

I COMPITI PRINCIPALI DI UNA SEMIOTICA DELL'INTELLIGENZA ARTIFICIALE^[1]

MASSIMO LEONE*

ENGLISH TITLE: *The Main Tasks of a Semiotics of Artificial Intelligence*

ABSTRACT: The article indicates the essential tasks of a semiotics of artificial intelligence: studying the way it simulates the expression of intelligence; the way it produces content that is creatively endowed; the ideological assumptions of artificial intelligence within the culture that produces it. Artificial intelligence is, from a semiotic point of view, the predominant technology of fakery in the current era. On the strength of its studies on the false, semiotics can therefore also be applied to the analysis of the fake that, in increasingly sophisticated forms, is produced through artificial intelligence and through the deep learning of neural networks. The article focuses on the adversarial ones, trying to highlight their ideological assumptions and cultural developments, which seem to indicate the entry of human societies and cultures into the 'realm of the absolute fake'.

KEYWORDS: Generative Adversarial Networks; Fake; Simulation; Undecidability of the Digital Fake

1. Lo studio degli artefatti simulativi

La semiotica dovrebbe concentrarsi sullo studio degli sforzi per simulare i comportamenti intelligenti dell'uomo attraverso dispositivi non

(1) Questo saggio è il risultato di un progetto finanziato dal Consiglio Europeo della Ricerca (ERC) nell'ambito del programma di ricerca e innovazione Horizon 2020 dell'Unione Europea (Accordo di sovvenzione n. 819649–FACETS).

* Università di Torino; Università di Shanghai; Università di Cambridge; Fondazione Bruno Kessler, Trento.

organici e non umani. Questa simulazione può avvenire a livello di espressione, di contenuto o di entrambi. A livello di espressione, l'attenzione si concentra sulla riproduzione inorganica dei segni che gli esseri umani associano all'intelligenza. Si veda, ad esempio, il progetto "Gaze", finanziato dalla Disney⁽²⁾, un robot che simula le espressioni umane e, inoltre, le emula in un'interazione faccia a faccia (Fig. 1).

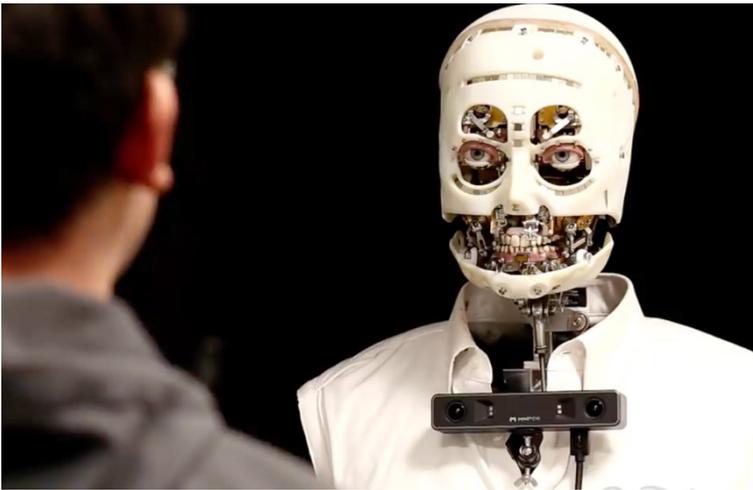


Figura 1. Il robot androide Gaze, finanziato dalla Disney, "interagisce" faccia a faccia con un essere umano; immagine di pubblico dominio.

In questa tecnologia, la materia essenzialmente inorganica viene organizzata, anche attraverso l'intelligenza artificiale, per trasmettere un'impressione di comprensione. Le espressioni facciali non sono solo cognitive, ma alcune lo sono e molte sono essenziali per comunicare la reciproca intelligibilità. A livello di espressione sono già stati fatti passi da gigante. Questo è evidente nel dominio del linguaggio verbale. L'espressione attraverso il linguaggio verbale è strettamente associata alle impressioni di intelligenza. Gli animali non umani tendono a non apparire dotati della stessa intelligenza e sensibilità degli esseri umani

(2) *Gaze*, un robot sviluppato da Walt Disney Imagineering con un gruppo di ricercatori dell'Università dell'Illinois a Urbana-Champaign e del California Institute of Technology, è il risultato di una ricerca avanzata nel campo della tecnologia e presenta una caratteristica molto interessante. È in grado di riprodurre alcune espressioni specifiche del volto umano. Ad esempio, è in grado di compiere piccoli movimenti della testa o di sbattere le palpebre.

