

Analisi e valorizzazione del patrimonio artistico mediante Intelligenza Artificiale

Giovanna Castellano¹, Gennaro Vessio²

¹Università degli Studi di Bari, Italia – giovanna.castellano@uniba.it

²Università degli Studi di Bari, Italia – gennaro.vessio@uniba.it

ABSTRACT

Soluzioni basate su Intelligenza Artificiale stanno già potenziando numerosi campi del sapere e dell'attività umana, inclusa l'arte. Infatti, i recenti progressi nel campo dell'Intelligenza Artificiale, insieme con la crescente disponibilità di collezioni di opere d'arte digitalizzate, stanno offrendo nuove opportunità agli specialisti in questi settori perché assistano la comunità artistica, ma anche semplicemente gli appassionati, con nuovi strumenti automatici e "intelligenti". Il presente articolo delinea attività di ricerca condotte nell'ambito del progetto PON AIM "Metodi computazionali per il patrimonio culturale", il cui fine è proporre nuovi metodi, basati su Intelligenza Artificiale, per una migliore analisi e valorizzazione del patrimonio artistico digitalizzato.

PAROLE CHIAVE

Digital Humanities, intelligenza artificiale, computer vision, deep learning, data science.

INTERVENTO

1. INTRODUZIONE

Il patrimonio artistico riveste un ruolo di importanza strategica per la crescita economica e culturale di un Paese. In tempi recenti, grazie al progresso tecnologico, è stato compiuto uno sforzo di digitalizzazione su larga scala, che ha portato alla crescente disponibilità di grandi collezioni di opere d'arte digitalizzate. Un esempio particolarmente noto è WikiArt, la versione "artistica" di Wikipedia. La disponibilità di tali basi di conoscenza, unita ai recenti progressi nel campo dell'Intelligenza Artificiale, ha aperto nuove opportunità agli specialisti in questi settori perché assistano storici dell'arte, curatori di musei, appassionati, ecc., nello studio e in una migliore fruizione del patrimonio artistico digitalizzato. La fruizione di tale patrimonio da parte di un pubblico sempre più vasto promuove la diffusione della cultura ([5]).

La capacità di riconoscere caratteristiche, similarità, ecc., in e fra opere d'arte digitalizzate ricade inerentemente nel dominio della percezione estetica umana. Giacché tale percezione è fortemente soggettiva, e influenzata da diversi fattori, fra cui la conoscenza storica relativa all'opera, o l'emozione che l'opera suscita in chi l'osserva, essa è estremamente difficile da codificare ([7]). Tuttavia, grazie a tecniche di apprendimento automatico e rappresentazione della conoscenza, come quelle su cui si basano modelli di reti neurali allo stato dell'arte ([10]), è oggi possibile dotare le macchine di capacità percettive, che permettano loro di estrarre automaticamente caratteristiche descrittive di opere d'arte dalla loro codifica elementare in pixel. Tali rappresentazioni possono essere di beneficio per automatizzare numerosi compiti rilevanti dal punto di vista artistico, come la categorizzazione di un dipinto in base ad artista, stile e genere, oppure il ritrovamento di opere simili a una data opera in ordine a caratteristiche visuali, testuali, ecc.

Il progetto PON AIM "Metodi computazionali per il patrimonio culturale", svolto all'interno del Laboratorio di Intelligenza Computazionale del Dipartimento di Informatica dell'Università degli Studi di Bari, sta contribuendo alla ricerca interdisciplinare in questo settore, oggi molto attiva e fertile, al fine di sviluppare e applicare nuovi metodi, basati su Intelligenza Artificiale, per una migliore analisi e fruizione del patrimonio artistico digitalizzato.

2. LINEE DI RICERCA

Il nostro contributo si è focalizzato su quattro linee di ricerca:

- il *visual link retrieval* e il *knowledge discovery* in database di opere d'arte digitalizzate;
- il *clustering* automatico di tali opere;
- l'integrazione di informazione "contestuale" in modelli di Computer Vision, codificata per mezzo di *knowledge graph* (KG);

- l'integrazione di tali modelli nel robot sociale Pepper per l'interazione uomo-robot all'interno di musei o gallerie d'arte.

Maggiori dettagli nei seguenti sottoparagrafi.

