descreve, ao final, algo a que podemos nos referir por um *espaço semântico* destas imagens, orientado por especificidades do programa de classificação de imagens empregado.

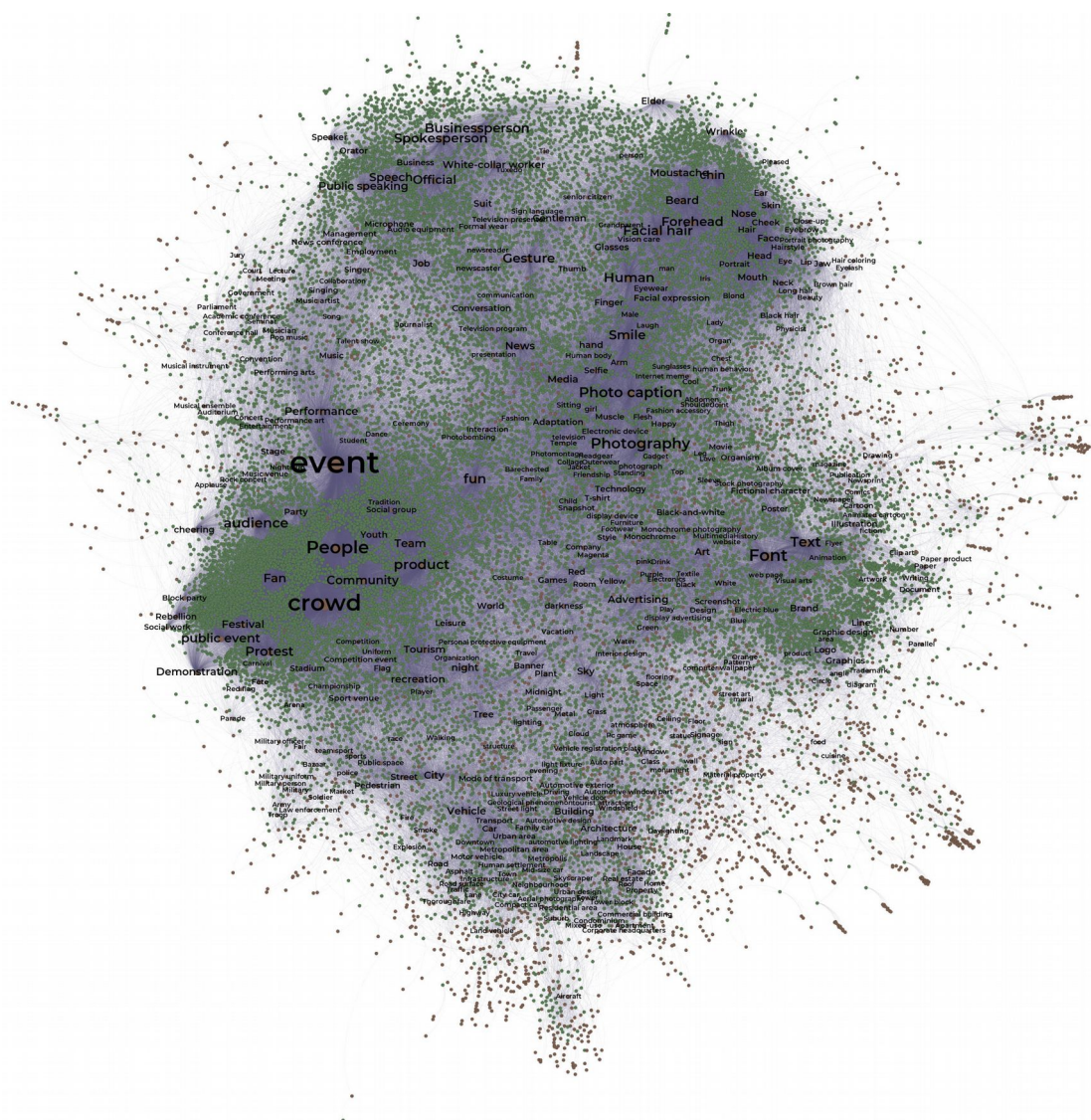


Gráfico 6: Visualização de rede bimodal imagem-etiqueta com dados da GVAPI. Rede composta por 21.092 nós, sendo que 18.326 (86,9%) representam imagens (cor verde) e 2.766 (13,1%) representam labels atribuídos pela GVAPI (cor marrom). Os labels são exibidos como etiquetas destes nós. Os tamanhos dos nós e das etiquetas são proporcionais ao grau de cada nó, em escala logarítmica. Nós com grau inferior a 50 tiveram suas etiquetas ocultadas, bem como componentes menores e desconectados do componente principal da rede (148 imagens e 5 labels). A rede é estruturada por 169.624 arestas não-direcionais que representam a atribuição de uma etiqueta a uma imagem. As arestas possuem peso equivalente à nota de probabilidade da atribuição da etiqueta, segundo definida pela API. Rede espacializada pelo algoritmo ForceAtlas2 (JACOMY et al., 2014), com pequenos ajustes manuais para restringir tamanho da figura.

Fonte: Visualização de elaboração própria produzida no software Gephi (GEPHI CONSORTIUM, 2017).

Fonte dos dados:

Imagens – API Streaming do Twitter e Web. Descrições – GVAPI (GOOGLE, 2017).

O Gráfico 6 é a visualização gerada por este protocolo para as imagens coletadas sobre a prisão do Lula. Em uma leitura visual (GRANDJEAN; JACOMY, 2019; VENTURINI; JA-

COMY; JENSEN, 2019), pode-se identificar aglomerações temáticas que oferecem descrições gerais sobre as imagens segundo as etiquetas atribuídas. Uma vez que o tamanho das etiquetas é proporcional, em escala logarítmica, ao seu número de conexões, as etiquetas mais atribuídas são facilmente identificáveis. No lado esquerdo inferior da rede, vemos uma aglomeração relativamente densa de imagens às quais foram atribuídas etiquetas como: *event* (evento), *crowd* (multidão), *people* (pessoas) e *protest* (protesto). Evidentemente, presume-se que esse conjunto contenha imagens em que aparecem muitas pessoas o que, no contexto do caso em análise, se deve à multidão que acompanhou a vigília de Lula em São Bernardo; ou à que aguardava a chegada o ex-presidente na sede da PF em Curitiba (entre ‘lulistas’ e ‘lavajatistas’).

Na parte superior, à esquerda, encontramos outra aglomeração, menos densa, em que dominam termos como: *business person* (pessoa de negócios, ou executivo/a), *spokesperson* (porta-voz), *speech* (discurso) e *public speaking* (fala em público). Para compreender estes termos, não basta que os tomemos de forma literal. Enquanto ‘porta-voz’, ‘discurso’ e ‘fala em público’ correspondem a, efetivamente, situações ocorridas durante o caso e que possivelmente repercutiram nas imagens no Twitter, ‘pessoa de negócios’ não indica uma categoria que efetivamente se aplicaria às imagens. No entanto, se consideramos que a atribuição de categorias responde a um processo de “aprendizado”, é possível imaginar as pessoas que tipicamente figuravam na base de treinamento para a categoria ‘pessoas de negócios’. Tipicamente, imagens de políticos seriam bastante similares às destas ‘pessoas de negócios’, inclusive considerando marcadores de raça e gênero. Esta suposição pode ser verificada nas visualizações geradas nos passos seguintes.

Ainda acima, à direita, uma terceira aglomeração, densa, é permeada por múltiplas etiquetas relativas a partes do corpo e, especialmente, do rosto, como *forehead* (testa), *head* (cabeça), *mouth* (boca), *nose* (nariz) e *cheek* (bochecha). Destacam-se, contudo, com maiores ocorrências, os termos *facial hair* (pêlos faciais), *beard* (barba) e *moustache* (bigode). Os termos de partes do rosto sugerem, a princípio, imagens em plano de detalhe, que dão prominência a estas partes a ponto de elas surgirem como categorias. A alta ocorrência de termos relativos à barba e pêlos faciais, por sua vez, sugere que muitas destas imagens de detalhe sejam do próprio ex-presidente Lula, já que a barba é um dos traços mais característicos de sua fisionomia. Embora uma predominância masculina, independente de se tratar ou não do ex-presidente, pudesse gerar um efeito similar.

Um estudo que antecede a pesquisa empreendida nesta tese e que faz uso de visualização similar a esta foi elaborado por Rykov *et al.* (2016). O estudo volta-se a imagens publicadas no Instagram na cidade de São Petersburgo, Rússia, e realizou um cruzamento entre os dados semânticos das imagens, obtidos via a GVAPI, e os dados de geolocalização das publicações. O objetivo da análise, portanto, era verificar a existência de correlação entre o conteúdo semântico das imagens (inferido a partir das etiquetas atribuídas pela API) e as regiões da cidade às quais eram vinculadas nas publicações. De certo modo, o estudo realiza, assim, um mapeamento da visualidade de regiões específicas da cidade segundo sua representação no Instagram. Evidentemente, os dados obtidos pelo protocolo que acabo de descrever poderiam subsidiar análises similares. Os metadados de geolocalização não são tão abundantes para os tuítes, mas seria possível perseguir, por exemplo, correlações temporais entre os dias e horários das publicações e as etiquetas mais frequentemente atribuídas às imagens daquele período.

Sem descartar o potencial desta abordagem, o interesse desta investigação é outro e se volta às etiquetas da GVAPI como apenas mais um tipo de inscrição possível para operar sobre as imagens colhidas para a análise. De modo importante, adota-se como princípio a complexidade das manifestações ontológicas das imagens, que não poderiam ser reduzidas a esta inscrição, isoladamente, como se o dado categórico valesse pela própria imagem. O objetivo, portanto, é mobilizar este dado como um operador para compreender as imagens, sem perder de vista outras formas de sua manifestação, em especial sua dimensão visual, que tem fundamental relevância se queremos compreender seu papel em dinâmicas comunicativas e midiáticas. Em que pesem suas muitas declinações, as imagens precisariam ainda ser *vistas* se queremos e reconhecer a necessária fragilidade que esforços que venham supor sua *omnitraduzibilidade* (DIDI-HUBERMAN, 2013a).

A estruturação dos dados em *rede* já favorece, em larga medida, a possibilidade de atingir estes objetivos. Trata-se de forma de representação dos dados que não desfaz as individualidades em massas agregadas de dados. Cada imagem segue, ali, como ponto na rede, oferecendo a possibilidade de retorno ao *individual*. Esta possibilidade proporcionada pela análise de redes foi observada por Latour *et al.* (2012) com muito otimismo quanto ao potencial de que este procedimento metodológico, da análise de redes, pudesse tornar mais difusa a distinção historicamente constituída, nas ciências sociais, entre níveis es-

truturais *macro* e *micro*, ou entre o *individual* e o *agregado*¹⁵⁹. Embora David Moats e Erik Borra (2018) tenham, posteriormente, reivindicado a possibilidade de articular estes níveis sem recorrer apenas à visualização em rede, trata-se de uma abordagem hoje bastante consolidada.

Como discuti em capítulo anterior, a oscilação entre estes níveis – que também guardariam analogia com as noções de *agência* e *estrutura* – encontra-se também entre os objetivos da TAR. O par *ator-rede* propõe, afinal, justamente uma vinculação ambígua entre o individual e o coletivo, como mutuamente constitutivos. Embora nem toda análise de redes e nem toda forma de visualização seja coerente com a TAR – como nos alertam Venturini, Jacomy e Munk (2018) – deve-se também reconhecer como esforços iniciais da TAR tomaram redes bastante similares a esta sobre a qual se debruça esta parte do estudo. Em volume que contém texto fundador da noção de *ator-rede* (CALLON, 1986), um dos eixos estruturantes das contribuições feitas ao livro é justamente a análise de redes co-palavra como instrumento metodológico para compreender os esforços de translação performados pelos diferentes atores no entorno de uma controvérsia sociotécnica (CALLON; LAW; RIP, 1986a). As redes co-palavra estruturam conteúdos verbais utilizando palavras como nós e desenhando conexões entre elas (arestas) a partir de suas relações de coocorrência em um mesmo corpo textual. Nos termos adotados pelos autores, àquele momento, eles visavam, por meio dos dados relacionais e destas visualizações – ainda rudimentares – elaborar uma *cientometria qualitativa* (CALLON; LAW; RIP, 1986b). Escrevem:

a abordagem ‘co-palavra’ não se satisfaz com a contagem como finalidade em si mesma. Nosso problema, de traçar as dinâmicas da ciência e da tecnologia, demanda o uso de métodos quantitativos, mas estes são melhor compreendidos como a busca do qualitativo por outros meios¹⁶⁰ (CALLON; LAW; RIP, 1986b, p. 107–108).

Esta formulação se aplica bem, me parece, aos desafios enfrentados por esta investigação. O objetivo de estudar imagens digitais *online* no contexto de um grande acontecimento midiático – poderia também ser um grande tema ou uma controvérsia sociotécnica – demanda que trabalhe com métodos de base computacional que, em

159 O olhar dos autores é particularmente informado pela teoria social de Gabriel Tarde, que Latour recupera, junto a outros autores, como uma perspectiva até então esquecida nas ciências sociais. Um de seus traços distintivos, em contraste com a linhagem sociológica fundada por Émile Durkheim, deve-se à não distinção entre dinâmicas sociais de menor e maior escala.

160 No original: “... the co-word approach does not indulge in counting for counting’s sake. Our problem, that of tracing the dynamics of science and technology, demands the use of quantitative methods, but these are best seen as the pursuit of the qualitative by other means”. Tradução minha.

alguma medida, serão sempre, também, quantitativos. Como discuti, esta demanda tende a se colocar mesmo que nos interessemos por *uma* imagem específica como foco da análise. Porém, na medida em que também compreendemos que as imagens não podem ser simplesmente traduzidas ou reduzidas a uma categoria ou a um índice estatístico, por exemplo, alguma possibilidade de retorno a sua apreensão visual precisaria ser resguardada. Como na formulação de Callon, Law e Rip, portanto, não se trata de uma investigação quantitativa em sentido estrito mas, sim, da busca pelo qualitativo por outros meios.

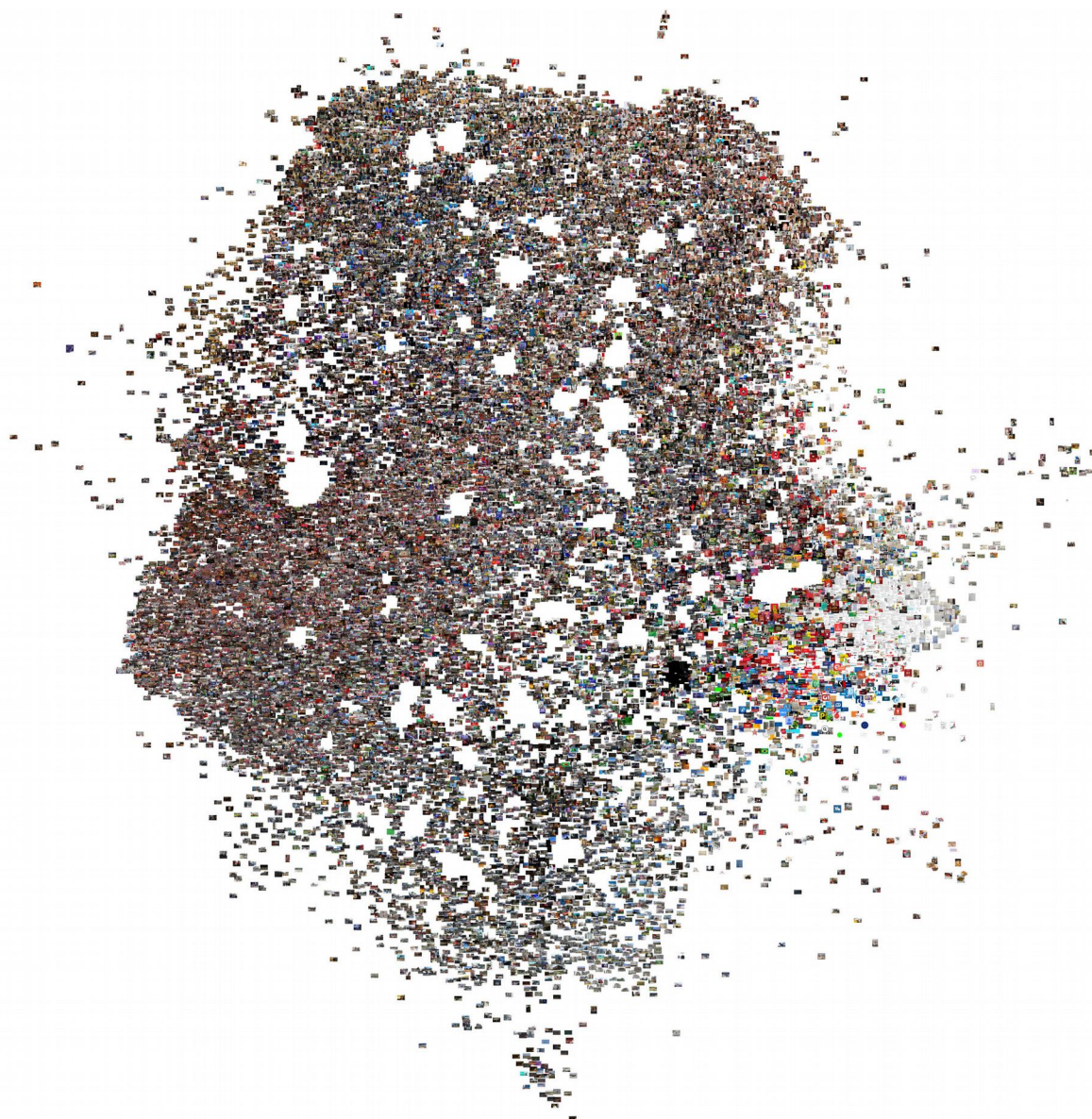


Gráfico 7: *Visualização de rede imagem-etiqueta com imagens plotadas. Trata-se da mesma rede apresentada no Gráfico 6, porém com as imagens correspondentes aos nós de imagem plotadas em sua posição relativa no grafo.*

Fonte: *Visualização de elaboração própria produzida por meio dos softwares Gephi (GEPHI CONSORTIUM, 2017), Image Network Plotter (MINTZ, 2018c) e Inkscape (ALBERT et al., 2018).*

Fonte dos dados:
*Imagens – API Streaming do Twitter e Web.
Descrições – GVAPI (GOOGLE, 2017).*

Avançando em relação à visualização da rede imagem-etiqueta, portanto, uma primeira estratégia para assegurar aos dados um retorno-imagem¹⁶¹ envolve procedimento similar ao adotado pelo grupo de Lev Manovich e a Analítica Cultural: *plotar* os arquivos de imagem no lugar de suas representações gráficas. No caso da Analítica Cultural, tipicamente, isto é feito em gráficos de dispersão. Neste caso, as imagens deverão substituir os nós que, na rede, as representam. Esta abordagem, especificamente no âmbito de redes imagem-etiqueta, foi adotada por Donato Ricci *et al.* (2017) em estudo voltado a figuras da natureza urbana na cidade de Paris. Baseio-me, aqui, na visualização produzida pelos autores. A posição relativa dos nós foi utilizada para plotar as imagens, mas os nós relativos às etiquetas foram excluídos da representação final (Gráfico 7). Gabriele Colombo (2019), aluno de Ricci, sugere este como um dos tipos de ‘imagens compostas’, como estratégia metodológica para estudar imagens “em grupos”. Por meio desta representação, os pesquisadores (RICCI *et al.*, 2017) voltaram-se a esforços de categorização com um distanciamento das categorias efetivamente atribuídas pela GVA-PI. O método que adotaram envolveu uma categorização colaborativa a partir de mapa similar que, assim, adotou categorias mais diretamente pertinentes ao caso em análise, segundo a perspectiva de olhares histórica e socialmente situados naquele contexto. A análise ganha, assim, uma autonomia relativa em relação ao vocabulário adotado pela Google na formulação de suas etiquetas.

De modo a sistematizar o método de plotagem das imagens na rede, desenvolvi um *script*, denominado *Image Network Plotter* (MINTZ, 2018c) baseado no procedimento adotado por Ricci *et al.* (2017). Essa ferramenta recebe como entrada o arquivo de rede (GEXF), especializado no Gephi, e gera um arquivo de imagem em vetor (SVG) com a posição relativa das imagens definida segundo o *layout* da rede produzido pelo Gephi (item 5 da Figura 22). O *script* consegue ser mais eficaz do que o procedimento utilizado por Ricci *et al.* que era limitado a cerca de duas mil imagens¹⁶². Além disto, incluí a possibilidade de estabelecer, para cada imagem, uma URL para funcionar como *hiperlink*, permitindo, a partir da visualização, acessar algum conteúdo *online* relativo à imagem

161 Faço, aqui, menção a uma formulação de Paul Virilio, quando discorre sobre o que chama de ‘imagens virtuais instrumentais’, ou seja, as representações internas à máquina, acerca das imagens: “Não esqueçamos, contudo, que a ‘imagem’ é aqui nada mais do que uma palavra vã, uma vez que a interpretação da máquina nada tem a ver (é o caso de o dizer!) com a visão habitual. A imagem eletrônica é, para o computador, nada mais do que uma série de impulsos codificados dos quais não podemos nem mesmo imaginar a configuração, uma vez que nesta ‘automação da percepção’, o *retorno-imagem não é mais assegurado*” (VIRILIO, 1994, p. 103)

162 Durante participação na Summer School da DMI em Amsterdã, em 2017, tive contato com a técnica utilizada por Ricci e Colombo para a produção destas visualizações. Ela, àquele momento, não era ainda computacionalmente automatizada e tinha limitações para trabalhar com quantidades na ordem dos milhares de imagens.

em questão – como, por exemplo, a publicação que é seu contexto original de circulação. Inicialmente desenvolvido em 2017, o *script* teve algumas revisões até o momento de escrita desta tese e foi utilizado em diversos estudos da imagem por meio de dados da GVAPI de que participei, nesse período (CICALI et al., 2018; MINTZ; SILVA et al., 2019; OMENA; RABELLO; MINTZ, 2017).

Neste estudo, porém, optei por ir além desta possibilidade experimentada em outros momentos da pesquisa e transpor a estes dados alguns dos procedimentos adotados para o modelo VGG19 (veja [4.2.4.VGG19](#)), com base no tutorial do projeto ML4A (REFSGAARD; TSENG; KOGAN, 2019). Especificamente, o objetivo era alcançar uma visualização de *grade* como a obtida ao final daquele processo em uma combinação dos algoritmos t-SNE (MAATEN; HINTON, 2008) e *Rasterfairy* (KLINGEMANN, 2019). Esta visualização de *grade*, embora se afaste das possibilidades derivadas da análise visual de redes (GRAND-JEAN; JACOMY, 2019; VENTURINI; JACOMY; JENSEN, 2019) possui algumas vantagens. Como indiquei na subseção anterior, estas incluem a otimização do uso do espaço, a não-sobreposição de imagens e a possibilidade de localizar imagens com maior precisão, inclusive entre diferentes manifestações do mapa (em tela, impresso etc.). Além disso, a visualização em uma *grade* regular evita destaques a imagens específicas e permite igualá-las, em certa medida, em níveis de relevância.