perché non hanno accesso allo stesso linguaggio. Ma il linguaggio verbale può essere simulato a livello di espressione senza necessariamente comportare una simulazione a livello di contenuto. Per essere più precisi, l'intelligenza artificiale può simulare il linguaggio verbale attraverso una padronanza sintattica del linguaggio verbale umano. Possiamo parlare con il nostro altoparlante IA, come Siri o Alexa. Ma stiamo davvero parlando? Il nome stesso, “speaker”, rivela che la tendenza ad attribuire comportamenti intelligenti umani ai dispositivi è antica. In passato, si trattava per lo più di una metafora. Un altoparlante tradizionale pre-IA non era in grado di parlare o di essere interpellato. Si limitava a riprodurre meccanicamente i suoni provenienti da una fonte. Ma un altoparlante di nuova generazione collegato all'intelligenza artificiale può, in una certa misura, interagire con gli esseri umani attraverso il linguaggio verbale. La padronanza sintattica del linguaggio verbale fornisce una solida infrastruttura per la simulazione dell'intelligenza. Tuttavia, è anche fondamentale che questa intelligenza ci parli attraverso una voce umana.

Una semiotica dell'intelligenza artificiale può applicarsi, in primo luogo, allo studio di tutti quei segni che vengono utilizzati per simulare l'intelligenza. A livello espressivo, i riferimenti al corpo umano sono fondamentali. Il progetto FACETS del Consiglio Europeo della Ricerca si concentra sul volto digitale per la sua attuale rilevanza sociale. Un aspetto di esso riguarda l'intelligenza artificiale. La simulazione del volto umano sta diventando sempre più centrale nella simulazione dell'intelligenza artificiale, soprattutto a livello espressivo. La simulazione di una pragmatica intelligente richiede la simulazione di un'estetica intelligente. Il fenomeno dell'*uncanny valley* (“la vallata inquietante”) misura la resistenza all'estetica dei volti simulati, ma ne indica anche la potenza. In futuro, la simulazione di un volto intelligente sarà una caratteristica centrale dell'intelligenza artificiale. L'industria dei videogiochi, come sempre, è in anticipo. Un nuovo strumento di *MetaHuman* offre già agli utenti la possibilità di creare un'immagine digitale fotorealistica di un volto umano all'interno di un browser (Fig. 2). Epic Games ha annunciato questo software basato su browser e alimentato dal suo Unreal Engine⁽³⁾.

(3) L'azienda ha condiviso diversi video che illustrano questa tecnologia.



Figura 2. Un esempio d'immagine fotorealistica di volto creata dal browser di MetaHuman; immagine di pubblico dominio.

In un futuro più lontano, la generazione di volti digitali potrebbe combinarsi con quella di volti biologici. Come fanno notare Kim *et al.* (2020):

la rigenerazione di tessuti persi o danneggiati è l'obiettivo principale dell'ingegneria tissutale. Le tecnologie di bioprinting 3D sono state ampiamente applicate in molte aree di ricerca sulla rigenerazione dei tessuti e sulla modellazione delle malattie, con una risoluzione spaziale senza precedenti e una complessità simile a quella dei tessuti. Tuttavia, l'estrazione dell'architettura dei tessuti e la generazione di cianografie ad alta risoluzione sono compiti impegnativi per la rigenerazione dei tessuti. Tradizionalmente, tali informazioni spaziali sono ottenute da una raccolta di immagini microscopiche e poi combinate insieme per visualizzare le regioni di interesse. Per fabbricare questi tessuti ingegnerizzati, le immagini microscopiche renderizzate vengono trasformate in codice per informare un processo di bioprinting 3D. Se questo processo viene incrementato con approcci basati sui dati e ottimizzato con l'intelligenza artificiale, l'identificazione di un blueprint [formula] ottimale può diventare un compito realizzabile per la rigenerazione funzionale dei tessuti.⁽⁴⁾

(4) "Regenerating lost or damaged tissue is the primary goal of Tissue Engineering. 3D bioprinting technologies have been widely applied in many research areas of tissue regeneration and disease modeling with unprecedented spatial resolution and tissue-like complexity.

A un livello più profondo, l'intelligenza artificiale è sempre più applicata alla genomica e all'ingegneria genetica. I limiti alla ricreazione di specifici volti umani attraverso l'ingegneria genetica potenziata dall'intelligenza artificiale sono attualmente più etici che tecnici. I geni possono essere modificati, come nel caso dell'editing genetico CRISPR, una tecnica d'ingegneria genetica in biologia molecolare che consente di modificare i genomi degli organismi viventi⁽⁵⁾. Per il momento, l'intelligenza artificiale e l'apprendimento profondo stanno rapidamente progredendo nella simulazione di volti fotorealistici. La semiotica deve studiare la retorica dell'intelligenza artificiale, cioè i segni espressivi che simulano un comportamento o un dispositivo d'intelligenza (il volto è il dispositivo espressivo primario dell'intelligenza umana).

