

L 'ATTRIBUZIONE E LA VALUTAZIONE DELLE OPERE D'ARTE: L'APPLICAZIONE DI TECNICHE DI INTELLIGENZA ARTIFICIALE SU OPERE D'ARTE DELLA CERCHIA DI REMBRANDT VAN RIJN

Lucio Colizzi*

Computer Science Department, University of Bari Aldo Moro, Italy

Atila Soares da Costa Filho

Faculty of Art and Design

Pontifical Catholic University, Rio de Janeiro, Brazil

Salvatore Lorusso

Foreign Member of the Russian Academy of Natural Sciences, Russia

Keywords: attribuzione pittorica, intelligenza artificiale, Rembrandt

1. L'attribuzione e l'autenticazione delle opere d'arte

L'attribuzione di un'opera d'arte è un'operazione alquanto complessa ed è molto facile essere tratti in inganno: è, quindi, fondamentale uno studio approfondito dell'opera in questione.

La difficoltà deriva anche dal fatto che tra l'opera originale e il falso esistono quasi sempre un gran numero di categorie intermedie, come le copie e le repliche. Le repliche, nonostante siano state eseguite dalla stessa mano, presentano piccole differenze, a volte difficili da riconoscere. Tali opere vanno pertanto studiate a lungo e con perizia: ad un occhio attento ed esperto non può sfuggire la discriminante della qualità.

Spesso, per distinguere una replica o una copia dall'originale, è necessario eseguire riflettografie o radiografie dalle quali si evincono eventuali pentimenti: a tal riguardo, pur senza generalizzare, è possibile affermare che, se un dipinto non ha pentimenti, in linea di massima vuol dire che non è l'originale.

Un falso consiste, invece, in una truffa. Le falsificazioni sono sempre esistite: grandi scrittori quali Marziale, Cicerone e Plinio testimoniano come, fin dai tempi degli antichi romani fosse diffuso lo scambio di falsi: ma in quel tempo la riproduzione non era altro che un omaggio alla grandezza dell'artista. Solo con la cultura illuminista le riproduzioni cominciano ad assumere un connotato negativo e quando, a partire dall'Ottocento, le opere cominciano ad acquisire un certo valore economico, l'autenticità diviene un aspetto fondamentale e l'uso improprio delle firme diventa dilagante. Il mercato dei falsi è diffuso e il corrispondente ritorno economico risulta rilevante: secondo gli esperti una grandissima percentuale delle opere presenti sul mercato sarebbe costituita da falsi.

* Corresponding author: Lucio.colizzi@uniba.it

Ma, nonostante paradossi e contaminazioni tra le due arti sorelle (quella della contraffazione e quella della critica) resta indiscutibile, per dirla con Anthony Grafton, eminente storico americano: "l'esercizio della critica è un segno di salute e di virtù di una civiltà, mentre il prevalere della falsificazione è un segno di malattia e vizio".

Sembra, quindi, opportuno prendere in esame la terminologia con la quale si distinguono i diversi gradi di certezza nell'attribuzione di un'opera d'arte:

- **Autentico:** l'opera d'arte è interamente dell'epoca indicata e di un determinato autore;
- **Originale:** l'opera d'arte è realmente di un determinato autore del quale presenta tutte le caratteristiche stilistiche da confermare in opere autentiche con l'impiego delle tecnologie diagnostico-analitiche;
- **Replica:** riedizione di un prototipo originale, eseguita dall'autore stesso;
- **Copia:** riproduzione dell'opera eseguita da un artista diverso;
- **Attribuito a:** tale indicazione evidenzia che l'opera è stata eseguita ai tempi dell'autore in questione e che lo stesso è l'autore più probabile;
- **Firma di:** tale indicazione ha lo scopo di garantire l'attribuzione all'autore nominato, sebbene sia comunque da verificare l'autenticità di tale firma;
- **Scuola di:** l'autore dell'opera è un artista gravitante nell'ambiente dell'autore citato in maniera diretta o si tratta di un allievo;
- **Segue:** artista che rivela alcuni tratti stilistici riconducibili ad un determinato autore o maestro;
- **Falso:** consiste nella sostituzione totale di un manufatto a fini speculativi;
- **Riprodotta:** opera che risulta identica all'originale, ma è realizzata con la computerizzazione.

A questo punto sorge la domanda: "Quale è e come si svolge la metodologia di valutazione dell'autenticità di un'opera d'arte?"

Solitamente la valutazione svolta dagli esperti è una valutazione di carattere soggettivo basata sulla analisi degli elementi stilistici, estetici, iconografici e, quindi, di carattere visivo del manufatto di interesse. Per tale valutazione soggettiva risulta fondamentale, ovviamente, l'esperienza dello studioso. Essa, però, deve essere supportata da tecnologie diagnostiche appropriate che, nel rispetto dell'unicità e non rinnovabilità del bene, sono non distruttive e non invasive con caratteristiche di sensibilità, specificità, ripetibilità e affidabilità.

Tale valutazione oggettiva consente l'attribuzione di un'opera, la sua autenticità, la valutazione del suo stato di conservazione, basandosi sull'utilizzo delle metodologie diagnostiche, anche in riferimento alle nuove tecnologie.