Gráfico 8: Visualização de grade das imagens com similaridade estimada pela GVAPI. A composição apresenta 18.329 matrizes. Similaridade foi computada segundo descrições do módulo Label Detection da GVAPI (GOOGLE, 2017). Foram encontrados 2.766 atributos, segundo os labels únicos atribuídos, que foram reduzidos por Análise de Componente Principal (PCA) e pelo algoritmo t-SNE (MAATEN; HINTON, 2008). Posicionamento em grade pelo algoritmo RasterFairy (KLINGEMANN, 2019). Método inspirado no tutorial Machine Learning for Artists (ML4A) (REFSGAARD; TSENG; KOGAN, 2019).

Fonte: Visualização de elaboração própria.

Fonte dos dados:
Imagens – API Streaming do Twitter e Web.
Descrições – GVAPI (GOOGLE, 2017).

Para os dados da GVAPI, portanto, dois procedimentos distintos foram adotados após a obtenção das etiquetas descritivas das imagens, como indicado na [Figura 22](#). O caminho se bifurca a partir do item 3. Os itens 4 e 5 foram descritos logo acima, para a elaboração do [Gráfico 7](#). A parte que descrevo agora compreende os itens 6 a 9. Nesta linha de abordagem, as etiquetas descritivas atribuídas pela GVAPI para as imagens foram tomadas como análogas aos vetores de 4.096 dimensões obtidos do modelo VGG19 pelo procedimento de *extração de características*. Para fazê-lo, contudo, foi preciso converter o modo de estruturação dos dados gerados pela GVAPI para a forma de vetores (item 6 da [Figura 22](#)). Isto foi feito considerando cada uma das 2.766 etiquetas atribuídas pela API às imagens como uma *dimensão* e tomando o índice de probabilidade vinculado às atribuições como valor para cada uma destas dimensões. Desse modo, as atribuições das etiquetas foram convertidas em um vetor, que passava a operar como inscrição computacional para cada imagem. Uma vez gerados estes vetores, o protocolo adotado foi o mesmo utilizado para o modelo VGG19 nos itens 2 a 4 da [Figura 19](#): Análise de Componente Principal, reduzindo os vetores a 500 componentes (item 7 da [Figura 22](#)); redução de dimensionalidade do espaço vetorial para uma visualização plana pelo algoritmo t-SNE (MAATEN; HINTON, 2008) (item 8); ajuste das coordenadas a uma grade regular pelo algoritmo *Rasterfairy* (KLINGEMANN, 2019) (item 9). Como resultado, este processo gerou uma segunda visualização em grade, desta vez baseada na leitura das imagens pela GVAPI ([Gráfico 8](#)).

Como para a abordagem do modelo VGG19, gostaria de sugerir nesse procedimento um *aprendizado por transferência* em uma reconfiguração humano-máquina. Uma diferença, contudo, é o estágio do processamento das imagens em que as características foram extraídas do modelo. Enquanto a rede neural do modelo VGG19 é aberta e permite extrair inscrições em uma fase intermédia do processo, o modelo da GVAPI não permite adotar estratégia similar. A utilização das *etiquetas* atribuídas pelo modelo é uma alternativa possível e que se adéqua ao objetivo final de obter a visualização por grade. Contudo, trata-se de uma classificação que já ocorre em um estágio final de convergência e que, como observado para o modelo VGG19, supõe não apenas uma redução do nível de detalhe das inscrições antecedentes como, também, já se aproxima mais de classes *semânticas* do que de atributos visuais das imagens.



Figura 23: Detalhe da visualização em grade com layout pelo modelo GVAPI. Seção extraída do Gráfico 8. □ GVAPI E:9-F:11.

Fonte: Elaboração própria.

A Figura 23 apresenta segmento da visualização que exemplifica bem os efeitos das *etiquetas* semânticas na disposição das imagens. Enquanto a visualização gerada por meio da extração de características de uma camada intermédia do modelo VGG19 tende a aproximar imagens visualmente similares, o posicionamento obtido a partir dos dados categóricos gerados pela GVAPI aproxima imagens com relações semânticas, mesmo que sejam visualmente distintas. No segmento apresentado, ao menos três categorias semânticas podem ser sugeridas a partir de imagens formalmente heterogêneas. Na parte superior da figura, várias imagens são relativas ao que poderíamos indicar como a categoria *comida*, indo desde planos de detalhe de pratos contendo receitas à base de lula (o molusco) até, por exemplo, um registro da ex-presidenta Dilma Rousseff em um restaurante popular, aparentemente. Logo abaixo, há uma faixa de imagens relativas a uma categoria que poderíamos identificar como *futebol*, porém formalmente diversas. Nem todas as fotografias enquadram o gramado, por exemplo. Nesta faixa também se encontram algumas imagens relativas ao que poderíamos indicar pela categoria mais ampla *esporte*. Na faixa inferior, encontramos vários registros de temas bastante distintos mas unidos por uma categoria formal, tornada semântica, que poderia ser descrita como *fotografia em preto e branco*. Em contraste, estes agrupamentos não são encontrados de

forma tão definida na visualização produzida a partir dos dados do modelo VGG19. Como o processamento foi interrompido antes de chegar à fase das categorias, a proximidade entre as imagens se deve mais a relações formais do que semânticas – embora o processo de otimização da rede neural tenda a dispersar as categorias semânticas por todo o modelo. Este exemplo demonstra, portanto, diferentes movimentos de leitura proporcionados por estes dois protocolos de processamento das imagens.

4.3 Compondo imagens-redes

Como ponto de chegada do percurso teórico e metodológico empreendido por esta tese, esta seção apresenta e discute um dispositivo metodológico constituído com o objetivo de apoiar o estudo das imagens baseado no que propus pela noção de *imagem-rede*. Um dos pontos principais desta hipótese conceitual é a compreensão da imagem como entidade ontologicamente instável, não apenas no sentido de sua ambiguidade interpretativa mas também – e fundamentalmente, neste caso – pela multiplicidade das inscrições e das práticas que performam estas imagens como objetos da investigação. Cada etapa adotada na coleta e no processamento dos dados e das imagens assume, como elaborei, diferentes olhares sobre as imagens e aquilo que as define. Cada operação produz uma nova inscrição ou transforma uma inscrição existente, tomando estes elementos como mediadores das operações realizadas sobre as imagens. Cada uma destas instâncias, que são transitoriamente tomadas *como* as imagens em cada etapa, contudo, não são inscrições neutras que poderíamos tomar como meras *intermediárias*. Elas são necessariamente informadas pelo tipo de operação a que se destinam. Deste modo, os esforços de análise, longe de estabilizarem o objeto analisado, acabam por multiplicá-lo, por decliná-lo em diferentes versões de si mesmo.

Esta condição não é exclusiva das imagens digitais, porém tende a se manifestar de forma mais intensa nesse contexto. Isto porque, se podemos afirmar uma *especificidade* das tecnologias digitais, temos que sua fundamentação em operações lógico-simbólicas e dinâmica de conectividade favorece à multiplicação das inscrições. Em certa medida, toda imagem digital atravessa estas múltiplas instanciações nas dinâmicas de circulação em rede, por exemplo, à medida em que é codificada, decodificada e transcodificada múltiplas vezes. Estas transformações são relegadas a um espaço negativo de nossa experiência cotidiana. São obscurecidas ou “encaixapretadas” pela eficácia dos sistemas

computacionais – exceto, claro, quando eles *falham* (Figura 1). Em contextos de pesquisa, estas transformações também tendem a ser relevadas. Seja por um distanciamento da materialidade das imagens digitais (não eram elas imateriais?), seja por uma confiança de que os dados que podemos extrair de sua superfície visível possam nos bastar como forma de conhecê-las. Ao propor a hipótese conceitual da *imagem-rede*, procuro, nesta tese, argumentar em favor de uma compreensão das imagens sem negar a instabilidade oriunda de sua multiplicidade ontológica. Em certo sentido, ao tomar esta hipótese conceitual como eixo da reflexão metodológica empreendida, busco elaborar um modo pelo qual esta multiplicidade pode ser *acolhida* e não mais *purificada* na investigação.

Na seção anterior, apresentei um conjunto de visualizações em que procuro exibir as matrizes dos agrupamentos de imagens atribuindo-lhes posições relativas de acordo com a leitura que delas provêm programas de classificação de imagens por redes neurais. Gabriele Colombo (2018) chama a este tipo de estratégia, de combinação e justaposição das imagens, de *imagens compostas*. A definição é simples: “a combinação de múltiplas imagens em um artefato singular”¹⁶³ (COLOMBO, 2018, p. 24). Uma ponderação que podemos fazer, contudo, é a relativa estabilidade conferida às imagens nesse processo, tomadas como unidades autossuficientes e não, como quero propor, como entidades materialmente relacionais, ontologicamente múltiplas. Se compreendemos a imagem desse modo, haveria um desafio anterior ao de combinar múltiplas imagens em um único artefato. Primeiro, seria necessário *compor* a imagem a partir de suas múltiplas instâncias e suas múltiplas inscrições. Este é, em suma, o desafio que coloco para esta seção, na qual busco me haver com o desenvolvimento de um dispositivo metodológico que poderia apoiar esse esforço de composição *ontográfica*.

Ao sugerir a ideia de uma *composição*, faço uma aproximação da defesa que Bruno Latour (2010) elabora de uma abordagem “composicionista”, que buscaria construir um *plano comum* para elementos heterogêneos e dispersos sem que, com isto, sejam homogenizados¹⁶⁴. Situando-se entre um *universalismo* e um *relativismo*, o que ele ironicamente propõe como *composicionismo* sugere a amplitude de um olhar geral, porém ciente das partes heterogêneas que inevitavelmente o compõem:

163 No original: “the combination of multiple images into a single artifact”. Tradução minha.

164 De modo mais estruturante do que nesta tese, voltei-me também proposição da composição por Bruno Latour em minha dissertação de mestrado (MINTZ, 2015).

Do universalismo, ele toma a tarefa de se construir um mundo comum; do relativismo, a certeza de que esse mundo comum deve ser construído com partes completamente heterogêneas que nunca comporão um todo, mas, no melhor dos casos, um material composto frágil, revisável e diverso¹⁶⁵ (LATOUR, 2010, p. 474).

Esta elaboração precisaria ser compreendida em um contexto mais amplo da obra do autor que, já àquele momento, voltava-se a um esforço de revisão do projeto moderno em resposta às demandas do que veio chamar de *regime climático* (LATOUR, 2018). Nesse contexto, a *composição* é pensada por Latour como um gesto político-epistêmico que visa agregar modos de existência heterogêneos segundo suas relações de interdependência¹⁶⁶. Evidentemente, minha aproximação à noção de composição tem sentido bem mais circunscrito e talvez fosse melhor compreendida como uma transposição conceitual.

A metáfora sugerida na descrição de Latour parece-me particularmente adequada para o esforço visado aqui, em sua sugestão de “um composto frágil, revisável e diverso”. Imagino uma criança coletando pedras, galhos, um punhado de areia e um punhado de lama; apertando essa mistura sobre o chão e vendo, nela, um ‘bolo’. Mas, tão logo retira sua mão, sua forma já começa a se desfazer, o galho tomba, a pedra rola. Sustentar a forma do ‘bolo’ demandará um trabalho constante de manutenção e recomposição mas, em todo caso, sempre será possível identificar ali as partes coletadas e forçosamente combinadas em um agregado. O esforço de conhecimento lançado, aqui, às imagens-redes, parece em larga medida assemelhar-se ao gesto da criança. Como se a tarefa metodológica fosse coletar arquivos de imagens distintos, *hashes* perceptivos, *etiquetas*, vetores de características formais. Tuítes, URLs, métricas. Juntar tudo entre as mãos e produzir, como “mágica”, um ‘bolo’. Grosso modo, estudos de métodos digitais assemelham-se, em maior ou menor grau, com mais ou menos rigor e sistematização, a gestos desse tipo. A questão é em que medida este gesto é consciente e, principalmente, em que medida a heterogeneidade é reconhecida na própria análise.

O objetivo desta seção é, enfim, constituir um instrumento analítico que proporcione uma experiência de exploração das imagens circuladas no entorno do acontecimento midiático da prisão. Este instrumento deveria auxiliar a compreensão tanto do conjunto

165 No original: “From universalism it takes up the task of building a common world; from relativism, the certainty that this common world has to be built from utterly heterogeneous parts that will never make a whole, but at best a fragile, revisable, and diverse composite material”. Tradução minha.

166 Em momento recente, tal compreensão é articulada em seu pensamento pela figura de Gaia (ARÈNES; LATOUR; GAILLARDET, 2018; LATOUR; LENTON, 2019).

das imagens quanto de sua individualidade, permitindo, portanto, oscilar entre leituras distantes e em profundidade. Ele deveria, ainda, oferecer uma entrada às múltiplas instâncias e inscrições que operam sobre aquelas imagens: os agrupamentos realizados segundo uma noção probabilística de identidade; as descrições elaboradas por classificadores de imagens; e aspectos dos tuítes e do contexto de publicação. Descrevo, a seguir, a constituição desse dispositivo de orientação, a que chamo de *Atlas*.

Como lembram Daston e Galison (2010, p. 23) e Amir Brito Cadôr (2016, p. 133), o nome do titã grego Atlas foi primeiro atribuído uma coleção de mapas em um volume que foi publicado por Gérard Mercator em 1585. Sua capa trazia a imagem do ser mitológico que carregava o mundo nas costas. Posteriormente, elabora Cadôr (2016, p. 133), no século XIX, o nome passou a designar “qualquer dispositivo tabular de conhecimento sistematizado: atlas de astronomia, anatomia, geografia, etnografia”. Gostaria de salientar, nesta proposição, a polivalência do termo e, também, dos próprios objetos que ele nomeia. Sem buscar vinculação estrita a nenhum atlas específico, gostaria de potencializar as remissões possíveis a estes muitos atlas e a suas diferentes implicações às práticas de conhecimento do mundo.

Em vista do percurso elaborado por esta tese, uma remissão evidente da denominação *Atlas* é o projeto inacabado de Aby Warburg (veja [2.3.1 O Atlas Mnemosyne e os veículos de imagem de Aby Warburg](#)). Escrevendo sobre o *Atlas Mnemosyne*, Warburg (2015) sugere ideias como uma “cosmologia das imagens”, ou como um “inventário das pré-formações” que diriam de uma “massa de impressões”, que constrangeria os artistas em seu trabalho de criação (WARBURG, 2015, p. 370). Cadôr (2016) ressalta a importância do *Atlas Mnemosyne* no contexto do ensino e pesquisa em história da arte. Ressalta, a esse respeito, o caráter eminentemente visual do projeto warburguiano e a possibilidade de constante revisão e reorganização dos registros¹⁶⁷. Interessado especificamente na reincidência da forma Atlas em uma investigação panorâmica sobre livros de artista¹⁶⁸, Cadôr indica ressonâncias do Mnemosyne em projetos desenvolvidos por diferentes artistas a partir dos anos 1960, como o alemão Gehrard Richter, que elaborou seu próprio Atlas, composto por cerca de 5.000 imagens. Citando Richter¹⁶⁹, escreve Cadôr (2016, p. 136): “A estrutura torna ‘tudo igualmente importante e igualmente desimportante’, de

167 Como consequência disto, inclusive, praticamente todas as pranchas originais e reproduções nelas utilizadas se perderam na mudança do Instituto Warburg de Hamburgo a Londres, no período de ascensão do nazismo.

168 Em resumo, livros de artista são obras em que o artista explora a forma *livro* como meio expressivo.

169 RICHTER, Gehrard. Notes. In: FERREIRA, Glória; MELLO, Cecilia Cotrim de. *Escritos de artistas: Anos 60/70*. Rio de Janeiro: Jorge Zahar, 2006. p. 113-119.

modo que ‘todas as partes se interpenetrem’”. Este ponto, quando trazido de volta a Warburg, me parece importante. Atlas de imagens como estes operam por meio de justaposições que permitem colocar objetos heterogêneos sobre um mesmo plano e, assim, compará-los. O Atlas de Richter, em particular, segue um rigor que tende à indiferenciação entre as imagens e seus formatos, todas dispostas em uma mesma grade homogênea. Certo modo, também o Atlas que proponho segue esta lógica de uniformização. O de Warburg, de outra maneira, parecia trabalhar conscientemente as distâncias, os tamanhos e as disposições espaciais. Em todo caso, ambos os Atlas, de Warburg e Richter, operam por tensões e choques entre as imagens como produtoras de um sentido relacional. Este é o aspecto que gostaria de trazer dos *atlas de imagens*.

Um outro sentido de atlas a que somos remetidos são justamente aqueles iniciados por Mercator: volumes encadernados que portam coleções de mapas. Estes podem ser referentes a diferentes regiões e territórios, mas nem sempre, pois atlas podem também dedicar-se a um território específico, trazendo múltiplas representações de uma mesma região. Mapas podem ser políticos, hidrográficos, demográficos, geomorfológicos etc. Este seria outro modo de materializar a proposição de Annemarie Mol (1999) de uma multiplicidade ontológica decorrente das diferentes práticas epistêmicas que, cada qual a seu modo, performam seus objetos e – de forma literal, neste caso – seus próprios *mundos*. Na linha do que discuti acerca das inscrições sociotécnicas, os atlas podem ser compreendidos como coleções de inscrições que operam, elas próprias, como seus veículos. Assemelham-se aos “centros de cálculo”, sugeridos por Latour (LATOURE; HERMANDT, 2004) para bibliotecas e coleções, mas condensados em volumes que funcionam como *centros de cálculo móveis*. Desse modo, atlas seriam artefatos que não garantiriam consistência a apenas uma inscrição, mas a um conjunto delas, comportando, desse modo, a multiplicidade ontológica de um território em uma rede de referências circulantes. Quero reter este ponto acerca deste outro sentido, cartográfico, dos atlas.

Haveria, ainda, um terceiro sentido, pelo qual podemos ser remetidos aos atlas científicos. Tipicamente, estes são grandes volumes que portam imagens dos objetos de determinadas disciplinas científicas para, entre outras finalidades, “treinar o olho” dos cientistas. Escrevem Daston e Galison (2010, p. 22):

Atlas são compilações sistemáticas de objetos de trabalho. Eles são os dicionários da ciência do olho. Tanto para iniciados quanto para neófitos, o atlas treina o olho para selecionar certos objetos como exemplares (por exemplo,

este fígado saudável “típico” em vez daquele com cirrose) e para observá-los de um modo específico [...]. Os atlas exercitam o olho do iniciante e refrescam o olho do experiente¹⁷⁰.

Sob este sentido, os atlas não são somente coleções de inscrições, até porque não são, necessariamente, os objetos estudados. Nos sentidos do que discuti acerca do aprendizado de máquina, eles seriam como “bases de treinamento” para observadores humanos. Suficientemente expostos às diferentes espécies de aves do cerrado, o ornitólogo torna-se mais apto a identificá-las *in loco*. Neste caso, portanto, os atlas científicos seriam materializações da constituição performativa das classes e categorias analíticas. Este é o ponto que quero reter deste terceiro sentido dos atlas.

O desafio de constituir um atlas para subsidiar um esforço metodológico como o proposto por esta tese encontra-se, principalmente, no modo de articulação das múltiplas inscrições produzidas a partir dos dados coletados. Estas deveriam ser combinadas em um “espaço de pensamento” que consiga, no sentido da *composição* que tomo de Latour (2010), articular a percepção de um espaço comum, mas sem homogenizar suas partes. Especificamente neste estudo de caso, um atlas *ideal* deveria reunir em um plano comum:

- a) Os agrupamentos de imagens tidas como provisoriamente *idênticas* umas às outras, “imagens únicas”, representadas por *matrizes* selecionadas (veja [4.2.1 Identificando imagens “únicas”](#)).
- b) Os tuítes que publicaram imagens destes agrupamentos, e seus metadados (veja [4.1.2 Tuítes](#)).
- c) Etiquetas atribuídas pela GVAPI para cada matriz dos agrupamentos de imagens (veja [4.2.5 API Cloud Vision da Google](#)).
- d) Características extraídas por meio do modelo VGG19 (veja [4.2.4 VGG19](#)).
- e) Os arquivos de imagem originalmente coletados, individualizados pelas URLs de sua proveniência, tidos como *instâncias* dos agrupamentos (veja [4.2.1 Identificando imagens “únicas”](#)).

170 No original: “Atlases are systematic compilations of working objects. They are the dictionaries of the sciences of the eye. For initiates and neophytes alike, the atlas trains the eye to pick out certain kinds of objects as exemplary (for example, this ‘typical’ healthy liver rather than that one with cirrhosis) and to regard them in a certain way [...]. The atlases drill the eye of the beginner and refresh the eye of the old hand”. Tradução minha.

Evidentemente, seria possível ampliar esta lista a partir de outros esforços analíticos. Por exemplo, em artigo publicado durante o doutorado, elaborei com Carlos d'Andréa um procedimento metodológico voltado para o estudo diacrônico da circulação entre plataformas de certas imagens, aplicando o módulo *Web Detection* da GVAPI (D'ANDREA; MINTZ, 2019). Abordagem similar é aplicada em corte sincrônico por Tarcízio Silva, Pedro Barciela e Pedro Meirelles (2018) com interesse específico em redes de desinformação eleitoral. Dados como estes podem ser, potencialmente, integrados em iterações futuras deste atlas¹⁷¹.

Contemplar com igual prioridade estes diferentes aspectos, contudo, constitui uma situação ideal. Hipoteticamente, um esforço mais extenso e aprofundado permitiria constituir múltiplos pontos de entrada à navegação que contemplariam de modo mais adequado as diferentes instanciações das imagens, a fim de proporcionar uma pluralidade epistêmica – como sugerem Rieder e Borra (2014) para o DMI-TCAT. Contudo, deixo esta possibilidade em um horizonte futuro e apresento aqui a constituição prototípica de um atlas que busca contemplar os itens indicados porém, em alguns casos, parcialmente. Dado o percurso teórico e os interesses iniciais da pesquisa, esta elaboração assume a opção de tomar os agrupamentos de imagens tidas como *imagens únicas* como ponto de entrada da exploração. Estas são inicialmente apresentadas segundo as *imagens compostas* que foram elaboradas pela planificação dos espaços vetoriais gerados a partir das classes semânticas e características extraídas das imagens. Considerando, portanto, os dois modelos de classificação de imagens aplicados nesta investigação – VGG19 e GVAPI (Figura 24). O Apêndice B apresenta instruções para acesso e navegação do *Atlas*, anexado à cópia física desta tese.

Tecnicamente, o Atlas é constituído por um conjunto de documentos de hipertexto simples, em linguagem HTML/CSS e JavaScript. Cada página do Atlas é, assim, um documento HTML individual e estático, gerado por um *script* escrito especificamente para este fim. Embora esta implementação seja menos eficiente sob a perspectiva do armazenamento, gerando um conjunto volumoso de arquivos (3,1 GB), optou-se por evitar páginas geradas dinamicamente, por linguagens de execução em servidor, para tornar mais simples o acesso e armazenamento do conteúdo, inclusive em modo *offline*. O Atlas pode ser aberto em qualquer navegador de internet contemporâneo. As linguagens utili-

171 Efetivamente, possuo dados desta circulação para algumas das imagens consideradas neste estudo, mas optei por não explorar esta frente na tese.

zadas são, também, bastante consolidadas, de uso difundido e fácil interpretação humana, o que tende a facilitar esforços de preservação digital e eventuais traduções a outros sistemas e linguagens.