A livello di contenuto, l'attenzione si concentra sulla riproduzione inorganica del significato che gli esseri umani associano all'intelligenza. Il significato prodotto dall'intelligenza artificiale è, a un primo livello, comportamentista, cioè l'uomo si rende conto che le macchine sono intelligenti perché producono comportamenti intelligenti. Una calcolatrice può produrre nuovi risultati sulla base di input noti, ma difficilmente potrebbe essere considerata un dispositivo d'intelligenza. Il comportamento intelligente deve andare oltre il calcolo. I suoi risultati devono essere sintetici, non semplicemente analitici. Nei termini di C.S. Peirce, una macchina produce un comportamento che può essere qualificato come intelligente se coinvolge non solo la deduzione e l'induzione, ma anche l'abduzione. Anche in questo caso, tuttavia, la credibilità dell'intelligenza artificiale dipende dalla sua estetica. Una vecchia calcolatrice e un robot odierno sono entrambi dispositivi computazionali, ma il secondo è dotato di un'interfaccia simile a quella umana (una voce, un volto).

However, the extraction of tissue architecture and the generation of high-resolution blueprints are challenging tasks for tissue regeneration. Traditionally, such spatial information is obtained from a collection of microscopic images and then combined together to visualize regions of interest. To fabricate such engineered tissues, rendered microscopic images are transformed to code to inform a 3D bioprinting process. If this process is augmented with data-driven approaches and streamlined with machine intelligence, identification of an optimal blueprint can become an achievable task for functional tissue regeneration." (Trad. nostra).

(5) Si basa su una versione semplificata del sistema di difesa antivirale batterico CRISPR-Cas9. Introducendo in una cellula la nucleasi Cas9 complessa con un RNA guida sintetico (gRNA), il genoma della cellula può essere tagliato nel punto desiderato, consentendo di rimuovere i geni esistenti e/o di aggiungerne di nuovi in vivo (negli organismi viventi).

Più in generale, bisogna distinguere tra simulazione dell'intelligenza e intelligenza artificiale vera e propria. La prima è soprattutto una questione di segni espressivi, mentre la seconda richiede creatività a livello di contenuti. Ma l'intelligenza artificiale è creativa? A questo livello, quello dello studio dell'intelligenza artificiale come tecnologia produttrice di contenuti significativi, ciò che conta è definire l'intelligenza, i diversi tipi d'intelligenza e i modi in cui essi generano creatività. Esistono diversi modi in cui i contenuti creativi vengono creati dalle macchine attraverso l'intelligenza artificiale (nello specifico, il deep learning). Il più comune oggi è il riconoscimento di modelli nei big data. L'intelligenza artificiale è in grado di riconoscere configurazioni che inizialmente non vengono individuate dai ricercatori a causa delle dimensioni dei database su cui può operare. In passato, il vantaggio dell'intelligenza artificiale sull'intelligenza umana era in gran parte di tipo computazionale. I campioni di scacchi continuavano a vincere sulle macchine perché erano in grado di identificare e pianificare schemi di gioco che venivano ignorati dalle macchine. Ma ora la quantità sta diventando qualità.

Negli anni '80, all'apice della sua carriera, il campione del mondo di scacchi Garry Kasparov sosteneva che non sarebbe mai esistito un programma di scacchi in grado di sconfiggerlo. E in effetti, nel 1989, giocò due partite contro il computer *Deep Thought* dell'IBM, vincendole entrambe. Nel 1996, Kasparov sconfisse il successore di *Deep Thought*, *Deep Blue*, in una sfida su sei partite per 4 a 2, ma è stato il primo campione del mondo di scacchi a perdere una partita in condizioni di torneo contro un programma di scacchi. L'anno successivo, Kasparov fu sconfitto da *Deep Blue* nella rivincita. *Deep Blue* sorprese il mondo con un gioco "istintivo" e superiore che sembrava creativo sotto molti aspetti. Kasparov diffuse la voce che IBM doveva aver barato. Oggi, in 100 partite, il campione del mondo di scacchi in carica Magnus Carlsen non otterrebbe una sola vittoria contro il miglior programma di scacchi del mondo. Carlsen ha attualmente un rating Elo di 2.845 (febbraio 2019), mentre Stockfish 9 ha un rating di 3.438 (questi rating non sono FIDE, ma il parco giocatori computazionali è molto più competitivo di quello degli umani, per cui teoricamente un rating FIDE per Stockfish 9 dovrebbe essere ancora più alto). Le recenti prestazioni