2.1 VISUAL LINK RETRIEVAL

Uno degli elementi costitutivi la maggior parte delle analisi nelle arti visive è la ricerca di relazioni di “similarità” tra opere di artisti e scuole pittoriche differenti. Queste relazioni possono aiutare gli storici dell'arte a scoprire o meglio comprendere le influenze artistiche e i cambiamenti culturali avvenuti da un movimento artistico a un altro. Infatti, gli esperti d'arte raramente analizzano le opere visive come creazioni isolate, ma studiano queste all'interno di contesti più ampi, che coinvolgono influenze e connessioni tra diverse scuole.

Tradizionalmente, questo tipo di analisi è condotta manualmente ispezionando grandi raccolte di foto annotate. Tuttavia, la ricerca manuale fra migliaia di immagini, distribuite tra diversi periodi e scuole pittoriche, è un processo difficile e oneroso. Lungo questa direzione, abbiamo proposto un metodo a supporto del *visual link retrieval*, che si basa sull'uso di una rete neurale convoluzionale *deep* per l'estrazione di “feature” significative dalle immagini delle opere digitalizzate e su di un meccanismo di *nearest neighbor* non supervisionato per ritrovare link tra tali opere ([2]). La ricerca di link “visuali” è completamente non supervisionata, rendendo il metodo particolarmente utile nei casi in cui eventuali metadati siano scarsi, non disponibili o difficili da raccogliere. Vale la pena notare che il metodo proposto non solo fornisce quelle immagini che sono più simili alla *query* in input (v. Fig. 1), ma consente anche di studiare *pattern* storici analizzando un “grafo delle influenze” costruito sui link ritrovati. Infatti, applicando misure di grafo sulla rete complessa costruita sui link ottenuti, il metodo proposto permette una forma di *knowledge discovery* sugli artisti.

Il metodo proposto può essere vantaggioso non solo per gli storici dell'arte. Gli appassionati, infatti, possono beneficiare del ritrovamento automatico di link durante la visita di collezioni digitalizzate di musei e gallerie d'arte online, favorendo la fruizione delle collezioni digitali.

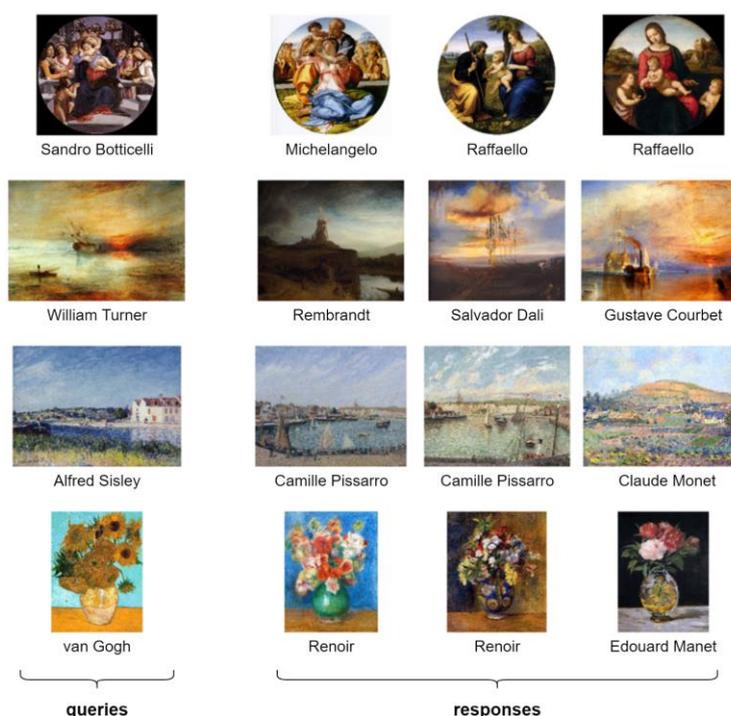


Figura 1. A sinistra, immagini in input al sistema; a destra, corrispondenti immagini restituite in output, considerate “visivamente” simili alle prime. L'algoritmo esclude, dal processo di *visual link retrieval*, opere prodotte dallo stesso artista della *query* in input.