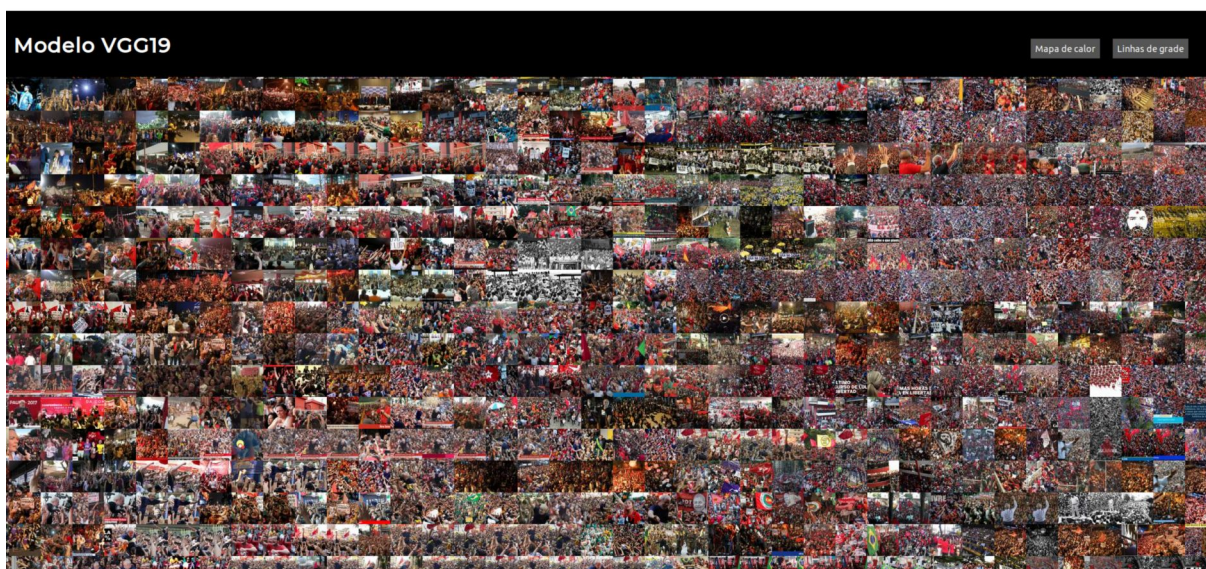


Figura 24: Impressão de tela do estado inicial do Atlas no layout VGG19.

Fonte: Impressão de tela de página Web de elaboração própria.

Esta visualização proporciona a possibilidade de uma navegação *espacial* e *visual* pelas imagens. Evidentemente, a noção de *visual*, neste caso, encapsula uma visualidade complexa que conjuga nosso olhar a uma visualidade maquínica, constituída pelas redes neurais e aquilo que identificam como similares. Este “modo de ver”, como discuti (veja [3.2 Aprendizado de máquina](#)), depende tanto da arquitetura de redes neurais quanto das imagens e das categorias utilizadas para o treinamento desses modelos. Ou seja, as máquinas de visão tampouco poderiam ser concebidas como ‘puramente máquinas’, distanciadas de qualquer olhar humano. Trata-se, de outro modo, de uma visualidade computacional constituída de forma híbrida e que, nesta interface, oferece-se, novamente, a um atrito com os olhos de um sujeito histórica e socialmente localizado. Caberá ainda a uma pessoa navegar por tal disposição e encontrar padrões emergentes interpretáveis à luz do caso específico a que as imagens se vinculam. Alguns dos aspectos pertinentes a cada uma destas visualizações são indicados em subseções precedentes (veja [4.2 Processamento das imagens](#)).

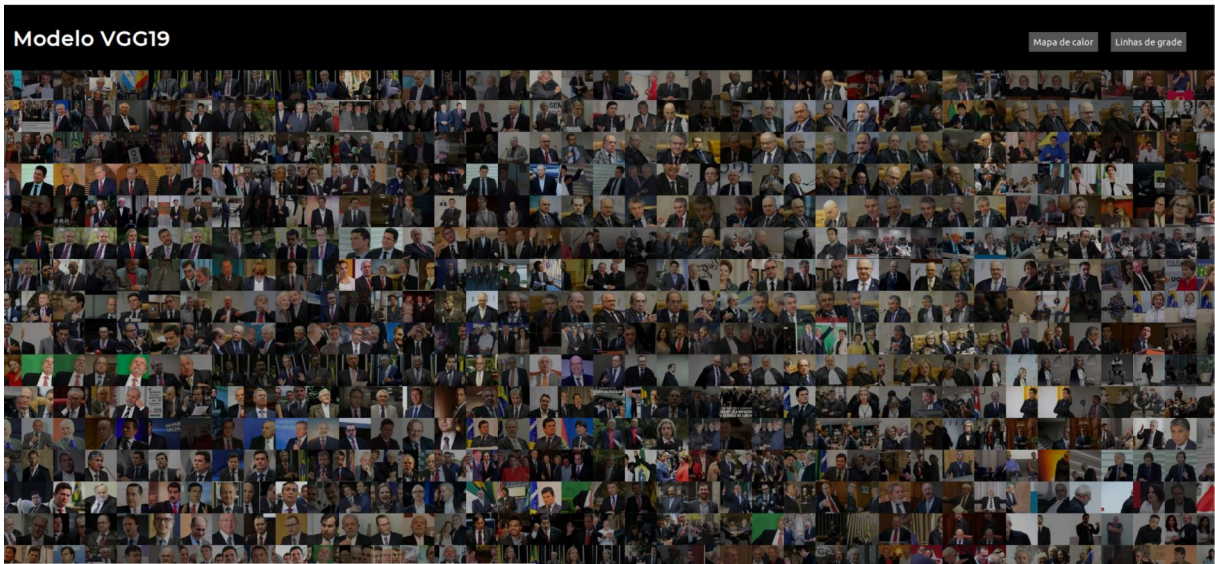


Figura 25: Impressão de tela do Atlas em layout VGG19 com mapa de calor .

Fonte: Impressão de tela de página Web de elaboração própria.

Como um complemento a esta visualização, um botão no canto direito superior da janela possibilita sobrepor às imagens uma visualização do tipo “mapa de calor” (*heat map*) (Figura 25). Tipicamente, trata-se de um tipo de visualização de dados que indica valores quantitativos como variações cromáticas de células de dados tabulares. Neste caso, o mapa de calor incide no valor de opacidade de retângulos pretos sobrepostos a cada imagem. O retângulo preto tem sua opacidade definida por um valor entre 0 e 0.5, inversamente proporcional, em escala logarítmica, à soma dos seguidores das contas que publicaram instâncias da imagem. Posto de outro modo, o retângulo torna-se tão mais transparente quanto maior for o público potencial daquela imagem. De modo que, com a sobreposição destes retângulos, as imagens que tiveram maior público potencial tornam-se mais visíveis, enquanto as demais tornam-se mas ‘apagadas’. Por meio deste elemento, pode-se compreender, já na navegação por esta visualização, imagens que possivelmente seriam mais relevantes a um olhar interessado em medidas quantitativas de sua circulação.

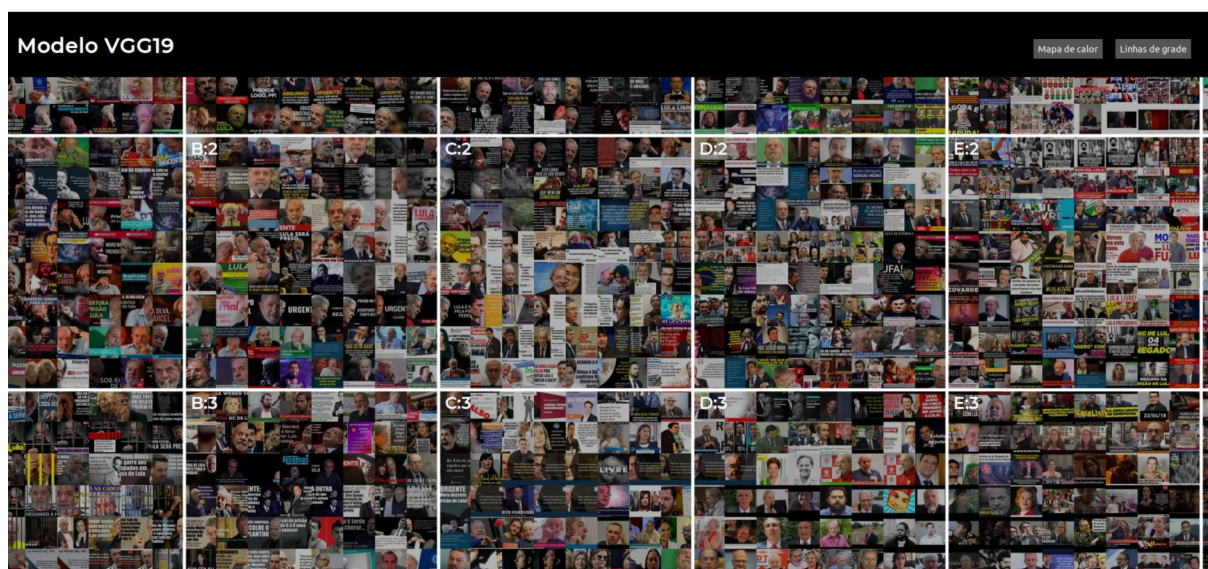


Figura 26: Impressão de tela do Atlas em layout VGG19 com linhas de grade.

Fonte: Impressão de tela de página Web de elaboração própria.

Como um último acréscimo a esta visualização, um segundo botão no canto direito superior (“Linhas de grade”) permite sobrepor ao mapa das imagens uma *grade* contendo coordenadas alfanuméricas (Figura 26). Esta camada tem por objetivo facilitar o registro da localização de imagens ou grupos de imagens específicos, bem como reencontrar imagens localizadas em outros momentos de navegação. A grade divide o mapa em 17 linhas e 17 colunas de tamanhos iguais, totalizando 289 quadrantes. Cada quadrante contém 8 linhas e 8 colunas de imagens, totalizando 64 imagens. Outro uso para esta grade é padronizar coordenadas a fim de permitir a transposição entre diferentes realizações do próprio mapa, já que ele também pode, potencialmente, ser observado em forma impressa. As coordenadas estabelecem um plano de consistência que habilita alternar entre diferentes instanciações do mapa.

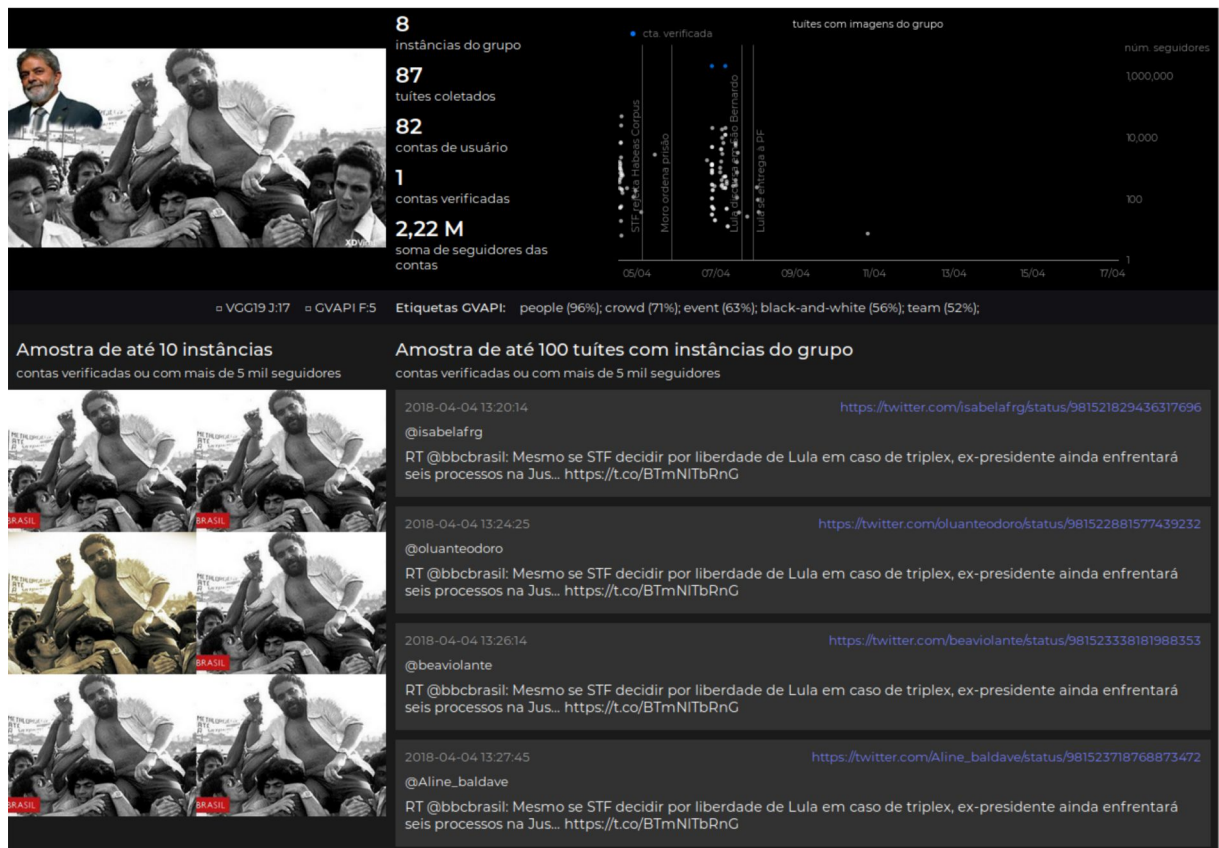


Figura 27: Impressão de tela do Atlas em visualização de detalhe.
 VGG19 J:17 | GVAPI F:5

Fonte: Impressão de tela de página Web de elaboração própria.

A partir desta visualização, também é possível clicar em uma imagem para navegar à sua página de detalhamento (Figura 27). Um dos objetivos destas páginas é desempacotar a identidade presumida de cada imagem entre as diferentes instâncias que a compõem. O canto esquerdo superior exibe o arquivo *matriz* que, como detalhado antes (4.2.1 Identificando imagens “únicas”), é a imagem de maiores dimensões em meio às instâncias que compõem o agrupamento. Imediatamente à direita, são apresentadas cinco métricas relativas ao agrupamento: (1) o número de arquivos de imagens (instâncias) agrupados; (2) o número de tuítes coletados contendo referência a alguma instância deste agrupamento; (3) o número de contas de usuário identificadas como autoras dos tuítes coletados com referência a instâncias do agrupamento; (4) o número de ‘contas verificadas’ pelo Twitter (cf. TWITTER, 2019e) dentre as que publicaram tuítes com referência a instâncias do agrupamento; (5) a soma absoluta do número de seguidores das contas que publicaram tuítes com referência a instâncias do agrupamento. À direita destas métri-

cas, apresenta-se um gráfico de dispersão em que cada tuíte coletado com referência a instâncias do agrupamento é representado por um ponto. A posição no eixo X indica data e hora da publicação do tuíte. Auxiliando a interpretação desta posição, linhas verticais indicam alguns acontecimentos-chave do período (cf. [Gráfico 4](#)). A posição no eixo Y indica o número de seguidores da conta que publicou aquele tuíte, em escala logarítmica. Em faixa estreita imediatamente inferior a esta seção, sob a matriz do grupo, são indicadas as coordenadas desta imagem nos dois mapas constituintes do atlas (VGG19 e GVAPI). À direita, nesta mesma faixa, são indicadas cinco das etiquetas atribuídas à imagem pela GVAPI, junto ao índice de probabilidade, segundo dados da API. Na porção inferior da página, o lado esquerdo apresenta uma amostra de até dez instâncias consideradas como parte do grupo, limitadas a arquivos de imagem publicados por contas verificadas ou com mais de 5 mil seguidores. Por fim, o lado direito da porção inferior apresenta uma amostra de até 100 tuítes coletados contendo referências a instâncias do agrupamento. Esta amostra é selecionada aleatoriamente, mas apenas tuítes publicados por contas verificadas e/ou com mais de 5 mil seguidores são apresentadas. As restrições colocadas aos arquivos de imagem e aos tuítes exibidos individualmente visam atender aos critérios estipulados como medidas de privacidade dos dados coletados (veja [4.2.2 Medidas de privacidade](#)).

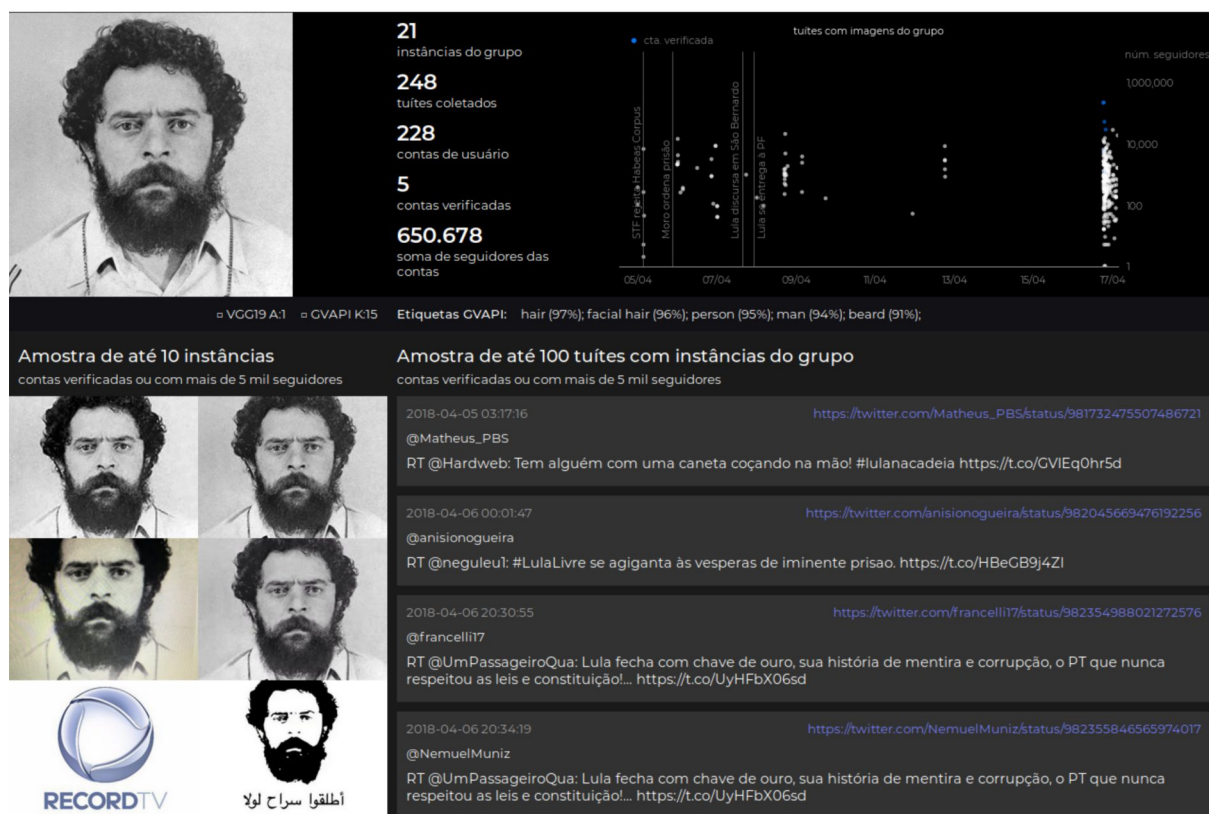


Figura 28: Impressão de tela com exemplo de agrupamento inconsistente
 VGG19 A:1 | GVAPI K:15

Fonte: Impressão de tela de página Web de elaboração própria.

Ao reunir, de forma articulada, as múltiplas inscrições geradas a partir dos dados coletados para este caso, uma das primeiras oportunidades analíticas proporcionadas pelo *Atlas* é a possibilidade de compreender aspectos dos métodos de processamento empregados. Um primeiro aspecto, especialmente relevante, diz respeito a algumas inconsistências eventuais do método de agrupamento das imagens. A Figura 28 exhibe um caso que ilustra o caráter complexo das relações de identidade e diferença performadas pela técnica empregada no agrupamento (veja 4.2.1 Identificando imagens “únicas”). O grupo, em sua maioria, apresenta diferentes instâncias de uma mesma fotografia: o retrato de identificação produzido quando da prisão de Lula em 1980. Na amostra das instâncias exibida na tela, as quatro primeiras imagens são claras variações do mesmo retrato, variando levemente tonalidades e enquadramentos. Já as duas instâncias localizadas na última linha não são correspondências diretas. A imagem da esquerda, a logo da TV Record, é um caso claro de agrupamento impertinente, equivocado. Já a imagem à direita é uma versão bastante modificada das demais imagens, mas que ainda é uma va-

riação sobre o retrato. Trata-se de uma ilustração entre várias que foram produzidas no entorno da prisão e que clamavam “Lula livre” em diferentes idiomas. Neste caso, reconhecemos que seria uma versão já bastante distinta, mas haveria certo sentido em agrupá-la, a depender do critério que quisermos adotar.

Deve-se compreender, em todo caso, que a inconsistência deriva da natureza probabilística do método de identificação via *hashing* perceptivo. Dois fatores devem ser levados em consideração, neste caso. Primeiro, que a técnica se baseia em uma amostragem dos *pixels* como forma de otimização do processo de codificação e comparação das imagens. Segundo, que o objetivo desta técnica não é encontrar arquivos absolutamente idênticos, mas garantir certo grau de variação para identificar variações de “uma mesma” imagem. Em função do primeiro fator, a técnica possui uma fragilidade específica que é o tratamento de imagens com grandes áreas sem variação cromática – com cor chapada – como é o caso da *logo* da TV Record, incluída no exemplo. O fundo branco, sem variação, aumenta a possibilidade de a imagem ser tida como muito próxima de outra imagem com esta característica. Possivelmente, sua inclusão no grupo se deveu à recorrência desta característica na ilustração derivada do retrato.

Como na discussão realizada no capítulo 2 (veja [2.2.3 Multiplicidade ontológica: declinações materiais da imagem](#)), percebe-se com este caso, mais uma vez, como a suposta indiferenciação do código digital, que se aplicaria a todo tipo de conteúdos e de imagens de forma indistinta, não se verifica. De modo similar à especialização do algoritmo de compressão JPEG a certos tipos de imagem, a técnica do *hashing* perceptivo também manifesta seu grau de especialização e viés no tratamento das imagens e, portanto, da elaboração performativa daquilo que compreende como identidade ou como diferença. Este tipo de problema é particularmente perceptível em alguns grupos de imagens provenientes de impressões de tela ou de desenhos gráficos. Especialmente nesses casos, a pertinência dos agrupamentos deve ser criticamente considerada para evitar sua consideração inadvertida como *imagens únicas* (veja [Figura 29](#)).



Figura 29: Montagem de matriz e instâncias de um agrupamento mal formado.