dell'intelligenza artificiale nel tipico gioco cinese del Go sono ancora più spettacolari. Il gioco del Go è stato a lungo considerato il più impegnativo dei giochi classici per l'intelligenza artificiale, a causa del suo enorme spazio di ricerca e della difficoltà di valutare le posizioni e le mosse della scacchiera. Silver *et al.* (2016) introducono un nuovo approccio al Go al computer che utilizza “reti di valore” per valutare le posizioni sulla scacchiera e “reti di politica” per selezionare le mosse. Queste reti neurali profonde sono addestrate da un'inedita combinazione di apprendimento supervisionato da partite di esperti umani e apprendimento per rinforzo da partite di auto-gioco. Senza alcuna ricerca in anticipo, le reti neurali giocano a Go al livello dei più avanzati programmi di ricerca ad albero Monte Carlo che simulano migliaia di partite casuali di auto-gioco. Silver *et al.* (2016) introducono anche un nuovo algoritmo di ricerca che combina la simulazione Monte Carlo con reti di valori e politiche. Utilizzando questo algoritmo di ricerca, il programma AlphaGo ha raggiunto una percentuale di vittorie del 99,8% contro altri programmi di Go e ha sconfitto il campione europeo umano di Go per 5 partite a 0, essendo questa la prima volta che un programma informatico ha sconfitto un giocatore professionista umano nel gioco completo di Go, un'impresa che in precedenza si pensava fosse lontana almeno un decennio.

2. Lo studio delle ideologie simulative

Una filosofia della comunicazione digitale orientata alla semiotica mira a leggere le tecnologie del senso nel lungo periodo della storia dei sistemi semiotici umani, per rivelare le ideologie implicite che stanno alla base della creazione di nuovi dispositivi, processi e artefatti di significato. L'intelligenza artificiale non fa eccezione, poiché il suo sviluppo è solitamente sostenuto da preconcetti specifici su cosa sia l'intelligenza, su come dovrebbe funzionare e su quali tipi di risultati dovrebbe generare nel mondo. Poiché l'intelligenza artificiale sembra essere, almeno in questa fase del suo sviluppo, una simulazione dell'intelligenza umana, la semiotica può studiarla come un caso particolare di falsificazione, un tema che questa disciplina conosce bene. Ogni cultura e ogni

epoca storica sono caratterizzate dalle specifiche modalità semiotiche che adottano nella produzione del falso; l'intelligenza artificiale sta diventando la modalità principale nell'attuale produzione digitale del falso. La specie umana è dotata di una capacità innata di dare origine a rappresentazioni che intenzionalmente non corrispondono ad alcuna realtà ontologica. Le tecnologie e i linguaggi del falso, tuttavia, cambiano nel tempo e nello spazio. Con la tecnologia digitale, la comunicazione telematica e, soprattutto, con l'intelligenza artificiale e il deep learning, la cultura umana del falso sta varcando una soglia decisiva.

Nel mondo digitale, le culture umane entrano nel regno del "falso assoluto". Ciò è dovuto, in primo luogo, alle caratteristiche materiali della tecnologia digitale: tutto ciò che può essere oggetto di rappresentazioni digitali, può anche essere oggetto di rappresentazioni digitali senza riferimento ontologico. Qualsiasi immagine digitale che verrà prodotta del mio volto invecchiato in un futuro la cui ontologia non esiste ancora, può essere ricostruita nel presente da una simulazione digitale. In secondo luogo, il dominio del falso assoluto è causato dal potere dell'accumulazione quantitativa: un'immagine del mio volto ringiovanito può circolare sui social media in modo così intenso e virale che finirà per rappresentare la mia identità sul web. In terzo luogo, il dominio del falso assoluto è causato dalle sue nuove modalità di creazione: in precedenza, la falsità era una questione giocata tra contraffattori e intenditori (ad esempio, nel campo dell'arte); ora questa partita è giocata sempre più da algoritmi con risultati ampiamente imprevedibili. L'intelligenza artificiale applicata alla creazione del falso è sempre stata praticata nei confronti di un oggetto particolare, ovvero il volto, che è l'interfaccia principale e il dispositivo umano più importante per la comunicazione interpersonale.