2.2 ARTWORK CLUSTERING

Sebbene l'approccio descritto nel sottoparagrafo precedente sia adeguato per trovare opere d'arte collegate visivamente, esso non è efficace per raggruppare le opere appunto in gruppi ben distinti, poiché i dati appaiono distribuiti uniformemente all'interno di un singolo *cluster* omogeneo nello spazio delle feature. Avere un modello in grado di raggruppare le opere d'arte senza dipendere da etichette difficili da raccogliere o da un giudizio umano soggettivo può essere utile per molte applicazioni di dominio. Ad esempio, tale modello potrebbe essere utilizzato per scoprire periodi diversi nella produzione

di uno stesso artista. Allo stesso modo, potrebbe aiutare gli esperti del dominio a classificare l'arte contemporanea, che non può essere riccamente annotata.

A tal fine, abbiamo proposto un metodo che utilizza una rete neurale convoluzionale *deep* pre-addestrata per l'estrazione delle feature, ma che utilizza anche una componente di *deep clustering*, basata su di un *auto-encoder*, per effettuare il clustering ([6]). La scelta di tale soluzione è stata motivata dalla difficoltà di applicare metodi tradizionali di *clustering* o *dimensionality reduction* sia allo spazio dei pixel in input, dalla dimensionalità intrattabile, che allo spazio delle feature risultante dall'*embedding* di una rete neurale, specialmente quando le immagini in input sono immagini artistiche molto complesse.

Risultati sperimentali quantitativi e qualitativi hanno mostrato che il metodo proposto è in grado di trovare *cluster* ben separati sia quando si considera un insieme di dati ampio ed eterogeneo che abbracci diversi periodi, sia quando ci si concentra su opere prodotte da uno stesso artista (v. Fig. 2). In particolare, da un punto di vista qualitativo, sembra che il modello guardi non solo agli stilemi per raggruppare le opere, ma anche soprattutto agli attributi semantici relativi al contenuto della scena rappresentata; in altre parole, al genere dell'opera. Questa capacità sembra essere promettente per affrontare il noto problema della *cross depiction*, che rappresenta ancora una sfida aperta per la comunità scientifica ([9]). In effetti, questa capacità potrebbe essere sfruttata per trovare somiglianze tra opere d'arte indipendentemente dal modo in cui queste sono rappresentate.

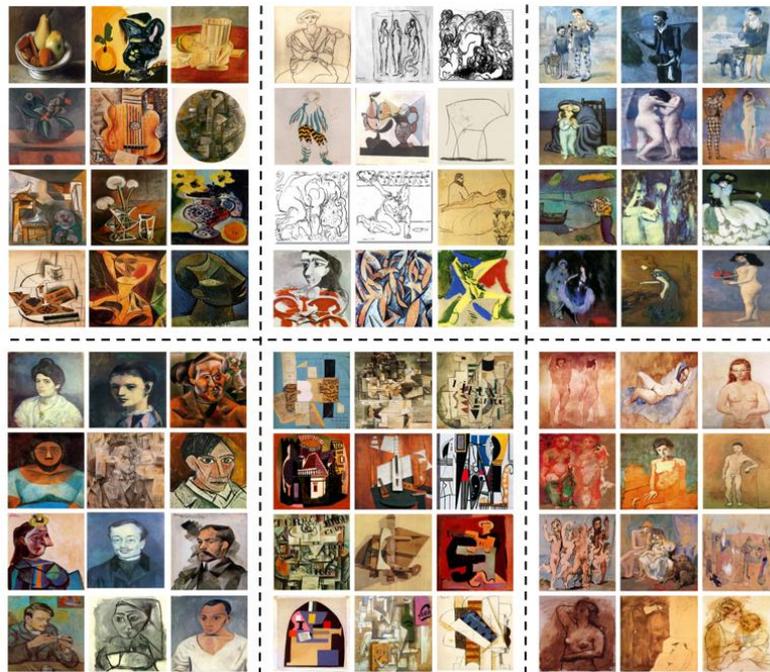


Figura 2. Cluster individuati automaticamente dal sistema tra opere prodotte da Pablo Picasso. I cluster contengono opere che condividono caratteristiche sia stilistiche che di contenuto.