□ VGG19 I:2 | □ GVAPI Q:10

Fonte: Elaboração própria.

A tolerância do método a variações, como indiquei para a Figura 28, contribui, por outro lado, a uma compreensão flexível da identidade das imagens. Esta flexibilidade permite considerar como *uma* imagem as variações produzidas a partir de uma fotografia, bem como montagens que depositam outras camadas de sentido a estas variações. O agrupamento exibido na Figura 27 também ilustra este aspecto, uma vez que são consideradas parte do mesmo grupo instâncias com variações pequenas, como variações de tratamento de cor, sobreposição de logo do veículo (BBC Brasil) ou, como na matriz do grupo, com pequenas montagens fotográficas. Embora se trate, claramente, de *outra* imagem, considerar tais instâncias conjuntamente permite que elas sejam tomadas como parte de um mesmo contexto, em que a imagem é colocada no centro de um embate discursivo.



Figura 30: Impressão de tela do Atlas com imagens similares consideradas distintas

□ VGG19 A:16-B:16 | □ GVAPI I:16-J:17

Fonte: Elaboração própria.

Em outros casos também há imagens que, embora pudessem ser consideradas as *mesmas*, para fins analíticos, acabam por aparecer dispersas em agrupamentos distintos (Figura 30). Esta dispersão revela outro aspecto das identidades e diferenças performa-

das pela técnica do *hashing* perceptivo, por uma alta sensibilidade a pequenos cortes e reenquadramentos das figuras, que afetam a consistência da amostragem dos *pixels* entre as diferentes versões. Isto é: os cortes parecem fazer com que as amostras extraídas para a identificação ocorram em porções distintas da imagem. Esta dispersão é, em larga medida, compensada pela disposição das imagens a partir das características extraídas pelo modelo VGG19 – efetivamente, elas se encontram lado a lado no mapa produzido por esse método.

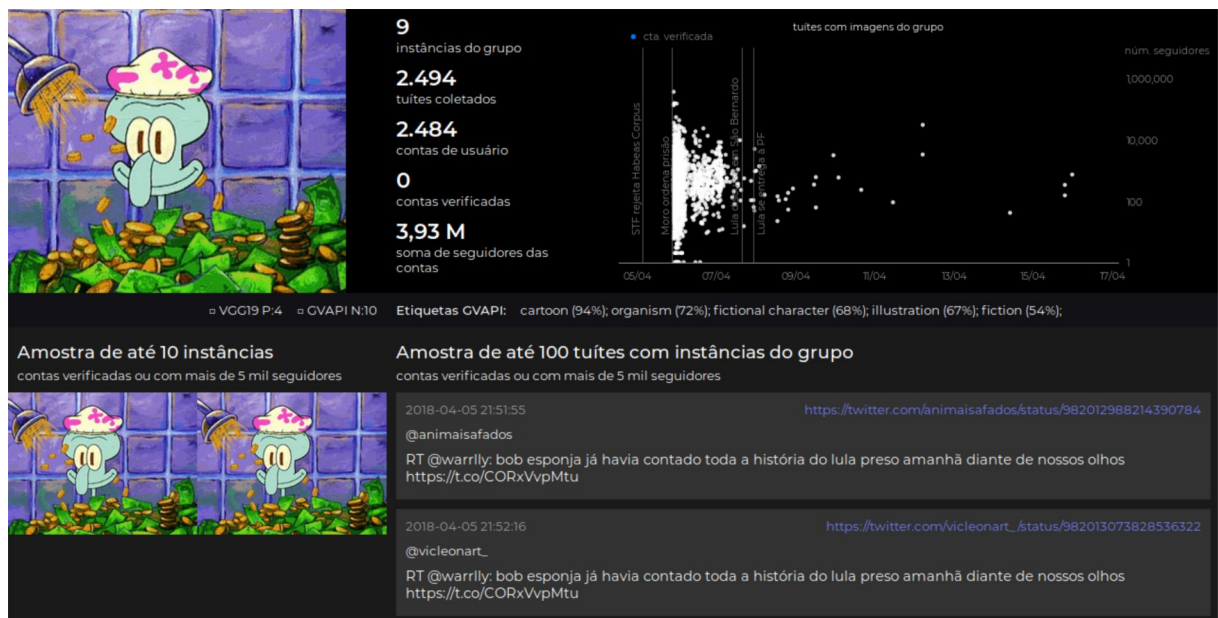


Figura 31: Impressão de tela com imagem do personagem Lula Molusco
□ VGG19 P:4 | □ GVAPI N:10

Fonte: Elaboração própria.

Ainda na página de detalhamento, podemos observar as etiquetas atribuídas à imagem pela GVAPI. Nos exemplos das Figuras 27 e 28, observa-se, por exemplo, uma classificação que, consistentemente, atém-se à descrição generalista de partes do rosto exibidas nas imagens – *hair* (cabelo), *facial hair* (pelo facial), *forehead* (testa), *person* (pessoa). Em um outro exemplo, trazido na Figura 31, temos uma imagem extraída da série animada *Bob Esponja*, com o personagem “Lula Molusco”. Imagens deste programa de televisão foram consistentemente utilizadas como parte do *meme* “Lula preso amanhã”, em antecipação da prisão do ex-presidente. O ponto que quero destacar nesse exemplo é a imprecisão das etiquetas atribuídas pela GVAPI, que não chegam a descrever o que o desenho representa, mas apenas: *cartoon* (cartum ou desenho animado), *fictional character* (personagem fictício), *illustration* (ilustração) e *fiction* (ficção). Apenas a etiqueta *orga-*

nism (organismo) chega mais próximo de uma descrição do “conteúdo” do desenho, mas mesmo ela é demasiado genérica. Este caso salienta, portanto, como a classificação das imagens pelo GVAPI volta-se principalmente a certos tipos de imagem em detrimento de outros – fotografias, não desenhos. Este é outro contraexemplo à consideração por Philippe Dubois (2017), que discuti anteriormente (veja [2.1 A imagem e o digital](#)), de que o digital trataria de forma indistinta todo tipo de imagem. Em todo caso, evidentemente, estas descrições, por si só, dizem pouco sobre as imagens. Mas a utilização dos dados da GVAPI como elementos que estabelecem relações entre elas permite um olhar para o conjunto das imagens, como indicado na seção anterior (veja [4.2 Processamento das imagens](#)).

O gráfico de dispersão apresentado na página de detalhamento, por sua vez, oferece indícios que permitem compreender a dinâmica de circulação da imagem, complementada pelo contexto dos tuítes originais, inseridos logo abaixo. A [Figura 32](#) apresenta um caso em que estes dados são fundamentais para compreender a imagem em análise. A fotografia que deriva no grupo de arquivos considerados é, certamente, o registro mais conhecido dos acontecimentos que antecederam a prisão e, como divulgado à época (cf. QUEIROGA, 2018), foi produzida por Francisco Proner Ramos, fotógrafo com então 18 anos, que encontrava-se no prédio do SMABC em função da proximidade entre sua mãe, Carol Proner, que é professora de direito da UFRJ, e o ex-presidente. A imagem, como indica o gráfico de dispersão da página de detalhamento ([Figura 32](#)) foi publicada intensamente no Twitter momentos após o registro, que foi feito ao final do discurso de Lula à frente do SMABC. A distribuição das publicações ao longo do tempo sugere, também, “ondas” de republicação ao longo do período coberto pelos dados, com muita intensidade nos dois primeiros dias, inclusive entre contas verificadas pelo Twitter. O número de publicações decai progressivamente ao longo da semana mas tem uma retomada por volta do dia 16 de abril. A amostra dos tuítes exhibe seu compartilhamento em diferentes idiomas – português, inglês, espanhol, francês e italiano – um aspecto que também se manifestou em mídias tradicionais. A fotografia ilustrou matérias jornalísticas sobre a prisão em jornais como o francês *Le Monde* e o estadunidense *The New York Times* (cf. FERNANDES, 2018; QUEIROGA, 2018).

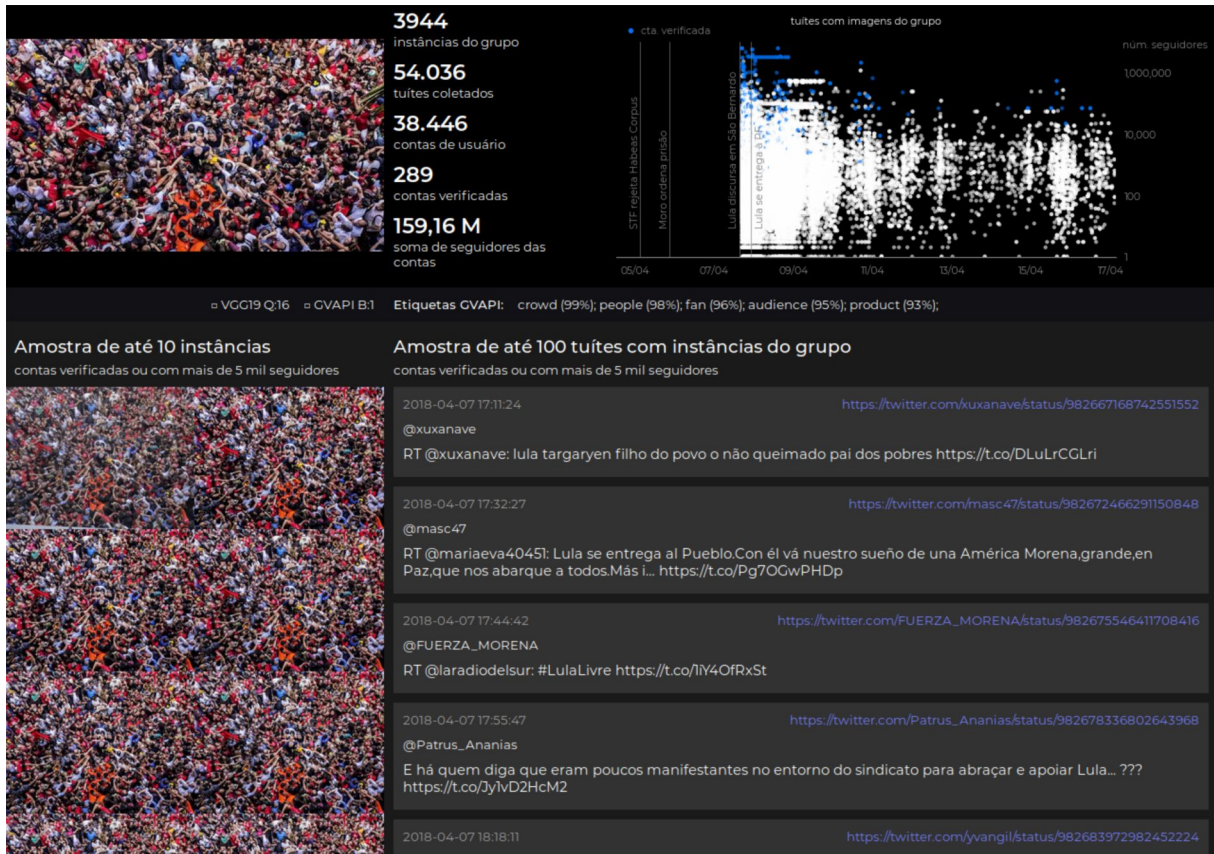


Figura 32: Impressão de tela do Atlas com fotografia de Proner Ramos.
□ VGG19 Q:16 | □ GVAPI B:1

Fonte: Elaboração própria.

Deve-se notar que o agrupamento detalhado na página considerada aqui não é o único relativo à fotografia, que aparece em diversos outros grupos situados em seu entorno nos mapas produzidos tanto pelo modelo VGG19 quanto pela GVAPI. A imagem também foi apropriada em formações meméticas que podem ser encontradas em outras partes dos mapas. A Figura 33 apresenta dois memes derivados desse registro. À esquerda, a fotografia de Lula carregado pela multidão é aproximada de um plano da série televisiva de ficção *Guerra dos Tronos* (*Game of Thrones*) do canal estadunidense HBO, em que a personagem Daenerys Targaryen é ovacionada por uma população escravizada que ela havia recém-libertado. À direita, realizou-se uma montagem que relaciona a fotografia à apresentação que o procurador responsável pela acusação que levou à condenação de Lula, Deltan Dallagnol, fez à imprensa. Especificamente, o diagrama sobreposto à imagem é extraído de um que foi utilizado por Dallagnol em coletiva à imprensa sobre a acusação imputada a Lula e que sofreu críticas por voltar-se mais um efeito retórico do que à apresentação de evidências.

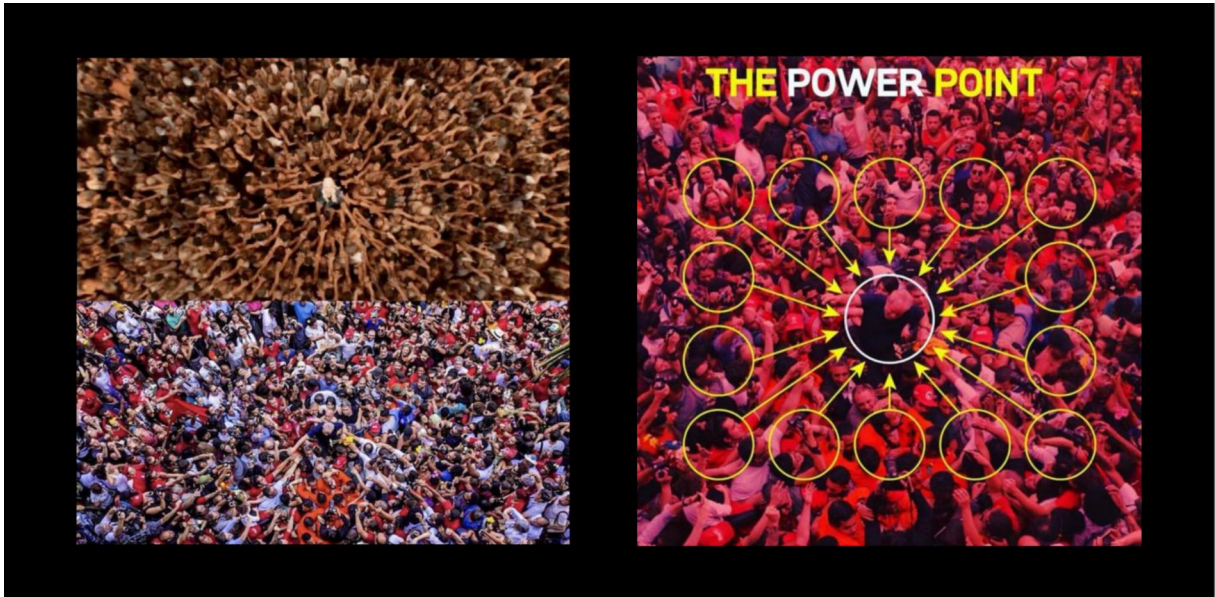


Figura 33: Memes derivados da fotografia de Proner Ramos.
 Esquerda: □ VGG19 Q:3 | □ GVAPI D:3. Direita: □ VGG19 Q:3 | □ GVAPI M:8

Fonte: Montagem de elaboração própria. Imagens obtidas pela API Streaming do Twitter.

Efetivamente, dada sua ampla circulação, a imagem de Francisco Proner Ramos tornou-se um registro emblemático do acontecimento a que se volta esse estudo de caso. Uma análise que se voltasse a questões empíricas específicas do caso deveria, desse modo, lhe dedicar uma boa parte do esforço de investigação. O interesse de minha discussão, contudo, é mais metodológico e, por isso, não vou me aprofundar muito nas especificidades do caso. O argumento que quero desenvolver, nesse sentido, é o de que uma breve incursão pelas imagens desse acontecimento, em especial esta fotografia, apontam para a necessidade de desenvolvermos habilidades de navegação pelas imagens e pelos dados que vão além tanto da consideração detida de uma imagem em isolamento quanto de seu tratamento agregado em índices estatísticos. De modo importante, o tratamento teórico que concederíamos à fotografia não poderia se bastar na consideração do instante crítico de seu registro e parece demandar, de outro modo, que percorramos os meandros de seu processo de circulação e as relações que vão sendo estabelecidas no caminho. A imagem é um *ator* importante desse processo, certamente, mas é importante observar que o percurso que a leva do momento do registro até sua impressão em meia página do *The New York Times* (cf. QUEIROGA, 2018) passa por uma série de outras mediações – inclusive por declinações da própria imagem em outras versões, como nos *memes* apresentados. Quero sugerir, portanto, que a fotografia precisaria ser compreendida segundo suas de-

pendências e segundo a multiplicidade de suas instanciações. Como uma *imagem-rede*, portanto, cujo mapeamento ontográfico é apoiado, neste exercício, pela reapropriação crítica de técnicas de aprendizado de máquina.

Um efeito do aprendizado de máquina que é possivelmente produtivo a um esforço analítico como o esboçado para este caso são as associações heurísticas que emergem na superfície do *Atlas* e que favorecem uma abordagem relacional das imagens. Desdobrar estas possibilidades em uma análise empírica vai além dos objetivos deste estudo de caso, em vista das questões perseguidas pela tese. Em todo caso, as Figuras 34 a 39 apresentam detalhes extraídos dos mapas do *Atlas* que ilustram alguns destes potenciais caminhos da análise que não cheguei a desenvolver textualmente aqui, mas que acredito serem elucidativos das dinâmicas que sugeri para este caso, ao início do capítulo, como um *iconoclash*. Como discuti mais acima, um *iconoclash*, segundo sugere Latour (2008), seria uma situação em que o ímpeto de destruir uma imagem de devoção se realiza pela a produção de outra imagem. Trata-se de uma tensão manifesta no caso da prisão do ex-presidente Lula em que, dada a antecipação do caráter histórico do acontecimento, travou-se uma disputa pelo controle da produção de seus registros. As imagens que compõem o *Atlas* oferecem, em alguma medida, testemunhos desse processo. Em particular, as associações construídas pela disposição das imagens e sua superfície, que-ro sugerir, auxiliam esforços de descrição das estratégias empregadas durante a produção e circulação das imagens, por temas reiterados e remissões formais a registros históricos.



Figura 34: Montagens de Lula atrás das grades.

□ VGG19 A:3.

Fonte: Detalhe do Gráfico 5.



Figura 35: Variações sobre o retrato de identificação de 1980.

□ VGG19 A:1-B:1

Fonte: Detalhe do Gráfico 5.



Figura 36: Depoimentos de testemunhas e réus da Lava Jato.

□ VGG19 E:7-F:7

Fonte: Detalhe do Gráfico 5.



Figura 37: Faixas da campanha pela liberdade de Lula.
□ VGG19 M:10-0:10.

Fonte: Detalhe do Gráfico 5.



Figura 38: Chegada de Lula na sede da PF em Curitiba.
Diversos registros feitos à distância de Lula descendo do helicóptero.
□ GVAPI B:9

Fonte: Detalhe do Gráfico 8.



Figura 39: Figurações de Lula na multidão.
À esquerda, esforços para atravessar a multidão para entregar-se à PF. No topo, ao centro, Lula como líder sindicalista, nos anos 1970 e 1980. No topo, à direita, e ao longo da linha inferior, Lula carregado pela multidão após discursar no dia da prisão. Na linha do meio, comícios e eventos de momentos diversos.
□ VGG19 I:17-J17

Fonte: Detalhe do Gráfico 5.

Evidentemente, esta proposição metodológica não “resolve” as incertezas do estudo de imagens por métodos computacionais, nem poderia pretender fazê-lo. Sugiro, de outro

modo, que ela seria um exercício para compreender de que forma os métodos computacionais e mediações algorítmicas poderiam ser trazidas a contextos de pesquisa em comunicação e mídia em um plano de *convivência* com essas incertezas e com as especificidades desse campo – isto é, sem subsumi-lo a perspectivas quantitativas ou positivistas. A partir deste ponto, a reapropriação crítica do aprendizado de máquina poderia proporcionar outros caminhos de reflexão que não se encontram necessariamente prescritos pelo tema da computacionalidade mas que, acredito, com ele se relacionam. Gostaria de salientar dois pontos que emergem a partir dos esforços empreendidos nas últimas páginas e que, ao que me parece, poderiam orientar análises realizadas por meio do *Atlas*.

Primeiro, indico as questões relativas ao tema da *circulação*, que constitui hoje, muito em função das plataformas *online*, um tópico de intensa discussão nos estudos de comunicação. Antônio Fausto Neto (2010, 2018) tem, no contexto brasileiro, dedicado esforços de sistematização teórica e conceitual sobre esse tema, inclusive destacando como as mídias digitais ofereceriam um contexto de possível mapeamento empírico da circulação. Laura Guimarães Corrêa (2017) também reflete sobre estas dinâmicas em um estudo empírico que se aproxima de algumas questões indicadas pelo estudo de caso apresentado neste capítulo, ao se voltar para imagens circuladas na mídia durante o processo de *impeachment* da ex-presidenta Dilma Rousseff. Em particular, ela analisa como elas conformaram figurações misóginas emblemáticas daquele acontecimento. Hito Steyerl (2009, 2013), em perspectiva bastante distinta, também afirmou em certo momento, a circulação como um dos aspectos mais salientes do modo de existência das imagens contemporâneas. Diria que estes diferentes enfrentamentos do tema da circulação poderiam ser elaborados empiricamente a partir do *Atlas* e de protocolos metodológicos similares.

Segundo, destaco a temática dos *memes*, que constitui outro tema bastante relevante para os estudos contemporâneos em comunicação e mídia (cf. CHAGAS et al., 2017; SHIFMAN, 2013; WIGGINS; BOWERS, 2015). Embora um *meme* não implique, necessariamente, conteúdo visual, já que eles podem ser verbais, por exemplo (SHIFMAN, 2013), é principalmente nessa forma que eles colocam os maiores desafios metodológicos às pesquisas. As discussões realizadas por esta investigação contribuem ao enfrentamento de alguns destes desafios, mas não chega a avançar muito em um nível técnico, acredito, dada a especialização de modelos de reconhecimento de imagens, como indiquei, para o

processamento de imagens fotográficas. *Memes* visuais são comumente formados pela justaposição de mais de uma fotografia e por montagens que incluem textos verbais, por exemplo. Contudo, uma relação importante entre a discussão que elaborei aqui e o tema dos *memes* diz respeito ao modo com que eles são também objetos ontologicamente ins-táveis, melhor descritos como efeitos de dinâmicas associativas complexas. Uma de suas consequências para a teoria da fotografia, segundo sugere Limor Shifman (2014) em uma discussão sobre *memes* fotográficos, seria a relativização de seu vínculo primordial com um referente, no passado, e a ênfase em sua capacidade de recombinação, no futuro. Este é outro caminho investigativo que seria sugerido a partir do *Atlas* e das discussões desta tese, mas que precisaria ser objeto de estudos futuros. Assim como o *Atlas* se constitui em um gesto de reapropriação, portanto, também ele poderia ser *reapropriado* em percursos derivados.