La semiotica è perfettamente attrezzata per condurre uno studio il cui oggetto si trova all'incrocio tra il falso, il volto e le rappresentazioni digitali costruite dall'intelligenza artificiale. Per quanto riguarda il falso, tutti i padri fondatori della semiotica si sono occupati dell'argomento (Ousmanova 2004): 1) Charles S. Peirce nella tradizione americana (Cooke 2014); 2) le principali voci della semiotica strutturale, da un numero speciale della rivista francese *Communication* dedicato al concetto di "vraisemblable": Tzvetan Todorov, Gérard Genette,

Christian Metz, Julia Kristeva, Gérard Genot, Roland Barthes e altri (Todorov 1968); Baudrillard è tornato sull'argomento (1987; 2000); più recentemente, una tavola rotonda su "Post-verità e democrazia" è stata organizzata da Jacques Fontanille durante il Congresso dell'Associazione Francese di Semiotica a Lione, dall'11 al 14 giugno 2019 (Di Caterino 2020); Umberto Eco ha scritto molto sulla falsificazione (1995), ha curato un numero speciale della rivista semiotica *Versus* su "Fakes, Identity, and the Real Thing" (1987; con saggi di Eco, Prieto, Calabrese, *et al.*) e ha trattato l'argomento in numerosi saggi e romanzi (*Il pendolo di Foucault, Il cimitero di Praga, Numero Zero*); 3) Jurij M. Lotman ha più volte affrontato la questione del falso (Andrews 2003, p. 101; Makarychev e Yatsyk 2017). Anche il volto è stato oggetto di ricerca fin dagli inizi dell'antropologia moderna (Leone 2021). In semiotica, dopo i libri di Patrizia Magli sulla fisiognomica, il progetto ERC FACETS ha costantemente sviluppato la letteratura su questo tema.

3. Caso di studio: i volti dell'intelligenza artificiale

La ricerca semiotica sulle rappresentazioni digitali del volto sta crescendo sempre di più, soprattutto per quanto riguarda la rappresentazione del volto da parte dell'intelligenza artificiale. Per sviluppare un'analisi delle ideologie semiotiche che stanno alla base della creazione di volti sintetici, bisogna però guardare all'origine degli algoritmi che, negli ultimi anni, hanno rivoluzionato le pratiche in questo campo. Bisogna tornare al loro testo fondante, un articolo che il giovane Ian J. Goodfellow ha pubblicato il 10 giugno 2014 — quando era dottorando all'Università di Montreal — con il titolo *Generative Adversarial Nets*. Da allora, questo ricercatore è diventato un guru mondiale dell'intelligenza artificiale e in particolare del deep learning, ricoprendo posizioni di primo piano nel settore, tra cui quella di direttore del dipartimento di machine learning nel gruppo dei progetti speciali di Apple.

Insieme a un gruppo di amici dottorandi in informatica, Ian J. Goodfellow ha proposto un nuovo framework per la stima di modelli generativi attraverso un processo contraddittorio, in cui due modelli vengono addestrati simultaneamente: un modello generativo, che

cattura la distribuzione dei dati, e un modello discriminativo, che stima la probabilità che un campione provenga dai dati di addestramento piuttosto che dal modello generativo. Il modello generativo antagonista ha portato ad applicazioni rivoluzionarie nell'intelligenza artificiale e nel deep learning, tra cui la creazione di "volti artificiali" (Leone 2021) e di *deepfakes*. La semiotica è già stata applicata allo studio dell'intelligenza artificiale. Tuttavia, ha guardato ai suoi risultati e ai suoi prodotti, mentre sarebbe essenziale esaminarne, con la medesima prospettiva, i presupposti ideologici e la struttura soggiacente al suo funzionamento.

Lo schema di produzione dell'intelligenza artificiale immaginato da Goodfellow consiste in un'opposizione tra due istanze; il quadro della semiotica strutturale può quindi contribuire alla sua intelligibilità. Nell'architettura astratta delle GAN compaiono due attanti principali. Il primo è un attante generatore che esamina una configurazione di dati e produce un testo che potrebbe essere emesso da questa stessa configurazione; il secondo è un attante discriminatore che esamina il testo così prodotto e valuta se esso provenga dalla configurazione di dati o dall'attante generatore. Da un punto di vista epistemico, quindi, l'attante generatore mira a "far sembrare" e quindi a "far passare per vero" ciò che non lo è, mentre l'agente discriminatore mira a "far apparire" e quindi a "smascherare come falso" ciò che non è vero.

In termini matematici, per apprendere la distribuzione del generatore p_g sui dati x , si definisce un a priori sulle variabili di rumore in ingresso $p_z(z)$, quindi si rappresenta una mappatura nello spazio dei dati come $G(z; \theta_g)$, dove G è una funzione differenziale rappresentata da un *perceptron* multistrato con parametri θ_g . Viene definito anche un secondo *perceptron* multistrato $D(x; \theta_d)$, che produce un singolo scalare. $D(x)$ rappresenta la probabilità che x provenga dai dati piuttosto che dal p_g . D è formato per massimizzare la probabilità di assegnare l'etichetta corretta sia agli esempi di addestramento sia ai campioni di G . Contemporaneamente, G è formato per minimizzare $\log(1 - D(G(z)))$ (Fig. 3).