2.3 KNOWLEDGE GRAPH

La ricerca ha poi mosso da un'altra considerazione: gran parte dei lavori in letteratura si basa esclusivamente sull'informazione codificata dai pixel caratterizzanti i dipinti digitalizzati. Sfortunatamente, questo approccio porta a ignorare una grande quantità di conoscenza di dominio, nonché relazioni e connessioni note tra opere d'arte e/o artisti, che potrebbero aumentare la qualità delle soluzioni esistenti ([8]). Le opere d'arte, infatti, non possono essere studiate solo sulla base del loro aspetto visivo, ma anche considerando vari altri fattori storici, sociali e "contestuali" che consentono di inquadrarle all'interno di un quadro più complesso. Pertanto, disporre di una base di conoscenza in cui non solo le opere d'arte, ma anche una ricca pletora di metadati, informazioni contestuali, descrizioni testuali, ecc., siano unificate all'interno di un unico framework strutturato può fornire una risorsa preziosa per sviluppare modelli più efficaci. Tale framework sarebbe utile non soltanto per gli utenti generici, che potrebbero sfruttare le informazioni codificate per navigare nella base di conoscenza, ma anche soprattutto per gli esperti d'arte, interessati a trovare nuove relazioni tra opere d'arte e/o artisti per una migliore comprensione dell'arte passata e contemporanea.

Per colmare questa lacuna, stiamo lavorando allo sviluppo di *ArtGraph*: un *knowledge graph* nel dominio artistico ([4]). Un KG fornisce una rappresentazione più espressiva e flessibile per incorporare relazioni di complessità arbitraria

tra entità concernenti l'arte, che non possono essere ottenute considerando solo il loro aspetto visivo. Il KG proposto integra le informazioni raccolte da WikiArt e DBpedia e sfrutta le potenzialità di un database NoSQL, Neo4j, che fornisce una modellazione basata su grafo altamente espressiva e un linguaggio di interrogazione molto potente. In questo modo, il database NoSQL aiuta già a fornire uno strumento per la scoperta di conoscenza senza addestrare esplicitamente un algoritmo di apprendimento. La conoscenza codificata in *ArtGraph* può essere poi integrata con feature apprese automaticamente per affrontare vari *task* tra cui la previsione di attributi, come stile e genere.

2.4 SOCIAL ROBOTICS

Man mano che le applicazioni degli algoritmi di Computer Vision a task nel dominio artistico diventano sempre più mature, una interessante implementazione di queste tecniche in contesti reali è incorporarle all'interno dei cosiddetti "robot sociali". Questi rappresentano un campo di ricerca emergente focalizzato sullo sviluppo di una "intelligenza sociale" che mira a mantenere l'illusione di avere a che fare con un essere umano ([1]). In questo contesto, i recenti progressi nella Computer Vision consentono ai ricercatori di dotare i robot di nuove e potenti capacità. Nella nostra ricerca stiamo usando un robot sociale, Pepper, come guida turistica di un museo ([3]). In particolare, stiamo sviluppando un approccio basato su visione per supportare i turisti durante la visita. Pepper è un robot semi-umanoide su ruote, dotato di diverse telecamere e sensori. Il modulo di visione consente a Pepper di percepire la presenza dei visitatori e di localizzarli nello spazio, stimandone età e genere. Inoltre, lo stesso modulo di *visual link retrieval* descritto in precedenza offre a Pepper la capacità di utilizzare l'immagine del dipinto osservato dal visitatore come una *query* visiva per cercare dipinti visivamente simili all'interno dello stesso museo. Il robot utilizza questi dati e altre informazioni acquisite durante il dialogo per fornire quindi ai visitatori consigli su opere d'arte simili a cui potrebbero essere interessati.

Progettare i comportamenti di un robot sociale che funga da guida museale richiede di dotarlo di diverse competenze che forniscano ai visitatori un'esperienza coinvolgente ed efficace durante la visita. Queste funzionalità hanno lo scopo di consentire al robot di rilevare e localizzare le persone nel museo, riconoscere l'opera d'arte che il visitatore sta guardando, profilare l'utente durante la visita al fine di generare raccomandazioni adeguate e, infine, coinvolgere le persone nell'interazione utilizzando adeguate capacità di conversazione. Abbiamo testato l'approccio proposto nel nostro laboratorio di ricerca ed esperimenti preliminari hanno dimostrato la sua fattibilità.