5 Considerações finais

Esta investigação assume, de início, uma orientação metodológica. Fundamentalmente, ela busca articular elementos para o início da constituição de um referencial teórico para o estudo de imagens por meio de sistemas de aprendizado de máquina por redes neurais. Em um diálogo contrastante com vertentes que tendem a essencializar os métodos e os dados computacionais, a perspectiva adotada nesta tese buscou enfatizar o caráter mediador dos métodos empregados e as instabilidades e incertezas dos dados que produzem. Justamente por isso, mais do que questões a serem definitivamente respondidas neste momento, quero compreender as perguntas geradoras da investigação como indagações a serem reiteradas em investigações baseadas nesses métodos. Considero que elas devem ser parte de uma abordagem autorreflexiva que é, ao final, aquela a que esta tese visa contribuir.

Enunciado à Introdução, o problema proposto para esta pesquisa interroga-se quanto: ao que fazemos quando estudamos imagens por métodos computacionais, em especial pelo aprendizado de máquina; ao modo com que estes métodos transformam e produzem suas imagens; e ao que podemos conhecer das imagens por meio deles. Evidentemente, qualquer resposta a estas questões deve reconhecer a circunstancialidade inerente a cada investigação. Logo, não quero sugerir, aqui, respostas definitivas. Porém, o percurso elaborado nestas páginas sugere a possibilidade de respostas em alguma medida perenes sobre aspectos particulares do aprendizado de máquina em sua constituição de *visualidades computacionais*. A este respeito, um ponto salientado na reflexão e que gostaria de destacar nestas considerações finais, diz respeito ao modo com que a operação do aprendizado de máquina transforma ativamente as imagens, em uma complexificação da multiplicidade ontológica desses objetos. Esta compreensão atravessa o arco argumentativo deste trabalho e ganha corpo, especificamente, com a hipótese da *imagem-rede*, uma formulação conceitual que, segundo quero argumentar, ganha força à medida que incide sobre o tratamento teórico das imagens em diferentes contextos.

Em uma primeira estância dessa elaboração, propus no capítulo 2 (Imagem-rede) uma revisão do tema da tecnicidade e da materialidade das imagens. O enfoque específico da

investigação são as imagens digitais, mas procurei também demonstrar como a aparente “crise” que a digitalidade provocaria sobre a questão talvez devesse ser encarada como o sintoma de um problema mais arraigado nas teorias das imagens. A noção de *imagem-rede*, nesse contexto, sugere uma compreensão mais complexa das mediações técnicas que participam da produção das imagens, sem subsumi-las por uma individualidade essencializada. A *rede*, nessa acepção, apoia-se em formulações da *Teoria Ator-Rede* (TAR), compreendendo o caráter distribuído da ação e incluindo os objetos técnicos no rol dos *atores* considerados (CALLON, 1986; LATOUR, 2001, 2005; LAW, 1999, 2009). A *imagem*, sob tal perspectiva, deveria ser compreendida enquanto *inscrição sociotécnica*, guardando uma vinculação com os objetos técnicos e os arranjos institucionais que a produzem (AKRICH, 1992; AKRICH; LATOUR, 1992; LATOUR, 1986, 2001). Mas, enquanto inscrição, a imagem também pode ser *exportada* para outros contextos, inclusive por meio de sua declinação em outras inscrições que, segundo suas características, proporcionariam diferentes gestos de conhecimento, segundo outras chaves interpretativas e outras operações de justaposição e recombinação de suas manifestações materiais. Em vista destas diferentes declinações, a imagem seria compreendida, portanto, por uma *multiplicidade ontológica*, o que implica tomá-la não por uma ontologia determinada mas, sim, como efeito emergente de múltiplos modos de existência (MOL, 1999). A *imagem-rede*, neste sentido, busca ressituar o objeto da investigação, que não seria um ponto determinado, mas um ente distribuído que precisaria ser mapeado em um exercício *ontográfico* (LYNCH, 2013; MENOTTI, 2019).

Em meio a esta revisão teórica, fundada em uma problemática mais voltada aos Estudos de Ciência e Tecnologia (STS), uma breve retomada de aspectos do trabalho de Aby Warburg buscou evidenciar as articulações desta discussão com temas mais diretamente relacionados às disciplinas que se voltam ao estudo das imagens (LISSOVSKY, 2014; MICHAUD, 2013; WARBURG, 2015). A proposta, nesse sentido, foi realizar uma leitura interessada do trabalho de Warburg em uma aproximação livre com noções advindas da TAR e dos STS. Gostaria de destacar três aspectos por reter desta aproximação. Primeiro, a concepção da imagem como um objeto relacional que emerge nos escritos e nos esforços metodológicos warburguianos, em especial no seu projeto do *Atlas Mnemosyne* e na noção, que discuti apenas tangencialmente, da fórmula de páthos (*Pathosformel*). Segundo, como seu olhar para a história das imagens enfatizou aspectos de sua constituição material, tratados como *veículos de imagem* (*Bilderfahrzeuge*), que com-

preende como fatores condicionantes da constituição de estilos de época e da formação de temas visuais recorrentes. Terceiro, como o próprio método warburgiano era devedor dos veículos de imagem de seu tempo, uma vez que o projeto inacabado do *Atlas Mnemosyne* constituía-se por meio de reproduções fotográficas de imagens que o estu-
dioso e seus colaboradores acumulavam e manipulavam. A retomada do método warburgiano constitui, assim, uma segunda estância de desenvolvimento conceitual, em que a noção de *imagem-rede* vem incidir sobre práticas situadas de investigação que, a seu próprio modo, produzem, por um lado, outras declinações materiais da imagem; e, por outro, composições situadas particulares que sugeri abordar pela noção de *ontografia*.

No capítulo 3 (*Visualidades computacionais*), as questões pertinentes a um arco conceitual abrangente da *imagem-rede* voltam a se particularizar. Desta vez, segundo os modos de processamento computacional da imagem. Assumindo a *multiplicidade ontológica* como um princípio metodológico, um primeiro movimento deste capítulo propõe uma revisão crítica da Analítica Cultural (MANOVICH, 2009, 2012; TIFENTALE; MANOVICH, 2015), indicando como seu tratamento da imagem tende a essencializar uma de suas declinações específicas: como *dado* computacional. Então, em uma terceira estância de desenvolvimento do conceito, a noção de *imagem-rede* sugere dinâmicas relacionais de constituição das imagens pela operação das técnicas de aprendizado de máquina por redes neurais. O processo de treinamento destas máquinas indutivas opera pela condensação de um grande volume de dados a fim de extrair um *modelo* que seria capaz de descrevê-los (BROUSSARD, 2018; CARDON; COINTET; MAZIÈRES, 2018; MACKENZIE, 2017). As imagens são tratadas, em seu domínio, como inscrições sociotécnicas particulares: como *vetores* que definem as coordenadas de posições relativas em um espaço vetorial multidimensional. O modelo, portanto, constitui-se por meio de procedimentos complexos de manipulação destas inscrições, produzidas a partir de milhares ou milhões de imagens. Tais procedimentos constroem uma espacialidade complexa em que atuam tanto as características destas imagens quanto aspectos da arquitetura das redes neurais empregadas. As *classes* constituídas por esse processo são, deste modo, condensações de múltiplas instâncias individuais – efeitos emergentes, poderíamos dizer, de uma rede de imagens. Ao longo das várias camadas e “neurônios” da rede neural, aspectos destas imagens e das classes a que são vinculadas, também se inscrevem, de forma distribuída, em características visuais abstratas. Por sua vez, o processo de reconhecimento de uma

imagem por meio de um tal modelo articula um novo dado visual às representações condensadas das imagens utilizadas no treinamento. A noção de *imagem-rede*, nesse contexto, aportaria um tratamento conceitual para a inscrição difusa das imagens ao longo das redes. Sugeriria uma compreensão complexa, também, para as classes atribuídas por um modelo a uma imagem específica, as quais derivam da atribuição, àquela imagem, de uma posição relativa no espaço vetorial “habitado” pelas máquinas indutivas.

Em outro eixo conceitual da tese, a discussão do tratamento computacional das imagens e das técnicas de aprendizado de máquina, em especial, é articulada à problemática das *visualidades* (FOSTER, 1999). A esse respeito, argumentei que o aprendizado de máquina poderia ser compreendido como instância concreta de operação de uma *visualidade computacional* contemporânea, na medida em que modelos de reconhecimento de imagens operariam, simultaneamente, como expressões condensadas da cultura visual *online*, considerando sua dependência das bases imagens utilizadas em sua base de treinamento; e, também, sua integração como instâncias de mediação algorítmica da datificação e da circulação das imagens. Esta implicação se adensa quando considerado o processo de *infraestruturalização* (DIJCK; POELL; WAAL, 2018; HELMOND; NIEBORG; VLIST, 2019; PLANTIN et al., 2016) desses modelos e das bases de treinamento utilizadas para construí-los. Neste processo, mais do que visualidades particulares, os modelos constituídos por técnicas de aprendizado de máquina operariam como reiterações de assimetrias sociais (BROUSSARD, 2018; BUOLAMWINI; GEBRU, 2018; SILVA, 2019), materializadas em uma concentração de recursos técnicos infraestruturais. Este aspecto é salientado como uma implicação política e epistêmica da utilização de modelos de reconhecimento de imagem, em um tensionamento da prática de sua *reapropriação* metodológica.

Esse percurso já permitiria algumas respostas às perguntas geradoras desta investigação. Ao estudarmos imagens por métodos computacionais, multiplicamos suas declinações materiais, em inscrições computacionais que atendem a demandas específicas e que, por vezes, se distanciam de nossa experiência das imagens em seus contextos de produção, circulação e observação. Contudo, estas transformações não são, necessariamente, exógenas a tais contextos. Em muitos casos, em especial no âmbito das plataformas *online*, afinal, elas são parte integrante, embora oculta, dos mesmos contextos de produção, circulação e observação. É importante se atentar, contudo, a como as imagens não atravessam incólumes a tais processos, que operam sobre elas por um mo-

vimento de *translação*, deslocando-as a programas de ação particulares, inscritos em sua forma de mediação. Mais do que uma simples *aplicação* metodológica, portanto, seria preciso assumir uma postura reflexiva em um gesto de *reapropriação*. Por meio dela, as ambivalências e contradições das mediações técnicas empregadas são tomadas como parte integrante da investigação. Por fim, ainda em uma remissão às perguntas geradoras da pesquisa, a reapropriação metodológica do aprendizado de máquina nos permitiria observar as imagens segundo uma visualidade maquina, articulando, de forma híbrida, diferentes “modos de ver” que habitam, contemporaneamente, o campo visual. Desde que, evidentemente, os dados do aprendizado de máquina não sejam tomados, em um gesto redutor, como representantes plenos das imagens estudadas mas, sim, como inscrições que compõem a multiplicidade ontológica de uma imagem-rede.

Como indiquei ao início destas considerações finais, estas respostas não se pretendem definitivas e, na verdade, elas descrevem mais *desafios* do que *respostas*. Em vista deles, o capítulo 4 (*Estudo de caso: imagens de uma prisão*) propõe dar maior concretude às discussões que lhe antecedem, em um exercício metodológico situado. Nesse esforço, alguns aspectos que não chegaram a ser antecipados pelas discussões precedentes ganham proeminência. Destaco, primeiramente, a complexidade das operações demandadas como passos preliminares para que fosse possível conceder às imagens uma primeira condição de existência em meio aos dados coletados do Twitter. Pois, em larga medida, os procedimentos de coleta e processamento podem ser lidos como esforços para contornar a constituição verbocêntrica da plataforma e dos métodos de coleta e análise disponíveis – um sintoma, por assim dizer, da constituição sociotécnica das mídias digitais. Em um segundo aspecto, o estudo também precisou se haver com o problema da *identidade* das imagens, dando concretude ao argumento sustentado por Lynch (2013) acerca de uma simetria entre *identidade* e *diferença*. Em um terceiro aspecto, o estudo de caso ofereceu uma situação específica para se observar os potenciais da mobilização de modelos de reconhecimento de imagem baseados em aprendizado de máquina como recursos metodológicos. Mesmo sem chegar a desenvolver uma análise aprofundada, o estudo permitiu observar o rendimento desta reapropriação metodológica como uma mediação que proporciona uma possibilidade de orientação em meio a uma grande quantidade de imagens. Especificamente, ao agrupar imagens semanticamente ou formalmente relacionadas, segundo as classes e características inscritas nos modelos utilizados.

De modo a aprofundar tal exercício metodológico, ainda no capítulo 4, a tese se voltou a uma proposição metodológica específica, denominada *Atlas para imagens-redes* (veja 4.3 *Compondo imagens-redes* e Apêndice B). Confluência das discussões elaboradas ao longo do texto, o *Atlas* é concebido como uma construção tentativa que busca se haver com os desafios colocados, pela pesquisa, ao estudo das imagens. O objetivo principal é constituir uma mediação capaz de proporcionar uma experiência de navegação heurística, porém evitando operar reduções abruptas dos objetos analisados. Nesse sentido, quero salientar a compreensão do *Atlas* como *confluência* dos diferentes meandros da discussão desta tese, mas não como um *sorvedouro*, que viria a forçar a submersão e, portanto, o obscurecimento, das incertezas do método – deixando apenas objetos definidos, e bem formados, à superfície. Este é o sentido que busco ao descrevê-lo como um gesto de *composição* (LATOURE, 2010). O *Atlas* viria ao auxílio de um olhar que busca recompor imagens-redes sem, porém, essencializá-las em individualizações bem definidas. Trata-se, nesse sentido, do exercício de uma abordagem *ontográfica*, que poderia inspirar procedimentos metodológicos similares mas que não se encerra como um método fechado e infinitamente replicável – como uma “receita de bolo”.

Embora, por sua denominação, o campo dos Métodos Digitais pareça se endereçar a um esforço de sistematização de *métodos* de pesquisa, relativamente estruturados e transponíveis a diferentes contextos de investigação, o modo com que esta proposta se desenvolve, desde sua formulação por Richard Rogers (2013), apresenta outros contornos. Como discuti ao longo da tese, os Métodos Digitais se caracterizam por uma abordagem metodológica reflexiva, que se volta simultaneamente ao “mundo” e às mediações digitais por meio das quais esse mundo toma forma e se permite conhecer. Devo admitir, permitindo certo grau de contradição ao argumento que construo neste momento, que boa parte da pesquisa desenvolvida pelo DMI, e iniciativas associadas, são conhecidas justamente pela elaboração programática de “receitas de bolo”. Refiro-me, especialmente, a protocolos de pesquisa replicáveis e ferramentas metodológicas como o DMI-TCAT (RIEDER; BORRA, 2014), ou o já inoperante Netvizz (RIEDER, 2013), que apresentam conjuntos “pré-fabricados” de operações de coleta e análise de dados. Porém, quero sustentar que talvez devamos considerar a aplicação meramente operacional destas ferramentas mais como uma distorção do que uma característica própria ao domínio dos Métodos Digitais. Uma leitura mais detida das investigações realizadas sob o arco desta denominação, inclusive dos artigos que acompanham algumas destas ferra-

mentas, permite ressitua-las como instâncias de materialização de uma reflexão crítica sobre as técnicas e os métodos empregados (cf. JACOMY et al., 2014; RIEDER et al., 2015; RIEDER; RÖHLE, 2012, 2017; VAN ES; WIERINGA; SCHÄFER, 2018; VENTURINI; ROGERS, 2019). Quero dizer que as ferramentas, bem além de simples operacionalizações técnicas, são construções intelectuais que condensam esforços reflexivos em um conjunto de provisões teóricas e metodológicas materializadas como ferramentas.

Inspiro-me nesta compreensão para sugerir, portanto, que embora boa parte dos protocolos sistematizados no capítulo 4 possam vir a subsidiar muitas válidas transposições a outras investigações, não é esse o sentido principal de seu desenvolvimento no âmbito desta tese. Eles não visam resolver os problemas ou as incertezas que permeiam o estudo das imagens no contexto das plataformas online. Quero situá-los, principalmente, como um movimento complementar de reflexão teórica e metodológica, com vistas deslocar as discussões entre diferentes ordens de concretude. Indiquei ao final do último capítulo como o *Atlas* poderia subsidiar estudos derivados. Em particular, segundo os temas da circulação e da memetização, que colocam questões hoje centrais aos estudos em comunicação e das imagens. Porém, estas articulações não supõem o *Atlas* como um instrumento metodológico neutro e, sim, como um operador teórico-metodológico que se vincularia a tais abordagens em um nível além do operacional. Espero que, desse modo, este trabalho possa auxiliar a elaboração de novos caminhos para nos havermos com os desafios colocados, hoje, ao estudo das imagens.

Referências

ABREU, Hortencia Nunes. *Estâncias de uma vida póstuma: função memorativa das imagens na arte contemporânea depois de Aby Warburg*. 177 f. Dissertação (Mestrado) – Escola de Belas Artes, Universidade Federal de Minas Gerais, Belo Horizonte, 2015. Disponível em: <<https://repositorio.ufmg.br/handle/1843/BUBD-A5FLRX>>. Acesso em: 29 set. 2019.

ABRIL, Gonzalo. Tres dimensiones del texto y de la cultura visual. *IC Revista Científica de Información y Comunicación*, Sevilha, n. 9, 2012.

AGAMBEN, Giorgio. *What is a paradigm?* 2002. Disponível em: <<http://www.egs.edu/faculty/giorgio-agamben/articles/what-is-a-paradigm/>>. Acesso em: 2 nov. 2013.

AGAMBEN, Giorgio. O que é um dispositivo? In: *O que é o contemporâneo? e outros ensaios*. Chapeco: ARGOS, 2009. a. p. 25–51.

AGAMBEN, Giorgio. Aby Warburg e a ciência sem nome. *Arte & ensaios*, Rio de Janeiro, v. 16, n. 19, p. 132–143, 2009. b.

AKRICH, Madeleine. The de-scription of technical objects. In: BIJKER, Wiebe E.; LAW, John (Eds.). *Shaping technology/building society: studies in sociotechnical change*. Cambridge: MIT Press, 1992. p. 205–224.

AKRICH, Madeleine; LATOUR, Bruno. A summary of a convenient vocabulary for the semiotics of human and nonhuman assemblies. In: BIJKER, Wiebe E.; LAW, John (Eds.). *Shaping technology/building society: studies in sociotechnical change*. Cambridge: MIT Press, 1992. p. 259–264.

ALBERT, Maximilian et al. *Inkscape*. v. 0.9.2, 2018.

ALLOA, Emmanuel (Org.). *Pensar a Imagem*. Trad. Marianna Poyares et al. Belo Horizonte: Autêntica, 2015.

ALPAYDIN, Ethem. *Machine learning: the new AI*. Cambridge: MIT Press, 2016.

ALZAMORA, Geane; ZILLER, Joana; D'ANDRÉA, Carlos. Mídia e dispositivo: uma aproximação. In: LEAL, Bruno; CARVALHO, Carlos Alberto; ALZAMORA, Geane (Orgs.). *Textualidades midiáticas*. Belo Horizonte: PPGCOM/UFMG, 2018.

ARÈNES, Alexandra; LATOUR, Bruno; GAILLARDET, Jérôme. Giving depth to the surface: An exercise in the Gaia-graphy of critical zones. *The Anthropocene Review*, Thousand Oaks, v. 5, n. 2, p. 120–135, 2018.

AUMONT, Jacques. *A imagem*. 7. ed. Campinas: Papirus, 2002.

BARTHES, Roland. *A câmara clara: nota sobre a fotografia*. Trad. Júlio Castañon Guimarães. Rio de Janeiro: Nova Fronteira, 1984.

- BATCHEN, Geoffrey. Enslaved sovereign, observed spectator: On Jonathan Crary, techniques of the observer. *Continuum: journal of media & cultural studies*, adington, v. 6, n. 2, p. 80–94, 1993.
- BAUDRY, Jean-Louis. Cinema: efeitos ideológicos produzidos pelo aparelho de base. In: XAVIER, Ismail (Org.). *A experiência do cinema: antologia*. São Paulo: Graal, 1983.
- BAZIN, André. Ontologia da imagem fotográfica. In: *O que é o cinema?* Trad. Hugo Mader. São Paulo: Cosac Naify, 2014. p. 27–34.
- BERRY, David M. (Org.). *Understanding digital humanities*. London: Palgrave Macmillan UK, 2012.
- BIJKER, Wiebe E.; LAW, John (Eds.). *Shaping technology/building society: studies in sociotechnical change*. Cambridge: MIT Press, 1992.
- BISHOP, Ryan; GANSING, Kristoffer; PARIKKA, Jussi. Across and beyond: post-digital practices, concepts, and institutions. In: BISHOP, Ryan et al. (Eds.). *Across & beyond: a transmediale reader on post-digital practices, concepts, and institutions*. Berlin: Sternberg Press, 2016. p. 11–23.
- BROUSSARD, Meredith. *Artificial Unintelligence: How Computers Misunderstand the World*. Cambridge: MIT Press, 2018.
- BRUNO, Fernanda. Controle, flagrante e prazer: regimes escópicos e atencionais da vigilância nas cidades. *Revista FAMECOS: mídia, cultura e tecnologia*, Porto Alegre, n. 37, 2008. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/revistafamecos/article/view/4799>>. Acesso em: 16 maio. 2018.
- BRUNO, Fernanda. Rastros digitais sob a perspectiva da teoria ator-rede. *Revista FAMECOS*, Porto Alegre, v. 19, n. 3, p. 681–704, 2012.
- BRUNO, Fernanda. *Máquinas de ver, modos de ser: vigilância, tecnologia e subjetividade*. Porto Alegre: Sulina, 2013.
- BRYANT, Levi; SRNICEK, Nick; HARMAN, Graham (Eds.). *The speculative turn: continental materialism and realism*. Melbourne: Re.Press, 2011.
- BUCHER, Taina. Objects of intense feeling: the case of the twitter api. *Computational Culture*, [s.l.], n. 3, 2013. a. Disponível em: <<http://computationalculture.net/objects-of-intense-feeling-the-case-of-the-twitter-api/>>. Acesso em: 13 ago. 2018.
- BUCHER, Taina. The friendship assemblage: investigating programmed sociality on Facebook. *Television & New Media*, Thousand Oaks, v. 14, n. 6, p. 479–493, 2013. b.
- BUCHER, Taina; HELMOND, Anne. The affordances of social media platforms. In: BURGESS, Jean; MARWICK, Alice; POELL, Thomas (Eds.). *The SAGE handbook of social media*. Thousand Oaks: SAGE Publications, 2018. p. 233–253.
- BUOLAMWINI, Joy; GEBRU, Timnit. Gender shades: intersectional accuracy disparities in commercial gender classification. In: CONFERENCE ON FAIRNESS, ACCOUNTABILITY

AND TRANSPARENCY. *Proceedings of Machine Learning Research 2018*. New York: Association for Computing Machinery, 2018. Disponível em: <<http://proceedings.mlr.press/v81/buolamwini18a.html>>. Acesso em: 15 jun. 2019.

CADÔR, Amir Brito. *O livro de artista e a enciclopédia visual*. Belo Horizonte: Ed. UFMG, 2016.

CALDWELL, Ben et al. Web content accessibility guidelines (WCAG) 2.0. *WWW Consortium (W3C)*, 00290, 2008. Disponível em: <<https://www.w3.org/TR/2008/REC-WCAG20-20081211/#text-equiv-all>>. Acesso em: 4 jul. 2019.