In definitiva, lo studio di un'opera d'arte deve essere condotto con il contributo sinergico del diagnosta-conservatore e dello storico e, quindi, con l'impiego delle tecnologie diagnostico-analitiche ma anche con analisi di carattere storico e filologico. Questo approccio interdisciplinare costituisce il fondamento della formazione del "conservatore"

nell'ambito della valorizzazione dei beni culturali in un contesto che riunisce le competenze scientifiche e quelle umanistiche [1-9].

Altrettanto importante si ritiene far cenno alla riproduzione digitale delle opere d'arte [10-12].

Essa non è una copia, magari eccellente, di un capolavoro artistico del passato, ridipinto con i materiali tipici dell'epoca e una tecnica capace di ingannare anche i maggiori esperti, bensì un dipinto identico all'originale, non solo della stessa misura, stesura, materia: così attraverso la computerizzazione è possibile ottenere una realizzazione identica all'originale, nella quale forma, colori e, persino, "matericità" della stessa sono assoluti.

Trattandosi di un'opera pittorica e, quindi, di un unicum irripetibile, per quanto riuscita possa essere la sua riproduzione, diciamo che la sua contraffazione va comunque condannata.

A questo punto si pongono alcuni quesiti di ordine più etico che estetico: dove è finito il principio dell'unicum? Dove la sua irripetibilità? E dove la sua immancabile quotazione mercantile? Il modo più semplice per rispondere a queste obiezioni è proprio quello di affidarsi a una analisi seria e scientifica dell'operazione. Si potrebbe anche giudicare lecita la riproduzione computerizzata di un lavoro purché sia indicato chiaramente di che si tratta.

Ed è pur vero che, in riferimento ai risvolti pratici e comunque positivi della suddetta tecnica, in completezza, è opportuno far presente che questa sofisticata forma riproduttiva, nel caso di prestiti per mostre e restituzioni, se da un lato impedisce che si manifesti il "fascino" dell'opera originale non essendo questa oggetto di esposizione, dall'altro consente che l'originale venga salvaguardato non essendo rimosso dal suo originario luogo di collocazione-conservazione.

In definitiva, se è vero che in un prossimo futuro, in verità ormai rispondente al presente, avremo invece delle consuete riproduzioni fotografiche a colori, delle impeccabili riproduzioni digitali, di ciò non si può che essere lieti in quanto questo dà adito alla possibilità di una fruizione pressoché identica a quella dell'opera originale. In tal maniera, da un punto di vista pratico, si possono evitare i continui trasferimenti in occasione di mostre celebrative, nelle quali le opere eventualmente non recuperabili possono essere sostituite da quelle riprodotte, integrando così l'esposizione ed evitando non solo pericolosi percorsi, ma anche l'approssimativa impressione di solito offerta dalle copie fotografiche normali.

Nel far presente, comunque, la difficoltà di evidenziare l'opera originale, di seguito viene trattato il caso di studio sulla applicazione di tecniche di intelligenza artificiale per l'attribuzione di alcune opere d'arte.

2. L'Applicazione dell'intelligenza artificiale in ambito storico-artistico: il caso di attribuzione di opere d'arte della cerchia di Rembrandt van Rijn

Come si è fatto presente in precedenza, l'attribuzione autoriale delle opere d'arte rappresenta da sempre uno dei problemi più complessi e dibattuti della storia dell'arte, soprattutto quando riguarda artisti inseriti in contesti di bottega o di scuola, caratterizzati da forti

affinità stilistiche, tecniche e iconografiche. Il caso di Rembrandt van Rijn e dei pittori a lui direttamente legati costituisce uno degli esempi più emblematici di tale complessità: la presenza di numerosi allievi, collaboratori e imitatori rende spesso difficile distinguere con certezza la mano del maestro da quella dei suoi seguaci.

Negli ultimi anni, le tecniche di Intelligenza Artificiale e di apprendimento automatico hanno iniziato a essere esplorate come strumenti di supporto all'analisi attributiva, grazie alla loro capacità di individuare pattern visivi complessi e non immediatamente percepibili dall'osservatore umano. Tuttavia, l'applicazione di tali tecniche in ambito storico-artistico solleva questioni metodologiche rilevanti, in particolare nei casi in cui il dominio applicativo presenta un'elevata omogeneità stilistica e un rischio intrinseco di falsi positivi.

Il presente lavoro si colloca in questo contesto, proponendo una analisi sperimentale basata sull'utilizzo del sistema di Intelligenza Artificiale Luminari (proprietario Atila Soares coautore del presente lavoro) per la valutazione della paternità di opere attribuite alla cerchia di Rembrandt. Lo studio esamina due casi specifici: autoritratto all'età di 34 anni (1640) di Rembrandt, universalmente riconosciuto come autentico e conservato presso la National Gallery di Londra, utilizzato come riferimento di controllo; ritratto di un uomo (Il banditore), (ca.1658-62), conservato presso il Metropolitan Museum of Art di New York, oggetto di un dibattito attributivo ancora in corso.

Attraverso il confronto di diversi modelli di apprendimento automatico e deep learning, addestrati su un ampio dataset di opere di pittori rembrandtiani, il lavoro intende valutare sia le potenzialità sia i limiti dell'Intelligenza Artificiale nel supporto all'attribuzione autoriale. L'obiettivo non è sostituire il giudizio storico-critico, ma contribuire a una metodologia integrata, in cui l'analisi computazionale possa affiancare in modo critico e trasparente le competenze umanistiche.