3. CONCLUSIONI

La crescente disponibilità di vaste collezioni di opere d'arte digitalizzate ha dato vita a una nuova intrigante area di ricerca in cui l'Intelligenza Artificiale e le arti visive si incontrano. Tale area di ricerca si inquadra all'interno dell'emergente Informatica Umanistica, che mira a unificare tecnologie digitali e discipline umanistiche. Le applicazioni sono innumerevoli e spaziano dal ritrovamento di informazioni in banche dati digitali alla generazione sintetica di nuove forme d'arte. Siamo fiduciosi che questo entusiasmante campo di ricerca sarà rafforzato sfruttando i rapidi progressi negli approcci di Deep Learning. Riteniamo che questi approcci continueranno a evolversi rapidamente, aprendo così la strada alla realizzazione di scenari sorprendenti in cui i sistemi informatici saranno in grado di analizzare e comprendere le "belle arti" autonomamente. Infatti, uno degli obiettivi finali di questa ricerca è la capacità delle macchine, quando opportunamente addestrate, di ricavare autonomamente una comprensione di ciò che la scena all'interno di un'opera rappresenta, qual è la metafora che ne è alla base, quali sono le possibili implicazioni storiche, ecc., senza alcuna supervisione umana.

Tuttavia, il dominio artistico è significativamente diverso dal dominio naturale/fotorealistico cui gli esperti di Intelligenza e Visione Artificiale sono avvezzi. In primo luogo, esiste una variabilità intrinseca tra le caratteristiche stilistiche e figurative dei due domini, nonché tra opere di artisti diversi appartenenti a uno stesso periodo, se non tra opere di uno stesso artista. Inoltre, i dataset con cui ora pre-addestriamo i modelli di Deep Learning sono affetti da "recentismo", e non sono cioè rappresentativi di situazioni, modi di essere/vestire, di scene iconografiche e mitologiche, ecc., del passato, che non sono mai esistiti oppure semplicemente non esistono più. In altre parole, il patrimonio artistico, dato il suo background storico evolutosi nel corso dei secoli, pone sfide scientifiche del tutto nuove e intriganti che, se affrontate, possono avanzare la comprensione automatica di una scena digitalizzata, per ora solo parzialmente ottenuta con i modelli attuali.

4. RINGRAZIAMENTI

Gennaro Vessio riconosce il sostegno finanziario del Ministero dell'Università e della Ricerca attraverso il progetto PON AIM 1852414.

BIBLIOGRAFIA

- [1] Castellano, Giovanna, Berardina De Carolis, Francesca D'Errico, et al. 2021. «PeppeRecycle: Improving Children's Attitude Toward Recycling by Playing with a Social Robot». *International Journal of Social Robotics* 13 (1): 97–111.
- [2] Castellano, Giovanna, Eufemia Lella, e Gennaro Vessio. 2021. «Visual link retrieval and knowledge discovery in painting datasets». *Multimedia Tools and Applications* 80 (5): 6599–6616.
- [3] Castellano, Giovanna, Nicola Macchiarulo, Bernardina De Carolis, e Gennaro Vessio. 2020. «Pepper4Museum: Towards a Human-like Museum Guide». *AVPCH@AVI*.
- [4] Castellano, Giovanna, Giovanni Sansaro, e Gennaro Vessio. 2021. «Integrating Contextual Knowledge to Visual Features for Fine Art Classification». *arXiv preprint* 2105.15028.
- [5] Castellano, Giovanna, e Gennaro Vessio. 2021. «Deep learning approaches to pattern extraction and recognition in paintings and drawings: an overview». *Neural Computing and Applications* 1 (20).
- [6] ———. 2021. «A deep learning approach to clustering visual arts». *arXiv* 2106.06234.
- [7] Cetinic, Eva, Sonja Grgic, e Lipic Tomislav. 2019. «A deep learning perspective on beauty, sentiment, and remembrance of art». *IEEE Access* 7.
- [8] Garcia, Nia, Benjamin Renoust, e Yuta Nakashima. 2020. «ContextNet: representation and exploration for painting classification and retrieval in context» *International Journal of Multimedia Information Retrieval* 9 (1): 17–30.
- [9] Hall, Peter, Cai Hongping, Wu Qi, e Tadeo Corradi. 2015. «Cross-depiction problem: Recognition and synthesis of photographs and artwork». *Computational Visual Media*, 1 (2): 91–103.
- [10] LeCun, Yann, Yoshua Bengio, e Geoffrey Hinton. 2015. «Deep learning». *Nature* 521: 436–44.