L'occhio della macchina

Silvia Bottani
28 Novembre 2019

Da domenica 9 giugno, Hong Kong è diventata teatro di scontri tra le forze governative e gli attivisti pro-democrazia che protestano contro un disegno di legge che faciliterebbe l'estradizione in Cina di cittadini accusati di reati gravi. La popolazione, preoccupata per le conseguenze che la legge avrebbe sul delicato equilibrio politico del paese, ha cominciato a manifestare pacificamente, in un crescendo di tensione e di scontri sempre più duri. Durante le manifestazioni, gli attivisti hanno invitato i partecipanti a utilizzare mascherine per coprirsi il volto e occhiali per rendersi irriconoscibili dalle forze di polizia e hanno chiesto ai reporter di non scattare immagini che potessero aiutare a identificare chi protesta. Le immagini dei civili che fronteggiano i poliziotti a colpi di puntatori laser, cercando di “accecare” i sistemi di riconoscimento facciale e di impedire agli agenti di mirare, hanno fatto il giro del mondo, segnando un cambio di passo nella pratica dei conflitti sociali.

Hong Kong, Photograph: Kevin On Man Lee/Penta Press/REX/Shutterstock.

L'“ansia da riconoscimento” dei manifestanti di Hong Kong trova una giustificazione nell'atteggiamento autoritario del governo cinese, fautore di una politica securitaria restrittiva, da cui discendono soprusi denunciati sistematicamente dalle associazioni di tutela dei diritti umani. La risposta più recente del governo cinese agli scontri in corso è l'atto legislativo che vieta l'utilizzo di maschere e durante le manifestazioni. Un provvedimento che gli attivisti hanno sfidato apertamente, continuando a scendere in strada a viso coperto.

Non è un caso che questo conflitto a bassa intensità si verifichi in un teatro come quello di Hong Kong, metropoli orientale che rientra in quegli scenari geo-politici ed economici che rappresentano dei veri e proprio laboratori del “capitalismo della sorveglianza”, definizione che circola da tempo e che trova una disamina puntuale in *The Age of Surveillance Capitalism* (2019), libro di Shoshanna Zuboff, appena uscito anche in Italia.

Per capire meglio come si articola questo regime scopico, si può visitare la mostra *Training Humans* di Kate Crawford e Trevor Paglen, aperta presso l'Osservatorio Prada di Milano. Si tratta della prima esposizione interamente dedicata ai set fotografici utilizzati per istruire le macchine, principalmente con scopi di riconoscimento facciale e per sviluppare modelli predittivi.

Sospeso all'ingresso, un ciclopico occhio umano (SDUMLA, Yilong Yin, Lili Liu, Xiwei Sun, 2011) accoglie lo spettatore che si accinge a entrare in un labirinto di immagini solo all'apparenza neutrali. Sul retro dell'opera è riprodotta un'impronta digitale presa da un set di dati multimodali sviluppati dall'Università di Shandong: occhi e impronta, due immagini icastiche che rappresentano con forza simbolica il tema della sorveglianza.

Nessuna fotografia d'arte è presente in mostra: l'esposizione raccoglie alcune collezioni di immagini basate su soggetti umani e utilizzate per allenare i sistemi di AI. Al piano d'ingresso dell'Osservatorio si trovano alcune tra le prime raccolte prodotte in ambito militare e scientifico, mentre al piano superiore sono esposti alcuni set prodotti durante l'era di Internet. Immagini raccolte e catalogate principalmente senza il consenso degli autori delle foto o di chi vi è ritratto, “rubate” per essere date in pasto agli algoritmi.

Addentrandosi nel percorso della mostra, si affronta un viaggio cronologico nella storia della computer vision attraverso i progetti che ne hanno permesso la nascita e lo sviluppo. Un tema tutt'altro che marginale, se si comprende il ruolo che questi dispositivi svolgono nel panorama politico ed economico odierno.