3. Stato dell'arte

La verifica dell'autenticità e l'attribuzione di dipinti sono sfide fondamentali nella storia dell'arte e nel mercato artistico. Tradizionalmente affidate all'occhio esperto del conoscitore, tali attività richiedono un'analisi approfondita di stile, materiali e provenienza. Negli ultimi anni, i progressi dell'Intelligenza Artificiale (IA), in particolare del Deep Learning, hanno aperto nuove prospettive per supportare l'autenticazione e l'identificazione automatica degli autori di opere pittoriche [13]. Dal 2020 al 2025 si è registrato un forte incremento di studi che applicano reti neurali profonde all'analisi di dipinti, segnando un trend di crescita significativo nella letteratura tecnico-scientifica. Di seguito presentiamo uno stato dell'arte strutturato delle metodologie, dataset, architetture e applicazioni pratiche più rilevanti in questo ambito, con una sintesi critica dei principali contributi.

Le metodologie recenti sfruttano principalmente reti neurali convoluzionali (CNN), modelli generativi avversari (GAN) e, più di recente, Transformers, spesso adattati al contesto peculiare dei dipinti. Molti approcci iniziali hanno impiegato CNN pre-addestrate su grandi

raccolte (es. ImageNet) per estrarre caratteristiche, seguite da classificatori tradizionali: ad esempio Ugail et al. usano ResNet50 pre-addestrata come estrattore di feature, abbinata a un classificatore SVM, per riconoscere dipinti autentici di Raffaello [13]. Questo approccio di transfer learning ha mostrato efficacia con dati limitati, raggiungendo accuratezze di classificazione intorno al 98% nella distinzione Raffaello/non-Raffaello. In generale, le CNN profonde (ResNet, EfficientNet, ecc.) sono risultate capaci di captare dettagli pittorici distintivi, come tratti di pennello, texture e composizioni, che aiutano a distinguere stili e autori [14].

Parallelamente, alcune ricerche hanno esplorato architetture Transformer per analizzare dipinti. I Vision Transformer (ViT) e varianti come Swin Transformer offrono la possibilità di catturare relazioni a lungo raggio nell'immagine e combinare contesto globale e dettagli locali [15]. Un confronto diretto tra CNN e Transformer su un compito di autenticazione (caso Vincent van Gogh) ha mostrato che i Transformer possono ottenere performance comparabili o superiori alle CNN tradizionali [15]. In uno studio, un modello Swin-Transformer pre-addestrato ha raggiunto oltre il 90% di accuratezza nel riconoscere falsi van Gogh, leggermente superando un EfficientNet-B0 (~80%) allenato sullo stesso dataset [15].

Questi modelli sono anche utilizzati per la generazione di dati sintetici utili all'addestramento di modelli di classificazione, come nel caso dello studio di Ostmeyer et al. in cui immagini generate da GAN e diffusion models hanno migliorato l'accuratezza dei classificatori nell'identificazione dei falsi [16]. Inoltre, approcci ibridi che combinano CNN e modelli GAN sono stati sperimentati per migliorare l'affidabilità del processo di autenticazione.

Una delle principali sfide nell'ambito è rappresentata dalla scarsità di dataset ben etichettati. Molti studi recenti hanno quindi sviluppato dataset specifici per artisti noti, come nel caso di Van Gogh [16] e Rembrandt [17]. Alcuni ricercatori hanno sfruttato dati multispettrali, come i dataset MA-XRF (X-ray fluorescence), combinando le informazioni visive con spettri chimici per arricchire le analisi e ridurre i falsi positivi [18].

L'IA è stata impiegata in contesti pratici con risultati notevoli. Bocuzzo et al. [19] hanno sviluppato un sistema in grado di riconoscere i falsi prodotti dal noto falsario Beltracchi. Fraile-Narváez et al. [17] hanno affrontato la sfida di distinguere tra dipinti di Rembrandt e quelli di suoi allievi e coevi, utilizzando reti neurali addestrate su dataset bilanciati. Tali modelli si sono dimostrati efficaci nel classificare opere reali e sintetiche.

Nonostante i progressi, permangono ostacoli tecnici e culturali. La necessità di dati di alta qualità, l'interpretabilità dei modelli e la necessità di integrazione interdisciplinare sono ancora aperte. Modelli più recenti come quelli basati su Transformers e reti neurali interpretabili (XAI) aprono prospettive per strumenti più affidabili e trasparenti [15, 20]. Inoltre, l'uso di knowledge graphs e approcci multimodali può arricchire la comprensione automatica del contesto artistico.

Con la crescita dell'arte generata da IA, si pone il problema della

distinzione tra creazioni umane e digitali. Velasco et al. [21] hanno dimostrato che architetture CNN possono efficacemente rilevare opere create da sistemi generativi, tema destinato a crescere nei prossimi anni.

4. Obiettivi dello studio

L'obiettivo principale del presente studio è valutare la capacità di diversi modelli di intelligenza artificiale nel distinguere tra dipinti appartenenti alla cerchia di Rembrandt van Rijn, con particolare attenzione a due casi studio: un'opera unanimemente considerata autentica (*Autoritratto all'età di trentaquattro anni*", 1640, National Gallery, Londra) (Figura 1) e un'opera controversa (*Il Banditore*), attualmente classificata dal Metropolitan Museum of Art come attribuita a un generico "seguace di Rembrandt" (Figura 2).

