

REMIX

# Umano, non troppo umano

*Training Humans*, la mostra di Kate Crawford e Trevor Paglen all'Osservatorio della Fondazione Prada.

di *Simona Arillotta* – 30 Settembre 2019



FaceApp è un'applicazione che permette di pre-vedere come sarà il nostro volto tra 20 o 30 anni. Il procedimento è semplice: gli utenti che desiderano scoprire quale sarà il loro aspetto tra qualche decennio, possono caricare una foto e attendere il verdetto. Non vi è nulla di scientifico, sia chiaro: non è un'analisi del DNA. **FaceApp è un gioco, un passatempo.** Attiva ormai da qualche anno, FaceApp è un'applicazione che ha raggiunto la sua massima diffusione solo negli ultimi mesi, arrivando a contare circa 80 milioni di nuovi utenti.

Un gioco, dicevamo. E i giochi, si sa: sono innocui, non fanno male a nessuno. O almeno così dovrebbe essere. Perché dopo una prima settimana di entusiasmo e condivisione massiccia delle proprie immagini modificate, il primo dubbio ha iniziato a serpeggiare tra gli utenti. La società che ha sviluppato l'applicazione ha sede a San Pietroburgo: che FaceApp sia un'operazione orchestrata da hacker russi per accedere ai nostri dati? In realtà, la società che ha creato l'app ha sì sede in Russia, ma non sembra aver nulla a che vedere con il Cremlino. Senza dare troppo peso al complottismo, e scongiurato l'attacco da parte degli hacker, è invece necessario sottolineare un aspetto

fondamentale. **FaceApp non è solo un'applicazione di fotografia: alle nostre foto non viene semplicemente applicato un filtro. Ciò che ci viene restituita è un'immagine modificata, elaborata grazie ad un sistema di Intelligenza Artificiale.**

*Training Humans*, la mostra ideata da Kate Crawford e Trevor Paglen e attualmente in corso presso l'Osservatorio della Fondazione Prada, **è un progetto espositivo dedicato alle immagini di training, ovvero ai set di fotografie utilizzate per "addestrare" le macchine a vedere, a identificare volti umani.** Per poter funzionare, infatti, i sistemi di computer vision e IA si servono di una base massiva di fotografie che, una volta etichettate e processate, ottimizzano il sistema che si vuole costruire. A partire dai primi casi di immagini realizzate per il riconoscimento facciale svolte negli anni sessanta da Woody Bledsoe per conto della Central Intelligence Agency americana, la mostra traccia una linea storica che racconta l'evoluzione nelle strategie di produzione di set di training: dagli anni novanta, quando la computer vision ha iniziato a prendere piede – è il caso del programma *FERET (Facial Recognition Technology)*, un programma di riconoscimento facciale per l'intelligence e le forze dell'ordine – fino ad arrivare ai giorni nostri e alla disponibilità di immagini garantita dal continuo afflusso di fotografie che quotidianamente riversiamo in rete attraverso i social network.

Osservando la mole di fotografie che ricoprono le pareti dello spazio espositivo, è possibile individuare due differenti tipologie di set: il primo, composto da ritratti fotografici – come il già citato *FERET*, ma anche il *Multiple Encounter Dataset-II*, un set realizzato per lo studio del riconoscimento facciale durante l'invecchiamento, composto da foto segnaletiche di stessi individui arrestati in diverse occasioni – riconducibile ai rilievi antropometrici e agli esperimenti di antropologia criminale condotti da Cesare Lombroso e Alphonse Bertillon, e dunque legati a **una dimensione di controllo e sorveglianza.** Il secondo gruppo di set è invece progettato per riconoscere le emozioni dell'uomo. Tra questi, il set di training *JAFFE (Japanese Female Facial Expression)* è composto da 213 immagini etichettate con sei espressioni: felicità, tristezza, paura, disgusto, rabbia, neutro. Un tentativo, quello di **leggere le emozioni del volto**, che trova un precedente nell'opera di Charles Le Brun il quale, già nel 1667, nel suo *Le figure delle passioni*, aveva provato a identificare e organizzare – secondo tavole fisiognomiche – un repertorio di espressioni umane.

Se dunque è possibile individuare una linea di continuità tra la tradizione tassonomica e fisiognomica e i primi esperimenti di riconoscimento facciale, le strategie di costruzione di set di training cambiano a partire dagli anni Duemila: le immagini vengono trovate tra le migliaia di fotografie che produciamo ogni giorno con i nostri telefoni, macchine digitali, e che vengono raccolte senza che sia necessario il nostro

consenso. **Ogni volta che postiamo una fotografia su Facebook o Instagram, partecipiamo** – inconsapevolmente – **all’ingrossamento di un sistema che classifica le immagini per età, sesso, razza, espressioni facciali e altri metadati.**

Nell’attraversare i due piani dell’Osservatorio, la prima associazione che la mia mente ha prodotto è stata con il funzionamento di FaceApp – da qui la scelta di utilizzare il caso dell’applicazione russa come *overture*. Ogni volta che l’immagine del nostro volto viene processata, questa subisce allo stesso tempo una serie di interpretazioni a partire da un algoritmo “addestrato” e una catalogazione in maniera automatica. Non solo: condividendo le nostre foto sull’applicazione, accettiamo implicitamente che queste stesse foto vengano immagazzinate in server e database e che, in futuro, possano essere ri-utilizzate da parte dell’azienda.