Tra la semplice raccolta di dati biometrici finalizzata allo sviluppo di un metodo di riconoscimento facciale della prima installazione, basata sugli studi finanziati dalla CIA (*A facial recognition project report*,Woodrow Wilson Bledsoe, 1963), e i modelli prodotti dai ricercatori di Facebook e Amazon pensati per stimare sesso, età e stato emotivo dei soggetti rilevati, che chiudono la mostra, è racchiusa una rivoluzione epocale la cui portata rimane forse ancora in parte oscura. Se consideriamo i due grandi momenti evolutivi nella storia umana quali l'invenzione dell'agricoltura durante il Neolitico e la rivoluzione industriale in epoca moderna, il terzo passaggio può essere individuato a pieno titolo nella rivoluzione informatica. È un passaggio che segna un cesura radicale con il mondo analogico, dalle molteplici implicazioni, e che rende possibile ipotizzare un passaggio di specie, da Homo Sapiens a qualcosa che ancora non ha nome ma che lo storico Yuval Noah Harari ha battezzato provocatoriamente Homo Deus.

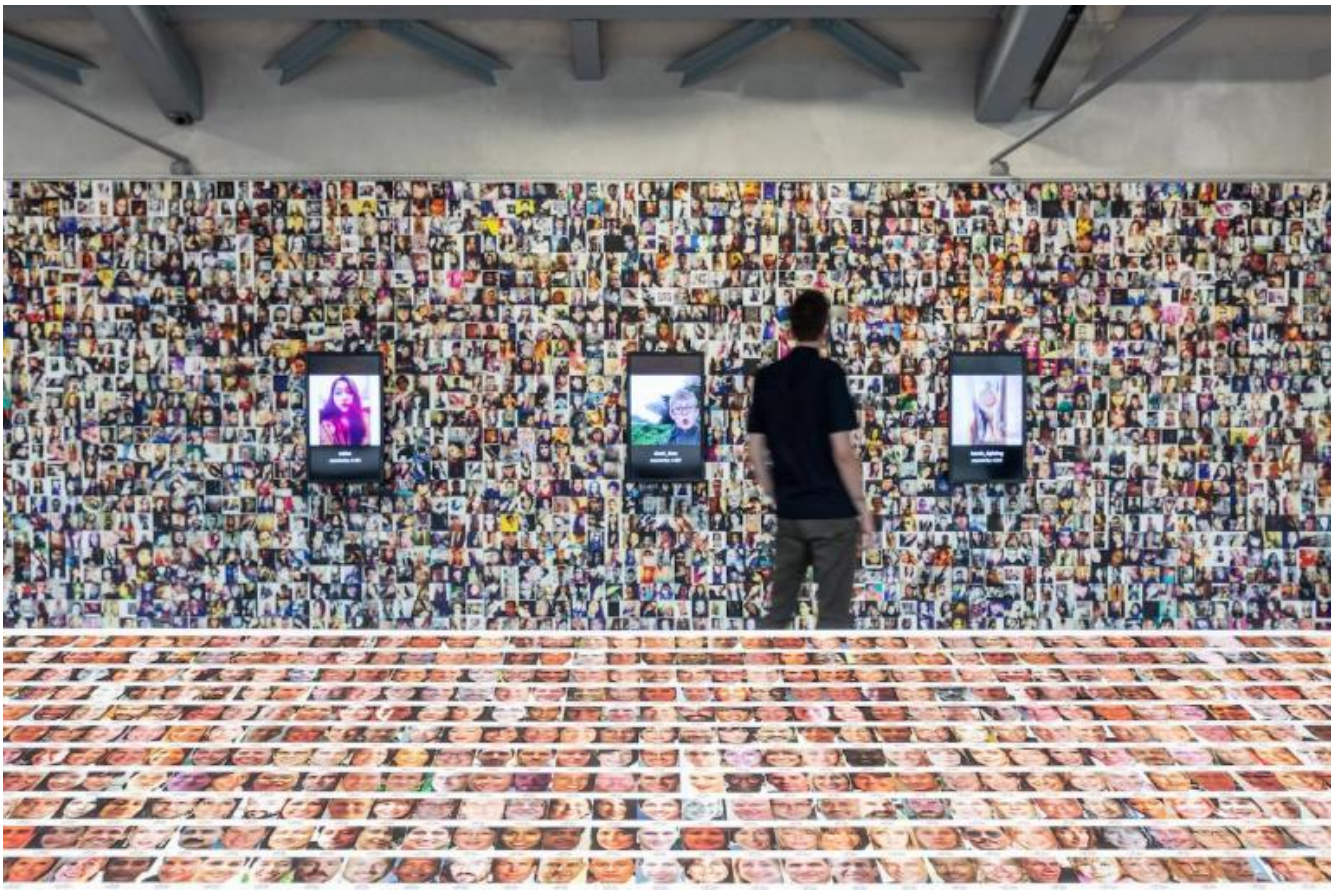
La mostra di Crawford e Paglen prende in esame un aspetto specifico di questa rivoluzione e si concentra su ciò che accade alla cultura visiva attraverso l'osservazione dello sviluppo e la diffusione della computer vision. Per la prima volta nella storia dell'umanità, gli esseri umani sono diventati oggetto di uno sguardo radicalmente altro, impegnato nella ricerca della materia di maggior valore sul mercato: i dati. Per il momento, tale sguardo è ancora supervisionato dal controllo umano, ma è lecito chiedersi fino a quando ciò accadrà e cosa comporterà l'eventuale emancipazione delle AI, alla luce dei processi istruiti affinché siano sempre più efficienti e degli elementi critici che emergono osservando proprio le immagini e i dispositivi che sono coinvolti in tali processi.