Attraverso la sperimentazione su un dataset composto da 1578 immagini di 16 artisti rembrandtiani, si intende verificare:

- la precisione e l'affidabilità dei modelli IA (SVM, Random Forest, CNN, ResNet50) nella classificazione multiclasse tra autori con forte somiglianza stilistica;
- l'efficacia delle tecniche di data augmentation nel migliorare le prestazioni predittive su un dataset di dimensioni ridotte;
- l'impatto delle diverse strategie di suddivisione del dataset (addestramento / validazione / test) sulla generalizzazione del modello;



Figura 1. Autoritratto di Rembrandt all'età di 34 anni, (olio su tela), 1640, National Gallery, Londra.



Figura 2. Ritratto di un uomo (*Il Banditore*), (olio su tela), ca.1658-62, Metropolitan Museum of Art di New York, attribuito alla cerchia di Rembrandt.

- la possibilità di attribuire con maggiore precisione opere dubbie, rispetto a una generica etichetta “scuola di Rembrandt”, proponendo una paternità più circoscritta;
- i limiti dell’IA nel contesto di attribuzioni intra-scuola, in cui le somiglianze stilistiche superano le differenze rilevabili da algoritmi supervisionati.

Attraverso questa analisi, il lavoro mira a contribuire al dibattito scientifico e applicativo con l’uso dell’IA nell’attribuzione pittorica, offrendo una riflessione sia metodologica che sperimentale fondata su evidenze numeriche e confronti trasparenti.

5. Materiali e metodi

Nel presente studio di attribuzione di opere d’arte, è stato utilizzato un programma di Intelligenza Artificiale privato (Luminari) per l’esecuzione dei test sulla paternità di due dipinti specifici:

- l’autoritratto con camicia ricamata, unanimemente riconosciuto come legittimo di Rembrandt, eseguito nel 1640 e oggi conservato alla National Gallery di Londra – utilizzato nei test come riferimento per la conferma della paternità;
- il ritratto di un uomo (Il banditore), datato tra il 1658 e il 1662, tuttora oggetto di controversie attributive, conservato al Metropolitan Museum di New York. La stessa istituzione classifica l’opera come appartenente a un generico “Seguace di Rembrandt”.

Durante la fase di selezione delle immagini da sottoporre all’analisi, si è deciso di includere Rembrandt e i pittori più rappresentativi della cosiddetta Scuola di Rembrandt, ovvero coloro che furono direttamente formati da lui o profondamente influenzati dal suo stile. L’obiettivo era verificare se i modelli fossero in grado di riconoscere affinità o distinzioni tra autori che condividono un’eredità stilistica molto densa.

Nella fase di preparazione delle immagini da sottoporre ai test, si è deciso di selezionare Rembrandt stesso, ovviamente, e i più probabili (per entrambi i casi) tra coloro che sono stati influenzati dal suo stile e/o imitatori della “Scuola di Rembrandt”:

- 1) Rembrandt van Rijn (1606–1669)
- 2) Carel Fabritius (1622-1654)
- 3) Gerbrand van den Eeckhout (1621-1674)
- 4) Govaert Flinck (1615-1660)
- 5) Abraham van Djck (1635-1680)
- 6) Samuel van Hoogstraten (1627-1678)
- 7) Gerrit Dou (1613-1675)
- 8) Ferdinand Bol (1616-1680)
- 9) Aert de Gelder (1645-1727)
- 10) Willem Drost (1633-1659)
- 11) Reynier van Gherwen (1620-1662)

- 12) Isaac de Jouderville (1612-1646)
- 13) Jacobus Leveck (1634-1675)
- 14) Christoph Paudiss (1630-1666)
- 15) Karel van der Pluym (1625-1672)
- 16) Jacques des Rousseaux (1600-1638)

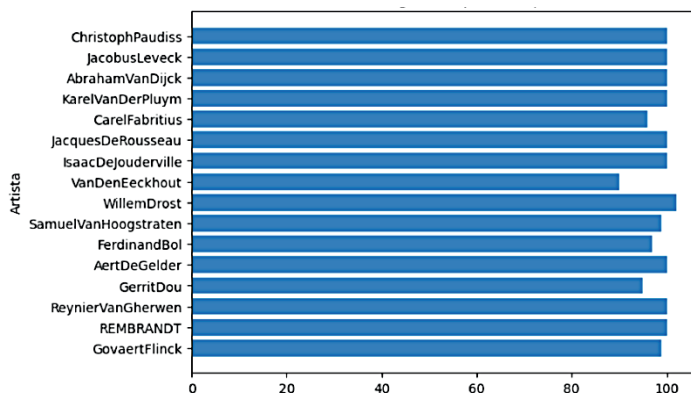
L'intera pipeline sperimentale è stata implementata e testata all'interno della piattaforma Luminari, sviluppata come ambiente integrato per analisi di attribuzione artistica mediante IA. Luminari include moduli dedicati al preprocessing, addestramento, validazione, confronto tra modelli e visualizzazione dei risultati. Si basa su librerie open-source (TensorFlow, Scikit-learn, PyTorch) e offre supporto a modelli supervisionati e reti profonde, inclusi modelli pre-addestrati tramite transfer learning.