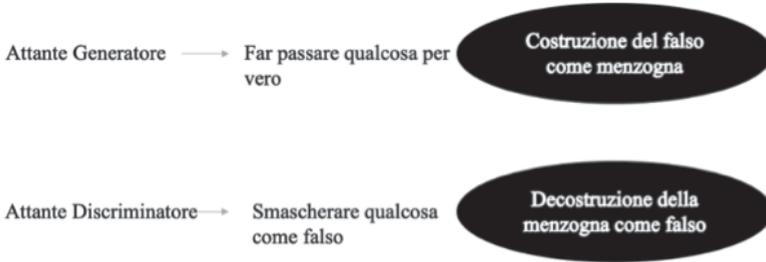


Figura 3. La dialettica attanziale nelle reti neurali avversariali.

L'attante discriminatore (D) è allo stesso tempo l'anti-soggetto dell'attante generatore (G), il suo destinatario e il suo coadiuvante. D sanziona i prodotti di G, designandoli come veri o falsi (cioè come provenienti o meno dall'insieme dei dati); quando la sanzione è positiva, però, determina ipso facto anche la sconfitta di D rispetto al suo anti-soggetto G, e viceversa: quando la sanzione è negativa, ciò porta alla sconfitta di G rispetto al suo anti-soggetto D. Ma in ogni caso D è sempre anche il coadiuvante di G, perché il secondo impara da ogni sanzione di D come ingannare meglio la sua controparte (Fig. 4).

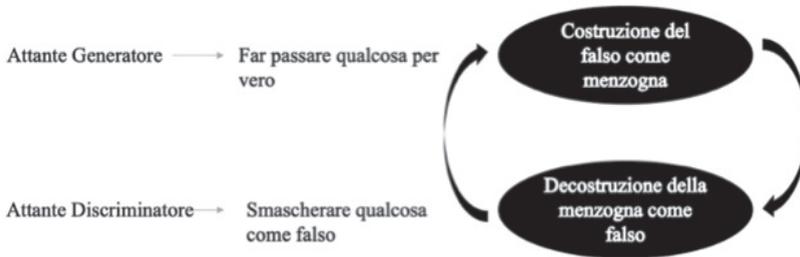


Figura 4. La spirale attanziale nelle reti neurali avversarie.

Quando si legge, attraverso la semiotica, l'articolo fondativo "Generative Adversarial Nets" (GAN), si rimane colpiti soprattutto da due elementi: 1) la concezione dell'intelligenza artificiale che esprime si basa sull'idea di antagonismo (né cooperazione, né semplice competizione); 2) la metafora che meglio spiega la nuova architettura del deep learning è quella del falsario e dell'intenditore (soprattutto nella

fabbricazione di denaro). Entrambi gli aspetti meritano un'ulteriore riflessione filosofica e semiotica, perché questa nuova architettura dell'intelligenza artificiale trova oggi applicazione in molti ambiti professionali e sociali, e in particolare nella creazione d'immagini e video sintetici di volti statici o in movimento, sempre più associati a teste, a corpi, oltre che a contesti sintetici, e che spesso comunicano attraverso molteplici sistemi di segni, come espressioni facciali, gesti, movimenti, frammenti di discorso verbale, canti, danze, ecc.

Il diagramma delle GAN può essere letto attraverso la metafora proposta dallo stesso Goodfellow nel 2014: D e G si comportano rispettivamente come un conoscitore e un falsario. Il falsario esamina la moneta in circolazione e cerca di produrne esemplari falsi; l'intenditore esamina la moneta prodotta dal falsario senza conoscerne l'origine e cerca di capire se si tratta di moneta falsa o autentica. Così facendo, però, il conoscitore fornisce al falsario informazioni che gli saranno utili per creare moneta falsa ancora più difficile da distinguere da quella autentica. Ma l'intenditore impara anche a discriminare sempre meglio tra denaro autentico e falso. Anche la metafora del mercato dell'arte può rendere l'idea di questa spirale di generazione e discriminazione: un falsario cerca di mettere in circolazione falsi Modigliani, mentre un conoscitore cerca di distinguerli dai Modigliani autentici; così facendo, però, il secondo fornisce al primo delle informazioni su come meglio falsificare le opere; viceversa, anche il primo impara dal secondo come falsificare le opere dell'artista italiano.