Le abitudini, i comportamenti e tutto il patrimonio di informazioni che compone il quadro delle vite delle persone rappresentano il “materiale raw” da cui le macchine possono estrarre ed elaborare le informazioni necessarie alla creazione di algoritmi sempre più raffinati. Per questo *“esistono set di training per qualunque biodato”*, come raccontano gli autori della mostra. Nell'ambito di questa rivoluzione in atto, *“la computer vision rappresenta una tipologia specifica di intervento sulla cultura visuale”*che porta con sé dei segni essenziali alla comprensione del presente.

Nel frattempo, i problemi che ci impone questo nuovo regime visivo si fanno sempre più concreti. Nella serie Multiple Encounter Dataset – II (Meds – II) del 2011, una serie di fotografie segnaletiche di persone decedute che hanno commesso reati reiterati nel tempo colpisce per la crudezza della ritrattistica. Non c'è alcun filtro artistico né pietà nelle foto ma solo il lavorio del tempo, che incide il dolore nella fisionomia dei volti fotografati, e l'evidenza di tragiche storie umane dimenticate negli archivi delle forze dell'ordine, a cui si aggiunge la consapevolezza che quelle foto contribuiscono a costruire un gigantesco database dell'umanità stessa, scandagliata e organizzata in una operazione tassonomica su vasta scala. Immagini per le quali nessuno dei soggetti ritratti ha dato il consenso all'utilizzo e che sono state scattate prima delle condanne, consegnate dall'FBI al National Institute of Standards, che ne ha disposto liberamente.

Da questi cataloghi umani quello che viene estratto è spesso un volto destinato a combaciare con un'altro volto ripreso magari da un circuito di sorveglianza, un'immagine di difficile lettura o di bassa qualità, utilizzata per individuare il responsabile di un reato. Un'idea seducente, sulla carta, ma che si è dimostrata

fallace. I set di training rispecchiano infatti i pregiudizi di chi li compone, e gli errori umani vengono amplificati dalle macchine, come racconta Angelique Carson su TechCrunch, che definisce il riconoscimento facciale uno “strumento di razzismo istituzionale”. I casi di procedure di riconoscimento effettuate dalla polizia statunitense in modo arbitrario, senza scientificità, sono documentati, così come è documentato il margine di errore dei sistemi di riconoscimento, che aumenta esponenzialmente nel caso di individui di colore o appartenenti a minoranze.