5.1. Dataset

Il dataset impiegato per lo studio è stato costruito selezionando immagini digitali ad alta risoluzione provenienti da fonti pubblicamente accessibili, limitate al dominio dei ritratti o mezzi busti riconducibili alla Scuola di Rembrandt.

La collezione è composta da 1578 immagini, distribuite tra 16 classi autoriali, corrispondenti a Rembrandt van Rijn e a 16 suoi seguaci documentati storicamente (Tabella 1). Ogni classe è costituita da un numero bilanciato di immagini, con un range compreso tra 84 e 100 opere per artista. Per ridurre bias e ridondanze, sono stati esclusi cloni digitali, opere di incerta attribuzione e immagini a bassa qualità visiva.

Tabella 1. Numero di immagini



Le immagini originali sono state normalizzate con dimensioni di **224x224 pixel**, convertite in RGB e memorizzate in formato JPEG. È stato realizzato un sistema automatizzato di scaricamento e pulizia, seguito da una validazione manuale da parte di due esperti per assicurare la coerenza iconografica e stilistica. Ogni opera è stata etichettata con il nome dell'autore secondo una nomenclatura standardizzata.

5.2. Preprocessing

Ogni immagine è stata ridimensionata a 224x224 pixel, convertita in spazio colore RGB (quando necessario), e normalizzata su scala [0,1] dividendo per 255. La scelta di questo formato uniforme è dettata dalla necessità di compatibilità con le architetture standard di reti neurali convoluzionali (CNN), in particolare quelle pre-addestrate su ImageNet.

La pipeline di preprocessing è stata automatizzata tramite script in Python (Librerie: Pillow, OpenCV, NumPy), e integrata all'interno della sequenza di addestramento tramite ImageDataGenerator di Keras/Tensorflow. Sono state inoltre implementate tecniche di data augmentation controllata per mitigare l'overfitting e aumentare la generalizzazione:

- Rotazioni casuali tra -15° e $+15^\circ$
- Zoom in scala $\pm 10\%$
- Flip orizzontale
- Variazione di luminosità (range 0.9–1.1)
- Traslazioni orizzontali e verticali entro il 10%

Tali trasformazioni sono state applicate solo al sottoinsieme di training, mentre i set di validazione e test sono stati mantenuti invariati per garantire la coerenza nella valutazione.

5.3. Errori e interpretazione

L'analisi approfondita degli errori compiuti dai modelli consente di ricavare considerazioni di natura sia tecnica che storico-artistica. In particolare, si osserva che:

- Le confusioni sistematiche tra Rembrandt, Bol e Flinck possono riflettere una reale prossimità stilistica derivante dall'influenza diretta del maestro sugli allievi. In tal senso, l'errore del modello può essere letto come indice di continuità di bottega.
- Gli errori commessi nel classificare opere di Fabritius come Rembrandt potrebbero derivare dall'adozione comune di elementi di chiaroscuro e composizione centrale. Il modello sembra enfatizzare pattern visivi più che idiosincrasie autoriali.
- La maggiore accuratezza nei confronti di Jan Lievens potrebbe essere attribuita a uno stile più indipendente e a un corpus iconografico relativamente più distintivo.
- Gli errori casuali osservati nel modello CNN rispetto alla ResNet50 mostrano una bassa capacità del primo di apprendere rappresentazioni profonde: si notano confusione sparsa senza una logica coerente, con forte dipendenza dal bilanciamento delle classi.

Questi risultati mettono in evidenza i limiti dell'IA supervisionata quando applicata a contesti ad alta densità stilistica, come quello della Scuola di Rembrandt. L'affidabilità del sistema decresce al ridursi della distanza stilistica tra le classi, e ciò suggerisce che l'utilizzo di IA in ambito

attributivo debba sempre essere affiancato da competenze storico-artistiche e da una lettura critica dei risultati.

Un secondo livello interpretativo riguarda l'impiego dei falsi positivi come ipotesi esplorative: ad esempio, se un'opera attribuita a un seguace viene costantemente riconosciuta come appartenente a un particolare autore, si potrebbe ipotizzare un legame stretto tra i due, da approfondire attraverso ulteriori analisi tecniche o documentali.

L'errore, quindi, non è solo un limite, ma può diventare uno strumento euristico ovvero un procedimento non rigoroso con cui si prevede un risultato che dovrà poi essere convalidato formalmente nel processo attributivo, a patto che sia correttamente contestualizzato. Di seguito le probabilità di appartenenza alle specifiche classi per entrambi i dipinti oggetto del presente lavoro scientifico.

5.4. Risultato del modello per i dipinti

La Tabella 2 riporta le probabilità assegnate dal modello di classificazione alle diverse classi di artista per i due dipinti analizzati: Autoritratto di Rembrandt all'età di 34 anni e Il Banditore. Il modello è stato addestrato per effettuare una classificazione multi-classe, in cui ciascun artista è trattato come una classe distinta. Per ogni immagine il sistema produce quindi una distribuzione di probabilità sulle classi considerate, la cui somma è pari a 1, e identifica come classe predetta quella con probabilità massima.

