# La Interpretazione delle Immagini:la Intelligenza Artificiale ha superato l'uomo?

**Data** 09 maggio 2021
**Categoria** Medicinadigitale

Vi sono importanti differenze tra i processi visivi degli umani ed i sistemi di interpretazioni delle immagini della intelligenza artificiale. Le riassumiamo sinteticamente per fornire una idea generale sui rispettivi limiti e potenzialità.

**UMANI: Il principio fondamentale della visione umana è che "vedere è ricordare e confrontare(1)".**
Le immagini vengono recepite dalla retina e trasmesse a vari centri nervosi che le organizzano in maniera integrata: la nostra percezione avviene con le modalità descritte dalla psicologia percettiva della Gestalt: figura-sfondo, vicinanza-lontananza, continuità-discontinuità, somiglianza-diversità, forme naturali ("buone forme") e innaturali, chiusura e apertura di linee, superfici, spazi, ecc(2). Si arriva così ad una visione di Insieme che i nostri centri nervosi catalogheranno secondo modelli e mappe da noi costruiti nel corso della nostra vita. Il processo è veloce ed in larga misura non consapevole. Un singolare fenomeno visivo, che sta alla base della visione umana e che non è presente nei sistemi di IA è la risoluzione del "Problema inverso dell'ottica(1)": immagini diverse hanno una medesima proiezione retinica ma appaiono ben distinte grazie al nostro sistema nervoso che le elabora con meccanismi molto efficienti anche se in larga parte sconosciuti. Il "problema inverso" ci espone tuttavia a imbarazzanti illusioni visive molte delle quali magistralmente illustrate dal grande artista Cornelis Escher.
Le illusioni visive sono tuttavia evidenziate e corrette dal nostro cervello attraverso un espediente semplice, un istintivo geniale meccanismo, la seconda occhiata: l'istinto, in una frazione di secondo ci dice che non è possibile e che dobbiamo rivedere l'immagine e verificarne la correttezza.
Le macchine super-intelligenti non possiedono questa facoltà a meno che ogni situazione inaspettata e/o paradossale non sia esplicitamente prevista dai programmatori.

**IMMAGINIDIGITALI**
L'immagine digitale è costituita da punti ottici (pixel), ciascuno dei quali è associato a coordinate numeriche registrate nella memoria del computer: il computer utilizza algoritmi che dalle coordinate numeriche formano l'immagine. Le informazioni estratte possono consentire diversi livelli di precisione: basso livello (come le statistiche sulla presenza dei vari toni di grigio o colori, sui bruschi cambiamenti di luminosità, ecc.); livello intermedio (caratteristiche relative a regioni dell'immagine ed a relazioni tra regioni) o alto livello (riconoscimento degli oggetti). Mentre per i livelli bassi e medi software ed algoritmi ci forniscono ottime prestazioni, per i livelli più elevati, quelli interpretativi, vi sono ancora grosse difficoltà: un interessante studio canadese ha dimostrato come la semplice introduzione di un finto elefante in una stanza abbia completamente disorientato il sistema di intelligenza artificiale che non era più in grado di riconoscere neppure gli oggetti precedentemente identificati(3).
Questo serio limite è una della maggiori preoccupazioni degli esperti dei sistemi di guida automatica delle auto, anch'essi disorientati da eventi non esplicitamente previsti dagli algoritmi. Ad esempio, una gallina che attraversa la strada disorienta completamente i sistemi di riconoscimento visivo.
I sistemi di riconoscimento visivo, se dobbiamo credere alle ditte che li producono, i sono pressoché infallibili. Sta di fatto che sono descritti e documentati clamorosi errori: da un poveruomo massacrato di botte dalla polizia di Denver, imprigionato e prosciolto, dopo un anno di detenzione, quando fu riconosciuto l'errore dell'algoritmo(4), a uno sveglissimo bambino di 10 anni che violò nientemeno che il sofisticato sistema di riconoscimento visivo di un pregiato iPhone della Apple e, per dimostrare che poteva a sua discrezione ri-violarlo, pubblicò il video su YouTube(5)..
Per gli esperti di visione artificiale c'è ancora molto lavoro da fare: "provaci ancora Sam"!!!

**Riccardo De Gobbi e Giampaolo Collecchia**

Bibliografia
1) Purves D, Brannon E, Cabeza R, Huettel S. Neuroscienze cognitive. Bologna: Zanichelli, 2014; 133.
2) Köhler Wolfgang: La psicologia della Gestalt Feltrinelli Edit. Milano 1998
3) Rosenfeld A, Zemel R, Tsotsos J. The Elephant in the Room (2018). 10.13140/ RG.2.2.15295.20642.
4) Fry H. Hello World. Milano: Bollati Boringhieri, 2018; 156.
5) Watch a 10-Year-Old's Face Unlock His Mom's iPhone X: https://www.wired.com/ story/10-year-old-face-id-unlocks-mothers-iphone-x

**Perapprofondire:**
**Collecchia G. De Gobbi R.: Intelligenza Artificiale e Medicina Digitale. Una guida critica. Il Pensiero Scientifico Ed. Roma 2020**
**pensiero.it/catalogo/libri/pubblico/intelligenza-artificiale-e-medicina-digitale**