CALLON, Michel. Struggles and negotiations to define what is problematic and what is not. In: KNORR, Karin D.; KROHN, Roger; WHITLEY, Richard (Eds.). *The Social Process of Scientific Investigation*. Dordrecht: Springer Netherlands, 1980. p. 197–219.

CALLON, Michel. Some elements of a sociology of translation: domestication of the scallops and the fishermen of st brieuc bay. *The Sociological Review*, Thousand Oaks, v. 32, n. 1 (suppl.), p. 196–233, 1984.

CALLON, Michel. The sociology of an actor-network: the case of the electric vehicle. In: CALLON, Michel; LAW, John; RIP, Arie (Eds.). *Mapping the Dynamics of Science and Technology: Sociology of Science in the Real World*. New York: Palgrave Macmillan, 1986. p. 19–34.

CALLON, Michel; LAW, John; RIP, Arie (Eds.). *Mapping the dynamics of science and technology: sociology of science in the real world*. New York: Palgrave Macmillan, 1986. a. Disponível em: <<https://link.springer.com/openurl?genre=book&isbn=978-1-349-07410-5>>. Acesso em: 20 mar. 2019.

CALLON, Michel; LAW, John; RIP, Arie. Qualitative scientometrics. In: CALLON, Michel; LAW, John; RIP, Arie (Eds.). *Mapping the dynamics of science and technology: sociology of science in the real world*. New York: Palgrave Macmillan, 1986. b. p. 103–123.

CARDON, Dominique; COINTET, Jean-Philippe; MAZIÈRES, Antoine. Neurons spike back. The invention of inductive machines and the artificial intelligence controversy. *Réseaux*, Paris, v. 211, n. 5, p. 173–220, 2018.

CHAGAS, Viktor et al. A política dos memes e os memes da política: proposta metodológica de análise de conteúdo de memes dos debates eleitorais de 2014. *Intexto*, v. 0, n. 38, p. 173–196, 2017.

CHOLLET, François. *How convolutional neural networks see the world*. 2016. Disponível em: <<https://blog.keras.io/how-convolutional-neural-networks-see-the-world.html>>. Acesso em: 5 set. 2019.

CHOLLET, François et al. *Keras*. v. 2.2.4, 2018.

CHUN, Wendy Hui Kyong. *Control and freedom: power and paranoia in the age of fiber optics*. Cambridge: MIT Press, 2006.

CICALI, Alessandra et al. *Femminicidio in Italian media and public debate*. Lisboa: Universidade Nova de Lisboa, 2018. Disponível em:

<<https://smart.inovamedialab.org/smart-2018/project-reports/project1/>>. Acesso em: 8 out. 2019.

COLOMBO, Gabriele. *The design of composite images: displaying digital visual content for social research*. 200 f. Tese (doutorado) – Politecnico di Milano, Milão, 2018.

COLOMBO, Gabriele. Studying digital images in groups: the folder of images. In: RAMPINO, Lucia; MARIANI, Ilaria (Eds.). *Advancements in design research: 11 PhD theses on design as we do in POLIMI*. Milão: FrancoAngeli, 2019. p. 185–195.

COMOLLI, Jean-Louis. Sob o risco do real. In: *Ver e poder: a inocência perdida: cinema, televisão, ficção*. Belo Horizonte: Ed. UFMG, 2008. p. 169–178.

CORRÊA, Laura Guimarães. O impeachment tem gênero? Circulação de imagens e textos sobre Dilma Rousseff na imprensa brasileira e britânica. In: CASTRO, Paulo César (Org.). *A circulação discursiva entre produção e reconhecimento*. Maceió: Edufal, 2017. p. 279–292.

COUCHOT, Edmond. *A tecnologia na arte: da fotografia à realidade virtual*. Porto Alegre: Ed. UFRGS, 2003.

CRAMER, Florian. What is “Post-digital”? In: BERRY, David M.; DIETER, Michael (Eds.). *Postdigital aesthetics*. London: Palgrave Macmillan UK, 2015. p. 12–26.

CRARY, Jonathan. *Techniques of the observer: on vision and modernity in the nineteenth century*. Cambridge: MIT Press, 1992.

CRAWFORD, Kate; JOLER, Vladan. *Anatomy of an AI system*. 2017. Disponível em: <<http://www.anatomyof.ai>>. Acesso em: 28 set. 2019.

CRAWFORD, Kate; PAGLEN, Trevor. *Excavating AI*. 2019. Disponível em: <<https://www.excavating.ai>>. Acesso em: 30 set. 2019.

CUBITT, Sean. Anecdotal evidence. *NECSUS. European Journal of Media Studies*, Amsterdam, v. 2, n. 1, p. 5–18, 2013.

CUBITT, Sean. *The practice of light: a genealogy of visual technologies from prints to pixels*. Cambridge: MIT Press, 2014.

CUBITT, Sean. *Finite media: environmental implications of digital technologies*. Reprint ed. Durham: Duke University Press, 2017.

D’ANDRÉA, Carlos Frederico de Brito. Cartografando controvérsias com as plataformas digitais: apontamentos teórico-metodológicos. *Galáxia*, São Paulo, n. 38, p. 28–39, 2018.

D’ANDREA, Carlos; MINTZ, André. Studying the live cross-platform circulation of images with computer vision api: an experiment based on a sports media event. *International Journal of Communication*, Los Angeles, v. 13, n. 0, p. 21, 2019.

DASTON, Lorraine; GALISON, Peter. *Objectivity*. Paperback ed. New York: Zone Books, 2010.

DAVIS, Antigone; ROSEN, Guy. Open-sourcing photo- and video-matching technology to make the internet safer. *Facebook Newsroom*, 2019. Disponível em: <<https://newsroom.fb.com/news/2019/08/open-source-photo-video-matching/>>. Acesso em: 4 ago. 2019.

DELEUZE, Gilles. *A imagem-movimento*. Trad. Stella Senra. São Paulo: Brasiliense, 1985.

DELEUZE, Gilles. *A imagem-tempo*. Trad. Eloisa de Araujo Ribeiro. São Paulo: Brasiliense, 1990.

DELEUZE, Gilles. ¿Qué es un dispositivo? In: BARBIER, Tienne et al. (Eds.). *Foucault, filósofo*. Barcelona: Gedisa, 1999. p. 155–163.

DELEUZE, Gilles. *Foucault*. Trad. Claudia Sant'Anna Martins. Sao Paulo: Brasiliense, 2006.

DENG, Jia et al. Imagenet: A large-scale hierarchical image database. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. *Proceedings... 2009*. Miami: IEEE, 2009.

DIAGNE, Cyril; BARRADEAU, Nicolas; DOURY, Simon. *Curator Table / t-SNE Map*. 2018. Disponível em: <<https://experiments.withgoogle.com/t-sne-map>>. Acesso em: 14 out. 2019.

DIDI-HUBERMAN, Georges. *Diante da imagem: questão colocada aos fins de uma história da arte*. Trad. Paulo Neves. São Paulo: Ed. 34, 2013. a.

DIDI-HUBERMAN, Georges. *A imagem sobrevivente: história da arte e tempo dos fantasmas segundo Aby Warburg*. Trad. Vera Ribeiro. Rio de Janeiro: Contraponto, 2013. b.

DIDI-HUBERMAN, Georges. Prefácio. In: MICHAUD, Philippe-Alain. *Aby Warburg e a imagem em movimento*. Trad. Vera Ribeiro. Rio de Janeiro: Contraponto, 2013. c. p. 17–28.

DIDI-HUBERMAN, Georges. *Diante do tempo: história da arte e anacronismo das imagens*. Trad. Vera Casa Nova; Marcia Arbex. Belo Horizonte: Ed. UFMG, 2015.

DIJCK, José Van; POELL, Thomas; WAAL, Martijn De. *The platform society: public values in a connective world*. New York: Oxford University Press, 2018.

DOMINGOS, Pedro. A few useful things to know about machine learning. *Communications of the ACM*, New York, v. 55, n. 10, p. 78, 2012.

DRAGONA, Daphne. What is left to subvert? Artistic methodologies for a post-digital world. In: BISHOP, Ryan et al. (Eds.). *Across & beyond: a transmediale reader on post-digital practices, concepts, and institutions*. Berlin: Sternberg Press, 2016. p. 184–196.

DRAINVILLE, Raymond A. *Algorithmic iconography: Intersections between iconography and social media image research*. 481 f. Tese (doutorado) - Manchester Institute for Research and Innovation in Art and Design (MIRIAD), Manchester Metropolitan University, Manchester, 2018. Disponível em: <<http://e-space.mmu.ac.uk/622714/>>. Acesso em: 7 mar. 2019.

DUBOIS, Philippe. Máquinas de imagens: uma questão de linha geral. In: *Cinema, vídeo, Godard*. Trad. Mateus Araújo Silva. São Paulo: Cosac Naify, 2004. p. 31–67.

DUBOIS, Philippe. *O ato fotográfico e outros ensaios*. Trad. Marina Appenzeller. 14. ed. Campinas, SP: Papyrus, 2012.

DUBOIS, Philippe. Da imagem-traço à imagem-ficção: o movimento das teorias da fotografia de 1980 aos nossos dias. *Discursos Fotográficos*, Londrina, v. 13, n. 22, p. 31–51, 2017.

EDJO LABS et al. *Image-Match*. v. 1.1.2, 2018. Disponível em: <<https://github.com/EdjoLabs/image-match>>. Acesso em: 7 out. 2019.

ELASTICSEARCH B.V. *Elasticsearch*. v. 2.2.1, 2016. Disponível em: <<https://www.elastic.co/pt/>>. Acesso em: 7 out. 2019.

FACEBOOK. *Earlier today, some people and businesses experienced trouble uploading or sending images, videos and other files on our apps and platforms. The issue has since been resolved and we should be back at 100% for everyone. We're sorry for any inconvenience.*@facebook, 2019. Disponível em: <<https://twitter.com/facebook/status/1146571015872552961>>. Acesso em: 4 jul. 2019.

FAROCKI, Harun. Phantom images. *Public*, Toronto, n. 29, 2004. Disponível em: <<https://public.journals.yorku.ca/index.php/public/article/view/30354>>. Acesso em: 3 jun. 2018.

FAULKNER, Simon; VIS, Farida; D'ORAZIO, Francesco. Analysing social media images. In: BURGESS, Jean; MARWICK, Alice; POELL, Thomas (Eds.). *The SAGE Handbook of Social Media*. Thousand Oaks: SAGE Publications, 2018. p. 160–178.

FAUSTO NETO, Antônio. As bordas da circulação. *Alceu. Rio de Janeiro*, v. 10, n. 20, p. 55–69, 2010.

FAUSTO NETO, Antônio. Circulação: trajetos conceituais. *Rizoma*, v. 6, n. 2, p. 08–40, 2018.

FAWCETT, Trevor. Visual facts and the nineteenth-century art lecture. *Art History*, Hoboken, v. 6, n. 4, p. 442–460, 1983.

FERNANDES, Marcos Lúcio. Le Monde critica ausência de “foto emblemática” de Lula em jornais brasileiros. *RFI*, [s. l.], 2018. Disponível em: <<http://br.rfi.fr/brasil/20180413-le-monde-critica-ausencia-foto-lula-jornais-brasileiros>>. Acesso em: 14 abr. 2018.

FERREIRA, Pedro P. Reticulações: ação-rede em Latour e Simondon. *Revista ECO-Pós*, Rio de Janeiro, v. 20, n. 1, p. 104–135, 2017.

FLORES, Luís Felipe Duarte. Reconhecer a imagem, perseguir a história: crítica da visibilidade técnica no cinema de Harun Farocki. In: XXV ENCONTRO ANUAL DA COMPÓS. *Anais... 2016*. Goiânia: Compós, 2016.

FLUSSER, Vilém. *Filosofia da caixa preta: Ensaios para uma futura filosofia da fotografia*. Rio de Janeiro: Relume Dumara, 2002.

- FOSTER, Hal (Ed.). *Vision and Visuality*. Seattle: Bay Press, 1999.
- FOUCAULT, Michel. *Microfísica do poder*. Rio de Janeiro: Graal, 1979.
- FOUCAULT, Michel. *Vigiar e punir*. Petrópolis, RJ, Brazil: Vozes, 1997.
- FROSH, Paul. Inside the image factory: stock photography and cultural production. *Media, Culture & Society*, Thousand Oaks, v. 23, n. 5, p. 625–646, 2001.
- FULLER, Matthew (Ed.). *Software studies: a lexicon*. Cambridge: MIT Press, 2008.
- GALISON, Peter. Images scatter into data, data gather into images. In: LATOUR, Bruno; WEIBEL, Peter (Eds.). *Iconoclasm: Beyond the Image Wars in Science, Religion and Art*. Cambridge: MIT Press, 2002. p. 300–323.
- GALLOWAY, Alexander R. The cybernetic hypothesis. *differences*, durham, v. 25, n. 1, p. 107–131, 2014.
- GEPHI CONSORTIUM. *Gephi*. v. 0.9.2, 2017. Disponível em: <<https://gephi.org/>>
- GIBBS, Martin et al. #Funeral and Instagram: death, social media, and platform vernacular. *Information, Communication & Society*, Abingdon, v. 18, n. 3, p. 255–268, 2015.
- GILLESPIE, Tarleton. The politics of ‘platforms’. *New Media & Society*, Thousand Oaks, v. 12, n. 3, p. 347–364, 2010.
- GINZBURG, Carlo. De A. Warburg a E. H. Gombrich: notas sobre um problema de método. In: *Mitos, emblemas, sinais: morfologia e história*. Trad. Federico Carotti. São Paulo: Companhia das Letras, 1999. p. 41–94.
- GOODFELLOW, Ian et al. Generative adversarial nets. In: ADVANCES IN NEURAL INFORMATION PROCESSING SYSTEMS. *Proceedings... 2014*. Montreal: NIPS, 2014.
- GOOGLE. *Google Cloud Vision API*. Mountain View: Google, 2017.
- GRAF, Alexander et al. *Instaloader*. v. 4.2.7, 2018. Disponível em: <<https://instaloader.github.io/>>
- GRANDJEAN, Martin; JACOMY, Mathieu. Translating networks: assessing correspondence between network visualisation and analytics. In: DIGITAL HUMANITIES CONFERENCE. *Proceedings... 2019*. Utrecht: Alliance of Digital Humanities Organizations, 2019. Disponível em: <<https://halshs.archives-ouvertes.fr/halshs-02179024>>
- GROHMANN, Rafael. A noção de engajamento: sentidos e armadilhas para a pesquisa em comunicação. *Revista FAMECOS*, Porto Alegre, v. 25, n. 3, 2018. Disponível em: <<http://revistaseletronicas.pucrs.br/ojs/index.php/revistafamecos/article/view/29387>>. Acesso em: 5 set. 2018.
- GRUSIN, Richard. *Premediation: affect and mediality after 9/11*. London: Palgrave Macmillan UK, 2010.
- GRUSIN, Richard (Ed.). *The nonhuman turn*. Minneapolis: Univ Of Minnesota Press, 2015.

GUATTARI, Félix. A paixão das máquinas. *Cadernos de Subjetividade*, São Paulo, v. 1, n. 1, p. 39–52, 2003.

GUIMARÃES, César. O novo regime do visível e as imagens digitais. In: VAZ, Paulo Bernardo; CASA NOVA, Vera (Eds.). *Estação imagem: desafios*. Belo Horizonte: Ed. UFMG, 2002. p. 147–161.

HARAWAY, Donna J. Manifesto ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX. In: TADEU, Tomaz (Org.). *Antropologia do ciborgue: as vertigens do pós-humano*. 2. ed. Belo Horizonte: Autentica, 2013. p. 35–118.

HARVEY, Adam. *MegaPixels*. 2019. Disponível em: <<https://megapixels.cc/>>. Acesso em: 13 jul. 2019.

HEIL, Axel; OHRT, Roberto (Eds.). *Aby Warburg Mnemosyne Bilderatlas: reconstruction – commentary – revision*. Karlsruhe: ZKM, 2016.

HELMOND, Anne. The platformization of the web: making web data platform ready. *Social Media + Society*, Thousand Oaks, v. 1, n. 2, 2015. Disponível em: <<http://journals.sagepub.com/doi/10.1177/2056305115603080>>. Acesso em: 23 abr. 2017.

HELMOND, Anne; NIEBORG, David B.; VLIST, Fernando N. Van der. Facebook's evolution: development of a platform-as-infrastructure. *Internet Histories*, v. 3, n. 2, p. 123–146, 2019.

HIGHFIELD, Tim; LEAVER, Tama. Instagrammatics and digital methods: studying visual social media, from selfies and GIFs to memes and emoji. *Communication Research and Practice*, Abingdon, v. 2, n. 1, p. 47–62, 2016.

HOELZL, Ingrid; MARIE, Rémi. *Softimage: towards a new theory of the digital image*. Bristol: Intellect, 2015.

HONORATO, Johanna et al. The color of the street: color as images visualization parameters of twitter pictures from Brazilians Manifestations of 2013. 00002, 2014. Disponível em: <<http://www.labic.net/wp-content/uploads/2015/09/The-Color-of-the-Street-DataWiz-2014.pdf>>

HONORATO, Johanna Inácia; CARREIRA, Lia Scarton; GOVEIA, Fábio Gomes. Análise de Big Data pelos Parâmetros de Características Visuais. In: XIX CONGRESSO DE CIÊNCIAS DA COMUNICAÇÃO NA REGIÃO SUDESTE. *Anais... 2014*. Vila Velha: INTERCOM, 2014. Disponível em: <<http://www.portalintercom.org.br/anais/sudeste2014/resumos/R43--1279--1.pdf>>

HUHTAMO, Erkki. *Illusions in motion: media archaeology of the moving panorama and related spectacles*. Cambridge: MIT Press, 2013.

HUSSAIN, Zaeem et al. Automatic understanding of image and video advertisements. In: IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION. *Proceedings... 2017*. Honolulu: IEEE, 2017. Disponível em: <http://openaccess.thecvf.com/content_cvpr_2017/html/Hussain_Automatic_Understanding_of_CVPR_2017_paper.html>. Acesso em: 19 ago. 2019.

IVINS, William M. *On the rationalization of sight: with the examination of three Renaissance texts on perspective*. New York: Da Capo Press, 1975.

JACOMY, Mathieu et al. Forceatlas2, a continuous graph layout algorithm for handy network visualization designed for the gephi software. *PLoS ONE*, San Francisco, v. 9, n. 6, 2014. Disponível em: <<http://dx.plos.org/10.1371/journal.pone.0098679>>. Acesso em: 4 out. 2017.

JOO, Jungseock et al. Visual persuasion: inferring communicative intents of images. In: *IEEE CONFERENCE ON COMPUTER VISION AND PATTERN RECOGNITION (CVPR). Proceedings... 2014*. Columbus: IEEE, 2014. Disponível em: <<http://ieeexplore.ieee.org/lpdocs/epic03/wrapper.htm?arnumber=6909429>>. Acesso em: 7 dez. 2018.

JPEG. *About JPEG*. [s.d.]. Disponível em: <<https://jpeg.org/about.html>>. Acesso em: 11 out. 2019.

KARPATHY, Andrej. *t-SNE visualization of CNN codes*. 2012. Disponível em: <<https://cs.stanford.edu/people/karpathy/cnnembed/>>. Acesso em: 14 out. 2019.

KLINGEMANN, Mario. *Rasterfairy*. [s.l.: s.n.]. Disponível em: <<https://github.com/Quasimondo/RasterFairy>>. Acesso em: 5 out. 2019.

KLINGER, Evan; STARKWEATHER, David. *phash*. 2010. Disponível em: <<https://www.phash.org/>>. Acesso em: 11 set. 2019.

KLUVER, Billy. *Um dia com Picasso: 29 fotografias de Jean Cocteau*. Rio de Janeiro: José Olympio, 2003.

KOGAN, Gene; MATHEWSON, Kyle. *Feature extraction and reverse image search*. 2018. Disponível em: <<https://github.com/ml4a/ml4a-guides/blob/master/notebooks/image-search.ipynb>>. Acesso em: 14 out. 2019.

KOGAN, Gene; OLDFIELD, James. *Image t-SNE*. 2018. Disponível em: <<https://github.com/ml4a/ml4a-guides/blob/master/notebooks/image-tsne.ipynb>>. Acesso em: 14 out. 2019.

KRAUSS, Rosalind E. *O fotográfico*. Barcelona: Gustavo Gili, 2002.

LATOUR, Bruno. Visualization and cognition. *Knowledge and society*, Greenwich, CT, v. 6, n. 1, p. 1–40, 1986.

LATOUR, Bruno. Where are the missing masses? The sociology of a few mundane artifacts. In: BIJKER, Wiebe E.; LAW, John (Eds.). *Shaping technology/building society: studies in sociotechnical change*. Cambridge: MIT Press, 1992. p. 225–258.

LATOUR, Bruno. *Jamais fomos modernos: ensaio de antropologia simétrica*. Trad. Carlos Irineu Da Costa. Rio de Janeiro: Ed. 34, 1994.

LATOUR, Bruno. On Recalling Ant. *The Sociological Review*, v. 47, n. 1, p. 15–25, 1999.

LATOUR, Bruno. *A esperança de pandora: ensaios sobre a realidade dos estudos científicos*. Trad. Gilson César Cardoso De Sousa. Bauru: EDUSC, 2001.

LATOUR, Bruno. *Reassembling the social: an introduction to actor-network-theory*. Oxford: Oxford University Press, 2005.

LATOUR, Bruno. O que é iconoclash? Ou, há um mundo além das guerras de imagem? *Horizontes Antropológicos*, v. 14, n. 29, p. 111–150, 2008.

LATOUR, Bruno. An attempt at a “compositionist manifesto”. *New Literary History*, v. 41, n. 3, p. 471–490, 2010.

LATOUR, Bruno. Networks, societies, spheres: reflections of an actor-network theorist. *International Journal of Communication*, Los Angeles, v. 5, n. 0, p. 15, 2011.

LATOUR, Bruno et al. ‘The whole is always smaller than its parts’ – a digital test of Gabriel Tardes’ monads. *The British Journal of Sociology*, London, v. 63, n. 4, p. 590–615, 2012.

LATOUR, Bruno. *An inquiry into modes of existence: an anthropology of the moderns*. Cambridge: Harvard University Press, 2013.

LATOUR, Bruno (Ed.). *Reset modernity!* Cambridge: MIT Press, 2016.

LATOUR, Bruno. *Facing gaia: eight lectures on the new climatic regime*. Trad. Catherine Porter. Cambridge: Polity, 2017.

LATOUR, Bruno. *Down to earth: politics in the new climatic regime*. Cambridge: Polity, 2018.

LATOUR, Bruno; HERMANDT, Emille. Redes que a razão desconhece: laboratórios, bibliotecas, coleções. In: PARENTE, André (Org.). *Tramas da rede: novas dimensões filosóficas, estéticas e políticas da comunicação*. Porto Alegre: Sulina, 2004. p. 39–63.

LATOUR, Bruno; LENTON, Timothy M. Extending the domain of freedom, or why gaia is so hard to understand. *Critical Inquiry*, Chicago, v. 45, n. 3, p. 659–680, 2019.