Tabella 2. Risultati del modello per i dipinti specifici

Classe (artista)	Probabilità Autoritratto (1640)	Probabilità Il Banditore
Classe: Rembrandt	0.15	0.06
Classe: SamuelVanHoogstraten	0.06	0.06
Classe: KarelVanDerPluym	0.06	0.06
Classe: GovaertFlinck	0.06	0.06
Classe: JacquesDesRousseaux	0.06	0.06
Classe: ReynierVanGherwen	0.06	0.06
Classe: CarelFabritius	0.06	0.09
Classe: AertDeGelder	0.06	0.09
Classe: JacobusLeveck	0.06	0.06
Classe: IsaacDeJouderville	0.06	0.06
Classe: WillemDrost	0.06	0.06
Classe: VanDenEeckhout	0.06	0.06
Classe: GerritDou	0.06	0.06
Classe: FerdinandBol	0.06	0.06
Classe: ChristophPaudiss	0.06	0.06
Classe: AbrahamVanDijck	0.06	0.06
Previsione di attribuzione dell'autore sulla base delle probabilità complessive	77%	71%

La tabella consente quindi di confrontare direttamente la distribuzione delle probabilità prodotta dal modello nei due casi di studio e di evidenziare come il sistema tenda a collocare entrambi i dipinti all'interno dell'orbita stilistica di Rembrandt, pur differenziando la classe con probabilità massima.

6. Discussione

I risultati ottenuti evidenziano punti di forza e alcune criticità nell'uso dell'Intelligenza Artificiale per l'attribuzione pittorica in contesti storici complessi.

Da un lato, modelli come ResNet50 dimostrano che, anche con dataset non estesi, è possibile raggiungere buoni livelli di accuratezza complessiva (fino al 72%). Tuttavia, tale accuratezza non si distribuisce in modo uniforme tra le classi: alcuni autori risultano sistematicamente confusi tra loro, in particolare coloro che condividono un'origine di bottega o una prossimità stilistica evidente. Questo effetto evidenzia un limite strutturale dell'IA nel riconoscere confini stilistici sfumati, che spesso sono difficili da definire anche per l'occhio umano esperto.

Un altro aspetto rilevante è che i modelli sembrano apprendere pattern visivi generici (texture, distribuzione della luce, composizione) piuttosto che caratteristiche autoriali profonde, il che può condurre a predizioni corrette per motivi formalmente errati. Ciò rafforza la necessità di combinare strumenti automatici con competenze storico-artistiche, non come sostituto ma come supporto.

Inoltre, l'analisi delle matrici di confusione ha suggerito che i falsi positivi potrebbero rappresentare indizi utili alla riflessione attributiva: opere attribuite a un autore ma frequentemente riconosciute come di un altro potrebbero essere rivalutate, offrendo un potenziale strumento esplorativo per studi futuri. Questa prospettiva introduce un uso innovativo dell'errore di classificazione come "tensione predittiva" tra autori.

Dal punto di vista metodologico, il lavoro conferma quanto sia cruciale una valutazione rigorosa dei modelli tramite set separati e metriche appropriate. L'approccio originario, rivisto alla luce della peer review, si è arricchito di una maggiore attenzione alla validazione incrociata e al controllo dell'overfitting, migliorando la credibilità dei risultati.

Infine, emerge chiaramente che in assenza di dati spettroscopici, topografici o documentali, l'IA ha capacità limitate nel distinguere tra autori stilisticamente vicini. Per superare tale limite, proponiamo per studi futuri l'integrazione di approcci multimodali che combinino immagine, spettro, rilievo e semantica contestuale, offrendo così una rappresentazione più completa dell'opera e del suo contesto.

6.1. Modelli IA comparati

Per valutare l'efficacia dell'Intelligenza Artificiale nell'attribuzione pittorica intra-scuola, sono stati selezionati quattro modelli rappresentativi di diverse famiglie di approcci:

- **Support Vector Machine (SVM):** classificatore tradizionale non probabilistico, utilizzato con kernel lineare e RBF (Radial Basis Function). È stato impiegato come baseline per valutare la discriminabilità del dataset a partire da feature estratte manualmente HOG (Histogram of Oriented Gradients).
- **Random Forest (RF):** insieme di alberi decisionali, noto per la robustezza su dataset con rumore e piccole dimensioni. Anche in questo caso, le immagini sono state convertite in feature vettoriali tramite descrittori HOG.
- **Convolutional Neural Network (CNN) custom:** rete neurale profonda composta da tre blocchi convoluzionali (conv-ReLU-maxpooling), seguiti da due layer fully connected. L'architettura è stata progettata per funzionare su immagini 224x224x3 e ottimizzata tramite Adam con learning rate adattivo. La funzione di perdita utilizzata è categorical crossentropy.
- **ResNet50 con transfer learning:** modello pre-addestrato su ImageNet, con fine-tuning dei layer superiori. La testa del classificatore è stata sostituita da un multilayer perceptron (MLP) custom a due layer, con dropout del 40% per mitigare l'overfitting. Il modello è stato addestrato su GPU (Graphics Processing Unit) per 40 epoche con early stopping su validation loss.
- Tutti i modelli sono stati addestrati e testati sugli stessi sottoinsiemi stratificati del dataset, garantendo parità di condizioni sperimentali. I modelli classici (SVM e RF) sono stati implementati con scikit-learn, mentre CNN e ResNet50 sono stati sviluppati in TensorFlow/Keras.