Dietro uno sterminato paesaggio di server che raccolgono quantità inimmaginabili di dati, si profila l'ombra dei *Tech Giants*, i grandi player che detengono le informazioni e che sono ormai titolati a trattare con i governi da pari. Da questa inedita concentrazione di potere nelle mani di pochi soggetti privati, trasformati in entità parastatali, discende quell'“asimmetria del potere” che interessa chiunque oggi abbia a che fare con le tecnologie digitali. La gestione opaca dei dati, spesso ceduti a terze parti senza che l'utente ne sia a conoscenza o possa esprimere una reale volontà di controllo, alimenta un sistema che ha generato lo scandalo di Cambridge Analytica e ha portato alla luce le ingerenze nelle elezioni

presidenziali statunitensi e nel processo della Brexit, rivelando il potenziale destabilizzante insito nella strumentalizzazione dei dati personali. È ormai evidente come il problema della privacy sia stato scavalcato da un problema di autonomia degli individui e di autodeterminazione, che riguarda tutti ma colpisce in maniera esponenziale chi è più fragile per via del proprio status sociale o perché appartenente a una delle tante minoranze. Essere costantemente sottoposti allo sguardo delle macchine sta cambiando il nostro modo di vivere, trasformandoci in sorvegliati perenni. La capilarizzazione del Panoticon di Bentham si è compiuta con il nostro benestare e gli spazi in cui possiamo agire senza essere identificati sono sempre più ristretti. Siamo diventati una piattaforma estrattiva senza rendercene conto, cedendo alla sirene della seduzione dell'esposizione socialmediatica e alla paranoia della sicurezza.

Il lavoro di Trevor e Paglen indica che è giunto il momento di interrogarsi sul carattere situato delle immagini utilizzate dall'AI, proprio perché all'apparenza esse si mostrano come neutre e “trasparenti”. La cultura che le ha prodotte, la nostra stessa cultura popolare, sembra averne perso il controllo e i soggetti delle immagini si ritrovano in una condizione di passività, incapaci di esercitare il diritto alla privacy.

The Japanese Female Facial Expression (Jaffe) Database è il nome di un controverso progetto basato sulle teorie di Paul Elkman, psicologo pioniere degli studi sul riconoscimento facciale che negli anni '70 individuava sei emozioni universali chiave che è possibile rilevare sul volto di una persona (paura, disgusto, rabbia, sorpresa, felicità tristezza). Malgrado le teorie di Elkman siano state ampiamente confutate, sono state comunque applicate da numerosi ingegneri impegnati nei progetti di machine learning. I due set presenti in mostra illustrano con chiarezza come l'attribuzione di sentimenti sul volto di alcune donne giapponesi sia del tutto opinabile, così come è evidente che ogni soggetto può aver simulato un'espressione senza averne provato i sentimenti reali. Il riconoscimento dei sentimenti, una sorta di evoluzione del riconoscimento facciale, comporta l'educare una macchina a riconoscere un codice espressivo che può non essere vero: al momento, la possibilità di commettere errori da parte del sistema AI è così alto da rendere quantomeno critico ogni utilizzo di queste tecnologie. Almeno fino a che sistemi di rilevazione dei dati biometrici non diventeranno così raffinati da poter raccogliere in un'istante valori incontrovertibili quali battito cardiaco, pressione sanguigna, dilatazione delle

pupille, sudorazione e così via.