LATOUR, Bruno; WEIBEL, Peter (Eds.). *Iconoclash: beyond the image wars in science, religion and art*. Cambridge: MIT Press, 2002.

LAUTENSCHLAEGER, Grazielle. Entre sensores e sentidos: sobre a materialidade da comunicação na artemídia. In: MENOTTI, Gabriel; BASTOS, Marcus; MORAN, Patrícia (Eds.). *Cinema Apesar da Imagem*. São Paulo: Intermeios, 2016. p. 179–200.

LAVAUD, Laurent. *L’image*. Paris: Flammarion, 1999.

LAW, John. After ant: complexity, naming and topology: *the sociological review*, thousand oaks, 1999. Disponível em: <<https://journals.sagepub.com/doi/abs/10.1111/j.1467-954X.1999.tb03479.x>>. Acesso em: 7 maio. 2019.

LAW, John. *After method: mess in social science research*. Abingdon: Routledge, 2004.

LAW, John. Actor network theory and material semiotics. In: TURNER, Bryan S. (Ed.). *The new Blackwell companion to social theory*. Hoboken: Wiley-Blackwell, 2009. p. 141–158.

LAW, John. STS as method. In: FELT, Ulrike et al. (Eds.). *The handbook of science and technology studies*. 4. ed. Cambridge: MIT Press, 2017. p. 31–57.

LE, Quoc V. et al. Building high-level features using large scale unsupervised learning. *arXiv:1112.6209 [cs]*, arXiv: 1112.6209, 2012. Disponível em: <<http://arxiv.org/abs/1112.6209>>. Acesso em: 8 jan. 2015.

LEVIN, Golan. Computer vision for artists and designers: pedagogic tools and techniques for novice programmers. *AI & Society*, Pittsburgh, v. 20, n. 4, p. 462–482, 2006.

LEWIS, Sarah. The racial bias built into photography. *The New York Times*, New York, 2019. Disponível em: <<https://www.nytimes.com/2019/04/25/lens/sarah-lewis-racial-bias-photography.html>>. Acesso em: 4 out. 2019.

LIEBERMAN, Zach et al. *OpenFrameworks*. v. 0.10.1, 2018. Disponível em: <<https://openframeworks.cc/>>

LISSOVSKY, Mauricio. A vida póstuma de Aby Warburg: por que seu pensamento seduz os pesquisadores contemporâneos da imagem. *Boletim do Museu Paraense Emílio Goeldi. Ciências Humanas*, v. 9, n. 2, p. 305–322, 2014.

LOVINK, Geert. *Dynamics of critical internet culture: (1994–2001)*. Amsterdam: Institute of Network Cultures, 2009.

LUKYANOVA, Olga; MINTZ, André. Deadartist.me: an experiment with networks and traps. *Transfers*, New York, v. 8, n. 2, p. 122–128, 2018.

LYNCH, Michael. Ontography: Investigating the production of things, deflating ontology. *Social Studies of Science*, Thousand Oaks, v. 43, n. 3, p. 444–462, 2013.

MAATEN, Laurens Van der; HINTON, Geoffrey. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, Brookline, v. 9, n. Nov, p. 2579–2605, 2008.

MACHADO, Arlindo. Anamorfozes cronotópicas ou a quarta dimensão da imagem. In: PARENTE, André (Org.). *Imagem-máquina: a era das tecnologias do virtual*. 3. ed. Rio de Janeiro: Ed. 34, 1993. p. 100–116.

MACHADO, Arlindo. *Arte e mídia*. Rio de Janeiro: Jorge Zahar, 2007.

MACIEL, Jane Cleide de Sousa. Atlas Mnemosyne e saber visual: atualidade de Aby Warburg diante das imagens, mídias e redes. *Ícone*, Recife, v. 16, n. 2, p. 191–209, 2018.

MACKENZIE, Adrian. The production of prediction: what does machine learning want? *European Journal of Cultural Studies*, Thousand Oaks, v. 18, n. 4–5, p. 429–445, 2015.

MACKENZIE, Adrian. *Machine learners: archaeology of a data practice*. Cambridge: MIT Press, 2017.

MALINI, Fábio et al. A viralização da revolta em redes sociais: genealogias de #vempruarua. In: MENDONÇA, Ricardo Fabrino; PEREIRA, Marcus Abílio; FILGUEIRAS, Fernando (Orgs.). *Democracia digital: publicidade, instituições e confronto político*. Belo Horizonte: Ed. UFMG, 2016.

MANOVICH, Lev. *The engineering of vision from constructivism to computers*. 211 f. Tese (doutorado) – University of Rochester, [s. l.], 1993. Disponível em: <<http://manovich.net/EV/EV.PDF>>. Acesso em: 12 set. 2013.

MANOVICH, Lev. *Cultural analytics: visualising cultural patterns in the era of “more media”*. 2009. Disponível em: <https://manovich.net/content/04-projects/063.../60_article_2009.pdf>. Acesso em: 30 set. 2017.

MANOVICH, Lev. How to compare one million images? In: BERRY, David M. (Ed.). *Understanding digital humanities*. London: Palgrave Macmillan UK, 2012. p. 249–278.

MANOVICH, Lev et al. *ImagePlot*. v.1.1, 2017. Disponível em: <<https://github.com/culturevis/imageplot>>

MANOVICH, Lev. The science of culture? Social computing, digital humanities and Cultural Analytics. *Journal of Cultural Analytics*, [s.l.], 2016. Disponível em: <<http://culturalanalytics.org/2016/05/the-science-of-culture-social-computing-digital-humanities-and-cultural-analytics/>>. Acesso em: 18 abr. 2018.

MARRES, Noortje; MOATS, David. Mapping controversies with social media: the case for symmetry. *Social Media + Society*, Thousand Oaks, v. 1, n. 2, 2015. Disponível em: <<http://journals.sagepub.com/doi/10.1177/2056305115604176>>. Acesso em: 13 ago. 2018.

MAURI, Michele et al. RAWGraphs: A visualisation platform to create open outputs. In: 12TH BIENNIAL CONFERENCE ON ITALIAN SIGCHI CHAPTER. *Proceedings... 2017*. Cagliari, Italy: ACM Press, 2017. Disponível em: <<http://dl.acm.org/citation.cfm?doid=3125571.3125585>>. Acesso em: 13 ago. 2018.

MCDONALD, Kyle. *How to recognize fake AI-generated images*. 2018. Disponível em: <<https://medium.com/@kcimc/how-to-recognize-fake-ai-generated-images-4d1f6f9a2842>>. Acesso em: 19 out. 2019.

MENDONÇA, Carlos Camargos; LEAL, Bruno. Ver a elas: mulheres trans e as dimensões políticas da cultura visual. In: LEAL, Bruno; CARVALHO, Carlos Alberto; ALZAMORA, Geane (Orgs.). *Textualidades midiáticas*. Belo Horizonte: PPGCOM/UFMG, 2018. p. 103–112.

MENOTTI, Gabriel. *Movie circuits: curatorial approaches to cinema technology*. Amsterdam: Amsterdam University Press, 2019.

MICHAUD, Philippe-Alain. *Aby Warburg e a imagem em movimento*. Trad. Vera Ribeiro. Rio de Janeiro: Contraponto, 2013.

MINTZ, André. *Visão computacional e visualidades contemporâneas: composições do ver e do visível entre a técnica, a ciência e a arte*. 215 f. Dissertação (Mestrado) – Faculdade de

Filosofia e Ciências Humanas, Universidade Federal de Minas Gerais, Belo Horizonte, 2015.

MINTZ, André. Máquinas que veem: visão computacional e agenciamentos do visível. In: MENOTTI, Gabriel; BASTOS, Marcus; MORAN, Patrícia (Orgs.). *Cinema Apesar da Imagem*. São Paulo: Intermeios, 2016. p. 157–175.

MINTZ, André. Internet landscapes: infraestrutura e espacialidade da imagem em rede. In: MENOTTI, Gabriel (Org.). *Curadoria, cinema e outros modos de dar a ver*. Vitória: EDUFES, 2018. a. p. 15–30.

MINTZ, André. *Memespector Python*. [s. v], 2019. Disponível em: <<https://github.com/amintz/memespector-python>>

MINTZ, André. *Image Network Plotter*. [s. v], 2019. Disponível em: <<https://github.com/amintz/image-network-plotter>>

MINTZ, André; SILVA, Tarcízio et al. *Interrogating vision APIs*. Lisboa: Universidade Nova de Lisboa, 2019. Disponível em: <<https://smart.inovamedialab.org/smart-2019/project-reports/interrogating-vision-apis/>>.

MINTZ, André Goes. Visualidade computacional e fissuras do pós-digital: uma aproximação às imagens invisíveis de Trevor Paglen. *Texto digital*, Florianópolis, v. 14, n. 1, p. 75–92, 2018. d.

MITCHELL, W. J. T. *Picture theory: essays on verbal and visual representation*. Chicago: University Of Chicago Press, 1995.

MOATS, David; BORRA, Erik. Quali-quantitative methods beyond networks: Studying information diffusion on Twitter with the Modulation Sequencer. *Big Data & Society*, Thousand Oaks, v. 5, n. 1, 2018. Disponível em: <<http://journals.sagepub.com/doi/10.1177/2053951718772137>>. Acesso em: 21 nov. 2018.

MOL, Annemarie. Ontological politics. A word and some questions. *The Sociological Review*, Thousand Oaks, 1999. Disponível em: <<https://journals.sagepub.com/doi/abs/10.1111/j.1467-954X.1999.tb03483.x>>. Acesso em: 7 maio. 2019.

MONDZAIN, Marie-José. A imagem entre proveniência e destinação. In: ALLOA, Emmanuel (Org.). *Pensar a Imagem*. Trad. Marianna Poyares et al. Belo Horizonte: Autêntica, 2015. p. 39–53.

MORDVINTSEV, Alexander; OLAH, Christopher; TYKA, Mike. *Inceptionism: Going Deeper into Neural Networks* Google Research Blog, 2015. a. Disponível em: <<https://web.archive.org/web/20150708233542/http://googleresearch.blogspot.com/2015/06/inceptionism-going-deeper-into-neural.html>>. Acesso em: 16 out. 2019.

MORDVINTSEV, Alexander; OLAH, Christopher; TYKA, Mike. *DeepDream – a code example for visualizing Neural Networks* Google Research Blog, 2015. b. Disponível em: <<https://web.archive.org/web/20150708233542/http://googleresearch.blogspot.co.uk/2015/07/deepdream-code-example-for-visualizing.html>>. Acesso em: 16 out. 2019.

MUSSO, Pierre. A filosofia da rede. In: PARENTE, André (Org.). *Tramas da rede: novas dimensões filosóficas, estéticas e políticas da comunicação*. Porto Alegre: Sulina, 2004.

NASA. *Revealing mars true colors*. 2004. Disponível em: <https://mars.jpl.nasa.gov/mer/spotlight/spirit/a12_20040128.html>. Acesso em: 8 set. 2019.

NIEBORG, David B.; POELL, Thomas. The platformization of cultural production: Theorizing the contingent cultural commodity. *New Media & Society*, Thousand Oaks, 2018. Disponível em: <<http://journals.sagepub.com/doi/10.1177/1461444818769694>>. Acesso em: 13 ago. 2018.

NIEDERER, Sabine. *Networked images: visual methodologies for the digital age*. Amsterdam: Amsterdam University of Applied Sciences, 2018.

NIEDERER, Sabine; COLOMBO, Gabriele. Visual methodologies for networked images: designing visualizations for collaborative research, cross-platform analysis, and public participation. *Diseña*, Santiago, n. 14, p. 40–67, 2019.

OMENA, Janna Joceli; RABELLO, Elaine; MINTZ, André. *Visualising hashtag engagement: imagery of political polarization on Instagram*. Amsterdam: Universiteit van Amsterdam, 2017. Disponível em: <<https://wiki.digitalmethods.net/Dmi/InstagramLivenessVisualisingengagement>>.

OMENA, Janna Joceli; RABELLO, Elaine Teixeira; MINTZ, André Goes. Digital methods for hashtag engagement research. *Social Media + Society*, Thousand Oaks, no prelo.

O'REILLY, Tim. *What Is Web 2.0*. 2005. Disponível em: <<http://www.oreilly.com/pub/a/web2/archive/what-is-web-20.html>>. Acesso em: 7 abr. 2017.

PAGLEN, Trevor. Operational Images. *e-flux Journal*, New York, v. 59, 2014. Disponível em: <<https://www.e-flux.com/journal/59/61130/operational-images/>>. Acesso em: 2 jun. 2018.

PAGLEN, Trevor. Invisible images (your pictures are looking at you). *The New Inquiry*, 2016. Disponível em: <<https://thenewinquiry.com/invisible-images-your-pictures-are-looking-at-you/>>. Acesso em: 30 maio. 2018.

PARIKKA, Jussi. *A geology of media*. Minneapolis: University of Minnesota Press, 2015.

PASQUINELLI, Matteo. *The thinking eye (draft)*. Karlsruhe, , 2017. Disponível em: <https://www.academia.edu/32191530/The_Thinking_Eye_draft_>. Acesso em: 15 jun. 2019.

PAUL, Christianne. The myth of immateriality: presenting and preserving new media. In: GRAU, Oliver (Ed.). *MediaArtHistories*. Cambridge: MIT Press, 2007. p. 251–274.

PAUL, Christianne. Genealogies of the digital: a post-critique. In: ERTAN, Ekmel (Ed.). *Dijital Sonrasi Tarihçeler = Histories of the post-digital*. Istanbul: Amber platform, 2015. p. 121–130.

- PEARCE, Warren et al. Visual cross-platform analysis: digital methods to research social media images. *Information, Communication & Society*, Abingdon, p. 1–20, 2018.
- PLANTIN, Jean-Christophe et al. Infrastructure studies meet platform studies in the age of Google and Facebook. *New Media & Society*, Thousand Oaks, 2016. Disponível em: <<http://journals.sagepub.com/doi/10.1177/1461444816661553>>. Acesso em: 23 abr. 2017.
- PUSCHMANN, Cornelius; GAFFNEY, Devin. Data collection on Twitter. In: WELLER, Katrin et al. (Eds.). *Twitter and society*. New York: Peter Lang, 2013. p. 55–67.
- QUÉAU, Philippe. O tempo do virtual. In: PARENTE, André (Org.). *Imagem-máquina: a era das tecnologias do virtual*. 3. ed. Rio de Janeiro: Ed. 34, 1993. p. 91–99.
- QUEIROGA, Louise. Jovem se surpreende com repercussão de foto que fez de Lula. *O Globo*, Rio de Janeiro, 2018. Disponível em: <<https://oglobo.globo.com/brasil/jovem-se-surpreende-com-repercussao-de-foto-que-fez-de-lula-22569898>>. Acesso em: 9 abr. 2018.
- RAJI, Inioluwa Deborah; BUOLAMWINI, Joy. Actionable auditing: investigating the impact of publicly naming biased performance results of commercial ai products. In: CONFERENCE ON ARTIFICIAL INTELLIGENCE, ETHICS AND SOCIETY. *Proceedings.. 2019*. Honolulu: AAI, 2019.
- REFSGAARD, Andreas; TSENG, Francis; KOGAN, Gene. *Machine learning for artists (ml4a)*. 2019. Disponível em: <<https://ml4a.github.io/>>. Acesso em: 29 set. 2019.
- RENAUD-ALAIN, Alain. L’image sans gravité: la forme image aux risques de l’information. *La revue d’esthétique*, Paris, v. 25, p. 11–22, 1994.
- RICCI, Donato et al. Designing Digital Methods to monitor and inform Urban Policy. The case of Paris and its Urban Nature initiative. In: 3RD INTERNATIONAL CONFERENCE ON PUBLIC POLICY (ICPP3). *Proceedings... 2017*. Cingapura: International Public Policy Association, 2017.
- RIEDER, B. What is in PageRank? A historical and conceptual investigation of a recursive status index. *Computational Culture*, 00064, v. 2, 2012. Disponível em: <<https://dare.uva.nl/search?identifier=0fae1f7f-0bf6-4b94-9b9e-8c4e23ee4c28>>. Acesso em: 10 jul. 2019.
- RIEDER, Bernhard. Studying Facebook via data extraction: the Netvizz application. In: 5TH ANNUAL ACM WEB SCIENCE CONFERENCE. *Proceedings... 2013*. Paris: ACM, 2013. Disponível em: <<http://dl.acm.org/citation.cfm?id=2464475>>. Acesso em: 8 ago. 2017.
- RIEDER, Bernhard et al. Data critique and analytical opportunities for very large Facebook Pages: Lessons learned from exploring “We are all Khaled Said”. *Big Data & Society*, v. 2, n. 2, 2015. Disponível em: <<http://journals.sagepub.com/doi/10.1177/2053951715614980>>. Acesso em: 20 nov. 2018.
- RIEDER, Bernhard. Examinando uma técnica algorítmica: o classificador de bayes como uma leitura interessada da realidade. *Parágrafo: Revista Científica de Comunicação Social da FIAM-FAAM*, v. 6, n. 1, p. 123–142, 2018.

RIEDER, Bernhard; BORRA, Erik. Programmed method: developing a toolset for capturing and analyzing tweets. *Aslib Journal of Information Management*, v. 66, n. 3, p. 262–278, 2014.

RIEDER, Bernhard; DEN TEX, Emille; MINTZ, André. *Memespector*. [s.l.: s.n.]. Disponível em: <<https://github.com/bernorieder/memespector>>

RIEDER, Bernhard; RÖHLE, Theo. Digital methods: five challenges. In: BERRY, David M. (Ed.). *Understanding Digital Humanities*. Londres: Palgrave Macmillan UK, 2012. p. 67–84.

RIEDER, Bernhard; RÖHLE, Theo. Digital methods: from challenges to bildung. In: SCHÄFER, Mirko Tobias; VAN ES, Karin (Eds.). *The datafied society: studying culture through data*. Amsterdam: Amsterdam University Press, 2017. p. 109–124.

RIEDER, Bernhard; SIRE, Guillaume. Conflicts of interest and incentives to bias: A microeconomic critique of Google's tangled position on the Web. *New Media & Society*, Thousand Oaks, v. 16, n. 2, p. 195–211, 2013.

ROBERTS, Lawrence G. *Machine perception of three-dimensional solids*, 1963. Disponível em: <<http://www.packet.cc/fles/mach-per-3D-solids.html>>. Acesso em: 8 jul. 2014.

ROGERS, Richard. *Digital methods*. Cambridge: MIT Press, 2013.

ROGERS, Richard. O fim do virtual: os métodos digitais. *Lumina*, Juiz de Fora, v. 10, n. 3, 2016. Disponível em: <<https://periodicos.ufjf.br/index.php/lumina/article/view/21353>>. Acesso em: 20 jul. 2019.

ROGERS, Richard. Foundations of digital methods: query design. In: SCHÄFER, Mirko Tobias; VAN ES, Karin (Eds.). *The datafied society: studying culture through data*. Amsterdam: Amsterdam University Press, 2017. p. 75–94.

ROGERS, Richard. Digital methods for cross-platform analysis. In: BURGESS, Jean; MARWICK, Alice; POELL, Thomas (Eds.). *The SAGE handbook of social media*. Thousand Oaks: SAGE Publications, 2018. a. p. 233–253.

ROGERS, Richard. Otherwise engaged: social media from vanity metrics to critical analytics. *International Journal of Communication*, Los Angeles, v. 12, p. 23, 2018. b.

ROSE, Gillian. *Visual methodologies: an introduction to researching with visual materials*. 4. Kindle ed. Thousand Oaks: SAGE Publications, 2016.

ROSENBLATT, F. The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, Washington, v. 65, n. 6, p. 386–408, 1958.

ROTH, Lorna. Looking at Shirley, the ultimate norm: colour balance, image technologies, and cognitive equity. *Canadian Journal of Communication*, Vancouver, v. 34, n. 1, 2009. Disponível em: <<https://www.cjc-online.ca/index.php/journal/article/view/2196>>. Acesso em: 7 set. 2019.

RUBINSTEIN, Daniel; SLUIS, Katrina. A life more photographic. *Photographies*, Abingdon, v. 1, n. 1, p. 9–28, 2008.

RYKOV, Yuri et al. Semantic and geospatial ,mapping of Instagram Images in Saint-Petersburg. In: AINL FRUCT 2016 CONFERENCE. *Proceedings... 2016*. Saint Petersburg: IEEE, 2016. Disponível em: <<http://ieeexplore.ieee.org/servlet/opac?punumber=7889413>>. Acesso em: 14 jan. 2018.

SALGADO, Tiago Barcelos Pereira. A virada não humana na comunicação: contribuições da teoria ator-rede e da ontologia orientada aos objetos. *Revista ECO-Pós*, Rio de Janeiro, v. 21, n. 2, p. 171–191, 2018.

SALGADO, Tiago Barcelos Pereira. Surpreendidos pela ação – mediação pelas Sociologias Pragmáticas Francesas. *Ação Midiática – Estudos em Comunicação, Sociedade e Cultura*, Curitiba, v. 18, p. 94–115, 2019.

SANTAELLA, Lucia. Do signo. In: *A teoria geral dos signos*. 3. ed. São Paulo: Iluminuras, 2001. p. 159–186.

SANTAELLA, Lucia; NÖTH, Winfried. Os três paradigmas da imagem. In: *Imagem: cognição, semiótica, mídia*. 3. ed. São Paulo: Iluminuras, 2001. p. 159–186.

SCHAEFFER, Jean-Marie. *A imagem precária: sobre o dispositivo fotográfico*. Campinas: Papirus, 1996.

SCHWARCZ, Lilia Moritz; STARLING, Heloisa Murgel. *Brasil: uma biografia*. São Paulo: Companhia das Letras, 2015.

SEKULA, Allan. The body and the archive. *October*, Cambridge, v. 39, p. 3–64, 1986.

SHAPIN, Steven; SCHAFFER, Simon. *Leviathan and the air-pump: Hobbes, Boyle, and the experimental life*. Reprint ed. Princeton: Princeton University Press, 2011.

SHIFMAN, Limor. *Memes in digital culture*. Cambridge: MIT Press, 2013.

SHIFMAN, Limor. The cultural logic of photo-based meme genres. *Journal of Visual Culture*, Thousand Oaks, v. 13, n. 3, p. 340–358, 2014.

SILVA, Tarcizio. Visão computacional e vieses racializados: branquitude como padrão no aprendizado de máquina. In: II COPENE – CONGRESSO DE PESQUISADORES/AS NEGROS/AS DO NORDESTE. *Anais... 2019*. João Pessoa: COPENE, 2019. Disponível em: <https://www.researchgate.net/publication/334263119_Visao_Computacional_e_Vieses_Racializados_branquitude_como_padrao_no_aprendizado_de_maquina>

SILVA, Tarcízio; BARCIELA, Pedro; MEIRELLES, Pedro. Mapeando imagens de desinformação e fake news político-eleitorais com inteligência artificial. In: 3º CONEC: CONGRESSO NACIONAL DE ESTUDOS COMUNICACIONAIS DA PUC MINAS POÇOS DE CALDAS – CONVERGÊNCIA E MONITORAMENTO. *Anais... 2018*. Poços de Caldas: PUC Minas, 2018. Disponível em: <<https://conec.pucpcaldas.br/wp-content/uploads/2019/06/anais2018.pdf>>. Acesso em: 2 jul. 2019.