Ci si deve interrogare sulla natura dell'attante osservatore di questa spirale. I prodotti del modulo generatore non sono infatti sanciti solo dal modulo discriminatore, ma anche da un destinatario umano, che coincide, almeno in prima istanza, con il destinatario delle GAN. I moduli sono programmati da un mittente umano, ma i loro "comportamenti" non sono del tutto prevedibili, anche a causa del disallineamento computazionale tra cognizione umana e intelligenza artificiale. Il programmatore umano è quindi sia il mittente che il destinatario dei prodotti dell'interazione tra il modulo generatore e quello discriminatore. Inoltre, oltre a questo attore osservatore professionale, ce n'è un altro che è costituito da coloro che riceveranno i prodotti del modulo generatore senza conoscerne l'origine. La spirale appena descritta

finisce con l'aumentare sempre più l'incertezza epistemica di questo attore osservatore non professionista.

Per dirla in termini più semplici nel contesto della prima metafora: la competizione tra falsari e conoscitori mette in circolazione denaro o opere d'arte che sono falsi, ma che sono sempre più difficili da riconoscere come tali, soprattutto da parte dell'osservatore che si trova fuori dalla spirale. La circolazione massiccia di un falso non più identificabile come tale finisce per gettare discredito epistemico anche sulle opere d'arte e sulla moneta autentica. In questo risiede, forse, il pericolo più importante della "spirale del falso". Alcuni ricercatori hanno gettato una luce positiva sui GAN, suggerendo che la loro dialettica interna dovrebbe piuttosto essere paragonata a quella tra insegnante e studente. Il modulo generatore sarebbe quindi come uno studente che cerca di produrre rappresentazioni credibili da un insieme di dati, mentre il modulo discriminatore sarebbe come un insegnante che esamina e valuta queste rappresentazioni. Questo è in parte vero, ma ciò che fa la differenza è che, nel mondo delle GAN, le rappresentazioni del modulo generatore iniziano a circolare senza riferimento al contesto di apprendimento. Inoltre, è come se le GAN non consentissero all'allievo di superare definitivamente il maestro, obbligando il primo a condividere con il secondo il modo in cui ci è riuscito.

Questa è anche la differenza tra il falso digitale e quello analogico. La specie umana è intrinsecamente capace di produrre intenzionalmente false rappresentazioni della realtà, cioè rappresentazioni che, pur essendo prive di origine indicale, ne simulano una creando un effetto di senso iconico. Questa capacità è stata probabilmente selezionata dall'evoluzione biologica della specie come adattativa, in quanto ha permesso agli esseri umani di sperimentare mentalmente situazioni potenzialmente pericolose senza doverlo fare empiricamente. Inoltre, ha permesso loro di proteggersi dai predatori o di intrappolare le prede. Si tratta di una capacità che non manca in altre specie, sia vegetali che animali. Una delle peculiarità più notevoli degli uccelli liriformi, ad esempio, è la loro capacità di imitare i suoni, come quelli di altri uccelli e di vari elementi naturali, ma anche quelli dell'ambiente umano, come lo scatto di una macchina fotografica, il rumore di una motosega, il suono di un allarme antincendio, il frastuono di un cilindro idraulico, ecc.

Nella specie umana, tuttavia, questa capacità, espressa nel linguaggio e attraverso di esso, ha dato luogo a una sorta di “esattamento”, consistente nella capacità di attribuire piacere e valore estetico a rappresentazioni intenzionalmente false, che ha innescato a sua volta un’enorme produzione di testi di finzione. La tecnologia digitale introduce un cambiamento qualitativo e quantitativo essenziale nella storia del rapporto della specie umana con il falso. Il digitale è dotato di una materialità proteiforme la cui manifestazione semiotica è completamente programmabile, cosa che non è mai avvenuta nella manifestazione dei testi predigitali. Ciò implica che qualsiasi rappresentazione digitale che abbia una relazione indicale con il suo oggetto può essere riprodotta in modo identico anche quando questa relazione è assente; la pittura può, ovviamente, simulare volti che non esistono, eppure lo scarto tra il volto ontologico e quello dipinto sarà sempre evidente, cosa che non accade nel digitale. Questo assorbe il senso di indicizzazione caratteristico della fotografia e lo riproduce in assenza di sostrato indicale; allo stesso tempo, introduce la piena programmabilità nella costruzione dell’immagine fotografica. Il quadro può rappresentare oggetti inesistenti ma non può far credere alla loro esistenza; la fotografia analogica può far credere all’esistenza degli oggetti che rappresenta, ma non può rappresentare oggetti inesistenti, almeno non efficacemente; la fotografia digitale può far credere all’esistenza degli oggetti inesistenti che essa rappresenta.