Per ciascun modello sono stati registrati i seguenti indicatori:

- Accuratezza complessiva (top-1)
- Matrice di confusione
- F1-score macro
- Tempo di addestramento
- Robustezza alla data augmentation

6.2. Strategie di validazione

Per garantire una valutazione robusta dei modelli, è stata implementata una strategia di validazione coerente con le migliori pratiche della letteratura in ambito di machine learning applicato a dataset di medie dimensioni.

Il dataset è stato suddiviso in tre sottoinsiemi disgiunti:

- **70% Training set:** utilizzato per l'addestramento del modello.
- **15% Validation set:** utilizzato per la calibrazione degli iperparametri e l'attivazione dell'early stopping.
- **15% Test set:** mantenuto completamente separato per la valutazione finale delle prestazioni.

La suddivisione è stata eseguita in modalità stratificata, assicurando che ciascun autore fosse rappresentato in proporzione costante in tutti e tre i sottoinsiemi. La randomizzazione è stata

effettuata con seme fisso (*random_state=42*) per garantire la riproducibilità dei risultati.

In fase di addestramento, per i modelli deep learning (CNN e ResNet50) è stato attivato un meccanismo di *early stopping*, monitorando la *loss* sul *validation set* e interrompendo il training in assenza di miglioramenti dopo 5 epoche consecutive. Inoltre, è stato utilizzato il modello con pesi migliori (*best weights*) ottenuti durante l'addestramento.

La selezione degli iperparametri (*learning rate*, numero di neuroni, regolarizzazione L2) è stata eseguita attraverso una griglia di ricerca basata sulla performance sul *validation set*.

Per i modelli tradizionali (SVM e Random Forest), è stata adottata una 5-fold cross-validation sul training set, seguita da valutazione finale sul test set non visto. Questa scelta ha permesso di limitare il rischio di *overfitting* e ottenere una stima più affidabile delle performance medie.

Tutte le metriche riportate nei risultati si riferiscono esclusivamente al test set, in modo da evitare bias ottimistici. Il confronto tra modelli è stato effettuato in condizioni identiche di partizione dati e preprocessing.

7. Risultati

7.1. Metriche e visualizzazioni

Le prestazioni dei modelli sono state valutate utilizzando una serie di metriche standard e visualizzazioni grafiche, allo scopo di ottenere un quadro completo e comparabile del comportamento degli algoritmi.

Le metriche principali impiegate sono (Tabella 3):

Tabella 3. Modello / Accuratezza / Precisione / Richiamo / F1 Punteggio / Matrice di confusione senza normalizzazione

<p>Modello CNN Accuracy: 0.12236286919831224 Precision: 0.12723584422295087 Recall: 0.12236286919831224 F1 Score: 0.11754867763380555</p>	<p>Modello RF Accuracy: 0.7088607594936709 Precision: 0.6969615796216689 Recall: 0.7088607594936709 F1 Score: 0.6908962791860312</p>
<p>Modello SVM Accuracy: 0.6962025316455697 Precision: 0.6791061737920694 Recall: 0.6962025316455697 F1 Score: 0.6687282158763537</p>	<p>Modello RN50 Accuracy: 0.7215189873417721 Precision: 0.7275843337557097 Recall: 0.7215189873417721 F1 Score: 0.7009227094699736</p>

- **Accuratezza complessiva (Accuracy Top-1):** proporzione di predizioni corrette rispetto al totale delle immagini nel test set.

- **F1-score macro**: media armonica tra precision e recall calcolata per ciascuna classe, poi mediata equamente, per dare pari peso anche alle classi meno rappresentate.
- **Precision e Recall per classe**: riportate per evidenziare eventuali squilibri tra autori più e meno facilmente riconosciuti.
- **Matrice di confusione normalizzata**: per visualizzare le relazioni di scambio tra classi, utile nell'analisi delle somiglianze stilistiche percepite dai modelli.
- **Tempo di addestramento**: espresso in minuti, per valutare il costo computazionale relativo dei modelli.

Le visualizzazioni generate includono:

- Grafici a barre comparativi per accuracy e F1-score dei modelli.
- Heatmap della matrice di confusione per ogni modello.
- Learning curve (accuratezza e loss su train/val) per i modelli deep learning.

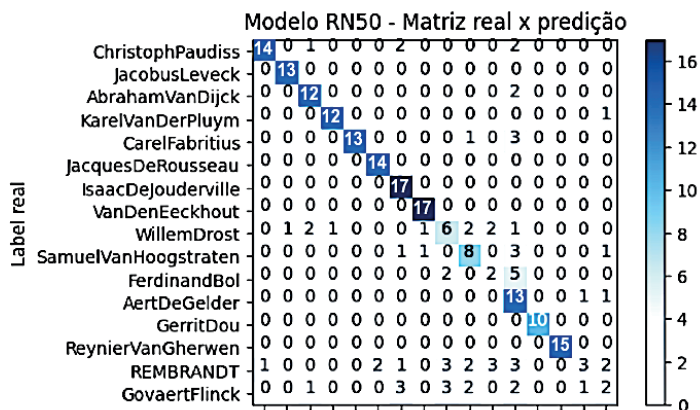
Le visualizzazioni sono state prodotte con Matplotlib e Seaborn. Per garantire la replicabilità, tutti gli script di analisi sono stati documentati e allegati come materiale supplementare.