Il compendio di volti che si susseguono nelle serie di UTK Face (Zhifei Zhang, Yang Song, Hairong Qi, 2017), 20.000 foto usate per classificare gli individui in base a razza, sesso ed età, evoca l'idea dell'uomo – cavia e la vertigine compilatoria tipica degli esperimenti di ingegneria sociale. Ancora più inquietante Image – Net (Li Fei – Fei, Kai Li, 2009), un progetto monumentale pensato allo scopo di mappare l'intero universo di oggetti e che include oltre 14 milioni di immagini suddivise in più di 21.000 categorie; in mostra sono presenti quasi un milione di fotografie suddivise in 2000 categorie che, alla lettura, sollevano molteplici interrogativi sulle attribuzioni di significato e sulle possibili derive interpretative di ogni immagine. Tra le categorie sono visibili alcune apertamente offensive, razziste o violente, abbinate a foto di persone ignare che hanno caricato i propri contenuti su piattaforme come Flickr e che sono all'oscuro di essere diventati un modello per una tipologia come “depresso”, “fallito”, “imbroglione”, “tossico” e così via. Inoltre, le classificazioni sono chiaramente frutto di una visione culturale, non di un processo scientifico, nel quale il confine labile tra osservazione e giudizio svanisce. Un punto evidenziato dall'opera Image – Net Roulette, creata dal Trevor Paglen Studio, che attraverso una rete neurale addestrata al riconoscimento di fotografie di persone dalle categorie del set di Image – Net, getta una luce sull'aspetto politico sotteso all'operazione di classificazione di soggetti umani da parte dei sistemi di AI.

La mostra si chiude con Age, Gender, and Emotion in the Wild (modelli di Gil Levi e Tal Hassner), un apparecchio realizzato dal Trevor Paglen Studio che propone un algoritmo di riconoscimento facciale sviluppato da Facebook e Amazon. Mentre mi sottopongo alla rilevazione della macchina, che analizza il mio volto e le espressioni, appaiono sullo schermo dei dati che mi riguardano: età approssimativa 25 – 32, femmina, emozione: paura. L'algoritmo deve avere una pessima vista, se è riuscito a individuare correttamente solo il mio sesso. O forse ci vede così bene da aver colto nelle micro espressioni del mio viso un'emozione nascosta di cui neanche io sono pienamente consapevole.

Mi torna in mente il romanzo di Jaques Spitz *L'occhio del purgatorio,* in cui un pittore cinico e depresso acquista casualmente, grazie a una pillola, una vista che gli permette di osservare il futuro, un futuro che si rivela in realtà un angosciante “presente invecchiato”. Se il protagonista del libro di Spitz acquisce il poco

invidiabile dono di vedere oltre la durata delle cose, il tempo a cui abbiamo assoggettato la nostra visione, quello di Internet, è un tempo per certi versi non meno sinistro, il tempo dell'infinito presente, dove tutto continua perennemente ad accadere e non conosce l'oblio. Le macchine che stiamo educando sembrano al momento essere il megafono di una visione umana distorta, inchiodata in un clamoroso fraintendimento. Un presente stravolto, perché non in progressione, come vorrebbe l'etimologia stessa della parola.

Come una GIF che si ripete all'infinito, l'occhio dell'AI fa proprio il nostro sguardo e lo disumanizza, riproponendone gli errori e le “sviste” su una scala incommensurabilmente ampia. Si sta completando il passaggio alla società panottica, dove siamo visti senza poter vedere, costantemente osservati senza avere coscienza dello sguardo sorvegliante. Proprio come auspicava Bentham, nelle parole di Foucault in *Sorvegliare e punire, “Lo schema panoptico, senza attenuarsi né perdere alcuna delle sue proprietà, è destinato a diffondersi nel corpo sociale; la sua vocazione è divenirvi funzione generalizzata.”*. Il passaggio dalla semplice condivisione di contenuti personali, trasformata nella spettacolarizzazione del privato, al restringimento degli spazi personali e alla compressione dell'esercizio dei diritti, è stato più rapido di quanto si potesse immaginare. Aprire la scatola nera tecnologica e comprenderne i meccanismi è la strada per riprendere il controllo dello sguardo e non cedere passivamente al voyeurismo tecnologico che ci costringe ad essere oggetti.

23_-_osservatorio_fondazione_prada_-_training_humans_23.jpg

Se continuiamo a tenere vivo questo spazio è grazie a te. Anche un solo euro per noi significa molto. Torna presto a leggerci e SOSTIENI DOPPIOZERO