4. Conclusioni

L’applicazione dell’intelligenza artificiale, e in particolare del deep learning delle GAN, alla produzione della manifestazione materiale delle rappresentazioni digitali le sottrae alla valutazione umana. Il falso è inseparabile dalla specie umana, eppure è la prima volta nella storia della specie che agenti non umani sono stati messi nella condizione di produrre un falso la cui valutazione sfugge sempre più all’uomo ed è affidata sempre più a un esame che viene a sua volta effettuato per mezzo dell’intelligenza artificiale. I falsi digitali possono ormai riprodursi e circolare con una facilità senza precedenti, e questo aspetto quantitativo si traduce anche in un cambiamento qualitativo: è come se l’arte autentica

dovesse difendersi da un numero infinito di contraffattori che lavorano incessantemente e molto velocemente alla produzione di copie.

Il falso digitale è destinato, a lungo andare, a essere indistinguibile dal “reale digitale”; nel caso dei volti, ad esempio, è solo questione di tempo prima che non si possa più sapere dalla foto digitale di un volto se la foto sia stata prodotta da un volto biologico e ontologico o se si tratti di un’immagine sintetica. La semiotica tende a problematizzare il concetto logico di verità come adeguatezza al reale, considerando piuttosto le condizioni semiotiche che producono un “effetto di realtà”. Tuttavia, spiegare la retorica dell’effetto di realtà senza postulare una realtà ontologica conduce a ineludibili aporie. Allo stesso modo, si può ben problematizzare l’effetto di realtà di una fotografia analogica, ma si deve anche riconoscere che l’avvento del digitale, e del deep learning digitale applicato alla creazione di immagini, mina la possibilità di distinguere tra un’immagine referenziale dotata di un effetto di realtà e un’immagine sintetica che produce lo stesso effetto.

Riferimenti bibliografici

- ANDREWS E. (2003), *Conversations with Lotman: Cultural Semiotics in Language, Literature, and Cognition* [Toronto Studies in Semiotics and Communication], University of Toronto Press, Toronto, Buffalo e Londra.
- BAUDRILLARD J. (1987), *Au-delà du vrai et du faux, ou le malin génie de l'image*, “Cahiers internationaux de sociologie”, nouvelle série, numero speciale su “Nouvelles images, nouveau réel”), gennaio–giugno, 82: 139–45.
- (2000), *The Vital Illusion*, The Wellek Library Lectures, New York.
- COOKE E.F. (2014), “Peirce and the ‘Flood of False Notions’”, in T. THELLEFSEN, B. SØRENSEN, C. DE WAAL (a cura di) (2014), *Charles Sanders Peirce in His Own Words: 100 Years of Semiotics, Communication and Cognition* [Semiotica, comunicazione e cognizione 14], De Gruyter Mouton, Boston, 325–31.
- DI CATERINO A. (2020), *Fake News: Une mise au point sémiotique*, online, “Actes Sémiotiques”, 123; disponibile nel sito <https://www.unilim.fr/actes-semiotiques/6445> (ultimo accesso il 14 maggio 2022).
- ECO U. (1984), *On Fish and Buttons: Semiotics and Philosophy of Language*, “Semiotica”, 48, 1–2: 97–118.

- (1987), *Fakes, Identity and the Real Thing*, numero speciale di *Versus*, 46, Bompiani, Milano.
- (1995), *Faith in Fakes: Travels in Hyperreality* (1986), Minerva, Londra.
- (1975), *Trattato di semiotica generale*, Bompiani, Milano.
- KIM J. *et al.* (2020), *Engineering Tissue Fabrication with Machine Intelligence: Generating a Blueprint for Regeneration*, “Frontiers in Bioengineering and Biotechnology”, 7, 443: online; DOI: <https://doi.org/10.3389/fbioe.2019.00443>.
- LEONE M. (a cura di) (2021), *Volte artificiali/Artificial Faces*, numero speciale di *Lexia*, 37–38, Aracne, Roma.
- MAKARYCHEV A.S., YATSYK A. (2017), *Lotman’s Cultural Semiotics and the Political: Reframing the Boundaries*, Rowman & Littlefield International, Londra.
- OUSMANOVA A. (2004), *Fake at Stake: Semiotics and the Problem of Authenticity*, “Problemos”, 66, 1: 80–101.
- SILVER D., HUANG A., MADDISON C. *et al.* (2016) *Mastering the Game of Go with Deep Neural Networks and Tree Search*, “Nature” 529, 484–89; DOI: <https://doi.org/10.1038/nature16961>.
- TODOROV T. (a cura di) (1968), *Recherches sémiologiques le vraisemblable*, numero speciale di *Communications*, 11.