SILVA, Tarcizio; ZANOTTI, Mariana. Evocando cartões postais no Instagram: estudo automatizado de imagens. In: 41º CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO. *Anais...* 2018. Joinville: INTERCOM, 2018.

SIMONDON, Gilbert. *El modo de existencia de los objetos técnicos*. 2. ed. Buenos Aires: Prometeo Libros, 2007.

SIMONYAN, Karen; ZISSERMAN, Andrew. Very deep convolutional networks for large-scale image recognition. *arxiv:1409.1556 [cs]*, 25384 arxiv: 1409.1556, 2014. Disponível em: <<http://arxiv.org/abs/1409.1556>>. Acesso em: 10 ago. 2019.

SMEULDERS, A. W. M. et al. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Piscataway, v. 22, n. 12, p. 1349–1380, 2000.

SONTAG, Susan. *Sobre fotografia*. Trad. Rubens Figueiredo. São Paulo: Companhia das Letras, 2004.

STEYERL, Hito. In Defense of the Poor Image. *e-flux Journal*, New York, v. 10, 2009. Disponível em: <<http://www.e-flux.com/journal/10/61362/in-defense-of-the-poor-image/>>. Acesso em: 24 mar. 2018.

STEYERL, Hito. Too much world: is the internet dead? *e-flux Journal*, New York, n. 49, 2013. Disponível em: <<http://www.e-flux.com/journal/49/60004/too-much-world-is-the-internet-dead/>>. Acesso em: 16 fev. 2017.

STOCKHAUSEN, T. Von. *Die Kulturwissenschaftliche Bibliothek Warburg – Architektur, Einrichtung und Organisation*. Hamburgo: Dölling un Gallitz Verlag, 1992.

SUCHMAN, Lucy. *Human-machine reconfigurations: plans and situated actions*. Kindle edition ed. Cambridge ; New York: Cambridge University Press, 2007.

SUWAJANAKORN, Supasorn; SEITZ, Steven M.; KEMELMACHER-SHLIZERMAN, Ira. Synthesizing Obama: learning lip sync from audio. *ACM Transactions on Graphics*, New York, v. 36, n. 4, p. 1–13, 2017.

TELLES, Marcio. Das materialidades às matérias-primas da comunicação: notas para uma perspectiva teórica geológica. In: XXV ENCONTRO ANUAL DA COMPÓS. *Anais...* 2016. Goiânia: Compós, 2016.

THE GUARDIAN. Facebook, Instagram and WhatsApp hit by media messaging outage. *The Guardian*, London, 2019. Disponível em: <<https://www.theguardian.com/technology/2019/jul/03/instagram-whatsapp-facebook-media-files-outage>>. Acesso em: 4 jul. 2019.

TIFENTALE, Alise. Art of the Masses: From Kodak Brownie to Instagram. *Networking Knowledge: Journal of the MeCCSA Postgraduate Network*, [s.l.], v. 8, n. 6, 2015. Disponível em: <<https://ojs.meccsa.org.uk/index.php/netknow/article/view/399>>. Acesso em: 31 jul. 2019.

TIFENTALE, Alise; MANOVICH, Lev. Selfiecity: Exploring photography and self-fashioning in social media. In: BERRY, David M.; DIETER, Michael (Eds.). *Postdigital Aesthetics*. London: Palgrave Macmillan UK, 2015. p. 109–122.

TURING, Alan. Computing machinery and intelligence. In: WARDRIP-FRUIN, Noah; MONTFORT, Nick (Eds.). *The new media reader*. Cambridge: MIT Press, 2003. p. 50–64.

TWITTER. *POST statuses/filter*. 2019a. Disponível em: <<https://developer.twitter.com/en/docs/tweets/filter-realtime/api-reference/post-statuses-filter.html>>. Acesso em: 19 jul. 2019.

TWITTER. *Como compartilhar e assistir a vídeos no Twitter*. 2019b. Disponível em: <<https://help.twitter.com/pt/using-twitter/twitter-videos>>. Acesso em: 5 out. 2019.

TWITTER. *Tweet objects*. 2019c. Disponível em: <<https://developer.twitter.com/en/docs/tweets/data-dictionary/overview/entities-object.html>>. Acesso em: 6 out. 2019.

TWITTER. *Getting started with cards*. 2019d. Disponível em: <<https://developer.twitter.com/en/docs/tweets/optimize-with-cards/guides/getting-started.html>>. Acesso em: 24 jul. 2019.

TWITTER. *Sobre contas verificadas*. 2019e. Disponível em: <<https://help.twitter.com/pt/managing-your-account/about-twitter-verified-accounts>>. Acesso em: 6 out. 2019.

URICCHIO, William. The algorithmic turn: photosynth, augmented reality and the changing implications of the image. *Visual Studies*, Abingdon, v. 26, n. 1, p. 25–35, 2011.

VAN DIJCK, José. *The culture of connectivity: a critical history of social media*. Oxford: Oxford University Press, 2013.

VAN DIJCK, Jose. Datafication, dataism and dataveillance: Big Data between scientific paradigm and ideology. *Surveillance & Society*, Kingston, v. 12, n. 2, p. 197–208, 2014.

VAN DIJCK, José. In data we trust? The implications of datafication for social monitoring. *MATRIZES*, São Paulo, v. 11, n. 1, p. 39, 2017.

VAN ES, Karin; WIERINGA, Maranke; SCHÄFER, Mirko Tobias. Tool Criticism: From Digital Methods to Digital Methodology. In: 2ND INTERNATIONAL CONFERENCE ON WEB STUDIES. *Proceedings...* 2018. New York: ACM, 2018. Disponível em: <<http://doi.acm.org/10.1145/3240431.3240436>>. Acesso em: 23 ago. 2019.

VELDEN, Daniel Van der; KRUK, Vinca; METAHAVEN (Eds.). *Black transparency: the right to know in the age of mass surveillance*. Berlin: Sternberg Press, 2015.

VENTURINI, Tommaso. Diving in magma: how to explore controversies with actor-network theory. *Public Understanding of Science*, Thousand Oaks, v. 19, n. 3, p. 258–273, 2010.

VENTURINI, Tommaso; JACOMY, Mathieu; JENSEN, Pablo. *What do we see when we look at networks. An introduction to visual network analysis and force-directed layouts*. SSRN

Scholarly Paper. 2019. Disponível em: <<https://papers.ssrn.com/abstract=3378438>>. Acesso em: 20 jul. 2019.

VENTURINI, Tommaso; LATOUR, Bruno. The social fabric: Digital traces and qualitative methods. In: FUTURE EN SEINE 2009. *Proceedings... 2010*. [s.l.: s.n.] Disponível em: <http://www.academia.edu/download/38150764/Venturini__Latour_-_2010_-_The_Social_Fabric_Digital_Traces_and_Quantitative_Methods.pdf>. Acesso em: 31 jul. 2017.

VENTURINI, Tommaso; MUNK, Anders; JACOMY, Mathieu. Ator-rede versus Análise de Redes versus Redes Digitais: falamos das mesmas redes? *Galáxia*, n. 38, p. 5–27, 2018.

VENTURINI, Tommaso; ROGERS, Richard. “API-based research” or how can digital sociology and journalism studies learn from the Facebook and Cambridge Analytica data breach. *Digital Journalism*, v. 7, n. 4, p. 532–540, 2019.

VIDAL JUNIOR, Ícaro Ferraz. Invisibilidade, superficialidade e plasticidade: três hipóteses sobre as câmeras inteligentes. *Galáxia*, n. 31, p. 156–167, 2016.

VIMIEIRO, Ana Carolina; BARGAS, Janine de Kássia Rocha. A virada computacional nas pesquisas em comunicação. In: XXVII ENCONTRO ANUAL DA COMPÓS. *Anais... 2018*. Belo Horizonte: Compós, 2018. Disponível em: <http://www.compos.org.br/data/arquivos_2018/trabalhos_arquivo_VH6KF9JHLB3B3AKCJSJD_27_6757_26_02_2018_12_26_14.pdf>. Acesso em: 26 ago. 2019.

VINCENT, James. *Facebook’s image outage reveals how the company’s AI tags your photos*. 2019. Disponível em: <<https://www.theverge.com/2019/7/3/20681231/facebook-outage-image-tags-captions-ai-machine-learning-revealed>>. Acesso em: 4 jul. 2019.

VIRILIO, Paul. A imagem virtual mental e instrumental. In: PARENTE, André (Org.). *Imagem-máquina: a era das tecnologias do virtual*. Rio de Janeiro: Ed. 34, 1993. p. 127–132.

VIRILIO, Paul. *A máquina de visão*. Rio de Janeiro: José Olympio, 1994.

VIS, Farida; GORIUNOVA, Olga (Orgs.). *The iconic image on social media: a rapid research response to the death of Aylan Kurdi*. Sheffield; Manchester; Londres: Visual Social Media Lab, 2015. Disponível em: <<http://visualsocialmedialab.org/projects/the-iconic-image-on-social-media>>. Acesso em: 24 maio. 2018.

WARBURG, Aby. *A renovação da antiguidade pagã: contribuições científico-culturais para a história do renascimento europeu*. Trad. Markus Hediger. Rio de Janeiro: Contraponto, 2013.

WARBURG, Aby. *Histórias de fantasma para gente grande: escritos, esboços e conferências*. Trad. Bárbara Lenin Bicudo. São Paulo: Companhia das Letras, 2015.

WATTENBERG, Martin; VIÉGAS, Fernanda; JOHNSON, Ian. How to Use t-SNE Effectively. *Distill*, 00149, v. 1, n. 10, 2016. Disponível em: <<http://distill.pub/2016/misread-tsne>>. Acesso em: 14 out. 2019.

WIENER, Norbert. *Cybernetics or control and communication in the animal and the machine*. Kindle ed. Cambridge: MIT Press, 2007.

WIGGINS, Bradley E.; BOWERS, G. Bret. Memes as genre: A structural analysis of the memescape. *New Media & Society*, v. 17, n. 11, p. 1886–1906, 2015.

WINNER, Langdon. Artefatos têm política? *Analytica. Revista de Filosofia*, Rio de Janeiro, v. 21, n. 2, p. 195–218, 2017.

WU, Shaomei. *How blind people interact with visual content on social networking sites* Facebook Research, 2016. Disponível em: <<https://research.fb.com/how-blind-people-interact-with-visual-content-on-social-networking-sites/>>. Acesso em: 8 jul. 2019.

YALE DIGITAL HUMANITIES LAB. *Neural neighbors: capturing image similarity*. 2017. Disponível em: <<https://dhlab.yale.edu/projects/neural-neighbors/>>. Acesso em: 14 out. 2019.

ZKM. *The Mnemosyne Atlas*. 2016. Disponível em: <<https://zkm.de/en/event/2016/09/aby-warburg-mnemosyne-bilderatlas/the-mnemosyne-bilderatlas>>. Acesso em: 29 set. 2019.

APÊNDICES

Apêndice A Termos de busca utilizados na coleta

Tabela 8: Termos de busca e métricas de desempenho de coleta no caso Lula.
Linhas em negrito indicam os 68 termos não redundantes na coleta.

termo	inicio da coleta	núm ocorrências	núm ocorrências isoladas
1 apoiotrf4	2018-01-19 18:51:00	0	0
2 cadeaprova	2018-01-19 18:51:00	86	31
3 cadeiaprolula	2018-01-20 20:06:00	17	0
4 cadeiasemlulaefraude	2018-01-20 20:06:00	724	0
5 carlos eduardo lenz	2018-01-18 22:29:00	29	0
6 carlos eduardo thompson	2018-01-18 22:29:00	46	2
7 carnalula	2018-01-23 10:19:00	958	4
8 clube bahamas	2018-04-07 20:25:00	29	27
9 comlulaempoa	2018-01-19 18:51:00	327	3
10 comlulaemportoalegre	2018-01-20 20:06:00	0	0
11 condenatrf4	2018-01-23 10:19:00	49	0
12 cristiano zanin	2018-01-19 00:53:00	5.022	787
13 dia24comlula	2018-01-21 21:23:00	0	0
14 dilma	2018-04-06 13:54:00	260.337	125.328
15 dona marisa	2018-04-07 00:10:00	13.352	5.620
16 eleiçãosemlulaéfraude	2018-01-19 18:51:00	118	0
17 estamosdeolhotrf4	2018-01-23 10:19:00	0	0
18 eusoulula	2018-04-07 16:09:00	146.005	0
19 francisco proner	2018-04-10 12:25:00	1.254	313
20 gebran neto	2018-01-18 22:29:00	105	24
21 iamlula	2018-04-07 16:09:00	140	0
22 instituto lula	2018-01-18 22:29:00	42.187	83
23 jamaisaprisionarãonossossonhos	2018-04-07 16:11:00	24.168	5.214
24 juiciodelula	2018-01-21 21:23:00	1	0
25 juiz moro	2018-01-24 10:26:00	103.621	38.254
26 julgamento lula	2018-01-18 22:29:00	65.214	16.515
27 julgamento lula	2018-01-23 10:19:00	1.028	6
28 justicaparalula	2018-01-23 10:19:00	0	0
29 justiceforlula	2018-01-23 10:19:00	14	0
30 justiciaporlula	2018-01-23 10:19:00	32	6
31 leandro paulsen	2018-01-18 22:29:00	873	12
32 leo pinheiro oas	2018-01-18 22:32:00	65	35
33 lula	2018-04-05 14:42:00	6.373.472	2.939.256
34 lula cadeia	2018-01-20 20:06:00	239.278	0
35 lula corrupção	2018-04-05 14:39:00	38.496	8
36 lula corrupto	2018-04-05 14:39:00	92.968	6

termo	inicio da coleta	núm ocorrências	núm ocorrências isoladas
37 lula felix fischer	2018-03-06 20:04:00	1.650	0
38 lula habeas corpus	2018-03-01 00:21:00	118.353	21.458
39 lula hc	2018-04-04 21:45:00	99.565	14.013
40 lula judgement	2018-01-21 21:29:00	117	51
41 lula juicio	2018-01-21 21:29:00	22.511	2.491
42 lula ladrão	2018-04-05 14:39:00	46.264	0
43 lula oas	2018-01-24 10:26:00	76.051	1.122
44 lula president	2018-01-24 10:26:00	482.991	5.135
45 lula presidente	2018-01-24 10:26:00	417.758	0
46 lula preso	2018-01-24 10:26:00	966.790	70.320
47 lula prisão	2018-01-24 10:26:00	605.673	23.075
48 lula silva	2018-01-24 10:43:00	339.374	28.848
49 lula stf	2018-03-01 00:21:00	197.949	24.273
50 lula stj	2018-03-06 20:04:00	22.470	609
51 lula trial	2018-01-21 21:29:00	1.374	77
52 lula triplex	2018-01-18 22:29:00	75.136	0
53 lula versuch	2018-01-21 21:29:00	8	2
54 lula2018	2018-01-20 20:06:00	5.491	305
55 lulaadore	2018-04-05 14:26:00	93	10
56 lulacondenado	2018-01-23 10:19:00	887	38
57 lulaforagido	2018-04-06 20:32:00	10.738	0
58 lulainocente	2018-01-21 21:23:00	9.196	384
59 lulalivre	2018-04-04 13:16:00	431.799	61.826
60 lulalivrejá	2018-04-16 12:29:00	156	0
61 lulamandela	2018-01-21 21:23:00	11	1
62 lulanacadeia	2018-01-19 18:41:00	126.985	189
63 lulanacadeiadia24	2018-01-19 18:51:00	5	0
64 lulanacadeiaem2018	2018-01-21 21:22:00	4.019	0
65 lulanacadeiasim	2018-01-21 21:29:00	76	0
66 lulanapapuda	2018-01-23 10:19:00	277	35
67 lulanaprisao	2018-01-23 10:19:00	48.497	1
68 lulanoprimeiroturno	2018-01-23 10:19:00	0	0
69 lulanotr4	2018-01-23 10:19:00	0	0
70 lulanuncamais	2018-01-19 18:51:00	250	11
71 lulapajaula	2018-01-19 18:41:00	1.935	50
72 lulapelobrasil	2018-04-05 14:42:00	576.295	0
73 lulaprajaula	2018-01-19 18:41:00	71	2
74 lulapreso	2018-01-23 10:19:00	232.200	205
75 lulapresoamanhã	2018-03-21 17:09:00	58.979	0
76 lulapresohoje	2018-04-05 14:07:00	53.829	0

termo	inicio da coleta	núm ocorrências	núm ocorrências isoladas
77 lulapresopolítico	2018-04-09 12:37:00	2.495	0
78 lulavalealuta	2018-01-23 10:19:00	160.856	15.281
79 marchadosbandidosemportoalegre	2018-01-23 10:19:00	0	0
80 marisa letícia	2018-04-07 00:10:00	15.583	4.796
81 molusco cadeia	2018-01-24 19:50:00	2.300	523
82 molusco julgamento	2018-01-24 19:50:00	126	85
83 molusco preso	2018-01-24 19:50:00	2.650	1.047
84 molusconacadeia	2018-01-24 19:50:00	720	1
85 moluscopresoamanhã	2018-03-21 17:09:00	1	0
86 ocupacuritiba	2018-04-07 16:12:00	31.345	5.699
87 ocupapoa	2018-01-20 20:06:00	20	12
88 ocupaportoalegre	2018-01-20 20:06:00	505	2
89 ocupasaobernardo	2018-04-06 12:01:00	92.792	33.328
90 ocupatrf4	2018-01-19 18:51:00	42	0
91 ocupecuritiba	2018-04-07 16:12:00	4	2
92 pixulecosday	2018-01-20 20:06:00	2	2
93 porto alegre lula	2018-01-18 22:29:00	8.351	80
94 primeiramentelulanacadeia	2018-01-21 21:23:00	0	0
95 prisao lula	2018-01-20 20:06:00	60.156	568
96 prisãodelula	2018-04-06 21:13:00	259	0
97 recurso lula	2018-03-21 17:09:00	40.741	5.701
98 sergio moro	2018-01-24 10:26:00	80.870	27.819
99 somoslula	2018-04-07 16:09:00	1.894	0
100 somoslulas	2018-04-07 16:09:00	19	0
101 somosmilhoesdelulas	2018-01-23 10:19:00	574	0
102 thompson flores lenz	2018-01-18 22:29:00	29	0
103 trf 4	2018-01-18 22:29:00	49.649	1.845
104 trf-4	2018-01-18 22:29:00	22.331	185
105 trf4	2018-01-19 18:44:00	25.280	195
106 trf4cadeaprova	2018-01-19 18:51:00	0	0
107 triplex guaruja	2018-01-18 22:29:00	929	394
108 triplexdoguaruja	2018-01-19 18:51:00	1	0
109 victor laus	2018-01-18 22:29:00	159	6

Apêndice B *Atlas para Imagens-Redes*

O *Atlas* encontra-se anexado à cópia física desta tese em um DVD. Para iniciar a navegação, basta acessar abrir o arquivo *index.html*, localizado na raiz do DVD. O arquivo pode ser aberto em qualquer navegador, mas os testes foram realizados no Mozilla Firefox. O carregamento do arquivo é lento, pois ele é composto de mais de 18 mil imagens. Por isto, *recomendo que, antes de abrir a página, encerre programas e processos que podem ocupar a memória RAM do computador.* Uma vez carregada a página, é possível navegar por meio do rolamento vertical e horizontal.

Utilizando o botão de rolamento vertical de um *mouse* ou o rolamento vertical com os dedos em um *trackpad*, é também possível realizar rolamento horizontal, alternando a este modo pelo pressionamento da tecla *Shift* de modo concomitante ao rolamento vertical. Em muitos navegadores, também é possível aumentar ou diminuir a exibição da página utilizando as teclas **Ctrl +** (para aproximar) e **Ctrl -** (para afastar).

Os botões situados no canto direito superior da tela permitem acionar camadas suplementares à visualização. O botão **Mapa de calor** aciona uma transformação aplicada às imagens em que elas são escurecidas de forma inversamente proporcional ao público potencial de sua circulação, estimado pela soma dos números de seguidores das contas que publicaram estas imagens. O botão **Linhas de grade** aciona a exibição de uma grade de coordenadas sobre o mapa, que tem o objetivo de facilitar a localização de imagens ou conjuntos de imagens entre diferentes representações desta visualização. *Quando a opção Linhas de grade estiver acionada, não será possível clicar nas imagens na visualização.*

Cada imagem disposta na visualização pode ser clicada. Ao fazê-lo, abre-se uma página de informações específicas sobre o grupo de imagens que aquela matriz representa. São indicadas métricas pertinentes àquele conjunto de imagens; as coordenadas daquela imagem em cada mapa (VGG19 e GVAPI); as etiquetas atribuídas à imagem pelo GVAPI; uma amostra das instâncias que compõem aquele conjunto; uma amostra de tuítes que compartilharam alguma das instâncias do conjunto; e um gráfico de dispersão que exhibe a distribuição dos tuítes que compartilharam instâncias do conjunto segundo o tempo e o número de seguidores da conta que publicou.

ANEXO

Anexo A Estrutura do modelo VGG19

Tabela 9: Sumário de camadas da rede neural do modelo VGG19.

Fonte: CHOLLET et al., 2018; SIMONYAN; ZISSERMAN, 2014.

Layer (type)	Output Shape	Param #
input_1 (InputLayer)	(None, 224, 224, 3)	0
block1_conv1 (Conv2D)	(None, 224, 224, 64)	1792
block1_conv2 (Conv2D)	(None, 224, 224, 64)	36928
block1_pool (MaxPooling2D)	(None, 112, 112, 64)	0
block2_conv1 (Conv2D)	(None, 112, 112, 128)	73856
block2_conv2 (Conv2D)	(None, 112, 112, 128)	147584
block2_pool (MaxPooling2D)	(None, 56, 56, 128)	0
block3_conv1 (Conv2D)	(None, 56, 56, 256)	295168
block3_conv2 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv3 (Conv2D)	(None, 56, 56, 256)	590080
block3_conv4 (Conv2D)	(None, 56, 56, 256)	590080
block3_pool (MaxPooling2D)	(None, 28, 28, 256)	0
block4_conv1 (Conv2D)	(None, 28, 28, 512)	1180160
block4_conv2 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv3 (Conv2D)	(None, 28, 28, 512)	2359808
block4_conv4 (Conv2D)	(None, 28, 28, 512)	2359808
block4_pool (MaxPooling2D)	(None, 14, 14, 512)	0
block5_conv1 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv2 (Conv2D)	(None, 14, 14, 512)	2359808
block5_conv3 (Conv2D)	(None, 14, 14, 512)	2359808

```

-----
block5_conv4 (Conv2D)      (None, 14, 14, 512)      2359808
-----
block5_pool (MaxPooling2D) (None, 7, 7, 512)        0
-----
flatten (Flatten)         (None, 25088)             0
-----
fc1 (Dense)                (None, 4096)              102764544
-----
fc2 (Dense)                (None, 4096)              16781312
-----
predictions (Dense)       (None, 1000)              4097000
=====
Total params: 143,667,240
Trainable params: 143,667,240
Non-trainable params: 0

```