Una prassi, quest’ultima, che non si differenzia molto dal libero utilizzo delle nostre foto di cui già dispongono Facebook e Instagram. Ciò che però è ancora più importante sottolineare è che **il funzionamento di FaceApp ci consente di osservare più da vicino quel processo rivoluzionario in atto ormai da molti anni, quel cambiamento che ha investito il campo del visibile e che ci costringe a ripensare tanto la nozione stessa di immagine, quanto il nostro rapporto con essa.**

In un articolo del 2016 dal titolo *Invisible Images (Your Picture Are Looking at You)*, Trevor Paglen identificava una particolarità nel rapporto tra visione umana e immagini digitali: queste ultime sono visibili per un periodo circoscritto, poi continuano a circolare assumendo una nuova forma. Anzi, una non-forma: diventano *invisibili*. Se, dunque, è possibile identificare un’origine, **la durata delle immagini digitali ha un tempo impossibile da stabilire, mentre la traiettoria è definita dalla migrazione in sistemi di dati e sistemi di calcolo.** Ciò comporta uno spostamento radicale: non solo le immagini non lavorano più secondo logiche tradizionali di rappresentazione e imitazione, adesso **sono le immagini a guardare noi.**

Ed è proprio all’interno di questo passaggio radicale che deve essere letta l’operazione di Kate Crawford e Trevor Paglen. *Training Humans* non è solo il tentativo di ricostruire le strategie di produzione di set di training, ma come queste fotografie lavorano all’interno dei sistemi di IA e computer vision. È chiaro: il dataset viene progettato per agire in maniera autonoma, ma si basa su input umani. Ovvero: **le categorie attraverso cui i sistemi raccolgono ed etichettano le immagini che saranno alla base del funzionamento della computer vision hanno una pretesa di oggettività, ma rispecchiano – e reificano – forme di potere.**

È il caso del set di training *Utk Face*, composto da 20.000 volti al fine di classificare gli individui per razza, genere ed età. Una catalogazione che prevede però un genere binario – maschio/femmina – mentre la razza si suddivide in bianco, nero, asiatico, indiano, *altro*. Una classificazione viziata da una ideologia politica precisa e che inevitabilmente produrrà una identificazione distorta. Torniamo ancora a FaceApp, tacciata di razzismo perché uno degli algoritmi trasformava tutti i volti secondo la classificazione di “bianco caucasico”, perché non era addestrato a riconoscere altre classificazioni razziali. O, ancora, il set di immagini *Image-Net*, creato dai ricercatori di Stanford e Princeton per «mappare l'intero universo degli oggetti». All'interno di questa catalogazione, un milione di immagini classificano 2000 tipologie di persone, categorie definite in modo arbitrario, storico, politico, non scientifico. **E così Diane Keaton è una *spinster* (un'anziana donna non sposata), John Waters è un *whoremaster*, così come un bambino mediorientale è *analphabet*.**

Ciò che emerge è uno scenario inquietante: ***Training Humans* mostra come i sistemi di IA si basino su una disparità di potere tra chi può controllare e programmare e chi invece non può che subire.** Solo pochi giorni fa la Unilever ha utilizzato il riconoscimento facciale per i colloqui di lavoro. I candidati hanno risposto a una serie di domande che hanno registrato sui loro dispositivi (laptop o telefono); il video è stato processato a partire da una serie di coordinate specifiche, volte all'individuazione del candidato perfetto. Il risultato sarà obiettivo? Oppure l'algoritmo, specificatamente addestrato, penalizzerà alcuni candidati e ne giudicherà altri in modo errato?

La mostra termina con due sistemi di riconoscimento: il primo *Image-Net Roulette*, fa uso di una rete neurale addestrata a riconoscere fotografie di persone dalla categorie del set *Image-Net*; il secondo *Age, Gender, and Emotions in the wild* è lo studio di modelli sviluppati dai ricercatori di Facebook e Amazon per stimare l'età, il sesso e la condizione emotiva dei volti che rileva. Io sono stata inserita nella categoria non fumatori (errato) e secondo *Image-Net Roulette* sarei un'infermiera (secondo errore). Il secondo sistema di riconoscimento ha decretato l'appartenenza al genere femminile, che la mia età si colloca tra i 25 e i 50 anni, e che le mie emozioni erano: rabbia per l'87% (molto probabilmente corretto), disgusto 15% (possibile), e paura per il restante 8%. Il riconoscimento è risultato – in parte – sbagliato. **Mi chiedo se lo scarto dell'errore non conceda ancora un margine di libertà.** Di certo, è un bene che io non fossi tra i candidati per il colloquio alla Unilever.

### Riferimenti bibliografici

K. Crawford, T. Paglen, *Training Humans*, Quaderni Fondazione Prada, Milano 2019.

T. Paglen, *Invisible Images (Your Pictures Are Looking at You)*, in “The New Inquiry”, 8

dicembre 2016.