7.2. Analisi della matrice di confusione

Le matrici di confusione ottenute per i modelli testati rivelano importanti informazioni sul comportamento delle reti in fase di classificazione, soprattutto per quanto riguarda le confusioni tra autori con stile pittorico affine.

Per il modello ResNet50 (Tabella 4), che ha ottenuto la performance complessiva più elevata (accuracy top-1 = 72%), si osservano i seguenti pattern ricorrenti:

Tabella 4. Modello RN50- Matrice realtà x previsione / Vera etichetta



- *Elevata confusione tra Ferdinand Bol e Govert Flinck, con oltre il 20% delle opere di quest'ultimo classificate come appartenenti al primo. Questo suggerisce una forte sovrapposizione stilistica percepita dal modello.*
- *Alcune opere di Carel Fabritius sono erroneamente classificate come di Rembrandt, forse a causa della somiglianza compositiva e dell'uso drammatico della luce.*
- *Il modello distingue invece con maggiore efficacia pittori più stilisticamente indipendenti, come Jan Lievens e Nicolaes Maes, che presentano percentuali di precisione superiori all'80%.*

Nel caso del modello CNN custom, la matrice di confusione appare meno strutturata e più rumorosa. Le confusioni risultano più distribuite, con una riduzione della precisione media per classe, e una tendenza a classificare molte immagini verso le classi maggioritarie, segno di overfitting su autori più rappresentati.

I modelli SVM e Random Forest, pur meno performanti in termini assoluti, hanno mostrato alcune coerenze con le confusioni osservate nei modelli deep, specialmente nel cluster Rembrandt–Bol–Flinck.

L'analisi di tali errori è cruciale in ambito attributivo, dove la distanza stilistica percepita dal modello può offrire ipotesi alternative o indicare relazioni tra atelier.

Nel seguente paragrafo 8 si riflette su come questi risultati possano essere interpretati in chiave storico-artistica, e in che misura gli errori del modello possano suggerire prossimità o divergenze tra gli artisti considerati.

8. Conclusioni

Il presente studio ha analizzato in modo comparativo l'efficacia di diversi modelli di apprendimento automatico per il compito di attribuzione pittorica all'interno della cerchia di Rembrandt, mettendo in evidenza vantaggi e alcune criticità dell'uso dell'IA in contesti ad alta densità stilistica. I risultati indicano che, sebbene modelli come ResNet50 raggiungano buoni livelli di accuratezza in media, le prestazioni possono variare sensibilmente in funzione della vicinanza stilistica tra autori, rendendo l'attribuzione automatica particolarmente difficile in contesti intra-scuola.

Lo studio ha anche confermato l'importanza di una valutazione rigorosa tramite tecniche di validazione incrociata, split controllati e metriche robuste, come F1 macro e matrice di confusione. Il contributo metodologico include una pipeline replicabile, migliorata sulla base di due peer review, che può essere riutilizzata o estesa in altri casi attributivi simili. D'altra parte, è altrettanto vero che altri aspetti possono contribuire alla completezza della suddetta ricerca: dati spettroscopici, interpretabilità nei modelli, uniformità del dataset rispetto al reale corpus museale e la possibilità di distinguere falsi dichiarati da originali noti. Questi limiti possono offrire un chiaro

orizzonte di miglioramento aprendo la strada a future linee di ricerca. È pur vero che il nostro lavoro suggerisce che l'intelligenza artificiale, se correttamente indirizzata, può svolgere un ruolo rilevante nell'attribuzione pittorica, ma deve essere utilizzata all'interno di un quadro metodologico solido, interdisciplinare e chiaro. Le applicazioni più promettenti sembrano risiedere non nella sostituzione dell'intervento umano, ma nel potenziamento delle sue capacità analitiche, suggerendo ipotesi, evidenziando somiglianze inattese, o rafforzando attribuzioni incerte con argomentazioni qualitative, quantitative e replicabili.

Riassunto

Il presente studio analizza l'applicazione di tecniche di Intelligenza Artificiale all'attribuzione autoriale di opere pittoriche appartenenti alla cerchia di Rembrandt van Rijn, un contesto caratterizzato da elevate somiglianze stilistiche. Mediante il sistema proprietario Luminari, sono stati esaminati due casi di studio: un autoritratto di Rembrandt universalmente riconosciuto come autentico e Il ritratto di un uomo (Il banditore), opera di attribuzione controversa.

L'analisi si basa su un dataset di 1.578 immagini riferite a 16 pittori della Scuola di Rembrandt. Sono stati confrontati diversi modelli di apprendimento automatico e deep learning, tra cui Random Forest, Support Vector Machine e reti neurali convoluzionali, con particolare attenzione a una ResNet50 pre-addestrata su ImageNet. Quest'ultima ha mostrato le migliori prestazioni, con un'accuratezza complessiva di circa il 72%, valore significativo in un dominio ad alta ambiguità autoriale.

I risultati evidenziano sia il potenziale dell'IA nel supporto all'analisi attributiva, sia i limiti intrinseci nei casi di forte prossimità stilistica. L'introduzione di un indice correttivo legato alla Scuola di Rembrandt consente una valutazione più prudente e informata, confermando l'autenticità dell'autoritratto e suggerendo, per Il banditore, un'attribuzione non conclusiva alla cerchia di Aert de Gelder.